Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

Functional and structural studies on the ribosome associated factors Trigger Factor and NAC

Julian Johannes Deeng

aus Ochsenhausen

2012

Erklärung
Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Roland Beckmann betreut.

Eidesstattliche Versicherung
Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 19.10.2012

Julian Deeng

Dissertation eingereicht am: 19.10.2012

1. Gutachter: Herr Prof. Dr. Roland Beckmann
2. Gutachter: Herr Prof. Dr. Klaus Förstemann

Mündliche Prüfung am: 20.11.2012

Table of Contents

Abbreviations

°C	degree Celsius
μl	10^{-6} L
A_{xxx}	absorption at x nm
AA	amino acid
AAP	argenine attenuator peptide
aeNAC	homodimeric archaeal NAC
Angström (Å)	0.1 nm
BCA	bicinchoninic acid
BiP	lumenal eukaryotic Hsp70
bp	base pair
BSA	bovine serum albumin
BTF3	basic transcription factor 3
BTT1	*S. cerevisiae* β_2 NAC
c	concentration
Ca	calcium
CAM	chloramphenicol
CC C	cross corelation coefficient
Cl	chlorid
Clp	*E. coli* Hsp100
CMV	cytomegalovirus stalling sequence
CORAN	correspondence analysis
CPU	central processing unit
cryo-EM	cryogenic electron microscopy
CTF	contrast transfer function
CV	column volume
D	diffusion coefficient
Da	1 g / mol
DNA	deoxyribonucleic acid
DnaK	*E. coli* Hsp70
dNTP/NTP	(deoxy) nucleotide triphosphate
DPAP	dipeptidyl aminopeptidase B
e^-	Electron
Egd1	*S. cerevisiae* βNAC
Egd2	*S. cerevisiae* αNAC
F	fluorescence
FA	fluorescence anisotropy
FCMR	focused competitive multi reference alignment
FCS	fluorescence correlation spectroscopy
fL	10^{-15} L
FSC	Fourier shell correlation
fwd	forward
G	Gibbs energy
Gal1/Gal10/Gal4	components of the Gal3 complex

GatD	galactitol-1-phosphate dehydrogenase
GroEL/ES	*E. coli* Hsp60
H	enthalpy
His$_{6/8}$-tag	6x/8x histidin tag
HPLC	high performance liquid chromatography
HspX	Heat shock protein X
HtpG	*E. coli* Hsp90
I	Intensity
IR	Infra-red
ITC	isothermal titration calorimetry
j	backdiffusion
K	potassium
K$_d$	dissociation constant
kDa	1,000 Dalton
kpsi	kilo pounds per square inch
LB	lysogeny broth
LED	light emitting diode
LSU	large ribosomal subunit
Lx	Large ribosomal subunit protein X
MBP	maltose binding protein
MDa	10^6 Da
mg	10^{-3} g
Mg	Magnesium
min	minute
mRNA	messenger RNA
MST	microscale thermophoresis
n.A.	not applicable
Na	sodium
NAC	nascent polypeptide associated complex
NBD	*N*-((2-(iodoacetoxy)ethyl)-*N*-Methyl)amino-7-Nitrobenz-2-Oxa-1,3-Diazole
NC	nascent polypeptide/chain
NEB	new england biolabs
ng	10^{-9} g
NHS	N-Hydroxysuccinimide
nM	10^{-9} M
NTA	nitrilotriacetic acid
OAc	acetate
OD$_{xxx}$	optical density at x nm
PCA	principal component analysis
PCC	protein conducting channel
PCR	polymerase chain reaction
pdb	protein data bank
PDF	peptide deformylase
pH	decimal logarithm of proton concentration
PPIase	prolyl-peptidyl cis-trans isomerase

PTC	peptidyl transferase center
PVDF	polyvinylidenfluorid
RAC	ribosome associated complex
RBD	ribosome binding domain
rev	reverse
RNA	ribonucleic acid
RNC	ribosome nascent polypeptide complex
rpm	revolutions per minute
rRNA	ribosomal RNA
RT	room temperature (23 °C)
S	Soret coefficient
s	second
SDS-PAGE	SDS-polyacrylamid gel electrophoresis
skNAC	skeletal αNAC isoform (*Mus musculus*)
SNR	Signal-to-noise ratio
SR	signal recognition particle receptor
SRP	signal recognition particle
SS/SA	signal sequence/anchor
Ssb1/2	cytosolic *S.cerevisiae* Hsp70
ssDNA	salmon sperm DNA
SSU	small ribosomal subunit
TCA	trichloroactic acid
TEV	tobacco etch virus
TF	Trigger Factor
T-jump	temperature jump
TnaC	tryptohanase leader peptide
tRNA	transfer RNA
tryp	tryptone
UAS1/2	universal adaptor site 1/2
v	molecule velocity
V	volt
YENB	yeast essential nutrient broth
YPD	yeast pepton dextrose
30S	prokaryotic SSU
40S	eukaryotic SSU
50S	prokaryotic LSU
60S	eukaryotic LSU
70S	prokaryotic ribosome
80S	eukaryotic ribosome

1   Introduction

1.1   Protein biosynthesis at ribosomes and folding of proteins

The biosynthesis of proteins is conducted by the ribosome. Ribosomes are macromolecular complexes which are composed mainly of ribosomal RNA (rRNA, 2/3 of the total mass) but also contain a variety of proteins. Ribosomes consist of two subunits: the large (LSU) and the small subunit (SSU). The LSU harbors the peptidyl transferase center where the central biochemical reaction, the formation of the new peptide bond, takes place. The small subunit (SSU) accommodates the mRNA and contains the mRNA decoding activity. Prokaryotic ribosomes are typically smaller (e.g. *Escherichia coli (E. coli)*: 2.4 MDa) than eukaryotic ribosomes (e.g. *Saccharomyces cerevisiae* (*S. cerevisiae*): 3.3 MDa).

The shape of the subunit is defined by rRNA molecules (Figure 1). The LSU contains two rRNAs in prokaryotes (5S, 23S) and three rRNAs in eukaryotes (5S, 28S, 5.8S) while the SSU always contains one rRNA (prokaryots: 16S, eukaryots: 18S). The rRNA architecture of archaeal ribosomes resembles that of the prokaryotic ribosomes on the basis of composition (LSU: two rRNAs (23S, 5S), SSU: 16S rRNA), its structure however resembles more the eukaryotic ribosome (Londei, 2010).

The peptidyl transferase reaction is performed by an integral ribozyme of the ribosome. The ribosome is a remarkably precise enzyme whose error rate is between $2x10^{-4}$ to $6x10^{-6}$ in eukaryotes (Stansfield et al., 1998). This precision comes at a cost. The speed of translation under optimal conditions was measured to range from only 20 amino acids per second (prokaryotes, Dennis and Nomura, 1974) to as few as five (eukaryotes, Olofsson et al., 1987) .

The number of ribosomal proteins varies from 54 in *E. coli* to 80 in humans (Melnikov et al., 2012). Recently, a common nomenclature for ribosomal proteins of all kingdoms of life was proposed that is used throughout this work (Ben-Shem et al., 2011).The ribosomal proteins are located mostly at the surface of the ribosome; however often contain extensions into the core of the ribosome. Amongst other functions, ribosomal proteins provide specific binding sites for ribosomal interactors.

The location of the peptidyl transferase center (PTC) in the interior of the ribosome raised the question, how the nascent polypeptide finally leaves the ribosome. Biochemical studies showed resistance of nascent polypeptides against proteolytic digestion (Malkin and Rich, 1967; Blobel and Sabatini, 1970). This and the discovery that the LSU contained a central tunnel suggested an model

where a tunnel in the ribosome leads the nascent chain from the PTC to the cytosol (Milligan and Unwin, 1986; Yonath et al., 1987) The localization of the PTC by X-ray crystallography directly at the beginning of this tunnel and ultimately the visualization of nascent polypeptides in the tunnel using high resolution cryo-EM proved this model right (Ban et al., 2000; Seidelt et al., 2009). While the majority of the tunnel surface is composed of rRNA (Lu et al., 2007), there is one region where two ribosomal proteins (L4 and L22) contact the nascent polypeptide and constrict the tunnel diameter relative to the upstream and downstream diameter (Ban et al., 2000).
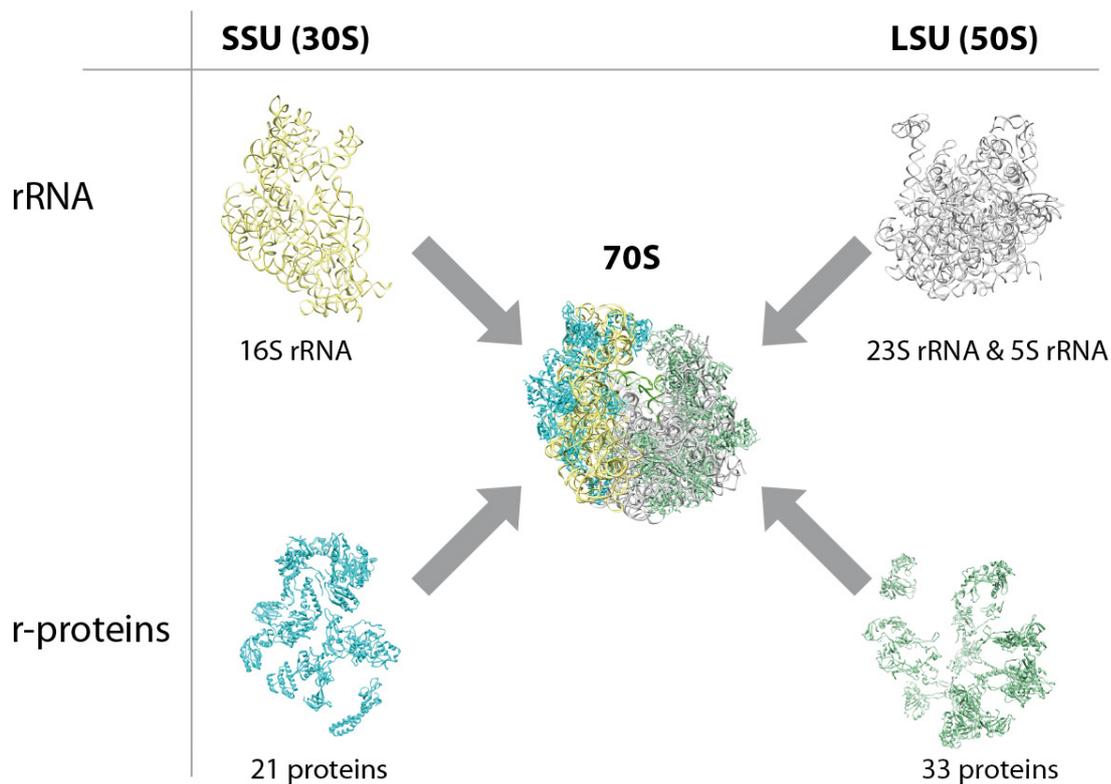
Figure 1: The *E. coli* ribosome is a prototypical prokaryotic 70S ribosome | Its 2.4 MDa mass is composed of three different rRNAs (16S, 23S and 5S rRNA) which make up 2/3 of the total mass and 54 protein s(Melnikov et al., 2012). pdb: LSU: 2wwq, SSU 2wwl, (Seidelt et al., 2009)

Biosynthesis of polypeptides on ribosomes leads to the vectorial synthesis of polypeptides. However, the functional state of proteins is a defined three dimensional fold of the polypeptide chain. The first tier of folding is the formation of secondary structures, α-helices and β-sheets that then interact to yield the final fold of single polypeptide chains, the tertiary structure.

The folding of polypeptides was experimentally investigated using chemically denatured proteins. Concepts that emerged from these *in vitro* studies were soon confronted with the mathematical impossibility of sampling all interactions (Levinthal, 1969). This led to the discovery of folding

landscapes that describe the formation of local low energy folds and the progression to the final fold (reviewed in Bryngelson et al., 1995 and Brunori et al., 2012).

The finding that dedicated proteins called chaperones further shape the folding pathways and the efficiency of protein folding towards physiological relevant yields and speed enhanced the understanding of protein maturation further. Chaperones play a major role in the formation of all the tiers of folding. They are proteins that either delay the folding of structures to prevent premature or unspecific interactions (holdases) or help in folding (foldases). Both actions optimize or are even necessary for high yielding protein folding.

The term molecular chaperones was used the first time when a protein was identified that prevented the aggregation of single histone proteins (Laskey et al., 1978). As chaperones are often upregulated under stress situations like increased temperature, this feature and the size of the induced protein product gave rise to their classification: Heat shock protein 70 (Hsp70) etc. There are five main classes of chaperones: Hsp20 (alpha-crystallin), Hsp60 (GroEL), Hsp70 (DnaK), Hsp90 (HtpG) and Hsp100 (Clp family). While most of the chaperones act post-translationally on polypeptides, some of these chaperones exert their function conjoined with the ribosome. Two of these chaperones will be discussed later (section 1.3 and 1.4).

The initiation of the folding process starts already in the tunnel. Principally, the diameter of the tunnel of the ribosome (10-20 Å) sterically allows the formation of α-helices but not that of β-sheets and tertiary structure elements. Indeed, FRET studies showed that the nascent polypeptide in the tunnel cannot only form α-helices but that this formation is induced by the ribosome (Woolhead et al., 2004). Further experiments showed that there are regions in the tunnel that allow formation of α-helices while other regions prevent this folding of the nascent polypeptide (Lu and Deutsch, 2005). The constriction created by L4 and L22 coincidences with this region. Finally, Cryo-EM studies succeeded to visualize the presence of α-helices in the nascent chain within the tunnel (Bhushan et al., 2010a).

1.2 The ribosome is more than a static platform for nascent polypeptide interactors

The nascent polypeptides that emerge from the ribosome are not in their functional state yet. A multitude of processing events modifying the nascent polypeptides and help them to obtain their final functionality. Additionally, proteins such as chaperones that delay folding to help reach the final fold, or proteins that define translocation pathways have to interact early with the nascent polypeptide to prevent the progression into non-productive pathways.

In prokaryotes for example, the charged tRNA that is used to decode the start AUG is a modified tRNA, the initiator tRNA (Marcker and Sanger, 1964, Adams and Capecchi, 1966). To insure proper translation and quench unfavorable side effects, its reactive N-terminal amino group is protected by formylation (Eisenstadt and Lengyel, 1966; Harvey, 1973). This formylated methionine is incorporated in the N-terminus of the nascent polypeptide. After emergence from the ribosome, it is removed directly at the ribosome by the enzyme peptidyl deformlyase (PDF) (Housman et al., 1972; Ball and Kaesberg, 1973).

To achieve proper access to the nascent polypeptide, divergent and non related proteins evolved similar strategies. The ribosomal tunnel exit is surrounded by proteins out of which four are conserved between the domains of life (L22, L23, L24, L29). Additionally, the region around the tunnel exit of eukaryotes contains additionally L19e, L31e and L39e (Ban et al., 2000; Harms et al., 2001; Lecompte et al., 2002).

A subset of these proteins is used for the recruitment of factors that mediate functions on the nascent polypeptide. Initially, a defined *locus* around the tunnel exit was shown to be important for this recruitment: the universal adaptor site (Figure 2). Its components L23 and L29 mediate the interaction between SRP, the translocon Sec61 (in prokaryotes also YidC and Oxa1), Erj1p, at least one co-translational chaperone – trigger factor (TF) – and the ribosome (Beckmann et al., 2001; Kramer et al., 2002; Halic et al., 2004; Blau et al., 2005; Bingel-Erlenmeyer et al., 2008; Kohler et al., 2009;).

Later, the binding sites of the ribosomal interactors SR, RAC and NAC where identified. They clustered around L22/L31e (Halic et al., 2006b; Peisker et al., 2008; Pech et al., 2010) (Figure 2). Pech et al. therefore suggested the presence of a new universal adaptor site on the ribosome, the UAS 2. Its conservation between the domains of life is illustrated by the fact that the ribosome binding domain of *E. coli* peptide deformylase (PDF) – a protein that only exists in prokaryots also interacts

with L22 (Bingel-Erlenmeyer et al., 2008). The way in which these interactors use the interaction with the ribosome to fulfill their function is not fully understood for all of them.
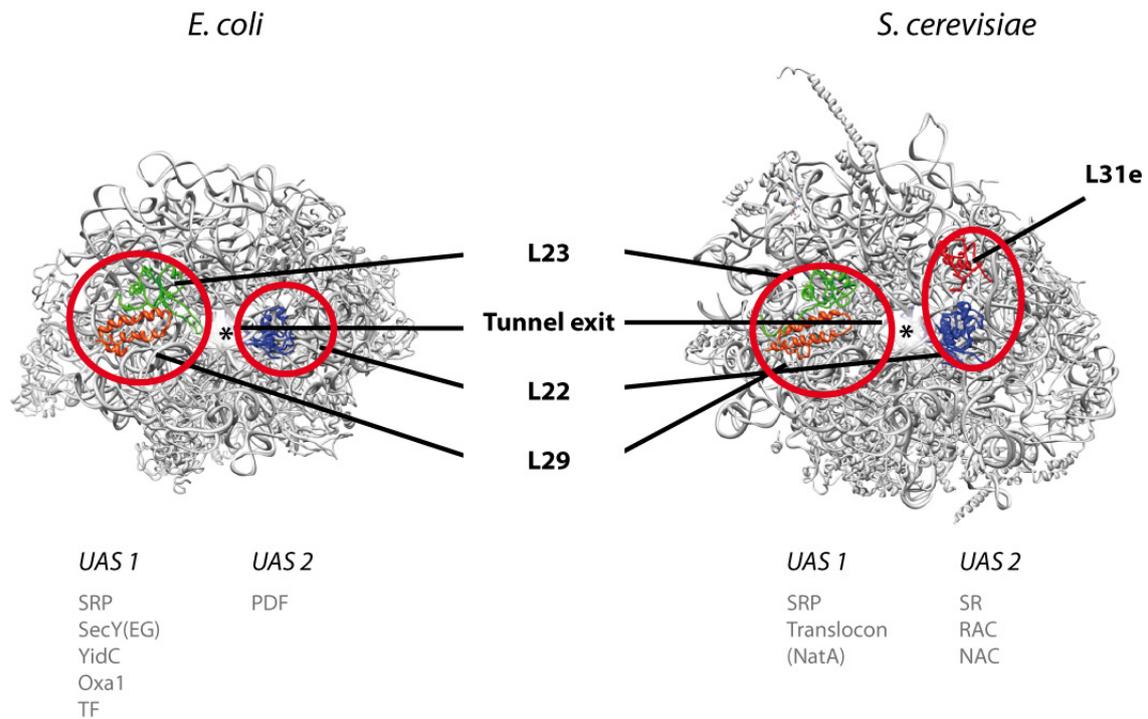
Figure 2: The universal adaptor sites 1 and 2 allow the interaction of ribosomal interactors and the nascent polypeptide | L23 and L29 provide a binding platform for SRP, SecYEG (Sec61), YidC/Oxa1 and TF. L17 and L31e are located on the opposite site of the tunnel exit and provide a second adaptor site for SR, NAC and RAC. The asterisk indicates the tunnel exit. The rRNA and proteins that are not part of one of the UAS' are colored in grey. L29 is colored in red-orange, L23 in green, L22 in blue and L31e in red. (Pech et al., 2010) developed the concept of the universal adaptor site 2. (*E. coli* pdb: 2wwq (Seidelt et al., 2009), *S. cerevisiae* pdb: 3o58 (Ben-Shem et al., 2011))

One critical step in the maturation of the polypeptide is the targeting to its final destination. In prokaryotes, some proteins have to cross the cell membrane. The situation in eukaryotes is more complex. Their cells are highly compartmentalized environments. The compartments are separated by one or more membranes that have to be crossed to reach certain compartements.

The translocation of nascent polypeptides into the ER-lumen (eukaryotes) or across the cell membrane (pro- and eukaryotes) using the protein conducting channel (PCC) is the best understood pathway that is initiated by ribosomal interactors. Proteins that are either secreted, integrated into the membrane or that have to be translocated into the lumen of the ER use this path (Rapoport, 2007).

Nascent polypeptides can either co- or post-translationally pass the PCC. During the co-translational translocation, the ribosome aligns its ribosomal tunnel with the channel of the PCC in a way that the elongation of the nascent polypeptide provides the necessary driving force for the nascent polypeptide to irreversibly cross the compartment boundary (Beckmann et al., 1997).

During posttranslational translocation, the driving force is either ATP hydrolysis by SecA (prokaryotic system) or diffusional trapping using the luminal chaperone BiP (a eukaryotic Hsp70 chaperone).

The co-translational translocation requires a precise orchestration of events from translation to translocation. The signal in proteins that decides in favor for the co-translational pathway is a N-terminal hydrophobic signal sequence or signal anchor (SS/SA) flanked by a downstream positive charge. Its emergence from the ribosome leads to a recruitment of the signal recognition particle (SRP) to the ribosome (Halic and Beckmann, 2005). The SRP interacts both with the SS/SA and the ribosome. In eukaryotes, this interaction reduces the translation rate and thus increases the time window for the recruitment of the SRP-RNC complex to the SRP-receptor (Mason et al., 2000; Lakkaraju et al., 2008).

To increase the likelihood of SRP to find its substrate, the ribosome serves as a mediator between the nascent polypeptide and SRP. Even before the SS/SA emerges from the ribosome, it increases the affinity for SRP to ribosomes in a yet unknown way and recruits it to the ribosome (Bornemann et al., 2008; Berndt et al., 2009).

The same behavior was also found for the interaction of TF to ribosomes. TF – a co-translational chaperone - was found to interact via L23 and L29 with the ribosome (see also section 1.4). It interacts with mainly unfolded nascent polypeptides. SRP competes with TF for access to SS/SA that emerge from the ribosome. Recent FRET and immuno-detection studies showed, that the presence of α-helix formation in the tunnel reduces the affinity of TF for ribosomes, before the nascent polypeptide emerges from the ribosome (Lin et al., 2012). Thus, the high concentration of TF in comparison to the low SRP concentration is counteracted by this intriguing mechanism to ensure proper access of all key players to nascent polypeptides.

These examples show that the ribosome is not only a static platform that passively allows the access of ribosomal interactors to the nascent polypeptide. On the contrary, it transfers information from the inside of the ribosomal tunnel to the surface in a yet unknown manner and thus optimizes the access of cytosolic and membrane located interactors to fine tune the basic biochemical activities associated with the growing nascent polypeptide. On top of the information transmission by the ribosome, the presence of two UAS principally allows co-binding and thus a functional interplay of more than one ribosomal interactor. The SRP-SR-PCC system is the prime example of this.

In a first step, SRP binds via L23/L29 (UAS 1) to the ribosome (Figure 3, panel C.). It either probes for the emergence of the SS/SA (see above) or directly interacts with the already emerged SS/SA. The interaction between the SS/SA and the SRP54 M-domain stabilizes this interaction further. The RNC-

SRP complex then diffuses to the membrane associated SRP-receptor SR (Halic et al., 2004). The SR interacts with the UAS 2 (via L31e) and SRP. The SRP moiety that was previously blocking the PCC interaction site (SRP54) via its interaction with the ribosome, gets rearranged and the PCC can now bind via L23 to the ribosome (Halic et al., 2006b, Figure 3 Panel D).
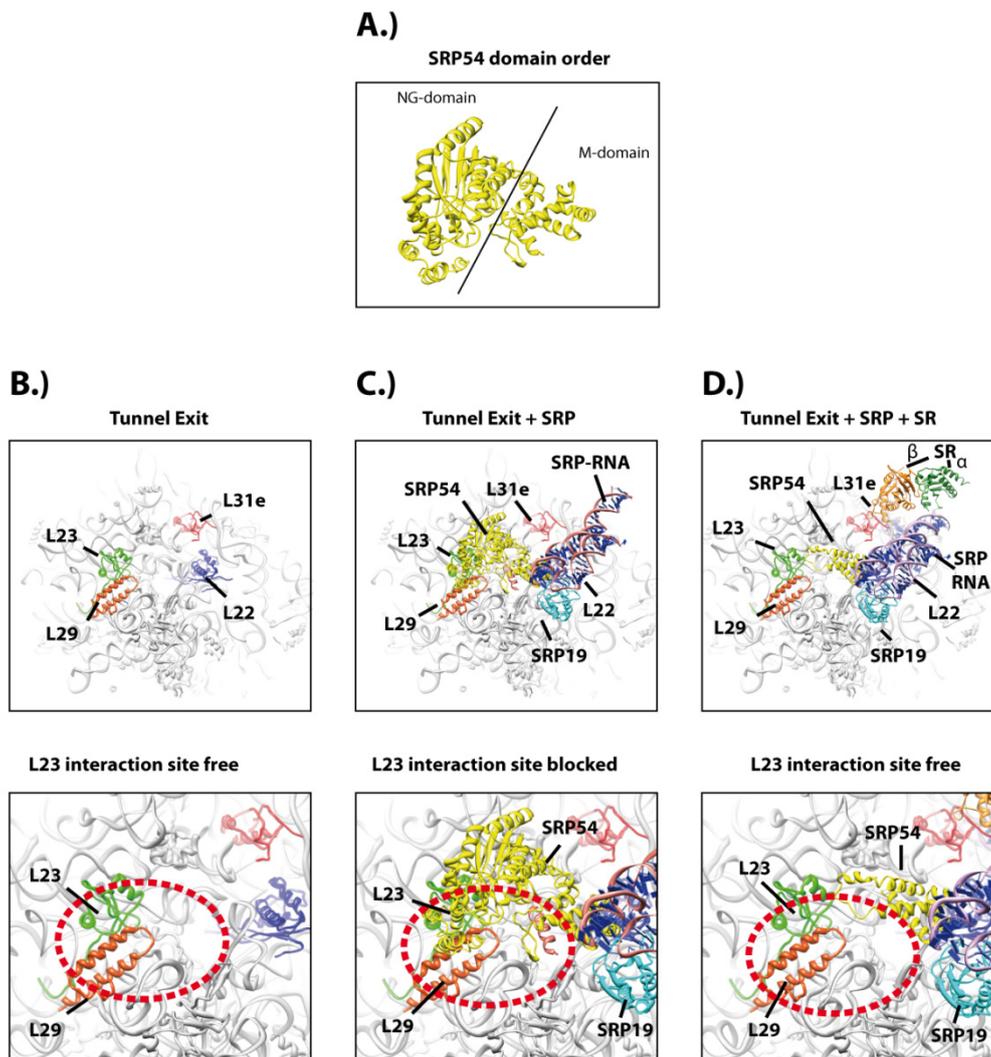
Figure 3: The SRP-SR interplay on the ribosome blocks and releases the UAS 1 interaction site for Sec61 | The molecular models of the *S. cerevisiae* ribosome (pdb: 3o58), the eukaryotic SRP-ribosome complex (pdb: 2go5) and the eukaryotic SRP-SR-ribosome complex (pdb: 2j37) illustrate the sequential blocking and release of the translocon interaction site (lower pictures, red circle). Panel A.): SRP54 consists of a NG and a M-domain. Panel B.) shows the access of the UAS1 in absence of SRP. Panel C.): Upon ribosome binding, the SRP binds to the ribosome and interacts via the NG domain with L23. Thereby, access to the UAS1 is restricted. Panel D.): In the RNC-SRP-SR complex, the NG-domain of SRP54 is not binding to the UAS 1 anymore. Cryo-EM therefore failed to fully resolve SRP54. Only the M-domain of SRP is still in direct contact with the ribosome. The UAS1 is now available for the interaction with the translocon. (Halic et al., 2004; Halic et al., 2006b; Ben-Shem et al., 2011) (green: L23, orange: L29, blue: L22, red: L31e, yellow: SRP54, cyan: SRP54, dark green: SR α, light orange: SR β).

This example shows the benefit of a complex model that allows interplay between different ribosomal interactors as opposed to a linear progression of events.

1.3 NAC: an abundant and poorly understood central ribosomal interactor

The nascent polypeptide associated complex (NAC) is a dimeric complex that interacts with ribosomes. Its stoichiometry to ribosomes ranges from 1:1 in *S. cerevisiae* up to a 10 fold excess in reticulocytes (George et al., 1998; Moller et al., 1998; Reimann et al., 1999). This high concentration makes it unique among the eukaryotic ribosomal interactors that are normally less abundant than the ribosome.

The discovery of NAC

In 1994, crosslinks between short nascent polypeptides that are still attached to eukaryotic ribosomes and an unknown protein complex were detected (Wiedmann et al., 1994). Since the release of the nascent polypeptide from the ribosome before crosslinking abolished this interaction, the significance of the ribosome interaction of this complex was clear.

Due to its interaction with the nascent polypeptide, the name nascent polypeptide associated complex (NAC) was used to describe this function. Using intensive biochemical fractionation, the complex was purified and characterized from bovine brain material (Wiedmann et al., 1994). Eukaryotic NAC consists of two subunits: αNAC and βNAC. Sequence comparison revealed, that in *S. cerevisiae* and humans NAC homologues were previously found in studies focusing on transcription regulation.

In *S. cerevisiae*, Egd1p (βNAC homologue) and Egd2p (αNAC homologue) were found to enhance the stability of the Gal4p-DNA complex (Parthun et al., 1992; Shi et al., 1995). Their names are acronyms of this function: Enhancer of Gal4 DNA binding 1 and 2. If this function was vital for the Gal4-DNA complex, their deletion should interfere *in vivo* with the responsiveness of cells to galactose.

However, the deletion of EGD1 alone did not abolish the transcriptional reaction of *S. cerevisiae* cells to galactose. It did however reduce the magnitude of the transcription response (Parthun et al., 1992). Later, a second βNAC subunit was found in *S. cerevisiae*: basic transcription factor three 1 (BTT1), (Hu and Ronne, 1994). In contrast to Egd1p, the protein levels of Btt1p are at least 100 fold lower. Deletion experiments of both βNAC derivatives in yeast demonstrated, that the double knockout showed a strong increase in the concentration of Gal1 and Gal10 mRNAs (Parthun et al., 1992). This was the opposite effect as expected.

Later it was found that the βNAC double deletion showed a reduced growth at 37°C in *S. cerevisiae*. The discovery that Egd1p and Egd2p form a complex and that the phenotype of the double deletion

strain of the βNAC subunits can be reverted to a wildtype phenotype by additional deletion of αNAC indicated a toxic effect of the αNAC subunit alone (Reimann et al., 1999). In a *S. cerevisiae* EGD1 deletion strain, BTT1 is indeed upregulated to compensate for the toxicity of free αNAC.

In human HeLa cells, βNAC was identified as BTF3 (basic transcription factor 3, (Zheng et al., 1990)). There are two known splicing variants of BTF3: BTF3a (β_2-NAC) and BTF3b (β_1-NAC) (Kanno et al., 1992). BTF3a contains at the N-terminus 44 additional AA. While both variants bind to RNA-Polymerase II, only BTF3a was found to play a role in the initiation of transcription (Kanno et al., 1992; Moncollin et al., 1986; Zheng et al., 1990; Zheng et al., 1987). This function was investigated carefully. A study that identified the minimal components necessary for transcription using the mammalian RNA-Polymerase II concluded, that BTF3a is not necessary for this function (Flores et al., 1992). Later, a skeletal isoform of αNAC (skNAC) was identified in mice that binds specifically to DNA sequences containing a consensus sequence (Yotov and St-Arnaud, 1996). While the function of this protein is unkown, this finding adds to the notion, that the NAC subunits have nuclear functions.

Despite intensive efforts, no NAC homologue was identified so far in eubacterias. In archaea however, genome wide studies indicated the presence of an αNAC ortholog (Makarova et al., 1999). A thorough biochemical analysis of aeNAC showed, that aeNAC behaves like eukaryotic NAC. It can be crosslinked to nascent polypeptides and competes with eukaryotic NAC for the same binding site (Spreter et al., 2005). The discovery of aeNAC led to the first crystal structure of a NAC (see section "The structure of NAC").

NAC's interaction with the ribosome

NAC's contribution to the regulation of transcription is still not well understood. However, the discovery that NAC mainly resides as a heterodimer in the cytosol stimulated the research on its cytosolic function (Reimann et al., 1999; Beatrix et al., 2000).

An analysis of human NAC in 2000 revealed, that the α- and βNAC subunits have different functions (Beatrix et al., 2000). While both subunits can be crosslinked to the nascent polypeptides, only the βNAC subunit interacts with ribosomes and 60S subunits. Based on this observation, the region of NAC that is responsible for the interaction with the ribosome could be determined. The deletion of the N-terminal 11 AA of βNAC prevents the cosedimentation of NAC with the ribosome (Franke et al., 2001; Pech et al., 2010). However, the first 23 AA are necessary to tether non-ribosome associated proteins like MBP efficiently to ribosomes. Additionally, a conserved motif in the N-terminal region consisting of the consensus motif $RRK(X)_nKK$ was proposed. Its mutation leads to a loss of ribosome interaction in pelleting assays (Wegrzyn et al., 2006). As this motif is located far

away of the portable ribosome location signal, its significance is under debate. It was suggested that its mutation by using AA-exchange to alanines forms an α-helical region in a physiological intrinsically disordered region and thus interferes with the behavior of the whole complex (Pech et al., 2010).

NAC can be crosslinked to amino acids of nascent polypeptides that are only located 17 AA away from the PTC (Wang et al., 1995). This argues for an interaction of NAC in close proximity to the tunnel exit as otherwise the small NAC subunits (Egd2p: 18,7 kDa and Egd1p: 17 kDa) could not reach that far into the tunnel. Crosslinking studies in *S. cerevisiae* indeed showed, that NAC can be crosslinked to L31e (Egd1p) and L22 (Egd2p), components of the UAS2 (see also section 1.2, (Pech et al., 2010)). Additionally, deletion of L31e leads to a complete loss of co-sedimentation of NAC with ribosomes.

These results are in contradiction with the observation that NAC interacts with *E. coli* ribosomes via L23 (Wegrzyn et al., 2006). This study also showed crosslinks between NAC and L23 in *S. cerevisiae*. The usage of N-terminally his-tagged αβNAC fusion protein in this study might explain this finding since it was later showed, that his-tagged NAC cosediments with *E. coli* ribosomes. Indeed, the removal of the tag abolishes this interaction (Pech et al., 2010).

The N-terminal part of Egd1p has a high propensity for α-helix formation. It has highly conserved amino acids. Especially positively charged amino acids clustered at one side of the predicted helix are conserved (see also Figure 21 p. 71 for a model of the helix). Other than skNAC, *S. cerevisiae* NAC was shown to interact with nucleic acids (DNA or RNA) irrespective of the sequence indicating a possible contribution of the rRNA to the interaction of NAC and the ribosome (Beatrix et al., 2000).

The structure of NAC

The X-ray structure of aeNAC revealed the presence of a new central domain (Spreter et al., 2005). This central domain, called NAC domain, is present in all NAC species and responsible for the dimerization of the α- and the βNAC subunit (Figure 4). So far, no structure for full length heterodimeric NAC is available. The regions of NAC that connect the functional elements are mainly intrinsically disordered which did not allow for crystal formation. The discovery of archaeal orthologs of αNAC (aeNAC) led to the first crystal structure of aeNAC (Spreter et al., 2005, Figure 6). In aeNAC, the intrinsically disordered regions between the NAC domain and the UBA-domain are shorter compared to eukaryotic αNAC. This allowed for the formation of crystals of aeNAC lacking only the N-terminal 18 AAs.

EGD1 (β-NAC)
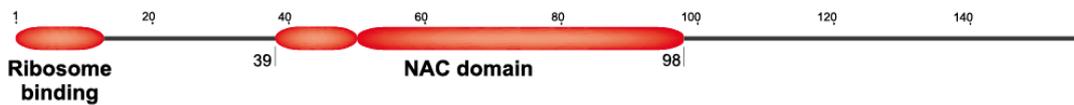
EGD2 (α-NAC)

Figure 4 : *S. cerevisiae* **NAC is composed of Egd1p and Egd2p. The dimerization between both subunits is achieved by the novel NAC-domain |** Identified domains (NAC and ubiquitin associated domain (UBA)) and functional elements of the *S. cerevisiae* NAC subunits are schematically depicted. The central NAC domain (~ 60AA) mediates the dimerization of the two subunits. The regions that connect the domains are mainly intrinsically disordered (figure: PhD-Thesis Markus Pech 2005)

archaeal NAC (aeNAC homodimer)

human NAC-domain (NAC heterodimer)
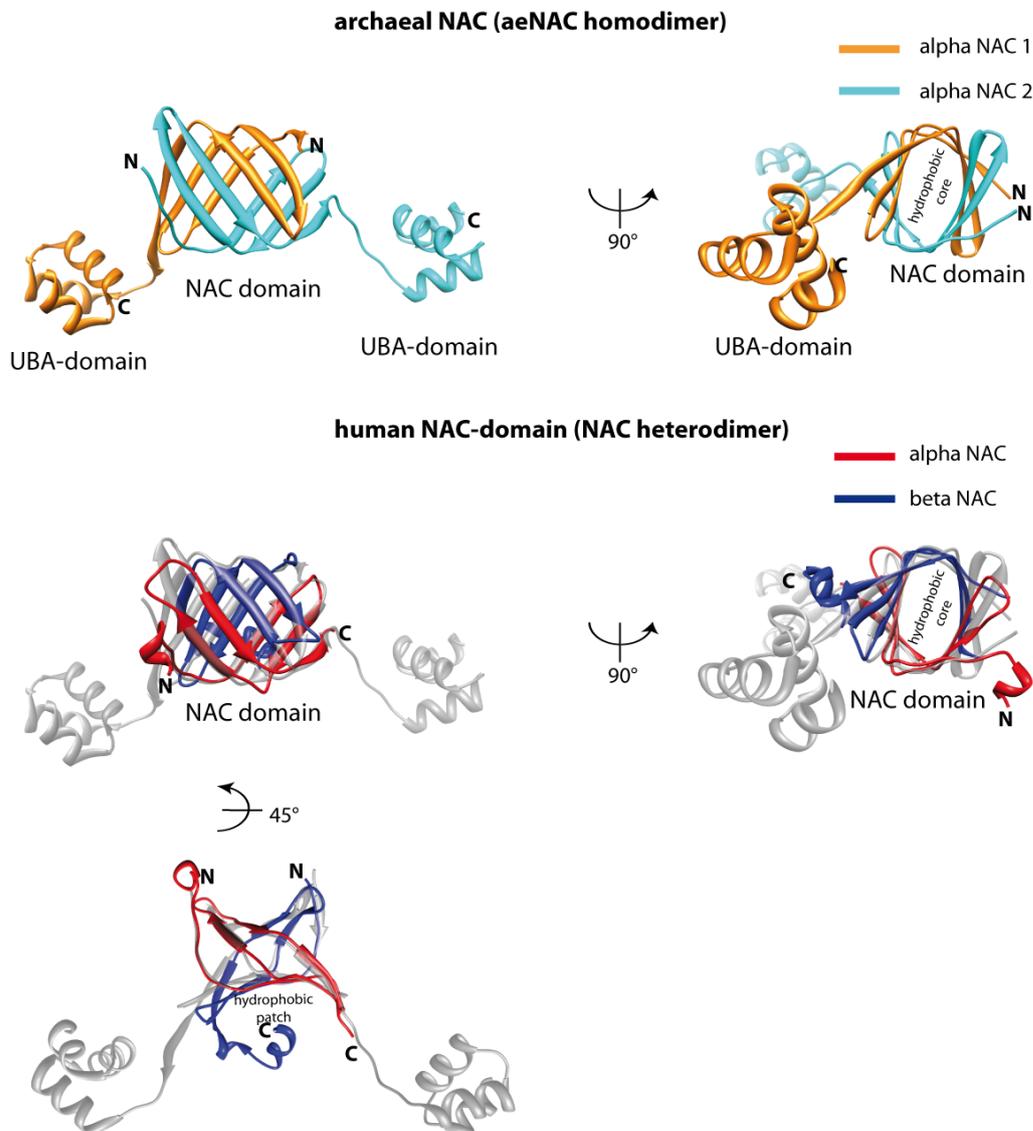
Figure 5 : The central NAC domain posses a unique β-barrel fold that surrounds a hydrophobic inner core and thereby increases stability significantly | The X-ray structure of homodimerac aeNAC (pdb: 1tr8; Spreter et al., 2005) unrevealed for the first time the fold of the NAC domain. A central, β-barrel rich region dimerizes and creates a hydrophobic core. This

explains the unusual stability of this complex (pdb: 2tr8). Later, the structure of the human NAC domain was published (Wang et al., 2010, pdb: 3mcb). The direct comparison of the NAC domains (red: human αNAC, blue: human βNAC, transparent: aeNAC) illustrates the remarkable conservation of the fold of the NAC domain. Additionally, the C-terminal region of the βNAC forms an α-helix that folds back on the surface of the NAC domain and creates a hydrophobic patch.

The NAC domain adopts a unique fold that is composed of four β-sheets. In the aeNAC homodimer, the β-sheets follow a regular pattern that creates a hydrophobic core. This core explains the high stability of this dimerization. The C-termini of the subunits fold into canonical UBA-domains that contain a hydrophobic patch (Figure 5). The high conservation between aeNAC and eukaryotic NAC allowed the formation of a homology model for the *S. cerevisiae* NAC (Spreter et al., 2005).

Later, the X-ray structure of the human NAC domain of heterodimeric αβNAC was published (Wang et al., 2010). As predicted, the NAC domain is highly conserved and hardly showed any deviation of the peptide backbone between aeNAC and human NAC. In human NAC, the C-terminal part of βNAC forms an α-helix that interacts with the surface of the NAC domain and creates a hydrophobic patch.

NAC's function at the ribosome

Accumulating experimental evidence suggests a cytoplasmic role of NAC in several cellular pathways. Directly after its discovery, *in vitro* studies showed that NAC prevents the inappropriate targeting of ribosome nascent polypeptide complexes (RNCs) without signal sequences to the PCC by blocking the interaction site of the PCC and the ribosome (Moller et al., 1998). NAC was also shown to regulate the access of SRP to nascent polypeptides and thereby ensuring the fidelity of SRP dependent targeting (Wiedmann et al., 1994). However, the verification of these observations *in vivo* is challenging, as central cellular networks dealing with key functions are often robust and redundant (Reimann et al., 1999).

Therefore, investigations of NAC's function in context of a reduced chaperone network were conducted (Koplin et al., 2010). By deletion of SSB1/2, a ribosome bound chaperone network was silenced. This network consists of the ribosome associated complex (RAC) and Ssb1/2p. RAC is a complex of an Hsp70 (Ssz1) and Hsp40 (Zuo1) homologue that interacts with Ssb1/2p (a Hsp70 homologue) and stimulates its ATPase activity (Gautschi et al., 2002). Additionally, it brings Ssb1/2p in proximity to the nascent polypeptide as RAC interacts with the ribosome. The deletion of the two SSB genes (SSB1 and SSB2) creates a pleiotropic phenotype. Yeast cells are hypersensitive to high salt concentrations, low temperatures, translation inhibitors like aminoglycosides and a wide range of cations. The importance of this system for proper protein production and folding is illustrated by the fact that deletion of SSB1/2 severely increases the fraction of aggregated proteins *in vivo* (Koplin

et al., 2010). Ssb1/2p's tethering to ribosomes is necessary for its function as the same phenotypes are observed by deletion of either of the RAC subunits (Gautschi et al., 2002; Hundley et al., 2002; Kim and Craig, 2005; Pfund et al., 1998). Furthermore, deletion of SSB1/2 lead to problems in ribosomal biogenesis (Albanese et al., 2010).

Interestingly, the penetrance of these phenotypes is enhanced if all three NAC subunits are additionally deleted (Koplin et al., 2010). Especially protein aggregation and a reduction of growth rate are significantly enhanced indicating a role of NAC and RAC in common pathways. The cellular pathways that are disturbed in the absence of NAC however are not clear.

NAC's functional interplay with the SRP targeting systems is still under investigation. A recent study showed that NAC, contrary to earlier *in vitro* studies, did not prevent interaction of SRP to all signal-less RNCs. Instead, it supported the early recruitment of SRP to RNCs whose signal sequences have not emerged from the ribosome yet (see also section 1.2; Zhang et al., 2012). This study also provided evidence for a trimeric complex of the ribosome containing a nascent polypeptide (RNCs), SRP and NAC. Moreover it showed, that in the absence of SRP, NAC can directly interact with signal sequences containing nascent chains in a stable, salt sensitive manner. The authors concluded that the observation of a trimeric complex of SRP, NAC and RNCs is only possible if both complexes bind on different regions on the ribosome as otherwise sterical clashes would destabilize this complex. This supports the finding that NAC interacts with the UAS2 of the ribosome.

A global screen of mRNAs associated with ribosomes in context of different ribosomal interactors (SRP and the NAC subunits) showed, that NACs specificity is wide and it is overlapping with the SRP interactome (del Alamo et al., 2011). A close inspection of the NAC subunits showed divergent preferences of the individual subunits. More importantly, the interactome of SRP is significantly modulated in the absence of NAC. While some SRP interactors are still bound to SRP even in the absence of NAC, others failed to do so. Interestingly, the hydrophobicity of the signal sequence does not correlate with the fate of the substrate. Finally, a different subset of proteins that is not interacting with SRP in wild type cells artificially interacts with SRP in the absence of NAC.

Of particular interest in this regard is the model substrate DPAP, as it is widely used as a SRP substrate to investigate the process of protein translocation. This protein falls into the category of SRP interactors whose *in vivo* association to SRP is NAC dependent. These observations might allow a new interpretation of previous biochemical observations. They show that the function of NAC in regulation of SRP-nascent polypeptide interaction is dependent on the identity of the nascent polypeptide.

However, despite the strong influence of the NAC deletion on SRP-binding, no phenotype in the translocation of nascent polypeptides or activation of the unfolded protein response was detected. Interestingly, the cell reacts with an increase in the folding capacity, especially the Ssb1/2p chaperones are upregulated. This supports the idea of a functional interplay between NAC and Ssb1/2p (del Alamo et al., 2011).

The contribution of NAC to the regulation of translocation to the ER and folding of nascent polypeptides is further suggested by the observation that down-regulation of αNAC in human cells induces an ER stress response. Reduction of αNAC using siRNA leads to an increase of ubiquitination of proteins and an increased BiP concentration in the ER-lumen, which is a sign for global misfolding in the cell (Hotokezaka et al., 2009). Finally, the same study suggests a role of αNAC in the regulation of apoptosis of cells in response to hypoxic stress.

Additionally, a NAC dependent novel translation coupled mitochondrial import mechanism based on *in vitro* studies was proposed (Fünfschilling and Rospert, 1999; George et al., 1998). Later, *in vivo* studies using fumarase as a reporter protein confirmed the contribution of NAC to this import mechanism (Yogev et al., 2007).

1.4 Trigger Factor: the first chaperone that interacts with the nascent polypeptide in prokaryotes

Trigger Factor's discovery, its structure and its binding mode with the ribosome

In eukaryotes, NAC is the first interactor of the nascent polypeptide with a broad range of functions. Later, a ribosome attached triad of Hsp40 and Hsp70 chaperones acts as the first dedicated chaperone system. In eubacteria, TF is the first interactor with the growing nascent polypeptide chaperoning the fate of the nascent polypeptide. It was originally identified as a factor that prevented folding of outer membrane proteins, thereby keeping them in a translocation competent form (Crooke and Wickner, 1987; Crooke et al., 1988a; Crooke et al., 1988b). Nevertheless, it soon became evident that its function is initiated by its interaction with the ribosome (Lill et al., 1988).

TF consists of three domains that fold independently: A N-terminal domain, a PPIase domain and a C-terminal domain (Figure 6; Hesterkamp and Bukau, 1996; Hesterkamp et al., 1997; Zarnt et al., 1997). The N-terminal domain was shown to be necessary and sufficient for the interaction with the ribosome while the PPIase domain is dispensable for its function (Genevaux et al., 2004).

The binding of TF to the ribosome is mediated via a conserved GFRxGxxP motif in the N-terminal ribosome binding domain (RBD) (Kramer et al., 2002). This motif interacts with L23 and L29 at the UAS1 of the eubacterial ribosome. It is located in an unfolded region of TF and its arginine interacts with a glutamic acid of L23 and a phosphate of the 23S rRNA (Ferbitz et al., 2004; Baram et al., 2005; Schlunzen et al., 2005). Additional contacts are observed between the second α-helix of the RBD and L29. However, these interactions are not as important as the interaction with L23 as deletion of L29 does not prevent the interaction of TF with mutant ribosomes (Kramer et al., 2002).

The X-ray structure of TF showed a surprising arrangement of the TF domains: the PPIase-domain is located at the opposite end of the molecule relative to the N-terminus while the C-terminus is located between those two entities (Figure 6). The co-crystallisation of the *E. coli* TF RBD with the *Haloarcula marismortui* 50S large LSUs and the availability of full length crystal structures of TF, allowed the modeling of TF onto the ribosome (Ferbitz et al., 2004). This unphysiological model was later confirmed by co-crystals of *Deinococcus radiodurans* TF bound to 50S LSUs of the same species (Baram et al., 2005; Schlunzen et al., 2005). The structure suggested, that TF provides a protected environment from the tunnel exit towards L24 which was further supported by protease digestion protection experiments (Hoffmann et al., 2006). TF's placement on the ribosome was later

confirmed by cryo-EM studies that showed full length TF crosslinked to a partially folded nascent polypeptide (Merz et al., 2008).
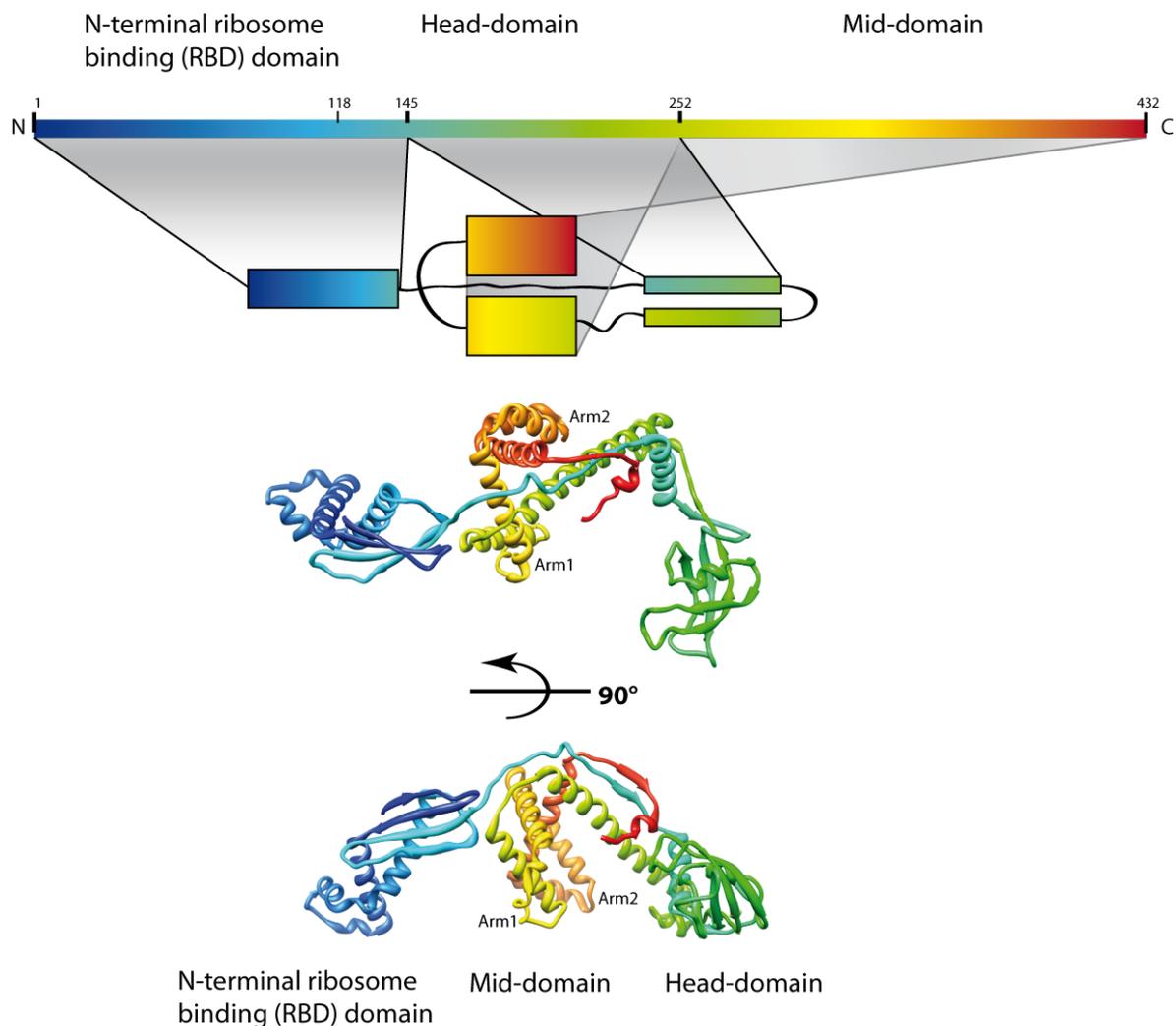
Figure 6 : The structure of TF | TF's order of domain as they are arranged in the primary sequence is not reflected in its 3D structure. The N-terminal RBD-domain and the Head-domain are located at the opposite ends of the molecule. The C-terminal domain folds into a form that provides two protruding arms and is located in between the Head- and the RBD-domain (Mid-domain) pdb:1w26, chain A (Ferbitz et al., 2004).

The same study followed the path of the nascent polypeptide along the interior of TF via crosslinking. According to these data, the nascent polypeptide migrates along the interior of TF in a vectorial manner until it reaches the Head domain. From the first contact of the nascent polypeptide to the TF, a total of 43 AA are necessary to contact the Head-domain. In an extended conformation this would translate into a distance of around 120 Å between the RBD and the Head-domain, which is in agreement with the cryo-EM structure and the crystal structure.
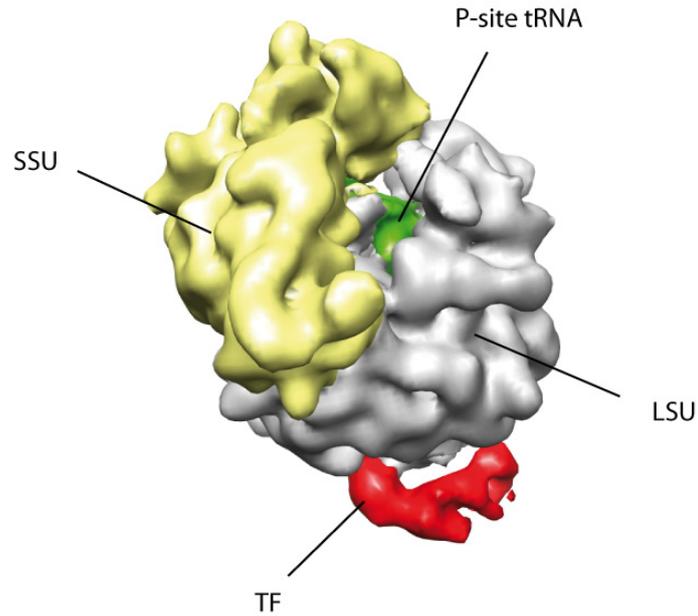
Figure 7 : A low resolution structure of TF bound to RNCs showed its orientation relative to the tunnel exit |TF was crosslinked to a SecM stalled nascent polypeptide that contained a folded SH3 domain. Additionally, an artificial TF binding site was placed directly at the tunnel exit. The resulting complex was subjected to cryo-EM and single particle reconstruction. The final volume showed full length TF spanning from the L23/L29 interaction site towards the other side of the tunnel exit with its interior facing towards the ribosome thereby providing an interaction site of 120 Å in length (emdb: 1499 (Merz et al., 2008), coloured according pdbs: 2wwq (LSU), 2wwl (SSU) (both: (Seidelt et al., 2009) and 2vrh (TF) (Merz et al., 2008)).

The molecular basis of the interaction between TF and the nascent chain is not yet fully understood. Upon binding to the 50S LSU, the binding domain of *D. radiodurans* TF was rearranged relative to the structure in solution (Baram et al., 2005; Schlunzen et al., 2005). This rearrangement could also be due to a crystal contact of the second α-helix that is elongated in D*. radiodurans* relative to *E. coli* TF. Since *E. coli* TF in complex with the archaeal *H. marismortui* 50S LSU was not resolved in these regions, the significance of this observation is not clear yet. It could however represent an activation of TF to accommodate the nascent polypeptide by opening a hydrophobic binding site. It was speculated that L29 interacts with the second α-helix of TF and thereby enables the interaction with the nascent polypeptide (Baram et al., 2005).

TF's function

While TF's contribution to the post-translational export of outer membrane proteins led to its discovery, other functions emerged only later. Its interaction with nascent polypeptides and ribosomes implied a function in co-translational chaperoning of nascent polypeptides. However, since its deletion did not result in a phenotype, this function remained speculative.

First proof for this hypothesis came from the observation that the co-deletion of TF together with DnaK led to synthetic lethality above 30°C due to protein aggregations (Deuerling et al., 1999; Genevaux et al., 2004; Teter et al., 1999). At lower temperature, where protein aggregation is physically reduced, the double deletion is viable. Furthermore, in TF deletion strains, the loading of the DnaK chaperone system with newly translated proteins is strongly increased indicating a reshaping of the progression along the folding pathways of newly synthesized polypeptides.

In contrast to DnaK, TF does not bind ATP and thus its regulation is still enigmatic. First insights into the mode of interaction with nascent polypeptides came from studies that used FRET and fluorescence quenching experiments (Kaiser et al., 2006). Upon ribosome binding, the distance between TF ribosome binding domain and the Head-domain is enlarged. As this effect occurs even in the absence of nascent polypeptides, it is a structural rearrangement that is induced by ribosome binding. Using quenching studies, the binding state of TF in context of actively translating ribosomes could be monitored. For some substrates, the interaction between TF and nascent polypeptide is longer than the interaction with the ribosome. This indicates a situation where TF leaves the ribosome but is still bound to the nascent protein.

The determination of the interaction duration of TF and different nascent polypeptides led to a qualitative description of TF's binding motif. The motif that correlated best with the residence time of TF on nascent polypeptides was a stretch of high hydrophobicity of at least 15 AA. Contrary to that, an earlier study employing immunoblotting of synthesized oligo-peptides (13mers) showed that TF interacts best with motifs of 8 AA that are enriched in basic and aromatic residues and carry a positive net charge (Patzelt et al., 2001). This contradiction reflects the different interaction modes. While the fluorescence quenching experiments investigated the co-translational interaction the latter analyzed TF's interaction in a ribosome free environment.

