Structural studies on protein scaffolds related to muscle physiology and disease:
The titin filament, its associated component MuRF-1 and nuclear LAP2α

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie vorgelegt vor
der Philosophisch-Naturwissenschaftlichen Fakultät der
Universität Basel

Von
Michael Christian Mrosek
aus Ingelheim am Rhein (Deutschland)

Basel, 2006
Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Olga Mayans
Prof. Dr. Ueli Aebi
Prof. Dr. Ulrich Baumann

Basel, den 24.10.2006,

Prof. Dr. Hans-Peter Hauri
Dekan
Declaration

I declare that I wrote this thesis, “Structural studies on protein scaffolds related to muscle physiology and disease: The titin filament, its associated component MuRF-1 and nuclear LAP2α”, with the help indicated and only handed it to the faculty of Science of the University of Basel and to no other faculty and no other university.
Abstract

The titin molecule has a length of over 1 µm and functions as a colossal protein scaffold in the muscle sarcomere. Up to 90% of its total mass is composed of repetitive immunoglobulin (Ig) and fibronectin (FnIII) domains that form linear tandems interspersed by unique sequences, among them a Ser/Thr kinase domain located at its C-terminus. The distinct pattern of Ig and FnIII motifs N-terminal to the kinase domain is conserved in other ‚giant kinases‘ and invertebrate titin homologues. In vertebrate titin, it is involved in the specific recruitment of the ubiquitin ligase MuRF-1 to the filament. MuRF-1 is involved in the pathological atrophy of skeletal and cardiac muscle.

We have determined the crystal structure of titin A168-A170 comprising two Ig and one FnIII domains and established its binding to MuRF-1 in solution. We analysed the structure with the aim to understand the interdomain relationships between repetitive Ig and FnIII subunits in titin as well as to shed light into the molecular determinants that confer specificity to ligand binding on the scaffold and in particular in the M-line interface to MuRF-1. A168-A170 shows an extended, rigid architecture. Its surface displays a shallow groove along its full length as well as a unique loop protrusion, both features conceivably mediating MuRF-1 binding. Moreover, our ITC data show that binding occurs with high affinity between residues 166-315 of MuRF-1. These data suggest that A168-A170 is of interest to attempt therapeutic inhibition of MuRF-1-mediated muscle turnover.

In addition we have elucidated the structure of the B-box domain of MuRF-1 to further investigate the role of MuRF-1 in homo- and hetero-oligomeric interactions at the M-line region. We found that MuRF-1 B-box adopts a RING-finger-like fold and exists in a dimeric state in solution. The domain possesses characteristic surface properties that are likely to mediate interactions of MuRF-1 with other sarcomeric components that are important in MuRF-1 function at the M-line.

Finally, we have also carried out the biophysical characterization of the nuclear adaptor protein LAP2α that interacts with the nuclear lamina scaffold. Conceptually, LAP2α and the nuclear lamina are closely related systems to MuRF-1 and titin. This characterization, whose ultimate finality is to understand the interaction of LAP2α with lamin A/C establishes now the basis for a future structure elucidation.
This work illustrates how scaffold protein systems, which are structural skeletons composed of multiple repetitive units, can become functionalized by the recruitment of specific shuttle proteins to their surface. Specific binding in such systems involves steric factors as well as the evolution of unique sequence inserts at defined locations. Recruited proteins often act as adaptors that, in turn, attract other cellular components. They often result in large, heterogeneous molecular assemblies that amplify the physiological response. In the case of titin, the potential formation of a signalosome assembly at its M-line, surrounding a kinase domain, is thought to mediate mechano-transduction pathways involved in the regulation of myofibril turn-over and, thereby, in the adaptative remodelling of muscle to mechanical load.
# Table of contents

1 INTRODUCTION .................................................................................................................. 1

1.1 Protein-protein interactions ................................................................................................. 1
    1.1.1 Network modules ........................................................................................................... 1
    1.1.2 Protein complexes ......................................................................................................... 3
    1.1.3 Advantages of complex formation .............................................................................. 5
    1.1.4 Scaffolding complexes enforce proximity ................................................................. 6
    1.1.5 Functions of Scaffold proteins .................................................................................... 6
    1.1.6 The Immunoglobulin superfamily (IgSF) ................................................................. 7

1.2 Titin’s M-line: a structural scaffold and signal transduction “module” ................................. 10
    1.2.1 Structure of the sarcomere .......................................................................................... 10
    1.2.2 Overall structure of the third filament of the sarcomere: titin ................................. 12
    1.2.3 Domain composition and interactions of the giant scaffold titin ............................... 13
    1.2.4 I-Band titin ................................................................................................................. 13
    1.2.5 A-band titin ................................................................................................................. 15
    1.2.6 M-line titin .................................................................................................................. 16
    1.2.7 Overview of existing structures of titin ...................................................................... 19
    1.2.8 Hereditary titin diseases ............................................................................................ 22
    1.2.9 MuRFs ......................................................................................................................... 23
    1.2.10 Functions of MuRF proteins ..................................................................................... 24
    1.2.11 The titin-interface with MuRF-1 ............................................................................. 26
    1.2.12 Function of MuRF-1 in M-line integrity and sarcomeric protein turnover .............. 27
    1.2.13 Multiple localisation of MuRF proteins and the SUMO pathway .......................... 28
    1.2.14 Titin as a sarcomeric stretch sensor: the kinase signalosome ............................... 29

1.3 The B-box, protein-protein interaction domain of the titin-associated protein MuRF-1 ...... 32
    1.3.1 MuRF-1 is a member of a protein family with the characteristic TRIM fold ............. 32
    1.3.2 TRIM components ..................................................................................................... 32
    1.3.3 Variable C-terminal domain of TRIM containing proteins ..................................... 36
    1.3.4 The TRIM as an integrated fold ................................................................................. 37

1.4 LAP2α, a nuclear scaffold protein involved in chromatin organization .............................. 38
    1.4.1 Nucleus, nuclear envelope and lamina ...................................................................... 38
    1.4.2 Structure and interactions of lamins ........................................................................... 40
    1.4.3 Structure and interactions of LAP2 isoforms α and β ............................................... 41
    1.4.4 Function of LAP2s and binding partners in nuclear structure and dynamics .......... 44
    1.4.5 Diseases associated with lamins ................................................................................. 45
1.5 Challenges of structural studies on filamentous proteins and protein scaffolds ..................... 47

1.6 Focus of the work .................................................................................................................................. 49

2 EVIDENCE FOR THE RECRUITMENT OF MURF-1 TO TITIN .......... 59

2.1 Materials and Methods .......................................................................................................................... 60
  2.1.1 Cloning ............................................................................................................................................ 60
  2.1.2 Protein production ......................................................................................................................... 60
  2.1.3 Crystallization of titin A168-A170 ................................................................................................. 63
  2.1.4 Collection of a native data set ....................................................................................................... 66
  2.1.5 Detection of non-crystallographic symmetry (NCS) .................................................................. 67
  2.1.6 Phasing attempts by Molecular Replacement .............................................................................. 69
  2.1.7 MAD data collection ...................................................................................................................... 70
  2.1.8 Data processing ............................................................................................................................ 71
  2.1.9 Determination of the sub-structure of anomalous scatterers ...................................................... 72
  2.1.10 Density modification .................................................................................................................. 74
  2.1.11 Model building and refinement ................................................................................................. 75
  2.1.12 Isothermal titration calorimetry (ITC) ..................................................................................... 77

2.2 Results .................................................................................................................................................. 79
  2.2.1 Structure of A168-A170 ................................................................................................................. 79
  2.2.2 Domain interfaces ....................................................................................................................... 80
  2.2.3 MuRF-1 binding ........................................................................................................................... 85

2.3 Discussion ............................................................................................................................................. 87

3 THE B-BOX 2 DOMAIN OF MURF-1 ......................................................................................... 90

3.1 Materials and Methods .......................................................................................................................... 91
  3.1.1 Cloning ............................................................................................................................................ 91
  3.1.2 Protein production ......................................................................................................................... 91
  3.1.3 Crystallization ............................................................................................................................... 94
  3.1.4 Data collection of the high resolution set .................................................................................... 95
  3.1.5 MAD data collection .................................................................................................................. 97
  3.1.6 Determination of heavy atom sub-structure and phase calculation ........................................ 99
  3.1.7 Model building ........................................................................................................................... 102
  3.1.8 Crystal packing and Dimeric arrangements ............................................................................... 106
  3.1.9 Crystallization and data collection of the C24D mutant of MuRF-1 B-box ................................. 107
  3.1.10 Phase determination and Model building ................................................................................ 108
  3.1.11 NMR experiments .................................................................................................................... 108
3.2 Results and Discussion........................................................................................................109
3.2.1 Crystal structure of MuRF-1 B2.........................................................................................109
3.2.2 Oligomeric state of MuRF-1 B2..........................................................................................111
3.2.3 Structure of the MuRF-1 B2 dimer......................................................................................114
3.2.4 Comparative analysis of B2 and RING finger-like motifs....................................................116
3.2.5 Final remark.......................................................................................................................119

4 BIOPHYSICAL CHARACTERIZATION OF LAP2α..........................................................122

4.1 Materials and Methods.......................................................................................................123
4.1.1 Cloning ..............................................................................................................................123
4.1.2 Protein preparation..............................................................................................................123
4.1.3 Mass spectrometry (MS).....................................................................................................124
4.1.4 Circular dichroism spectroscopy (CD)................................................................................124
4.1.5 Analytical Ultracentrifugation (AUC)..............................................................................125
4.1.6 Transmission electron microscopy (TEM).........................................................................125
4.1.7 Size exclusion chromatography with multi-angle light scattering (SEC-MALS).................125

4.2 Results..................................................................................................................................127
4.2.1 Characteristics of LAP2α constructs..................................................................................127
4.2.2 Protein preparation of LAP2α410-693..................................................................................128
4.2.3 Overexpression and purification of full-length LAP2α1-693..............................................130
4.2.4 Mass spectrometry on LAP2α410-693.................................................................................131
4.2.5 Circular dichroism spectroscopy on LAP2α410-693 and LAP2α1-693..............................132
4.2.6 Size exclusion chromatography with multi-angle light scattering on LAP2α410-693........134
4.2.7 Analytical Ultracentrifugation on LAP2α410-693............................................................135
4.2.8 Transmission electron microscopy (TEM) of LAP2α410-693 and LAP2α1-693................136
4.2.9 Oxidation of LAP2α protein solutions.................................................................................138
4.2.10 Carboxymethylation of LAP2α410-693 with iodo-acetic acid..........................................140
4.2.11 Generation of a cysteine-to-serine mutant of LAP2α410-693...........................................143

4.3 Discussion............................................................................................................................143

5 CONCLUSIONS..................................................................................................................147

6 APPENDIX..............................................................................................................................150

6.1 Sequence alignments of A-band Ig and FnIII domains.........................................................150

6.2 Sequence alignments of B-box B1 and B2 subtypes............................................................153
6.3 Sequence alignment of human and mouse LAP2α ................................................................. 155

6.4 Calibration curve of Superdex 200 (16/60PG) ........................................................................ 156

6.5 Manuscript in preparation: Homo-oligomerization of LAP2α requires its C-terminal domain and is unaffected by a disease-causing mutation ................................................................. 156
List of Figures

Figure 1.1: Concept of hierarchical modularity in biology. ................................................................. 2
Figure 1.2: Homo-oligomeric protein-protein interactions are highly abundant................................. 5
Figure 1.3: Sub-classification of the Ig-fold according to connectivity and strand topology................. 9
Figure 1.4: EM micrograph of striated muscle and schematic representation of sarcomere............... 11
Figure 1.5: Domain structure of human cardiac titin and characterized titin ligands............................ 14
Figure 1.6: M-line titin, a complex, modular protein-protein interaction network............................... 18
Figure 1.7: Available high-resolution structures of titin domains....................................................... 21
Figure 1.8: Overview of the titin binding site for MuRF-1 and MuRF-1 domains.................................... 25
Figure 1.9: The M-line signal transduction module of titin.................................................................... 29
Figure 1.10: Overview of human TRIM family members and domains.................................................. 34
Figure 1.11: Structurally characterized components of the TRIM motif.............................................. 35
Figure 1.12: Components of the nuclear envelope and nuclear interior associated with lamins............. 39
Figure 1.13: Available structural information for lamin and associated............................................... 42
Figure 1.14: Localisation and domain organization of mammalian LAP2 isoforms............................ 43
Figure 2.1: Size exclusion chromatograms of titin A168-A170 and MuRF-1 constructs....................... 62
Figure 2.2: Refined crystals of His-tagged titin A168-A170 after macroseeding................................. 63
Figure 2.3: Diffraction pattern of crystals from C-terminally His-tagged titin A168-A170..................... 64
Figure 2.4: Crystals of untagged titin A168-A170................................................................................ 66
Figure 2.5: Diffraction pattern of titin A168-170 crystals..................................................................... 67
Figure 2.6: Graphical representation of the self-rotation function calculated using POLARFN......... 68
Figure 2.7: Real-space analysis of non-crystallographic symmetry using GETAX......................... 69
Figure 2.8: Sequence alignment of domain homologues used in molecular replacement.................... 70
Figure 2.9: Real (f’) and imaginary (f”) components of the anomalous signal from tantalum........... 71
Figure 2.10: Anomalous and isomorphous difference Patterson map at Harker section z=0.33........... 73
Figure 2.11: Projection of z-sections of electron density maps contoured at 1 σ................................. 75
Figure 2.12: Improvement of the quality of electron density maps during structure solution........... 76
Figure 2.13: Ramachandran diagram of the two NCS-related copies of titin A168-A170................... 77
Figure 2.14: Crystal structure of A168-A170....................................................................................... 79
Figure 2.15: Structure-based sequence alignment of titin Ig and FnIII of known structure.............. 80
Figure 2.16: Structural superimposition of Ig domains from titin....................................................... 81
Figure 2.17: Domain interfaces in the titin A168-A170 structure....................................................... 82
Figure 2.18: Ig-FnIII domain interfaces in the titin A-band............................................................... 84
Figure 2.19: Surface features of titin A168-A170 and model for MuRF-1 interaction......................... 85
Figure 2.20: MuRF-1 domain composition and constructs used in ITC experiments........................ 86
Figure 2.21: ITC data on MuRF-1:A168-A170 binding........................................................................ 87
Figure 3.1: Purification of MuRF-1 B-box wildtype and C210D mutant.............................................. 92
Figure 3.2: 1D-NMR spectra of soluble MuRF-1 B-box 2 wildtype and refolded C210D mutant........ 93
Figure 3.3: Refined conditions of MuRF-1 B-box crystals................................................................. 95
Figure 3.4: Diffraction pattern of MuRF-1 B-box crystals................................................................. 96
Figure 3.5: Real ($f'$) and the imaginary ($f''$) components of the anomalous scattering at zinc. .......... 98
Figure 3.6: Experimental anomalous difference Patterson maps at P6,22 Harker sections............... 101
Figure 3.7: Schematic description of the option warpNtrace in ARP/wARP........................................ 103
Figure 3.8: Electron density map after RESOLVE and final 2F$_{obs}$-F$_{calc}$ map.................................. 104
Figure 3.9: Ramachandran plot for three molecules of MuRF-1 B-box in the asymmetric unit......... 105
Figure 3.10: Crystallographic and NCS interactions in MuRF-1 B-box crystals.............................. 106
Figure 3.11: Refined condition of C$_{210}$D mutant of MuRF-1 B2 crystals........................................ 107
Figure 3.12: Structure of B2 and B1 boxes ......................................................................................... 110
Figure 3.13: The dimeric structure of MuRF-1 B2.............................................................................. 114
Figure 3.14: Comparative analysis of RING finger-like zinc-binding motifs...................................... 117
Figure 4.1: Schematic representation of LAP$_{\alpha}$.......................................................................... 128
Figure 4.2: Size exclusion chromatography and final purity of LAP$_{2\alpha}^{410-693}$.......................... 129
Figure 4.3: Size exclusion chromatography of full-length LAP$_{2\alpha}^{1-693}$. ...................................... 131
Figure 4.4: CD-spectra of LAP$_{2\alpha}^{410-693}$ and LAP$_{1\alpha}^{1-693}$. ......................................................... 133
Figure 4.5: Thermal denaturation of LAP$_{2\alpha}^{410-693}$ and LAP$_{1\alpha}^{1-693}$. ............................... 134
Figure 4.6: Analysis of LAP$_{2\alpha}^{410-693}$ with SEC-MALS.............................................................. 135
Figure 4.7: Molar masses of LAP$_{2\alpha}^{410-693}$ calculated by sedimentation equilibrium.............. 136
Figure 4.8: TEM images after low-angle rotary shadowing of LAP$_{2\alpha}^{410-693}$............................. 137
Figure 4.9: TEM images after low-angle rotary shadowing of full LAP$_{2\alpha}^{1-693}$. ........................ 138
Figure 4.10: LAP$_{2\alpha}^{410-693}$ protein solutions show a propensity to oxidate............................. 140
Figure 4.11: ESI-MS results after alkylation of LAP$_{2\alpha}^{410-693}$ by iodo acetic acid (IAA). .......... 142
Figure 6.1: Sequence alignment of titin Ig domains of A-band short super-repeat.......................... 150
Figure 6.2: Sequence alignment of titin Ig domains of A-band long super-repeat.......................... 150
Figure 6.3: Sequence alignment of titin FnIII domains of A-band short super-repeat.................... 151
Figure 6.4: Sequence alignment of titin FnIII domains of A-band short super-repeat.................... 152
Figure 6.5: Sequence alignment of B-boxes of subtype B2............................................................. 153
Figure 6.6: Sequence alignment of B-boxes of subtype B1............................................................. 154
Figure 6.7: Sequence alignment of human and mouse isoforms of LAP$_{2\alpha}$. ............................... 155
Figure 6.8: Calibration curve of size exclusion column superdex 200 (16/60 PG). ......................... 156
List of Tables

Table 1.1: MuRF proteins and their interaction partners (modified from Gregorio et al., 2005)........ 26
Table 2.1: Crystallization conditions for titin A168-A170.......................................................... 65
Table 2.2: Native data processing statistics. .................................................................................. 66
Table 2.3: Euler and Polar rotational NCS relations as determined using AMORE. ...................... 68
Table 2.4: MAD data processing statistics on [TaBr12]2+ derivatized titin A168-170 .................... 71
Table 2.5: Estimation of anomalous signal with XDS. .................................................................. 72
Table 2.6: Positions (in fractional coordinates) of heavy atom clusters as refined by SHARP......... 73
Table 2.7: Phasing statistics prior to density modification calculated using SHARP.................. 74
Table 2.8: Final Refinement statistics of titin A168-A170. ............................................................ 77
Table 2.9: Domain arrangement in poly-domain fragments from titin........................................... 81
Table 3.1: X-ray data of the high resolution set ......................................................................... 96
Table 3.2: Calculation of the Matthews coefficient. .................................................................. 97
Table 3.3: X-ray data of the MAD set. ....................................................................................... 98
Table 3.4: Estimation of anomalous signal with XDS. .............................................................. 99
Table 3.5: Real space positions of the six Zn-sites in the asymmetric unit................................. 100
Table 3.6: Phasing statistics prior to solvent flattening. ............................................................ 100
Table 3.7: Phasing statistics after solvent flattening with RESOLVE ........................................ 102
Table 3.8: Final Refinement statistics. ...................................................................................... 105
Table 3.9: Overview of data collection on B-box mutant C24D. ................................................. 108
Table 3.10: Refinement of MuRF-1 B-box C21D mutant. .......................................................... 108
Table 4.1: LAP2α constructs used in this study. ........................................................................ 127
Table 4.2: Ultracentrifugation experiments carried out on LAP2α410-693 .............................. 136
Table 4.3: ESI-MS results after carboxymethylation of LAP2α410-693 with iodo acetic acid (IAA). 141
Table 4.4: Mass mapping results for LAP2α410-693 after carboxymethylation......................... 142
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>AUC</td>
<td>Analytical ultracentrifugation</td>
</tr>
<tr>
<td>B1</td>
<td>B-box domain subtype I</td>
</tr>
<tr>
<td>B2</td>
<td>B-box domain subtype II</td>
</tr>
<tr>
<td>BAF</td>
<td>Barrier-to-autointegration factor</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic acid</td>
</tr>
<tr>
<td>β-ME</td>
<td>β-mercapto-ethanol</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CC</td>
<td>Coiled-coil</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CDR</td>
<td>Complementary determining region</td>
</tr>
<tr>
<td>CHAPS</td>
<td>3-[(Cholamidopropyl)dimethyl-ammonio]-2-propanesulfonate</td>
</tr>
<tr>
<td>DCM</td>
<td>Dilated cardio myopathy</td>
</tr>
<tr>
<td>DM</td>
<td>Density modification</td>
</tr>
<tr>
<td>DNA</td>
<td>Desoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EDMD</td>
<td>Emery Dreifuss muscular dystrophy</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>Ethyleneglycoltetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmatic reticulum</td>
</tr>
<tr>
<td>ESI-MS</td>
<td>Electron spray ionization mass spectroscopy</td>
</tr>
<tr>
<td>ESRF</td>
<td>European synchrotron radiation facility</td>
</tr>
<tr>
<td>Fn</td>
<td>Fibronectin</td>
</tr>
<tr>
<td>FOM</td>
<td>Figure of merit</td>
</tr>
<tr>
<td>GMEB-1</td>
<td>Glucocorticoid modulatory element binding protein-1</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-Hydroxyethyl)-piperazine-1-sulfonic acid</td>
</tr>
<tr>
<td>HMERF</td>
<td>Hereditary myopathy with early respiratory failure</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>IAA</td>
<td>Iodo acetic acid</td>
</tr>
<tr>
<td>IEG</td>
<td>Immediate early genes</td>
</tr>
<tr>
<td>IF</td>
<td>Intermediate filaments</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>INM</td>
<td>Inner nuclear membrane</td>
</tr>
<tr>
<td>ISOT-3</td>
<td>Isopeptidase T3</td>
</tr>
<tr>
<td>LEM</td>
<td>Lap Emerin Man</td>
</tr>
<tr>
<td>LMW</td>
<td>Low molecular weight standard</td>
</tr>
<tr>
<td>MAD</td>
<td>Multiple anomalous dispersion</td>
</tr>
<tr>
<td>MES</td>
<td>2-Morpholino-ethane-sulfonic acid</td>
</tr>
<tr>
<td>MFC</td>
<td>MuRF conserved region</td>
</tr>
<tr>
<td>MID-1</td>
<td>Midline-1</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-Morpholino-propanesulfonic acid</td>
</tr>
<tr>
<td>MPD</td>
<td>2-Methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>MuRF</td>
<td>Muscle specific RING finger</td>
</tr>
<tr>
<td>MyBP-C</td>
<td>Myosin binding protein</td>
</tr>
<tr>
<td>NCS</td>
<td>Non-crystallographic symmetry</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>Nuclear overhauser effect</td>
</tr>
<tr>
<td>NPC</td>
<td>Nuclear pore complexes</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>ONM</td>
<td>Outer nuclear membrane</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycole</td>
</tr>
<tr>
<td>PEVK</td>
<td>Proline Glutamate Valine Lysine</td>
</tr>
<tr>
<td>PHD</td>
<td>Plant homeo domain</td>
</tr>
<tr>
<td>PQS</td>
<td>Protein quaternary server</td>
</tr>
<tr>
<td>PVP</td>
<td>Polyvinylpyrrolidone</td>
</tr>
<tr>
<td>RBCC</td>
<td>RING B-box coiled-coil</td>
</tr>
<tr>
<td>RDC</td>
<td>Residual dipolar coupling</td>
</tr>
<tr>
<td>RFP</td>
<td>Ret finger protein</td>
</tr>
<tr>
<td>RING</td>
<td>Really interesting new gene</td>
</tr>
</tbody>
</table>
RNA  Ribonucleic acid
pRb  Retinoblastoma protein
SAXS  Small angle X-ray scattering
SDS-PAGE  Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC-MALS  Size exclusion chromatography with multi-angle light scattering
SLS  Swiss light source
SMT3b  Small ubiquitin related modifier type 3
SRF  Serum response factor
SUMO  Small ubiquitin related modifier type 3
TEM  Transmission electron microscopy
TEV  Tobacco etch virus
TK  Titin kinase
TMD  Skeletal muscular dystrophy
TRIM  Tripartite motif
TRIS  Tris-(hydroxymethyl)-aminomethane
UV  Ultra-violet
XNF-7  *Xenopus* nuclear factor-7
Y2H  yeast two hybrid screen
Acknowledgements

Research described in this thesis was carried out at the Biozentrum of the University of Basel. I want to particularly thank Prof. Olga Mayans for giving me the opportunity to work in her group, where I was able to explore my research interests and ideas in an excellent scientific environment. Her constant guidance, advice and support as well as her personal attention to the progress of the work created a motivating and constructive atmosphere throughout my entire PhD, which will also be inspiring for me beyond this thesis. I would like to thank Dietmar and Siegfried Labeit, both world-leading authorities in the titin field, for their result-oriented and effective collaboration. Many thanks also to Roland Foisner, our collaborator in the challenging LAP2α project. Here in Basel, especially Erik Hedbom, Zöhre Ucurum and Sebastian Meier for their contribution to the work on MuRF-1 B-box. Without their personal and scientific interest in the topic, these results would have never been possible. My sincere thanks go also to Heiko Herrklotz for his involvement in the ITC experiments. Also, Rene Hemmig and Markus Kroemer from Novartis and Laurent Kreplak, Ariel Lustig, Bohumil Maco, Manuela Schätzle, Kitaru Suda from the Biozentrum made important contributions at various stages of the work and helped me with numerous, scientific discussions. I want to thank Prof. Ulrich Baumann for kindly accepting to coreferee this thesis. Thanks a lot of course to my office, lab, train, mensa and Freiburg mates Rainer and Marco for giving me company and support all the way through. Thank you Darko, Elli and Pilar for the time inside and outside the Biozentrum. My friends Tilman, Vale, Christoph, Eckart, Christina and Bernd have also personally contributed to the success of this work in a non-scientific way. Thank you Mireia for all your understanding, patience and love. You were always there despite our spatial distance. Last but not least I would like to dedicate this thesis to my parents who always believed in me and supported me wherever they could.
dedicated to my parents

für meine Eltern
Chapter 1

Introduction
1 Introduction

1.1 Protein-protein interactions

1.1.1 Network modules

The cell as the basic unit of life is able to move, communicate, metabolise, grow, reproduce and dynamically remodel its shape in response to external and developmental stimuli. For these processes to occur, a precise spatiotemporal order is absolutely essential. This is reflected by the concept of hierarchical organization of the cellular machinery from atoms to macromolecular assemblies to organelles and beyond to tissues, organs and organisms. A constant interplay across all organization levels extends from events that happen on the timescale of the whole cell, e.g. cell motility, reproduction and development to the most rapid events operating at the timescale of macromolecules, e.g. protein synthesis or catalysis. As a result, life is inevitably orchestrated by chemical reactions between all cellular components, which thus form metabolic, protein-protein or protein-DNA networks.

Current post-genomic projects aim at unravelling the relationships between cellular protein components (Alloy and Russell, 2006; Joyce and Palsson, 2006). The study of protein-protein interactions in pathways, complexes or even complete organisms has become a new paradigm in protein biology (Butland et al., 2005; Stelzl et al., 2005; Uetz, P. et al., 2000). These studies revealed that most cellular proteins are tightly embedded into biological networks and function often within the context of multimeric or supramolecular assemblies. Prominent, due to their stability well characterized examples are the spliceosome (Neubauer et al., 1998), the proteasome (Verma et al., 2000) or the nuclear pore complex (Rout et al., 2000). Other examples include more transient interactions between proteins and protein complexes like tyrosine kinase signalling cascades that are composed of scaffold protein complexes and various regulators and effectors (Csiszar 2006).

According to the network modularity principle introduced by Hartwell et al. (Hartwell et al., 1999) these molecular assemblies contain so-called “functional modules” as critical level of their biological organization. These modules are discrete entities whose functions are separable from other modules. The underlying molecular components (small molecules, protein, RNA, DNA) and their interactions collectively
contribute to the function of the module (Figure 1.1). Nevertheless, the function of the module cannot be predicted from the isolated underlying components. By computational analysis it has become clear that most relevant functional modules in biological networks are found in the meso-scale format, that is, they consist of 5-25 genes/proteins per module (Spirin et al., 2003).

**Figure 1.1: Concept of hierarchical modularity in biology.**

a) Modularity on the protein level: Uridylate kinase (1ukz; Mueller-Diekckman & Schulz, 1995) contains a single NTP-hydrolase domain. In the multi-domain transcription factor protein EF-TU (1exm; Hilgenfeld et al., 2000) a domain from the same superfamily is present (in red).

b) Modularity at the cellular level as exemplified by three distinct types of functional modules:
   
i) a metabolic pathway - the mevalonate pathway.
   
ii) a signalling pathway - mating response MAPK pathway in yeast (Schwartz and Madhani, 2004).
   
iii) a protein complex - ATP synthase (1e79; Gibbons et al., 2000) Six chains (red and orange) that all contain an NTP-hydrolase domain are assembled into a ring together with additional domains (Figure modified from: Pereira-Leal et al., 2004).
Two types of modules have been classified: i.) stable protein complexes such as the splicing machinery, protein export-, protein transport-complexes, transcription factors; and ii.) dynamic functional units like signalling and metabolic cascades, cell cycle regulation modules (Spirin and Mirny, 2003). The latter ones do not require physical contact of all components at all times or physical contact at all. Rather they are connected by a succession of interactions in the case of an information-processing signalling cascade or by the product of one enzyme being the substrate of the next in a metabolic cascade (Pereia-Leal et al., 2004). By definition, the module is a discrete entity whose function is distinct from other modules. The underlying isolation principle can either be in the case of complexes spatial location or in the case of dynamic functional modules their chemical specificity and temporal existence.

The modular organization of protein-protein interaction networks facilitates the combinatorial generation of complexity as well as functional diversity, both of which are important factors in the context of evolution biology (Kirschner and Gerhart, 1998). Evolution re-wires modules instead of altering individual modules, following the concept of hierarchical modularity applicable also to many other scientific disciplines including computer science and organization theory (Cohen and Tong, 2001). Also in protein structure, the concept of modularity is also an established principle (Figure 1.1). Proteins are formed by autonomously folding units called domains that display strong connections within, but weaker connections outside the domain, a defining feature that is also shared with modules in cellular networks (Pereia-Leal et al., 2004).

1.1.2 Protein complexes

Protein complexes act as well-defined functional modules in cellular networks. They are formed by cohesive and strong interactions and often can be reconstituted in a functional form independently of the rest of the protein interactions network. Two types of complexes can be distinguished: i) Obligate protein complexes, in which the individual protomers in isolation do not form stable structures; ii) non-obligate complexes, where the protomers exist and function also as individual units independently of complex formation (Nooren and Thornton, 2003)

The stability of the complex depends strongly on the physiological environment and cannot be predicted a priori. Nevertheless, the transient or stable nature of the
complex can often be inferred from the biological context, i.e. temporary interaction in signalling (weak linkage) or stable constitutive interaction in structural function (strong linkage). An advantage of protein-protein interactions via weak linkage is an increased “evolvability” of the system (Kirschner and Gerhart, 1998). Together with compartmentation and redundancy it leads to a reduction of the interdependence of each components and facilitates the accommodation to novelty through accumulation of non-lethal mutations. This principle is reflected by the fact that proteins interacting with many other proteins in strong, structural interactions like for example active sites in enzymes, histones, actin and tubulin monomers have changed little during evolution.

Within a complex, protein-protein interactions can be established between identical or non-identical chains, thus forming homo- or hetero-oligomeric arrangements. Frequently observed motifs of homo-oligomeric interactions are coiled-coil motifs (Burkhard et al., 2001). In an isologous association both partners interact via the same surface on both monomers, whereas in an heterologous association assemblies use different interfaces, which potentially can lead to higher order aggregation. This is exemplified by the pathogenetic amyloid formation leading to Alzheimers disease (Sunde et al., 1997).

Homo-oligomeric complexes are highly abundant (Figure 1.2). For example calculations utilizing the typical macromolecular composition of the E. coli cytosol revealed that the average oligomeric state of soluble proteins is four, with 15% forming higher order assemblies (Goodsell et al., 1991). Analysis of the Protein Quaternary Structure (PQS) database (Henrick and Thornton, 1998) has revealed that over 70 % in a non-redundant set of protein complexes contained interactions between identical proteins (Pereia-Leal et al., 2004), i.e. homo-oligomerisation contributes to a large extent to protein complex formation.

Given its frequent occurrence, homo- and hetero-complex formation must stand under substantial evolutionary pressure and have a number of structural and functional advantages.
Introduction: Protein-protein interactions

Figure 1.2: Homo-oligomeric protein-protein interactions are highly abundant.

Shown is the result of an analysis of the protein quaternary structure database (PGS) (Henrick and Thornton 1998). Structures on the top represent examples of the complex type: homotrimer of carbonic anhydrase (1qr; Iverson et al., 2000), 2-oxoisovalerate dehydrogenase heterotetramer (1ps0; Valencia et al., 2004) and hetero-dimer of the cap-binding protein (1n52; Calero et al., 2002). Interactions are show as red (homomeric) and green (heteromeric) lines between the nodes of a two-dimensional representation of each complex (Figure taken from: Pereira-Leal et al., 2004).

1.1.3 Advantages of complex formation

Protein complex formation leads to synergistic effects. Clustering of molecules in signalling complexes for example ensures a rapid information exchange and eliminates delays that would occur as a result of cytosolic crowding and diffusion limitation. Complex formation can also provide a unique, local physicochemical environment that leads to an increase in specificity, affinity and, at times, potentiated activity through cooperativity; a prominent example are the ubiquitin ligases that frequently self-assembly into higher order structural assemblies with compartment character (Kentsis et al., 2002; Reymond et al., 2001; see chapter 1.3). Furthermore DNA binding proteins are frequently components of large functional modules and bind DNA as homo-dimers like the type II restriction enzyme (Pingoud and Jeltsch, 2001) or transcription factors like the glucocorticoid response element (Cairns et al., 1991).

In addition, formation of higher-order assemblies provides an increase in mechanical and chemical stability, an advantage most clearly envisioned in the homo-oligomeric filamentous assemblies of the cytoskeleton. Intermediate filaments (IFs), for
example, are resistant against treatments with high concentrations of urea or detergents and constitute stress-bearing fibres. Yet cytoskeletal filaments are dynamic structures that can be assembled and disassembled rapidly. In this case, the process of self-assembly can reduce the genome size and prevents problems emerging from the folding of large single-chain proteins (Marianayagam et al., 2004).

1.1.4 Scaffolding complexes enforce proximity

One method to influence components, subcellular localization and activity of protein ensembles involves scaffold proteins. These proteins influence protein-protein interactions without modifying their targets and can recruit two or more molecules simultaneously into a global complex (Vondriska et al., 2004), a concept known as “forced proximity” (Ferrell and Cimprich, 2003). Accordingly, scaffold proteins have been described as signal transduction “catalysts”: they mediate the pre-localization of all participants in a pathway, so that immediate and rapid access to the up and downstream substrates in the cascade can be achieved. The regulation of scaffolders is crucial for the coordinated function of the cell. Hence, scaffold proteins can themselves be subjected to post-translational modification. Additional regulatory mechanisms are their oligomerization state, their ability to undergo conformational changes and their susceptibility to irreversible protease cleavage or degradation by ubiquitin pathways (Csizar, 2006).

1.1.5 Functions of Scaffold proteins

Scaffold proteins constitute structural formations in distinct cellular locations with a number of specific advantages. Through combinatorial control, components of a complex might be displaced or added, deactivated or activated depending on competitive/non-competitive interactions with other elements of the complex (Burack et al., 2002). Conformational changes, alterations in the interaction surface or post-translational modifications trigger these processes. A varying combination of a discrete set of molecules also increases the number of distinct signalling modules for distinct phenotypes.

Scaffolding can also increase the specificity of the processes by preventing interactions with components of other pathways, hence conferring insulation to the
signalling module. Depending on sub-cellular location with its local environment and together with the sequential assembly process on the scaffold, different kinetic activation thresholds can regulate signal propagation downstream of the signalling complex. In contrast to diffusion controlled signalling events, tethering of transducers to protein scaffolds can lead to attenuation rather than amplification of responses, a fact that might help the cell to prevent undesirable amplification of signals in the absence of appropriate stimuli. On the other hand, scaffolding can clearly potentiate activity by increasing the local concentration of the effector molecule and the close proximity of the signal processing machinery (Burack et al., 2002). Taken together, anchoring, targeting and isolation are the major regulatory features of scaffold proteins.

To fulfil these functions, the “scaffolder” needs to provide a rigid folding unit that is capable of minimizing the conformational entropy loss of both interaction partners during the complexation event. Thus, a well-defined protein core that substantially contributes to the energy of domain folding is critical (Skerra, 2000). A structurally well-characterized example involved in protein modification can be found in the ubiquitin ligases of cullin type (Zheng et al., 2002). Scaffolding building blocks include for example phospho-tyrosine binding (PTB), Src homology (SH2, SH3) or pleckstrin homology (PH) domains in the case of tyrosine kinase signalling complexes (Csizar, 2006)) or the zinc binding RING finger domain (Borden, 2000; chapter 1.2 and 1.3).

Scaffolding subunits can be grouped according to their secondary structure content into α-helical proteins, into small structures with few secondary structure elements and irregular α/β architecture and finally into predominantly β-sheet containing scaffolding motifs (Hosse et al., 2006. Among the all-β-motifs, clearly the Ig as well as the related fibronectin type III (FnIII) domain are highly abundant protein-protein scaffolding platforms (Holt et al., 2003; Koide et al., 1998).

