Structure of trigger factor in complex with the ribosome defines the molecular environment of the emerging nascent polypeptide chain

Author(s):
Ferbitz, Lars

Publication Date:
2005

Permanent Link:
https://doi.org/10.3929/ethz-a-005064101

Rights / License:
In Copyright - Non-Commercial Use Permitted
Structure of trigger factor in complex with the ribosome defines the molecular environment of the emerging nascent polypeptide chain

A dissertation submitted to the Swiss Federal Institute of Technology Zürich

for the degree of Doctor of Sciences

presented by

Lars Ferbitz
Dipl.–Biochem. Univ.

born October 14, 1974

from Neunkirchen (Germany)

Prof. Dr. Nenad Ban, Examiner
Prof. Dr. Rudolf Glockshuber, Coexaminer
Dr. Elke Deuerling, Coexaminer

2005
For my parents.
i. Summary

Every protein in a cell is synthesised by an organelle-sized machine, the ribosome, which translates information contained in mRNA into polypeptides. After synthesis each protein has to adopt its unique three-dimensional structure. This folding process is an error prone process and requires a sophisticated network of helper proteins, so called molecular chaperones, to ensure productive *de novo* folding and to prevent non-native proteins from misfolding, aggregation or premature degradation. Some of these chaperones interact with nascent chains already during their synthesis on the ribosome (co-translationally), while others come into play after the release of the protein from the ribosome (post-translationally). The first chaperone that co-translationally interacts with all proteins in bacteria is the trigger factor. The trigger factor is the only chaperone in bacteria that directly associates with the ribosome. A significant amount of biochemical data has been obtained since its first discovery almost 20 years ago. Nevertheless, the mechanism of action that allows the trigger factor to assist the folding of proteins remained unknown. We concluded that structural studies of the trigger factor would provide an important contribution to our understanding of the co-translational folding events.

In this thesis crystallisation attempts with different trigger factor variants from *Thermus thermophilus* or *Escherichia coli* are described. The first diffraction data of the trigger factor crystals were obtained with the *Escherichia coli* protein either using a deletion mutant lacking the protease sensitive 47 N-terminal amino acids or the full-length protein. The optimisation of the crystallisation conditions lead to the determination of the full-length protein structure at 2.7 Å resolution. The trigger factor adopts a unique conformation resembling a crouching dragon with separated domains forming the N-terminal ribosome-binding “tail”, the peptidyl-prolyl isomerase (PPIase) “head”, the C-terminal “arms”, and connecting regions building up the “back”. The finding that the trigger factor specifically associates with the large ribosomal subunit from all three kingdoms of life allowed an experiment in which a heterologous complex between the *Escherichia coli* trigger factor and the large ribosomal subunit from *Haloarcula marismortui* was formed to obtain additional structural data from the trigger factor acting at the ribosome. The structure of the ribosome-binding domain of the bacterial chaperone in complex with the large ribosomal subunit from *Haloarcula marismortui* was solved at 3.5 Å resolution. From its attachment point on the ribosome, the trigger factor projects its extended domains over the exit of the ribosomal tunnel creating a protected folding space where nascent polypeptides are shielded from proteases and aggregation. This
study sheds new light on the understanding of co-translational protein folding and suggests an unexpected mechanism of action for ribosome-associated chaperones.
ii. Zusammenfassung


grosen ribosomalen Untereinheit aus Haloarcula marismortui wurde bei 3.5 Å Auflösung bestimmt. Von der Bindungsregion am Ribosom erstreckt der Trigger Faktor seine ausgedehnten Domänen über den Ausgang des ribosomalen Tunnels, so dass ein geschützter Faltungsraum entsteht, in dem die naszierende Kette vor Proteasen und Aggregation geschützt ist. Diese Studie wirft ein neues Licht auf das Verständnis ko-translationaler Proteinfaltung und schlägt einen unerwarteten Mechanismus der Wirkungsweise von Ribosomen bindenden Chaperonen vor.
### iii. Table of Contents

1. Summary ................................................................................................................................. 3
2. Zusammenfassung ...................................................................................................................... 5
3. Table of Contents .................................................................................................................... 7
4. List of Figures .......................................................................................................................... 11
5. List of Tables .......................................................................................................................... 13
6. Abbreviations ........................................................................................................................ 14

#### 1 Introduction .................................................................................................................... 16
1.1 Protein synthesis by ribosomes ............................................................................................ 16
1.2 Protein folding and the role of molecular chaperones .......................................................... 17
1.3 Protein translocation ............................................................................................................. 20
1.4 Periplasmic chaperones ......................................................................................................... 21
1.5 ATP-dependent chaperones .................................................................................................. 22
1.6 Ribosome-associated chaperones ......................................................................................... 24
1.7 Trigger factor ......................................................................................................................... 26
1.8 Goals for the thesis project ................................................................................................. 30

#### 2 Materials and Methods ................................................................................................... 31
2.1 DNA Methods ....................................................................................................................... 31
  2.1.1 Agarose gel electrophoresis ............................................................................................ 31
  2.1.2 Design of PCR primers .................................................................................................. 31
  2.1.3 Polymerase chain reaction ............................................................................................. 31
  2.1.4 Restriction enzyme digest of DNA used for subcloning .................................................. 32
  2.1.5 DNA ligation .................................................................................................................. 32
  2.1.6 Transformation of chemically competent E. coli cells ....................................................... 32
  2.1.7 Transformation of electrocompetent E. coli cells ............................................................. 32
  2.1.8 Plasmid DNA purification ............................................................................................... 33
  2.1.9 DNA sequencing ............................................................................................................ 33
    2.1.9.1 ABI sequencer ......................................................................................................... 33
    2.1.9.2 Sequencing by Synergene Biotech ......................................................................... 33
  2.2 Protein expression and purification ....................................................................................... 33
    2.2.1 Test expression in LB medium ..................................................................................... 33
    2.2.2 Large scale expression in LB medium ........................................................................... 34
    2.2.3 Expression of trigger factor variants ............................................................................. 34
    2.2.4 Expression of selenomethionine derivatised trigger factor ........................................... 35
    2.2.5 Cross-linking experiments ............................................................................................. 35
  2.3 Growth of Haloarcula marismortui cells .............................................................................. 36
  2.4 Protein purification methods ............................................................................................... 36
2.4.1 Cell lysis ..........................................................................................................................36
2.4.2 TEV cleavage .........................................................................................................................36
2.4.3 Metal affinity purification .....................................................................................................37
  2.4.3.1 Column purification .........................................................................................................37
  2.4.3.2 Batch purification ............................................................................................................37
2.4.4 Anion exchange chromatography .........................................................................................37
2.4.5 Size exclusion chromatography ..........................................................................................38
2.5 Protein analysis .......................................................................................................................38
  2.5.1 Denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) ......38
  2.5.2 Coomassie staining .........................................................................................................38
  2.5.3 Silver staining ..................................................................................................................39
  2.5.4 Determination of concentration ......................................................................................39
  2.5.5 Electro-blotting and N-terminal sequencing .....................................................................39
  2.5.6 Mass spectroscopy .........................................................................................................40
2.6 Ribosome purification ............................................................................................................40
2.7 Crystallisation methods ...........................................................................................................41
  2.7.1 Vapour diffusion ..............................................................................................................41
    2.7.1.1 Trigger factor ..............................................................................................................41
    2.7.1.2 Large ribosomal subunit (H. marismortui) ....................................................................41
  2.7.2 Cryo-protection of crystals ...............................................................................................41
  2.7.3 Soaking crystals with heavy atom compounds or ligands ..................................................42
2.8 X-ray data collection ..............................................................................................................42
  2.8.1 Crystal mounting under cryo-conditions .........................................................................42
  2.8.2 Data collection ..................................................................................................................43
    2.8.2.1 Trigger factor ..............................................................................................................43
      2.8.2.1.1 Multiwavelength anomalous dispersion (MAD) experiment .................................43
    2.8.2.2 Large ribosomal subunit (H. marismortui) ....................................................................43
2.9 Computational methods .........................................................................................................44
  2.9.1 Sequence manipulation and analysis ..............................................................................44
  2.9.2 Computer software for structure determination ...............................................................44
3 Results .........................................................................................................................................46
  3.1 Crystallisation of the trigger factor, crystal handling and data collection ............................46
    3.1.1 Initial screening conditions ...........................................................................................46
    3.1.2 Thermus thermophilus trigger factor ..............................................................................47
    3.1.3 E. coli trigger factor constructs ......................................................................................48
      3.1.3.1 Δ47 trigger factor construct .......................................................................................48
      3.1.3.2 Full-length trigger factor .........................................................................................50
  3.2 Optimising and characterising E.coli trigger factor crystals ..................................................52
    3.2.1 Refinement of crystallisation conditions at 4º C ..............................................................52
    3.2.2 Crystal composition .......................................................................................................53
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.3</td>
<td>Derivatisation attempts of native crystals</td>
<td>54</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Crystal characterisation</td>
<td>55</td>
</tr>
<tr>
<td>3.3</td>
<td>Structure determination</td>
<td>56</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Selenomethionine derivatisation of the protein</td>
<td>56</td>
</tr>
<tr>
<td>3.3.1.1</td>
<td>Expression and purification</td>
<td>57</td>
</tr>
<tr>
<td>3.3.1.2</td>
<td>Crystallisation</td>
<td>58</td>
</tr>
<tr>
<td>3.3.1.3</td>
<td>Data collection</td>
<td>59</td>
</tr>
<tr>
<td>3.3.2</td>
<td>MAD phasing</td>
<td>60</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Model building and refinement</td>
<td>64</td>
</tr>
<tr>
<td>3.3.4</td>
<td>The trigger factor fold</td>
<td>66</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Crystal packing</td>
<td>69</td>
</tr>
<tr>
<td>3.3.6</td>
<td>Structural homologues</td>
<td>71</td>
</tr>
<tr>
<td>3.4</td>
<td>Cloning, expression and purification of trigger factor constructs</td>
<td>73</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Expression and purification of trigger factor fragments</td>
<td>73</td>
</tr>
<tr>
<td>3.5</td>
<td>Co-crystallisation and soaking of H. marismortui 50S with trigger factor fragments</td>
<td>74</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Binding of trigger factor to H. marismortui 50S</td>
<td>75</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Crystallisation of the trigger factor:50S complex</td>
<td>76</td>
</tr>
<tr>
<td>3.5.3</td>
<td>Data collection</td>
<td>77</td>
</tr>
<tr>
<td>3.5.4</td>
<td>Data processing and scaling</td>
<td>78</td>
</tr>
<tr>
<td>3.5.5</td>
<td>Model refinement</td>
<td>78</td>
</tr>
<tr>
<td>3.5.6</td>
<td>Structure determination of the N-terminal fragment of trigger factor in complex with H. marismortui 50S</td>
<td>79</td>
</tr>
<tr>
<td>3.5.7</td>
<td>Loop mediated interaction of trigger factor on the ribosome</td>
<td>80</td>
</tr>
<tr>
<td>3.5.8</td>
<td>Sequence conservation of the trigger factor docking site</td>
<td>81</td>
</tr>
<tr>
<td>3.5.9</td>
<td>Full-length trigger factor bound to the ribosome</td>
<td>82</td>
</tr>
<tr>
<td>4</td>
<td>Discussion</td>
<td>84</td>
</tr>
<tr>
<td>4.1</td>
<td>The trigger factor fold</td>
<td>84</td>
</tr>
<tr>
<td>4.2</td>
<td>Trigger factor acting at the ribosome</td>
<td>86</td>
</tr>
<tr>
<td>4.3</td>
<td>Trigger factor assisted co-translational folding</td>
<td>89</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Co-translational protein folding in prokaryotes</td>
<td>89</td>
</tr>
<tr>
<td>4.3.2</td>
<td>A model for the mechanism of trigger factor function on the ribosome</td>
<td>90</td>
</tr>
<tr>
<td>4.4</td>
<td>Interplay with other factors involved in folding or targeting of newly synthesised proteins</td>
<td>91</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Interplay of trigger factor and DnaK</td>
<td>91</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Trigger factor and the secretory pathway</td>
<td>92</td>
</tr>
<tr>
<td>4.4.2.1</td>
<td>L23 constitutes a general docking site on the ribosome</td>
<td>92</td>
</tr>
<tr>
<td>4.4.2.2</td>
<td>SRP-dependent export</td>
<td>93</td>
</tr>
<tr>
<td>4.4.2.3</td>
<td>Sec-dependent export</td>
<td>94</td>
</tr>
<tr>
<td>5</td>
<td>Future prospects</td>
<td>95</td>
</tr>
<tr>
<td>6</td>
<td>References</td>
<td>96</td>
</tr>
<tr>
<td>7</td>
<td>Appendix</td>
<td>110</td>
</tr>
</tbody>
</table>
7.1 General chemicals

7.1.1 Buffers and solutions

7.1.2 Media for bacterial growth

7.1.3 Enzymes

7.1.4 Gel electrophoresis molecular weight markers

7.1.5 General equipment

7.1.6 Centrifuges

7.1.7 FPLC equipment

7.1.8 FPLC and resins

7.2 Crystallisation equipment

7.2.1 General equipment

7.2.2 X-ray sources and detection

7.2.3 Crystallisation screens

7.2.3.1 Commercially available crystallisation screens

7.2.3.2 Grid screens: Ammonium sulphate, PEG, MPD vs pH

7.2.3.3 Additive screens

7.3 Plasmid maps

7.3.1 pH6-tig

7.3.2 pH6-1-111-TEV

7.4 Trigger factor amino acid sequence

7.5 Multiwavelength anomalous dispersion (MAD) experiments

8 Acknowledgements

9 Curriculum Vitae
iv. List of Figures

Figure 1: The protein synthesis machinery in bacteria .............................................................. 17
Figure 2: Protein folding pathways in bacteria .......................................................................... 19
Figure 3: Chaperones involved in post-translational folding ..................................................... 23
Figure 4: Domain architecture of the trigger factor and the TF signature motif ....................... 29
Figure 5: Trigger factor variants used for crystallisation .......................................................... 47
Figure 6: *Thermus thermophilus* trigger factor crystals .......................................................... 47
Figure 7: Δ47 crystallisation ...................................................................................................... 48
Figure 8: Refined condition for Δ47 crystals .............................................................................. 49
Figure 9: Δ47 trigger factor crystals diffracted to a maximum resolution of 8 Å ...................... 50
Figure 10: Initial screening of the *E.coli* trigger factor crystals ............................................. 51
Figure 11: Refined conditions for trigger factor crystals .......................................................... 53
Figure 12: Biochemical analysis of trigger factor crystals ........................................................ 53
Figure 13: Diffraction pattern of full-length *E. coli* trigger factor crystals ............................ 55
Figure 14: Self-rotation function showing section $x = 180^\circ$ .............................................. 56
Figure 15: Superdex200 purification of trigger factor .............................................................. 57
Figure 16: Mass spectroscopic analysis of selenomethionine derivatised trigger factor ....... 58
Figure 17: Crystals of selenomethionine derivatised trigger factor ....................................... 59
Figure 18: Plot of the real $(f')$ and the imaginary $(f'')$ components of the anomalous scattering around the K edge of selenium as a function of the X-ray energy ........................................ 59
Figure 19: Experimental and predicted Patterson Maps of Harker section $y = 0$ ................. 62
Figure 20: Peaks identified in anomalous difference Fourier maps ....................................... 63
Figure 21: Electron density maps used for building the structure of the trigger factor ............ 65
Figure 22: Ramachandran Plot of the two trigger factor molecules in the asymmetric unit ..... 66
Figure 23: Trigger factor folds into a unique dragon shape ..................................................... 67
Figure 24: The domain organisation of the trigger factor fold ............................................... 68
Figure 25: Trigger factor forms a large hydrophobic cradle .................................................... 69
Figure 26: Crystal contact sites of the full-length trigger factor ............................................ 70
Figure 27: Superimposition of the two full-length TF molecules in the crystallographic asymmetric unit ........................................................................................................... 71
Figure 28: The trigger factor fold shows a modular architecture ............................................ 72
Figure 29: Purification of trigger factor fragments ................................................................... 74
Figure 30: Trigger factor binding to *H. marismortui* 50S.......................... 76
Figure 31: Diffraction pattern of 50S co-crystallised with trigger factor fragments .......... 77
Figure 32: Fo-Fo difference Fourier map calculated for TF144 bound to *H. marismortui* 50S 79
Figure 33: Structure of the N-terminal trigger factor domain in complex with 50S........... 81
Figure 34: Sequence conservation of L23 and ribosomal RNA at the TF binding site .......... 82
Figure 35: Model of full-length trigger factor bound to the ribosome............................. 83
Figure 36: The cradle fold, trigger factor and SurA..................................................... 85
Figure 37: The trigger factor cradle could accommodate entire protein domains .............. 89
Figure 38: Schematic representation of the model for the mechanism of trigger factor assisted co-translational protein folding......................................................... 91
Figure 39: Full-length trigger factor bound to the ribosome side-by-side shown with EM map representations of ribosome-bound SRP and Sec61 ................................................. 93
v. List of Tables

Table 1: Names, sequences and annealing temperatures of used primers ......................... 31
Table 2: Strains and plasmids used for protein expression .................................................. 34
Table 3: Pipetting scheme for denaturing sodium dodecyl sulfate-polyacrylamide gels ........ 38
Table 4: All program packages and input scripts used during structure determination and   
representation are listed ................................................................................................. 44
Table 5: Fine screening of the initial crystallisation condition ................................................. 52
Table 6: Calculation of the Matthews coefficient and the corresponding number of trigger   
factor molecules in the asymmetric unit (ASU) ............................................................... 55
Table 7: Data collection statistics of the selenomethionine derivative ..................................... 60
Table 8: Heavy atom search performed with CNS ................................................................. 61
Table 9: Refinement statistics of the trigger factor structure .................................................. 64
Table 10: Data collection statistics of 50S co-crystallised with TF144 or of 50S crystals
soaked with TF111 ........................................................................................................... 78
Table 11: Refinement statistics of H. marismortui 50S crystals ............................................... 78
Table 12: Refinement statistics of the structure of the N-terminal trigger factor domain in   
complex with 50S ............................................................................................................. 80
vi. Abbreviations

Δ47  TF fragment with the first 47 amino acids deleted
°C  Degrees Celsius
1 Da  1 g / mol
50S  50 Svedberg units
Å  Angstroem (10 nm)
ADP  Adenosine diphosphate
AIDS  Autoimmune deficiency syndrome
APS  Ammonium persulfate
ASU  Asymmetric unit
ATP  Adenosine triphosphate
ATPase  ATP hydrolase
BSA  Bovine serum albumine
BPMI  Benzophenone-4-maleimide
CV  Column volume
ddH₂O  Bidistilled water
ddNTP  Dideoxy nucleotide triphosphate
DNA  Deoxy ribonucleic acid
DNase  DNA hydrolase
dNTP  Deoxy nucleotide triphosphate
dsDNA  Double stranded DNA
DTT  Dithiothreitol
E. coli  *Escherichia coli*
EG  Ethyleneglycol
EM  Electron microscopy
ER  Endoplasmatic reticulum
Ffh  Fifty-four homologue
FKBP  FK506 binding proteins
FOM  Figure-of-merit
g  g-force
GTPase  GTP hydrolase
h  Hour
H. marismortui  *Haloarcula marismortui*
HPLC  High pressure liquid chromatography
Hsp  Heat shock protein
IgG  Immunoglobulin G
k  10³
K  Kelvin
l  Liters
L23  Large ribosomal protein 23
LB  Luria-Bertani
MAD  Multiwavelength anomalous dispersion
MALDI-TOF  Matrix-assisted laser desorption and ionisation time of flight
MES  2-N-morpholino-ethanesulphonic acid
mg  Milligramm
ml  Millilitter
MPD  2-methyl-2,4-pentanediol
MR  Molecular replacement
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ms</td>
<td>Millisecond</td>
</tr>
<tr>
<td>NAC</td>
<td>Nascent-chain associated complex</td>
</tr>
<tr>
<td>NCS</td>
<td>Non-crystallographic symmetry</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometer (10^{-9} m)</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>OmpA</td>
<td>Outer membrane protein A</td>
</tr>
<tr>
<td>Pa</td>
<td>Pascal</td>
</tr>
<tr>
<td>PC</td>
<td>Patterson correlation</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein database</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethyleneglycol</td>
</tr>
<tr>
<td>pmol</td>
<td>Picomol (10^{-12} mol)</td>
</tr>
<tr>
<td>PPIase</td>
<td>Peptidyl-prolyl-cis/trans-isomerase</td>
</tr>
<tr>
<td>PT</td>
<td>Peptidyl transferase</td>
</tr>
<tr>
<td>RB</td>
<td>Ribosome-binding domain</td>
</tr>
<tr>
<td>RF</td>
<td>Rotation function</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>RNA hydrolase</td>
</tr>
<tr>
<td>RNC</td>
<td>Ribosome nascent chain</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SIR/ MIR</td>
<td>Single/multiple isomorphous replacement</td>
</tr>
<tr>
<td>SR</td>
<td>SRP receptor</td>
</tr>
<tr>
<td>SRP</td>
<td>Signal recognition particle</td>
</tr>
<tr>
<td>Tat</td>
<td>Twin-arginine translocation</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris-(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TF</td>
<td>Trigger factor</td>
</tr>
<tr>
<td>TF144</td>
<td>TF fragment of the first 144 N-terminal amino acids</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting temperatures</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td>U</td>
<td>Unit</td>
</tr>
<tr>
<td>UBA</td>
<td>Ubiquitin associated</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
<tr>
<td>v/v</td>
<td>Volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight per volume</td>
</tr>
<tr>
<td>µm</td>
<td>Micrometer (10^{-6} m)</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Protein synthesis by ribosomes

In all kingdoms of life proteins are synthesised by a huge cellular machine (~20nm), the ribosome, which uses amino acyl tRNAs as substrates to polymerase amino acids into proteins based on messenger RNA templates. In bacteria, the ribosome consists of a large (50S) and a small (30S) subunit, which associate into the active 70S ribosome during translation. Each subunit is made of ribosomal RNA (rRNA) and proteins. The small subunit contains one rRNA (16S rRNA) and 21 proteins. The 50S subunit consists of a large (23S rRNA) and a small (5S rRNA) rRNA and 35 proteins. The high-resolution structures of both the small and large subunits of bacterial and archaeal ribosomes have provided first structural insights into the protein synthesis machinery (Ban et al., 2000; Pioletti et al., 2001; Wimberly et al., 2000). The decoding centre of the 30S subunit mediates the molecular recognition between mRNA codons and the corresponding tRNA anticodons. The large subunit catalyses the formation of peptide bonds between amino acids delivered to the active site as amino acylated tRNAs (Figure 1). The growing nascent peptide is then directly synthesised into a 100 Å long exit tunnel in the large subunit which can accommodate about 35 amino acids in an extended conformation (Hardesty and Kramer, 2001). Its diameter of ~15 Å prohibits formation of tertiary protein interactions and folding of entire protein domains. Although a helix-like state has recently been observed early in the tunnel for integral inner membrane proteins (Woolhead et al., 2004) and further compaction inside the tunnel has been suggested based on the cryo-EM imaging of stalled ribosomes (Gilbert et al., 2004), the growing nascent chain leaves the ribosomal tunnel in a mostly unfolded state (Nissen et al., 2000).
Chapter 1

Introduction

Figure 1: The protein synthesis machinery in bacteria
The bacterial 70S ribosome is composed of a small (30S) and large (50S) ribosomal subunit. Proteins are coloured yellow (50S) and orange (30S), ribosomal RNA is shown in light grey (23S and 5S rRNA) and dark grey (16S rRNA). The 30S subunit contains the decoding centre, the 50S subunit the peptidyl transferase centre. Three tRNA molecules mark their corresponding amino acyl (A, red), peptidyl (P, blue) and exit (E, green) sites on the ribosome. The mRNA is coloured violet.

In the active site of the bacterial ribosome amino acids are attached to the C-terminus of the growing polypeptide chain with a speed of 5-20 residues per second (Sorensen and Pedersen, 1991). Thus a protein consisting out of 100 amino acids is synthesised within 5 s in a bacterial cell under optimal conditions. In sum it is estimated that about 20,000 ribosomes synthesise approximately 30,000 proteins per minute in \textit{E. coli} (Bukau et al., 2000) and each of these newly synthesised proteins must fold into their correct three dimensional structure.

1.2 Protein folding and the role of molecular chaperones

Classic \textit{in vitro} experiments demonstrated that the entire information necessary for the folding process is encoded in the amino acid sequence of a protein (Anfinsen, 1973; Baker, 2000) and that small stable proteins can spontaneously fold within milliseconds (Dobson and Karplus, 1999). However, \textit{in vivo} the situation is more complex and protein folding is a highly error prone process for the following reasons: (i) The protein is synthesised in a vectorial
manner where the N-terminal portion is leaving the ribosomal tunnel whereas the C-terminal segment is not yet synthesised. Thus during translation the sequence available for folding is always incomplete and transient and premature folding has to be prevented until at least an entire protein domain sequence is available outside the ribosomal tunnel. In addition, considering the relatively slow rate of the translation process, the emerging nascent chain will remain in an unfolded, non-native state for a considerable amount of time. (ii) Furthermore, the cytosol is an unfavourable environment for the folding process since high concentrations of proteins and other macromolecules (300 mg ml$^{-1}$ in *Escherichia coli*) can favor unproductive interactions of the newly made proteins exposing hydrophobic segments in their unfolded state. Such incorrect inter- or intramolecular interactions promote aggregation or premature degradation of the unfolded protein (Ellis, 2001). Therefore, cells have to develop a system of helper proteins, the so-called molecular chaperones, that assist the folding of its newly synthesised proteins.