TF's interaction with hydrophobic stretches might prevent a hydrophobic collapse of newly synthesized proteins or the interaction with hydrophobic stretches in nascent polypeptides of adjacent ribosomes in polysomes. Fluorescence data indicate, that the major shielding function is exerted by the N- and C-terminal domain while the Head-domain – dispensable for the chaperone function and interacting with different motifs – functions on a different subset of substrates (Kaiser et al., 2006).

The development of ribosomal footprinting allowed a global analysis of the TF interactome. TF was crosslinked to RNCs, the TF-RNC complexes isolated, unprotected mRNA digested and the copurified mRNA fragments sequenced (Oh et al., 2011). Consistent with its first discovery, outer membrane β-

barrel proteins are the most important substrates. Deletion of TF leads to deregulation of the post-translation translocation pathway and aberrant co-translational translocation of this class of proteins. Subsequently, this leads to a failure in the outer membrane architecture.

While cytosolic and outer membrane proteins are enriched in the TF-interactome, inner membrane proteins are under-represented. These data are in agreement with *in vitro* data which showed, that SRP and TF – even though sterically capable of simultaneously binding to the ribosome – bind in a mutually exclusive manner to nascent polypeptide. SRP outcompetes TF in binding to its substrate (Ullers et al., 2006). Interestingly, TF recruitment to RNCs *in vivo* was only observed once the nascent polypeptide reached a length of 100 AA. This was in contrast to the current view where TF is believed to scan idle ribosomes before emergence of the nascent polypeptide (Lin et al., 2012).

Insights into TF's function, independent of the ribosome, were achieved by the discovery that *Thermotoga maritima* TF copurifies with a broad range of proteins (Martinez-Hackert and Hendrickson, 2009). Crystallisation and X-ray structure determination revealed an atomic picture of the promiscuous interaction of TF and its substrates. In line with the idea that TF promotes folding and provides a protected folding environment, this study showed a X-ray structure of a TF dimer and a folded substrate, the ribosomal protein S7. The unsusal large interaction surface (4,500 Å²) and a poor packing between the bound substrate and TF allowed a wide range of interactions.

The significance of this observation became evident when the interaction between TF and S7 was investigated in more detail. TF shielded significant parts of S7 (40%) that are buried in the ribosome. This shielding keeps S7 competent for its assembly into 30S particles. Subsequent analysis showed that under high temperature stress TF deletion results in a defined ribosome assembly defect of the large subunit (Martinez-Hackert and Hendrickson, 2009).

1.5 Microscale Thermophoresis: a sensitive detector for complex formation

Thermophoresis describes the migration of molecules along a temperature gradient. In nature, this phenomenon is thought to play an important role in the accumulation of precursor molecules of biomolecules in hydro thermal vents (Baaske et al., 2007). Thermophoresis in pores of the hydrothermal vents is capable of concentrating biomolecules like single nucleotides up to 10^8 times which could provide the critical environment for the initiation of life.

Carl Ludwig observed thermophoresis the first time in 1856 when he noticed, that in a solution where the temperature is unevenly distributed, the ion concentration varies as well (Ludwig, 1856). 23 years later, this effect was systematically analyzed by Charles Soret (Soret, 1879). This led to the development of new processes of concentration of atoms and isotopes and a first theoretical description of this effect (Clusius, 1938; Debye, 1939).

The change in concentration of molecules as an effect of a change in temperature is a function of two counteracting processes. The change in temperature induces a linear thermophoretic migration. The molecule velocity of this migration is linearly dependent on the temperature gradient. The migration leads to a concentration gradient that induces a backdiffusion. Even though the majority of the so far investigated molecules showed a decreased concentration in reaction to a positive temperature gradient, in other systems an increase in the concentration was also observed (Jerabek-Willemsen et al., 2011).

The thermophoretic behavior of a system changes upon altered properties of the molecule. Interestingly, apart from changes in size or masses due to binding events, changes of the hydration shell and of the conformation also have a strong influence on this parameter. Thus, they are observable by a changed thermophoretic mobility. Thereby, even minute changes due to small molecules such as binding of protease inhibitors to proteases lead to a change in the thermophoretic migration behavior.

This led to the development of a technique called Microscale Thermophoresis (MST, Nanotemper Technologies GmbH). An infrared laser (IR-laser) establishes a micro gradient within a glass capillary. The thermophoretic movement of the molecule to be measured is observed via fluorescent tags. This allows the detection of minute amounts of materials in small volumes as low as 5 µl. Typically, 5 to 100 nM of the labeled partner are used and the unlabeled ligand is titrated.

The thermophoretic properties of the labeled molecule are deduced by the analysis of the fluorescence intensity in the hot spot created by the IR-laser over time. However, not only the thermophoretic migration is reflected in this time course, but also the reaction of the fluorescent dye to the change in temperature (temperature jump). While the thermophoretic movement is rather slow, the reaction of the dye happens practically instantly together with the temperature change. As the IR-laser needs ~ 1 s to reach the final temperature, the T-jump shows a dynamic change within that timeframe (Figure 8).
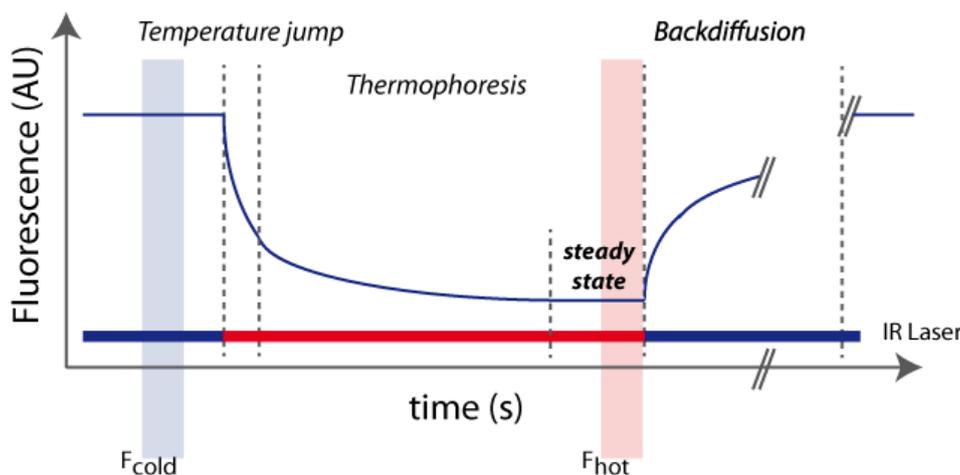
Figure 8 : Using a MST device, the fluorescence over time of a labeled sample is recorded. This time trace represents various different physical reactions of the system | After the IR-laser is switched on, an immediate decrease (an increase for some dyes) in the fluorescence of the dye shows the changed microenvironment. As soon as the IR Laser established the steady state temperature in the hot spot, further changes in the fluorescence are due to the thermophoretic migration of labeled molecules until a steady state is reached. Once the IR laser is switched of, the temperature in the hot spots equilibrates with the environment leading to a fast cooling of the recorded area and thus the reverse of the temperature jump. Finally, backdiffusion of labeled molecules leads to an even distribution of molecules. For analysis, F_{cold} and F_{hot} are recorded.

Once the temperature in the hotspot reaches the steady state level, further changes in fluorescence are only due to migration in and out of the hotspot. As soon as the steady state level is reached, the IR-laser is switched off and a reverse temperature jump can be observed. Finally, backdiffusion abolishes the concentrations gradient that was established due to the thermophoretic movement.

To analyze the data, the average fluorescence at the cold (F_{cold}) and at the hot steady state level (F_{hot}) are measured and a relative value (F_{hot}/F_{cold}) is calculated. Using direct titration series, the relative values for the receptor alone and the saturated receptor-ligand system can be measured. The relative value of a given ratio between receptor and ligand can then be used to calculate the complex formation (see section 4.1.3.2, p. 58 and Formula 1, p. 60).

Apart from the changed migration behavior of the labeled receptor upon complex formation with an unlabeled ligand, the temperature jump can also be affected by the binding state. If the dye is in proximity of the binding site, its altered microenvironment might change its quantum yield, its

emission or absorption maximum or its susceptibility to temperature change. Conformational changes can also influence the physical nature of the dye. As this reaction might support or counteract the signature of the thermophoresis in the fluorescence trace, it is necessary to choose the F_{cold} region carefully. This technique can also be used to determine the competition of factors that share a common binding site.

1.6 Cryo-EM: a powerful tool and its limitations

For structure determination of macromolecules like ribosomes, cryo-EM and single particle analysis is the technique of choice. In comparison to NMR or X-ray crystallography, it requires only small amounts of materials. While the crystallization of proteins typically happens in the high mg/ml concentration regime, a typical cryo-EM sample has a protein content of 0.2-0.5 mg/ml at a volume as low as 10 µl per grid. Additionally, it does not require any ordered packing of the molecules like in crystals. Therefore, the buffer can be set to physiological conditions. Finally, due to their high content of ordered rRNA, ribosomes are exceptionally good molecules to detect in cryo-EM.

Biological specimens consist mainly of carbon, oxygen, nitrogen and phosphor with very low electron absorption. Conventional transmission electron microscopy enhances the signal of the sample by using heavy metals like uranium. In contrast in cryo-EM, the sample is flash frozen using liquid ethane (cooling rate: 100,000°C/s). Thereby, amorphous ice is formed whose water molecules are randomly orientated similar to liquid water. This ice has a lower electron absorption and scattering in comparison to cubic ice and allows therefore for the direct detection of signals arising from the sample. However, the signal to noise ratio (SNR) of the individual particles is very low. Additionally, the samples are very prone to radiation damage. To reduce the radiation damage, electron-micrographs are therefore taken under low dose conditions ($20e^-/Å^2$). This reduces the SNR even further

As a result of the preparation, the particles are randomly orientated on the grid. The image acquisition results in the generation of two-dimensional projections of the three-dimensional sample molecules. These particle-images are then digitalized and subjected to the single-particle analysis work-flow. During this process, the individual particles are aligned against a three-dimensional reference. Thereby, the orientation of the respective sample molecules can be deduced (three euler angles and shifts of the particles relative to the reference). Many particules (10,000 - >1,000,000) are then combined into a single backprojection. Thereby, the 2D information of the particles and the determined angles and shifts are integrated into a 3D volume. To improve the resolution, this process is iterated with increasing angular accuracy.

The final resolution of a cryo-EM density volume (EM-map) is dependent on the size of the dataset (number of individual particles) and the quality of the sample (homogeneity). Due to the low SNR, the information of the particles originating from a homogenous population of sample molecules has to be averaged. Thereby, sub-nanometer resolved 3D-maps can be calculated. Currently the best

resolved ribosomal cryo-EM map reaches a resolution of 5.5 Å (Armache et al., 2010). However, compositional heterogeneity of the sample interferes with a high resolution reconstruction. This can be caused by several sources: substoichiometric binding of ligands and different conformations of molecules are the most common ones. In such cases, the particles are aligned according to the strongest EM-signal and the signal for the weaker parts is less resolved or even absent.

To obtain high-resolution maps of heterogeneous datasets, different techniques are available that define subdatasets that contain particles that share common features. The most widespread technique - competitive multi reference alignment - aligns the individual particles against different references and calculates cross correlation coefficients (CC C) based on similarities to the reference. The particles are iteratively assigned to the better fitting references until equilibrium. In 2D, clustering e.g. using K-means algorithms was applied to generated class averages that are then backprojected with the relevant class averages of neighbouring projection groups (reviewed in van Heel et al., 2000). Finally, 3D bootstrapping of subdatasets transforms the 2D eigenimage analysis into 3D using 100,000s of volumes analyzed by 3D eigenvolumes (Penczek et al., 2006b). All these approaches worked on some heterogeneous datasets, but none is universally applicable. The need for the evolution of the currently applied techniques to push the envelope further is evident as more and more flexible complexes are investigated.

2   Aims

NAC's binding to the ribosome has not been described yet quantitatively. Its high abundance makes it one of the key interactors of ribosomes. In order to describe the regulation of the access of ribosomal ligands to nascent polypeptides, it is therefore mandatory to understand the binding dynamics of this interactor first.

Thus, we aimed at the establishment of a suitable technique that allows the measurement of such binding events with the lowest possible sample requirements. We chose microscale thermophoresis since it combines low sample-need with the necessary sensitivity. This technique has various advantages of established methods but was never used for studying the interaction between ribosomes and ligands. To strengthen our understanding of NAC's behavior at ribosomes further, we then aimed at a description of the conformational changes of NAC upon ribosome binding using environmental sensitive dyes. After we described the NAC-ribosome interaction in a fair detail, the competition between ribosomal interactors of the UAS two was analyzed further. By comparison with known *in vivo* data, the applicability of simplified *in vitro* models to describe *in vivo* situations is tested.

In the second part of this thesis, the interaction of TF with nascent polypeptides and the ribosome is structurally investigated. While there is already a cryo-EM structure of TF bound to RNCs available, the low resolution of this structure prevents a molecular understanding of this interaction. Moreover, it shows an unphysiological situation where TF is crosslinked to an artificial binding site. Since TF is highly flexible, the study used a folded SH3-domain in the nascent polypeptide to restrict TF flexibility. As TF is believed to interact with unfolded nascent polypeptides in a flexible manner, this setup does not reflect its physiological activity. We therefore first aimed at the generation of physiological substrate RNCs, the generation of TF-RNC complexes and subsequently at a high resolution cryo-EM structure that allows the molecular interpretation of the interaction between TF and the nascent polypeptide.

Finally, we plan to investigate the predicted dynamics of TF using cryo-EM data to provide a first description of TFs conformational space on ribosomes bound to substrates.

3 Materials and Methods

3.1 Materials

3.1.1 Devices and Kits

Company	Device
Beckmann Coulter	OptimaMAX Ultra 130k
	SW32 rotor
	SW40 rotor
	TLA100
	TLA110
	Type 70Ti (Ti70)
	Ultra-centrifuge Optima L-70
BioRad	MJ Mini
	Mini-Protean Tetra System
	Trans-Blot SD
Eppendorf	Biophotometer
	Table centrifuge 5415D/5471R
GE healthcare	Äkta purifier
	S75 16/60
life technologies	Poros HQ20 4.6mmDx100mL
	Poros HS20 4.6mmDx100mL
Nanotemper	hydrophilic capillaries
	Monolith NT-015T Blue Gree
	NT.115 Protein Labeling Kit Green-NHS
Pierce	BCA Protein Assay Kit
Qiagen	QIAprep Spin Miniprep Kit
	QIAquick Gel Extraction Kit
	QIAquick PCR Purification Kit
Sorvall	Evolution RC
	SLC6000
	SS34
Thermo Scientific	Nanodrop 1000

3.1.2 Chemicals

All chemicals, unless stated otherwise, were obtained from Merck, Invitrogen, Roth, Sigma-Aldrich, VWR and Fluka.

Company	Special chemicals
Clontech	Talon®
life technologies	Alexa Fluor 555 C_5 Maleimide
	Alexa Fluor 568 C_5 Maleimide
	IANBD Ester
Macherey-Nagel	Protino Ni-NTA Agarose
New England BioLabs	Phusion High-Fidelity PCR MasterMix
	T4-DNA Ligase
	Restriction enzymes

3.1.3 *E. coli* strains

XL1-Blue (Stratagene):

endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB$^+$ lacI^q Δ(lacZ)M15] hsdR17(r$_K^-$ m$_K^+$)

ER2566 (New England Biolabs):

F- λ- fhuA2 [lon] ompT lacZ::T7 gene 1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::miniTn10-TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm]

Rosetta(DE3) (Novagen):

F$^-$ ompT hsdS$_B$(R$_B^-$ m$_B^-$) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)

KC6 (Seidelt et al., 2009)

Rna-19 gdhA2 his-95^{a}relA1 spoT1 metB1 ΔendA met$^+$ ΔtonA ΔspeA ΔtnaA ΔsdaA ΔsdaB ΔgshA ΔsmpB ΔssrA

3.1.4 Oligonucleotides
Oligonucleotides used for cloning of NAC derivatives

Mutation Egd1p-I3C/Egd2p to wt Egd1p/Egd2p

```
fwd: GTATAAGAAGGAGATATACATATGCCGATTGATCAAGAGAAACTGGCTAAACTG

rev: CAGTTTAGCCAGTTTCTCTTGATCAATCGGCATATGTATATCTCCTTCTTATAC
```

Mutation wt Egd1p/Egd2p to Egd1p-A16C/Egd2p

```
fwd: GCTAAACTGCAGAAACTGTCGTGCAATAACAAAGTCGGAGGTACACGTCG

rev: CGACGTGTACCTCCGACTTTGTTATTGCACGACAGTTTCTGCAGTTTAGCCAGTTTCTCTTGATCAATC
```

Mutation wt Egd1p/Egd2p to Egd1p-A16C/Egd2p-V166C

```
fwd: CCATCAAAGCCCTGAAAGCTCACAATGGTGACTTATGCAACGCTATCATGTCCTTG

rev: CAAGGACATGATAGCGTTGCATAAGTCACCATTGTGAGCTTTCAGGGCTTTGATGG
```

Mutation wt Egd1p/Egd2p to Egd1p/Egd2p-V97C

```
fwd: CCATCTAACGAAGACTGCGCTACCAAGTCCCCA

rev: TGGGGACTTGGTAGCGCAGTCTTCGTTAGATGG
```

Blunt end ligation Egd1p-A16C/Egd2p to wt A16C-Egd1p/Egd2p-delUBA

```
fwd: P-TAAGGTGACTTGAATAAGGATGATATC

rev: P-ATCAACTTCACCTTCTTCGTCATC
```

Oligonucleotides used for cloning of MBP fusion proteins

Egd1p(1-39) MBP cloning into pET28a

```
fwd: AAAATCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACC ATGCCAATTGACCAAGAAAAATTAGCTAAG

rev:  GGGGCTCGAG TTAAGTCTGCGCGTCTTTCAGGG
```

Mutation Egd1p(1-39) MBP to Egd1p(1-39) A16C derivative

```
fwd: GCTAAGCTACAAAAGTTGTCTTGCAACAACAAAGTTGGTGGTACTAG

rev: CTAGTACCACCAACTTTGTTGTTGCAAGACAACTTTTGTAGCTTAGC
```

Mutation Egd1p(1-39) MBP to Egd1p(1-39) K240C MBP derivative

```
fwd: GCATGGTCCAACATCGACACCAGCTGCGTGAATTATGGTGTAACGGTACTG

rev: CAGTACCGTTACACCATAATTCACGCAGCTGGTGTCGATGTTGGACCATGC
```

Mutation MBP to MBP(K240C)

```
fwd: GCATGGTCCAACATCGACACCAGCTGCGTGAATTATGGTGTAACGGTACTG

rev: CAGTACCGTTACACCATAATTCACGCAGCTGGTGTCGATGTTGGACCATGC
```

Oligonucleotides used for cloning of templates for RNC preparation

GatD I-35 truncated

```
fwd: AAGGAGATATACCAATGCATCATCATCATCATCATTACCCATACGATGTTCCAGATTACGCTGTTGCCTGTGTGCCGTTATTACCCT

rev: AACTTTTCGCTTAACGACAATATATTCAG
```

Generation of 80AA TF GatD I-35 TnaC via blunt end ligation based on 8H 20AA GatD I-35 TnaC

```
fwd: P- AGCGTAATCTGGAACATCGTATGGGTAATGATGATGATGATGATGGTGGTGAC
rev: P-
ATGAAATCAGTGGTGAATGATACTGATGGTATCGTGCGCGTTGCAGAAAGCGTCATTCCTGAAATTAAACATCAGGATGAGGTGCGGGTAAAAATTGCCAG
```

3.1.5 Plasmids

Complex to be expressed	Proteins expressed by this plasmid complex	Plasmid backbone	Reference/ Cloning procedure
NAC-derivatives			
Egd2p-I3C NAC	Egd1p-I3C-His$_6$ Egd2p	pET-Duet-1	newly synthesized
wt NAC	Egd1p-His$_6$ Egd2p	pET-Duet-1	Mutation I3C-NAC to wtNAC
Egd1p-A16C NAC	Egd1p-A16C-His$_6$ Egd2p	pET-Duet-1	Mutation wtNAC to A16C
Egd1p-A142C NAC	Egd1p-A142C-His$_6$ Egd2p	pET-Duet-1	Mutation wtNAC to A142C
Egd1p-C-terminal Cystein	Egd1p-Cter-C-His$_6$ Egd2p	pET-28a	(Pech, 2006)
Egd2p-V97C NAC	Egd1p-His$_6$ Egd2p-V97C	pET-Duet-1	Mutation wtNAC to V97C
Egd2p-V166C NAC	Egd1p-His$_6$ Egd2p-V166C	pET-Duet-1	Mutation wtNAC to V166C
Egd1p-(1-39)-MBP	fusion protein Egd1p-(1-39)-(GGGS)$_4$-MBP	pET-28a	this study Insert from (Pech, 2006)
Egd1p-(1-39; A16C)-MBP	fusion protein Egd1p-(1-39;A16C)-(GGGS)$_4$-MBP	pET-28a	Mutation wtNAC to A16C
Egd1p-(1-39)-MBP (K240C)	fusion protein Egd1p-(1-39)-(GGGS)$_4$-MBP(K240C)	pET-28a	Mutation wtMBP to K240C
MBP(K240C)	MBP(K240C)	pMAL-c2	Mutation wtNAC to K240C
MBP(K240C)-UBA	fusion protein MBP(K240C)-(GGGS)$_4$-UBA	pMAL-c2	Master-Thesis Steffen Fliehmann
Trigger Factor			
Trigger Factor	Trigger Factor-His$_6$	pPROEX-Hta	(Kaiser et al., 2006)
Trigger Factor (T150C)	Trigger Factor (T150C) -His$_6$	pPROEX-Hta	(Kaiser et al., 2006)
Templates for RNC generation			
His$_6$ 20AA GatD I-35 TnaC	TnaC stalled RNC template: His$_6$_HA_GatD(61-103)_TnaC	pBAT4	Newly Synthesized
His$_8$ 20AA GatD I-35 TnaC	TnaC stalled RNC template: His$_8$_HA_GatD(61-103)_TnaC	pBAD	Eli van der Sluis
His$_8$ 80AA GatD I-35 TnaC	TnaC stalled RNC template: His$_8$_HA_GatD(1-103)_TnaC	pBAD	Insert of 60AA in His$_6$ 20AA GatD I-35 TnaC

His$_6$ FtsQ 84 TnaC	TnaC stalled RNC template: His$_6$_HA_FtsQ(1-61)_TnaC	pBAT4	Newly Synthesized
	others		
SR	α subunit of S. cerevisiae SR	pET28a	Diploma-thesis Marco Gartmann
SR	β subunit (31-244) *S. cerevisiae* SR N-terminal His$_6$ Tag	pET21a	Diploma-thesis Marco Gartmann
MBP	MBP	pMAL-c2	Purchased from NEB

3.1.6   Antibodies

Mouse HA-probe Antibody (F-7) IgG   (Santa Cruz Biotechnology, Inc.)

goat-anti-mouse IgG-HRP   (Santa Cruz Biotechnology, Inc.)

3.2   Molecular biology techniques

3.2.1   Transformation

3.2.1.1   *Preparation of electrocompetent E. coli*

Escherichia coli cells (XL1 Blue) from an overnight culture in LB medium were inoculated 1/300 in 250 ml of a low salt medium (YENB) and incubated at 37°C, 120 rpm until an OD$_{600}$ of 0.7. The cultures were cooled on ice (10 minutes) and the cells pelleted (GSA rotor, 6.000 rpm, 4°C, 15 minutes). The cells were washed once in 100 ml sterile water and twice in 10 ml 10% glycerol. Afterwards, the cells were resuspended in 2 ml 10% glycerol, aliquoted, flash frozen in liquid nitrogen and stored at -80°C until used.

YENB-medium
0.75 % Bacto-Yeast extract
0.8% Nutrient Broth
pH 7.5

3.2.1.2   *Bacterial transformation by electroporation*

40 µl electro competent XL1 Blue cells (see Section 3.3.1.1.1) were thawn on ice and mixed with 1 µl of the ligation reaction in precooled cuvettes (1 mm slit width). The electroporation was performed at 1,800 V. Afterwards, the cells were immediately resuspended in 500 µl LB medium, incubated at 37°C for 45 min and plated on appropriate selection plates

LB- medium
1% tryptone
1% NaCl
0.5 % yeast extract

3.2.1.3 Preparation of chemical competent E. coli

Escherichia coli cells (XL1 Blue, ER2566, Rosetta (DE3)) from an overnight culture in LB medium were inoculated 1/300 in 100 ml LB medium and incubated at 37°C, 120 rpm until an OD_{600} of 0.7. In case of the Rosetta (DE3) strain chlorampenicol was added to a final concentration of 34 mg/l to the media. The cultures were cooled on ice (15 minutes) and the cells pelleted (GSA rotor, 6,000 rpm, 4°C, 15 minutes). The pellets were washed once in 50 ml sterile 0.1 M $CaCl_2$ and resuspended in 5 ml 15% glycerol, 0.1 M $CaCl_2$. After aliquoting, the competent cells were flash frozen and stored at -80°C.

3.2.1.4 Bacterial transformation by heat shock

100 ng of plasmid was incubated with 100 µl chemical competent cells (see section 3.3.1.1.3) for 30 minutes on ice. Heat shock was performed at 42°C for 90 seconds. Following 5 minutes incubation on ice, 500 µl of LB medium were added to the cells and the transformed bacteria were incubated at 37°C cells for 45 minutes. Selection for positive transformation was done on LB agar plates with the appropriate selection markers (LB-medium see section 3.2.1.2.

3.2.1.5 Long term storage of bacterial cultures

Bacterial cells at logarithmic phase were mixed 1:1 with sterile glycerol, aliquoted, flash frozen and transferred to -80°C

3.2.2 Plasmid purification in E. coli

One colony of a plasmid containing *E. coli* strain was used for inoculating 5-10 ml of LB medium with the appropriate antibiotic at 37°C over night. The DNA was then purified using the QIAprep®Spin Miniprep Kit (Qiagen) according to the manufacturer's protocol. For large scale purification of plasmids, 500 ml LB medium with the appropriate antibiotics was inoculated with a single colony and grown at 37°C over night. The DNA was then purified using the QIAprep®Spin Maxiprep Kit (Qiagen) according to the manufacturer's protocol.

3.2.3 Expression

2-12 liter LB medium were inoculated with 1/300 of a preculture of the protein expressing strain and incubated at 37°C on a shaker (New Brunswick Scientific Co, New Jersey) at 120 rpm until an OD_{600} of 0.5. Expression was induced by adding IPTG to a final concentration of 1 mM for 2h. Cells were harvested (SLC6000, 10 min, 5,000 rpm, 4°C) and resuspended in buffer A (20 ml per 12 g cells). The cells were either directly used or flash frozen in liquid nitrogen and stored -80°C until further usage.

Buffer A
50 mM Tris-Cl, pH 8.0
500 mM NaCl

3.2.4 PCR based Techniques

3.2.4.1 Amplification of DNA fragments

For amplification or mutagenesis of double stranded DNA molecules, polymerase chain reaction (PCR) was used. For all applications, except colony PCR, the Phusion Flash High-Fidelity Master Mix (Finnzymes) was used according to the manual. Typically, 1 ng of a PCR product or 100 ng of a plasmid were mixed with 0.5 µM of each primers and the appropriate amount of master mix. Initial denaturation was performed at 98 °C for 80 s while in the individual cycles the denaturation time was reduced to 5 s if as template a PCR product was used or to 60 s if a plasmid was used. The annealing temperature was adjusted between 55 °C and 60 °C according to the primer's sequence, and annealing was allowed for 20 s. Extension was performed at 72°C for 30 s per kbp product length. Typically, 30 cycles were run per PCR reaction. Colony PCR was used as a quick tool to visualize presence of the desired DNA-fragment in bacterial cells. One freshly grown colony was used as template and a 10 µl reaction using 0.1 µl Taq-Polymerase, 1 µl 10x Taq-Buffer, 0.5 µl fwd Primer (10 µM), 0.5 µl rev Primer, and dNTPs was prepared. The cells were opened by incubation at 95°C for 5 min, and amplification of the DNA sequence was accomplished by denaturation at 95°C for 10 s, annealing at 55°C for 10 s and extension at 72°C for 1 min per kbp. Typically, 30 cycles were run. All oligonucleotides that were used in this study were ordered from metabion international AG (Munich) or biotez GmbH (Berlin). Primers longer than 80 bases were HPLC purified to ensure that only full length products were used for the experiments.

3.2.4.2 Site specific mutagenesis

To mutate up to 6 bp in a plasmid, site specific mutagenesis was used. Primers that contained the desired base pair exchange centrally, flanked by pairing sequences of 25 bp were designed. A PCR on 40 ng of the Plasmid in a total volume of 50 µl was performed. Deviating from the protocol in section 3.2.4.1, the reaction was split up before adding the primers. On half was supplemented with the forward primer and the other half with the reverse primer. Both sub reactions were then subjected to four PCR cycles and then combined for 14 more cycles. The original, not mutated DNA was then digested using DpnI, the mutated DNA precipitated and transformed using chemical competent cells.

3.2.4.3 Blunt end ligation

Blunt end ligation used the ability of T4 DNA Ligase to ligate blunt DNA ends if the 5' OH auf the molecules are phosphorylated. This technique allows the high efficient addition of up to 100 nucleotides in plasmids. Primers are designed that bind on the plasmid in opposed directions The sequence to be added is 5' of the annealing sites of the primer on the plasmids. If the primer are 5' phosporylated, the generated linear DNA with blunt ends on both sides can then be ligated by the T4-DNA-Ligase.

40 ng of the starting plasmid were used as a template per 50 µl reaction. Two 5' phosphorylated primers that yielded a full length product of the plasmid plus the insertion were added (2.5 µl of a 10 µM stock per 50 µl) and a PCR reaction (3.2.4.1) was carried out. Two identical 50 µl PCR reactions yielded enough DNA for further steps. The template was then digested using DpnI and the products were ligated for 2 h at 25°C. Heat shock transformation yielded the ligated product with a high efficiency.

3.2.5 Restriction of DNA fragments

For restriction of DNA, restriction enzymes from New England Biolabs were used. Typically, 20 µg of DNA was diluted with water to 50 µl. 7 µl of the appropriate buffer, eventually 0.7 µl of a BSA standard (10 mg/ml, NEB) and 20 units of the restriction enzymes were added. The reaction was filled up to 70 µl using water. Digestion was allowed for 5 h at 37°C to ensure full cleavage of the DNA fragments.

3.2.6 Ligation of DNA fragments

DNA fragments containing compatible sticky ends were ligated using T4 DNA Ligase (NEB). In 16 µl, 20 times molar excess of the insert was added to the backbone, 2 µl 10x T4 DNA ligase buffer and 2 µl T4 DNA Ligase were added. Ligation was performed over night at 16°C. The ligated DNA-molecules were transformed using electroporation.

3.2.7 Sequencing

For sequencing, 1.5 µg of Plasmids or 200 ng of PCR products and 0.3 µl of the appropriate primers at a concentration of 50 µM were sent to eurofins MWG biotech.

3.2.8 *In vitro* Transcription

2.5 µg of a PCR DNA template containing a T7-Promotor was incubated in the presence of 8 mM ATP, 8 mM CTP, 8 mM UTP, 8 mM GTP, and 5 mM DTT in T7 buffer in 96 µl. Addition of 4 µl T7-DNA Polymerase (prepared by technicians) started the transcription. The reaction was performed for 120 min. The RNA was then precipitated in 0.8 M LiCl for 5 h at -20°C and pelleted at 20,000 g for 30 min at 4°C. The RNA pellet was then washed using 1 ml ice-cold 70% ethanol and pelleted again at 20.000 g for 10 min at 4°C. Next, the RNA was per reaction resuspended in 20 µl RNAse free water and stored at -80°C until further use.

T7-Buffer
40 mM Tris-CL, pH 7.9,
26 mM MgCl$_2$,
2.5 mM spermidine,
0.01% (v/v) Trition X-100

3.3 Biochemical techniques

3.3.1 Cell lysis

The harvested cells were resuspended in resuspension buffer and mechanically disrupted using a microfluidizer (Microfluidics) by three passes with 17 kpsi each (for yeast cells: 5 x 20 kpsi). The cell lysate was cleared (SS34, 15 min, 13,000 rpm, 4°C) and the supernatant centrifugated again to pellet cell debris (Ti45, 30 minutes, 33,000 rpm, 4°C). The lysate was either flash frozen and stored at -80 °C or directly used.

Resuspension buffer (for downstream Ni-NTA purification)
50 mM Tris-Cl, pH 8.0
500 mM NaCl,
0.1 mM PMSF

3.3.2 Protein purification

3.3.2.1 NAC and derivatives

Codon optimized EGD1 with a C-terminal TEV-cleavable His$_6$ tag and wildtype EGD2 was cloned into pET-Duet-1. Co-expression of the NAC-subunits in *E.coli* ER2566 of the subunits was performed in 2.5 l LB medium according to section 3.2.3. For Cystein derivatives, all Buffers were supplemented with 2 mM DTT. The cells were lysed and the cleared lysate incubated with 6 ml pre equilibrated Ni-NTA for 1 hour at 4°C. The column was washed five times with 7 column volume (CV) buffer A plus 5 mM imidazole and the complex eluted in 6 steps of one CV buffer A plus 500 mM imidazole each. The peak fractions were pooled and together with 300 µl TEV protease (7 A$_{280}$/ml) dialyzed against 4 l of buffer B at 4°C. One third of the dialysate was used for further purification.

Anion exchange column:

The dialysed sample was supplemented with 1 mM ATP to activate the ATP cycle of chaperones. Thereby chaperones bound to the intrinsically disorded region are released and NAC could be purified to homogeneity. The sample was incubated for 15 minutes on ice and then loaded on a Poros-HQ20 column (4.6 ml, life technologies) using an äkta purifier system. Unbound sample was washed out with two CV buffer HQ without salt and elution was performed via a gradient from 0-1 M KCl in buffer HQ. The gradient consisted of two CVs from 0 to 5 mM KCl in buffer B followed by 14 CVs from 5 to 1,000 mM KCl in buffer HQ. NAC eluted between 335-550 mM KCl The fractions containing NAC (335-550 mM KCl) were pooled and diluted to 100 mM KCl in buffer HQ.

Cation exchange column:

The pooled fraction were then loaded on a Poros-HS20 column (4.6 ml), unbound sample was washed out with two CVs buffer HQ without salt and elution was performed via a gradient from 0-1 M KCl in buffer HQ. The gradient consisted of two CVs from 0 to 5 mM KCl in buffer B followed by 35 CVs from 5 to 1000 mM KCl. NAC eluted between 270-420 mM KCl. The peak fractions (270-420 mM KCl) were pooled.

Gel permeation chromatography column:

The pooled fractions were concentrated by filtration to 0.6 ml (Amicon ultra 10 k, Millipore) and loaded onto a S75 16/60 gel permeation chromatography column (GE healthcare, Buffer B). The fractions containing pure NAC were pooled, the protein concentration determined using the BCA assay and the sample was flash frozen and stored at -80 °C until further use.

When Cystein derivatives were purified, all buffers contained 2 mM DTT.

Buffer A	*Buffer HQ/HS*
50 mM Tris-Cl, pH 8.0	20 mM Hepes, pH 7.4
500 mM NaCl	0/1000 mM KOAc

Buffer B
50 mM Tris-Cl, pH 8.0
50 mM NaCl
1 mM EDTA

3.3.2.2 MBP-Fusion proteins

MBP-Fusion proteins were cloned in either pMAL-c2 or pET28a plasmids. Expression for all derivatives was best in *E. coli* Rosetta (DE3) and expression of a 2 liter culture was done according to section 3.2.3 (p. 38) with the modification that the LB medium was supplemented with 1 % glucose to repress expression of amylase. The cells were lysed and clear by ultracentrifugation (see section 3.3.1) and the lysate incubated with 2 ml amylose resin (equilibrated in MBP-200 buffer) per liter starting culture. The flow-through was reloaded on the resin once. Unbound sample was washed away with one CV MBP-200 buffer. The column was washed 7 times with two CVs each of MBP buffer with increasing NaCl concentration (200 mM, 300 mM, 500 mM, 700 mM, 800 mM, 1,000 mM, 1500 mM). Before elution, the column was equilibrated to low salt by washing with two CVs of MBP-200. The MBP Fusion proteins were eluted in 4 steps using two CVs of buffer MBP-200 plus 10 mM maltose each. Protein concentration of the MBP fusion proteins was determined by measuring

A_{280} (see also section 3.4.1). The proteins were concentrated, aliquoted, flash frozen and stored at -80 °C until further use.

MBP-X buffer
20 mM Hepes, pH 7.4
1 mM EDTA
X mM NaCl

3.3.2.3 Soluble SRP-receptor

The α subunit of *S. cerevisiae* SRP-receptor in the pET28a vector and a truncated version of the β subunit (Δ31-244) in frame with a PreScission-protease cleavable N-terminal His_6-tag in pET21a were provided by Marko Gartmann (AG Beckmann, Gene center LMU Munich). The plasmids were co-transformed into *E. coli* ER2566 and expression in 12 liter LB medium plus Ampicillin and Kanamycin was done according to section 3.2.3 (p. 38). Purification based on the His_6-tag was done according to section 3.3.2.1 with 7 ml Ni-NTA per 12 liters medium. Dialysis was done in the presence of 40 Units of PreScission-protease against SR-Buffer to cleave the N-terminal His_6 tag.

SR buffer
20 mM Tris-Cl, pH 7.5
150 mM NaCl
0.5 mM EDTA

Cation exchange column:

The dialyzed sample was loaded onto a Poros HS20 column (4.6 ml, life technologies). A gradient from 0-1 M KOAc in HQ buffer over 20 CVs was applied to elute the complex. The peak fractions (240-500 mM KOAc) were pooled, concentrated by filtration (Amicon ultra 10 k, Millipore), flash frozen and stored at -80 °C until further use. The concentration was determine using the BCA-assay.

Buffer HQ/HS
20 mM Hepes, pH 7.4
0/1000 mM KOAc

3.3.2.4 TF

Plasmids coding for wild type TF or the single cystein mutant R150C in pPROEX-Hta were kindly provided by Ulrich Hartl, MPI Biochemie in Martinsried (Kaiser et al., 2006). The plasmid was transformed into *E. coli* ER2566 and expression in 2 Liter LB-medium plus Ampicillin was performed according section 3.2.3 (p. 38). The cells were lysed. For Cystein derivatives, all Buffers were supplemented with 2 mM DTT.

Purification based on the His-tag was done according section 3.3.2.1 with 1 ml Ni-NTA resin per 2 liters medium. To cleave the His$_6$-tag, the elution of the Ni-NTA column was incubated at 34°C for 2h in the presence of 40 µl TEV-protease (7,8 A$_{280}$/ml). TF was then dialysed over night against 50 mM Tris, 50 mM NaCl.

Anion exchange column:

The dialysed sample was loaded onto a Poros HQ20 column (4.6 ml, life technologies). A two step gradient from 0-1M KOAc in 20 mM Hepes, pH 7.4 over 17 CVs was applied to elute the complex (0-50 mM KOAc, 2 CV and 50-1,000 mM KOAc, 15 CV). TF eluted at 170-280 mM KOAc. The peak fractions (170-280 mM KOAc) were pooled, concentrated by filtration (Amicon ultra 10 k, Millipore), flash frozen and stored at -80 °C until further use.

Gel filtration column:

The pooled fractions were concentrated to 0.6 ml (Amicon ultra 10 kDa cut off, Millipore) and loaded onto a S200 26/60 gel permeation chromatography column in 20 mM Hepes, 100 mM KOAc, pH 7.4. The fractions containing only full length TF (elution volume 78-90 ml) were pooled, the protein concentration determined using the absorption at A$_{280}$ and the sample was flash frozen and stored at -80 °C until further use.

3.3.3 Protein labeling techniques

3.3.3.1 Unspecific labeling

To unspecifically label primary amines, the GREEN-NHS labellig kit of Nanotemper was used and the proteins labeled according the manufacturer's instructions.

3.3.3.2 Cystein specific labeling

Typically, 10 nmol of the protein to be labeled was used. First, a 10 times molar excess of TCEP (pH 7.0) was added to reduce the cysteines (10 mins at 25 °C). The reduced protein was then mixed with the dye (NBD: 10x molar excess, Alexa dyes: 5x molar excess), and the reaction performed at 25°C for 2 h in the dark. The reaction was quenched by adding β-mercaptoethanol at a final concentration of 10 mM. Then, the reaction was filled up to 1 ml with labeling buffer and purified using a NAP-10 column according the manufacturer's instructions. While the TF-150C derivative and the A16C NAC derivatives (Egd1p-A16C NAC and Egd1p(1-39)-A16C-MBP) were labeled with a high efficiency, other NAC- and MBP-derivatives were only substoichiometrically labeled. Therefore, the NAP-10 eluate was incubated with 250 µl activated thiolsepharose-6B for 2h at 4°C in the dark. The unlabelled proteins bound to the matrix and the supernatant contained only labeled protein. A final NAP-10

buffer exchange step using labeling buffer removed 2-thiopyridone that is released from the matrix upon binding of the protein to the matrix.

Labelling Buffer
20 mM Hepes, pH 7,4
100 mM KOAc
10 mM MgOAc

3.3.4 Ribosome preparation

3.3.4.1 Pelleting

Cells were collected at the logarithmic phase (OD_{600} 1-1.2). The cells were washed in cold ddH_2O and afterwards in 1% KCl. The cell wall was destabilized by incubation in 100 mM Tris-Cl pH 8.0, 10 mM DTT at RT (15 min). The cells were then resuspended in lysis-buffer and mechanically disrupted using a microfluidizer (Microfluidics) by five passes with 20 kpsi each. The lysate was then centrifugated in a SS34 at 15,500 rpm for 15 mins (4°C). The supernatant was cleared in a Ti70 rotor for 28', 36,000 rpm (4°C). The top layer (white, lipids) was discarded and only the clear fraction used (S100). The S100 was then adjusted to 500 mM KOAc and incubated at 4°C for 30 mins. The ribosomes were pelleted through a 1 M sucrose cushion (200 µl, 75', 90,000 rpm, TLA110) and resuspended in MST-buffer (final ribosome concentration: 0.9 A_{260}/µl = 18 µM).

Lysis-buffer   *Sucrose cushion*
20 mM Hepes, pH 7,4   20 mM Hepes-KOH pH 7.5
100 mM KOAc   500 mM KOAc
4 mM DTT   20 mM Mg(OAc)2
150 mM sucrose   1 mM DTT
0,5 mM PMSF   1 M sucrose
1x Pill / ml   1x PI

MST-buffer
20 mM Hepes, pH 7.4
100 mM KOAc
10 mM MgOAc
2 mM DTT

3.3.4.2 80S-purification

Pelleted ribosomes were mixed with the same volume of 2x Puromycin-buffer for 30 mins at 25°C. The ribosomes were then spun through a 10-40% sucrose gradient (SW40, 2.5 h, 40,000 rpm, 4°C). The gradient was fractionated under constant recording of A_{256}. The peak fractions were combined and spun through a 1 M sucrose cusion for concentration (45 min, 100,000 rpm, TLA110, 4°C). The ribosomes were then resuspended in MST buffer (see section 3.3.4.1).

2x Puromycin buffer
40 mM Hepes-KOH, pH 7.5
1 M KOAc
25 mM Mg(OAc)2
2 mM DTT
5 mM RNasin
2 mM Puromycin

Sucrose-cushion
20 mM Hepes-KOH pH 7.5
500 mM KOAc
20 mM Mg(OAc)2
1 mM DTT
1 M Saccharose
1x PI

3.3.4.3 Bacterial polysomal profile

Bacterial cells were grown to an OD_{600} of 0.4-0.5. Chloramphenicol was added to a final concentration of 0.1 mg/ml to the culture to stall the bacterial ribosomes during translation. After three minutes of incubation, the cells were cooled on ice and the cells were pelleted (SLC6000, 5,000 rpm, 10 min, 4°C). The medium was removed and the cells resuspended in polysomal-buffer (0.5 ml). 50 µl of lysozyme (10 mg/ml) was added. For lysis, the cells were flashfrozen in liquid nitrogen and thaw in a cool water bath (three times). The lysis was completed by adding 15 µl 10% Na-deoxycholate. The lysate was cleared by centrifugation in a Eppendorf tabletop centrifuge, 14,000 rpm, 10 min (4°C). The supernatant was loaded onto a 10-40% sucrose gradient and centrifugated (SW40, 40,000 rpm, 2.5 h, 4°C). The gradient was fractionated under constant recording of A_{256}.

Polysomal buffer
20 mM Tris-Cl, pH 7.5
100 mM NH_4Cl
10 mM $MgCl_2$

3.3.4.4 E. coli in vivo RNCs

His_8 - tagged, TnaC stalled nascent polypeptides were encoded on pBAD Plasmids. The plasmids were freshly transformed into *E. coli* KC6. 6-12 Liter of LB and Amp were inoculated with a preculture (1/300) and grown to an OD_{600} of 0.6. Expression of the stalling peptide was induced by 0.2 % Arabinose for 1 h. The cells were pelleted (SLC 6,000, 5,000 rpm, 10 min, 4°C) and resuspended in 40 ml 250 buffer (plus 1 mM tryptophane, 2% (v/v) pill/ml, 1/1,000 Chloramphenicol (34 mg/ml)) per 40 g Cells resuspended.

The cells were mechanically disrupted using a microfluidizer (Microfluidics) by three passes with 17 kpsi each. The cell lysate was cleared (SS34, 25 min, 16,000 rpm, 4°C) and the ribosomes pelleted (Ti45, 2 h, 42,000 rpm, 4°C). The pelleted ribosomes were resuspended in 2 ml 250 buffer per liter starting medium (plus 1 mM tryptophane, 2% (v/v) pill/ml, 1/1,000 Chloramphenicol (34 mg/ml)). The RNCs were bound to 300 µl of a Co^{2+} chelating matrix (Talon®, Clontech) per liter of culture (1h, 4°C), washed with 10 CVs of 250 (plus 1 mM tryptophane, 2% (v/v) pill/ml, 1/1,000 Chloramphenicol (34 mg/ml)), 5 CVs of 500 buffer (plus 1/50 pill/ml, 1/1,000 Chloramphenicol) and eluted with three

CVs of 250/100 buffer. The RNCs were then gradient purified on a 10-40 % sucrose gradient (SW40, 3h, 40,000 rpm, 4°C) and the fraction containing the RNCs were pooled for a final pelleting step (TLA110, 4h, 40,000 rpm, 4°C). 150 µl of 250 buffer per liter of starting volume was used for resuspension of the RNCs (45 min, 4°C). The quality of the RNCs was assessed by western blotting (HA tag).

250 Buffer
50 mM HEPES 7,2
250 mM KOAc
25 mM Mg(OAc)
250 mM Sucrose
250 µg/ml CAM
0.1 % DDM
0.1 % (v/v) pill/ml

500 Buffer
50 mM HEPES 7,2
500 mM KOAc
25 mM Mg(OAc)$_2$
250 mM Sucrose
250 µg/ml CAM
0.1 % DDM
0.1 % (v/v) pill/ml

Gradient (pH adjusted to 7.2)
50 mM HEPES 7.2
25 mM Mg(OAc)$_2$
50 mM KCl
10 mM NH$_4$Cl
10/40 % (w/v)

Sucrose cushion
50 mM HEPES 7.2
500 mM KCl
25 mM Mg(OAc)$_2$
1.4 mM ß-ME
0.75 M Sucrose
250 µg/ml CAM
0.1 % (v/v) pill/ml

250/100 Buffer
50 mM HEPES 7,2
250 mM KOAc
25 mM Mg(OAc)$_2$
250 mM Sucrose
100 mM Imidazol
250 µg/ml CAM
0.1 % DDM
0.1 % (v/v) pill/ml

3.3.4.5   E. coli *in vitro RNCs – truncated mRNA*

In vitro translation was performed using the *in vitro E. coli* T7 S30 Extract for Circular DNA (Promega) supplemented with *in vitro* transcribed mRNA lacking a stop codon (GatD I-35). The purification was done according to (Halic et al., 2006a).

3.3.4.6   E. coli *TnaC stalled in vitro RNCs*

A translation competent S30 extract was prepared as previously published and used for a coupled *in vitro* transcription/translation reaction of 10 ml (Seidelt et al., 2009). pBAT4 based plasmids coding for the stalling peptides His$_6$-FtsQ 84 TnaC and His$_6$-GatDI-35 TnaC were added to start the reaction. The translation was stopped after 30 min by adding chloramphenicol to a final concentration of 100

μg/ml. The ribosomes of the translation reaction were pelleted through a sucrose cushion (TLA110, 3h, 40,000 g, 4°C) and resuspended in 250 Buffer. The resuspended ribosomes were subjected to the same RNC purification protocol as in *vivo st*alled RNCs (see above). The Buffers used were the same as in the *in vivo* RNC protocol.

3.4   Protein characterization techniques

3.4.1   Protein concentration determination
Pierce BCA Protein assay Kit was used for determination of concentration of NAC and SR preparations according to the manufacturer's protocol. For all other proteins, the A_{280} of the preparation was measured and the protein concentration was calculated based on extinction coefficient.

Protein	*extinction coefficient (M^{-1} cm^{-1} at 280 nm)*
TF/TF150C	17,420
MBP-derivatives	66,350 (non derivative contained additional tryptophane, tyrosine or cysteines that change the Extinction coefficient)

3.4.2   SDS –Polyacrylamid electrophoresis
Protein samples were separated using SDS-Polyacrylamid electrophoresis (SDS-PAGE, Laemmli, 1970). The sample was mixed with 5 x sample loading buffer, denatured for 5 minutes at 95 °C and loaded onto a discontinuous 15 % gel in a Mini-Protean II electrophoresis (Bio-Rad) chamber. Electrophoresis was performed at 230 V for 45 minutes in standard SDS running buffer.

5x sample loading buffer
250 mM Tris/HCl pH 8,0
12% SDS
40% (v/v) Glycerin
0.6% bromphenole blue
5% β-mercaptoethanol

3.4.3   Staining of SDS-PAGE gels
The SDS-PAGE gels were stained and fixed for 1 hour at room temperature under shaking using 40 ml Coomassie staining solution per gel. Unspecific binding of the dye was removed by incubating the gel three times with 40 ml destaining solution under shaking. The stained gel was afterwards equilibrated in TBS and scanned using an Epson perfection 4490 PHOTO scanner.

Coomassie staining solution	*destaining solution*
0.25% Coomassie Blue R250	40% methanol
50% methanol	10% acetic acid
10% acetic acid	

3.4.4 Western Blotting

For western blotting, semi dry western blotting was used. SDS-PAGE gels were equilibrated for 10 mins in blotting buffer and blotted onto nitrocellulose or polyvinylidenfluorid membranes (45 min, 75 mA per gel). The membrane was stained for 5 min with amido black staining solution. Unspecific bound dye was washed away using destaining solution and the membrane was scanned. After one wash step using TBS, the membrane was subsequently blocked for 2 h at room temperature using 5 % milk powder in TBS for all antibodies used in this study. After blocking, the membranes were washed three times for 15 minutes TBS-T and the primary antibody was added over night at 4°C (mouse anti-HA: 1:500 in 3 % BSA in TBS). Unbound antibody was washed away using three wash steps, each 15 minutes (TBS-T) and the secondary antibody was added for 1 hour at room temperature (goat anti mouse HRP: 1:2,000 in 5% milk in TBS). Unbound secondary antibody was washed away (3 x 15 minutes) and the membrane was equilibrated in TBS.Detection of bound secondary antibodies was done using the Chemiluminescent Detection Kit (AppliChem) and films (Amerskam Hyperilm ECL) according to the manufacturer's protocol.

amido black staining solution	*TBS*
0,1% Naphthol Blue Black	20 mM Tris-Cl pH 7.5
7,5% acetic acid	150 M NaCl
20% methanol	
destaining solution	*TBS-T*
40% methanol	20 mM Tris-C pH 7.5
10% acetic acid	150 M NaCl
	0.1 % Tween-20

3.4.5 Ribosome binding assays

200 pmol of the potential ribosomal interactor were incubated with 20 pmol ribosomes in 20 µl appropriate buffers. After 15 minutes incubation at room temperature, the samples were loaded on top of 150 µl of a 750 mM sucrose cushion in TLA100 tubes. If *S. cerevisiae* ribosomes were used, pelleting of the ribosome was done for 45 minutes at 75,000 rpm, 4°C. *E. coli* ribosomes were pelleted at 50,000 rpm for 1 hour, 20 minutes. The first 50 µl were defined as the supernatant fraction, the last 20 µl the pellet fraction. The fractions were then TCA precipitated and 1/2 was loaded on SDS-PAGE gels.

Sucrose cushion
50 mM HEPES 7.2
500 mM KCl
25 mM Mg(OAc)$_2$
1,4 mM ß-ME
0.75 M Saccharose
250 µg/ml CAM

3.4.6 Complex-reconstitution

3.4.6.1 Ribosome nascent polypeptide-TF reconstitution

A tenfold molar excess of TF was incubated with RNCs for 10 min at RT. The concentration of the components varied in dependence on the later use of the complex. The RNC concentration of samples that were subsequently used for cryo-EM analysis was 80 nM while the concentration of RNCs in binding assays was 500 nM. After complex formation, the samples were kept on ice.

Final buffer conditions
20 mM Hepes, pH 7.4
100 mM KOAc
10 mM MgOAc
2 mM DTT

3.5 Biophysical techniques

3.5.1 Fluorescence quenching

The environement sensitive dye IANBD Ester (N-((2-(iodoacetoxy)ethyl)-N-Methyl)amino-7-Nitrobenz-2-Oxa-1,3-Diazole) was covalently linked to Cystein mutants (see section 3.3.3.2). For measurements dedicated fluorescence spectrometers as well as the nanotemper thermophoresis device were used. In the fluorescence spectrometers IANBD was excited at 481 nm and recorded at 538 nm (slit width 2 nm). When using the nanotemper device, IANBD was measured using the "blue" settings with an LED power of 80% or 200 units.

3.5.2 Fluorescence anisotropy

The fluorescence anisotropy measurements were conducted in the facility of AG Gaul, Genzentrum der LMU München. 100 nM of the NBD labeled protein was excited with a 488 nm laser (polarized). The intensity of the parallel and perpendicular emitted light was measured simultaneously using a beam splitter. The fluorescence anisotropy was calculated according to Formula 4 (p. 68) with the constant G=1.

3.5.3 Microscale thermophoresis measurements

The MST-measurements were conducted using a Monolith NT-115 Instrument (blue/green) in MST-Buffer. Ribosomes or RNCs were titrated against either 25 nM of the Alexa labeled and the unspecific labeled or against 550 nM of the NBD labeled proteins. For competition experiments, 200 nM of the NBD labeled proteins were used. Before the measurement, the sample was incubated at RT in the dark for 10 min to equilibrate the complex formation. 10 μl of the sample were loaded in hydrophilically coated microcapillaries. The LED intensity was adjusted to yield ~ 100 fluorescence counts. The IR-Laser intensity was set to 2.0 V / 80%. In dependence on the sample, the sample was heated between 20 s and 100 s until a equilibrium was reached.

3.6   Structure determination techniques

3.6.1   Cryo-electron microscopy

80 nM reconstituted complexes were applied on 2 nm pre-coated Quantifoil R3/3 holey carbon supported grids and vitrobotted using a Vitrobot Mark IV (FEI Company). The grids were stored in liquid nitrogen until further use. Data was collected under low dose cryo-conditions (20 electrons/Å²) on a Tecnai F30 field emission gun electron microscope at 300 keV (Berlin, Ultra Structure Network). The Tecnai F30 micrographs were scanned on a Heidelberg drum scanner (pixelsize 1.2375 Å per pixel). Pictures were taken at different defocus values ranging from -1.2 µm to -3.5 µm.

3.6.2   Single particle analysis and 3D reconstruction

Processing of the datasets was done using the SPIDER package, a FORTAN based environment (Frank et al., 1996). In a first step, micrographs were imported into the SPIDER environment and transformed into SPIDER files. The defocus values of individual micrographs were calculated using CTFFIND (Mindell and Grigorieff, 2003). Then, the power spectra of micrographs were inspected visually using WEB. Only micrographs showing power spectra of high quality (continuous, symmetrical rings that indicate no shift while recording, and high resolution) were used.

Individual particles on the micrographs were automatically detected using the SIGNATURE software (Chen and Grigorieff, 2007) and the coordinates were used to cut out single particles from the original micrographs. These particles were filtered to a resolution of 40 Å and five times decimated for classification using MAPPOS software (http://arxiv.org/abs/1112.3173v2). To this end, a training dataset consisting of 400 good and 400 bad particles that spanned the whole defocus range was generated by visual inspection. MAPPOS classified the unseen particles according to this subdataset and only good particles were used for further processing. Initial alignment using appropriate references was done using determination of the cross correlation between each good particles and projections of the reference. These projections were based on quasi-evenly spaced projection directions with an angular increment of 15°. Best matches were assigned as initial projections and based on the projections, individual angles were calculated. After initial alignment, the organization of the dataset was changed from micrograph based to defocus group based to improve speed and ease of handling. Therefore groups of micrographs were assigned, covering a defocus span of 200 nm. Based on the initial assignment of projections, initial backprojections using interpolation in Fourier space of the individual defocus groups were calculated. The subvolumes were then combined and ctf corrected using Wiener filtering. Resolution was calculated using the Fourier Shell Correlation technique (FSC). The dataset was split into two subdatasets that were individually backprojected. The Fourier transformed volumes were then compared along shells in 3D. An average

crosscorrelation coefficient per shell was calculated. As a criterion for resolution a FSC of 0.5 was chosen. This is equivalent to a signal to noise ratio of 2. This volume was then used for iterative alignment and backprojection to improve resolution. Therefore, the filter parameters for the reference were gradually adjusted and the angular increments narrowed. To generate homogenous sub-datasets out of a heterogenous dataset, sorting was necessary. Competitive cross correlation assignement of individual particles of two to four references yielded good results for large differences, but for smaller ligands the massive ribosome dominated the cross correlation determination. A new sorting regime called focused competitive multi-reference alignment was used in such cases. After alignment, the global cross correlation coefficient as calculated by the alignment algorithm was substituted by a local coefficient. This was achieved by i) creation of a 3D mask consisting of zeros and ones that enveloped the region of interest, ii) individual backprojections from 3D to 2D for all references and the mask for each particle with the determined angles of the particle, iii) multiplication of all 2D projections of references with the corresponding 2D projection of the mask and iv) calculation of the new cross correlation coefficient of only the region of interest. The particle was then assigned to the subdataset with the highest local cross correlation.