1.1.6 The Immunoglobulin superfamily (IgSF)

The Immunoglobulin (Ig) fold is probably the most widespread architectural motif in multidomain proteins, among them matrix proteins, receptors, chaperones and enzymes. This is reflected in the wealth of functional, structural and sequence diversity among the Ig domains (Barclay, 2003). Nevertheless, a common function of all Ig–like domains is their involvement in diverse binding interactions. To date, no single Ig
domain has been reported to contain enzymatic activity (Bork et al., 1994). Structurally, the Ig domain consists of seven to nine anti-parallel β-strands forming a β-sandwich with a Greek key.

The available structural information lead to a topological sub-classification of the Ig fold into four distinct sub-classes based on number of the strands and their connectivity (Figure 1.3a). Most conserved are the strands termed B, C, E, F, which define a two-by-two structural core. Nevertheless, the edges of the sheets are conformationally flexible (strands A, G, C’, C’’) and the location of the strands C’/D defines the four subtypes. The C-type (Constant) represents the classical seven-stranded topology, whereas the S-type (Switched) also contains seven strands, with strand D and C’ switched. The H-type (Hybrid) of Igs can be envisioned as hybrid between C- and S-type with eight strands and finally the nine-stranded V-type (Variable) is observed in the variable domain of immunoglobulins (Bork et al., 1994). Two additional topologies are realized in the Ig domains of human lamin A/C (Dhe-Paganon et al., 2002; Krimm et al., 2002) and the Ig domain of Telokin (Harpaz and Chotia, 1994), which both form deviations of the canonical topologies through the presence of extra strands at the edges of the sheets (Figure 1.3b).

Binding of Ig ligands is not restricted to the loop regions like in the complementary determining region (CDR) of the immunoglobulins variable domain, but can also be mediated by interactions within the sheets. Often consecutive regions in more than one domain constitute the binding interface (Huber et al., 1994).

The fibronectin type III fold (FnIII) found in numerous extracellular matrix (ECM) proteins resembles the S-type Ig fold (Figure 1.3b). Fibronectins’s function depends on its ability to bind ECM components including adhesion molecules, cytokine receptors, collagens and integrins. Its extended domain arrangement of independently folded units leads to a “beads on a string” appearance in electron micrographs (Erickson et al., 1981). Three different kinds of Fn modules are realized in the subtypes I, II and III (Ruoslhati 1988), which are all structurally characterized (Type I: Williams et al, 1994; Type II: Constantine et al., 1992; Type III: Leahy et al., 1992). The Type III module is the biggest module and characterized by a consensus sequence of approximately 90
Introduction: Protein-protein interactions

Figure 1.3: Sub-classification of the Ig-fold according to connectivity and strand topology.

a) The two sheets of the domains are displayed as if they are in an open book lying face down. Closing of the book can reproduce the fold. Four distinct subtypes are classified. β-strands A, B, C, E, F and G are common to all Ig-like domains (Bork et al., 1994). The strand connectivity and presence of accessory strands (grey) distinguishes between Ig domain subtypes. The fibronectin type III (FnIII) fold possesses similar topology like the S-type Ig. b) Two different types of Ig domains relevant for the work in hand are the lamin (L-type; Dhe-Phaganon et al., 2002) and the titin subtype (I-type; Harpaz & Chotia, 1994), both characterized by extra strands. The L-type represents a variation from the S-type Ig fold, whereas the I-type represents a variation from the V-type Ig. The 3D structure of Ig 11 from titin I-band (Mayans et al., 2001) and of FnIII A-71 (Muhle-Goll et al., 1996) is included for illustration. In the lamin fold the extra strands are not named.

residues (Patthy, 1991) and contains unique features that are necessary for cell adhesion activity. These features include the Arg-Gly-Asp (RGD) loop between strand F and G and an additional region on an adjacent FNIII modules termed synergy region. Compared to the Igs it shows a distinct hydrophobic core and only low sequence homology (Main et al., 1992).
1.2 Titin’s M-line: a structural scaffold and signal transduction “module”

1.2.1 Structure of the sarcomere

The sarcomere or contractile unit of striated vertebrate muscle is considered to be the most highly ordered and functionally coordinated supra-molecular structure known to date. Its remarkable paracrystalline arrangement of myofilaments is directed to generate force in a rapid and directed way.

The striated pattern that is visible in the light microscope lead to a division of the 2-3 µm long sarcomere into several zones (Figure 1.4a). It is bordered at each end by a dark, narrow line known as the Z-disc, that bisects a lighter region, the so-called I-band (isotropic in polarized light), which is shared by two adjacent sarcomeres. At the center of the sarcomere, a dark zone called the A-band (anisotropic in polarized light) is again divided by a less dense region called the H-zone. Within the H-zone, a narrow band of higher density called the M-line defines the geometrical middle of the sarcomere. The M-line also contains a fine-structure of up to five dense and less-dense lines that varies in number among different muscle types (review Squire et al., 2005).

The striation pattern results from the precisely ordered arrangement of the two contractile filaments actin and myosin as deduced from ultrathin sections of muscle investigated by electron microscopy (Hanson and Huxley, 1953). These two cytoskeletal filaments are the basic active-force generating proteins forming the thin (actin-based) filament of the I-band and the thick (myosin-based) filament of the A-band. Capping and cross-linking proteins are necessary for the correct polymerization of both filaments and construction of a registered, three dimensional lattice. Intermediate filaments at the Z and M-line reinforce sarcomere structure through the sarcolemma providing a scaffold of force transduction, maintain sarcomeric registry and connect adjacent sarcomeres.
**Introduction: The M-line of Titin**

**Figure 1.4: EM micrograph of striated muscle and schematic representation of sarcomere.**

**a)** Individual myofibrils are aligned and consist of alternating dark and light stripes. The dark bands are called A bands (labelled A), while the light regions are termed I bands (labelled I). The A bands are themselves striped, consisting of a lighter central band called the H zone (labelled H) flanked by two darker regions. At the centre of the H zone is a thin dark line, the M line (labelled M). A similar line, the Z line (labelled Z) is visible in the centre of the I band. The sarcomere is defined as the region between two Z lines (picture taken from Engel & Franzini-Armstrong, Myology, 3rd edition).

**b)** Principle components of the sarcomere are the I-band, the A-band, Z-disc and M-line. Four distinct filament systems (titin, nebulin, actin and myosin) contribute to sarcomere integrity. At the Z-disc the N-terminus of titin is bound to telethonin (T-cap). Each of the four filaments has a distinct domain composition. Titin contains several unique regions among them an unstructured region (PEVK) and a Ser/Thr kinase domain. In the A-band a super-repeat pattern of alternating immunoglobulin and fibronectin type III domains is believed to mediate interaction myosin binding protein C (labelled C-protein). (Figure taken from Gregorio et al., 1999).
1.2.2 Overall structure of the third filament of the sarcomere: titin

The giant protein titin (formerly known as connectin; (Maruyama, 1976) is the third muscle specific filament system and has essential structural, contractile and regulatory functions in the sarcomere. It spans half of the sarcomere from the Z-disc to the M-line (Figure 1.4b).

The titin gene contains 363 exons coding for 38138 residues (4200 kDa) and multiple splice isoforms have been identified. The expressed protein exists as a single polypeptide chain with an isoform-dependent molecular composition between 27000 and 33000 residues corresponding to a molecular weight of 2970 kDa in the cardiac and 3700 kDa in the soleus skeletal muscle isoform (Bang et al., 2001). With its length of over 1 µm it is the largest protein known to date (Swiss-Prot/TrEMBL accession numbers Q10465, Q10466, and Q8WZ42, Labeit and Kolmerer, 1995).

Its N-terminus resides in the Z-disc interacting with a protein called T-cap or telethonin. Telethonin gets phosphorylated by the kinase domain of titin (Mayans et al., 1998) and is essential for sarcomere formation in the differentiated myocyte (Gregorio et al., 1998). After the Z-line titin runs parallel to the actin filament the I-band region. This region of titin contains tissue-specific isoforms with distinct mechanical properties as well as intrinsically unstructured regions functioning as entropic springs (Labeit and Kolmerer, 1995; Freiburg et al., 2000).

In the A-band region titin binds to meromyosin and MyBP-C, serving as a template or “blueprint” during fibrillogenesis (Trinick, 1994). Due to this strong interactions the A-band of titin is stiff under physiological conditions. By contrast, the I-band region is an elastic spring that creates the passive and restoring force during sarcomere stretch and compression, key aspects of mechanical behaviour of the myofibrill (Linke et al., 1994).

Titin’s C-terminus overlaps with an adjacent titin molecule at the M-line creating a continuous system throughout the sarcomere. Through embedding into transversal cytoskeletal networks at the M-line, it ensures that equal forces are developed in the two halves of the A band (Agarkova et al., 2003).

Several titin-related molecules have also been found in invertebrate muscles (Bullard et al., 2002). The conservation of Fn-Ig modules and the kinase domain in twitchin and two insect projectins suggest an important function for this part of the molecule. Stretchin-MLCK (stretchin myosin light chain kinase) is a protein present in
Introduction: The M-line of Titin

*Drosophila*, which is predicted to have a tandem Ig composition as well as unique proline, glutamate valine and lysine rich sequences termed PEVK and a C-terminal kinase domain. By contrast, the tandem-Ig containing protein kettin, which is found in insects, *C.elegans* and crayfish, contains no FnIII modules and no kinase domain. It shows high similarity to the titin I-band region.

1.2.3 Domain composition and interactions of the giant scaffold titin

Similar to cell adhesion and extracellular matrix proteins (fibronectin, L1-CAM, N-CAM; Gerrow and El-Husseini, 2006) titin consists of a modular array of up to 166 copies of immunoglobulin (Ig) and 132 copies of fibronectin type III (FnIII) domains in soleus skeletal muscle (Figure 1.5). As a result, it appears as a string-of-beads in electron micrographs (Trinick et al., 1984).

The filament utilizes the nature of the Ig fold as a protein-protein interaction domain to interact with a majority of sarcomeric proteins, other filaments and cytosolic signalling molecules of highly diverse functions and cellular location (Lange et al., 2006). Currently, the amount of titin-associated proteins identified by yeast two hybrid techniques (Piehler 2005) is constantly growing (Miller et al., 2004; Lange et al., 2006).

Structural interactions that contribute to sarcomere integrity are found in the Z-disc, A-band and M-line. Proteins with proposed signalling or regulatory function cluster at three “signalling hot spots” with extensive cross-talk between them: the Z-disc, the beginning of the I-Band (N2-PEVK region; next chapter) and the M-line (Granzier and Labeit, 2004).

1.2.4 I-Band titin

The I-band region of titin contains tandem Ig domains and unique sequences. The Ig domains are arranged in two sections, proximal and distal I-band to the Z-disk which are intervened by a unique PEVK region of variable length. Passive tension and sarcomeric slack length are largely determined by I-band titin through tissue-specific and developmental expression of proximal Ig and PEVK splice isoforms (Bang et al., 2001). Cardiac spliceoforms include the stiff N2B (short spring) isoform and the more compliant N2BA isoform (long spring). A variation of I-band molecular mass of 0.7
Introduction: The M-line of Titin

MDa in cardiac and 1.5 MDa in soleus muscle was observed (Labeit and Kolmerer, 1995). Accordingly, Wu et al. have shown that upon exposure to elevated heart pacing levels leading to pace-induced cardiac failure, the sarcomere changes its titin composition on the timescale of weeks in favour of the more stiffer N2B isoform (Wu et al., 2002). Hence via alternative splicing of titin and together with likewise isoform-dependent calcium and actin binding properties of titin the sarcomere can rapidly adjust its mechanical properties and modulate a broad range of elasticity (Granzier and Labeit, 2004). Furthermore, titin phosphorylation in the cardiac N2B and N2BA region was found to reduce passive force. The reduction of passive force was found independent of the association between titin and the thin filament, as thin filament-extracted preparations showed comparable, solely titin-based decrease in stiffness.

Figure 1.5: Domain structure of human cardiac titin and characterized titin ligands.

Titin consists to > 90% of its mass of tandem Ig (dark grey) and FnIII (light grey) modules. 19 unique sequences including a PEVK (blue) and kinase (orange) region have been identified. Shown are cardiac N2B (yellow) and N2A (red) spliceforms. Various titin ligands have been identified by Y2H screens. They cluster in three distinct positions: i) Z-disc, ii) I-band N2A region and iii) M-line. Shown are T-cap (red), α-actinin (orange), DRAL/FHL2 (light blue), cariac ankyrin protein (CARP; yellow), myopalladin (blue), p94 (calpain, green ), Muscle specific RING-finger protein (MURF; brown) and myomesin (pink). The A-band shows a super-repeat pattern of alternating Ig and FnIII domains (Figure taken from: Miller et al., 2004).

A potential alteration of native structures within the N2B region resulting in slack length gain was proposed, but the exact mechanism is currently unknown.

Despite its mechanical involvement in muscle passive tension, both I-band regions N2B and N2A have been shown to constitute targets for signalling molecules. N2B was found to interact with the zinc-binding LIM family member DRAL/FHL-2,
which in turn binds to the metabolic enzymes creatinine-, adenylate- and phosphofructo-kinase (Lange et al., 2002). Hence a function in compartmentalisation of metabolic functions was proposed for the central I-band region to assure the ATP supply needed for muscle contraction. For the N2A element present in both cardiac and skeletal muscle several ligands have been identified. Two of its Ig domains (Ig82/83) interact with the calpain protease p94 (Ono et al., 2004). Its binding to titin is thought to regulate p94 protease activity.

Additional ligands include the two homologous ankyrin repeat proteins cardiac ankyrin repeat protein (CARP) and diabetes ankyrin repeat protein (DARP). CARP in turn interacts with myopalladin, an actin associated scaffold found also in the nucleus (Otey et al., 2005). All ankyrin repeat protein were found up-regulated after cardiac injury, muscle denervation or during recovery after metabolic efforts suggesting their involvement in stress response pathways (Kuo et al. 1999; Kemp et al., 2000; Ikeda et al., 2003). This is further corroborated by their presence in the nucleus and interaction with transcriptional regulators (Miller et al., 2003).

1.2.5 A-band titin

The A-band section of titin is an integral part of the myosin filament. The latter can be subdivided based on visual appearance in electron micrographs into three regions termed D-, C- and P-zone (Sjöström and Squire, 1977).

Strikingly, titin FnIII domains are exclusively found in the A-band region forming two defined super-repeat patterns. A seven domain superrepeat (Ig-Fn-Fn-Ig-Fn-Fn) is found at the beginning (D-zone), which is repeated six times and a 11 domain super-repeat (Ig-Fn-Fn-Ig-Fn-Fn-Ig-Fn-Fn-Fn) repeated 11 times in the region of the A-band (Figure 1.5). Near the M-line, the unique Ig-Ig-Fn-Fn-Ig-Ig-Fn motif directly precedes the kinase domain (P-zone). Corresponding positions of Igs and FNIIIs in the super-repeats were found to have substantially higher sequence homology (Amodeo et al., 2001). With a length of 4-4.5 nm per domain this super-repeat pattern is roughly 43 nm long, a distance found also in the intrinsic helical repeat of the myosin filament with a pitch of 42.9 nm made up of three 14.3 nm-spaced crowns (Squire et al., 1998). Freiburg & Gautel have shown in vitro via dot blot assays that recombinantly expressed fragments of titin’s C-zone, in particular the first Ig domain of the long superrepeat, are involved in binding to the last three C-terminal domains of myosin-binding protein C
Introduction: The M-line of Titin

(MyBp-C) (Freiburg and Gautel, 1996). These are currently the only data on a direct titin interaction with MyBp-C, suggesting a direct correlation between super-repeat pattern and MyBp-C binding.

Labeit et al. further suggested an interaction of the FnIII domains of the titin A-band with myosin (Labeit et al., 1992). Muhle-Goll demonstrated via co-sedimentation and solid state binding assays that the poly-domain constructs A77-A78, A80-A82 and A84-A86 from the titin A-band are able to bind to myosin, light meromyosin portion and subfragment S1 (Muhle-Goll et al., 2001). This stands in contrast to previous electron microscopic studies, in which binding of titin only to the light meromyosin region of myosin was detected (Houmeida et al., 1995). Further analysis is needed to elucidate the interaction between the two filaments.

In summary, it is speculated that the titin A-band performs a template function regulating the thick filament assembly (Okagaki et al., 1993; Freiburg et al., 1996). Moreover, it is thought to keep thick filaments in register in the relaxed and activated state of the sarcomere and creates restoring forces to unequal contraction through its integration into the M-band lattice (Horowits and Podolski, 1987).

1.2.6 M-line titin

Towards the C-terminus of titin, a unique region of titin harbours a catalytic Ser/Thr kinase domain between the A-band and the M-line (TK; Labeit et al., 1992; Mayans et al., 1998). The arrangement and sequence of kinase and adjacent Ig/FnIII domains is homologous to other so-called giant kinases like Myosin-Light-Chain Kinase (MyLCK) and the invertebrate titin homologues projectin and twitchin (Bullard et al., 2002).

TK consists of a catalytic and a regulatory domain. From the crystal structure (Mayans et al., 1998; Figure 1.7) it became clear that the active site is blocked by the regulatory domain and gets activated upon phosphorylation by a to date unknown kinase and through Ca\(^{2+}\)/Calmodulin binding. As a substrate of TK, telethonin (T-cap) was identified (Mayans et al., 1998; Gregorio 1998; Mues et al., 1998), which is located remotely at the N-terminus of titin in the differentiated sarcomere. Hence, the authors suggested a role of TK in the early stages of myofibrillogenesis. The exact function of TK in existing or yet unknown myofibrillar signalling pathways remains to be determined. Interestingly, TK has been detected as a critical component of a signalling
module at the M-line potentially involved in a mechanotransduction pathway (Lange et al., 2005; chapter 1.2.14).

Following the kinase in C-terminal direction a set of ten Ig modules (M1-M10) extends towards the very C-terminus of titin. These Ig domains are interconnected by unique insertions of various lengths and unknown function (Obermann et al., 1997).

Titin is embedded into the M-line matrix consisting of the poly-Ig/poly-FnIII proteins myosin, myomesin and M-protein (Luther and Squire, 1978). They are involved in transversal and longitudinal organization and connection of overlapping filament ends providing a continuous filament that is necessary for rapid and efficient contraction of the sarcomere (Fürst et al., 1999).

The carboxy-terminal Ig M4 of titin interacts with a region of Ig/FnIII repeat protein myomesin (domains My4-My6) in a phosphorylation dependent manner. This was proven using recombinantly expressed, purified myomesin poly-domain constructs that were tested in a solid state overlay assay against individual titin M-line Igs (M1-M10). A minimum of three Fn domains were necessary for binding suggesting a cooperative binding through a large number of weak interactions (Obermann et al., 1997). A similar observation was found for the interaction of A-band titin with MyBP-C (Freiburg and Gautel, 1996). The interaction cross-links titin with the myosin filaments and has implications during sarcomere formation, regeneration and turnover. It was suggested that M-band proteins are responsible for lateral alignment of the thick filament, while titin provides the centring of the A-band during relaxation (Agarkova et al., 2003). Furthermore, recent studies by Weinert et al. show that complete knockout of titin’s M-line (A169-M7) through a germline recombination approach does not influence sarcomere assembly, but severely interferes with sarcomere strength and lateral growth, which eventually leads to sarcomere disassembly. As a result of the knockout titin’s C-terminal region was found not integrated into the A-band structure and early embryonic lethality was observed (Weinert et al., 2006). These experiments clearly further corroborate the role of M-line titin in the structural integrity of the sarcomere.

Additional to the mentioned structural proteins a set of soluble metabolic, signalling and regulatory factors was found to target the M-line components and titin (Figure 1.6). Often they are not exclusively present at the M-line, but can also be found at the Z-disc, the nucleus or in a cytoplasmatic pool (Lange et al. 2006). Among them
are enzymes involved in energy metabolism like muscle creatinin kinase (M-CK, Stolz et al., 1998) and LIM domain proteins DRAL/FHL-2, which were found in various locations of the sarcomere including I-band, M-band and nucleus (McLaughlin et al., 2002).

Recently, also members of the muscle specific Ring finger family of proteins (MuRF-1 and MuRF-2) have been identified as M-line components (Centner et al., 2001) and moreover act as dynamic shuttles between M-line, Z-disc and the nucleus (McElhinny et al., 2002). The muscle specific cysteine protease calpain (p94) was found to bind to titin’s M-line Ig domain M9 and the adjacent unique insertion Is-7 (Sorimachi et al., 1995). Since the C-termini of two titin molecules overlap at the M-line, the binding sites for calpain and MuRF-1 could come close in space and a functional interaction between the E3-ligase MuRF-1 and calpain (p94) was speculated. Potentially MuRF-1 is a substrate for the proteolytic activity of p94 (Gregorio et al., 1999).
1.2.7 Overview of existing structures of titin

Extensive structural studies have been carried out on the titin filament using diverse microscopy techniques (Kellermeyer and Grama, 2002; Squire 1997). In contrast to the wealth of microscopic data, the analysis of the filamentous molecule titin by methods providing higher resolution like e.g. NMR-spectroscopy or X-ray crystallography is severely hampered by its large size, structural complexity and elastic nature. Hence, high-resolution structural studies are restricted only to small polydomain segments or single individual molecular building blocks due to experimental constraints (chapter 1.5). In principle, the overall organization of the filament can be reconstituted from its building blocks in a “divide and conquer approach” using complementary information provided for example by small angular X-ray scattering ((SAXS); Vazina et al., 2006; Marino et al., 2005; Improta et al., 1998), electron microscopy or by in silico modelling approaches (Amodeo et al., 2001; Muhle-Goll, 2001). The latter approach needs a reliable ensemble of representative structures of high sequence similarity comprising both individual building blocks as well as transitions between filament subunits that are necessary to exactly model long-range effects on a larger scale. In the case of titin with only a handful high-resolution structures of individual domains this prerequisite does not exist yet. Figure 1.7 summarizes the currently available structural data.

The molecular structures of several individual domains from titin have been determined by NMR spectroscopy and X-ray crystallography (for references see Figure 1.7b). These structures are representatives for different regions of titin, i.e. the Z-disc (Z1Z2), the proximal (I1) and distal I-band (I91), the A-Band (A71), titin kinase and the M-line (M1, 5). Currently, the study on the very N-terminus of titin, the Z1Z1 Ig tandem, remains the only poly-domain high-resolution structure of titin (Zou et al., 2006; Marino et al., 2005).

The titin Ig domains all posses a characteristic β-sandwich-fold of the intermediate set of the immunoglobulin superfamily (Harpaz and Clothia, 1994). The only fibronectin domain structure (Muhle-Goll et al., 1997) shows high homology to the fibronectin archetype (Leahy et al., Science 1992).

Homology modelling identified surface features that are distinct for the titin regions. Z-disc Ig have the lowest polarity among all titin Igs, most of which are
Introduction: The M-line of Titin

generally more polar than charged. A-band Ig modules have larger non-polar surfaces potentially resulting from filament interactions.

The analysis of structural data by various authors has identified characteristic features of representatives of constitutively and differentially expressed I-band Ig domains (Fraternali et al., 1999; Witt et al. 1998; Marino et al., 2005).

The first class comprises the constitutively expressed Ig-domains with shortened BC and FG elements and a minimal number of prolines at the N-terminal loop cluster. As a result, the shape of the molecule is slightly less elongated and the expected inter-domain flexibility is higher.

An N-terminal loop cluster of the differentially expressed group is strictly conserved, while these region appears more variable in the constitutionally expressed I-band. The second class, N-conserved, differentially expressed type of I-band Igs are characterized by a long BC β-hairpin harbouring a PP motif and long FG turns harbouring a NxxG motif as well as the prevalence for a PP motif at the N-terminus within strand A (for strand nomenclature see Figure 1.3 and Figure 1.7a). These two proline residues are expected to be inserted into the groove formed by the BC and FG loops, conferring rigidity to the N-conserved, differentially expressed isoforms.

Studies aiming at elucidation of the relative domain arrangements in the constitutive I-band by a combined approach of NMR and SAXS on a four-domain construct (Impronta et al., 1998) and six-domain construct (Marino et al., 2005) have confirmed the string-of-beads model determined by electron microscopy suggesting a bending of 155° and twisting of 90° degrees with low interdomain mobility between successive domains. As a result a shortening of 10-17% compared to fully extended conformation was found. For the skeletal I-band Igs, this conformational rigidity and extended domain arrangement was proposed to be linker-independent. This is important since titin’s Ig-based elasticity is mainly governed by inter-domain mobility, which in turn depends on linker length and interface residues. The differentially expressed I-band region as well as titin’s A-band are predicted to be even stiffer due to a conserved super-repeat pattern (Tskhovrebova and Trinick, 2004). Concerning titin’s A-band FnIII domains homology modelling approaches utilizing the representative structure of A71 (Muhle Goll et al., 1998) as well as homologous structures from related proteins have identified potential hydrophobic residues in linker regions and on the surface (Muhle-Goll et al., 2001; Amodeo et al., 2001).
Introduction: The M-line of Titin

Figure 1.7: Available high-resolution structures of titin domains.

a) Representative molecular structures of individual domains have been obtained for several regions of titin. The schematic layout shows all domains of titin including cardiac spliceoforms of the I-band. Strands are assigned for Ig I1 according to Mayans et al., 2001 and for FnIII A71 according to Leahy et al., 1994 (Figure modified from: Lange et al., 2006)

Together with the fact that the second position of the long super-repeat was considerably more conserved than the rest (values range from 42 to 64% of total...
surface area) a potential myosin-binding model was proposed and subsequently tested in \textit{in vitro} binding experiments (see previous chapter). Generally, the D-zone shows a lower degree of surface conservation (between 29 to 41\% of total surface area).

Recently, the structure of the N-terminal tandem Ig domains Z1Z1 in complex with telethonin was determined (Zou et al., 2006) (\textbf{Figure 1.7}). The structure reveals a unique, palindromic arrangement of the two overlapping titin N-termini, that is tightly hold together by telethonin. From the structure it became evident that the N-terminus of titin provides a rigid anchoring scaffold linking together two giant titin molecules form the same sarcomere (Lee et al., 2006). The arrangement might be required to avoid uncontrolled disassembly during muscle-contraction. Furthermore additional, mobile molecules might target the Telethonin/Z1Z2 complex (Kontrogianni-Konstantopoulos and Bloch, 2003) and integrate it into the Z-disk stretch sensor machinery (Knoll et al., 2002).

Finally, also the PEVK region has been investigated by CD-spectroscopy in combination with NMR (Ma et al. 2004.) A 28-mer corresponding to human fetal titin contained distinct features of poly-proline II helices interrupted by flexible spacer regions. These regions could therefore act as entropic spring elements.

\section*{1.2.8 Hereditary titin diseases}

The immense size of titin makes it a prominent target for disease causing mutations. Most prominent are mutations leading to dilated cardiomyopathy (DCM). They are localized in titin exon 18 and 326. Mutation in exon 326 is causing a frameshift leading to a 2000 kDa truncated version of titin, which is additionally sensitive to proteolysis. Mutation in exon 18 encodes an Ig domain and the mutation was thought to disrupt the Ig structure (Gerull et al., 2002).

In the case of skeletal muscular dystrophy (TMD) a mutation in exon 365 was found to destabilize the fold of the corresponding Ig domain. Mutations in exon two and 14 (Z-disc) and exon 49 (N2B) element lowered the affinity for telethonin and $\alpha$-actinin (Zou et al., 2006). In addition, a DCM causing mutation of TK was found to interfere with a potential signalling complex at titin's M-line (Lange et al. 2005; \textbf{chapter 1.2.14}). In summary, disease-causing mutations in titin are so far either involved in perturbing individual domains or interfere with protein-protein interactions.
1.2.9 MuRFs

The MuRF family of proteins consists of three, highly conserved isoforms coded by three distinct genes. MuRFs were first described as Striated Muscle RING Zinc finger (SMRZ, Dai & Liew 2001) and sequentially identified by yeast to hybrid screens (Y2H-screens). To date, they are only reported in vertebrates (Centner et al., 2001). MuRF-1 and MuRF-2 share 62%, MuRF-1 and MuRF-3 share 77%, MuRF-2 and MuRF-3 share 65% sequence identity. All three posses a highly conserved N-terminal region (up to 85% within the first 140 residues) containing a zinc binding RING-finger domain (RING), a MuRF Conserved region (MFC) and a zinc binding B-box type two (B2) domain (Figure 1.8). The B-box domain is directly followed by a predicted coiled-coil (CC) region (Centner et al., 2001). A C-terminal acidic tail with high glutamic acid content is the region of strongest divergence within the family. Three additional spliceoforms of MuRF-2 and one of MuRF-3 were identified and are characterized mostly by the presence of C-terminal extensions of unrelated structural nature. The small, cardiac specific isoform of MuRF-2, MuRF-2p27 lacks the coiled-coil regions in its C-terminus (Pizon et al., 2002).

According to domain arrangement and composition, MuRFs are members of the growing family of RBCC (RING, B-box, coiled-coil) or tripartite motif family of proteins (Reddy et al., 1992; Freemont 2000; and chapter 1.3). Further names for MuRF-1, 2, 3 are RING finger protein (RNf28, 29, 30) or Tripartite Motif protein TRIM63, 54, and 55 (http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/search-genes.pl). As a characteristic to TRIM-fold containing proteins, all three MuRFs have the potential to homo- and hetero-oligomerize via their CC regions (Spencer et al., 2000; Centner et al., 2001; McElhinny et al., 2002). However, like for many TRIM proteins, the exact oligomeric state of the MuRFs is not known to date, since in addition to the CC domain, also the N-terminal RING domain has the propensity to form homooligomeric interactions. This oligomerization propensity of TRIM proteins can eventually lead to macromolecular, meso-scale assemblies (Borden, 2000; Kentsis and Borden, 2004). Higher order oligomers are necessary for the correct cellular functioning of the TRIM proteins (Reymond et al., 2001). Furthermore, many TRIM proteins contain canonical features of ubiquitin-ligases (Freemont, 2000; Joaeiro et al. 2000).
1.2.10 Functions of MuRF proteins

MuRF-1 was found constitutively expressed throughout development in three distinct locations in myocytes: in the Z-disc, in the M-line and in a cytoplasmatic pool. Among the best-characterized functions of MuRF-1 is its involvement in muscular atrophy. Atrophy is characterized by muscle protein degradation upon a variety of stimuli including denervation, immobilization, glucocorticoid treatment, space flight and aging (Bodine et al., 2001). Several lines of evidence point towards a role of MuRF-1 in these pathologic conditions:

MuRF-1 possesses intrinsic ubiquitin ligase activity (Bodine et al., 2001). In addition, MuRF-1 null mice are resistant to atrophy and MuRF-1 levels are upregulated in skeletal muscle during the atrophic state. Hence, MuRF-1 is thought to be involved in ubiquitin-dependent degradation of muscle protein (Bodine et al., 2001). Accordingly, in neonatal rat cardiomyocytes MuRF-1 targets troponin I for degradation via its E3-ubiquitin ligase activity (Kedar et al., 2004). Finally, the binding of MuRF-1 to nuclear glucocorticoid modulatory element binding protein-1 (GMEB-1) involved in transcription and up-regulation in response to glucocorticoid exposure, might link MuRF-1 function in atrophy to gene expression profiles in response to external stimuli (McElhinny et al., 2002).

A second function common to all three MuRFs is connected to their association with the cytoskeletal microtubule system. Short and Cox (Short and Cox, 2006) have identified a sequence motif with in the TRIM proteins (MID-1, MID-2, TRIM9, TNL, TRIM36, TRIFIC and the MuRFs) termed COS (C-terminal subgroup One Signature) Box. It is located C-terminally to the CC of the TRIM fold (residues 267 to 326 in MuRF-1), right before the C-terminal acidic tail region and is predicted to contain two α-helical coiled-coils. The authors demonstrate with the help of transfection studies and mutational analysis that members of this distinct TRIM subfamily associate with microtubules in a COS-box dependent manner.

To date, microtubule association has been demonstrated for MuRF-2 and MuRF-3 (Table 1.1), which show different expression pattern and subcellular localization than MuRF-1. MuRF-2 was found to be down-regulated during development, being only present in fetal heart (Centner et al., 2001). It also co-localizes transiently with glutaminated microtubules, is present as essential component of the M-line and was found enriched in nuclei after mechanical inactivity (Pizon et al., 2002; Lange et al.
2005, see chapter 1.2.14). MuRF-2-labelled microtubules transiently associate with myosin and A-band titin in the nascent myofibril. After maturation, labelled microtubules disappear from the sarcomere suggesting an active, microtubule-mediated transport of myosin. The temporal order of MuRF-2 localization (Z-disc, I-band and A-band) furthermore suggests an involvement of MuRF-2/microtubule associations in the stretching and straightening of titin during myofibrillogenesis (Pizon et al., 2002). According to the authors MuRF-2 might function as transient adaptor between the sarcomeric proteins titin and myosin and the microtubular system. MuRF-3 was found to be up-regulated during development and necessary for microtubule stability. It was detected both at the Z-disc and M-line (Spencer et al., 2000).

![Figure 1.8: Overview of the titin binding site for MuRF-1 and MuRF-1 domains.](image)

The central region of MuRF-1 was found to interact with titin’s A-band Ig domains A168-A169 in close proximity the kinase domain of titin. MuRF-1 belongs to the RBCC (RING, B-box, coiled-coil) or tripartite motif (TRIM) family of proteins. Three distinct genes have been observed. Alternative splicing creates isoforms of MuRF-2 and MuRF-3 with differences mostly in the C-terminal tail region (Figure taken from Gregorio et al., 2005). A small cardiac specific isoform of MuRF-2\textsuperscript{27} (not shown) is missing the coiled-coil regions and terminates after the B-box.

Co-localization of myofibrillar-based MuRF-1 and MuRF-2 might link together myofibrillar and microtubular pathways since MuRF-2 co-localizes both with MuRF-1 and microtubular-based MuRF-3. This suggests an interdependence of both systems. Critical to these processes is the hetero-oligomerization potential among the MuRFs...
Introduction: The M-line of Titin (Centner et al., 2001), an ability that is rarely seen across different TRIM containing proteins (Reymond et al., 2001).

In summary, the binding partners of MuRF-1 are either nuclear factors or myofibrillar proteins (Witt et al., 2005; Gregorio et al. 2005; Table 1.1; Figure 1.9). An additional class of proteins interacting with MuRF-1 are metabolic enzymes involved in energy production Witt et al., 2005).

Table 1.1: MuRF proteins and there interaction partners (modified from Gregorio et al., 2005).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Expression</th>
<th>Localization</th>
<th>Interactions</th>
<th>Proposed function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MuRF-1</td>
<td>constitutive during development</td>
<td>Myofibril (Z-disc, M-line)</td>
<td>titin A168-A169, MuRF-1,2,3</td>
<td>Ubiquitin ligase activity</td>
</tr>
<tr>
<td></td>
<td>up-regulated in response to atrophy</td>
<td>Nucleus</td>
<td>SUMO-3,GMEB-1, Ubc-9, IsoF-3, ATP Synthase, creatine kinase</td>
<td>Degradation of muscle proteins during atrophy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rack1, Nebulin, N-RAP, TnT</td>
<td>Thick filament assembly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MLC-2</td>
<td>M-line integrity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Myotilin, Teletethionin (T-cap)</td>
<td>myofibril maintenance/turnover</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cardiac troponin I (cTpi)</td>
<td></td>
<td>Regulator of energy metabolism</td>
</tr>
<tr>
<td>MuRF-2</td>
<td>down-regulated during development</td>
<td>Stable microtubules (Glu-MTs)</td>
<td>titin A164-A169, MuRF-1,2,3</td>
<td>Stability of stable microtubules and intermediate filaments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myofibril (M-line)</td>
<td></td>
<td>adaptor between titin and microtubules</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nucleus (upon mechanical inactivity)</td>
<td>p62/SQSTM1, Nebulin, N-RAP</td>
<td>critical for myoblast fusion and myofibrillogenesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TnT, MLC-2</td>
<td>M-line integrity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Myotilin, Teletethionin (T-cap)</td>
<td>Component of M-line stretch sensor</td>
</tr>
<tr>
<td>MuRF-3</td>
<td>up-regulated during development</td>
<td>Stable microtubules (Glu-MTs)</td>
<td>MuRF-1,2,3</td>
<td>establish/maintain stable Glu-MT array</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myofibril (Z-disc, M-line)</td>
<td></td>
<td>required for myoblast fusion and myogenesis, gene expression and differentiation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Microtubules</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.2.11 The titin-interface with MuRF-1

Inspired by the strong homology of the titin kinase region with other “giant kinases” like myosin-light chain kinase or invertebrate titin homologues, Centner et al. have carried out Y2H-screens using the titin A167-kinase-M2 region (P-zone, Figure 1.8) as bait (Centner et al., 2001). MuRF-1 was found to interact with titin at the two Ig domains A168 and A169, which are located N-terminally to the kinase domain, ~100 nm away from the centre of the M-line in the M-line periphery (Figure 1.6).

On the MuRF-1 side, 144 amino acids of the central region (residues 74-218) including the B-box and parts of the adjacent coiled-coil have been identified by Y2H screens as necessary and sufficient for binding. The studies were also complemented by pull-down assays (Centner et al. 2001). Whereas MuRF-1 was found to be the only
MuRF that interacts with titin, Pizon et al. (Pizon et al. 2002) have also demonstrated a direct interaction of MuRF-2 with titin domains A164-A169. Further studies will be needed to clarify this discrepancy.

More recently, after the initiation of the study in hand, Witt et al. (Witt et al., 2005) have further characterized the titin binding region on MuRF-1 by SPOTS blots using a set of overlapping peptides of MuRF-1 and a titin A168-A170 construct. They identified a C-terminal region at the coiled-coil periphery as primary binding site (residues 284-315). Notably, this region overlaps with the COS-Box motif that was found to mediate microtubule association of certain TRIMs (Short and Cox, 2006).