Many chaperones are called heat shock proteins (Hsp) as they were initially identified in cells exposed to high temperatures and were named according to their molecular weights. Later on, it became evident that these proteins are also present under normal growth conditions to maintain the cell’s protein repertoire in a native, functional state. Some of them rescue cellular proteins from denaturation, while others guide newly made proteins to their predestined fate (Figure 2). Importantly, chaperones interact transiently with substrates and do not remain associated with their final product (Becker and Craig, 1994; Bukau and Horwich, 1998; Hartl, 1996; Johnson and Craig, 1997). Molecular chaperones are found in all species and kingdoms of life, but so far they are best studied in the bacterial model system *Escherichia coli*.
Figure 2: Protein folding pathways in bacteria
Nascent chains first interact with the ribosome-associated chaperone trigger factor. Depending on their future functions proteins either fold inside the cytosol or are exported through or inserted into the inner membrane. In the cytosol, proteins might fold without further assistance or their folding is facilitated by additional chaperones, mainly DnaK with its co-chaperone DnaJ and the GroEL/ES chaperonin. In the periplasm, a different group of chaperones welcomes the protein. ClpB breaks apart protein aggregates, holdases Hsp31, Hsp33 or LbpA/B reduce the workload on cytosolic chaperones during stress conditions.
1.3 **Protein translocation**

At least 10% of all bacterial proteins do not remain in the cytosol for folding, but are translocated through or inserted into the membrane in an unfolded state. The main secretion pathway in all kingdoms of life involves a heterotrimeric protein complex, a translocon, that allows unfolded proteins to traverse the membrane (de Keyzer et al., 2003; Rapoport et al., 1996; Van den Berg et al., 2004). Polypeptides carrying hydrophobic cleavable signal sequences or transmembrane segments are co-translationally targeted to the translocon via a signal recognition particle (SRP) (Figure 2).

In *E. coli*, the majority of proteins are secreted post-translationally via the Sec-dependent pathway. The homotetrameric SecB chaperone (Xu et al., 2000) associates with ribosome-bound nascent chains with more than ~150 residues in length (Hardy and Randall, 1991) to recognise stretches of about nine amino acids enriched in aromatic and basic residues (SecB-binding motif) (Knoblauch et al., 1999; Xu et al., 2000). SecB maintains the preprotein in a translocation-competent state and transfers its cargo to SecA. The cytoplasmatic ATPase SecA cycles between a SecYEG bound and free form. In the membrane bound form it acts as the motor component of the translocase while undergoing repeated ATP-driven conformational changes (Bu et al., 2003; Economou and Wickner, 1994; Eichler and Wickner, 1997). Although being dimeric in solution SecA is discussed to function as a monomer during translocation (Or et al., 2005). The translocation process is assisted by a couple of membrane proteins (SecG, SecD, SecF, YajC) and the proton gradient across the membrane. In addition to its role in protein export, SecB is considered to serve as a general chaperone as it interacts both with secretory and cytosolic polypeptides *in vitro* and its overexpression can compensate for defects observed in strains lacking both the trigger factor and DnaK (Ullers et al., 2004).

The SRP-mediated targeting of secretory and membrane proteins to the plasma membrane in prokaryotes or the endoplasmatic reticulum (ER) in eukaryotes has been extensively studied (Luirink and Sinning, 2004). SRP binds to a hydrophobic N-terminal signal sequence as it emerges from the ribosome. The SRP-RNC (ribosome nascent chain) complex associates with the SRP receptor (SR) that facilitates the delivery of the RNC to the translocon. The whole process is controlled by the GTPase activities of the SRP and the SR.

In eukaryotes SRP consists of six proteins bound to a 7S RNA forming the two functional domains of the protein-RNA complex: The S-domain with its binding site for the signal
peptide is involved in ribosome binding and SR interaction (Wild et al., 2004a; Wild et al., 2004b) and the Alu domain is responsible for translation arrest which resumes upon delivery of the ribosome to the Sec61 translocon (Halic et al., 2004; Keenan et al., 2001; Walter and Blobel, 1981).

In Archaea the SRP consists of a 7S RNA bound by protein homologues of SRP54 and SRP19 (Eichler and Moll, 2001). In Eubacteria we can find the minimal particle required for SRP function: A 4.5S RNA is bound by a single protein, the fifty-four homologue (Ffh) (Bernstein et al., 1989; Romisch et al., 1989). In all kingdoms SRP54 is a three domain protein with a methionine rich M domain, involved in RNA and signalpeptide binding, a G domain with GTPase activity and an N-terminal N domain, involved in ribosome and SR binding (Lu et al., 2001).

The SRP receptor, either heterodimeric (Tajima et al., 1986) (SRα / SRβ in eukaryotes) or monomeric (Luirink et al., 1994) (FtsY or its homologue in Eubacteria or Archaea, respectively), mediates the GTP-dependent release of SRP from the signal peptide (Connolly and Gilmore, 1989) to allow the translocation of the nascent chain. Upon release of the RNC SRP and SR dissociate from another, so that they are ready for the next translocation cycle (Connolly and Gilmore, 1989; Connolly et al., 1991).

Secretory proteins in many prokaryotes, containing a twin-arginine consensus motif (S-R-R-x-F-L-K), can also be exported in a fully folded conformation via the twin-arginine translocation (Tat) pathway. Tat is composed of three transmembrane proteins TatA, TatB, and TatC. TatC recognises the protein substrate (Alami et al., 2003) while TatA is discussed to form the tranlocation pore (Gohlke et al., 2005). TatB might be involved in substrate delivery to the translocase (Muller, 2005). The entire process of the Tat pathway is still poorly understood.

1.4 Periplasmic chaperones

Interestingly, the export of proteins through the translocation pore into the periplasm in their unfolded state resembles the exit of cytosolic proteins emerging from the ribosomal tunnel. Consequently, within the periplasm an arsenal of chaperones specialised for the oxidative, ATP depleted conditions awaits the unfolded protein (Baneyx and Mujacic, 2004). The periplasm harbours, on the one hand, a set of generic chaperones like Skp or FkpA with broad substrate specificity. On the other hand some chaperones like SurA (outer membrane proteins), the PapD-like superfamily (P pili biosynthesis) or FimC (type 1 pili biosynthesis) are specialised for specific substrates. FkpA combines PPIase and chaperone activity (Arie et
al., 2001) in contrast to SurA which contains two parvulin-like PPIase domains, but its chaperone activity does not depend on proline isomerisation (Behrens et al., 2001). The Dsb family of proteins, a set of thiol-disulfide oxidoreductases, ensures correct disulfide bond formation (Kadokura et al., 2003).

1.5 ATP-dependent chaperones

In the cytosol of prokaryotic and eukaryotic cells as well as in organelles of eukaryotic cells two major classes of ATP-dependent chaperones are present, the Hsp70s and the cylindrical chaperonin complexes (Hsp60s) (Figure 2). The Hsp70s consist of two functionally coupled domains: The 44 kDa N-terminal domain mediates ATP binding (Flaherty et al., 1990), the 18 kDa C-terminal domain binds to short hydrophobic polypeptide stretches (Rudiger et al., 1997; Zhu et al., 1996) (Figure 3). In the ATP-bound state, Hsp70 binds and releases substrates rapidly, whereas in the ADP-bound form the exchange is significantly slower (Bukau and Horwich, 1998; Hartl, 1996). In E.coli the cycling between different nucleotide bound states of the Hsp70 homologue DnaK is mediated by two cofactors, the Hsp40 homologue DnaJ and GrpE. The 41 kDa protein DnaJ itself acts as a chaperone by binding to unfolded polypeptides (Langer et al., 1992; Szabo et al., 1996). In addition DnaJ binds DnaK to stimulate its ATPase activity. DnaK in its ADP-bound form stably interacts with the polypeptide substrate that is bound in an extended conformation in the peptide-binding cleft of its N-terminal domain (Figure 3). The 23 kDa protein GrpE acts as a nucleotide exchange factor by distorting the nucleotide binding pocket of DnaK causing the release of bound ADP. The chaperoning cycle of DnaK ends with rebinding of ATP and release of the polypeptide substrate, which now can continue folding or is delivered to other chaperones in the cytosol such as the chaperonin complex.
Figure 3: Chaperones involved in post-translational folding

A. The crystal structure of the substrate-binding domain of DnaK is shown in complex with a 7 amino acid long polypeptide substrate (upper part). The close-up view of the substrate-binding region of DnaK shows the polypeptide bound in an extended conformation in the peptide-binding cleft. Selected side-chains of DnaK in the substrate-binding region are depicted as stick representations. B. The crystal structure of the asymmetric GroEL-GroES complex is shown in the upper part (GroEL: orange, GroES: yellow). GroEL’s folding cavity shown underneath allows the binding of proteins smaller than 60 – 65 kDa. During refolding of nonnative proteins the folding space is closed by GroES which functions as a lid.

Chaperonins are large cylindrical protein complexes of two stacked rings (Saibil, 2000). Each ring is composed of seven to nine subunits. Group I chaperonins (GroEL in *E.coli*, Hsp60 in mitochondria and chloroplasts) act in conjunction with a ring-shaped co-chaperonin (GroES or Hsp10) that functions as a lid by closing the folding space the polypeptide is bound in (Figure 3) (Bukau and Horwich, 1998; Hartl, 1996; Johnson and Craig, 1997). This cage allows the binding of proteins smaller than 60 – 65 kDa (Xu et al., 1997). In the case of the group II chaperonins, which are only found in Archaea and in Eukarya, no cofactor is identified yet. Again the cycle of substrate binding and release is ATP driven. The equatorial domains of GroEL harbour the ATP-binding pocket; the apical domains contain a patch of
hydrophobic residues facing the interior of the cavity. GroEL does not bind linear peptides, but interacts with proteins that still have hydrophobic residues exposed in their nonnative state. In the ATP bound form the apical domain of GroEL interacts with GroES, which induces a conformational change that causes substrate release into the chaperonin cavity now lined by hydrophilic residues. The substrate remains in the cavity for up to 15 s until it reaches the folded, native state with now too few exposed hydrophobic residues to interact with GroEL.

Only 15 % of all newborn proteins interact with DnaK/J (Deuerling et al., 1999; Teter et al., 1999) and 10 % with GroEL/ES (Ewalt et al., 1997; Houry et al., 1999). On the contrary, all proteins are thought to interact with a group of small chaperones that bind to the ribosome to form a “welcoming committee” for nascent chains and presumably assist protein folding during ongoing synthesis. Unlike DnaK and GroEL they act in an ATP independent manner and are grouped in a special class of Hsp70 related, ribosome-associated, chaperones.

1.6 Ribosome-associated chaperones

The bacterial ribosome-associated chaperone trigger factor was the first chaperone of this class described (Hesterkamp et al., 1996). Originally Wickner and co-workers identified the trigger factor in an attempt to search for new factors involved in protein translocation. A single ~60 kDa protein in the cytoplasmic fraction of E. coli lysate (100,000g supernatant S100) was shown to trigger the transport of a precursor of the outer membrane protein A (pro-OmpA) by supporting the translocation competent protein conformation of pro-OmpA (Crooke and Wickner, 1987). Consequently, the discovered factor was named the trigger factor (TF). Finding that the trigger factor stimulates pro-OmpA export in vitro was initially interpreted to indicate a secretion-specific chaperone function for the trigger factor during the shuttling of secretory precursor proteins. However, it was shown thereafter that E. coli cells depleted for the trigger factor did not show any secretion defects. Nevertheless, an important role for the trigger factor was predicted as the tig gene is highly conserved among bacteria and even included in the genome of Mycoplasma genitalium which is thought to contain the minimum set of genes required for life (Fraser et al., 1995). The role of the trigger factor in vivo was enigmatic for many years, until it was discovered that the ribosome-associated trigger factor plays an important role in the folding of newly synthesised proteins (Deuerling et al., 1999; Teter et al., 1999).

The trigger factor is exclusively found in bacteria and chloroplasts, but not in the cytosol of eukaryotic cells. However, evidence exists that ribosome-associated chaperones represent a
highly conserved principle in eukaryotes and prokaryotes, even though the involved components differ between species. They all share the ability to bind in a 1:1 stoichiometry to ribosomes and to interact with emerging nascent polypeptides during the translation process. In the unicellular eukaryote *Saccharomyces cerevisiae* the Hsp70 homologue Ssz1p was identified to form a complex with Zuotin, an Hsp40 homologue, which is able to bind ribosomes. Together, they form the so-called ribosome-associated complex or RAC. Later on, with Ssb1/2p, another Hsp70 homologue was identified that also binds Zuotin and is able to associate with ribosomes. Its expression is co-regulated with ribosomal proteins (Lopez et al., 1999). The association of Ssb with nontranslating ribosomes is salt-sensitive, but becomes salt-resistant once it interacts with the emerging nascent chain (Pfund et al., 1998). Therefore, in yeast, the heterodimeric RAC and Ssb1/2 seem to act as a chaperone triad on emerging nascent chains (Gautschi et al., 2002) with Zuotin and Ssb1/2 forming the primary chaperone (Huang et al., 2005). Recently, a similar ribosome-associated chaperone complex was also identified in higher eukaryotes. In humans the Zuotin ortholog Mpp11 associates with ribosomes and forms a stable complex with Hsp70L1, a distantly related Ssz1p homologue. The heterodimeric complex is able to rescue growth defects in yeast cells lacking RAC (Otto et al., 2005). Mpp11 has also been shown to recruit the multifunctional Hsp70 chaperone Hsc70, a homologue of the yeast Hsp70 protein Ssa, to nascent polypeptide chains (Hundley et al., 2005).

A second ribosome-associated chaperone complex has been identified both in yeast and in higher eukaryotes, the nascent chain-associated complex (NAC) (Lauring et al., 1995a; Lauring et al., 1995b; Rospert et al., 2002; Wiedmann and Prehn, 1999). The heterodimeric NAC is formed by proteins αNAC and βNAC (Btf3). Both subunits bind the nascent chain, but only βNAC directly interacts with the ribosome (Beatrix et al., 2000). In yeast two βNAC homologues exist that are encoded by the *EGD1* and *BTT1* genes. The deletion of all three genes (EGD1:α-subunit, EGD2:β1-subunit, BTT1:β3-subunit) in the unicellular eukaryote does not cause a drastic phenotype (Reimann et al., 1999). However, the crucial role of the ribosome-associated chaperone for cell viability is emphasised by the embryonically lethal phenotype of NAC mutants in *Drosophila melanogaster* and *Caenorhabditis elegans* (Deng and Behringer, 1995; Markesich et al., 2000). Furthermore, the NAC levels have been observed to change significantly in the context of several human diseases like AIDS, Alzheimer or brain tumours (Bloss et al., 2003; Kim et al., 2002; Kroes et al., 2000; Scheuring et al., 1998). The heterodimeric NAC binds the ribosome in close proximity to the site where the emerging nascent-chain exits the ribosomal tunnel and this way the NAC might
prevent premature interactions of newborn proteins with cytosolic factors and might shield the proposed membrane-attachment site on the ribosome to regulate the translocation into the ER and into mitochondria (George et al., 1998; Lauring et al., 1995a; Rospert et al., 2002; Wang et al., 1995). The described chaperone-like function of the NAC is accompanied by the activity of its individual subunits which are rather involved in transcriptional than translational control (Franke et al., 2001; Moreau et al., 1998; Yotov et al., 1998) suggesting a dual function of the NAC protein (Rospert et al., 2002).

In archaebacteria a protein with sequence homology to the eukaryotic NAC was discovered (Makarova et al., 1999). The archaeal “NAC” is a single protein that contains a domain with homology to the conserved NAC domain also present in eukaryotic αNAC and βNAC and an UBA (ubiquitin associated) domain only present in αNAC. The conserved NAC domain seems to be crucial for dimerisation, while the role of the UBA domain is unknown (Spreter et al., 2005). The archaeal NAC homologue is able to associate with ribosomes and interacts with the nascent chain. This interaction can be competed with the human NAC (Spreter et al., 2005).

1.7 Trigger factor

The trigger factor is an abundant cellular protein present in 2-3 times molar excess over ribosomes. However, trigger factor depleted cells have no defects in growth or protein folding (Deuerling et al., 1999; Teter et al., 1999). Similarly, the deletion of the dnaK gene, encoding the major Hsp70 gene in E. coli, does not affect protein folding or cell viability at permissive growth temperatures (Bukau and Walker, 1989; Deuerling et al., 1999; Hesterkamp and Bukau, 1998; Teter et al., 1999). However, the simultaneous deletion of the DnaK and trigger factor genes results in synthetic lethality at 30°C and above (Deuerling et al., 1999; Teter et al., 1999). It was concluded that these two chaperones cooperate in the folding of newly synthesised proteins. In co-immunoprecipitation experiments the direct association of DnaK with nascent polypeptides was shown: In wild type E. coli cells about 5-18% of newly synthesised proteins interact with DnaK. This level increases 2 to 3-fold (26-36%) in trigger factor depleted cells indicating a more intensive use of the DnaK chaperone (Deuerling et al., 1999; Teter et al., 1999). Moreover, the deletion of the trigger factor gene induces a heat shock response in E. coli causing an increase in the steady state levels of chaperones including DnaK (Deuerling et al., 2003). Therefore, the trigger factor and DnaK chaperones seem to cooperate in the folding of nascent polypeptides with DnaK providing a back-up system for trigger factor depleted cells.
These findings were very surprising as the trigger factor and DnaK use completely different mechanisms to assist protein folding. DnaK interacts with the regulatory proteins DnaJ and GrpE and the substrate binding and release is regulated by nucleotide binding and hydrolysis (as described before). On the contrary, the trigger factor interaction with unfolded substrates is highly dynamic (Maier et al., 2001) and there is no evidence for cofactors or a regulation by nucleotide binding. However, the screening of potential binding sites of the trigger factor and DnaK in natural protein substrates revealed that DnaK and the trigger factor share a high overlap (77%) in binding peptides (Deuerling et al., 2003) consistent with the overlapping features of the substrate binding motifs of both chaperones (Patzelt et al., 2001).

The trigger factor binds to peptides enriched in aromatic and basic amino acids, whereas peptides with acidic amino acids are disfavored (Patzelt et al., 2001). Within the trigger factor binding motif no preference for the positioning of basic and aromatic residues became evident. Interestingly, prolyl residues do not contribute to the binding of the trigger factor both to peptides (Patzelt et al., 2001) and to the protein substrate RNase T1 (Scholz et al., 1998). These results were unexpected since the trigger factor is a PPIase and was shown to catalyse the cis/trans isomerisation of peptidyl-prolyl peptide bonds in peptides and protein substrates (Hesterkamp et al., 1997; Scholz et al., 1998).

The DnaK binding motif, 5 hydrophobic residues flanked by positively charged residues, is related but not identical to the trigger factor binding motif. The trigger factor prefers aromatic residues while DnaK prefers aliphatic residues. Basic residues in the trigger factor binding motif are randomly distributed while in the DnaK binding motif they are flanking the hydrophobic core. In both cases negative charges are excluded (Patzelt et al., 2001; Rudiger et al., 1997; Zhu et al., 1996). The overlap in binding specificities might be a prerequisite that both chaperones, despite their different mechanisms of action, can promote folding of the same substrate pool by protecting similar hydrophobic stretches in unfolded protein species. In wild-type cells this interaction is likely to follow a hierarchical order imposed by the positioning of the trigger factor adjacent to the polypeptide exit tunnel on the large ribosomal subunit (see below) (Kramer et al., 2002b).

Proteolytic fragments of the E. coli trigger factor (432 aa) revealed that this chaperone is a three-domain protein (Figure 4A). A 102 amino acids long fragment (Glu148 to Thr249), homologous to PPIases of the FKBP family, displays PPIase activity (Callebaut and Mornon, 1995; Hesterkamp et al., 1996). The cis/trans isomerisation of peptidyl-prolyl bonds is often the rate limiting step in protein folding (Kiefhaber et al., 1990) as shown in several in vitro folding studies (Gothel and Marahiel, 1999). Three classes of PPIases can be distinguished:
The cyclophilins, the FK506-binding proteins (FKBPs) and the parvulins. FKBPs and parvulins adopt a similar fold of their PPIase domains, which has been called the FKBP fold (Sekerina et al., 2000). Cyclophilins and FKBPs, collectively called immunophilins, are targets for the immunosuppressive drugs cyclosporin A and FK506, respectively. In vivo, this class of PPIases seems to rather act specifically on unique targets than to play a general role in protein folding since its enzymatic activity was shown to be dispensable for viability in yeast (Dolinski et al., 1997). The trigger factor belongs to a subfamily of the FKBP-type PPIases as its PPIase activity cannot be inhibited by the immunosuppressant FK506 (Stoller et al., 1995). Surprisingly, the function of the trigger factor to assist the folding of newly synthesised proteins seems to be independent of its prolyl isomerisation activity and the role of its PPIase domain in vivo is still enigmatic (Huang et al., 2000; Kramer et al., 2004b). The second proteolytic fragment corresponding to the N-terminal 14 kDa of the protein was identified as a ribosome-binding (RB) domain with comparable affinity for the ribosome as observed for the full-length trigger factor (Hesterkamp et al., 1997; Stoller et al., 1996). The C-terminal region could not be studied in isolation as it is highly susceptible to proteolysis (Hesterkamp and Bukau, 1996; Stoller et al., 1996). Additionally, this domain has no sequence homology with any other protein and, therefore, its function remained unknown.

In the cytosol the trigger factor is in monomer-dimer equilibrium with two thirds of it being dimeric. Only the monomeric form binds the ribosome with ~1 \( \mu \text{M} \) affinity (Patzelt et al., 2002). The trigger factor is present in 2 - 3 fold molar excess over ribosomes, so that 90 % of the ribosomes can be estimated to be in complex with the chaperone. The affinity for peptides in solution is low, so that efficient binding can only occur in the context of the ribosome (Patzelt et al., 2001).

The trigger factor could be efficiently cross-linked to nascent-chains derived from both secretory and cytosolic proteins (Valent et al., 1995). The interaction with translating ribosomes is salt-resistant (Hesterkamp et al., 1996), but a salt-sensitive association with the isolated large ribosomal subunit was observed (Lill et al., 1988). This lead to the speculation that the trigger factor might bind in the close proximity of the nascent chain exit site of the large ribosomal subunit to accomplish its duty as a general chaperone required in the early stages of protein synthesis.

With the aim to characterise the binding of the trigger factor to the large ribosomal subunit cross-linking and mutational studies were performed. In a cross-linking experiment using a trigger factor variant coupled with a UV-activatable crosslinker two adjacent ribosomal proteins L23 and L29, both located in proximity to the exit tunnel, were identified as potential
interaction partners of the chaperone (Kramer et al., 2002b). Using mutant cell lines with a deletion in the L29 gene it was shown that only ribosomal protein L23 is essential for the interaction with the trigger factor. Bacterial glutamate 18 in L23 positioned adjacent to the exit site of the emerging nascent chain was identified as the critical interaction partner.

Sequence comparison of bacterial trigger factor homologues revealed a very well conserved sequence of about 17 residues, the TF signature motif, within the N-terminal region of the trigger factor (Kramer et al., 2002b) (Figure 4B). A mutation of residues Phe44, Arg45 and Lys46 to alanine, within that region, significantly reduced trigger factors affinity for the ribosome.

In summary, ribosomal protein L23 and a flexible loop in the ribosome binding domain of the trigger factor were discovered as the contact points between the trigger factor and the ribosome, which facilitates its important interactions with nascent polypeptide chains and its function as a chaperone in vivo (Kramer et al., 2002b).

Figure 4: Domain architecture of the trigger factor and the TF signature motif
A. Linear domain architecture of the trigger factor. Red: Ribosome-binding domain with the conserved TF signature motif. Yellow: PPlase domain. Blue: C-terminal domain of unknown function. B. Sequence-logo representation of aligned bacterial sequences within the region comprising the TF signature motif. The TF signature motif (GFRxGxxP) is highly conserved among bacteria.
1.8 Goals for the thesis project

The first folding helper that interacts with all nascent proteins in bacteria is the trigger factor. A significant amount of biochemical data has been obtained since its first discovery almost 20 years ago: The trigger factor is a three domain protein, it preferentially binds hydrophobic and basic residues of nascent chains in the context of the ribosome with ribosomal protein L23 being its docking site.

Nevertheless, prior to this work, it was still unknown how the trigger factor accomplishes its function in vivo where it protects newly synthesised polypeptide chains from misfolding, aggregation and proteolysis. In particular the function of its C-terminal domain remained enigmatic due to its extreme sensitivity to proteases and absence of sequence homology to other proteins. With an aim to reveal the mechanism of trigger factor action we have decided to conduct structural studies on the trigger factor protein and on the trigger factor in complex with the large ribosomal subunit.

The first goal was to obtain the crystal structure of the full-length trigger factor protein either from *Thermus thermophilus* or *Escherichia coli*. Alternatively, deletion variants of *E. coli* trigger factor such as the isolated trigger factor domains or a deletion mutant lacking the protease sensitive 47 N-terminal amino acids should be tested in crystallisation trials. The optimisation of crystallisation conditions should lead to a structure determination of the full-length *E. coli* trigger factor or any variant tested. The second goal was to get structural information for the TF-ribosome complex. To this end co-crystallisation attempts of trigger factor variants with the large ribosomal subunit (50S) from *Halobacula marismortui* should be carried out. Alternatively, soaking experiments of the trigger factor into already grown 50S crystals should be performed.

