DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES
DER FAKULTÄT FÜR CHEMIE UND PHARMAZIE
DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Investigation of conformational dynamics using single-molecule FRET: Following different steps on the path to protein expression

vorgelegt von
Lena Freiin Voith von Voithenberg
aus
Bielefeld, Deutschland

2016

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Don C. Lamb, PhD, betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den

(Unterschrift des Autors)

Dissertation eingereicht am 10.05.2016

1. Gutachter: Prof. Don C. Lamb, PhD
2. Gutachter: Prof. Dr. Christoph Bräuchle

Mündliche Prüfung am 10.06.2016

Abstract

Over the last two decades, single molecule fluorescence methods have become widely used to tackle biological questions. The advantage of detecting individual subpopulations offers the possibility to study heterogeneous solutions of molecules, molecular binding, and dynamics. In this thesis, fluorescence correlation spectroscopy as well as single-pair and single molecule three-color Förster resonance energy transfer (FRET) were applied for the observation of cell biological processes. The conformational changes of proteins and nucleic acids involved in protein expression were studied.

On the level of transcription initiation, spFRET was applied to investigate the conformation of DNA upon protein binding. The transcription initiation factor TATA-box binding protein (TBP) induced a dynamic switching behavior of the DNA between three observable FRET states. Further investigations explained the dissociating effect of different cofactors on TBP from DNA. The interplay of these transcription initiation factors regulates gene expression.

Once fully transcribed, pre-mRNA is further processed inside the nucleus. The U2 auxiliary factor plays an important role in the recognition of 3'splice sites for intron splicing. Its large subunit, U2AF65, was shown to undergo large scale conformational rearrangements in the absence and presence of RNA. U2AF65 is intrinsically highly dynamic and switches between a closed and an open conformation. The addition of polypyrimidine tracts of increasing strengths, gradually shifts the equilibrium of the dynamics towards the open state. Extension of the large subunit U2AF65 by the addition of its small partner, U2AF35, further enhances this effect. The equilibrium between the closed and open conformation correlates with the binding affinity of the different constructs. The shift in equilibrium therefore offers an explanation on the molecular mechanism underlying different splicing efficiencies from Py-tracts of varying strengths.

Mature mRNA is translated into protein in the cytosol of eukaryotic cells. Many nascent proteins require the aid of molecular chaperones for *de novo* folding. Hsp70 chaperones, which often already bind to the proteins cotranslationally, were investigated by single molecule three-color FRET. The simultaneous detection of three distances makes it possible to distinguish coordinated from non-coordinated motion. Hsp70s are known to allosterically communicate between their domains. We therefore analyzed the coordination of their dynamic behavior. The comparison of three Hsp70 chaperones from different organelles and organisms revealed differences in their coordinated motion and therefore point to different functionalities and evolutionary development of these chaperones.

The substrate of one of these chaperones, the prokaryotic Hsp70 DnaK, was investigated in a next step. DnaK binds to its substrate, the transcription factor Sigma32, to keep it in an inactive form. The conformational transition of Sigma32 upon DnaK binding was probed using spFRET. Conformational rearrangements of the protein bound to DnaK and in its free

form were measured and can explain the activation of Sigma32 upon release from DnaK. Once activated, Sigma32 binds to RNA polymerase and promoter regions of heat shock induced genes.

Table of contents

Abbreviations

A	adenine
ADP	adenosine diphosphate
AMP-PNP	adenosine 5'-(β,γ-imido)triphosphate
ATP	adenosine triphosphate
AOTF	acousto-optical tunable filter
APD	avalanche photodiode
BBP	branch-point binding protein
BiP	binding-immunoglobulin protein, Hsp70 from endoplasmic reticulum
bp	base pair
BSA	bovine serum albumin
C	cytosine
C-terminus	carboxy terminus
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulation factor
DM	dichroic mirror
DNA	desoxyribonucleic acid
DnaK	bacterial Hsp70
DNase	desoxyribonuclease
dNTP	desoxynucleoside triphosphate
dsDNA	double-strand DNA
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DTT	dithiothreitol
EBNA1	Epstein–Barr nuclear antigen 1
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EF	emission filter
EJC	exon junction complex
EMCCD	electron multiplying charge-coupled device
ER	endoplasmic reticulum
FCCS	fluorescence cross-correlation spectroscopy
FCS	fluorescence correlation spectroscopy
FRET	Förster resonance energy transfer
G	guanine
H2B	histone 2B
Hepes	2-(4-(2-hydroxyethyl)-1-piperazinyl)-ethansulfonic acid
HMGA1	high-mobility group protein
HMM	hidden Markov modeling
HPLC	high pressure liquid chromatography

Hsp70	heat shock protein, molecular size 70 kDa
IAA	iodoacetamide
IRF	instrument response function
L	lens
LB	Luria broth culture medium
M	mirror
MFD	multiparameter fluorescence detection
Mge1	nucleotide exchange factor of Ssc1
M. jannaschii	*Methanocaldococcus jannaschii*
MLE	maximum likelihood estimator
M. musculus	*Mus musculus*
Mot1	modifier of transcription 1, cofactor of TBP
mRNA	messenger RNA
NADH	nicotinamide adenine dinucleotide
NBD	nucleotide binding domain
NC1	negative cofactor 1, cofactor of TBP
NC2	negative cofactor 2, cofactor of TBP
NEF	nucleotide exchange factor
NMR	nuclear magnetic resonance
NPC	nuclear pore complex
N-terminus	amino terminus
OD	optical density
ORC	origin recognition complex
P	pinhole
PAGE	polyacrylamide gel electrophoresis
PAM	presequence translocase-associated motor
PAP	poly-A polymerase
PBS	polarizing beam splitter
PC4	positive cofactor 4, cofactor of TBP
PCR	polymerase chain reaction
PDA	probability distribution analysis
PDB	protein data bank accession numbers
PE	phosphatidylethanolamine
PEG	poly(ethylene glycol)
PEP	phosphoenol pyruvate
PIE	pulsed interleaved excitation
PM	polychroic mirror
PMSF	phenylmethylsulfonylfluorid
pre-mRNA	precursor mRNA
PrK	N-propargyl-lysine
PSF	point-spread function
RNA	ribonucleic acid

RNase	ribonuclease
RT	room temperature
SBD	substrate binding domain
S. cerevisiae	*Saccharomyces cerevisiae*
Ssc1	mitochondrial Hsp70 from *S. cerevisia*
SD	standard deviation
SDS	sodium dodecyl sulphate
SMF	single mode fiber
Snf2/Swi2 ATPase	switch/sucrose non-fermenting ATPase
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleic protein
spFRET	single pair Förster resonance energy transfer
STED	stimulated emission depletion
T	thymine
TBP	TATA-box binding protein
TBTA	tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl) methyl]amin)
TCEP	tris(2-carboxyethyl)phosphin
TCSPC	time-correlated single photon counting
TDP	transition density plot
TEAB	triethylammonium bicarbonate
TFA	trifluoroacetic acid
TIM	translocon of the inner mitochondrial membrane
TIRF	total internal reflection
Tris	tris(hydroxymethyl)aminomethane
u	uracil
v/v	volume per volume
w/v	mass per volume
WT	wildtype

1 Introduction

Over the past few decades, many biological questions were tackled by combining tools from physics and biology. Large achievements in the field are structural approaches using X-ray crystallography or nuclear magnetic resonance. The discoveries of protein or nucleic acid structures combined with biochemical investigations can often reveal a functional mechanism of the molecule of interest. In some cases, however, even though biological processes have been studied extensively by biochemical and structural means, a detailed molecular understanding of the functional mechanism of individual proteins is still missing. The application of fluorescence methods provides a great opportunity to approach these open questions. The attachment of fluorescent probes usually causes little structural effects and is non-invasive. The molecules of interest are therefore not disturbed in their proper mechanism and can be followed in real-time *in vitro* or *in vivo*.

Almost all cellular mechanisms involve proteins which undergo conformational changes to perform their function. A heterogeneous mixture of different conformations is therefore often present and individual populations can only be distinguished by single molecule approaches. Single molecule methods offer great advantages over ensemble experiments in separating individual populations from heterogeneous solutions. For static heterogeneous mixtures of molecules, an *a priori* separation may be possible. For molecules undergoing conformational changes during the experiment, a separation only becomes possible using single molecule approaches. In the last years, multiple methods in this field were developed, ranging from force measurements using optical tweezers (Perkins et al., 1994) or atomic force microscopy (Binnig et al., 1986), over torsional approaches using magnetic tweezers (Smith et al., 1992) to single molecule fluorescence experiments. A variety of reviews exist which describe these methods and their applications to biological systems (Deniz et al., 2008; Joo et al., 2008; Walter et al., 2008).

1.1 Single molecule fluorescence

The phenomenon of fluorescence is omnipresent; to determine when its first observation took place is therefore a difficult task. One of the first documentations was the observation of wood extract by Nicolas Monardes. In the middle of the 19^{th} century the phenomenon was then termed “fluorescence” by George Gabriel Stokes and Herschel (1845) first documented the observation of fluorescence in a solution of quinine sulfate in great detail. Nowadays, the concept of fluorescence is explained by the description of molecules which are raised into their excited state by the absorption of light. When they return to the ground state by emitting a photon on the timescale of 10^{-9} to 10^{-7}s this phenomenon is termed fluorescence. The use of fluorescence for biological applications emerged during the 19^{th} century and Stanislav von Provazek was probably the first who, by staining cells, documented fluorescence as an application to biology in 1914. Fluorescence microscopy was introduced by August Köhler, who constructed the first ultraviolet microscope in 1904, and Oskar Heimstädt, who built the first fluorescence microscope in 1910. From this timepoint onwards and especially with the discovery of fluorescent proteins and the establishment of confocal microscopy by Marvin

Minsky in 1957, fluorescence methods became a widely used technique in the biological sciences.

Using fluorescence microscopic methods, a multitude of cell biological and molecular processes can be investigated and visualized. Many biological processes, however, do not only contain a single homogeneous species. When using ensemble fluorescence methods an averaging of species occurs. To detect the behavior of individual molecules, fluorescence methods were developed in the second half of the 20th century. The detection of single molecules poses the great advantage of revealing individual subpopulations, distributions of states, and rare species. To achieve the detection of individual molecules, high signal intensities and low background are key requirements.

The first single molecule fluorescence experiments were performed by Rotman (1961), who observed single enzymes separated in droplets and Hirschfeld (1976) investigating individual molecules which carried 80-90 fluorophores each. It was only around 15 years later that the detection of individual fluorophores became possible (Brooks Shera et al., 1990). Necessary prerequisites for single fluorophore detection were technical advances in the field of detectors and the suppression of background by the reduction of the observation volume. Two microscopy methods are especially well suited in providing a small detection volume. The restriction of the detection volume down to femtoliters by a pinhole in confocal microscopy gives the opportunity to follow individual fluorophores in solution (Chapter 2.10.2) (Pawley, 2006). Recently, the volume has been even further decreased by stimulated emission depletion (STED) microscopy to allow for higher sample concentrations while still detecting individual molecules (Hell and Wichmann, 1994). Scanning of the single confocal spot has to be used to image a full field-of-view, however. Total internal reflection fluorescence microscopy, on the other hand, allows for the observation of a complete field-of-view by limiting the excitation by an evanescent wave to ~200 nm below the surface of total internal reflection (Chapter 2.10.3). For detection, the fluorophores need to be close to this surface and therefore are often attached to the passivated prism or coverglass (Chapter 2.9.2).

Interactions, association, and dissociation of molecules are widely studied by fluorescence correlation methods. Fluorescence correlation spectroscopy (FCS) uses temporal intensity fluctuations in the fixed focal volume of a confocal microscope (Chapter 2.10.1) (Rigler and Elson, 2001). Dependent on the nature of the fluctuations, different parameters can be extracted. Intensity fluctuations based on the diffusion of molecules due to Brownian motion contain information about their diffusion coefficient and concentration. On shorter timescales, photophysical conversions of the fluorophores to non-fluorescent triplet states and rotational diffusion can be quantified.

Structural information about the molecules of interest can be obtained by distance measurements using Förster resonance energy transfer (FRET). Single-pair FRET was described by Ha et al., (1996), Deniz et al., (1999), and Weiss (1999) (Chapter 2.10.2). A donor fluorophore in its excited state transfers its energy to an acceptor in close proximity by dipole-dipole interactions. The spectral overlap and the distance between the donor and acceptor fluorophores determine the value of the energy transfer. This method allows

conformational changes to be observed in a distance range of ~20-80 Å, a distance well suited for single molecule studies (Stryer, 1978). The great advantage of distance measurements using FRET is the possibility to not only observe distances, but also dynamic changes of this parameter (Weiss, 2000).

1.2 Following conformational changes of proteins along the way to protein expression

With the exception of structural proteins, many cellular proteins undergo conformational changes. Various steps during protein expression necessitate conformational changes of proteins for their proper functionality. These conformational transitions and their dynamic interconversion were investigated in this thesis by spFRET experiments and correlation spectroscopic methods.

Using these methods, the conformational changes of a variety of proteins along the path to protein expression were followed (Figure 1.1). The journey starts in the nucleus where we investigated HMGA1 and its Epstein-Barr-Virus mimick EBNA1 (Chapter 3). HMGA1 has been ascribed multiple cellular roles with its most important function lying in the early steps of transcription initiation. Recently, it was found to be involved in DNA replication. Its function in DNA and RNA binding and associated conformational changes of the nucleic acids were investigated. HMGA1 is not the only protein associated with the early phase of transcription. Due to the great importance of the correct regulation of gene expression, a multitude of transcription factors are known. Especially critically is the regulation of transcription initiation. TBP, a subunit of the transcription initiation factor IID, is the first protein that binds to promoters to initiate transcription. Chapter 4 analyzes the role of TBP cofactors on transcription initiation. After successful transcription, the pre-mRNA is processed still inside the nucleus. In Chapter 5, we have a look at U2 auxiliary factor, a protein complex necessary for the early phase of spliceosome assembly at the 3'splice site. In the absence and presence of Py-tract RNA, U2AF65 shows large scale conformational changes of its domains. The dynamic switching between these states is investigated in this chapter. After successful transcription and translation, the last step of protein production involves the folding of the proteins into their native structure. Hsp70 chaperones co- and posttranslationally act on freshly produced proteins to aid in their folding. The conformational cycle of three of these chaperones from different organelles was analyzed using three-color FRET (Chapter 6). We analyzed the allosteric communication between the nucleotide binding domain and the substrate binding domain of the Hsp70s in the absence and presence of substrate and nucleotide exchange factors. The comparison of the three Hsp70s is of evolutionary interest considering the bacterial origin of the different organelles. The protein Sigma32, described in Chapter 7, closes the cycle of gene expression from final protein folding to the early phase of transcription. This bacterial transcription factor is kept in its inactive form when bound to the Hsp70 DnaK. Upon heat-shock, it is released and binds as an activating factor to promoters of heat-shock induced genes. The conformational changes necessary for it to gain its activity are described in this chapter.

Even though, proteins from different organisms were investigated, the overall cycle of gene expression describes their place and function inside the cell reasonably well. Following the chapters of this thesis can therefore be considered as taking a stroll through the cell along the path of protein expression.

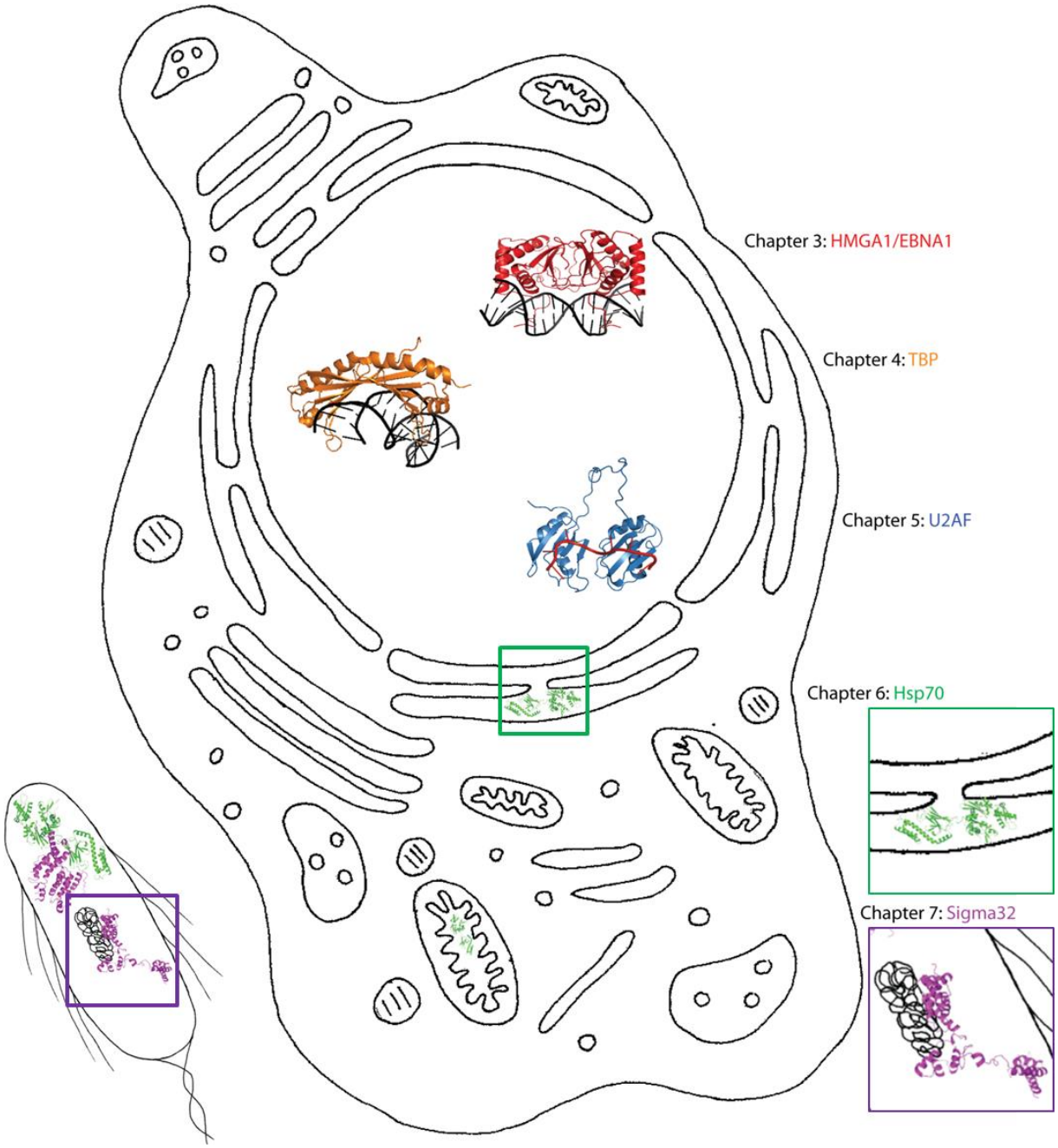


Figure 1.1 Overview of proteins investigated in this thesis schematically depicted inside their respective compartments in the cell or bacterium (not to scale). Chapter 3 describes binding studies of HMGA1 and EBNA1. Both proteins are involved in replication. In Chapter 4 the effects of cofactors on the binding of TBP to promoter start sites are shown. Conformational dynamics of U2AF binding to pre-mRNA are analyzed in Chapter 5. Chapter 6 deals with the allosteric conformational effects of Hsp70 chaperones from different organelles. Sigma32 can interact with the Hsp70 from *E. coli* and is required as a transcription factor (Chapter 7).

2 Materials and Methods

This section describes the materials used during this work (Chapters 2.1-2.5). Furthermore, the methods applied to investigate the projects of this thesis are outlined in Chapters 2.6-2.11.

2.1 Chemicals and Solutions

Chemicals and solutions used in this work were ordered from AppliChem GmbH (Darmstadt, Germany), Atto-Tec GmbH (Siegen, Germany), Avanti Polar Lipids Inc. (Alabaster, USA), Bio-Rad Laboratories GmbH (München, Germany), Carl Roth GmbH (Karlsruhe, Germany), Clontech Laboratories Inc. (Mountain View, USA), GE Healthcare (München, Germany), Jena Bioscience GmbH (Jena, Germany), Laysan Bio Inc. (Arab, USA), Life Technologies GmbH (Darmstadt, Germany), Merck-Millipore (Darmstadt, Germany), New England Biolabs GmbH (Frankfurt, Germany), Qiagen (Hilden, Germany), SiChem GmbH (Bremen, Germany), Sigma-Aldrich (Taufkirchen, Germany), Thermo Fisher Scientific (Waltham, USA), and WTW GmbH (Weilheim, Germany).

2.2 Vectors and constructs

Plasmid constructs used in this work were provided by different laboratories (Appendix, Table 2.2) or cloned during this thesis (Appendix, Table 2.3).

2.3 Oligonucleotides

All oligonucleotides used in this work and listed in Table 2.4 (Appendix) were synthesized at and ordered from Metabion International AG (Planegg, Germany) and IBA GmbH (Göttingen, Germany) except noted otherwise.

2.4 Proteins

The proteins used in this work were either recombinantly expressed as described or kindly provided by our collaboration partners. Christoph-Erik Mayer and Julia Damaschke from the group of Dr. Aloys Schepers (Helmholtz Zentrum München, Germany) provided HMGA1 and EBNA1. We obtained Mot1 from the group of Professor David Auble (University of Virginia, USA), NC2, and PC4 from the lab of Professor Michael Meisterernst (Westfälische Wilhelms-Universität Münster, Germany). For the collaborative project investigating the splicing subunit U2AF65, proteins were provided by Carolina Sanchéz Rico from the group of Professor Michael Sattler (Technische Universität München, Germany). Sigma32 was cooperatively investigated with Roman Kityk and Professor Matthias Mayer (Ruprecht-Karls-Universität Heidelberg, Germany).

2.5 Bacterial cells

Microbiological work was performed in *Escherichia coli* (*E. coli*) TOP10 for cloning and plasmid preparation purposes and *E. coli* BL21-DE3 RIL and BL21AI strains for protein expression.

2.6 Microbiological Methods

2.6.1 Chemically competent bacterial cells

To render bacteria competent for plasmid uptake, approximately 50 mL *E. coli* cells were cultured over night at 37°C. They were then diluted 1:100 and grown for 1-2 hours until the culture reached an optical density OD_{600nm}=0.5-0.6. Bacteria were transferred sterilely into precooled falcons and collected by centrifugation at 5000 rpm and 4°C for 10 minutes. They were resuspended in 50 mL 100 mM $MgCl_2$ (0-4°C, sterile) and incubated for 20-30 min on ice. After collection of cells at 4000 rpm, 4°C for 10 min, the bacteria were resuspended in 10 mL 100 mM $CaCl_2$, 15% glycerol (0-4°C, sterile) and, after quick-freezing, stored as aliquots at -80°C.

2.6.2 Transformation of *Escherichia coli*

For plasmid transformation, chemically competent *E. coli* cells were thawed on ice. After addition of ~20 ng plasmid DNA or, for cotransformation of multiple vectors, 50 ng each, the cells were incubated on ice for 30 minutes. Following DNA attachment, plasmid uptake was induced by heating the bacteria to 42°C for 45 s. Cells were then grown in SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) at 37°C for 1 h, before spreading them onto LB-agar plates (1% (w/v) tryptone/peptone, 0.5% (w/v) yeast extract, 0.09 M NaCl, 1% agar, autoclaved) containing the respective antibiotic. The bacteria were incubated at 37°C overnight.
For electroporation, competent bacterial cells were thawed, 2 ng plasmid were added and uptake was ensured by an electrical voltage of 2.1 kV, 100 Ω, and 25 µF for 5 ms. Immediately afterwards, the cells were suspended in 200 µL SOC medium and incubated under vigorous agitation at 37°C for 1 hour.

2.6.3 Amplification and purification of plasmid DNA

For plasmid amplification, *E. coli* cultures were inoculated in 3-5 mL LB-medium (1% (w/v) tryptone/peptone, 0.5% (w/v) yeast extract, 0.09 M NaCl, autoclaved) comprising the respective antibiotic (100 µg/ml ampicillin or 50 µg/ml ampicillin and 33 µg/ml chloramphenicol) and cultured at 37°C overnight. Plasmid DNA was either purified using a QIAprep Spin Miniprep Kit (Qiagen GmbH, Hombrechtikon, Switzerland) or as described in the following procedure. The bacteria were collected by centrifugation, resuspended in 300 µl P1-solution (50 mM Tris pH 8.0, 10 mM EDTA pH 8.0, 100 µg/ml RNase), and lysed by adding 300 µl P2-solution (0.2 M NaOH, 1% SDS). Subsequently, 300 µl of P3-solution (3 M potassium acetate, pH 5.5 with acetic acid) were added for neutralization. After a 10 min incubation period on ice, cell debris was removed by centrifugation at 16000g and 4°C for 15 min. The plasmid DNA from the supernatant was precipitated using 600 µl isopropanol and centrifugation at 16000g, 4°C for 15 min. The plasmid was washed with 500 µl 70% ethanol. After votalization of the ethanol, the plasmid was dissolved in 50 µl 10 mM Tris-HCl. The concentration of plasmid DNA was determined by measuring the absorption at

260 nm. To account for the dependancy of the extinction coefficient on the DNA length, the value for the optical density OD_{260nm}=1 corresponds to 50 µg/mL of pure dsDNA.

2.6.4 Screening of clones (Colony PCR)

Clones with correctly inserted point mutations were found by sequencing of the amplified plasmid segment. To this end, a polymerase chain reaction mix (coral load buffer, 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.5 µM primer forward, 0.5 µM primer reverse, 2.5 units *Taq* DNA polymerase) was prepared accordingly for the number of expected clones. Transformed *E. coli* cells were selected from agar plates and each colony was mixed with the polymerase reaction mixture and at the same time plated for additional selection onto LB-agar plates with the respective antibiotic. The polymerase chain reaction was performed as described (Chapter 2.7.1) and amplified fragments were sent for DNA sequencing (Chapter 2.7.5).

2.7 Molecular Biological Methods

2.7.1 Polymerase chain reaction (PCR) and detection of DNA fragments

The amplification of DNA was achieved using polymerase chain reactions (PCR) using *Taq* DNA polymerase according to the protocol of the manufacturer (Qiagen, Hilden, Germany). The standard program for polymerase chain reactions included an initial denaturation step at 95°C for 3 min, which was followed by 35 cycles of denaturation (95°C, 45 s), annealing (54°C-61°C, 45 s), and elongation (72°C, 1 min) followed by a final elongation step at 72°C for 7 min.

For detection of PCR fragments, 1% agarose gels in 40 mM Tris, 20 mM acetate, 1 mM EDTA, pH 8.6 (TAE) were used. The samples were mixed with a 6x DNA loading buffer and loaded onto the gel. For size comparison of PCR fragments, a 1 kB DNA ladder (GeneRuler 1kB, Thermo-Fisher Scientific) was used. Electrophoresis on agarose gels was performed at 100 V for approximately 1 hour. The DNA was stained using a solution of SybrGold nucleic acid gel stain (Thermo-Fisher Scientific) and imaged on a gel documentation system (Gel Doc, Bio-Rad Laboratories, München, Germany).

2.7.2 Cloning strategies I: Phusion site-directed mutagenesis

For introduction of point mutations, the phusion site-directed mutagenesis kit (Thermo-Fisher Scientific) was used. 5'-phosphate containing primers were chosen in such a way that either the forward or the reverse primer contained base mismatches to introduce point mutations in the middle flanked by around 10-15 matching nucleotides on each side. The corresponding second primer was placed directly adjacent to the first one pointing in the opposite direction. Mutations were introduced by amplification of the whole plasmid with these primers. The reaction was performed in Phusion High Fidelity buffer containing 200 µM of each dNTP, 0.5 µM of the forward and the reverse primer, 50 ng of plasmid template and 0.02 U/µl Phusion Hot Start II DNA polymerase. Initial denaturation was achieved by heating the sample to 98°C for 30s, followed by 25 cycles of denaturation (98°C, 10 s), annealing (68-72°C, 30 s), and elongation (72°C, 3 min 20 s) and a final elongation at 72°C for 7 minutes.

The plasmid was then ligated using 25 ng of the PCR product in Quick Ligation buffer containing T4 DNA ligase at 25°C for 10 min. Transformation into *E. coli* TOP10 cells for selection of clones was performed as described (Chapter 2.6.2).