1.2.12 Function of MuRF-1 in M-line integrity and sarcomeric protein turnover

MuRF-1 was found in Z-disc, M-line location was well as in a cytosolic pool and nuclei of myocytes and a dynamic, controlled translocation between the locations was suggested (Centner et al., 2001; McElhinny et al., 2002). Overexpression of full MuRF-1, its central region (comprising the B-box and adjacent coiled-coil regions) or its complementary titin binding site A168-A170, resulted in a severe disruption of the M-line region and the organization of the thick filament. MuRF-1 was found in aggregates of varying size and intensity. However, other regions of titin like the Z-disc or the thin filaments are not affected as verified by site-specific immunofluorescence studies and might be stabilized by Z-disc and thin filament specific components (McElhinny et al., 2002). A dominant-negative phenotype caused by the truncation MuRF-1 was excluded by the fact that overexpression of full-length MuRF-1 also lead to the same phenotype furthermore suggesting a tight regulation of MuRF-1. Hence, the constitutive expression of MuRF-1 seems to be essential for myofibril stability throughout development. Nevertheless, reducing the levels of MuRF-2 by anti-sense oligonucleotides results in similar impairment of myofibrillogenesis with disrupted M-line, perturbation of microtubule populations and the intermediate filaments desmin and vimentin (McElhinny et al. 2004).

Studies by Weinert et al. (Weinert et al., 2006) show that complete knockout of the M-line region of titin (ΔA169-M7) including the kinase domain and M-line interfaces to MuRF-1, myomesin and calpain does not influence sarcomere assembly. However, it severely interferes with sarcomere strength and lateral growth at later stages eventually leading to sarcomere disassembly. A similar observation was made
upon deletion of the kinase and MuRF-1 binding region (Gotthardt et al., 2003). Surprisingly, knockout of MuRF-1 in mice showed no primary effect in myofibril integrity, but animals were resistant to muscle atrophy (Bodine et al., 2001).

The E3 ubiquitin-ligase activity of the MuRF-1 furthermore suggests a role in sarcomeric protein turnover. Witt et al. have detected ubiquitination of titin, nebulin, myosin light chain-2 (MLC-2) and troponin cTnI in skeletal and cardiac muscle by staining of Western Blots with anti-ubiquitin antibodies (Witt et al., 2005). This suggests that besides titin, the myofibril itself could be a target of MuRF-1 dependent ubiquitin transfer. The authors localized the ubiquitination epitope on titin to the A-band region. However, in a MuRF-1 knockout, the target proteins were still ubiquitinated suggesting that other ubiquitin-ligases are involved in the ubiquitination process or that ubiquitination may not be an exclusive property of MuRF-1. Possibly they could also be carried out by MuRF-2 and3, which where not affected by the knockout.

### 1.2.13 Multiple localisation of MuRF proteins and the SUMO pathway

Y2H-screens revealed that all MuRF's interact via their RING domain with small ubiquitin like modifier SMT3b/SUMO-3 (Dai and Liew, 2001) and isopeptidase T-3 (ISOT-3). Ubiquitination and SUMOylation are related, but distinct mechanism of post-translational modification of target proteins (Lange et al., 2006).

Nuclear and cytoplasmatic localisation is characteristic for sumoylated proteins (Melchior, 2001, Melchior et al., 2003). Accordingly, binding of MuRF-1 to ubiquitin-modifying enzyme-9 (Ubc-9), which acts as a SUMO E2 transferase (Tong et al., 1998; McElhinny et al. 2002) was detected. Nuclear import, influence on gene expression and subcellular localisation of MuRF family members might be controlled by SUMOylation. However, to date it is not clear if MuRF-1 is indeed SUMOylated or merely interacts with sumoylated proteins.

Furthermore, SUMO-proteins are components of cellular response pathways (Saitoh and Hichey 2000). Hence MuRF-1 SUMOylation may also be a response to the atrophic state of muscle. SUMOylation does not lead to degradation, but instead regulates protein-protein interactions, intracellular localization and protects modified targets from ubiquitin-dependent degradation (Pichler and Melchior, 2002). Its interaction with the nuclear transcription factor GMEB-1, which is involved in
transcriptional control in response to altered glucocorticoid levels, probably represents the nuclear part of a signalling module involved in stress detection and muscular atrophy. It clearly highlights the link between gene expression and myofibrillar signalling.

1.2.14 Titin as a sarcomeric stretch sensor: the kinase signalosome

Titin might regulate MuRF activity in a stretch-dependent fashion. Recently, a signalling module at the titin M-line was detected potentially involved in mechanotransduction (Lange et al. 2005; Figure 1.9). It links the titin kinase domain via two multi-compartment zinc finger proteins, Nbr1 and p62 to MuRF-2 and furthermore to the nuclear transcription factor serum response factor (SRF).

Figure 1.9: The M-line signal transduction module of titin.

A functional module that is proposed to act as a sarcomeric stretch sensor. Mechanically induced formation of a signalling cascade consisting of titin kinase, the scaffolding components NBR1, p62, MuRF-2 and the nuclear transcriptional regulator SRF translates mechanical into transcriptional stimuli. (Figure modified from Lange et al., 2005).

Key prerequisite for formation of this signalling cascade is the presence of a kinase lacking parts of the C-terminal autoinhibitory segment. Based on modelling by force-probe molecular dynamics (Gräter et al., 2005), the authors refer to this conformation as a conformation that resembles a mechanically activated titin kinase. However, currently both in vivo and in vitro data do not exist that proof the existence of
such a conformation. Several constructs designed based on these in silico data were screened for potential interaction partners in Y2H-screens.

A scaffolding protein called next to breast cancer gene-1 (Nbr1) was found to interact with the truncated TK via its N-terminal PB1 domain. On the contrary, wildtype TK did not lead to any interaction. Structurally, the PB-1 domain (~85 residues) resembles the ubiquitin-like β-grasp fold (Müller et al., 2006). It mediates polar heterodimeric interactions by two types of PB1 domain termed A and B, which are of opposite charge on their surface (Noda et al., 2003).

Subsequent Y2H-screening lead to the identification of p62, which together with the PB domain of Nbr1 is a prominent scaffolding unit of large signalling complexes (Pawson and Scott, 1997). MuRF-2 was identified as ligand of p62 and consistently, all three proteins, Nbr1, p62 and MuRF-2 were detectable at the M-band by immunofluorescence microscopy. Both Nbr1 and p62 contain a zinc binding ZZ domain that possess strong similarities to the RING-finger like domain of MuRF-1 (Legge et al., 2004). Moreover, both proteins are involved in homo-oligomeric interactions.

Upon mechanical arrest by chemical agents or denervation, a substantial accumulation of MuRF-2 was detectable in the nucleus. Nuclear MuRF-2 was found to interact via its RING/B-box domain to the trans-activation domain (TA) of serum response factor (SRF) (Lange et al., 2005), a protein involved in myogenic transcription events (Li et al., 2004). SRF belongs to the MADS box family of transcription factors. The N-terminal half of the MADS motif determines DNA binding specificity, while the C-terminal part is involved in dimerization. SRF controls expression of immediate early genes (IEG), for example the prototypic IEG c-fos (Murphy et al, 2002), most of which are by themselves transcription factors or DNA binding proteins and represent a rapid response mechanism to various cellular stimuli (Davis et al., 2003). It acts as a nuclear target of kinases and Rho signalling networks directing cell activities as diverse as cell-cycle progression, neuronal synaptic transmission and muscle cell differentiation (Mercado et al., 2003).

Upon overexpression of the small cardiac isoforms of MuRF-2 (missing the C-terminal CC region), a substantial suppression of c-fos activity was detectable, which is indicative of a strong influence of mechanically induced nuclear localization of MuRF-2 on transcription events controlled by SRF. Conversely, overexpression of TK could
rescue the mechanically arrest-dependent low transcription levels of c-fos. Thus, the presence of the NBR1-p62-MuRF-2 protein complex clearly reduced the nuclear pool of MuRF-2 leading to unhindered SRF activity.

Finally, a mutation in titin kinase regulatory domain (R279W) that leads to hereditary myopathy with early respiratory failure (HMERF) was found to inhibit binding of TK to NBR-1. As a result, the structural organization of the sarcomere was impaired (Edstrom et al., 1990) highlighting the importance of TK-NBR1 interaction for correct sarcomere function.

Taken together, the mechanically induced formation of the TK-NBR1-p62-MuRF-2-SRF module might form a sarcomeric stretch-sensor that could translate mechanical stimuli into biological response, i.e. affecting muscle gene expression.

The above signalling cascade represents an alternative pathway to the MuRF-1/SUMO-3/GMEB-1 centred signalling network mentioned by McElhinny et al. (McElhinny et al. 2002). Clearly, further studies are needed to elucidate the role of titin and associated proteins in sarcomeric signalling events to unravel the molecular basis of myofibrillogenesis, mechanical properties of the sarcomere and pathological situations leading for example to muscle atrophy.
1.3 The B-box, protein-protein interaction domain of the titin-associated protein MuRF-1

1.3.1 MuRF-1 is a member of a protein family with the characteristic TRIM fold

TRIM or RBCC proteins are eukaryotic multidomain proteins that share a common tripartite domain architecture at their N-terminus. The TRIM fold is characterized by the presence of a Really Interesting New Gene (RING or A-box; Reddy et al. 1992), one or two B-box domains and an associated coiled-coil domain (Freemont et al., 1993). Hence proteins containing this motif are named TRIM for Tripartite Motif or RBCC proteins standing for RING, B-box, Coiled-Coil. Currently, 65 members of the TRIM family are known in man and less than 20 members in flies and worms (Short and Cox, 2006; Meroni & Diez-Roux 2005). An overview can be found in Figure 1.10a and in the web at (http://www.informatics.jax.org/mgihome/nomen/genefamilies/trim.shtml).

In contrast to the RING domain, which is found in many other proteins, the B-box is the critical determinant of the TRIM motif and some proteins have been termed TRIM despite the absence of the RING domain (Reddy et al., 1992). Following the CC domain, the C-terminus of TRIM proteins shows strong divergence in domain composition and sequence length (Figure 1.10c).

Throughout the genome, TRIM/RBCC proteins are highly dispersed, which suggests their independent evolution (Reymond et al., 2001). The RING domain is not found in prokaryotes consistent with the lack of the ubiquitination system in these organisms (see below). B-box domains are absent in prokaryotes, viruses and fungi (Meroni and Diez-Roux, 2005). RING and B-boxes are characterized through unique sequence consensi (Figure 1.10c).

1.3.2 TRIM components

The best-characterized domain of the TRIM fold is the RING domain. Currently over 17 structures of RING domains have become available in the RCSB protein database (for overview see prosite entry PS00518).

The first determined RING structure was the one of the immediate early gene of herpes simplex virus (IEEV) determined by Barlow et al. (Barlow et al., 1994). The
RING domain binds two zinc atoms by means of eight conserved cysteine or histidine ligands thereby adopting a characteristic topology known as „cross-brace“ (Figure 1.11a). In this motif the first and third metal ligation pair bind the first zinc atom, while the second and fourth metal ligation pair bind the second zinc atom (Freemont, 1993). Two main types of RING fingers have been described: The H2- and C2-subtypes depend on the nature (cystidine or histidine) of the metal ligand at position five (Freemont, 2000). Notably, only C2-type RINGs are found in the TRIM/RBCC family.

Specific functions of the RING domain include a critical role in ubiquitination events (Joazeiro et al., 2000). Ubiquitin ligases like MuRF-1 (known as E3s) are involved in the last step of a three-enzyme cascade involving so-called ubiquitin activating (E1) and ubiquitin conjugating enzymes (E2) (Pickart et al., 2001). The E3 mediates the transfer of ubiquitin from the E2 to the substrate by formation of an isopeptide bond between ubiquitin carboxy-terminus and specific lysine chains on the substrate. The interface forms a shallow groove on the domain surface. Apart from binding of the E2 moiety, it additionally confers specificity to the reaction. Borden and co-workers refer to a second function of the RING domain in the RING-dependent self-assembly of large macromolecular assemblies through homo-oligomeric protein-protein interactions (Borden, 2000; Kentsis and Borden, 2002).

Like the RING, the B-box is also a zinc-binding domain with a similar yet distinct pattern of metal binding residues. The B-box motif has been sub-classified by Freemont and others (Freemont, 1993) into a B-Box 1 (B1) and a B-box 2 (B2) subtype. The two variants mainly differ in the nature of the second metal ligand, which is a cysteine in B1 and a histidine in B2, and in the slightly longer insertions between the metal binding ligands in B1 (for consensus sequences see Figure 1.10c). Two high-resolution B-box structures are available, one of a B1 and the other of a B2 B-box domain (Figure 1.11b and Figure 1.11c).
Introduction: The B-box domain of MuRF-1

Figure 1.10: Overview of human TRIM family members and domains.

a) Human TRIM proteins can be sub-classified according to their domain composition. The MuRF family constitutes a distinct subclass. Identity, alternative names, chromosomal location and sub-class based on domain composition are shown. b) Sub-classification based on C-terminal variable domains according to structure elucidation of the B2 domain from the TRIM protein Xenopus nuclear factor-7 (XNF7) by means of NMR–spectroscopy (Borden et al., 1995) in combination with titration experiments (Borden et al., 1993) revealed that of the eight conserved metal ligands in the B-box motif, only four are involved in binding of one structural zinc ion, leaving four potential metal binding ligands unoccupied. Interestingly, the recently determined B1 solution structure of the TRIM protein midine-1 (MID-1, TRIM18), which became available during formation of this manuscript, shows a remarkably different topology (Massiah et al., 2006). While the B1 domain adopts a RING-finger-like fold by ligation of two structural zinc ions, the XNF7 structure folds into a unique topology through ligation of only one zinc atom. Apart from general

b) Sequence Consensus

RING (R)
C-x2-C-x11-16-C-x-H-x2-C-x2-C-x7-74-C-x2-[CD]

B-box 1 (B1)
C-x2-C-x-x12-16-C-x2-C-x2-C-x2-[CH]x3-4-H-x4-9-H

B-box 2 (B2)
C-x2-H-x7-9-C-x2-[CDHE]x4-C-x2-C-x16-12-H-x2-4-[CH]
involvement in protein-protein interactions, a specific function has only been attributed to the B-Box B1 of MID-1 so far. It was found to bind directly to Alpha4, a subunit of the catalytic domain of microtubular phosphatase 2A (PP2Ac) (Liu et al., 2001).

In all TRIM containing proteins the B2 domain is always invariably followed by a predicted coiled-coil region of approximately 100 residues, which is composed of two or more segments (Torok et al., 2001). However, currently no structural data exists on any coiled-coil region of TRIMs and the exact location of the coiled-coil segments can only been inferred from predictions.

Functionally, the coiled-coil was found to be the critical determinant of homo-oligomer formation leading to the formation of large macromolecular assemblies. Through homo-/hetero-oligomerization the TRIM is able to target specific cellular structures including cytoskeletal filaments, chromatin and also defines its own structures of compartment character like for example nuclear bodies (Reymond et al.,
2001). In contrast to Borden (Borden, 2000), Reymond et al. did not detect any RING-dependence during formation of higher order assemblies for a broad range of TRIMs in studies with recombinantly expressed proteins and size exclusion chromatography for determination of the oligomeric state. The molecular basis of TRIM oligomerization is currently unknown.

1.3.3 Variable C-terminal domain of TRIM containing proteins

More recently Cox et al. have sub-classified TRIM protein based on their unique C-terminal domains and TRIM composition.

The most common motifs in the C-terminal region of TRIM proteins comprise the B30.2 domain also known as ret finger protein-like domain (RFP-like; Henry et al., 1998), NHL repeat domains (Slack and Ruvkun, 1998) and a PHD-BROMO domain tandem (Torrok et al., 2001; Figure 1.10b). But also Ig, FnIII and Armadillo-repeats (ARF) can be found in C-terminal position to the TRIM. These domains are not unique to TRIM proteins and are found independently in many different eukaryotic and prokaryotic proteins with a broad functional spectrum. Nevertheless, also unique C-terminal sequences with no homology to any known motif as well as spliceoforms are observed. Recently, Short and Cox have identified an approximately 60 residues long sequence motif in the MuRF family members and other TRIM proteins that is responsible for microtubule association (Short and Cox, 2006; chapter 1.2.10). It is located at the C-terminal coiled-coil periphery.

TRIM-based diseases are often caused by mutations in the coiled-coil and the C-terminal TRIM-associated domains or result from recombination of the TRIM motif with specific kinases (Torok, 2000). As a primary effect in many disease models, an incorrect oligomerization state of the TRIM protein can lead to a failure of identifying its specific protein target. Second, given the E3 ligase activity of the RING, a direct link between pathological phenotype and E3-activity of the TRIM proteins might exist in many cases (Meroni and Diez-Roux, 2005).

Probably the most prominent example of a TRIM-based disease is caused by mutations in the C-terminal rfp-like motif as well as in the TRIM motif of the E3 ubiquitin ligase MID-1 leading to Opitz/BBB syndrome. The mutations result in an impairment of MID-1 binding to microtubules (DeFalco et al., 2003) As a secondary effect, a diminished proteolysis of the catalytical subunit of microtubule associated
phosphatase 2A is thought to lead to abnormal phosphorylation events in microtubule-associated proteins and hence pathologically altered microtubule dynamics (Trockenbacher et al., 2001).

1.3.4 The TRIM as an integrated fold

A striking feature of the TRIM fold is its conserved domain structure. The spacing between as well as the order of RING, B-box and coiled-coil domain is absolutely conserved in all TRIMs and throughout evolution (Reymond et al., 2001). Rearrangements within the TRIM cause severe cellular dysfunctions (Torok et. al. 2001). This clearly highlights the functional relevance of the TRIM fold as an integrated structural unit rather than a collection of individual motifs. On the primary sequence level, the entire family is characterized by a low sequence homology apart from the highly conserved metal binding ligands of RING and B-box domains and the heptad repeats necessary for the “knobs-into-wholes” coiled-coil packing interactions.

Peng et al. have demonstrated by mutational analysis for the KRAB-KAP-1 transcriptional repression module that each subdomain of the TRIM protein KAP1 (now termed TIF1-β or TRIM28) contributes to binding to the KRAB domain of the transcription factor KOX-1. The RBCC domain either provides direct interaction determinants or indirectly mediates formation of the appropriate homo-oligomeric state of KAP-1 (Peng et al., 2000).

In the case of TRIM motifs containing two B-boxes, both are always of different subtype B1 and B2, with B1 exclusively preceding B2. When only one B-box domain is present in the motif, it is always of B2 type, suggesting a strong association of B2 and coiled-coil domain potentially forming a single structural module within the TRIM motif. A role of B2 in orientation of the coiled-coil moieties domain was proposed (Cao et al. 1997). Furthermore, through intensive domain swapping experiments Peng et al. show that RING, B-boxes and coiled-coils from highly homologues TRIMs cannot functionally substitute each other. All these observations clearly suggest that the TRIM fold serves as an integrated, functional scaffold involved in protein-protein interactions either in the context of the TRIM proteins itself or to other cellular components.
1.4 LAP2α, a nuclear scaffold protein involved in chromatin organization

1.4.1 Nucleus, nuclear envelope and lamina

The nucleus is the defining feature of eukaryotic cells. Within this compartment a number of important functions are carried out like DNA replication, RNA transcription, RNA processing and ribosome assembly. Considering the high concentration of macromolecules inside the nucleus, a high degree of spatial order is one essential prerequisite (Dechat et al., 1998). At the periphery of the nucleus, organization is provided by the two structurally and functionally distinct membrane systems of the inner and outer nuclear membrane. The outer nuclear membrane is continuous with the endoplasmic reticulum (ER) and covered by ribosomes. Together with the perinuclear space it is involved in protein translation and modification. The inner nuclear membrane contains integral membrane proteins that link the membrane to the underlying filamentous meshwork of the nuclear lamina and to other nuclear components. Nuclear pore complexes (NPCs) at membrane junctions mediate macromolecular transport across these two boundaries. Membranes, NPCs and the nuclear lamina together form the nuclear envelope (NE), which separates cytoplasmatic and nucleic activities during interphase and controls their interactions (Figure 1.12) However, it is primarily the nuclear lamina, which contributes to the structural integrity of the NE (Dechat et al., 1998; Gant et al., 1999; Gerace and Burke, 1988; Moir et al., 1995).

The lamina consists mostly of a set of self-assembling proteins called lamins that form are type V intermediate filament (IF) proteins. These nucleo-specific IF-proteins line the nucleo-plasmatic face of the inner nuclear membrane with a two-dimensional, quasi-tetragonal meshwork, which is supposed to provide size, shape and mechanical stability of the NE (Stuurman, 1998). The reversible process of filament formation is required for a proper disassembly and reassembly of the lamina during open mitosis in eukaryotes and is controlled by mitose-specific phosphorylation (Dechat et al., 1998; Nigg, 1992). However, single cell eukaryotes and plants do not have a nuclear lamina.
Figure 1.12: Components of the nuclear envelope and nuclear interior associated with lamins.

Inner and outer nuclear membrane (INM, ONM) together with the peripheral nuclear lamina (in white) and nuclear pore complexes constitute the nuclear envelope. The protein content of the inner nuclear membrane consists of membrane and membrane anchored proteins, which provide attachment sites for the lamina and chromatin. LAP2α is found associated with intranuclear lamins, chromatin and retinoblastoma protein (pRB) (Figure taken from: Foisner, 2001).

Lamins are classified as type A and type B according to homology in sequence, expression pattern, biochemical properties and subcellular location. A-type lamins are expressed in differentiated cells, have a neutral isoelectric point and are completely soluble in the cytoplasm during mitosis. B-type lamins are constitutively expressed throughout development in every cell, have acidic isoelectric points and remain membrane associated during mitosis through C-terminal farnesylation. There are two B-type lamins encoded on separate genes in vertebrates and one lamin A gene. Alternative splicing generates at least three lamin B and four lamin A isoforms (Gerace and Burke, 1988).

Lamins directly bind to chromosomes, they are required for DNA replication and are involved in chromatin (Dechat et al., 1998) and nuclear pore organization (Markiewicz et al., 2002). Moreover, they are early targets of proteolytic degradation by specific caspases during apoptosis (Gotzmann et al., 2000). Upon cleavage, the lamina undergoes specific morphological changes that are typical of apoptotic nuclei. A characteristic feature of lamins and other IF proteins is their chemical resistance against denaturing agents and high salt washes.
More recently, intranuclear structures containing only lamin A have been detected during interphase (Dechat et al., 1998; Goldberg et al., 1999; Broers et al., 1999; Broers et al., 2004). Besides providing mechanical stability, this tubular network in the nuclear interior might also be involved in higher order chromatin organization (Dechat et al., 1998; Gotzmann and Foisner, 1999; Broers et al., 2005).

Several lamina-associated proteins (LAPs) have also been implicated in the structural organization of peripheral and intranuclear lamin (Figure 1.12). Among these, LAP1A, 1B, 1C (Foisner and Gerace, 1993; Martin et al., 1995) and LAP2 (Furukawa et al., 1995) are type II integral membrane proteins of the inner nuclear membrane, which share a 43-residue homologous LEM domain near their N-termini with two other inner nuclear membrane proteins named emerin and MAN1 (LAP2-EMERIN-MAN).

LAP2, now called LAP2β (Vlcek et al., 1999), has been proven to bind directly to lamin B at the nuclear periphery and most likely indirectly to chromosomes and therefore might be involved in attaching nuclear envelope and chromosomes (Dechat et al., 1998). LAP2α, the only non-membrane bound isoform of the LAP2 family of proteins has been proven to associate preferentially with A type lamins in the nuclear interior (Dechat et al., 2000).

1.4.2 Structure and interactions of lamins

Similar to many cytosolic intermediate filaments (e.g. vimentin, desmin, keratin, etc.) the nuclear lamins consist of a central rod domain comprising up to 350 residues and contain a typical hydrophobic heptad repeat pattern. The central rod is a double stranded α-helical coiled-coil flanked by two non α-helical domains, termed N-terminal head and C-terminal tail domain. The rod and the tail domain together are responsible for the myosin-like shape of the parallel lamin dimer in electron micrographs. Phosphorylation sites can be found on both ends of the central rod domain. Mitose-specific phosphorylation induces conformational changes leading to a complete disassembly of the lamina (Gerace and Burke, 1988; Stuurman et al., 1998). The next level of lamin assembly involves longitudinal association of dimers into head to tail polymers (in contrast to cytoplasmatic IFs) and finally lateral association into filaments or eventually paracrystalline arrays (Stuurman et al., 1998). It is noteworthy that globular head and tail domains are present in over one million copies in the fully
assembled lamin meshwork providing a huge surface for protein-protein interactions (Zastrow et al., 2004).

The atomic structure of the C-terminal globular domain (residues 428-547) has been determined by NMR (Krimm et al., 2002), and by X-ray crystallography (Dhe-Paganon et al., 2002). The domain adopts a novel immunoglobulin fold (L-subtype) composed of seven beta strands that are classical for s-type Igs (Bork et al., 1994) and two lamin specific extra strands. The nine strands form two \( \beta \)-sheets of five and four strands that are tilted against the axis of the two sheets by a large angle of 45° (Figure 1.13d).

More recently, one fragment of the 310 residues long lamin A rod domain, termed coil 2B has been characterized structurally by X-ray crystallography (Strelkov et al., 2004; Figure 1.13e). Furthermore, the rod domain is sufficient for binding to mitotic chromosomes. Lamin A has been found to interact with DNA by interactions with the minor groove (Luderus et al., 1994; ). A large number of lamin binding partners have been discovered (for a recent review see: Zastrow et al., 2004).

### 1.4.3 Structure and interactions of LAP2 isoforms \( \alpha \) and \( \beta \)

The LAP family of proteins can be divided into two main groups: LAP1 and LAP2 isoforms. LAP1 isoforms 1A, 1B and 1C are membrane-bound forms exclusively localized at the nuclear rim in a continuous pattern. They associate with lamin A that together with lamin B forms the major component of the peripheral nuclear lamina (Foisner and Gerace, 1993).

LAP2\( \beta \), the most extensively characterized isoform of LAP2, binds to the lamin B tail domain in a phosphorylation dependent manner (Dreger et al., 1999). The human LAP2 isoforms \( \beta \) (51 kDa), \( \delta \), \( \epsilon \), \( \gamma \) and \( \zeta \) (39 kDa) are structurally related (Figure 1.14). They all possess an N-terminus that is oriented towards the nucleoplasm, a single trans-membrane domain (apart from LAP2\( \zeta \)) and a C-terminal domain located in the luminal space between inner and outer nuclear membrane (Dechat et al., 2000b). Isoforms \( \delta \), \( \epsilon \), \( \gamma \) and \( \zeta \) have only been reported as mouse mRNA transcripts so far (Dechat et al., 2000).

All LAP2 isoforms possess a highly conserved LEM domain within the first 187 amino acid at their N-terminus. This region is termed LAP2 constant region and
structural information has become available for residues 1-50 and 111-153 by solution NMR (Laguri et al., 2001). The constant region seems to be composed of two structurally independent, non-interacting globular domains connected by a highly flexible linker. Both regions are composed of two helices connected by a long loop and tilted 45° with respect to each other and are proposed to tumble independently in solution (Figure 1.13 a and Figure 1.13 b).

The N-terminal “LEM-like” region is involved in direct DNA binding through positively charged residues. The LEM domain (amino acids 67-137) interacts directly with the DNA-bridging protein BAF (barrier to autointegration factor), which is proposed to block the auto-integration of retroviral DNA by compacting it into a rigid structure. Binding to BAF takes place via complementary, hydrophobic interactions (Laguri et al., 2001).

Figure 1.13: Available structural information for lamin and associated.

a) LEM-like domain of LAP2 involved in direct DNA binding
b) LEM domain of LAP2 involved in BAF binding (Figure a and b, pdb: 1GJJ, Cai et al., 2001)
c) C-terminal domain of MAN-1 involved in DNA binding (Caputo et al. unpublished)
d) Structure of the C-terminal tail Ig of human lamin A/C (Dhe-Paganon et al., 2002)
e) Structure of human lamin A/C coil 2B (Strelkov et al., 2004).
The largest isoform of the LAP2 proteins is LAP2α. It is more distantly related to the other isoforms and together with LAP2ζ, the only non-membrane bound gene product distributed in the nuclear interior (Figure 1.14). Apart from the constant 1-187 amino acids found in all LAPs, it possesses a unique 506 residues long C-terminal α-specific region that contains the proposed binding site for lamin A between residues 615 and 693 (Dechat et al., 2000a). No structural information exists on the lamin binding domains of any of the LAPs.

**Figure 1.14: Localisation and domain organization of mammalian LAP2 isoforms.**

**a)** The N-terminus is conserved in all alternatively spliced isoforms and consists of a LEM and LEM-like domain separated by a flexible linker region. Sequence deviations occur in the nucleoplasmatic domains (green); LAP2α is more distantly related to the other isoforms and has a unique nucleoplasmatic domain (blue). On the protein level, only LAP2α, β, γ have been detected in humans. The other isoforms are reported as mRNA transcripts in mouse; IM: inner nuclear membrane; OM: outer nuclear membrane; NLS, nuclear localization signal. **b)** Binding partners of LAP2α and their reported binding region; LEM, domain homologous in the nuclear envelope proteins LAP, Emerin and Man-1.

Furthermore, the unique C-terminus harbours a chromosome interaction domain (between 270-615) that is functionally distinct from the N-terminal DNA interaction region and essential and sufficient for targeting LAP2α to chromosomes (Vlcek et al., 1999).
More recently an interaction of LAP2α with retinoblastoma protein (pRb) has been described (Markiewicz et al., 2002). This protein is involved in regulation of cell cycle progression by binding to the transcription factor E2F. It has a well-defined domain structure consisting of an N-terminal oligomerization domain and three C-terminal “pocket” domains. LAP2α has been found to bind to the pocket C domain of Rb and anchors the protein inside the nucleus through selective retention. Binding takes place within the α-specific region residues 415-693 (Dorner et al., 2006), which are unique to LAP2α and contains functionally significant motifs important for chromosome targeting and lamin binding. This suggests the presence of trimeric LAP2α, pRb, lamin A/C complexes that have been detected in coimmunoprecipitation experiments. A function of this complex lies in the regulation of cell proliferation and differentiation in adult stem cells (Markiewicz et al., 2002).

1.4.4 Function of LAP2s and binding partners in nuclear structure and dynamics

All higher eukaryotes undergo an open mitosis. Upon phosphorylation of lamins and lamin binding proteins the nuclear envelope completely disassembles. The condensing chromosomes detach from the nuclear lamina. Lamin A/C complexes stay cytoplasmatic, lamin B structures remain membrane-associated and lamina associated proteins loose their affinity for lamins.

By subcellular fractionation studies LAP2α has been found to be in soluble form and to be localized in the nuclear interior during interphase (Dechat et al., 2000). It co-fractionated with LAP2β in lamina enriched fractions after extraction with salt, urea and nucleases. After reassembly it localized to distinct sites around decondensing chromosomes prior to LAP2β. In vitro dephosphorylation lead to its integration into insoluble, chromatin associated structures. Two perfect and five minimal consensus sites were found for p34\(^{cdk2}\) kinase (Gajensko et al., 2004).

Like LAP2β it associates also with the chromosomal surface with a region that is not located within the first 187 amino acids. Hence the functional diversity of LAP2α and LAP2β in postmitotic chromosome association depends on the isoform-specific nuclear targeting domain (Vlcek et al., 1999). There is no need for active transport into the nucleus since membranes are not entirely formed when LAP2α accumulates at nuclear structures through selective retention.
LAP2α might have important functions in early, post-mitotic, nuclear organization. The findings suggest that LAP2α binds decondensing chromosomes and forms a structural scaffold for higher chromatin organization. At first, LAP2α associates with its α-specific region to chromosomes prior to LAP2β. In a second step the N-terminal binding sites might recruit BAF, chromosomal proteins and DNA to organize a higher order structure of the chromatin. Binding of the LEM-like domain to DNA can only take place when LAP2α is a dimer. The temporal coordination of these events can be maintained through phosphorylation and/or oligomerization (Dechat et al., 2000b).

More recently, it became clear that LAP2α has distinct functions in nuclear assembly during mitosis and in the progression of cells into S-phase. The inhibition of nuclear assembly by overexpression of C-terminal nuclear targeting domain interferes with endogenous LAP2α and shows that the N-terminal, constant region is necessary for nuclear assembly, thereby competing for binding sites (dominant negative effect). In contrast, cell-cycle arrest was caused by overexpression of full and C-terminal LAP2α, which indicates a specific function of the C-terminus in this process.

LAP2α could control G1-S phase progression through binding to transcription factors and pRb, which represses transcription of S-Phase genes. Furthermore it can manipulate higher order chromatin organization and prevent DNA replication (Markiewicz et al., 2002).

LAP2α has also a function during apoptosis since the lamina provides an attachment site for the apoptotic machinery. LAP2α, LAP2β and lamins are early targets of apoptotic caspases. Mislocalized lamins might trigger apoptosis (Gotzmann et al., 2000).

1.4.5 Diseases associated with lamins

Mutations in genes encoding for nuclear lamina proteins have been found to cause diseases called laminopathies, i.e. severe disorders of muscle and adipose tissue. The X-linked form of Emery-Dryfuss muscular dystrophy is caused by mutations in the LEM-domain containing protein Emerin. The autosomal dominant variety of EDMD is caused by mutations in the lamin A/C. Since there is a reported direct interaction of Emerin with lamin A/C, the molecular mechanism of the disease might involve the disturbance
in lamin A/C-Emerin complexes caused by mutations. Because LAP2α has also been described to interact directly with lamins, it is believed that the same is true for LAP2α-lamin A/C complexes.
1.5 Challenges of structural studies on filamentous proteins and protein scaffolds

Filamentous proteins are challenging targets for structural studies employing X-ray crystallography. To date, it remains the method of choice for structure determination of protein complexes at atomic and sub-atomic resolution (Russel et al., 2004). Difficulties when working with filamentous, structural proteins are encountered during all stages of structure elucidation including construct design, isolation and purification, biophysical characterization, crystallization and finally structure determination:

i.) Filamentous structural proteins are by definition elongated, multi-modular assemblies of large size and molecular weight (in the case of titin over 1 µm in length and up to 3.8 MDa) and cannot be analysed on the full protein level. The design of working constructs for structural studies must carefully balance between the amount of information gained and experimental constraints like solubility, chemical stability, homogeneity, conformational rigidity and correct determination of domain boundaries of all protein partners under investigation. Often initial biological interaction data are present based on yeast two-hybrid screen (Y2H), immuno-fluorescence microscopy or affinity tag methods (review Piehler, 2005). These methods serve as initial screening tools for protein-protein interactions, but have a number of problems with both false positives and negatives. The incidence of false-positive results particularly increases as complexes become less stable and biophysical verification by other techniques like mass-spectroscopy, calorimetry, analytical ultracentrifugation or light-scattering techniques is indispensable. A particular problem when working with scaffolding proteins is their high degree of molecular association and tight cellular regulation. As a result, overexpression of modified constructs often leads to a so-called dominant-negative phenotype, i.e. the overexpressed material interferes with endogenous protein thereby disturbing the system. As a result the protein-protein interaction leading to the observed phenotype might not be physiologically relevant and represent an experimental artefact.

ii.) Many structural proteins are only stable in their natural context, i.e. in the context of the filament, in association with binding partners or under specific physiological conditions. Thus, isolation and purification is only applicable to the most robust complexes. Crystallography requires milligram quantities of pure and
monodisperse material and heterologous expression systems are a key factor to efficiency. Post-translational modifications like for example phosphorylation, ubiquitination, SUMOylation or glycosylation lead to an increase in complexity of the system. Eukaryotic expression systems with their complex protein modification and folding machinery might be required to express the target construct in large amounts and functional form that is amenable to structural investigation. Furthermore, a key question is, whether all partners of the complex need to be over-expressed simultaneously or separate in vitro reconstitution of the complex is possible.

iii.) The fact that structural proteins rarely possess intrinsic activity, requires in depth biophysical characterization using for example spectroscopic methods like CD- or IR-spectroscopy or de novo design of functional assays to assess if the material is correctly folded and active under the in vitro conditions. In many cases the oligomeric state of the protein is directly linked to its functional activity and therefore of particular interest for biophysical characterization. Frequently during in vitro characterization, a disruption of the macromolecular assembly can occur during isolation steps, affinity purification and biophysical analysis due to unsuitable buffer conditions or mechanical forces. Thus, the experimental conditions need to be carefully established prior to the measurement and the results critically scrutinized. Self-association of scaffolding proteins can also be a major problem when working at high protein concentrations. Unspecific aggregation leading to non-specific heterogenous precipitate is frequently observed upon concentration of the protein sample.

iv.) Finally, elongated shape and high molecular weight leads to large unit cell dimensions and high solvent content, both parameters critically affecting the resolution of the diffracting experiment. Furthermore, observed oligomers present in the crystalline state might represent nonphysiological oligomeric states as a result of the specific aggregation events necessary for the crystallization process. To discriminate between crystal contacts and real physiologically relevant protein interfaces might not be trivial and requires additional complementary approaches employing for example NMR spectroscopy, database comparison employing the quaternary structure database or site directed mutagenesis in combination with functional assays.
1.6 Focus of the work

A particular focus of this work lies in the characterization of the protein complex between titin M-line domains A168-A170 and the ubiquitin-ligase MuRF-1. Currently no structural data are available of this interaction, which is believed to have crucial regulatory roles in myofibrillar protein turnover, sarcomere integrity and consequently in the pathological events leading to muscular atrophy. With the aim to reveal the molecular determinants of the recruitment of MuRF-1 to titin we have decided to conduct structural studies on titin domain A168-A170.