3.6.3 Analysis of the dataset

Due to large heterogeneities that could not be subsorted, additional methods were applied to yield the maximum information of the calculated densities. These methods contained particles based techniques like 2D classification, density map based techniques like 3D bootstrapping and local resolution determination as well as pdb-based techniques like normal mode analysis.

3.6.3.1 2D Classification

According to their angular distributions, particles were classified into projection groups with similar angles. The particles were filtered and rotated. A CORAN on all particles of one projection group followed by a k-means clustering with the expected number of classes yielded projection specific groups of particles based on selected Eigenvectors. The groups of neighbouring projection groups were compared to each other and based on the cross correlation coefficient of the average 2D image joined into subdatasets. Finally, backprojections based on the newly created selectfiles yielded the new volumes.

3.6.3.2 3D bootstrapping – finding regions of high variance

The total variance can be calculated using the bootstrap variance voxel per voxel and the background variance (Penczek et al., 2006a; Penczek et al., 2006b). To calculate the bootstrap variance, a sub data set of 16,000 particles was generated. To speed up the back projection, CTF correction was done on the particle level. To this end, 2D particles were fourier transformed, multiplied with the CTF function of their defocus value and the particles reversly fourier

transformed. For fast backprojection, BP 3N was used. 80,000 subdata sets using bootstrapping were generated and back projected. Both variances were calculated using the va3r command of the SPIDER package. The average background variance (summed over all voxels divided by the amount of voxels) was then substracted from each of the bootstrap variance voxels. The background variance was determined by circular shifting of the particles by half the window size in x and y directions. These particles were then used for bootstrap based backprojection of 500 volumes.

3.6.3.3 *Local resolution determination*

To accurately calculate the local resolution of cryo-EM density maps, the map was subdivided into smaller boxes. These boxes were not smaller than 10^3 voxels and overlapping. For each of the box, the resolution was calculated independently in the fourier space using the fourier shell correlation algorithm. The boxes were padded into bigger boxes to improve accuracy of resolution determination. Additionally, linear estimation of the FSC 0.5 point was applied to further improve accuracy (Figure 53, p. 104). The resulting local resolution numbers were then written in a volume based on the top left coordinate of the box it was calculated from. Padding of the resulting volume into a volume of the size of the input volume allowed for subsequent overlaying of the input volume with the local resolution map.

3.6.3.4 *Normal mode analysis*

Normal mode analysis was done using the nomad-ref web page (http://lorentz.immstr.pasteur.fr/nma/submission.php)

4 Results

4.1 NAC's interaction with the ribosome

Although NAC's interaction with ribosomes is known since 1996, its molecular mode of interaction with the ribosome is still not understood. To investigate this interaction in more detail, a fluorescently labeled NAC that binds wild-type-like to ribosomes was needed. Different strategies were applied to reach this goal. Microscale thermophoresis (MST), fluorescence quenching and fluorescence anisotropy measurements for ribosome – ligand interactions were established. With these tools in hand we then continued to investigate the dynamic behavior of NAC with the ribosome and intra molecular interactions of NAC subunits.

4.1.1 Purification of NAC

EGD1 contains a significant number of codons that are defined as rare codons in *E. coli*. Thus, a codon optimized EGD1 was synthesized (MWG Eurofins). A His$_6$ tag was fused via a TEV-cleavage site to the C-terminus of Egd1p. The His$_6$ tagged EGD1 and EGD2 were co expressed in *E. coli* ER 2566. After an initial metal chelat affinity chromatography (Ni-NTA), His$_6$-tagged NAC contained significant impurities. To remove the His$_6$-tag, the eluate of the Ni-NTA column was dialysed against a low salt buffer in the presence of TEV-protease (Figure 9, panel 1).

The cleaved NAC was then subjected to a cation exchange column followed by an anion exchange column. This yielded a highly pure NAC preparation (Figure 9, Panel 2 and 3). To assess the quality of the purification, a gel permeation chromatography was performed (Figure 9, Panel 4). The monodispers peak after the void volume indicated a homogenous population of NAC molecules which was verified using SDS-PAGE.

All NAC derivatives used in this study were purified according to this scheme to ensure reproducibility and a clean preparation. The yields of the purification scheme were between 1 and 2 mg/liter culture.
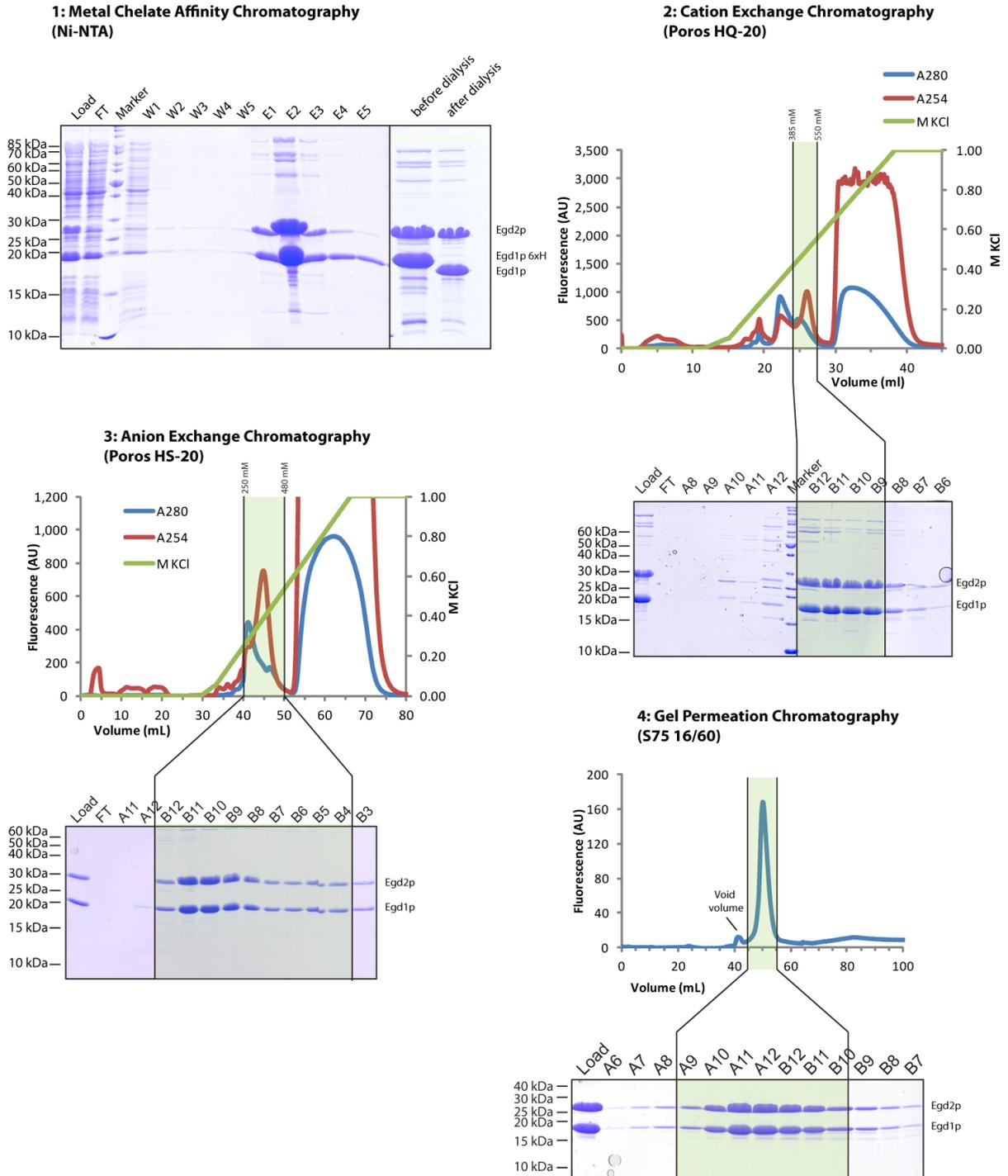
Figure 9 : The purification of heterologously expressed NAC yieled highly pure NAC | N-terminally his-tagged Egd1p and Egd2p were co-expressed in *E. coli* ER2566. After a capture step using metal chelate affinity chromatography (Panel 1), the His$_6$ tag was cleaved during dialysis against low salt buffer. NAC was purified to homogeneity using i) a cation exchange column (Panel 2) followed by ii) an anion exchange column (Panel 3) and a final polishing step using a gel permeation chromatography (S75 16/60, Panel 4).

4.1.2   Unspecific labeled NAC (NAC-547)

The prerequisite for studies using fluorescence as readouts are labeled molecules that behave like the unlabeled protein. In a first approach, an N-Hydroxysuccinimide (NHS) based coupling of dyes with the primary amines of the protein was applied. 4.5 nmol of wild type NAC in 20 mM Hepes, pH 7.4 (KOH), 100 mM KOAc, 10 mM $Mg(OAc)_2$ were labeled using the GREEN-NHS kit.

To assess the efficiency of labeling, the protein concentration was determined using a BCA assay and the concentration of the dye was measured using a nanodrop device (molecular absorption coefficient of the dye is 150,000 M^{-1} cm^{-1} at 547 nm). 2.6 nmol NAC were labeled with 4.4 nmol dye yielding an average of 1.73 dye molecules per molecule NAC.

Since the MST-experiments were done on the nanotemper device that uses capillaries for measurements, the buffer conditions had to be optimized to minimize adsorption of the NAC derivative and ribosomes to the micro-capillaries. Using fluorescently labeled 80S ribosomes and NAC-547, no measurable adsorption was detectable at the buffer conditions that were used in the subsequent analysis (data not shown).

Direct titration microscale thermophoresis measurements were performed according to section 3.6.3. To this end, 20 nM of NAC-547 was kept constant and titration series with idle yeast ribosomes were performed in MST-Buffer. The ribosome concentration was varied between 0 to 3.5 µM final concentration. The infrared laser was set to 100% and the LED excitation (green) to 80%. To calculate the relative mobility of NAC-547, the relative change in fluorescence upon the establishment of the temperature gradient was calculated. To do so, the average fluorescence of 50 time increments of 0.1 s each was measured and calculated at RT equilibrium (F_{cold}). Then, the IR-laser was switched on and the equilibrium was allowed to form for 15 s. Finally, the average fluorescence of 50 time increments was measured and calculated (F_{hot}). The relative mobility can be visualized by normalizing F_{hot} to F_{cold}. Figure 10 shows a logarithmic depiction of the relative mobility of NAC-547 in presence of different ribosome concentrations.
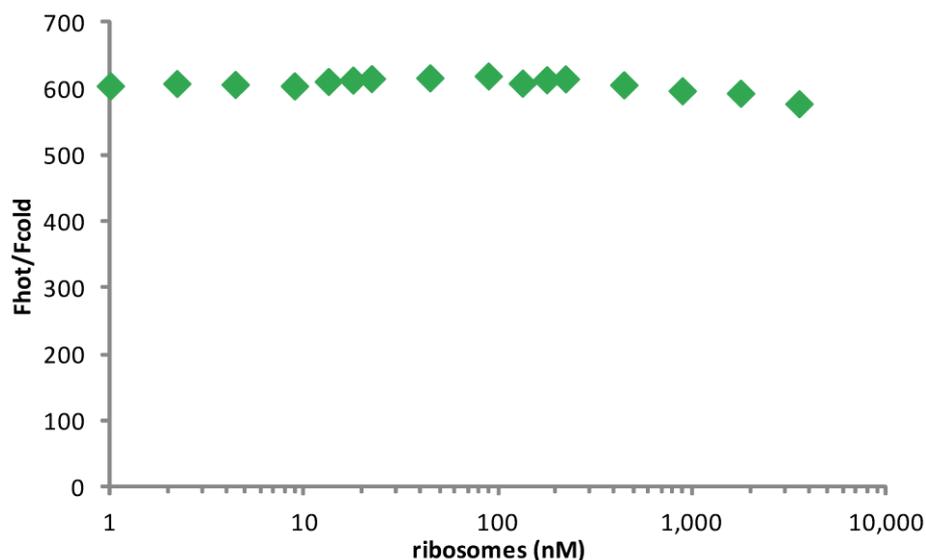
Figure 10 : NAC-547 does not bind to yeast ribosomes (MST) | Yeast ribosomes were titrated against 20 nM NAC-547. Between 2 and 3,500 nM ribosomes, no change in the thermophoretic mobility could be detected. The measurement was performed in MST-Buffer, IR-Laser 100% and the green setting.

The absence of a change in the thermophoretic mobility of NAC-547 upon titration of ribosomes hints towards an inability of this NAC-variant to bind to ribosomes. This observation was verified by binding assays that did not show any co-pelleting (data not shown).

4.1.3   Cysteine specific labeled NAC

As NAC-547 did not bind to ribosomes, we next used a different approach to label NAC. Since NAC does not code for any cysteines, cysteine reactive dyes were chosen. By placing single cysteines at defined positions, the location where the dye is linked to the protein was precisely controlled.

NACs unusual structural features – large, intrinsically disordered segments combined with rigid domains –hindered a rational identification of sites that were suited to report binding without interfering it. Therefore, four derivatives of NAC (Egd1p-I3C, Egd1p-A16C, Egd2p-V97C, Egd2p-V166C) that contained single cysteins at different functional regions were constructed. The subunits were co-expressed and natively co-purified. Additionally, two already purified, cystein containing complexes (Egd1p-C-ter containing a C-terminal cysteine and Egd1p-A142C) were used to cover all important parts of NAC.

4.1.3.1 Cysteine specific labeling of NAC

The proteins were labeled according to section 3.4.3.2. Depending on the position of the cystein, the labeling efficiencies varied. Table 1 gives an overview of the labeling efficiency.

Table 1: Overview of NAC derivatives and the labeling efficiency

				efficency of labeling
EGD1	I3C	}	Binding motif Pech et *al.*	80%
	A16C			90%
	A142C	}	C-terminus of EGD1	70%
	C-ter cysteine			20%
EGD2	V97C	—	intrinsically disorderd region	50%
	V166C	—	UBA-domain	50%

To further increase the concentration of the labeled over the unlabeled fraction, a thiol sepharose incubation step was added to the labeling procedure. The thiol sepharose matrix establishes disulfide bonds with unlabeled, free cysteines. This efficiently removes the majority of unlabelled proteins. Thereby, a maximum fluorescence signal of a given protein concentration was achieved. This additional step increased the labeled fraction up to 90% for almost all derivatives, except the Egd2p-V166C NAC variant.

4.1.3.2 MST of cysteine specific labeled NAC

The next step was to determine the dissociation constant (K_d) of the cysteine mutants. Since the Alexa dyes were best compatible with the fixed setting of the thermophoresis device, the Alexa568 (Egd1p-A16C, Egd1p-A142C) and Alexa555 (Egd1p-Cter-Cystein, Egd2p-V97C, Egd2p-V166C) derivatives were measured first. Ribosomes were titrated against 25 nM of the labeled protein and changes in the thermophoretic mobility were detected (Figure 11 shows exemplarily the measurement of Egd1p A16C-Alexa568 NAC).

The change in fluorescence can be subdivided into two parts. Immediately after the infrared laser established the temperature gradient, the dye can react fast with a change in the quantum yield or changed absorption/emission properties. This phenomenom is called temperature jump, as it reflects a direct reaction of the dye on the temperature change. After this reaction, that happens

within hundereds of milliseconds, the major determinant of fluorescence change is the migration of labeled molecules due to thermophoresis.
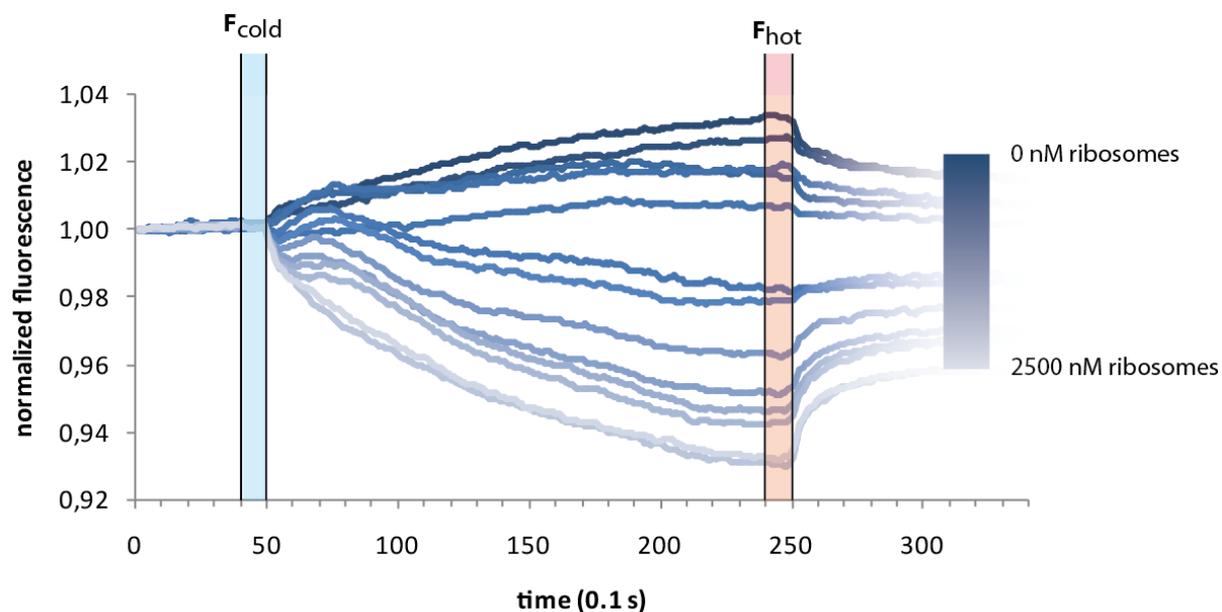
Figure 11 : Egd1p A16C-Alexa568 NAC changes its thermophoretic mobility upon titration of ribosomes | The measurement were performed in MST-Buffer with 25 nM of Egd1p A16C-Alexa568 NAC. F_{cold} was defined as the average fluorescence of 10 time increments of each 0.1 s after 4 s of the start of the experiment, F_{hot} was defined as the average fluorescence of 10 increments of 0.1 s each after 19 s of thermal gradient formation. The infrared laser was set to 2.0 V, the excitation by LED (green) to 100%.

Under low ribosome concentrations (dark blue curves in Figure 11), the temperature jump does not lead to a changed fluorescence. This is reflected in a slow increase in fluorescence at the same rate as can be seen in the later section of the measurement. This increase in fluorescence is the migration of labeled NAC into the temperature gradient. However under high ribosome concentrations (light curves in Figure 11), the temperature jump leads to a decrease in fluorescence. The initial fast decrease in fluorescence is faster than the decrease in the fluorescence due to thermophoresis and thus represents a clear reaction of the dye upon ribosome binding.

The change in the thermophoresis follows a different pattern. Free NAC (dark blue curves, Figure 11) shows an increase in fluorescence upon establishment of the temperature gradient. This is due to the thermophoretic movement of labeled molecules into the temperature gradient. However, if NAC binds to ribosomes, the fluorescence drops after the IR-Laser established the gradient (light blue curves, Figure 11). The NAC-ribosome complexe moves thermophoretically away from the hotspot.

To further analyze the binding behavior, the change in fluorescence ($F_{norm}=F_{hot}/F_{cold}$) for each individual measurement was calculated. From F_{norm} the fraction of bound molecules x can be calculated using Formula 1.

$$fraction\ of\ bound\ molecules\ (x) = \frac{F_{norm}(x) - F_{norm}(ligand\ alone)}{F_{norm}(complex) - F_{norm}(ligand\ alone)}$$

Formula 1: The fraction of bound molecules can be calculated based on the Fnorm | If the normalized fluorescence of the ligand alone (F_{norm} (ligand alone) and the normalized fluorescence of the complex (F_{norm} (complex)) is known, the fraction of bound molecules can be calculated.

Once the fraction of bound molecules is known for each point of a titration series, a binding curve can be plotted. Based on this plot, an equation that describes the fraction of molecules bound as a function of the titrated partner can be fitted to the dataset (Formula 2 - (5)). The concentration of the labeled molecule C and the concentration of the titrated molecule x are known, while the variable K_d is adjusted.

For most of the derivatives, a K_d could be calculated. Figure 12 shows the binding curve of selected Egd1p mutants (Egd1p-A16C NAC, Egd1p-A142C NAC, Egd1p-C-ter Cys NAC) to demonstrate the differences in dissociation affinities between different cysteine mutants and how this translates into distinctive binding curves.

$$K_d = \frac{c(A) * c(T)}{c(AT)}$$

(1)

$$c(A) = c_0(A) - c(AT) * n_A \qquad c(T) = c_0(T) - c(AT) * n_T \qquad n_A = n_T = 1$$

(2)

$$K_d = \frac{\big(c_0(A) - c(AT)\big) * \big(c_0(T) - c(AT)\big)}{c(AT)}$$

(3)

$$fraction\ bound\ (A) = \frac{c(AT)}{c_0(A)}$$

(4)

$$fraction\ bound\ (A) = \frac{(A + T + K_d - \sqrt[2]{A + T + K_d)^2 - 4 * A * T})}{2A}$$

(5)

Formula 2: The fraction of bound molecules of T can be calculated based on the dissociation constant K_d, the concentration of the constant partner A and the concentration of the molecule to be titrated T | The dissociation constant K_d is the ratio of the product of the concentrations of free A times free T divided by the concentration of the complex AT (1). The concentration of free A is the starting concentration of A (which is known) minus the concentration of the complex AT (which is known). The same holds true for the concentration of free T (2). nX depicts the stoichiometry of the complex: n=1 indicates a 1:1 stoichiometry. This information transforms the starting Kd formula in an quadratic equation (3) that can be solved using the information that the fraction of bound A is the concentration of the complex divided by the starting concentration of A (4). This formula is the possible solution for the quadratic equation (5).
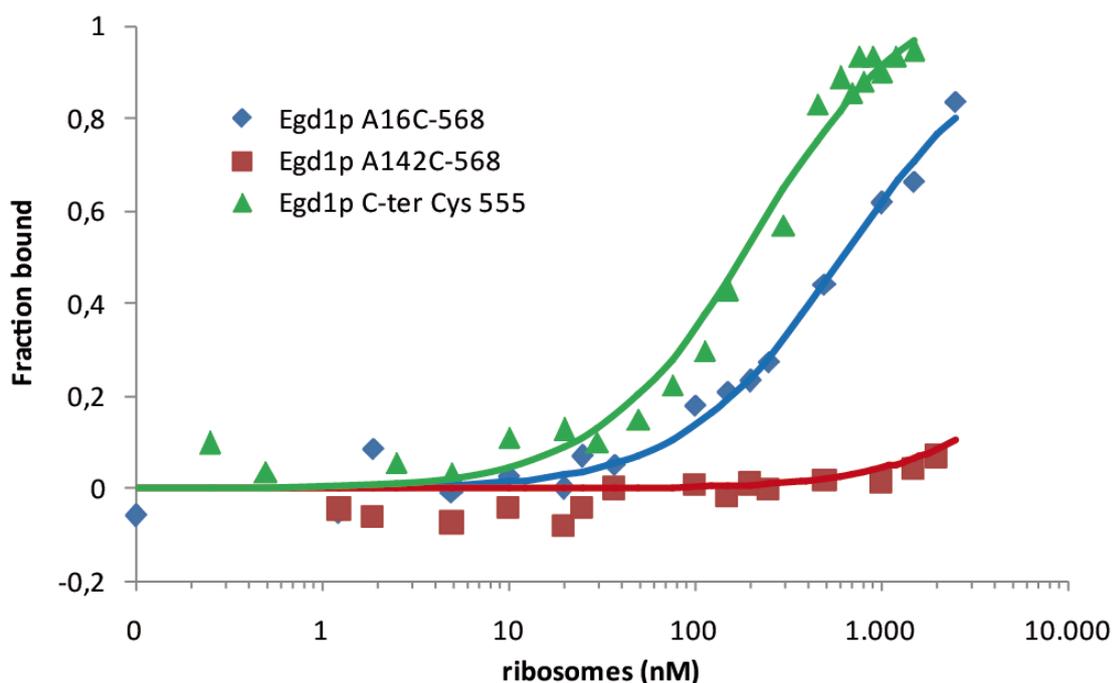
Figure 12 : Binding curves of selected cysteine mutants (MST) | Yeast ribosomes were titrated against different NAC variants. The dissociation constants were calculated using automated curve fitting of Formula 4. For visualization, not the actual F_{hot}/F_{cold} ratio but the fraction bound is plotted to account for the difference of thermophoretic mobility of the individual NAC derivatives. The calculated K_ds are 608 nM for Egd1p-A16C Alexa568, 21 µM for Egd1p A142C-Alexa568 and 200 nM for Egd1p C-ter Cys-Alexa555. All experiments were conducted in MST-Buffer.

Egd1p-A16C Alexa568 NAC exhibits a K_d of 605 nM, while Egd1p-A142C Alexa568 NAC and Egd1p-C-ter Cys Alexa555 NAC bind to ribosomes with a K_d of 21 µM and 200 nM, respectively.

The difference between the K_ds of different NAC-derivatives raised the question which of them behaves like the unlabelled, wild-type NAC. To systematically address this question, we measured the Egd1p-A16C NAC variant, labeled with a different dye with distinct properties - NBD.

While Alexa568 is a bulky dye and due to the sulfonation negatively charged and hydrophilic, NBD is in comparison a rather small, uncharged dye. If the dye at position A16C has an influence on the binding to ribosomes, such drastic changes in physical properties of the dye moiety might translate into a changed binding behavior.

NBD is due to its spectral properties not as suitable for the MST-setup as the Alexa dyes. Therefore, higher concentrations of the labeled NAC derivative had to be used. To measure the affinity of Egd1p A16C-NBD NAC to ribosomes, ribosomes were titrated against 550 nM Egd1p A16C-NBD NAC. F_{hot} and F_{cold} were measured within time spans of 10 time increments of 0.1 seconds each. Equilibrium was reached 19 s after the temperature gradients was established. In contrast to the Egd1p-A16C Alexa568 NAC setup, the fluorescence decrease in every measurement, irrespective of the ribosome

concentration. Additionally, every measurement showed a clear temperature jump. The amplitude seems to be dependent on the ribosome concentration.
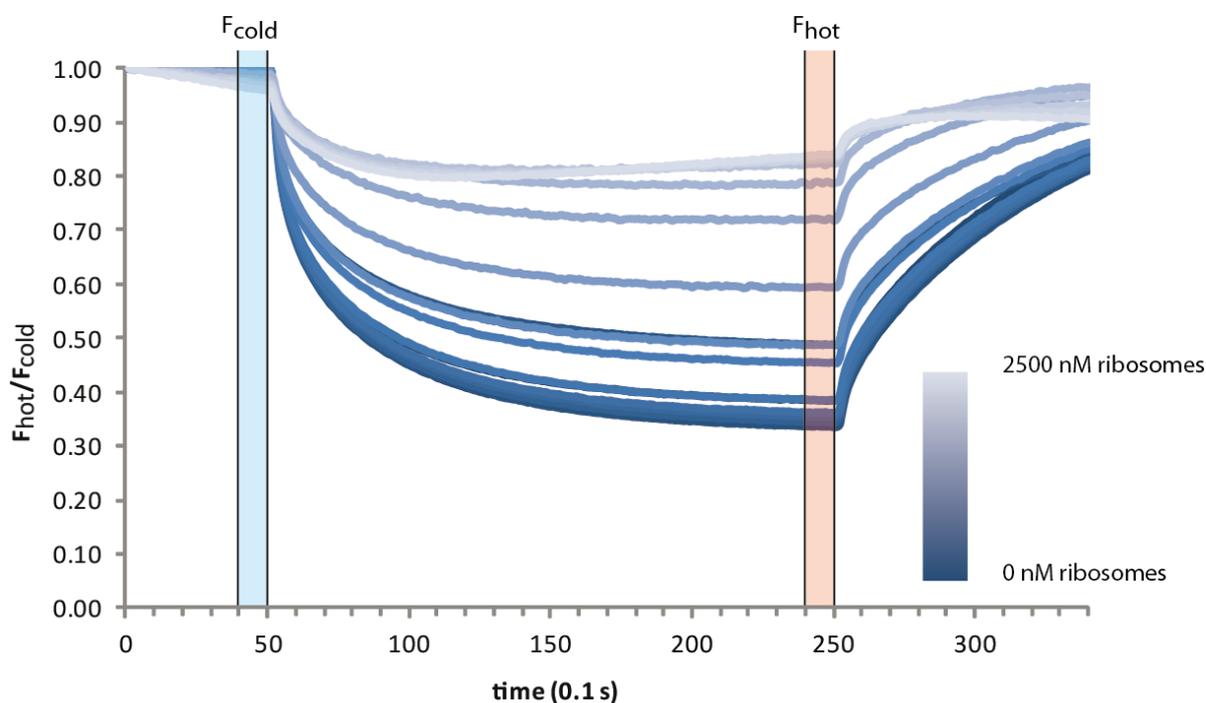
Figure 13 : Egd1p A16C-NBD NAC changes its thermophoretic mobility upon titration of ribosomes (MST) | The signal of this NAC-derivative is around 6 times higher in comparison to Egd1 A16C-Alexa568. Additionally, the change in the temperature jump is higher. The measurements were performed in MST-Buffer. F_{cold} was defined as the average fluorescence of 10 time increments of each 0.1 s after 4 s of the start of the experiment, F_{hot} was defined as the average fluorescence of 10 increments of 0.1 s each after 20 s of thermo gradient formation. The infrared laser was set to 2.0 V, the excitation by LED (green) to 100%.

Initial plotting of the normalized fluorescence (F_{hot}/F_{cold}) against the ribosome concentration showed, that the point where 50% of Egd1p-A16C NBD NAC is bound to ribosomes is very close to half of the constant concentration of Egd1p-A16C NBD NA. This hinted towards a K_d that was significantly lower than the constant concentration of labeled NAC. The deduction of K_ds from such setups is very prone to errors. To increase the confidence to this setup, independent triplicates were measured and combined into one binding curve (Figure 14).
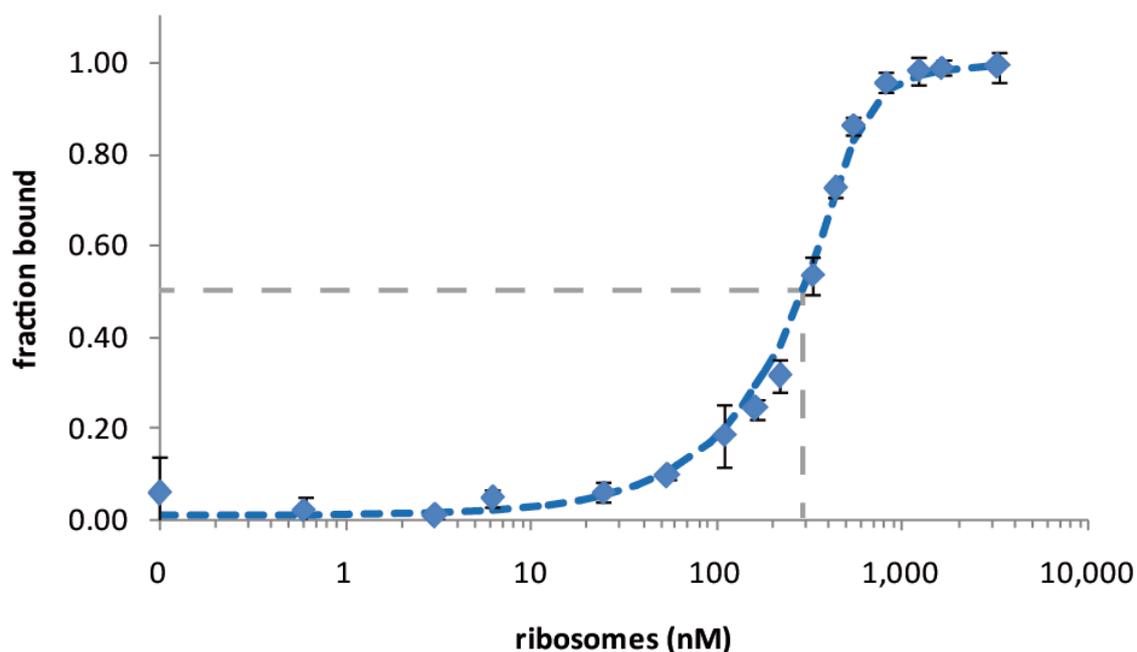
Figure 14 : Egd1p A16C-NBD NAC binds with a high affinity to idle ribosomes (MST) | *S. cerevisiae* ribosomes were titrated against 550 nM Egd1p-A16C NBD NAC in MST-Buffer. Each measurement was independently repeated three times. For better visualization, not the actual F_{hot}/F_{cold} ratio but the fraction bound is plotted. The fraction bound was based on the normalized fluorescence and calculated using Formula 1. The dissociation constant was calculated using automated curve fitting of Formula 2. The curve fit (dashed blue lines) revealed a dissociation constant of 20 nM. The grey dashed lines indicate the ribosomal concentration, were 50 % Egd1-A16C-NBD NAC is complexed (290 nM).

The K_d of Egd1p-A16C NBD NAC to idle ribosomes is 20 nM. This is 30 times higher than the affinity of Egd1p-A16C Alexa568 NAC to ribosomes (600 nM, Figure 12). Since the only difference between these two NAC derivatives is the chemical nature of the dye, the influence of dye to the binding to ribosome seems to be stronger then the influence of the mutation to cysteins.

Next, we determined the K_ds of all the constructed NAC mutants (Table 1) to see, if a population of derivatives that behaves similar exists. This would hint towards a wildtype like behaviour. The dissociations constants of the NAC derivatives covered a range of over three orders of magnitude (Egd1p-A16C NBD: 20 nM, Egd1p-A142C Alexa568: 21 µM; Figure 15). For the Egd2p derivatives, no dissociation constant could be determines as we could not see binding of the labeled derivatives.
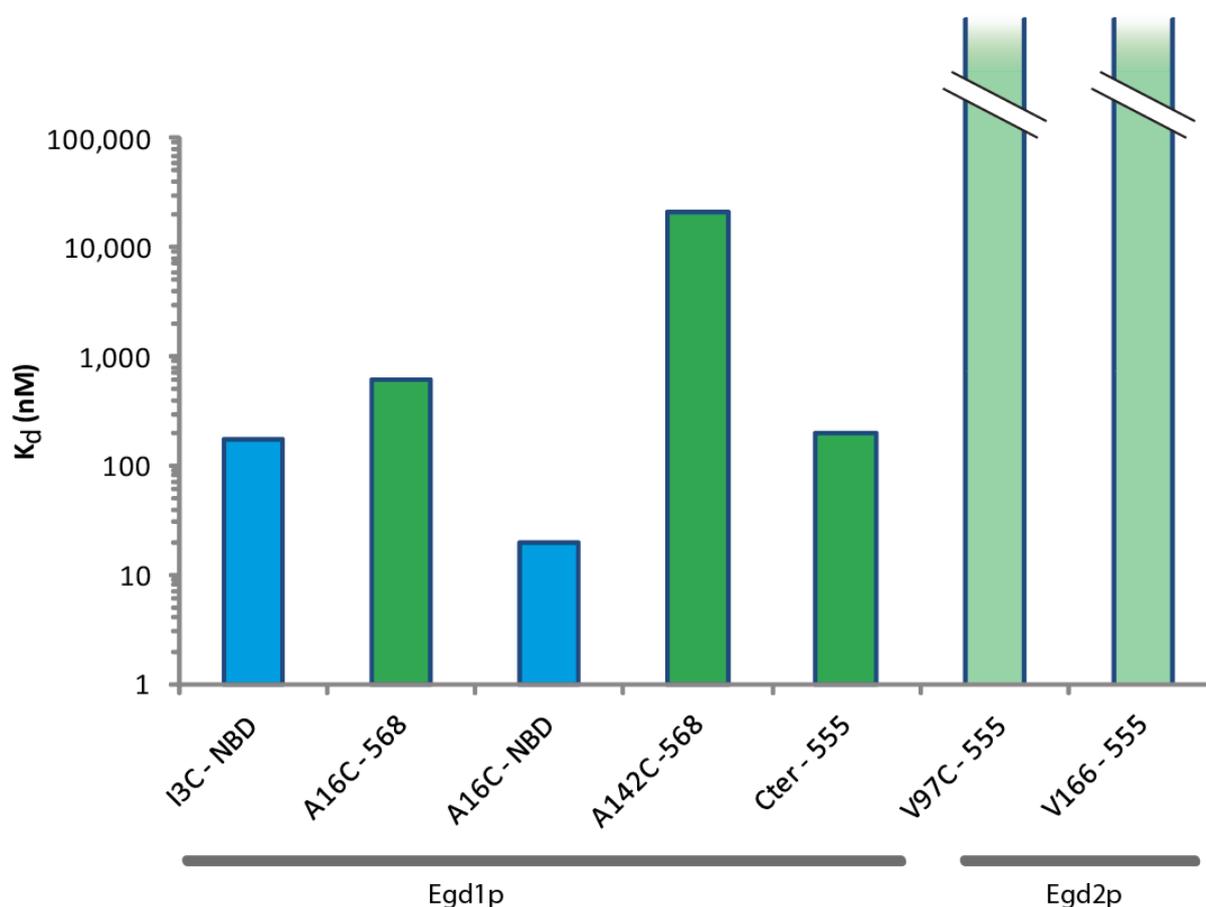
Figure 15 : Overview of Affinities of different NAC-derivatives (MST) | The bars indicate the MST-determined K_ds of the respective derivative. The color of the bars indicate which setup of the MST device was used to determine the binding affinities.

4.1.3.3 Verifying the wild type like behavior of Egd1p-A16C NAC NBD

The divergences of the dissociation constants lead to a change of strategy to identify wild type like behavior. Instead of recording the direct titration of ribosomes against NAC derivatives, we measured the displacement of labeled NAC by unlabeled NAC from ribosomes. To do so, a complex of the labeled NAC derivatives and ribosomes was allowed to form. Then, wild type NAC was titrated against this complex.

Roehrl et al. 2004 developed the theory for these competition assays (Roehrl et al., 2004). The fraction of bound, labeled molecules can be calculated using the known concentration of labeled ligand (=NAC, C), unlabelled ligand (=NAC, U), unlabelled receptor (=ribosome, x) and the dissociation constants of the labeled and the unlabeled ligand to the receptor (K_{d1} and K_{d2}, respectively) (Formula 3).

$$fraction\ of\ bound, labeled\ molecules\ (U) = \frac{2\sqrt[2]{(d^2 - 3e)}\cos(\theta/3) - d}{3K_{d1} + 2\sqrt[2]{(d^2 - 3e)}\cos(\theta/3) - d}$$

with

$$d = K_{d1} + K_{d2} + C + U - x$$

$$e = (U - x)K_{d1} + (C - x)K_{d2} + K_{d1}K_{d2}$$

$$f = -K_{d1}K_{d2}x$$

$$\theta = \cos^{-1}\left[\frac{-2d^3 + 9de - 27f}{\sqrt[2]{(d^2 - 3e)^3}}\right]$$

Formula 3: The fraction of labeled, bound molecules in a competition experiment can be calculated based on the dissociation constants Kd, the concentration of the labeled (C) and the unlabelled partners (U) and the concentration of ribosomes (x)| This formula describes a complete competitive binding model in which the ribosome (x) either binds the labeled (C) or the unlabeled (U) ligand. K_d1 is the dissociation constant for the system C-x and K_d2 for the system U-x. The nonspecific immobilization of the unlabeled ligand was too weak to influence the outcome of the calculation, and thus this effect was not taken into account.

In our setup, we know all variables except K_{d2}. We can now use this formula and fit it, as a function of the unlabeled ligand, to the competition data. If labeled NAC behaves like wild type NAC, The formula should describe the data if K_{d1} equals K_{d2}.

Since we saw a strong, negative effect of Alexa568 dyes on the affinity of NAC to ribosomes, we started with the analysis of Egd1p-I3C NBD NAC. 400 nM of Egd1p-I3C NBD and 400 nM idle ribosomes were preincubated in MST-Buffer. After 10 min incubation at RT the same volume of a 2 x master mix of wild type NAC in the same buffer was added to the experiment and the mixture was incubated for 10 min at RT in the dark.

The infrared laser was set to 100% and the LED excitation (blue) to 100%.. F_{hot} and F_{cold} were measured within time spans of 10 time increments of 0.1 seconds each. The Equilibrium was reached 19 s after the temperature gradient was established. Egd1p-I3C NBD NAC bound to ribosomes was displaced by wt NAC (Figure 16).

200 nM Egd1p-I3C NBD NAC and 200 nM ribosomes form 80 nM of Egd1p-I3C NBD NAC – ribosome complex (Fraction bound: 0.4). Addition of 1 µM wt NAC leads to an almost complete displacement of Egd1p-I3C NBD NAC from the ribosome.

The blue curve in Figure 16 represents a fit of Formula 3 to the data where the K_d of Egd1p-I3C NBD NAC to ribosomes was set to 170 nM and K_{d2} was the fitting parameter. The curve only describes the data, if the K_d of wt NAC to ribosomes is set to 20 nM. The green curve is a simulation of a situation where $K_{d1} = K_{d2} = 170$ nM.
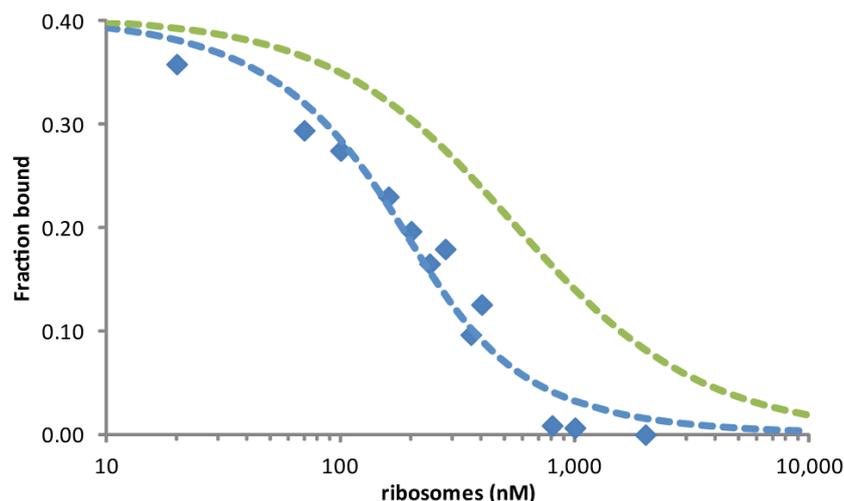
Figure 16 : Egd1p-I3C NBD NAC is displaced by wt NAC (MST) | 200 nM of Egd1p-I3C NBD NAC and 200 nM idle yeast ribosomes were mixed in MST-Buffer. After 10 min, wt NAC was added and the reaction incubated for 20 min at RT. The thermophoretic mobility and temperature jump reaction was recorded by normalizing the average fluorescence of 10 time increments of 0.1 s each after establishment of the equilibrium (F_{hot}) to the average fluorescence of 10 time increments of 0.1 s each before the IR-Laser was switched on. The affinity of Egd1p-I3C NBD NAC to empty ribosomes is known (170 nM). This allowed the definition of the normalized fluorescence in the absence of wt NAC to 0.40 fraction bound. The relative mobility of free Egd1p-I3C NBD NAC is also known. This is defined as Fraction bound of 0. The individual normalized fluorescence can then be assigned to a fraction bound using an adjustment of Formula 3. The blue curve is the fit of Formula 5 to the data; K_d1 was set to 170 nM, C to 200 nM and x to 200 nM. K_d2 was calculated to be 20 nM. The green curve is the simulation of Formula 5 where K_{d1} = K_{d2} = 170 nM.

The finding that the green curve did not explain the measured data makes it unlikely that Egd1p-I3C NBD NAC behaves like wtNAC. The information that the blue curve, where the K_d for wt NAC had to be set to 20 nM to explains the measurements, however is useful. It indicates that wt NAC binds to empty ribosomes with a K_d of 20 nM.

Since the K_d of Egd1p-A16C NBD NAC to empty ribosomes is also 20 nM (Figure 14), we next tested the displacement of Egd1p-A16C NBD NAC by wt NAC. The experimental details were identical to the Egd1p-I3C NBD NAC setup. Due to the higher affinity to empty ribosomes, 200 nM Egd1p-A16C NBD NAC and 200 nM ribosomes form in the absence of wt NAC 146 nM of the NAC-ribosome complex (fraction bound 0.73).
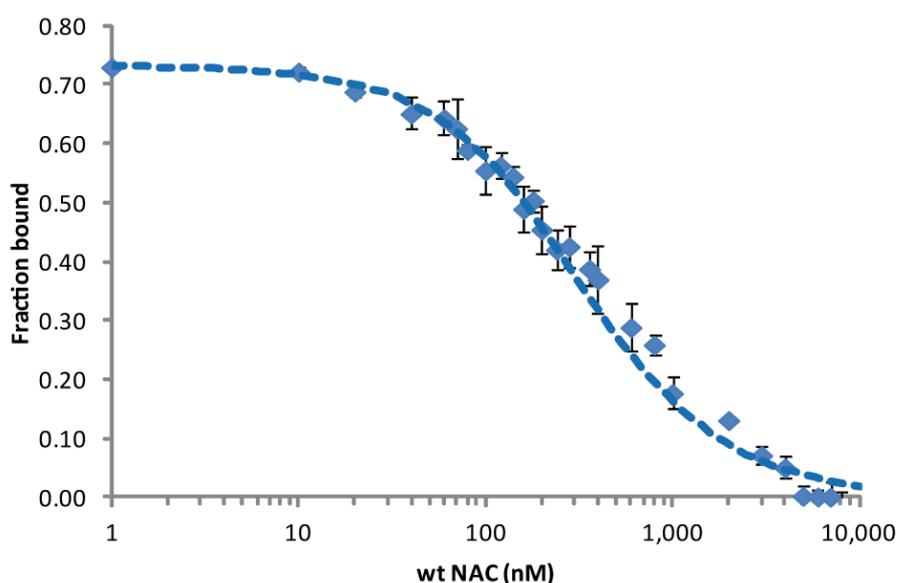
Figure 17 : Egd1p-A16C NBD NAC is displaced by wt NAC (MST) | 200 nM of Egd1p-A16C NBD NAC and 200 nM idle yeast ribosomes were mixed MST-Buffer. After 10 min, wt NAC was added and the reaction incubated for 20 min at RT. The thermophoretic mobility and temperature jump reaction was recorded by normalizing the average fluorescence of 10 time increments of 0.1 s each after establishment of the equilibrium (F_{hot}) to the average fluorescence of 10 time increments of 0.1 s each before the IR-Laser was switched on. The affinity of Egd1p-A16C NBD NAC to empty ribosomes is known (20 nM). This allowed the definition of the normalized fluorescence in the absence of wt NAC to 0.73 fraction bound. The relative mobility of free Egd1p-A16C NBD NAC is also known. This is defined as Fraction bound of 0. The individual normalized fluorescence can then be assigned to a fraction bound using an adjustment of Formula 3. All measurements were carried out as independent triplicates. The blue curve is the fit of Formula 3 to the data; K_d1 was set to 20 nM, C to 200 nM and x to 200 nM. K_d2 was calculated to be 25 nM.

The blue curve in Figure 17 is the fit of Formula 3 to the Egd1p-A16C NBD NAC data. It explains the measured data in a scenario where the K_d of the labeled and unlabeled ligand to the ribosome is the same: 20 nM. This indicates, that Egd1p-A16C NBD NAC i) binds to empty ribosome with an affinity comparable to wt NAC and ii) it competes with wt NAC for the binding site and thus makes it unlikely that the dye leads to an artificial binding with the same affinity.

4.1.3.4   *Fluorescence anisotropy measurements with Egd1p-A16C NBD NAC*

Microscale thermophoresis is a recently developed method to measure complex formation in a variety of setups. However, no studies measuring the interaction of proteins with macromolecules like ribosomes using MST are available so far. Additionally, the used concentrations of Egd1p-A16C NBD NAC (550 nM in the direct titration and 200 nM in the competition experiments) are relatively high for high affinity systems like the NAC-ribosome interactions. To verify the finding that Egd1p-A16C NBD NAC binds with a K_d of 20 nM to idle ribosomes, to detect systematical errors in the setup and to measure lower concentrations of Egd1p-A16C NBD NAC, we applied fluorescence anisotropy measurements to this system.

S. cerevisiae ribosomes were titrated against 100 nM Egd1p-A16C NBD in MST-Buffer. The Egd1p-A16C NBD NAC dye was excited using a polarized light source (488 nm). The intensity of the parallel ($I_{parallel}$) and the perpendicular ($I_{perpendicular}$) emission was recorded. These intensities can be transformed to anisotropy (Formula 4).

$$anisotropy = \frac{(I_{parallel} - I_{perpendicular} * G)}{I_{parallel} + 2 * I_{perpendicular} * G}$$

Formula 4: The fluorescence anisotropy is based on the parallel and perpendicular intensity of polarly excited fluorophores | As the emission of photons upon excitation is not immediate, the tumbling of the fluorophore between absorption of polarized photons and emission leads to a reduction in the polarization of the emitted light. The faster the fluorophore tumbles, the higher is the intensity of $I_{perpendicular}$ and the lower $I_{parallel}$. The complex formation between Egd1p-A16C NBD NAC and ribosome leads to an increase in mass and thus a decrease in tumbling speed.
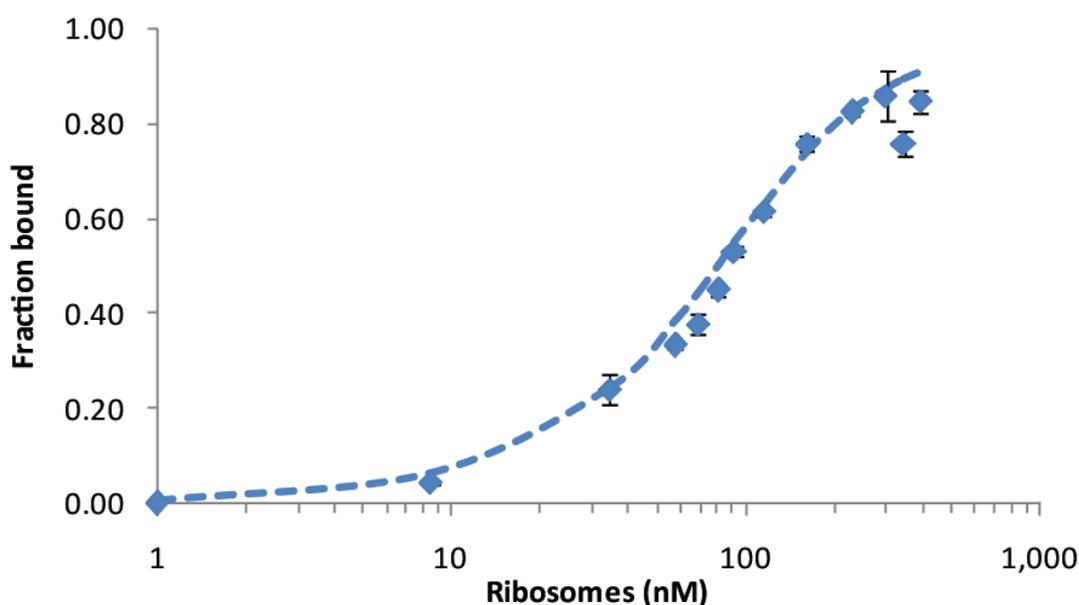
Figure 18: Fluorescence anisotropy measurements of the binding of Egd1p-A16C NBD NAC to ribosomes | Ribosomes were titrated against 100 nM Egd1p-A16C NBD NAC in MST-Buffer. Upon excitation with polarized light, the Intensity of the parallel ($I_{parallel}$) and the perpendicular fluorescence was measured. The anisotropy is calculated using Formula 4. For visualization, not the fluorescence anisotropy is plotted, but the calculated fraction bound. The fraction bound was calculated using a derivative Formula 1 where F_{norm} was replaced by the fluorescence anisotropy values.

The binding curve indicates a K_d of 30 nM (Figure 18). Thus, the binding affinities of the MST- and of the fluorescence anisotropy analysis of Egd1p-A16C NBD NAC support each other.

4.1.4 Fluorescence quenching measurements

The MST- and fluorescence anisotropy experiments used the Egd1p-A16C NAC derivative labeled with NBD. This derivative has a cysteine placed directly C-terminally of an α-helix that is believed to

mediate binding to the ribosome. This vicinity to the binding event might lead to a change in the microenvironment of the dye that could be detected by altered fluorescence intensity. This effect has been used to study the interaction of TF to prokaryotic ribosomes in previous studies (Kaiser et al., 2006).

We recorded intensity of 550 nM Egd1p-A16C NBD NAC as a function of ribosome concentration (Figure 19). A marked decrease in fluorescence upon ribosome titration was measured. However, a strong auto fluorescence signal from the ribosomes overlaid the primary signal at ribosome concentrations > 1 µM.
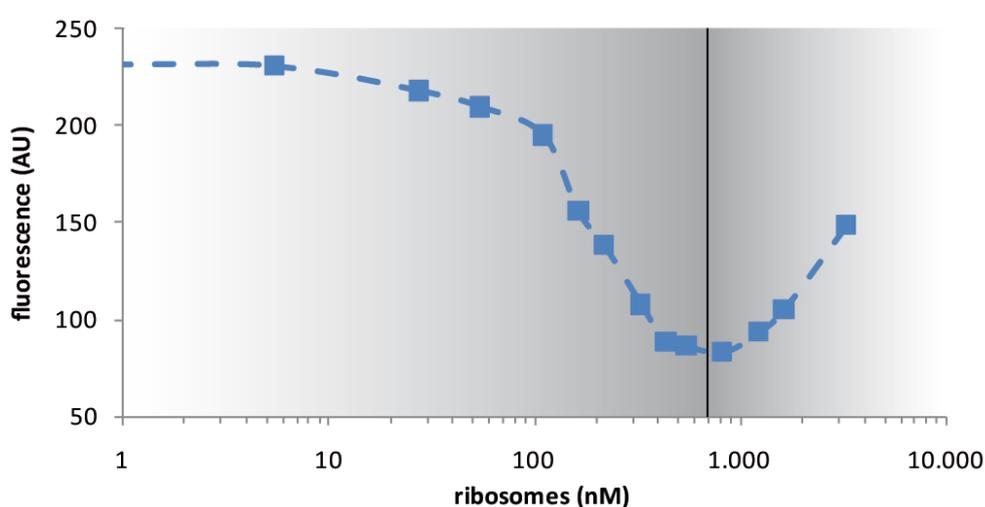
Figure 19: The fluorescence of Egd1p-A16C NBD NAC changes upon ribosome titration | Ribosomes were titrated against 550 nM Egd1p-A16C NBD NAC in MST-Buffer. The fluorescence intensity decreases with increasing ribosome concentrations until a plateau. Then the auto fluorescence intensity of ribosomes is stronger than the fluorescence of the NAC derivative.

The ribosomal auto-fluorescence did not allow for a quantitative interpretation of the binding event, but gave new insights into the microenvironment of the dye after binding to ribosomes.

NBD is an environmentally sensitive dye that changes its quantum yield and thus fluorescence intensity in response to its microenvironment. A decrease in fluorescence can be achieved by either active quenching or loss of activation of the fluorophore upon binding. To discriminate between these two hypothesis, a setup was developed that compared the fluorescence intensity of 200 nM free NBD and of 200 nM Egd1p-A16C NBD NAC under different Tween-20 concentrations (0.01 – 2 % Tween-20 (80 µM – 16 mM)). Additionally, the influence of 200 nM ribosomes (fraction bound: 0.73) on the fluorescence intensity was measured.

Tween-20 provides a hydrophobic environment for the dye and an activation of free NBD was clearly visible (green bars, Figure 20). However, Egd1p-A16C NBD NAC's fluorescence was already relatively

high under 0.01 % Tween-20 (red bars). Only high concentrations of Tween-20 showed a slight increase in the fluorescence intensity.

Under low Tween-20 concentration, the fluorescence intensity of Egd1p-A16C NBD NAC in the presence of ribosomes was markedly reduced, however still higher then NBD in the free form (pink bars). In complex with ribosomes, the fluorescence of Egd1p-A16C NBD NAC reacts to higher Tween-20 concentration with a increased fluorescence intensity only at Tween-20 concentrations higher than 0.5 %.
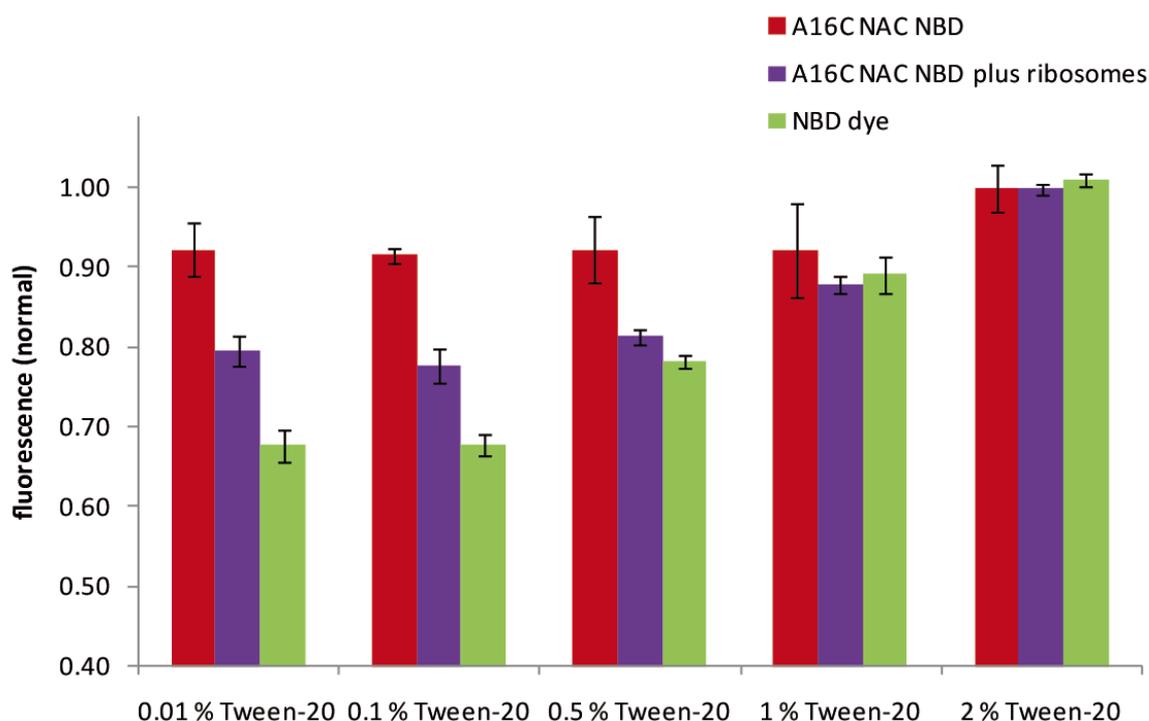
Figure 20: The fluorescence intensity of free NBD dye gets activated by Tween-20, while Egd1p-A16C NBD NAC does not react to Tween-20 titration | All measurements were performed in triplicates. The samples (200 nM Egd1p-I3C NAC NBD, 200 nM ribosomes) were incubated in MST-Buffer and the indicated Tween-20 concentration. Due to the quenching effect in the presence of ribosome, the fluorescence intensity for each measurement series was normalized to 2 % Tween-20. Free NBD was quenched with 10 mM ß-mercaptoethanol before use.