2.7.3 Cloning strategies II: In-Fusion PCR cloning

The genes of interest, *DnaK*, *Ssc1*, and *BiP*, were introduced into a different vector (pBAD-intein-12His) containing a C-terminal histidine-tag using the In-Fusion Advantage PCR cloning reaction (Clontech Laboratories, Mountain View, USA). The gene of interest was amplified by a PCR reaction using primers with approximately 15 bases overhang homologous to the site of entry in the pBAD vector. Initial denaturation was achieved by heating the sample to 95°C for 3 min, followed by 35 cycles of denaturation (95°C, 45 s), annealing (61°C, 45 s), and elongation (72°C, 2 min) and a final elongation at 72°C for 5 minutes. The final vector pBAD was linearized by restriction digestion with NcoI at 37°C for 40 min. NcoI recognizes and cuts the C↓CATGG site. The linearized vector and the amplified gene of interest were purified using a PCR purification protocol (QIAquick PCR purification kit protocol, Qiagen). The In-Fusion reaction leading to homologous recombination between the DNA ends of the vector and the insert was performed using 100 ng of purified PCR fragment, 100 ng of linearized vector, and an In-Fusion enzyme mixture. The reaction proceeded at 37°C for 15 min followed by 50°C for 15 min and was stopped on ice. The reaction was diluted in TE buffer (10 mM Tris, 1 mM EDTA, pH 8) before 5 µL of the reacted sample were used for transformation of competent *E. coli* (Chapter 2.6.2).

2.7.4 Restriction digestion and detection of DNA fragments

To determine the correct insertion of the desired insert into the vector, colonies of *E. coli* were grown in liquid media and their plasmid DNA was amplified and purified (Chapter 2.6.3). A restriction digestion of the purified plasmid DNA was performed using NcoI FastDigest enzyme in FastDigest buffer (Thermo-Fisher Scientific) by incubation at 37°C for 20 min. DNA fragments were detected as described (Chapter 2.6.3) (Figure 2.1) and plasmids with correctly inserted inserts were sent for sequencing (Chapter 2.7.5).

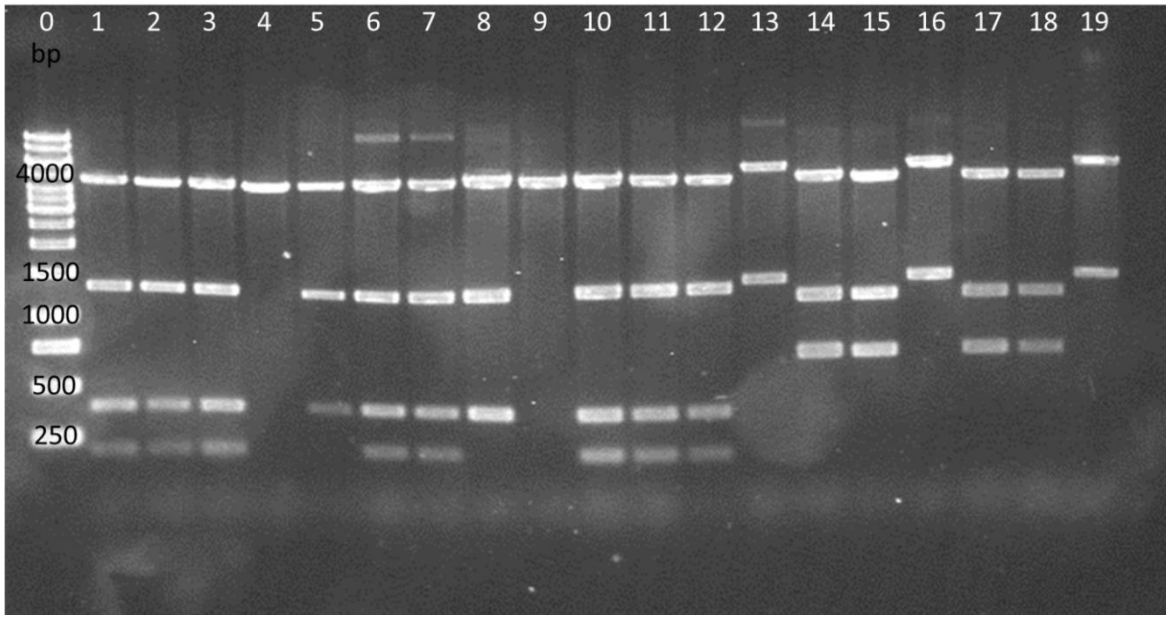


Figure 2.1 Exemplary agarose gel of DNA fragments obtained from restriction digestion of DNA from *E. coli* colonies 1-19. Lane 0: 1 kb DNA ladder. Lanes 1-4: BiP-C167-C519-C638. Lanes 5-8: BiP-UAA167-C519-C638. Lanes: 9-12: BiP-C167-UAA519-C638. Lanes 13-16: DnaK-C425-C563. Lanes 17-19: DnaK-C318-UAA425-C563.

2.7.5 Sequencing of DNA

DNA fragments were sequenced using fluorescence-based Sanger sequencing at the Sequencing Service of the Genomics Service Unit (Fakultät für Biologie, Ludwig-Maximilians-Universität München). Samples were prepared in a total volume of 7 µl 10 mM Tris, HCl (pH 8.5) including 20-40 ng of the DNA fragment and 0.7 µM of a sequencing primer with annealing temperature of 52-60°C. Fragments were analysed regarding their sequence using the pairwise sequence alignment tool EMBOSS Needle of the EMBL-EBI and regarding their sequence strength using Chromas Lite (Technelysium Pty Ltd, South Brisbane, Australia).

2.7.6 Hybridization of complementary nucleotide strands

Single-stranded complementary nucleotide strands (RNA or DNA) were hybridized at concentrations of 100 µM in annealing buffer (10 mM Tris, 100 mM NaCl, 1 mM EDTA, pH 7.5) by denaturation at 95°C, initial hybridization at 62°C, and cooling at -0.6°C/min.

2.8 Biochemical Methods

2.8.1 Detection of proteins on denaturing polyacrylamide gels

For SDS-PAGE under denaturing conditions, a separating gel (12.5% gels: 4.7 mL 30% acrylamide, 2.8 mL 1.5 M Tris/HCl pH 8.8, 3.75 mL H_2O, 13 µl TEMED, 45 µl 10% APS) and a stacking gel (4.5% gels: 0.8 ml 30% acrylamide, 1.2 mL 0.5 M Tris/HCl pH 6.8, 2.9 mL H_2O, 6 µl TEMED, 30 µL 10% APS) were poured succeedingly.

Samples prepared with 5x Laemmli buffer (40% glycerine, 9% SDS, 250mM Tris pH 6.8, bromophenol blue, 20% β-mercaptoethanol) and heated to 95°C as well as prestained protein

marker were loaded. For visualization of proteins, polyacrylamide gels were stained with Coomassie staining solution (80 mg Coomassie G250 in 1 L H_2O, 35 mM HCl).

2.8.2 Electromobility Shift Assay

Electromobility Shift Assays (EMSA) were performed to detect interaction of nucleic acids and proteins. Native polyacrylamide gels (6% acrylamide/bisacrylamide, 0.025% glycerol, 500 µM DTT, 10 mM magnesium acetate, 0.09% APS, 0.0009% TEMED in TG buffer) were prepared and pre-run in the absence of proteins in TG buffer (25 mM Tris, 192 mM glycine, 1 mM EDTA, pH 8.3) at 100 V for 1 hour. 5 pmol of DNA mixed with increasing concentrations of DNA-binding protein in 20% glycerol were loaded onto the gel and run at 150 V and 4°C for 1 h. The gel was imaged on a gel scanner (Typhoon, GE Healthcare) with excitation wavelengths specific for the fluorophores attached to the DNA and to the protein of interest.

2.8.3 Recombinant expression and purification of TBP

TBP was recombinantly expressed using a pET-15b vector in *E. coli* BL21-DE3-RIL cells by growing a preculture of 100 mL LB (1% (w/v) tryptone/peptone, 0.5% (w/v) yeast extract, 0.09 M NaCl) containing 100 µg/mL ampicillin and 33 µg/mL chloramphenicol at 30°C overnight. Autoinducing medium was prepared by addition of NPS buffer (final concentrations: 25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM Na_2HPO_4), 5052 solution (final concentrations: 0.5% glycerol, 0.05% glucose, 0.2% α-lactose), and 1 mM $MgSO_4$ to ZY medium (1% (w/v) trytone, 0.5% (w/v) yeast extract) with a final volume of 2 L. The preculture was used to inoculate bacteria in 2 L autoinducing medium containing the respective antibiotics. The bacteria were cultured at 30°C for 3 h and at 18°C for an additional 18 h. Cells were harvested by centrifugation at 4500 rpm and 4°C for 20 min and resuspended in 30 mL lysis buffer (25 mM Tris, 500 mM NaCl, 50 mM ammonium acetate, 10% glycerol, pH 8 (acetic acid)) containing 20 mM TCEP and 20 mM PMSF. Cells lysis was achieved by sonication (duty cycle 30%, output 8) on ice for 20 min. Soluble protein was separated from cell debris by centrifugation at 16000 rpm and 4°C for 45 min. For purification by affinity chromatography via the histidine-tag, the supernatant was incubated with 3-4 mL Ni-NTA suspension at 4°C for 1 h while rotating in a polypropylene column. The Ni-NTA column was then washed with 10 column volumes lysis buffer, followed by 4 column volumes wash buffer (25 mM Tris, 500 mM NaCl, 50 mM ammonium acetate, 10% glycerol, 5 mM imidazole, pH 8) and 4 column volumes lysis buffer. Proteins containing a histidine tag were eluted by addition of 2 column volumes elution buffer (25 mM Tris, 500 mM NaCl, 50 mM ammonium acetate, 10% glycerol, 250 mM imidazole, pH 8). After equilibration of a heparin column with low-salt buffer buffer (20 mM Tris, 250 mM NaCl, 50 mM ammonium acetate, 10% glycerol, pH 8) in the fast protein liquid chromatography system (ÄKTA, GE Healthcare), eluted proteins were diluted with equal volumes of low-salt and high-salt buffer (20 mM Tris, 1 M NaCl, 50 mM ammonium acetate, 10% glycerol, pH8), filtered, and loaded onto the heparin column. The protein was eluted from the column by a 30 column volume gradient from low-salt to high-salt buffer. The TBP elution was found in fractions containing

600-800 mM NaCl. The TBP containing fractions were chosen by detection on SDS-PAGE, combined and purified in a third step by gel filtration. To this end, a superose 6 gel filtration column was equilibrated with TBP buffer (5 mM Hepes, 40 mM ammonium sulfate, 10 µM $ZnCl_2$, 20% glycerol, pH 7.3), the TBP containing fractions were loaded onto the column, and proteins were separated according to size. Eluted fractions containing TBP (determined by SDS-PAGE) were again combined and concentrated in Amicon ultracentrifugation columns. TBP was either directly fluorescently-labeled (Chapter 2.8.9) or frozen in liquid nitrogen and stored at -80°C.

2.8.4 Recombinant expression and purification of DnaK

DnaK was purified using affinity chromatography to the nucleotide exchange factor Mge1 from *S. cerevisiae*. Therefore, Mge1-His was expressed recombinantly in a first step. A preculture of pQU60-mGrpEΔ43 containing *E. coli* BL21-DE3-RIL was inoculated in 50 mL TB medium (1.2% tryptone, 2.4% yeast extract, 0.5% glycerol, 89 mM potassium phosphate buffer) containing 100 µg/mL ampicillin and incubated at 37°C with continuous agitation overnight. The preculture was used to inoculate 2 L of culture in TB medium. Protein expression was induced by the addition of 1 mM IPTG after approximately 45 min of incubation at 37°C (OD_{600}(*E. coli*)=0.5). The bacteria were grown for an additional 4 hours at 37°C before being harvested by centrifugation at 4000 rpm for 20 min. Cells were resuspended in 10 mL Mge1 lysis buffer (50 mM Tris, 250 mM KCl, 5 mM $MgCl_2$, 5% glycerol, 0.2 mM TCEP, 5 mM imidazole, 1 mM PMSF) and lysed by sonication (duty cycle 50%, output 3). Soluble protein was separated from cell debris by centrifugation at 16000 rpm and 4°C for 1 h. Mge1 was bound to equilibrated Ni-NTA sepharose via its histidine-moiety by constant agitation at 4°C for 90 min. The Ni-NTA column was washed with 10 column volumes lysis buffer, 10 column volumes wash 1 buffer (lysis buffer containing 10 mM imidazole), 2 column volumes elution buffer (lysis buffer containing 10 mM imidazole, 2 mM ATP), 5 column volumes wash 1 buffer, and 10 column volumes wash 2 buffer (lysis buffer containing 20 mM imidazole) and stored at 4°C. Before use, the Mge1-Ni-NTA column was again equilibrated with lysis buffer.

DnaK was recombinantly expressed in a pET-Duet1 vector in *E. coli* BL21-DE3-RIL cells or in BL21AI cells cotransfected with DnaK-PrK-pET-Duet1 and pEvol $tRNA^{Pyl}$ $PylRS^{WT}$ for introduction of a non-natural amino acid (Chatterjee et al., 2013; Milles et al., 2012a; Plass et al., 2011; Tyagi and Lemke, 2013, 2015). A preculture of bacteria inoculated in 50 mL LB containing either 100 µg/mL ampicillin (DnaK-pET-Duet1 expression without non-natural amino acid) or a combination of 50 µg/mL ampicillin and 33 µg/mL chloramphenicol (expression of DnaK-PrK-pET-Duet1) was grown overnight. Further amplification was achieved by incubation of the preculture in 1 L LB medium containing antibiotics for approximately 45 min. For expression of DnaK with non-natural amino acids, 1 mM N-propargyl-lysine in 0.1 M NaOH was added to the culture. The inducing agents 1 mM IPTG and 0.02% L-(+)-arabinose (for BL21AI only) were added when the bacteria reached an OD_{600nm} of 0.5. Cells were harvested as described and lysed in 5 mL Mge1 lysis buffer. After

separation from cell debris, the soluble protein was incubated on Ni-NTA sepharose at 4°C for 70 min. The unbound fraction of proteins was directly loaded onto a Mge1-Ni-NTA column and incubated by rotation at 4°C for 1 h. The column was then washed with 5-10 column volumes Mge1 lysis buffer and 30 column volumes wash 1 buffer. DnaK was eluted in 3 column volumes elution buffer, concentrated and stored in storage buffer (20 mM Hepes, 100 mM KCl, 5 mM $MgCl_2$, 5% glycerol, pH 7.4).

2.8.5 Recombinant expression and purification of Ssc1

The recombinant expression of Ssc1 was performed in *E. coli* BL21-DE3-RIL in the presence of Hep1 for its stabilization by coexpression of both proteins in the vector pET-Duet1. When Ssc1 constructs for the introduction of non-natural amino acids were expressed, *E. coli* BL21-AI cells were used and the Ssc1-PrK-Hep1-pET-Duet1 construct was coexpressed with pEvol $tRNA^{Pyl}$ $PylRS^{WT}$ (Chatterjee et al., 2013).

Purification of Ssc1 was performed as described for DnaK (Chapter 2.8.4). An exemplary SDS-PAGE showing the protein fractions during Ssc1-PrK purification is shown in Figure 2.2. Hep1 containing a histidine tag was removed from the eluate by affinity chromatography to the Ni-NTA sepharose. Similar to DnaK, Ssc1 was then purified by an affinity chromatography to its nucleotide exchange factor Mge1.

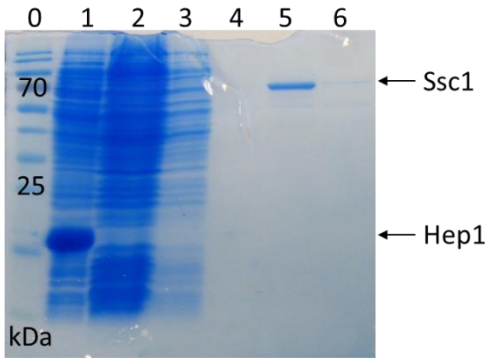


Figure 2.2 Exemplary SDS-PAGE of different protein fractions during an Ssc1-PrK purification. Lane 0: Protein ladder. Lane 1: Proteins bound to Ni-NTA sepharose beads including Hep1; first purification step to remove Hep1 from the eluate. Lane 2: Flow-through of proteins without affinity for Mge1 from an Mge1-Ni-NTA column. Unspecifically bound proteins were removed from the Mge1-Ni-NTA by washing with Mge1 lysis buffer (lane 3) and wash 1 buffer (lane 4). Ssc1-PrK was eluted using elution buffer (lane 5). An additional wash with wash 2 buffer did not elute any additional protein (lane 6).

2.8.6 Recombinant expression and purification of BiP

BiP-pProEx or BiP-PrK-pProEX was recombinantely expressed in *E. coli* BL21-DE3-RIL or BL21-AI (for coexpression with pEvol $tRNA^{Pyl}$ $PylRS^{WT}$ plasmid), respectively (Chatterjee et al., 2013). An overnight preculture of the bacteria in TB medium was used to inoculate 1 L of TB medium containing antibiotics suitable to the specific vector. For the expression of protein with non-natural amino acids, 1 mM N-propargyl-lysine was added after an initial incubation at 37°C for approximately 30 min (OD_{600nm}(*E.coli*)=0.2). When the optical density reached a value of 0.5, 1 mM IPTG (and 0.02% L-(+)-arabinose for BL21-AI) was added as inducing reagent. The bacteria were grown under constant agitation at 37°C for 4 hours. Cells were harvested by centrifugation at 4000 rpm for 20 min and resuspended in 5 mL BiP lysis buffer

(50 mM Hepes, 300 mM KCl, 0.2 mM TCEP, 5 mM imidazole, 1 mM PMSF, pH 7.5). Sonication on ice (duty cyle 50%, output 3) released soluble protein, which was then separated from cell debris by centrifugation at 16000 rpm, 4°C for 1 hour. Equilibrated Ni-NTA sepharose was loaded with the supernatant containing soluble proteins and incubated at 4°C for at least 1 h. The column was washed with 5 column volumes BiP lysis buffer, 5 column volumes wash 1 buffer (BiP lysis buffer containing 10 mM imidazole), and 5 column volumes wash 2 buffer (BiP lysis buffer containing 25 mM imidazole). BiP was eluted by addition of 2 column volumes elution buffer (BiP lysis buffer containing 500 mM imidazole). In filtration columns, the buffer was exchanged and the protein was concentrated in HKM buffer (50 mM Hepes, 150 mM KCl, 10 mM $MgCl_2$, pH 7.5). The histidine-tag for affinity purification of BiP was removed by site-specific cleavage of the peptide sequence Glu-Asn-Leu-Tyr-Phe-Gln↓Gly using TEV protease (1:100 (w/w) protease:BiP, 1 mM DTT, HKM) by dialysis at 4°C overnight. The histidine-tag containing protease was removed by an additional affinity chromatography to Ni-NTA sepharose, where the unbound fraction of protein was collected and contained BiP without histidine-tag. After concentration of the protein in filter columns, further purification was done by size exclusion chromatography on a superose 6 gel filtration column equilibrated in HKM. The protein was then concentrated and stored in the same buffer.

2.8.7 Determination of protein concentration

Protein concentrations were determined by either using their extinction at 280 nm or by a biuret reaction with copper detection.

Absorption of the protein at 280 nm mainly by its aromatic amino acids was measured on a Nanodrop 1000 (Thermo-Fisher Scientific). With known extinction coefficients of the protein ($\varepsilon_{280\text{nm}}$) (Table 2.1), the concentration was calculated using the Lambert-Beer equation:

$$A = cd\varepsilon_{280\text{nm}}, \tag{2.1}$$

with the absorption A at 280 nm and the cuvette thickness d.

Table 2.1 Extinction coefficients of the proteins under investigation.

Protein	Approximated extinction coefficient $\varepsilon_{280\text{nm}}$ [1/(M·cm)]
Cys-HMGA1-Cys	128
EBNA1	24325
TBP	14565
Mot1	149060
NC2	10345
U2AF65(RRM1,2)	8940
DnaK	15993
Ssc1	15930
BiP	29005

The extinction coefficients were estimated from the amino acid sequence of the proteins.

For proteins with low extinction coefficients, the BCA Protein Assay (Thermo-Fisher Scientific) was used to measure the total protein concentration. The assay is based on the reduction of Cu^{2+} to Cu^{+} by the protein. Cu^{+} is then used for the reaction of bicinchoninic

acid. Its absorption was detected at 562 nm. The assay was performed according to the protocol of the manufacturer using a standard concentration series of bovine serum albumin (BSA) and multiple dilutions of the protein of interest for each individual concentration measurement. Protein concentration was then determined by comparison to the absorption of known BSA samples.

2.8.8 Fluorophore attachment via Click Chemistry

Fluorophores were attached to propargyl-lysine site-specifically by copper-catalyzed Click reactions (Figure 2.3). These [3+2] cycloadditions are described in Kolb et al. (2001). The reaction with the unnatural amino acid offers the advantage of placing the fluorophore at one very specific site in the protein. Furthermore, the small alkyne tag of the unnatural amino acid does not cause large structural changes and the reaction proceeds bioorthogonally not affecting any naturally occurring macromolecules (Chatterjee et al., 2013; Milles et al., 2012a; Tyagi and Lemke, 2013).

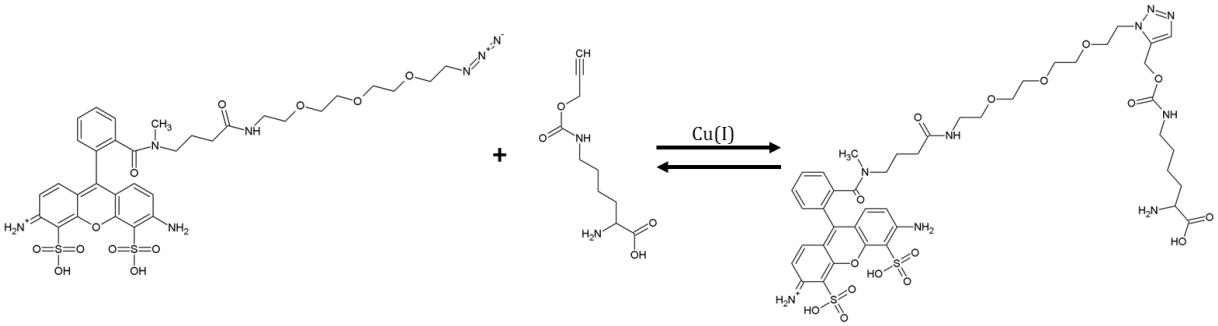


Figure 2.3 [3+2]cycloaddition of Atto488-azide and propargyllysine in the presence of copper. The azide moiety attached to the conjugated ring-system of the fluorophore by a linker reacts with the alkyne group of propargyllysine to form a 1,2,3-triazine ring structure.

The optimal concentrations of each reagent needed for the reaction were determined for each protein individually. 5 nmol BiP-PrK reacted with 25 nmol azide-coupled fluorophore (e.g. Atto488-azide) in the presence of 2 mM $CuSO_4$, 4 mM TBTA, 0.02 mM TCEP, 10 mM fresh sodium ascorbate and protease inhibitor cocktail (Roche) in HKM buffer at room temperature for 2 hours. DnaK-PrK and Ssc1-PrK were reacted in Hepes storage buffer by addition of 500 µM $CuSO_4$, 1 mM TBTA, 2.5 mM sodium ascorbate and fivefold molar excess of azide-coupled fluorophore.

The reaction was stopped by addition of 5 mM EDTA and excess dye was removed by repeated centrifugation in centrifugal filters with a cutoff of 10 kDa (Corning, Sigma-Aldrich).

2.8.9 Fluorophore attachment via sulfhydryl-maleimide chemistry

The attachment of fluorophores to cysteine residues was performed via sulfhydryl-maleimide chemistry (Figure 2.4). This reaction places the fluorophores specifically on cysteine residues in a protein. Using this amino acid is advantageous since proteins usually only carry a few natural cysteines (2.3% of all amino acids in mammals), which can be replaced by other polar amino acids under certain circumstances. Additional cysteines can thus be placed within the protein of interest to site-specifically choose the position of the attached fluorophores.

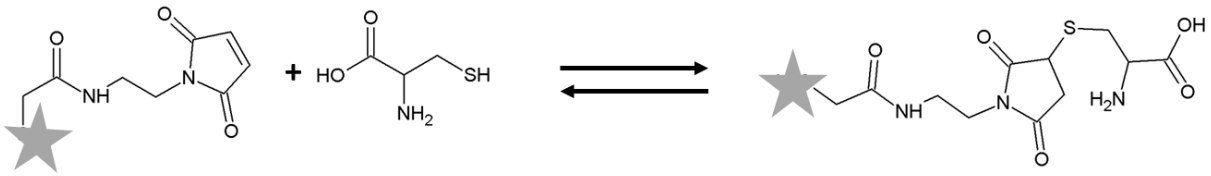


Figure 2.4 Sulfhydryl-maleimide chemistry. A maleimide moiety attached to a fluorophore (gray star) reacts with the sulfhydryl group of a cysteine at pH 6.5-7.5.

Cysteines were reduced by incubating the proteins in their respective storage buffer (adjusted to pH 7-7.5) containing 1 mM DTT. The reducing agent was removed by buffer exchange into oxygen-free storage buffer containing 50 µM TCEP. After removal of the reducing agent, the labeling reaction was performed by the addition of two- to threefold molar excess of fluorophore-maleimide conjugates (Atto488-Maleimide, Atto532-Maleimide, Atto565-Maleimide, Atto647N-Maleimide (AttoTec, Siegen, Germany), Alexa647-Maleimide, Cy5-Maleimide (Life Technologies, Carlsbad, United States)) in storage buffer (50 µM TCEP, deprived of oxygen, pH 7-7.5) at room temperature for two hours. The attachment of fluorophores to the protein was confirmed on a denaturing polyacrylamide gel using UV-excitation. Unbound fluorophores were removed by a size-dependent filtration with storage buffer containing DTT. The labeling efficiency was determined from the absorption of each of the fluorophores and the protein concentration determined photometrically by a bicinchoninic acid assay. Fluorescence correlation spectroscopy (FCS) measurements validated the labeling and the removal of unbound fluorophores.

2.8.10 ATPase activity assay

For the Hsp70 chaperones, the correct functioning after recombinant expression, purification, and labeling could be confirmed using an ATPase activity assay. To this end, a coupled ATP-NADH reaction assay was performed at room temperature and the absorption of NADH was followed (Figure 2.5). The rate of NADH oxidation is linearly coupled to the hydrolysis of ATP. 2 µM of the respective protein were mixed with 3 mM phosphoenol pyruvate, 50 µM NADH, and 60 units/mL pyruvate kinase/lactate dehydrogenase in measurement buffer. 1 mM ATP was added during the measurement after an initial reading in the absence of ATP. The absorption at 340 nm was measured every minute for up to 1 hour.

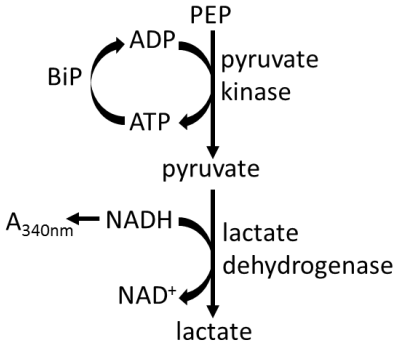


Figure 2.5 Schematic representation of the ATPase activity assay reactions. After hydrolysis of ATP by the protein of interest, pyruvate kinase uses phosphoenol pyruvate (PEP) and ADP as a substrate for the reaction to pyruvate and ATP. ATP is therefore reconstituted. This back-reaction is then coupled to the oxidation of NADH by lactate dehydrogenase. The decreasing absorption by NADH at 340 nm is followed in a spectrophotometer.

2.8.11 Protein digestion for mass spectrometry

The Hsp70 chaperones were digested into small peptides for mass spectrometry (Chapter 2.8.12). 1 µg of protein was denatured by addition of 6 M urea, 2 M thiourea, and 50 mM dithiothreitol (DTT) and incubation at 20°C for 2h. For the reaction of free sulfhydryl groups, the alkylating agent iodoacetamide (IAA) was added to a concentration of 50 mM and incubated for 30 min. Afterwards, 50 mM DTT were added and the sample was again incubated for 30 min at room temperature. Urea was diluted to a concentration below 1 M by adding 50 mM triethylammonium bicarbonate (TEAB). Trypsin digestion was performed overnight at 25°C by addition of 1:100 (trypsin/protein mass). Trifluoroacetic acid (TFA) was then added to a final concentration of 1% and the samples were lyophilized by vacuum centrifugation in aliquots of 0.5 µg and afterwards resuspended in 10% TFA.
Alternatively, the protein was digested from a polyacrylamide gel. The piece of gel containing the protein of interest was washed twice in water at room temperature for 10 min each. It was then incubated twice in 40 mM ammonium hydrogen carbonate for 10 min before the addition of 50% acetonitrile. Following the last 10 min incubation step, the piece of gel was incubated in 100% acetonitrile for 10 min and dried for an additional 5 min. The gel was incubated in 40 mM ammonium hydrogen carbonate, 10 mM DTT for 1 hour, followed by an incubation period in 40 mM ammonium hydrogen carbonate and 55 mM IAA for 30 min. The gel was washed three times with 40 mM ammonium hydrogen carbonate and twice with acetonitrile for 10 min each before drying it for 5 min. The digestion was performed with trypsin in 40 mM ammonium hydrogen carbonate at 37°C with constant agitation (600 rpm) overnight. The peptides were dissolved in 40 mM ammonium hydrogen carbonate (50 µL), 50% acetonitrile, and 0.1% TFA (100 µL). 100% acetonitrile (100 µL) was added before evaporation of the solvents. Digested samples were resuspended in 0.1% TFA.