The first goal was to elucidate the crystal structure of the interface region. Distinct surface features as well as the overall domain orientation can be utilized to prompt further studies directed to block the interaction with MuRF-1 and potentially cause resistance to atrophy. The second goal was to characterize the binding site of titin A168-A170 on MuRF-1, which was reported at the time to include the B-box 2 and the adjacent predicted coiled-coil domain. Initial binding experiments between A160-A170 and MuRF-1 B-box carried out in our laboratory did not support an interaction of titin with isolated MuRF-1 B-box. Consequently, our studies are focussed on the predicted coiled-coil region, MuRF-1$^{146-341}$, with the aim to quantify the interaction by means of isothermal titration calorimetry. A168-A170 currently represents the largest poly-domain construct of titin that is characterized on the molecular level. It comprises an Ig-Ig as well as an Ig-FnIII transition that can serve as representative transitions and aid in homology modeling of other titin domains and their interfaces. We could not determine binding of MuRF-1 B-box to titin. Given the ill-characterized function of individual B-box domains together with the vicinity to the believed titin binding region of MuRF-1, we also aimed at solving the structure of MuRF-1 B-box. The frequent integration of the B-box 2 subtype N-terminal of coil-coil homo-oligomerization motifs that are involved in protein targeting should imply specific structural features that are unique to this B-box subtype.

We also investigated the nuclear scaffold protein LAP2α, which recruits various protein and DNA components during cell cycle progression and focused our attention in particular on the C-terminal lamin A/C binding functionality. Detailed biophysical characterization was directed on the determination of its homo-oligomerization state as
well as on the assessment of the crystallization propensity of this disease-related nuclear scaffold.
Chapter 2

Evidence for the recruitment of MuRF-1 to titin featured by the poly-domain structure of A168-A170

Michael Mrosek¹, Dietmar Labeit³, Heiko Heerklotz², Siegfried Labeit³, Olga Mayans¹*

¹Division of Structural Biology and ²Chemical Biophysics, Biozentrum, University of Basel, Klingelbergstr. 70, CH-4056 Basel, Switzerland; ³Institut für Anästhesiologie und Operative Intensivmedizin, Universitätsklinikum Mannheim, Mannheim 68167, Germany

Plasmids containing coding sequences for A168-A170 as well as MuRF-1 were provided by PD Dr. Dietmar Labeit and Prof. Siegfried Labeit. Isothermal titration calorimetry was done in collaboration with PD Dr. Heiko Heerklotz. All other work is my own.
2 Evidence for the recruitment of MuRF-1 to titin

Abstract

Titin forms a sarcomeric filament system in vertebrate striated muscle that has both elastic and signaling properties. Near its C-terminus and directly preceding a Ser/Thr kinase domain, titin contains a conserved pattern of Ig and FnIII modules (Ig\textsuperscript{A168}-Ig\textsuperscript{A169}-FnIII\textsuperscript{A170}, hereby A168-A170) that recruit the ubiquitin-ligase MuRF-1 to the filament. This interaction is believed to regulate myofibril turnover and the trophic state of muscle. We have elucidated the crystal structure of A168-A170 and established its binding to MuRF-1 in solution. A168-A170 shows an extended, rigid architecture where Ig-Ig and Ig-FnIII interfaces reveal conserved principles of interdomain arrangements across titin. Its surface displays a shallow groove along its full length as well as a unique loop protrusion, both features conceivably mediating MuRF-1 binding. ITC data show that A168-A170 binds MuRF-1 with high affinity. In MuRF-1, residues 166-315 are sufficient for this interaction. These data suggest that this region of titin is of interest to attempt therapeutic inhibition of MuRF-1-mediated muscle turnover.
Chapter 2: Evidence for the recruitment of MuRF-1 to titin

2.1 Materials and Methods

2.1.1 Cloning

The expression plasmid for A168-A170 was provided by Prof. Siegfried Labeit and PD Dr. Dietmar Labeit at the Institut für Anästhesiologie und Operative Intensivmedizin, Universitätsklinikum Mannheim, Germany. Coding sequences for domains A168-A170 from human titin (residues 24431-24731, EMBL X90568) had been cloned into the pETM-11 vector at restriction sites Nco-I and Mlu-I. pETM-11 (EMBL vector collection) is a variant of pET-24d (Novagen) including an N-terminal His6-tag and a TEV (tobacco etch virus) protease cleavage site prior to the inserted gene.

An initial plasmid containing the sequence of full-length MuRF-1 was provided by the collaborators mentioned above. Fragments from MuRF-1 (Swiss-Prot Q969Q1), hereby termed MuRF-1166-341 and MuRF-1166-315, were PCR-amplified using the following primer pairs:

MuRF-1166-315 forward: 5’-CGGGGTACCTTAATCTGTCCCAAAGTC-3’
MuRF-1166-315 reverse: 5’-CGGGGTACCTTACTGGTGTCCTTCTTCC-3’
MuRF-1166-341 forward: 5’-CATGCCATGGTGGTGGCGGGGAATG-3’
MuRF-1166-341 reverse: 5’-CGGGGTACCTTACTGGTGTCCTTCTTCC-3’.

Both constructs were inserted into pETM-11 via Nco-I and Kpn-I restriction sites and correspond to a C280S mutated variant designed to prevent unspecific aggregation of the samples due to oxidation during storage. Site-directed mutagenesis used the Qiagen protocol and the following primers:

forward 5’-GCTTCCAAGGGCTCCCCAGCTGGGGAAGACAGAGC-3’,
reverse 5’-GCTCTGTCTTCCCCAGCTGGGAGCCCTTGGAAGC-3’.

1.1.2 Protein production

Overexpression of A168-A170 was carried out in the E. coli strain BL21(DE3) Rosetta (Novagen). Cultures were grown at 37°C up to an OD600 of 0.6 in Luria Bertani medium supplemented with 25 µg/ml kanamycin and 34 µg/ml chloramphenicol. Expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cultures were grown at 25°C for approximately 18
additional hours. Cells were harvested by centrifugation at 2800* g and 4°C during 40 minutes. Bacterial pellet was resuspended in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, containing a protease inhibitor cocktail (Roche) and DNAse. Lysis was induced by sonication (Branson sonifier) for 2 minutes at 4°C in the presence of lysozyme. The homogenate was clarified by centrifugation at 15000* g at 4°C for 40 minutes. The supernatant was applied to a Ni²⁺-chelating HisTrap column (GE Healthcare) equilibrated in lysis buffer, with elution using 250 mM imidazole. The eluent was dialyzed against 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT in the presence of TEV protease. The protease digest, incubated for 16 h at 4°C and a 1:50 molar ratio TEV:protein, resulted in 90% efficient tag removal. Given that A168-A170 still interacted with the Ni²⁺-chelating resin after tag removal, the digested mixture was further purified by gel filtration on a Superdex 200 Hiload 16/60PG column (GE Healthcare) equilibrated in dialysis buffer. The samples were finally dialyzed against 25 mM MOPS pH 8.0, 150 mM KCl, 1 mM DTT, 1 mM EGTA and stored at 4°C for further use. It was observed that the presence of 20 mM NaCl in the medium was necessary to prevent protein precipitation. The expression yield of soluble titin A168- A170 in E. coli was approximately 10 mg of pure protein per 1 L culture. The protein purity was assessed by SDS-PAGE.

MuRF-1 fragments were prepared as above, but purification of the protease digest used subtractive affinity chromatography prior to gel filtration. The yield of both MuRF-1 constructs was approximately 5 mg of pure sample per 1 L culture for MuRF-1166-315 and 3 mg per 1 L culture for MuRF-1166-366.

All three proteins eluted in one single peak in gel filtration chromatography, indicating chemical homogeneity and a stable oligomeric state. Figure 2.1 shows sample preparations for all three constructs.
Chapter 2: Evidence for the recruitment of MuRF-1 to titin

Figure 2.1: Size exclusion chromatograms of titin A168-A170 and MuRF-1 constructs.

All runs were in 50 mM Tris pH 8.0, 100 mM NaCl and used untagged samples. The signal shown corresponds to A$_{280nm}$. a) Titin A168-170 (M$_w$=33746 Da); b) MuRF-1$_{166-341}$ (M$_w$=20429 Da); c) MuRF-1$_{166-315}$ (M$_w$=17293 Da). The final purity of samples according to SDS-PAGE is shown on the right. Lane M contains the LMW standard (GE Heathcare).
2.1.3 Crystallization of titin A168-A170

Initial crystallization conditions for A168-A170 were sought using protein sample supplied by Dr. Dietmar Labeit, Mannheim. This contained a non-cleavable C-terminal His-tag and was in 25 mM MOPS pH 8.0, 150 mM KCl, 1 mM DTT and 1 mM EGTA and at a concentration of 12 mg/ml (based on A280 values). Crystallization trials were performed using vapor diffusion, where 1 µl of titin A168-A170 was mixed with 1 µl of reservoir solution to form a 2 µl hanging drop. Each droplet was equilibrated against 500 µl of reservoir solution in VDX plates (Hampton Research). For initial screening, the sparse matrix kits Hampton Research CS-I & II, Emerald Wizard I & II, Emerald Wizard Cryo I & II, Molecular Dimension PSS-I & II were used.

Crystals grew in the form of thin, multiple plates with approximate dimensions of 100x40x10 µm³ after 4 days from a solution containing 10 % PEG 6000, 0.1 M HEPES pH 7.5 and 5 % MPD. Refinement of these conditions required varying PEG and MPD content followed by macroseeding with crystalline fragments. This yielded single crystals (Figure 2.2) suitable for diffraction tests using synchrotron radiation. Best crystals, grown from 7% PEG 6000, 0.1M HEPES pH 7.5 and 7% MPD, were frozen in mother liquor supplemented with 20% glycerol for cryoprotection and subjected to diffraction tests at beamline X06SA at the Swiss Light Source (SLS), Villigen, Switzerland.

Figure 2.2: Refined crystals of His-tagged titin A168-A170 after macroseeding.

a) Thin plates with approximate dimensions of 100x40x10 µm³ grew from 10 % PEG 6000, 0.1 M HEPES pH 7.5 and 5 % MPD after macroseeding with crystalline fragments; b) Close-up view of a).
The crystals diffracted anisotropically, with an approximate isotropic limit of 4 Å resolution (Figure 2.3). The diffraction pattern could be indexed in an orthorhombic cell (P222) with dimensions a=68.11 Å, b=97.66 Å, c=353.25 Å (α=β=γ=90°), where the c axis is exceptionally long. No further attempts were made to improve the diffraction quality of this crystal form since, at that point, crystals more amenable to structure elucidation were obtained from untagged A168-A170 samples.

Figure 2.3: Diffraction pattern of crystals from C-terminally His-tagged titin A168-A170.

Data were recorded at λ=1.00007 Å (12.3974 keV) on a MAR CCD165 detector exposing for 2 s. An oscillation range of 1° was used.

Further crystallization trials were conducted with an untagged version of titin A168-A170. Like the tagged sample, this was in 25 mM MOPS pH 8.0, 150 mM KCl, 1 mM DTT and 1 mM EGTA. The screening for initial crystallization conditions used sparse matrix preparations from commercial suppliers as described above and protein stocks concentrated to 12 mg/mL. This resulted in a significantly higher number of initial hits. Table 2.1 summarizes all conditions in which titin A168-A170 crystals were observed. Most crystalline formations exhibited needle or rod-like morphology, hexagonal in cross-section, and were inter-grown or associated in clusters.
Table 2.1: Crystallization conditions for titin A168-A170.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Solution content</th>
<th>Morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSI-35</td>
<td>0.8 M NaH₂PO₄, 0.8 M KH₂PO₄, 0.1 M Na-HEPES, pH 7.5</td>
<td>needles</td>
</tr>
<tr>
<td>CSI-48</td>
<td>2 M NH₄H₂PO₄, 0.1M Tris-HCl pH 8.5</td>
<td>needles</td>
</tr>
<tr>
<td>CSII-1</td>
<td>10% PEG 6000, 2M NaCl</td>
<td>hexagonal plates</td>
</tr>
<tr>
<td>CSII-23</td>
<td>1.6 M (NH₄)₂SO₄, 10% Dioxane, 0.1 M MES pH 6.5</td>
<td>needles</td>
</tr>
<tr>
<td>CSII-25</td>
<td>1.8 M (NH₄)₂SO₄, MES pH6.5, 0.01 CoCl₂</td>
<td>spherulites</td>
</tr>
<tr>
<td>CSII-32</td>
<td>1.6 M (NH₄)₂SO₄, 0.1M Na-HEPES 7.5, 0.1 M NaCl</td>
<td>needle cluster</td>
</tr>
<tr>
<td>CSII-42</td>
<td>1.5 M (NH₄)₂SO₄, 0.1 M Tris-HCl pH 8.5, 12 % Glycerol</td>
<td>spherulites</td>
</tr>
<tr>
<td>WizI-38</td>
<td>2.5 M NaCl, Na-Acetate pH 4.5, 0.2 M Li₂SO₄</td>
<td>needles</td>
</tr>
<tr>
<td>WizI-20</td>
<td>0.4 M NaH₂PO₄, 1.6 M KH₂PO₄, Imidazole pH 8.0, 0.2 M NaCl</td>
<td>needles, single</td>
</tr>
</tbody>
</table>

CS: Crystal screen (Hampton Research); Wiz: Wizard screen (Emerald biosciences). The condition shown in bold led to structure solution.

The most promising condition, Wizard-II-38 (Emerald Biosciences), resulted in crystal growth after two days and in the form of multiple, long needles. This condition was subsequently used as basis for a refined search by varying precipitant concentration, pH, salt type and salt concentration as well as employing various methods to slow down vapor diffusion. The latter included the addition of polyvinylpyrrolidone (PVP) and glycerol at various concentrations, as well as the covering of the reservoir with mixtures of silicon and paraffin oil (Chayen 1997) that led in cases to crystals of larger size and improved morphology. The modified condition that finally resulted in crystals suitable for data collection (Figure 2.4) contained 2.1 M NaCl, 0.6 M Li₂SO₄ and 0.1 M MES pH 5.5. Crystals grew over an extended period of time (2-3 months) to a final size of 600x50x50 μm³. It is likely that during this time desiccation of the drop facilitated the formation of large, single crystals. The yield of best crystals, however, was minimal and only two optimal specimens were obtained, both in one same drop (Figure 2.4b).

The crystals obtained from this single drop were used in the collection of native data as well as in a three-wavelength MAD experiment required for phasing. To perform the latter, a crystal was derivatized with the heavy atom cluster [Ta₆Br₁₂]²⁺ (kindly donated by PD Dr. Joerg Stetefeld), which interacts with proteins via ionic bonds and is a very strong scatterer (856 electrons) (Knabein et al., 1997). After soaking the crystals for 30 min in solutions containing the [Ta₆Br₁₂]²⁺ cluster at saturation, an intensive green color was observed indicative of the incorporation of the compound into the protein lattice. Crystals for native and derivative data acquisition were shock-frozen in liquid nitrogen after being cryo-protected in mother liquor supplemented with 20% glycerol. The cryo-protectant solution of the derivatized crystal contained in addition [Ta₆Br₁₂]²⁺ at saturation.
Chapter 2: Evidence for the recruitment of MuRF-1 to titin

2.1.4 Collection of a native data set

A native data set was collected from a best crystal of titin A168-170 at 100 K on the synchrotron beamline ID-29-1, ESRF (Grenoble), which is equipped with a Quantum ADSC Q210 detector. Data were collected in a series of 1° non-overlapping oscillations and processed up to 2.9 Å resolution using the XDS/XSCALE package (Kabsch, 1993). The lattice was identified as P3\(_{1}(x)21\) (x=1 or x=2) with unit cell dimensions a=b=125.16 Å, c= 134.54 Å (\(\alpha=\beta=90^\circ, \gamma=120^\circ\)). A diffraction pattern is shown in Figure 2.5 and data processing statistics are given in Table 2.2.

**Table 2.2: Native data processing statistics.**

<table>
<thead>
<tr>
<th>Xray source</th>
<th>ID29-1 (ESRF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detector</td>
<td>ADSC Quantum Q210</td>
</tr>
<tr>
<td>Wavelength</td>
<td>0.9792</td>
</tr>
<tr>
<td>Resolution</td>
<td>15-2.9 (3.0-2.9)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>26792 (2619)</td>
</tr>
<tr>
<td>R-factor</td>
<td>7.3 (52.8)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>5.4 (5.4)</td>
</tr>
<tr>
<td>Completeness</td>
<td>97.5 (98.6)</td>
</tr>
<tr>
<td>I/(\sigma) (I)</td>
<td>21.4 (3.6)</td>
</tr>
</tbody>
</table>
Figure 2.5: Diffraction pattern of titin A168-170 crystals.

Diffraction was observed up to a maximum of 2.7 Å resolution, but complete data of quality suitable for structure elucidation could only be obtained to 2.9 Å resolution. Images displayed with the program MOSFLM (Leslie, 1992).

2.1.5 Detection of non-crystallographic symmetry (NCS)

The analysis of the Mathews coefficient (Matthews, 1968) revealed that the asymmetric unit of this crystal form must contain two to four copies of titin A168-A170 ($V_M=4.5\ \text{Å}^3/\text{Da}$ for two molecules per asymmetric unit leading to a solvent content of 72% and $V_M=2.2\ \text{Å}^3/\text{Da}$ for four molecules per asymmetric unit leading to a solvent content of 46%). In the subsequent steps of structure elucidation it became clear that only two molecules are present in the asymmetric unit of this crystal form, resulting in an elevated solvent content. This possibly explaining the limited resolution of the diffraction obtained from these crystals.

For the detection of the rotational component of the non-crystallographic symmetry (NCS) axis a self-rotation function was calculated using the program, AMORE (Navaza, 1994). The method correlates sets of self-vectors from two Patterson maps. The solutions are given in Table 2.3.
Chapter 2: Evidence for the recruitment of MuRF-1 to titin

Table 2.3: Euler and Polar rotational NCS relations as determined using AMORE.

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>ω</th>
<th>φ</th>
<th>κ</th>
<th>Correlation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOLUTIONRS</td>
<td>120.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>120.0</td>
<td>100 (crystallographic symmetry)</td>
</tr>
<tr>
<td>SOLUTIONRS</td>
<td>180.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>180.0</td>
<td>60</td>
</tr>
<tr>
<td>SOLUTIONRS</td>
<td>60.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>60.0</td>
<td>60</td>
</tr>
</tbody>
</table>

The self-rotation function was calculated between a resolution of 20 and 5 Å.

![Figure 2.6: Graphical representation of the self-rotation function calculated using POLARFN](image)

Stereographic projection of section \( \kappa = 180^\circ \). The presence of NCS is shown by additional 2-fold axes indicated by red arrows. The central peak indicates an NCS axis co-aligned with the crystallographic three-fold along \( c \).

A two-fold (\( \kappa = 180^\circ \)) NCS axis was found to be co-aligned with the crystallographic three-fold axis along \( c \) (\( \omega = \phi = 0 \)) (Figure 2.6). Peaks at \( \phi = 30^\circ \) and \( \phi = 90^\circ \) (red arrows in Figure 2.6) result from the fact that the NCS axis is perpendicular to the \( ab \) plane, thus generating additional two-fold NCS axes perpendicular to the crystallographic axes \( a \) and \( b \) and indicating that the NCS axis must intersect \( a \) and \( b \) at their origin. Thus, the NCS axis generates a pseudo hexagonal symmetry as indicated by peaks in section \( \kappa = 60^\circ \) resembling a P622 point group symmetry.

The position of the NCS axis in real-space was confirmed at a later stage by analyzing initial electron density maps with the program GETAX (Vonrhein & Schulz, 1999). This showed that the NCS axis passed through the unit cell origin, perpendicular to the \( ab \) plane (\( a = b = 0 \)) and parallel to \( c \) (Figure 2.7). Given that self-rotation correlation values for this NCS axis were 60% at low resolution, this axis could be expected not to be a perfect two-fold. In effect, further steps in the structure elucidation process revealed that slightly divergent NCS operators related individual domain pairs across molecular copies A and B. Thus, the two molecular copies do not exhibit...
identical overall conformations, with their C-terminal domains A170 diverging most pronouncedly from a two-fold relation.

Figure 2.7: Real-space analysis of non-crystallographic symmetry using GETAX.

a) Density correlation search at 15-5 Å resolution and using a sphere radius of 20 Å. The position of the NCS axis is indicated by accumulation of density along c (black). Shown are the molecular copies A (green and red) and B (blue and magenta) of the final model of titin A168-170. Domain A170 is coloured red and magenta in copies A and B, respectively; b) Unit cell diagram showing the location of crystallographic three-fold axes and “pseudo” two-fold NCS axes (marked by red and hollow obloids). Red obloids show the locations were the local two-fold correlation calculated using GETAX is higher as a result of the contribution from domains A168-A169 that diverge only slightly from a two-fold symmetry. The hollow obloid represents a symmetrical copy of the NCS axis that displayed no detectable correlation in GETAX (see a.) due to the local influence of domains A170 that diverge more notably from a two-fold symmetry. Red arrows indicate additional 2-folds leading to the hexagonal "pseudo"symmetry in the rotation function.

2.1.6 Phasing attempts by Molecular Replacement

Titin A168-170 is composed of two Ig and one FnIII domain, which are very abundant folds in nature. Given the structural conservation of these folds, structure solution by molecular replacement was attempted using the program PHASER (Storoni et al., 2004). The program carries out a double three-dimensional search (three rotations, three translations) to maximize the overlap of target and model self-vectors based on maximum-likelihood calculations.

Structures of homologous domains were identified in the Protein DataBase by sequence alignment with BLAST (Altschul et al., 1990). Three structures were chosen
based on sequence identity levels (Figure 2.8), namely the crystal structure of telokin at 2.0 Å resolution (1FHG, 30% sequence identity to titin Ig\textsuperscript{A168}; Holden et al., 1992), an NMR structure of Ig domain I18 from twitchin (1WIT, 30% sequence identity to titin Ig\textsuperscript{A169}; Fong et al., 1996) and the NMR structure of A71 from the titin A-band (1BPV, 40% identity to titin Fn\textsuperscript{A170}, Muhle-Goll et al., 1998). All models were trimmed to fit the side chain composition of titin A168-170.

![Sequence alignment of domain homologues used in molecular replacement.](image)

**Figure 2.8:** Sequence alignment of domain homologues used in molecular replacement.

The crystal structure of 1FHG (Telokin, 30% identity) and the solution structures of 1WIT (I18 from twitchin, 30% identity) and 1BPV (A71 from titin, 40% identity) were trimmed to match the side-chain composition of titin A168-170 in volume. Trimmed residues are indicated in cyan, identical residues in grey.

Multiple trials were carried out in PHASER considering space groups P3\textsubscript{1}21 and P3\textsubscript{2}12, different number of molecules per asymmetric unit and variations in input models. However, this did not result in any potential solutions. The failure might be possibly due to the decrease of diffraction data quality at medium resolution together with the use of NMR structures as search models. Next, MAD experimental phasing was pursued.

### 2.1.7 MAD data collection

MAD data were collected at the beamline ID23-1, ESRF (Grenoble) on crystals derivatized with the electron-rich compound [Ta\textsubscript{6}Br\textsubscript{12}]\textsuperscript{3+} (for details on the derivatization process see chapter 2.1.3). A three wavelength MAD experiment was carried out at the L\textsubscript{III}-edge of tantalum, using energies of E\textsubscript{p} = 9.88256 keV (\(\lambda\text{_{peak}} = 1.2546 \, \text{Å}\)) for the peak corresponding to the maximal f'; E\textsubscript{p} = 9.87887 keV (\(\lambda\text{_{reflection}} = 1.2550 \, \text{Å}\)) for an edge point representing the minimal f' and E\textsubscript{p} = 12.300 keV (\(\lambda\text{_{remote}} = 1.008 \, \text{Å}\)) for a high-energy remote point selected outside the L-edge (Figure 2.9). The data were collected in the order peak, inflexion and remote wavelengths aiming at maximum completeness and
redundancy of the Bijvoet pairs. A total rotation of 120° was recorded per wavelength in 1° oscillations and two seconds exposure time per frame. Data collection was carried out at a maximum “edge-of-detector” resolution of 3.2 Å.

Figure 2.9: Real ($f'$) and imaginary ($f''$) components of the anomalous signal from tantalum.

Theoretical plot of $f'$ and $f''$ versus energy. For MAD data collection, wavelengths were selected experimentally based on a fluorescence scan (data not shown). The individual wavelengths used for data collection are indicated.

2.1.8 Data processing

Data were processed using XDS (Kabsch et al., 1993) by treating the Bijvoet pairs $F'^+$ and $F'^-$ as separate reflections. Unmerged data collected at edge and remote energies were scaled against the peak set as reference wavelength using XSCALE. For derivative-to-native scaling the program FHSCAL (Kraut et al., 1962) was used applying Kraut’s scaling procedure. Data processing statistics are provided in Table 2.4.

Table 2.4: MAD data processing statistics on [Ta$_6$Br$_{12}$]^{2+}-derivatized titin A168-170

<table>
<thead>
<tr>
<th>X-ray source</th>
<th>ESRF-ID-23-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detector</td>
<td>MAR CCD 225</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>20-3.35 (3.4-3.35)</th>
<th>15-3.45 (3.5-3.45)</th>
<th>15-3.5 (3.55-3.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.2546 (9.88256 keV)</td>
<td>1.2550 (9.87887 keV)</td>
<td>1.008 (12.300 keV)</td>
</tr>
<tr>
<td>No. Bijvoet reflections</td>
<td>33994 (1556)</td>
<td>31059 (1347)</td>
<td>29580 (1216)</td>
</tr>
<tr>
<td>$R_{sym}$ (%)</td>
<td>11.1 (42.6)</td>
<td>10.6 (41.6)</td>
<td>12.9 (47.5)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.9 (3.9)</td>
<td>3.7 (3.7)</td>
<td>3.9 (3.9)</td>
</tr>
<tr>
<td>Completeness</td>
<td>98.8 (99.7)</td>
<td>98.2 (99.4)</td>
<td>98.2 (98.9)</td>
</tr>
<tr>
<td>$I/\sigma$ (I) (%)</td>
<td>9.8 (3.5)</td>
<td>10.1 (3.5)</td>
<td>9.2 (3.1)</td>
</tr>
</tbody>
</table>
The anomalous scattering contribution of \([\text{Ta}_6\text{Br}_{12}]^{2+}\) clusters was initially estimated in XDS, as an indication of whether the compound had become specifically bound to the protein (Table 2.5). The cluster is a regular octahedron consisting of six metal atoms with 12 bridging bromine atoms along the 12 edges of the octahedron. Its radius is about 4.3 Å. At low resolution (approximately <6 Å) all atoms of the cluster scatter in phase and the cluster can be treated as a super-atom (Knäblein et al., 1997).

<table>
<thead>
<tr>
<th>Resolution</th>
<th>S_norm/S_ano (Peak)</th>
<th>Remote</th>
<th>Inflex</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.00</td>
<td>2.31</td>
<td>1.55</td>
<td>2.00</td>
</tr>
<tr>
<td>13.00</td>
<td>1.99</td>
<td>1.26</td>
<td>1.51</td>
</tr>
<tr>
<td>12.00</td>
<td>1.51</td>
<td>1.13</td>
<td>1.17</td>
</tr>
<tr>
<td>10.00</td>
<td>1.48</td>
<td>1.08</td>
<td>1.18</td>
</tr>
<tr>
<td>6.00</td>
<td>1.20</td>
<td>1.02</td>
<td>1.06</td>
</tr>
<tr>
<td>5.00</td>
<td>1.03</td>
<td>0.99</td>
<td>1.01</td>
</tr>
</tbody>
</table>

S_norm: mean value of Sigma(I) for acentric reflections assuming Friedel’s law is valid.
S_ano: mean value of Sigma(I) for acentric reflections assuming Friedel’s law is violated.
Anomalous scattering contributions to the intensities are indicated by S_norm/S_ano > 1

2.1.9 Determination of the sub-structure of anomalous scatterers

In order to obtain experimental phases, the position of anomalous scatterers in the unit cell was determined using the program CNS (Brunger et al., 1998) on anomalous difference Patterson maps calculated between 12-6.0 Å resolution and derived from peak-wavelength data (Figure 2.10a). The features of this map were consistent with those of isomorphous difference maps (Figure 2.10b) and anomalous difference Patterson maps derived from high-energy remote data. The clear predominance of peaks at sections \(w=0.33\) and \(w=0.66\) indicated that these were Harker sections and, thus, that a screw symmetry axis was present along \(c\). This confirmed the space group to be either \(P3_121\) or \(P3_212\), as previously deducted from the systematic absence of diffraction indices. Two anomalous scatterers were located (Figure 2.10a). Their positions were first refined in Patterson space using both CNS and VECREF (CCP4) and plotted onto the Patterson maps with VECTOR (CCP4) for visual inspection. Phases were then calculated and refined in SHARP (De la Fortelle and Bricogne, 1997) using native data up to 3.0 Å resolution and derivative data to 4.5 Å resolution. The refined positions of the heavy atom substructure are listed in Table 2.6. As a cross-validation, an iterative approach was followed were phases were calculated from one single scatterer (in this
case the primary site, site 1 in Table 2.6) and the location of the second scatterer confirmed through anomalous and isomorphous difference Fourier maps. Neither the analysis of difference Patterson or Fourier maps revealed any additional sites for $[\text{Ta}_6\text{Br}_{12}]^{2+}$ scatterers.

![Figure 2.10](image.png)

**Figure 2.10: Anomalous and isomorphous difference Patterson map at Harker section z=0.33.**

**a)** Anomalous difference Patterson map of the peak wavelength using data between 15 and 6 Å. Contouring starts at 2σ, with contour increments representing 0.5σ. The numbering is as Table 2.6. **b)** Isomorphous difference Patterson map between inflexion and the native Contouring starts at 1σ, with contour increments representing 0.5σ.

Given the ambiguity in space group symmetry remaining at that point (P3121 and P321 enantiomeric lattices) and the intrinsic handness uncertainty of the Patterson map, both direct and inverted positions of heavy atom clusters where assayed in phasing. Visual inspection of the resulting Fourier maps clearly revealed this crystal form to belong to the space group P3121. For better assessment of phase quality, the phasing statistics prior to density modification are summarized in Table 2.7. Phasing based on the two cluster sites yielded electron density maps, which had clearly identifiable solvent-protein regions and partially interpretable features upon solvent flattening using the SOLOMON procedure (Abrahams and Leslie, 1996) implemented in SHARP and assuming a solvent content of 70%.

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>0.5131</td>
<td>0.6076</td>
<td>0.011</td>
</tr>
<tr>
<td>Site 2</td>
<td>0.1490</td>
<td>0.1710</td>
<td>0.0105</td>
</tr>
</tbody>
</table>
Table 2.7: Phasing statistics prior to density modification calculated using SHARP.

<table>
<thead>
<tr>
<th></th>
<th>Peak</th>
<th>Inflexion</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phasing power (iso/ano)</td>
<td>1.46/1.58</td>
<td>1.34/0.96</td>
<td>1.13/0.87</td>
</tr>
<tr>
<td>R-cullis (iso/ano)</td>
<td>0.66/0.71</td>
<td>0.67/0.86</td>
<td>0.71/0.86</td>
</tr>
<tr>
<td>Figure of Merit</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(resolution < 4.5 Å)

2.1.10 Density modification

Skeletonization and subsequent skeleton editing of the initial electron density map calculated using SHARP allowed determining the orientation and position of the individual domains FnIII\textsuperscript{A170} and Ig\textsuperscript{A169}. The domain Ig\textsuperscript{A168} could not be located at that time, but its position was estimated based on that of the neighboring domain Ig\textsuperscript{A169}. In order to improve the quality of the electron density maps, density modification techniques were applied by imposing theoretical constraints to the density in real space and by calculating a combined and improved phase estimate (Cowtan and Zhang, 1999). Two density modification approaches have been particularly helpful in solving the structure of titin A168-A170: solvent flattening and NCS averaging. Solvent flattening benefited from the high solvent content in these crystals and used Wang’s method for the automatic calculation of a solvent mask (Wang, 1985). NCS averaging exploited the presence of local symmetry in the asymmetric unit. As a complement, histogram matching was applied in conjunction with these approaches. This adjusted the experimental protein density distribution to that expected for proteins at this resolution (Zhang and Main, 1990).

Density modification was carried out with the program DM using initial phases computed with SHARP (Cowtan, 1994). As a pre-requirement for NCS averaging, operators for individual domains where calculated using the least square superposition routine of the program O (Jones, 1991) and three NCS masks covering the individual domains were created in MAMA (Uppsala Software Suite). The NCS-operators were refined based on electron density correlation using IMP (Uppsala Software Suite). Figure 2.11 shows a comparative cross-section of the electron density obtained by solvent treatment within SHARP and after density modification in DM (solvent flattening, NCS averaging, histogram matching and phase extension to 3.0 Å).
Figure 2.11: Projection of z-sections of electron density maps contoured at 1 σ.

a) Electron density map after solvent flipping in SOLOMON (SHARP). A solvent-protein boundary is detectable.

b) DM map after solvent flattening, histogram matching and NCS averaging. After the density modification steps, a Figure of merit of 57 % (resolution shell 10-3 Å) was achieved. Both projections include 50 sections along the z-axis. Maps were displayed with mapslicer (CCP4).

2.1.11 Model building and refinement

As a first step towards model building, poly-alanine variants of homologous structures were manually positioned in skeletonized electron density maps (homologues are described in chapter 2.1.6) and progressively edited in iterative cycles of model building and refinement. The structure was refined against native data between 15-2.9 Å resolution in CNS (Brunger et al., 1998). The R-free was used as a cross-validation indicator (Brunger 1993), for which reflections were partitioned into a working and a free set using FREERFLAG (CCP4). Model refinement included overall anisotropic B-factor scaling, bulk solvent correction, conjugate gradient minimization of atomic coordinates and grouped B-factor refinement with two B-factor values per residue (one for the main chain and one for the side chain). Tight NCS restraints were applied during refinement, where a weight of 500 kcal/(mol*Å^2) was applied to geometrical restraints and B-factor values were restrained to two sigma differences. For the refinement of initial models, a maximum-likelihood target function taking experimental phases into
account was employed – such phases were excluded at later stages. Solvent atoms were identified with the water-pick routine of CNS and visually validated in O. In the last cycles of refinement, NCS restraints were decreased to a weight of 150 kcal/mol*A² and excluded residues 142-146 (Ig A169), which were close to a lattice contact point and adopted different conformations in copies A and B. The final model contained 292 protein residues per chain and 45 solvent molecules. 12 C-terminal residues were not identified as they were disordered in both NCS copies, while two residues in –1 and –2 position, remnants of the TEV protease cleavage sequence, were modeled. Figure 2.12 shows a progression of the improvement in electron density maps during refinement. The Ramachandran plot (Ramachandran et al., 1963) is given in Figure 2.13 and final refinement statistics are summarized in Table 2.8.

Figure 2.12: Improvement of the quality of electron density maps during structure solution.

a) Electron density map calculated using experimental MAD phases prior to solvent flattening
b) Solvent flattened electron density map after SOLOMON.
c) Electron density map after density modification in DM.
d) (2Fcalc–Fobs)αcalc map of the refined model.
All maps are contoured at 1 σ and the final model is shown.
Chapter 2: Evidence for the recruitment of MuRF-1 to titin

Table 2.8: Final Refinement statistics of titin A168-A170.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of reflections in working/free set</td>
<td>25645 / 1147</td>
<td></td>
</tr>
<tr>
<td>Number of protein residues / solvent molecules</td>
<td>584 / 45</td>
<td></td>
</tr>
<tr>
<td>R-factor / R-free (%)</td>
<td>21.8 / 27.6</td>
<td></td>
</tr>
<tr>
<td>rmsd bond length (Å) / bond angle (°)</td>
<td>0.010 / 1.59</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.13: Ramachandran diagram of the two NCS-related copies of titin A168-A170.
Plot generated using PROCHECK (Laskowski et al, 1993).

2.1.12 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry is frequently used for the characterization of macromolecular interactions in solution. It does not require any chemical modification or immobilization of the interacting components. The released or absorbed heat upon binding provides the basis for the determination of the binding constant (K_a), the standard enthalpy/entropy changes (ΔH_0/ΔS_0) in the binding reaction and the stoichiometry of the reaction (Doyle, 1997).

In this study all data were recorded with the power-compensation calorimeter VP-ITC (Microcal) at 25°C. The device consists of a reference cell (filled with buffer) and a sample cell (filled with the macromolecule) both situated in an adiabatic jacket. The instrument measures the electrical power, which is required to maintain the same temperature in the sample cell as in the reference cell upon titration with the ligand from a syringe.

All samples were previously dialyzed against 50 mM Tris-HCl pH 8.0, 20 mM NaCl. Titrations consisted of 5-7 µl injections of A168-A170 concentrated to 1.4 mg/ml into MuRF-1 166-341 or MuRF-1 166-315 solutions at 1 mg/ml and 0.7 mg/ml, respectively (as
determined by BCA assay; Brown, 1989). The time interval between injections was 200-300 s. The data were corrected for the heat of dilution by subtraction of the small constant heat obtained at the end of the titration, where no further binding occurs. All measurements were done in collaboration with PD Dr. Heiko Heerklotz (Biozentrum Basel).
2.2 Results

2.2.1 Structure of A168-A170

The structure of A168-A170 from human titin has been elucidated at 2.9 Å resolution using X-ray crystallography. A168-A170 comprises two Ig domains (IgA168-IgA169) and one FnIII (FnA170). The molecule adopts an extended conformation with an end-to-end distance of ~115 Å, where the two N-terminal Ig lie almost perfectly coaxial, but the FnIII domain is bent away from the molecular axis (Figure 2.14). The crystal form used in this study contains two A168-A170 copies in its asymmetric unit.

Figure 2.14: Crystal structure of A168-A170.

Ribbon representation. The insertion loop in IgA169 is displayed in cyan. The inset shows the molecular conformation schematically.

They are essentially identical in conformation (rmsd 0.68 Å for 288 Ca atoms of the whole molecule, calculated with SPDBV; Guex and Peitsch, 1997), suggesting that this fragment of titin has a well-defined long-range order and a high degree of stiffness.

Domains IgA168 and IgA169 belong to the I (intermediate)-set of Ig folds (Harpaz and Chothia, 1994). They share 20% sequence identity and a high structural similarity (rmsd 0.99 Å for 82 matching Ca atoms, SPDBV). Both belong to the “N-conserved” type of
Ig from titin (Marino et al., 2005), characterized by a N-terminal loop cluster comprising proline residues from β-strand A, a PxP motif in the BC loop and an extended FG β-hairpin hosting an NxxG sequence.