These data support a model in which the NBD moiety of the Egd1p-A16C NBD NAC derivative is located in a hydrophobic environment when NAC is not bound to ribosomes. This environment activates NBD and leads to a high fluorescence intensity. Upon ribosome binding, the N-terminus von Egd1p is not in contact with this environment. NBD's fluorescence intensity reacts to the changing environment in by addition of Tween-20.

To exclude the possibility of artifacts introduced by high Tween-20 concentrations, a competition experiment between Egd1p-A16C NBD NAC and wt NAC to ribosomes in MST-Buffer plus 2 % Tween

was performed. The K_d of NAC to ribosomes in the presence of 2% Tween-20 is identical to 0.01 % Tween-20 (data not shown).

We further extended this analysis to the Egd1p-I3C NBD NAC derivative. Although we showed that this derivative does not bind to ribosomes like wt NAC, it has an interesting features that helps to investigate this system in more detail. A structure of the N-terminus of Egd1p is not available, however structure prediction programs predict for this region an α-helix. Modeling of an α-helix with the amino acid sequence of Egd1p 1-16 showed, that the side chains of I3C and A16C point towards opposite directions (Figure 21).
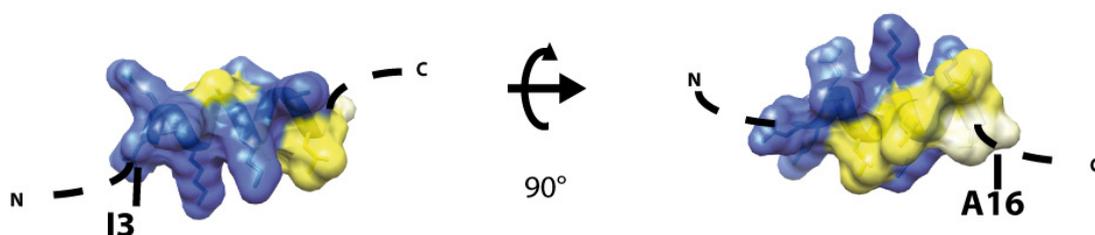
Figure 21: The N-terminus of Egd1p forms an α-helix where the amino acids I3 and A16 point towards different directions | Using Chimera, an α-helix consisting of the first 16 amino acids was modeled. I3 and A16 point towards different directions. Yellow indicates hydrophobic and blue hydrophilic side chains.

Thus, Egd1p-I3C NBD NAC reports on the microenvironment of the α-helix on one side and Egd1p-A16C NBD NAC on the other. Again, we measured the reaction of 200 nM NBD or 200 nM Egd1p-I3C NBD to increasing concentrations of Tween-20. As the K_d of Egd1p-I3C NBD NAC to ribosomes is higher than Egd1p-A16C NBD NAC, the fraction bound of Egdp-I3C NBD NAC to ribosomes is reduced to 0.4 (Figure 15).

NBD alone (green bars) and Egd1p-I3C NBD NAC bound to ribosomes react similar to Tween-20 like Egd1p-A16C NBD NAC bound to ribosomes. However, Egd1p-I3C NBD NAC without ribosomes showed a fluorescence intensity change to increasing Tween-20 concentrations of the same magnitude like free NBD (Figure 22). These data indicate that the NBD fluorophore of Egd1p-I3C NBD NAC is not preactivated like Egd1p-A16C NBD NAC.

To identify the region that activates the NBD dye of Egd1p-A16C NBD NAC in the unbound state, the possibility that the immediate vicinity of the dye interferes with the NBD dye at A16C had to be excluded first. Therefore, a fusion between the 39 N-terminal amino acids of Egd1p and maltose binding protein (MBP) as an inert carrier via a linker was constructed. The Alanin 16 of this fusion was mutated to Cystein and labeled with NBD (Egd1p[1-39] NBD MBP). In this setup, the NBD dye

experiences the same direct surrounding like in the Egd1p-A16C NBD NAC derivative but lacks possible long distance interactions with the remaining of the NAC protein.
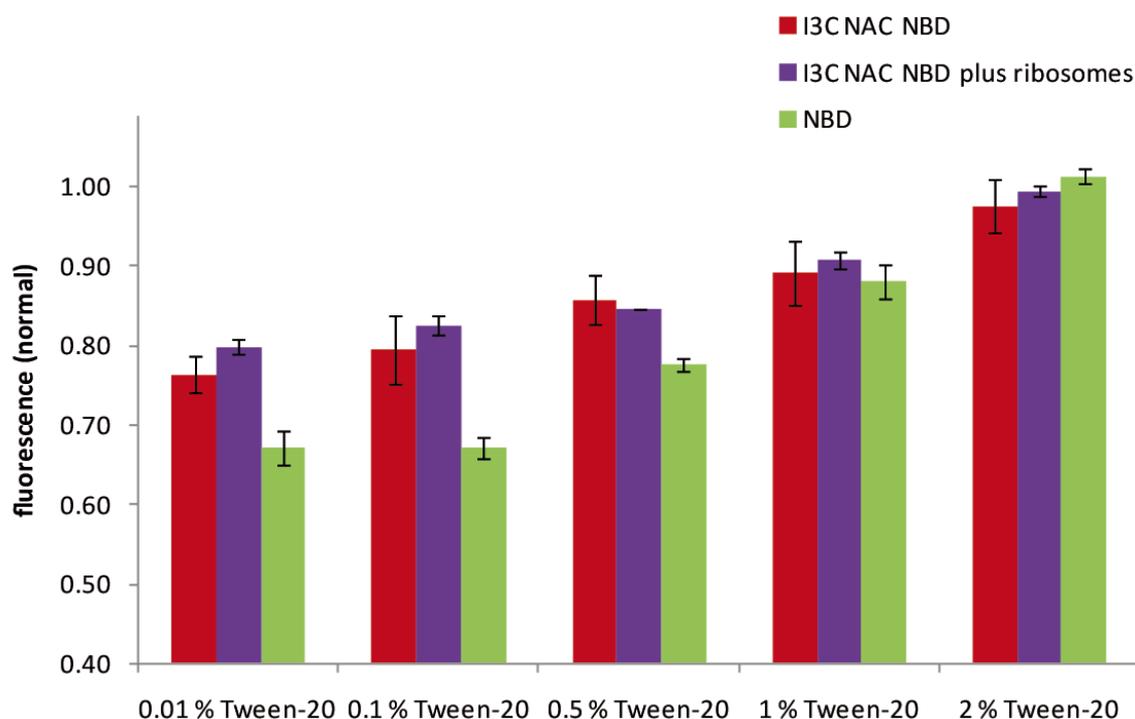
Figure 22: The fluorescence intensity of free NBD and Egd1p-I3C NBD NAC in the presence and absence of ribosomes gets activated by Tween-20 | The samples (200 nM Egd1p-I3C NAC NBD, 200 nM ribosomes) were incubated in MST-Buffer and the indicated Tween-20 concentration. Due to the quenching effect in the presence of ribosome, the fluorescence intensity for each measurement series was normalized to 2 % Tween-20. Free NBD was quenched with 10 mM ß-mercaptoethanol before use. All measurements were performed in triplicate.

The effect of Tween-20 on the fluorescence intensity of unbound and bound Egd1p[1-39] NBD MBP was then measured. 200 nM of Egd1p[1-39] NBD MBP (and 200 nM ribosomes) in MST Puffer with different Tween-20 concentration were used. In the ribosome bound state 46% of the fusion protein was bound to ribosomes (see section 4.1.5).

The fluorescence-intensity of the fusion protein changes similar to Egd1p-I3C NBD NAC – as a function of Tween-20 in the unbound and the bound state (Figure 23). This means that the NBD dye is not preactivated and the cause for the preactivation of Egd1p-A16C NBD NAC must be caused by the remaining NAC portions.

The UBA-domain of Egd2p exhibits a hydrophobic surface patch that could be responsible for the activation for Egd1p-A16C NBD NAC. To test this hypothesis, a deletion mutant of the UBA-domain based on the Egd1p-A16C NAC derivative was constructed and labeled with NBD. The reaction of the fluorescence intensity to titration of Tween-20 was recorded.
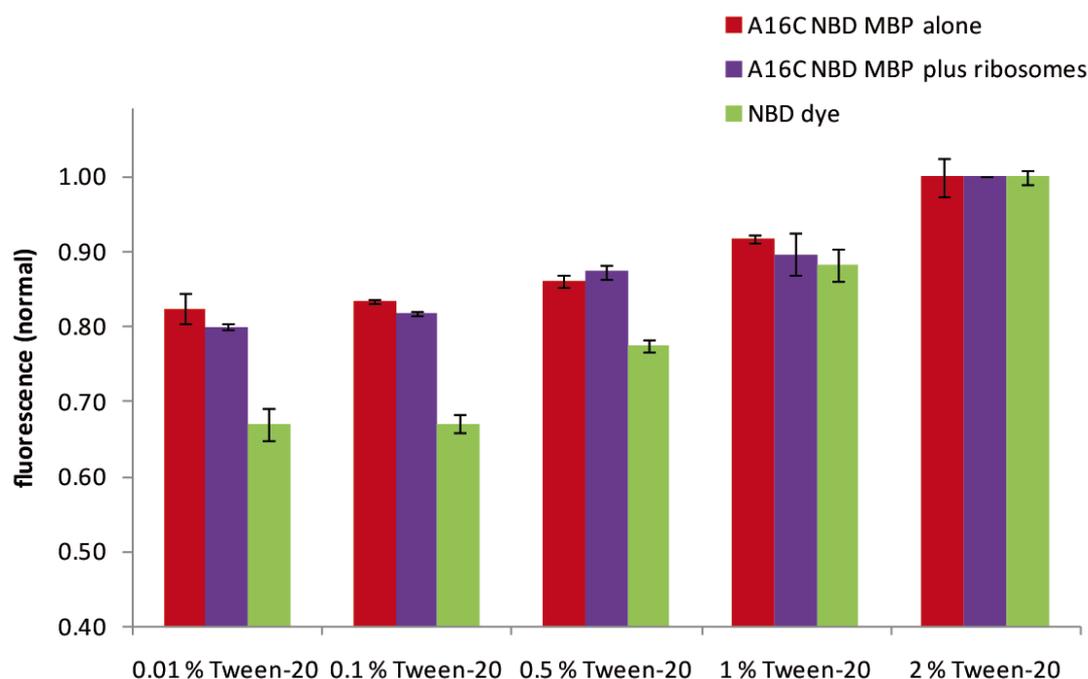
Figure 23: Free NBD dye and Egd1p (1-39)-A16C NBD MBP get activated by Tween-20 | The fluorescence intensity of 200 nM of the fluorescently labeled Fusion protein alone and with 200 nM ribosome in MST-Buffer under changing Tween-20 concentrations was recorded. The measurements were conducted in duplicates. The fluorescence was normalized to the fluorescence at 2% Tween-20.
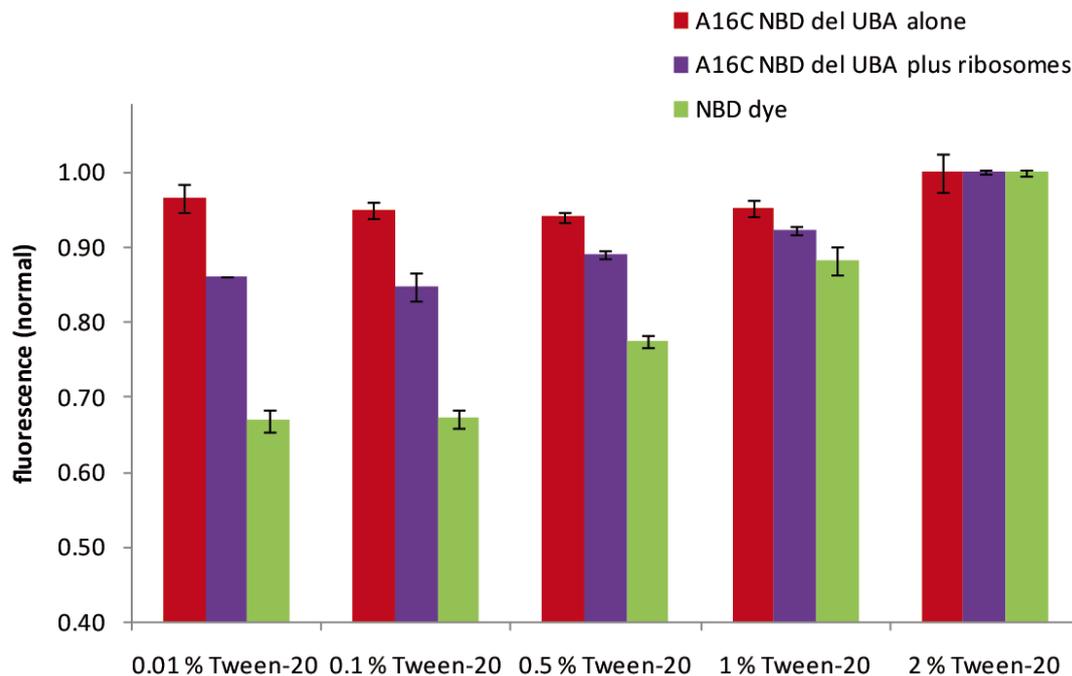
Figure 24: The fluorescence intensity of free NBD dye is increased by the presence of Tween-20, while Egd1p-A16C NBD NAC ΔUBA does not react to Tween-20 titration | The samples (200 nM Egd1p-A16C NAC NBD, 200 nM ribosomes) were incubated in MST-Buffer and the indicated Tween-20 concentration. Due to the quenching effect in the presence of ribosome, the fluorescence intensity for each measurement series was normalized to 2 % Tween-20. Free NBD was quenched with 10 mM ß-mercaptoethanol before use. All measurements were performed in triplicates.

The fluorescence change of unbound Egd1p-A16C NBD ΔUBA NAC was minimal, while the presence of ribosomes lead to a profound increase in the fluorescence. This effect resembles the effect for full length Egd1p-A16C NBD NAC (Figure 20). The activation of NBD at position Egd1 A16C must be therefore mediated by a different region of NAC. Another potential activator of NBD at position A16C are exposed hydrophobic side chains at the surface of the NAC domain which is composed of Egd1p and Egd2p. Due to the complex architecture of this pocket, the contribution of this region to the activation and thus a potential intramolecular interaction site could not be tested yet.

4.1.5 Assessing the contribution of different regions of NAC to the ribosome interaction

It was shown that the N-Terminus of Egd1p is sufficient to tether MBP to ribosomes (Pech et al., 2010). To assess if other parts of NAC also contribute to the ribosome binding, the MST-analysis was extended. First, a fusion consisting of the N-terminal 39 AA, a Linker $(GGGS)_4$ and MBP was used. MBP does not contain any cysteines. A single cystein was introduced either at position 16 of the N-terminus of Egd1p (analogous to the Egd1p-A16C NAC) or at the surface exposed amino acid 240 of MBP. The A16C derivative was labeled with NBD, while the K240C derivative was labeled with Alexa568. Both derivatives were measured using direct titration experiments with ribosomes. In case of Egd1p[1-39]-A16C NBD - MBP, 550 nM of the fusion protein was used due to NBDs low fluorescence in the MST setup (see also section 4.1.3.2). 80 nM of Egd1p[1-39] − MBP (K240C-Alexa568) derivative were sufficient to produce a strong signal.

The steady state level of the thermophoretic movement of the MBP fusion proteins with or without ribosomes was reached only after 90 seconds of the temperature gradient formation. Therefore, F_{hot} was calculated 95 s after temperature gradient formation. 50 time increments of 0.1 s were averaged to correct for the high noise of the signal. The dissociation constant of Egd1[1-39]-A16C NBD - MBP and Egd1[1-39] − MBP (K240C-Alexa568) were identical (230 nM) confirming the hypothesis that the A16C mutation labeled with NBD behaves like the unlabeled derivative (Figure 25). The affinities were however significantly lower than for Egd1p-A16C NBD NAC. This suggests an additional contribution of other parts of NAC to ribosomes binding. A result, that is supported by the fact that an additional crosslink was reported between Egd2p and the ribosomal protein L22 (Pech et al., 2010).
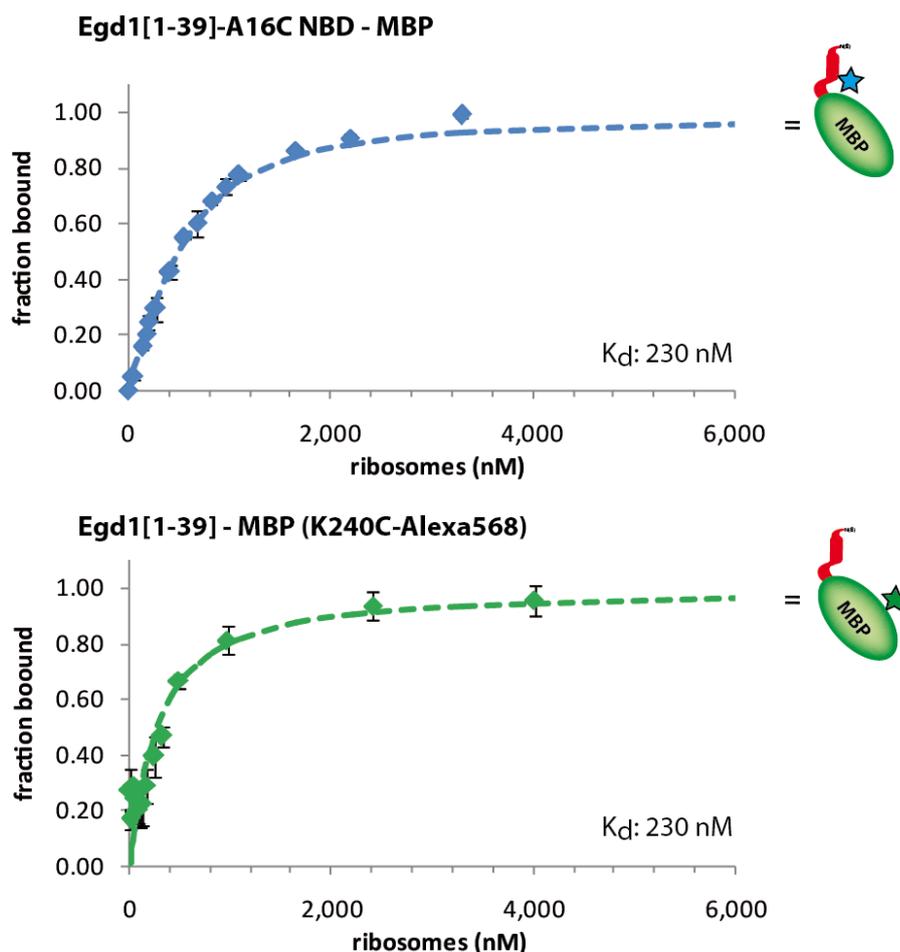
Figure 25: A fusion protein consisting of the N-Terminus of Egd1p and MBP directs MBP to the ribosome with an K_d of 230 nM | Ribosomes were titrated against 550 nM of the Egd1[1-39]- A16C NBD - MBP and 80 nM Egd1p[1-39]-MBP (K240C-Alexa568), respectively. All measurements were performed in MST Buffer. The Laser was set to 100%. F_{hot} was recorded for 50 time increments of 0.1 s each 95 seconds after the IR-Laser established the temperature gradient. F_{cold} was recorded for 50 time increments before the temperature gradient was established. The K_d of the NBD and the Alexa568 derivative (dashed lines) were identical. The measurements were performed in duplicates. The experiments were performed by Steffen Fliehmann.

To test a possible contribution of the UBA-domain of Egd2p, the UBA-domain was C-terminally fused protein to MBP (K240C). The protein was purified and labeled with Alexa568. Binding affinities of the fusion construct to ribosomes was measured using the MST setup. Upon ribosome titration, a weak reduction of the thermophoretic mobility of the MBP-UBA fusion protein was observed (dark green rectangles, Figure 26). This indicates a binding of the MBP-UBA fusion protein to empty ribosomes.

Technical constraints limited the analysis of the complex formation to 6 µM ribosomes. At this ribosome concentration the binding potential of the fusion protein is not saturated. Thus, the mobility of the MBP-UBA complex cannot be determined. However, a qualitative description of the data is possible – namely that the UBA domain mediates a weak interaction with the ribosome.

Figure 26 also shows the raw data for the MBP only negative control (light green triangles). Titration of ribosomes changes the thermophoretic mobility at extremely high concentrations of > 2 µM

slightly. At this concentration of ribosomes, the density of ribosomal particle reaches more than 8 mg/ml. These concentrations increase the viscosity and thus decrease the thermophoretic mobility. In direct comparison between the MBP - (K240C - Alexa568) negative control and the MBP (K240C)-Alexa568 UBA, the contribution of the ribosome binding of MBP (K240C) Alexa568 UBA to the changed migration behavior is nevertheless obvious.
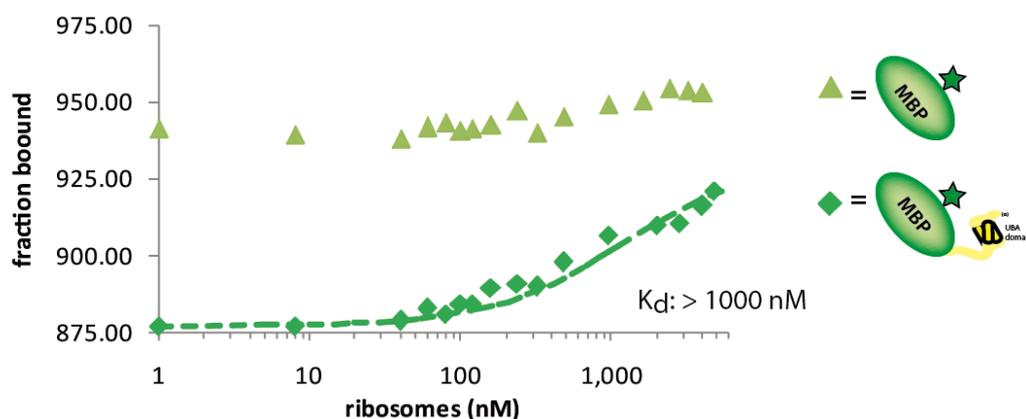
Figure 26: The UBA domain of NAC mediates a weak recruitment of MBP to ribosomes | Ribosomes were titrated against 80 nM of either the negative control MBP (K240C-Alexa 568) or the Fusion MBP (K240C-Alexa568) – UBA. All measurements were performed in MST Buffer. The Laser was set to 100%. F_{hot} was recorded for 50 time increments of 0.1 s each 95 seconds after the IR-Laser established the temperature gradient. F_{col}d was recorded for 50 time increments before the temperature gradient was established. A clear however weak signature of recruitment of the MBP-UBA fusion to the ribosome was observed. The experiments were performed by Steffen Fliehmann.

It is known that NAC interacts with DNA and RNA in an unspecific manner. It is however not known, which part of NAC mediates this interaction and what the significance of this observation for NAC's interaction with ribosomes is. To investigate this problem, the affinity of NAC to tRNAs as a model for structured RNA was measured. tRNA was titrated against 1.8 µM Egd1p-A16C NBD NAC in MST Buffer. A clear binding of NAC to tRNA with a K_d of 11 µM was observed (Figure 27).
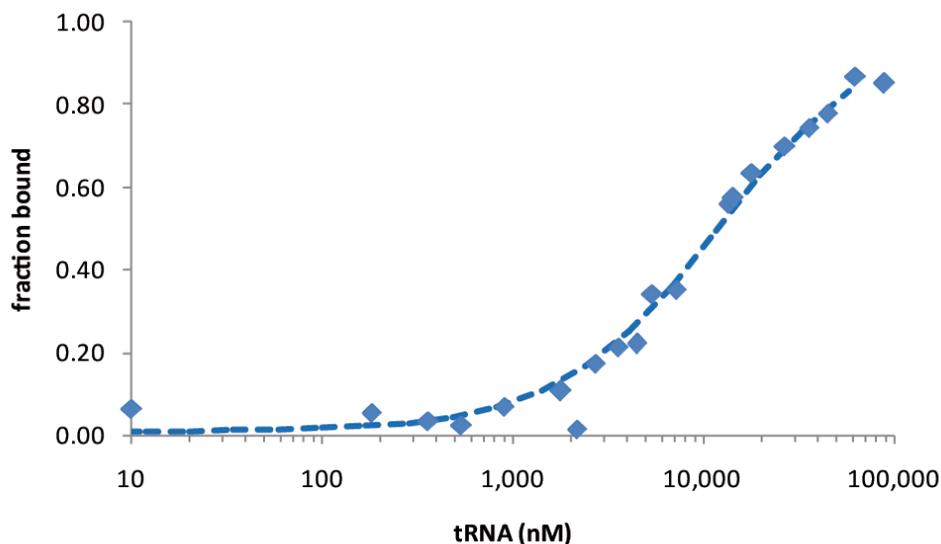
Figure 27: Egd1p-A16C NBD NAC binds with a K_d of 11 µM to tRNAs | Yeast tRNAs were titrated against 1.8 µM Egd1p-A16C NBD NAC. The affinity of Egd1p-A16C NAC NBD to tRNAs is 11 µM. 1.8 µM Egd1p A16C NBD NAC were kept constant and tRNA was titrated in MST-Buffer.

Since NBD's fluorescence is dependent on the microenvironment, it can be used as a reporter of the properties of this environment. We could detect in section 4.1.4 a quenching of the dye located at A16C of Egd1p. This indicated that the N-terminus of Egd1p is located in a hydrophilic environment when NAC is bound to the ribosome. A quenching of the dye at position Egd1p-A16C in NAC was also observed when NAC binds to tRNA (Figure 28). This hints towards an interaction between the N-terminus of Egd1p and rRNA. Interestingly the Egd1[1-39]- A16C NBD – MBP fusion also bound to tRNAs (section 9.3, p. 154).
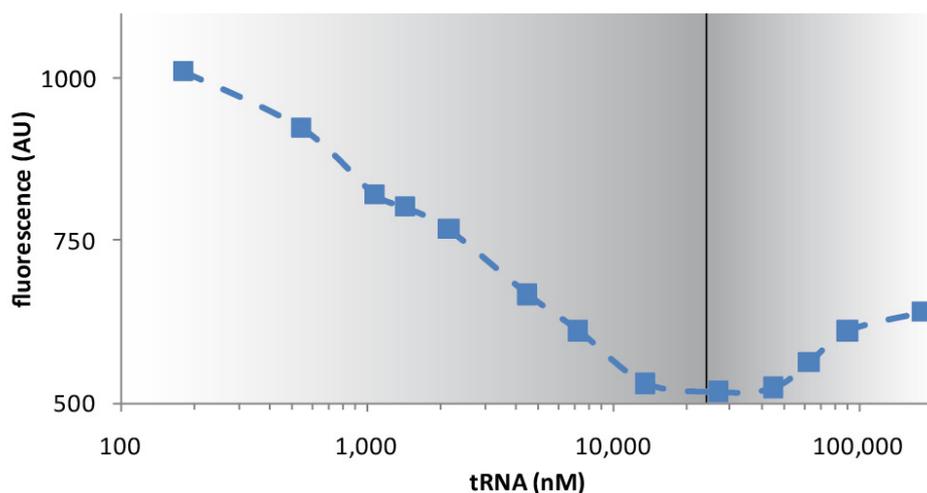
Figure 28: The fluorescence of Egd1p-A16C NBD NAC is quenched upon tRNA titration | tRNAs were titrated against 1.8 µM Egd1p-A16C NBD NAC in MST Buffer. Increasing tRNA concentration leads to a decrease in the fluorescence. At 30 µM, the autofluorescence of the tRNAs lead to increase in fluorescence.

Finally, we wanted to test whether the N-terminus of Egd1p when bound to ribosomes is an hydrophilic environment. To so, we had to use a NAC derivative that does not enhance the fluorescence of NBD located at position A16 in an unbound state. Otherwise, the loose of activation and the gain of repression is indistinguishable. Since we could not identify the region that enhances the fluorescence, we chose Egd1p(1-39)-A16C NBD MBP as a model since it does not show the fluorescence enhancement feature of Egd1p-A16C NBD NAC. The data of the previous quenching experiment were reanalyzed and the change in the fluorescence intensity upon binding to the ribosome was calculated (in contrast to the change in fluorescence intensity upon the titration of Tween-20, Figure 23 p. 73). A clear reduction of the fluorescence upon ribosome binding was observed indicating the transfer of the NBD-dye at position A16 into a hydrophilic environment (Figure 29).
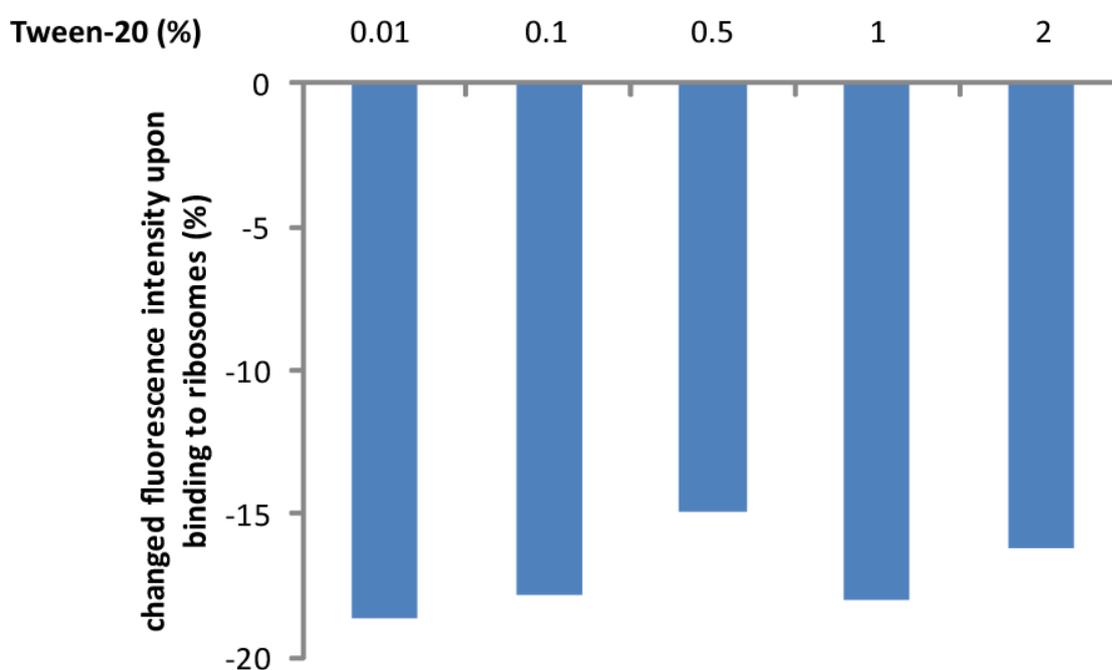
Figure 29 : NBD at position A16 at the N-terminus of Egd1p gets repressed upon binding to the ribosome | The data of **Figure 23** was reanalyzed to show the change in fluorescence intensity of the fusion protein Egd1p-A16C NBD MBP upon binding to the ribosome. Irrespective of the Tween-20 concentration in the buffer, a reduction in the fluorescence is visible indicating a transfer of the dye into a region of higher hydrophilic properties.

4.2 Competition between ribosomal interactors

After the MST and fluorescence approach gave new insights on the NAC-ribosome interaction, we extended our investigation to the dynamic interplay between different ribosomal interactors such as SR, RAC and NAC. Since there are finite binding sites at the ribosome – especially in vicinity of the tunnel exit – the binding mode of action evolved to optimize the interaction of these ligands with particular nascent polypeptides in the context of other ligands. Previous findings suggested that they both bind to the same universal adaptor site like NAC.

The MST competition approach is particularly suitable since it allows the determination of dissociation constants of unlabeled ligands that compete with NAC.

4.2.1 NAC-SR

First, we investigated the binding behavior of the SR to idle ribosomes. A truncated version of the SR α/β heterodimer lacking the hydrophobic membrane anchor was used. 4.8 nmol of protein were unspecifically labeled using the Monolith GREEN-NHS kit. To assess the labeling efficiency, the protein concentration was calculated based on the A_{280} (68,675 M^{-1} cm^{-1}) and the concentration of the dye was measured based on the A_{575} (150,000 M^{-1} cm^{-1}). 3.6 nmol SR were labeled with 4.32 nmol of the fluorophor yielding an average of 1.2 dyes per complex.

Initial experiments showed that the labeled SR derivative (SR-547) did not unspecifically bind to the capillaries in MST-Buffer and is thus compatible with the NAC-derivatives. Subsequently, a direct titration experiment in these buffer conditions was performed using 25 nM of SR-547. The thermophoretic migration behavior allowed the determination of F_{hot} and F_{cold} as described earlier (4.1.3.2). SR-547 bound to ribosomes with a K_d of 100 nM (Figure 30).
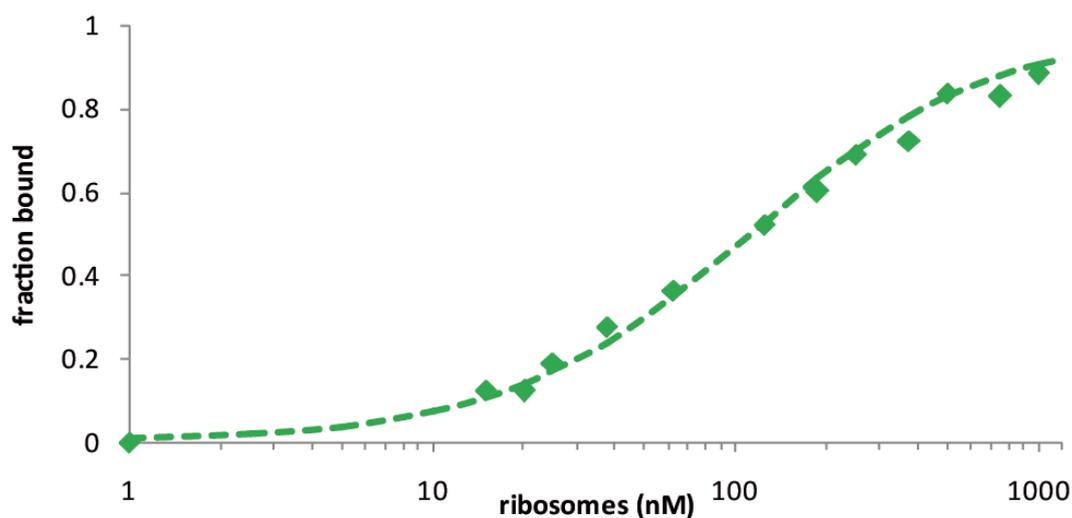
Figure 30: SR-547 binds to idle ribosomes with a K_d of 100 nM | Ribosomes in MST-Buffer were titrated against 25 nM of SR-547. The dashed line represents the fit of Formula 2 to the data with a K_d of 100 nM.

Since a binding of SR-547 to ribosomes was observed, competition experiments were performed next. Since the affinity of SR-547 to ribosomes is relatively low and the high fluorescence of the dye allowed measurements of low concentrations of SR-547, a setup with 50 nM SR-547 and 100 nM ribosomes was chosen. This resulted in a complex formation of 22 nM. Egd1p-A16C NBD NAC was titrated against the preformed complex.
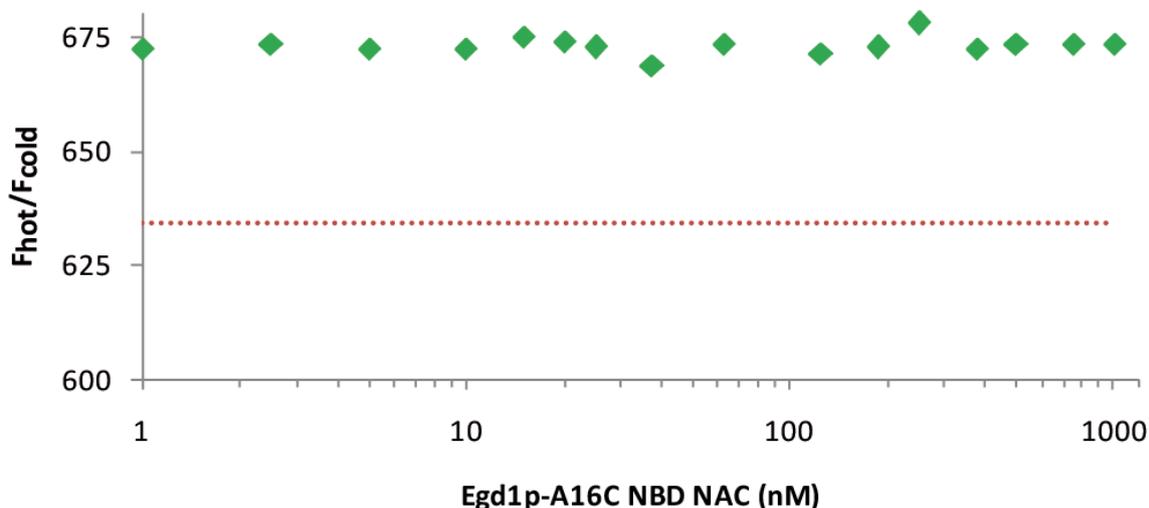
Figure 31: Egd1p-A16C NBD NAC does not compete with SR-547 for binding to ribosomes | The red line marks the relative mobility of SR-547 in the absence of ribosomes. The data indicate, that SR-547 is bound by the ribosome. Titration of Egd1p-A16C NBD NAC does not compete with SR-547 for binding to the ribosome. 50 nM SR-547 and 100 nM ribosomes were incubated in MST-Buffer. Egd1p-A16C NBD NAC was titrated against the preformed complex.

The low relative mobility indicated a complex formation of SR-547 and the ribosomes (green dots, Figure 31). For comparison, the relative mobility of SR-547 in the absence of ribosomes is shown by the red dotted line. Egd1p-A16C NBD NAC did not compete with SR-547's binding to the ribosome.

The simultaneous use of labeled NAC and labeled SR enabled us to investigate NACs binding to the ribosome in the same experiment due to the fact that different wave lengths were used for readout of the two dyes. In addition, only NBD showed a quenching of the fluorescence upon binding to ribosomes.

Since Egd1p-A16C NBD NAC's is titrated and thus the fluorescence in the blue channel varies significantly, a quantitative description of the fraction that is bound to the ribosome is not possible via MST. However, a qualitative description based on the fluorescence intensity is possible.

As it was shown in section 4.1.4 the fluorescence intensity of Egd1p-A16C NBD NAC is dependent on its binding status. Free Egd1p-A16C NBD NAC has a higher fluorescence per molecule NAC then ribosome bound Egd1p-A16C NBD NAC. We therefore analyzed the fluorescence intensity of Egd1p-A16C NBD NAC in the titration experiment and calculated the fractional fluorescence of Egd1p-A16C NBD NAC. The fractional fluorescence intensity was determined using Formula 7. It describes the average fluorescence intensity generated by 1 nM of Egd1p-A16C NBD NAC.

$$fractional\ fluorescence\ intensity = \frac{F_{x\ nM\ NAC} - F_{0\ nM\ NAC}}{C_{NAC}}$$

Formula 5: The fractional fluorescence describes the fluorescence intensity generated per nM Egd1p-A16C NBD NAC | The recorded fluorescence in the absence of Egd1p-A16C NBD NAC (F 0 nM NAC) is subtracted from the data points (x nM NAC) to correct for background fluorescence. This normalized fluorescence is then divided by the concentration of NAC (C_{NAC}).

If titrated Egd1p-A16C NBD NAC is directly bound to the ribosome, the fractional fluorescence intensity should be low, as the binding to the ribosome quenches the fluorescence intensity of the NBD dye. Higher Egd1p-A16C NBD NAC concentrations lead to a saturation of the binding site of the ribosomes and the additional, unbound NAC possesses an higher fractional fluorescence intensity. The average fractional fluorescence intensity should increase. If Egd1p-A16C NBD NAC does not bind to ribosomes in the presence of SR-547, no change in the fractional fluorescence should be detectable.

The fractional fluorescence intensity of Egd1p-A16C NBD NAC changed as a function of NAC concentration (Figure 32). A clear increase in the fractional fluorescence intensity can be observed indicating a scenario where first, Egd1p-A16C NBD NAC is binding to the ribosome. Thereby, the fluorescence of one Egd1p-A16C NBD NAC molecule is weak due to the quenching via the ribosome. Once the ribosomes are saturated, the high fractional fluorescence intensity of additional Egd1p-A16C NBD NAC that does not find unoccupied ribosomes increases the average fractional intensity.
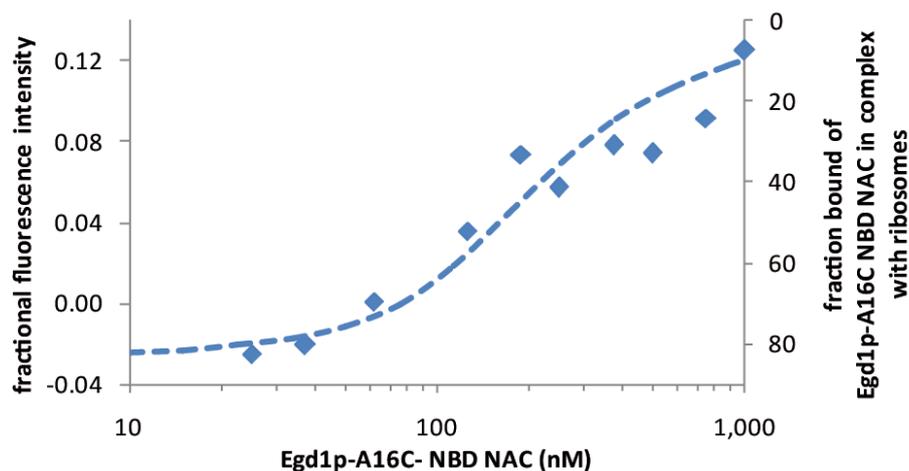
Figure 32: Fractional fluorescence intensity analysis: Egd1p-A16C NBD NAC is bound to ribosome-SR-547 complexes | Egd1p-A16C NBD NAC was titrated against a preformed complex of 50 nM SR-547 and 100 nM ribosomes in MST-Buffer. The fractional fluorescence intensity of Egd1p-A16C NBD NAC was calculated using Formula 5. A clear increase in the fractional fluorescence intensity indicates the binding of Egd1p-A16C NBD NAC to ribosomes. Low NAC concentration leads to a relative high fraction of bound NAC to ribosomes. This leads to a relative low fractional fluorescence intensity due to the quenching effect of the ribosome binding. Additional titration of NAC results in i) a complex formation between NAC and the ribosome until the ribosome is saturated and ii) additional NAC thus cannot bind to the ribosome. Therefore, a relative decrease of fraction bound of NAC to ribosomes is observed. Since the fluorescence of free NAC is higher than the fluorescence of bound NAC, the fractional intensity of the NAC Pool increases. The blue dashed line indicates a simulation of fraction bound of NAC using Formula 6.

In contrast to the previous direct titration experiments, here the complex status of the partner that is titrated is measured. This leads to a situation where at low NAC concentrations, a relative high percentage of NAC is in complex with the ribosome. Increasing the concentration of NAC leads to a saturation of the ribosome. After the saturation of the ribosomal binding site, additional NAC stays in a unbound state. The fraction of NAC that is bound to ribosome is therefore high under low NAC concentrations and low under NAC concentrations that significantly exceeds the ribosome.

This fraction can be calculated using Formula 6. The blue, dashed line in Figure 32 indicates the fraction bound of Egd1p-A16C NBD NAC (note the inverse display of the fraction bound).

$$fraction\ of\ bound\ molecules_{Egd1p-A16C\ NBD\ NAC} = \frac{fraction\ bound_{ribosome} * c_{ribosomes}}{c_{Egd1p-A16C\ NBD\ NAC}}$$

Formula 6: The fraction of bound molecules of the titrated partner can be calculated | This formula describes the fraction of bound molecules of the titrated competition partner. Based on Formula 4, the concentration of the NAC-ribosome complex can be calculated using the K_d as a fitting parameter and the concentration of the ribosome and NAC. This concentration divided the total concentration of the titrated partner gives the fraction of bound molecules of the titrated partner.

These data indicate, that SR-547 and Egd1p-A16C NBD NAC can bind simultaneously to idle ribosomes. To confirm this hypothesis, we performed a competition experiment where a complex between Egd1p-A16C NBD NAC and ribosomes was preformed. The unlabelled SR was titrated against this complex. The thermophoretic mobility of Egd1p-A16C NBD NAC was used as a read out

to deduce the complex formation of NAC. 200 nM of Egd1p-A16C NBD NAC and 200 nM ribosomes were incubated in MST-Buffer (complex concentration 146 nM). Unlabeled SR was titrated against this preformed complex.

The red dotted line in the plot of the competition experiment indicates the mobility of free Egd1p-A16C NBD NAC (Figure 33). The high F_{hot}/F_{cold} ratio and thus low mobility of Egd1p-A16C NBD NAC indicates a proper complex formation with the ribosome. Titration of up to 6 µM SR did not significantly compete with Egd1p-A16C NBD NAC's binding to ribosomes verifying the absence of a competition between Egd1p-A16C NBD NAC and SR.
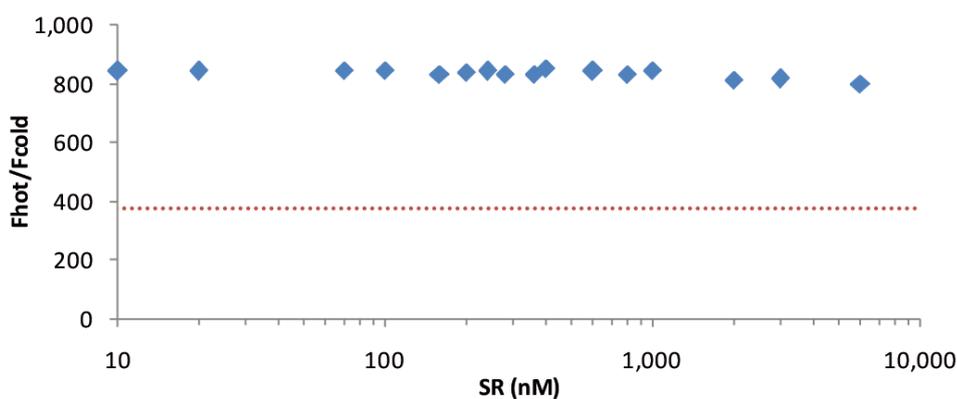
Figure 33: Egd1p-A16C NBD NAC does not compete with SR | 200 nM Egd1p-A16C NBD NAC and 200 nM ribosomes were incubated for 5 min in MST under RT. Unlabeled SR was titrated against this complex. The red dotted line indicates the mobility of free Egd1p-A16C NBD NAC. Since the mobility of Egd1p-A16C NBD NAC does not change as a function of SR and Egd1p-A16C NBD NAC is bound to ribosome, SR and NAC do not compete for binding to the ribosome.

4.2.2 NAC-RAC

The data of section 4.2.1 showed that SR and NAC can bind simultaneously to the ribosome. Next, the binding of the ribosome associated complex RAC was investigated. RAC was also cross-linked to L31e and thus makes it another possible candidate for competition with NAC for binding to the ribosome (Peisker et al., 2008).

So far, the affinity of RAC to ribosomes is not known. Direct MST titration experiments using labeled RAC and titration of ribosomes did not yield dissociation constants (personal communication, C. Leidig, Gene Center, LMU Munich). Therefore, competition experiments using unlabelled RAC and Egd1p-A16C NBD NAC were performed. 200 nM Egd1p-A16C NBD NAC and 200 nM ribosomes were preincubated and competition with increasing concentrations of RAC was measured. At concentration > 1 µM, RAC, NAC and ribosomes precipitated. Therefore, only data points up to 1 µM RAC were considered for the analysis. A plateau phase were all NAC was displaced could therefore not be reached. However, the data of the direct titration of Egd1p-A16C NBD NAC against ribosomes

allowed for an assignment of the fraction bound of Egd1p-A16C NBD NAC based on the relative thermophoretic mobility of NAC.

A clear competition for ribosomes between NAC and RAC was observed (Figure 34). RAC binds with an K_d of 20 nM ribosomes which is the same range as the K_d of NAC to ribosomes. Furthermore, it competes with NAC for the binding site.
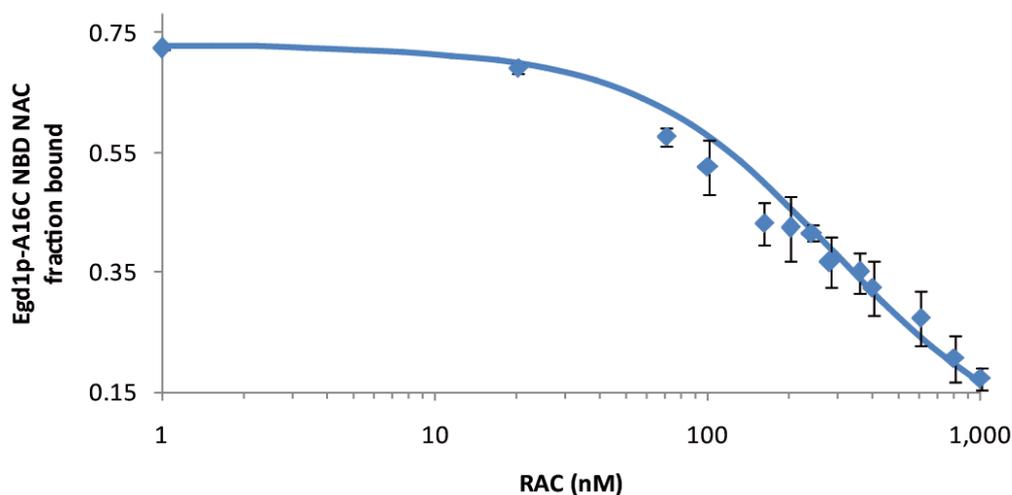
Figure 34: RAC competes with Egd1p-A16C NBD NAC for ribosome binding | 200 nM ribosomes and 200 nM Egd1p-A16C NBD NAC were preincubated in MST-Buffer. Unlabeled RAC was titrated against this preformed complex. The affinity of RAC for the ribosome was calculated to be 20 nM.

In addition to the quantitative analysis of NACs binding behavior to ribosomes, a structural analysis of TF binding to translating ribosomes using cryo-EM and single particle analysis was performed during this thesis.

4.3 Trigger Factor

Trigger Factor (TF) forms together with the *E. coli* Hsp70 system (DnaK/J, GrpE) a redundant and crucial first chaperone system that helps newly synthesized nascent polypeptide to either reach their native fold or prevent the hydrophobic collapse and thus allows transfer to downstream acting systems one systems such as GroEL/ES.

TF is the ribosome associated part of this chaperone network and thus principally accessible for cryo-EM and single particle analysis. Although a cryo-EM structure and a variety of X-ray structures and fragments are available, none of the projects so far in shows TF interacting with an endogenous substrate on the ribosome at a resolution that allowed unambiguous interpretation of TFs mode of action.

TFs unusual structure, molecular size and high flexibility required the implementation of new algorithms to visualize and describe this chaperone in complex with the ribosome.

4.3.1 Reconstitution of a physiological TF – ribosome complex

4.3.1.1 Preparation of ribosome nascent polypeptide complexes

To investigate the interaction between TF and a physiological substrate, purification of homogenous populations of programmed ribosomes carrying a defined nascent polypeptide (ribosome nascent polypeptide complex, RNC) was the first necessity.

The *E. coli* model substrate galactitol-1-phosphate dehydrogenase (GatD) was used because its interaction with TF was analyzed previously (Kaiser et al. 2006). Since a correlation between the number of hydrophobic amino acids in the primary sequence and the interaction time with TF was observed, we speculated that the presence of a hydrophobic stretch directly at the tunnel exit stabilizes TF at the ribosome and its interaction with the nascent polypeptide.

GatD contains two hydrophobic stretches (AA 81-91 and AA 165-185; Figure 35). The first stretch was used for the design of the nascent polypeptide as it does not code for any α-helix. It is therefore likely that the nascent polypeptide resides in an unfolded state outside the tunnel. This is important since the TF-RNC complex was *in vitro* reconstituted from individually purified RNCs and TF substrates are unfolded nascent polypeptides.
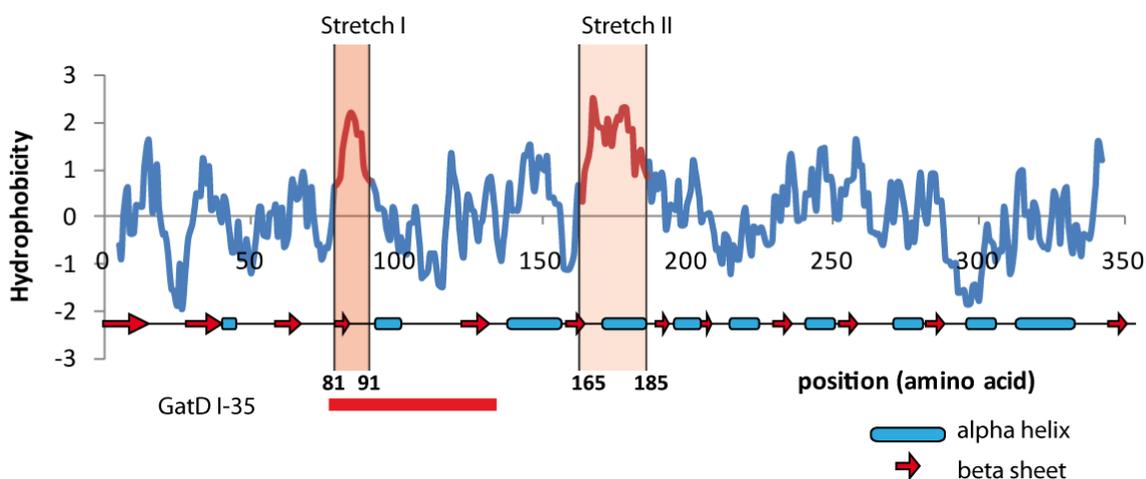
Figure 35: GatD contains two hydrophobic stretches | The analysis of the hydrophobicity of GatD (Kyte and Doolittle algorithmus, windows size 9 AA) revealed the presence of two hydrophobic stretches (positive values: high hydrophobicity). Since no formation of α-helical secondary structures was predicted for the first hydrophobic stretch (AA 81-91), this part was used for the further analysis. The red line indicates the sequence of GatD that was used for the generation of the GatD I-35 nascent polypeptide.

The length of the ribosomal tunnel (100 Å) allows for the accommodation of up to 35 AA of an extended nascent polypeptide. Therefore, 35 amino acids of the GatD sequence C-terminally of T91 were used to place the AA 81-91 directly at the tunnel exit. This also placed a predicted α-helix composed AA 93 to 102 in a region of the tunnel that allows for the formation of α-helices of the nascent polypeptide in the tunnel. For affinity purification and detection, a His$_6$ tag followed by an HA tag was engineered N-terminally of V81 (GatD I-35, Figure 36).
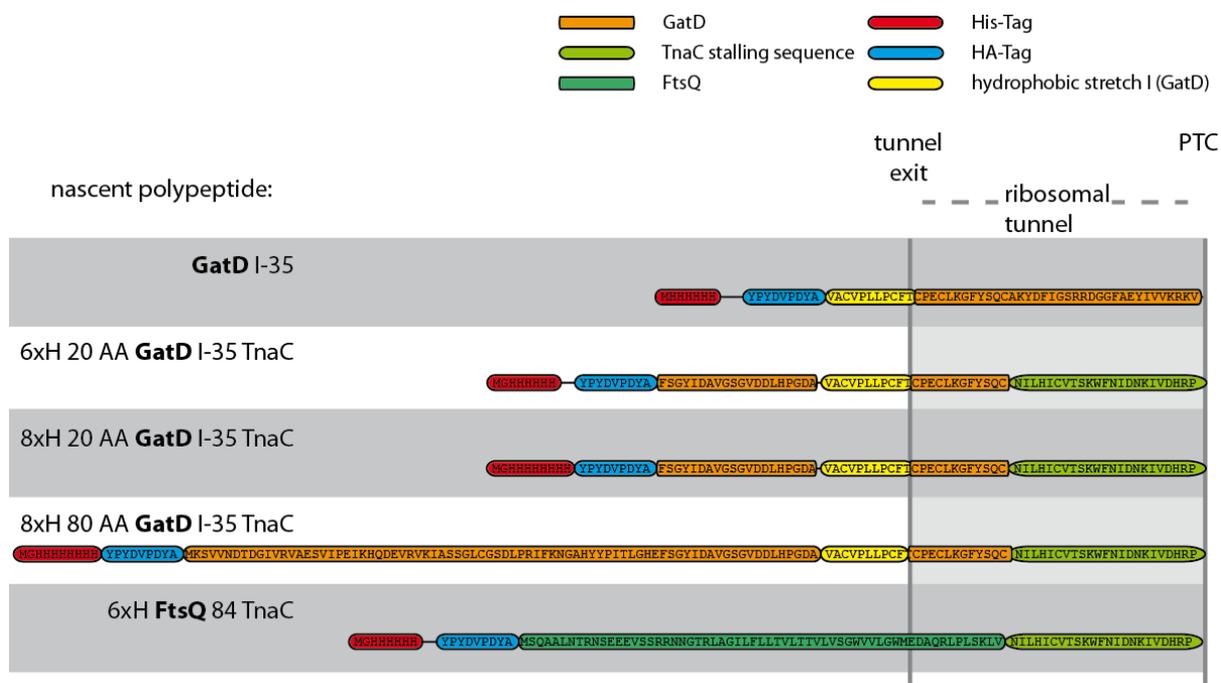
Figure 36: Overview of the nascent polypeptides used in this study | The templates differ in numbers of histidins for purification, the number of AA of the protein used as a substrate (FtsQ, GatD), the stalling mechanism (GatD I-35 TnaC truncated, the others TnaC stalled) and the length of the nascent polypeptide.