2.8.12 Mass spectrometry (MALDI and LC-MS/MS)

Mass spectrometry of the digested protein sample was performed after separation on a C18 peptide purification column. The purification columns were prepared in pipette tips. The material was conditioned in methanol. Wetting was done using 0.1% TFA, 80% acetonitrile. The columns were then equilibrated in 0.1% TFA and the sample was loaded twice to increase the yield. The column was washed twice with 0.1% TFA and the peptide sample was eluted by three times adding 0.25% TFA, 80% acetonitrile. The solvent was evaporated by vacuum centrifugation and resuspended in 50% acetonitrile, 0.1% TFA saturated with α-cyano-4-hydroxycinnamic acid. Resuspended samples were loaded onto a stainless steel target plate and air-dried for MALDI-TOF or used for LC-MS/MS. For MALDI-TOF experiments, proteins were ionized in a first step by a laser and then their time of flight (TOF) was analyzed (Voyager-DE STR, Zentrallabor für Proteinanalytik, München, Germany). LC-MS/MS experiments were performed by Ignasi Forné (Zentrallabor für Proteinanalytik, München, Germany), Roman Körner (Professor Franz-Ulrich Hartl, Max-Planck-Institut für Biochemie, München, Germany), or Marina Daake (Professor Johannes Buchner, Technische Universität München, Germany).

2.8.13 High pressure liquid chromatography

High pressure liquid chromatography (HPLC) was used for the separation of peptides of ArgC digested Hsp70 chaperones. The chaperones were digested by incubation with 8 ng ArgC (in incubation buffer: 50 mM Tris, 5 mM $CaCl_2$, 2 mM EDTA, pH 7.8) and 3 mM DTT in activation buffer (5 mM Tris, 5 mM DTT, 2 mM EDTA, pH 7.8) at 37°C for 3 hours. The peptides were separated on a 6.5% Tris/Tricine gel at 100 V. HPLC of the protein fragments was performed by Wolfgang Rödl (Professor Ernst Wagner, Ludwig-Maximilians-Universität, München, Germany). The HPLC system (Waters, P-900 gradient pump, 996 photodiode array detector) equipped with a C18 column was used for the analysis of protein fragments separated by a water/acetonitrile gradient containing 0.1% TFA at the wavelengths 280 nm, 536 nm, and 642 nm.

2.9 Preparation of samples for fluorescence microscopy

2.9.1 Preparation of teflon sample holders

Teflon sample holders were used for fluorescence correlation measurements in small sample volumes down to 5 µL (Chapter 2.10.1). The sample chambers were prepared by a cleaning procedure of boiling and sonication in 2% non-fluorescent detergent (Hellmanex II, Hellma Analytics, Müllheim, Germany) followed by the same steps in ultrapure water. Coverslips were attached to the teflon holders by placing them on a ring of melted paraffin surrounding the sample chamber. Covering the chamber by sticky tape after the placement of the sample prevented evaporation of the solution.

2.9.2 Preparation of quartz prisms

Quartz prisms were used as the reflective surface as well as the sample chamber for total internal reflection fluorescence experiments. Thus, they needed to be cleaned and passivated for sample adsorption on the surface (Schluesche, 2007; Schluesche et al., 2007). To this end, the prisms were first washed with ultrapure water, and then cleaned with ethanol and acetone. They were then again washed by alternatingly boiling in water and non-fluorescent detergent (2% Hellmanex II). The quartz surface was prepared for reactions by placing the prisms in a mixture of hydrogen peroxide and concentrated HCl for 1 hour. The prisms were again washed by boiling in water and detergent and, after a final wash with water, dried under a stream of nitrogen. Any additional dust was removed by flaming the surface of the prisms. Silanization of the surface was achieved by addition of 2% (3-aminopropyl)triethoxysilane in acetone and incubation for 30 min. The prisms were then again rinsed with water and dried using nitrogen. A mixture of polyethylene glycol (PEG-5000-SVA) and biotinylated PEG (3-4%) in 0.1 M sodium bicarbonate buffer (pH 9.0) was incubated on the silanized surface of the prisms for 30 min. A layer of PEG prevents unspecific binding of proteins to the surface, while biotin allows for the specific attachment of molecules of interest using biotin-streptavidin-biotin linkages. After attachment of the PEG layer, prisms were rinsed with water and dried under a stream of nitrogen. The fluorescence background was reduced by photobleaching of the prism surface using a UV lamp (254 nm and 366 nm).

Coverslips were prepared by washing and boiling with water, detergent, and again water. They were then rinsed with water and dried using nitrogen. The surface of the coverslips was cleaned using the flame of a Bunsen burner.
Sample channels prepared from Nescofilm by laser cutting technology (laser power P29, V5, Technische Universität Garching, Germany) were placed on the prism surface and covered by a coverslip. This sample chamber was heated to melt the Nescofilm and attach it to both surfaces by gently pushing on top of the chamber. These surface functionalized prism chambers were stored at -20°C until use.

2.9.3 Preparation of flow chambers for dissociation experiments

To analyze the dissociation behavior of proteins from nucleic acid chains attached to the surface, prisms were prepared as sample chambers as described (Chapter 2.9.2). Simple flow channels were prepared by fixation of a tubing with 580 µm inner diameter to the openings of the sample chamber using pipette tips. The tubing at the entrance point of the channel was connected to a liquid reservoir, while the tubing exiting the sample chamber was connected via a needle to a syringe. Addition of a sample during the course of the measurement was thus possible by pulling in a solution containing the molecule of interest without defocussing of the chamber surface.

2.9.4 Encapsulation of proteins in lipid vesicles

Lipid vesicles were used to attach proteins to the surface for prism-type total internal reflection fluorescence measurements. A mixture of 10 mg/mL 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 250 ng/mL biotinylated phosphatidylethanolamine was prepared in chloroform and divided into volumes of 30 µL in reaction tubes. After evaporation of the solvent under a stream of nitrogen, the lipids covered the sides of the reaction glasses and were completely dried in a vacuum chamber for 2 hours. The lipid mixture was then stored under dry conditions at -20°C until use. For the preparation of protein containing lipid vesicles, the protein of interest was diluted to a final concentration of approximately 400 nM to encapsulate on average one protein per liposome. Additional unlabeled cofactors or binding partners were added at concentrations higher than their binding affinity (e.g. 1 mM ATP/ADP, 5 or 20 µM RNA, 1 µM cofactor) to a total volume of 300 µL in the measurement buffer of the respective protein. The protein suspension was then added to the dry lipid mixture and incubated at room temperature for 30 min under constant agitation. During the incubation period, the lipids were dissolved into the suspension as micelles by careful mixing on a vortex shaker every 10 min. An extruder was prepared by washing with detergent and ultrapure water. Wetting of the membrane was performed with water and equilibration was done in measurement buffer. Lipid vesicles with a defined diameter of 200 nm were formed by extrusion of the protein-lipid suspension. The salt concentration in the buffers inside and surrounding the lipid vesicles was kept constant to avoid any osmotic changes. Free protein was removed by multiple centrifugation steps in filter columns (Amicon Ultra) with a cutoff of 100 kDa. A 1:10 dilution of the lipid vesicles was then used for spFRET experiments on a total internal reflection fluorescence microscope (Chapter 2.10.3). The integrity of the protein

after encapsulation was simultaneously monitored by performing spFRET experiments of the vesicles in solution (Chapter 2.10.2).

2.9.5 Buffered solutions for microscopic measurements

For fluorescence microscopic experiments, the molecules of interest were diluted in their respective measurement buffers. Measurements with DNA were performed in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). HMGA1/EBNA1 experiments were performed in a buffered solution of 20 mM Hepes, 40 mM KCl, 0.1 mM EDTA, 0.5 mM $MgCl_2$, 1 mM DTT, 0.15 mg/mL BSA, and 0.5 mM ATP. TBP was diluted in 10 mM Tris, 10% glycerol, 0.25 mg/mL BSA, 2.5 mM DTT, 60 mM KCl, and 5 mM $MgCl_2$ at pH 8.2. U2AF was measured in a buffered solution of 20 mM KH_2PO_4, 20 mM K_2HPO_4, and 50 mM NaCl at pH 6.5. The Hsp70 chaperone BiP was diluted in 50 mM Hepes, 150 mM KCl, and 10 mM $MgCl_2$ at pH7.5 and Ssc1 and DnaK were measured in a solution of 20 mM Tris, 80 mM KCl, and 5 mM $MgCl_2$ at pH 7.5. Single-pair FRET experiments of Sigma32 were performed in 10 mM Hepes, 100 mM KCl, and 10% glycerol at pH 7.9.

2.10 Fluorescence Microscopy Methods

2.10.1 Fluorescence correlation spectroscopy

Fluorescence correlation spectroscopy (FCS) analyzes the intensity fluctuations of a sample to extract information about the fluctuation causing processes. It was first used to obtain information about the number and diffusion time of freely diffusing molecules (Elson, 2011; Elson and Magde, 1974; Haustein and Schwille, 2007; Krichevsky and Bonnet, 2002). When freely diffusing, fluorescently labeled molecules transverse the static focus of a confocal microscope, they emit photons after their excitation by a laser. Due to their diffusive behavior, the numbers of fluorescent molecules in the focus varies, leading to fluctuating intensities. In addition, intensity fluctuations do not solely occur due to the variable number of molecules in the focal volume, but they can also be caused by photophysical properties of the fluorophores such as singlet-triplet state interconversion, by rotational diffusion of the molecules, or conformational dynamics.

FCS experiments were performed here to investigate the diffusive behavior and the conformational dynamics of proteins. They were used to determine effective fluorophore-attachment to proteins of interest, or binding coefficients of cofactors. For a known point-spread function (PSF) of the microscope, additional information such as the diffusion coefficient and average molecule concentrations could be obtained. To this end autocorrelation functions of the intensity fluctuations over time were derived from a single detection channel:

$$G(\tau) = \frac{\langle \delta F(t) \cdot \delta F(t+\tau) \rangle}{\langle \delta F(t) \rangle^2} = \frac{\langle F(t) \cdot F(t+\tau) \rangle}{F(t)^2} - 1. \quad (2.2)$$

The PSF of the confocal microscope was approximated as a three-dimensional Gaussian with lateral and axial dimensions of ω_r and ω_z which give the distances from the center of the

point-spread function (PSF) to the position at which the intensity has decayed to the value of $\frac{1}{e^2}$. These dimensions were determined for each individual experiment using measurements of fluorophores with known diffusion coefficients (D(Atto655, 25°C)=426 µm²/s, D(Cy5, 25°C)=360 µm²/s) (Müller et al., 2008). The parameters for the point spread function were then determined by fitting the autocorrelation function where the individual diffusion coefficients were held fixed to their respective values. When the size of the PSF was known, the diffusion coefficient D was determined as

$$D = \frac{\omega_r^2}{4\tau_D}, \tag{2.3}$$

with τ_D being the average diffusion time of the molecules through the focal volume.
Autocorrelation functions were then fit with a one component diffusion model including intensity fluctuations due to an additional triplet state:

$$G(\tau)=\frac{\gamma}{N}\cdot\left(1+\frac{T}{1-T}\cdot e^{-\frac{\tau}{\tau_T}}\right)\cdot\frac{1}{1+\frac{4D\tau}{\omega_r^2}}\cdot\frac{1}{\sqrt{1+\frac{4D\tau}{\omega_z^2}}}+y_0, \tag{2.4}$$

where N is the number of particles in the focus, D the diffusion coefficient, T the fraction of photophysical fluctuations (triplet fraction), τ_T the triplet time, and y_0 is an offset to correct for small deviations of the baseline. Errors of the diffusion coefficients were estimated from the Jacobians of the fit.
Figure 2.6 shows exemplary FCS experiments of fluorophores and fluorescently-labeled proteins in solution. The amplitude of the autocorrelation function decreases for increasing fluorophore concentrations (Figure 2.6B). An increase in the diffusion time is observed as a shift of the autocorrelation function to longer timescales (Figure 2.6C).

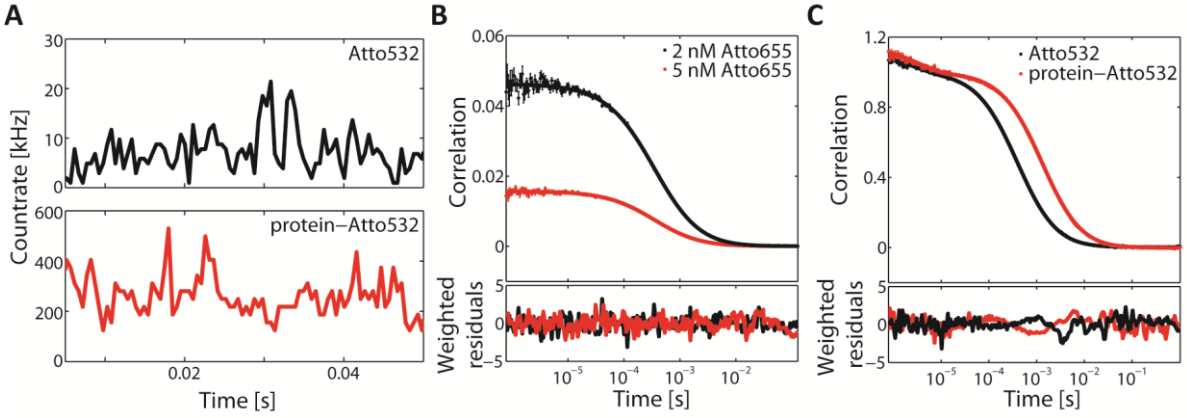


Figure 2.6 Fluorescence correlation spectroscopy. (**A**) Intensity time trace of a solution of Atto532 fluorophores (black) and of a solution of fluorescently-labeled protein (red). (**B**) Fluorescence correlation function of Atto655 fluorophores of different concentrations (2 nM in black, 5 nM in red). The amplitude of the correlation function is inversely proportional to the number of fluorescent molecules. (**C**) Fluorescence correlation function of Atto532 fluorophores (black) compared to fluorescently-labeled protein (red). The diffusion time of the protein is increased in comparison to the one of the free fluorophore.

When assuming a spherical shape of the molecules, the hydrodynamic radii can be calculated using the Stokes-Einstein relation:

$$D = \frac{k_B T}{6\pi\eta r_h} \tag{2.5}$$

with $T = 294$ K and $\eta = 9.8 \cdot 10^{-4}$ Pa·s. The diffusion time or diffusion coefficient contains information about the labeling efficiency, possible aggregation of molecules, or the interaction with binding partners. For experiments with multiple species, models with multiple diffusing components were used for fitting the data and therefore binding rates could be determined by comparing the fractions of two different species in the presence of different concentrations of an unlabeled binding partner.

For the analysis of fluorescence cross-correlation spectroscopy (FCCS) data, two channels i and j were correlated:

$$G(\tau) = \frac{\langle F_i(t) \cdot F_j(t+\tau) \rangle}{\langle F_i(t) \cdot F_j(t) \rangle}. \tag{2.6}$$

FCCS contains information on the co-diffusive behavior of multiple fluorophores and thus interactions between different molecules (Bacia and Schwille, 2007; Foo et al., 2012; Schwille et al., 1997). The amplitude of the cross-correlation function, compared to the individual autocorrelation functions carries information on the number of dual-labeled complexes and the degree of interaction between the molecules. Figure 2.7 displays the autocorrelation and cross-correlation functions of green and red fluorophores and dual-labeled DNA. Small cross-correlation amplitudes introduced by crosstalk can be eliminated when using pulsed interleaved excitation (PIE) (Müller et al., 2005) and the crosstalk-free cross-correlation function can be used to determine the degree of dual-labeled complexes.

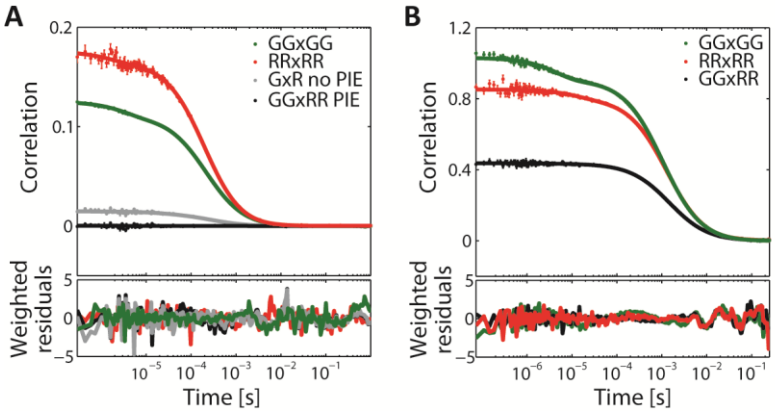


Figure 2.7 Fluorescence cross-correlation spectroscopy. (**A**) Fluorescence correlation function of a mixture of Atto532 and Atto655 fluorophores. The autocorrelation functions of Atto532 (green) and Atto655 (red) and cross-correlation function between Atto532 and Atto655 fluorophores with continuous green and red excitation (gray) or illuminated by PIE (black) are shown. The residual amplitude in the cross-correlation function of the mixture of independently diffusing fluorophores for continuous illumination results from crosstalk of the donor fluorophore into the acceptor channel, while crosstalk can be separated by alternating laser sources, filtering of the donor signal, and direct excitation of the red fluorophore using PIE. (**B**) Fluorescence correlation function of dual-labeled DNA. The autocorrelation functions of the Atto532 (green) and Atto647N (red) fluorophore and the crosstalk-free cross-correlation function with the use of PIE (black) are displayed.

The timescales of conformational dynamics were investigated using FRET-FCS, for which the signals of the donor and the acceptor fluorophores were auto- and cross-correlated. For molecules undergoing a conformational transition during a burst, which leads to a change in FRET efficiency, an anti-correlated signal in the cross-correlation function should be observed (Price et al., 2011; Torres and Levitus, 2007). The timescales of this switching

behavior were determined by fitting the correlation functions globally to a diffusion model with an additional term for the anti-correlated motion τ_t:

$$G(\tau) = \frac{1}{\sqrt{8}N}\left(1 - Ae^{\frac{-\Delta\tau}{\tau_t}}\right)\frac{1}{1+\frac{\Delta\tau}{\tau_D e^{-6}}}\cdot\frac{1}{\sqrt{1+\frac{1}{p^2}}}\cdot\frac{\Delta\tau}{\tau_D e^{-6}} + y_0. \quad (2.7)$$

The diffusion term was fit for a time lag τ between $2\cdot10^{-5}$ s and 1 s. For fixed values of the diffusion time, an additional triplet term was then fit for time lags beginning from $2\cdot10^{-7}$ s.

Experimental Section

For FCS experiments, proteins, DNA, or other molecules of interest were diluted in buffered solutions required for the individual molecules (Chapter 2.9.5) to concentrations in the range of around 1 nM and 25-50 µL of sample were added to the wells of LabTek #1.0 chamber slides (Thermo Scientific Nunc). FCS measurements with fluorophores in the red and green excitation range (Atto532, Atto647N, Atto647, Alexa647, Cy5) were performed for few minutes up to one hour using a home-built confocal microscope (Figure 2.8) (Kudryavtsev et al., 2012). The fluorophores were excited alternatingly by pulsed 532 nm (PicoTA 530 Picoquant, Berlin, Germany) and 640 nm lasers (LDH-D-C640, Picoquant, Berlin, Germany) at a repetition rate of 26.66 MHz and laser powers of 10-100 µW (measured at the entrance point of the beam into the objective). Pulsed interleaved excitation offers crosstalk-free detection of intensity fluctuations due to assignment of emitted photons to their excitation source (Hendrix et al., 2013; Müller et al., 2005). Fluorescence was collected by a 60x water immersion objective (Plan Apo IR 60x/1.27 WI, Nikon, Düsseldorf, Germany) and focused onto a pinhole with diameter d=75 µm for confocal detection, confining the focal volume to a few femtoliters. A polarizing beam-splitter was used to separate the fluorescence according to polarization (PBS3, Thorlabs, Dachau, Germany). The green fluorescence was then reflected and the red fluorescence was transmitted by a dichroic mirror (640DCXR, AHF Analysentechnik, Tübingen, Germany). After passing an emission filter (Brightline HQ582/75 and HQ700/75 for green and red photons respectively, AHF Analysentechnik, Tübingen, Germany), the signals were detected on four single-photon-counting avalanche photodiodes (Perkin Elmer, Hamburg, Germany). The data were recorded using time-correlated single-photon-counting cards (SPC-154, Becker and Hickl, Berlin, Germany) and data analysis was performed with a custom written Matlab program (The MathWorks, Ismaning, Germany) by Matthias Höller and Waldemar Schrimpf (Professor Don C. Lamb, München, Germany). Whenever FCS was measured on a protein fluorescently labeled with Atto488 or Fluor Azide 488 (Jena Bioscience), a confocal scanning microscope, described in Hendrix et al. (2013), was used.

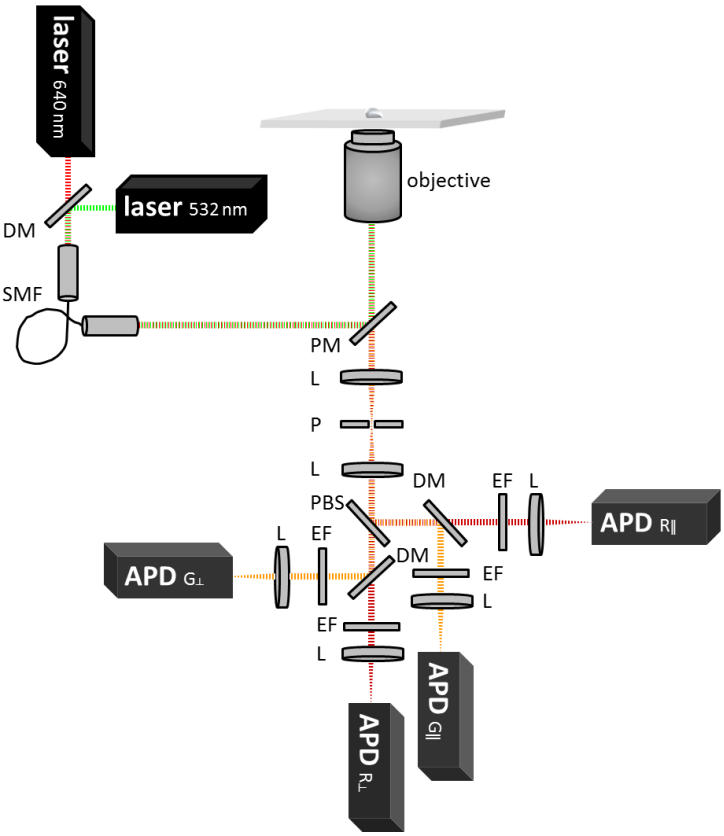


Figure 2.8 Schematic representation of the two-color confocal multiparameter fluorescence detection microscope. The sample is alternatingly excited by two pulsed lasers combined into a single-mode fiber (SMF). The fluorescence collected by the objective is separated from the excitation by a polychroic mirror (PM), focused onto the pinhole (P) by a lens (L) and separated according to polarization by a polarizing beam splitter (PBS). The fluorescence is further separated according to wavelengths by dichroic mirrors (DM) and after passing an emission filter (EF) focussed onto the detecting avalanche photodiodes (APD).

2.10.2 Single-pair and single molecule three-color FRET experiments in solution

When performing single-molecule FRET, information on subpopulations and individual molecules can be obtained that might be hidden in ensemble FRET experiments. SpFRET experiments in solution rely on the detection of individual bursts achieved by a small detection volume and low concentrations of the fluorescent species. In the analysis these photon bursts have to be separated from background noise and assigned to individual molecules. Different parameters such as the FRET efficiency were determined by burst analysis using either a software written in Matlab (by Matthias Höller, Wolfgang Kügel, Anders Barth, Waldemar Schrimpf, Professor Don C. Lamb, München, Germany) or a software written in Labview (by Volodymyr Kudryavtsev, Professor Claus Seidel, Düsseldorf, Germany) (Kügel, 2012). Bursts of photons were selected by an all-photon burst search (APBS) algorithm (Eggeling et al., 2001; Nir et al., 2006). Photons were regarded as belonging to a burst when their interphoton time (time between arriving photons) fell below a certain threshold distinguishing photons belonging to a burst from background photons. They were considered for further analysis when the local count rate obtained by a sliding time window exceeded a certain threshold. Bursts with a minimum number of 100 photons and a count rate of at least 10 photons after green and red excitation in a time window of 500 µs

were considered for further analysis. To further refine the selected species of molecules during data extraction, a dual channel or triple channel burst search (DCBS or TCBS respectively) algorithm was used. By choosing the minimum number of photons in each detection channel only dual- or triple-labeled molecules were selected for further analysis (Figure 2.9).

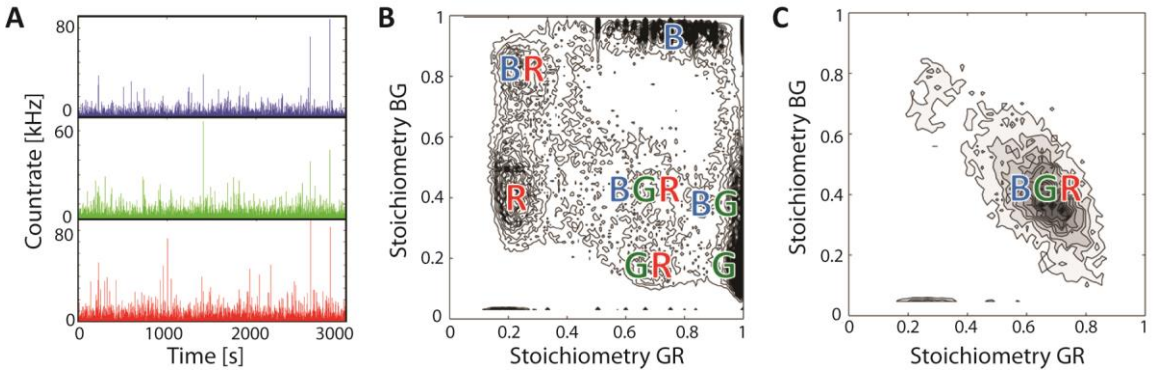


Figure 2.9 Burst analysis of single-molecule three-color FRET data. (**A**) Exemplary intensity traces from the three channels blue, green, and red extracted from an spFRET experiment in solution using PIE. (**B**) A two-dimensional stoichiometry histogram after an APBS of the data. Besides the triple-labeled population of molecules (BGR), singly-labeled (B, G, R) and dual-labeled populations (BR, GR, BG) of molecules are selected with this burst search algorithm. (**C**) A two-dimensional stoichiometry histogram after a TCBS of the data. Only molecules emitting in all three channels are selected by this burst search algorithm.

Since every molecule was excited by multiple excitation cycles (on average several hundred), multiple parameters could be calculated burstwise. The calculation of these parameters is described here for spFRET experiments. The proximity ratio or uncorrected FRET efficiency is defined as the number of acceptor photons (F_{DA}) divided by the the total number of photons after donor excitation:

$$E_{PR} = \frac{F_{DA}}{F_{DD} + F_{DA}}. \tag{2.8}$$

Similarily, the value for the stoichiometry describes the number of all photons after green excitation divided by all photons, after both, green and red, excitation:

$$S_{PR} = \frac{F_{DD} + F_{DA}}{F_{DD} + F_{DA} + F_{AA}}. \tag{2.9}$$

However, for a quantitative analysis, several correction factors have to be considered. The fluorescence background was determined from a measurement of the respective measurement buffer under the same conditions as in the actual experiment. Common intensity values for the background in the different detection channels were BG_{GG}=1.1 kHz, BG_{GR}=0.4 kHz, and BG_{RR}=0.5 kHz. Furthermore, the donor laser has a certain probability of exciting the acceptor fluorophore directly (direct excitation):

$$\alpha = \frac{F_{DA}^{A(0)}}{F_{AA}^{A(0)}}, \tag{2.10}$$

where $F_{DA}^{A(0)}$ and $F_{AA}^{A(0)}$ are the signals in the acceptor channel for an acceptor-only species after donor excitation or acceptor excitation, respectively. Additionally, the donor fluorophore can be detected in the acceptor channel (crosstalk):

$$\beta = \frac{F_{DA}^{D(0)}}{F_{DD}^{D(0)}} \quad . \tag{2.11}$$

Both values can be determined using the donor-only and acceptor-only population and the uncorrected stoichiometry and FRET efficiency from each experiment (Kudryavtsev et al., 2012):

$$\alpha = \frac{S_{PR}^{A(0)}}{1-S_{PR}^{A(0)}} \; , \beta = \frac{E_{PR}^{D(0)}}{1-E_{PR}^{D(0)}} \; . \tag{2.12}$$

For two-color experiments, an APBS was performed to determine the correction factors for direct excitation and crosstalk as shown in Equation 2.12. We usually obtained values of α=0.1 for the direct excitation and of β=0.03 for the spectral crosstalk. It is further necessary to account for the differences in detection efficiency η of the red and green detection channels as well as the differences in quantum yield Φ of the donor and acceptor fluorophore. The correction factor γ incorporates these variations in detection efficiency and quantum yield:

$$\gamma = \frac{\Phi_A \eta_A^{\lambda_{emA}}}{\Phi_D \eta_D^{\lambda_{emD}}} . \tag{2.13}$$

γ was determined by minimizing the differences in stoichiometry when multiple dual-labeled subpopulations were present or by taking into account the lifetime information of the donor. Typical values for this correction factor were approximately γ=0.7-0.85 (Atto647N as acceptor fluorophore) or γ=0.45-0.55 (Alexa647 as acceptor dye). The exact correction factors were determined for each measurement individually by analysis of the donor- and acceptor-only populations and background measurements with water and buffer. Using these, the corrected FRET efficiency and stoichiometry were obtained as follows:

$$E = \frac{F_{DA} - \alpha F_{AA} - \beta F_{DD}}{F_{DA} - \alpha F_{AA} - \beta F_{DD} + \gamma F_{DD}} , \tag{2.14}$$

$$S = \frac{F_{DA} - \alpha F_{AA} - \beta F_{DD} + \gamma F_{DD}}{F_{DA} - \alpha F_{AA} - \beta F_{DD} + \gamma F_{DD} + F_{AA}} . \tag{2.15}$$

As an example for the selection criteria of individual bursts see Figure 2.10. In a first step, single dual-labeled molecules were selected by removal of multimolecule, acceptor blinking, and photobleaching events using the mean macroscopic arrival time of the photons after donor and acceptor excitation. This value was chosen to lie between -0.1 ms < T_{DD+DA}-T_{AA} < 0.1 ms. Dual-labeled molecules were either selected by DCBS or by restricting the stoichiometry to values of 0.3 < *S* < 0.8.