Ultimatively, the structural information on the TF-ribosome complex in combination with the structure of the full-length uncomplexed trigger factor should provide the first insights into the trigger factor structure bound to the ribosome and its mode of action as a chaperone for nascent polypeptides.
2 Materials and Methods

2.1 DNA Methods

2.1.1 Agarose gel electrophoresis

DNA was analysed and purified on 0.8 – 1.2 % (w/v) agarose gels. Solid agarose was boiled in 50 ml of 1 x TBE. 2 µl of a 1 µg ml\(^{-1}\) ethidiumbromide solution were added to the melted agarose before casting the gel.

2.1.2 Design of PCR primers

The PCR primers were between 33 and 39 bases long with melting temperatures predicted by a web-based oligonucleotide property calculator that uses nearest neighbour and thermodynamic calculations described by Breslauer et al. with optimised thermodynamic parameters (http://www.basic.northwestern.edu/biotoools/oligocalc.html) (Breslauer et al., 1986; Sugimoto et al., 1996). Primer names, sequences and melting temperatures (\(T_m\)) are listed in Table 1.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Description</th>
<th>(T_m^1) [ºC]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF111fw</td>
<td>5’ TTC CAG GGA TCC CAA GTT TCA GTT GAA ACC ACT 3’</td>
<td>PCR forward primer for TF111 subcloning into the ORF of pRET3a</td>
<td>45.0</td>
</tr>
<tr>
<td>TF111bw</td>
<td>5’ TCT CGA GGT ACC TTA AAC TTC AAA CTC TAC AGA GTA AGT 3’</td>
<td>PCR reverse primer for TF111 subcloning into the ORF of pRET3a</td>
<td>46.0</td>
</tr>
</tbody>
</table>

\(^1\): The initial annealing temperature for the first rounds of PCR

Table 1: Names, sequences and annealing temperatures of used primers

The melting temperatures were predicted by a web-based oligonucleotide property calculator (http://www.basic.northwestern.edu/biotoools/oligocalc.html).

2.1.3 Polymerase chain reaction

DNA for subcloning was amplified by the polymerase chain reaction (PCR). 50 ng of genomic or plasmid DNA were incubated on ice in a 500 µl Eppendorf tube with two PCR primers (0.5 µM each) and 0.2 mM dNTP, 1 x Pwo reaction buffer, 1.5 mM MgSO\(_4\) and 12.5 U Pwo DNA polymerase. The reaction mix was heated to 90º C for 20 s in a thermocycler equipped with a heated lid. In 30 to 35 cycles, the reaction mix was brought to 90º C for 30 s, incubated for 1 min at a temperature 5º C below the predicted annealing temperature of the primers, followed by incubation at 72º C for 1 min per 1000 base pairs length of the expected product. At the end, the sample was incubated at 72º C for 5 min and put on hold at 4º C. The PCR products were analysed and purified by DNA gel electrophoresis.
2.1.4 Restriction enzyme digest of DNA used for subcloning

In a total reaction volume of 50 µl, 0.1 µg to 10 µg of DNA were mixed with 5 µl of appropriate 10x reaction buffer (NEB). After addition of 2 to 5 U restriction enzyme (NEB) per µg of DNA the reaction was incubated at 37º C for 3 hours. Plasmid DNA was treated with 1 U calf intestine phosphatase for 30 min at 37º C. DNA was analysed and purified by agarose gel electrophoresis. The product was extracted from the gel using a gel extraction kit (Qiagen) and procedures recommended by the manufacturer. The DNA was eluted in 30 µl of ddH₂O and stored at -20º C.

2.1.5 DNA ligation

In a total volume of 10 µl, 1 - 2 µl of digested DNA plasmid were incubated with 5 µl of appropriately digested insert DNA, 1 x T4 DNA ligase buffer (NEB), and 400 units of T4 DNA ligase. In order to identify the fraction of religated plasmids, a test reaction lacking insert DNA was prepared. The ligation reactions were incubated over night at 19º C or at room temperature (RT).

2.1.6 Transformation of chemically competent *E. coli* cells

This procedure was mainly used for retransformation of plasmid DNA: 1 µl of diluted plasmid DNA (10 – 20 ng) was added to a suspension of 2 µl of chemically competent *E. coli* cells in 50 µl of 0.1 M CaCl₂. The cell suspension was kept on ice for 10 min followed by incubation at 42º C for 2 min. After addition of 500 µl of LB medium, the heat-shocked cells were shaken with 180 revolutions per minute (rpm) in a 14 ml culture tube (Greiner) at 37º C for 45 min. 50 µl to 100 µl were spread on an LB plate containing the appropriate antibiotic and the plate was incubated over night at 37º C.

2.1.7 Transformation of electrocompetent *E. coli* cells

This procedure was mainly used for cloning: 1 – 2 µl of the ligation mix was added to 50 µl of electrocompetent *E. coli* DH5α cells. The suspension was transferred into a precooled electro-cuvette and placed into the MicroPulser (Biorad) module with its capacitor charged to 1.8 kV. The current is then discharged into the sample with a typical time constant τ of 4.5 to 5 ms (the time over which the voltage declines to e⁻¹ (~37 %) of the peak value). After immediate addition of 1 ml LB medium cells were regenerated for 1 h at 37º C by shaking
with 180 rpm. 100 µl were plated on LB agar containing the appropriate antibiotic to be incubated over night at 37º C.

2.1.8 Plasmid DNA purification

Plasmid DNA was purified from *E. coli* DH5α cells by an alkaline lysis protocol and subsequent purification steps involving an anion exchange chromatography resin. This procedure was carried out using DNA Mini- or Midi-Prep kits (Qiagen), according to the manufacturers recommendation. DNA was stored in ddH₂O at -20º C.

2.1.9 DNA sequencing

2.1.9.1 ABI sequencer

Double stranded DNA (dsDNA) was sequenced by the dideoxy chain termination method (Sanger et al., 1977) using dideoxy nucleotide triphosphate terminators (ddNTP) carrying base specific fluorescent labels with one colour for each base. In a PCR machine, 0.2 µg of dsDNA in 20 µl of 80 mM Tris-HCl pH 8.8, 2 mM MgCl₂, and 2 pmol of sequencing primer were amplified using a thermostable DNA polymerase, dNTP, labelled ddNTP as supplied by the manufacturer. Within 25 cycles the DNA was denatured at 96º C for 10 s, the primer annealed at 50º C for 5 s and the DNA polymerised for 4 min at 60º C. After PCR the DNA was ethanol precipitated by addition of 2 µl of 3 M NaOAc pH 5.2 and 50 µl of absolute ethanol and centrifugation at 13,000 rpm for 30 min at room temperature. The supernatant was aspirated and the pellet was washed with 300 µl of 80 % ethanol, centrifuged at 13,000 rpm for 5 min and air-dried for 5 min. The air-dried pellet was resuspended in 30 µl of ddH₂O and loaded on an ABI PRISM automated sequencing machine.

2.1.9.2 Sequencing by Synergene Biotech

Alternatively, 15 µl of a 150 ng µl⁻¹ dsDNA solution containing 1 µl of a 10 pmol µl⁻¹ primer solution were sent to Synergene Biotech (www.synergene-biotech.com) for sequencing.

2.2 Protein expression and purification

2.2.1 Test expression in LB medium

Cells from a single colony were transferred into a 14 ml Greiner tube containing 3 ml of LB medium with the required antibiotic and incubated over night at 37º C under constant
shaking at 150 rpm. In a 500 ml Erlenmeyer flask 100 ml of LB medium with antibiotic were inoculated with 1 ml of the preculture. The inoculum was shaken at 150 rpm at 30º C or 37º C. The optical density (OD) of the inoculum was measured regularly at a wavelength of 600 nm. When the cells reached a density of 0.6 to 1.0 absorption units, expression was induced by adding IPTG to a final concentration of 1 mM. 1 ml aliquots were taken before and every hour after induction with IPTG and centrifuged at 13,000 rpm for 1 min. After removal of the supernatant the remaining pellet was resuspended in 1 ml PGLB and boiled for 5 min at 95º C. The samples were analysed with SDS-PAGE.

### 2.2.2 Large scale expression in LB medium

In a 500 ml Erlenmeyer flask, 200 ml of LB with antibiotic were inoculated with cells from a single colony. The preculture was shaken over night with 150 rpm at 37º C before six times 15 ml were transferred into six 5 l Erlenmeyer flasks, each containing 1.5 l LB medium and antibiotic. Cells had been grown to an OD$_{600}$ of 0.6 to 1.0 before protein-overexpression was induced by adding IPTG to a final concentration of 1.0 mM. 3 hours after induction, cells were harvested by centrifugation at 4000 rpm for 15 min at 4º C. The pellets were resuspended in 80 ml to 100 ml NiNTA lysis buffer and flash frozen in liquid nitrogen to be stored at -20º C.

### Table 2: Strains and plasmids used for protein expression

<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
<th>Plasmid</th>
<th>Growth conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (pRep4)</td>
<td>$F, \text{ dcm, ompT, hsdS (rø mB), gal}$</td>
<td>pH6-1-144 pH6-Δ47 pH6-tig pH6-1-111-TEV</td>
<td>LB medium, 100 µg ml$^{-1}$ Ampicillin, 50 µg ml$^{-1}$ Kanamycin, 37º C, 150 rpm</td>
</tr>
<tr>
<td>DH5α (pZA4)</td>
<td>$F, \Phi 80\Delta lacZ\Delta M15, \Delta (lacZYA-argF)$</td>
<td>pH6-tig pH6-tig-V49C</td>
<td>LB medium, 100 µg ml$^{-1}$ Ampicillin, 50 µg ml$^{-1}$ Spectinomycin, 37º C, 150 rpm</td>
</tr>
<tr>
<td>MC4100 (Δtig::Kan) (pZA4)</td>
<td>$F araD139 \Delta(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR$</td>
<td>pH6-tig</td>
<td>M9 minimal medium 100 µg ml$^{-1}$ Ampicillin, 50 µg ml$^{-1}$ Spectinomycin, 50 µg ml$^{-1}$ Kanamycin, 37º C, 120 rpm</td>
</tr>
</tbody>
</table>

In each case, growth conditions are indicated.

### 2.2.3 Expression of trigger factor variants

The trigger factor expression constructs in plasmid pDS56/RBSII (Bujard et al., 1987; Hesterkamp et al., 1997) were obtained from the groups of Prof. Bernd Bukau (B.B.) and Dr. Elke Deuerling (E.D.) (ZMBH Heidelberg). The trigger factor expression constructs were ligated into the single BamHI site of the expression vector pDS56/RBSII, 6 × His resulting in
expression products that have four additional amino acids at the amino terminus (MRGS) and eight additional amino acids at the carboxyl terminus (RSHHHHHH) (Fountoulakis et al., 1990). For purposes of protein expression, plasmids were transformed into *E. coli* DH5α cells containing pZA4 which encodes lacIq to ensure a controlled trigger factor expression (Tomoyasu et al., 2001). We have transformed pDS56/RBSII expression constructs into BL21 (pREP4) cells for test expression or large scale expression in LB medium (as described in sections 2.2.1 or 2.2.2).

### 2.2.4 Expression of selenomethionine derivatised trigger factor

MC4100 (∆tig::Kan) (pH6-tig/pZA4) cells were propagated in 50 ml LB<sub>Amp,Kan,Sp</sub> medium at 37º C to be washed and resuspended in M9 minimal medium containing all aminoacids (50 mg l<sup>-1</sup>) except methionine. 10 ml of the resuspended preculture was then added to 1 l of M9 minimal medium containing Kanamycin, Ampicillin and Spectinomycin. Cells were grown at 37º C with constant shaking at 120 rpm. At OD<sub>600</sub> = 0.5, solid L-selenomethionine (50 mg l<sup>-1</sup>) was added and the cells were grown for another 30 min before protein expression was induced by the addition of 1 ml of 1 M IPTG per liter of culture. Cells were incubated for another hour before they were harvested at 4º C. Pellets were transferred into 50 ml Greiner tubes and flash frozen in liquid nitrogen.

### 2.2.5 Cross-linking experiments

The native trigger factor protein does not contain any cystein residues and was therefore perfectly suited for site-specific cross-linking experiments by engineering amino acid residues to cysteine at appropriate positions. This allowed the cystein specific attachment of UV-activatable cross-linkers such as BPIA to the modified trigger factor protein. The maleimide group within BPIA can be covalently linked to free thiol groups. Its photoreactive benzophenone group generates free radicals upon irradiation with UV light that readily react with proteins in their vicinity. For cross-linking experiments to ribosomal proteins in the group of Prof. Bernd Bukau and Dr. Elke Deuerling (ZMBH Heidelberg) trigger factor mutants were generated that contained cystein mutations within their ribosome-binding region (Ferbitz et al., 2004; Kramer et al., 2002b). In the case of *H. marismortui* 50S the trigger factor mutant V49C proved to be most efficient in the cross-linking reaction (in contrast to *E. coli* ribosomes where the mutation D42C was more efficient). In each case 60 µM trigger factor in 750 µl 20 mM Hepes pH 7.0, 50 mM NaCl, 0.1 mM EDTA was reduced for 30 min at RT with 50 µl 10 mM TCEP. A UV activatable maleimide was covalently linked to the
introduced cystein sidechain via a Michael addition using 2 µl of 100 mM BPIA to be incubated for 2h at 30° C in the dark. The reaction was quenched by 4-fold molar excess of β-mercaptoethanol and dialysed against 20 mM Heps pH 7.0, 50 mM NaCl, 0.1 mM EDTA overnight. The efficiency of the reaction was analysed by mass spectrometry and proved to be ~100 % efficient. The activated protein was added in 2-fold molar excess over H. marismortui 50S, incubated for 30 min at 30° C in the presence of 0.6 M NaCl (Kramer et al., 2002b) and irradiated with UV light on ice for 5 min (365 nm, 100 W, Ultraviolet Products, Model B-100AP) at a distance of 5 cm. Ribosomal complexes were separated by centrifugation through sucrose cushions (Hesterkamp et al., 1997). Protein identification was done by in-gel-digest of the crosslinking products with trypsin followed by mass spectrometry (Kramer et al., 2002b).

2.3 Growth of Haloarcula marismortui cells

Cells were grown in a variation of the ATCC medium 1230 (http://www.lgcromochem.com/atcc/) at 30º C. A 100 ml preculture of OD_{600} = 1.0 – 1.5 was used to inoculate 6 l of medium, which was stirred with 100 rpm and air aerated. Usually after 3 - 4 days the cells reached the logarithmic growing phase and were harvested with an OD_{600} = 1.2 – 1.6 to be stored at -80º C.

2.4 Protein purification methods

2.4.1 Cell lysis

Frozen cell pellets were thawed at room temperature and the volume was adjusted to 100 ml with NiNTA lysis buffer. After the addition of 50 µg of DNasel the suspension was incubated on ice for 30 min and passed once through the cell cracker at a pressure of 15,000 psi. Alternatively, cells were lysed at 20,000 psi (150,000 kPa) in an Emulsiflex-C5 high-pressure homogeniser (Avestin), which was precooled on ice. The lysate was cleared by centrifugation for 1 h at 20,000 rpm in a SS-34 rotor at 4º C. The supernatant was immediately applied to a NiNTA resin for purification of the histidine tagged proteins in a metal affinity purification procedure.

2.4.2 TEV cleavage

TF111 cloned into the plasmid pRET31H6TRX contained a TEV cleavage site, which allowed the removal of N-terminally fused His$_6$-tagged thioredoxin after affinity purification. 250 U of TEV protease (Invitrogen) were incubated with 0.5 mg protein overnight at 4º C.
After the reaction the His\textsubscript{6}-tagged protease and still uncleaved substrate was removed by metal affinity purification using NiNTA agarose. The successfully cleaved product was isolated in the supernatant.

2.4.3 Metal affinity purification

2.4.3.1 Column purification

10 ml of NiNTA Superflow resin (Amersham Biosciences) was equilibrated in NiNTA lysis buffer prior to the addition of the cleared lysate. The resin was washed with 5 column volumes (CV) of NiNTA lysis buffer and 10 CV of NiNTA wash buffer at 4\(^\circ\) C. Protein bound to the NiNTA resin was eluted with NiNTA lysis buffer containing 1 M imidazole (NiNTA elution buffer) using a linear or step gradient. All fractions were analysed with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing the desired protein were pooled for further purification.

2.4.3.2 Batch purification

Six to eight millilitres of Ni-NTA agarose (Qiagen) was equilibrated with 5 volumes of cold French press buffer in a suction filter. The supernatant from the centrifugation step was supplemented with 10 mM MgCl\textsubscript{2}, incubated with the equilibrated Ni-NTA agarose on ice for 15 min, and then passed through the suction filter. After being washed with at least 0.5 litre of ice-cold washing buffer (50 mM Tris-HCl, 20 mM imidazole, 500 mM NaCl, pH 7.5) and 5 volumes of ice-cold low-salt buffer (50 mM Tris-HCl, 20 mM imidazole, 25 mM NaCl, pH 7.5), the protein was eluted from the Ni-NTA agarose with 3 volumes of cold elution buffer (50 mM Tris-HCl, 250 mM imidazole, 25 mM NaCl, pH 7.5).

2.4.4 Anion exchange chromatography

The eluate from the affinity purification procedure was directly applied to an anion exchange column (ResourceQ6; Pharmacia) at 4\(^\circ\)C, and eluted with a salt gradient (going from a low-salt buffer consisting of 50 mM Tris-HCl [pH 7.5], 25 mM NaCl, and 1 mM EDTA to a high-salt buffer consisting of 50 mM Tris-HCl [pH 7.5], 1 M NaCl, and 1 mM EDTA). Under these conditions, the trigger factor elutes at 150 mM NaCl (15 % high-salt buffer). The times of elution of the trigger factor fragments or mutants may vary, and all elution fractions were analysed with SDS-PAGE. Appropriate fractions were pooled and dialysed against 2 litres of storage buffer (20 mM Tris-HCl [pH 7.5], 100 mM NaCl, 1 mM EDTA) overnight at 4\(^\circ\)C.
2.4.5 Size exclusion chromatography

For a final purification step the protein was dialysed against the respective running buffer, concentrated to 10 - 15 mg ml$^{-1}$ and 200 - 500 µl samples were injected onto a Superdex 200 HR 10 x 300 mm (Amersham Biosciences) column. At a flowrate of 0.5 ml min$^{-1}$ 0.5 – 1 ml fractions were collected and analysed with SDS-PAGE. The desired protein fractions were pooled and finally concentrated to 10 – 50 mg ml$^{-1}$.

2.5 Protein analysis

2.5.1 Denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

This analytical procedure was employed to estimate the amount and purity of a protein sample. SDS causes proteins to unfold and binds non-specifically with a ratio of 1.4 g SDS to 1 g of protein. As a result, proteins treated with SDS will have a net negative charge and migrate to the anode of the electrophoretic system. In the polyacrylamide gel proteins will be separated by their mass. 2x SDS loading buffer was added to an equivalent volume of protein sample. The mix was boiled at 95º C for 2 min and immediately loaded on a polyacrylamide gel. The gel was run at room temperature with a current of 20 mAmp in 1 x SDS running buffer for 1 h. After electrophoresis, protein bands were visualised either by silver or Coomassie staining.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Volume [ml] Separating gel</th>
<th>Volume [ml] Stacking gel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower gel buffer (pH 6.8)</td>
<td>1.34</td>
<td>1.34</td>
</tr>
<tr>
<td>Upper gel buffer (pH 8.8)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>1.65</td>
<td>1.00</td>
</tr>
<tr>
<td>30 % acrylamide:bisacrylamide (37.5:1)</td>
<td>2.00</td>
<td>2.50</td>
</tr>
<tr>
<td>25 % APS</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>TEMED</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 3: Pipetting scheme for denaturing sodium dodecyl sulfate-polyacrylamide gels

In dependence of the molecular weight of the protein to be separated, resolving SDS-polyacrylamide gels were prepared with acrylamide:bisacrylamide (37.5:1) concentrations of 12% or 15%.

2.5.2 Coomassie staining

Following SDS-PAGE, protein bands were visualised by staining the gel with Coomassie staining solution for 30 min and destaining for 45 min in destaining solution at room temperature (solutions are listed in Appendix 7.1.1).
2.5.3 Silver staining

After SDS-PAGE, the gel was immersed in a mixture of 40 % methanol and 10 % acetic acid for 30 min. Proteins were stained with reagents of the Silver stain Kit (BioRad) according to the manufacturer’s recommendation. After fixation in methanol/acetic acid, the gel was incubated in oxidising solution for 5 min and washed with water. The gel was transferred in silver reagent for 20 min followed by a quick wash with water. Protein and nucleic acid bands became visible after treating the gel with developer solution. The reaction was stopped with 5 % acetic acid.

2.5.4 Determination of concentration

Protein concentrations were determined by UV-absorption spectroscopy. The optical density at 280 nm (OD\textsubscript{280}) was measured and the protein concentration was calculated using the corresponding extinction coefficient (ε) and molecular weight as determined with the program ProtParam (http://www.expasy.ch/tools/protparam.html). Alternatively, the protein concentration was measured by Bradford assay using the Bio-Rad protein assay solution (Biorad) according to the manufacturer’s recommendation. After the addition of the acidic dye absorption was measured at 595 nm. The comparison to a BSA standard curve provided a relative measurement of the protein concentration.

Ribosome concentrations were determined by UV-absorption spectroscopy at 260 nm (OD\textsubscript{260}). 1 OD\textsubscript{260} corresponds to 60 mg ml\textsuperscript{-1} 50S with a molecular weight of M\textsubscript{w} = 1.6 \cdot 10^6 g mol\textsuperscript{-1}. In almost all crystallisation attempts samples with 10 mg 50S in 1 ml solution were used which corresponds to a 50S concentration of 6.25 µM.

2.5.5 Electro-blotting and N-terminal sequencing

After SDS-PAGE the protein bands were electrotransferred to PVDF-membranes (Biorad) using a semi-dry transfer unit (Hoefer TE 70, Amersham Pharmacia Biotech). Towbin buffer (Tris-HCl 25 mM, glycine 192 mM, SDS 3.5 mM, 15 % methanol) was used as blot running buffer. The membrane was pre-incubated in methanol for 15 s, washed with ddH\textsubscript{2}O and soaked with blot running buffer prior to transfer. Electroblotting was performed at 50 V for 2 hours. The PVDF membrane was stained with staining solution (0.5 % Poinceau S, 1 % acetic acid) to check for blotting efficiency and determination of the size of the blotted protein. For sequencing the corresponding band was cut out and submitted to the Protein Service Laboratory of the Department of Biology (ETH Zurich) for N-terminal sequencing by Edman degradation and HPLC. For the detection of His\textsubscript{6}-tagged proteins, the membrane was washed
twice for 10 min with TBS buffer and incubated in blocking buffer for 1 hour. After washing twice with TBS-TWEEN/Triton and once with TBS for 10 min each time the primary Penta-His antibody (Qiagen) was added to the membrane in 1/1000 dilution (0.2 µg ml\(^{-1}\)) in blocking buffer to be incubated for at least 1 hour or overnight. After that, the membrane was washed once again with TBS-TWEEN/Triton and TBS before the secondary antibody, anti-mouse IgG conjugated with horseradish peroxidase, was added as 1/20,000 dilution (0.04 µg ml\(^{-1}\)) in blocking buffer. After washing four times with TBS-TWEEN/Triton for 10 min each, the tagged protein was detected by a chemiluminescent reaction using the NBT/BCIP solution (Roche) according to the manufacturers recommendation.

2.5.6 Mass spectroscopy

Protein samples were dialysed against low salt buffer or ddH\(_2\)O to be submitted for matrix-assisted laser desorption and ionisation time of flight (MALDI-TOF) mass spectrometry analysis to the Protein Service Laboratory of the Institute of Molecular Biology (ETH Zurich).

2.6 Ribosome purification

The whole purification procedure was carried out at 6º C as the high salt content of the solutions used results in the formation of salt crystals at lower temperatures. These might damage the centrifugation tubes especially at high speeds of the ultracentrifugation runs. Cells were lysed in 70S buffer (30 ml for 10 g cell pellet) at 20,000 psi (150,000 kPa) in the EmulsiFlex-C5 high-pressure homogeniser (Avestin). Usually cells had to be passed 3 to 4 times to ensure complete lysis. The cell lysate was centrifuged at 12 k for 40 min in a SS34 rotor (Sorvall) prior to ultracentrifugation at 30 k for 40 min using a Ti70 rotor (Beckman). The cleared lysate was underlied with a cushion of 6 ml 30 % sucrose containing 100 mM MgCl\(_2\) and centrifuged for 15.5 h at 50 k in a Ti70 rotor. Afterwards, the glassy pellet was washed with 1 ml of the dissociation buffer and resuspended in 1 ml of the dissociation buffer by shaking on ice for 1 – 2 hours. The resuspended ribosomal pellet was centrifuged for 15 min at 13 k and loaded on a 6 – 30 % sucrose gradient in the dissociation buffer. For a proper separation of the 30S and 50S subunits not more than 16 mg of ribosomes at a concentration of less than 25 mg ml\(^{-1}\) were loaded. The gradients were centrifuged at 22 k for 17 hours in a SW28 rotor (Beckmann). Next, the tubes were pierced from the bottom to allow the collection of the 1.2 ml samples using a peristaltic pump. During the collection the absorption of the sample at 260 nm was monitored to identify the ribosome containing fractions. The fractions
comprising the first peak of the elution profile contained the large ribosomal subunit and were dialysed. Afterwards, 50S was concentrated to 10 mg ml\(^{-1}\) and 200 µl aliquots were precipitated with half volume of 30 % (w/w) PEG6000 for storage at 4º C. Samples were kept for no longer than 1 – 1.5 weeks before being used for crystallisation.