In a first approach, prevention of termination of translation und subsequent release of the nascent polypeptide was achieved by truncation of the mRNA. Thereby no stop codon is present at the C-terminus. Without a stop codon, the release factors cannot hydrolyze the peptidyl-tRNA.

This mRNA was used as a template for *in vitro* test translation. The translation product was probed by Western Blot either against the HA-tag or the His-tag. No translation product was detectable (data not shown). However, radioactive labeling of the nascent polypeptide showed translation. Therefore, GatD I-35 RNCs were purified *in vitro*. The yields of the *in vitro* purification and their programming rates as judged by cryo-EM and single particle reconstruction were too low for further structural analysis of the TF-RNC complex (data not shown).

Therefore, the stalling technique was changed. In 2009, an *in vitro* TnaC stalling technique was successfully applied to obtain high amounts of highly programmed RNCs (Seidelt et al., 2009). To establish the techniques, an FtsQ based, TnaC stalled nascent polypeptide was used (FtsQ 84 TnaC, Figure 36). The translation of a truncated FtsQ nascent polypeptide was previously successfully detected by western blot and is therefore an ideal positive control for the establishment of a new protocol. FtsQ however is not a TF substrate due to its signal anchor.

For *in vitro* translation, an *E. coli* KC6 based translation extract was prepared. Transcription coupled translation in 10 ml extract was initiated by addition of 150 μg of the template encoding vector (pFtsQ84TnaC). Translation was stopped after 30 mins by addition of chloramphenicol.

To stabilize the TnaC based stalling, the ribosomal fraction was then in presence of 1 mM tryptophan and 0.1 mg/ml CAM pelleted and the RNCs purified. The monosomal fraction of the eluted RNCs was separated from the disomal fraction using a 10-40 % saccharose gradient (blue line, Figure 37). To verify the identity of the monosomal peak, an *E. coli* polysomal fraction was fractionated using the same conditions (red line, Figure 39).

The ribosomes of the monosomal fraction of Figure 37 were then pelleted through a sucrose cushion for concentration and removal of released peptides. A western blot analysis of the pellet that contained the ribosomes and the supernatant confirmed the presence of the peptidyl tRNA and thus programming of the ribosomes with the FtsQ 84 TnaC nascent polypeptide (Figure 38). We used a mild centrifugation force to pellet the ribosomes since *E. coli* ribosomes are prone to damage. Under this conditions, ribosomal subunits resulting from damaged RNCs after the gradient do not pellet but stay in the supernatant. That these are not just ribosomes that do not pellet is visible by the complete absence of a signal for the peptidy-tRNA. The total yield of the in vitro RNC prep was 296 pmol (14.8 A_{260}) from 10 ml in vitro translation extract.
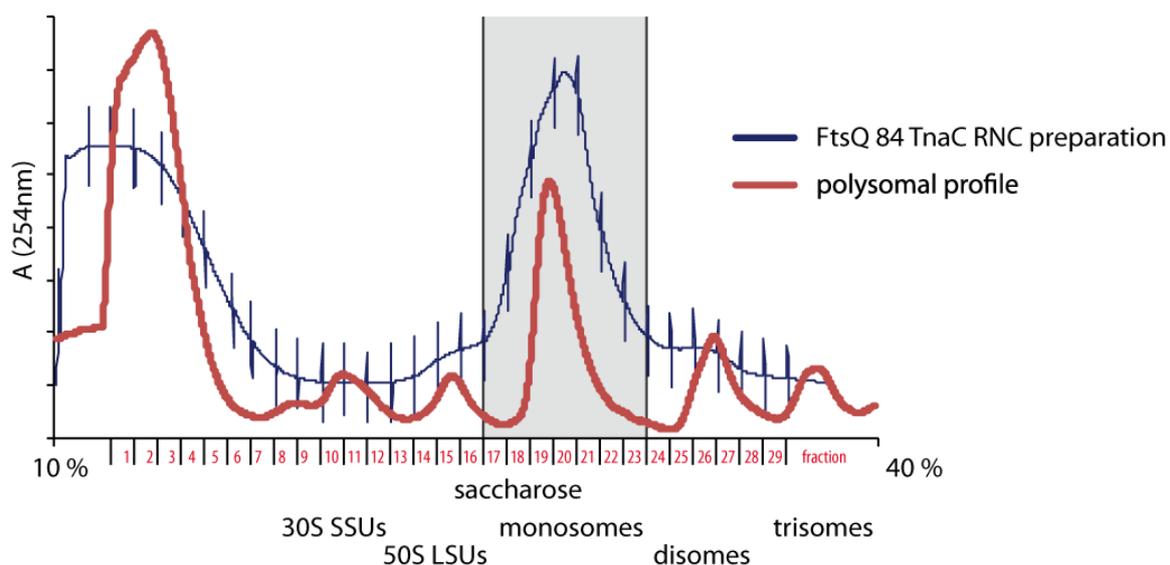
Figure 37: FtsQ 84 TnaC RNCs eluted from the talon column contained mainly monosomes | The elution of the Talon® affinity chromatography was loaded onto a 10-40 % saccharose gradient (SW32, 4h, 30,000 rpm, blue lines). The fractions 17 – 23 were pooled and used for further analysis. As a standard, an *E. coli* polysomal preparation containing 30S, 50S, 70S and polysomes was also fractionated on a identical 10-40% saccharose gradient. This profile is plotted using a red line.
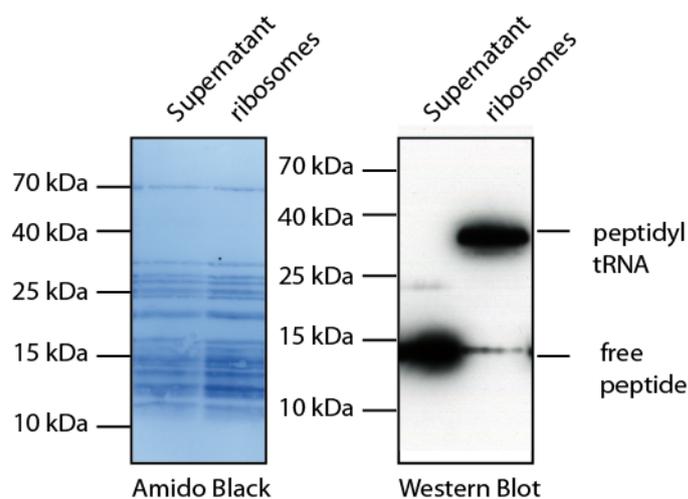
Figure 38: FtsQ 84 TnaC RNCs show a strong signal for the peptidyl tRNA | The monosomal fractions of Figure 37 (fraction 17-23) were pooled and the ribosomes pelleted. 2% of the supernatant fraction and 2% of the pellet fraction (ribosomes) were loaded on a SDS-PAGE and blotted on a PVDF membrane (45 min, 75 mA). The membrane was stained with amido black, blocked with 5 % milk in TBS and decorated with anti HA- antibody (1:1,000, mouse). The secondary antibody (1:500, goat anti rabbit, labeled with HRP) was detected for 15 seconds.

Next, we tried to purify GatD RNCs using this protocol. To introduce the TnaC stalling in the GatD I-35 substrate, the C-terminal 23 AA of GatD I-35 (starting with A104) were replaced by the TnaC stalling sequence. During the course of this work, a study was published that suggested that TF contacts more than 40 AA of the nascent polypeptide (Merz et al., 2008). However, the GatD I-35 RNCs are predicted to displays only 27 AA at the surface of the ribosome. Therefore, 20 AAs of the GatD sequence (F61-A80) were introduced N-terminally of the hydrophobic stretch (20AA GatD I-35

TnaC, Figure 36) to extend the available polypeptide chain. However, all *in vitro* approaches and optimizations to purify 20AA GatD I-35 TnaC failed. Translation was detectable, but binding to the Talon® material was too weak.

Therefore, *in vivo* expression of His$_6$ 20AA GatD I-35 TnaC using a T7-promotor based expression plasmid (pBAT4) for the in *E. coli* ER2566 was tested (Figure 39). After a lag phase of 6 h, growth of the expression strain accelerated. Upon induction with 1 mM IPTG, no HA-tagged translation product was detectable. Purification of His$_6$ 20AA GatD I-35 TnaC RNCs failed. Later, a TnaC based *in vivo* expression protocol by L. Bischoff and E. van der Sluis (both AG Beckmann, Genzentrum der LMU) that used an arabinose inducible promoter, His$_8$ Tag and the *E. coli* KC6 strain as expression strain was developed. We adapted this protocol to the GatD substrate. This resulted in the His$_8$ 20 AA GatD I-35 TnaC substrate (Plasmid from E. van der Sluis). Uninduced growth was normal and the induction led to a strong induction of the transgene (Figure 40).
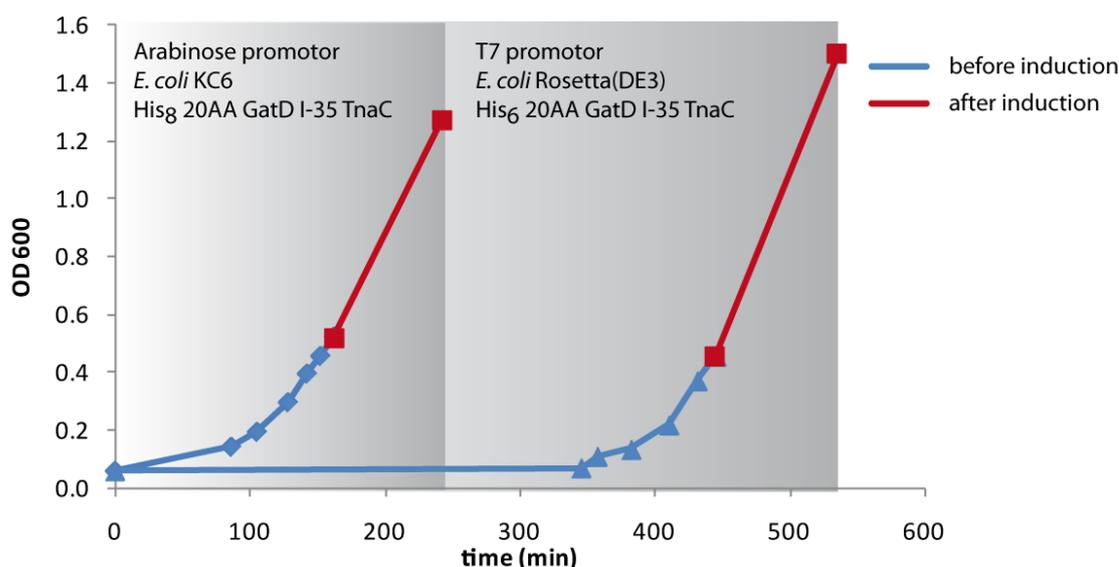
Figure 39: The use of T7 promotor controlled His$_6$ 20AA GatD I-35 TnaC template plasmids in *E. coli* Rosetta (DE3) lead to an extended lag phase, whereas arabinose promoter controlled His$_8$ 20AA GatD I-35 TnaC template plasmids in *E. coli* KC6 lead to a normal growth | Freshly transformed *E. coli* KC6 or Rosetta(DE3) colonies were diluted to an OD$_{600}$ of 0.06 in LB-medium plus the appropriate antibiotics. Expression was induced at an OD$_{600}$ of 0.5 with either 1 mM IPTG (Rosetta(DE3) or 0.2 % Arabinose (KC6). While *E. coli* KC6 plus pBAD-His$_8$ 20AA GatD I-35 TnaC showed a normal exponential growth, *E. coli* Rosetta(DE3) plus pBAT4-His$_6$ 20 AA GatD I-35 TnaC showed a lag phase of 6h before logarithmic growth started.

This setup allowed for the expression and purification of His$_8$ GatD I-35 TnaC derivatives (Figure 36, 20 AA GatD I-35 TnaC and 80 AA GatD I-35 TnaC). Figure 40 shows the western blot analysis of the total cell lysate before purification of the RNCs and the final pelleting step of the RNC purification. The stability of the *in vivo* RNCs is reflected in the absence of any released peptide in the pelleting step.
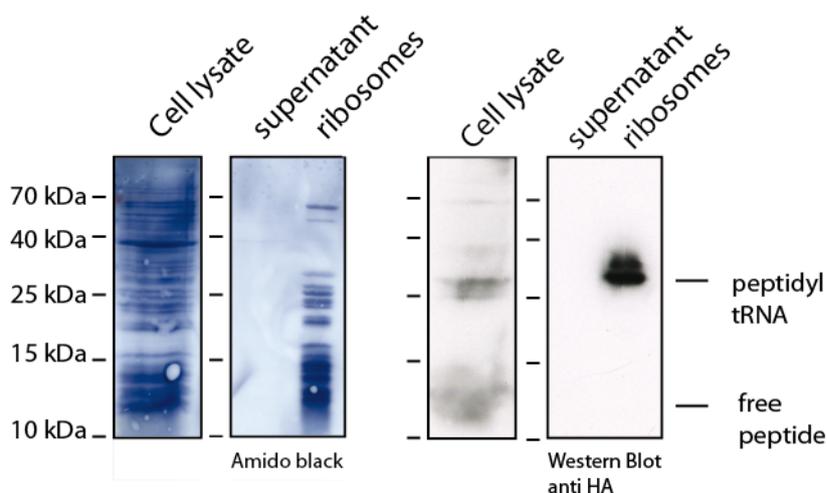
Figure 40: GatD derivatives containing his₈ tag efficiently stabilize *in vivo* RNC and allow for purification of GatD RNCs | pBAD – His$_8$ 57 AA GatD I-35 TnaC was transformed in *E. coli* KC6. 2h after induction, the harvested cells were lysed using a french press (cell lysate). The cleared cell lysate was then incubated with Talon® beads. After washing of the beads, the elution was loaded onto a 10-40% saccharose gradient. The monosomal fraction was collected and pelleted to concentrate the ribosomes. The cell lysat, supernatant and the pelleted ribosomes were loaded on a SDS-PAGE and blotted on a PVDF membrane (45 min, 75 mA). The membrane was stained with amido black, blocked with 5 % milk in TBS and decorated with anti HA- antibodys (1:1,000, rabbit). The secondary antibody (1,500, goat anti rabbit, labeled with HRP) was detected for 10 seconds.

In comparison to published SecM based *in vivo* purification protocols, the yield of the TnaC *in vivo* purification was very high (Evans et al., 2005). 1 L starting culture with an OD$_{600}$ of 1.2 yielded 120 pmol (His$_8$ 20 AA GatD I-35 TnaC) and 80 pmol (His$_8$ 80 AA GatD I-35 TnaC) RNCs.

4.3.1.2 Purification of TF

TF was expressed in *E. coli* ER2566 from a pPROEX-HTa based expression plasmid (AG Hartl, MPI Biochemie, Martinsried) coding for a TEV cleavable His$_6$ tagged TF. The cells were lysed and TF purified using NiNTA affinity chromatography. After dialyses and proteolytic cleavage of the His$_6$ tag, a cation exchange chromatography (Poros HQ20) followed by a gel permeation chromatography (S200 26/60) yielded pure, untagged TF (Figure 41).
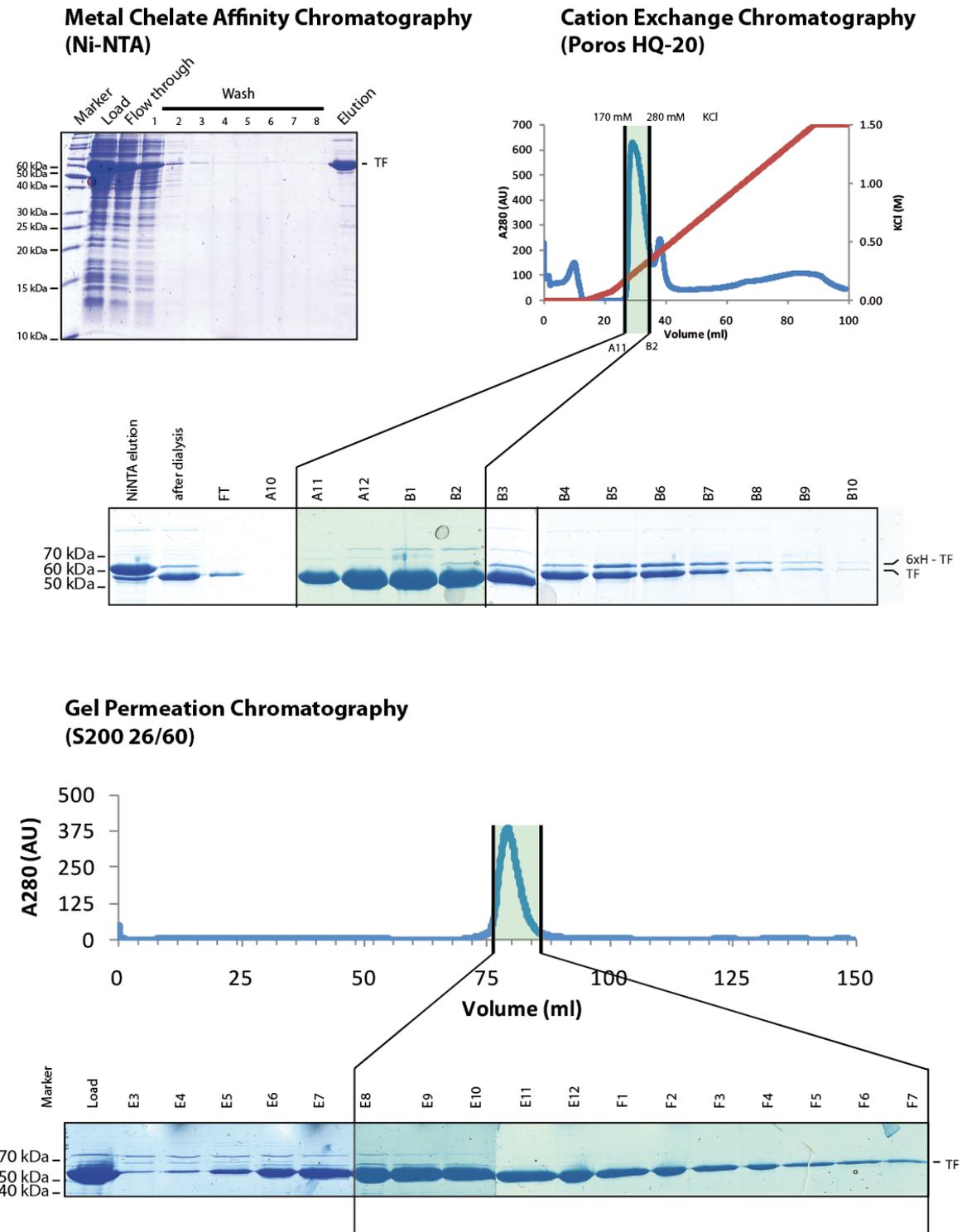
Figure 41: Purification of TF | TF was expressed in *E. coli* ER2566 and purified using a NiNTA column. The elution was then dialysed in the presence of TEV-protease to cleave the 6xH tag. The dialysed, cleaved fraction was then subjected to a cation ion exchange purification at pH 7.4 (Poros HQ20). Fractions A11-B2 were pooled, concentrated and used for gel permeation (S200 26/60). A monodispers protein fraction (E8-F7) was obtained and used for further analysis.

4.3.1.3 In vitro reconstitution of a TF-RNC complex

The complex formation between TF and ribosomes in context of nascent polypeptides was investigated by biophysical and biochemical assays previously. TF and ribosomes robustly form a complex under a variety of buffer conditions. To verify that the purified RNCs and TF indeed form a complex, co sedimentation assays were performed. 10 pmol of ribosomes and 100 pmol TF were incubated in MST-Buffer. After an incubation time of 10 min at RT, the ribosomes were pellet through a 20 % sucrose cushion. The samples were then TCA precipitated and half of the volume analyzed by SDS-PAGE (Figure 42).

Ribosomes alone did not show any protein band corresponding to TF (around 60 kDa). TF alone was exclusively found in the supernatant. A preformed complex of TF and ribosomes pelleted through the cushion and based on the band intensities a stoichiometric comigration of TF with ribosomes can be assumed.

The double band in TF was caused by the TCA precipitation as the protein preparation of TF did not show this band (Figure 41).
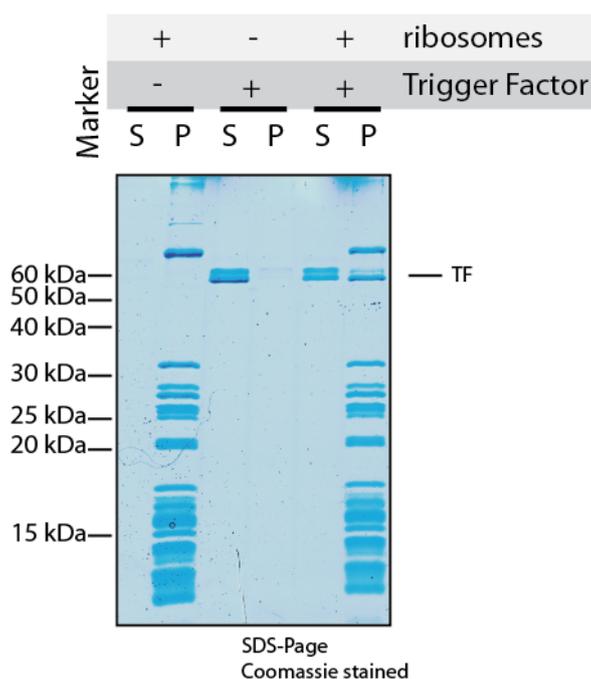
Figure 42: TF is found in the pellet after pelleting through a sucrose cushion only if ribosomes are also present | TF and ribosomes were preincubated at RT and the spun through a 20% sucrose cushion. Only in presence of ribosomes, TF was found in the pellet fraction indicating a complex formation. (S: supernatant, P: pellet)

4.3.2 Single particle cryo-EM reconstruction of the TF – RNC complex

80 nM His$_8$ 20AA GatD I-35 TnaC RNCs were incubated with 800 nM TF. This preformed complex was then applied on Quantifold cryo-EM grids and vitrified using the vitrobot (FEI).

A dataset of 278 micrographs was collected under low electron dose conditions and scanned. 20.5% of the micrographs were discarded after visual inspection due to low quality data and the remaining 221 micrographs used for further processing. Ribosomal particles were isolated and subjected to quality control using the MAPPOS algorithm (http://arxiv.org/ftp/arxiv/papers/1112/1112.3173.pdf). 478,378 particles were classified as ribosomes and used for initial alignment.

The particles were aligned against an idle *E. coli* ribosome low-pass filtered to 25 Å. The first backprojection already showed a clear density for the presence of the P-site tRNA. After twelve initial refinement rounds the pixelsize was decreased from 3.712 Ångström/pixel to 2.475 Ångström/pixel. After that, a back-projection in real space (bprp) yielded a starting reference for subsequent sorting.

Previous cryo-EM and single particle reconstruction projects used a hierarchical sorting procedure based on competitive two-reference alignments. Using different references and masks, sorting was done according to decreasing significance of structural features. Which reference and masks were used strongly defined the outcome. The presence of more than two, equally significant features often obstructs this approach. Therefore, multi-reference alignments with more than two distinct but almost identical references were used throughout the sorting process. This resulted in various improvements. The choice of the reference did not determine the outcome as the references developed genuinely bias-free into the final volume. Moreover, no prior knowledge of existing subpopulation in the dataset was necessary. By increasing the number of unbiased references until additional adding of references did not yield in qualitatively different maps, the existing subpopulations were exhaustively detected.

We could not sort for the presence and the absence of the P-site tRNA. All ribosomal particles were assigned to programmed ribosomes. Instead, the most significant difference was the presence or absence of the E-site tRNA accompanied by distinct L1-stalk confirmations. The multi-reference strategy worked for major differences like ribosomes versus non ribosomal particles and at the same time generated two distinct ribosomal subdatasets (presence and absence of E-site tRNA; Figure 43).
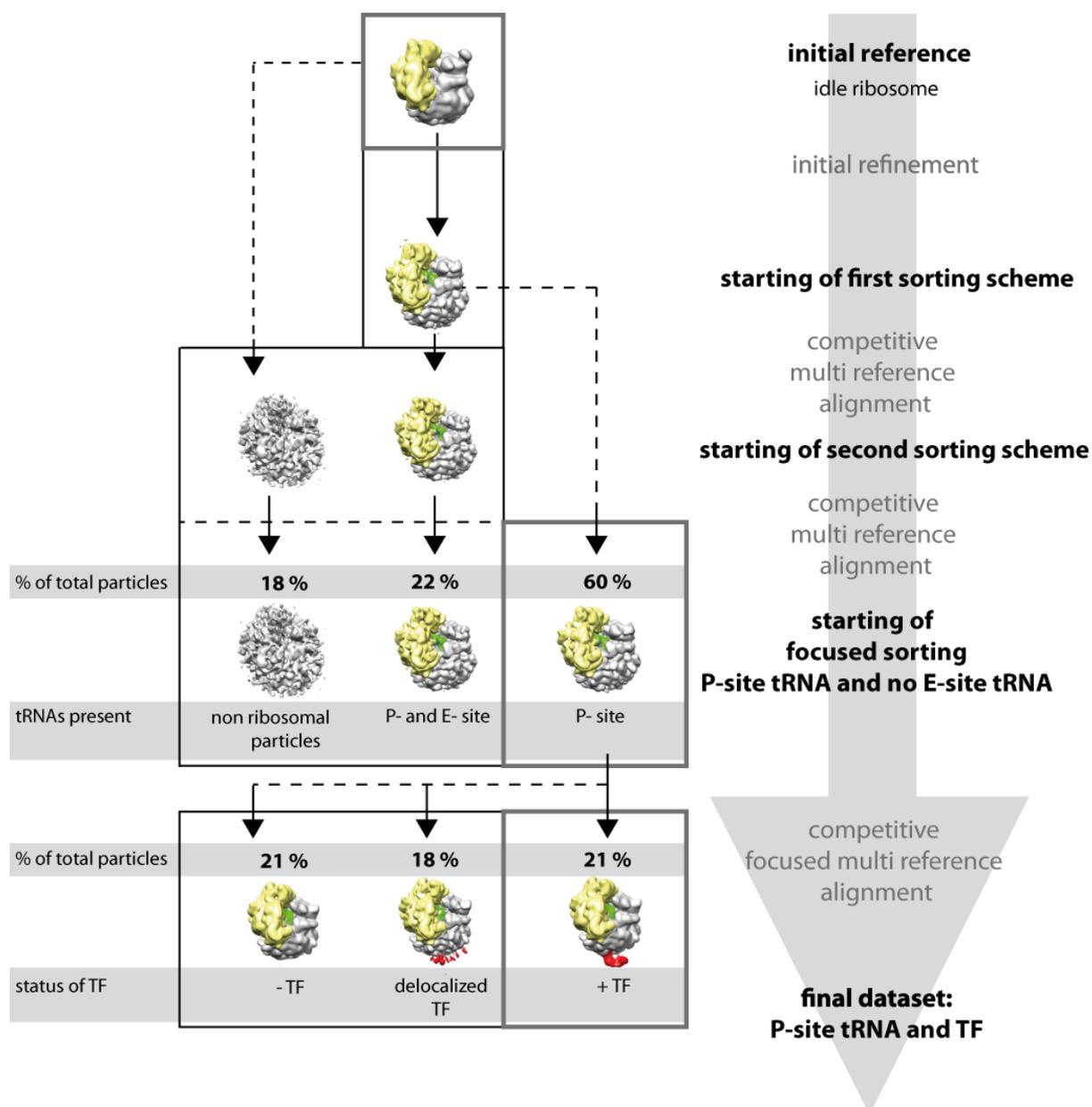
Figure 43: Overview of the processing workflow for the TF-His$_8$ 20 AA GatD I-35 TnaC dataset | The inital dataset of 478,378 particles was aligned against an idle *E. coli* ribosomes. After twelve rounds of refinement, competitive multi reference alignment using the refined volume and the initial reference yielded three stable and distinct datasets. Two ribosomal volumes that both showed a defined density for the P-site tRNA differed in the occupancy of tRNAs in the E-site and in the L1-stalk position. Particles that contained only P-site tRNA were used for further analysis (60% of the total dataset). Using a competitive multi reference alignment that employs only a subset of information defined by a binary 3D mask allowed for definition of three subclasses of P-site tRNA containing ribosomal particles (competitive focused multi reference alignment); 1.) RNCs with defined TF, 2.) RNCs with delocalized TF and 3.) RNCs without TF. Group 1.) was used for further analysis. (Yellow depicts the small, grey the large ribosomal subunit. tRNAs are colored green and density of TF red; dashed arrows represent the origin of references while solid lines represent succession of densities).

The presence or absence of the E- site tRNA created a more significant difference than the proteinaceous TF. Therefore the further structure calculation was based on the homogenous subdataset that contained only P- site tRNAs as this was the largest dataset (60 % of all particles, 288,484 particles in total). TF was only at low contour levels at the tunnel exit visible, indicating

either a high flexibility or substoichiometric binding of TF. To obtain a homogenous TF population, global cross correlation coefficient based approaches for competitive multi reference alignment were used. In this scheme, the cross correlation coefficient (CC C) between the particle and each of the references is calculated and the particles are assigned according to the better CC C.

This approach failed as the signal created by the proteinaceous 60 kDa TF was small in comparison by the signal of the large ribosome. To increase the contribution of the TF signal to the CC C, a focused sorting scheme was implemented (focused competitive multi reference alignment; FCMR alignment). In this scheme, the particles are aligned to all references and individual angles are assigned to each of the references. A binary 3D mask was then used, that covered the region of interest. To calculate the CC C, for each particle each reference is projected into 2D using the new angles. Additionally, the 3D mask is projected for each particle according to the new angles into 2D. The CC C is then calculated for each reference only inside the specific 2D mask. This local CC C then replaces the global CC C that is determined during the alignment process (Figure 44).

To reduce the likelihood of artifacts due to wrong references, the same volume (plus p-site, no E-site) filtered under different conditions was used as a reference. Using the FCMR alignment, the P-site tRNA containing subdataset could be further sorted into two populations of particles. 21 % of the particles did not contain density for TF, while 39 % did (of all particles). The TF containing particles were further subdivided into a set of particles where TF was in a defined position (21 %) and a second set, where (i) the ribosomal binding domain was shifted relative to the first set and (ii) the TF density was delocalized around the tunnel exit.

The FCMR alignment might lead to accumulation of noise if a small 3D mask is used. Therefore the defined TF dataset was investigated in more detail. An extra density near a region of the tunnel exit that is composed of L23 and L29 was visible. Since it was previously shown that TF interacts with L23 and L29 and the region where the ribosomal binding domain of TF interacts with the ribosome is known, the extra density clearly originated from TF.
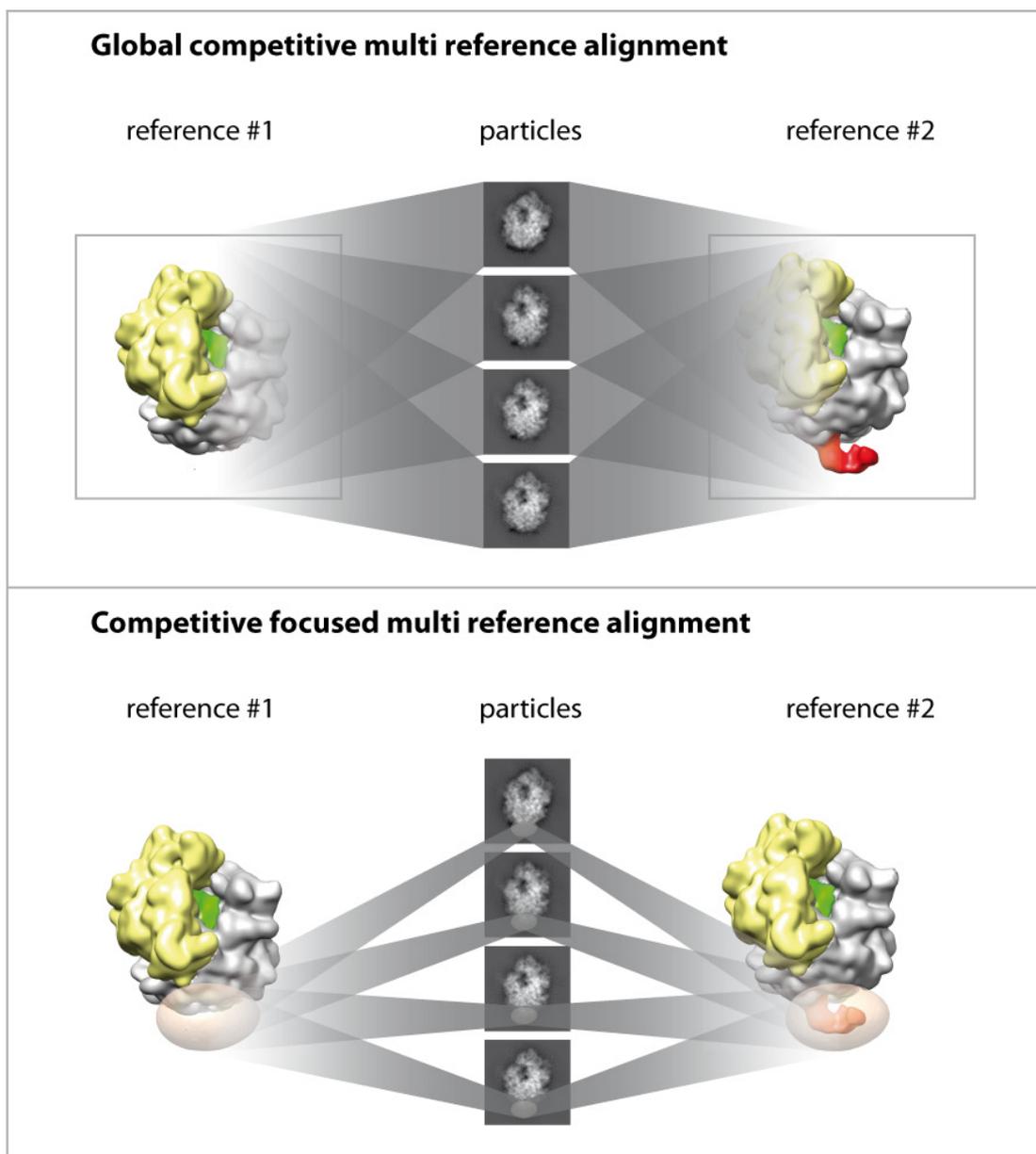
Figure 44: Focused competitive multi reference alignment (FCMR) | To assign the particles based on TF-density, not the entire information of the particle and the reference is used, but the information within a 2D projected binary 3D mask. The new CC C replaces the original, from the global alignment derived CC C and thus contains more of the ligand derived information.

The dataset was further refined to a final resolution of 7.7 Å according the FSC 0.5 criteria (Figure 45 and Figure 46). Ribosomal details were resolved to a degree that was expected for a resolution of 7.7 Å: α-helices were clearly resolved and the rRNA grooves visible. The TF density of the other hand was reduced to a small fraction of the RBD (red density in Figure 45). While the majority of the density was blurred out, the RBD showed resolution of secondary structure elements. The density distribution of the TF density indicated a near stoichiometrical occupation of the RNCs with TF. This hints towards a highly flexibility of the middle- and the head- domain.
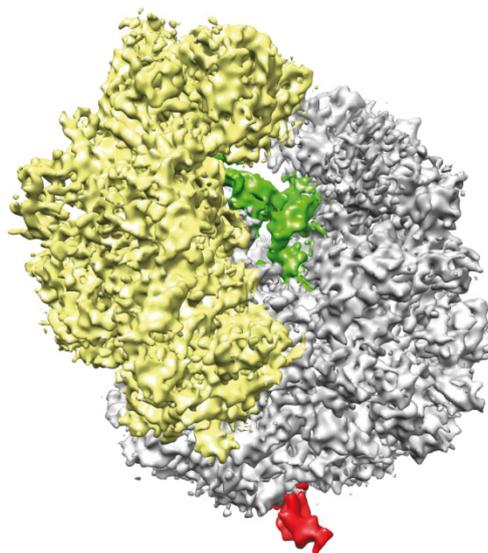
Figure 45: Overview of the final 7.7 Å density map (volume) of the TF-RNC complex | 100,931 particles were refined to a ribosomal volume with a clearly occupied P-site (green) and no E-site tRNA. The visible density for TF (red) was reduced compared to the final density in the FCMR alignment (Figure 45).
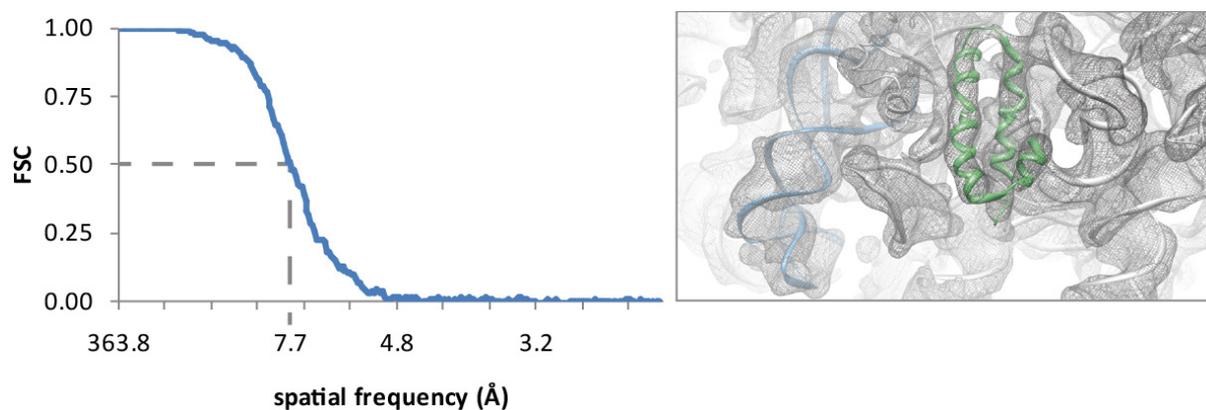
Figure 46: The resolution of the TF-RNC complex according to the FSC 0.5 criterion reaches 7.7 Å| The FSC plot in the left panel shows that the FSC 0.5 criteria is met at a spatial frequeny of 7.7 Å. The right panel shows a closeup of L29 (green) and the 23S rRNA (nts 1382-1420 and 1575-1604, in turquois) overlaid with a meshed density of the TF-RNC complex. The map features structural information like proper distinction of α-helices and resolved RNA helices.

4.3.3 Structural interpretation of the TF – RNC reconstruction

The improvement of the resolution of the TF RNC-map reduced the size of the additional TF density at the tunnel exit. As the occupancy of the TF-RBD judged by the density intensity in comparison to ribosomal proteins was very high, flexibility of TF could be an explanation of this phenomenon. If TF is flexible and this is true, low pass filtering of the cryo-EM maps at different resolutions should lead

to a partial recovery of TF. Figure 47 shows the same TF-RNC map filtered to different resolutions using the Fourier based buterworth filter. At resolutions lower 15 Å, the mid domain of TF is visible, while the Head-domain is never visible. At a resolution of 10 Å, only the TF-RBD is visble.
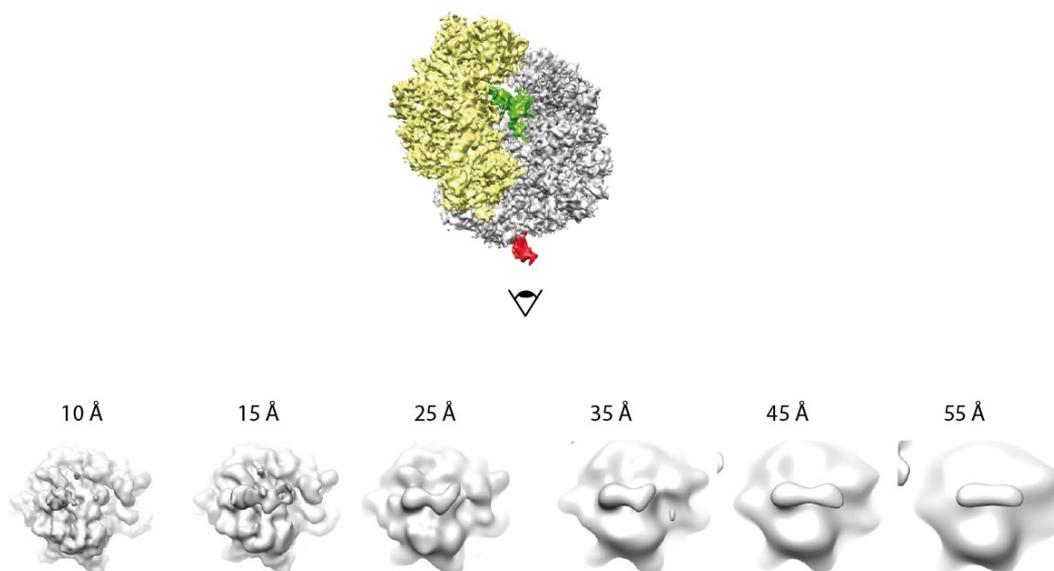
Figure 47: The visibility of TF changes as a function of resolution | The cryo-EM map of the TF-RNC complex was filtered to different resolutions. At high resolution, only the TF-RBD is visible, while under low resolution the mid domain becomes visible, too.

This behavior hints towards a highly flexible molecule that is bound to the ribosome via a rigid RBD. To investigate the dynamic behavior of TF, a normal mode analysis using the online available software NOMAD-Ref was performed (Figure 48). This analysis allows for a fast and simple approximation of possible movements of domains within protein structures. The normal mode analysis of a TF in solution (pdb:1w26a; Ferbitz et al., 2004) indeed showed a high flexibility in the Head-domain (normal mode 7 and 8) and lower but still strong flexibilities in the region of the Mid-domain (normal mode 9). As low normal mode frequencies (normal mode 7, 8) stand for higher contribution to the overall flexibility opposed to the high frequencies (normal mode 9 and following), these results are in agreement with the cryo-EM maps.
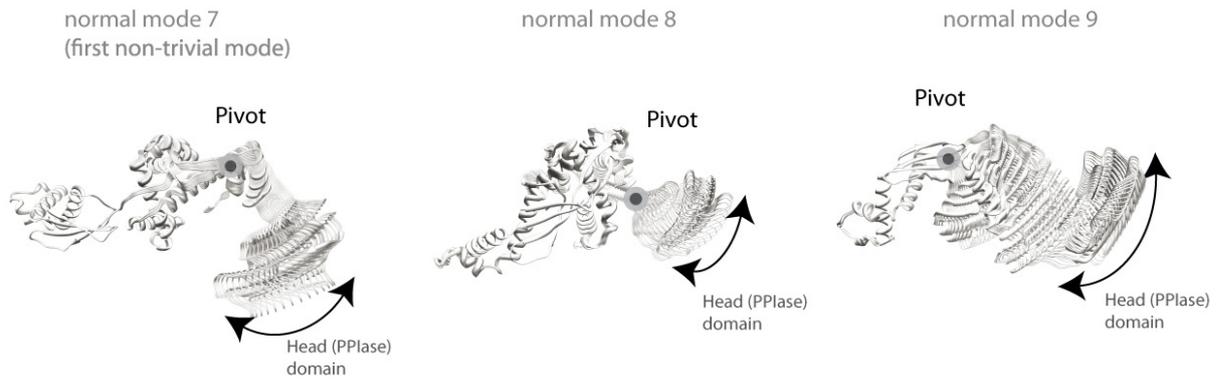
Figure 48: TF is a highly flexible protein | A normal mode analysis of the possible domain movements of TF revealed its high flexibility in the head regions, but also in the mid domain. The analysis was carried out using the NOMAD-Ref algorithm using pdb 1w26a (Ferbitz et al., 2004).

TF interacts with the nascent polypeptide. To investigate this interaction, the visualization of the nascent polypeptide is the first step. The TF-RNC structure has a high occupancy of the P-site tRNA indicating that the majority of the particles and thus the structure have a nascent polypeptide in the tunnel exit.

To verify this assumption, a section of the TF-RNC cryo-EM map along the path of the nascent polypeptide was cut and compared with a published structure of a TnaC-stalled ribosome (Figure 49, Seidelt et al., 2009)). The TnaC moiety of the nascent polypeptide in the TF-RNC map (colored in blue) adopts the same conformation and follows the same path as the published TnaC nascent polypeptide. N-terminal of the TnaC stalling sequence, additional density for the GatD segment is visible. Since the resolution of the TF-RNC map does not allow for a molecular interpretation of the nascent polypeptide, this direct comparison with a 5.8 Å map allowed for a clear assignment of the TnaC part of the nascent polypeptide.
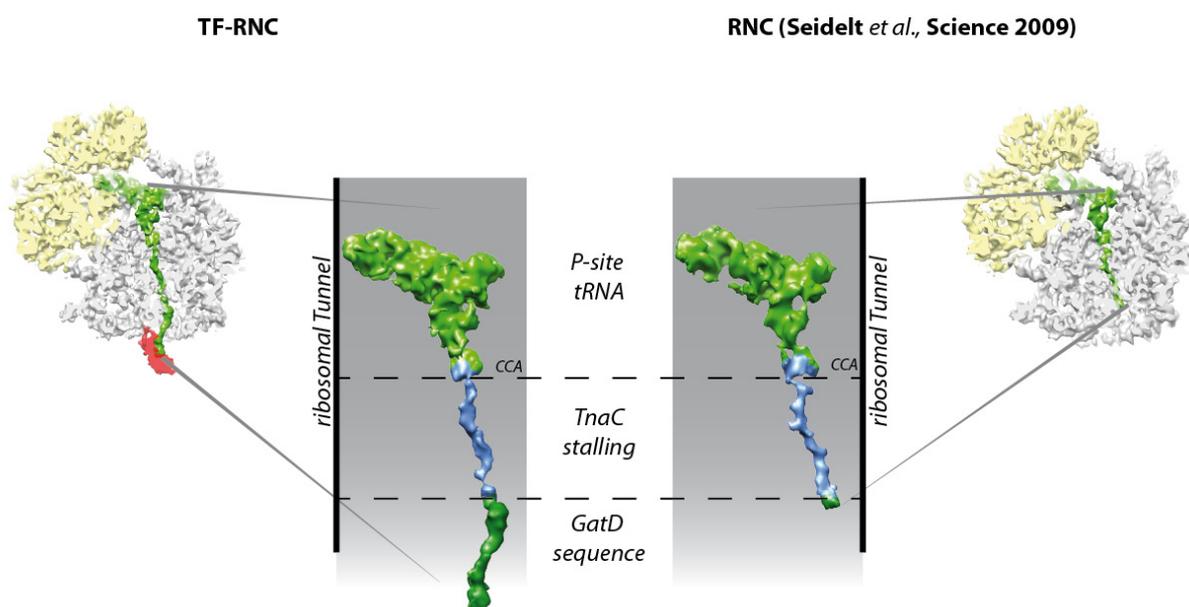
Figure 49: The TnaC-GatD nascent polypeptide of the TF-RNC is visible in the cryo-EM map | The nascent polypeptide in the ribosomal tunnel directly adjacent to the P-site tRNA follows the same path as the nascent polypeptide in a previously published structure of in vitro TnaC stalled ribosomes. Downstream of the TnaC stalling sequence, the polypeptide can be traced until it leaves the ribosome. Additionally, a compaction of the nascent polypeptide downstream of the constriction site of the tunnel indicates a helix formation in this region of the tunnel. (green: P-site tRNA, blue: TnaC stalling sequence, darkgreen: GatD sequence)

To further investigate the conformation of the nascent polypeptide and the location of the hydrophobic stretch, the density for the nascent polypeptide of the TF-RNC map was isolated and compared with the designed nascent polypeptide. The compaction of the nascent polypeptide N-terminal of the TnaC stalling sequence coincides with the presence of an α-helix in the GatD structure (AAs 93-102, Figure 50). This supports the direct comparison of the nascent polypeptide in Figure 49 and allows for a clear assignment of the hydrophobic stretch.

Moreover, the extra density that was assigned for TF showed a contact to the hydrophobic stretch at a distance that was in agreement with crosslinking experiments indicating that we indeed visualized the interaction between TF and the nascent polypeptide (Figure 50; Merz et al., 2008).
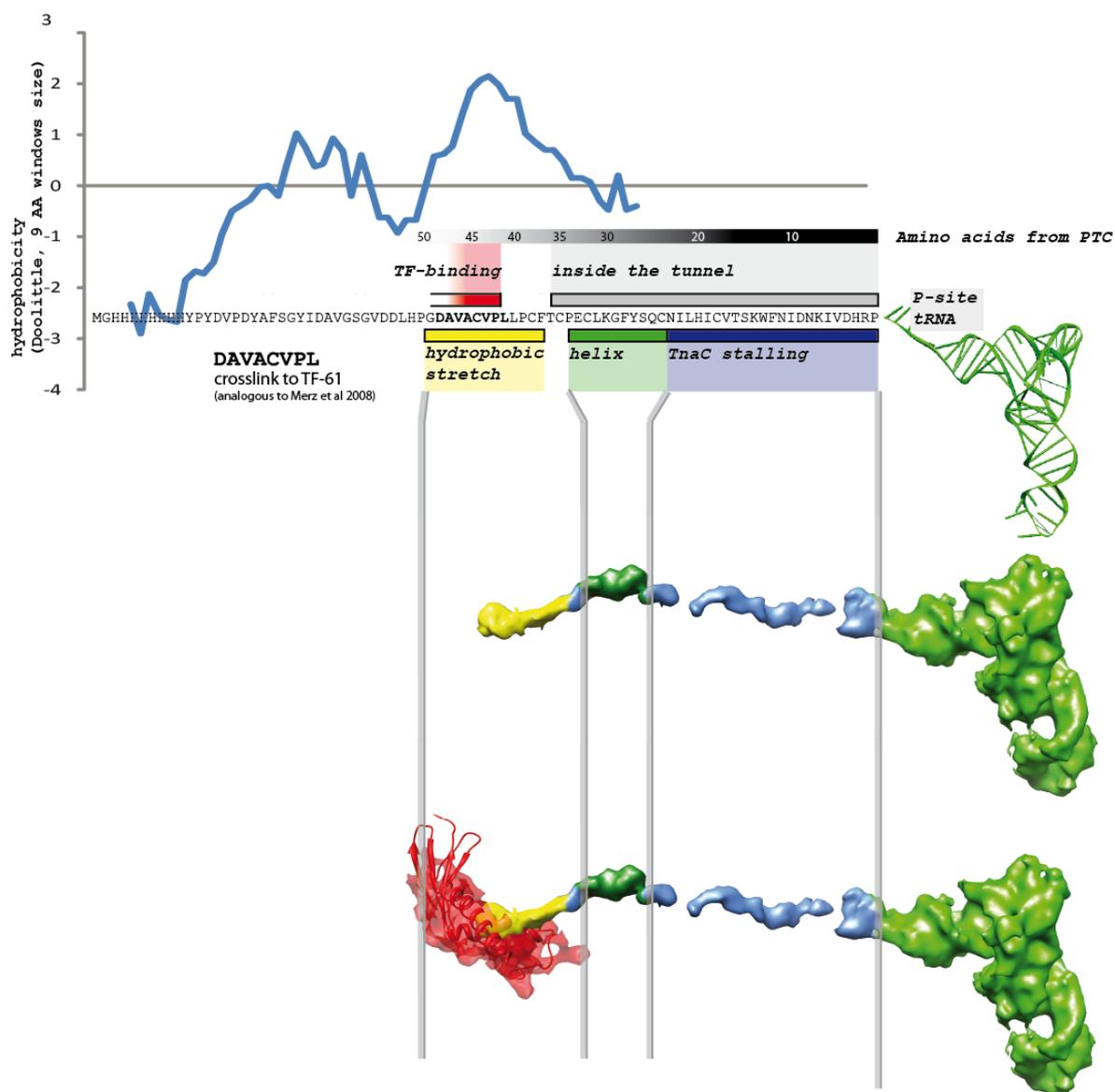
Figure 50: The compaction of the nascent polypeptide downstream of the TnaC stalling sequence allows for a unambiguous assignment of the hydrophobic stretch I and its interaction with the TF RBD | The helix in GatD that is composed of AA 93-102 is located downstream of the constriction site and thus can form a compaction (dark green). This compaction allows the verification of the assignment of the TnaC stalling sequence (blue) and a clear assignment of the hydrophobic stretch directly at the tunnel exit (yellow). The TF-density (red) contacts the hydrophobic stretch at a distance from the PTC which is in agreement with crosslinking data.

After the assignment of the TnaC stalling sequence and the position of the GatD sequence, modeling of the TF RBD was the next step. The ribosome binding domain of a crystal structure of *E. coli* TF (pdb: 1OMS; Kristensen and Gajhede, 2003) was used as a starting model. The RBD was placed on the ribosome according a X-ray structure of *D. radiodurans* TF RBD bound to ribosomes of the same species (pdb: 2D3O; Schlunzen et al., 2005).

The structure did not fit. Helix 1 and helix 2 had to be rearranged relative to each other. Since the loop connecting these helices was not resolved unambiguously, it was modeled according (i) the EM

density, (ii) the loop in structure 2D3O and (iii) the known interactions between TF and the ribosome (TF_Arg45 interacts with L23_Glu18 and 23S_A1392, TF_Arg57 with L29_Ala33 and TF_Gln56 with L29_Gln36) (Baram et al., 2005; Ferbitz et al., 2004; Schlunzen et al., 2005). This model was then regularized using the WinCoot algorithms (Emsley et al., 2010).
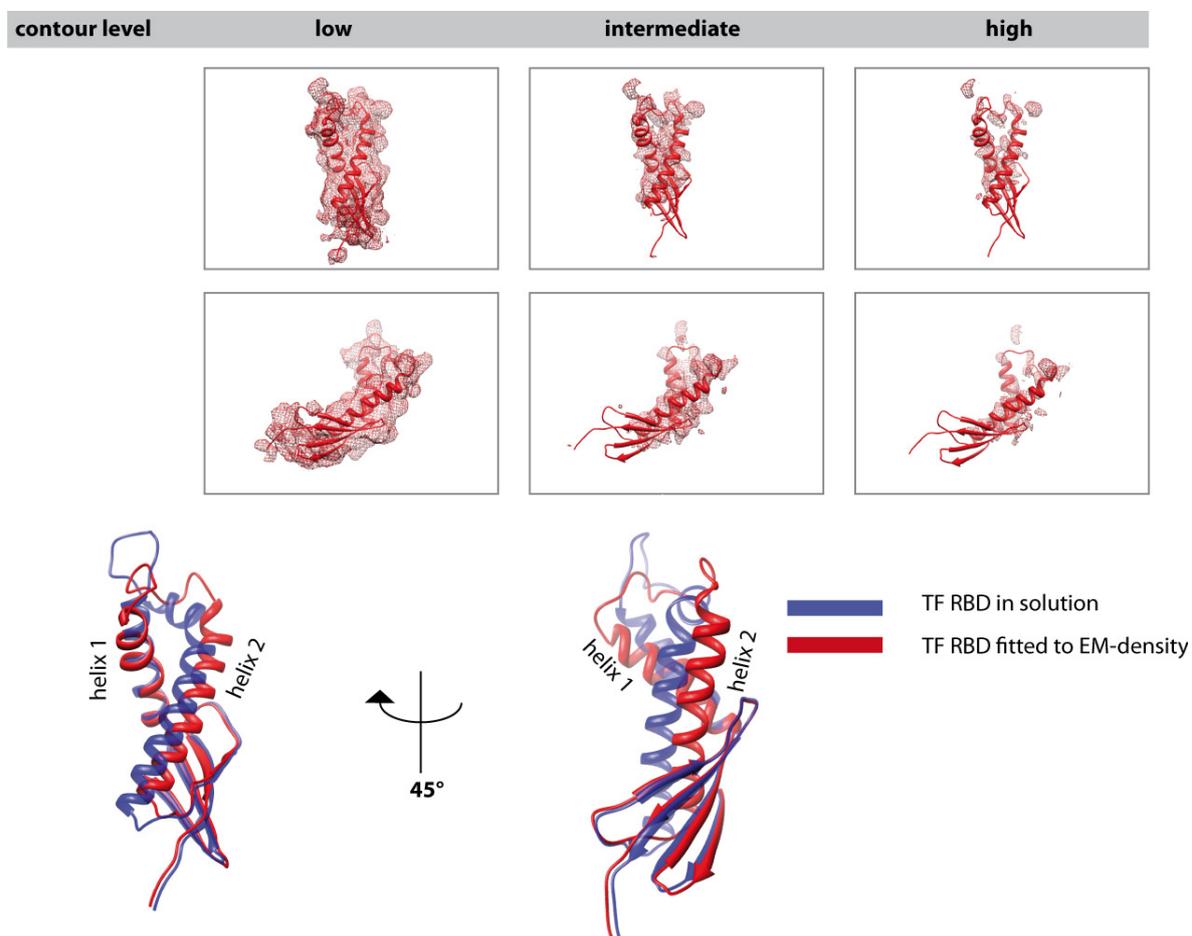
Figure 51: Modeling of the TF binding domain | A in solution structure of an *E. coli* TF RBD (1OMS, chain A, (Kristensen and Gajhede, 2003)) was placed in the TF density according to the X-ray structure of a co-crystal between a homologous system of *D. radiodurans* 50S and its TF-RBD (2D3O). Clear density for the two helices allowed the unambigous positioning of these parts. The connecting loop was only partially resolved and thus modeled based on the cryo-EM map density, the *D. radiodurans* TF-RBD and known interactions between TF and the ribosome. Finally, the model was regularized using WinCoot algorithms.