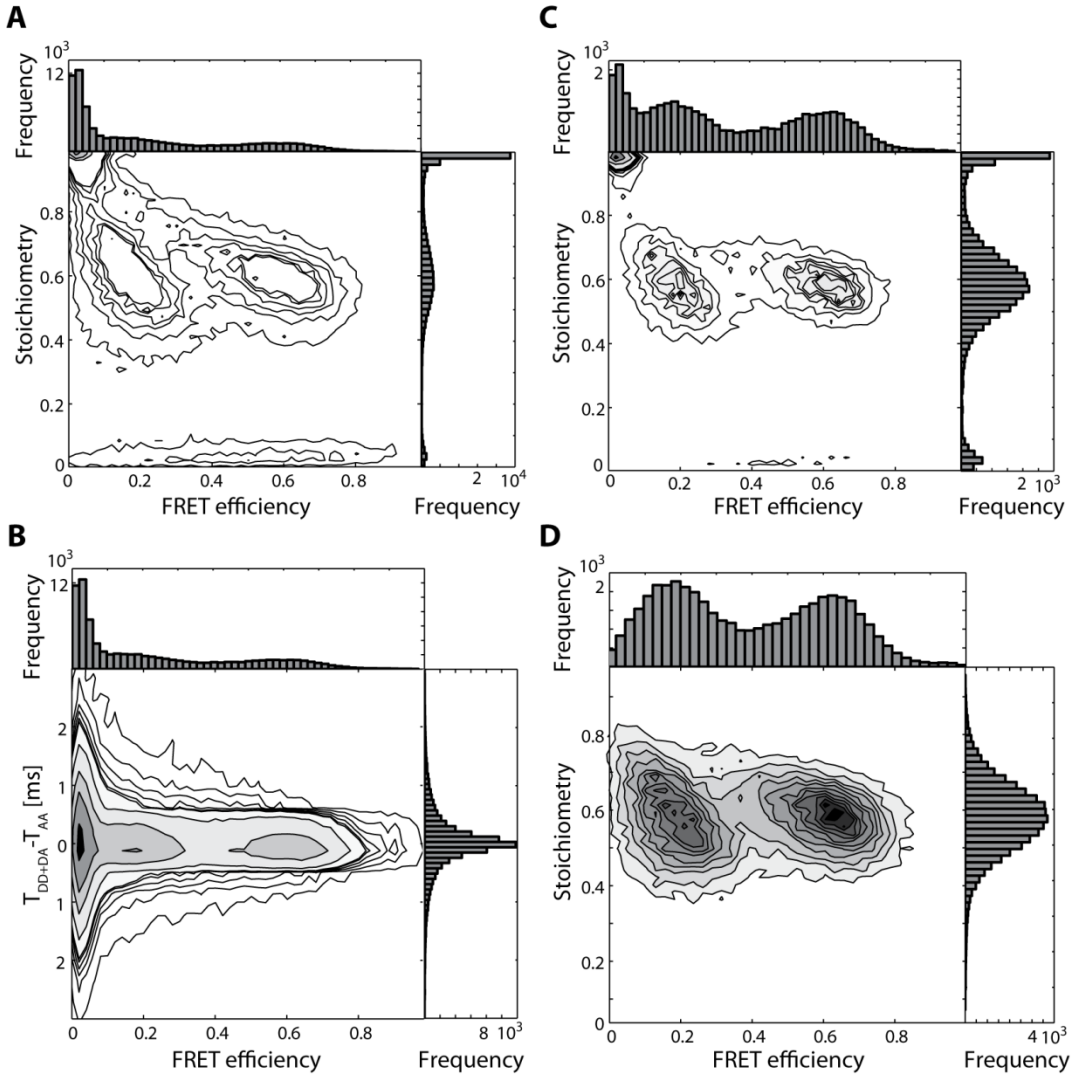


Figure 2.10 Selection criteria for an spFRET experiment of a mixture of two dual-labeled DNA species. (**A**) Histogram of FRET efficiency versus stoichiometry of molecules selected by an APBS. The donor-only and acceptor-only species are visible as well as two populations of double-labeled DNA molecules. Trailing from the dual-labeled population to the donor-only population can be caused by multimolecule events, acceptor blinking or photobleaching. (**B**) Histogram of FRET efficiency versus T_{DD+DA}-T_{AA}. A selection of bursts can be performed regarding the mean macroscopic arrival time of the photons after donor and acceptor excitation to remove molecules for which the acceptor photobleached during their residence time in the confocal volume. (**C**) Histogram of FRET efficiency versus stoichiometry as in (A) after removal of molecules using the selection criterion -0.1 ms < T_{DD+DA}-T_{AA} < 0.1 ms. Trailing populations between the dual-labeled and the donor-labeled species disappear. (**D**) Histogram of FRET efficiency versus stoichiometry after additional removal of molecules by selection of 0.3 < stoichiometry < 0.8 and correction of the populations for crosstalk, direct excitation and the detection efficiency of the different channels.

For three-color experiments, the FRET efficiencies were calculated correspondingly as described (Clamme and Deniz, 2005; Gambin and Deniz, 2010; Lee et al., 2010; Milles et al., 2012b) and corrected for background (BG_{BB}=2.38 kHz, BG_{BG}=1.21 kHz, BG_{BR}=0.38 kHz, BG_{GG}=1.03 kHz, BG_{GR}=1.29 kHz, BG_{RR}=0.36 kHz), crosstalk (β_{BG}=0.054, β_{BR}=0.005, β_{GR}=0.135), direct excitation (α_{BG}=0.129, α_{BR}=0.033, α_{GR}=0.104), and differences in detection efficiency (γ_{BG}=0.74, γ_{BR}=0.54, γ_{GR}=0.78).

The fluorescence lifetime was determined by a maximum likelihood estimator described in Kudryavtsev et al. (2012) and Schaffer et al. (1999). The decay of the fluorescence intensity was deconvolved with the experimentally determined instrument response function (Chapter 2.10.2) and fit by a single exponential decay. We used the relationship between the donor lifetime and the FRET efficiency to obtain information about conformational dynamics on the millisecond timescale (Figure 2.11A). Because the FRET efficiency is species-averaged and the lifetime is an intensity-weighted average, the FRET efficiency is related to the donor lifetime in the absence of dynamics by:

$$E_{\text{static}} = 1 - \frac{\tau_{\text{D}}}{\tau_{\text{D(0)}}}. \tag{2.16}$$

Taking into account the flexibility of the linkers used for attachment of the fluorophores to the protein, the static FRET line is described by an empirically determined third-order polynomial:

$$E = 1 - \frac{a + b\langle\tau_{\text{DA}}\rangle + c\langle\tau_{\text{DA}}\rangle^2 - d\langle\tau_{\text{DA}}\rangle^3}{\langle\tau_{\text{D(0)}}\rangle} \tag{2.17}$$

with $\tau_{D(0)}$=3.8 ns, and a=-0.0421, b=0.508, c=0.2691, d=0.03589 for the donor-acceptor pair Atto532-Atto647N or a=-0.0178, b=0.6226, c=0.2188, d=0.0312 for Atto532-Alexa647 (Gansen et al., 2009; Kudryavtsev et al., 2012).

Whenever a molecule undergoes dynamic conformational changes leading to a different FRET efficiency when transversing the focal volume, it will deviate from the static FRET curve due to the intensity-weighted average of the donor lifetime. Molecules deviating from this line can be described by a dynamic FRET curve, given here for a two-state system:

$$E_{\text{dynamic}} = 1 - \frac{\tau_1\tau_2}{\tau_{\text{D(0)}}\left(\tau_1 + \tau_2 - \langle\tau\rangle_f\right)}, \tag{2.18}$$

where τ_1 and τ_2 are the respective donor lifetimes of both FRET states between which the transition occurs.

The relationship between the donor and acceptor lifetimes and the FRET efficiency are displayed in Figure 2.11A. While the donor lifetime varies with FRET efficiency, the acceptor lifetime is not expected to change and can thus be used to obtain information about photophysical processes such as acceptor quenching.

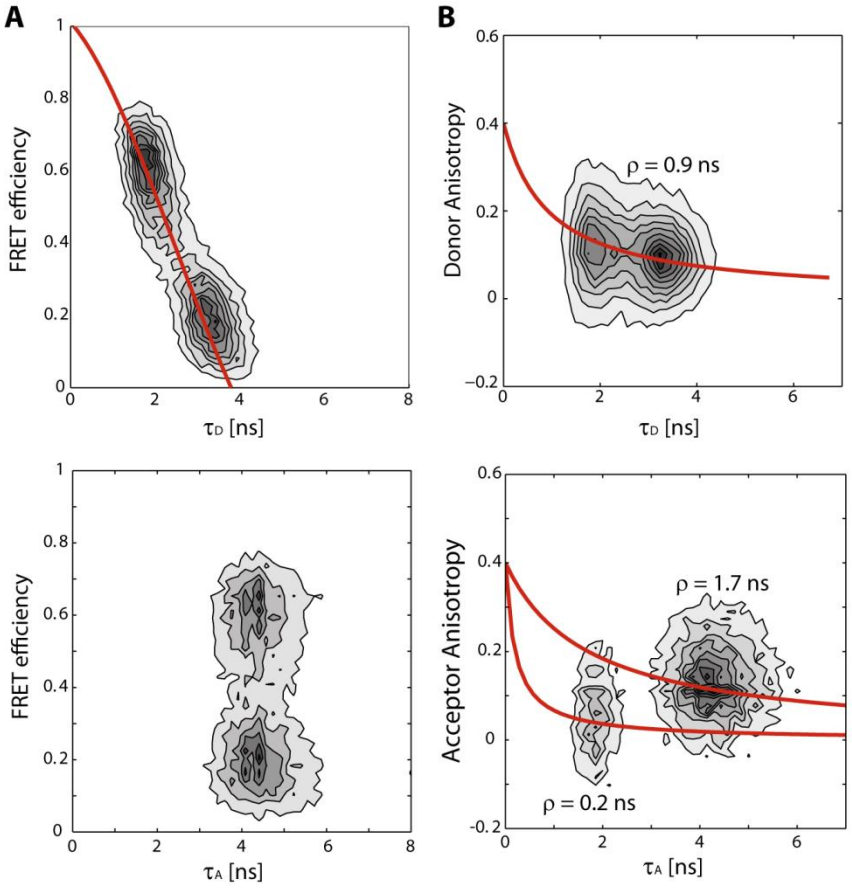


Figure 2.11 Histograms of additional parameters obtained from a MFD-PIE experiment on a mixture of two populations of dual-labeled DNA. (**A**) Two-dimensional histogram of FRET efficiency versus the lifetime of the donor (upper graph) or the acceptor (lower graph). The red line corresponds to the static FRET curve (Equation 2.16). Both populations of DNA clearly fall onto this line. The acceptor lifetime is not expected to change depending on the FRET efficiency. However, cases can arise, where the acceptor lifetime changes due to acceptor quenching. This histogram can thus be used to detect acceptor quenching to shorter lifetimes. (**B**) Histograms of anisotropy versus lifetime for the donor (upper graph) and the acceptor (lower graph) fluorophores. The populations can be described by Perrin equations with different rotational correlation times ρ (red line) (Equation 2.21). For the histogram of the acceptor anisotropy versus lifetime, a measurement of Atto532-Atto647N dual-labeled DNA was performed in the presence of free Atto655 fluorophores. Both acceptor species can be separated using the information obtained from these histograms.

As a fourth parameter, the fluorescence anisotropy was calculated from the ratio of photons detected in the parallel and perpendicular detection channels. As we corrected for the differences in detection efficiency of the different colors, a correction factor G was used for the polarization. The different detection efficiencies of the channels perpendicular ⊥ and parallel ∥ to the excitation were corrected by:

$$G = \frac{\eta_{\perp}}{\eta_{\parallel}}, \tag{2.19}$$

while the correction factors l_1 and l_2 accounted for the depolarizing effects of the objective (Koshioka et al., 1995):

$$r = \frac{GF_{\parallel} - F_{\perp}}{(1 - 3l_2)GF_{\parallel} + (2 - 3l_1)F_{\perp}}. \tag{2.20}$$

The anisotropy contains information on the rotational freedom of the fluorophores which is important for the calculation of the correct FRET efficiencies, when considering effects of κ^2,

as well as the rotational correlation time of the whole molecule, which can be of interest in binding studies. The relationship of the rotational anisotropy r(τ) to the fluorescence lifetime τ can be described by the Perrin equation

$$r(\tau)=\frac{r_0}{1+\frac{\tau}{\rho}}, \tag{2.21}$$

with the rotational correlation time ρ and the fundamental anisotropy r_0.
Figure 2.11B displays exemplary anisotropies for the donor and acceptor fluorophores of a mixture of dual-labeled DNA and an additional red fluorophore.

The underlying distance distributions of spFRET and single molecule three-color FRET experiments were calculated from proximity ratio histograms by photon distribution analysis (PDA) as described in Antonik et al. (2006). This analysis was extended to the description of three-color FRET experiments by Anders Barth (Professor Don C. Lamb, München, Germany; manuscript in preparation). Triple-labeled species were selected and their proximity ratios for all three combinations of fluorophores were used for photon distribution analysis. In a first step, the one-dimensional proximity ratio histogram for the green-red population was fit by binomial distributions including information about noise and background as described for PDA (Antonik et al., 2006). The analysis was then extended to two-dimensional and three-dimensional descriptions of the data by binomial and trinomial distributions. An exemplary fit of the three-dimensional data and its one-dimensional projections is shown in Figure 2.12.

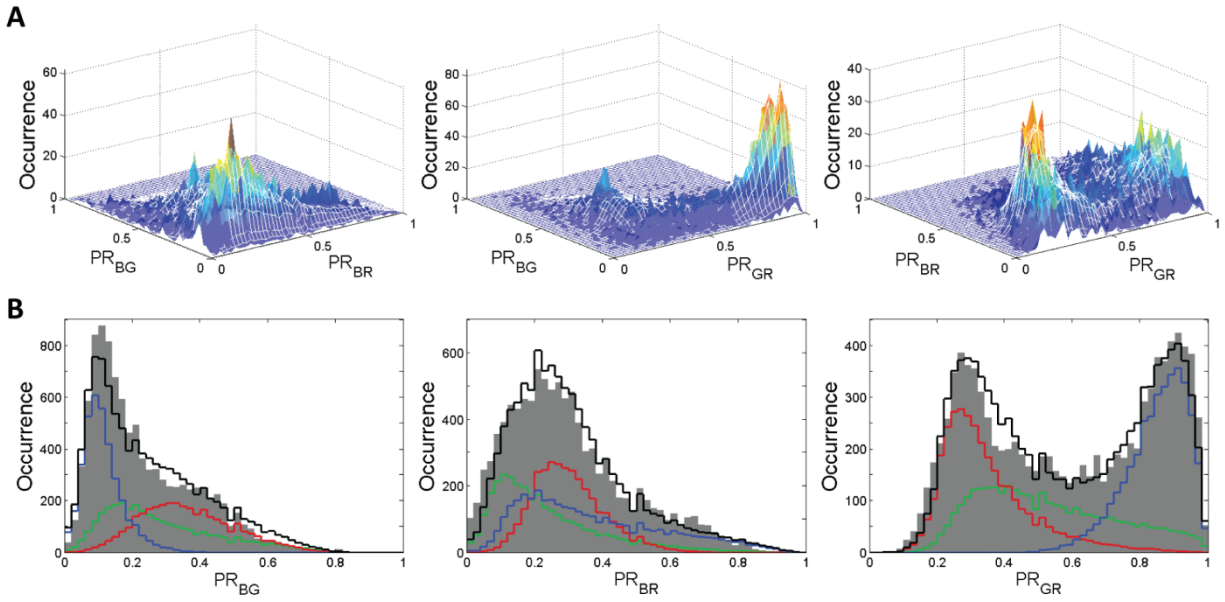


Figure 2.12 Exemplary three-dimensional PDA fit for three populations of three-color FRET data of a triple-labeled Hsp70 protein. (**A**) Three-dimensional histograms of the data (color) with their PDA fit (white lines). (**B**) One-dimensional projections of the histograms with the overall fit (black) and the individual fits for each ot the three populations (blue, green, red).

Dynamic conformational changes were investigated quantitatively by performing lifetime-filtered FCS on the data obtained from single molecule burst experiments (Böhmer et al., 2002; Felekyan et al., 2012). A low FRET and a high FRET efficiency species were selected from the $E(\tau_D)$ graphs and filters were calculated according to the fluorophore lifetime. In a

correlation analysis between these two species, molecules undergoing conformational dynamics introduced an anti-correlation term in the cross-correlation function. The timescales of conformational dynamics could be determined by fitting the correlations to the diffusion model described for FRET-FCS.

Experimental Section

Dual-color single-pair FRET experiments were performed on a custom-built confocal microscope as described (Chapter 2.10.1, Figure 2.8). The microscope was equipped with multiparameter fluorescence detection, which allows for the simultaneous determination of not only FRET efficiency, but also stoichiometry, fluorescence lifetime, and anisotropy. Pulsed interleaved excitation (or alternatively alternating laser excitation (ALEX) (Kapanidis et al., 2004, 2005; Lee et al., 2005)) makes it possible to calculate the values for stoichiometry (Kudryavtsev et al., 2012; Müller et al., 2005; Rüttinger et al., 2006). Information on fluorescence lifetime can only be obtained by the combined application of pulsed lasers and synchronized time-correlated single-photon-counting (TCSPC) detection where the exact arrival time of each photon with respect to its exciting laser is determined. The use of polarized excitation and detection additionally offers the possibility to gather information on anisotropy and therefore rotational correlation of the molecules.

Fluorescently labeled proteins were diluted to concentrations of 20-50 pM in their respective measurement buffer (Chapter 2.9.5) to ensure the detection of individual molecules (1-5 fluorescent bursts per second). LabTek #1.0 chamber slides (Thermo Scientific Nunc) were coated with 1 mg/mL BSA to suppress the unspecific attachment of proteins to the sample chamber and washed with the respective measurement buffer before the diluted sample was added to the chamber. Single molecule experiments were performed for 3-8 hours. The instrument response function of the system was obtained by additional measurements with ultrapure water (30 min) and the fluorescent background was determined using 30 min measurements of buffer in BSA-coated sample chambers.

For the fluorophore combinations Atto565/Atto647N and Atto565/Alexa647 or three-color FRET experiments (Atto488/Atto565/Atto647N or Atto488/Atto565/Alexa647) a similar confocal microscope was used. A schematic representation of the microscope is found in Figure 2.13. The confocal microscope was equipped with the laser excitation lines 485 nm (LDH-D-C-485), 565 nm by a frequency-doubled fiber-based femtosecond pulsed laser (FemtoFiber FFSOFT_TSHG, Toptica Photonics, Gräfelfing, Germany) or LDH-D-TA-560, and 641 nm (LDH-D-C-640) (Picoquant, Berlin, Germany). Laser powers of 110 µW for the blue laser, 80 µW for the green laser, and 40-80 µW for the red laser were measured at the entrance point of the beam into the objective. Fluorescence was collected by a 60x water immersion objective (60x/1.27 WI, Nikon, Düsseldorf, Germany), separated from the exciting wavelengths by a polychroic mirror (zt405/488/561/633), and focussed onto a 75 µm pinhole. It was then split according to polarization (PBS251, Thorlabs, Dachau, Germany), and wavelength (dichroic mirrors BS560 and 640DCXR) and filtered by the emission filters for blue (ET525/50), green (ET607/36), and red (ET670/30) fluorescence (AHF Analysentechnik). The signals were detected on six avalanche photodiodes (Count-100B:

Becker and Hickl, Berlin, Germany, SPCM-AXR-14, SPCM-AXRH-14: Perkin Elmer, Hamburg, Germany). The data were recorded using time-correlated single-photon counting cards (SPC-154, Becker and Hickl, Berlin, Germany) or a HydraHarp 400 system (Picoquant, Berlin, Germany). Data analysis was performed using a custom written Matlab program (The MathWorks, Ismaning, Germany) by Wolfgang Kügel and Anders Barth.

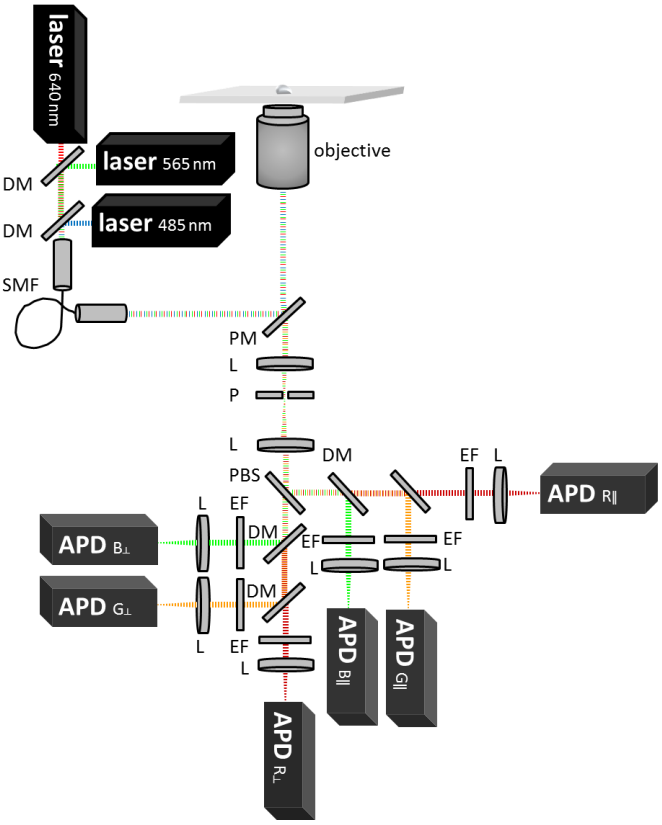


Figure 2.13 Schematic representation of the three-color confocal multiparameter fluorescence detection microscope. The sample is alternatingly excited by three pulsed lasers combined into a single-mode fiber (SMF). The fluorescence collected by the objective is separated from the excitation by a polychroic mirror (PM), focused onto the pinhole (P) by a lens (L) and separated according to polarization by a polarizing beam splitter (PBS). The fluorescence is further separated according to wavelengths by dichroic mirrors (DM) and after passing an emission filter (EF) focussed onto the detecting avalanche photodiodes (APD).

2.10.3 SpFRET experiments on the surface using total internal reflection fluorescence microscopy

SpFRET experiments in solution on a confocal microscope are limited by short observation times (few ms) of individual molecules. Surface-based single-pair FRET experiments can overcome this limitation to extend the observation period of individual molecules to several seconds up to one minute solely dependent on the photobleaching of the fluorophore. SpFRET experiments of immobilized molecules were analyzed using a program written in Matlab by Gregor Heiss (Professor Don C. Lamb, München, Germany) (Heiss, 2011). To detect individual molecules showing fluorescence in two colors, in a first step, a transformation matrix between the green and red detection channel was determined. To correct for differences in the detection channels, reference points in the channels were used for mapping. To this end either single particles or, for a more uniform distribution, the holes of zero-mode waveguides were identified in both detection channels. Their locations in both

channels were mapped onto each other to determine their transformation coordinates. These mapped coordinates were used for the single molecule experiments. To extract the intensities of individual molecules the 5% brightest pixels of an image averaged over several frames (usually 50) were chosen. The intensity was extracted by defining a circular region around these molecules and averaging the intensity value. Background values were extracted from the pixels surrounding this mask in a circular fashion. Whenever multiple molecules overlapped or the ratio between signal and background was limited, the molecules were excluded. Individual particles were manually selected according to their intensity trace over time. Single molecules were only considered for further analysis when they showed a single bleaching step. The background was subtracted as described and further corrections could be calculated for each molecule (Figure 2.14). The correction factor for spectral crosstalk was determined from the ratio of acceptor to donor fluorescence after photobleaching of the acceptor:

$$\beta = \frac{\langle I_{\mathrm{A}} \rangle_{\mathrm{bl}}}{\langle I_{\mathrm{D}} \rangle_{\mathrm{bl}}}. \tag{2.22}$$

Typical values for the correction factor were around β=0.035 (Figure 2.14B). Using intensity traces in which the acceptor fluorophore photobleached first, the correction factor for the relative detection efficiency γ was calculated as

$$\gamma - \beta = \frac{\Delta I_{\mathrm{A}}}{\Delta I_{\mathrm{D}}} \tag{2.23}$$

with ΔI_{A} being the difference in acceptor intensity before and after photobleaching and ΔI_{D} the difference in donor intensity after and before photobleaching. Typical values of the γ-factor in spFRET experiments of immobilized molecules were γ=0.65-0.7 (Figure 2.14C).

Individual molecules were selected according to a variety of properties. Molecules were chosen to be static, when their donor and acceptor signal as well as the total intensity stayed at a constant value (within the noise fluctuations) until one of the fluorophores photobleached. The intensities in both channels were used to calculate the FRET efficiency, which could be analyzed either framewise or as an average over all frames in one molecule (moleculewise):

$$E = \frac{F_{\mathrm{A}} - \beta F_{\mathrm{D}}}{F_{\mathrm{A}} - \beta F_{\mathrm{D}} + \gamma F_{\mathrm{D}}}. \tag{2.24}$$

Molecules were considered to be dynamic whenever they showed fluctuating anticorrelated intensities in their donor and acceptor fluorescence while the total intensity stayed constant (Figure 2.15B). At least two transitions between different FRET states were necessary for a molecule to fall into this category. In this case, the FRET efficiency was determined framewise (Figure 2.15C). For dynamic molecules, the correction factor γ was determined by minimizing the fluctuations in the total intensity:

$$\frac{d}{d\gamma} \sqrt{\sigma(I_{\mathrm{A}} + \gamma I_{\mathrm{D}})} = 0. \tag{2.25}$$

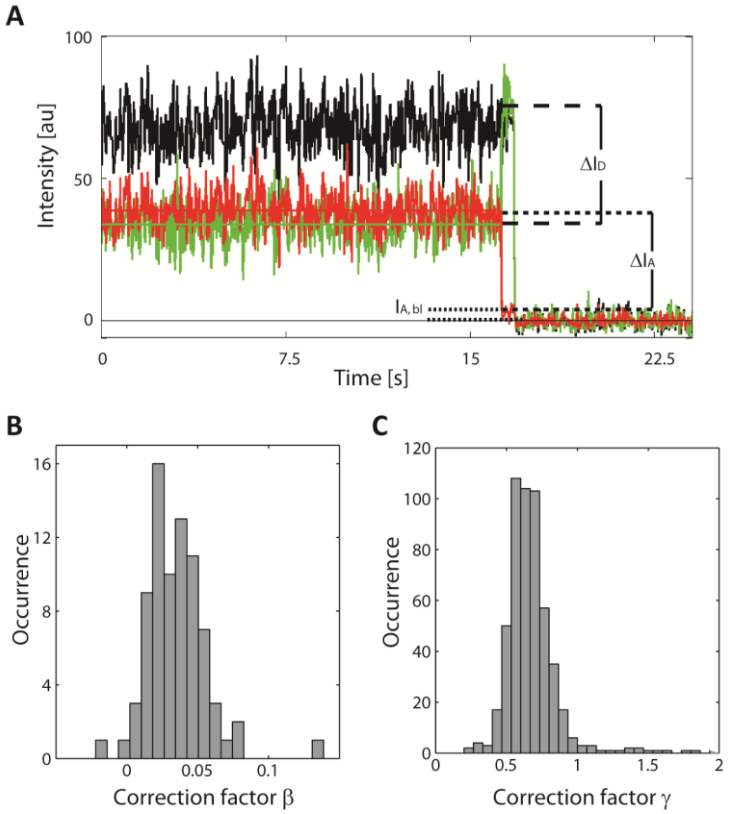


Figure 2.14 Correction factors for crosstalk and differences in detection efficiency for spFRET TIRF experiments. (**A**) Exemplary intensity time trace of a dual-labeled immobilized DNA-protein complex. The intensity of the donor (green), the intensity of the acceptor (red) as well as the corrected total intensity (black) are displayed. The acceptor fluorophore photobleaches after around 17s before photobleaching of the donor fluorophore. The correction factor γ can therefore be determined as described in Equation 2.23. The crosstalk is corrected from the intensity after photobleaching (Equation 2.22). (**B**) Exemplary distribution of the correction factor for spectral crosstalk β obtained from 78 molecules. (**C**) Exemplary distribution of the correction factor for differences in detection efficiency γ obtained from 525 molecules.