**2.7 Crystallisation methods**

**2.7.1 Vapour diffusion**

**2.7.1.1 Trigger factor**

2 to 4 µl of the protein solution were mixed with an equal volume of reservoir buffer in a Cryschem sitting drop crystallisation plate (Hampton Research). The drop was equilibrated against 1 ml of reservoir buffer at 4º C or 19º C. The plate was sealed with crystal clear duct tape. Alternatively, a linbro plate (Hampton Research) for hanging drop crystallisation was used. Sample and buffer were mixed on a siliconised 22 mm circle cover slide (Hampton Research). The slide was sealed to a well containing 1 ml of corresponding reservoir buffer using vacuum grease. Crystal screen formulations are listed in the Appendix.

**2.7.1.2 Large ribosomal subunit (*H. marismortui*)**

The purified large ribosomal subunits (50S) were concentrated to 10 mg ml\(^{-1}\) (6.25 µM) and precipitated with PEG6000 by adding half volume of 30 % (w/w) PEG6000 stock solution to the sample and incubating 1 hour on ice. The white precipitate was pelleted for 1 min at 13,000 g at 4º C and resuspended in final buffer to yield a ribosomal concentration of 20 mg ml\(^{-1}\). After pre-incubation for at least 1 hour at RT, the sample was either directly used for crystallisation or was mixed with 1 volume of reservoir buffer and centrifuged for 2 min at 13,000g at RT. The saturated ribosome solution was then used for vapour diffusion crystallisation either in Cryschem sitting drop crystallisation plates (Hampton Research) or linbro hanging drop crystallisation plates (Hampton Research).

**2.7.2 Cryo-protection of crystals**

Cryo-crystallography reduces radiation damage and permits data collection with higher signal-to-noise ratios (Pflugrath, 2004). Unfortunately, in many cases, if a crystal is flash frozen directly from its growth solution the water inside the crystal might form ice causing cracks or lattice distortions. One way to avoid ice formation during crystal freezing is to
exchange the mother liquor with a suitable cryoprotectant. Cryoprotectants are either bifunctional molecules (like glycerol, MPD, PEG) or salts (LiCl). Wells containing crystals of suitable quality for data collection were opened and the drop solution was successively exchanged with increasing concentrations of cryoprotectant usually ending with a final concentration of 20 % to 25 % of cryoprotectant. Finally, the reservoir buffer was exchanged with the final cryoprotectant buffer and the wells were sealed again. All crystals were stabilised at 4º C. After at least overnight equilibration, the stabilised crystals were flash frozen in liquid propane.

2.7.3 Soaking crystals with heavy atom compounds or ligands

Stabilised crystals were transferred on a siliconised 22 mm circle cover slide (Hampton Research) into a 10 µl drop of final stabilisation solution containing an excess of the desired heavy atom compound or, in case of ribosome crystals, the ligand of interest and were incubated for at least 12 hours at 4º C.

2.8 X-ray data collection

2.8.1 Crystal mounting under cryo-conditions

Crystals, pre-equilibrated in cryo-solution and chilled to 4º C, were mounted on cryoloops (Hampton Research) of sufficient diameter under visual inspection using a light microscope. The loops were either glued to Yale or Hampton Research cryo-pins. All crystals were immediately transferred into a vial filled with liquid propane and stored in liquid nitrogen. For data collection, the loop containing the frozen crystal was mounted on a goniometer head, where a cold nitrogen gas stream kept the temperature at 100 K. In the case of the crystals from the large ribosomal subunit, loops were first glued onto a thin fishing line to allow the bending of the loop by almost 90º. In this manner the plate-like crystals can be aligned along the long crystal axis without being dependent on a diffractometer system with kappa geometry. In addition, the data collection of ribosomal crystals critically depends on the use of a large detector, so that all spots are sufficiently separated. The mar345 Image Plate Detector System (MAR Research) provided the largest detection area of the available detectors with its diameter of 345 mm but resulted in longer collection times due to longer read-out cycles than can be obtained with CCD-detectors. The data was collected with a minimal 90 degree sweep from one major zone in the orthorhombic crystal to the next in order to reduce the overall exposure time and avoid crystal decay during the data collection.
procedure. This strategy was necessary since the long axis of the crystal (575 Å) had to be aligned with the rotation axis to prevent spot overlap at high resolution.

2.8.2 Data collection

2.8.2.1 Trigger factor

The diffraction data was collected at the Swiss Light Source (SLS) X06SA PX beamline (Paul Scherrer Institut, Villigen). All crystals were frozen in liquid propane and the crystallographic data was collected at cryogenic temperature (100 K). Native data was collected in 1º oscillations using X-rays with 1.0 Å wavelength. Diffraction patterns were recorded with a 165 mm marCCD detector (MAR Research) and processed with the HKL software package (Otwinowski Z., 1997).

2.8.2.1.1 Multiwavelength anomalous dispersion (MAD) experiment

Trigger Factor crystals soaked with an anomalously scattering compound or derivatised with selenomethionine were mounted onto the goniometer head and an X-ray absorption spectrometry scan was performed to determine the exact absorption edge of the anomalous scatterer. Anomalous scattering factors were plotted against the beam energy. Based on the diagram, three different X-ray energy values were chosen for data collection; at the peak and at the inflection point of the absorption edge, and at a wavelength that is remote to these wavelengths. For each wavelength, the crystal was exposed in 1° oscillations for 4 s using two aluminium filters to attenuate the beam to 23 % of its maximal intensity or for 1.2 s without attenuation. Up to 240 oscillation frames per wavelength were measured from the same crystal under cryogenic temperatures.

2.8.2.2 Large ribosomal subunit (\textit{H. marismortui})

Diffraction data was collected either at the Swiss Light Source (SLS) X06SA PX beamline (Paul Scherrer Institut, Villigen) or at the Swiss Norwegian Beamline (SNBL) at the ESRF in Grenoble. Crystals were exposed at wavelengths of 0.9 Å or 1.0 Å during 1° oscillations until a complete dataset was obtained. Diffraction patterns were recorded with a 345 mm mar345 image plate detector and processed with the HKL software package (Otwinowski Z., 1997).
2.9 Computational methods

2.9.1 Sequence manipulation and analysis

DNA and protein sequences were analysed and manipulated using the DNAMAN software package (Lynnon, 2001) and sequences were aligned with CLUSTAL X 1.81 (Thompson et al., 1997). DNA sequences from ABI PRISM automated sequencing runs where converted to plain text or FASTA formats using the program Chromas v.1.45 (Mc Carthy, 2001). Physicochemical properties of proteins were calculated with PROTPARAM, database searches performed with GenBank, Swiss-Prot and TrEMBL using the ExPASy (Expert Protein Analysis System) proteomics server (Gasteiger et al., 2003).

2.9.2 Computer software for structure determination

All program packages and input scripts that have been used during structure determination and representation are listed together with a short description and the corresponding reference (Table 4).

Table 4: All program packages and input scripts used during structure determination and representation are listed

<table>
<thead>
<tr>
<th>Package</th>
<th>Routine/Script</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNS 1.1</td>
<td>analyse.inp</td>
<td>Diffraction data analysis for amplitudes, intensities and phases</td>
<td>(Brunger et al., 1998)</td>
</tr>
<tr>
<td></td>
<td>bindividual.inp</td>
<td>Restrained, individual B-factor refinement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>combine.inp</td>
<td>Calculation of phases with optional phase combination from different sources</td>
<td></td>
</tr>
<tr>
<td></td>
<td>composite_omit_map.inp</td>
<td>Calculates a composite annealed omit map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>contact.inp</td>
<td>Determination of contacting residues between two sets of atoms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>densityModify.inp</td>
<td>Density modification to improve phases</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fo-fo_map.inp</td>
<td>Calculates an electron density map using phase information from a model and two experimental amplitude data sets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fourier_map.inp</td>
<td>Computes FOM-weighted Fourier synthesis and difference Fourier maps</td>
<td></td>
</tr>
<tr>
<td></td>
<td>generate.inp</td>
<td>Generates structure files for protein, DNA/RNA, water, ligands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>get_ncs_matrices.inp</td>
<td>Determination of NCS operators between two molecules</td>
<td></td>
</tr>
<tr>
<td></td>
<td>heavy_search.inp</td>
<td>Heavy atom search and Patterson correlation refinement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mad_phase.inp</td>
<td>Computes phase probability distributions from a MAD experiment and refines anomalous scatterer parameters against MAD data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>make_cv.inp</td>
<td>Sets up a test array for cross-validation using a random set of data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>matthews_coeff.inp</td>
<td>Calculates the Matthews coefficient and estimates the solvent content</td>
<td></td>
</tr>
<tr>
<td></td>
<td>merge.inp</td>
<td>Merges datasets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>model_phase.inp</td>
<td>Generates calculated amplitudes, phases, Hendriksen-Lattman coefficients, FOM and bulk solvent correction from model(s) crystallographic model statistics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>model_stats.inp</td>
<td>Computes an annealed omit map followed by a single cycle of electron density averaging</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ncs_average_map.inp</td>
<td>Defines non-crystallographic constrains and restraints</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ncs.def</td>
<td>Improves NCS operators for density averaging</td>
<td></td>
</tr>
<tr>
<td></td>
<td>optimize_ncsop.inp</td>
<td>Optimises R-weight (X-ray/B-restraints weight)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>optimize_rweight.inp</td>
<td>Native and difference Patterson map and Harker section calculation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>patterson_map.inp</td>
<td>Patterson map and Harker section prediction from a set of heavy atom sites</td>
<td></td>
</tr>
<tr>
<td></td>
<td>predict_patterson.inp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4 continued

<table>
<thead>
<tr>
<th>Package</th>
<th>Routine/Script</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNS 1.1</td>
<td>qgroup.inp</td>
<td>Grouped, unrestrained occupancy refinement</td>
<td>(Brunger et al., 1998)</td>
</tr>
<tr>
<td></td>
<td>qindividual.inp</td>
<td>Individual, unrestrained occupancy refinement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>refine.inp</td>
<td>Combines crystallographic simulated annealing, energy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rigid.inp</td>
<td>minimisation, individual B-factor refinement and map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rmsd.inp</td>
<td>calculation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scale.inp</td>
<td>Crystallographic rigid body refinement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sdb_to_pdb.inp</td>
<td>Calculates rms coordinate difference between two structures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>xtal_pdbsubmission.inp</td>
<td>Converts a heavy atom site database file into a pdb coordinate file</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Produces partial information required for pdb submission</td>
<td></td>
</tr>
<tr>
<td>Dali 2.0</td>
<td>Dali</td>
<td>Three-dimensional, sequence independent homologous structure search</td>
<td>(Holm and Sander, 1993)</td>
</tr>
<tr>
<td>Delphi 1.1</td>
<td>Delphi</td>
<td>Program package that calculates electrostatic potentials in and around macromolecules</td>
<td>(Rocchia et al., 2002)</td>
</tr>
<tr>
<td>Dino 0.90</td>
<td>Dino</td>
<td>Real-time 3D visualisation program for structural biology data</td>
<td>(Philippsen, 2003)</td>
</tr>
<tr>
<td>Pymol 0.97</td>
<td>Pymol</td>
<td>Structure visualisation</td>
<td>(DeLano, 2002)</td>
</tr>
<tr>
<td>RAVE</td>
<td>MAMA, MAPMAN</td>
<td>Mask generation and manipulation</td>
<td>(Kleywegt and Jones, 1999)</td>
</tr>
<tr>
<td>Raster3D</td>
<td>Raster3D</td>
<td>Manipulation of map and mask formats and parameters</td>
<td></td>
</tr>
<tr>
<td>SOLVE/ RESOLVE 2.06</td>
<td>solve.com</td>
<td>Program for generating high quality raster images</td>
<td>(Merritt and Bacon, 1997)</td>
</tr>
<tr>
<td></td>
<td>solve.setup</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>resolve.com</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weblogo 2.81</td>
<td>Weblogo</td>
<td>Web based application to generate sequence logos</td>
<td>(Crooks et al., 2004)</td>
</tr>
</tbody>
</table>

Table 4: References for each program package as well as a short description of each routine or script are given.
3 Results

3.1 Crystallisation of the trigger factor, crystal handling and data collection

The structure determination of macromolecules by X-ray diffraction analysis depends on the availability of crystals that diffract to suitably high resolution. For most crystallographic projects crystallisation is still the bottleneck in the process of structure determination due to its unpredictability. In order to increase the chance of succeeding in the structure determination of the trigger factor it was decided to attempt crystallisation of several trigger factor constructs from *Thermus thermophilus* or *Escherichia coli*. The identification of initial crystal growth conditions and their refinement to yield single, well diffracting crystals are described in the following chapters.

3.1.1 Initial screening conditions

Several different expression constructs of the trigger factor (Material and Methods section 2.2.3) were tested in the initial screening conditions (Figure 5) comprising the N-terminal domain (TF144), a construct lacking the protease sensitive 47 N-terminal amino acids (Δ47), the full-length trigger factor mutant TF-AAA (GFR_{45} -> AAA_{45}) or the full-length wild-type protein. Trigger factor variants purified according to the batch purification procedure (Material and Methods section 2.4.3.2) were obtained in collaboration with BB and ED or alternatively trigger factor variants were purified according to the column purification procedure (Material and Methods section 2.4.3.1). In all cases the protein purity was estimated by SDS-PAGE. In the first set of experiments on all trigger factor constructs the following commercially available screens were used (formulations are listed in the Appendix section 7.2.3.1): Hampton Crystal Screens I and II (Cudney et al., 1994; Jancarik and Kim, 1991), Hampton Nucleic acid Sparse Matrix Screen (Natrix) (Scott et al., 1995), Hampton Crystal Screen Light and Hampton Nucleic Acid Mini Screen (Adams et al., 2002; Berger et al., 1996). In addition several grid screens were tested: PEG400 vs. pH, PEG4000 vs. pH, ammonium sulphate vs. pH and MPD vs. pH (McPherson, 1999) (Appendix section 7.2.3.2).
Chapter 3

Results

Figure 5: Trigger factor variants used for crystallisation
The trigger factor expression constructs used for crystallisation comprised the N-terminal ribosome-binding domain (TF144), a construct lacking the protease sensitive 47 N-terminal amino acids (Δ47), the full-length trigger factor mutant TF-AAA (GFR45 -> AAA45) or the full-length wild-type protein.

3.1.2 Thermus thermophilus trigger factor

The full-length, wild-type trigger factor protein was concentrated to 10 mg ml\(^{-1}\) and subjected to initial screening conditions. As a result of these experiments thin, needle-like, crystals up to 50 µm in length appeared when sodium formate was used as a precipitant (Figure 6). Several attempts to reproduce these crystals including fine screening around the initial conditions failed and consequently the investigation was continued with the remaining trigger factor constructs.

Figure 6: Thermus thermophilus trigger factor crystals
Crystals were obtained with 2.0 M sodium formate, 0.1 M sodium acetate at pH 4.6 (Hampton Crystal Screen I, condition 41) at 19º C, 10 mg ml\(^{-1}\) protein.
3.1.3 *E. coli* trigger factor constructs

The *E. coli* trigger factor expression constructs TF144, Δ47, TF-AAA and the wild-type protein were tested for crystallisation. All purified protein samples were subjected to initial crystallisation screens at a concentration of 10 mg ml\(^{-1}\). In the case of the N-terminal domain and the TF-AAA mutant no crystals were observed under any of the tested conditions.

3.1.3.1 Δ47 trigger factor construct

The first *E. coli* trigger factor crystals were observed from a construct that was shortened by 47 amino acids at its N-terminus. In the initial screening a condition was identified under which plate-like crystals appeared after 3 to 4 weeks from a drop that initially showed strong phase-separation (Figure 7). These initial crystals were rather small and irregular.

![Figure 7: Δ47 crystallisation](image)

Crystals were obtained in 30 % PEG4000, 0.1 M Tris HCl pH 8.5, 0.2 M sodium acetate (Hampton Crystal Screen 1, condition 22) at 19º C, with 10 mg ml\(^{-1}\) protein.

The quality of the initial crystals was improved by screening a pH range between pH 7.9 and 9.0 in steps of 0.2 pH units for PEG4000 concentration between 20 % and 34 % in 2 % steps. For each condition crystal trays were incubated at 4º C and 19º C using 10 and 25 mg ml\(^{-1}\) of protein. After 1 month at 19º C crystals with defined edges appeared in 20 % to 24 % PEG4000, 0.1 M Tris-HCl pH 7.9 - 8.5 in the presence of 0.2 M sodium acetate using 25 mg ml\(^{-1}\) protein (Figure 8).
Chapter 3

Results

Figure 8: Refined condition for Δ47 crystals
Thin plates of up to 100 μM in two dimensions were obtained in 22 % PEG4000, 0.1 M Tris HCl pH 8.5, 0.2 M sodium acetate at 19º C, with 25 mg ml⁻¹ protein.

Crystals were first transferred to 4º C and stabilised in ethyleneglycol, glycerol or MPD by step-wise addition of increasing amounts of cryoprotectant. The final concentrations used were 20 % ethyleneglycol, 15 % glycerol or 25 % MPD. Crystals were equilibrated in the final cryoprotectant solution overnight before flash freezing in liquid propane.

Several crystals were tested for diffraction at the X06SA PX beamline (Paul Scherrer Institute, Villigen, Switzerland). Diffraction patterns were recorded with a 165 mm marCCD detector at a wavelength of λ = 1 Å and exposure times of 1.5 s. Crystals stabilised in either ethyleneglycol or MPD showed no or very few diffraction spots at very low resolution. Crystals stabilised in 15 % glycerol diffracted to 8 Å resolution (Figure 9). These crystals severely suffered from radiation damage and only very few images per crystal could be recorded. The obtained diffraction pattern showed spots from multiple crystal lattices, but could be indexed on a primitive tetragonal lattice with unit cell dimensions of 71 Å x 71 Å x 188 Å. The volume of the measured unit cell was 9.477·10⁵ Å³ and the number of trigger factor molecules accommodated by the unit cell was estimated to be eight using the method proposed by Matthews (Matthews, 1968). This would result in one or two molecules in the asymmetric unit, depending on the crystallographic symmetry of the space group. In this case the solvent content of the unit cell, having a volume to mass ratio (Matthew’s coefficient) of $V_M = 2.72 \text{ Å}^3 \text{Da}^{-1}$, is 53 %. Despite extensive attempts to refine the crystallisation conditions the diffraction properties of the obtained crystals could not be improved.
Figure 9: Δ47 trigger factor crystals diffracted to a maximum resolution of 8 Å
The crystal lattice was determined to be primitive tetragonal with unit cell dimensions of 71 Å x 71 Å x 188 Å.

3.1.3.2 Full-length trigger factor

Full-length trigger factor crystals were previously grown from 10 % iso-propanol, 0.1 M Hepes pH 7.5 and 20 % PEG4000 at 19º C (H. Patzelt personal communication), but it was not possible to reproduce these conditions in spite of screening different protein concentrations (10 mg ml\(^{-1}\) and 30 mg ml\(^{-1}\)) and various additives (Appendix 7.2.3.3). Since at 19º C no crystals were obtained in any of the screens tested, the crystallisation screening was also conducted at 4º C. At this temperature and using highly concentrated protein (50 mg ml\(^{-1}\)) needle-like crystals were observed (Figure 10A).
The size of initially obtained crystals was significantly improved by screening with additives (Figure 10B-C). Various types of alcohols proved to be most successful: 1,3-butanediol; 1,6-hexanediol; 2,5-dimethyl-2,5-hexanediol and 1,4-butanediol. Furthermore, an increase of the iso-propanol concentration to 14% provided crystals with sharp edges and of sufficient size to be tested for diffraction at the synchrotron (Figure 10D).

Needles of 100 x 10 x 10 µM were stabilised in either 25% MPD, 20% ethyleneglycol or 15% glycerol by step-wise addition of increasing amounts of cryoprotectant. In addition, glycerol was tested as a cryoprotectant at a constant PEG4000 concentration (20%) and in combination with an increase of the polyethylene glycol concentration to 25%. All crystals were equilibrated overnight in the final cryoprotectant concentration before flash freezing in liquid propane. Overnight addition of ethyleneglycol dissolved the crystals.

Crystals stabilised in glycerol or glycerol with increased PEG concentration were severely damaged as became evident in their diffraction pattern with a resolution limit of 6 Å. Crystals stabilised in MPD diffracted anisotropically to 4.0 Å resolution. They showed diffraction spots from multiple lattices probably due to crystal damage during stabilisation or during the freezing, presumably since they were very thin and sensitive to manipulation. In addition, they severely suffered from radiation damage during the data collection. Therefore, in order to collect a complete dataset at high resolution, the crystals had to be improved both regarding their size and their diffraction properties.

Figure 10: Initial screening of the *E.coli* trigger factor crystals
A. No additive, 10% (v/v) iso-propanol. B. 3% (v/v) ethyleneglycol, 10% (v/v) iso-propanol. C. 4% (v/v) 1,4-butanediol, 10% (v/v) iso-propanol. D. No additive, 14% (v/v) iso-propanol. All crystals were grown in the presence of 0.1 M Hepes pH 7.5 and 20% PEG4000 at 4°C.
Chapter 3  

3.2 Optimising and characterising *E.coli* trigger factor crystals  

3.2.1 Refinement of crystallisation conditions at 4º C  

In an attempt to produce better diffracting crystals the most promising additives 1,3-butanediol, 1,6-hexanediol, 2,5-dimethyl-2,5-hexanediol and 1,4-butanediol were tested at varying concentrations in the presence of 10 % iso-propanol, 0.1 M Hepes pH 7.5 and 20 % PEG4000. The crystal quality improved most in the presence of 2.4 % (v/v) 1,6-hexanediol, 3 % (v/v) 2,5-dimethyl-2,5-hexanediol and 6 % (v/v) 1,4-butanediol. The optimised concentration of all three additives was further evaluated at varying concentrations of iso-propanol and PEG4000 at different pH (Table 5).

| PEG4000 (w/v) | 16 – 26 % |
| Iso-propanol (v/v) | 8 – 14 % |
| pH | 7.0, 7.25, 7.5, 7.75, 8.0 |

Table 5: Fine screening of the initial crystallisation condition

Different PEG4000 and iso-propanol concentrations were tested at 4º C in the presence of 0.1 M Na Hepes at varying pH with either 2.4 % (v/v) 1,6-hexanediol, 3 % (v/v) 2,5-dimethyl-2,5-hexanediol or 6 % (v/v) 1,4-butanediol as additives. In all cases 50 mg ml⁻¹ protein were used. In the final screen iso-propanol was replaced with 1,4-butanediol to be tested in the same manner, but without any additives.

Single crystals with defined edges appeared after four to six weeks. Unfortunately, the size of the crystals did not increase. The additive screen where the highest concentration of 1,4-butanediol was used yielded largest crystals. Consequently, iso-propanol, which was initially identified as a suitable precipitant, was completely exchanged with 1,4-butanediol to be tested in the same manner (Table 5). After four weeks first needle-like crystals appeared in the presence of 14 % 1,4-butanediol, 24 % and 26 % PEG4000 and 0.1 M Na Hepes at pH 7.0 – 7.25. These crystals grew to form large rods of 1000 µm x 100 - 150 µm x 50 - 100 µm over a period of two to three months (Figure 11).
Figure 11: Refined conditions for trigger factor crystals
Trigger factor crystals obtained in 14 % 1,4 butanediol, 0.1M Hepes, 4°C, 50 mg ml$^{-1}$, 26 % PEG4000 at pH 7.0 (A) or pH 7.25 (B).

3.2.2 Crystal composition

Considering that more than one month elapsed before the first crystals appeared in the crystallisation experiments it was possible that the protein might have been degraded and that only a fragment of the full-length protein formed the observed crystals. Therefore, the composition of the obtained crystals was analysed by SDS-PAGE and mass spectrometry (Figure 12). For that purpose several crystals were transferred into their corresponding growth solution, carefully washed with the same solution and then solubilised in ddH$_2$O. This analysis confirmed that the crystals contained a single protein with a measured size of 49971 Da and that therefore the crystallised protein corresponds to the full-length trigger factor with a calculated mass of 49690 Da.

Figure 12: Biochemical analysis of trigger factor crystals
A. Silver stained SDS-PAGE of washed, resuspended crystals. B. The entire mass spectrum of the resuspended crystals C. Zoom into the region of the peak that confirms that the crystals contain protein with a measured mass of 49971 Da, which corresponds well with the calculated mass of 49690 Da for the trigger factor.
3.2.3 Derivatisation attempts of native crystals

The most widely used techniques to solve the phase problem in protein X-ray crystallography are the multiple isomorphous replacement method (MIR), the multiple wavelength anomalous dispersion method (MAD) and the molecular replacement method (MR). The MR method was not feasible as the structural information available for the trigger factor was restricted to the NMR structure of the *Mycoplasma genitalium* PPIase domain (Vogtherr et al., 2002). Therefore, the MAD method was considered to be most promising as it allows the use of a single crystal for the complete data collection, which eliminates problems with non-isomorphism in the MIR method (Appendix 7.5). Furthermore, the trigger factor does not contain cystein residues (Appendix 7.4), which are particularly useful for derivatisation with mercury compounds in MIR experiments.