This Ig type predominates in the Z-disc, the skeletal I-band and the A-band fractions of titin, but not in its constitutive cardiac I-band. Domains IgA168 and FnA170 show no significant deviation from other equivalent modules across titin, as reflected by a structure-based sequence alignment of single domains of titin with known structure (Figure 2.15).

Figure 2.15: Structure-based sequence alignment of titin Ig and FnIII of known structure.

β-Strand composition is shown in yellow, residues at Ig-Ig and Ig-FnIII interfaces are in red and cyan, respectively. Features characteristic of “N-conserved” Ig (A168, A169, I1, Z1, Z2 and M5) are given in blue. I91 was formerly I27. Strand nomenclature for Ig as in Harpaz and Chothia, 1994, and for FnIII as in Leahy et al., 1992. The distinct loop of IgA169 is shown in grey.

IgA169, however, shows a unique 9-residue loop protrusion between β-strands A and A’ with sequence PKTLEGMA and adopting in part an α-helical conformation (Figure 2.15 and Figure 2.16). A loop insertion at this position is not detectable in any other Ig of titin.

2.2.2 Domain interfaces

The domains of A168-A170 form a tight tandem. They are connected through one-residue linkers in extended conformation that effectively result in the continuation of structural elements from one domain into the next. The extended arrangement of the molecule is characterized by domain interfaces with unusually small buried areas (Table 2.9) and engaging a minimal number of residues in each domain. Both Ig-Ig and Ig-FnIII interfaces are void of specific interactions and only involve small hydrophobic clusters.
Superimposition of Ig domains from titin for which a structural model is available, namely IgA169 (cyan), IgA168 (green), I1 (magenta; PDB accession code 1G1C), M5 (yellow; 1NCT), Z1 (salmon; 2A38) and Z2 (grey; 2A38). The unique loop insertion of IgA169 is clearly identifiable. The Cα atoms of its component residues are displayed as solid spheres and the sequence given. I91 (formerly I27), which does not belong to this Ig subgroup and has shorter loop clusters, is excluded from the superimposition to ease visualization. As the other Ig, I91 does not comprise loop insertions in its N-terminal β-strand region.

Table 2.9: Domain arrangement in poly-domain fragments from titin.

<table>
<thead>
<tr>
<th>Domain Arrangement</th>
<th>Opening anglea</th>
<th>Torsionb</th>
<th>Distancec</th>
<th>Buried surface area (Å²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgA168IgA169</td>
<td>176°</td>
<td>+89°</td>
<td>43.3 Å</td>
<td>212 (~1.4%)</td>
</tr>
<tr>
<td>IgA169FnA170</td>
<td>127°</td>
<td>--</td>
<td>40.2 Å</td>
<td>308 (~2.0%)</td>
</tr>
<tr>
<td>IgZ1IgZ2XTALd</td>
<td>146°</td>
<td>+75°</td>
<td>49.4 Å</td>
<td>160 (~1.4%)</td>
</tr>
<tr>
<td>IgZ1IgZ2RDC/SAXSd</td>
<td>136°</td>
<td>+85°</td>
<td>44.5 Å</td>
<td>165 (~1.4%)</td>
</tr>
<tr>
<td>IgZ1IgZ2Tele</td>
<td>167°</td>
<td>+48°</td>
<td>52.4 Å</td>
<td>235 (~2.0%)</td>
</tr>
</tbody>
</table>

a Angle defined by the primary axes of inertia (longitudinal) of consecutive domains; b Angle defined by the projection of secondary axes of inertia (cross-sectional) of consecutive domains onto a common plane; c Distance between centers of mass of individual domains; d Crystallographic coordinates for free Z1Z2 (IgZ1IgZ2XTAL) derive from PDB entry 2A38 (Marino et al., 2006); domain arrangement of Z1Z2 in solution was calculated from SAXS data and NMR residual dipolar couplings (IgZ1IgZ2RDC/SAXS) as reported (Marino et al., 2006) and crystal coordinates of Z1Z2 complexed to telethonin (IgZ1IgZ2Tele) derive from 1AY5 (Zou et al., 2006).
Domains Ig\textsuperscript{A168}\cdot Ig\textsuperscript{A169} interact through hydrophobic groups that surround the linker strand dorsally and ventrally (Figure 2.17a). Given their coaxial orientation, the interacting residues originate from the linker region (V92, aliphatic portion of E93, V94) and the C-terminus of β-strand A’ (aliphatic portion of R13, Y14) in the preceding domain Ig\textsuperscript{A168} and the conserved elements of this Ig type in the following Ig\textsuperscript{A169}, namely the FG turn (F180) and the PxP motif of the BC loop (P128, K127) (Figure 2.15). A comparison of Ig\textsuperscript{A168}\cdot Ig\textsuperscript{A169} and Z1Z2 (the only other Ig doublet of titin structurally characterized) (Zou et al., 2006; Marino et al., 2006) reveals certain similarities.

Figure 2.17: Domain interfaces in the titin A168-A170 structure.

a) Ig\textsuperscript{A168}\cdot Ig\textsuperscript{A169} and b) Ig\textsuperscript{A169}\cdot Fn\textsuperscript{A170} interfaces. Domain color code as in Figure 2.14. Hydrophobic contributions are shown in green. Selected van der Waals surfaces are displayed. The C\textalpha-atoms of residues in the NxG motif of the FG β-turns of both Ig\textsuperscript{A169} and Fn\textsuperscript{A170} are shown as spheres. Hydrogen bonds are indicated by dotted lines.

Also Z1Z2 exhibit an extended conformation, lack specific interdomain contacts and include a small hydrophobic cluster between modules. As that of Ig\textsuperscript{A168}\cdot Ig\textsuperscript{A169}, the cluster of Z1Z2 comprises residues C-terminal to β-strand A’, the BC loop and linker residues (Figure 2.15). Despite, Z1Z2 displays a different domain orientation (Table 2.9) and it has a longer, three-residue linker that allows a moderate modular dynamics (Marino et al., 2006). It can then be concluded that the mere presence of elements characteristic of the “N-conserved” Ig type does not determine domain orientations in these doublets. Since the identified interface residues are not conserved in Ig across titin (Witt et al., 1998), it cannot be predicted how closely the observed conformations represent those of other tandems. Yet, the overall structural principles of
Ig arraying revealed by these studies are likely to be generic to the titin filament (Marino et al., 2005).

The interface of domains Ig\textsuperscript{A169}-Fn\textsuperscript{A170} presents a hydrophobic component as that of Ig\textsuperscript{A168}-Ig\textsuperscript{A169} (Figure 2.17b). Similarly, contacts involve β-strand A’ (H111) in the initial domain Ig\textsuperscript{A169} and the NxxG motif of β-turn FG in Fn\textsuperscript{A170} (K272, F273). However, in contrast to Ig, the FG turns in FnIII of titin are highly conserved in sequence (alignment used CLUSTALW; Higgins et al., 1994). FnIII following an Ig domain within A-band repeats consistently host a conserved residue pair - a charged group followed by an aromatic (Y or F) - in the variable positions of their NxxG motif (Figure 2.18). Such segregation is not observed in the 7-domain super-repeats of the preceding D-zone.

In the current crystal structure, the lateral amino group of K272 is not involved in interactions, suggesting its availability for binding to other sarcomeric proteins. Myosin-binding protein C (MyBP-C) binds to titin through the first Ig domain of every super-repeat at the C-zone, thereby possibly aiding the regular polymerization of myosin filaments (Freiburg and Gautel, 1996). Given that the FG β-turn of Fn\textsuperscript{A170} is located towards the inter-modular space, its charge might be speculated to contribute to MyBP-C docking. Interestingly, the FG turns of FnIII domains at other positions within repeats of the D- and C-zones also exhibit certain conservation of their sequence motifs (Figure 2.18), suggesting that these might be generically important for FnIII-FnIII arrangements and/or interactions to other A-band components.
Figure 2.18: Ig-FnIII domain interfaces in the titin A-band.

Ig and FnIII domains are shown as filled boxes, where Ig domains flanked by FnIII are in brown, FnIII following an N-terminal Ig in grey, other FnIII in white and Ig followed by another C-terminal Ig in blue. Fn\textsuperscript{A71}, the only other FnIII of titin with known structure (1BPV), is in cyan. The two residues in variable positions of the NxxG b-turn of FnIII domains are displayed. Domain super-repeats are indicated. The pattern of alternating charges is displayed, where red and blue indicate negative and positive charges, respectively.
2.2.3 MuRF-1 binding

The surface of A168-A170 reveals a shallow groove at the concave side of the molecule that spans its full length (Figure 2.19a). Based on SPOTS blots of MuRF-1 peptides, the latter has been proposed to bind titin through a sequence just prior to its acidic C-terminal tail (Figure 2.20, Witt et al., 2005).

![Figure 2.19: Surface features of titin A168-A170 and model for MuRF-1 interaction.](image)

According to our own analysis (Jpred software; Cuff et al., 1998), the C-terminal half of MuRF-1 (including the proposed binding sequence) has tendency to $\alpha$-helical formation, with its middle section predictably following a coiled-coil association (Figure 2.20) (COILS; Lupas et al., 1991).

Both manual examination and docking prediction software (PatchDock; Schneidman et al., 2005) indicate that the dimensions and geometry of the surface groove in A168-A170 can well accommodate an $\alpha$-helix along its length. In particular, a best docking is obtained if the helix exhibits a long-range bending of its axis as that of components of a dimeric coiled-coil (Figure 2.19b), like human lamin 2B (Strelkov et
This suggests that MuRF-1 interacts with titin through an extensive contact area, involving more than one module. This is in agreement with biochemical data that show how domains Ig$^{A168}$, Ig$^{A169}$ and Fn$^{A170}$ in isolation are not able to secure MuRF-1 binding (Center et al., 2001). Ig$^{A169}$ might, however, be central to the interaction. Its unique loop insertion, located towards the middle point of the groove (Figure 2.19a and Figure 2.19b), could be speculated to play a role in MuRF-1 binding or function.

We have confirmed the interaction between A168-A170 and the helical fraction of MuRF-1 using soluble, recombinant samples. This study used two N-terminally truncated MuRF-1 constructs, MuRF-1$^{166-341}$ and MuRF-1$^{166-315}$ (Figure 2.20), where the latter lacks in addition the acidic C-terminal tail. Binding monitored by isothermal calorimetry (ITC) showed that both MuRF-1 variants interacted strongly with A168-A170, yielding heat release values of 14 and 28 kcal/mol and $K_d$ affinities of 35 and 37 nM, respectively (Figure 2.21). The finding that saturation occurs at a molar ratio of $\sim$0.1 indicates that the MuRF-1 species were only partially active (possibly due to aggregation) and/or that two or more MuRF-1 molecules bind one A168-A170. Since $K_d$ values are similar for both constructs, it can be concluded that the acidic tail of MuRF-1 does not influence binding.
a.

![Figure 2.21: ITC data on MuRF-1:A168-A170 binding.

a) Experimental pattern of injections for MuRF-1_{106-341}; b) Heats of binding for titrations of A168-A170 into MuRF-1_{106-341} (filled circles) and MuRF-1_{116-315} (open triangles) recorded in 50 mM Tris-HCl pH 8.0, 20 mM NaCl at 25°C. Both curves show the same $K_d$ and $\Delta H$ and differ only slightly in the active fraction of MuRF-1.]

2.3 Discussion

The muscle filament titin, despite its colossal dimensions, has a simple and repetitive architecture consisting of linear tandems of Ig and FnIII modules. The structure of A168-A170 reveals the molecular details of Ig-Ig and Ig-FnIII interfaces, whose global features are likely to be generic to domain arrangements across titin (Marino et al, 2005). A168-A170 has a pronouncedly extended conformation and exhibits a well-defined long-range order with limited conformational freedom. The latter seems to result from extremely short linker sequences and the presence of small hydrophobic clusters that restrict the modular orientations. Specific contacts are not found across any of these domain pairs. Hydrophobic interactions include groups from similar structural elements in Ig and FnIII modules (Figure 2.15), with residues around the linker region and the FG $\beta$-turn being central in both cases. The observed Ig-Ig contacts are poorly conserved (Figure 2.15). In agreement, domain orientations in Ig A168-Ig A169 and Z1Z2 differ somewhat (Table 2.9). On the contrary, sequence
conservation at the Ig-FnIII interface indicates that this modular conformation might be common to equivalent pairs from the A-band. It is yet to be established whether the conservation of interface groups in this case, in particular the FG b-turn of FnIII domains, might play an additional role in interactions to other sarcomeric proteins – such as MyBP-C (Freiburg & Gautel, 1996).

A168-A170 recruits MuRF-1 to the M-line region of titin. The specificity of the binding might be achieved through two idiosyncratic features of this fragment of titin: i) the topography of a shallow groove spanning the length of its surface and defined by the long-range domain arrangement of this distinct tandem; and ii) the unique loop insertion of IgA169 (Figure 2.19).

ITC data show that the titin/MuRF-1 interaction is of high affinity, where the putatively a-helical region of MuRF-1 (residues 166-315) is sufficient for binding. The groove in A168-A170 could host up to 70% of the maximal hypothetical coiled-coil length achievable by MuRF-1, if its full helical region was to form such motif. These data suggest that the A168-A170 region of titin could serve as a potential therapeutic target against muscle atrophy, where binding of small molecules to its distinctive structural features could block access of MuRF-1.
Chapter 3

The B-box 2 domain of MuRF-1 is a RING-finger variant with specific dimerization properties and surface features. Comparative analysis of “cross-brace” zinc-binding domains

Michael Mrosek\textsuperscript{1}, Sebastian Meier\textsuperscript{1}, Zöhre Ucurum\textsuperscript{1}, Dietmar Labeit\textsuperscript{2}, Stephan Grzesiek\textsuperscript{1}, Siegfried Labeit\textsuperscript{2}, Olga Mayans\textsuperscript{1}

\textsuperscript{1}Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland; \textsuperscript{2}Institut für Anästhesiologie und Operative Intensivmedizin, Universitätsklinikum Mannheim, Mannheim 68167, Germany.

A plasmid containing the coding sequence for full-length MuRF-1 was provided by PD Dr. Dietmar Labeit and Prof. Siegfried Labeit. The NMR study has been carried out in collaboration with Dr. Sebastian Meier and Prof. Stephan Grzesiek. Zöhre Ucurum contributed to the study of a C21D mutated variant. All other work is my own.
3 The B-box 2 domain of MuRF-1

Abstract

The B-box motif is the defining feature of the TRIM family of proteins, characterized by a RING-finger/B-box/coiled-coil tripartite fold. We have elucidated the crystal structure of the B-box type 2 (B2) from human MuRF-1 at 1.9 Å resolution. MuRF-1 is an E3 ubiquitin ligase involved in myofibril turnover and the regulation of the trophic state of muscle. The structure of MuRF-1 B2 shows that this motif has an a/b architecture, which coordinates two zinc ions in a cross-brace topology, thus, belonging to the RING finger-like superfamily. MuRF-1 B2 adopts a distinct dimerization pattern as revealed by crystallographic and NMR data, which differs from that of other dimeric RING finger-like structures. Our findings represent the molecular basis of B2 contribution to TRIM self-assembly. By analysis of surface features we identify a long shallow groove encircling the C-terminal metal binding site that contains a group of conserved hydrophobic residues which potentially mediate protein-protein interactions of this motif. Finally, we have comparatively analyzed the RING finger-like folds including RING-fingers, PHD, U-box, FYVE, ZZ domains, B1 and B2 boxes. This reveals the extent of shared structural features among these, otherwise highly variable, folds and allows inferring the distinctive features of the B2 class of B-boxes.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

3.1 Materials and Methods

3.1.1 Cloning

The coding sequence for human MuRF-1 B-box (residues 105-149) was cloned into the vector pETM11, a variant of pET24d (Novagen) containing an N-terminal His$_6$-tag and a TEV protease cleavage site prior to the cloned gene. A B$_2$C$_{21}$D mutated variant was engineered in pETM-11 using the QuikChange protocol (Stratagene) and primers 5’-CATCTACTGTCTCACGGACGAGGTGCCCACCTGC-3’ (forward) and 5’-GCAGGTTGGGCACCTCGTCCGTGAGACAGTAGATG-3’ (reverse).

3.1.2 Protein production

Protein production used *E. coli* BL21(DE3) Rosetta (Novagen). Cultures were grown at 37°C up to an OD$_{600}$ of 0.6 in Luria Bertani medium supplemented with 25 µg/ml Kanamycin and 34 µg/ml chloramphenicol. Protein production was induced with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and growth continued for further 18 hours at 25°C. For successful protein folding, media were supplemented with 50 µM ZnCl$_2$ upon induction. Cultures were harvested by centrifugation at 2800 g and 4°C for 40 minutes. Pellets were resuspended in 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM β-mercaptoethanol in the presence of a protease inhibitor cocktail (Boehringer). Lysis was by addition of 0.1 mg/ml lysozyme and French pressing at 1500 psi and 4°C. DNase and RNase were added to a final concentration of 5 µg/ml. The homogenate was clarified by centrifugation and the supernatant applied to a Ni$_2$-chelating HisTrap column (GE Healthcare) equilibrated in lysis buffer supplemented with 20 mM imidazole. Elution used 250 mM imidazole. The eluant was dialyzed against 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM β-mercaptoethanol in the presence of TEV protease at 4°C for 16 h. Subsequent purification was by reverse metal affinity chromatography on a HisTrap column followed by gel filtration on a Superdex 75 HiLoad 16/60PG column (GE Healthcare) equilibrated in lysis buffer (Figure 3.1a and Figure 3.1b). This procedure yielded 1-3 mg of soluble, pure protein per litre of bacterial culture. Finally, samples were concentrated to 10 mg/ml as determined by BCA assay. It should be noted that zinc was only added during expression and that no zinc was present in the solutions during the purification and crystallization steps.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Figure 3.1: Purification of MuRF-1 B-box wildtype and C21D mutant.

a) Size exclusion run of B-box proteins on Superdex75 (16/60PG) equilibrated in 50 mM Tris, 150 mM NaCl and 10 mM β-Mercaptoethanol. A large portion of the soluble protein elutes in the void and is highly aggregated (black curve). The two peaks in the profile correspond to a dimeric protein (peak1) and an unfolded species (peak2) as identified by NMR spectroscopy (Figure 3.2).

b) SDS-PAGE of wildtype B-box. Fractions 1-6 correspond to protein in the void. Fraction 7-10 correspond to the first peak, fraction 11-14 to the second peak in the elution diagram shown in a). A polypeptide standard is shown in lane M (MW(WT)=5010 Da).

Interestingly, both peaks1 and peak2 in Figure 3.1a contained B-box peptide (black curve). Since The B2 domain of XNF-7 was reported to be monomeric (Borden et al., 1995), fractions of peak2 were pooled and used in crystallization experiments due to the lower apparent oligomerization state. The protein so obtained crystallized in initial crystallization screens.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger.

Figure 3.2: 1D-NMR spectra of soluble MuRF-1 B-box 2 wildtype and refolded C21D mutant.

Expression of the B2C21D variant (as described for the wild-type) yielded mostly insoluble material. Only a small soluble fraction of B2C21D eluted as one peak from the size exclusion column at an elution volume corresponding to peak2 of wildtype (blue curve in Figure 3.1). Thus, bacterial pellet resuspended in lysis buffer (see above) supplemented with 8 M urea was subjected to affinity purification under denaturing conditions. Eluted fractions were dialysed overnight into a refolding buffer containing 50 mM Tris pH 8.0, 150 mM NaCl, 10 mM β-mercaptoethanol, 1 mM ZnCl2. Refolded protein was applied to a gel filtration column equilibrated in refolding buffer. Pure samples were then concentrated to 10 mg/ml.

1H NMR spectroscopic analysis of both peak fractions of the refolded material revealed that peak1 represents a well-folded dimeric protein (Figure 3.2c), whereas peak2 corresponds to an unfolded protein population (Figure 3.2b). Subsequent NMR investigation of protein corresponding to WT peak1 indicated a well-folded species.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger (Figure 3.2a). Most likely, the pooled protein fraction of WT peak2 used initially to crystallize the protein still contained enough material of peak1 to form crystals. In line with the predominantly unfolded material present, large amounts of amorphous precipitate have been observed in hanging drops. Moreover, pooled fractions from C\textsuperscript{21D} peak2 did not result in crystal formation, while screens with peak1 containing fractions yielded a higher number of initial hits.

3.1.3 Crystallization

Untagged protein was present in 0.05 M Tris pH 8.0, 0.15 M NaCl and 10 mM β-mercaptoethanol at a concentration of 10 mg/mL as determined by the BCA assay. The reducing agent was present to avoid protein oxidation by the sulfhydryl groups involved in zinc ligation. DTT was avoided since the bipartite thiol groups in DTT can efficiently compete with the protein for a loosely bound zinc atom. Initial crystallization conditions were found using the Personal structure screen PSS-I-32 (Molecular Dimensions). Crystals grew from solutions containing 1.8 M ammonium sulfate, 0.1 M Tris pH 8.5 in hanging-drops composed of 1μl protein and 1μl reservoir solutions suspended over 500 μl mother liquor in Linbro trays. Hexagonal crystals with typical dimensions of 200x200x100 μm\textsuperscript{3} grew within two days at 4 °C (Figure 3.3). These were often affected by macro-twinning, which was detectable by visual inspection. To avoid cluster formation, the rate of diffusion within the drop was reduced by addition of varying amounts of glycerol together with lowering the temperature to 4°C. A large fraction of the protein aggregated in form of amorphous precipitate (see above). After refinement, the reservoir condition contained 1.8 M Ammoniumsulfate, 0.1 M Tris pH 8.5, 15 % glycerol. Crystals were carefully selected by visual inspection to avoid macro-twinning and finally yielded interpretable diffraction patterns.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Figure 3.3: Refined conditions of MuRF-1 B-box crystals.
Single hexagonal plates of 200x200x100 µm\(^3\) grew within three days under conditions 1.8 M Ammoniumsulfate, 0.1 M Tris pH 8.5, 15 % glycerol at 10 mg/mL protein concentration.

3.1.4 Data collection of the high resolution set

Cryo-cooled crystals used for structure determination diffracted up to 1.9 Å in a synchrotron beam at the Swiss Light Source beamline PX1. Diffraction patterns were recorded on a marCCD detector (MAR Research). Due to the presence of structural Zn ions, the “pseudo” native data set was collected at a wavelength of \(\lambda_{\text{Native}}=1.008\) Å (12.300 keV) that was well above the absorption edge of the Zn to avoid anomalous contribution and increased radiation damage.

During data collection often split spot profiles were observed and the presence of a second crystal lattice was suspected. Several crystals were tested and finally it was possible to index the crystal in space group \(P6_122\) with unit cell dimensions of \(a=76.92\) Å, \(b=76.92\) Å and \(c=146.92\) Å. All data were integrated, scaled and merged with the XDS package (Kabsch, 1993). The mosaicity as estimated by the program was 0.135. The presence of a 6\(_1\) or 6\(_5\) screw axis was indicated by systematic absences in 00l reflections. Table 3.1 describes the statistics of the high-resolution set and Figure 3.4 shows a representative diffraction image.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Table 3.1: X-ray data of the high resolution set

<table>
<thead>
<tr>
<th>Native set</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray source</td>
<td>SLS-PX1</td>
</tr>
<tr>
<td>Detector</td>
<td>MAR-CCD</td>
</tr>
<tr>
<td>Resolution</td>
<td>18-2.0 (2.1-2.0)</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.008 (12.300 keV)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>17628 (2333)</td>
</tr>
<tr>
<td>R-factor</td>
<td>6.7 (25.5)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>10.4 (10.7)</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.2 (99.3)</td>
</tr>
<tr>
<td>I/σ (I)</td>
<td>22.74 (11.31)</td>
</tr>
</tbody>
</table>

Figure 3.4: Diffraction pattern of MuRF-1 B-box crystals.

Crystals of MuRF-1 B-box diffracted to 1.9 Å in a synchrotron beam. The space group was identified as P6322 with unit cell parameters of a=b=76.21 Å, c=146.59 Å, α=β=90°, γ=120°.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Calculation of the Matthew coefficient led to the assumption that the asymmetric unit of this crystal must contain three to six copies (Table 3.2), most likely four based on a value of \( V_M = 2.8 \). A search for non-crystallographic relationships was carried out using the programs AMORE (Navaza 1994), and POLARFN (CCP4) between a resolution of 15-4 Å.

Table 3.2: Calculation of the Matthews coefficient.

<table>
<thead>
<tr>
<th>Number of molecules</th>
<th>( V_M ) [Da(^3)/Da]</th>
<th>Solvent content [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.7</td>
<td>66.7</td>
</tr>
<tr>
<td>4</td>
<td>2.8</td>
<td>55.6</td>
</tr>
<tr>
<td>5</td>
<td>2.2</td>
<td>44.5</td>
</tr>
<tr>
<td>6</td>
<td>1.9</td>
<td>33.4</td>
</tr>
</tbody>
</table>

Space group and Unit cell dimensions: \( P6_{1,5} \); \( a=b=76.21 \) Å, \( c=146.59 \) Å, \( \alpha=\beta=90^\circ, \gamma=120^\circ \).

The Matthews coefficient \( V_M \) is calculated with \( V_M = V/M_w \times Z \times N \).

- \( V \): unit cell volume (737324.812 Å\(^3\))
- \( M_w \): molecular weight of a B-box monomer (5010 Da)
- \( N \): number of molecules/ASU, \( Z \): ASU/unit cell
- \( Z \): 6 for space group \( P6 \)

A partial specific volume for proteins of 0.74 cm\(^3\)*g\(^{-1}\) can be assumed.

A distance distribution was calculated with the program MOLEMAN (CCP4) from the existing model of XNF B-box yielding a maximum distance in the monomer of 31 Å. Thus a Patterson radius of integration of 20 Å was used as parameter in for the self-rotation function. No clear NCS peaks could be identified with both programs. The small overall dimensions of the molecule might lead to a weak correlation in self-rotation searches.

### 3.1.5 MAD data collection

Anomalous diffraction data were collected at the ESRF beamline ID-23, exploiting the anomalous scattering properties of the B-box’s endogenous zinc ions. An X-ray absorption spectrometry scan was performed prior to the data collection to determine the absorption edge of the anomalous scatterer. A clear fluorescence signal indicated the presence of the anomalous scatters. A plot of anomalous scattering factors against the beam energy yielded the exact wavelengths for the experiment (data not shown). Figure 3.5 shows a theoretical plot. Based on the diagram, a three wavelength anomalous dispersion experiment was carried out at the Zn K-edge in the order peak, high energy remote and inflexion (\( \lambda_{peak} = 1.28270 \) Å (9.6602 keV), \( \lambda_{inflection} = 1.28315 \) (9.6624 keV), \( \lambda_{remote} = 1.21555 \) (10.200 keV). The exposure time was increased during
The B-box 2 domain of MuRF-1 is a dimeric RING-finger data collection ranging from 0.1 s for the peak, 0.15 s for the remote and 0.2 s for the inflexion wavelength.

![Graph](image)

**Figure 3.5: Real (f') and the imaginary (f'') components of the anomalous scattering at zinc.**

From the plot the wavelengths for the MAD experiment can be derived. For the MuRF-1 B-Box crystal the wavelengths were used: \( \lambda_{\text{peak}} = 1.28270 \text{ Å} \) (9.6602 keV), \( \lambda_{\text{inflexion}} = 1.28315 \text{ (9.6624 keV)}, \lambda_{\text{remote}} = 1.21555 \text{ (10.200 keV)}. \) Wavelengths used in the MAD experiment are indicated.

An oscillation range of \( \Delta \phi = 0.75^\circ \) was chosen for all three sets and for every wavelength 105° degree were collected in 140 frames aiming at maximal redundancy.

**Table 3.3** describes the crystallographic statistics.

<table>
<thead>
<tr>
<th></th>
<th>X-ray source</th>
<th>Detector</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ESRF-ID-23-1</td>
<td>ADSC Quantum Q210</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Peak</th>
<th>Inflexion</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>20-3.35 (3.4-3.35)</td>
<td>15-3.45 (3.5-3.45)</td>
<td>15-3.5 (3.55-3.5)</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.2546 (9.6602 keV)</td>
<td>1.28315 (9.6624 keV)</td>
<td>1.21555 (10.200 keV)</td>
</tr>
<tr>
<td>No. Bijvoet reflections</td>
<td>33994 (1556)</td>
<td>31059 (1347)</td>
<td>29580 (1216)</td>
</tr>
<tr>
<td>( R_{\text{sym}} ) (%)</td>
<td>11.1 (42.6)</td>
<td>10.6 (41.6)</td>
<td>12.9 (47.5)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.9 (3.9)</td>
<td>3.7 (3.7)</td>
<td>3.9 (3.9)</td>
</tr>
<tr>
<td>Completeness</td>
<td>98.8 (99.7)</td>
<td>98.2 (99.4)</td>
<td>98.2 (98.9)</td>
</tr>
<tr>
<td>( I/\sigma(I) ) (%)</td>
<td>9.8 (3.5)</td>
<td>10.1 (3.5)</td>
<td>9.2 (3.1)</td>
</tr>
</tbody>
</table>

Friedel pairs are treated as different reflections

All data were collected under cryogenic conditions and processed with the XDS package (Kabsch, 1993). A first estimate of the anomalous difference signal was obtained after scaling the data in XDS. A comparison between the mean intensity error of symmetry related reflections assuming the situation that Friedel’s law is valid versus the situation that symmetry-related reflections do not have the same intensity due to
anomalous contribution revealed the presence of anomalous scatterers in the crystal (Table 3.4).

Table 3.4: Estimation of anomalous signal with XDS.

<table>
<thead>
<tr>
<th>Wavelength</th>
<th>S_norm/S_ano$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native (reference)</td>
<td>1.11 (1.01)$^b$</td>
</tr>
<tr>
<td>Peak</td>
<td>1.27 (1.02)</td>
</tr>
<tr>
<td>Remote</td>
<td>1.22 (1.02)</td>
</tr>
<tr>
<td>Inflexion</td>
<td>1.16 (1.01)</td>
</tr>
</tbody>
</table>

$^a$S_norm: mean value of Sigma(I) for acentric reflections assuming Friedel’s law is valid.
S_ano: mean value of Sigma(I) for acentric reflections assuming Friedel’s law is violated.

Anomalous scattering contributions to the intensities are indicated by S_norm/S_ano >1

$^b$total between 16-2.1 Å; in brackets highest shell 2.2-2.1 Å.

3.1.6 Determination of heavy atom sub-structure and phase calculation

The subsequent steps in the structure elucidation of MuRF-1 B-box aimed at the determination of heavy atom positions to obtain experimental phase information. Thus, a heavy atom search was carried out using the program SOLVE (Terwilliger, 1999). Input files for the three wavelengths contained the raw, unmerged intensities of the three wavelength MAD experiment. The program SOLVE searches for heavy atoms using difference Patterson methods and scores potential solutions on the basis of agreement with the Patterson map, the difference Fourier map, the presence of solvent and protein regions in a native electron density map, and the Figure of merit of phasing. Given the presence of at least three to six molecules per asymmetric unit as determined from the Matthew coefficient a search for eight Zn sites was carried out using data between 18 to 2.3 Å resolution. Theoretical f’ and f” values were used at the corresponding wavelengths. The program unambiguously determined the coordinates of six Zn ions in the asymmetric unit that are summarized in Table 3.5. Figure 3.6 shows the harker peaks in anomalous difference Patterson maps.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger.

Table 3.5: Real space positions of the six Zn-sites in the asymmetric unit.

<table>
<thead>
<tr>
<th>ATOM</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Peak height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Zn</td>
<td>40.895</td>
<td>56.067</td>
<td>5.419</td>
<td>45.81</td>
</tr>
<tr>
<td>2 Zn</td>
<td>11.910</td>
<td>8.324</td>
<td>0.480</td>
<td>51.77</td>
</tr>
<tr>
<td>3 Zn</td>
<td>44.075</td>
<td>41.235</td>
<td>10.850</td>
<td>45.79</td>
</tr>
<tr>
<td>4 Zn</td>
<td>10.545</td>
<td>24.737</td>
<td>9.954</td>
<td>40.07</td>
</tr>
<tr>
<td>5 Zn</td>
<td>20.903</td>
<td>25.421</td>
<td>6.534</td>
<td>43.73</td>
</tr>
<tr>
<td>6 Zn</td>
<td>42.283</td>
<td>35.630</td>
<td>1.606</td>
<td>35.13</td>
</tr>
</tbody>
</table>

Table 3.6 shows the phasing statistics after the SOLVE run, which yielded the initial electron density maps.

Table 3.6: Phasing statistics prior to solvent flattening.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Total</th>
<th>7.81</th>
<th>5.09</th>
<th>4.03</th>
<th>3.44</th>
<th>3.04</th>
<th>2.76</th>
<th>2.55</th>
<th>2.37</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflections</td>
<td>11772</td>
<td>678</td>
<td>1014</td>
<td>1268</td>
<td>1454</td>
<td>1618</td>
<td>1779</td>
<td>1908</td>
<td>2053</td>
</tr>
<tr>
<td>Figure of Merit</td>
<td>0.81</td>
<td>0.88</td>
<td>0.86</td>
<td>0.90</td>
<td>0.88</td>
<td>0.84</td>
<td>0.78</td>
<td>0.76</td>
<td>0.68</td>
</tr>
</tbody>
</table>
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Figure 3.6: Experimental anomalous difference Patterson maps at P6_522 Harker sections.

The Maps were calculated with FFT (CCP4) and heavy atom coordinates superimposed with the program vector (Evans, 1990). A contour level of 1.5 $\sigma$ is shown with an increment of 0.5 $\sigma$. Peaks are found at Harker sections w=0.33 and w=0.17 as well as three sections below Harker level at w=0.30.

The initial maps were density modified assuming a solvent content of 50% with the program RESOLVE (Terwilliger, 2000; Terwilliger, 2002). Resolve uses a
maximum-likelihood approach to do solvent flattening, while other methods modify a map to meet expectations and the obtained new phases are recombined with experimental phases. In RESOLVE the correlation between the experimentally available and the expected map is maximized in a simultaneous manner by statistical methods. A process called prime and switch phasing is applied to reduce model bias (Terwilliger, 2000). Resolve also carries out an automatic model building. At first elements of secondary structure are built, which are then extended by addition of tri-peptide fragments followed by the building of side chains according to a given sequence. The so obtained model (after RESOLVE; Table 3.7) was directly used for automated model building and refinement with the program ARP/wARP (Lamzin and Wilson, 1997) and REFMAC (Murshudov, 1997). Non-crystallographic symmetry was neglected during the building.

<table>
<thead>
<tr>
<th>Acentric reflections</th>
<th>DMIN</th>
<th>FOM-Prior</th>
<th>FOM-Map</th>
<th>FOM-Total</th>
<th>CC-Prior-Map</th>
<th>Fraction Prior</th>
<th>&lt;m&gt; total</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL:</td>
<td>0.83</td>
<td>0.70</td>
<td>0.89</td>
<td>0.72</td>
<td>0.66</td>
<td>0.89</td>
<td>9356</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>0.92</td>
<td>0.88</td>
<td>0.98</td>
<td>0.82</td>
<td>0.60</td>
<td>0.98</td>
<td>306</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>0.91</td>
<td>0.88</td>
<td>0.97</td>
<td>0.82</td>
<td>0.57</td>
<td>0.97</td>
<td>1172</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>0.91</td>
<td>0.87</td>
<td>0.97</td>
<td>0.84</td>
<td>0.61</td>
<td>0.97</td>
<td>1551</td>
<td></td>
</tr>
<tr>
<td>2.9</td>
<td>0.87</td>
<td>0.76</td>
<td>0.93</td>
<td>0.78</td>
<td>0.67</td>
<td>0.93</td>
<td>1615</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>0.79</td>
<td>0.62</td>
<td>0.86</td>
<td>0.67</td>
<td>0.67</td>
<td>0.86</td>
<td>2878</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>0.71</td>
<td>0.49</td>
<td>0.77</td>
<td>0.53</td>
<td>0.68</td>
<td>0.77</td>
<td>1834</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3.7: Phasing statistics after solvent flattening with RESOLVE

#### 3.1.7 Model building

The structure of MuRF-1 B-box domain was automatically built with the program ARP/wARP (Lamzin and Wilson, 1997) using the warpNtrace procedure shown in Figure 3.7.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger.

Figure 3.7: Schematic description of the option warpNtrace in ARP/wARP.

This procedure initially positions unconnected atoms into the electron density and then searches for protein-like patterns by applying stereochemical constraints to build the main-chain. In an iterative procedure the existing model is updated by the density-based addition of more and more atoms in real space and fitting of the calculated and observed structure factor in reciprocal space.

The structure was built in 10 consecutive cycles of automatic model building and refinement using REFMAC against a data set with a resolution of 1.9 Å. Three monomers were build in the asymmetric unit. The progress of rebuilding and refining was monitored by the decrease in the R-factor and the R-free as well as by inspection of $\sigma_a$-weighted $2F_{\text{obs}} - F_{\text{calc}}$ maps. Finally the solvent model was build in ARP/wARP running in solvent mode and was manually completed in O in a $F_{\text{obs}} - F_{\text{calc}}$ map that was contoured at 3.5 $\sigma$. Finally the data were refined against a data set that contained data to 1.9 Å resolution. The presence of two zinc ions in each of the three NCS-related monomers in the asymmetric unit is evident in an anomalous difference Fourier map.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger after RESOLVE contoured at 4 σ using the peak data (Figure 3.8a). The quality of the final 2F_{obs}-F_{calc} is shown in Figure 3.8b.

![Electron density map after RESOLVE and final 2F_{obs}-F_{calc} map.](image)

**Figure 3.8: Electron density map after RESOLVE and final 2F_{obs}-F_{calc} map.**

a) Electron density at 2.3 Å resolution after RESOLVE is shown in grey contoured at 1.5 σ. An anomalous difference Fourier map contoured at 4 σ confirms the presence of two zinc ions in the fold. The metal ligating residues are shown as stick and zinc ions are shown as spheres. b) Final 2F_{obs}-F_{calc} map at 1.9 Å resolution contoured at 1.5 σ.