Dynamic properties of the molecules were further investigated by a hidden Markov model analysis (Baum and Petrie, 1966; Baum et al., 1970; Durbin, 1998; Rabiner and Juang, 1986). Early applications of HMM to biophysical questions include electrophysiological experiments (Qin et al., 2000) and fluorescence measurements with TCSPC detection (Andrec et al., 2003; Schröder and Grubmüller, 2003; Yang, 2003). Here, time-binned FRET values were described by a hidden Markov process (Heiss, 2011; McKinney et al., 2006; Sikor, 2011) included in the HMM Matlab Toolbox (Murphy, 1998). As model parameters of the HMM, the FRET efficiencies of a predefined number of states were used and their noise was approximated by Gaussians with a standard deviation σ (Dahan et al., 1999; McKinney et al., 2006). A transition probability matrix calculated the probability of transitions between different states and the Viterbi algorithm was used to determine the most likely sequence of states (Viterbi, 1967). Optimization of the parameters to a maximum probability was achieved using the Baum-Welch algorithm (Baum et al., 1970). The transitions were visualized as two-dimensional Gaussian distributions in a transition density plot (TDP) (Figure 2.15D). The widths of each point of this distribution contained the standard deviation of the noise found by HMM analysis. Populations transiting from one state to another were chosen from these TDPs and the kinetic rates of the dynamics were obtained by fits of the dwell time histograms

(number of molecules still in the initial state over time) to single or biexponential decay functions (Figure 2.15E).

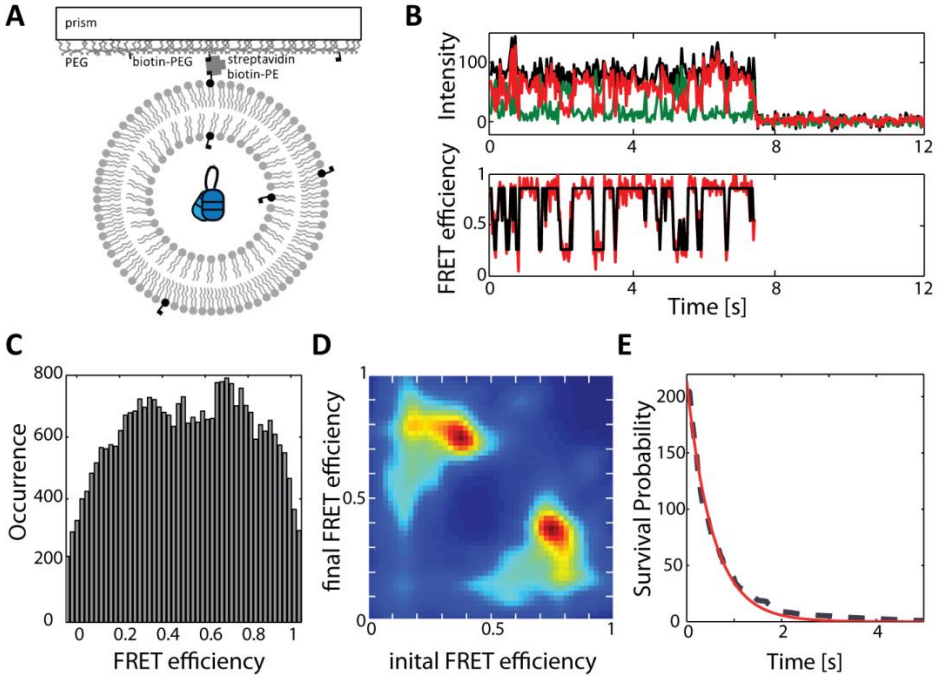


Figure 2.15 Illustration of the HMM analysis procedure. (**A**) Schematic representation of a lipid vesicle attached to the prism surface via biotin-streptavidin linkage (not to scale). Proteins, in this case U2AF65, were immobilized on the surface by encapsulation into these lipid vesicles. (**B**) Exemplary intensity time trace of a dual-labeled, encapsulated U2AF65 protein showing conformational dynamics. Upper graph: The donor (green) and acceptor (red) signal show an anticorrelated behavior when the protein undergoes conformational changes, while the total intensity (black) stays constant. The FRET efficiency (lower graph, red) switches between two states. The black curve (lower graph) represents the most likely path as determined from the HMM analysis. (**C**) Framewise FRET efficiency histogram of dynamic proteins. (**D**) Transition density plot obtained by HMM analysis. The proteins switch between a low FRET (~0.4) and a high FRET (~0.8) state. (E) Survival probability of molecules undergoing a transition from one state to the other. Transition rates were extracted by fitting the transition probability to a single exponential decay function.

Experimental Section

SpFRET experiments of surface-immobilized molecules were performed on a custom-built prism-type total internal reflection fluorescence microscope (TIRFM). Preparation of the prisms was performed as described (Chapter 2.9.2). The quartz prism served as part of the sample chamber as well as to create an angle of total internal reflection. A detailed description of the construction of the quartz prism and integration into the microscope to achieve total internal reflection can be found in Heiss (2011), Schluesche (2007), or Sikor (2011). A schematic illustration of the microscope is shown in Figure 2.16. Excitation of the fluorophores was achieved by illumination of the prism at 532 nm (GCL-100-L, CrystaLaser, Reno, USA replaced by Cobolt Samba 100, 532 nm, Solna, Sweden) and at 633/647 nm (helium-neon laser, Polytec Laser 640 nm, München, Germany replaced by Cobolt MLD, 647 nm). An acousto-optical tunable filter (AOTFnC.400-650-PV-TN, Pegasus Optik, Germany) combined with a single-mode polarization maintaining fiber was used for switching and intensity regulation of the lasers. Fluorescence of the sample was collected by a water immersion objective (CFI Plan Apochromat 60xWI, NA 1.2, Nikon, Düsseldorf, Germany) and separated by a dichroic mirror (630DCXR, AHF, Analysentechnik, Düsseldorf,

Germany). After passing the emission filters HQ550/88 or HQ715/150 (AHF, Analysentechnik) respectively, the fluorescence of the donor and acceptor fluorophores was detected on two different regions of the electron multiplying charge coupled device (EMCCD camera, Andor iXON, Andor Technology, Belfast, Great Britain). Measurements were performed either by directly immobilizing biotinylated RNA or DNA on the surface or by encapsulation of the protein complex in lipid vesicles (Chapter 2.9.4). The quartz prisms were therefore washed with PBS, incubated with 0.3 mg/mL streptavidin in PBS for 5-10 min, washed again with PBS and the respective buffer for each individual protein (Chapter 2.9.5). The background fluorescence was photobleached by exposition of the prism to the 488 nm laser. The biotinylated vesicles or RNA complexes were then immobilized via biotin-streptavidin-biotin linkage on the surface of the quartz prisms (Sikor et al., 2013). The fluorophores were excited by the 532 nm laser or alternatively by alternating both, the red and green laser, with laser powers of 3-10 mW measured before the entrance point into the prism. An exposure time of 15-100 ms was chosen depending on the experiment. For single molecule experiments the gain of the camera was usually set to a factor of 300. Fluorescently-labeled polystyrene particles coated with avidin (nile red fluorescent polystyrene particles, Kisker Biotech, Steinfurt, Germany) visible in both detection channels or zero-mode waveguides were measured for calibration of the channels. The analysis was performed in a custom written Matlab program (The MathWorks, Ismaning, Germany) by Gregor Heiss.

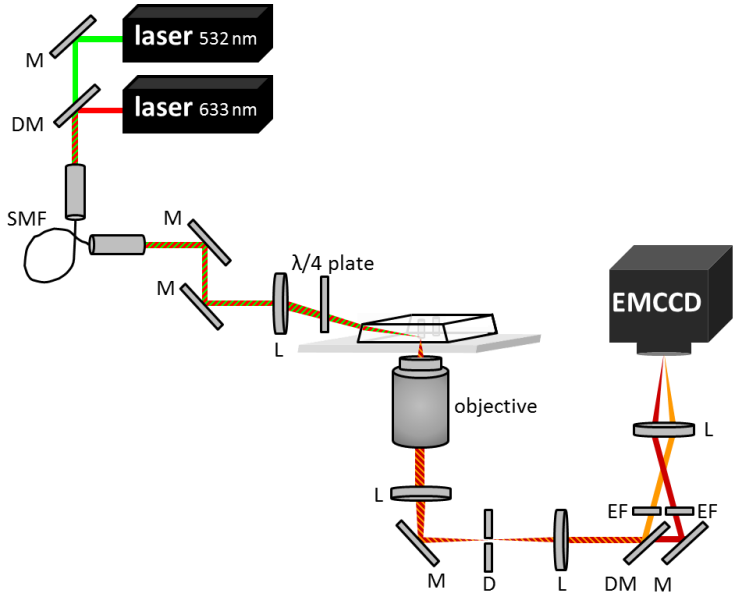


Figure 2.16 Schematic representation of the prism-type total internal reflection microscope for spFRET experiments. The sample is attached to the surface of a quartz prism by biotin-streptavidin linkage. At the same time, the prism is used to create an angle of total internal reflection with the incoming lasers (combined by dichroic mirrors (DM) and mirrors (M) into a single-mode fiber (SMF) and focussed onto the prism by a lens (L)). The objective collects the fluorescence, which is then focussed onto a diaphragm (D) and separated according to wavelengths by dichroic mirrors. After selection of the wavelengths of interest by emission filters (EM), these are focussed onto different regions of an electron multiplying charge-coupled device (EMCCD).

For multicolor experiments, the two-color TIRF microscope was adjusted by Sushi Madhira. The additional laser lines 561 nm (Cobolt Jive-50) and 488/491 nm (argon-krypton ion laser, Stabilite 2018-RM, Coherent, Santa Clara, USA; later replaced by Cobolt Calypso-50)

replaced the excitation at 532 nm. The fluorescence emission was split after the diaphragm by dichroic mirrors 630DXCR, 560DXCR, and 660DXCR before filtering of the emission by HQ525/50, HQ595/50, or 705/100 (AHF, Analysentechnik).

2.10.4 Determination of dissociation rates on a TIRF microscope

The rate of dissociation of nucleic acid binding proteins from their nucleic acid binding sequence on DNA or RNA was determined using TIRF microscopy. The nucleic acids were incubated with their binding partners and immobilized on the surface of a quartz prism at a concentration of 50 pM. The excitation laser power was chosen such that the bleaching rate of the fluorophores was significantly lower than the time of dissociation. Control experiments of samples which did not show any dissociation were used to extract the rate of photobleaching. The number of fluorescent molecules observed over time was fit by single-exponential decay functions. The difference between the obtained decay rate associated with photobleaching and the decay in the number of fluorescent molecules due to a combination of photobleaching and dissociation was explained by dissociation of fluorescently-labeled proteins from immobilized nucleic acids. We used this approach to determine the rate of dissociation of TBP from TATA-box containing DNA and of U2AF65 from pyrimidine-tract RNA.

2.11 In silico analysis

2.11.1 Protein sequence comparison

Primary sequences of proteins were compared using Clustal W (Larkin et al., 2007; Thompson et al., 1994). FASTA sequences were used as input parameters and pairwise alignments were calculated for all combinations of sequences. These were converted to a distance matrix to quantify the homology content and evolutionary distances.

2.11.2 Secondary structure comparison of proteins

The secondary structure of proteins was predicted using a known crystal structure of a homologous protein and the primary sequence of the structurally unknown protein (SWISS-MODEL (Arnold et al., 2006; Biasini et al., 2014; Guex et al., 2009; Kiefer et al., 2009)). The PyMOL software (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC) was used to structurally align homologous proteins. We performed structural alignments using the known crystal structure of DnaK (protein data bank accession numbers (PDB) 2KHO and 4B9Q) and the primary sequences of Ssc1, BiP, and cpHsp70 for comparison.

2.11.3 Accessible volume calculations

To calculate the accessible volumes of fluorophores attached to the proteins of interest for spFRET experiments, the FRET positioning software by the group of Prof. Seidel was used (Kalinin et al., 2012). It was used for the accessible volume calculations of fluorophores attached to RRM1,2 (PDB: 2YH0) and RRM1,2 attached to U8-RNA (PDB: 2YH1). A linker length of 20 Å, linker width of 4.5 Å, and a radius of the fluorophores of 3.5 Å were used for the calculations.

3 The role of HMGA1a/nucleic acid interactions during replication

The high mobility group protein HMGA1 plays a highly important role in cell proliferation, growth, and differentiation. More recently, it was associated with the initiation of DNA replication. To perform its function, HMGA1 binds to the minor groove of AT-rich DNA via three AT-hook DNA binding motifs. The Epstein-Barr virus nuclear antigen EBNA1 uses a similar mechanism of DNA recognition to induce replication of its genome. Using FCS and spFRET, the binding of both proteins to DNA and RNA was studied. Both proteins showed a high affinity for the nucleic acids and formed higher order complexes, while large scale conformational rearrangements of short DNA sequences have not been observed.
We investigated HMGA1 and EBNA1 binding to DNA in collaboration with Dr. Aloys Schepers, Christoph-Erik Mayer, and Julia Damaschke (Helmholtz Zentrum München, Germany).

3.1 Introduction

HMGA1 is known to be involved in a variety of cellular processes. It has been ascribed a major role in the regulation of gene expression by rearrangement of the chromatin structure. Besides this, HMGA1 is involved in nuclear processes such as replication and recombination (Reeves, 2001). Due to its diverse functions, it is associated with many diseases, including an overexpression of the protein in almost every tumor or HMGA1-associated inhibition of nucleotide excision repair in Xeroderma pigmentosum (Benecke et al., 2015). In recent years, HMGA1 was found to play an important role in the recruitment of proteins recognizing DNA replication origins (Norseen et al., 2008; Thomae et al., 2008).

3.1.1 DNA replication

For the maintenance of the genome during cell division, DNA replication is an essential process. DNA duplication takes place during the S-phase of the cell cycle (Alberts, 2015). The whole eukaryotic DNA in a cell is replicated starting from multiple origins of replication (Bielinsky and Gerbi, 2001; DePamphilis, 1999; Gilbert, 2001). These mostly AT-rich regions are recognized by the origin recognition complex (ORC). The expression of one subunit, ORC1, is regulated by the cell cycle, while a complex of subunits 2-5 is constitutively present as a stable complex. Several regulatory steps ensure that replication initiation only starts in S-phase (Kearsey and Cotterill, 2003; Kelly and Brown, 2000). The binding of helicases, phosphorylation of the proteins by cell cycle dependent kinases, and loading of the DNA polymerase lead to initiation of replication. Both strands serve as matrices for the synthesis of new DNA. While the leading strand DNA is continuously synthesized by DNA polymerase in the 5'→3' direction, a discontinuous synthesis occurs on the lagging strand and these so called Okazaki fragments are then joined and the replication ends at the telomeres of eukaryotic chromosomes.

3.1.2 Regulation of ORC recruitment by HMGA1

The origin replication complex has a high DNA affinity, which however does not induce any sequence specificity. Therefore, additional factors have been proposed to play a role in ORC binding to the DNA. The high mobility group protein HMGA1a has been associated with ORC recruitment to AT-rich heterochromatic regions (Norseen et al., 2008; Thomae et al., 2008). HMGA1 belongs to a family of chromosomal proteins important in transcripton, DNA replication, and recombination with isoforms ranging from 10.6-12 kDa (Disney et al., 1989; Lund et al., 1983). It carries three palindromic AT-hooks for the binding of AT-rich minor grooves, which are highly conserved carrying the amino acid sequence PRGRP (Figure 3.1) (Aravind and Landsman, 1998; Fonfría-Subirós et al., 2012; Reeves and Nissen, 1990, 1995; Solomon et al., 1986; Strauss and Varshavsky, 1984). Binding to the DNA was shown to occur in a structure- and sequence-specific manner (Geierstanger et al., 1994; Huth et al., 1997; Reeves and Nissen, 1990). Some publications propose an unchanged DNA structure upon binding (Bewley et al., 1998; Dragan et al., 2003; Huth et al., 1997), while others showed a bending of the DNA by widening of the minor groove (Fonfría-Subirós et al., 2012). Coimmunoprecipitation and bimolecular fluorescence complementation experiments detected the interaction of HMGA1a and ORC6 (Thomae et al., 2008) and Thomae et al. (2008) as well as Norseen et al. (2008) showed that these interactions are mediated by RNA and depend on RGG motifs.

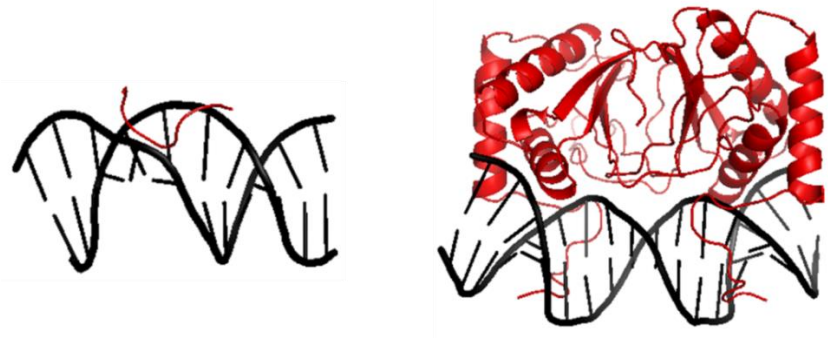


Figure 3.1 Structures of the AT-hook of HMGA1 and EBNA1 bound to DNA. The binding of a single AT-hook of HMGA1 to the minor groove of the DNA is displayed on the left (PDB: 3UXW). On the right EBNA1 binding to DNA is shown (PDB: 1B3T). The DNA is shown in black and both proteins are displayed in red.

This RNA-dependent interaction is also used by the Epstein-Barr virus encoded EBNA1 (Norseen et al., 2009). Its viral genome is maintained separately inside the cell, hijacking the cell's replication initiation machinery (Sarkari et al., 2009; Schepers et al., 2001). The N-terminal domain of EBNA1 thereby mimicks the human HMGA1a (Hung et al., 2001; Sears et al., 2004). The EBNA1 DNA binding domain consists of a core domain and a flanking region (Bochkarev et al., 1996). While the core domain does not contact the DNA, a helical region of the flanking domain binds to the major groove and an extended peptide chain contacts the minor groove. It binds to palindromic DNA sequences as a dimer and induces structural changes in the DNA (Figure 3.1) (Bochkarev et al., 1996; Frappier and O'Donnell, 1991; Hearing et al., 1992; Hsieh et al., 1993). A compression of the binding domain is caused by overwinding of the DNA (Bochkarev et al., 1996, 1998). These conformational

changes then lead to the separation of both strands for the initiation of replication of the viral genome.

3.2 Results

The proteins HMGA1 and EBNA1 were studied regarding their mechanism in DNA and RNA binding. Structurally, HMGA1 is disordered, forming random coils, in its free form, while upon binding to DNA, it assumes its secondary structure (Reeves, 2010).

In a first step, we investigated the binding of the proteins to DNA and RNA by FCS. To this end, short DNA stretches of 30 and 40 nucleotides in length were used. Electromobility shift assays showed effective binding of HMGA1 to these DNA sequences (Figure 3.2A). The addition of gradually increasing concentrations of unlabeled HMGA1 to these DNAs in FCS experiments showed the formation of DNA/HMGA1 complexes. We found the complex to reach a first plateau in the binding curve at an HMGA1 concentration of ~200 nM (Figure 3.2B, C). This corresponds to a high affinity of HMGA1 for the DNA. When further increasing the concentration, however, we again observed increasing diffusion times suggesting the formation of larger complexes rather than complexes with a 1:1 ratio of DNA:protein. We observed the same rise for DNAs labeled with different fluorophores, thus excluding an artificial effect of the fluorophore on the formation of larger complexes.

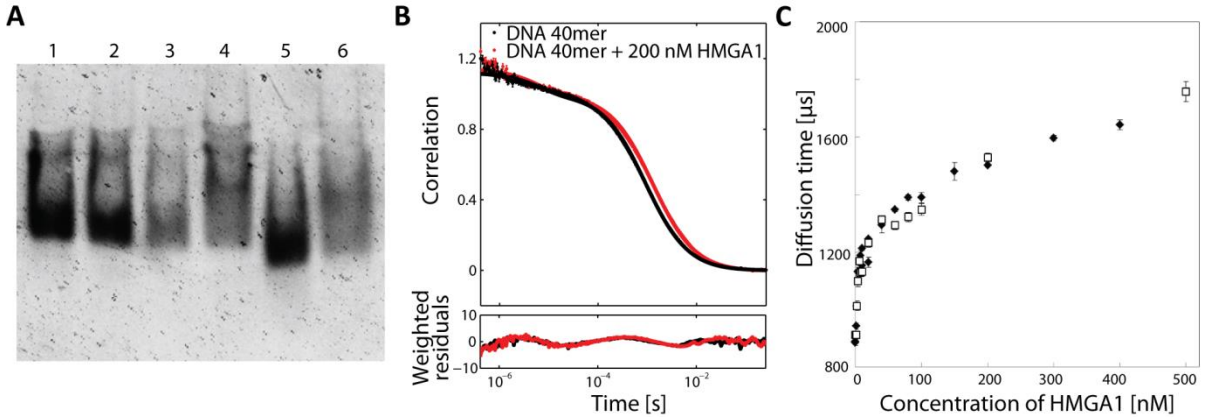


Figure 3.2 Binding of HMGA1 to short stretches of DNA. (**A**) EMSA of HMGA1 binding to DNAs, 30 and 40 nucleotides in length. In a gel scanner, the labeled DNA was excited at 633 nm and its fluorescence was detected after passing a 670BP30 filter. 3 nM 40mer (lanes 1-4) in the absence of HMGA1 (lane 1) and in the presence of 3 nM HMGA1 (lane 2), 30 nM HMGA1 (lane 3), and 300 nM HMGA1 (lane 4). 3 nM 30mer (lanes 5-6) in the absence of HMGA1 (lane 5) and in the presence of 300 nM HMGA1 (lane 6). (**B**) Exemplary autocorrelation functions of the 40mer in the absence (black) and presence of 200 nM HMGA1 (red). (**C**) Diffusion times of the 40mer dependent on the concentration of HMGA1 as obtained from FCS. Two different measurement series were displayed as squares and diamonds, respectively. The error bars represent the errors of the fit to the data of the autocorrelation function.

We also analyzed the behavior of EBNA1 binding to DNA. EBNA1 induces the formation of large complexes of varying size, which were observed as aggregates causing large spikes of fluorescence on the confocal microscope. Any combination of DNA/EBNA1 complexes in the absence or presence of additional RNA causes the formation of these complexes (Figure 3.3A). Even the addition of EBNA1 to a mixture of DNA and RNA at the beginning of the measurement induced the formation of large complexes. In general, in the presence of additional RNA, the observed complexes showed a lower molecular brightness than for direct

complex formation of DNA with EBNA1 followed by the addition of EBER1-RNA during the experiment (Figure 3.3A). This might indicate the formation of complexes with a competitive binding of DNA and RNA. Therefore, an exchange of fluorescent DNA units for RNA occurred over the course of the measurement.

The binding of EBNA1 to the DNA was shown to be salt-dependent. The addition of high NaCl concentrations during complex formation drastically decreased the amount of aggregates observed (Figure 3.3B). For salt concentrations above 200 mM NaCl, complexes of high molecular brightness were not observed. However, even for these salt concentrations, EBNA1 binding to the DNA was not generally inhibited. With salt concentrations higher than 200 mM NaCl, we still observed longer diffusion times for the complexes than for free DNA. A similar aggregation-prone behavior of EBNA1 binding to fluorescently-labeled RNA was observed (Appendix). Here as well, the aggregation behavior was inhibited by 250 mM NaCl, although, complex formation still occurred.

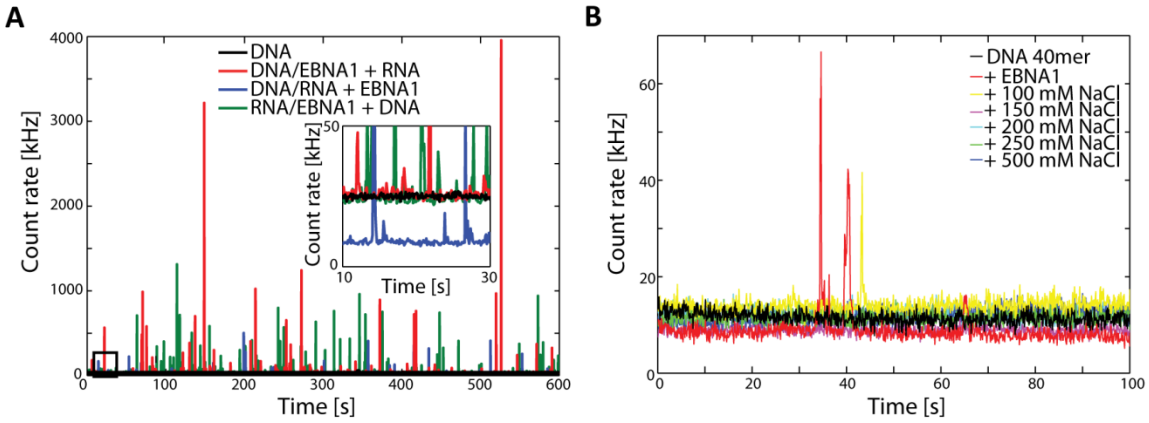


Figure 3.3 Salt-dependent binding of EBNA1 to DNA. (**A**) Competitive DNA 40mer and EBER1-RNA binding of EBNA1. Fluorescence intensity trace over the course of the measurement. DNA in the absence of EBNA1 (black) or upon complex formation with EBNA1 and RNA. DNA/EBNA1 complexes upon RNA addition (red), addition of EBNA1 to a mixture of DNA and RNA (blue) and DNA added to RNA/EBNA1 complexes (green). The inset shows a magnification of a short time window and count rates up to 50 kHz (black square). (**B**) EBNA1 binding to DNA in the presence of different NaCl concentrations. Fluorescence intensity of a short time period of the experiments.

Since the effect of HMGA1 on the conformation of the DNA is still under debate (Bewley et al., 1998; Chase et al., 1999; Dragan et al., 2003; Falvo et al., 1995; Fonfría-Subirós et al., 2012; Huth et al., 1997), we used spFRET with PIE-MFD to study DNA/HMGA1 complexes in solution. Even though binding of the proteins to DNA was clearly observed in FCS experiments, burst analysis of dual-labeled DNA did not show any conformational changes upon EBNA1 or HMGA1 binding. Two different dual-labeled DNAs, a 30mer and a 40mer with fluorescent labels on both ends, were tested. The mean FRET efficiency of the 30mer in the absence of protein as well as in the presence of HMGA1 or EBNA1 was 0.07 (Figure 3.4A). The elongated DNA fragment, 40 nucleotides in length, showed an average FRET efficiency of 0.06 in the absence and in the presence of both proteins (Figure 3.4B). An additional small subpopulation of the DNA was observed, which showed higher FRET efficiencies of around 0.78. Since it was also found in the DNA sample without protein, this subpopulation might be attributed to incomplete complementary strand binding.