Nevertheless, initially, in parallel with the preparation of the MAD experiment (see below) crystals were soaked with platinum (10 mM K$_2$PtCl$_4$) and gold (20 mM NaAuCl$_4$ or 20 mM KAu(CN)$_2$) compounds that are sometimes observed to bind methionine and histidine residues. Osmium (10 mM K$_2$OsO$_4$) and Ytterbium (50 mM YbCl$_3$) were also chosen as they can bind to charged regions on the protein as an anion or cation, respectively. Trimethyllead (25 mM trimethyllead acetate) was shown to bind to small hydrophobic pockets, either within the protein or formed by protein-protein contacts in the crystal lattice, or it is sometimes coordinated by two adjacent carboxylate groups (Holden and Rayment, 1991). Additionally derivatisation with a Tantalum bromide heavy atom cluster was attempted (3 mM [Ta$_6$Br$_{12}$]$^{2+}$), which interacts with proteins via ionic bonds and is a very strong scatterer (856 electrons) (Schneider and Lindqvist, 1994). In most mentioned cases the diffraction properties of crystals were destroyed by the addition of the heavy atoms. Only a crystal derivatised with [Ta$_6$Br$_{12}$]$^{2+}$ diffracted to 2.8 Å resolution (Figure 13) but statistical analysis of the collected data indicated that the heavy atom was not bound and it could not be used for phasing purposes. However, this ‘native’ data was used for characterisation of the trigger factor crystals.
Chapter 3

Figure 13: Diffraction pattern of full-length *E. coli* trigger factor crystals

Crystals of full-length *E. coli* trigger factor diffracted to 2.8 Å resolution. The space group was identified as P2₁ with unit cell parameters of \(a = 100\) Å, \(b = 47\) Å, \(c = 114\) Å and \(\beta = 113^\circ\).

3.2.4 Crystal characterisation

The data was processed with the HKL package (Otwinowski Z., 1997) and analysed using the program CNS (Brunger et al., 1998). The collected data revealed that the reciprocal lattice could be indexed in a primitive monoclinic space group. Systematic absences were found for serial reflections 0k0 with \(k = 2n +1\), conditions that is indicative of the presence of a \(2_1\) screw axis. The number of molecules in the asymmetric unit was estimated to be two by calculating the Matthews coefficient \(V_M\) (Kantardjieff and Rupp, 2003; Matthews, 1968) (Table 6).

<table>
<thead>
<tr>
<th>molecules / ASU (N)</th>
<th>Matthews coefficient ((V_M))</th>
<th>solvent content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.0</td>
<td>75.6</td>
</tr>
<tr>
<td>2</td>
<td>2.5</td>
<td>51.2</td>
</tr>
<tr>
<td>3</td>
<td>1.7</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Table 6: Calculation of the Matthews coefficient and the corresponding number of trigger factor molecules in the asymmetric unit (ASU)

The Matthews coefficient \(V_M\) \([\text{Å}^3/\text{Da}]\) is calculated with \(V_M = \frac{V}{M_w Z N}\) \((V = \text{unit cell volume}, N = \text{number of molecules} / \text{ASU}, Z = \text{ASU} / \text{unit cell}, Z = 2 \text{ for space group P2}_1\). For proteins a partial specific volume of 0.74 cm\(^3\) g\(^{-1}\) can be assumed. The most frequent value for \(V_M\) is 2.2, corresponding to \(\sim 50\%\) solvent content. Therefore, it is likely that there are two trigger factor molecules in the ASU.
Assuming that two trigger factor molecules may be present in the asymmetric unit the data was analysed by a self-rotation function to reveal if a two-fold relationship between molecules might exist (Figure 14). No strong peak was identified that could have resulted from non-crystallographic 2-fold symmetry (NCS).

![Figure 14: Self-rotation function showing section \( \kappa = 180^\circ \)](image)

Peaks on this section show positions of 2-fold symmetry in the diffraction pattern. The central peak marks the crystallographic 2-fold axis. No peaks are identified to be positions of non-crystallographic 2-fold axes. Weak peaks at the edge of the plot can be considered as background peaks.

### 3.3 Structure determination

#### 3.3.1 Selenomethionine derivatisation of the protein

Since attempts to derivatise the protein by heavy atom soaks into existing crystals failed, it was decided to produce selenomethionine derivatised protein for crystallisation. The ratio of one ordered Se atom per 10 kDa of protein is assumed to be enough for phase determination in anomalous dispersion methods. As two molecules were expected in the asymmetric unit a maximum of 24 sites out of a total of 888 residues will be occupied by selenomethionine.
Therefore, this strategy was considered to be suitable for phase determination in a MAD experiment (Appendix 7.5).

### 3.3.1.1 Expression and purification

Selenomethionine derivatised protein was produced following the modified inhibition protocol (E.D. and see Materials and Methods). In the final step the protein was purified by size exclusion chromatography and analysed with SDS-PAGE (Figure 15).

![Figure 15: Superdex200 purification of trigger factor](image)

The column was equilibrated with storage buffer (Appendix). Fractions B5 to C4 were analysed by SDS-PAGE (‘inset’) and contained relatively pure selenomethionine derivatised protein. The fractions were pooled and concentrated to 25 mg ml\(^{-1}\).

The amount of selenomethionine incorporated into the protein was determined by mass spectroscopy (Figure 16). For each methionine substituted with selenomethionine the overall molecular weight of the protein increases by 46.8 Da. The molecular weight of the derivatised trigger factor molecule was determined to be 50398.5 Da compared to the native protein with a measured molecular weight of 49836.9 Da. The overall increase by 561.6 Da corresponds to 12 incorporated selenomethionines. Therefore, all methionine sidechains in the protein were successfully derivatised.
Figure 16: Mass spectroscopic analysis of selenomethionine derivatised trigger factor

Only a part of the complete spectrogram is shown, as no other protein peaks were detectable. A. The mass spectroscopic analysis of selenomethionine derivatised protein showed a single peak at a molecular weight of 50398.5 Da. B. For comparison native trigger factor protein was analysed in parallel and showed a single peak at a molecular weight of 49836.9 Da. The mass difference of 561.6 Da corresponds to 12 incorporated selenomethionines proving the derivatisation to be 100% efficient.

3.3.1.2 Crystallisation

As it is usually observed the incorporation of selenomethionine caused the protein to be less soluble (Doublie, 1997) and, therefore, could only be concentrated to 25 mg ml\(^{-1}\). Crystallisation was performed under the same conditions that proved successful for the native protein. The obtained crystals were significantly smaller than the native with dimensions of 300 - 400 µm x 50 µm x 50 µm (Figure 17). Crystals were cryoprotected by adding MPD to 25% and in some cases soaked with 1 mM DTT for 1 hour prior to freezing in liquid propane in an attempt to completely reduce the incorporated selenomethionine.
Results

Figure 17: Crystals of selenomethionine derivatised trigger factor
Crystals were obtained with 25 mg ml\(^{-1}\) protein in 14 % 1,4 butanediol at 4\(^\circ\) C in the presence of either A. 26 % PEG4000 in 0.1M Hepes pH 7.0 or B. 24 % PEG4000 in 0.1M Hepes, pH 7.25. The average size of the crystals was 300 – 400 \(\mu m\) x 50 \(\mu m\) x 50 \(\mu m\).

3.3.1.3 Data collection

An X-ray absorption scan was recorded in the range of \(E = 12625 - 12695\) eV (corresponding to the K edge of selenium at 12657.8 eV) in order to choose the exact wavelengths for the MAD experiment (Figure 18): The peak wavelength (\(E_{\lambda,1} = 12669.45\) eV, \(\lambda_1 = 0.97879\)), the wavelength at the inflection point (\(E_{\lambda,2} = 12660.25\) eV, \(\lambda_2 = 0.97931\)) and the remote wavelength (\(E_{\lambda,3} = 126769.45\) eV, \(\lambda_3 = 0.97095\)).

Figure 18: Plot of the real (f') and the imaginary (f'') components of the anomalous scattering around the K edge of selenium as a function of the X-ray energy
The X-ray absorption spectrum was recorded from a trigger factor crystal that had been soaked with 1 mM DTT for 1 hour prior to freezing in liquid propane. Nevertheless, the plot still shows two f'' peaks corresponding to the reduced and oxidised form of selenomethionine. Data sets were collected in a sweep of 240\(^\circ\) at the absorption peak with the highest energy (\(\lambda_1 = 1.2550474\) Å), at the inflection point (\(\lambda_2 = 1.2555\) Å) and at a remote point (\(\lambda_3 = 1.2301465\) Å) in 1\(^\circ\) oscillation steps.
For all three wavelengths data was collected in 1º steps with a total sweep comprising 240º resulting in highly redundant datasets (Table 7). The exposure time was 1.2 s without attenuation of the beam.

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>Peak</th>
<th>Inflection</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Parameters (Å)</td>
<td>$a = 100.2$</td>
<td>$b = 47.4$</td>
<td>$c = 114.8$</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97879</td>
<td>0.97931</td>
<td>0.97095</td>
</tr>
<tr>
<td>Mosaicity (°)</td>
<td>0.434</td>
<td>0.383</td>
<td>0.550</td>
</tr>
<tr>
<td>Total reflections</td>
<td>233040</td>
<td>233386</td>
<td>262814</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>26467</td>
<td>27749</td>
<td>19726</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50 - 2.7</td>
<td>50 - 2.7</td>
<td>50 - 3.0</td>
</tr>
<tr>
<td>$R_{\text{sym}}$ (%)</td>
<td>6.3 (58.8)</td>
<td>4.4 (69.1)</td>
<td>4.4 (70.0)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (99.5)</td>
<td>100 (91.4)</td>
<td>99.9 (99.4)</td>
</tr>
<tr>
<td>$I / \sigma I$</td>
<td>15.8 (2.8)</td>
<td>26.1 (1.8)</td>
<td>23.0 (2.6)</td>
</tr>
<tr>
<td>$f'$</td>
<td>-7.9</td>
<td>-10.02</td>
<td>-4.8</td>
</tr>
<tr>
<td>$f''$</td>
<td>4.92</td>
<td>3.1</td>
<td>3.8</td>
</tr>
<tr>
<td>Phasing power</td>
<td>1.535</td>
<td>1.922</td>
<td>1.278</td>
</tr>
<tr>
<td>Combined FOM</td>
<td>0.322</td>
<td>0.379</td>
<td>0.286</td>
</tr>
<tr>
<td>Number of sites</td>
<td>19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Data collection statistics of the selenomethionine derivative
At three different wavelengths data sets were recorded from a selenomethionine derivatised trigger factor crystal. The crystal suffered from radiation damage as evident from the reduced resolution of diffraction at the ‘remote’ wavelength.

During initial analysis the data of the peak wavelength was scaled and postrefined using the anomalous flag in HKL (Otwinowski Z., 1997). After carefully adjusting the error model ($\chi^2$ values of 1 in all resolution shells) the comparison of the intensities of Bijvoet pairs resulted in $\chi^2$ values in the range of 1.8 and 3.0 for the data between 4 – 100 Å indicating the presence of an anomalous signal. Therefore, some of the selenomethionine was probably incorporated into the protein within well-ordered regions. For subsequent heavy atom search procedures data sets from all three wavelengths were indexed and scaled with the HKL package (Otwinowski Z., 1997).

3.3.2 MAD phasing

The subsequent calculations aimed at determining the positions of heavy atoms and phasing were conducted with CNS (Brunger et al., 1998). An anomalous difference Patterson
map was calculated to inspect the Harker sections for anomalous signals. A number of outstanding peaks were visible in Harker section \( y = 0 \) using data from the peak wavelength of the MAD experiment for the map calculation (Figure 19).

Heavy atom positions were searched with difference Patterson methods (Grosse-Kunstleve and Brunger, 1999) using the combined real and reciprocal space searches in CNS. The search procedure starts with a single heavy atom that is translated through the entire asymmetric unit and by calculating a correlation coefficient between the observed differences and the calculated heavy atom amplitudes (\( F_2F_2 \)). During the determination of additional sites the first atom is fixed at the highest value of the function and so forth. Heavy atom positions were searched using the 100 highest peaks in the Patterson search map as initial trial sites and the resulting correlation coefficients were compared (Table 8).

![Heavy Atom Search](image)

**Table 8: Heavy atom search performed with CNS**
The 100 heavy atom configurations found in 100 independent trials are shown with their corresponding correlation coefficients. Trial 13 shows a significantly higher correlation coefficient as any other trial and was therefore further evaluated in subsequent phasing procedures.

Only one of the solutions showed an outstanding correlation coefficient of 0.5243 and its predicted Patterson map agreed well with the observed Patterson map (Figure 19) providing confidence that the heavy atom search yielded a correct solution for the positions of heavy atoms.
Figure 19: Experimental and predicted Patterson Maps of Harker section $y = 0$

Patterson maps were calculated with the experimental data and with the predicted sites from the best solution of the heavy atom search procedure. Maps are contoured at $2.0 \sigma$ with increments of $0.5 \sigma$. The experimental anomalous difference Patterson map agrees well with the predicted Patterson map.

Phases were calculated with the original heavy atom configuration and its inverse image to resolve the enantiomorph ambiguity. Phases resulting from both choices of heavy atom positions were used to calculate figure-of-merit (FOM) weighted Fourier synthesis maps. A solvent boundary was visible only when the maps were computed with the initially determined heavy atom coordinates. Although mass spectroscopy results indicated that 12 selenium atoms were incorporated into the derivatised trigger factor, and preliminary crystallographic analysis suggested that there are two trigger factor molecules within the crystallographic asymmetric unit (see previous chapter), the heavy atom search procedure revealed only 17 heavy atom sites. This is not entirely surprising since two out of the 12 selenomethionine residues within the trigger factor are located in the flexible N- and C-terminal extensions that were introduced as a result of the cloning procedure (Appendix 7.4). Subsequently, using phases calculated from the 17 initially identified heavy atom sites coordinates for two additional heavy atoms were identified in difference Fourier and log-likelihood gradient maps bringing the total number of heavy atoms to 19 (peaks 18 and 19 in Figure 20).
Solvent-flipping was used to improve the quality of electron density maps calculated from the positions of heavy atoms. The estimated solvent content in the crystal, as predicted from the Mathews coefficient for two molecules in the asymmetric unit (Table 2), is 51%. Nevertheless, solvent flipping procedures resulted in the best improvement of the computed electron density maps, as judged by visual inspection, assuming 57% solvent content. Using these maps it was possible to fit in the already known crystal structure of the N-terminal domain of the trigger factor (Kristensen and Gajhede, 2003) and to build most parts of the C-terminal domain using the program O (Jones et al., 1991). Once a partial structure was built it was possible to re-evaluate the coordinates for the heavy atoms used in phasing. At this point it became evident, that one of the initially determined sites did not correspond to a real heavy atom position and was therefore omitted. In addition, the so far omitted peak at position 21 in the positive peak list of the difference Fourier map corresponded to a predicted selenomethionine position within a loop region of the C-terminal domain and was therefore included for phasing (Figure 20). Electron density maps computed from the corrected 19 sites showed more features of the protein and allowed the placement of the two PPlase domains of the trigger factor (Figure 21A). The 20th expected selenium site, which was never identified in heavy atom search procedures, turned out to be part of a poorly ordered loop within the N-terminal domain in one of the two molecules in the crystal. At this point of model building the calculated electron density maps were of sufficient quality to allow building of 85% of the
trigger factor sequence with the exception of the most flexible regions of the trigger factor corresponding to loops and inter-domain linkers.

### 3.3.3 Model building and refinement

All aspects of model building was carried out with program O (Jones et al., 1991) and refinement with program CNS (Brunger et al., 1998). Data from crystals derivatised with selenomethionine were used for refinement due to non-isomorphism with the native data. Using the almost complete model, several cycles of maximum likelihood refinement with amplitudes and experimental phase probability distribution were combined with model rebuilding into prime-and-switch composite omit maps calculated with RESOLVE (Terwilliger and Berendzen, 1999b) (Figure 21B,C). This strategy finally allowed the building of two full-length trigger factor molecules into the electron density. The structure was refined to 2.7 Å resolution with R/R\text{free} = 0.241/0.324 and with an average B-value of 74 Å² (Table 9).

<table>
<thead>
<tr>
<th>Refinement</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r.m.s.d. bonds (Å)</td>
<td>0.007</td>
</tr>
<tr>
<td>r.m.s.d. angles (°)</td>
<td>1.31</td>
</tr>
<tr>
<td>Luzzati error * (Å)</td>
<td>0.43</td>
</tr>
<tr>
<td>Monomers / ASU</td>
<td>2</td>
</tr>
<tr>
<td>Number of protein residues</td>
<td>864</td>
</tr>
<tr>
<td>Number of protein atoms</td>
<td>6787</td>
</tr>
<tr>
<td>Number of water molecules</td>
<td>89</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>56</td>
</tr>
<tr>
<td>\langle B_{total} \rangle (Å²)</td>
<td>74</td>
</tr>
</tbody>
</table>

* Luzzati coordinate error (5.0 – 2.7 Å)

**Table 9: Refinement statistics of the trigger factor structure**

The model was built using the program O (Jones et al., 1991) and refined with the program CNS (Brunger et al., 1998). Maximum likelihood refinement with amplitudes and experimental phase probability distribution was applied during model building. The structure was refined to 2.7 Å resolution with R/R\text{free} = 0.241/0.324 and with an average B-value of 74 Å².
Figure 21: Electron density maps used for building the structure of the trigger factor
Selenium sites are marked in red as anomalous difference Fourier maps contoured at 4 $\sigma$. A. Experimental electron density contoured at 1.4 $\sigma$ showing regions of the N- and C-terminal domains of the trigger factor. B. Experimental electron density map within a poorly defined region of the PPIase domain. The main chain and selenomethionine sidechains are shown as stick representations. C. The same region showing a prime-and-switch composite omit map contoured at 1.4 $\sigma$. 

65
Chapter 3

Results

The atomic coordinates of the structure were deposited to the protein database (PDB) with the accession code 1W26. The geometry of the protein structure was analysed with PROCHECK (Laskowski et al., 1993). The majority of the residues lay within the most favoured or acceptable values of the Ramachandran plot with only 12 residues (1.6%) being in the generously allowed region. (Morris et al., 1992) (Figure 22).

![Ramachandran Plot of the two trigger factor molecules in the asymmetric unit](image)

All protein residues are located within the coloured areas marking allowed values for $\phi$ and $\psi$ angles. The plot was calculated with PROCHECK.

### 3.3.4 The trigger factor fold

Trigger factor folds into an unusual shape resembling a crouching dragon with overall dimensions of 122x63x59 Å³, much larger than would be expected from a tightly packed 48 kDa protein (Figure 23). The N-terminal ribosome-binding domain (RB) is located at the “tail” (red in Figure 23). Curiously, the second domain in the linear sequence harbouring the peptidyl-prolyl isomerase activity (PPIase) is located on the opposite site of the molecule and forms the “head” (yellow) of the trigger factor. It is connected to the N-terminal domain via a long linker extending along the “back” of the trigger factor. The mostly $\alpha$ helical C-terminal domain finds itself in between the N-terminal and the PPIase domain and besides contributing to the “back” of the trigger factor forms two extended “arms” (“arm” 1 green, “arm” 2 blue).
Figure 23: Trigger factor folds into a unique dragon shape

Top: Ribbon diagram of the trigger factor fold. The loop within the N-terminal domain known to be responsible for ribosome binding is marked with the corresponding amino acid sequence. Bottom: Domain arrangement in a linear sequence representation. The positions of the ribosome-binding TF signature (residues 43-50) and domain borders are indicated. In all parts the N-terminal ribosome binding “tail” is shown in red, the PPIase “head” in yellow and “arm” 1 and “arm” 2 in green and blue, respectively. Surprisingly, the tertiary structure does not represent the linear domain order. The overall shape of the fold resembles that of a crouching dragon (insert).

Secondary structure elements are numbered according to their appearance in the primary sequence as indicated in Figure 24B. The N-terminal “tail” domain of the trigger factor is built up of a four-stranded antiparallel β sheet (b1,2,4,3) and two long, protruding α helices (a1,3) linked by an extended loop containing the short α helix a2. The loop region between helices a1 and a2 contains the TF signature motif, which is essential for ribosome binding (Kramer et al., 2002b). The linker between the N-terminal and the PPIase domain starts with strand b5 forming a parallel sheet with b13 of the C-terminal domain followed by a short α helix (a4) and a two-stranded twisted antiparallel β sheet (b6,12). The PPIase domain is only
connected to the remaining protein via the 15 \AA long linker region around strands b6,12 without any further interdomain contacts and displays a slight domain movement when two molecules in the crystal are compared (Figure 27). It adopts a typical FKBP-binding protein fold (Vogtherr et al., 2002) and is not essential for trigger factor activity \textit{in vivo} (Genevaux et al., 2004; Kramer et al., 2004a). The C-terminal domain extends from the PPlase domain backward toward the N-terminal domain through the long \(\alpha\) helix a6 in the back of the molecule. From this helix a short linker leads into “arm” 1 (green in Figure 24) which is built up of helices a7 and a8 with a connecting loop at the tip of the “arm”. “Arm” 1 is directly connected to “arm” 2, composed of helices a9 and a11 with the short connecting helix a10 at the tip.

![Figure 24: The domain organisation of the trigger factor fold](image)

**A.** Ribbon diagram of the trigger factor fold. **B.** Schematic representation of the domain organisation.

A large cradle, lined by \(\alpha\) helices, is formed between the N-terminal ribosome-binding “tail” and the C-terminal “arms” of the molecule (Figure 25). This region is characterised by an accumulation of hydrophobic residues and is extensively involved in forming crystal contacts as can be seen by analysing the packing of trigger factor molecules in the crystal lattice (Figure 25B).
Chapter 3

Results

Figure 25: Trigger factor forms a large hydrophobic cradle
A. Solvent accessible surface of the trigger factor, coloured by electrostatic potential (blue = positive, red = negative). The approximate dimensions of the cradle are indicated. B. Neighbouring trigger factor molecules in the crystal lattice contact the C-terminal arms and either the region behind the “arms” leading to the PPIase domain or the region between the “arms” and the N-terminal domain. These contact sites are formed by both hydrophobic and charged amino acids. Positions of residues involved in crystal contacts are shown as light and dark spheres for the contacts with two distinct symmetry mates of the molecule.

3.3.5 Crystal packing

Trigger factor forms an extensive network of intermolecular contacts through intertwining of the “arms” and the inner portion of the ribosome-binding domain. (Figure 26).
Figure 26: Crystal contact sites of the full-length trigger factor

A network of trigger factor molecules is formed within the crystal lattice. The tight packing of the trigger factor is shown approximately along the 2-fold screw axis. Each trigger factor molecule is coloured differently.

The crystal contact sites cluster on the inner surface of the N-terminal “tail”, in between the “arms” of the C-terminal domain and on the cradle portion behind the “arms” of the trigger factor leading to the PPIase domain (Figure 25B). These interactions involve both hydrophobic contacts and salt links. A very large surface area of 3500 Å² is buried upon formation of these contacts, which is more extensive than usually observed in interfaces of oligomeric proteins.

Although there are two molecules in the asymmetric unit of the crystal, no isolated dimer can be distinguished among the network of trigger factor molecules (Figure 26). Therefore, it is not clear if some of these contacts formed in the crystal structure may represent the contacts that occur upon dimerisation of the trigger factor in in vitro experiments (Patzelt et al., 2002).

Superimposition of the two molecules in the crystallographic asymmetric unit shows no difference in the overall fold (Figure 27). Structural flexibility is limited to extended loop regions. The ribosome-binding loop adopts different conformations and probably becomes
fully ordered only upon ribosome binding. Slight domain movements are observed for the PPIase domain and “arm” 2 of the C-terminal domain.

Figure 27: Superimposition of the two full-length TF molecules in the crystallographic asymmetric unit
Molecule A is shown in blue, molecule B in red. Apart from local flexibility in extending loop regions the two molecules show no differences in their overall fold.

Structures of a C-terminally truncated trigger factor from *Vibrio cholerae* (Ludlam et al., 2004), in which the second “arm” is completely disordered, and of the isolated ribosome-binding (Kristensen and Gajhede, 2003) and PPIase domains (Vogtherr et al., 2002) closely resemble the corresponding portions of the full-length trigger factor protein.

### 3.3.6 Structural homologues

The trigger factor molecule shows a modular architecture being composed of three individual domains. All three domains are structurally homologous to domains found in other proteins implicated in assisting protein folding (Figure 28).
The trigger factor fold shows a modular architecture

The trigger factor fold is compared to the known structures of Hsp33 and SurA and to the conserved FKBP fold to demonstrate its modular architecture. The corresponding parts of the structure are coloured accordingly. The loop within the N-terminal domain known to be responsible for ribosome binding is marked with the corresponding amino acid sequence.

The ribosome-binding domain is similar to parts of the ATP-independent chaperone holdase Hsp33 (Kristensen and Gajhede, 2003), which is activated during severe oxidative stress. During its activation process the reduced, monomeric Hsp33 is converted to a fully oxidised dimer with two potential polypeptide-binding sites lined by hydrophobic residues. One of the binding sites is formed by an intersubunit 10-stranded β sheet "saddle" with a largely uncharged or hydrophobic surface (Graumann et al., 2001; Janda et al., 2004; Vijayalakshmi et al., 2001).