The final R_{factor} was 21.6 and the R_{free} was 25.6 excluding 853 (5% of the unique reflections as test set). The refinement statistics can be found in Table 3.8.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger.

Figure 3.9: Ramachandran plot for three molecules of MuRF-1 B-box in the asymmetric unit.

All protein residues are located within the colored areas (apart from Glu47) indicating allowed values for \( \phi \) and \( \psi \) angles. The plot was calculated with the program PROCHECK.

Table 3.8: Final Refinement statistics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of reflections in working/ free set</td>
<td>19667 / 853</td>
</tr>
<tr>
<td>Number of protein residues / solvent molecules</td>
<td>139 / 145</td>
</tr>
<tr>
<td>Number of Zn atoms</td>
<td>6</td>
</tr>
<tr>
<td>R-factor / R-free (%)</td>
<td>21.2 / 25.5</td>
</tr>
<tr>
<td>rmsd bond length (Å) / bond angle (°)</td>
<td>0.012 / 1.47</td>
</tr>
</tbody>
</table>

The structure was analyzed with the program PROCHECK. Only one residue was found to be in disallowed regions in all three NCS copies (Figure 3.9).
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

3.1.8 Crystal packing and Dimeric arrangements

After determining the position of the heavy atoms with SOLVE, it became clear that three molecules per asymmetric unit are present in MuRF-1 Box 2 crystals. These NCS copies are referred to as A, B, C. Given the proposed dimerization propensity of XNF B-box (Borden 1995), crystal packing and NCS were analyzed for potential and physiologically relevant dimeric interactions. Based on the orientation, distance and interface area of molecules A and B in the asymmetric unit a NCS related dimerAB was identified. A second interaction resulted from the crystallographic relation between NCS copies C and B’ across different asymmetric units, termed dimerBC. Figure 3.10 shows the interrelationship of crystallographic and NCS-related dimer.

Figure 3.10: Crystallographic and NCS interactions in MuRF-1 B-box crystals.

Two dimeric arrangements based on orientation, distance and interface area of monomers are characterized by either non-crystallographic two-fold symmetry between NCS copies A and B or crystallographic 2-fold symmetry between NCS monomers C and B’ of adjacent asymmetric units (dotted lines). NCS copies in the asymmetric unit: A (green), B (cyan) and C (magenta).
3.1.9 Crystallization and data collection of the C24D mutant of MuRF-1 B-box

Protein crystals of mutated B-box grew in form of thin long needles of 400x20x20 µm³ in condition PSS-I-32 (Molecular dimension) containing 1.8 M Ammoniumsulfate, 0.1 M Tris pH 8.5 (Figure 3.11). The final condition included additionally 5 % glycerol. Crystals appeared after five days at 4°C.

![Figure 3.11: Refined condition of C²⁴D mutant of MuRF-1 B2 crystals.](image)

Single long needles of 400x20x20 µm³ grew within five days under conditions: 1.9 M Ammoniumsulfate, 0.1 M Tris pH 8.5, 5 % glycerol at 10 mg/mL protein concentration.

Data on MuRF-1 B-box crystals were collected in house using an Eliot GX20 rotating disc anode. This source yielded X-rays with a wavelength of λ=1.5418 Å. The detector was a MAR345 (X-ray Research) image plate detector. Crystals were exposed for 1h per frame in 1° oscillations. In total 120 frames were recorded at cryogenic temperature (100K). The data were processed with the XDS package (Kabsch, 1993). The crystal lattice could be indexed assuming a primitive orthorhombic cell with cell dimensions of a= 60.8 Å, b=66.9 Å, c=77.4 Å and α=β=γ=90°. The diffraction limit was 3.4 Å. Table 3.9 shows the overall statistics of data collection on B-box C24D crystals.

Calculation of the Matthew coefficient indicated that between three (Vₘ= 4.8 [Å³/Da]; 74 % solvent content) and eight (Vₘ= 1.8 [Å³/Da]; 31 % solvent content) copies can be expected in the asymmetric unit.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

| X-ray source | Eliot GX20 |
| Detector    | MAR-345 image plate |
| Resolution  | 15-3.4 (3.5-3.4) |
| Wavelength  | 1.5418 Å (8.040 keV) |
| Unique reflections | 4542 (365) |
| R-factor | 16.5 (28.3) |
| Multiplicity | 3.4 (3.5) |
| Completeness | 97.2 (96.6) |
| I/σ (I)    | 7.8 (4.75) |

### Table 3.9: Overview of data collection on B-box mutant C24D.

#### 3.1.10 Phase determination and Model building

The structure of the B-box was phased by molecular replacement with the program PHASER (Storoni et al., 2004) using the NCS related dimerAB as search model. The PHASER output model was subsequently refined with CNS (Brunger et al., 1998) in the absence of zinc ions against the 3.4 Å data set. The model for refinement was characterized by an R\textsubscript{factor} of 35 % and and R\textsubscript{free} of 38 % (4.5 % reflection in test set). Conjugate gradient and grouped B-factor refinement against a maximum likelihood target function improved the model to an R\textsubscript{factor} of 27.4 % and an R\textsubscript{free} of 35.8 %. Inspection of 2F\textsubscript{obs} - F\textsubscript{calc} electron maps for the presence of two zinc was carried out with the program O (Jones et al., 1991). Table 3.10 summarizes the statistics of refinement.

### Table 3.10: Refinement of MuRF-1 B-box C21D mutant.

| Number of reflections in working/ free set | 3988 / 208 |
| Number of protein residues / solvent molecules | 285 / - |
| Number of Zn atoms | 12 |
| R-factor / R-free (%) | 30.3 / 36.6 |
| rmsd bond length (Å) / bond angle (°) | 0.009 / 1.52 |

#### 3.1.11 NMR experiments

Uniformly $^{15}$N-labelled wild-type B2 protein was derived from M9 minimal medium using $^{15}$NH$_4$Cl as the sole nitrogen source. Otherwise, expression protocols were as above. In this case, only insoluble sample was obtained presumably due to the formation of Zn$_3$(PO$_4$)$_2$ and the absence of soluble pools of zinc from the medium. Thus, refolding and purification protocols as those described above for B2$^{C21D}$ were followed.
NMR spectra were recorded at 3 mM sample concentration in 20 mM Tris-HCl at pH 7.0, 75 mM NaCl, 10 mM β-mercaptoethanol. The appearance of $^1\text{H}-^{15}\text{N}$ HSQC NMR spectra was consistent with a well-structured sample. The oligomeric state of the sample was estimated from $^1\text{H} T_2$ spin echo relaxation measurements as well as $^{15}\text{N} T_1, T_2$ and $^{15}\text{N}$-heteronuclear NOE data (Barlow et al., 1994). The protein backbone $^{15}\text{N}$ and $^1\text{H}$ resonances were assigned with $^{15}\text{N}$ separated NOESY and TOCSY spectra. Residual dipolar couplings (RDCs) were measured in a solution of 1 mM protein dissolved in 20 mM Tris pH 7.0, 75 mM NaCl and 10 mM β-mercaptoethanol supplemented with Pf1 phage (ASLA Biotech) (Hansen et al., 1998) to a final phage concentration of 15 mg/ml. The electrostatic alignment was tuned down by the addition of NaCl to a final concentration of 150 mM NaCl. Inter-subunit NOEs to residue F32 were identified with a 2D NOESY spectrum with a mixing time of 100 ms. All spectra were recorded on a BRUKER DRX 600 equipped with a TXI probe and on a BRUKER DRX 800 spectrometer equipped with a TCI cryoprobe. Data were processed with NMRPipe (Delaglio et al., 1995) and analyzed with PIPP (Garrett et al., 1991).

### 3.2 Results and Discussion

#### 3.2.1 Crystal structure of MuRF-1 B2

The crystal structure of the B2 box from human MuRF-1 has been elucidated at 1.9 Å resolution. The crystal form used in this study contains three molecular copies in its asymmetric unit (chains A, B and C) that are basically identical (average rmsd is 0.30 Å for 43 matching Ca atoms as calculated using SPDBV; Guex and Peitsch, 1997). MuRF-1 B2 consists of a structural core formed by a α-helix (α1) that packs against a three-stranded anti-parallel β-sheet (β1-β3). This core supports three dominant loop regions (L1, L2, L3) involved in zinc binding (Figure 3.12). Anomalous difference Fourier maps unambiguously revealed that each molecular copy of B2 binds two zinc ions. Metal ligation is mediated by a “cross-brace” topology closely related to that of RING finger domains (Kentsis and Borden, 2000), where zinc-binding site I (ZnI) is formed by the first and third ligand pairs and zinc-binding site II (ZnII) by the second and fourth ligand doublets.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Figure 3.12: Structure of B2 and B1 boxes.

a) Structure of MuRF-1 B2, where metal binding ligands have been displayed. The N- and C-terminal zinc ions are shown in blue and red, respectively. Secondary structure elements and loop regions are labelled; b) Superposition of MuRF-1 B2 (cyan), ATDC B2 (PDB entry C2SV) (green) and MID1 B1 (PDB entry 2FFW) (blue). The B2 boxes show a close structural similarity. Loop insertions in the B1 class with respect to B2 sequences are shown in red; c) Structure-based sequence alignment of B-boxes of known structure. Secondary structure elements and loop regions are indicated and coloured as in a). Metal ligands are highlighted in black, numbered and coloured according to the metal site they form. The cross-brace topology is indicated.

In MuRF-1 B2, site ZnI is constituted by residues C6xxH9 from the N-terminal L1 loop and C26xxC29 spanning the C-terminus of strand β2 and N-terminus of helix α1. Site ZnII involves C18xxC21 and H35xxC38 located in L2 and L3, respectively (Figure 3.12a and Figure 3.12c). The C-terminal cysteine ligand, C38, is a unique feature of human MuRF proteins. This residue is a conserved histidine in all other B2 boxes, but the possible functional implications of this substitution are currently unknown. Overall, MuRF-1 B2 constitutes a compact, spherical structure with an inter-zinc distance of ~11 Å, currently the shortest among RING-related domains.
The fold and zinc ligation properties of MuRF-1 B2 resemble those of MID1 B1 (Massiah, 2006) and ATDC B2 (PDB 2CSV) (Figure 3.12b), but differs strongly from those described for B2 from XNF7 (Borden et al., 1995). The latter adopts a topology unique among zinc-binding folds and coordinates one single zinc ion leaving four potential metal ligands unoccupied. The B2 boxes of XNF7 and MuRF-1 share 24% sequence identity and the position of all zinc ligands is conserved (Figure 3.12c). Thus, the structural differences in these domains are surprising. In order to investigate whether such divergence could derive from the presence of an unusual zinc-ligand aspartate (D20) in XNF7 B2 leading to altered zinc binding properties, we analyzed a C21D mutated variant of MuRF-1 B2 (B2C21D). This was expected to mimic the zinc ligation properties of XNF-7. However, 1H-NMR experiments showed that both MuRF-1 wild-type B2 and its B2C21D variant shared a high spectral similarity indicative of common structural features. This was confirmed by a subsequent crystallographic analysis of B2C21D, which demonstrated that an aspartate residue at this position does neither disrupt the zinc binding properties of the B2 fold nor affects its topology. This conclusion is supported by the recent model of ATDC B2 (PDB 2CSV), which also contains an aspartate residue at this position and exhibits structural features closely related to those of MuRF-1 B2 (Figure 3.12b). Thus, a further examination of XNF7 B2 directed to elucidate the source of this discordance is prompted.

3.2.2 Oligomeric state of MuRF-1 B2

The crystal forms of wild-type MuRF-1 B2 and its variant B2C21D contained identical dimeric formations in their asymmetric units. In order to investigate whether these represent the oligomeric state of the B2 motif in solution, we analyzed both samples by NMR-spectroscopy. The correlation time τc at 25 °C estimated from 1H T2 as well as a complete set of 15N relaxation data was 6 ns, corresponding to a molecular weight (Mw) of ~12 kDa, approximately twice that of a B2 monomer (Mw = 5.5 kDa) confirming that the molecule is dimeric in solution.

To assess whether the structure of the dimer in solution corresponds to that in the crystalline state, models were compared to experimental 15N-1H residual dipolar couplings (RDCs) recorded from weakly aligned samples in a solution containing Pf1 phage. RDC data were fitted to i) monomeric models, ii) the dimer observed in the crystallographic asymmetric unit and iii) one additional dimeric form that arises in the
The B-box 2 domain of MuRF-1 is a dimeric RING-finger lattice by effect of crystallographic two-fold symmetry (Figure 3.13a). RDC measurements agreed excellently with the tertiary structure of the monomer (quality factor $Q=0.21$; Q-factor definition as in Cornilescu et al., 1998) indicating that the crystal structure of MuRF-1 B2 closely represents that of the protein in solution. The fitting of RDCs to dimeric models unambiguously identified the dimer in the asymmetric unit ($Q=0.22$), and not that generated by crystallographic symmetry ($Q=0.64$), to agree with the quaternary structure of the sample in solution. It should be noted that the Q-factor value of this dimer was nearly identical to that calculated from fits to a single protomer. Furthermore, monomeric spectral species could not be observed in a dilution series of the sample monitored by NMR. Since the sample remained fully assembled at the minimal concentration assayed of 20 µM, the $K_d$ of the interaction must be well below that value. Equivalent results were obtained from $B_2^{C21D}$ samples, proving that the substitution of the zinc-ligand at that position does not affect the dimerization properties of the motif. The stable dimerization of this domain as measured by NMR points to a physiological role of the B2 motif in the self-assembly of MuRF-1.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger.
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Figure 3.13: The dimeric structure of MuRF-1 B2

a) Fitting of $^1$H$^1$N RDC data to the crystal structures of the monomer (upper panel), the dimer in the asymmetric unit (middle) and the dimer formed by two-fold crystallographic symmetry (bottom). Q-factors below 0.25 are expected for fittings of high resolution crystal structures (Cornilescu, 1998); b) Intersubunit NOEs (16 in total) measured between protons of F32 and neighbouring hydrophobic residues are shown as dashed lines; c) Surface representation. The groove encircling the metal site ZnII is indicated. The exposed side chains of hydrophobic residues are coloured green and the conserved aromatic Y17 is labelled. d) Interface contacts in the crystal structure of dimeric MuRF-1 B2. Hydrophobic groups are shown in green and highlighted using van der Waals spheres. Polar interactions are given as dashed lines.

3.2.3 Structure of the MuRF-1 B2 dimer

The dimeric structure of MuRF-1 B2 exhibits a two-fold symmetry, where the $\alpha_1$ helix of one subunit packs against a concave depression formed by the $\alpha$-helix and $\beta$-sheet of the other (Figure 3.13b). The interface contains both hydrophobic and polar interactions, which are mediated by groups of sequential residues that form segregated clusters in space (Figure 3.13d). Inter-subunit polar contacts are primarily established by a charged sequence segment, D11-N15, located at the N-terminal L1 loop and just prior to strand $\beta_1$ (Figure 3.12d). This generates a salt bridge, D11-K13, and a hydrogen bonded pair, E12-N15. One additional polar contact is that established by the mutual packing of residues S27 from the $\alpha_1$ helices of each subunit. Hydrophobic interactions involve residue I16 in strand $\beta_1$ and the sequence V40-A41-P42 in strand $\beta_3$ of one subunit, which pack against the hydrophobic side of the amphipathic helix $\alpha_1$, comprising M28 and the C-terminal flanking group F32, in the other. These interactions were confirmed by NMR measurements that identified 16 inter-subunit NOEs between F32 and I16, V40, A41, P42 (Figure 3.13b), thereby corroborating the structure of the dimer in solution.

The hydrophobic groups of strands $\beta_1$ and $\beta_3$ as well as those of helix $\alpha_1$ are well conserved across members of the B2 family (according to an analysis of 52 B2 sequences). Remarkably, positions S27 and F32 are subject to compensatory sequence variations in the family so that, in fact, position 27 commonly hosts a hydrophobic group - often an aromatic – and position 32 is most frequently occupied by a serine residue. Also a preference to host charged residues at the L1 loop exists in the family, although no strict sequence conservation can be observed in this region. Thus, members of the B2 family have the potential to dimerize following the same pattern of MuRF-1 B2, indicating that self-assembly might be a generic property of this fold. On the other
The B-box 2 domain of MuRF-1 is a dimeric RING-finger (Meroni and Diez-Roux, 2005). Contrary to the short loops of B2 motifs, which are conserved in length, B1 boxes contain two long variable regions following the second and seventh metal binding ligand (Figure 3.14c). In MID1 B1 the C-terminal insertion adopts a bulky, “lasso-like” structure stabilized by conserved proline residues that results in an unusual Hx,H metal binding pair that is not present in any known B2 (Figure 3.12b). This long, charged loop insertion would prevent the formation of a dimeric arrangement as that of MuRF-1 B2 by causing steric clashes across subunits. Thus, a differential propensity to dimerization specific to B2 but not B1 boxes might be inferred. This is in agreement with the evolutionary invariability of the B2 domain directly preceding the CC region that suggests a functional division of the TRIM fold into RING, B1 and B2-CC domain (Reymond et al., 2001), where the latter possibly acts as an integrated module in self-association securing the correct quaternary structure of the TRIM protein.

The B-box motif is known to mediate heterologous protein-protein interactions, but structural data on these complexes are yet to become available. In an attempt to reveal the possible determinants of B2-mediated interactions, we have examined the surface of MuRF-1 B2. Each protomer exhibits a long, shallow groove that encircles the C-terminal metal binding site ZnII formed by loops L2 and L3 (Figure 3.13c and Figure 3.12c). In one side of the semicircle, the groove includes a hydrophobic cluster centered on the solvent exposed surface of the b-sheet and containing a highly conserved aromatic residue, Y17 (Figure 3.13c). It can be speculated that this groove and its hydrophobic features might confer MuRF-1 B2 the capability to mediate interactions to MuRF-1 binding partners. Alternatively, this feature might play a role in the domain organization of MuRF-1 or in its formation of high-order oligomeric states. It should be noted that the characteristics of this groove, which appear conserved throughout the B2 family, do not resemble those defining ubiquitin ligase activity in clefts of RING finger domains (Zheng et al., 2000). In the latter the groove is shaped by long C-terminal loop insertions, which exist only as short sequences in B2 motifs (Figure 3.14a and Figure 3.14c). Thus, a possible involvement of MuRF-1 B2 in ubiquitination events cannot be inferred from its structure.
The dimerization pattern of MuRF-1 B2 is unique among those of RING finger-like domains. The RING finger of RAG1 (Bellon et al., 1997), the heterodimeric BRCA1/BARD1 tumor suppressor (Brzovic, et al., 2001), the homodimeric FYVE domain from EEA1 (Dumas et al., 2001) and the U-box of protein CHIP (Xu et al., 2006) self-assemble via additional α-helical fractions, mostly C-terminal to the zinc-binding domain. In those cases where additional direct contacts across the zinc-binding motifs are observed, these commonly involve β-sheet components and/or its neighbouring loop region. Helix α1 invariably lies in the outside, opposite to the dimer interface. In the absence of their additional helical extensions, the self-association of these motifs appears weak, as confirmed by recent data on Prp19 U-box (Vander Kooi et al., 2006). In contrast, the association of B2 is comparatively more stable, suggesting that its contribution to overall protein self-assembly might be more manifest than that mediated by those other motifs.

### 3.2.4 Comparative analysis of B2 and RING finger-like motifs

The structures of the B2-boxes from MuRF-1 and ATDC show that this motif belongs to the RING finger superfamily of zinc-binding motifs. RING (Barlow et al., 1994), PHD (Capili et al., 2001), FYVE (Misra et al., 1999), ZZ (Legge et al., 2004), U-box (Ohi et al., 2003) and B-boxes B1 (Massiah et al., 2006) and B2 are structurally related (Figure 3.14a). Each of these folds exhibits a high variability and, apart from a few residues defining the hydrophobic core and the metal binding groups, little conservation characterizes the respective classes (Figure 3.14c).
The B-box 2 domain of MuRF-1 is a dimeric RING-finger

Figure 3.14: Comparative analysis of RING finger-like zinc-binding motifs.

a) Compendium of representative RING-like folds structurally characterized to date. PDB accession codes are given in brackets. Throughout, N- and C-terminal zinc ions are shown in blue and red, respectively; b) Superposition of all domains shown in a. The shared structural core consisting of helix α1 and strands β1 and β2 is revealed (coloured). The most structurally similar region encompasses the metal site ZnI, while the architecture surrounding site ZnII is largely variable; c) Comparison of consensus sequences for the fold classes displayed in a. Metal binding residues are numbered and coloured according to the metal site they form. The position of hydrophobic residues is indicated by Θ and highlighted in yellow. It should be noted that in MuRF-1 B2, hydrophobic residues do not form part of a motif core but that they are mainly solvent exposed. Other conserved residues within the classes are indicated. Consensus sequences for individual folds have been adapted from: RING and PHD domain (Capili et al., 2004); B1 (Massiah et al., 2006); ZZ (Meroni et al., 2005); FYVE (Stenmark et al., 2002); U-box Ohl et al., 2003).

Despite, a similar architectural α/β-core is shared by most of these motifs. It consists of an N-terminal, two-stranded, antiparallel β-sheet that leads to a α-helical formation of longer or shorter length (Figure 3.14b). Conserved hydrophobic residues within the classes are primarily located around the fifth metal binding residue, in the sequence connecting strand β2 to the α-helix (Figure 3.14c). They play a role in the packing of the β-sheet against the helix. Interestingly, the hydrophobic residues of MuRF-1 B2 at this position do not form part of a motif core, but of the dimerization
The B-box 2 domain of MuRF-1 is a dimeric RING-finger interface. In fact, MuRF-1 B2 does not possess a defined hydrophobic core beyond the interaction of V40 with the aliphatic chain of K30, indicating that this motif might be stabilized mostly by its metal binding and dimerization. The shared α/β-topology supports in all cases the ligation of two zinc ions in a “cross-brace” fashion, as first described for the RING finger domain (Barlow et al., 1994). Remarkably, the U-box domain, which is closely related to the RING motif, does not bind zinc. Instead, it houses a decentralized network of hydrogen bonds and salt bridges (Ohi et al., 2003). The different folds become individualized through the nature of their metal binding ligands and the length and composition of the spacing sequences. Strictly conserved spacing across folds is observed only between the first two metal ligand residues, which are separated by 2-3 variable residues. Other spacings can vary largely (Figure 3.14c). Of all “cross-brace” zinc-binding motifs known to date, the B2 domain adopts the most compact structure as a consequence of the short sequence insertions between its metal binding residues. This results in an inter-zinc distance of ∼11 Å, similar to that of the FYVE motif, but shorter than that of ZZ (∼13 Å), RING and PHD domains (∼14 Å). The variable loop insertions confer substantial structural plasticity to this group of motifs and allow for a broad range of protein interactions and, hence, cellular roles to be supported. RING, PHD and U-box show E3-ligase activity but, apart from the mediation of protein-protein interactions, no other common function has yet been attributed to the remaining motifs.

The structural α/β-core of these folds maintains the architecture of the zinc-binding site ZnI. This constitutes the most structurally similar region across motifs. In contrast, the loops forming the C-terminal site ZnII – and in particular loop L3 - display enormous structural variability across motifs. Such C-terminal sequences often form additional secondary structure elements such as an extra β-strand (β3) that extends the β-sheet (Figure 3.14a). In fact, the length and composition of these loops is likely to determine the specific domain functionality by affecting the groove surface topography, which is allegedly involved in protein-protein interactions in these folds. Metal site ZnII has been shown to have a lower affinity for zinc in RING fingers and, in some cases, does not even seem required for primary folding (Kentsis et al., 2002). Thus, this area might be under lower structural pressure and has higher evolvability. While the ZnII site architecture of MuRF-1 B2 is closely related to that of the ZZ domain, both showing a very short loop that leads into a third strand β3, that of the U-box resembles RING
The B-box 2 domain of MuRF-1 is a dimeric RING-finger domains. The C-terminal regions of PHD and FYVE motifs significantly differ from those of any other class, with the FYVE fold representing a convolution of RING topologies where site ZnII shows an own RING α/β architecture.

3.2.5 Final remark

The structure of MuRF-1 B2 aids our understanding of the TRIM integrated fold, whose biological function relies in its ability to self-assemble into higher-order oligomers. We found that isolated MuRF-1 B2 adopts a highly stable dimeric state, suggesting that this motif is likely to participate in self-association together with the coiled-coil fraction. Given the defining nature of the B2 domain within the TRIM fold, its dimerization might be critical for the higher order assembly of these proteins and/or the structural integrity of this fold. The B2 motif might contribute to ensure the stoichiometry and/or registry of the CC fraction. Although CC are established oligomerization motifs in proteins, their association properties are not always uniquely defined and they can be found in proteins in combination with other motifs that influence their association. A prominent example is that of the trimeric CC of fibrin, which is followed by a C-terminal extension of ~30 residue length that folds into a b-hairpin (Tao et al., 1997). This intertwines with equivalent b-hairpins from other subunits, being essential to achieve stable CC trimerization in that case. An additional example is that of Dystrophia Myotonica Kinase, whose assembly into dimers is essential for activation. Although its self-association was thought to be mediated by a C-terminal CC domain, it has been recently shown that this CC in isolation forms trimeric assemblies, possibly unrelated to the physiological state of the kinase, while sequence segments flanking the catalytic kinase fraction itself are responsible for the dimeric state of this protein (Garcia et al., 2006). Thus, it should be investigated to which extent the assembly properties of the B2 under study are essential for MuRF-1 to achieve a functional quaternary structure. Furthermore, it should be examined whether the B2 motif of MuRF-1 plays a central physiological role in muscle by being involved in the recruitment of MuRF-1 binding partners to this signalling spot in the sarcomere.
Chapter 4

Biophysical characterization of LAP2α

Michael Mrosek¹, Barbara Korbei², Bohumil Maco¹, Roland Foisner², Olga Mayans¹

¹Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland; ²Max F. Perutz Laboratories, department of Medical Biochemistry, Medical University of Vienna, Dr. Bohrgasse 9, A-1030 Vienna, Austria

Initial clones for LAP2α were provided by Barbara Korbei and Prof. Roland Foisner. Bohumil Maco contributed the electron microscopy study. All other work is my own.
Chapter 4: Biophysical characterization of LAP2α

Abstract

The non-membrane bound scaffolding protein LAP2α is critically involved in the recruitment of retinoblastoma protein as well as chromatin to the filamentous intranuclear structures of lamin A. Binding to these factors is mediated through a unique C-terminal domain only present in the LAP2α isoform. To gain an insight into the architecture of specific nuclear structures formed by LAP2α as well as to assess its behavior with respect to further crystallization attempts, we have initiated biophysical characterization on recombinantly expressed, purified full length protein and on a C-terminal construct LAP2α_410-693 that contained the proposed lamin A binding region.

We demonstrate based on size exclusion chromatography that full length LAP2α forms higher order homo-oligomeric structures. By contrast, size exclusion chromatography in connection with multile-angle light scattering as well as ultracentrifugation revealed that C-terminal LAP2α_410-693 exists in a homo-dimeric form in solution. Furthermore, LAP2α_410-693 possesses a high degree of α-helical secondary structure content as deduced from CD-spectroscopy. Both constructs show a pronounced oxidation tendency due to exposed surface cysteines and were found to be associated with DNA fragments.
4.1 Materials and Methods

4.1.1 Cloning

Plasmids containing coding sequences for full-length human LAP2α and LAP2α^410-693 were provided by Prof. Roland Foisner (Vienna Biocenter). Those sequences had been cloned into the plasmid pET-23a(+) via Nhe-I and Xho-I as described (Vlcek et. al., 1999). A cystein-to-serine mutant (LAP2α^410-693Cys-->Ser) was generated by Barbara Korbei (Vienna Biocenter) by means of site-directed mutagenesis, where C517, C560, C569, C657 and C683 became exchanged by serine residues. All exchanges were validated by DNA sequencing.

4.1.2 Protein preparation

Expression of LAP2α constructs was carried out in E. coli strain BL21 (DE3) Rosetta (Novagen). Cultures were grown at 37 °C up to an OD₆₀₀ of 0.6 in Luria Bertani medium supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol. Induction of LAP2α^1-693 was achieved by addition of 0.6 mM isopropyl-β-D-thiogalactopyranoside (IPTG), while the truncated variant LAP2α^410-693 required 0.05 mM or total absence of IPTG in order to reduce inclusion body formation. Media were replenished with ampicilline after initiation of induction and cultures were grown at 25°C for approximately 18 additional hours. Cells were harvested by centrifugation at 2800×g and 4°C during 40 minutes. Bacterial pellets were resuspended and lysated in 0.1 M phosphate buffer pH 7.5, 0.1 M NaCl, 10 mM β-mercaptoethanol containing a commercial protease inhibitor cocktail (Boehringer). Lysis was achieved by sonication (Branson sonifier) at 4°C in the presence of lysozyme. DNAse and RNAse were added to a final concentration of 5 µg/ml. The homogeneate was clarified by centrifugation at 15000×g at 4°C during 40 minutes. The supernatant was applied to a Ni²⁺-chelating HiTrap column (GE Healthcare) equilibrated in lysis buffer with additional 40 mM imidazole. The eluate was dialyzed against a buffer containing 100 mM sodium phosphate pH 7.5, 100 mM NaCl and 10 mM DTT at 4°C for 16h, concentrated to about 20 mg/ml (according to A₂₈₀) and applied to a size exclusion chromatography column (Superdex 200 HiLoad 16/60PG; GE Healthcare) equilibrated in dialysis buffer. The protein so obtained was pure according to SDS-PAGE and mass spectrometry. All
purification procedures were carried out at room temperature. The protein was stored at 4°C at concentrations up to 15 mg/ml. SDS-PAGE was performed using 6% stacking and 12% resolving gels.

Overexpression of a cysteine-to-serine mutated variant, LAP2α<sup>410-693Cys→Ser</sup>, followed protocols as described for the corresponding wild type construct.

### 4.1.3 Mass spectrometry (MS)

The molecular mass of purified LAP2α<sup>410-693</sup> was determined on a TOF-mass spectrometer (TSQ-7000, Finnigan, San Jose, CA) with ESI-MS. The protein at a concentration of 10 mg/ml in 100 mM Na-phosphate pH 7.5, 100 mM NaCl, 10 mM DTT was diluted 1:100 with water prior to the measurement. Mass determination was carried out by Paul Jenö (Biozentrum, Basel).

For chemical modification of LAP2α<sup>410-693</sup>, protein solutions were adjusted by appropriate dilution with buffer to 1 mg/ml and 1-2 mM DTT. The buffer contained iodo-acetic acid (IAA) in various concentrations between 5 and 50 mM. Samples were incubated at room temperature or 37°C for 1 h followed by quenching with an at least double molar excess of DTT and subsequently digested with trypsin. The fragments were analysed by ESI-MS (peptide mapping). The proteases trypsin, Asp-N, Lys-C and V8 (Sigma Aldrich) were used to generate diverse peptide fragments to resolve the position of all cysteine residues. As a positive control a reaction mix was supplemented with 8 M urea before quenching and incubated for additional 1 h at 37°C. Urea containing samples were dialyzed against water overnight, digested with trypsin and analyzed by LC-MS.

### 4.1.4 Circular dichroism spectroscopy (CD)

Purified, concentrated LAP2α constructs were extensively dialysed against a buffer solution containing 100 mM sodium phosphate pH 7.5, 100 mM NaCl, 5 mM DTT and diluted prior to the experiment to a final concentration of 0.5-1.0 mg/ml with similar buffer but containing only 1 mM DTT. The far-UV circular dichroism (CD) spectra were measured at 4°C and 22°C on a circular dichroism spectropolarimeter (62A DS, AVIV) equipped with a temperature-controlled quartz cell of 1 mm path. Data were averaged from four scans and normalized for concentration, buffer and path.
length. Thermal denaturation curves were recorded by monitoring the change of ellipticity as a function of temperature at a fixed wavelength of $\lambda=230$ nm. The heating rate was 30 °C/h and data points were acquired for every degree centigrade.

4.1.5 Analytical Ultracentrifugation (AUC)

Sedimentation velocity and equilibrium measurements were performed on a Optima XL-A analytical ultracentrifuge (Beckman Instruments) equipped with an An-60 Ti Rotor at 20°C. Samples were dialyzed against final conditions or taken directly after size exclusion chromatography. Concentration was monitored by $A_{280}$ against reference buffer. A protein partial specific volume of 0.73 ml/g was used, while solution density was assumed to be 1.003 g/ml.

Sedimentation velocity data were obtained by centrifuging LAP2α samples at 54000 rpm for 46 min recording radial scans at $\lambda=277$ nm every 10 min. The data were analyzed with the program SEGAL21 (Machaizde and Lustig, 2006). All measurements were done in collaboration with Ariel Lustig (Biozentrum, Basel).

4.1.6 Transmission electron microscopy (TEM)

Purified recombinant LAP2α410-693 was taken from current preparations and diluted to appropriate protein concentrations with different buffer/salt solutions. Samples were mixed with glycerol to a final concentration of 30% directly before the experiment and sprayed onto freshly cleaved mica using an airstream. The dried specimens were rotary shadowed with platinum and carbon using an electron gun. The replicas were floated on distilled water and collected on formvar-coated grids. The replicas were examined in a Zeiss 9/2 120 kV transmission electron microscope. Samples were prepared and analyzed by Bohumil Maco.

4.1.7 Size exclusion chromatography with multi-angle light scattering (SEC-MALS)

The oligomeric state of LAP2α410-693 in solution was determined via SEC-MALS measurements performed on an ÄKTA explorer 10 system (GE Healthcare) connected to a tri-angle light scattering detector and a differential refractometer (miniDAWN
Tristar and Optilab, respectively; Wyatt Technology). A Superdex 75 HR 10/30 column (GE Healthcare) equilibrated in 0.1 M Na-Phosphate pH 7.5, 0.1 M NaCl, 10 mM DTT at a flow rate of 0.5 ml/min was used. A sample volume of 100 µl was injected at a concentration of 2 mg/ml. Data were processed using ASTRA software (Wyatt Technology) assuming a specific refractive index increment (dn/dc) of 0.185 ml/g. To determine the detector delay volumes and the normalization coefficients for the MALS detector, a BSA sample (Sigma A-8531) was used as reference. Neither despiking nor band broadening correction was applied and molecular masses were calculated based on the refractive index.
4.2 Results

4.2.1 Characteristics of LAP2α constructs

The work focused on the analysis of full-length LAP2α and a truncated, C-terminal construct, LAP2α<sup>410-693</sup>, hosting a lamin A/C binding domain between residues 615-693 (Dechat et al., 2000). Table 4.1 summarizes the biophysical properties of these protein constructs. All LAP2α forms contained cysteine residues that led to aggregation via formation of intermolecular disulfide bridges resulting from the oxidation of surface sulfhydryl groups. Hence, to prevent oxidation, the protein was kept under reducing conditions at all times during preparation. Figure 4.1 shows the number and position of cysteine residues potentially involved in oxidation events Table 4.1. The full-length protein contains ten cysteines and the C-terminal truncated domain LAP2α<sup>410-693</sup> contains six. Given that LAP2α is localized in the nuclear interior in vivo, all cysteine residues could be expected to be in a reduced state under native conditions.

The sequences of human and mouse LAP2α share 72% sequence identity (for an alignment see appendix 6.3). The mouse isoform contains four additional cysteines that are not present in the human form. Residue C683 in the human isoform is substituted by a histidine in the mouse sequence, whereas H640 is a cysteine in the mouse isoform. We hypothesize that these residues could be involved in the formation of a metal, possibly zinc, binding motif. However, no canonical metal binding motif could be identified using prediction tools on available sequences.

<table>
<thead>
<tr>
<th>Protein</th>
<th>aa</th>
<th>M&lt;sub&gt;r&lt;/sub&gt; [Da]</th>
<th>pI</th>
<th>No. Cys</th>
<th>Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAP2α&lt;sup&gt;410-693&lt;/sup&gt;</td>
<td>296</td>
<td>32477</td>
<td>6.7</td>
<td>6</td>
<td>C-term. His&lt;sub&gt;6&lt;/sub&gt;</td>
</tr>
<tr>
<td>LAP2α&lt;sup&gt;1-693&lt;/sup&gt;</td>
<td>703</td>
<td>76670</td>
<td>7.3</td>
<td>10</td>
<td>C-term. His&lt;sub&gt;6&lt;/sub&gt;</td>
</tr>
<tr>
<td>LAP2α&lt;sup&gt;410-693Cys→Ser&lt;/sup&gt;</td>
<td>296</td>
<td>32397</td>
<td>6.7</td>
<td>1</td>
<td>C-term. His&lt;sub&gt;6&lt;/sub&gt;</td>
</tr>
</tbody>
</table>
Chapter 4: Biophysical characterization of LAP2α

4.2.2 Protein preparation of LAP2α410-693

The yield of pure, soluble protein that was obtained by overexpression at low IPTG levels and temperature was around 5-10 mg per 1 L of E. coli culture. After purification, the soluble fraction was homogeneous as evidenced by SDS-PAGE and size exclusion chromatography (Figure 4.2). The solubility limit was reached at a concentration of about 20 mg/ml in PBS buffer.

In size exclusion chromatography, LAP2α410-693 elutes at 70.3 ml. Assuming a globular shape of the molecule, this is indicative of a trimeric state of the sample or potentially of an elongated dimer (see appendix 6.4 for calibration curve). To further investigate the oligomeric state of LAP2α410-693 we used analytical ultracentrifugation as well as size exclusion chromatography in connection with multi-angle static light scattering (SEC-MALS; see chapter 4.2.6), a technique that is independent of the molecular shape.
Figure 4.2: Size exclusion chromatography and final purity of LAP2α\textsuperscript{418-693}.

a) Size-exclusion chromatogram from a Superdex 200 (PG 16/60) column. The protein elutes as one single species at 70 mL. The small pre-peak at 60 mL contains a fraction of LAP2α and the \textit{E. coli} chaperone DnaK as identified by mass-spectroscopy. b) SDS-PAGE of elution fractions from size exclusion chromatography. A black bar indicates the composition of the main peak. Lane 1 shows the applied sample, lane 9 the low molecular weight standard. Lane 2 corresponds to the pre-peak shown also in a). LAP2α\textsuperscript{418-693} appears at an apparent molecular mass of 35 kDa. Lane 2 shows the presence of the chaperone DnaK at an apparent molecular mass of 69 kDa as determined by mass-mapping of gel bands.