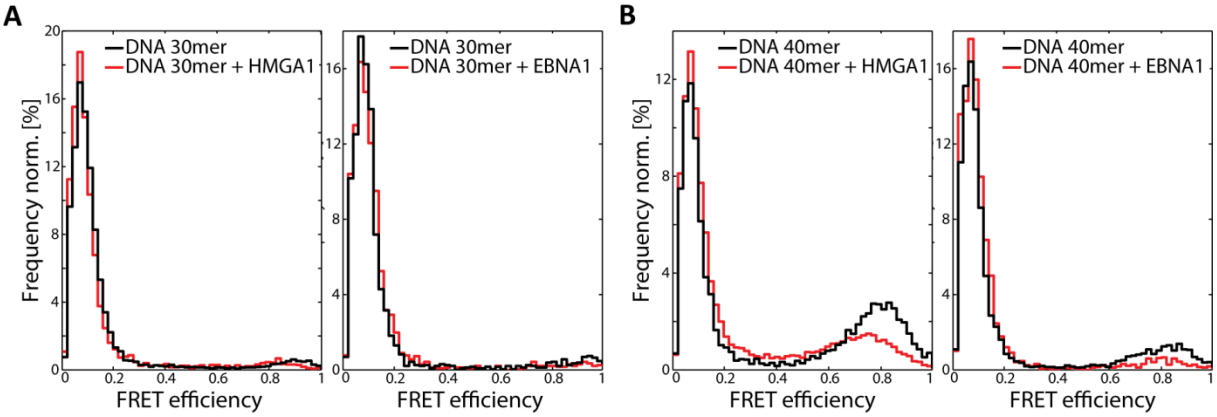


Figure 3.4 FRET efficiency distributions of dual-labeled DNA obtained from spFRET experiments with PIE-MFD. (**A**) DNA 30mer in the absence (black) and presence of 200 nM HMGA1 (red, left) or EBNA1 (red, right). (**B**) DNA 40mer in the absence (black) and presence of 200 nM HMGA1 (red, left) or EBNA1 (red, right).

We further labeled Cys-HMGA1a at a single N-terminal cysteine with different fluorophores (Atto488, Atto532, and Atto647N) (Chapter 2.8.9). The cysteine did not affect the binding of HMGA1 to the DNA as we conclude from the diffusion coefficients in FCS experiments (Figure 3.5A). Similarly, the diffusion coefficient decreased upon complex formation of DNA in the presence of fluorescently labeled HMGA1 (Figure 3.5B). However, although binding was still observed in the presence of labeled Cys-HMGA1, the cross-correlation amplitude between the DNA and the labeled protein was always zero, suggesting that unlabeled Cys-HMGA1 was responsible for the change in diffusion time, while the labeled Cys-HMGA1 did not bind to the DNA (Figure 3.5C). The cross-correlation analysis is shown exemplarily for Cys-HMGA1-Atto647N. For the attachment of other fluorophores (Atto532 and Atto488), however, the cross-correlation amplitude between the DNA and HMGA1 was likewise negligible. Therefore, a decrease in the binding affinity due to one specific fluorophore can be excluded. Most likely, the interaction between HMGA1 and the DNA was affected due to a modification at the specific N-terminal position.

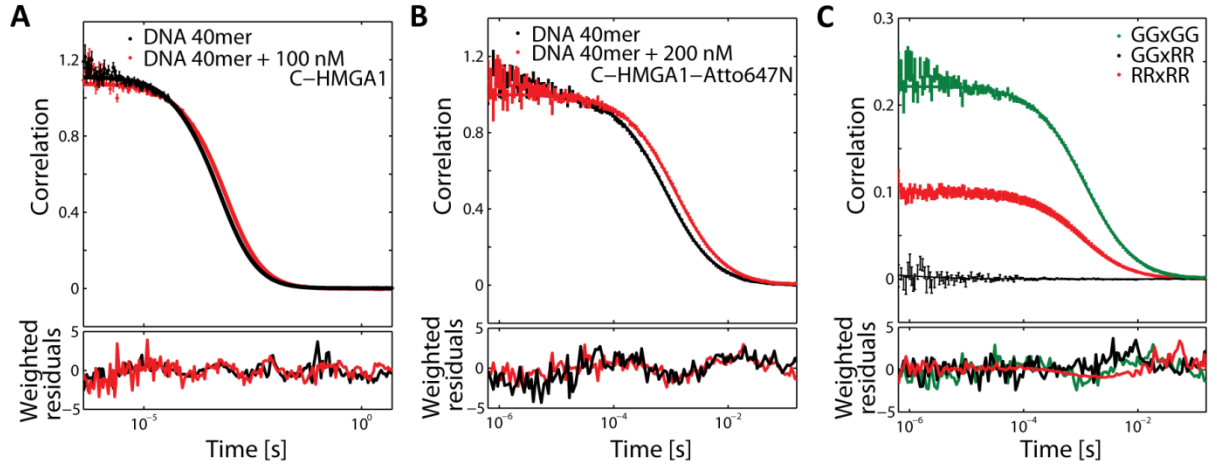


Figure 3.5 Effects of HMGA1 labeling on DNA-binding studied by FCS and FCCS. (**A**) Autocorrelation functions of DNA (black) and DNA/C-HMGA1 complexes (red). (**B**) Autocorrelation curves of DNA (black) and DNA in the presence of C-HMGA1-Atto647N (red). (**C**) FCCS of DNA-Cy3 in the presence of C-HMGA1-Atto647N. Autocorrelation functions of DNA (green), C-HMGA1-Atto647N (red), and the cross-correlation function between them (black).

3.3 Discussion and prospects

The binding of HMGA1 to DNA by the insertion of its three AT-hooks into the minor grooves of DNA has been shown by NMR studies (Geierstanger et al., 1994; Huth et al., 1997) and biochemical experiments combined with structural predictions (Reeves and Nissen, 1990). We observed a high binding affinity of HMGA1 to the DNA in FCS experiments. In these experiments, the AT-hooks of HMGA1 bound to a palindromic sequence of AATATT segments. Earlier experiments had shown specific binding of the AT-hook 3 of HMGA1 to one of two AATT sequences only (Fonfría-Subirós et al., 2012). In our experiments, a single complex with stable diffusion coefficient was not observed even with a large excess of HMGA1. Rather, we found the diffusion time to gradually increase with increasing HMGA1 concentrations. Due to its palindromic sequence, two HMGA1 molecules might possibly bind to DNA inducing the additional increase in diffusion time above a concentration of 200 nM. Furthermore, multiple DNAs might be bound by a single HMGA1. This could be investigated in additional FCCS experiments with multiple differently labeled DNAs. The binding and cross-linking of multiple DNA strands would agree well with a described cross-linking effect of HMGA1 on the DNA in *in vitro* experiments and in cells overexpressing the protein (Vogel et al., 2011) and could explain the observed formation of aggregates.

Interestingly, for EBNA1, we observed large DNA/protein complexes, which dissociate in the presence of high NaCl concentrations. This indicates that the interactions leading to aggregate formation are of electrostatic nature. EBNA1 is known to dimerize and to bind DNA as a homodimer. Here again, complexes of higher molecular weight and brightness might be formed due to a cross-linking effect of the protein as described for HMGA1 (Vogel et al., 2011).
Furthermore, EBNA1 was shown to introduce a localized change in DNA structure, which leads to the separation of both DNA strands (Bochkarev et al., 1996; Frappier and O'Donnell, 1991; Hearing et al., 1992; Hsieh et al., 1993). Our spFRET experiments do not suggest any conformational rearrangement of DNA. The FRET efficiency, however, was very low in general and far away from the Förster radius, where small changes in the interdye-distance are detected most easily. Very small conformational changes or changes for which the overall distance between the fluorophores was not affected, might thus not have been observed due to the lower sensivity of FRET in this distance range.

Many different structural rearrangements of the DNA were described for DNA binding of the small protein HMGA1. While for some DNA stretches, straightening of preformed bends was described (Falvo et al., 1995), other studies observed bending (Chase et al., 1999), loop formation and chromatin remodeling (Benecke et al., 2015). Fonfría-Subirós et al. (2012) showed the bending of a short stretch of DNA and widening of the minor groove by binding of a single AT-hook peptide. Bewley et al. (1998), Dragan et al. (2003), and Huth et al. (1997) proposed a stable B-form DNA even upon HMGA1 binding. Using the results from our spFRET experiments, we can exclude bending of the DNA with a bending angle below 95°. At this bending angle, the distance between the fluorophores would reach a value of 75 Å for

a DNA of 30 nucleotides in length and could therefore be detected by a change in the FRET efficiency. We cannot exclude, however, any small conformational changes in the DNA or RNA upon HMGA1 binding.
A bending effect was easily visualized, however, for elongated DNAs, where larger changes in the positions of both ends can occur. An interesting next step would be to perform spFRET experiments with PIE-MFD on elongated DNAs, where larger repositioning of the dyes is observed or with fluorophores positioned directly next to the HMGA1 and EBNA1 binding site, so that the mean FRET efficiency is closer to the Förster radius. Alternatively, pairs of fluorophores with a larger Förster radius such as Atto565/Atto647N could be used on the DNA. As a next step, active fluorescently labeled proteins could be prepared by repositioning of the fluorophores to different attachment points. These could then be analyzed in complex with dual-labeled DNA by three-color FRET. Three-color FRET and FCCS might additionally be utilized to study the competitive binding of DNA and RNA. An interesting aspect would be to investigate whether RNA can induce specificity in HMGA1-binding to the DNA as proposed by Norseen et al. (2008) and Thomae et al. (2008). This effect could be analyzed by performing FCS experiments of DNAs with different affinities and fluorescent labels in the presence of HMGA1 and upon complex formation with RNA or by observation of FRET between DNA and HMGA1 in the absence and presence of RNA.

Due to the ubiquitous expression and manifold functions of HMGA1 in cell maintenance and differentiation, it is highly important to understand the molecular mechanism of the protein in DNA recognition and binding. As described for other DNA-binding proteins such as TBP (Chapter 4) (Heiss, 2011; Schluesche et al., 2007), spFRET offers great advantages in investigating conformational changes and dynamics of proteins bound to DNA or the DNA itself. Using a combined effort of varying the FRET pairs, adjusting the labeling position on HMGA1 and optimizing the DNA lengths, the project can be further pursued to reveal the binding mechanism of HMGA1 and any associated conformational rearrangements in the DNA.

4 The role of cofactors in the regulation of the DNA/TBP interaction during transcription initiation

Protein expression is a highly regulated process in eukaryotic cells. The main regulation takes place at the early steps of transcription. Transcription is initiated by the binding of TBP to TATA-box containing promoters. This interaction of DNA with TBP is tightly controlled by a multitude of transcription cofactors. The effect of a few of these cofactors on the conformation of the DNA/TBP complex was investigated by spFRET experiments. We found a stable DNA/TBP complex. Dissociation of the complex occurred in the presence of TBP's cofactor Mot1, while other cofactors such as NC2 and PC4 did not induce TBP dissociation from the DNA. A detailed investigation of Mot1 revealed its remodeling role on the DNA/TBP complex and its two-step mechanism of dissociating TBP from the DNA. We additionally analyzed the orientation of TBP on the DNA and the influence of DNA twist on the bending of DNA upon TBP binding.

The projects were investigated in collaboration with Professor David Auble, Ramya Viswanathan (University of Virginia, USA), and Professor Michael Meisterernst (Westfälische Wilhelms-Universität Münster, Germany).

4.1 Introduction

Protein expression is a necessary process for cell survival, adaptation to environmental changes, and cellular differentiation (Jacob and Monod, 1961). It is a highly regulated process in the cell because any alterations can lead to severe diseases associated with uncontrolled cell growth and tumor formation (Golub, 1999; Johnson et al., 2003b). This ubiquitous process can be divided into two major steps. Initially, the DNA is transcribed into pre-mRNA in the nucleus of eukaryotic cells. Secondly, matured mRNA is translated into protein by ribosomes in the cytosol of the cell. Unlike in bacteria, these processes are spatially separated in eukaryotic cells.

Protein expression is highly regulated in each individual cell. While short-term regulations of protein content are performed by protein modification, degradation, or translational regulation, for long-term adjustments the regulation takes place during the early steps of transcription (Curtis et al., 1995; Epstein and Beckwith, 1968; Jaenisch and Bird, 2003; Lu et al., 2007). By regulating the initiation of DNA transcription, energy costs for the cell are kept to a minimum (Berger, 2007; Chen and Rajewsky, 2007; Mitchell and Tjian, 1989).

4.1.1 Initiation of DNA transcription

DNA transcription, the first step of protein biosynthesis, takes place in the nucleus, where the information of a gene is transcribed into RNA. The process of transcription can be divided into three steps: initiation, elongation, and termination. These steps are mainly driven by RNA polymerase, which is, in turn, strongly regulated by a multitude of additional proteins. In eukaryotes, three types of RNA polymerases exist, of which the RNA-polymerase II is responsible for the transcription of messenger RNA (mRNA).

For the initiation of transcription, protein-coding genes are flanked by a promoter region. It is responsible for the binding of a variety of transcription factors. Promoters close to the transcription start site are the binding site of the general transcription factor TFIIB 31-26 basepairs upstream of the start site, and the initiator motif and downstream core promoter element for binding of TFIID (Alberts, 2015). A frequently occurring core promoter element is the TATA-box, which can form a promoter on its own or in combination with the other two sequences. Besides these core promoter elements, additional upstream promoter regions exist for the regulation of DNA transcription, including the CAAT-box and the GC-box (Alberts, 2015). At the beginning of transcription in eukaryotic cells, a pre-initiation complex is formed at these sites. Formation of this complex starts with the binding of TATA-box binding protein (TBP), a subunit of the transcription factor complex TFIID, to the promoter region. Binding of TBP to the TATA-box induces a bend in the DNA (Masters et al., 2003; Parkhurst et al., 1996; Wolner and Gralla, 2001). Upon this conformational change, additional transcription factors are able to recognize the promoter and prime it for transcription. TBP-associated factors can regulate the formation of the TFIID complex, although they are not always required. Subsequent binding of TFIIA stabilizes the DNA/TBP complex and promotes transcription. Additional transcription factors and RNA polymerase bind to the DNA in the order displayed in Figure 4.1. The transcription initiation ends with the phosphorylation and the thus induced conformational change of the C-terminal domain of RNA polymerase II and the RNA synthesis begins (Nikolov et al., 1995; Roeder, 1996, 2005; Thomas and Chiang, 2006). The DNA matrix strand is read by the RNA polymerase traveling along the non-coding strand in the 3’ to 5’ direction and complementing the nucleotides with a rate of 30-60 bases per second. Palindromic sequences forming hairpin structures terminate the transcription in eukaryotic cells.

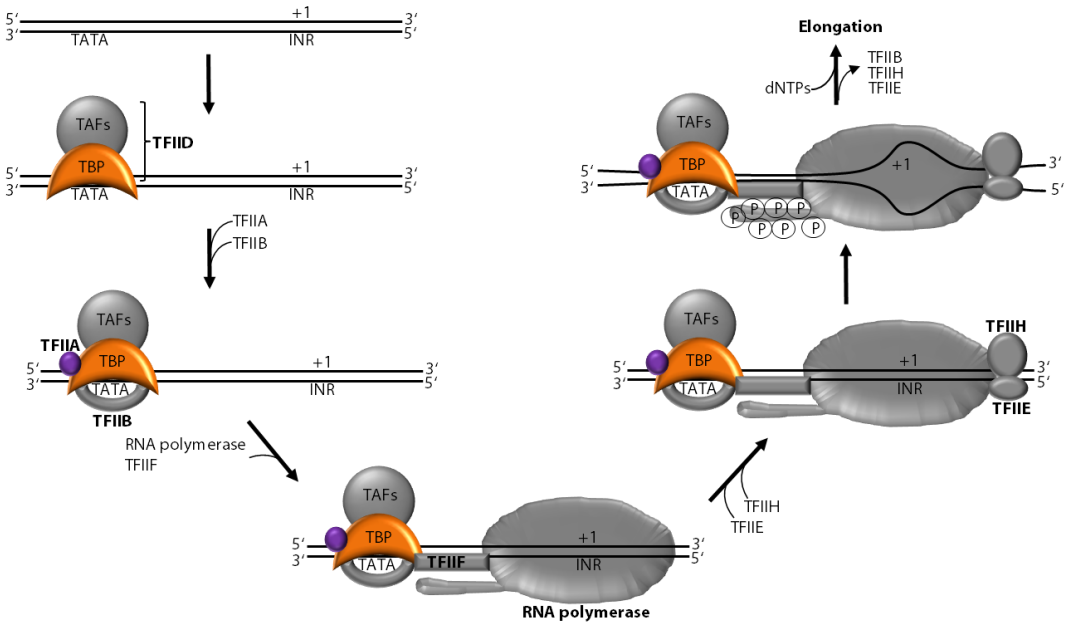


Figure 4.1 Schematic representation of transcription initiation. The first steps of initiation involve the binding of TBP to the TATA-box. Together with TBP, TBP-associated factors form the transcription factor IID. Subsequently, TFIIA and TFIIB bind to the DNA and stabilize the interaction with TBP. TFIIF and the RNA polymerase are recruited in the next step followed by TFIIH, including the helicase subunit, and TFIIE. Upon phosphorylation of the C-terminal domain of the RNA polymerase, the transcription starts and the RNA polymerase is released from the transcription start site and moves along the DNA.

4.1.2 Regulation of transcription initiation

Longterm regulation of protein expression is mainly performed on the level of transcription initiation. The strength of the promoter sequences and conformational rearrangements of the chromatin structure have regulatory roles. The earliest regulation can take place during binding of the promoter associated protein (TBP). TBP binds to the minor groove of the TATA-box promoter introducing an 80° bend in the DNA in yeast (102° in humans) (Kim and Burley, 1994; Kim et al., 1993; Masters et al., 2003; Parkhurst et al., 1996; Wolner and Gralla, 2001), which coincides with transcriptional activity (Wu, 2001). This conformational rearrangement leads to the recruitment of additional transcription factors. Many regulatory processes take place on this level of DNA recognition by TBP. In addition to the aforementioned TBP-associated factors, other positive (transcription activating) and negative (inhibiting) cofactors for TBP exist. Some cofactors act to destabilize the bending of the DNA (Schluesche et al., 2007), while others lead to the removal of TBP (Auble and Hahn, 1993). However, the concrete molecular mechanism of many of these cofactors is still unknown to this day. Besides cofactors regulating the formation of the pre-initiation complex, the DNA carries enhancer and silencer sequences up to 1000 basepairs away from the transcription start site, which are recognized by different cofactors. By folding back of the DNA, these can directly influence the pre-initiation complex.

Figure 4.2 Interaction of TBP with DNA. Structure of TBP (orange) bound to a stretch of DNA (black) with the respective schematic represenation. PDB: 1NH2.

4.1.3 Effects of transcriptional cofactors of TBP

A multitude of cofactors, including the negative NC2 or modifier of transcription (Mot1) and positive cofactors PC1-PC4, that regulate the binding of TBP to the promoter site is known (Kaiser and Meisterernst, 1996). The exact regulatory mechanism of many of these cofactors is yet to be described.

Mot1 is a 210 kDa large protein that is highly conserved among all eukaryotes. Its C-terminal Snf2/Swi2 ATPase domain binds to the minor groove of the DNA upstream of the TATA-box introducing conformational rearrangements (Auble and Hahn, 1993; Auble et al., 1994; Darst et al., 2001; Davis et al., 1992; Gumbs et al., 2003; Moyle-Heyrman et al., 2012; Sprouse et al., 2006; Wollmann et al., 2011). Contacts to TBP are mediated by several HEAT repeats in the N-terminal region (Butryn et al., 2015; Wollmann et al., 2011). A structural loop between the HEAT repeats 2 and 3 can block DNA/TBP association by binding to TBP. Mot1 can bind to TBP in its free form or to DNA/TBP complexes and induce its dissociation from the DNA by ATP hydrolysis (Auble et al., 1994; Davis et al., 1992; Gumbs et al., 2003; Poon et al., 1994). Stable DNA/TBP/Mot1 complexes could thus not easily be isolated. However, the

addition of NC2 can stabilize the complex. Recently, a stable DNA/TBP/Mot1/NC2 complex could be isolated from chromatin extracts (van Werven et al., 2008) and a crystal structure of the complex from *Encephalitozoon cuniculi* was obtained (Butryn et al., 2015).
NC2 forms a heterodimer of the subunits α (22 kDa) and β (20 kDa), which are structurally very similar to the histone complex H2A/H2B (Kamada et al., 2001). It has a high affinity for DNA/TBP complexes, binding to TBP with its C-terminal α-helices as well as its histone-like N-terminus (Meisterernst et al., 1991). It can inhibit transcription initiation by a sterical hindrance for the binding of TFIIB and possibly TFIIA (Masson et al., 2007). SpFRET experiments showed, that NC2 can induce the diffusion of TBP along the DNA without dissociating it from the nucleotide sequence (Schluesche et al., 2007).

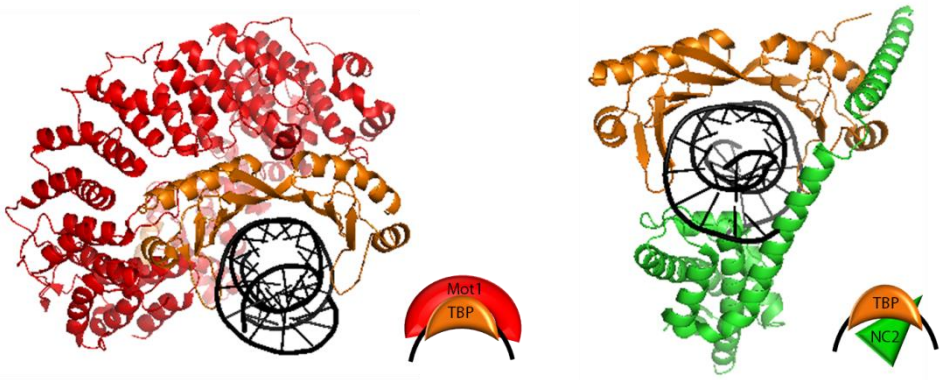


Figure 4.3 Structures and schematics of the cofactors Mot1 and NC2 bound to DNA/TBP complexes. Left: N-terminal domain of Mot1 from *Encephalitozoon cuniculi* (red) bound to a DNA (black)/TBP (orange) complex. PDB: 4WZS. Right: A complex of *Homo sapiens* DNA (black), TBP (orange), and NC2 (green). PDB: 1JFI.

Both cofactors have been associated with positive and negative roles in transcription initiation. Due to its TBP-dissociating effect, Mot1 was proposed to have inhibitory effects on transcription and might be responsible for the redistribution of the transcription initiation factor on the DNA (Auble and Hahn, 1993; Darst, 2003; Zentner and Henikoff, 2013). A redistributory function of Mot1 for TBP towards low-affinity sites was proposed by Collart (1996). Distributing TBP along the DNA can have activating as well as inhibitory effects dependent on the site of alteration. In addition, an analysis of gene transcription levels suggested, that Mot1 might work as an activator *in vivo* (Madison and Winston, 1997).
Mot1 as well as NC2 could also inhibit transcription by sterically hindering the assembly of the preinitiation complex (Auble, 2009). The aforementioned diffusion of TBP along the DNA in the presence of NC2 can lead to negative or positive effects for transcription, dependent on the direction of diffusion (Auble, 2009; Schluesche et al., 2007). In agreement with this, biochemical experiments with both proteins, Mot1 and NC2, suggest a repressing function of these cofactors on gene expression (Dasgupta et al., 2005; Davis et al., 1992; Geisberg et al., 2001, 2002; Spedale et al., 2012). On the other hand, chromatin-immunoprecipitations found NC2 and Mot1 primarily on actively transcribed promoters (Andrau et al., 2002; Creton, 2002; Geisberg et al., 2002; van Werven et al., 2008) and for one family of promoters an activation of preinitiation complex assembly by NC2 was observed (Masson et al., 2007). Hsu et al. (2008) investigated the inhibiting role of Mot1 and NC2 on TATA-containing promoters and activating effects on DPE-dependent genes and

Zhou et al. (2009) postulated that Mot1 and NC2 can work as a promoter remodeling complex to position TFIID at different promoter elements.

The wide variety of approaches used to investigate these cofactors revealed many facets and roles of these proteins in transcription initiation. Their combined effect can differ for the multitude of existing promoter sequences in eukaryotes. The basic molecular mechanism of the individual cofactors on TBP, however, can be investigated *in vitro* and give a general understanding of their function. Here, NC2's and Mot1's function on TBP bound to TATA-containing promoters was analyzed using spFRET of immobilized DNA molecules. Peter Schlüsche could already describe the underlying molecular mechanism of NC2 on TBP using this approach (Schluesche, 2007; Schluesche et al., 2007). Gregor Heiss performed subsequent analysis on the conformations of ternary DNA/TBP/Mot1 complexes (Heiss, 2011). He found a conformational rearrangement of DNA/TBP upon the binding of Mot1 in the absence of ATP (Figure 4.4A). Upon addition of ATP, the original FRET efficiency of the DNA/TBP complex was restored, while Mot1 stayed bound to the complex. By performing spFRET experiments of the complex in the presence of ATPγS, the binding of ATP was shown to suffice to induce the conformational change. Gregor Heiss showed, that on the single molecule level, the hydrolysis of ATP by a single Mot1, either bound to the complex or freely diffusing in solution, was not sufficient to induce dissociation of TBP from the complex (Figure 4.4B). In both cases the survival of fluorescent molecules on the surface did not differ from the rate of photobleaching. The basic molecular mechanism of Mot1's function in TBP dissociation is investigated in more detail in this chapter.

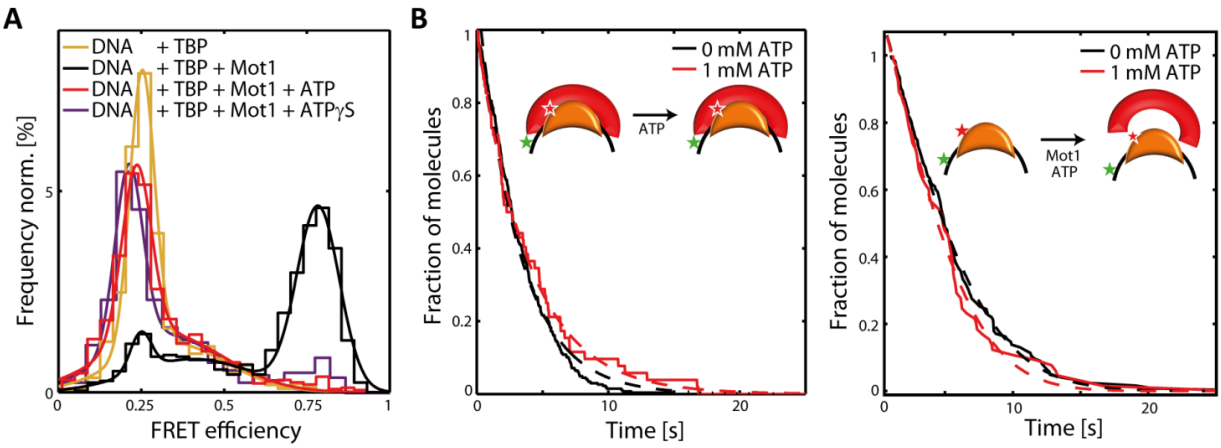


Figure 4.4 The effect of Mot1 on the DNA/TBP complex. (**A**) Moleculewise FRET efficiency distributions of DNA/TBP complexes in the absence (yellow) and presence of the cofactor Mot1 (black) and the ternary DNA/TBP/Mot1 complex upon addition of ATP (red) or ATPγS (magenta). (**B**) Fraction of fluorescently-labeled TBP molecules detected over the course of a measurement. Either a ternary DNA/TBP/Mot1 complex was formed (left, red) and ATP was flown into the sample chamber during the experiment or DNA/TBP complexes were bound to the surface and a mixture of ATP and Mot1 was given into the chamber (right, red). To account for photobleaching, these survival probability histograms were compared to control experiments without the addition of molecules during the course of the measurement (black). Data by Gregor Heiss (Heiss, 2011).

4.2 Results

4.2.1 Dissociation of TBP from DNA/TBP complexes upon cofactor addition

The early transcription initiation complex was investigated using spFRET between a DNA containing the histone 2B (H2B) promoter (DNA18, Chapter 2.3) and TBP-C61 labeled with the fluorophores Atto488, Atto532, or Atto647N. The lifetimes of stable DNA/TBP complexes were described to lie in the range of 15 minutes to 1 hour (Patikoglou et al., 1999; Perez-Howard et al., 1995; Sprouse et al., 2008). We ensured TBP activity after attachment of the fluorophores by investigating its binding to the TATA-box in electrophoretic mobility shift assays. The gradually increasing amount of DNA/TBP complex formation by the addition of increasing concentrations of the protein was observed (Figure 4.5). While the number of complexes increased (red arrow in Figure 4.5), the amount of free DNA decreased in the presence of TBP. The fluorescently labeled TBP thus effectively bound to TATA-containing DNA.