The rearrangement of α-helix 1 and 2 of TF leads to an opening of a hydrophobic patch between these two α-helices relative to the structure in solution (marked by a red box in Figure 52). This hydrophobic patch is placed directly at the tunnel exit and thus is potentially an interaction site between TF and hydrophobic regions of nascent polypeptides. Noteworthy, the hydrophobic stretch of the GatD nascent polypeptide can be traced as far as it reaches into the hydrophobic patch thereby supporting this notion.

Apart from the RBD, no reliable molecular interpretation of the RNC-bound TF could be obtained. However, low resolution information is available that can be used for subsequent analysis.
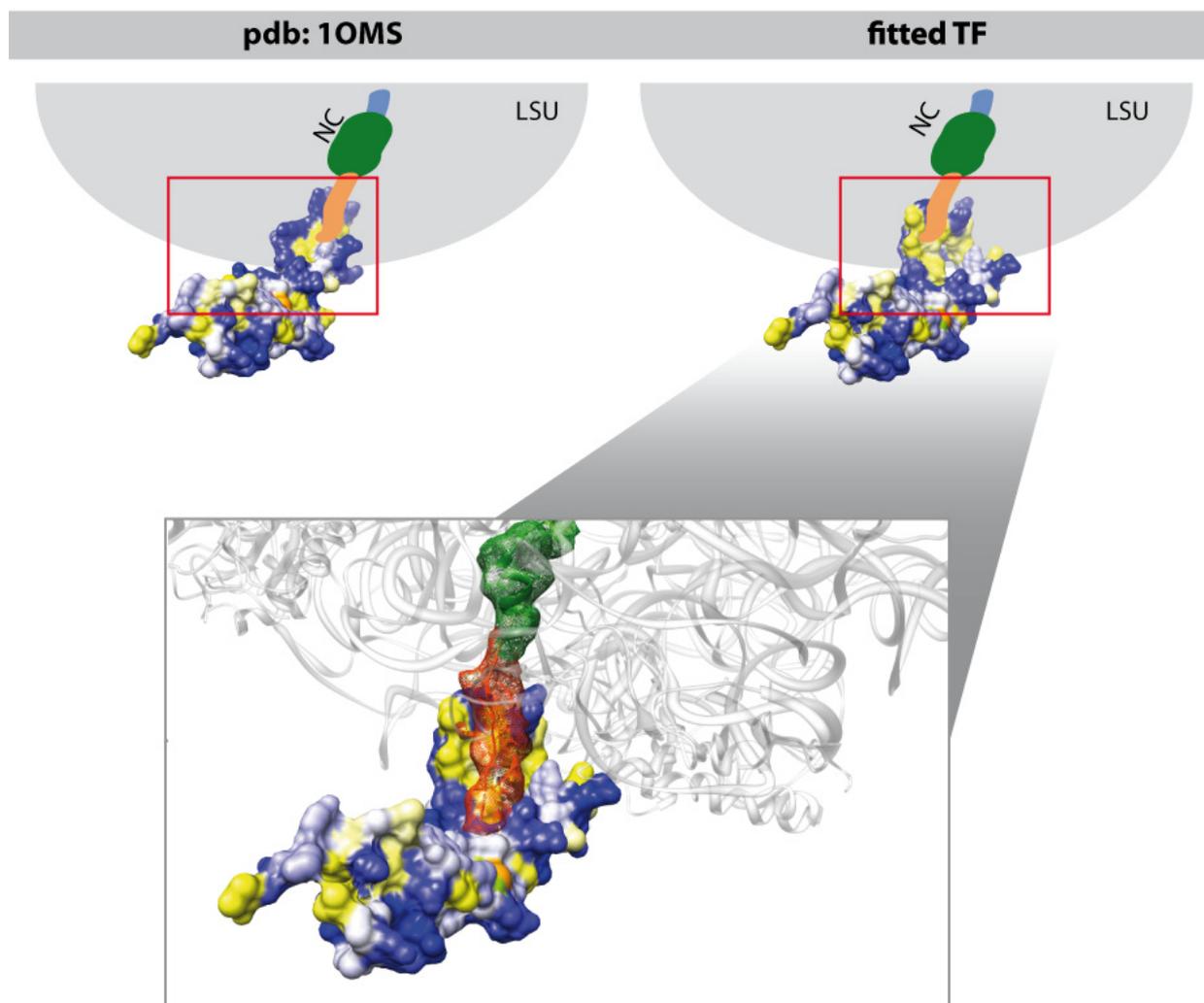
Figure 52: The rearrangement of helices 1 and 2 lead to an opening of a hydrophobic patch of TF-RBD directly at the tunnel exit that is perfectly located to interact with nascent polypeptides | The surfaces of a TF-RBD X-ray structure (1OMS, (Kristensen and Gajhede, 2003)) and the fitted TF were colored according to the hydrophobicity (blue: hydrophile, yellow hydrophob, red: methionin, LSU: large ribosomal subunit, NC: nascent polypeptide). The fitted TF RBD contained a widely open hydrophobic stretch that is directly located at the tunnel exit. It is thus perfectly placed to engage the growing nascent polypeptide. The lower panel shows a molecular model of the fitted TF placed into the density of the cryo-EM map (molecular model of the 50S subunit: pdb:2wwq, (Seidelt et al., 2009)). The nascent polypeptide was modeled according to the cryo-EM density. The color of the nascent polypeptide and the molecular surface reflects the functional parts (blue: TnaC stalling, green: GatD helix, orange: the hydrophobic stretch). For better visualization, the hydrophobic stretch in the nascent polypeptide is colored in orange in this figure.

The interpretation of a cryo-EM map is limited by its resolution. The assumption, that the resolution is constant over the whole map however does not hold true. For example, regions of high flexibility or substoichiometric complex formation are reflected in regions of lower resolution. This needs to be accounted for in the molecular interpretation as otherwise over-interpretation of low resolution regions might lead to wrong conclusions.

To determine the limit of interpretation, a new procedure that calculates the local resolution within small fractions of the map was used. The map was subdivided into smaller maps for which the resolution was calculated. As the resolution determination was conducted in the Fourier space, the resolution can be determined at a sampling rate that is dependent on the size of the local map. To enhance the sampling rate and the robustness of the local resolution determination, the original sub-map was padded prior to the resolution determination into a volume of larger size.
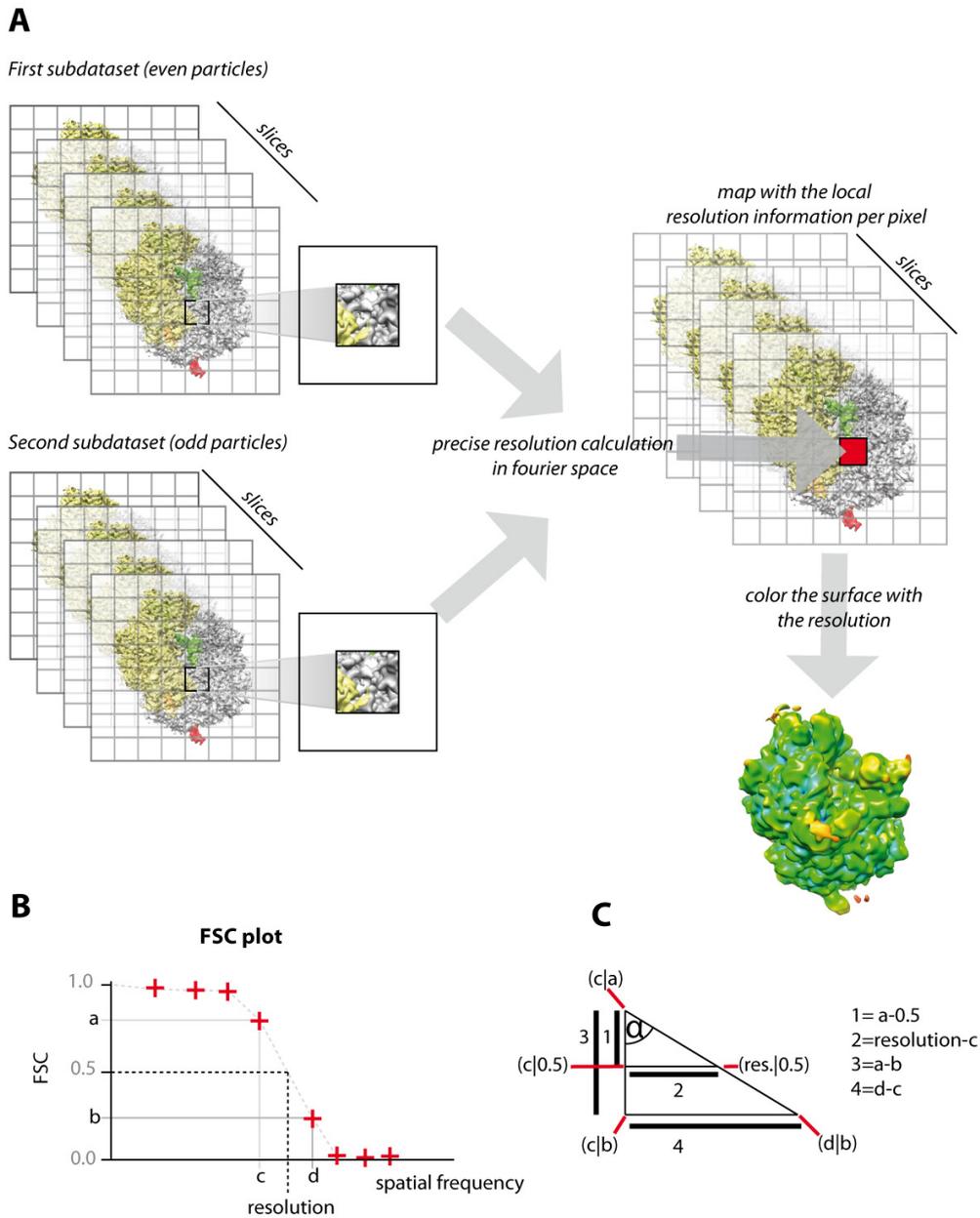
Figure 53: The local resolution determination scheme calculates the resolution within a subvolume and allows visualization of this information on the surface of cryo-EM maps| A: The dataset is split into two equally large subvolumes and two volumes are backprojected. The subdata volumes are boxed into subvolumes and the resolution is calculated using the FSC 0.5 approach. The calculated resolutions are then placed into a 3D file. The resolution can be displayed by coloring the surface of a cryo-EM map. B: The FSC plot sampling along the spatial frequency is limited by the dimension of the subboxes. To increase the accuracy of the resolution determination, the boxes are padded into higher dimension boxes. C: The resolution (the spatial frequency where the FSC is 0.5) can be calculated based on the cotangent (Formula 7).

Almost all coordinates (spatial frequency | FSC) are defined and only the resolution had to be calculated. For simplicity, the figure only shows a 2D chart of the scheme

Further improvement of the resolution determination was achieved by using a linear approximation of the FSC 0.5 criterion in the background of a low sampling rate (Figure 53, Panel B and C and Formula 7). The triangle created by the points in the FSC plot (c|a), (c|b) and (d|b) with the angle α has the same angle α as the triangle (c|a), (c|0.5) and (res|0.5). Therefore, the ratio between side 1 and 2 and side 3 and 4 are the same due to the common cotangent (α). Replacement of the sides by the expressions in Figure 53, Panel C leads to the final expression in Formula 7. Since the plot of the FSC as a function of spatial frequency is linear, but the conversion from spatial frequency to resolution is exponential (resolution = spatial frequency^{-1}) this improves the resolution determination in Å significantly.

$$\cot(\alpha) = \frac{1}{2} = \frac{3}{4}$$

with

1 = a - 0.5	3 = a - b
2 = resolution – c	4 = d - c

leads to

$$resolution = \frac{(a - 0.5) * (d - c)}{a - b} + c$$

Formula 7: The resolution in Figure 55, Panel B and C can be approximated using a cotangental expression | The cotangent of α (Figure 53, Panel C) describes the ratio between the sides 2 and 1 as well as between 3 and 4. This allows the calculation of the length of 2 and thus of the resolution. Since the spatial frequency is linear in Å^{-1}, this approximation increases the precision of the resolution determination significantly. It does, however, assume a constant slope between the point before and the one directly after the FSC 0.5 criterion.

With this tool in hand, the reliability of the TF RBD interpretation could now be assessed. The final map of the RNC with the defined TF was used and the local resolution within a box of 10^3 voxels (10x10x10 pixels) was calculated after padding the 10^3 boxes into 45^3 boxes. Since the pixel size of the final volume was 1.2375 Å/pixel, the local resolution was calculated at a sampling rate of 1.2375 nm on the original volume. The boxes were shifted by 2 pixels relative to each other leading to a high resolution local resolution map of 143^3 datapoints.

The inter-subunit space between the ribosomal subunits is solvent filled and thus not as well resolved as the core of the ribosome surrounding it. The resolution in this region should be lower relative to the core of the ribosome. Thus this allows the proof of principal of the local resolution determination method presented here.

The 3D-resolution map was therefore analyzed along a 2D-plane (Figure 54, panel A). The resolution distribution of central plane (plane 72 of 143 generated planes along the x-axis) along the central line (72 of 143 along the y-axis) was plotted against the distance to the edge of the box. Along the distribution of the core ribosomal density, the resolution was significantly higher in comparison to the surrounding. The peak of low resolution at the center coincidences with the position of the inter subunit space. As this feature is visible in all the planes that cross the inter subunit space, this observation verifies the information provided by the local resolution determination.
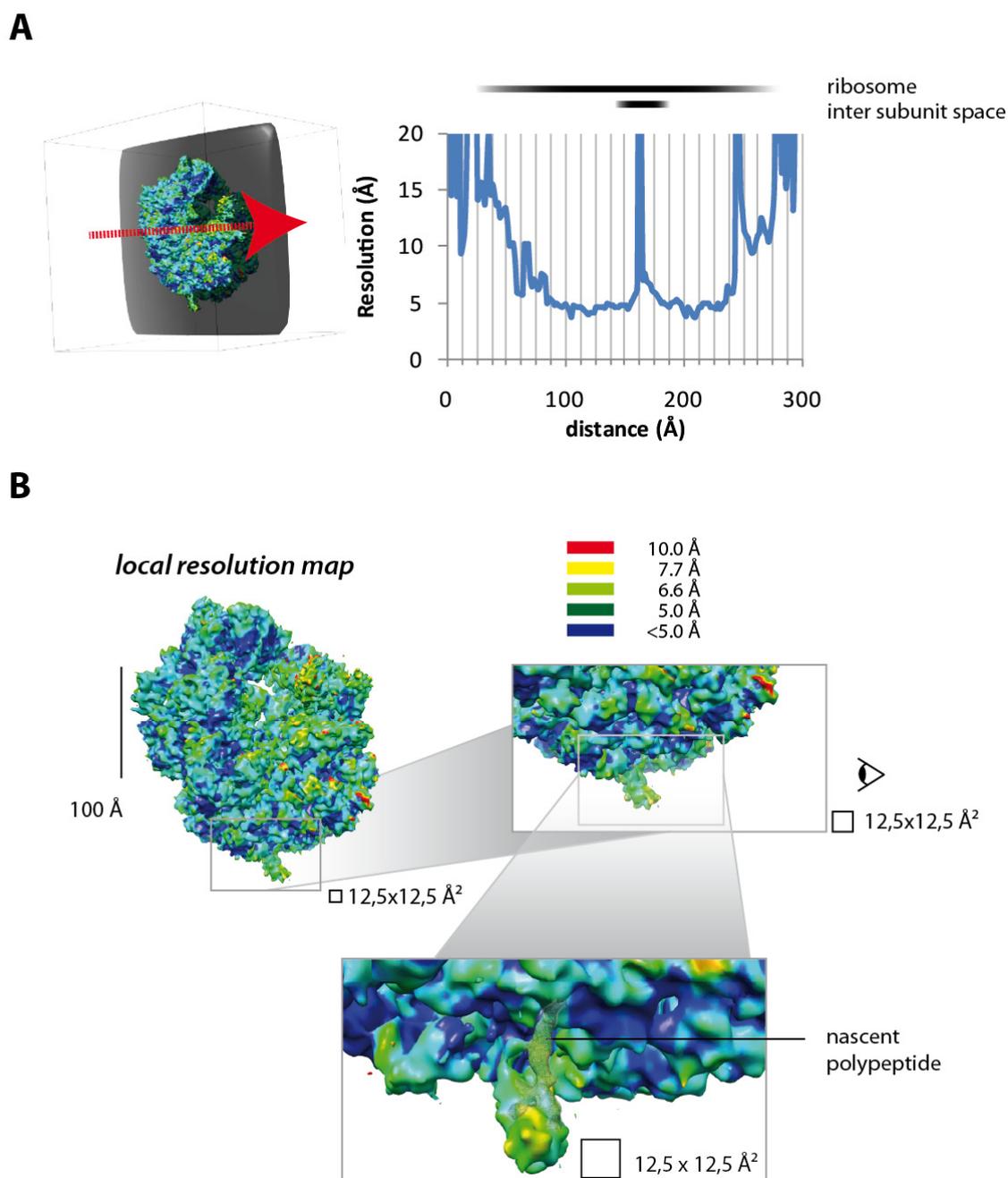
Figure 54: The local resolution map showed a highly resolved TF-RBD | A local resolution map based on boxes of 12.5³ Å overlapping by 2.5 Å with the next box was calculated. Before resolution determination, the subboxes were padded to 55³

Å maps. A: The local resolution map was analyzed in 2D to see the distribution along a central line in the central plane of the 3D map (red arrow). A drop in the resolution in the interior of the ribosome due to the inter subunit space was observed. The boxsize of the subboxes (12.5 Å) is indicated by the minor sampling lines in the graph. B: The local resolution map with 143^3 data pointes was then used to color the high resolution map. A close up of the TF-RBD shows a sub nanometer resolution for the TF-RBD and a resolution of better than 7 Å for the nascent polypeptide.

The local resolution map was then used to assess the resolution distribution on the TF-RNC map. The resolution is indicated by different colors (Figure 54, panel B). The map showed a local resolution that is higher than the previously determined 7.7 Å (Figure 46). The close-up of the TF-RBD in complex with the ribosome shows, that this region is resolved beyond the global resolution of 7.7 Å. Thus the TF-fitting (Figure 51) is reliable. Additionally, the nascent polypeptide is well resolved supporting the interpretation of the interaction between TF and the nascent polypeptide.

4.3.4 Analyzing the interaction of TF with ribosomes in context of different nascent polypeptides

The high flexibility of TF in the presence of nascent chains that were just long enough to traverse the interior of the TF raised the questions whether (i) longer nascent chains might stabilize or strengthen the interaction between TF and the nascent polypeptide and (ii) subsequently lead to a reduction in the flexibility of TF. The first question can be approached using the MST setup.

In sections 4.1 and 4.2, MST was applied to investigate the interaction between NAC and other ribosomal interactors to idle ribosomes. However, the investigation of the interaction between RNCs and these interactors was hampered by the difficulty to achieve yields in the RNC preparation that were high enough to use them for this biophysical analysis. Additionally, the stability and programming rate of the RNCs as judged by Western Blot analysis was not sufficient enough. The availability of *in vivo* expressed and stalled RNCs in *E. coli* in high quantities allowed a more quantitative analysis of the interaction of TF with RNCs carrying nascent polypeptides of different lengths.

Due to difference in size between TF and RNCs and the absence of any endogenous cysteines, a strategy using single cysteines at a defined site analogous to the NAC project was chosen. The availability of an X-ray structure of a full length protein and its defined domain and global structure allowed for a rational identification of suitable positions for the cysteine. Based on previous studies, T150 that is located between the Head- and the Mid-domain was mutated to a cystein (Figure 55).
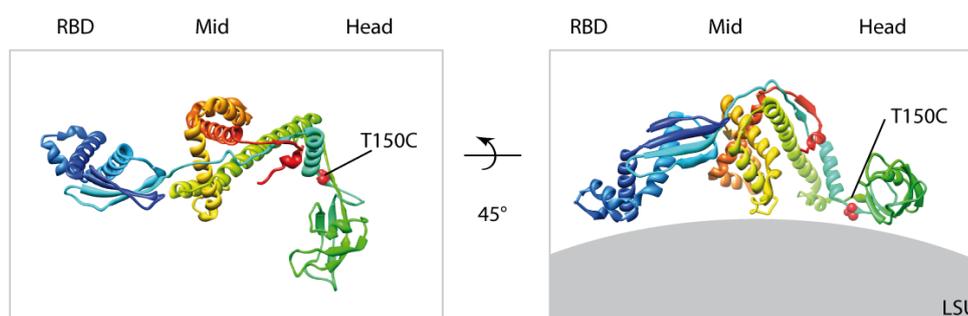
Figure 55: The threonine 150 of E. coli TF is located between the Head-domain and the Mid-domain | A study showed that this mutation does not affect the chaperone activity of TF (Kaiser et al., 2006). Thus this position was used to insert a single cysteine for subsequent labelling with Alexa568.

The TF150C derivative was then labeled with Alexa568 followed by the activated Thiol-Sepharose incubation step to enrich for labeled TF. Starting from 10 nmol, 4.5 nmol TF150C labeled with 4.4 nmol Alexa568 were recovered after the labeling procedure.

A template for RNC that was N-terminally extended to the endogenous N-terminal methionine was cloned and used for in vivo expression and purification (His$_8$ 80 AA GatD I-35 TnaC, total length: 144 AA, Figure 38).

The affinity between idle 70S ribosomes, His$_8$ 20AA GatD I-35 TnaC RNCs (total 82AA) and His$_8$ 80AA GatD I-35 TnaC (total: 144 AA) to TF150C was measured using the direct titration approach. The ribosomes were titrated against 20 nM of the labeled TF. The thermophoretic equilibrium was reached only slowly. Therefore, F$_{hot}$ was measured 95 second (950 time increments) after the IR-Laser was switched on (100% intensity). Due to the low signal intensity, the noise in the time traces was relatively high. Therefore, 100 time increments of 0.1 s each were integrated to yield F$_{hot}$.

Contrary to published data, we could not detect any binding of TF150C-Alexa568 to idle ribosomes (data not shown). However, a binding event between this TF derivative and RNCs was clearly detectable (Figure 56). The K$_d$ for TFs interation to RNCs carrying the short nascent polypeptide (82AA) was 110 nM. An increase of the length of the nascent polypeptide by 60AA increased the affinity of TF to these RNCs by a factor of 5 (22 nM). The absence of binding of TF150C-Alexa568 to idle ribosome indicates, that the binding between TF150C to the ribosome is challenged by the presence of the fluorescent dye.
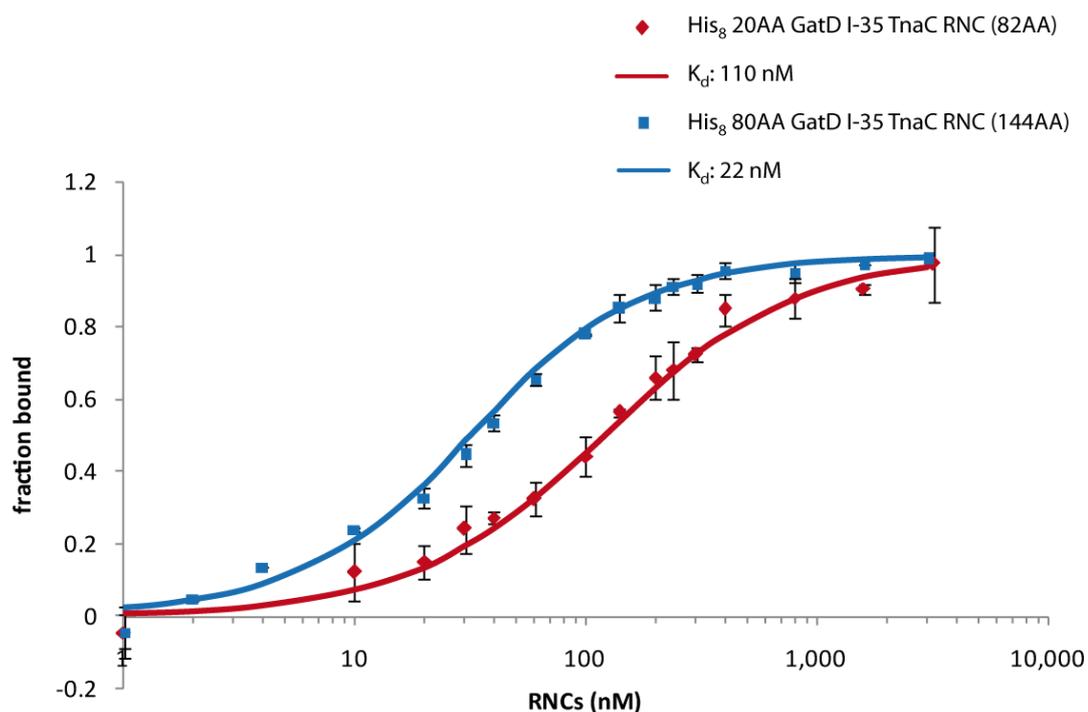
Figure 56: TF150C-Alexa568 binds with a high affinity to RNCs. This affinity is modulated by the length of the nascent polypeptide | His_8 20 AA GatD I-35 TnaC RNCs and His_8 80 AA GatD I-35 TnaC RNCs were titrated against 20 nM of TF150C in MST-Buffer. For better visualization, not the actual F_{hot}/F_{cold} ratio but the fraction bound is plotted. The fraction bound was based on the normalized fluorescence and calculated using Formula 3. The dissociation constant was calculated using automated curve fitting of Formula 4. The curve fit (solid blue and red lines) revealed a dissociation constant of 110 nM for the short nascent polypeptide and 22 nM for the long one. F_{hot} was integrated for 100 time increments 95 seconds after establishment of the temperature gradient.

The absence of binding of TF150C-Alexa568 to idle ribosomes indicates, that the binding between TF150C to ribosomes is challenged by the presence of the fluorescent dye. The increase in the K_d of TF to ribosomes as a function of the nascent polypeptide length in this context allows a qualitative understanding of the contribution of this functional element.

5 Discussion

5.1 NAC

While the function of NAC on ribosomes is still far from understood, this work provides insights into the interaction that is assumed to be crucial for its function – the binding to ribosomes. Using different fluorescence based techniques; a quantitative description of the NAC-ribosome interplay was achieved. Subsequently, this analysis was extended to intra-molecular rearrangements and finally a first step towards a systemic description of events at the tunnel exit was taken.

5.1.1 Measuring interaction between ribosomal ligands and ribosomes

To measure the dissociation constants (K_ds) for various ribosomal interactors to ribosomes, a multitude of techniques is principally available. The final goal of this study is to measure the binding between those interactors and RNCs. The low yield of available eukaryotic RNC preparation protocols makes it mandatory to first identify techniques that provide the highest sensitivity, thus using the lowest amounts of RNCs. Therefore, the technique that will be chosen, as well as the strategy to measure binding constants, has to be optimal with respect to sensitivity.

Currently, one of the most widespread techniques used to determine K_ds in biological systems is surface plasmon resonance (SPR). This technique uses an immobilized partner that is flushed with the ligand and exploits strongly enhanced electromagnetic surface waves on thin metal films (Kretschmann, 1971; Otto, 1968). These waves are altered by the dielectric properties of the surface. Moreover, their properties change as a reaction to the layers of surface coupled molecules, and more importantly their mass. The binding and dissociation of the partner protein can be directly monitored by a changed SPR signal. Thereby, on and off rates, and subsequently K_ds can be calculated. The drawback of this technique is the difficult establishment of a new assay, making it hard to optimize for high throughput measurements and the necessity to immobilize one binding partner.

To measure binding, features of the binding event can be directly exploited to deduce the complex formation of a given ratio between the interactors. Calorimetric changes upon binding or altered physical properties of the complex like tumbling rate in fluorescence anisotropy (FA) measurements, thermophoretic movement or changed diffusion constants can be used.

Isothermal titration calorimetry (ITC) is a veracious label free method and thus insensitive to artifacts of labeling. Additionally, it allows the direct measurement of the K_d, the enthalpic change (ΔH) and the binding stoichiometry. Furthermore, these parameters allow the direct calculation of the changes in the entropy ΔS) and in Gibbs free energy (ΔG) (Wiseman et al., 1989). However, it requires a measurable enthalpic change upon binding. Therefore, typically high amounts of the constant partner (in the mg/ml regime) are used to create a significant heat signature. This high consumption of material is the major drawback of this method. Additionally, it is very sensitive to changes in buffer conditions.

While ITC measures physical properties of the unlabeled molecule, the other techniques – FA, fluorescence correlation spectrometry (FCS) and thermophoresis – measure changed migrational features via fluorophores attached to the molecule of interest. FCS records the fluctuation of fluorescence in diluted buffer due to Brownian motions of the molecules (Magde et al., 1972). The fluctuation allows the calculation of diffusion constants and thereby the size of molecules. Comparing different ratios between ligand and receptors, this allows for the calculation of K_ds and even the extension of the analysis into the single molecule regime.

FA on the other hand records the tumbling of molecules in solution. The sample fluorophore is first excited with polarized light. Then, the polarization state of the emitted light is recorded (parallel and perpendicular to the excitation) (Weber, 1953). If the fluorophore tumbles between absorption and emission, the polarization of the emitted light is more and more perpendicular to the excitation. The fluorescence anisotropy is a measurement of the ratio between parallel and perpendicular emitted light by the fluorophore relative to the excitation light. As the time between absorption and emission for a given fluorophore can be assumed to be constant, a change in fluorescence anisotropy of emitted light is only due to changed tumbling rates. The tumbling rates in turn are a function of the molecule size, which is governed by the complex formation.

FCS and FA both suffer from the fact, that the signature in the recorded data that indicates binding, is dependent on the difference in size. While this generally holds true for a system where the ribosomal interactor is labeled while the ribosome is not, the initial plan was to label the ribosome – which is a constant variable through all the interaction studies – and use unlabeled ribosomal interactors.

Recently, MST extended the toolbox of available techniques using labeled molecules. In contrast to FA and FCS it does not primarily record the size of the labeled complex but rather small changes in features like the hydration shell. It has been used to monitor changed physical properties of proteins

upon binding of small molecules (Pollack et al., 2011). In principal, this should allow for the usage of labeled ribosomes and unlabeled ligands to standardize measurements. It is a sensitive technique that has only low material requirements. A typical experiment is performed in volumes as small as 5 µl with the concentration of the constant partner in the low nM range. Additionally, by using temperature gradients, it allows for the measuremens of temperature dependent effects on fluorophores as a function of the binding status (temperature jump).

Initial experiments using the protein conducting channel (PCC) and GFP labeled eukaryotic ribosomes indeed showed the potential of MST. Using labeled ribosome, the binding of unlabeled proteins could be recorded (A. Hiekel, personal communication). The extension to smaller interactors (notably without a micelle that increases the signal of PCC bound to ribosomes), however failed. NAC was too small to influence the MST signal of labeled ribosomes.

Due to the low material requirements, easy handling procedures and fast measurements of MST in comparison to the other methods, we extended the study to labeled NAC molecules. Thereby, the MST signal was enhanced since the smaller interactor was labeled. The affinity of NAC to ribosomes varied greatly over several orders of magnitude with respect to its dependence on nature and the position of the dye (Figure 15, p. 64).

That these different affinities are indeed a result of the labeling was verified by binding assays. While unlabeled NAC showed a stoichiometric binding to ribosomes, Egd2p V166C-Alexa555 NAC hardly comigrated with ribosomes through sucrose density cushions (data not shown). This observation supports the general applicability of MST to such systems and shows that weaker binding constants can be measured using this technique. However, it does not provide any physiological insights into which of the derivatives behave like wildtype. A thorough analysis of the NAC derivatives finally identified a species that behaved wild type like. This could be subsequently used to investigate NACs binding properties in more detail.

Finally, we could show in the last section that differences in affinity between ribosomal interactors and various RNCs can be measured. Due to the lack of sufficient amounts of eukaryotic RNCs, a pilot experiment was carried out in a prokaryotic system. The observation that the affinity of TF to ribosomes is governed by the length of the nascent chain is in agreement with published data. This supports the usability of the MST approach for this more complicated biological system. Thus, we state that the herein described approach using MST and various ribosomal interactors is suited for the thorough analysis of the interplay at the tunnel exit.

5.1.2 NAC's binding mode to ribosomes

To investigate the interaction between NAC and ribosomes in a physiological context, a labeled NAC derivative that binds like wildtype NAC to ribosomes had to be studied.

Artifacts of binding can arise due to two principal components of the derivatives: the placement of the labeling site and the nature of the dye. In a first approach, we wanted to reduce the impact of the placement of the dye and used an unspecific labeling strategy that modifies primary amines (lysines and the N-terminus). The chemical conditions were set such that on average 1.5 dyes were covalently linked to NAC. If interaction sensitive residues would be modified, these labeled molecules would be part of a small population as there are a total of 35 possible modification sites (33 lysines and 2 N-termini) (Figure 57). Under the assumption, that all primary lysines are labeled with the same efficiency, the majority of NAC molecules would be labeled on residues that do not interfere with the binding. Thus, this approach would dilute the effect of the labeling site on the binding event.
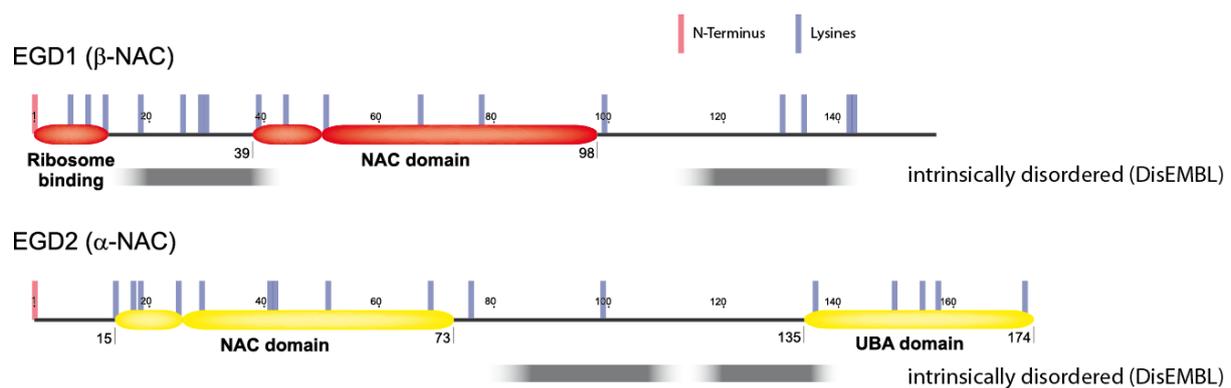
Figure 57: Overview of the presence of lysines in yeast NAC | In total, 33 lysines and 2 N-termini provide modification sites for the unspecific labelling of NAC. The lysines are marked with blue lines, the N-termini with red ones.

However, no binding of the unspecific labeled NAC to ribosomes was observed (even in co-pelleting assays). It is therefore tempting to speculate that some of the amines are preferentially modified and those sites contribute to the binding to ribosomes. As no further data of the site of labeling on NAC is available, the identification of these sites is not possible. However, in the light of the data that was acquired with the site specific labeling, this issue will be discussed later again.

Since the unspecific labeling approach failed, we screened different positions of NAC for its suitability to be labeled by placing unique cysteines at different functional regions (Table 1, p. 58). To account for the high sensitivity that is necessary for the later application of this setup, we first used NAC derivatives labeled with either Alexa 555 or 568. They are perfectly suited for a use in the MST device, as their absorption emission properties perfectly match the fixed setup of the device. These

dyes absorb in the green and emit in the orange electromagnetic sprectrum. We reasoned, that the choice of dyes that emit in this range would also reduce the interference with the auto-fluorescence of ribosomes. As little as 5 nM of these dyes were reliably detected (data not shown).

We expected to find a consistent appearance of a distinct affinity and a few deviations as a result of accidental modification of important residues. However a wide range of affinities without an obvious pattern was found (Figure 15, p. 64). As it is unlikely that all these residues contribute to the binding to ribosomes, we reasoned that the chemical nature of the used dyes might interfere with the binding irrespective of the positioning.

The Alexa568 fluorophore is a negatively charged, planar molecule (Figure 58; the same holds true for Alexa555). Alexa dyes are sulfonated to increase stability, brightness and to reduce the sensitivity to pH (Panchuk-Voloshina et al., 1999). This sulfonation leads to a negative charge which might interfere with the association to the ribosome, as the surface of the ribosome is negatively charged due to the phosphate groups of the rRNA backbone. We therefore extended the analysis to a smaller dye that is less charged, NBD (also called IANBD-ester).

IANBD-ester

N-((2-(iodoacetoxy)ethyl)-N-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole
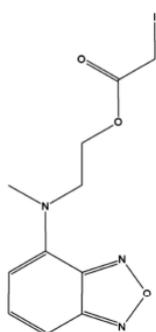
MW: 406 Da
Abs: 472 nm
Em: 536 nm

Alexa-568

Benzenedicarboxylic acid, [1,10-dihydro-2,2,10,10-tetramethyl-4,8-bis(sulfomethyl)-2H-pyrano[3,2-g:5,6-g']diquinolin-6-yl]-, compd. with N-ethyl-N-(1-methylethyl)-2-propanamine (1:1)
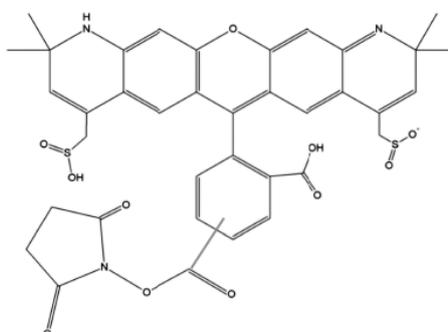
MW: 880 Da
Abs: 575 nm
Em: 600 nm

Figure 58: Comparison between IANBD-ester and Alexa-568 | Left panel: IANBD-ester is a derivative of the NBD dye that employs haloalkyl chemistry to link the NBD dye to cysteines. NBD is an environmental sensitive dye whose quantum yield varies in dependence of the microenvironment. Right panel: Alexa-568 is a sulfonated derivative of rhodamin. The sulfonation leads to a higher brightness and lower pH sensitivity. The sulfonation renders Alexa568 negatively charged.

To systematically address the question whether the chemical nature of the dye influences the binding properties on a given labeling site, we compared the binding properties of the Egd1p-A16C NAC labeled with Alexa568 and NBD. As this labeling site is directly next to the known binding site of

NAC to ribosomes, we expected to see an effect if the dye interferes with the binding. Indeed, we measures a 30 fold lower affinity for the Alexa568 derivative in comparison to the NBD derivative (Egd1p-A16C NBD NAC: 20 nM, Egd1p-A16C Alexa-568 NAC: 600 nM, Figure 15, p. 64). To investigate the influence of the positioning of the label on the binding, we extended the analysis further by using the same dye (NBD) at a different position (Egd1p-I3C NBD NAC). This cysteine is located, like A16C, next to the predicted binding helix (AA 4-15 in Egd1p, Figure 21 p.71). While I3C is N-terminal of that helix, A16C is C-terminal. Additionally, both side chains point in opposite directions with respect to the helix surface. An artifact introduced by the dye should be more pronounced in one derivative as the spatial orientation of the dye is different in the derivatives. On the other hand, if the affinity of both derivatives is the same, this would support the hypothesis that NBD does not interfere with the binding.

Unfortunately, we measured an 8.5 fold lower affinity of Egd1p-I3C NBD NAC for ribosomes as compared to Egd1p-A16C NBD NAC (Egd1p-I3C NBD NAC: 175 nM). This clearly indicates that at least one of these derivatives did not bind wildtype like. We therefore changed the strategy to identify the right derivative. So far, we tried to indirectly show wild-type like binding behavior by identifying derivatives that share common affinities. As artifacts should be randomly distributed, this would indicate a common, and thus wild-type like behavior.

Next, we tried to directly show wild-type like behavior. For this, we used a competition strategy. As we knew from co-pelleting studies that NAC interacts strongly with ribosomes, we started with the NAC derivatives that displayed the highest affinities: Egd1p-I3C NBD NAC and Egd1p-A16C NBD NAC. Unlabeled wild-type NAC was titrated against a preformed complex of labeled NAC and unlabeled ribosomes. Due to the steady state dissociation and association of the complexes over time, the labeled and the unlabelled NAC are bound to ribosomes as a function of their relative concentration. If however the affinities are different for the two NAC versions, the distribution is not equal, but one version is favored over the other.

To conduct this experiment, two parameters of the system have to be known: the affinity of labeled NAC to ribosomes and the relative mobility of free and complexed labeled NAC. The affinity of the labeled NAC to ribosomes is necessary to calculate the fraction of the preformed complex. The relative mobility of labeled NAC allows the determination of complex formation of labeled NAC as a function of the total NAC concentration.

The analysis is complicated by the fact that both NAC derivatives contribute to the complex formation. The effective NAC concentration is thus $c_{NAC-labeled} + c_{NAC-unlabeled}$. On the experimental side,

it is important to note that we are not able to record the bound fraction in total (fraction bound $_{total}$) but only the bound fraction that is labeled (fraction bound $_{labeled}$). In a competition experiment where the unlabeled partner is titrated against a preformed complex, this leads to a reduction of the fraction bound labeled as it is titrated away from the binding site at the ribosome.

$$fraction\ bound_{labeled} = \frac{fraction\ bound_{total} * c(labeled\ NAC)}{c(total\ NAC)}$$

Formula 8: In a classical competition experiment with identical proteins, the fraction bound$_{total}$ can be derived from the fraction bound $_{labeled}$ and the concentration of the proteins | The fraction bound $_{labeled}$ is dependent on the fraction bound of the total pool corrected for the ration of labeled NAC to total NAC. The fraction bound $_{total}$ in turn can be calculated using Formula 2 (5) p.60.

The application of Formula 8 to the competition for Egd1p-I3C NBD NAC with wild type NAC failed (Figure 17, p.67), the bound fraction of labeled NAC dropped faster than expected. We therefore applied a model established by Roehrl et al. that describes a situation where the affinities of the individual components to one binding site are independent and the binding is exclusive (Roehrl et al., 2004). We could determine that the dissociation constant of the unlabeled NAC to ribosomes is 20 nM The higher affinity of unlabeled NAC effectively competed with Egd1p-I3C NAC NBD at the ribosome.

Egd1p-A16C NBD NAC and wt NAC exhibit the same K_d (20 nM); this suggested that the Egd1p-A16C NBD NAC behaved wild type like. We could confirm this observation by competition experiments using Egd1p-A16C NBD NAC and wt NAC (Figure 17, p. 67). Furthermore the competition excluded a possible unspecific binding of the labeled NAC to ribosomes as we observed a clear competition.

The MST approach needed relatively high concentrations of 200 nM of the dye. As this concentration renders the K_d determination prone to error, special care had to be taken to not over-interpret the results. The K_d of wt NAC to ribosomes was calculated in two different competition experiments with the same outcome, a first hint that the strategy produces reproducible results. To further verify the quantitative description of the NAC-ribosome interaction, we used a different technique – fluorescence anisotropy – to calculate the affinity of NAC to ribosomes. Due to the setup, we could use lower concentrations of Egd1p-A16C NBD NAC. We could confirm the previously measured affinity (Figure 18, p. 68).

In conclusion we provide data that show that Egd1p-A16C NBD NAC is a suitable tool to investigate NACs behavior on ribosomes and that the MST-technique provides robust data when used to measure the interaction with ribosomes.

With this tool in hand we then investigated the binding mode of NAC in more detail. The binding site for NAC on the ribosomes is principally identified as being L31e for βNAC and L22 for αNAC (Pech et al., 2010). However, so far no structure for a ribosome-NAC complex is available. Principally, ribosomal proteins or rRNA can be the interaction partner of NAC. While there is evidence in the literature, that the L31e is part of the binding site and it directly interacts with NAC, the role of rRNA interaction was not addressed yet.

The N-terminus of Egd1p that mediates the interaction with the ribosome contains a possible rRNA interaction site. A number of positively charged lysines in a predicted α-helix might be a motif for the interaction with rRNA (Figure 59). They are predicted to show in the same direction with respect to the surface of the helix (Pech, 2006). Positively charged amino acids in an α-helix are a hallmark of several proteins that interact unspecifically with RNAs via the negatively charged RNA-backbone (see appendix 9.2, p. 152).
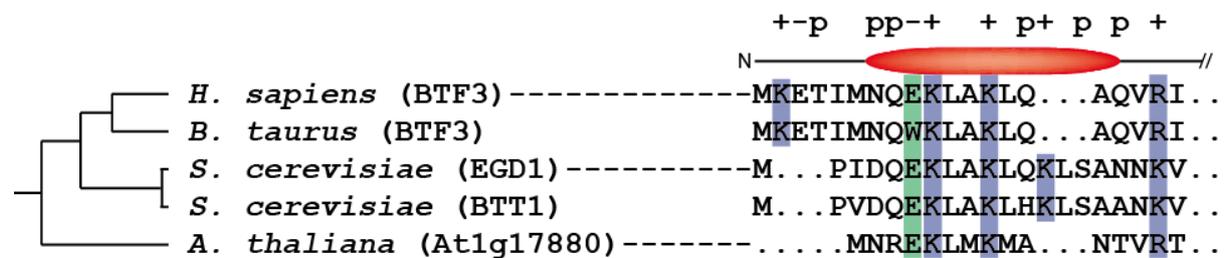
Figure 59: The N-terminus of βNAC is enriched in polar and positively charged amino acids | The N-terminus of βNAC contains conserved positively charged and polar AA that might mediate an unspecific interaction with the negatively charged backbone of nucleic acids. The relation between the species is not drawn to scale (*H.*=homo, *B.*=Bos, *A.*=Arabidopsis; +=positively charged AA, - = negatively charged AA, p=polar AA).

As it is known that NAC interacts with nucleic acids in an unspecific manner, we conducted a qualitative analysis of the interaction between NAC and tRNAs using EMSA (section 9.3, p. 154). tRNAs were used as they contain tertiary structure elements and are thus suitable to simulate rRNA. NAC interacted with tRNAs leading to a complex formation that was saturable (Figure 74, p.154). The fusion of the N-terminal region of βNAC (the first 39 AAs) was sufficient to bind MBP, that does not interact with RNA, to tRNA. To exclude the possibility, that the linker between the N-terminus and MBP mediates the binding to RNAs, we performed the same experiment with the fusion between MBP and the UBA domain of Egd2p. No complex formation was observed (data not shown).

While this analysis showed, that the N-terminus of Egd1 is sufficient to mediate the interaction with nucleic acids, we also observed aggregation of Egd1 (1-39) MBP in the presence of tRNAs. As this indicates most likely the formation of a complex where more than one fusion protein is bound per tRNA, a quantitative description of this interaction was not possible. Indeed, a MST analysis of this

system showed that we were not able to saturate the tRNA using this fusion protein. The long linker (4*(GGGS) 4.5 nm) reduces the steric restriction that occurs in wt NAC upon binding to tRNAs (data not shown).

While we show here that the N-terminus is able to interact with nucleic acids that contain a tertiary fold, this does not directly show that the interaction with rRNA plays a role in the binding of NAC to ribosomes. One more line of evidence however supports a model that includes binding of NAC via rRNA to ribosomes.

We saw a strong variability of the affinity of NAC to ribosomes as a function of magnesium concentration (data not shown). Doubling the concentration of Mg^{2+} ions from 10 to 20 mM reduces the affinity by a factor of 10. Reduction of the Mg^{2+} concentration on the other hand increases the affinity. As these experiments were performed at constant salt concentrations, this is a direct effect of this bivalent cation. While we do not understand the underlying mechanism, Mg^{2+} generally influences the RNA architecture (especially the tertiary structure) and might interfere with the binding of NAC by rigidification of the rRNA core (Batey et al., 1999; Celander and Cech, 1991).

In this context, it is also important to note, that we do not see any binding to ribosomes, when primary amines are labeled (see the unspecific labeling approach). An elegant explanation might be, that the same amines that are important for binding to the ribosome are surface exposed and thus preferentially labeled.

Finally, we investigated the influence of different domains of NAC on its recruitment to ribosomes. The N-terminal 39 AA of Egd1p when fused to MBP recruits the latter to ribosomes with a K_d of 230 nM. This intermediate affinity indicates that secondary, so far unidentified, binding elements must be present in NAC. We reasoned, that the UBA domain, whose function is enigmatic, might aid with the binding to ribosomes. UBA domains were found to be responsible for protein-protein interaction (see section 9.2, p. 153).

A C-terminal fusion between MBP and the *S. cereviasiae* Egd2p UBA domain did show weak binding to ribosomes, but the affinity was too low to be quantified. While this does not rule out the possibility that the UBA domain contributes to the binding to ribosomes, its contribution is minor at least. NAC is recruited to the ribosome via the interaction of the N-terminus of Egd1p and L31e. Due to the vicinity of the UBA domain to the ribosome upon this event, the local concentration of the UBA binding site is increased. Thus, already a low affinity for the ribosome is sufficient that the interaction occurs. While such a binding mode does not significantly increase the strength of the interaction, it might orient NAC ideally for the subsequent activity. Alternatively, the central NAC

domain could contribute to the overall binding of NAC to ribosomes. Upon an initial binding of NAC via the N-terminus of Egd1p to ribosomes, this domain is also brought in proximity of the ribosome.

5.1.3 NAC's conformation change upon binding to ribosomes

During the course of the MST experiments that used the NBD dye, we realized that the fluorescence of Egd1p-A16C NBD NAC was quenched in the presence of ribosomes. This dye indeed shows sensitivity to its microenvironment and it was previously used to monitor the hydrophobicity of the surrounding of the dye (Kaiser et al., 2006). Due to the auto-fluorescence of ribosomes, it was not possible to use this effect for binding studies. However, the quenching provided insights into the environment of the dye that allowed a qualitative understanding of the events that take place upon binding. As hydrophilic environments quench and hydrophobic environments enhance the fluorescence of NBD, three models of the change of the environment are principally possible (Figure 60).
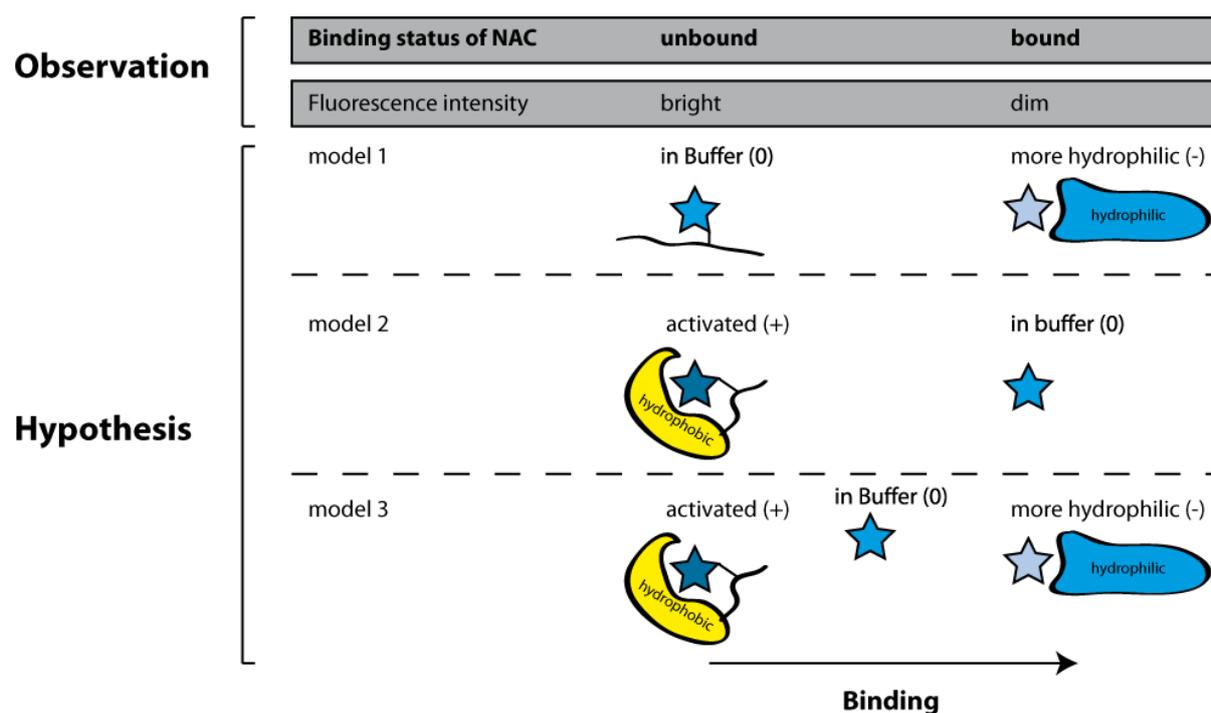
Figure 60: Three principal models can explain the change in the fluorescence of A16C NBD NAC upon ribosome binding | In the unbound state, the NBD fluorescence was relatively high, while in the bound state is was quenched. Each of the depicted models describe such an effect. If the dye was surrounded by buffer in the unbound state and upon binding gets transferred into the vicinity of a more hydrophilic environment (e.g. rRNA), the fluorescence would be moderately quenched (model 1). Alternatively, in the unbound state, the dye might be located in a hydrophobic environment and gets upon ribosome binding in contact with buffer. Thereby, the loss of activation translates into a reduction of the fluorescence upon binding (model 2). Finally, both effects can contribute to the quenching: In the unbound state, the dye is activated by a hydrophobic environment. Upon binding, its microenvironment changes and it is finally located in a hydrophilic environment (model 3).

The first model describes a situation where the dye is first surrounded by the buffer. Upon binding, it resides in a hydrophilic environment. The fluorescence intensity gets quenched. In the second model the quenching is actually a loss of activation. The dye resides in a hydrophobic environment that activates the fluorescence. Upon binding, it loses contact to the hydrophobic environment and simultaneously the activation. Finally, both effects may contribute to the observation that the preactivated dye loses the activation upon binding and furthermore gets quenched upon ribosome binding.

To distinguish between the three models, fluorescence intensity assays were performed. Tween-20 simulates a hydrophobic environment, and an increase in Tween-20 leads to an increase in the fluorescence of free NBD. When Egd1p-A16C NBD NAC was measured, the increase was only small (Figure 20, p. 70). This indicates that the dye was already almost maximally activated at minimal Tween-20 concentrations. In the presence of ribosomes, an increase of the detergence concentration increased the fluorescence of the complex consisting of Egd1p-A16C NBD NAC and ribosomes. Therefore, upon ribosome binding, the activation gets lost and an external activation is possible. Since control experiment showed that the affinity of NAC for ribosome is unchanged by the addition of up to 2% Tween-20, this effect was solely due to the changed microenvironment of the dye and not due to a changed complex formation.

These data exclude the first model for the Egd1p-A16C NBD NAC interaction. The second model implies, that upon ribosomes binding the dye is accessible by the buffer. The Egd1p-A16C NBD NAC fluorescence intensity was quenched upon ribosome binding irrespective of the Tween-20 concentration. Only at Tween-20 concentrations higher than 0.5%, an increase in the fluorescence intensity was visible. At the same time, 0.5% or less detergence activates free NBD. Therefore NBD in Egd1p-A16C NBD NAC is not accessible by buffer upon ribosome binding. But does the ribosome repress the fluorescence of NBD at position A16C of Egdp1p actively?

To address this question, we had to find a way to investigate the fluorescent properties of Egd1p-A16C in the absence of the fluorescent enhancement in the unbound form. We therefore aimed at identifying the region that enhances the fluorescence of NBD at position A16. The N-terminus of Egd1p folds into an amphipatic α-helix. This could in principal be sufficient to explain the observed activation (Figure 21, p. 71). We used a MBP fusion protein that contains the N-terminus of Egd1p (1-39). The dye was placed at position A16C to recapitulate the situation of the Egd1p-A16C NBD NAC. As this derivative showed a strong dependency of the fluorescence intensity as a function of the Tween-20 concentration, the activation site for NBD must be located further in other domains of NAC (Figure 25, 73).
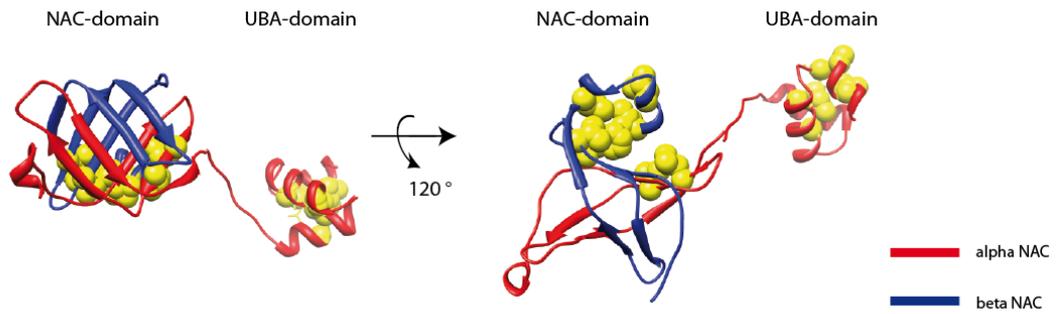
Figure 61: The NAC heterodimer contains two hydrophobic patches that are surface accessible | Apart from the central hydrophobic core of the NAC domain, two surface accessible areas allow the interaction with hydrophobic entities. A hydrophobic pocket consisting of the region of βNAC that is C-terminally of the NAC domain folds into a helix and binds back to the NAC surface. It thereby provides a relatively large binding site. A second hydrophobic patch is preset on one side of the UBA domain. The model shown here is built based on the heterodimeric human NAC domain and the aeNAC UBA domain (pdbs: human: 3mcB, aeNAC: 1tr8).

NAC features two distinct surface regions with hydrophobic properties (Figure 61). Both regions can be reached by the N-terminus of βNAC via the 20 AA long intrinsically disordered region that can span distances up to 60 Å.

The deletion of the hydrophobic patch on the surface of NAC is technically challenging and due to its complex nature very prone to artifacts. Therefore, we first deleted the UBA domain in the Egd1p-A16C NDB NAC derivative. In the unbound state, Egd1p-A16C NBD NAC ΔUBA did not change its fluorescence upon titration of Tween-20 while in the bound form it did (Figure 24, p.73). This indicates that the N-terminus resides in the same environment like in the full length version.