The trigger factor PPIase domain adopts the typical FK506 binding protein (FKBP) fold (see introduction) (Vogtherr et al., 2002), but was shown to be unable to bind FK506 (Stoller et al., 1995).
The C-terminal domain shares no sequence homology to any other protein, but surprisingly, upon submission to the DALI program (Holm and Sander, 1993), the periplasmic chaperone SurA turned out to be a structural homologue. SurA is a chaperone that participates in the folding of outer membrane proteins in gram-negative bacteria (Bitto and McKay, 2002; Rouviere and Gross, 1996). Trigger factor and SurA share several features: 1) Both the trigger factor and SurA exhibit a modular architecture composed of substrate binding domains and distinct catalytically active PPIase domains. 2) Variants of both proteins lacking their PPIase domains still have chaperone activity (Behrens et al., 2001; Kramer et al., 2004b) and, therefore, the C-terminal domain of the trigger factor and the structurally homologous region of SurA are involved in chaperone function. 3) The regions implicated in chaperone activity are extensively involved in the formation of crystal contacts (Bitto and McKay, 2002).

In contrast to the trigger factor with its broad substrate specificity for polypeptides enriched in basic and aromatic residues (Patzelt et al., 2001), SurA recognises a peptide motif that is characteristic for integral outer membrane proteins (Bitto and McKay, 2003).

### 3.4 Cloning, expression and purification of trigger factor constructs

In order to better understand the mechanism of the trigger factor as a ribosome associated chaperone we have decided to study the structure of a complex between the trigger factor and the large ribosomal subunit. Several alternative experimental strategies involving co-crystallisation and soaking of the full-length trigger factor or of its fragments into the pre-existing crystals of the large ribosomal subunit were pursued in parallel.

#### 3.4.1 Expression and purification of trigger factor fragments

The DNA sequence coding for the N-terminal 111 amino acids of the trigger factor (12.2 kDa fragment TF111) was amplified by PCR and cloned into the plasmid pRET31H6TRX (kindly provided by Imre Berger, unpublished) using its KpnI / BamHI restriction sites (Appendix 7.3.2). The correctness of the obtained construct was verified by ABI-sequencing. The DNA sequence coding for the N-terminal 144 amino acids of the trigger factor cloned into plasmid pDS56/RBSII was used as the second expression construct and was obtained from the groups of Prof. Bernd Bukau and Dr. Elke Deuerling (ZMBH Heidelberg).

All constructs were expressed in BL21 pREP4 or DH5α cells as described (Materials and Methods). Trigger factor fragments were purified by affinity chromatography followed by
anion exchange and size exclusion chromatography. As all proteins were prepared for co-crystallisation or soaking experiments with *H. marismortui* 50S the buffer was exchanged to the high-salt final buffer for 50S crystallisation during the size exclusion chromatography. In each step of the procedure products were analysed with SDS-PAGE. In the case of the TF111 construct the N-terminal His$_6$-tag was cleaved with TEV protease after the affinity column and further purified (Figure 29). The final products were tested for integrity by N-terminal sequencing or by immunoblotting using Penta-His antibodies (Material and Methods).

**Figure 29: Purification of trigger factor fragments**

All expression constructs contain His$_6$-tags and were purified by affinity chromatography. Final products were obtained after anion exchange and size exclusion chromatography. The size exclusion chromatography was done in the final buffer used for 50S crystallisation. A. Size exclusion chromatography of TF111 after TEV protease cleavage. TF111 before (2) and after (1a/b) protease cleavage was analysed with SDS-PAGE. Upon cleavage, the molecular weight of 26.3 kDa is reduced to 12.3 kDa (Appendix 7.3.2). B. Size exclusion chromatography of TF144. Again, purity was monitored with SDS-PAGE analysis.

### 3.5 Co-crystallisation and soaking of *H. marismortui* 50S with trigger factor fragments

Binding of the trigger factor to the large ribosomal subunit from *E. coli* was shown *in vitro* using assays in which ribosomes and any associated factors are pelleted through a sucrose cushion while soluble proteins that are not in complex with ribosomes remain in the supernatant (Kramer et al., 2002b). Using such binding assays trigger factor residues within the TF-signature motif were identified as crucial for this interaction. Furthermore, cross-linking experiments demonstrated that the trigger factor binds to the ribosome in close proximity of ribosomal proteins L23 and L29 (Kramer et al., 2002b). We anticipated that the crystal structure of the complex between the ribosome-associated chaperone and the large
ribosomal subunit would provide important structural insights into co-translational protein folding. Since the large ribosomal subunit from *H. marismortui* is more suitable for crystallisation experiments it was decided to first investigate if it is possible to form a specific and functional complex between the bacterial chaperone and the archaeal ribosome. Binding tests and crosslinking experiments, similar to those described for *E. coli* ribosomes, demonstrated that the binding site on the archaeal ribosome is conserved and that a specific complex can be formed.

### 3.5.1 Binding of trigger factor to *H. marismortui* 50S

Ribosomes were purified as described (Material and Methods). The obtained 50S sample was used in binding tests as for *E. coli* ribosomes (E.D. and B.B.) (Figure 30). Briefly, purified *H. marismortui* 50S was incubated either with the trigger factor wild type (TF WT) or with the trigger factor mutant deficient in ribosome binding (TF-AAA). After re-isolation of the *H. marismortui* 50S chaperone complex the sample was analysed by SDS-PAGE and Coomassie staining. Except TF-AAA all trigger factor variants associated with the *H. marismortui* 50S subunit and were detectable in the ribosomal pellet. Since TF-AAA does not interact with the ribosome it is reasonable to assume that the other two trigger factor constructs are forming specific interactions. In addition, the V49C mutant of the trigger factor, labelled with BPIA, was shown to crosslink to ribosomal protein L23 when bound to archaeal ribosomes. Therefore, the bacterial trigger factor specifically binds the *H. marismortui* 50S subunit and can be crosslinked to the ribosomal protein L23, identically as observed for the homologous complex (Kramer et al., 2002b).
Chapter 3

Results

Figure 30: Trigger factor binding to *H. marismortui* 50S

In the experiment all trigger factor variants were added in a physiological twofold molar excess over ribosomes and therefore are always detectable in the supernatant (S). Trigger factor wild type (TF WT), a trigger factor mutant deficient in ribosome binding (TF-AAA) and the trigger factor mutant TF-V49C were used for the binding assay. The purity of the protein was analysed with SDS-PAGE as is shown for TF WT (I). *H. marismortui* ribosomes were dissociated into their subunits to be separated by sucrose gradient centrifugation (II). 50S was incubated with the different trigger factor versions and chaperones associated with the ribosome were separated from the unbound pool and analysed with SDS-PAGE. Except TF-AAA all trigger factor variants associated with the *H. marismortui* 50S subunit and were detectable in the ribosomal pellet (P). Upon UV irradiation TF-V49C labelled with BPIA crosslinks to ribosomal protein L23 (lane 9) as identified by mass spectrometry.

3.5.2 Crystallisation of the trigger factor:50S complex

Attempts to obtain crystals of the TF-ribosome complex involved both soaking of the small TF111 fragment into the pre-existing *H. marismortui* 50S crystals and co-crystallisation experiments. The published *H. marismortui* 50S crystallisation procedure (Ban et al., 2000) was slightly modified, so that the equilibration during crystal growth takes place against a higher salt concentration of 1.7 M KCl. During co-crystallisation experiments the large ribosomal subunit was mixed with 10 to 12.5x molar excess of the full-length trigger factor or of its fragments to compensate for the low binding affinity of the chaperone, 1.2 µM for the *E. coli* complex (Maier et al., 2003; Patzelt et al., 2002). Unfortunately, the large amount of the trigger factor protein added to the sample severely interfered with the crystal growth. After 1 to 2 days first plate-like crystals appeared but were usually very thin and of irregular shape. Only few crystals, obtained in the presence of the TF144 fragment, showed sharp edges and finished growing to their maximal size of 500 µm x 500 µm x 100 µm after 3 weeks. After stabilisation, crystals were kept in a cryoprotectant solution containing 10 to 12.5x molar excess of the corresponding ligand for at least 24 hours. For soaking experiments already stabilised 50S crystals were transferred to cryoprotectant solutions containing 10 to 12.5x
molar excess of TF111. Crystals were kept in the cryoprotectant solution containing the ligand for at least 48 hours before being flash frozen in liquid propane.

### 3.5.3 Data collection

Crystals were mounted on a bent loop (Materials and Methods) and tested for diffraction either at the Swiss-Norwegian Beamline (SNBL at the ESRF, Grenoble) or at the X06SA PX beamline (Paul Scherrer Institute, Villigen, Switzerland). At the SNBL a diffraction pattern of 50S co-crystallised with TF144 was recorded with a mar345 detector at a wavelength of $\lambda = 0.9$ Å (Figure 31). The read-out diameter of the image plate was 300 mm with a pixel size of 100 µm. The average exposure time was ~1 min per frame. The exact exposure time was dependent on the actual beam intensity (‘dose mode’). At the SLS a diffraction pattern of a 50S crystal which was soaked with TF111 was recorded with a mar345 detector at a wavelength of $\lambda = 1.0$ Å. The read-out diameter of the image plate was 345 mm with a pixel size of 100 µm. The exposure time was 3.5 s.

![Diffraction pattern](image.png)

**Figure 31: Diffraction pattern of 50S co-crystallised with trigger factor fragments**

50S was co-crystallised with TF144. Plate-like crystals diffracted to 3.5 Å resolution. Reflections within the resolution range that was affected by the ice powder diffraction were selectively removed from the data after scaling.
3.5.4 Data processing and scaling

The data was processed with the HKL package (Otwinowski Z., 1997). All data sets were collected in 1° steps with a total of 158 degrees for TF144 and 130 degrees for TF111 to obtain redundant data (Table 10).

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>TF144</th>
<th>TF111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Parameters (Å)</td>
<td>$a = 211, b = 299, c = 574$</td>
<td>$a = 215, b = 302, c = 579$</td>
</tr>
<tr>
<td>Space group</td>
<td>C2221</td>
<td>C2221</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Mosaicity (°)</td>
<td>0.79</td>
<td>0.50 - 0.60</td>
</tr>
<tr>
<td>Total reflections</td>
<td>4572477</td>
<td>3595459</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>241873</td>
<td>227772</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50 – 3.5</td>
<td>50 – 3.9</td>
</tr>
<tr>
<td>$R_{	ext{sym}}$ (%)</td>
<td>6.8 (69.1)</td>
<td>15.6 (63.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.8 (94.2)</td>
<td>99.5 (99.7)</td>
</tr>
</tbody>
</table>

Table 10: Data collection statistics of 50S co-crystallised with TF144 or of 50S crystals soaked with TF111

3.5.5 Model refinement

Two datasets of the large ribosomal subunit co-crystallised (TF144) or soaked (TF111) with the trigger factor were collected. For both datasets it was possible to use the published structure of the large ribosomal subunit from *H. marismortui* (Protein data bank entry 1JJ2) for phasing without molecular replacement searches since the original space group and cell dimensions of the native crystals were preserved. Refinement against the measured amplitudes was performed using a maximum likelihood target in CNS (Grosse-Kunstleve and Adams, 2001). The refined model was obtained by 30 steps of rigid body refinement followed by 100 steps of conjugate gradient minimisation refinement (Table 11).

<table>
<thead>
<tr>
<th></th>
<th>TF144</th>
<th>TF111</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$, $R_{free}$</td>
<td>0.312, 0.284</td>
<td>0.318, 0.313</td>
</tr>
<tr>
<td>r.m.s.d. bonds (Å)</td>
<td>0.008</td>
<td>0.007</td>
</tr>
<tr>
<td>r.m.s.d. angles (°)</td>
<td>1.318</td>
<td>1.180</td>
</tr>
<tr>
<td>Monomers / ASU</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>55</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 11: Refinement statistics of *H. marismortui* 50S crystals
Crystals were either obtained by co-crystallisation with TF144 or by soaking large ribosomal subunit crystals with TF111.
3.5.6 Structure determination of the N-terminal fragment of trigger factor in complex with \textit{H. marismortui} 50S

For both the co-crystallised and soaked TF-ribosome complex datasets bound chaperone fragment was initially visualised in difference Fourier maps using the observed diffraction amplitudes and structure factor amplitudes from native data (Fo-Fo) (Figure 32). Since the TF144 dataset extended to higher resolution it was used for all subsequent calculations (Table 10).

Visual inspection of the difference electron density revealed that the part of the trigger factor in complex with the ribosome involves the ribosome-binding loop of the trigger factor containing the TF “signature motif” (\textit{GFRxGxxP\textsubscript{50}}) and the binding helices \(\alpha1\) (aa 20-39) and \(\alpha2\) (aa 50-59). From its anchor point on the ribosome, the N-terminal domain is visible up to a distance of 20 Å above the surface of the 50S subunit before the density fades out due to increasing temperature factors. Although flexible in the uncomplexed form, it was possible to obtain an unambiguous rigid body fit of one of the previously observed conformations of the
ribosome-binding loop of the trigger factor (Protein Data Bank: 1OMS_B) (Kristensen and Gajhede, 2003) into the difference electron density. A kink in one of the helices and bulky residues (Phe 45 and Arg 46) were used as structural points of reference (Figure 33). The final structure of the ribosome-binding domain of the trigger factor was built into improved maps calculated as prime-and-switch composite omit maps around the attachment region (Figure 33A) using the program RESOLVE (Terwilliger and Berendzen, 1999a). Starting from the structure of the 50S subunit (Protein Data Bank: 1JJ2) and the fitted ribosome-binding loop of the trigger factor (aa 25-59) rigid body fitting followed by positional and temperature factor optimisation in CNS (Brunger et al., 1998) yielded an overall model of the 50S:trigger factor complex with an R-factor of 0.195 and R\text{free} of 0.262 (Table 12). The structure of the attachment region was further rebuilt into 2Fo-Fc difference Fourier maps to take into account small local conformational changes in the area of the chaperone attachment, in particular, ribosomal proteins L23, L29 and domain III of the 23S rRNA. The atomic coordinates of the structure were deposited to the protein database (PDB) with the accession code 1W2B.

<table>
<thead>
<tr>
<th>Refinement</th>
<th>R, R\text{free}</th>
<th>0.195, 0.262</th>
</tr>
</thead>
<tbody>
<tr>
<td>r.m.s.d. bonds (Å)</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>r.m.s.d. angles (°)</td>
<td>1.236</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Refinement statistics of the structure of the N-terminal trigger factor domain in complex with 50S

The completely refined structure was obtained with rigid body fitting followed by positional and temperature factor optimisation in CNS (Brunger et al., 1998).

3.5.7 Loop mediated interaction of trigger factor on the ribosome

The trigger factor binds the ribosome at the triple junction between rRNA and proteins L23 and L29, proximal to the exit of the ribosomal nascent chain tunnel (Figure 33B). Overall approximately 1227 Å² surface area is buried upon complex formation, which is divided into 423 Å², 622 Å² and 182 Å² for the contacts of the trigger factor with 23S rRNA, L23 and L29, respectively. Since the proteins at the interface form a ridge, only residues at the tip of the ribosome-binding loop of the trigger factor (Arg45, Gly43 and possibly Arg40) extend far enough to reach the rRNA binding sites. In contrast to the mostly polar interactions with the rRNA, the interaction with L23 is hydrophobic and primarily involves three residues of the trigger factor (Phe44, Pro50 and Ile53) surrounding residue Met16 on L23 (Figure 33C). A second interaction site involves hydrophobic contacts between the C-terminus of L23, which was not visible in the 50S structure alone, and the trigger factor helix a3 (Figure 33C).
Interestingly, important interactions are mediated by Glu13 on L23, which was shown to be essential for trigger factor binding to 50S (Kramer et al., 2002b). This residue has an altered sidechain conformation in this structure and positions, with two hydrogen bonds, Arg45 of the trigger factor to form a hydrophobic $\pi-\pi$ stacking with the unpaired RNA base of A1501 in the H. marismortui ribosome (A1392 in the E. coli ribosome) and orients its guanidine group for salt interactions with the phosphate of A1501 (Figure 33C). Contacts with ribosomal proteins L29 and L19 are tangential and are unlikely to contribute to the trigger factor binding affinity in agreement with biochemical data (Kramer et al., 2002b). Furthermore, ribosomal protein L19 does not have a prokaryotic homologue and therefore would not be able to contribute to the contacts in case of the homologous trigger factor ribosome complex.

![Figure 33: Structure of the N-terminal trigger factor domain in complex with 50S](image)

**A.** Trigger factor fragment TF144 bound to the 50S ribosomal subunit: Simulated annealed omit map around residues 26-59 of the bound trigger factor fragment shown together with a surface representation of 50S; key interacting proteins are labelled. For clarity only selected side chains are shown. **B.** View from a distance; trigger factor is shown as a red surface representation and the peptide exit tunnel is denoted with an asterisk. In **A.** and **B.** rRNA is coloured bisque; non-interacting proteins grey, L29 cyan, L23 green and L19 palegreen. **C.** Key interactions between the trigger factor (TF) fragment (red), L23 (green) and 23S rRNA. Selected residues are shown together with C$_{\alpha}$-traces. Hydrogen bonds are indicated only for the interaction between key contact residues E13 (L23), R45 (TF) and A1501 (23S rRNA).

### 3.5.8 Sequence conservation of the trigger factor docking site

Although not present in archaebacteria and eukaryotes the trigger factor can specifically bind to the large ribosomal subunit from H. marismortui (as described before) and can even
function in yeast (Ito, 2005; Rauch et al., 2005). Sequence alignment of 24 archaeal and 160 bacterial L23 proteins was performed around the trigger factor binding region. In addition, the sequence and secondary structure of the *H. marismortui* and *E. coli* ribosomal RNA around A1501 were compared. Both the sequence of L23 and the ribosomal RNA turned out to be highly conserved between bacteria and archaeabacteria (Figure 34). Obviously, due to a small interaction area with the ribosome that generally has a high level of conservation, the trigger factor is able to bind ribosomes from all three kingdoms of life.

![Sequence conservation of L23 and ribosomal RNA at the TF binding site](image)

**Figure 34: Sequence conservation of L23 and ribosomal RNA at the TF binding site**

A. Sequence-logo representation of aligned sequences of 24 archaeal and 160 bacterial L23 proteins. The *H. marismortui* and *E. coli* L23 sequences are given in the inset. B. *H. marismortui* and *E. coli* ribosomal RNA sequences around the trigger factor interacting A1501 in secondary structure representation. In A. and B. asterisks indicate residues in contact with the trigger factor.

### 3.5.9 Full-length trigger factor bound to the ribosome

The structure of the full-length trigger factor together with the structure of its N-terminal ribosome-binding region in complex with the 50S permits accurate placement of the entire chaperone on the ribosome through superposition of the ribosome-binding region visible in both structures. Remarkably, the trigger factor hunches over the polypeptide exit of the ribosome and extends its hydrophobic inner face of the C-terminal “arms” and the N-terminal “tail” towards the area where the nascent polypeptides appear. Judging from the increasing mean atom displacement (temperature factors) of the helices belonging to the ribosome-binding “tail”, the full-length trigger factor bound to the ribosome could swing approximately 10° in all directions around its attachment point. No other region of the chaperone, except for the “tail”, contacts the ribosome (Figure 35). The closest approach from the surface of the ribosome to the tips of the C-terminal “arms” is less than 10 Å while the distance to the bottom of the cradle is 40 Å (Figure 35A).

The PPIase domain is tethered to the main body of the trigger factor via a double linker and is located peripherally to the region forming the main substrate-binding cradle of the
trigger factor and to the exit of the ribosomal tunnel. Dictated by the length of the tether, and/or possible conformational changes, the PPIase activity will be limited to the region above the “arms” and behind the “back” of the trigger factor where partially folded proteins will appear only after they have escaped from underneath the cradle. The PPIase domain will therefore interact with nascent chains at later stages of their synthesis restricting its catalytic and potential chaperone activity to selected sites, which remain accessible at this folding stage. These findings agree with biochemical data demonstrating that trigger factor lacking the PPIase domain still binds peptides enriched in hydrophobic residues and shows wildtype function \textit{in vivo} (Kramer et al., 2004b).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{model.png}
\caption{Model of full-length trigger factor bound to the ribosome}
\textbf{A.} Overview of the trigger factor:50S complex. Full-length trigger factor positioned by superimposition onto the ribosome-bound trigger factor fragment TF144 is shown as Cα-trace together with a slice of 50S along the peptide exit tunnel (for clarity, further cavities peripheral to the tunnel are not shown) with a modelled nascent chain in magenta, extending from the peptidyl transferase centre (PT). Colouring as in Figure 23. Inset: Dimension of the cradle in Å. The exit of the peptide tunnel marks an asterisk. Close-up \textbf{B.} side and \textbf{C.} top views of the complex shown in \textbf{A.} without a nascent chain.
\end{figure}
4 Discussion

In this thesis structural studies of the bacterial trigger factor and of the bacterial trigger factor in complex with the large ribosomal subunit from *Haloarcula marismortui* were performed with the hope to provide an important contribution to our understanding of the function of ribosome-associated chaperones. The project started with crystallisation attempts with different trigger factor variants from *Thermus thermophilus* or *Escherichia coli* finally leading to the determination of the full-length *Escherichia coli* trigger factor structure at 2.7 Å resolution. The trigger factor structure was surprising not only because of its unusually extended fold but also because of the structural homology of its C-terminal domain to the chaperoning part of the periplasmatic protein SurA although the corresponding regions in both proteins share no sequence homology. The structure of the ribosome-binding domain of the trigger factor from *Escherichia coli* in complex with the large ribosomal subunit from *Haloarcula marismortui* solved at 3.5 Å resolution allowed the accurate placement of the entire chaperone on the ribosome through superposition of the ribosome-binding region visible in both structures. Unexpectedly, the trigger factor projects its extended domains over the exit of the ribosomal tunnel creating a protected folding space where nascent polypeptides are shielded from proteases and aggregation.

4.1 The trigger factor fold

The trigger factor adopts an unusual extended and open fold (Figure 36A). The structures of the N-terminal ribosome-binding domain as well as of the PPIase domain have been studied in isolation (Kristensen and Gajhede, 2003; Vogtherr et al., 2002). In the context of the full-length trigger factor the fold of individual domains is preserved. The C-terminal domain has never been studied in isolation as it is highly susceptible to proteolytic cleavage. Obviously, its overall fold is critically stabilised by the long linker connecting the N-terminal domain to the PPIase domain. Recently, the structure of a C-terminally truncated trigger factor from *Vibrio cholerae* has been solved, in which the second “arm” is completely disordered. The fold of its N-terminal and PPIase domains are preserved. Nevertheless, the missing “arm” causes a collapse of the C-terminal domain, so that the relative positioning of the N-terminal and the PPIase domains changes the overall shape of the molecule (Figure 36A). The feature of the extended, open fold with the long linker region connecting the RB and PPIase domains is lost in the truncated trigger factor structure as the linker is no longer stabilised by the C-terminal end of the trigger factor molecule (Figure 36B, top). This causes
a movement of the first “arm” closer to the N-terminal domain (Figure 36B, bottom). Consequently, the fold of the first “arm” of the C-terminal domain is preserved in the *Vibrio cholerae* trigger factor structure, but not the unique cradle, which is observed in the *E.coli* trigger factor and SurA (see chapter 3.3.6) structures (Figure 36A,C).

**Figure 36: The cradle fold, trigger factor and SurA**
The trigger factor RB domain and the C-terminal SurA extension are coloured red, “arm” 1 and “arm” 2 of the cradle fold are coloured blue and cyan, respectively. The PPIase domain is shown in yellow. **A.** Overview of the trigger factor fold: Top, full-length *E.coli* trigger factor. Bottom: Truncated trigger factor from *Vibrio cholerae*. In the case of the truncated trigger factor variant the feature of the extended fold is lost. **B.** Superimposition of the structures obtained with the truncated and full-length trigger factor protein. The missing C-terminal end results in a less extended linker region between the RB and PPIase domain (top) and causes a movement of the first “arm” closer to the N-terminal domain (bottom). **C.** The cradle fold found in the *E. coli* trigger factor and SurA protein structures. In the case of the truncated *Vibrio cholerae* trigger factor the fold of the first arm is preserved.
The superposition of the two full-length \textit{E. coli} trigger factor molecules in the crystal shows no evidence for possible domain rearrangements (Figure 27). Flexible parts of the molecule are exclusively localised within peripheral loops. Therefore, in spite of the modular design, the extended structure of the full-length trigger factor seems to be relatively rigid.

All three trigger factor domains have structural homologies to domains found in other proteins (as described in section 3.3.6). The functional relevance of the RB domain being homologous to parts of Hsp33 remains to be elucidated as its substrate-binding site is not yet characterised. The periplasmatic chaperone SurA and the trigger factor resemble each other both regarding their overall architecture and their preference for binding sequences enriched in aromatic residues (as described in section 3.3.6). For both chaperones the binding to the unfolded polypeptides is independent of whether they contain proline residues. But their analogy extends further. SurA stabilises secreted OMPs during their passage through the periplasm and, just like the trigger factor, acts in an ATP-independent manner. Interestingly, SurA null strains are viable, but characterised by outer membrane perturbations, while double knock-outs of SurA and either of the periplasmic chaperones DegP or Skp exhibit synthetic lethality (Krojer et al., 2002; Rizzitello et al., 2001). These chaperones might act at different stages of OMP maturation but with partially redundant functionality (Rizzitello et al., 2001), potentially representing a periplasmatic analogy to the cytosolic trigger factor, DnaK, GroEL system.

\section*{4.2 Trigger factor acting at the ribosome}

The role of the trigger factor in the folding of newly synthesised proteins critically depends on its association with the ribosome (Kramer et al., 2002b). It was shown that in the cytosol trigger factor is in a monomer-dimer equilibrium, but at the ribosome it is monomeric (Patzelt et al., 2002). The \textit{in vivo} function of the trigger factor dimer remains elusive and the presented crystallographic data shows no evidence of a perfect dimer formation. Presumably, in the dimeric form the substrate-binding cradle is hidden and only becomes accessible when bound to the ribosome in the monomeric form. Therefore, dimerisation as observed in solution may help to prevent unwanted substrate interactions of the trigger factor in the cytosol (Patzelt et al., 2002).