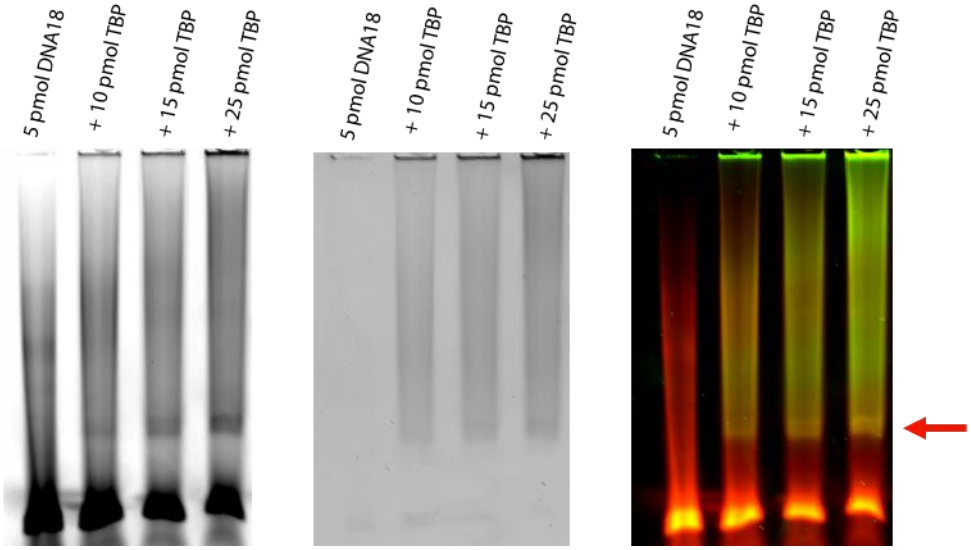


Figure 4.5 Binding of TBP to DNA observed by an electrophoretic mobility shift assay. The binding of different concentrations of TBP-Atto532 to DNA18 labeled with Atto647N was investigated. Left: Fluorescence of the DNA was visualized by excitation at 633 nm in the detection window of a bandpass filter with a mean transmission wavelength of 670 nm and a width of 30 nm (670/30). Middle: Fluorescence of TBP after excitation at 532 nm and detection 580/30. Right: Merged illustration of both fluorescence images. TBP in green and DNA in red. The red arrow marks the band, where DNA/TBP complexes are observed.

In spFRET experiments on a TIRF microscope, the DNA/TBP complex exhibited a low FRET efficiency of 0.25-0.3. Binding of Mot1 to the complex induced conformational rearrangements, which led to a FRET efficiency of around 0.78. Whenever Mot1 bound to ATP or a non-hydrolyzable ATP analogue, the original conformational state of the DNA/TBP complex was reinforced, suggesting the binding of ATP to restore the original conformation (Figure 4.4A) (Heiss, 2011). Three-color experiments showed that Mot1 was still bound to the complexes in this low FRET state (Heiss, 2011). In contrast to ensemble experiments on DNA/TBP/Mot1 complexes where ATP induced the removal of TBP from the DNA (Darst et al., 2001; Gumbs et al., 2003; Viswanathan and Auble, 2011), the binding and hydrolysis of

ATP by Mot1 was not sufficient in these single molecule experiments for the dissociation of TBP (Figure 4.4).

The effect of other TBP cofactors on the conformation of the DNA/TBP complex was investigated using spFRET on a TIRF microscope. Besides Mot1, none of the cofactors NC2, TFIIA, or PC4 introduced a similarly large conformational change observable by spFRET between the DNA and TBP. NC2 and TFIIA, as well as their combined action, left the FRET efficiency of DNA/TBP largely unaffected with a value of around 0.25-0.3 (Figure 4.6A). The binding of PC4 to the DNA/TBP complex induced a second small subpopulation with a FRET efficiency of around 0.5 (Figure 4.6B), while the major population with a FRET efficiency of 0.25-0.3 was still present. Interestingly, PC4 affected the conformation of DNA/TBP/Mot1 complexes. The subpopulation at FRET efficiency around 0.3 was largely unaffected by Mot1, when PC4 was present, and the population with a FRET efficiency of 0.78 represented only approximately 14% (Figure 4.6B). PC4 thus either inhibited the binding of Mot1 to the complex or its addition to the DNA/TBP/Mot1 complex induced a conformational rearrangement otherwise only observed by ATP binding to Mot1.

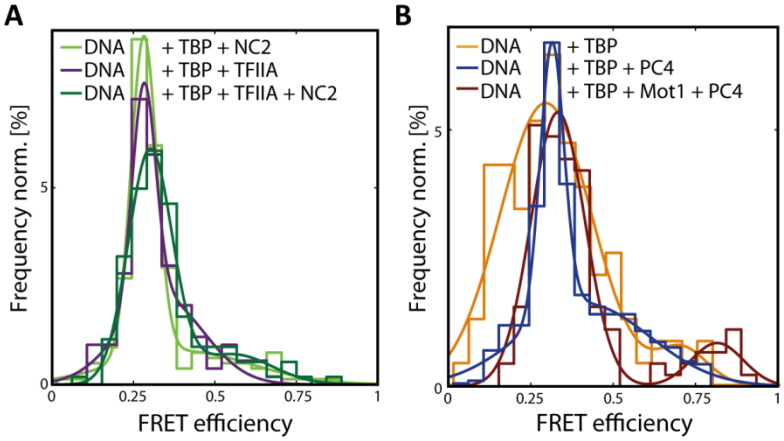


Figure 4.6 Moleculewise FRET efficiency histograms of DNA/TBP complexes in the presence of different transcriptional cofactors of TBP. (**A**) DNA/TBP complexes in the presence of NC2 (light green) and (dark green)/or (violet) TFIIA. (**B**) DNA/TBP complexes (yellow) upon addition of PC4 alone (blue) or PC4 in combination with Mot1 (dark red).

In vitro ensemble experiments showed the role of Mot1 in the dissociation of TBP from the TATA-box in the presence of ATP (Auble and Hahn, 1993; Auble et al., 1994; Darst, 2003; Darst et al., 2001), while the complex did not disassemble in spFRET experiments upon addition of Mot1 or ATP alone (Figure 4.4B). Instead, a conformational rearrangement of the complex was observed (Figure 4.4A). The ATPase activity of a single Mot1 on the complex was not sufficient to dissociate TBP. Gregor Heiss observed the dissociation of TBP from the ternary DNA/TBP/Mot1 complex only in the presence of Mot1 and ATP in solution, meaning that two Mot1 molecules are required for efficient displacement of TBP from the DNA (Figure 4.7A) (Heiss, 2011). This behavior was further investigated here by analyzing the fluorescence intensity of TBP labeled with Atto647N and bound to DNA18-Atto532 in TIRF experiments. Excitation by the green laser in the first few frames ensured that the TBP under investigation was bound to DNA, while the intensity of labeled TBP over time was analyzed

by direct excitation of the donor. Nucleotides and proteins were added to the solution during the course of the experiment by the use of a simple flow system. Dissociation rates for complexes with a lifetime shorter than the rate of photobleaching, analyzed separately in each experiment, could be determined.

The observed dissociation in the presence of Mot1 on the complex and in solution was shown to be dependent on the hydrolysis of ATP by Mot1 rather than on the binding of ATP alone. When comparing the survival of molecules on the surface, the addition of ATP and Mot1 in solution induced dissociation observed as a difference from the survival probability by photobleaching. ADP or ATPγS in combination with Mot1, on the other hand, were not sufficient to induce the dissociation of TBP from the DNA (Figure 4.7B, C). These results agree well with the previouly observed necessity of ATP hydrolysis for effective complex disassembly (Auble et al., 1994). We further investigated increasing concentrations of Mot1 and ATP in solution. The dissociation rate linearly correlated to the concentration of Mot1 and ATP in solution, suggesting the local limitation of these factors at the DNA/TBP complexes on the surface (Figure 4.7C, D). To determine, whether other cofactors of TBP similarly induced the dissociation of TBP from ternary DNA/TBP/Mot1 complexes, NC2 and PC4 were flown into the sample chambers of preformed complexes in combination with ATP. These cofactors, as well as additional freely diffusing TBP were not able to dissociate TBP from the DNA (Figure 4.7D). Additional control measurements in the presence of DNA and additional proteins not associated with DNA binding can be found in the Appendix.

We could show, that for the disassembly of DNA/TBP complexes a Mot1 molecule bound the complex as well as additional Mot1 and ATP in solution are required. The cofactors NC2 and PC4 cannot induce dissocation from preformed DNA/TBP/Mot1 complexes.

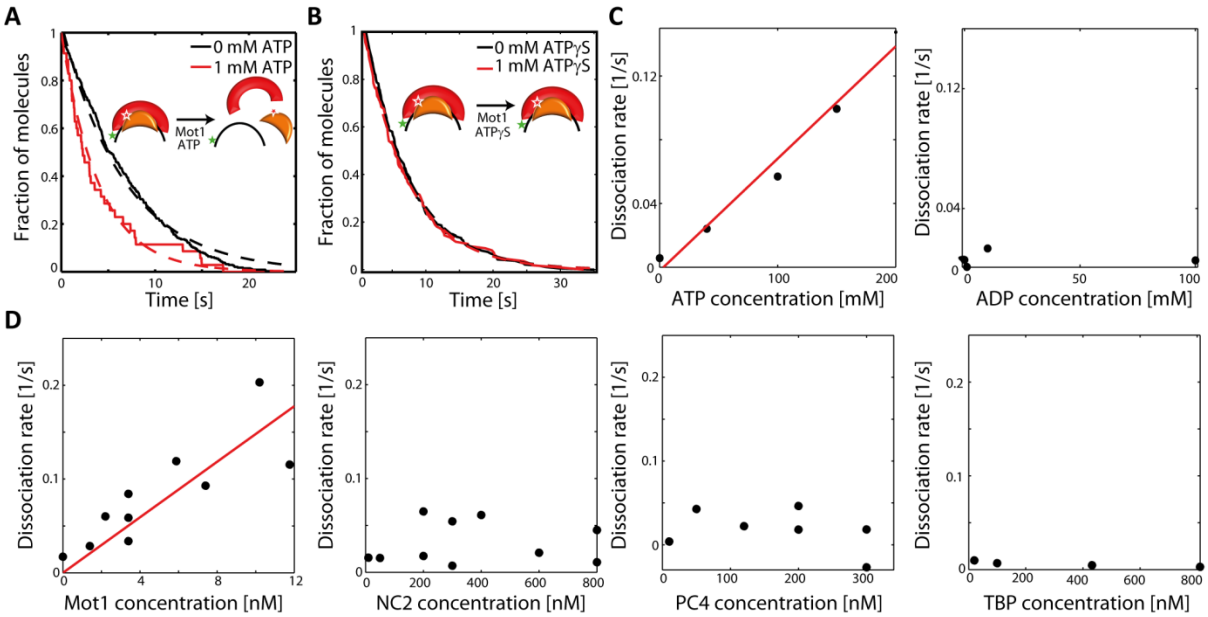


Figure 4.7 Dissociation of TBP from the DNA/TBP/Mot1 complex upon protein and ATP addition and hydrolysis. (**A**) Survival time histogram of TBP bound to DNA upon flow of Mot1 and ATP (red) compared to the photobleaching rate (black). (**B**) Survival time histogram of TBP bound to DNA upon Mot1 and ATPγS addition (red) in comparison to photobleaching (black). (**C**) Dissociation rate of TBP from DNA for different ATP (left) or ADP (right) concentrations added in combination with 3.4 nM Mot1. (**D**) Dissociation rates of TBP dependent on the concentration of the different proteins Mot1 (left), NC2 (second from left), PC4 (second from right), and TBP (right) in solution in combination with 1 mM ATP.

We further inquired, at which step the hydrolysis of ATP is performed by Mot1. To this end, ATP and Mot1 were sequentially added to the ternary DNA/TBP/Mot1 complex in the flow chambers. The sequential addition of ATP in the first round and Mot1 in a second step led to the dissociation of TBP in the second step (Figure 4.8A). Since excess nucleotides and proteins were removed after each step, this implies, that ATP is bound and hydrolyzed by the DNA/TBP-bound Mot1 rather than Mot1 in solution. After the addition of ATP and Mot1 most TBP molecules had dissociated from the DNA, some complexes, however, remained to which a second round of ATP was added. The second addition of ATP did not lead to any further dissociation of TBP, which is well in agreement with previously observed results for single additions of ATP without freely diffusing Mot1.

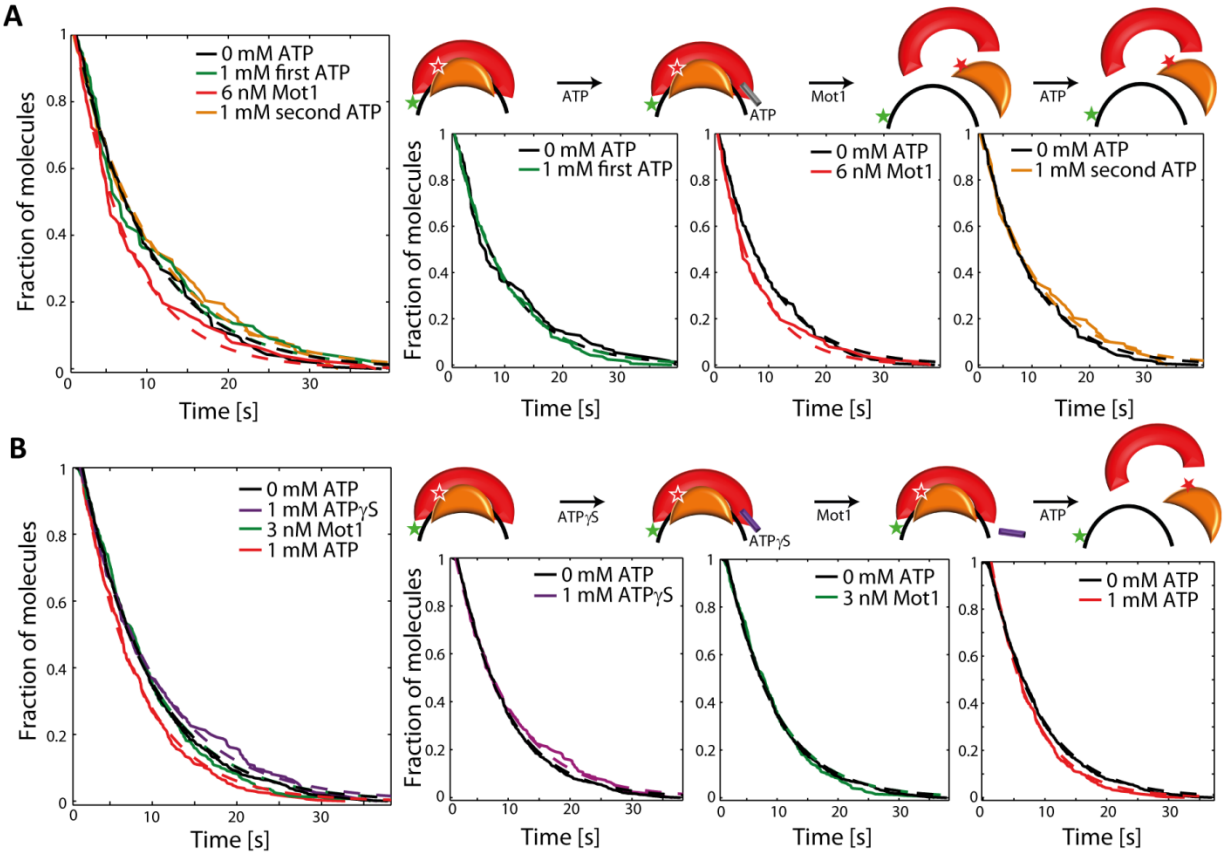


Figure 4.8 Dissociation of TBP from DNA/TBP/Mot1 complexes by hydrolysis of ATP in the presence of Mot1 in solution. (**A**) Sequential addition of ATP, Mot1, and ATP to the ternary DNA/TBP/Mot1 complex. Shown is the fraction of TBP molecules remaining bound to the DNA upon addition of ATP (green), Mot1 (red), and again ATP (yellow). (**B**) Fraction of TBP molecules bound to DNA upon the stepwise addition of ATPγS (magenta), Mot1 (green), and ATP (red).

Experiments in the presence of ATPγS stressed the importance of ATP hydrolysis during the process of TBP removal. When this non-hydrolyzable ATP analogue was added to preformed DNA/TBP/Mot1 complexes, dissociation could not be observed even when in a second step Mot1 was added (Figure 4.8B). A following addition of ATP led to a minimally increased dissociation rate possibly due to the exchange of Mot1-ATPγS to Mot1 free to bind ATP and thus freely dissociating Mot1.

These results suggest, that ATP hydrolysis by Mot1 bound to the ternary complex is required for the effective dissociation of TBP from the DNA. The simple binding of ATP to DNA/TBP/Mot1 complexes or the possible hydrolysis of ATP by Mot1 in solution are not sufficient for complex disassembly. Additionally, the hydrolysis of ATP by freely diffusing Mot1 is highly unlikely to preserve low levels of energy consumption in the cell.

Previous studies described the TBP-dissociating role of Mot1. Interestingly, it was recently observed, that a single Mot1 bound to the DNA/TBP complex even upon ATP hydrolysis did not induce this disassembly, but rather forced the complex into a different conformational state (Heiss, 2011). Preformed DNA/TBP/Mot1 complexes could be dissociated when around 10 nM Mot1 in combination with ATP were present in solution. The additional Mot1 was either important to induce the hydrolysis of ATP by Mot1 bound to the complex or the disassembly of the primed complex with Mot1 having already hydrolyzed ATP. Other cofactors such as NC2 or PC4, or TBP itself were not able to initiate the dissociation of TBP from ternary DNA/TBP/Mot1 complexes. This clearly indicates the importance of Mot1's function on transcription initiation.

4.2.2 Dynamics of DNA bound to TBP

TBP binding to the DNA introduces a bending angle between 80° and 102° for different eukaryotic organisms (Kim and Burley, 1994; Kim et al., 1993; Masters et al., 2003; Parkhurst et al., 1996; Wolner and Gralla, 2001). Recently, this bending was also observed in spFRET studies (Blair et al., 2012; Gietl et al., 2014; Heiss, 2011). Heiss (2011) observed a population of dual-labeled DNA molecules undergoing conformational switching between states of different FRET efficiency in the presence of TBP. Two subpopulations, each transversing between three different, but distinct FRET states, were found. The experiments suggested the described bidirectional binding of TBP to the DNA (Cox et al., 1997; Kays and Schepartz, 2000). This hypothesis was supported by the effect of TFIIA which shifted the equilibrium between both subpopulations to TBP bound primarily in the transcription-inducing orientation well in agreement with Cox et al. (1997), Kays and Schepartz (2000), and Spencer and Arndt (2002). We repeated the experiments of dual-labeled DNA with unlabeled TBP and obtained 55% dynamic DNA molecules. Molecules exhibiting static FRET could not be distinguished from free DNA by the FRET efficiency; with high likelihood these might thus not be bound by TBP (Parkhurst et al., 1996).
Here, an H2B promoter complex dual-labeled with a FRET pair upstream and downstream of the TATA-box (DNA 25.3) was used in combination with Atto488-labeled TBP. We first tested the complex formation of DNA 25.3-Atto565/Atto647N with TBP-Atto488 using three-color FRET experiments in solution (Figure 4.9A). The low concentration of labeled TBP necessary to perform spFRET experiments in solution was below the binding affinity for TBP to the DNA. From the stoichiometry information S_{GR} and S_{BG}, a significant amount of triple-labeled complexes could not be detected in these measurements. The FRET efficiency E_{GR} of all dual-labeled DNA molecules in the presence of these low concentrations of TBP did not deviate from the one of the DNA in the absence of TBP (Figure 4.9A). The dynamics

of the DNA were investigated using immobilized DNA/TBP complexes on the TIRF microscope. The same Atto532/Atto647N-labeled DNA as in Heiss (2011) was used. The complexes were illuminated by 488 nm excitation in the first 50 frames, followed by excitation at 532 nm. Single molecules were detected in the green and red detection channel. The concentration of TBP-Atto488 had to be lowered when compared to experiments with unlabeled TBP. This reduced the number of dynamic DNA complexes to 28%. An exemplary intensity time trace of one of these complexes is displayed in Figure 4.9B. The donor-acceptor intensity signals are clearly anticorrelated, while the total intensity stays constant, leading to three distinct FRET states. A framewise FRET efficiency histogram of more than 600 dynamic complexes showed at least two states with FRET efficiencies of around 0.6 and 0.75 (Figure 4.9C). HMM analysis revealed two populations with three states each for the excitation of the complexes by the 532 nm laser corresponding to the observations in Heiss (2011), although for some molecules only two of the states were populated. These two subpopulations are displayed in blue and green, respectively, in a transition density plot (Figure 4.9D). The FRET efficiency distributions for the individual subpopulations are displayed in Figure 4.9E. When the proximity ratio between Atto532 and Atto647N on the DNA after TBP-Atto488 excitation was analyzed, a slight shift in the mean proximity ratio could be observed when comparing both subpopulations (Figure 4.9F). This shift could be explained by the closer proximity of Atto488 to Atto647N, when TBP was bound to the DNA in the transcription-competent conformation. Due to FRET from Atto488 to Atto647N, more photons were therefore emitted in the red channel for this subpopulation. At the same time, however, one has to consider, that due to this orientation of TBP, Atto488 has a larger distance to Atto532, therefore the energy transfer between these dyes and thus the energy transfer from Atto532 to Atto647N is lower. Furthermore, the signal of the blue fluorophore is detected in the green detector due to their small wavelength separation. These aspects can explain why the shift observed between the two populations was very small. Since the signal after blue excitation also exhibited dynamic switching between three states, the interpretation of this small shift in proximity ratio between these populations has to be taken with care. However in combination with the data obtained by adding TFIIA to the DNA/TBP complex, where the stabilization of the complex with TBP in the transcription-inducing orientation (blue subpopulation) was observed (Heiss, 2011), it provides a clear hint for the TBP-orientation specific subpopulations of the DNA.

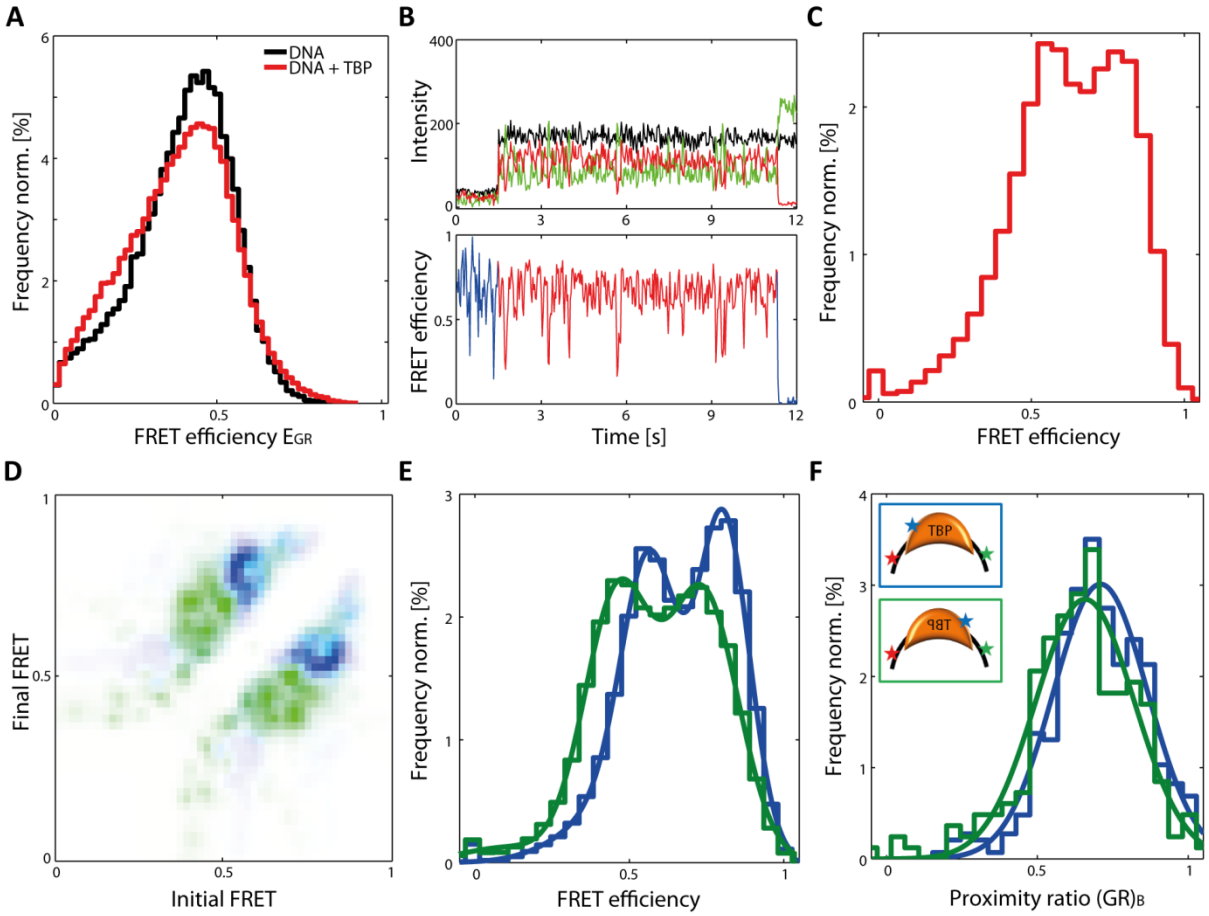


Figure 4.9 Three-color experiments of dual-labeled H2B promoter DNA with TBP-Atto488. (**A**) SpFRET efficiency distribution of DNA 25.3-Atto565/Atto647N in the absence and presence of TBP-Atto488 measured by PIE-MFD. (**B-F**) SpFRET experiment of DNA25.3-Atto532/Atto647N in the presence of TBP-Atto488 performed by TIRF microscopy. (**B**) Exemplary intensity trace over time of the DNA after 488 nm excitation ($t \leq 1.5$ s) or 532 nm excitation ($t > 1.5$ s). (**C**) Framewise FRET efficiency histogram of 616 dynamic DNA molecules. (**D**) Transition density plot of the molecules from (C) color-coded in blue (393 molecules) and green (223 molecules) for the subpopulations undergoing transitions between FRET states around 0.3/0.6/0.75 or 0.25/0.45/0.7, respectively. (**E**) Framewise FRET efficiency distributions of the individual subpopulations defined in (D). (**F**) Histogram of the moleculewise proximity ratio GR from the first 1.5 s after 488 nm excitation for both populations. The populations are associated with TBP bound to the DNA in two different orientations (schematic representations as inlet).

4.2.3 Outlook on the effect of twisting of the DNA during bending by TBP

Several previous studies described the bending of DNA by TBP (Kim and Burley, 1994; Kim et al., 1993; Masters et al., 2003; Parkhurst et al., 1996; Wolner and Gralla, 2001). In this paragraph, we investigated whether twisting along the strands of the DNA plays a role in the bending process induced by TBP binding. To analyze the DNA twisting, we used a DNA composed of complementary strands, which was however not ligated at one position five bases upstream of the TATA-box. The previously used sequence, DNA18, was compared to an additional promoter sequence, DNA B11(nick). These two DNAs thus differed in their nucleotide sequence. DNA18(nick) and DNA B11(nick) carried the fluorescent label and the site of the strand break at the same position with respect to the TATA-box. DNA18 and DNA18-2(nick) shared the same nucleotide sequence and were both fluorescently-labeled five bases downstream of the TATA-box. They only differed in the existence of a non-ligated

region in one of the DNA strands. Schematic representations of the DNAs are found in Figure 4.10A.

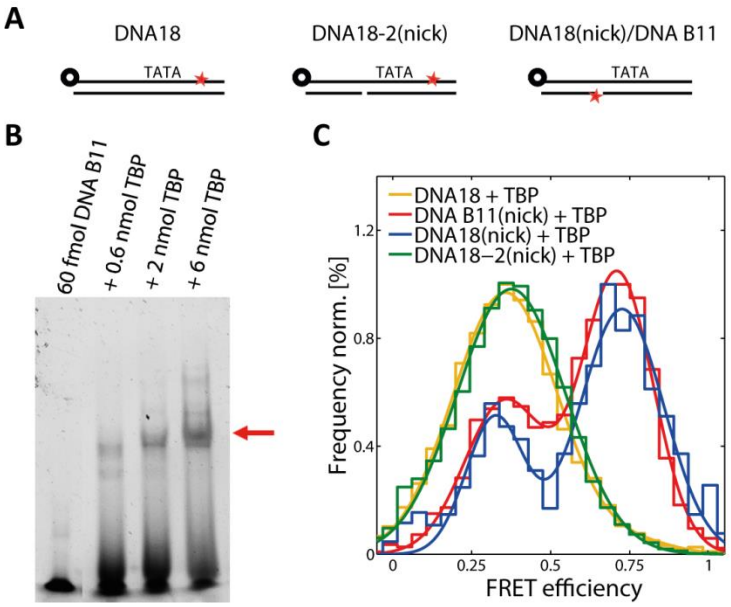


Figure 4.10 Investigation of the influence of DNA twist on DNA bending upon TBP binding. (**A**) Schematic representations of the different DNA constructs under investigation. DNA18 always shared the same nucleotide sequence. DNA18 and DNA18-2(nick) carried the fluorescent label at the same position downstream of the TATA-box, while it was placed upstream for DNA18(nick) and DNA B11(nick) (red star). DNA18-2(nick), DNA18(nick), and DNA B11(nick) contained a non-ligated nucleotide upstream of the TATA-box. The 5'biotin is shown in black. (**B**) EMSA of TBP binding to DNA B11 labeled with Cy5. The DNA/TBP complex is marked by a red arrow. (**C**) Framewise FRET efficiency histograms of TBP bound to the different DNAs DNA18 (yellow), DNA18-2(nick) (green), DNA B11(nick) (red), and DNA18(nick) (blue).