A main impurity at an apparent molecular mass of 69 kDa in SDS-PAGE Figure 4.2b) was identified by mass-mapping of gel bands to be the \textit{E. coli} chaperone DnaK. The latter influences the partitioning of functional protein between soluble and insoluble fraction (Gonzalez-Montalban et al., 2006). It potentially binds to LAP2α populations to enhance their solubility. No further characterization was carried out in this respect since DnaK could be effectively removed by size exclusion chromatography.
The purified and concentrated protein showed a pronounced temperature dependence of solubility. At 4°C the solubility limit in a buffer containing 50 mM Tris pH 7.5, 100 mM NaCl, 5 mM DTT decreased to approximately 5 mg/ml and the protein gradually precipitated upon storage. Since this could result from a pH shift of the solution with temperature due to the properties of the Tris buffer, this was replaced with phosphate buffer. This alleviated to a certain extent the solubility of the sample in cold temperatures.

4.2.3 Overexpression and purification of full-length LAP2α\(^{1-693}\)

The amount of pure, soluble LAP2α\(^{1-693}\) was approximately 1 mg per 1 L E. coli culture – significantly less quantity than that obtained from LAP2α\(^{410-693}\). Only a small amount of LAP2α\(^{1-693}\) was found in the insoluble fraction after cell lysis and the full protein did not precipitate upon cooling or storage.

In size exclusion chromatography (Figure 4.3), LAP2α\(^{1-693}\) elutes at a volume corresponding to a high apparent molecular mass, possibly larger than 400 kDa (MW\(_{monomer}\) = 76670.3 Da). Although molecular weight estimates at elution volumes close to the column void are prone to large errors and cannot be interpreted reliably. This indicates that the full-length protein has a higher oligomeric state than its truncated version LAP2α\(^{410-693}\). Potentially it could also represent unspecifically aggregated material. The result stands in contrast to cross-linking experiments followed by semi-native PAGE that suggest that LAP2α\(^{1-693}\) forms dimers, trimers or tetramers in solution (Snyers et al., submitted manuscript; see appendix 6.5). No further attempts were made to determine the molecular mass of full-length LAP2α at that time.

Frequently, LAP2α\(^{1-693}\) yields a second, variable elution peak in size exclusion chromatography. This elutes close to the exclusion limit of the column (600 kDa) and indicates the presence of very high molecular mass species in the sample (Figure 4.3). The peak showed lower absorption at \(\lambda=280\) nm than at \(\lambda=260\) nm. Most likely, it contained DNA fragments resulting from nuclease treatment during cell lysis that have propagated through previous purification and dialysis steps and remained associated with LAP2α. To further investigate this we used electron microscopy (chapter 4.2.8).
Figure 4.3: Size exclusion chromatography of full-length LAP2α<sup>1-693</sup>.

a) Size exclusion chromatogram obtained on a Superdex 200 (16/60PG) column. Peak 1 has a high absorption at λ=260 nm and contains LAP2α based on detection by SDS-PAGE. Peak 2 contains LAP2α<sup>1-693</sup> multimers with an apparent molecular mass of > 400 kDa; b) SDS-PAGE. LAP2α<sup>1-693</sup> has an apparent molecular weight of 75 kDa. Lane 1 shows the low molecular mass standard, lane 10 the applied sample.

4.2.4 Mass spectroscopy on LAP2α<sup>410-693</sup>

ESI-MS was applied to investigate the exact length of the LAP2α protein products obtained by bacterial overexpression. The results indicate that LAP2α<sup>410-693</sup> had a molecular mass of 32351.0 Da. The measured mass is inferior, ΔM<sub>exp/theo</sub>=126.8, to that calculated from sequence data, M<sub>w</sub>=32477.8 Da. This can be speculated to be due to the removal of the N-terminal methionine (M<sub>Met</sub>=131.2 Da) in the E. coli system. Initiator methionines are frequently cleaved by methionine aminopeptidases (MetAPs)
if they are followed by small and uncharged residues (Alanine and serine in the case of LAP2α_{410-693}).

Full-length LAP2α^{1-693} could not be analyzed this way since the protein did not elute from the liquid chromatography (LC) column employed prior to sample injection. A reason for this might be its high oligomeric state and/or unspecific interaction with the column material (Paul Jenö, personal communication).

### 4.2.5 Circular dichroism spectroscopy on LAP2α_{410-693} and LAP2α^{1-693}

The far-UV CD spectrum (the peptide region) of a globular protein primarily reports on its secondary structure. Accordingly, CD-spectroscopy was used to gain information about the folding status of recombinant LAP2α products. The CD-spectra of LAP2α_{410-693} under native and denaturing conditions is shown in Figure 4.4a.

The position and intensity of the CD-bands reflect spectral features that are characteristic of a protein with a high α-helical content, i.e. a minimum at $\lambda=222$ nm and $\lambda=208$ nm.

\[
%\alpha-helix = \frac{[\theta_{mrw}]_{208\text{nm}} + 4000}{(-29000)} \times 100\%
\]

\[
%\alpha-helix = \frac{[\theta_{mrw}]_{222\text{nm}} - 3000}{(-39000)} \times 100\%
\]

**Equation 1:** Degree of α-helicity as derived from the far-UV CD-spectra.

$[\theta_{mrw}]$ : molar mean residue ellipticity at given wavelength

Using **Equation 1**, the α-helical content of the protein was determined to be 47 % (based on $[\theta_{mrw}]$ at $\lambda=208$ nm) and 57 % based on $[\theta_{mrw}]$ at $\lambda=222$ nm).
Figure 4.4: CD-spectra of LAP2\textsuperscript{410-693} and LAP2\textsuperscript{1-693}.

a) Far-UV CD-spectrum of native (black curve) and denatured (6 M urea, red curve) LAP2\textsuperscript{410-693}. Shown is the molar ellipticity after buffer correction. The insert shows the final protein purity (SDS-PAGE; lane M corresponds to the LMW-standard); b) Far-UV CD-spectrum of native (black curve) and denatured (6 M urea, red curve) LAP2\textsuperscript{1-693}. All spectra were recorded in 100 mM phosphate buffer pH 7.5, 100 mM NaCl, 1 mM DTT in a 1 mm cuvette. The samples were measured at a protein concentration of 0.5 mg/ml at 5°C.

Similar to LAP2\textsuperscript{410-693}, the full-length protein also shows an absorption minima at $\lambda=208$ and $\lambda=222$ nm (Figure 4.4b). The content of $\alpha$-helix for the full protein is about 15 % determined at $\lambda=208$ nm or 23 % at $\lambda=222$ nm. These percentages are equivalent to those of the shorter C-terminal construct, if the latter were recalculated to account for the full protein sequence, suggesting that no additional $\alpha$-helical regions are present in the protein (apart from the ~90 residues of N-terminal LEM-like and LEM domain). This is indicative of potentially unstructured regions.

In addition, to gain an insight into the stability of the protein, a heat denaturation experiment was recorded at $\lambda=230$ nm. Figure 4.5 illustrates that the heat unfolding of both LAP2\textsuperscript{1-693} and LAP2\textsuperscript{410-693} could be described by a two state model based on the sigmoidal shape of the transition curve. Two clear plateaus were reached at the beginning and the end of the experiment. However, for both proteins the transition to the unfolded state was irreversible and resulted in insoluble aggregates. Refolding under those conditions compatible with CD recording was probably troubled by the presence of the cysteine groups. The apparent transition temperatures recorded were 53°C for LAP2\textsuperscript{410-693} and 57°C for the full-length protein, indicating a high stability of the samples.
b) Temperature denaturation curves of LAP2α<sub>410-693</sub> (black curve) and LAP2α<sub>1-693</sub> (red curve) were recorded at a wavelength of λ=230 nm. b) LAP2α<sub>410-693</sub> (black curve) shows an apparent transition temperature of 53°C, whereas the full protein (red curve) denatures at 57°C (first derivative shown). All spectra were recorded in 100 mM phosphate buffer pH 7.5, 100 mM NaCl, 1 mM DTT in a 1 mm cuvette path.

4.2.6 Size exclusion chromatography with multi-angle light scattering on LAP2α<sub>410-693</sub>

Static light scattering in connection with refractive index measurements yield an accurate determination of the molecular mass of proteins in solution, which is independent of molecular shape or hydrodynamic parameters. The intensity of the scattered light is directly proportional to the weight–averaged molar protein mass and concentration based on Zimm’s formalism of the Rayleigh-Debye-Gans model (Zimm, 1948). The measurements are carried out online, directly after a size-exclusion column, which provides a fractionation of the sample population and aids assessing the possible presence of polydisperse states.

Analysis of LAP2α<sub>410-693</sub> by size exclusion chromatography combined with multi-angle light scattering (SEC-MALS) reveals a molecular mass of 66.3 ± 1 kDa. This is within excellent agreement of the calculated theoretical mass of 65 kDa for a dimer of this construct (Figure 4.6). Only one single oligomeric species could be detected in solution, indicating a high monodispersity of LAP2α<sub>410-693</sub> populations forming homo-dimers.
Figure 4.6: Analysis of LAP2α\textsuperscript{410-693} with SEC-MALS.

The normalized refractive index signal of the peak region of a size exclusion chromatogram (Superdex 75 HR 13/30) performed in 0.1 M Na-Phosphate pH 7.5, 0.1 M NaCl, 10 mM DTT is shown as solid line. The weight-averaged molecular mass measured inside the peak area at volume intervals is displayed as filled squares. A black line indicates the theoretical mass of a dimer.

4.2.7 Analytical Ultracentrifugation on LAP2α\textsuperscript{410-693}

In addition to SEC-MALS, the oligomeric state of LAP2α\textsuperscript{410-693} solutions was analyzed using ultracentrifugation. Samples were investigated at protein concentrations between 0.6 mg/ml and 2.5 mg/ml in Tris and phosphate buffer conditions. Table 4.2 summarizes the measured Mw. Samples at concentrations of 1.2 and 0.6 mg/ml showed a very low signal-to-noise ratio due to the interference of reducing agent in the sample. Due to their poor interpretability, those data were discarded.

The average molecular mass over all runs was 74 kDa ± 5 kDa (Figure 4.7). This indicates the presence of a dimer of LAP2α\textsuperscript{410-693} in solution and is in agreement with SEC-MALS data. No significant dependence on concentration was found between 1.2 and 2.5 mg/ml, i.e. the protein seems to form an isolated, non-aggregating dimeric particle. The sedimentation coefficient $S_{w,20^\circ C}$ of $4.2 \times 10^{-13}$s\textsuperscript{-1} resembles that of BSA ($S_{w,20^\circ C}= 4.3 \times 10^{-13}$s\textsuperscript{-1}) and points towards a globular overall shape.
Chapter 4: Biophysical characterization of LAP2α

Table 4.2: Ultracentrifugation experiments carried out on LAP2α<sub>410-693</sub>.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Buffer</th>
<th>Concentration [mg/mL]</th>
<th>Speed [rpm]</th>
<th>Mol. Mass [kDa]/Sedimentation coefficient [Svedberg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAP2α&lt;sup&gt;410-693&lt;/sup&gt;</td>
<td>20 mM phosphate pH 7.5 100 NaCl 5 mM DTT</td>
<td>1.2 0.6 0.6</td>
<td>15000 15000 54000</td>
<td>101&lt;sup&gt;a&lt;/sup&gt; 99&lt;sup&gt;a&lt;/sup&gt; 4.2 S</td>
</tr>
<tr>
<td>LAP2α&lt;sup&gt;410-693&lt;/sup&gt;</td>
<td>100 mM phosphate pH 7.5 100 NaCl 5 mM DTT</td>
<td>1.3 1.3 2.5</td>
<td>8000 13000 18000</td>
<td>77&lt;sup&gt;b&lt;/sup&gt; 68 75</td>
</tr>
<tr>
<td>LAP2α&lt;sup&gt;410-693&lt;/sup&gt;</td>
<td>100 mM Tris pH 7.5 500 NaCl 10 mM DTT</td>
<td>2.5 1.5 2.5 500 NaCl 10 mM DTT 20 mM CHAPS&lt;sup&gt;c&lt;/sup&gt;</td>
<td>13000 13000 18000 18000</td>
<td>70 68 75 75</td>
</tr>
</tbody>
</table>

<sup>a</sup> treated as outliers; not shown.

<sup>b</sup> interference optics were used.

<sup>c</sup> five times the critical micellar concentration (4mM).

Figure 4.7: Molar masses of LAP2α<sub>410-693</sub> calculated by sedimentation equilibrium.

LAP2α<sub>410-693</sub> shows a molecular mass of 68-77kDa at all concentrations indicating the presence of a dimer in solution (linear fit in red).

4.2.8 Transmission electron microscopy (TEM) of LAP2α<sub>410-693</sub> and LAP2α<sub>1-693</sub>

The goal of our TEM experiments was to monitor the dispersity of LAP2α samples in solution for future crystallization trials. Samples were investigated in different buffer conditions and at concentrations up to 2 mg/ml. Under all conditions assayed, LAP2α<sub>410-693</sub> showed a same globular shape and defined appearance (Figure 4.8a). Hence, the protein solutions appeared homogeneous, free of aggregate...
formation, and suitable for crystallization trials. These data, however, do not preclude that a different behavior of the protein might occur at concentrations higher than 2 mg/ml as routinely employed in crystallization (>10 mg/ml).

Occasionally, long rope-like objects with an approximate diameter of 5-8 nm could be seen among the globular LAP2α<sup>410-693</sup> specimens (Figure 4.8b). In some cases, branched or overlapping structures could also be detected, but did not co-localize with LAP2α<sup>410-693</sup>. Visualization of full-length LAP2α<sup>1-693</sup> revealed the presence of similar objects. In this case, they appeared much more pronounced, distributed over a larger area and apparently co-localizing with LAP2α<sup>1-693</sup> (Figure 4.9b). Purified material that was lysed in the presence of 500 mM NaCl did not contain such fibrous objects, indicating an ionic interaction. However, the LAP2α<sup>1-693</sup> sample so produced could be visualized as globular particles, inhomogeneous in their size distribution and exhibit certain clustering (Figure 4.9a), indicating that this sample is unlikely to be amenable to crystallization.

**Figure 4.8: TEM images after low-angle rotary shadowing of LAP2α<sup>410-693</sup>.**

a) The protein is in 20 mM sodium phosphate pH 7.5, 100 mM NaCl, 5 mM DTT at 1.7 mg/ml and appears homogeneously distributed on the grid; b) protein in 50 mM Tris pH 7.5, 100 NaCl, 5 mM DTT at 1.0 mg/ml. Several filamentous objects could be observed.

Lane 1: DNA-Marker (100 bp ladder; BIORAD)
Lane 2: DNA-Marker (Marker X, BIORAD)
Lane 3: LAP2α<sup>410-693</sup> (7 mg/ml)
Lane 4: LAP2α<sup>410-693</sup> (20 mg/ml)
Lane 5: BSA (10 mg/ml)

**Figure 4.9:**
Given that LAP2α is thought to complex DNA, we speculated that the filamentous formations observed by TEM might be DNA chains. Thus, we analyzed purified LAP2α^{410-693} using 5% CTG-agarose gel stained with ethidium-bromide (Figure 4.8c). This dye intercalates between aromatic base pairs forming the DNA polymer and gives rise to an increased fluorescence upon UV excitation. The pronounced fluorescence of LAP2α^{410-693} protein solutions is indicative of the presence of DNA fragments, which were probably generated during cell lysis, carried through affinity and size exclusion steps and eventually concentrate together with the protein. Yet, it remains to be determined if the protein binds specifically to them, particularly in its full-length form.

![Figure 4.9: TEM images after low-angle rotary shadowing of full LAP2α^{1-693}.

a) The protein was present in 50 mM Tris pH 8.0, 100 mM NaCl, 5 mM DTT at 1.0 mg/mL. 500 NaCl was used in the lysis step. LAP2α appears as locally clustered, globular object of approximately 20 nm.
b) The protein was present in 20 mM Phosphate pH 8.0, 250 mM NaCl, 5 mM DTT at 1.9 mg/mL.](image)

### 4.2.9 Oxidation of LAP2α protein solutions

During purification and biophysical characterization of LAP2α it became clear that both full LAP2α^{1-693} and the C-terminal region LAP2α^{410-693} form higher order, covalent oligomers via oxidation of their solvent-exposed sulphydryl groups in the absence of reducing agents. Overnight dialysis in the absence of reducing agent led to severe oxidation of the C-terminal domain LAP2α^{410-693} as shown by size exclusion chromatography and the fact that a large fraction of the protein precipitated. The elution
peak of the supernatant subjected to gel filtration shifted from 71 ml under reducing conditions close to the void volume of the size exclusion column \( (V_0 = 40 \text{ mL}) \), thus indicating a molecular mass higher than 600 kDa (Superdex 200; Figure 4.10a). Clearly two or more cysteine residues must be exposed on the protein surface to form such higher order aggregates.

In order to detect protein oxidation by means of SDS-PAGE, it was necessary to avoid reducing conditions during electrophoresis, since denaturing conditions together with the reducing agent reduced the oligomers to their monomeric components and the detection of higher order, covalent aggregates became troubled. Hence, LAP2α\(^{410-693}\) samples were dialysed against a buffer containing 50 mM Tris pH 7.5, 100 mM NaCl and an identical buffer containing additionally 20 mM β-ME. The protein precipitated slightly after 16 h dialysis under both conditions. After centrifugation, the protein pellets were washed with dialysis buffer in the absence of a reducing agent, re-dissolved in buffer plus urea and loaded onto the gel after mixing with a sample buffer containing β-ME or with a buffer containing nβ-ME (Figure 4.10b). The protein pellet resulting from precipitation after four days of non-reductive dialysis was also checked under the same conditions. In both pellet fractions covalent aggregation was detectable. Accordingly, aggregation disappeared in the presence of reducing agent in the sample buffer. The same behaviour was observed in LAP2α\(^{1-693}\) (data not shown).
Chapter 4: Biophysical characterization of LAP2α

Figure 4.10: LAP2α<sub>410-693</sub> protein solutions show a propensity to oxidate.

a) Size exclusion chromatography of LAP2α<sub>410-693</sub> on a Superdex 200 16/60PG column after overnight dialysis against 100 mM sodium phosphate pH 7.5, 100 mM NaCl (non-reducing condition). The elution peak is shifted from 71 ml to the void volume of the column indicating a high oxidation state. 

b) SDS-PAGE of LAP2α<sub>410-693</sub> pellet fractions under reducing (lane 1-3) and non-reducing (lane 4-7) conditions. The pellet fractions contain high molecular mass aggregates that can be reduced.

Lane 1: pellet after non-reductive dialysis
Lane 2: pellet after reductive dialysis
Lane 3: supernatant after reductive dialysis
Lane 4: LMW-marker
Lane 5: pellet after non-reductive dialysis (4 days) and no β-ME in sample buffer
Lane 6: pellet after non-reductive dialysis (1 day) and no β-ME in sample buffer
Lane 7: supernatant after reductive dialysis and no β-ME in sample buffer

Spontaneous disulfide cross-linking is a distinct source of heterogeneity in LAP2α samples, that will most likely interfere with its crystallization since this involves timescales in the days-to-weeks range, perhaps even months. In an attempt to overcome this problem by cysteine-to-serine residue exchanges, we pursue the identification of exposed cysteines involved in the oxidation process by chemical modification with an alkylating agent followed by mass-spectroscopic analysis.

4.2.10 Carboxymethylation of LAP2α<sub>410-693</sub> with iodo-acetic acid

Upon incubation of LAP2α<sub>410-693</sub> with the carboxy-alkylation reagent iodo-acetic acid (IAA) at room temperature, the molecular mass of the protein increased by approximately 179 Da according to full mass determination. This was indicative of a triple alkylation of accessible cysteine residues (Mw<sub>carboxymethyl</sub> = 59 Da). At 37°C a mixed
population of three and four (approximate mass increase 230 Da) modified residues was observed (Table 4.3). These data clearly suggest the presence of at least three exposed sulfhydryl groups involved in protein oxidation.

Table 4.3: ESI-MS results after carboxymethylation of LAP2α<sub>410-693</sub> with iodo acetic acid (IAA)

<table>
<thead>
<tr>
<th>Protein</th>
<th>IAA concentration</th>
<th>Molecular mass (mass shift) [Da]&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Interpretation&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAP2α&lt;sub&gt;410-693&lt;/sub&gt;, 1 mg/ml</td>
<td>20 mM IAA</td>
<td>32530 (Δ179)</td>
<td>3x modified</td>
</tr>
<tr>
<td></td>
<td>1 h, RT standing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAP2α&lt;sub&gt;410-693&lt;/sub&gt;, 1 mg/ml</td>
<td>25 mM IAA</td>
<td>32580 (Δ230)</td>
<td>4x modified</td>
</tr>
<tr>
<td></td>
<td>1 h, 37°C shaking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAP2α&lt;sub&gt;410-693&lt;/sub&gt;, 1 mg/ml</td>
<td>10 mM IAA</td>
<td>32526 (Δ174)</td>
<td>3x modified</td>
</tr>
<tr>
<td></td>
<td>1 h, 37°C shaking</td>
<td>32582 (Δ232)</td>
<td>4x modified</td>
</tr>
</tbody>
</table>

<sup>a</sup>The theoretical mass of the protein is 32477 Da, the observed experimental mass is 32351 Da (cleavage of N-terminal methionine). <sup>b</sup>Each modification contributes with 59 Da.

To identify those cysteine residues exposed and accessible to IAA, carboxymethylated LAP2α<sub>410-693</sub> was subjected to protease digestion followed by mass-mapping of resulting proteolytic fragments. Four separate reactions with different IAA concentrations were tested as well as a control probe with urea-denatured protein (Figure 4.11). The data indicate that C517, C560, C569 and C657 are modified by the reagent and, hence, most likely exposed. The two cysteines, C560 and C569, initially could not be resolved since both were located in a same tryptic fragment. Mass estimations indicated that one of them must be exposed. In a separate reaction and using different proteases, both residues were found to be modified, so that it was not possible to identify the residue with higher propensity to oxidation. By contrast, C628 was never found modified. In a series of reactions, C683 was only detected once as modified. In a control reaction in the the presence of urea, all cysteines were found to be modified Figure 4.11 and Table 4.4.
Figure 4.11: ESI-MS results after alkylation of LAP2α410-693 by iodo acetic acid (IAA).

Indicated are modified, unmodified and not determined cysteines as deduced from fragment mass after alkylation and proteolytic digest. They were generated in four reactions at various IAA concentrations (Table 4.4). A cross indicates the detection of modified cysteine containing fragments, unmodified fragments are marked by a check.

Table 4.4: Mass mapping results for LAP2α410-693 after carboxymethylation

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Molecular Mass</th>
<th>IAA Conc</th>
<th>Protease</th>
<th>modified</th>
<th>unmodified</th>
<th>unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>32526 Da</td>
<td>10 mM</td>
<td>trypsin</td>
<td>C517</td>
<td>C628</td>
<td>C657</td>
</tr>
<tr>
<td></td>
<td>Δ174 (3x)</td>
<td>1 h</td>
<td></td>
<td>C560 or C569</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32582 Da</td>
<td></td>
<td></td>
<td>C683</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Δ232 (4x)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5 mM</td>
<td>1 h</td>
<td>Asp-N</td>
<td>C517</td>
<td></td>
<td>C683</td>
</tr>
<tr>
<td></td>
<td>Lys-C</td>
<td></td>
<td>V8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>20 mM</td>
<td>1 h</td>
<td>Asp-N</td>
<td>C517, C569</td>
<td>C560</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys-C</td>
<td></td>
<td>V8</td>
<td>C517</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>50 mM</td>
<td>1 h</td>
<td>Asp-N</td>
<td>C517</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lys-C</td>
<td></td>
<td>V8</td>
<td>C560, C569</td>
<td></td>
<td></td>
</tr>
<tr>
<td>control</td>
<td>5 mM</td>
<td></td>
<td>Asp-N</td>
<td>C517, C657, C683</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 M Urea</td>
<td></td>
<td>Lys-C</td>
<td>C517, C560, C569</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 h</td>
<td></td>
<td>V8</td>
<td>C628, C517, C628, C657</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The full intact molecular mass was only determined for the first reaction.
4.2.11 Generation of a cysteine-to-serine mutant of LAP2α^{410-693}

To avoid protein oxidation a mutated variant of LAP2α^{410-693}, LAP2α^{410-693Cys→Ser}, was created in which five of the six cysteine residues were substituted by serine. Given that C628 did not appeared exposed in alkylation studies, it was not modified. Overexpression of LAP2α^{410-693Cys→Ser} was carried out under conditions similar to those of applied to the wild-type. However, a dramatic decrease of solubility was observed in this case. Only a small amount of the protein was present in solution after affinity purification (data not shown). One reason for this might be that mutation of five cysteine residues to serine potentially impaired the formation of a possible metal binding motif leading to the observed protein insolubility. Alternatively serine substitution might have interfered with protein folding in a non-metal dependent way, if for example buried or partially buried positions required mutation into alanine residues instead.

4.3 Discussion

LAP2α is one of six alternatively spliced products of the LAP2 gene, which share a common N-terminal region. In contrast to the other isoforms, which also share domains in their C-termini, LAP2α has a large unique C-terminal region that contains binding sites for chromatin, A-type lamins, and retinoblastoma protein. Full-length LAP2α^{1-693} and its C-terminal region LAP2α^{410-693} could be produced in soluble form through recombinant expression in *E. coli*. Furthermore, samples of a high degree of purity could be obtained by chromatographic procedures, making their biophysical characterization feasible. LAP2α forms higher order structures containing multiple molecules *in vivo* and complex formation is primarily mediated by its C-terminus (see manuscript in appendix 6.5). We have reliably determined the oligomerization state of LAP2α^{410-693} by ultracentrifugation and multi-angle light scattering. Consistently, both techniques have proven LAP2α C-terminal fragment to be dimeric. Furthermore, recombinant full-length LAP2α forms high molecular weight oligomers *in vitro* in agreement with data from cell preparations (appendix 6.5).

CD-spectroscopy revealed that C-terminal LAP2α^{410-693} is well folded and mainly composed of α-helical secondary structure elements under buffer conditions close to physiological. An apparent transition temperature of 55-57°C indicates that both the
full-length sample and its truncated C-terminus are stable protein products. Despite this apparent fold stability, the truncated construct LAP2α410–693 showed a high tendency to precipitation upon storage in solution and at cold temperatures if buffers other than sodium phosphate were employed. This could have resulted from a pH shift of the solutions upon temperature changes when buffers, such as Tris-HCl, were assayed. An alternative reason for a phosphate-dependent solubility might be a strong affinity of the protein for phosphate groups, which might provide additional charges thereby increasing the protein solubility. Binding of phosphate groups seems likely given the proposed direct interaction of LAP2α with chromatin (Vlcek et al., 1999). In line with this observation and using electron microscopy, preparations of both constructs were found to contain filamentous formations, most likely consisting of DNA. In agreement, LAP2α410–693 solutions showed a strong fluorescence after reaction with ethidium-bromide, indicating the presence of DNA in the preparations. Solutions of full-length LAP2α appeared to be even more populated by DNA fibers as indicated by the more abundant and longer filaments observed in TEM images compared to C-terminal fragment. Taken together, these data suggest that phosphate/DNA binding sites in the scaffold LAP2α become occupied in vitro either by buffer ions or by DNA fragments mimicking the native conditions of the protein in the cell nucleus.

Both LAP2α constructs under investigation contain cysteine residues that lead to aggregation by oxidation. High molecular mass aggregates were detected in size exclusion chromatography after non-reducing dialysis. They also became evident by the presence of SDS-resistant species in SDS-PAGE under non-reducing conditions. The addition of reducing agents like β-mercaptoethanol and DTT as well as the degassing of buffers and their subsequent flushing with nitrogen, effectively prevented aggregate formation and was indispensable for protein preparation and storage. These procedures, however, can only maintain a reducing environment for a limited period of time and thus, they cannot ensure the long-term monodispersity of the sample expected to be required for crystallization. Thus, to overcome oxidation on permanent basis, sample variants were produced, where cysteine residues were exchanged by serines. However, the overexpression of mutated variants did only result in minimal amounts of soluble protein, suggesting a disruption of the fold. Although the basis of such disruption is currently unknown, the possible alteration of a zinc-binding motif cannot be discarded even if such motif could not be predicted from LAP2α sequence data.
The initial objective of this work was to obtain a crystallographic model for the C-terminal, lamin binding domain of LAP2α. For successful crystal growth, monodisperse protein solutions that are stable over time at sufficiently high concentration are desirable. Sources of protein heterogeneity such as chemical impurities, multiple oligomeric states and conformational flexibility resulting from poorly ordered regions are likely to impair the formation of well-ordered crystals. LAP2α suffered from several of these problems in combination and accordingly crystallization attempts to this date have been unsuccessful. The current characterization, however, will serve as basis for the further refinement of protocols directed to achieve LAP2α samples amenable to crystal formation.
Chapter 5

Conclusions
5 Conclusions

Titin is the longest molecule known to date and pivotal component of the sarcomere, a cellular compartment that requires a unique biochemical and structural composition to adapt to an extreme biomechanical environment. The molecule utilizes the unique mechanical and protein interaction properties of the Ig and FnIII fold to constitute spatially isolated, functional modules involved in structural integrity (Z-disc, M-line), elasticity (I-band), molecular templating (A-band) and signaling (Z-disc, I-band).

Similar to poly-domain structures from the cell adhesion proteins fibronectin (Leahy et al., 1996, Sharma et al., 1999), tenasin (Lundell et al., 2004) to the Z1Z1 Ig tandem of titin (Marino et al., 2006; Zou et al., 2006) individual subunits of A168-A170 are arranged in an extended manner. We attributed the limited conformational freedom to hydrophobic clustering at the interface and to the absence of particular linker residues. From our studies, several molecular features can be identified that exemplify how the titin scaffold confers specificity to its distinct regions. At first, the potential binding site for MuRF-1 on titin A168-A170 involves cooperative interactions of all three domains in form of a surface groove. Thus, the inter-domain orientation is a critical parameter in the formation of a specific interface. In addition, the presence of a unique helical insertion between β-strand A and A` in IgA169, i.e. in the central region of the putative interface, seems to be directly involved in MuRF-1 binding. By contrast, the only other structurally characterized interface of titin, the Z1Z2-telethonin complex, despite also composed of two Ig domains in an extended conformation, possesses different domain orientation and moderate conformational freedom across domains as a result of a longer linker length. Apart from that, binding to telethonin is mediated through β-strand cross-linking and involves exclusively sequence-independent main-chain interactions. Yet again, specificity is created by cooperative binding of telethonin to both domains. Thus, although composed of highly conserved Ig/FnIII frameworks, both interface regions are structurally unrelated and represent “insulated” components in two distinct functional modules of the Z-disc and the M-line.

What is more, titin utilizes a super-repeat pattern covering different regions of the filament to modulate its mechanical properties as well as most likely its binding specificity. Through sequence comparison we have been able to correlate the presence of a sequence motif in the FG turn of the highly conserved IgA169-FnIII A170 interface that
seems to be generic for the whole A-band C-zone and is absent in the D- and P-zone. Potentially, this observed repetitive pattern is involved in directing the binding of MyBp-C and consequently the myosin filament to a particular A-band region. Moreover, MuRF-1 is a member of the TRIM family of proteins that is known to form homo-/heterooligomers. By means of ITC, we have mapped the binding site of titin to the central region of the predicted coiled-coil region on MuRF-1.

We have also determined the crystal structure of MuRF-1 B-box and its oligomeric state. Unexpectedly we found that, structure adopts RING-finger-like topology and exists in a dimeric form in solution. Clearly, an additional level of specificity of the interaction to titin can be induced if oligomerization of the titin ligand MuRF-1 is required for binding. Both, the CC and the B2 are involved in homo-oligomerization of MuRF-1. Given the serious consequences of MuRF-1 ubiquitination on its sarcomeric targets, all of the above factors aid in increasing the selectivity of interactions, thereby prohibiting unspecific binding events along the titin scaffold.

Finally we also show that the scaffolding events at the nuclear platform LAP2α also critically depend on homo-oligomerization of the protein.
Chapter 6

Appendix
6 Appendix

6.1 Sequence alignments of A-band Ig and FnIII domains

Figure 6.1: Sequence alignment of titin Ig domains of A-band short super-repeat.

Secondary structure elements (based on: Kabsch and Sander, 1983) are indicated for structure of Ig A168. Colour coding based on equivalence (highlighted in red are 100% identical residues). Alignment created with the program ESPRIPT (Gouet et al., 1999; http://espript.ibcp.fr/ESPript/ESPript/).

Figure 6.2: Sequence alignment of titin Ig domains of A-band long super-repeat.

Secondary structure elements and colour coding as in Figure 6.1.
Figure 6.3: Sequence alignment of titin FnIII domains of A-band short super-repeat.

Secondary structure elements are indicated for FNIII A170. Colour coding as in Figure 6.1.
Figure 6.4: Sequence alignment of titin FnIII domains of A-band short super-repeat.
Secondary structure elements are indicated for FNIII A170. Colour coding as in Figure 6.1.
6.2 Sequence alignments of B-box B1 and B2 subtypes

Figure 6.5: Sequence alignment of B-boxes of subtype B2.

Colour coding as follows: red, negatively charged residues (D, E); blue, positively charged residues (K, R); green, hydrophobic residues (L, I, V, M, A, F, W, P). Secondary structure elements as in MuRF-1 B2. The “cross-brace” zinc ligation scheme is indicated underneath. In bold: structurally characterized B2 domains.
Figure 6.6: Sequence alignment of B-boxes of subtype B1.

Colour coding as in Figure 6.5. Secondary structure elements as in MID-1 B1 (Massiah et al., 2006).
6.3 Sequence alignment of human and mouse LAP2α

Figure 6.7: Sequence alignment of human and mouse isoforms of LAP2α.

LEM-like and LEM domain are indicated as determined by Cai et al., 2001. The LAP2α unique C-terminus starts at residue 187 (indicated by an arrow). The full protein contains ten cysteines in the mouse isoform. The C-terminal construct LAP2α<sub>410-693</sub> (indicated by an arrow) contains six cysteines in the mouse isoform. Highlighted in cyan are histidine/cysteine substitutions.
6.4 Calibration curve of Superdex 200 (16/60PG)

Figure 6.8: Calibration curve of size exclusion column superdex 200 (16/60 PG).
Indicated are molecular mass estimates for LAP2α\textsubscript{1-693} and LAP2α\textsubscript{410-693} (chapter 4.2.2). The chart equation is given.

6.5 Manuscript in preparation: Homo-oligomerization of LAP2α requires its C-terminal domain and is unaffected by a disease-causing mutation
HOMO-OLIGOMERIZATION OF LAMINA-ASSOCIATED POLYPEPTIDE 2-ALPHA REQUIRES ITS C-TERMINAL DOMAIN AND IS UNAFFECTED BY A DISEASE-CAUSING MUTATION.

Luc Snyers¹, Sylvia Vleen², Thomas Dechat³, Michael Mrosek³, Barbara Korbei², Andreas Gajewski³, Olga Mayans³, Christian Schöfer¹ and Roland Foisner²,⁴

From the ¹Center for Anatomy and Cell Biology, Medical University of Vienna, Schwarzenplanerstrasse 17, A-1090 Vienna, ²Max F. Perutz Laboratories, Medical University Vienna, Dr. Bohr-Gasse 9, A-1030 Vienna, Austria, and ³Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland.

Running Title: LAP2α self-interaction

²Address correspondence: Dr. Roland Foisner, Max. F. Perutz Laboratories, Department of Medical Biochemistry, Medical University of Vienna, Dr. Bohrgasse 9, A-1030 Vienna, Austria, Tel 43-1-4277-61680, FAX 43-1-4277-9616, E-mail roland.foisner@meduniwien.ac.at

The nucleoplasmic protein, Lamina-associated polypeptide (LAP) 2α, is one of six alternatively spliced products of the LAP2 gene, which share a common N-terminal region. In contrast to the other isoforms, which also share domains in their C-termini, LAP2α has a large unique C-terminal region that contains binding sites for chromatin, A-type laminas, and retinoblastoma protein. By immunoprecipitation analyses of LAP2α complexes from cells expressing differently tagged LAP2α proteins and fragments, we demonstrate that LAP2α forms higher order structures containing multiple LAP2α molecules in vivo, and that complex formation is mediated by the C-terminus. In vitro cross-linking of LAP2α complexes revealed the existence of stable homo-oligomeric structures of 200-250 kDa. In vitro binding studies and multi-angle light scattering of purified LAP2α fragments showed homo-dimerization of LAP2α, involving the most C-terminal fragment comprising residues 410-693. Finally we show that, in contrast to the LAP2α-lamin A interaction, its self-association is not affected by a disease-linked single point mutation in the LAP2α C-terminus.

Key Words: chromatin, laminas, nuclear organization, homo-oligomerization

INTRODUCTION

The nuclear envelope comprises the inner and outer nuclear membranes, the nuclear pore complexes and the nuclear lamina, which underlies the inner nuclear membrane (1,2). The nuclear lamina is the major structural framework in the nucleus of multicellular eukaryotes and is composed of a filamentous meshwork of type V intermediate filament proteins, the laminas. B-type lamins, encoded by two human genes (LMNB1 and LMNB2) are essential for cell viability. In contrast, the four A-type lamins (A, C, C2 and Adelta10), representing splicing isoforms of the LMNA gene, are dispensable for viability of individual cells but have crucial functions in development and tissue organization after birth (3,4).