Since we do not know, which part of NAC activates this fluorophore, we used the Egd1p-A16C NBD MBP fusion to probe, if the ribosome quenches NBD attached to position A16 (Figure 23, p. 73). It binds to the same region of the ribosome, but lacks the activation domain as Egd1p-A16C NBD NAC. If Egd1p-A16C NBD MBP is bound to ribosomes, the fluorescence intensity is reduced irrespective of the Tween-20 concentration of the buffer (Figure 29, p. 78). This is even more remarkable since the addition of ribosomes leads to an increase in the fluorescence due to their autofluorescence. Thus we can conclude, that we see both: an activation of NBD in the unbound state and a repression in the ribosome bound state (model 3). This means, that the N-terminus of Egd1p is upon binding to the ribosome from a more hydrophobic into a more hydrophilic environment relative to the buffer used in this study.

While we cannot prove that the N-terminus of Egd1p folds back on the hydrophobic surface, the absence of any other detectable hydrophobic surface strongly suggests this conformation in the unbound state (Figure 62). This region consists of the surface of the NAC domain itself and an α-helix

that is located C-terminally of the NAC domain of Egd1p. Upon binding, the N-terminus dissociates from this region and interacts with the ribosome.

Additionally, quenching experiments showed, that the fluorescence of Egd1p-I3C NBD NAC is not preactivated (Figure 22, p. 72). Since the dye of this derivative points to a different side of the helix in comparison to Egd1p-A16C NBD NAC, this information allows for a placement of the α-helix. If the α-helix of is placed in a way that the hydrophobic AA interact with the hydrophobic stretch, A16 (notably also a hydrophobic AA) points towards this stretch (see also Figure 21, p. 71).

What is the function of this conformational change? Assuming that NAC interacts also with hydrophobic sequences of nascent polypeptides such as signal anchors, it needs hydrophobic interaction sites. However, such sites are entropically difficult to maintain in solution due to the hydrophilic nature of the surrounding. If now the N-terminus of Egd1p blocks this site by interacting with a hydrophobic motif, NAC's solubility is stabilized.

Upon binding to ribosomes, the N-terminus of NAC binds to ribosomes and opens the interaction site for nascent polypeptides. The hydrophobic side of the amphipatic N-terminus of Egd1 is then shielded by the ribosome.

Another important point is the protection of the integrity of NAC. While the function of the intrinsically disordered regions is not understood, they are potentially accessible for proteases. If the N-terminus of βNAC folds back on the NAC domain, at least the N-terminus of βNAC might be less accessible for proteases.
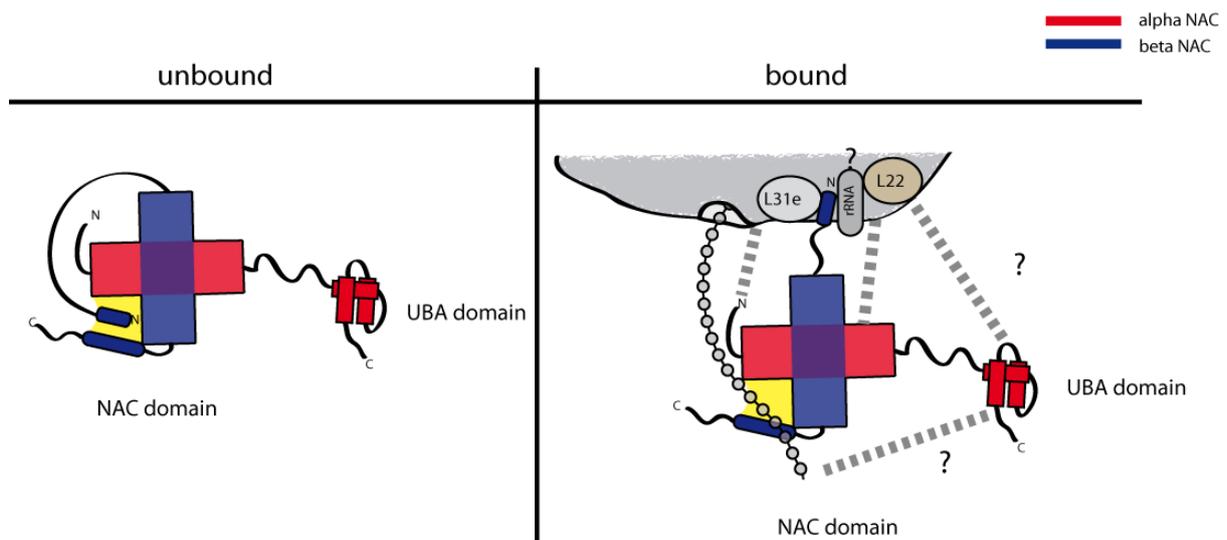
Figure 62: In solution, The N-terminus of Egd1p folds back on the NAC domain. Upon binding to ribosomes, a hydrophobic interaction site is provided that allows the interaction of hydrophobic nascent polypeptides with NAC | The hydrophobic patch created by a α-helix C-terminal of the NAC domain in βNAC is in solution occupied by the RBD of NAC. Upon binding to the ribosomes, the RBD releases this site and interacts with L31e and maybe rRNA. Additional interactions

between the NAC domain, the UBA-domain or the N-terminus of α-NAC and the ribosome are likely (Crosslinking studies showed, that L22 interacts with a yet unidentified part of Egd2p (Pech et al., 2010)).

In the ribosome bound state, NAC contacts the ribosome not only via the N-terminus of Egd1p at L31e, but also via the αNAC subunit which was crosslinked to L22 (Pech et al., 2010). How this interaction is established is not known yet, but since homodimerc aeNAC (two αNAC subunits) binds to ribosomes and crosslinks between αNAC and L22 are known, its role in the binding is evident.

5.1.4 Orchestration of ribosomal interactors at the exit site

The access of ribosomal interactors to nascent polypeptides has to be precisely regulated by the cell to obtain high fidelity of the maturation and translocation of the protein. Secreted proteins for example, not only play a role in shaping the environment of single cells (secreted enzymes or biofilm proteins) but also signal to adjacent or even distant cells. In multicellular eukaryotes for example, the developmental fate of tissues and steady state levels of various physiological parameters is regulated by secreted paracrine and endocrine acting factors – often proteins. Therefore, mistargeting of such proteins due to aberrant interaction with SRP not only reduces the fitness of the whole organism but also might lead to deregulation and subsequent death of the organism. There are a multitude of ribosomal interactors that are partly already introduced in section 1.2 (p. 11). Figure 63 gives an overview of the complexity of this interactome.
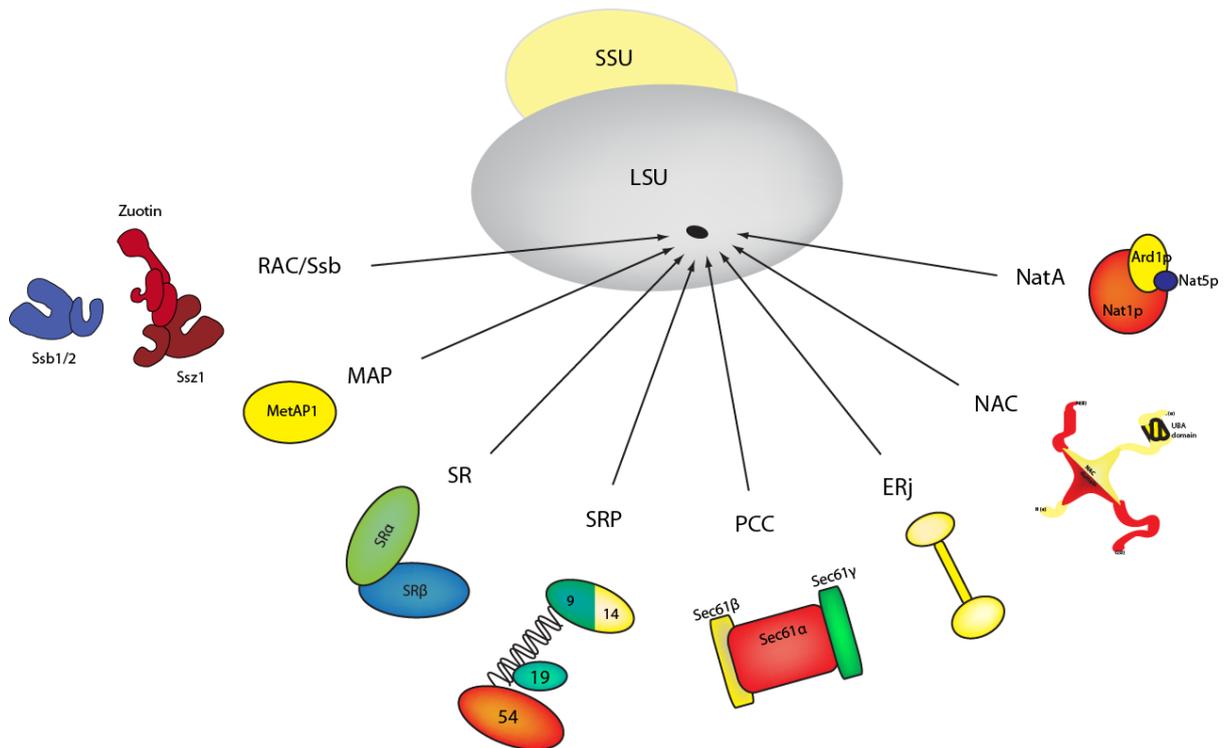
Figure 63 : A number of ribosomal interactors – chaperones, binding modulator proteins and enzymes – interact with the ribosome and the nascent polypeptide | Chaperones such as the Ssb1/2-RAC system, NAC and the ERj as well as nascent polypeptide modifying enzymes like MAP or NatA interact with the ribosome. Additionally, binding modulator proteins such as SRP, SR or the protein conducting channel PCC use the ribosome to interact with nascent polypepitdes (figure created in collaboration with Christoph Leidig, AG Beckmann).

Section 5.1.2 focused on the understanding of a highly reduced system of a ribosomal interactor and the ribosome. The goal of this project was to extend the analysis to the regulation of the access of different interactors to the ribosome. While the experimental data provided here are limited to a small subset of proteins in the absence of nascent polypeptides, this study lays the ground work for a more global investigation of this intricate system. The presence of additional layers of complexity by the chemical nature of the nascent polypeptide and more importantly the effects of ongoing translation in contrast to a static inspection shows the hurdles that have to be taken until a systemic understanding can be achieved.

The availability of a system where we could measure the affinity of NAC to ribosomes and a mathematical framework that describes the (complete) competition between two ligands to the same binding site allowed potentially the extension of the analysis to interactors of the UAS2. We conducted competition experiments between NAC and RAC and between NAC and SR in the absence of nascent polypeptides. Our data show that NAC and RAC cannot co-bind to ribosomes while NAC and SR can (NAC-SR, Figure 31 p. 80; SR-NAC, Figure 33 p. 83; NAC-RAC, Figure 34 p. 84).

While a competition can be shown directly, the absence of a competition has to be shown indirectly. We could clearly see a complex formation of labeled NAC and ribosomes according the MST-mobility that was not impaired by titration of SR. Therefore, the absence of a competition was obvious (Figure 33, p. 83). In an inverted setup, we preformed a complex between ribosome and labeled SR and titrated labeled NAC. Again, the complex formation of SR-ribosomes was not perturbed by increasing concentrations of NAC (Figure 31, p. 80). To directly show, that the titrated NAC indeed forms a complex with ribosomes in the presence of SR-ribosomes complexes, we investigated the fluorescence of labeled NAC in the second setup in more detail. The analysis of the complex formation was complicated by the fact that an increasing concentration of the labeled NAC increased the total fluorescence intensity. However, since the ribosome concentration (and therefore the autofluorescence), was constant we could use the fluorescence quenching of Egd1p-A16C NBD NAC upon binding to ribosomes as a measure for complex formation.
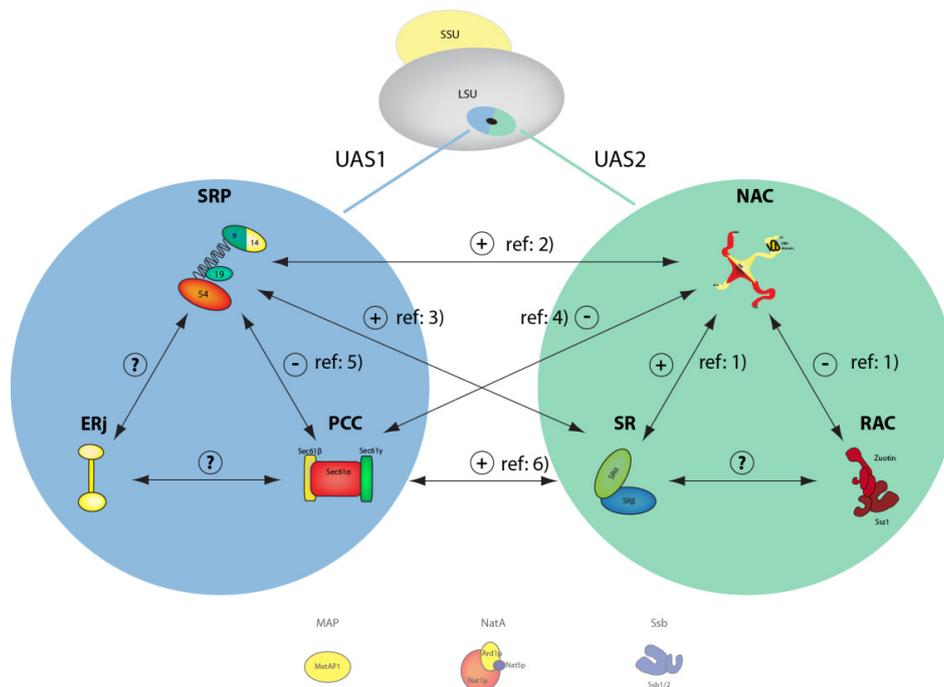
Figure 64 : The interactors of Figure 64 sorted according to their interaction platform show multiple interdependencies | Qualitative descriptions of the competition between ribosomal interactors are already available in the literature or can be derived from structural studies. An overview of the so far investigated interactions is shown in this figure. (ref: 1) this work, ref: 2): (del Alamo et al., 2011; Zhang et al., 2012), ref: 3): , ref: 4): (Moller et al., 1998), ref: 5): (Beckmann et al., 2001; Halic et al., 2004), ref: 6): (Angelini et al., 2005; Moller et al., 1998), + indicates cobinding, - competition and ? no data available, figure created in collaboration with Christoph Leidig, AG Beckmann)

Using the fractional fluorescence (total fluorescence intensity normalized to the labeled NAC concentration) we could see that under low concentration of NAC, most of NAC is bound to ribosomes. Under high concentration (>300 nM) additional NAC resides mainly in an unbound state

in the test tube due to the mostly occupied NAC binding sites (Figure 32, p.82). Thus, we concluded that SR and NAC can co-bind without steric clashes.

Based on these data and previous studies, a first step towards a more global description of the steady state competition between ribosomal interactors in the absence of nascent polypeptides can be taken. Figure 64 shows the so far known interdependencies on a qualitative scale. If a competition between two interactors –like NAC and RAC – was measured *in vitro*, the *in vivo* situation has to be carefully extrapolated. To understand the real situation, these interdependencies have to be seen in context of the real concentrations in the cell.

Table 2: Overview of the concentration of ribosomal interactors in yeast | The number of the molecules are from (Raue et al., 2007). The volume of an average yeast cells was assumed to be 39.5 fL (in the literature, two average cell volumes of *S. cerevisiae* cells grown at 30°C in YPD are available: 37 fL (Tyson et al., 1979), 42 fL (Jorgensen et al., 2002) (37 fL+42 fL) / 2=39.5 fL). The cytosol accounts in a human pathogen yeast (*Exophialia dermatitidis*) for 43-53% of the total cell volumes (Biswas et al., 2003). Due to lack of other sources, we therefore assumed that the cytosol in *S. cerevisiae* accounts for 50% of the total cell volume (= 19.75 fL). No molecule numbers are available for SR, the Translocon and ERj. (remarks: C=cytosolic protein, M=membrane bound protein)

	Number of molecules/ cell	molecules/ fL	Cytosolic concentration (µM)	Remarks
NAC	391,000	19,797	32.9	C
ribosomes	315,000	15,949	26.5	C/M
Ssb1/2	280,000	14,177	23.5	C
RAC	86,100	4,359	7.2	C
Map1	21,100	1,068	1.8	C
SRP	7,850	397	0.7	C
NatA	7,630	386	0.6	C
Map2	6,210	314	0.5	C
SR	?	?	?	M
Translocon	?	?	?	M
Erj	?	?	?	M

Table 2 shows the copy number of of various ribosomal interactors in *S. cerevisiae* cells as estimated by semi-quantitative western blotting (Raue et al., 2007). It is however important to note, that there are big discrepancies in the literature regarding the cellular NAC concentration. A study that used GFP-tagged proteins to quantify the number of molecules per cell estimated the cellular concentration of Ssb1/2p, RAC, Map1, SRP, NatA and Map2 in the same range as the study by Raue et al (Ssb1/2p: 274,000, RAC: 86,400, Map1: 20,000, SRP: 8,000, NatA: 7,600, Map2: 1,080 molecules per cell) (Ghaemmaghami et al., 2003). However, the numbers of Egd1p molecules per cell was estimated as 18,000 and 38,000 for Egd2p. Since the study that used semi-quantitative western

blotting as a read-out measured the concentration of untagged proteins in one experiment simultaneously, for the following analysis we refer to these data.

Based on these numbers, we calculated the cellular concentration by using a volume of an haploid *S. cerevisiae* cell of (39.5 fL) (Jorgensen et al., 2002; Tyson et al., 1979). Since the proteins are located in the cytosol, the fraction of the whole cell covered by the cytosol had to be known. No study in *S. cerevisiae* addressed this question to our knowledge. However an electron microscopy study determined this fraction of the human pathogen yeast *Exophialia dermatitidis* (Biswas et al., 2003). In this organism, the cytosol accounts for 43-53 % of the total volume. We therefore assumed that the cytosol accounts for 50% of the whole cell volume.

For membrane bound interactors like ERj, SR or the PCC the nature of membrane proteins makes it difficult to integrate them into complex formation calculations. Since membrane bound proteins do not diffuse freely in the cytosol but along the 2D plane of the surface of the respective membrane environment, their effective concentration deviates significantly from their actual concentration.

With these cellular concentrations, we can now investigate the steady state levels of different complexes in situations where all players are present or one is removed. Due to our limited information, we can only simulate the artificial situation where all ribosomes are idle and the interactor pool is completely available for interaction with the ribosome. Table 3 shows the result of these calculations. We can calculate the respective complex formations using the affinities determined in section 4.1.5 (Figure 34, p. 84). While in the presence of NAC, 1/3 of the total RAC concentration is unbound by the ribosome, in the absence of NAC, virtually all RAC is bound to ribosomes.

To show the robustness of the analysis, we calculated a scenario where only half of the ribosome is accessible for interaction with NAC and RAC. While this scenario also does not reflect the *in vivo* situation, it allows us to see, how the system reacts if one of the components is changed significantly. The outcome of this scenario (Table 3, B) shows that under these conditions, the absence of NAC has an even higher impact. The pool of free RAC is under low ribosome concentrations significantly higher as under high ribosome concentrations (66% of all RAC vs 33% in the latter). The removal of NAC in this system again leads to an almost complete sequestering of RAC by the ribosome.

Table 3: Simulations of the binding status of NAC and RAC and ribosomes in an isolated system | A.): Based on the cytosolic concentrations of NAC (32.9 µM), RAC (7.2 µM) and ribosomes (26.5 µM) and the K_ds for NAC and RAC to ribosomes (20 nM) two situations are simulated. Under standard conditions, all interactors are present at the respective concentrations. Using Formula 3 (p. 65), the ribosome bound NAC and RAC fractions can be calculated The second situation describes the binding status of RAC in the absence of ribosomes using

Formula 2 (p. 60). **B.):** A modified version of scenario A.) where only half of the ribosomes are accessible to binding to NAC/RAC is calculated. The same formulae are used. The sums have been rounded up to the closest decimal.

A.)		**free**			**ribosome bound**	
ribosome conc.= 26.5 µM	**ribosomes**	**NAC**	**RAC**		**NAC**	**RAC**
standard conditions	0.1 µM	11.1 µM	2.4 µM		21.7 µM	4.7 µM
% of total protein		**33.7%**	**33.3%**		**66.0%**	**65.3%**
% of ribosomes	0.4%				81.9%	17.7%
in the absence of NAC	19.3 µM	n.A.	~ 0 µM		n.A.	~7.2 µM
% of total protein		**n.A.**	**~ 0 %**		**n.A.**	**~ 100%**
% of ribosomes	72.8%				n.A.	27.2%

B.)		**Free**			**ribosome bound**	
ribosome conc.= 13.25 µM	**ribosomes**	**NAC**	**RAC**		**NAC**	**RAC**
standard conditions	10 µM	14.1 µM	4.8 µM		10.8 µM	2.3 µM
% of total protein		**42.9%**	**66.7%**		**32.8%**	**31.9%**
% of ribosomes	0.1%				81.5%	17.4%
in the absence of NAC	6.1 µM	n.A.	~ 0 nM		n.A.	~7.2 µM
% of total protein		**n.A.**	**~ 0 %**		**n.A.**	**~ 100%**
% of ribosomes	45.7%				n.A.	54.3%

The outcome of the calculation can now be compared with experimental data. The aforementioned study by the Rospert group not only determined the number of molecules per yeast cell, but it also quantitatively measured the occupancy of ribosomes by ribosomal interactors (Raue et al., 2007). Moreover, they separately looked at non-translating and at randomly translating ribosomes in a semi-quantitative manner. Table 4 shows the concentration of the respective ribosome-ribosomal interactor complexes. These data support our calculations as the non translating ribosomes showed a complex formation with NAC in the range of 23 µM (in our simulation 21.7 µM) and with RAC 4.9 µM (in our simulation 4.7 µM). Additionally, these data show that while the association of NAC with ribosomes is independent of the translation statues, RAC (and also Ssb1/2) gets recruited by active translation.

In summary we could show that a theoretical description of complex formation in complex situations like multiple binding sites (also exclusive binding sites) based on reduced *in vitro* setups can reflect the *in vivo* situation. Thus, based on this initial study, a systemic approach to describe the events at the tunnel exit site is promising. However, problems that increase exponentially with the grade of complexity might hamper the systemic approach. Special care has to be taken as the investigation of reduced systems is prone to failure to detect unintuitive functional interactions. Moreover, our system operates so far under the assumption, that the binding is 100% exclusive. While this

approach works perfectly fine for the NAC-RAC competition, it might not describe the situation of other partly competing binders. Finally, the mathematical difficulty to describe more complex systems remains challenging.

Table 4: Measured concentrations of complexes between ribosomes and ribosomal interactors in yeast | Based on the percentage of ribosomes occupied with ribosomal interactors and the ribosomal concentration in yeast cells (26.5 µM), the concentration of ribosomal complexes can be determined. Nontranslating ribosomes were obtained by carbon source starving of yeast cells, while randomly translating ribosomes were stalled using cycloheximide. The data is from (Raue et al., 2007)

	ribosomes	
Complex between ribosomes and:	**Non translating**	**randomly translating**
NAC	23.4 µM	23.5 µM
Ssb1/2	4.0 µM	8.4 µM
RAC	4.9 µM	9.4 µM
Map1	0.5 µM	1.0 µM
SRP	0.0 µM	0.2 µM
NatA	0.5 µM	1.0 µM
Map2	0.0 µM	0.5 µM

What does this now mean for the cell? The sequestering of RAC by ribosomes in the absence of NAC might lead to problems in RACs function. The relatively high concentration of RAC, however, might provide a large enough cytosolic pool that allows for efficient enough screening of newly translated nascent polypeptides.

In a different setup, the SRP-NAC-ribosome system, the effect might by much stronger. SRP's low concentration makes it potentially much more prone to disturbance of the balance of cytosolic vs. ribosome bound SRP. Indeed, deletion studies showed that in *S. cerevisiae* deletion of NAC results in an aberrant interaction of SRP with non-substrates (del Alamo et al., 2011). This interaction might be a concentration driven interaction with non-substrates. However, this deregulation must be repressed downstream of the SRP-NC interaction, as no obvious translocation defect was detected. As in *S. cerevisiae*, the majority of protein translocation is posttranslational and SRP independent, a comparison with the situation in higher eukaryotes might lead to new insights. Indeed, the deletion of NAC in higher eukaryotes leads to embryonical lethality (Bloss et al., 2003; Deng and Behringer, 1995; Markesich et al., 2000). While no data is available on why this happens, it is tempting to speculate, that deregulation of secretion leads to this phenotype.

Interestingly, not only were aberrant interactions between SRP and non substrates observed, but also a loss of interaction with a specific subset of SRP-substrates. A simple model, where SRP is sequestered by the ribosome in the absence of NAC (similar to the proposed RAC-NAC-ribosome

model) leads to exactly this phenotype, if the lost interactions between SRP and the RNCs are substrates with weak signal sequences. Due to low concentration of SRP, the steady state levels of SRP interacting with low affinity substrates would be reduced. However, the global study conducted by the Frydman lab provided data, that the deregulation of SRP interaction is not that simple (del Alamo et al., 2011). Some nascent polypeptides with which SRP interacts newly in the absence of NAC indeed show a low hydrophobicity in their signal sequence. Therefore, the aberrant interaction might be a concentration driven event. However, the lost substrates do not show consistently weak SRP interaction motifs as expected. Therefore, the regulation of the specificity of SRP via NAC is not only a simple concentration dependent regulation of free SRP, but a direct effect of the interaction between NAC-SRP the nascent polypeptide and ribosomes as also suggested by others (del Alamo et al., 2011; Zhang et al., 2012).

To fully understand the system in a physiological system (in actively translating ribosomes), a first approximation of this setup would be the usage of randomly translating ribosomes as the binding partner of NAC and RAC. Cycloheximide stalled polysomes where the connecting mRNAs are digested by RNAses would replace in this set of experiments the specifically stalled RNCs. While this would not allow for the identification of primary interaction partners, it would give insights into the steady state levels of the global interactome. Additionally, we could compare our findings again with quantitative Western Blotting studies and subsequently extend this analysis in the absence of working protocols for RNC preparations to other interactors.

5.2 Trigger Factor

The second part of this thesis dealt with a different layer of organization: structural aspects of the interaction of ribosomal binders on the ribosomes. A new study directly showed, that TF works as a foldase and a holdase, but needs the interaction with the ribosome for this activity (Hoffmann et al., 2012). The way in which binding to the ribosome supports this function is not clear. It could be that binding to the ribosome puts TF simply in the vicinity of the substrate. Alternatively, more complicated mechanisms that lead to structural rearrangements might activate TF and allow proper binding of the substrate.

While a number of X-ray high resolution studies provided insights into TF unbound state, TF's arrangement on the ribosome is poorly understood (Ferbitz et al., 2004; Baram et al., 2005; Schlunzen et al., 2005). A cryo-EM structure confirmed previous studies that suggested L23/L29 as the binding site, but due to the low resolution, it did not provide new insights into the interaction between TF and the nascent polypeptide (Merz et al., 2008). Additionally, TF was crosslinked to a partially folded nascent polypeptide that is not a physiological substrate. X-ray structures of ribosomes soaked with the TF RBD were either difficult to interpret due to the heterologous origins of the components or the TF RBD was hardly resolved (Baram et al., 2005; Ferbitz et al., 2004; Schlunzen et al., 2005). Additionally, none contained nascent polypeptides. However, it was evident, that the TF-RBD was rearranged upon binding. The two homologous *D. radiodurans* co-crystal structures showed a rearrangement of the α-helices of the TF-RBD that potentially opened a hydrophobic surface that would allow for binding to the nascent polypeptide (Baram et al., 2005; Schlunzen et al., 2005) (Figure 65). However, the conformation differed. In the first structure (pdb: 2D3O), an α-helix that covers β-sheets of the TF-RBD is located closer to the ribosome in comparison to an *E .coli* (1w26) structure in solution (orange). In the second structure, the same α-helix together with the β-sheet, is bent away from the ribosome due to interactions with L29 (pdb: 2AAR). Both rearrangements lead to the opening of a hydrophobic surface; the underlying mechanism is however different.

To further examine the ribosome bound TF-RBD, we aimed at (i) providing a high resolution cryo-EM snapshot of the interaction between TF, an endogenous nascent polypeptide as a substrate and the ribosome and at (ii) understanding the flexibility of TF in a quantitative manner.
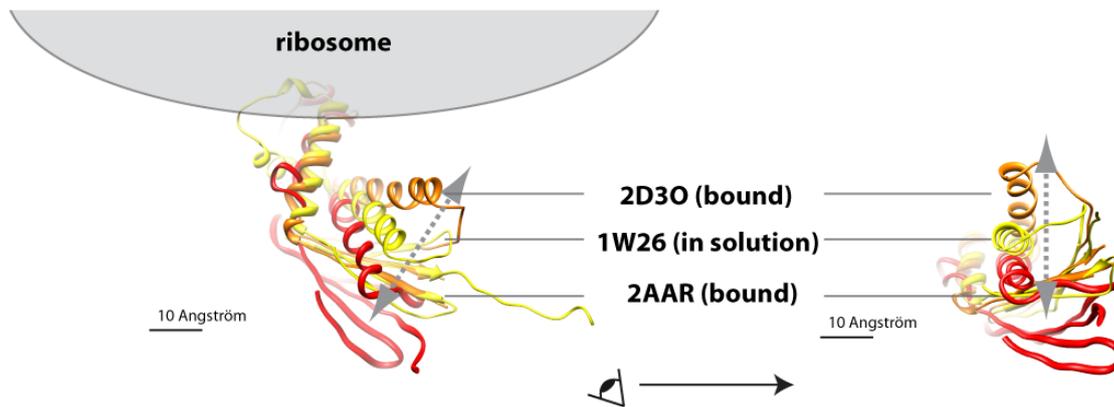
Figure 65: The TF-RBDs of different studies showed different conformations on the ribosome |In the first structure, an α-helix that covers the β-sheets of the TF-RBD is moving towards the ribosomes (orange, pdb:2D3O) (Schlunzen et al., 2005). The β-sheets stay in place. In the second structure, the α-helix together with the β-sheets are moved away from the ribosome (red, pdb:2AAR) (Baram et al., 2005). For comparison, the *E. coli* in solution structure is shown in yellow, occupying an intermediate conformation. (1W26) (Ferbitz et al., 2004). These spatial rearrangements are in a scale that can potentially be resolved by cryo-EM.

5.2.1 Towards high yielding RNC preparations

To observe physiological interactions of TF with its substrate, we chose to place a strong TF interaction sequence near the tunnel exit. The interaction time of TF and nascent polypeptides correlates with the presence of long hydrophobic stretches (Kaiser et al., 2006). We therefore positioned a 10 AA long hydrophobic stretch of GatD (AA 81-91 of GatD) 35 AA away from the PTC. Since the ribosomal tunnel can accommodate up to 35 extended AA, we reasoned that this would locate the potential interaction motif directly at the tunnel exit. Thereby, we aimed at stabilizing TF by natural means at the tunnel exit. The second, longer hydrophobic stretch of GatD was not used since it has a high propensity for α-helix formation (Figure 35, p. 86). As TF presumably interacts with unfolded nascent polypeptides we reasoned that this would be a suboptimal TF substrate. This first hydrophobic stretch was not used in its isolated form but in context of its natural surrounding primary sequence.

Typically, mRNAs lacking a stop codon are used to program ribosomes for structural and functional studies in eukaryotes and prokaryotes (Halic et al., 2004; Halic et al., 2006b; Berndt et al., 2009; Frauenfeld et al., 2011). In initial experiments using a GatD sequence however we saw that cytosolic nascent polypeptides in *in vitro* translation extracts are not as efficiently stalled as signal sequence containing nascent polypeptides such as FtsQ.

Over the last years more and more stalling sequences derived from different organisms were used to stall ribosomes. While in eukaryotic systems the human cytomegalovirus gp48 upstream open reading frame 2 and the arginine attenuator peptide AAP were structurally investigated, in

prokaryotes the TnaC, SecM and ErmC stalling sequences were used (Bhushan et al., 2011; Bhushan et al., 2010b; Seidelt et al., 2009; Yao et al., 2008). The TnaC stalling sequence was shown to be more longlasting than the SecM stalling sequence. Hence, we opted for it and constructed nascent polypeptides that were stabilized by this 23 AA long peptide (Figure 36, p. 86). The C-terminal 23 AAs of the GatD I-35 construct were replaced by the TnaC stalling sequence.

Based on a previous publication, we generated *in vitro* translation extracts and tried different conditions to purify a TnaC stalled GatD substrate (Seidelt et al., 2009). We started with a signal anchor containing positive control and obtained high yields. Starting from 10 ml of a self-made *in vitro* translation reaction, ~ 300 pmol of RNCs were obtained. The quality of the RNCs was reflected in the virtual absence of any disomal or trisomal contaminations indicating a population of ribosomes who contain a homogenous nascent polypeptide (Figure 37, p. 88). Polysomes appear, if the translation of a single stalling mRNAs is initiated more than once. Only the first initation event results in a proper stalling. Secondary and tertiary initiating ribosomes are stopped by the first, stalled ribosome and contain intermediate nascent polypeptides. Additionally, only minute amounts of free peptides were present in the final preparation, indicating a stable stalling of the nascent polypeptides (Figure 38, p. 88). We then purified *in vivo* TnaC stalled RNCs carrying a GatD derived nascent polypeptide.The yield of the RNC preparation (120 pmol per liter of the expression medium) not only allowed the use of these RNCs for structural analysis but also for subsequent MST measurements.

5.2.2 High resolution structural analysis of a TF-RNC complex

The complex formation between TF and ribosomes is quite robust. Since TF binds stoichiometrically in co-pelleting assays already to empty ribosomes, the additional contribution of the presence of a nascent polypeptide is hard to detect using this technique (Figure 42, p. 92). Using microscale thermophoresis, the additional contribution of the nascent polypeptide to TF's recruitment was shown (Figure 56, p. 109). Next, we formed a complex between TF and GatD RNCs (20AA GatD I-35 TnaC) and subjected this sample to cryo-electron microscopy and single particle reconstruction.

Starting with 480,000 particles, we were able to subsort the heterogenous dataset into homogenous subdatasets. While normally substoichiometric programming of the RNCs makes it a necessity to first sort for idle and programmed ribosomes, the quality of the RNC preparation was reflected in the absence of any unprogrammed ribosome in the dataset. 18% of the dataset aligned to a random distribution of density. These particles were either non ribosomal particles that escaped the preselection of suitable particles or misaligned ribosomal particles. 82% of the dataset were ribosomes that could be further subsorted according to the E-site tRNA occupancy. The bigger

subpopulation (presence of P-site absence of E-site) was then used for further subsorting of the dataset into different TF states (Figure 43, p. 94).

While the competitive multi reference alignment approach yielded good results in sorting for strong features like presence and absence of the E-site tRNA (accompanied by differences in the L1-stalk), the sorting for TF conformations was more difficult. This is due to two factors: density and conformational heterogeneity. In cryo-EM maps, RNA produces (due to the electron denser phosphate backbone) in comparison to proteinaceous parts a higher signal. Additionally, the RNA of the core structure is more rigid. These features lead to a strong signature of RNA defined substates and hence, the stable sorting according to presence/absence of tRNA.

TF on the other hand is a protein of intermediate size. Additionally, it shows huge degrees of flexibilities since it contacts the ribosome only via a single contact point. Even in the absence of any conformational flexibility of the protein itself, the decentralized anchor of the protein on the ribosome leads to leveraged movements of distant ends of the molecule.

First, we tried 2D-classification based techniques and 3D bootstrapping to identify regions of heterogeneity and to backproject subpopulations. However, both approaches failed due to the weak signal of TF in comparison to the ribosome (section 9.4 p. 155 and section 9.5 p. 157). Then, we managed to sort for TF-defined subpopulations. We succeeded doing this by (i) reducing the heterogeneity of the ribosome that is independent to TF and (ii) by applying new sorting algorithms. The first factor was achieved by biochemical (homogenous, highly programmed RNC population) and bioinformatical means (classical, stringent sorting according to movements of integrated but flexible parts of the ribosome).

The new sorting scheme (FCMA) employed, other than the classical global approach, CC C's of defined subregions. These were defined by binary 3D masks projected into 2D. Thereby, the minute signal arising from the presence/absence of TF was weighted higher as it would have been in a global approach. The reliability of this approach was shown in a parallel project where this scheme was also used (personal communication Christoph Leidig, AG Beckmann, Gene center, LMU Munich). A binary mask that was smaller than the ligand yielded density of the full length ligand even beyond the boundaries of the mask. This showed that the use of limited information for sorting yields meaningful data.

5.2.3 Flexibility and cryo-EM

Extra density for TF was visible in all data processing steps. However, the strength of additional density decreased with increasing resolution, a feature of flexible proteins. That this is indeed a

reflection of TF's properties is obvious, as we also saw this effect when the same volume was filtered at different resolutions (Figure 47, p. 98). To support our finding with a simplified simulation, a normal mode analysis of TF alone in solution was carried out (Figure 48, p. 99). We saw multiple resonance frequencies that resulted in large scale movements of individual domains. Notably, the head domain showed a relative movement to the mid-domain. Additionally, the mid domain showed movements relative to the ribosome binding domain.

The overlay of the first seven non trivial resonance frequencies (the first six being rotation along three axis and translation in x,y,z) with the high resolution TF map showed, that the movement of the MID domain indeed happens within the density distribution of extra TF density in the low resolution map. The movement of the head domain is too large to result in any detectable density beyond background noise. Figure 66 shows the direct comparison of the movement of TF and the density of the TF map.
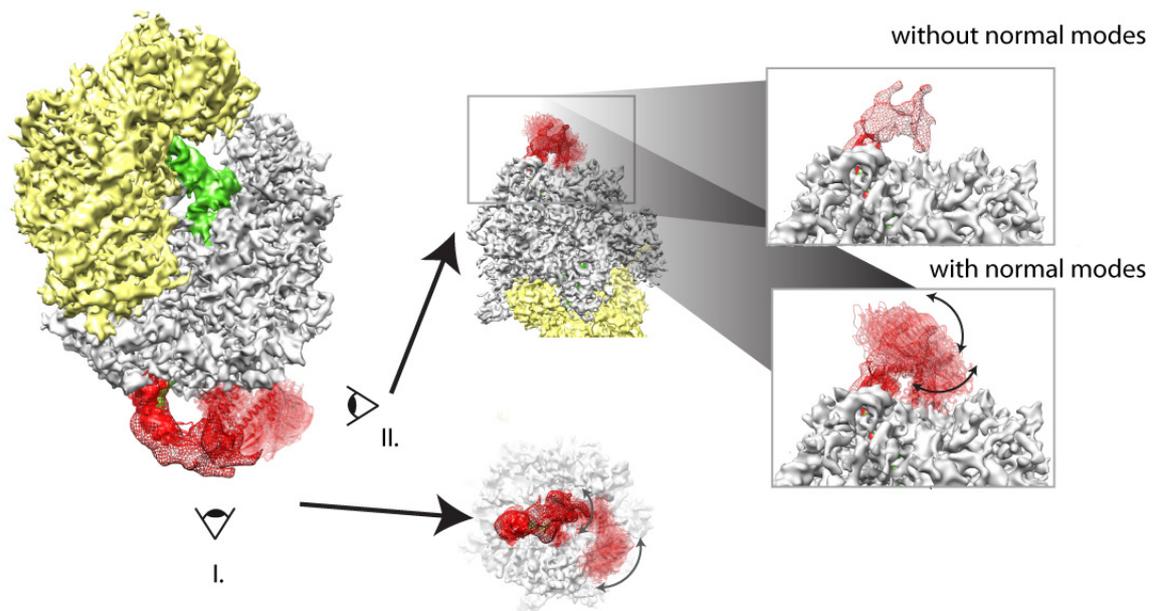
Figure 66: The first seven normal modes overlaid with the TF-cryoEM map show the high flexibility of TF on the ribosome | The 7.7 Å resolution map (solid) and the TF density of the low resolution map (meshed red density) were overlaid. The RBDs of the first seven normal modes of TF (pdb 1w26a) were aligned to the fitted TF RBD. The delocalized density is illustrated by a low opacity of the models. Thereby, states that are highly populated by TF are more densely illustrated in comparison to low populated states. The major movement trajectories are indicated by the arrows.

Flexibility of proteins that translate into a density distribution as observed in the TF map can be due to several qualitatively different conformational spaces. In one scenario, two or more energetically stable states are populated by the majority of the individual molecules. A small subfraction however is recorded during the transition between those states. Alternatively a multitude of energetically similar states exits along the trajectory of the movement that are similarly populated.

The first scenario is common in large scale rearrangements that are energetically optimized to adopted states that differ in functional activity. In context of ribosomes, the highly organized ratchet movement of the SSU relative to the LSU (however with more than two substates) or the movement of the L1-stalk are examples for this organizational scheme (Frank and Agrawal, 2000; Gao et al., 2003). While these movements fall into the category of flexibility (the proteins are inherent flexible or externally induced), the static nature of the rearrangement makes the substates similar to compositional heterogeneity. In dependence of the complexity of the trajectory, classical and focused cross-correlation coefficient based competitive multireference alignment techniques (for low complexity) or maximum likelihood based techniques are sufficient to subsort into the substates.

On the other hand, molecules that contain flexibilities of the second scenario do not oscillate between distinct states. Extreme cases are domains that are intrinsically disordered. While more and more studies suggest, that some domains of this class obtain a fold upon activation, the general significance is not understood. Their flexibility is beyond a scale that allows structural investigation using the averaging of signals of multiple individual proteins like cryo-EM with single particle reconstruction and diffraction techniques. More interesting in this context are proteins, that are folded, however display huge degrees of flexibility in between or within domains without defined substates.

Figure 67 visualizes the problem in a simplified context. A model protein whose two domains are connected via a hinge, can either follow scenario 1 (Panel A) or scenario 2 (Panel B). In the first scenario, particles that fall in the different subdataset are similar to each other. That is, they are close together at the distance axis in Figure 67. The particle distribution further stabilizes the two subgroups. The challenge in such a scenario is to define the number of classes. If fewer classes than actual subgroups are assigned, two similar subgroups are fused into one; if too many are assigned groups are artificially split. The overlay of different classes leads to a cloud like density distribution. Artificially split classes on the other hand are indistinguishable but split the dataset into two parts. Therefore, sorting according to low intensity features like random accumulation of noise can lead to the introduction of artifacts especially when sorting according to high frequency information. In the second scenario, neither the particle distribution nor the similarity of particles is stabilizing subgroups. Depending on the signal intensity of the moving part, the dataset is artificially split either into subgroups of similar particles or according to random noise.
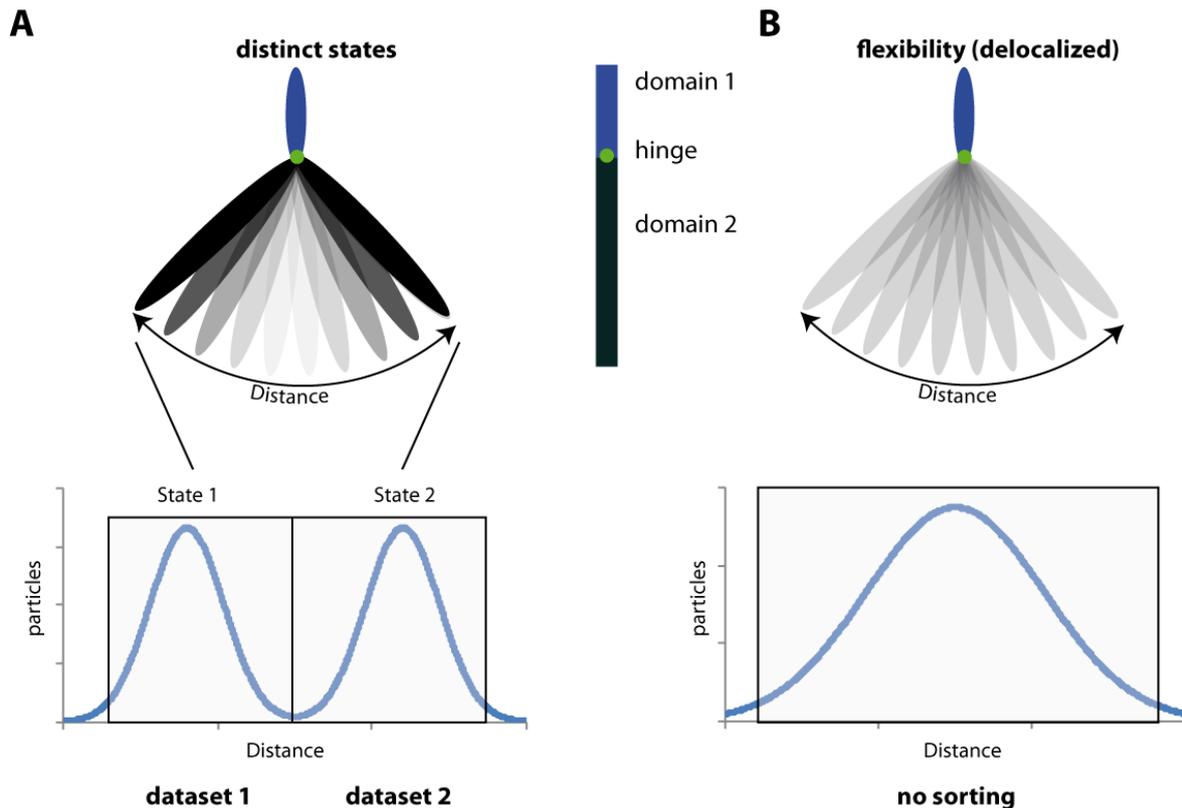
Figure 67: Flexible proteins can be classified into different-state or in continuously-moving proteins | A simulated two-domain protein can either oscillate between stable states with a subpopulation being in the transition (Panel A) or adopt a wide range of conformations (Panel B). In Panel A, the particle distribution along a function of distance between 1D conformation is plotted. The energetically stable states are reflected by accumulations of particles. Multi-reference competitive alignment can lead to the definition of two subdatasets that can independently refined. In Panel B, the simplified movement along the same axis of variability leads to an even distribution. Under the assumption of an interconversion between the states this leads to a gaussian distribution of particles. Subsorting might lead to artificially grouping similar particles. However, due to the weak signal and the weak signal intensity of the averaged particles, the influence of noise is mostly dominant.

The failure of the competitive multi-reference alignment approach to subsort TF in combination with the normal mode analysis suggested, that TF is inherently flexible without energetically favored states (Scenario 2, Panel B, Figure 67). The almost complete absence of the Head-domain and the low information of the MID-domain in the reconstruction indicated a vectorial increase of the flexibility as a function of the distance. However, this feature alone is not sufficient to explain the strong flexibility of TF in the cryo-EM maps. An additional hinge in between the Head and the MID-domain must be present to increase the flexibility (see also Figure 48, p. 99).

To gain further insights into the space that the Head-domain is sampling and its location, we investigated the cryo-EM map further. Therefore we used the information provided by the local resolution map. This scheme was developed to support the structural interpretation of the map. However, it turned out that flexible parts of proteins that cannot be visualized by cryo-EM, leave their trace in this representation. Figure 68 shows, that the delocalized HEAD-domain reduces the

local resolution at the tip of the mid domain. While we cannot visualize the Head-domain, this region of low resolution is a *bona fide* trace left by this domain. While the rest of the map shows a linear decrease of the resolution as a function of distance from the ribosome (see also Figure 53, p. 104), this region shows a robust deviation from that trend.
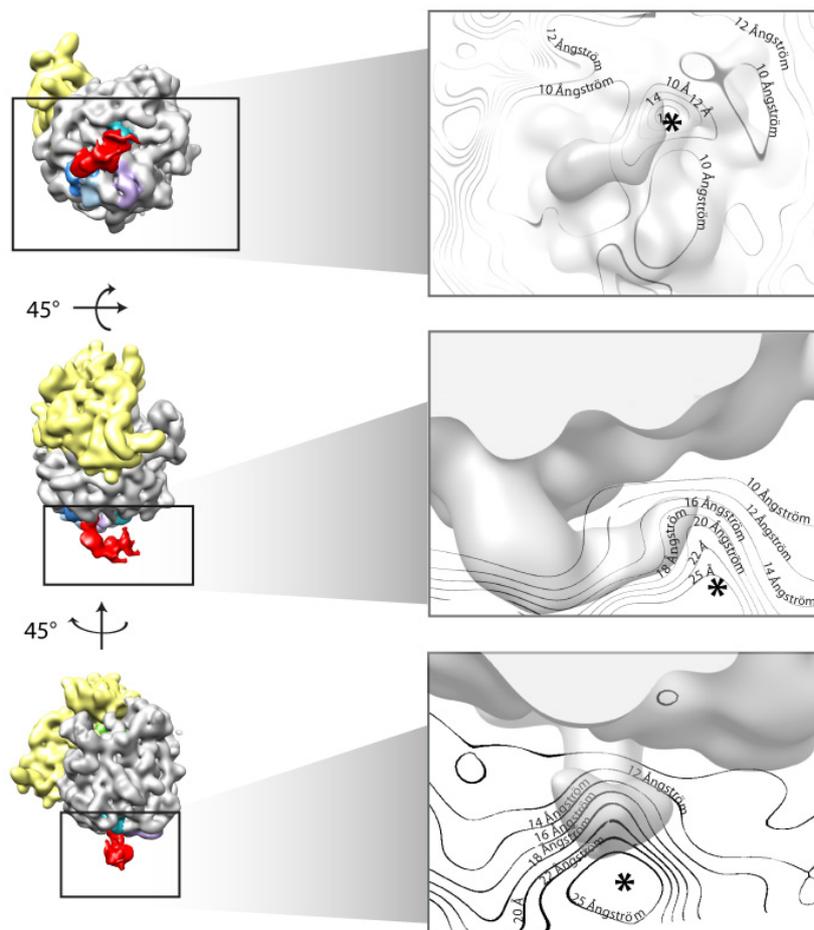
Figure 68: TF's head domain pertubs the resolution at the tip of the MID-domain | The local resolution map of high resolution TF-RNC structure showed a decrease in the resolution at the tip of the MID domain where the Head-domain is delocalized. This island of low resolution is the centre of gravity of the movement of the Head-domain (marked with an asterisk).

We speculate, that this feature is due to two factors: (i) higher density than the surrounding and (ii) huge variation in the conformation. The higher density (to be more precise the higher probability of the Head-domain being in this region) gives a relative high signal in the two sub-boxes of the local resolution determination scheme (Figure 53, p. 104). This, together with the conformational plurality gives a strongly deviating Fourier transformation in this region and subsequently a low resolution. While this technique does not allow for assignments of and medium to high resolution features, it provides additional information for the general arrangement of the domains.

With this information in hand, a starting model for a molecular dynamic simulation can be built and simulated in a ribosome bound form in the presence of nascent polypeptides. The average density distribution along the trajectories of the movement of TF can then be calculated and compared with the actual density of the cryo-EM map. Thereby, we hope to achieve a quantitative description of this flexibility.

This however will only explain, how flexible TF is, not why it is so flexible. The simplest answer would be that the flexibility and TF's function is not interconnected. Bacterial genomes are often tailored to maximize the growth rate. This, together with the primarily asexual propagation, enormous numbers of individuals and high selective pressure in the practical absence of genetic drift leads to a maximal reduction of biological macromolecules and pathways. TF's flexibility might just be a sign for the absence of any advantage of a rigid TF. TF interacts with the nascent polypeptide along a stretch of > 100 Å and to stiffen a protein of this size, higher structural cost due to the need of more structural elements are probable.

The alphaproteobacterium *Candidatus pelagibacter ubique* is one of the smallest, free living organism that is known (Rappe et al., 2002). Due to its abundance (up to 25% of the microbial plankton cells), it is an ideal candidate to investigate this hypothesis. Indeed, its genome is streamlined and has the smallest genome of autonomous cells known so far (Giovannoni et al., 2005). Additionally, it has the fewest number of genes and the shortest intergenic regions known to date. However, most of the metabolic pathways including the complete biosynthetic pathways for all 20 AA are present. Indeed it encodes for a TF homologue (gene ID:3517003 tig, 28% identity, E-value 1e-51). The *C. p. ubique* TF is even larger than the *E. coli* TF supporting the assumption that it is indeed a maximally optimized protein (*C. p. ubique*: 477 AA, *E. coli 432* AA).

This does not prove, that no functional relationship between the function and the flexibility exists. Quite on the contrary, its flexibility might be a prerequisite of its function and is therefore maintained even under the strongest selection. Further support for this hypothesis came from a study that showed that TF does not only act as a holdase but also as an unfoldase on the ribosome (Hoffmann et al., 2012). A prefolded protein can be unfolded by TF as long as it is still in contact with the ribosome. A flexible TF has more degrees of freedom to sample the surface of folded proteins while still being attached to the ribosome. Thereby, it could find the weakest point of the folded protein and initiates unfolding. This function however cannot be sufficiently important as the DnaK/J-GrpE Hsp70 system does not possess such an activity but still can complement TF's deletion. We therefore think, that the flexibility is mainly maintained since it is the most energy-saving architecture with which a long interaction between the nascent polypeptide and TF can be obtained.

5.2.4 The role of the TF-RBD in the interaction between TF and the NC

The flexibility of the TF MID- and Head-domain did not allow for a high resolution analysis of these parts. The TF-RBD on the other hand was well resolved. Since the ultimate goal of this project was to understand the interaction between TF and nascent polypeptides, molecular models for both parts had to be built. The presence of additional density within the ribosomal tunnel indicated a solid programming of the ribosomes. A comparison with published TnaC stalled nascent polypeptide maps confirmed the specific stalling via this peptide (Figure 49, p. 100, (Seidelt et al., 2009)). The low resolution in this region however did not allow for an AA specific assignment of the nascent polypeptide. We therefore used a feature of the nascent polypeptide as a molecular ruler: the nascent polypeptide coded N-terminal of the TnaC stalling sequence and C-terminal of the potential interaction site for an α-helix. The ribosomal tunnel allows limited formation of α-helices downstream of the constriction site (Bhushan et al., 2010a; Lu and Deutsch, 2005). We therefore hoped to see a compaction of the nascent polypeptide. This compaction would separate the TnaC-stalling sequence and the putative TF-binding site. Therefore, an unambiguous assignment of regions of the NC would be possible. Indeed, a compaction was visible. Interestingly, this compaction was downstream of the NC-density of the previous published TnaC-map and therefore definitely not part of the TnaC stalling sequence. We therefore could locate the hydrophobic stretch and finalize a model for the NC.

The model for the TF-RBD was built based on published structures, known interactions at the ribosome and information provided by the density map (Figure 51, p. 102). In the vicinity of the ribosome, two α-helices were well resolved. Further away from the ribosome, they fused with the extra density of the nascent polypeptide. At the distant end of the TF-RBD relative to the ribosome, the α-helices were rearranged similar to pdb 2AAR. However, directly at the ribosome the α-helix is located further away from L29 as compared to the pdb:2AAR (Baram et al., 2005).

Having the model for the nascent polypeptide and the TF-RBD ready, we could test if previously published crosslinks between TF and the nascent polypeptide would be possible in this conformation (Merz et al., 2008). The molecular model of the nascent polypeptide contacts the TF-RBD in close vicinity of AA 61 which is the first known contact point between TF and nascent polypeptides (Figure 69, Crosslink 1). Interestingly, the hydrophobic stretch of GatD is placed in between the two rearranged α-helices of the TF-RBD.

Moreover, the rearrangement of the TF-RBD opens exactly at the place where the NC contacts TF a hydrophobic interaction site that is perfectly located to accommodate nascent polypeptides upon their emergence from the ribosome. In agreement with previous structures we therefore suggest

that TF is primed upon binding to ribosomes for the shielding of hydrophobic stretches of the nascent polypeptide. This interaction with the ribosome is necessary for TF to obtain the conformation since the in solution structure does not exhibit this interaction site.
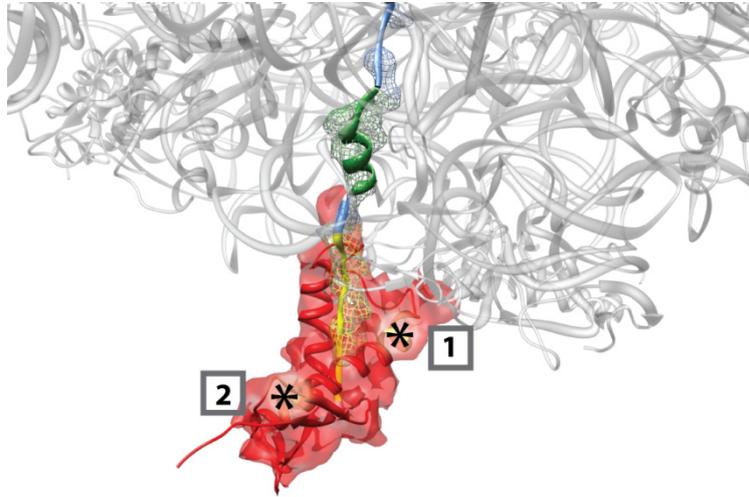
Figure 69: Reported crosslinks between the TF-RBD and the nascent polypeptide | Upon emergence from the ribosome, the nascent polypeptide contacts first AA 61 of TF (first cross link) and then AA76 (second crosslink). The molecular models of the nascent polypeptide and of the TF-RBD allow this interaction.(Merz et al., 2008)

It was suggested earlier that TF interaction with L29 is the cause for this rearrangement (Baram et al., 2005). Since we cannot visualize a time course of the event but steady state levels, we only can compare the endpoints of this binding event: TF in solution versus the bound state.

In the unbound state, the TF-RBD contains three helices that are connected by flexible linkers. In the ribosome bound form, we saw only two helices. Arg57 interacts with L29 and thereby bends helix 2 by 90° in line with helix three. This rearrangement results in (i) the unfolding of AA 51 to 56 (ii) the fusion of helices 2 and 3 to helix 2/3 and (iii) a repositioning of the fusion helix 2/3 relative to helix 1. This movement leads to the exposure of hydrophobic AA that in the in solution form stabilize the interaction between helix 1 and 2/3 (yellow region between helix 1 and 2/3, Figure 70, Panel B).