The overall architecture of the trigger factor offers a perfect protective environment for the newborn protein with its hydrophobic residues still exposed in its unfolded, nonnative state. The slight domain motions within loop regions of the trigger factor, shown by
superimpositions of the two molecules in the crystal lattice, might be important for the binding of trigger factor’s various nascent peptide substrates.

The model of the full-length trigger factor bound to the ribosome is supported by several pieces of evidence.

First, the interaction with the ribosome in the presented structure is characterised by a delicate loop mediated interaction of the RB domain with L23 and 23S rRNA on the ribosome: (i) Only very few residues are involved in the interaction including Glu13 on L23 and Arg45 belonging to the chaperone’s signature motif. Ribosomal proteins L29 and L19 are close but do not contribute to the interaction. No other part of the trigger factor apart from the RB domain contacts the ribosome. These results agree very well with biochemical results. L23 was shown to be the essential docking site on the ribosome (Kramer et al., 2002b). If the Glu13 residue on L23 and its two neighbouring residues are mutated to alanine the \textit{E.coli} ribosome showed severely decreased binding of the trigger factor. The \textit{E. coli} trigger factor mutant TF-AAA (44FRK46 \rightarrow 44AAA46) is impaired in ribosome binding (Kramer et al., 2002b). The RB domain alone is sufficient for ribosome binding (Kramer et al., 2004b). (ii) The recently solved structure of the homologous complex of the N-terminal RB domain of the trigger factor from \textit{Deinococcus radiodurans} with the large ribosomal subunit (Baram et al., 2005) shows identical structural features as we observed for the heterologous complex: The binding of the trigger factor is characterised by delicate loop mediated interactions of the TF signature motif with L23 and 23S rRNA on the ribosome. The trigger factor hunches over the polypeptide exit of the ribosome to perfectly position its hydrophobic inner face for the interaction with the emerging nascent chain. (iii) Trigger factor has inherent, salt sensitive affinity (K_D=1.2 µM, (Maier et al., 2003; Patzelt et al., 2002)) for ribosomes, which is increased when they display nascent peptides (Hesterkamp et al., 1996). The fact that the interactions between the signature motif of the trigger factor and the ribosome involve hydrogen bonds and salt links explains why this interaction is sensitive to increasing salt concentrations. By maintaining a precise balance of binding affinities, the trigger factor dissociates from vacant ribosomes within seconds (Maier et al., 2003), but may become stabilised by hydrophobic interactions with an emerging, unfolded nascent peptide. The broad distribution of hydrophobic residues within the trigger factor cradle may allow extensive but transient contacts of the trigger factor with the emerging chain which is consistent with the observation that the association of the trigger factor with the translating ribosomes becomes salt insensitive (Maier et al., 2003).
Second, the major substrate-binding site of the trigger factor is located within its N-terminal and C-terminal domain. (i) There is significant hydrophobicity over the entire inner surface of the cradle providing a perfect environment for the interaction with its substrates. (ii) The cradle is involved in forming an extensive network of crystallographic intermolecular contacts in the crystal lattice, which might resemble the way the trigger factor interacts with its substrates. (iii) The structurally similar domain of the SurA chaperone has substrate-binding properties. It has been demonstrated that a 150 amino acid long N-terminal fragment of SurA, which does not contain the two PPIase domains of the native protein, interacts with model peptides as observed for the wildtype protein (Webb et al., 2001). (iv) The fused N-terminal and C-terminal domains of the trigger factor, lacking the PPIase domain, exhibit almost wildtype-like chaperone activity in vivo and in vitro as well as peptide binding activity (Genevaux et al., 2004; Kramer et al., 2004b).

The role of the PPIase domains both in the trigger factor and SurA remains elusive. In agreement with these findings, the PPIase domain is tethered to the main body of the trigger factor via a double linker peripherally to the main substrate-binding cradle of the trigger factor and to the exit of the ribosomal tunnel. As a consequence, the PPIase activity will be limited to the region above the “arms” and behind the “back” of the trigger factor where partially folded proteins will appear only after they have escaped from the trigger factor cradle. The PPIase domain will therefore interact with nascent chains at later stages of their synthesis, restricting its catalytic activity to selected sites, which remain accessible at this folding stage. The PPIase domain in isolation was shown to have a 500 times reduced rate of its catalytic activity to refold RNAseT1 compared to its activity when fused to the N- and C-terminal domains of the trigger factor (Zarnt et al., 1997). Obviously, the N- and the C-terminal domains of the trigger factor are able to present substrates that require peptidyl-prolyl isomerisation for optimal interaction with the active site of the PPIase domain for the isomerisation reaction (Bang et al., 2000).

So far it was assumed that the 48 kDa trigger factor can only interact with small stretches of unfolded peptides. With the structure of the trigger factor at the ribosome it became evident that the small chaperone is able to provide a relatively large shielded folding space for the newborn protein. The folding space formed between the trigger factor and the ribosome at the exit site is sufficiently large to accommodate an average size protein domain demonstrated by the fitting of the 14 kDa large lysozyme between the chaperone and the ribosome (Figure 37). These results suggest an unexpected mode of trigger factor action at the ribosome.
4.3 Trigger factor assisted co-translational folding

4.3.1 Co-translational protein folding in prokaryotes

*In vivo* experiments demonstrated that co-translational protein folding can occur in the cytosol of both prokaryotic and eukaryotic cells (Nicola et al., 1999). However, the absolute proportion of co-translationally folding proteins remains unclear. In fact, in bacteria, but not in reticulocyte lysate, folding of several large model proteins occurs mainly post-translationally (Agashe et al., 2004; Netzer and Hartl, 1997), suggesting that the eukaryotic system has a higher capacity for efficient co-translational folding of multi-domain proteins than the bacterial one. Perhaps, the high speed of bacterial translation per se disfavours co-translational folding (Netzer and Hartl, 1997). Still, the preferred folding mode of a protein may depend on its particular folding kinetics, its domain-sizes and the condition in the cell. A plausible hypothesis assumes that the trigger factor may effectively coordinate the stepwise co-translational folding of multi-domain proteins and limit unproductive inter- and intramolecular interactions during the early folding steps.
4.3.2 A model for the mechanism of trigger factor function on the ribosome

The formation of a 1:1 complex between the trigger factor and ribosomes is required for the interaction of the trigger factor with nascent chains (Kramer et al., 2002b; Maier et al., 2003; Patzelt et al., 2002). In vivo, the trigger factor is present at a three-fold molar excess over ribosomes, so that 90% of the ribosomes are occupied by the chaperone. The remaining non-ribosome bound fraction of the trigger factor in the cytosol is in a monomer-dimer equilibrium (Patzelt et al., 2002). The trigger factor is able to interact both with vacant ribosomes and with ribosome nascent chain complexes as soon as the nascent peptides reach sufficient length to contact the cradle of the chaperone (Hesterkamp et al., 1996; Valent et al., 1997). The cradle is ideally suited to promote co-translational folding of domains by providing a shielded environment in which the folding is initially postponed through hydrophobic contacts (Figure 38, step 1). The trigger factor preferentially binds peptide stretches enriched in basic and aromatic residues (Patzelt et al., 2001) which would allow the chaperone to prevent unproductive intra-molecular interactions which could cause misfolding and aggregation of the newly synthesised protein. The binding of the trigger factor to unfolded proteins postpones their folding, as shown by in vitro refolding experiments (Maier et al., 2001), which is then allowed to proceed in a self-promoted fashion when sufficient sequence information becomes available for the generation of a folded core. Small proteins or domains might adopt their native fold within the trigger factor cradle resulting in the burial of hydrophobic residues and subsequent destabilisation and dissociation of the RNC-TF complex (Figure 38, step 2). In the case of multi-domain proteins trigger factor rebinding would be favoured upon appearance of newly synthesised unfolded portions of the nascent chain exposing hydrophobic residues (Figure 38, step 3). Protein domains larger than the size of the cradle must either escape from the cradle or promote dissociation of the trigger factor from the ribosome before productive folding. Interestingly, no stable association of the trigger factor with short, unfolded, nascent chains has been observed upon their induced release from the ribosome (Hesterkamp et al., 1996). Nevertheless, during the production of a multi-domain firefly luciferase, exchange of the trigger factor bound to the RNCs has been observed (Agashe et al., 2004), indicating that the ongoing translation of a large multi-domain protein indeed expels nascent chain-bound trigger factor from the ribosome.
Figure 38: Schematic representation of the model for the mechanism of trigger factor assisted cotranslational protein folding

The small ribosomal subunit is shown in bisque, the large ribosomal subunit in grey. The peptidyl transferase centre (PTC) is indicated in orange, L23 in green and the trigger factor in pink. The trigger factor provides a shielded environment for the emerging nascent chain. Its affinity for the ribosome-nascent chain complex (RNC) is enhanced by the presence of an emerging protein exposing hydrophobic residues (1). Once sufficient sequence information is available, the first domain of the growing polypeptide adopts its native fold (2). During the folding process, hydrophobic residues become buried inside the folded domain, consequently the trigger factor loses affinity for the RNC and dissociates from the ribosome. In the case of multi-domain proteins the trigger factor can rebind to assist the folding of the second domain (3).

4.4 Interplay with other factors involved in folding or targeting of newly synthesised proteins

4.4.1 Interplay of trigger factor and DnaK

Mediated by its co-chaperone DnaJ, DnaK binds in an ATP-dependent manner to hydrophobic stretches flanked by basic amino acid residues in unfolded proteins. Its binding specificity overlaps with that of the trigger factor (Patzelt et al., 2001; Rudiger et al., 1997; Rudiger et al., 2001) and they share a substrate pool of about 340 proteins in vivo (Deuerling et al., 2003). In contrast to the trigger factor, DnaK is not localised on the ribosome and mainly associates with newly synthesised proteins post-translationally, except for some longer ribosome-bound nascent chains, preferentially larger than 20 kDa (Hesterkamp et al., 1996; Kramer et al., 2002a; Teter et al., 1999). In total, only 10 - 15 % of all newly synthesised proteins interact transiently with DnaK in vivo. In trigger factor deficient strains this number increases to 40 % and DnaK uncharacteristically also interacts with shorter nascent chains (Deuerling et al., 1999; Teter et al., 1999). A competitive advantage of the trigger factor over DnaK in binding to short nascent chains has been demonstrated directly (Deuerling et al., 2003). Recently, it has been shown that the trigger factor and DnaK are required for efficient folding of a multi-domain model protein during in vitro translation and to cooperatively
decrease the folding speed for the studied proteins, thereby shifting the folding towards a post-translational mode. However, it remains unclear if these observations represent a general rule or are specific for the studied proteins (Agashe et al., 2004). Based on the model of trigger factor action, the main function of DnaK is seemingly to assist the folding of larger substrates, which cannot be accommodated in the trigger factor cradle.

4.4.2 Trigger factor and the secretory pathway

The trigger factor is the first protein to interact with almost all nascent chains since, at physiological conditions, it is bound to most ribosomes in the cell. Due to the spatial and temporal overlap of its action with the early events in protein targeting, the trigger factor apparently combines two opposing functions: On the one hand, it provides shielding for nascent polypeptides in the space defined by its cradle and the widening of the ribosomal tunnel. On the other hand, it has to provide access to the nascent chain, such that periplasmic and membrane-proteins are recognised early enough to be targeted to the translocon by the SRP- or Sec-dependent pathways in a translocation-competent unfolded state as outlined in Figure 2. The trigger factor has no involvement in post-translational Tat-mediated export beyond its general chaperone function (Jong et al., 2004).

4.4.2.1 L23 constitutes a general docking site on the ribosome

The crystal structure of the RB domain-ribosome complex identified, in agreement with cross-linking studies (Kramer et al., 2002b; Ullers et al., 2003), protein L23 as the main interaction site of the trigger factor. Remarkably, the same protein participates in contacts with ribosome-associated factors involved in protein export. The eukaryotic homologue of L23 mediates a prominent contact with SRP54 (Halic et al., 2004), the subunit of the SRP particle responsible for nascent chain binding (Figure 39B). Based on a 15 Å cryo-EM reconstruction, L23 also mediates a major interaction of the eukaryotic translocon, Sec61, with the ribosome (Figure 39C) (Halic et al., 2004). Such contacts have also been observed for the bacterial SecYEG-RNC complex (C. Schaffitzel, personal communication). The importance of L23 for the docking of protein export factors even extends to the structurally distinct 55S mitochondrial ribosome (Sharma et al., 2003). Here, the Oxa1 translocase, facilitating protein insertion into the inner mitochondrial membrane, efficiently cross-links to the mitochondrial L23 homologue (Jia et al., 2003).
The observed interaction of the trigger factor in the crystal structure of the RB domain-ribosome complex will be useful in designing smaller ribosome binding domains of the SRP, SR or the translocon to be used in the future crystallisation experiments similar to what has been done with the trigger factor fragment. The experience accumulated regarding the crystallisation strategies can be applied to all of these factors. It seems that the strategy that involves soaking of the factor into the crystals of the large ribosomal subunit has many advantages over co-crystallisation experiments since the complex between the ribosome and its ligand does not have to be formed prior to the crystallisation experiment. Furthermore, co-crystallisation experiments with these factors would require that they are present in a significant excess over the ribosomes due to their relatively low affinity for vacant ribosomes. Finally, large excess of the protein in the crystallisation mixture may prevent crystallisation completely or it requires extensive re-screening of the crystallisation conditions.

4.4.2.2 SRP-dependent export

Surprisingly, simultaneous binding of SRP and the trigger factor to the ribosome has been observed in spite of the fact that they both use L23 as a contact site and occupy sterically exclusive regions on the ribosomal surface (Buskiewicz et al., 2004; Raine et al., 2004). These observations suggest that conformational rearrangements must take place to permit binding of both factors at the same time. Biochemical data further suggests that binding of FtsY, the bacterial SRP receptor homologue, to the SRP-TF-ribosome complex excludes the
trigger factor (Buskiewicz et al., 2004). SRP exhibits high on/off-rates, enabling it to sample all ribosomes for signal anchor sequences, though it is present at substoichiometric amounts \textit{in vivo} (Buskiewicz et al., 2004). Competition of binding to signal anchor sequences between the trigger factor and SRP has been demonstrated by crosslinking studies, indicating a clear advantage for SRP over the trigger factor (Ullers et al., 2003). However, the presence of nascent-chain interacting trigger factor prevents the binding of SRP to the less hydrophobic signal sequences of proteins determined for Sec-dependent export and thus increases SRP’s specificity for its cognate substrates (Beck et al., 2000; Lee and Bernstein, 2001).

\subsection*{4.4.2.3 Sec-dependent export}

Deletion of the trigger factor accelerates Sec-dependent secretion and reduces the requirement for SecB as a targeting factor such that translocation occurs even in its absence and perhaps already in a co-translational mode. Based on these observations and the fact that the trigger factor has a primary access to nascent chains, the likely function of SecB in translocation is to receive polypeptides from the trigger factor and keep them in an unfolded translocation-competent state for the subsequent post-translational transport through the membrane (Lee and Bernstein, 2002). The transfer of nascent chains might be guided by the presence of low affinity binding of SecB to free trigger factor in solution (Ha et al., 2004). The general ability of SecB to prevent the premature folding of nascent chains has recently become evident by the fact that a 10-fold over-expression of the already abundant SecB can suppress the growth defects of a strain lacking the trigger factor and DnaK. Its co-translational binding to short nascent cytosolic and secretory proteins is indiscriminately suppressed by the presence of the trigger factor (Ullers et al., 2004). However, the fact that its primary role is in protein translocation and that strong over-expression of SecB is required for suppression of the TF / DnaK-phenotype indicates that it does not promote folding as efficiently as the combination of the trigger factor and DnaK.
5 Future prospects

The trigger factor, apparently, requires two conflicting properties to fulfil its role in the early co-translational phase of protein folding: It has to interact and shield nascent chains in order to facilitate co-translational folding while it also has to permit and potentially control access of other factors to the nascent chains. The unique dragon-shaped fold of the trigger factor seems perfectly tailored to match both requirements. On the one hand the cradle has a hydrophobic inner surface, which enables the trigger factor to interact with hydrophobic patches of nascent chains preventing them from getting trapped in unproductive folding intermediates. On the other hand, the trigger factor, due to its unusual open shape and relatively weak affinity to the ribosome mediated through delicate contacts, is in a position to precisely control the hand over of the nascent chains to downstream post-translational folding factors such as DnaK, SecB and GroEL/ES or to the co-translational protein targeting machinery. Remarkably, even in the presence of the trigger factor, bacterial SRP is able to gain access to the nascent chain and to bind to hydrophobic signal anchor sequences. The competition of the trigger factor and SRP for the binding of less hydrophobic signal secretion sequences enhances the specificity of SRP. The trigger factor and SRP together with further factors involved in protein secretion, e.g. the translocon and the mitochondrial Oxa1 protein, utilise ribosomal protein L23 (or its homologue) as a general interaction platform.

Structural information presented here will serve as a basis for functional experiments aimed at addressing the mechanism of protein folding aided by ribosome-associated chaperones. Molecular details of the interplay between the trigger factor and post-translational folding factors and between the trigger factor and the components of the co-translational protein translocation machinery remain to be elucidated and significant contributions can be expected from interdisciplinary structural studies of individual complexes involved in these processes.
6 References


(SRP) complexes and is excluded by binding of the SRP receptor. *Proc Natl Acad Sci USA*, **101**, 7902-7906.


Lynnon. (2001) DNAMAN. Quebec, Canada.


Sekerina, E., Rahfeld, J.U., Muller, J., Fanghanel, J., Rascher, C., Fischer, G. and Bayer, P. (2000) NMR solution structure of hPar14 reveals similarity to the peptidyl prolyl
cis/trans isomerase domain of the mitotic regulator hPin1 but indicates a different functionality of the protein. *J Mol Biol*, **301**, 1003-1017.


# Appendix

## 7.1 General chemicals

Chemicals were purchased from Applichem, Fluka or Sigma if not stated otherwise.

### 7.1.1 Buffers and solutions

<table>
<thead>
<tr>
<th>Buffer/Solution</th>
<th>Composition</th>
<th>Buffer/Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>70S buffer</td>
<td>3.4 M KCl 0.5 M NH₄Cl 100 mM MgCl₂ 10 mM Tris-HCl 6 mM β-mercaptoethanol pH 7.6</td>
<td>Blocking buffer</td>
<td>TBS 3 % BSA</td>
</tr>
<tr>
<td>Coomassie destaining solution</td>
<td>40.0% (v/v) Ethanol 10.0% (v/v) Acetic acid</td>
<td>Coomassie staining solution</td>
<td>0.25 % (w/v) Coomassie Brilliant Blue R-250 40.0 % (v/v) Ethanol 10.0 % (v/v) Acetic acid</td>
</tr>
<tr>
<td>Dialysis buffer (50S)</td>
<td>1.2 M KCl 0.5 M NH₄Cl 20 mM MgCl₂ 10 mM Tris-HCl 5 mM β-mercaptoethanol pH 7.5</td>
<td>Dissociation buffer</td>
<td>2.7 M KCl 0.45 M NH₄Cl 1 mM MgCl₂ 20mM Tris-HCl 6 mM β-mercaptoethanol pH 8.0</td>
</tr>
<tr>
<td>6 x DNA loading dye</td>
<td>0.25 % (w/v) Bromophenol blue 30 % (w/v) Glycerol</td>
<td>Final buffer (50S)</td>
<td>1.2 M KCl 0.5 M NH₄Cl 20 mM MgCl₂ 100 mM NaAcetate 10 mM Tris-HCl 10 mM MES pH 7.1</td>
</tr>
<tr>
<td>French press buffer</td>
<td>50 mM Tris-HCl 200 mM NaCl 20 mM imidazole 1 mM EDTA 1 mM PMSF pH 7.5</td>
<td>High-salt buffer (Resource Q6)</td>
<td>50 mM Tris-HCl 1 M NaCl 1 mM EDTA pH 7.5</td>
</tr>
<tr>
<td>Low-salt buffer (Resource Q6)</td>
<td>50 mM Tris-HCl 25 mM NaCl 1 mM EDTA pH 7.5</td>
<td>NiNTA elution buffer</td>
<td>50 mM NaH₂PO₄ 300 mM NaCl 1 M Imidazole pH adjusted to 8.0 with NaOH</td>
</tr>
<tr>
<td>NiNTA lysis buffer</td>
<td>50 mM NaH₂PO₄ 300 mM NaCl 10 mM Imidazole pH adjusted to 8.0 with NaOH</td>
<td>NiNTA wash buffer</td>
<td>50 mM NaH₂PO₄ 300 mM NaCl 20 mM Imidazole pH adjusted to 8.0 with NaOH</td>
</tr>
<tr>
<td>2 x PGLB</td>
<td>125 mM Tris-HCl pH 6.8 1.4 M β-mercaptoethanol 10 % (w/v) SDS 0.02 % (w/v) Bromophenol blue 20 % Glycerol</td>
<td>10 x SDS running buffer</td>
<td>250 mM TRIZMA base 1.9 M Glycine 10 % (w/v) SDS pH not adjusted (~8.3)</td>
</tr>
</tbody>
</table>
### 7.1.2 Media for bacterial growth

<table>
<thead>
<tr>
<th>Medium</th>
<th>Composition (per litre)</th>
</tr>
</thead>
</table>
| ATCC medium 1230 *H. marismortui* | 2.0 g KCl  
0.8 g MgSO₄ (anhydrous)  
3.5 g Na Citrate·2H₂O  
250.0 g NaCl  
0.2 g CaCl₂ (anhydrous)  
50 ml 10% Yeast extract  
100 ml 0.5M Tris-HCl (pH 7.2)  
20 ml 20% Glucose  
2.0 ml Trace elements |
| Luria-Bertani (LB) medium (liquid) | 10 g Bacto-tryptone  
5 g Bacto-yeast extract  
10 g NaCl |
| LB medium plates | 10 g Bacto-tryptone  
5 g Bacto-yeast extract  
10 g NaCl  
15 g Bacto-agar |
| 5 x M9 salts | 64 g Na₂HPO₄·7H₂O  
15 g KH₂PO₄  
2.5 g NaCl  
5.0 g NH₄Cl |
| M9 minimal medium | 200 ml 5 x M9 salts  
stereile ddH₂O to 1 litre  
1 ml 1M MgSO₄  
100 ml of 4% (w/v) Glucose  
0.1 ml of 0.5% (w/v) Thiamine  
1 ml of 4.2 g l⁻¹ Fe(II)SO₄ |
| Trace elements *H. marismortui* | 0.218 g MnCl₂·4H₂O  
4.860 g FeCl₃·6H₂O |
| Amp, Cam, Kan, Sp, Tet | Medium supplemented with Ampicillin (100 µg/ml), Chloramphenicol (25 µg/ml), Kanamycin (50 µg/ml), Spectinomycin (50 µg/ml), Tetracyclin (30 µg/ml). |

### 7.1.3 Enzymes

Enzymes were purchased from NEB, Applichem or MBI and used with the recommended buffers.
### 7.1.4 Gel electrophoresis molecular weight markers

<table>
<thead>
<tr>
<th>Marker</th>
<th>Supplier</th>
<th>Marker band size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peqgold 100 bp DNA ladder</td>
<td>Peqlab</td>
<td>100, 200, 300, 400, 600, 800, 1031, 1200, 1500, 2000, 3000 base pairs</td>
</tr>
<tr>
<td>Peqgold 1kb DNA ladder</td>
<td>Peqlab</td>
<td>250, 500, 750, 1000, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10000 base pairs</td>
</tr>
<tr>
<td>Low molecular weight (LMW) protein marker for SDS-PAGE</td>
<td>Amersham</td>
<td>14.4, 20.1, 30.0, 45.0, 66.0, 97.0 kD</td>
</tr>
</tbody>
</table>

### 7.1.5 General equipment

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Supplier</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>-20° C Freezer</td>
<td>Liebherr</td>
<td>Sample storage</td>
</tr>
<tr>
<td>C 19 Waterbath</td>
<td>HAAKE</td>
<td>Heat incubation</td>
</tr>
<tr>
<td>DB-2D Dri-Block</td>
<td>Techne</td>
<td>Heating block</td>
</tr>
<tr>
<td>Electrophoresis power supply EPS 3501</td>
<td>Amersham Biosciences</td>
<td>Gel electrophoresis</td>
</tr>
<tr>
<td>Emulsiflex-C5 high pressure homogeniser</td>
<td>Avestin</td>
<td>Cell lysis (&quot;cell cracker&quot;)</td>
</tr>
<tr>
<td>Fridge (4° C)</td>
<td>Affinox</td>
<td>FPLC equipment cooling</td>
</tr>
<tr>
<td>Fridge (4° C)</td>
<td>Liebherr</td>
<td>Sample storage</td>
</tr>
<tr>
<td>Hoefer HE mini submarine unit</td>
<td>Amersham Biosciences</td>
<td>Agarose gel casting/electrophoresis</td>
</tr>
<tr>
<td>Hoefer Mighty Small II SE 250</td>
<td>Amersham Biosciences</td>
<td>Gel electrophoresis</td>
</tr>
<tr>
<td>HT multitron II</td>
<td>Infors</td>
<td>Cell shaker</td>
</tr>
<tr>
<td>KM-2</td>
<td>Edmund Bühler</td>
<td>Rotating platform</td>
</tr>
<tr>
<td>Microwave NN-E202</td>
<td>Panasonic</td>
<td>Heating</td>
</tr>
<tr>
<td>MILLI-RO 10 plus</td>
<td>Millipore</td>
<td>H₂O ultrafiltration</td>
</tr>
<tr>
<td>MR 3001 K</td>
<td>Heidolph</td>
<td>Stirring and/or heating solutions</td>
</tr>
<tr>
<td>Rabbit pump</td>
<td>Made as ordered</td>
<td>Gradient collection</td>
</tr>
<tr>
<td>REAX control</td>
<td>Heidolph</td>
<td>Vortexing</td>
</tr>
<tr>
<td>ROSS pH electrode</td>
<td>Orion</td>
<td>pH measurement</td>
</tr>
<tr>
<td>Thermostatically controlled cabinets ET 618 and ET 626</td>
<td>Lovibond</td>
<td>Crystallisation</td>
</tr>
<tr>
<td>Ultrospec 2000</td>
<td>Amersham Biosciences</td>
<td>UV/VIS spectrophotometrie</td>
</tr>
<tr>
<td>Unimax 1010</td>
<td>Heidolph</td>
<td>Rotating platform</td>
</tr>
</tbody>
</table>