In addition to the lamins, the nuclear lamina contains a number of integral membrane proteins of the inner nuclear membrane, the best characterized of which are the Lamin B receptor, Lamina-associated polypeptide (LAP)1 and the three LEM-domain containing proteins LAP2β, emerin and MAN1 (5,6). All these proteins interact with lamin A/C and/or B and contribute to anchorage of the nuclear membrane to the lamina. The LEM-domain, a conserved 40 amino acid motif located near the N-terminus of the LEM family proteins, interacts with the DNA-binding protein Barrier-to-Autointegration Factor (BAF) and thus mediates the binding of these proteins to chromatin (7). In LAP2 proteins a LEM-like segment at the very N-terminus has been shown to interact with DNA directly (8).

The family of LAP2 proteins includes six alternatively spliced isoforms derived from the same gene (9). Most LAP2 isoforms are closely related structurally and functionally and are localized to the inner nuclear membrane, such as LAP2β. In contrast, LAP2α shares only the N-terminal 187 amino acids with the other
isoforms, including the LEM and LEM-like domains, but otherwise possesses a unique 506 amino acid C-terminal region without a transmembrane domain (Fig. 1A), encoded by one large exon found only in mammals (10).

LAP2α is exclusively located in the nucleoplasm in interphase and interacts with lamin A/C (11) and hypophosphorylated retinoblastoma protein (pRb) via distinct C-terminal domains (12,13). The LAP2α-lamin A/C-pRb complex is thought to regulate cell proliferation and differentiation in adult stem cells (3,12). During mitosis, LAP2α dissociates from chromosomes in a phosphorylation-dependent manner and is redistributed throughout the mitotic cytoplasm, like most nuclear lamina components (14,15). However, during anaphase, LAP2α associates with the telomeres of separated sister chromatids and subsequently forms stable structures associated with decondensing chromatin before the nuclear envelope is formed (14). While LAP2α can interact with BAF and DNA via its common LEM and LEM-like motifs, its C-terminus was shown to be essential and sufficient for chromatin association during mitosis (16,17). Intriguingly, a mutation causing an amino acid substitution (Arg690 to Cys) near the C-terminus of LAP2α has been associated with dilated cardiomyopathy (DCM) (18), a condition also known to be caused by mutations in the LMNA gene (5). The mutation altered the observed LAP2α interaction with A-type lamins in vitro and may represent a rare cause of DCM.

In this study, we show that LAP2α is engaged in homo-oligomerizaton via its unique C-terminal domain and that the C-terminal half of this domain forms stable dimers in vitro. These homo-oligomeric units might constitute the basis for the formation of higher-order structures containing LAP2α and other proteins.

**EXPERIMENTAL PROCEDURES**

Reagents - Mouse monoclonal anti-LAP2 (6E10) and rabbit polyclonal anti-myc antibodies were obtained from Abcam. Rabbit anti-LAP2α antiserum and monoclonal anti-LAP2α antibody 15/2 and 12 were described (16,17). Polyclonal anti-GFP antibodies were from Clontech Laboratories Inc; polyclonal anti-DsRed antibody from BD Biosciences; anti-Protein C affinity matrix (mouse monoclonal antibody HPC4, immobilized) and mouse monoclonal anti-myc antibody (9E10) from Roche Diagnostics; protein G-agarose conjugate from Sigma.

**Vector construction and expression of recombinant proteins - cDNA coding for monomeric Red Fluorescent Protein 1 (mRFP1) was amplified from pRSETB-mRFP1 (a kind gift of Dr. R. Tsien, Howard Hughes Medical Institute, University of California, San Diego) and inserted into the vector pCI (Promega). The resulting plasmid was used to create N-terminally tagged mRFP1-LAP2α and mRFP1-LAP2α(188-693), by ligating amplified fragments of LAP2α in frame with mRFP1. mRFP1-LAP2α and mRFP1-LAP2α(188-693) were then excised and ligated blunt into the EcoRV site of eukaryotic expression vector pEF-puro.PL3 (19). Plasmids pEF.mRFP1-LAP2α and pEF.mRFP1-LAP2α(188-693) were transfected into HeLa cells using lipofectamine (Invitrogen), and clones were selected in 2 μg/ml puromycin.

For the construction of PC-myc-mRFP1-LAP2α(188-693) and PC-myc-LAP2α(188-693), a sequence coding for Protein C epitope EDQVDPRLDGK fused to the myc-tag EKLISEEDL was inserted into the EcoRV site of pEF-puro.PL3. The resulting plasmid was linearized with EcoRV and ligated in frame with mRFP1-LAP2α(188-693) or a fragment encoding LAP2α(188-693).

For GFP-LAP2α carrying the DCM mutation, vector gAG43 was constructed by shuttling wild-type LAP2α cDNAs from gAG41 (18) into GFP-destination vector pcDNA-DEST53 by the LR-reaction (Invitrogen). The vector pAG39 containing mutated LAP2α cDNA was generated from pSV5 using the Quick-ChangeTM site-directed mutagenesis kit (Stratagene) using the following primers:

5'-GGAGGAGAAGTATGTCAAGTTAATGAAAGTGGAAATAAACAC-3', and 5'-GTGTTAAAACCATTTTAAATATTGAATTCGATCTTCCTCCGC-3'. The bacterial expression vectors encoding His-tagged LAP2α fragments are described elsewhere (16).

For bacterial expression, proteins were expressed in E. coli strain BL21(D3) using the inducible T7 RNA polymerase-dependent pET vector system as described previously (11,16). Protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside for 3 h. Bacteria were harvested by centrifugation at 4,000 rpm for 5 min (Heraeus Megafuge, 1.0R) and lysed in one-
tenth of the original culture volume of Tris buffer (20 mM Tris-HCl, pH 8, 500 mM NaCl, 5 mM imidazol, 1 mM dithiothreitol, protease inhibitors) by freezing and thawing and addition of 0.1 mg/ml lysozyme, 0.1% Triton X-100, 10 mM MgCl₂, 50 µg/ml DNase, and 20 µg/ml RNase. Following a 30 min incubation at 30°C the samples were centrifuged for 10 min at 14,000 rpm and pellets resuspended in one-tenth of the original culture volume of Tris buffer plus 7 M urea and incubated for 1h at room temperature. Cell lysates were centrifuged at 45,000 rpm –for 30 min, and supernatants were aliquoted and stored at -20°C. If fragments were soluble, urea was added directly to the cell extract prior to centrifugation at 14,000 rpm. Renaturation of recombinant proteins was achieved by dialyzing twice against KHM buffer (78 mM KCl, 50 mM HEPES, pH 7.4, 8.4 mM CaCl₂, 10 mM EGTA, 4 mM MgCl₂, 1 mM dithiothreitol) and cleared by centrifugation at 4000 rpm for 5 min. For light scattering, expression of LAP2α (410-693) was induced with 0.05 mM IPTG at 25°C for 18 h. Bacterial pellets were resuspended in 100 mM phosphate buffer, pH 7.5, 100 mM NaCl, 10 mM betamercaptoethanol, incubated in the presence of a commercial protease inhibitor cocktail (Boehringer), lysozyme, DNase and RNase and finally sonicated. The homogenate was centrifuged at 15,000g at 4°C for 40 min and the supernatant applied to a Ni²⁺-chelating HiTrap column (GE Healthcare) equilibrated in lysis buffer plus 40 mM imidazole. Elution used 200 mM imidazole. The elutant was dialyzed against 100 mM sodium phosphate pH 7.5, 100 mM NaCl, 10 mM DTT at 4°C and further subjected to gel filtration (Superdex 200 HiLoad 16/60PG; GE Healthcare). The protein so obtained was pure according to SDS-PAGE and mass spectrometry.

**Immunoblotting** - Cells extracts in 20 mM Tris-HCl, pH 8.0, 130 mM NaCl, 1% Triton X-100, and the protease inhibitors aprotinin, leupeptin and pepstatin (10 mg/ml each, Sigma) or Protease Inhibitor Cocktail Tablets (Roche Diagnostics) were centrifuged at 13,000g for 10 min and pellet and supernatant fractions were analyzed on 10% polyacrylamide gels and electrotransferred to nitrocellulose (Schleicher and Schuell Inc.). Blots were blocked for 60 min in Tris-buffered saline, pH 8.0 0.1% bovine serum albumin, and 5% non-fat dry milk, incubated with primary antibody in Tris-buffered saline, 0.1% bovine serum albumin for 60 min, and with anti-mouse (Sigma) or anti-rabbit (DAKO) alkaline phosphatase-conjugate, and visualized using BM Purple AP Substrate, Precipitating (Roche Diagnostics).

Semi-native gels contained SDS only in the running buffer (0.1% SDS); no SDS was present in the sample buffer or the polyacrylamide gels themselves.

**Immunoprecipitations** - Confluent cell monolayers in 10 cm dishes were lysed in TNCT (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM Ca²⁺, 1% Triton X-100 and protease inhibitors) and centrifuged for 10 min at 13,000g. Supernatants were incubated with 25 ml anti-Protein C matrix for 1h. Bound complexes were washed 4 times with TNCT and eluted by boiling in SDS-PAGE sample buffer. Samples were processed for immunoblotting. Alternatively, supernatants were divided in two parts; one half was incubated on ice with monoclonal anti-myec (2 mg) or monoclonal anti-LAP2 (15 mg) antibody and the other half with a control monoclonal antibody (mouse monoclonal anti-HRV2, a kind gift of Dr. Blaas, Max Perutz Laboratories, Vienna) for 30 min. Immunocomplexes were bound to 12.5 ml Protein G-agarose for 1h, washed with TNCT and eluted in sample buffer.

**Immunofluorescence microscopy** - Cells on coverslips were fixed with 3% paraformaldehyde in PBS and quenched in 50 mM NH₄Cl. They were permeabilized in 0.1% Triton X-100, blocked in 5% FCS and incubated with monoclonal anti-mye antibody at 1 mg/ml in PBS containing 1% FCS. The secondary antibody was anti-mouse Alexa 488 (Molecular Probes). Preparations were mounted using Citifluor AP1 (PLANO) and examined with a Nikon Eclipse 800 fluorescent microscope.

**Chemical Cross-linking** - HeLa cells were resuspended in KHM buffer and homogenized on ice by pressing the suspension 10-15 times through a metal ball cracker (EMBL, Heidelberg). Cell lysates and dialyzed recombinant proteins were mixed with various concentrations of cross-linking agent Dithiobis(succinimidylpropionate) (DSP, Pierce) for 2h on ice, and the reaction was stopped by quenching free active groups with 50 mM Tris-HCl, pH 6.8, for 1h on ice. Samples were
analyzed by SDS-PAGE in the presence or absence of 100 mM DTT.

**Blot overlay assays** - Recombinant polypeptides were resolved by SDS-PAGE and blotted to nitrocellulose membranes as described (11). Membranes were stained with Ponceau S, washed with PBS containing 0.05% Tween 20 and incubated in overlay buffer (10 mM Heps, pH 7.4, 100 mM NaCl, 5 mM MgCl₂, 2 mM EGTA, 0.1 % Triton X-100, 1 mM DTT) for 1 h. After blocking with 2 % BSA in overlay buffer, membranes were probed 3 h with 35S-methionine-labeled polypeptides diluted 1:50 in overlay buffer plus 1% BSA. Membranes were washed extensively with overlay buffer and bound proteins were detected by autoradiography.

**In vitro transcription/translation and GFP-immunoprecipitation** - 150μl protein A sepharose were coupled overnight with 6 μl polyclonal anti-GFP antibodies and 150μl protein G sepharose with 1 ml monoclonal anti-LAP2 antibody 12. [35S]-labeled wild-type or mutated LAP2α were expressed by in vitro transcription/translation using the TNT Quick Coupled Transcription/Translation reaction mix (Promega) either alone or together with GFP-tagged wild-type LAP2α (pSV5, pSV5 and gAG43, pAG39, pAG39 and gAG43). After incubation at 30°C for 3 h binding buffer (50 mM HEPES pH 7.4, 50 mM NaCl, 5 mM MgCl₂, 1 mM EGTA, 0.1 % Triton, 1 mM DTT and 1 mM PMSF) was added to 225 μl. For pre-clearing samples were incubated with 20 μl protein A sepharose or protein G sepharose for 15 min and centrifuged at 1000 rpm for 3 min. Supernatants were incubated with antibody-coupled beads (see above) for 2 h at 4°C and beads centrifuged through a 30 % sucrose cushion. Supernatants (S) and beads (P) were mixed with SDS sample buffer and analyzed by gel electrophoresis and autoradiography.

**Size exclusion chromatography combined with multi-angle light scattering (SEC-MALS)** - The oligomeric state of LAP2α(410-693) in solution was determined by SEC-MALS measurements performed on an ÄKTA explorer 10 system (GE Healthcare) connected to a tri-angle light scattering detector and a differential refractometer (miniDAWN Tristar and Optilab, respectively; Wyatt Technology). A Superdex 75 HR 10/30 column (GE Healthcare) was used equilibrated in 0.1 M Na-Phosphate pH 7.5, 0.1 M NaCl, 10 mM DTT at a flow rate of 0.5 ml/min. A sample volume of 100 ml was injected at a concentration of 2 mg/ml. Data were processed using ASTRA software (Wyatt Technology) assuming a specific refractive index increment (dn/dc) of 0.185 ml/g. To determine the detector delay volumes and the normalization coefficients for the MALs detector, a BSA sample (Sigma A-8531) was used as reference. Neither despiking nor band broadening correction was applied.

**RESULTS**

The C-terminus of LAP2α is involved in the formation of oligomeric complexes in vivo – Our previous studies revealed the formation of stable, chromatin-associated LAP2α structures during early nuclear assembly stages when many other nucleoskeleton- and nuclear envelope proteins were still cytoplasmic (14,17), indicating oligomerization of the protein. To analyze whether LAP2α is indeed capable of forming oligomeric structures that contain several LAP2α molecules in vivo and to determine the domain of LAP2α responsible for oligomerization we generated constructs containing full-length LAP2α, or the LAP2α-specific C-terminal domain, LAP2α(188-693), fused to the C-terminus of monomeric Red Fluorescent Protein 1 (mRFP1) (Fig. 1A), and stably expressed these proteins in HeLa cells. Unlike GFP and DsRed, mRFP1 remains entirely monomeric, and therefore precludes misinterpretations of LAP2α oligomerization due to dimerization or tetramerization of the tag (20). Immunoblot analyses of cell extracts prepared from stable cell clones revealed the presence of fusion proteins of expected sizes, in amounts comparable to that of the endogenous LAP2α (Figure 1B, left panel). Both constructs reacted with polyclonal antibodies against LAP2α-C-terminus (anti-LAP2α and antibodies against Red Fluorescent Protein (anti-RFP), whereas a monoclonal antibody directed against the N-terminal region of LAP2α that is common to all LAP2 isoforms recognized only mRFP1-LAP2α plus the endogenous protein (anti-LAP2). Both mRFP1-LAP2α and mRFP1-LAP2α(188-693) were readily extracted in buffer containing 1% Triton X-100 and 130mM NaCl (Fig 1B, left panel). However, while a small fraction of endogenous LAP2α and of mRFP1-LAP2α remained in the
pellet fraction, mRFP1- LAP2α(188-693) was completely soluble, indicating that full-length LAP2α is more tightly bound to nuclear complexes than the LAP2α C-terminus.

Next, we examined the cellular distribution of the tagged proteins at various stages of the cell cycle, using fluorescence microscopy. As shown in Fig. 1C, both mRFP1- LAP2α(188-693) and mRFP1- LAP2α localized to the nucleoplasm in interphase, to the cytoplasm in metaphase, and to chromatin in anaphase/telophase as reported previously for the endogenous protein and GFP-LAP2α (14,16,17,21). Unlike the full-length protein, a fraction of mRFP1- LAP2α(188-693) remained in the cytoplasm during anaphase (Fig. 1C, arrow), again indicating a more soluble protein and a less efficient targeting of the C-terminus to chromatin as compared to full-length protein. Nevertheless, these data show that the C-terminus is capable of interacting with chromatin or nuclear lamina components during nuclear assembly in vivo.

To investigate whether the C-terminus can also be incorporated into higher order LAP2α structures we developed a co-immunoprecipitation approach. The mRFP1- LAP2α(188-693) fusion protein was modified by adding a Protein C-epitope and a myc-tag at the N-terminus (PC-myc-mRFP1-LAP2α(188-693), Fig. 1A), facilitating efficient isolation of the fusion protein from cell extracts. In addition, we generated a fusion protein containing Protein C-epitope and myc-tag fused directly to the C-terminus of LAP2α (PC-myc- LAP2α(188-693)). Upon stable transfection into HeLa cells, the fusion constructs were expressed at similar levels as endogenous LAP2α as revealed by immunoblot analyses of cell extracts with anti-LAP2α and anti-myc antibodies (Fig. 1B, right panel). Both proteins behaved like mRFP1- LAP2α(188-693) upon Triton extraction (Fig. 1B) and in immunofluorescence microscopy (Fig. 1C). PC-myc-mRFP1-LAP2α(188-693) and PC-myc-LAP2α(188-693) were precipitated from Triton X-100 soluble cell fractions using the anti-Protein C matrix (Fig. 2A). A protein of about 80 kDa reacting with the anti-LAP2 common domain antibody was consistently found in the immunoprecipitates from these clones, but not in non-transfected HeLa cells or clones expressing unrelated constructs (Fig. 2A and data not shown). As the region of LAP2α recognized by the monoclonal antibody is not present in the recombinant proteins, the 80-kDa protein could be unambiguously identified as the endogenous LAP2α protein. The other LAP2 isoforms were not co-immunoprecipitated (data not shown). Similarly, when the tagged LAP2α C-terminus was precipitated with anti-myc antibody, full-length LAP2α was detected in the immunoprecipitates, whereas a control antibody did not bring down any of those proteins (Fig. 2B). Likewise, unlike control antibodies, the anti-LAP2 common domain antibody co-precipitated PC-myc-mRFP1-LAP2α(188-693) and PC-myc-LAP2α(188-693) with the endogenous protein, as detected using the anti-myc antibody for immunoblotting (Fig. 2C).

Overall, our data demonstrate that the LAP2α-C-terminus associated with full-length LAP2α in vivo, and indicate a role of the C-terminal region in higher order structure organization of LAP2α complexes.

LAP2α forms stable complexes of defined molecular weight - In order to obtain further evidence for the existence of oligomeric LAP2α complexes, we chemically cross-linked protein complexes in a total HeLa cell lysate using Dithiobis(succinimidylpropionate) (DSP). DSP-mediated cross-links are stable in non-reducing conditions but can be removed by addition of reducing agents. HeLa cell extracts treated with different concentrations of DSP were analyzed by reducing and non-reducing SDS-PAGE and immunoblotting using a monoclonal antibody against the LAP2α-specific C-terminus (Fig. 3A). Depending on the concentration of the cross-linking agent, LAP2α was detected in a complex with an apparent molecular mass larger than 200 kDa. When the cross-linking agent was cleaved by addition of dithiothreitol, complexes were dissociated yielding monomeric LAP2α. Intriguingly, semi-native, non-denaturing electrophoresis of HeLa cell lysates also revealed a LAP2α-complex bigger than 200 kDa (Fig. 3A, left lane). Hence, LAP2α may exist in a stable complex of approximately 200-250 kDa in vivo.

To test whether LAP2α is able to form oligomeric complexes of similar sizes in vitro, we cross-linked highly enriched recombinant, bacterially expressed LAP2α. Immunoblot analyses revealed a LAP2α-complex of more than 200 kDa in non-reducing conditions (DTT), while mostly the monomeric protein of ~80 kDa was detected in reducing conditions (Fig. 3B). Intriguingly, the C-terminus of LAP2α(188-693)
also formed larger complexes in non-reducing conditions, while the N-terminal region (1-187) mostly remained monomeric under these conditions. Since both, the N- and C-terminus contain lysine residues, which can be targeted by the cross-linker, we concluded that LAP2α forms higher order, oligomeric complexes through self-association of its C-terminal region.

To further show the homo-oligomerization of LAP2α we generated LAP2α by in vitro translation in a reticulocyte lysate and analysed the complex by semi-native PAGE (Fig. 3C). Interestingly, a major band between 200 and 250kDa was detected, in addition to the 80-kDa monomer. Since the reticulocyte lysate does not contain any nuclear proteins, which may be specific interaction partners for LAP2α, it is very likely that the >200kDa complex represented homo-oligomeric LAP2α complexes.

The LAP2α self-interaction is mediated by the C-terminal region - Our data suggest a direct interaction of the LAP2α C-terminal region with itself mediating homo-oligomerization of LAP2α. To further show that direct self-interaction of LAP2α polypeptides can occur via the C-terminus, we performed additional in vitro binding assays. Bacterially expressed full length LAP2α(1-693), LAP2α C-terminus (188-693) and LAP2 N-terminus (1-187) were separated by SDS-PAGE, blotted onto a membrane and overlayed with radioactively labelled in vitro translated full-length LAP2α or LAP2α C-terminus. Autoradiography revealed binding of both labelled proteins to full-length LAP2α and to the C-terminus, while the N-terminus did not interact (Fig. 4A). Thus, only the C-terminus of LAP2α can mediate self-interaction of the protein. To narrow down the self-interaction domain within the LAP2α C-terminus, we expressed different fragments of LAP2α in bacteria and performed solid phase overlay assays with radioactively labeled, in vitro translated, full-length LAP2α or LAP2α C-terminus. Results presented in Figure 4B and summarized in Figure 4C show that removal of 78 residues from the C-terminus of LAP2α only moderately affected binding of the truncated protein to LAP2α and LAP2α C-terminus, removal of the C-terminal 279 residues significantly reduced binding, and deletion of the last 439 residues abolished binding completely. Concomitantly, the C-terminal 284 residues of LAP2α (410-693) showed a binding, although interaction was slightly weaker than that of the entire C-terminus (188-693). Vimentin as a negative control did not interact with LAP2α.

In order to establish the level of assembly of the LAP2α C-terminus and to confirm its interaction in solution, we performed size exclusion chromatography combined with multi-angle static light scattering (SEC-MALS) and refractive index measurements, which yields an accurate determination of the molecular mass independent of molecular shape or hydrodynamic parameters. Size exclusion chromatograms revealed the presence of one single oligomeric species in solution, indicating a high monodispersity of LAP2α(410-693) populations. Further analysis by SEC-MALS yielded a molecular mass of 66.3 kDa, which is in excellent agreement with the calculated theoretical mass of 65.0 kDa for a dimer of this construct (Fig 5).

In summary, these in vitro data demonstrate that the LAP2α C-terminus can specifically and directly bind to itself, primarily forming homodimers. The self-interaction domain is located between residues 410 to 693, but domains upstream of amino acid 410 may also contribute to stable interactions. In contrast, the interaction of LAP2α’s C-terminus with lamin C is restricted to the last C-terminal 78 amino acids in LAP2α (111), and Fig. 4B).

The self association of LAP2α is not affected by a disease linked mutation - A mutation in the extreme C-terminus of LAP2α has recently been linked to dilated cardiomyopathy in humans (18). The molecular mechanism of this disease is currently unknown, but we have shown that the disease-linked mutation in LAP2α decreased its binding affinity for lamins A and C. Therefore, we wondered whether the mutation could also interfere with the self-association of LAP2α molecules and whether this could contribute to the cellular defect in patient cells. To test this hypothesis, GFP-LAP2α and untagged wild type or disease-linked LAP2α variants were in vitro translated simultaneously using a reticulocyte extract, and LAP2α complexes were immunoprecipitated with anti-GFP antibody. Immunoblot analyses using LAP2α antibodies detected untagged LAP2α in the immunoprecipitates only when GFP-LAP2α was present in the samples, whereas untagged protein alone was not precipitated by anti GFP antibodies (Fig. 6). These data clearly support
the direct self-interaction of LAP2α molecules. Both, wild type and mutated LAP2α co-precipitated with GFP-LAP2α with similar efficiencies, indicating that the disease-linked mutation in LAP2α does not interfere with its self-association.

**DISCUSSION**

The studies presented here revealed a so far unappreciated biochemical property of LAP2α, the formation of homo-oligomeric structures. Co-immunoprecipitation of endogenous LAP2α with exogenously expressed full-length LAP2α or different fusion proteins containing the α-specific C-terminal domain demonstrated that higher order LAP2α structures occur in living cells, are stable and require the protein’s unique C-terminus. *In vitro* overlay studies and size exclusion chromatography of purified proteins combined with multi-angle light scattering showed that LAP2α can form homo-dimers by direct interaction of its C-terminus.

**LAP2α C-terminus is involved in several interactions** - *In vitro* overlay experiments and light scattering of purified protein complexes in solution clearly showed that LAP2α can form homo-oligomers through direct interaction of the C-terminus. While the last 284 amino acids were sufficient for dimerization, other more upstream regions may also contribute to the stabilization of this interaction. The C-terminus has been shown to contain several additional binding domains: The last 78 amino acids of LAP2α are involved in the interaction with lamin C (11), while further upstream regions mediate the interaction with pRb (12). Although we cannot rule out that the binding regions of LAP2α for pRb and for self-association overlap and influence each other, we argue that self-assembly of LAP2α may occur at the same time as interactions with laminA/C and pRb. By this LAP2α structures can act as a platform supporting higher order chromatin structure and transcriptional regulation.

Intriguingly, mutated LAP2α (R690C) expressed in the human disease DCM showed reduced binding to lamin A tail *in vitro* (18), while LAP2α self-association was not impaired by the R690C substitution. This observation is consistent with the results of overlay blots indicating a relatively broad interface of self-interaction between LAP2α monomers potentially involving several regions within its unique C-terminus.

**Size of stable homo-oligomeric LAP2α structures** - Analysis of HeLa cell extracts after chemical cross-linking or in semi-native non-denaturing gels showed the presence of stable LAP2α-containing complexes of molecular masses between 200-250 kDa. Complexes of the same size were detected after cross-linking of bacterially expressed LAP2α and by semi-native PAGE of LAP2α expressed in reticuloocyte lysate, indicating that LAP2α is the only protein in the 200-250 kDa complex. According to their size, these complexes might accommodate three LAP2α molecules of 75-80 kDa. However, light scattering analyses of highly purified bacterially expressed LAP2α C-terminal fragment revealed a high monodispersity of the sample strongly arguing for a stable dimeric organization. Therefore we consider it very unlikely that LAP2α forms trimers and favour the idea that the > 200kDa complex of LAP2α either represents a homodimer, which shows unusual mobility in gel electrophoresis due to anisometric features, or a compact homotetramer with slightly increased mobility.

**Functional implication of the LAP2α self interaction** - By fluorescent microscopy, LAP2α C-terminal domain stably expressed in HeLa cells showed a cellular distribution identical to endogenous LAP2α, except for anaphase-telophase, where it showed a slightly less efficient association with chromosomes. This observation supports previous findings, showing that the C-terminus is required and sufficient for chromosome association of LAP2α during nuclear assembly (17). The less efficient chromatin binding of LAP2α C-terminus to chromatin also indicates that, although the C-terminus is sufficient for targeting LAP2α to chromosomes, its stable association requires also the N-terminal common domain. However, based on the data presented here, we cannot rule out completely that the C-terminal fragment associated with chromosomes mainly through its interaction with full length endogenous LAP2α during nuclear assembly *in vivo*.

In any case, the engagement of LAP2α in a self-interaction has important implications for its reported functions. The formation of LAP2α
homo- or hetero-oligomeric complexes has several consequences for its chromatin binding properties. First, by bringing various LEM- and LEM-like domains together in a LAP2α oligomer and thus potentially generating multiple binding sites for BAF and DNA in the oligomeric complex, its affinity for chromatin could be significantly increased. Secondly, the N-termini of the oligomeric complex could interact simultaneously with several DNA fibers and thus perform efficient cross-linking of chromatin regions. In line with this model it was shown that full length LAP2α and BAF are essential components of the preintegration complex (PIC) of retroviruses (22). LAP2α’s proposed role in stabilizing the interaction of BAF with PICs required both the N-terminal common domain and the -specific C-terminal domain of LAP2α. Our results, showing that the C-terminus is the determinant for oligomerization of LAP2α provide a possible explanation for this observation.

In conclusion, our data suggest a model where a homo-dimeric core complex of LAP2α, formed via self-association of the C-terminal domain, serves as the building block for higher-order homo and hetero-oligomeric structures of LAP2α.

REFERENCES
FOOTNOTES
ACKNOWLEDGMENTS
This study was supported by grants from the Austrian Science Research Fund (FWF P17871) to RF and from the Austrian National Bank (grant 9840) to CS. MM was supported by a grant from the Swiss National Foundation (3100A0-100852) to OM.

ABBREVIATIONS
BAF, Barrier-to-Autointegration Factor; DCM, dilated cardiomyopathy; LAP, Lamina-associated polypeptide; LEM, Lamina-associated polypeptide2-Emerin-MAN1; PAGE, polyacrylamide gel electrophoresis; pRb, retinoblastoma;
FIGURE LEGENDS

FIGURE 1. Stable expression of LAP2α and the α-specific C-terminus in HeLa cells. A, schematic presentation of fusion proteins of full-length LAP2α and LAP2α(188-693) with mRFP1, Protein C epitope and myc-tag, and for comparison, the domain organizations of LAP2α and LAP2α are shown. The light gray region represents the LAP2α-specific C-terminus; gray and black boxes show positions of LEM-, LEM-like-, and transmembrane domains, respectively. Arrow indicates the epitope recognized by the monoclonal anti-LAP2 common domain antibody. B, untransfected HeLa cells (Hela) or HeLa cells stably expressing the indicated constructs described in A were lysed in 1% Triton X-100 and total cell lysates (T), or pellet (P) and supernatant (S) fractions following centrifugation were analyzed by immunoblotting using polyclonal anti- LAP2α (anti- LAP2α), monoclonal anti-LAP2 common domain (anti-LAP2) and polyclonal anti-DsRed (anti-RFP) or polyclonal anti-myc antibodies, as indicated. Numbers show molecular masses in kDa. C, cells expressing mRFP1- LAP2α or mRFP1- LAP2α(188-693), or PC-myc- LAP2α(188-693) as indicated, were analyzed at various cell cycle stages by fluorescence microscopy. Left images show mRFP1 or anti-myc staining, right images DAPI-stained DNA. Arrow indicates anaphase cells with remaining cytoplasmic staining. Bar, 10 μm.

FIGURE 2. Co-immunoprecipitation of endogenous LAP2α with expressed LAP2α C-terminus from HeLa cell extracts. A, soluble cell extracts of non-transfected HeLa cells, two HeLa clones expressing PC-myc-mRFP1- LAP2α(188-693), and one clone expressing PC-myc- LAP2α(188-693) were incubated with anti-Protein C matrix and immunoprecipitates were analyzed by immunoblotting using anti-LAP2 common domain antibody. B, cell extracts were incubated with monoclonal anti-myc (+) or a control monoclonal antibody (-) and immunoprecipitates probed with anti-LAP2 antibody. C, extracts were incubated with monoclonal anti-LAP2 (+) or with control (-) antibody and immunoprecipitates probed with anti-myc antibody.

FIGURE 3. Analysis of LAP2α complexes by chemical crosslinking and seminative gel electrophoresis. A, HeLa cell lysates were analyzed on semi-native polyacrylamide gels (Endogenous complex, left panel) or incubated with the indicated amounts of DSP for 2h, dissolved in sample buffer containing (+) or lacking (-) DTT and analyzed by SDS-polyacrylamide gel electrophoresis and immunoblotting using monoclonal anti- LAP2α, antibody 15/2. B, bacterially expressed full-length LAP2α and LAP2α fragments, as indicated, were incubated with 1mM DSP, mixed with sample buffer with (+) or without (-) DTT and analyzed by immunoblotting using monoclonal anti- LAP2α antibody 15/2 (left and central panels) or anti-LAP2 antibody 12 (right panel). C, in vitro translated, [35S]methionine-labeled LAP2α was analyzed by seminative PAGE and autoradiography. Numbers denote molecular masses in kDa, arrows indicate position of monomeric proteins and arrowheads position of complexes.

FIGURE 4. LAP2α self-associates via its α-specific C-terminus in vitro. A, in vitro translated [35S]methionine-labeled LAP2α or the α-specific domain LAP2α(188-693) were overlayed onto immobilized bacterially expressed full-length LAP2α, LAP2α(188-693) and LAP2(1-187). An autoradiogram of an SDS-polyacrylamide gel of in vitro translated proteins ( Autoradiogram), a Coomassie blue-stained gel of bacterially expressed proteins (Coomassie), and autoradiograms of overlays (Overlay) are shown. B, transblotted vimentin and recombinant LAP2α fragments, as indicated, were overlayed with in vitro translated [35S]methionine-labeled recombinant full-length LAP2α or the α-specific domain LAP2α(188-693). Ponceau S-stain of blotted LAP2α fragments and autoradiograms of overlays are shown. Numbers show molecular masses in kDa. C, schematic diagrams of immobilized LAP2α and LAP2α fragments used in B are shown. Numbers denote amino acid position. The right panel indicates binding of the in vitro translated proteins to immobilized proteins, based on signal intensities on autoradiograms of the overlays (++++, strongest interaction; +, weakest interaction; -, no interaction). The calculated molecular masses (Mr) of LAP2α and LAP2α fragments are shown in kDa.
FIGURE 5: Analysis of purified LAP2α C-terminal fragment 410-693 by size exclusion chromatography combined with multi-angle light scattering (SEC-MALS).

The refractive index of the corresponding peak region of a size exclusion chromatogram (Superdex 75 HR 13/30) performed in 0.1 M Na-Phosphate, pH 7.5, 0.1 M NaCl, 10 mM DTT is shown as solid line. The weight-averaged molecular mass measured inside the peak area at volume intervals is displayed as filled squares. The theoretical molecular mass is indicated by the continuous horizontal line.

FIGURE 6: A disease (DCM)-causing mutation in LAP2α does not affect its self-interaction.

Wild type LAP2α or DCM mutant LAP2α were in vitro translated and labeled with [35S]methionine in a reticulocyte extract either alone or with GFP-LAP2α, and complexes were immunoprecipitated using anti-GFP antibody. Input samples and supernatants (S) and pellet (P) fractions were analyzed by SDS PAGE and autoradiography. Numbers indicate molecular masses in kDa.
Figure 1, Snyers et al.
Figure 2, Snyers et al.
Figure 3, Snyers et al.
Figure 4, Snyers et al.
Figure 5, Snyers et al.
Figure 6, Snyers et al.
References
7 References


McElhinny, A.S., Kakinuma, K., Sorimachi, H., Labeit, S. and Gregorio, C.C. (2002) Muscle-specific RING finger-1 interacts with titin to regulate sarcomeric M-line and thick filament structure and may have nuclear functions via its interaction...


Taking geometry to its edge: fast unbound rigid (and hinge-bent) docking.

*Proteins, 52*, 107-112.


Curriculum Vitae

Personal details
Surname: Mrosek
Forename: Michael
Date of birth: 12.09.74
Nationality: German
Marital status: unmarried
Languages: German (native), English (fluent), Latin

Scientific qualifications
Ph.D. thesis in protein X-ray crystallography
Department of structural biology and biophysics, Biozentrum, University of Basel, Switzerland, Group of Assist. Prof. Dr. Olga Mayans.
Title: “Structural studies on protein scaffolds related to muscle physiology and disease: the titin filament, its associated component MuRF-1 and nuclear LAP2α” (in English).
Grade: summa cum laude.

10/1995 - 04/2002
Study of chemistry and diploma thesis
Department of organic chemistry and biochemistry, University of Freiburg, Germany, Group of Prof. Dr. Georg Schulz.
Title: “Mutagenesis, preparation and X-ray crystallographic investigations on surface-mutants of the ring-opening Maltodextrinase from Flavobacterium sp. Nr. 92” (in German).
Grade: 1.5 (very good-good).

Foreign exchange
09/1997- 06/1998
Erasmus/Sokrates exchange student
Department of physical chemistry, University of Liverpool, United Kingdom, Group of Dr. Richard Nichols.

Practical skills
Molecular Biology

Biophysical methods
CD/NMR-spectroscopy, static/dynamic light scattering, ultracentrifugation, electron microscopy, calorimetry, binding assays.

Structural biology
Protein crystallization, organize/perform synchrotron data collection (SLS/ESRF), experimental MAD/MIR phasing.

Computational
Active working knowledge: XDS, MOSFLM, CCP4, CNS, SOLVE, SHARP, SHELX, COOT, O, other: MacOS X, Unix, MS-Office.

Teaching
Tuition of undergraduate students in X-ray crystallography and organic/inorganic chemistry, practical and theoretical sessions.
Publications


Conferences, workshops

10/2006 Talk at the 20th REGIO Meeting on biological crystallography, Konstanz, Germany. Title: “Evidence for the recruitment of a disease-related muscle specific factor onto titin”.

05/2006 8th International School on the Crystallography of Biological Macromolecules, Como, Italy, poster presentation.

10/2005 Practical course at the Paul Scherer Institute (PSI) at the Swiss Light Source (SLS), Villingen, Switzerland. Topic: “Synchrotron Data Acquisition Techniques in Macromolecular Crystallography”. Techniques: Beamline automatization, micro-crystals, SAD/MAD.

10/2003 Practical course at the University of Zürich, Switzerland. Topic: “Characterization and Crystallization of Soluble and Membrane proteins”. Techniques: CD-, IR-spectroscopy, DLS, HPLC.


08/2003 Talk at the Amersham Biosciences seminar: “From Gene to Functional Protein”, Biozentrum, Switzerland. Title: “Preparation and characterization of LAPα, a protein from the nuclear matrix”.


Further personal data

08/1994 - 10/1995 Civil service


Personal interest Mountaineering (member of the german alpine club since 1994)

References Assis. Prof. Dr. Olga Mayans (Olga.Mayans@unibas.ch)
Prof. Dr. Ulrich Baumann (Ulrich.Baumann@ibc.unibe.ch)

Basel,