The local resolution information allows us to judge, which of these interpretations is supported by our data and which is beyond the information content of the cryo-EM map. The local resolution of the TF-RBD was beyond 7.7 Å (Figure 54, p. 106). The repositioning of helix two and its fusion with helix three is a secondary structure rearrangement that indeed can be visualized at such resolutions (the fusion helix is 10 Å longer as helix 3). The repositioning of the fusion helix relative to helix one however is a small scale event that is at the limit of our resolution. Due to spatial constraints namely the known anchoring sites and the stable β-sheets that are C-terminal of helix 3, it is however likely that this repositioning is happening upon binding to the ribosome.
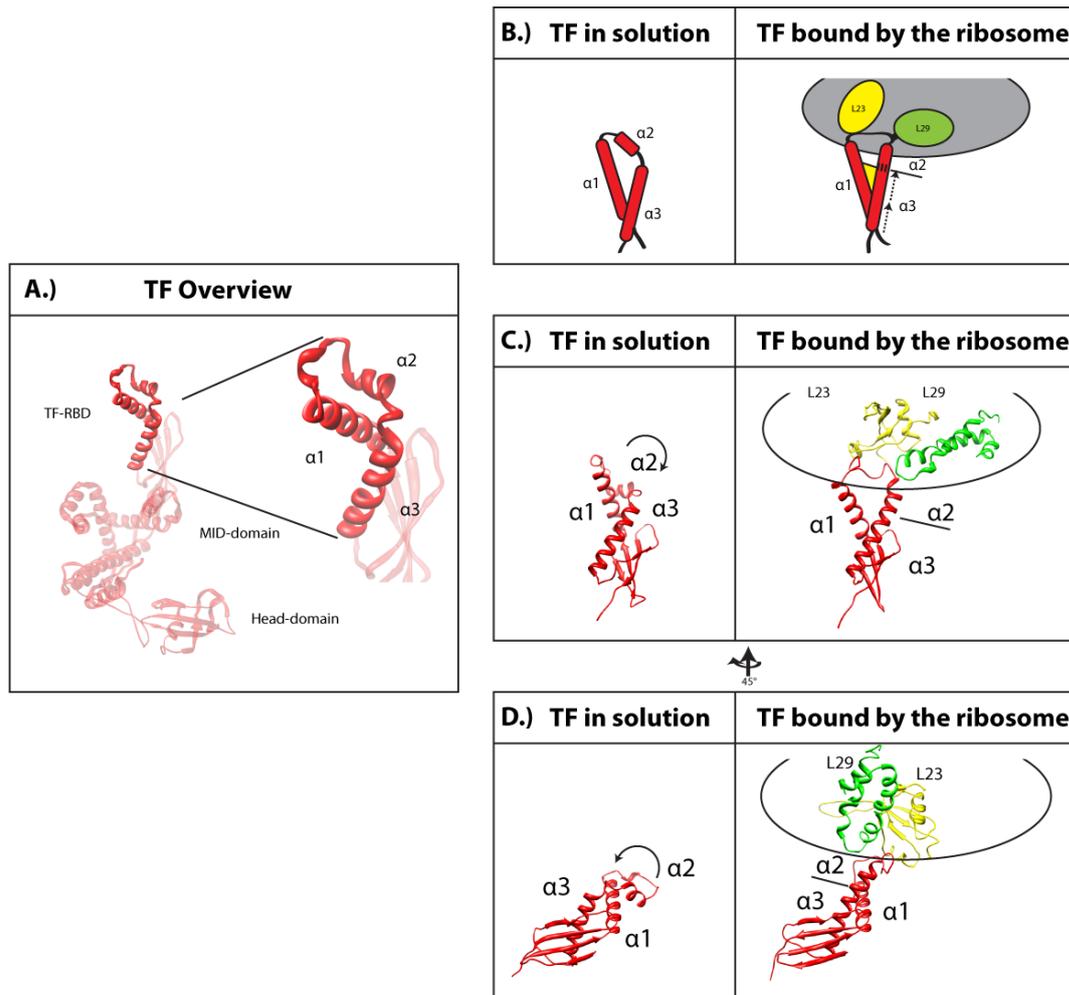
Figure 70: TF's RBD is rearranged upon binding to the ribosome; this rearrangement is most likely due to interaction between L29 and TF | In the unbound form, the TF-RBD adopts a fold of three α-helices (1-3, Panel A). However, upon binding helices 2 and 3 fuse together (Panels B-D). Parts of the second helix are unfolded upon ribosome binding. Additionally, the fused helix is repositioned relative to helix one. Thereby, a hydrophobic patch is opened that is a potential interaction site for hydrophobic substrates (yellow region in Panel B).

The cryo-EM map presented here shows the first time how the nascent polypeptide that just left the ribosomal tunnel enages with the TF-RBD. A rearrangement of the TF-RBD opens an interaction site the shieldw hydrophobic stretches of the nascent polypeptide. The side of the Mid-domain that faces the ribosomes also contains a hydrophobic surfaces that – together with the TF-RBD – creates a large interaction site that allows productive folding of newly translated proteins.

6 Summary

The regulation of the access to substrates by competing interactors is governed by a multitude of physical and chemical properties that are inherently encoded by the structure of the protein. *In vivo*, high and low affinity interactions are balanced by additional factors like abundance of the proteins, dynamic regulations of the interactors via biochemical means, sequence features of the nascent polypeptide and competition by progression into alternative pathways to optimize the functionality of the cell. Thereby, evolution built robust and highly efficient functional networks. The access of proteins that use the ribosome for access to the nascent polypeptide is one example of such a highly integrated network. So far two universal adaptor sites were identified. Via those sites, a multitude of proteins and protein-complexes use the ribosome as a platform to interact with the nascent protein. The dynamic properties of translation create a rapidly changing environment. Here, the access of proteins that *e.g.* have to interact with the emerging N-terminus like MAP/NAT or SRP has to be extremely efficiently regulated to guarantee optimal protein folding yields and targeting.

We successfully applied microscale thermophoresis and other techniques to determine *in vitro* the binding properties of two ribosomal interactors to the ribosome. Furthermore, we detected and described conformational changes of one of this interactors – NAC – upon ribosome binding. Finally, we extended this analysis to describe the competition properties of this subsystem of the ribosome interactome. In a next step, we integrated the acquired knowledge into a physiological context and simulated the *in vivo* situation and predicted the respective complex formation in wild-type and mutant situations. We showed that in a highly simplified model – no translation and two interactors that completely compete with each other against one binding site – the data is in agreement with *in vivo* data.

In the second part we showed, that cryo-EM and single particle reconstruction can provide insights into highly flexible proteins like TF. We applied different techniques: ranging from 2D clustering, to 3D bootstrapping and local cross correlation coefficient based sorting algorithms. Finally, we could visualize the TF-RBD at a resolution that allowed a molecular interpretation of the interaction between TF and the NC. We could visualize the NC contacting the TF-RBD via a hydrophobic interaction site. However, the majority of the additional density was delocalized due to the inherent flexibility of TF. While the Mid-domain was visible at low resolution, the Head-domain was too flexible to detect. Using the information provided by a local resolution map, we were able to detect the center of gravity of this highly flexible region of TF and thereby provide a starting model for subsequent molecular dynamic simulations.

7 References

Adams, J.M., and Capecchi, M.R. (1966). N-formylmethionyl-sRNA as the initiator of protein synthesis. Proceedings of the National Academy of Sciences of the United States of America *55*, 147-155.

Albanese, V., Reissmann, S., and Frydman, J. (2010). A ribosome-anchored chaperone network that facilitates eukaryotic ribosome biogenesis. The Journal of cell biology *189*, 69-81.

Angelini, S., Deitermann, S., and Koch, H.G. (2005). FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. EMBO Rep *6*, 476-481.

Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., et al. (2010). Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome. Proceedings of the National Academy of Sciences of the United States of America *107*, 19754-19759.

Baaske, P., Weinert, F.M., Duhr, S., Lemke, K.H., Russell, M.J., and Braun, D. (2007). Extreme accumulation of nucleotides in simulated hydrothermal pore systems. Proceedings of the National Academy of Sciences of the United States of America *104*, 9346-9351.

Ball, L.A., and Kaesberg, P. (1973). Cleavage of the N-terminal formylmethionine residue from a bacteriophage coat protein in vitro. Journal of molecular biology *79*, 531-537.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science (New York, NY) *289*, 905-920.

Baram, D., Pyetan, E., Sittner, A., Auerbach-Nevo, T., Bashan, A., and Yonath, A. (2005). Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. Proceedings of the National Academy of Sciences of the United States of America *102*, 12017-12022.

Batey, R.T., Rambo, R.P., and Doudna, J.A. (1999). Tertiary Motifs in RNA Structure and Folding. Angewandte Chemie (International ed in English) *38*, 2326-2343.

Beatrix, B., Sakai, H., and Wiedmann, M. (2000). The alpha and beta subunit of the nascent polypeptide-associated complex have distinct functions. The Journal of biological chemistry *275*, 37838-37845.

Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997). Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. Science (New York, NY) *278*, 2123-2126.

Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell *107*, 361-372.

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 A resolution. Science (New York, NY) *334*, 1524-1529.

Berndt, U., Oellerer, S., Zhang, Y., Johnson, A.E., and Rospert, S. (2009). A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. Proceedings of the National Academy of Sciences of the United States of America *106*, 1398-1403.

Bhushan, S., Gartmann, M., Halic, M., Armache, J.P., Jarasch, A., Mielke, T., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2010a). alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. Nature structural & molecular biology *17*, 313-317.

Bhushan, S., Hoffmann, T., Seidelt, B., Frauenfeld, J., Mielke, T., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2011). SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. PLoS Biol *9*, e1000581.

Bhushan, S., Meyer, H., Starosta, A.L., Becker, T., Mielke, T., Berninghausen, O., Sattler, M., Wilson, D.N., and Beckmann, R. (2010b). Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. Molecular cell *40*, 138-146.

Bingel-Erlenmeyer, R., Kohler, R., Kramer, G., Sandikci, A., Antolic, S., Maier, T., Schaffitzel, C., Wiedmann, B., Bukau, B., and Ban, N. (2008). A peptide deformylase-ribosome complex reveals mechanism of nascent chain processing. Nature *452*, 108-111.

Biswas, S.K., Yamaguchi, M., Naoe, N., Takashima, T., and Takeo, K. (2003). Quantitative three-dimensional structural analysis of Exophiala dermatitidis yeast cells by freeze-substitution and serial ultrathin sectioning. Journal of electron microscopy *52*, 133-143.

Blau, M., Mullapudi, S., Becker, T., Dudek, J., Zimmermann, R., Penczek, P.A., and Beckmann, R. (2005). ERj1p uses a universal ribosomal adaptor site to coordinate the 80S ribosome at the membrane. Nature structural & molecular biology *12*, 1015-1016.

Blobel, G., and Sabatini, D.D. (1970). Controlled proteolysis of nascent polypeptides in rat liver cell fractions. I. Location of the polypeptides within ribosomes. The Journal of cell biology *45*, 130-145.

Bloss, T.A., Witze, E.S., and Rothman, J.H. (2003). Suppression of CED-3-independent apoptosis by mitochondrial betaNAC in Caenorhabditis elegans. Nature *424*, 1066-1071.

Bornemann, T., Jockel, J., Rodnina, M.V., and Wintermeyer, W. (2008). Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. Nature structural & molecular biology *15*, 494-499.

Brunori, M., Gianni, S., Giri, R., Morrone, A., and Travaglini-Allocatelli, C. (2012). Morphogenesis of a protein: folding pathways and the energy landscape. Biochem Soc Trans *40*, 429-432.

Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins *21*, 167-195.

Celander, D.W., and Cech, T.R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. Science (New York, NY) *251*, 401-407.

References

Chen, J.Z., and Grigorieff, N. (2007). SIGNATURE: a single-particle selection system for molecular electron microscopy. Journal of structural biology *157*, 168-173.

Clusius, K.a.D., G. (1938). Neues Verfahren zur Gasentmischung und Isotopentrennung. Naturwissenschaften *26*, 546.

Crooke, E., Brundage, L., Rice, M., and Wickner, W. (1988a). ProOmpA spontaneously folds in a membrane assembly competent state which trigger factor stabilizes. The EMBO journal *7*, 1831-1835.

Crooke, E., Guthrie, B., Lecker, S., Lill, R., and Wickner, W. (1988b). ProOmpA is stabilized for membrane translocation by either purified E. coli trigger factor or canine signal recognition particle. Cell *54*, 1003-1011.

Crooke, E., and Wickner, W. (1987). Trigger factor: a soluble protein that folds pro-OmpA into a membrane-assembly-competent form. Proceedings of the National Academy of Sciences of the United States of America *84*, 5216-5220.

Debye, P. (1939). Zur Theorie des Clusiusschen Trennverfahrens. Annalen der Physik *36*, 284-294.

del Alamo, M., Hogan, D.J., Pechmann, S., Albanese, V., Brown, P.O., and Frydman, J. (2011). Defining the specificity of cotranslationally acting chaperones by systematic analysis of mRNAs associated with ribosome-nascent chain complexes. PLoS Biol *9*, e1001100.

Deng, J.M., and Behringer, R.R. (1995). An insertional mutation in the BTF3 transcription factor gene leads to an early postimplantation lethality in mice. Transgenic research *4*, 264-269.

Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature *400*, 693-696.

Eisenstadt, J., and Lengyel, P. (1966). Formylmethionyl-tRNA dependence of amino acid incorporation in extracts of trimethoprim-treated Escherichia coli. Science (New York, NY) *154*, 524-527.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta crystallographica Section D, Biological crystallography *66*, 486-501.

Evans, M.S., Ugrinov, K.G., Frese, M.A., and Clark, P.L. (2005). Homogeneous stalled ribosome nascent chain complexes produced in vivo or in vitro. Nat Methods *2*, 757-762.

Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. Nature *431*, 590-596.

Flores, O., Lu, H., and Reinberg, D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. The Journal of biological chemistry *267*, 2786-2793.

Frank, J., and Agrawal, R.K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature *406*, 318-322.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. Journal of structural biology *116*, 190-199.

Franke, J., Reimann, B., Hartmann, E., Kohlerl, M., and Wiedmann, B. (2001). Evidence for a nuclear passage of nascent polypeptide-associated complex subunits in yeast. Journal of cell science *114*, 2641-2648.

Frauenfeld, J., Gumbart, J., Sluis, E.O., Funes, S., Gartmann, M., Beatrix, B., Mielke, T., Berninghausen, O., Becker, T., Schulten, K., et al. (2011). Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. Nature structural & molecular biology *18*, 614-621.

Funfschilling, U., and Rospert, S. (1999). Nascent polypeptide-associated complex stimulates protein import into yeast mitochondria. Molecular biology of the cell *10*, 3289-3299.

Gao, H., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S.M., Van Roey, P., Agrawal, R.K., Harvey, S.C., Sali, A., et al. (2003). Study of the structural dynamics of the E coli 70S ribosome using real-space refinement. Cell *113*, 789-801.

Gautschi, M., Mun, A., Ross, S., and Rospert, S. (2002). A functional chaperone triad on the yeast ribosome. Proceedings of the National Academy of Sciences of the United States of America *99*, 4209-4214.

Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. EMBO Rep *5*, 195-200.

George, R., Beddoe, T., Landl, K., and Lithgow, T. (1998). The yeast nascent polypeptide-associated complex initiates protein targeting to mitochondria in vivo. Proceedings of the National Academy of Sciences of the United States of America *95*, 2296-2301.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., et al. (2005). Genome streamlining in a cosmopolitan oceanic bacterium. Science (New York, NY) *309*, 1242-1245.

Grant, T.D., Snell, E.H., Luft, J.R., Quartley, E., Corretore, S., Wolfley, J.R., Snell, M.E., Hadd, A., Perona, J.J., Phizicky, E.M., et al. (2012). Structural conservation of an ancient tRNA sensor in eukaryotic glutaminyl-tRNA synthetase. Nucleic acids research *40*, 3723-3731.

Halic, M., Becker, T., Pool, M.R., Spahn, C.M., Grassucci, R.A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. Nature *427*, 808-814.

Halic, M., and Beckmann, R. (2005). The signal recognition particle and its interactions during protein targeting. Current opinion in structural biology *15*, 116-125.

Halic, M., Blau, M., Becker, T., Mielke, T., Pool, M.R., Wild, K., Sinning, I., and Beckmann, R. (2006a). Following the signal sequence from ribosomal tunnel exit to signal recognition particle. Nature *444*, 507-511.

Halic, M., Gartmann, M., Schlenker, O., Mielke, T., Pool, M.R., Sinning, I., and Beckmann, R. (2006b). Signal recognition particle receptor exposes the ribosomal translocon binding site. Science (New York, NY) *312*, 745-747.

References

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell *107*, 679-688.

Harvey, R.J. (1973). Growth and initiation of protein synthesis in Escherichia coli in the presence of trimethoprim. Journal of bacteriology *114*, 309-322.

Hesterkamp, T., and Bukau, B. (1996). Identification of the prolyl isomerase domain of Escherichia coli trigger factor. FEBS letters *385*, 67-71.

Hesterkamp, T., Deuerling, E., and Bukau, B. (1997). The amino-terminal 118 amino acids of Escherichia coli trigger factor constitute a domain that is necessary and sufficient for binding to ribosomes. The Journal of biological chemistry *272*, 21865-21871.

Hoffmann, A., Becker, A.H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012). Concerted Action of the Ribosome and the Associated Chaperone Trigger Factor Confines Nascent Polypeptide Folding. Molecular cell *48*, 1-12.

Hoffmann, A., Merz, F., Rutkowska, A., Zachmann-Brand, B., Deuerling, E., and Bukau, B. (2006). Trigger factor forms a protective shield for nascent polypeptides at the ribosome. The Journal of biological chemistry *281*, 6539-6545.

Hotokezaka, Y., van Leyen, K., Lo, E.H., Beatrix, B., Katayama, I., Jin, G., and Nakamura, T. (2009). alphaNAC depletion as an initiator of ER stress-induced apoptosis in hypoxia. Cell death and differentiation *16*, 1505-1514.

Housman, D., Gillespie, D., and Lodish, H.F. (1972). Removal of formyl-methionine residue from nascent bacteriophage f2 protein. Journal of molecular biology *65*, 163-166.

Hu, G.Z., and Ronne, H. (1994). Yeast BTF3 protein is encoded by duplicated genes and inhibits the expression of some genes in vivo. Nucleic acids research *22*, 2740-2743.

Hundley, H., Eisenman, H., Walter, W., Evans, T., Hotokezaka, Y., Wiedmann, M., and Craig, E. (2002). The in vivo function of the ribosome-associated Hsp70, Ssz1, does not require its putative peptide-binding domain. Proceedings of the National Academy of Sciences of the United States of America *99*, 4203-4208.

Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. Assay and drug development technologies *9*, 342-353.

Jorgensen, P., Nishikawa, J.L., Breitkreutz, B.J., and Tyers, M. (2002). Systematic identification of pathways that couple cell growth and division in yeast. Science (New York, NY) *297*, 395-400.

Kaiser, C.M., Chang, H.C., Agashe, V.R., Lakshmipathy, S.K., Etchells, S.A., Hayer-Hartl, M., Hartl, F.U., and Barral, J.M. (2006). Real-time observation of trigger factor function on translating ribosomes. Nature *444*, 455-460.

Kanno, M., Chalut, C., and Egly, J.M. (1992). Genomic structure of the putative BTF3 transcription factor. Gene *117*, 219-228.

Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., and Leberman, R. (1996). The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution. Nature *379*, 511-518.

Kim, S.Y., and Craig, E.A. (2005). Broad sensitivity of Saccharomyces cerevisiae lacking ribosome-associated chaperone ssb or zuo1 to cations, including aminoglycosides. Eukaryot Cell *4*, 82-89.

Kohler, R., Boehringer, D., Greber, B., Bingel-Erlenmeyer, R., Collinson, I., Schaffitzel, C., and Ban, N. (2009). YidC and Oxa1 form dimeric insertion pores on the translating ribosome. Molecular cell *34*, 344-353.

Koplin, A., Preissler, S., Ilina, Y., Koch, M., Scior, A., Erhardt, M., and Deuerling, E. (2010). A dual function for chaperones SSB-RAC and the NAC nascent polypeptide-associated complex on ribosomes. The Journal of cell biology *189*, 57-68.

Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002). L23 protein functions as a chaperone docking site on the ribosome. Nature *419*, 171-174.

Kretschmann, E. (1971). Die Bestimmung optischer Konstanten von Metallen durch Anregung von Oberflächenplasmaschwingungen. Zeitschrift für Physik *241*, 313-324.

Kristensen, O., and Gajhede, M. (2003). Chaperone binding at the ribosomal exit tunnel. Structure *11*, 1547-1556.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lakkaraju, A.K., Mary, C., Scherrer, A., Johnson, A.E., and Strub, K. (2008). SRP keeps polypeptides translocation-competent by slowing translation to match limiting ER-targeting sites. Cell *133*, 440-451.

Laskey, R.A., Honda, B.M., Mills, A.D., and Finch, J.T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. Nature *275*, 416-420.

Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., and Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic acids research *30*, 5382-5390.

Levinthal (1969). How to Fold Graciously. Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois, 22-24.

Lill, R., Crooke, E., Guthrie, B., and Wickner, W. (1988). The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane. Cell *54*, 1013-1018.

Lin, K.F., Sun, C.S., Huang, Y.C., Chan, S.I., Koubek, J., Wu, T.H., and Huang, J.J. (2012). Cotranslational Protein Folding within the Ribosome Tunnel Influences Trigger-Factor Recruitment. Biophysical journal *102*, 2818-2827.

Londei, P. (2010). Archaeal Ribosomes. eLS John Wiley & Sons Ltd, Chichester.

Lu, J., and Deutsch, C. (2005). Folding zones inside the ribosomal exit tunnel. Nature structural & molecular biology *12*, 1123-1129.

Lu, J., Kobertz, W.R., and Deutsch, C. (2007). Mapping the electrostatic potential within the ribosomal exit tunnel. Journal of molecular biology *371*, 1378-1391.

References

Ludwig, C. (1856). Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösungen. Sitzungsber Akad Wiss Wien: Math-Naturwiss 20, 539.

Magde, D., Elson, E., and Webb, E.E. (1972). Thermodynamic Fluctuations in a Reacting System—Measurement by Fluorescence Correlation Spectroscopy. Physical review letters 29, 705-708.

Makarova, K.S., Aravind, L., Galperin, M.Y., Grishin, N.V., Tatusov, R.L., Wolf, Y.I., and Koonin, E.V. (1999). Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome research 9, 608-628.

Malkin, L.I., and Rich, A. (1967). Partial resistance of nascent polypeptide chains to proteolytic digestion due to ribosomal shielding. Journal of molecular biology 26, 329-346.

Marcker, K., and Sanger, F. (1964). N-FORMYL-METHIONYL-S-RNA. Journal of molecular biology 8, 835-840.

Markesich, D.C., Gajewski, K.M., Nazimiec, M.E., and Beckingham, K. (2000). bicaudal encodes the Drosophila beta NAC homolog, a component of the ribosomal translational machinery*. Development (Cambridge, England) 127, 559-572.

Martinez-Hackert, E., and Hendrickson, W.A. (2009). Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. Cell 138, 923-934.

Mason, N., Ciufo, L.F., and Brown, J.D. (2000). Elongation arrest is a physiologically important function of signal recognition particle. The EMBO journal 19, 4164-4174.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nature structural & molecular biology 19, 560-567.

Merz, F., Boehringer, D., Schaffitzel, C., Preissler, S., Hoffmann, A., Maier, T., Rutkowska, A., Lozza, J., Ban, N., Bukau, B., et al. (2008). Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. The EMBO journal 27, 1622-1632.

Milligan, R.A., and Unwin, P.N. (1986). Location of exit channel for nascent protein in 80S ribosome. Nature 319, 693-695.

Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. Journal of structural biology 142, 334-347.

Moller, I., Jung, M., Beatrix, B., Levy, R., Kreibich, G., Zimmermann, R., Wiedmann, M., and Lauring, B. (1998). A general mechanism for regulation of access to the translocon: competition for a membrane attachment site on ribosomes. Proceedings of the National Academy of Sciences of the United States of America 95, 13425-13430.

Moncollin, V., Miyamoto, N.G., Zheng, X.M., and Egly, J.M. (1986). Purification of a factor specific for the upstream element of the adenovirus-2 major late promoter. The EMBO journal 5, 2577-2584.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., et al. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell 147, 1295-1308.

Olofsson, S.O., Bjursell, G., Bostrom, K., Carlsson, P., Elovson, J., Protter, A.A., Reuben, M.A., and Bondjers, G. (1987). Apolipoprotein B: structure, biosynthesis and role in the lipoprotein assembly process. Atherosclerosis 68, 1-17.

Otto, A. (1968). Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. Zeitschrift für Physik 216, 394-410.

Panchuk-Voloshina, N., Haugland, R.P., Bishop-Stewart, J., Bhalgat, M.K., Millard, P.J., Mao, F., and Leung, W.Y. (1999). Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. J Histochem Cytochem 47, 1179-1188.

Parthun, M.R., Mangus, D.A., and Jaehning, J.A. (1992). The EGD1 product, a yeast homolog of human BTF3, may be involved in GAL4 DNA binding. Molecular and cellular biology 12, 5683-5689.

Patzelt, H., Rudiger, S., Brehmer, D., Kramer, G., Vorderwulbecke, S., Schaffitzel, E., Waitz, A., Hesterkamp, T., Dong, L., Schneider-Mergener, J., et al. (2001). Binding specificity of Escherichia coli trigger factor. Proceedings of the National Academy of Sciences of the United States of America 98, 14244-14249.

Pech, M. (2006). Identifizierung und Charakterisierung der Bindungsstelle des Nascent Polypeptide-Associated Complex am Ribosom. Dissertation.

Pech, M., Spreter, T., Beckmann, R., and Beatrix, B. (2010). Dual binding mode of the nascent polypeptide-associated complex reveals a novel universal adapter site on the ribosome. The Journal of biological chemistry 285, 19679-19687.

Peisker, K., Braun, D., Wolfle, T., Hentschel, J., Funfschilling, U., Fischer, G., Sickmann, A., and Rospert, S. (2008). Ribosome-associated complex binds to ribosomes in close proximity of Rpl31 at the exit of the polypeptide tunnel in yeast. Molecular biology of the cell 19, 5279-5288.

Penczek, P.A., Frank, J., and Spahn, C.M. (2006a). A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation. Journal of structural biology 154, 184-194.

Penczek, P.A., Yang, C., Frank, J., and Spahn, C.M. (2006b). Estimation of variance in single-particle reconstruction using the bootstrap technique. Journal of structural biology 154, 168-183.

Pfund, C., Lopez-Hoyo, N., Ziegelhoffer, T., Schilke, B.A., Lopez-Buesa, P., Walter, W.A., Wiedmann, M., and Craig, E.A. (1998). The molecular chaperone Ssb from Saccharomyces cerevisiae is a component of the ribosome-nascent chain complex. The EMBO journal 17, 3981-3989.

Pollack, S.J., Beyer, K.S., Lock, C., Muller, I., Sheppard, D., Lipkin, M., Hardick, D., Blurton, P., Leonard, P.M., Hubbard, P.A., et al. (2011). A comparative study of fragment screening methods on the p38alpha kinase: new methods, new insights. Journal of computer-aided molecular design 25, 677-687.

Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450, 663-669.

Rappe, M.S., Connon, S.A., Vergin, K.L., and Giovannoni, S.J. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 418, 630-633.

Raue, U., Oellerer, S., and Rospert, S. (2007). Association of protein biogenesis factors at the yeast ribosomal tunnel exit is affected by the translational status and nascent polypeptide sequence. The Journal of biological chemistry *282*, 7809-7816.

Reimann, B., Bradsher, J., Franke, J., Hartmann, E., Wiedmann, M., Prehn, S., and Wiedmann, B. (1999). Initial characterization of the nascent polypeptide-associated complex in yeast. Yeast (Chichester, England) *15*, 397-407.

Roehrl, M.H., Wang, J.Y., and Wagner, G. (2004). A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. Biochemistry *43*, 16056-16066.

Schlunzen, F., Wilson, D.N., Tian, P., Harms, J.M., McInnes, S.J., Hansen, H.A., Albrecht, R., Buerger, J., Wilbanks, S.M., and Fucini, P. (2005). The binding mode of the trigger factor on the ribosome: implications for protein folding and SRP interaction. Structure *13*, 1685-1694.

Seidelt, B., Innis, C.A., Wilson, D.N., Gartmann, M., Armache, J.P., Villa, E., Trabuco, L.G., Becker, T., Mielke, T., Schulten, K., et al. (2009). Structural insight into nascent polypeptide chain-mediated translational stalling. Science (New York, NY) *326*, 1412-1415.

Shi, X., Parthun, M.R., and Jaehning, J.A. (1995). The yeast EGD2 gene encodes a homologue of the alpha NAC subunit of the human nascent-polypeptide-associated complex. Gene *165*, 199-202.

Soret, C. (1879). Sur l'état d'équilibre que prend, du point de vue de sa concentration, une dissolution saline primitivement homogène, dont deux parties sont portées à des températures différentes. Archives de Genève *3e période*, 48.

Spreter, T., Pech, M., and Beatrix, B. (2005). The crystal structure of archaeal nascent polypeptide-associated complex (NAC) reveals a unique fold and the presence of a ubiquitin-associated domain. The Journal of biological chemistry *280*, 15849-15854.

Stansfield, I., Jones, K.M., Herbert, P., Lewendon, A., Shaw, W.V., and Tuite, M.F. (1998). Missense translation errors in Saccharomyces cerevisiae. Journal of molecular biology *282*, 13-24.

Tachiwana, H., Osakabe, A., Shiga, T., Miya, Y., Kimura, H., Kagawa, W., and Kurumizaka, H. (2011). Structures of human nucleosomes containing major histone H3 variants. Acta crystallographica Section D, Biological crystallography *67*, 578-583.

Teter, S.A., Houry, W.A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F.U. (1999). Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. Cell *97*, 755-765.

Tyson, C.B., Lord, P.G., and Wheals, A.E. (1979). Dependency of size of Saccharomyces cerevisiae cells on growth rate. Journal of bacteriology *138*, 92-98.

Ullers, R.S., Houben, E.N., Brunner, J., Oudega, B., Harms, N., and Luirink, J. (2006). Sequence-specific interactions of nascent Escherichia coli polypeptides with trigger factor and signal recognition particle. The Journal of biological chemistry *281*, 13999-14005.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., et al. (2000). Single-particle electron cryo-microscopy: towards atomic resolution. Q Rev Biophys *33*, 307-369.

Wang, C.C., Morales, A.J., and Schimmel, P. (2000). Functional redundancy in the nonspecific RNA binding domain of a class I tRNA synthetase. The Journal of biological chemistry *275*, 17180-17186.

Wang, L., Zhang, W., Zhang, X.C., Li, X., and Rao, Z. (2010). Crystal structures of NAC domains of human nascent polypeptide-associated complex (NAC) and its alphaNAC subunit. Protein & cell *1*, 406-416.

Wang, S., Sakai, H., and Wiedmann, M. (1995). NAC covers ribosome-associated nascent chains thereby forming a protective environment for regions of nascent chains just emerging from the peptidyl transferase center. The Journal of cell biology *130*, 519-528.

Weber, G. (1953). Rotational Brownian motion and polarization of the fluorescence of solutions. Advances in protein chemistry *8*, 415-459.

Wegrzyn, R.D., Hofmann, D., Merz, F., Nikolay, R., Rauch, T., Graf, C., and Deuerling, E. (2006). A conserved motif is prerequisite for the interaction of NAC with ribosomal protein L23 and nascent chains. The Journal of biological chemistry *281*, 2847-2857.

Wiedmann, B., Sakai, H., Davis, T.A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. Nature *370*, 434-440.

Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Analytical biochemistry *179*, 131-137.

Woolhead, C.A., McCormick, P.J., and Johnson, A.E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. Cell *116*, 725-736.

Yao, S., Blaustein, J.B., and Bechhofer, D.H. (2008). Erythromycin-induced ribosome stalling and RNase J1-mediated mRNA processing in Bacillus subtilis. Mol Microbiol *69*, 1439-1449.

Yogev, O., Karniely, S., and Pines, O. (2007). Translation-coupled translocation of yeast fumarase into mitochondria in vivo. The Journal of biological chemistry *282*, 29222-29229.

Yonath, A., Leonard, K.R., and Wittmann, H.G. (1987). A tunnel in the large ribosomal subunit revealed by three-dimensional image reconstruction. Science (New York, NY) *236*, 813-816.

Yotov, W.V., and St-Arnaud, R. (1996). Differential splicing-in of a proline-rich exon converts alphaNAC into a muscle-specific transcription factor. Genes & development *10*, 1763-1772.

Zarnt, T., Tradler, T., Stoller, G., Scholz, C., Schmid, F.X., and Fischer, G. (1997). Modular structure of the trigger factor required for high activity in protein folding. Journal of molecular biology *271*, 827-837.

Zhang, Y., Berndt, U., Golz, H., Tais, A., Oellerer, S., Wolfle, T., Fitzke, E., and Rospert, S. (2012). NAC functions as a modulator of SRP during the early steps of protein targeting to the ER. Molecular biology of the cell.

Zheng, X.M., Black, D., Chambon, P., and Egly, J.M. (1990).
Sequencing and expression of complementary DNA for the
general transcription factor BTF3. Nature *344*, 556-559.

Zheng, X.M., Moncollin, V., Egly, J.M., and Chambon, P. (1987).
A general transcription factor forms a stable complex with RNA
polymerase B (II). Cell *50*, 361-368.

8   Acknowledgments

PERSONAL PROFILE

Name	Julian Deeng
Date of birth	21.02.1984
Place of birth	Ochsenhausen, Germany
Citizenship	German

EDUCATIONAL AND PRACTICAL TRAININGS

2008-2012	PhD thesis at the Gene center der LMU München
2007-2008	Diploma thesis at the Max-Planck-Institute for Biochemistry, Martinsried
2003-2008	studies in biology, University of Würzburg
2003	Abitur at Gebhard-Müller-Schule, Biberach

9 Appendix

9.1 Unspecific interaction between proteins and RNAs

The *S. cerevisiae* glutaminyl tRNA synthetase was shown to contain two domains that mediate an unspecific interaction with RNAs. Both domains are enriched in lysines (the first contains four lysines, the second six (Wang et al., 2000)). The crystal structure of the N-terminal 187 AA that contain the first domain revealed, that these lysines are placed in an α-helix that positions them optimal to interact nonspecifically with the RNA backbone (Grant et al., 2012).
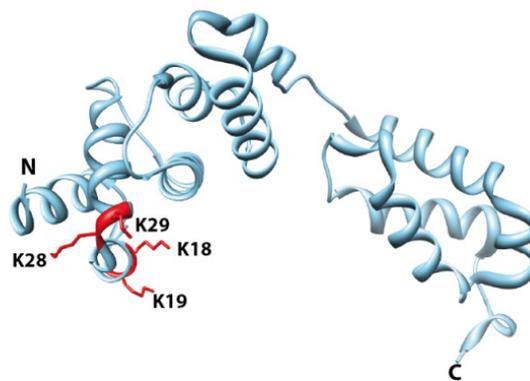
Figure 71: The N-terminus of Gln4 (yeast glutaminyl-tRNA synthetase) contains a lysine rich α-helix that mediates unspecific interaction with tRNAs | (pdb: 3tl4) (Grant et al., 2012)

The prime example for unspecific interaction with nucleic acids is the interaction between histones and DNA. Amongst others, basic AA located in an α-helix mediate the interaction with the phosphate backbone of the DNA.
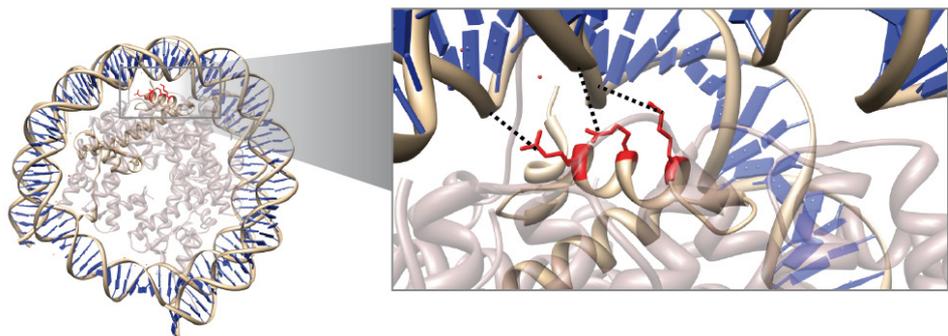
Figure 72: Histone proteins interact nonspecifically with DNA using basic AA that are located in an α-helix| pdb: 3av2 (Tachiwana et al., 2011)

9.2 EF-Tu EF-Ts interaction, a UBA-domain mediated interaction

The N-terminal domain of EF-Ts contains a UBA domain (also called TS-N domain). This domain mediates in concert with other domains the interaction of EF-Tu with EF-Ts and is thus an example how UBA domains are involved in protein-protein interactions (Kawashima et al., 1996).
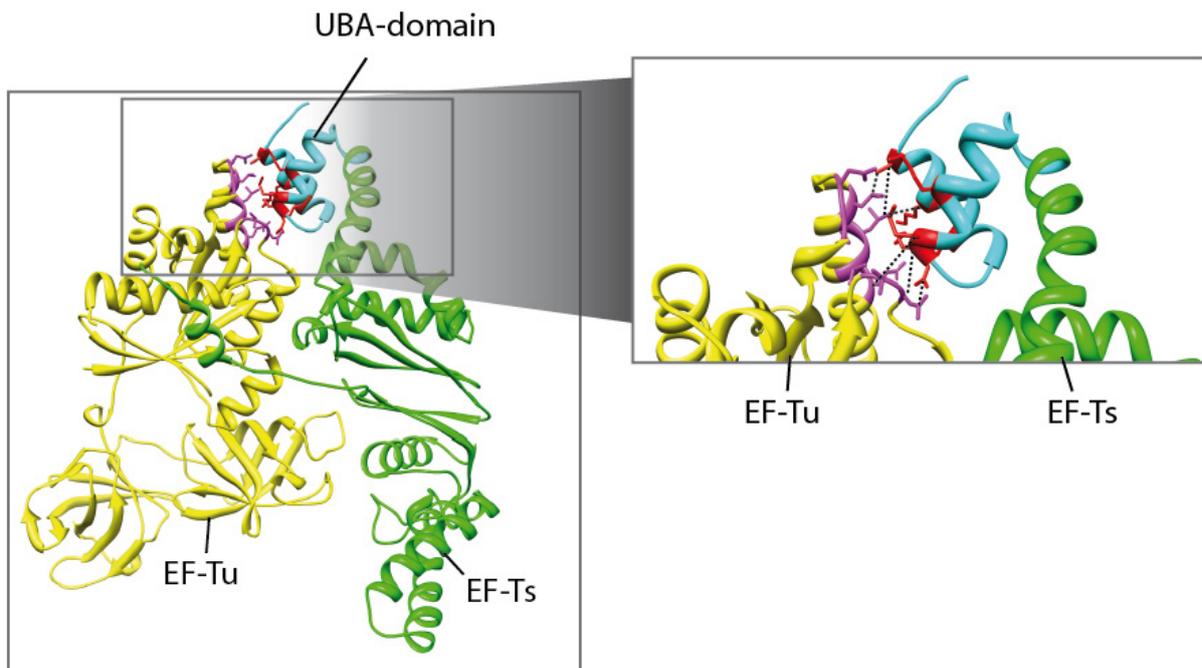
Figure 73: The N-terminal UBA-domain of EF-Ts mediates protein-protein interactions | The N-terminal 25 AA of EF-Ts (green) fold into a UBA domain (cyan). Amongst other interactions, this domain is responsible for the dimerization of EF-Tu and EF-Ts. Pdb: 1efu, (Kawashima et al., 1996)

9.3 Analysis of the interaction between NAC derivatives and tRNAs using EMSA

Increasing concentrations of full length yeast NAC, a fusion protein containing the N-terminus of Egd1p or MBP alone were incubated with 20 pmol tRNA in 15 µl. After complex formation (15 mins), the complexes were separated using Agarose gel electrophoresis (0.5% Agarose). The tRNA was visualized using GelRed.
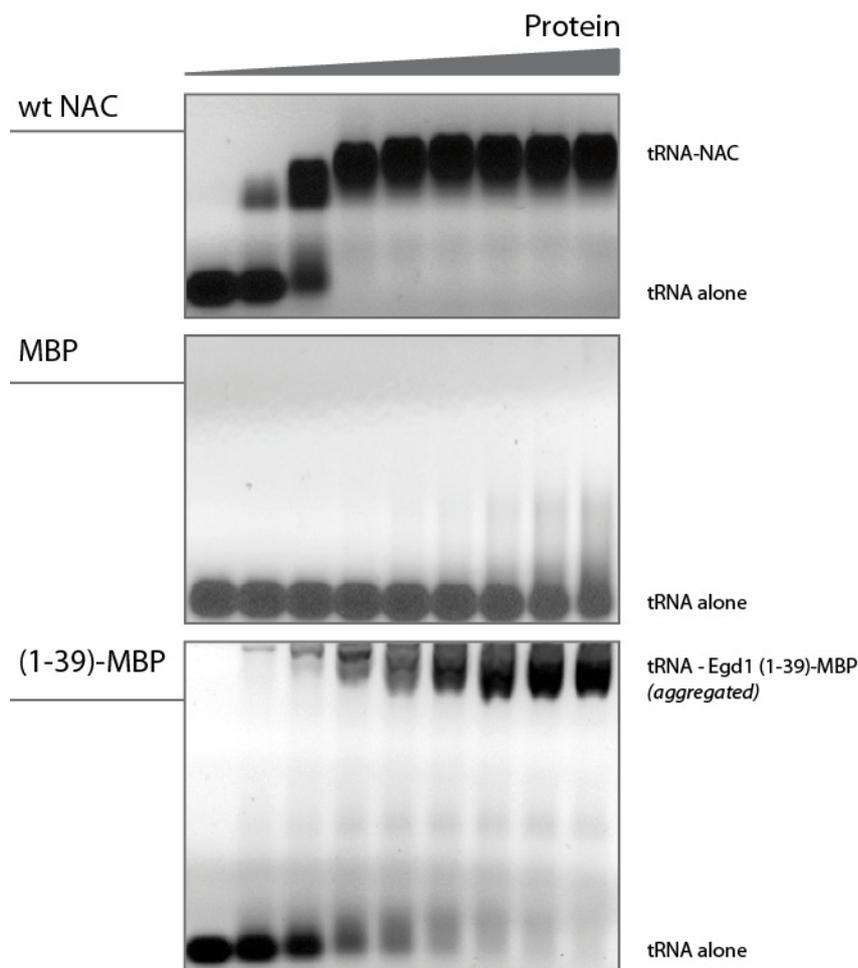
Figure 74: NAC and the fusion protein Egd1 (1-39) MBP interacts with tRNAs in EMSA assays| 20 pmol of tRNA were incubated with increasing concentrations of the respective protein (0-150 pmol) in 15 µl. NAC and the fusion protein form a complex with tRNAs while MBP alone does not. (Experiments performed by Steffen Fliehmann)

9.4 Correspondence analysis on a Cryo-EM dataset

We first tried to used a 2D classification strategy to group particles of a given projection group into subgroups with defined TF conformations. Therefore, we used aligned particles (containing P-site and no E-site) and grouped them into 195 quasi-evenly spaced projection directions (SPIDER command vo ea, angular increment 10°) and chose projection direction 45 since it represents a view where the TF binding site is visible (Figure 75, Panel C).
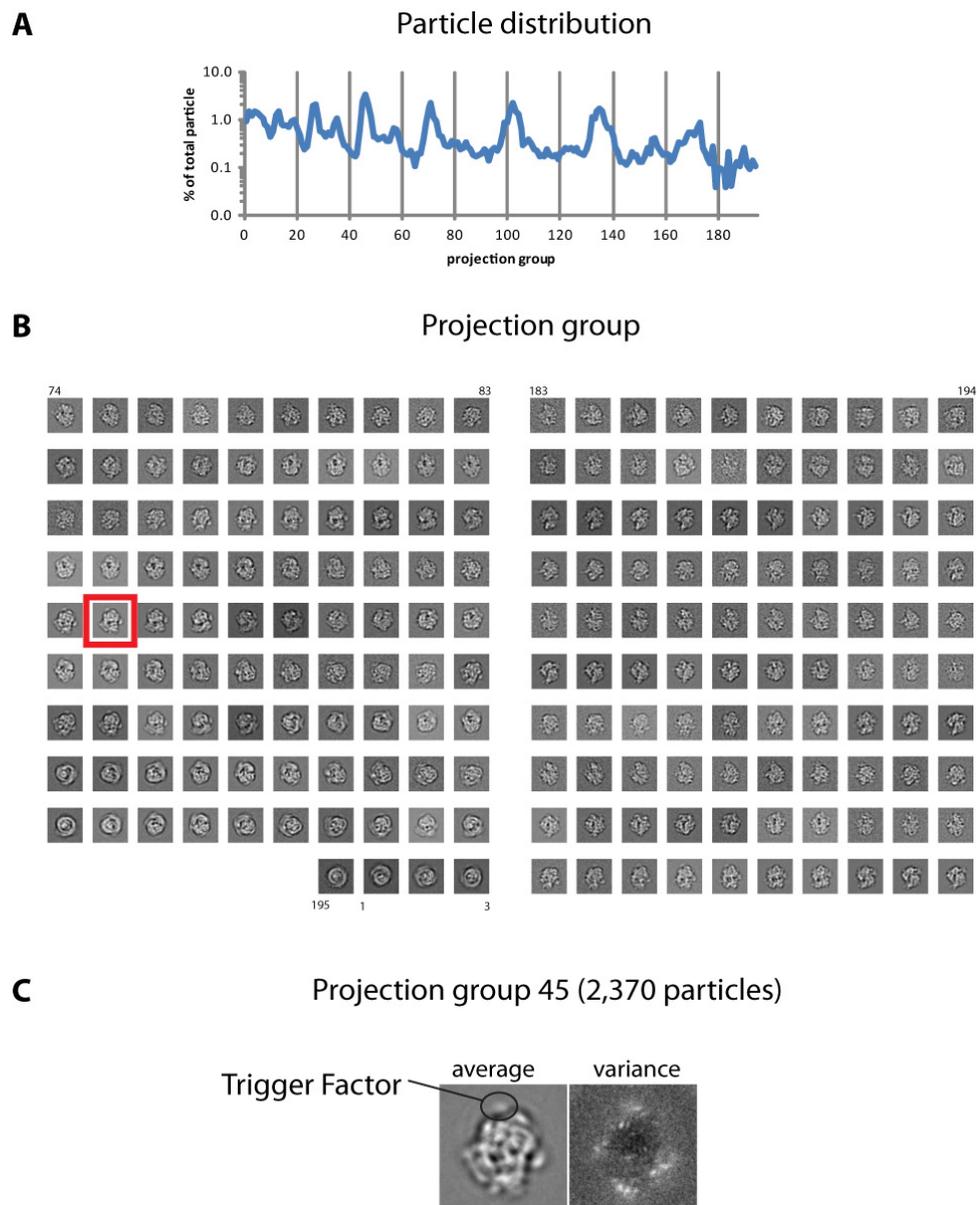
Figure 75 : Aligned particles that all contained P-site tRNAs were grouped according their euler angles into 195 projection directions | The particle showed the for ribosomes typical uneven distribution over the projection space (Panel A). Projection group 45 contained 2,370 particles and showed a weak density for TF (Panel C).

The analysis of the dataset using a correspondence analysis (with 8 factors) showed, that only the first Eigenimage contained any significant inter-image variance (Figure 76, panel A). Eigenimage is the dimension, Eigenvalue the value that describes the particle with respect to the dimension of the Eigenimage. A plot of Eigenimage 1 against Eigenimage 2 showed the presence of two distinct populations (Figure 76, panel B). The reconstruction of the Eigenpictures however showed, that this Eigenimage did not carry any information in the region where TF binds (Figure 75), but information that is present all over the ribosome (Figure 76, panel C Eigenimage 1). The low signal intensity and/or the absence of distinct conformational states of TF might obstruct the application of this technique to the TF-RNC project. It was therefore not pursued further.
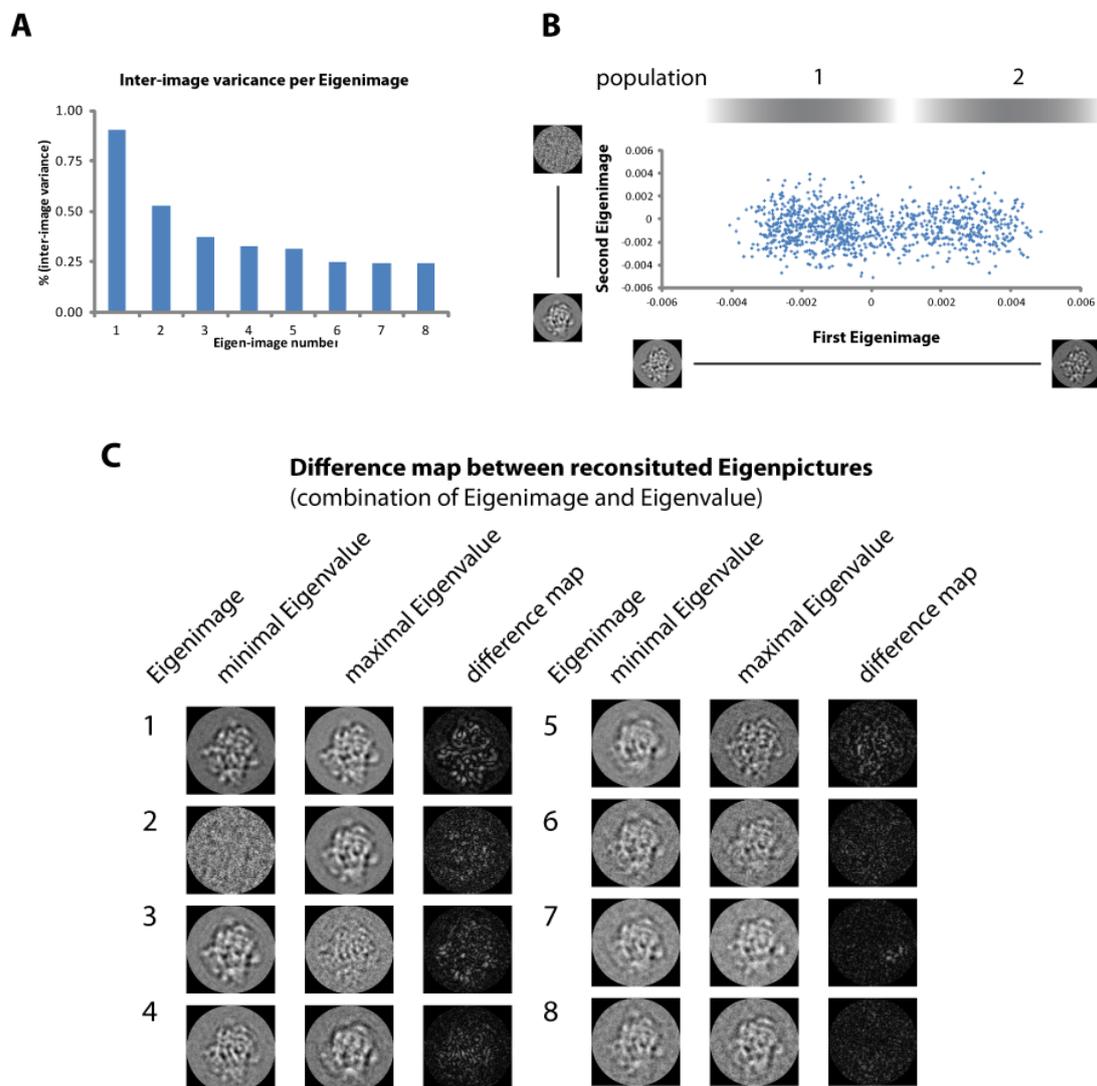
Figure 76 : The Eigenimage analysis of the particles did not allow for sorting according to the TF state | Projection direction group 45 (Figure 75) was subjected to a correspondence analysis (CORAN). Only the first eigenimage coded for significant inter-image variation (Panel A). The plotting of the first against the second Eigenimage showed the presence of two particle distribution (Panel B). However, no signal with respect to TF was encoded in any of the Eigenimages as shown by the reconstruction of particles using eigenvalues present in the dataset (CA SR command in spider, Panel C).

9.5   3D bootstrapping

Using the bootstrapping technique, Penczek and Co-workers succeeded to define unbiasely regions of high variance (Penczek et al., 2006a; Penczek et al., 2006b). We applied this technique on the TF dataset to define the region that is sampled by TF. The basis therefore is a collection of 3D volumes that are subdatasets of one dataset. To optimize the procedure, we first tested different defocus correction techniques and backprojection algorithms for the ideal combination of speed and accuracy.

Using 14,501 particles with a pixel size of 2.475 Å/pixel, we first compared (1) backprojection without defocus correction; (2) backprojection in defocus groups with wiener filtering; (3) with weightening of the defocusgroups according to the particle distribution and (4) multiplication of the individual particles with the CTF-function in the Fourier space prior to the backprojection in one group. The resolution of the third strategy was superior to the others. Surprisingly, the fourth strategy yielded the second best result at a significantly higher speed. The time saved by this strategy is dependent on the architecture of the dataset. The more defocus groups are used in the third strategy, the higher is the time-benefit of the fourth technique. The second variable is the backprojection algorithm. We tested four fourier space based algorithms that are implemented in spider and found that BP 3N was the fastest and the best.

To further optimize the procedure, outliers had to be removed. If accidentally assigned, they artificially create signals that are not reflections of the compositional or conformational heterogeneity. This is achieved by comparing the average CC C of the particles of the total dataset with the individual CC C of particles. Particles that do not fall within a certain CC C region are removed.

> 10,000 volumes were backprojected and the variance was calculated voxel for voxel. In a pilot experiment, we were able to visualize heterogeneity created by substoichiometric occupancy of the P-site tRNA but failed to visualize heterogeneity caused by Trigger Factor. Therefore, this technique was not further pursued. A diagram of the workflow is shown in Figure 77.
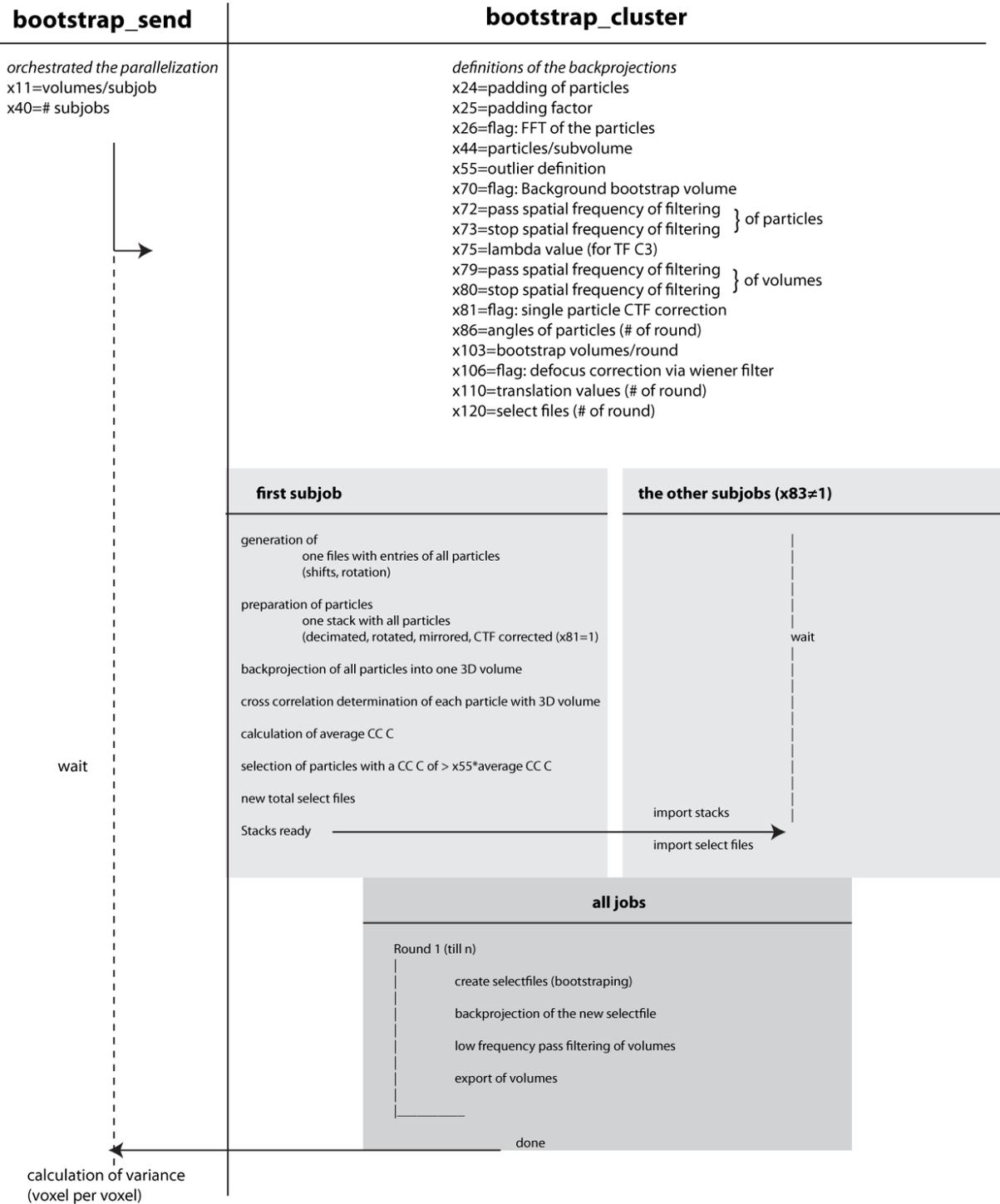
Figure 77: Overview of the 3D bootstrapping generation | Using bootstrapping, n selectfiles are generated from a total_select file. The first subjob organizes the particles to optimize speed by (i) generating a stack with all particles, (ii) elimination of outliers and (iii) CTF correction on single particles. Once this preliminary work is done, m subjobs generate the volumes based on the rearranged dataset. Finally, the variance voxel per voxel is calculated