### 7.1.6 Centrifuges

<table>
<thead>
<tr>
<th>Centrifuge</th>
<th>Supplier</th>
<th>Rotor</th>
</tr>
</thead>
<tbody>
<tr>
<td>5415 D (table top)</td>
<td>Eppendorf</td>
<td>F-45-24-11, fixed angle</td>
</tr>
<tr>
<td>5417 R (table top)</td>
<td>Eppendorf</td>
<td>F-45-30-11, fixed angle</td>
</tr>
<tr>
<td>5810 R (table top)</td>
<td>Eppendorf</td>
<td>A-4-62, swinging bucket</td>
</tr>
<tr>
<td>EvolutionRC</td>
<td>Sorvall</td>
<td>Same as for RC 50 plus and SLC 6000, fixed angle</td>
</tr>
<tr>
<td>RC 50 plus</td>
<td>Sorvall</td>
<td>SS-34, fixed angle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GSA, fixed angle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SLA-1500, fixed angle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SLA-3000, fixed angle</td>
</tr>
<tr>
<td>Optima L-80 XP</td>
<td>Beckman</td>
<td>SW28, SW32, Ti70</td>
</tr>
</tbody>
</table>
7.1.7 FPLC equipment

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Model</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKTA FPLC fraction collector</td>
<td>Frac-900</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>AKTA FPLC mixer</td>
<td>M-925</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>AKTA FPLC monitor</td>
<td>UPC-900</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>AKTA FPLC system pump</td>
<td>P-920</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>AKTA FPLC valve</td>
<td>INV-907</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>Empty columns</td>
<td>HR and XK series</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>UNICORN software package</td>
<td>4.10</td>
<td>Amersham Biosciences</td>
</tr>
</tbody>
</table>

7.1.8 FPLC and resins

<table>
<thead>
<tr>
<th>FPLC resin</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>NiNTA agarose</td>
<td>Qiagen</td>
</tr>
<tr>
<td>NiNTA Superflow resin</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Resource Q6</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>Superdex 200</td>
<td>Amersham Biosciences</td>
</tr>
</tbody>
</table>

7.2 Crystallisation equipment

7.2.1 General equipment

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Supplier</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayer Silicon mittelviskos</td>
<td>Bayer</td>
<td>Seal cover slides to Linbro plate</td>
</tr>
<tr>
<td>Cryo-loops (0.05 to 0.5 mm)</td>
<td>Hampton Research</td>
<td>Mounting crystals</td>
</tr>
<tr>
<td>Cryo-vials</td>
<td>Corning</td>
<td>Freezing, storage, transportation of crystals</td>
</tr>
<tr>
<td>Cryschem plate</td>
<td>Hampton Research</td>
<td>Sitting drop crystallisation</td>
</tr>
<tr>
<td>Crystal clear tape</td>
<td>Hampton Research</td>
<td>Sealing Cryschem plates</td>
</tr>
<tr>
<td>Dewars</td>
<td>Lovibond</td>
<td>Freezing, storage and transportation of crystals</td>
</tr>
<tr>
<td>Linbro plate</td>
<td>Hampton Research</td>
<td>Hanging drop crystallisation</td>
</tr>
<tr>
<td>Microscopes</td>
<td>Nikon</td>
<td>Visualisation</td>
</tr>
<tr>
<td>XYZ short goniometer head</td>
<td>Hampton Research</td>
<td>Crystal mounting and alignment</td>
</tr>
<tr>
<td>Yale magnetic pins</td>
<td>Made as ordered</td>
<td>Mounting crystals</td>
</tr>
<tr>
<td>22 mm glass circle cover slides</td>
<td>Hampton Research</td>
<td>Hanging drop crystallisation</td>
</tr>
<tr>
<td>- siliconised</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.2.2 X-ray sources and detection

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Supplier</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>MarCCD detector</td>
<td>MAR Research</td>
<td>X-ray detection</td>
</tr>
<tr>
<td>Mar345 image plate detector</td>
<td>MAR Research</td>
<td>X-ray detection</td>
</tr>
<tr>
<td>Synchrotron</td>
<td>Swiss Light Source (SLS)</td>
<td>X-ray source</td>
</tr>
<tr>
<td></td>
<td>Beamline: X06SA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ESRF (Grenoble)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beamline: Swiss-Norwegian Beamline (SNBL)</td>
<td></td>
</tr>
</tbody>
</table>
7.2.3 Crystallisation screens

7.2.3.1 Commercially available crystallisation screens

All screens were purchased from Hampton Research: Sparse Matrix Crystallisation Screen I, Sparse Matrix Crystallisation Screen II, Nucleic Acid Mini Screen, Nucleic Acid Sparse Matrix Screen.

7.2.3.2 Grid screens: Ammonium sulphate, PEG, MPD vs pH

Ammonium sulphate grid screen:

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>0.1 M NaOAc pH 4.0, 0.32 M (NH₄)₂SO₄</th>
<th>C1</th>
<th>0.1 M NaOAc pH 4.0, 0.96 M (NH₄)₂SO₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>0.1 M NaCitrate pH 5.0, 0.32 M (NH₄)₂SO₄</td>
<td>C2</td>
<td>0.1 M NaCitrate pH 5.0, 0.96 M (NH₄)₂SO₄</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>0.1 M NaCitrate pH 6.0, 0.32 M (NH₄)₂SO₄</td>
<td>C3</td>
<td>0.1 M NaCitrate pH 6.0, 0.96 M (NH₄)₂SO₄</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>0.1 M Tris-HCl pH 7.0, 0.32 M (NH₄)₂SO₄</td>
<td>C4</td>
<td>0.1 M Tris-HCl pH 7.0, 0.96 M (NH₄)₂SO₄</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>0.1 M Tris-HCl pH 8.0, 0.32 M (NH₄)₂SO₄</td>
<td>C5</td>
<td>0.1 M Tris-HCl pH 8.0, 0.96 M (NH₄)₂SO₄</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>0.1 M Tris-HCl pH 9.0, 0.32 M (NH₄)₂SO₄</td>
<td>C6</td>
<td>0.1 M Tris-HCl pH 9.0, 0.96 M (NH₄)₂SO₄</td>
<td></td>
</tr>
</tbody>
</table>

Ammonium sulphate grid screen:

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>0.1 M NaOAc pH 4.0, 5% PEG</th>
<th>C1</th>
<th>0.1 M NaOAc pH 4.0, 15% PEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>0.1 M NaCitrate pH 5.0, 5% PEG</td>
<td>C2</td>
<td>0.1 M NaCitrate pH 5.0, 15% PEG</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>0.1 M NaCitrate pH 6.0, 5% PEG</td>
<td>C3</td>
<td>0.1 M NaCitrate pH 6.0, 15% PEG</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>0.1 M Tris-HCl pH 7.0, 5% PEG</td>
<td>C4</td>
<td>0.1 M Tris-HCl pH 7.0, 15% PEG</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>0.1 M Tris-HCl pH 8.0, 5% PEG</td>
<td>C5</td>
<td>0.1 M Tris-HCl pH 8.0, 15% PEG</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>0.1 M Tris-HCl pH 9.0, 5% PEG</td>
<td>C6</td>
<td>0.1 M Tris-HCl pH 9.0, 15% PEG</td>
<td></td>
</tr>
</tbody>
</table>

PEG grid screen (PEG 400 or PEG4000):

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>0.1 M NaOAc pH 4.0, 10% PEG</th>
<th>C1</th>
<th>0.1 M NaOAc pH 4.0, 20% PEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>0.1 M NaCitrate pH 5.0, 10% PEG</td>
<td>C2</td>
<td>0.1 M NaCitrate pH 5.0, 20% PEG</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>0.1 M NaCitrate pH 6.0, 10% PEG</td>
<td>C3</td>
<td>0.1 M NaCitrate pH 6.0, 20% PEG</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>0.1 M Tris-HCl pH 7.0, 10% PEG</td>
<td>C4</td>
<td>0.1 M Tris-HCl pH 7.0, 20% PEG</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>0.1 M Tris-HCl pH 8.0, 10% PEG</td>
<td>C5</td>
<td>0.1 M Tris-HCl pH 8.0, 20% PEG</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>0.1 M Tris-HCl pH 9.0, 10% PEG</td>
<td>C6</td>
<td>0.1 M Tris-HCl pH 9.0, 20% PEG</td>
<td></td>
</tr>
</tbody>
</table>
### MPD grid screen:

| A1 | 0.1 M NaOAc pH 4.0, 5% MPD | C1 | 0.1 M NaOAc pH 4.0, 15% MPD |
| A2 | 0.1 M NaCitrate pH 5.0, 5% MPD | C2 | 0.1 M NaCitrate pH 5.0, 15% MPD |
| A3 | 0.1 M NaCitrate pH 6.0, 5% MPD | C3 | 0.1 M NaCitrate pH 6.0, 15% MPD |
| A4 | 0.1 M Tris-HCl pH 7.0, 5% MPD | C4 | 0.1 M Tris-HCl pH 7.0, 15% MPD |
| A5 | 0.1 M Tris-HCl pH 8.0, 5% MPD | C5 | 0.1 M Tris-HCl pH 8.0, 15% MPD |
| A6 | 0.1 M Tris-HCl pH 9.0, 5% MPD | C6 | 0.1 M Tris-HCl pH 9.0, 15% MPD |
| B1 | 0.1 M NaOAc pH 4.0, 10% MPD | D1 | 0.1 M NaOAc pH 4.0, 20% MPD |
| B2 | 0.1 M NaCitrate pH 5.0, 10% MPD | D2 | 0.1 M NaCitrate pH 5.0, 20% MPD |
| B3 | 0.1 M NaCitrate pH 6.0, 10% MPD | D3 | 0.1 M NaCitrate pH 6.0, 20% MPD |
| B4 | 0.1 M Tris-HCl pH 7.0, 10% MPD | D4 | 0.1 M Tris-HCl pH 7.0, 20% MPD |
| B5 | 0.1 M Tris-HCl pH 8.0, 10% MPD | D5 | 0.1 M Tris-HCl pH 8.0, 20% MPD |
| B6 | 0.1 M Tris-HCl pH 9.0, 10% MPD | D6 | 0.1 M Tris-HCl pH 9.0, 20% MPD |

### 7.2.3.3 Additive screens

#### Additive screen I:

| A1 | 11 mM CdCl<sub>2</sub> | B1 | 11 mM EuCl<sub>3</sub> | C1 | 10 mM Urea | D1 | 1.5% 1,6-Hexanediol |
| A2 | 16 mM CaCl<sub>2</sub> | B2 | 5 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> | C2 | 0.5% Guanidinium-HCl | D2 | 1.5% PEG 400 |
| A3 | 11 mM CuSO<sub>4</sub> | B3 | 10 mM LiCl | C3 | 10 mM Sucrose | D3 | 0.5% PEG 4000 |
| A4 | 5 mM MgCl<sub>2</sub> | B4 | 10 mM FeCl<sub>3</sub> | C4 | 19 mM Glucose | D4 | 5 mM CoCl<sub>2</sub> |
| A5 | 5 mM ZnSO<sub>4</sub> | B5 | 10 mM KCl | C5 | 10 mM Maltose | D5 | 1.5% Methanol |
| A6 | 14 mM MnCl<sub>2</sub> | B6 | 10 mM NiSO<sub>4</sub> | C6 | 13 mM DTT | D6 | 2% Acetone |

#### Additive screen II:

| A1 | 2 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> | B1 | 15 mM Maltose | C1 | 1.5% PEG400 | D1 | 1.5 mM Spermine |
| A2 | 5 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> | B2 | 20 mM Maltose | C2 | 3.0% PEG400 | D2 | 2.0 mM Spermine |
| A3 | 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> | B3 | 1% 1,6-Hexanediol | C3 | 6.0% PEG400 | D3 | 1% Acetone |
| A4 | 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> | B4 | 5% 1,6-Hexanediol | C4 | 12% PEG400 | D4 | 2% Acetone |
| A5 | 5 mM Maltose | B5 | 10% 1,6-Hexanediol | C5 | 0.5 mM Spermine | D5 | 4% Acetone |
| A6 | 10 mM Maltose | B6 | 12% 1,6-Hexanediol | C6 | 1.0 mM Spermine | D6 | 8% Acetone |

#### Additive screen III:

| A1 | 2% Fructose | B1 | 2% Sucrose | C1 | 2% Trehalose | D1 | 2% Glycerin |
| A2 | 5% Fructose | B2 | 5% Sucrose | C2 | 5% Trehalose | D2 | 5% Glycerin |
| A3 | 7% Fructose | B3 | 7% Sucrose | C3 | 7% Trehalose | D3 | 7% Glycerin |
| A4 | 10% Fructose | B4 | 10% Sucrose | C4 | 10% Trehalose | D4 | 10% Glycerin |
| A5 | 12% Fructose | B5 | 12% Sucrose | C5 | 12% Trehalose | D5 | 12% Glycerin |
| A6 | 15% Fructose | B6 | 15% Sucrose | C6 | 15% Trehalose | D6 | 15% Glycerin |
7.3 Plasmid maps

7.3.1 pH6-tig

Trigger factor genes were cloned into pDS56/RBSII using its BamHI site (Hesterkamp et al., 1997). The plasmid map of pH6-tig containing the full-length trigger factor gene tig (red) is shown together with selected restriction sites. The shorter fragments were cloned in the same way to yield plasmids pH6-1-144 and pH6-Δ47.

7.3.2 pH6-1-111-TEV

Trigger factor fragment TF111 was cloned into the pET3a derivative pRET3aHisTrxRAP74IAD using its BamHI and KpnI sites while removing the iad gene (green), which is coding for an HIV protein. The cloning procedure results in a trigger factor gene fused to an N-terminally His₆-tagged thioredoxin (Trx) gene (orange), which is linked to a TEV cleavage site. Upon TEV cleavage the His₆-TRX-fusion protein is removed resulting in an untagged TF111 protein with the two additional amino acids glycine and serine at its N-terminus.
The expressed protein sequence (26.3 kDa) is shown below with the His₆-tag marked with bold letters, the TEV recognition site is depicted in green and the TF111 protein sequence (12.3 kDa) in red. The TEV protease cleaves between amino acids Gln and Gly.

```
MGSSHHHHHHHHHHHHHHHSDMVPRVVTQGRRLVTVSPSNQSVVVQIVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAVAV
7.4 Trigger factor amino acid sequence

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>21</td>
<td>31</td>
<td>41</td>
<td>51</td>
</tr>
<tr>
<td>MRGMQSVE</td>
<td>TQGLGRVVT</td>
<td>ITIAADSIET</td>
<td>AVKSELVNVA</td>
<td>KKVRIDGFRK</td>
<td>GKVPMNIVAQ</td>
</tr>
<tr>
<td>61</td>
<td>RYGASVRQDV</td>
<td>LQDLMSRNFI</td>
<td>DAIKEKINP</td>
<td>AGAPTVPGE</td>
<td>YKLGEDFTYS</td>
</tr>
<tr>
<td>121</td>
<td>LQGLEAEIEVE</td>
<td>KPIVEVTDAD</td>
<td>VDGMLTLRK</td>
<td>QQATWKEDK</td>
<td>GAVEAEDRTVI</td>
</tr>
<tr>
<td>181</td>
<td>FEGGKASDFV</td>
<td>LAMGQGRMIP</td>
<td>GFEDG1KGHK</td>
<td>AGEEPTIDVT</td>
<td>FPEEYHAENL</td>
</tr>
<tr>
<td>241</td>
<td>LKKEERLIP</td>
<td>ELTAEFIKRF</td>
<td>GVEDGSVEGL</td>
<td>RAERVKNNER</td>
<td>ELKSAIRNRV</td>
</tr>
<tr>
<td>301</td>
<td>ANDIDVPAAL</td>
<td>IDSEIDVLR</td>
<td>QQARQFNGNE</td>
<td>KQAELPLREL</td>
<td>FEEQAKRRVV</td>
</tr>
<tr>
<td>361</td>
<td>TNLKADDEER</td>
<td>VRGLIEEMAS</td>
<td>AYEDPKIEVE</td>
<td>FYSKNNELMD</td>
<td>NMRNVAEEQ</td>
</tr>
<tr>
<td>421</td>
<td>VTEKETTENL</td>
<td>MNQQARSHH</td>
<td>HHHH++</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of amino acids: 444 (during the cloning procedure additional codons coding for 12 amino acids shown in bold were added to the wild type trigger factor gene)

Molecular weight: 49690.2

Theoretical pI: 5.06

Total number of negatively charged residues (Asp + Glu): 84
Total number of positively charged residues (Arg + Lys): 63

Formula: C_{2180}H_{3514}N_{612}O_{688}S_{12}
Total number of atoms: 7006

Extinction coefficient (280 nm): 15930 M^{-1} cm^{-1}
Abs 0.1% (= 1 g l^{-1}): 0.321
(Conditions: 6.0 M guanidium hydrochloride, 0.02 M phosphate buffer, pH 6.5)

7.5 Multiwavelength anomalous dispersion (MAD) experiments

With the better availability of synchrotron beam lines and the common use of cryoprotection for crystals MAD has become a routinely used approach for phase determination (Moffat and Ren, 1997). All diffraction data can be collected from a single derivatised crystal guaranteeing perfect isomorphism of the datasets if the crystal does not decay during the X-ray exposure. The essential requirements for successful MAD experiments like the tuneability to a set of wavelengths with great precision, cryocooling of the sample, high flux on the sample in a range of energies are inherent characteristics of synchrotron beam lines. The collection of anomalous diffraction data at a single wavelength (SAD or SIR data) does not provide unique phase choices (SAD phase ambiguity) and electron-density maps obtained from SAD phases are rather difficult to interpret. The MAD method solves the SAD phase ambiguity and produces interpretable electron-density maps by collecting additional data near the absorption edge of the anomalous scatterer. The differences between diffraction amplitudes is then used to obtain unique phase choices (Hendrickson and Ogata, 1997).
$F_{\text{anom}}$ can be split into a wavelength independent normal atomic scattering coefficient $f^0$ and a wavelength dependent term $f(\lambda)$. Absorption effects of the heavy atom results on the one hand in a reduction of the scattering power ($f'(\lambda)$) as well as in an addition of an imaginary term to its phase ($if''(\lambda)$). Both the dispersive ($f'(\lambda)$) as well as the anomalous differences ($if''(\lambda)$) are proportional to the structure factor amplitudes of the heavy atom used.

$$F_{\text{anom}} = f_0 + f(\lambda) = f_0 + f'(\lambda) + if''(\lambda)$$

The appropriate wavelengths for the MAD experiment have to be selected based on X-ray absorption spectra where the real ($f'$) and imaginary term ($f''$) of the anomalous scattering factor ($F_{\text{anom}}$) are plotted as a function of the X-ray energy in the range of interest. The interaction of the scattering atom with its chemical neighbours significantly influences the scattering behaviour, so that the experimental values for the scattering factors differ from the Cromer-Liberman theoretical values (Cromer and Liberman, 1981). Therefore, each time at the synchrotron an X-ray absorption scan in the near-edge region of the used heavy atom compound has to be performed which relates directly to the imaginary component of the anomalous scattering factors ($f''$). The corresponding $f'$ spectrum is then calculated by the Kramers-Kronig transformation (Kronig and Kramers, 1928) which relates $f'$ to $f''$.

In theory two datasets of the same crystal, one of them collected at the absorption edge, would be sufficient for phase determination (Gonzalez et al., 1999). Nevertheless, usually three datasets are collected: The first at the wavelength with maximal $f''$ (peak wavelength $\lambda_1$), the second at a wavelength with minimal $f'$ (wavelength at the inflection point $\lambda_2$) and the third at a wavelength far from the absorption edge with a small anomalous signal (remote wavelength $\lambda_3$). This way both the dispersive ($\lambda_2$ and $\lambda_3$) and anomalous (Bijvoet) differences are maximised ($\lambda_1$) (Gonzalez et al., 1999).
8 Acknowledgements

I am particularly thankful to Prof. Nenad Ban for guiding me through my PhD thesis over the past four years. I was allowed to explore my research interests in an excellent environment. He was always there for support, personal attention and motivation throughout the project.

I would also like to thank our collaborators Dr. Elke Deuerling, Prof. Bernd Bukau and their group members. Without their help the project would never have succeeded. I am especially grateful to Elke, who was always there in case of any trouble (both with good advice and fast delivery of urgently needed protein samples) and critically helped, together with Bernd, in preparing the manuscript; and Holger, who observed the first trigger factor microcrystals.

Now I would like to thank Prof. Rudolph Glockshuber and Dr. Elke Deuerling for their effort in reviewing my thesis and for accepting to be my coexaminer.

Thanks a lot to the members of the Ban Lab for all the help and discussions not only concerning the project, but also for the excellent working atmosphere! I am especially grateful to: Snježana, who re-established not only the growth of Haloarcula marismortui cells, but also managed to obtain the corresponding 50S crystals for the first time in this lab. She spent a lot of time with me at the synchrotron and always tried to make the time there as enjoyable as possible; Timm, who worked day and night with me to finally put the project into a publishable form and to solve any problem in structure determination; and finally Marco who invested his time in growing Haloarcula marismortui cells and crystals of 50S.

I am also grateful to Clemens Schulze-Briese, Takashi Tomizaki and Armin Wagner at the Swiss Light Source (SLS) of the Paul Scherrer Institut (PSI) in Villigen and Phil Pattison and Silvia Capelli at the Swiss Norwegian Beamline (SNBL) at the ESRF (Grenoble). Without their outstanding efforts, these experiments would not have been possible.
This work was supported by the Swiss National Science Foundation (SNSF), the NCCR Structural Biology program of the SNSF, grants of the Deutsche Forschungsgemeinschaft to B. Bukau and E. Deuerling and a Young Investigator grant from the Human Frontier Science Program to N. Ban and E. Deuerling.

Thank you Simone for all your help and love especially in the busy, final stages of my thesis. You always believed in me to succeed in whatever I have been trying, so that there was no place for failing.

Last but not least, I would like to thank my parents to whom this work is dedicated. They have been close to me through their love and support during my studies despite the spatial distance between us.
9 Curriculum Vitae

Lars FERBITZ

Date of Birth 14. October 1974
Birthplace Siegen
Nationality German

Business address Institute of Molecular Biology and Biophysics
Swiss Federal Institute of Technology
Schafmattstr. 20
ETH Honggerberg
Building HPK D3
CH-8093 Zurich
Tel.: +41(0) 44 6333148

Home address Ottostr. 8
8005 Zurich
Phone: +41(0) 43 3213225

EDUCATION

01/2001 – 08/2005 Doctoral studies
Institute for Molecular Biology and Biophysics
Swiss Federal Institute of Technology (ETH) Zurich,
Switzerland, in the group of Prof. Dr. Ban

Diploma degree, diploma thesis 1 in the group of Prof. Dr. Hegemann

10/1998 – 07/1999 Erasmus exchange program
Diploma in Biochemistry with distinction at the
University of Kent at Canterbury, research project 2 in the
group of Prof. Dr. Jarvis

08/1985 – 06/1994 Matura degree
Gymnasium Neunkirchen and Gymnasium Siegen,
Germany

1 Title:"Detection of promotor activity in Chlamydomonas reinhardtii using a synthetic reporter gene that codes for the Renilla reniformis gene”

2 Title:"Subcloning of a putative nucleoside transporter in aid of its expression in a mammalian cell line”
WORK EXPERIENCE

01/2001 – 08/2005  Cloning of homologous genes using degenerated primers and ‘step-down’ PCR; protein purification techniques; ribosome purification (ultracentrifugation, density gradients); protein crystallography (crystallisation, data collection, computational techniques)

06/2000 – 11/2000  Gene synthesis; working with an unicellular green alga: Cultivation, transformation; optimisation of long term and high throughput measurements of promoter activities using luminescence assays

02/1999 – 04/1999  Working with a mammalian cell line (COS-1): Cultivation and transfection; measuring uptake activities of membrane transporters using radioactively labelled substrates

SELECTED PRESENTATIONS

Selected for an oral presentation at the EMBO conference on “Structures in Biology”, 11th November 2004, EMBL Heidelberg, Germany

Oral presentation at the 18th Regional Meeting on “X-ray Crystallography of Biomacromolecules”, 23th September 2004, Einsiedeln, Switzerland

Poster prize for the presentation at the meeting “Structural Biology at Crossroads: From Biological Molecules to Biological Systems”, 15th - 18th September 2004, EMBL-Hamburg, Germany

PUBLICATIONS