Effective binding of DNA B11 by TBP was ensured in an EMSA (Figure 4.10B). We then compared the FRET efficiencies of the completely ligated DNA18 to its counterparts carrying an non-ligated region. When comparing the DNAs with fluorescent labels at the same position with respect to the TATA-box (DNA18 versus DNA18-2(nick) and DNA B11 versus DNA18(nick)), the FRET efficiency histograms did not differ significantly (Figure 4.10C). DNA18/TBP complexes as well as DNA18-2(nick)/TBP complexes both showed a mean FRET efficiency of ~0.4. DNA B11(nick)/TBP as well as DNA18(nick)/TBP complexes were observed in two states with mean FRET efficiencies of ~0.35 and ~0.7. The insertion of an unligated nucleotide pair in one of the strands therefore did not affect the bending behavior of TBP on the DNA. If twisting of the strands by binding of TBP occurs, it could not be observed by a change in the FRET efficiency when only a single bond in the DNA is missing. Additional experiments with strand breaks in different positions of the expected bending region or with multiple strand breaks could be performed to further investigate whether twisting of the DNA plays an important role in the DNA/TBP interaction. Additionally, the effect of twisting in the DNA strand could be observed by spFRET of dual-labeled DNA with unligated regions.

A comparison between DNA B11(nick) and DNA18(nick) showed that the FRET efficiency distributions did not differ with respect to the nucleotide sequence upstream and downstream of the TATA-box. The two populations with FRET efficiencies of ~0.35 and ~0.7 were

obtained for both complexes, DNA18(nick)/TBP and DNA B11(nick)/TBP, with similar amplitudes (Figure 4.10C).

4.3 Discussion

Cofactors of TBP are highly important in the regulation of transcription initiation. The process is strictly controlled to regulate protein expression at low energy costs for the cell. Many different effects have been described for the TBP-binding cofactors NC2 and Mot1 regarding their overall influence on transcription (Auble, 2009). A few studies were performed that investigated their structure and functional mechanism (Butryn et al., 2015; Kamada et al., 2001; Moyle-Heyrman et al., 2012; Wollmann et al., 2011). However, a detailed understanding of NC2's and Mot1's molecular function on individual promoters was only recently offered by studying DNA/TBP complexes by spFRET (Heiss, 2011; Schluesche, 2007; Schluesche et al., 2007). Moreover, the experiments revealed that Mot1 bound to the ternary DNA/TBP/Mot1 complex needs to hydrolyze ATP in the presence of additional Mot1 for effective dissociation of TBP to occur. We could show that the ATPase activity of Mot1 alone is not sufficient for removal of TBP from the complex. Mot1 bound to ATP does, however, prime the complex for disassembly, possibly by weakening its interactions. We could exclude the necessity of multiple rounds of ATP hydrolysis due to high concentrations of ATP in solution, except if the hydrolysis of ATP needs to be induced by an additional factor. Further investigations are needed to determine the exact timing of ATP hydrolysis. It is possible, that ATP is already hydrolyzed in the DNA/TBP/Mot1 complex in the absence of freely diffusing Mot1, although a difference in the FRET efficiency distribution between ATP- and ATPγS-bound complexes was not observed (Figure 4.4). The induction of hydrolysis by the presence of at least two Mot1 molecules in proximity of the complex, however, is equally possible and can explain the necessity of additional Mot1 in solution. Regarding the cellular context, the concentration of free Mot1 could indicate the metabolic state of the cell or its time point in the cell cycle. The regulation of ATP hydrolysis by the concentration of free Mot1 in the nucleus could therefore link the cellular conditions to the level of transcription initiation. It was shown that salt-dependent changes in the arrangement of nucleosomes limit the binding of TBP to the TATA-box (Hieb et al., 2014). Before nucleosome remodeling can occur, other factors thus have to be responsible for prior TBP removal from the DNA.

Interestingly, the combined addition of Mot1 and PC4 to DNA/TBP complexes diminished the effect of Mot1 on the conformation of the complex. The main population was still found in the low FRET conformation, with only 14% of the molecules shifting to the high FRET state observed for DNA/TBP/Mot1 complexes. As for other cofactors of TBP, PC4 has been associated with activatory as well as inhibitory roles in transcription (Liao et al., 2011; Malik et al., 1998; Wu, 1998; Wu and Chiang, 1998). The opposing action of PC4 and Mot1 offers an interesting glimpse on the cooperative actions of different cofactors. Further experiments are required for a detailed understanding of the combined molecular mechanism of a combination of cofactors. Future experiments can involve the addition of a multitude of cofactors up to whole nuclear extracts to spFRET experiments.

We further investigated the conformational state of the DNA upon TBP and cofactor binding. TBP binding to the DNA induces, via two intermediate structures, a bend in the DNA (Kim and Burley, 1994; Kim et al., 1993; Parkhurst et al., 1999). The DNA bending is attributed to the insertion of a set of phenylalanines between the base pairs at each side, introducing an angle of 45° each (Kim and Burley, 1994; Kim et al., 1993; Wu, 2001). Surface-tethered motion experiments on the complex revealed a three-step binding of TBP to the DNA (Tolić-Nørrelykke et al., 2006). A first intermediate was observed after the insertion of a phenylalanine residue at the upstream region of the TATA-box. This insertion is followed by placement of the phenylalanines at the downstream end. With this, one of the TBP domains rotates relative to the other and van der Waals contacts between the DNA and TBP are formed (Kim and Burley, 1994; Tolić-Nørrelykke et al., 2006). Additional intermediate steps were observed in spFRET experiments (Schluesche et al., 2008). Interestingly, drastic differences between the binding of eukaryotic and archaeal TBP (Gietl et al., 2014) and the one of yeast and human TBP (Masters et al., 2003; Whittington et al., 2008) were observed, with human TBP introducing a single homogeneously bent DNA (Blair et al., 2012). We performed preliminary experiments to investigate the role of DNA twisting in the bending reaction. To this end, the FRET efficiency between TBP and the DNA upstream and downstream of the TATA-box was analyzed. DNAs that were not completely ligated should show differences in their bending behavior and thus the FRET efficiency should change, if twisting of the DNA is important for the bending reaction. In our experiments, a twisting could not be observed by the difference in the FRET efficiency of DNAs completely ligated compared to those carrying a single strand break.

Additionally, in the here described experiments, stably bound TBP was shown to introduce conformational dynamics on dual-labeled DNA on the timescale of seconds (Heiss, 2011). Two subpopulations of complexes, switching between three different FRET states each, were observed. These subpopulations could be attributed to orientation non-specific binding of TBP. This phenomenon is already known from literature (Cox et al., 1997; Kays and Schepartz, 2000). The addition of TFIIA stabilizes the transcription-competent orientation of TBP (Cox et al., 1997; Kays and Schepartz, 2000; Spencer and Arndt, 2002), and was observed in the spFRET experiments to increase the fraction of molecules in one of the subpopulations (Heiss, 2011). In this chapter, results on the dynamics of dual-labeled DNA in the presence of fluorescently labeled TBP were described. The attachment of fluorophores to TBP did not hinder orientation-independent TBP binding or the dynamics of the DNA and therefore dynamic complexes switching between the same three FRET states were found in two subpopulations, although some of the molecules only showed transitions between two of the three FRET states (Figure 4.9). The results presented small evidence, in form of the proximity ratio after excitation of the fluorophore on TBP, on the orientation of TBP binding for the two subpopulations. Only in combination with results obtained in the presence of TFIIA, however, can the data be clearly linked to the orientation of TBP on the DNA.

5 Dynamics of U2AF65 in mRNA recognition for intron splicing

After transcription of pre-mRNA, RNA maturation occurs inside the nucleus before translation into protein is achieved in the cytosol. mRNA maturation includes the modification of both ends of the RNA as well as intron splicing. The process of intron splicing is highly regulated by a large RNA- and protein-containing splicing complex, which is very selective and yet flexible due to constant interchange of its subunits. As one of the early steps during splicing, the 3'splice site is recognized by the U2 auxiliary factor (U2AF). The large and small subunits of the U2AF heterodimer bind the polypyrimidine (Py) tract and the AG dinucleotide, respectively. The minimal RNA-binding region of the large subunit, U2AF65(RRM1,2), exists in an open conformation when bound to this Py-tract region, while it adopts a closed state in the absence of RNA (Mackereth et al., 2011). The conformations of U2AF constructs were investigated here using spFRET. We found conformational dynamics, which were intrinsic to RRM1,2 and not induced by the binding of RNA. The RNA substrates, however, were able to shift the equilibrium of the dynamics between these two conformational states. An increasing amount of time spent in the open conformation corresponded to a higher affinity of the U2AF construct or the Py-tract itself. The correlation between the conformation of RRM1,2 and the Py-tract affinity revealed the molecular mechanism of Py-tract recognition. Differences in the equilibrium for strong and weak Py-tracts gave a glimpse on the importance of U2AF for selecting *bona fide* 3'splice sites in the context of alternative splicing.

We investigated the molecular mechanism of U2AF in collaboration with Professor Michael Sattler, Carolina Sanchéz Rico, Hyun-Seo Kang, and Lisa R. Warner (Technische Universität München, Germany).

5.1 Introduction

Protein expression is a necessary process for cell survival and cellular differentiation. This ubiquitous process can be divided into two major steps. Initially, the DNA is transcribed into precursor (pre)-mRNA in the nucleus of eukaryotic cells (Chapter 4). Secondly, matured mRNA is translated into protein by ribosomes in the cytosol of the cell. This chapter investigates the processing of the pre-mRNA into a mature mRNA ready to be transported into the cytosol for translation.

5.1.1 mRNA maturation

Stretches of DNA carrying genetic information for the expression of proteins, are transcribed into pre-mRNA in the eukaryotic nucleus. This RNA is then directly bound by heterogeneous ribonucleoproteins and serine-arginine-rich proteins and further processed inside the nucleus. RNA processing includes three modifications of the pre-mRNA, namely 5'-capping, 3'-polyadenylation, and intron splicing. As a first step after the polymerisation of around 25 nucleotides, the 5'-end of the RNA is capped by a modified guanine as a protection of the

RNA from degradation and as a ribosome-recognition signal as well as to distinguish mRNA from other types of RNA (Alberts, 2015). During the process of transcription, the RNA is cleaved internally by the spliceosome (Chapter 5.1.2) and finally the 3'-end of the RNA is processed after recognition of the termination signal. The enzyme complexes cleavage stimulation factor (CstF) and cleavage and polyadenylation specificity factor (CPSF) move along with the RNA polymerase until they specifically recognize and bind to the termination sequence, where the RNA is cleaved. Modifications at this 3'-end of the RNA are dominated by the addition of multiple (around 200) adenines by the poly-A polymerase (PAP) without the requirement of a template strand. Polyadenylation including the binding of poly-A-binding proteins aids in the export of the RNA from the nucleus as well as the protection from degradation. RNA processing automatically involves a change of the proteins bound to the RNA, whereby the cell can recognize completion of the RNA maturation. Cap-binding complexes, exon junction complexes, and poly-A binding proteins mark mature mRNAs. Only fully mature mRNA is exported from the nucleus through nuclear pore complexes (NPCs) where it is recognized by transport receptors and actively transported (Zhao et al., 2002). All other RNA segments remaining in the nucleus are degraded by the nuclear exosome.

5.1.2 RNA splicing

mRNA processing is dominated by RNA splicing. Pre-mRNAs consist of a sequence of coding, expressed (exons) and non-coding intervening (introns) sequences. These were first described as mosaic genes by Richard Roberts and Phillip Sharp in 1977 (Berk and Sharp, 1977; Berget et al., 1977; Chow et al., 1977). On average, a eukaryotic pre-mRNA sequence contains 8000 nucleotides, while a corresponding final eukaryotic mRNA has a size of 1200 nucleotides (Alberts, 2015). This clearly demonstrates the importance of RNA splicing. Introns play an important role in the evolutionary process of the emergence of new proteins as well as in the increased coding potential of eukaryotic genomes due to alternative splicing, which occurs in 75-90% of all human genes (Smith and Valcárcel, 2000; Graveley, 2001; Black, 2003; (Johnson et al., 2003a) Blencowe, 2006; Kim et al., 2006). Their size can vary from 10 to over 100000 nucleotides, requiring highly specific and flexible splicing reactions. The exon size, on the other hand, is more uniform with a size of about 150 nucleotides. Both, exons and introns are recognized and marked during RNA elongation. The actual splicing reaction however can occur posttranscriptionally.

5.1.3 Spliceosome-assisted splicing

The recognition of splice sites occurs inside the nucleus, while the RNA polymerase still processes the pre-mRNA towards the 3'end. The RNA polymerase II carries components of the spliceosome bound to its phosphorylated tail. These proteins are directly transferred to the RNA during synthesis to correctly recognize all introns and prevent exon skipping. The splicing reaction, by which a single intron is removed, can take place co- or posttranslationally. The actual splicing reaction is achieved by two phosphoryl-transfer reactions (S_N2-type transesterifications) (Figure 5.1). In a first step, the 2'-hydroxyl group of

an adenosine in the branch point sequence attacks the phosphodiester bond at the 5'splice site leaving a free 5'exon and an intron lariat structure bound to the 3'exon. The 3'hydroxyl group of the 5'exon then attacks the phosphodiester bond of the 3'splice site excising the intron and joining the exons.

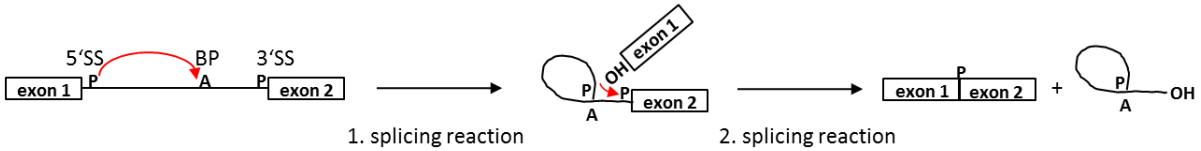


Figure 5.1 S_N2-type transesterification reactions during pre-mRNA splicing by the spliceosome. In a first step the phosphate group (P) of the 5'splice site attacks the adenosine at the branch point to form a lariat (red arrow in first step). In a second step, the free hydroxyl group of the 5'splice site attacks the 3'splice site at the phosphodiester bond (red arrow in second step). Both exons are ligated and the intron is removed. Exons are displayed as boxes and introns are represented by black solid lines.

Splicing is highly regulated and involves a complex machinery of five uridine-rich RNAs (U1, U2, U4, U5, U6) and over 200 proteins (Nilsen, 2003; Wahl et al., 2009; Will and Luhrmann, 2011). The large number of splicing factors is required for the correct and at the same time highly flexible recognition of intron sequences. Splicing factors recognize introns at three specific conserved regions: the 5'splice site, the branch point sequence, and the 3'splice site (Wahl et al., 2009). The five specialized small nuclear RNAs (snRNAs), which in complex with their binding proteins form the core of the spliceosome, recognize these sequences. While U1snRNP contacts the 5'splice site and the branch point binding protein (SF1/BBP) binds to the branch point, the 3'splice site is recognized by U2 auxiliary factor (U2AF) during the early steps of spliceosome assembly (Zamore et al., 1992; Zorio and Blumenthal, 1999). During splicing, the spliceosome complex is constantly altered exchanging a variety of subunits and most importantly creating an active catalytic center. Multiple RNA-RNA interactions are exchanged during this process (e.g. U1 is replaced by U6) which ensures a high splicing accuracy due to multiple sequence verifications. The assembly of this large multisubunit complex requires energy to break RNA-RNA interactions, which is provided by ATP hydrolysis.

An overview of intron removal by the spliceosome is given in Figure 5.2. The recognition of the splice sites is achieved in a first step by the binding of U1 snRNP to the 5'splice site (Hoskins et al., 2011). The base pairing at this site and the 5'end of the U1 RNA is stabilized by proteins. SF1/BBP and U2AF bind cooperatively to the branch point sequence and Py-tract, respectively. U2AF35 additionally contacts the conserved AG dinucleotide of the 3'splice site (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999). This early 5' and 3'splice site recognizing spliceosome is called complex E. Upon binding of U2, SF1/BBP leaves the complex and frees the branch point (BP) sequence for U2 snRNA base pairing. As soon as the U2 snRNP uses ATP to form interactions with the branch point sequence, which are stabilized by a protein complex of SF3a and SF3b as well as the RS-rich domain of U2AF65, the complex A is formed (Gozani et al., 1996; Valcárcel et al., 1996). SF3b 155 is responsible for the interaction with the C-terminal RRM of U2AF65 (Gozani et al., 1998). In a following step, a preassembled complex of U4/U5/U6 snRNPs is recruited to the intron to form a catalytically inactive complex B (Hoskins et al., 2011). Only upon

activation by release of a total of around 35 proteins including U1 and U4 snRNP and binding of twelve other factors, the spliceosome becomes reactive (complex B*) (Fabrizio et al., 2009). After the first transesterification, the spliceosome is called complex C. The interaction of the AG dinucleotide with U2AF35 is replaced by additional splicing factors (Umen and Guthrie, 1995). Subsequent rearrangements lead to the second splicing reaction after which the spliceosome dissociates (Konarska et al., 2006). Once a splicing reaction is completed, snRNPs are disassembled from the lariat by ATP hydrolysis and proteins form an exon junction complex (EJC) at the site of exon junction.

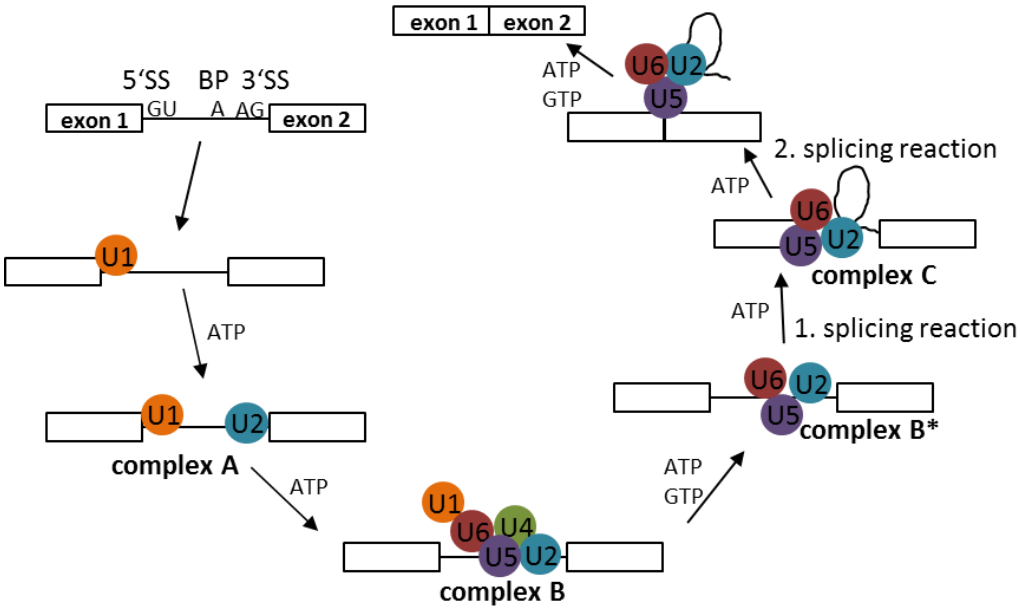


Figure 5.2 Removal of an intron by the spliceosome. U1 recognizes the 5'splice site, followed by U2 binding to the 3'splice site and SF1/BBP for the branch point sequence. After base pairing, the complex A is formed. U4/U5/U6 assemble at the complex and activation is achieved by release of U1 and U4. Both transesterification reactions are performed by rearrangements to the active catalytic forms of the spliceosome complex required. After removal of the intron as a lariat structure and exon ligation, the splicing subunits dissociate from the RNA. Exons are displayed as boxes and introns are shown as black lines. Steps in which energy is required are marked in a simplified manner by the nucleotides used for hydrolysis.

A few reoccurring motives become apparent when investigating mRNA splicing. Firstly, a delicate network of RNA-RNA interactions is formed, where the 3' and 5'splice sites and the BP are recognized by multiple binding partners. Secondly, these direct interactions are often weak and stabilized by a combination of multiple interactions. Thirdly, as a consequence of these interactions, other interactions are broken by the hydrolysis of nucleotides and the rearrangement of the spliceosome is achieved by the replacement of subunits. A detailed description of all proteins involved in each of the complexes as well as the secondary and tertiary structure interactions between RNAs is given in Will and Luhrmann (2011).

Complex eukaryotes carry a second set of snRNPs (U12-type spliceosome) that are responsible for splicing from a small number of intron sequences. These are reviewed in detail in Patel and Steitz (2003).

5.1.4 Self-splicing introns

Self-splicing introns are highly interesting in the context of the early evolution of life in the "RNA world". The enzymatic activity of RNA in combination with its potential to carry genetic information makes it the ideal molecule to achieve both the transfer of information and the performance of reactions. Self-splicing introns perform the splicing reaction in the absence of proteins or other RNAs. These introns are found, for example, in some mitochondrial and chloroplast genes. Two major classes of self-splicing introns are known (Cech, 1986). Group I intron sequences attack the phosphodiester bond at the 5'splice site by an activated guanosine nucleotide. Group II introns, on the other hand, use a reactive adenosine to attack the 5'splice sequence. Without the aid of additional proteins, they can adopt an active three-dimensional structure by folding, in which the reactive groups of the intron are exposed (Toor et al., 2008). In both cases, the free hydroxyl group then induces the second transesterification reaction. Evolutionary, it is very likely that the spliceosome evolved from self-splicing group II introns, where the domains V and VI show high similarity to the U2 and U6 snRNAs (Toor et al., 2008).

5.1.5 The role of U2AF in intron splicing

While in spliceosome-assisted splicing, U1snRNP contacts the 5'splice site and SF1/BBP binds to the branch point, U2AF specifies the upstream exon boundary by binding to the Py-tract and the AG dinucleotide.

U2AF is a heterodimer consisting of a large subunit (U2AF65) and a small subunit (U2AF35) (Figure 5.3,Figure 5.6). Both subunits contain conserved RNA-recognition motifs (RRM), the most abundant eukaryotic RNA-binding sequences. U2AF65 carries two of these RRMs embedded by a C-terminal U2AF homology motif (UHM), an N-terminal UHM-ligand motif (ULM), and an arginine/serine rich (RS) domain. U2AF35 has one RRM, which is framed by two CCCH-type zinc finger motifs, an RS domain, and a glycine tract (Figure 5.3,Figure 5.6) (Kellenberger et al., 2002; Mollet et al., 2006). The tandem RRM1,2 domains of U2AF65 contact the Py-tract at the 3'splice site.

Dimerization between the large and small subunits occurs between the ULM region of U2AF65 and the UHM of U2AF35 (Kielkopf et al., 2001; Corsini et al., 2007). The interaction of U2AF35 with the large U2AF subunit was shown to stabilize binding to the RNA (Zuo and Maniatis, 1996; Rudner et al., 1998). The conserved AG dinucleotide at the 3'splice site is recognized by U2AF35 with high specificity (Wu et al., 1999; Yoshida et al., 2015, 2011; Zorio and Blumenthal, 1999). U2AF65, on the other hand, binds to the Py-tract sequence located at the 3'splice site via its RRM domains.

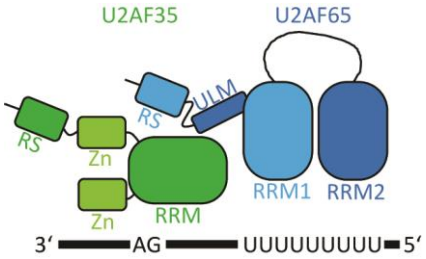


Figure 5.3 Schematic representation of the recognition of the 3'splice site by U2AF. The polypyrimidine tract is recognized by U2AF65 (blue), while the conserved AG dinucleotide is bound by U2AF35 (green).

The Py-tract can be highly degenerate ranging from high affinity (strong) Py-tracts with around eleven uridines to Py-tracts of lower affinity where the sequence of pyrimidines can even be alternated by purine residues. The U2AF complex is therefore required to correctly recognize a wide variety of pyrimidine sequences. The efficiency of the 3'splice site recognition depends on the length and the number of uridines present in the sequence (Zamore et al., 1992; Singh et al., 1995). A difference in the binding affinity of the Py-tract correlates with varying efficiency of splicing, whereby U2AF65's role in alternative splicing is of high importance (Coolidge et al., 1997; Johnson et al., 2003a; Mackereth et al., 2011; Sickmier et al., 2006). The Py-tract thus determines the strength of the intron for inducing the splicing reaction. U2AF65 is sufficient for splicing of certain introns *in vitro*, while U2AF35 is required for splicing of introns with weak (low affinity) Py-tracts and is essential *in vivo* (Guth et al., 1999; Banerjee, 2004; Pacheco et al., 2006). The interaction of U2AF35 with the AG dinucleotide can stabilize the interaction of U2AF65 with weak Py-tracts. The high flexibility in intron recognition is needed for alternative splicing, which in eukaryotes occurs in over 60% and in humans in 75-90% of all genes (Johnson et al., 2003a; Kim et al., 2006; Wang et al., 2008).

5.1.6 Structure and conformation of U2AF

The structural composition of the minimal tandem RRM1,2 construct of U2AF65 was studied in detail by nuclear magnetic resonance (NMR) (Mackereth et al., 2011) and X-ray crystallography (Kielkopf et al., 2001). This minimal region of amino acids 148-342, which has the same structure as the same region in the full-length protein, recognizes the Py-tract at the 3'splice site. Mackereth et al. (2011) found RRM1,2 to exist in two distinct conformations dependent on the presence of a Py-tract sequence. In the absence of RNA, RRM1,2 adopted a closed state in which the RNA-binding surface of RRM1 is hidden by RRM2 (Figure 5.4). In this state, only RRM2's RNA-binding interface is accessible, which allows binding of short Py-tracts (U4) without a conformational change. Whenever a Py-tract exceeding four uridines was added, RRM1,2 was found in an open conformation. In this case, the two domains presented the RNA-binding surface in an extended conformation (Figure 5.4). In a follow-up study, the closed conformation was shown to consist of an ensemble of many different unbound conformations, which U2AF65 adopted (Huang et al., 2014).

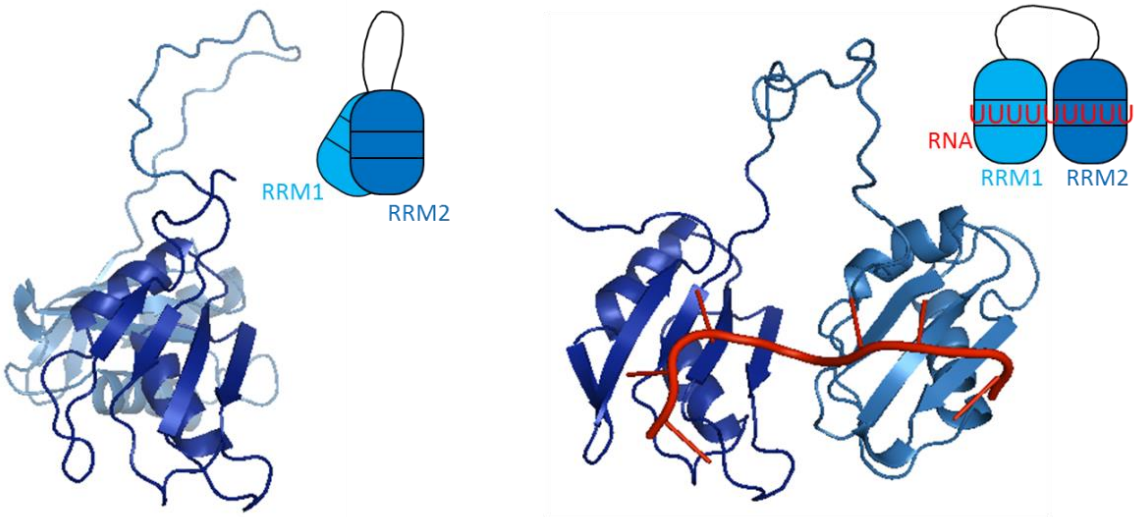


Figure 5.4 Structure and schematics of the RRM1,2 construct in the closed conformation in the absence of RNA (left) and in the open conformation when bound to a Py-tract (right). RRM2 is displayed in dark blue, RRM1 is shown as light blue and the RNA is displayed in red. PDBs: 2YH0, 2YH1.

The interaction of U2AF65 with U2AF35 occurs via U2AF65's N-terminal ULM region (residues 85-112). A crystal structure of U2AF35 bound to the N-terminal polyproline region of U2AF65 was described in Kielkopf et al. (2001) (Figure 5.5). U2AF35(RRM) was found to be unstructured in solution and only assumed its tertiary helix-containing conformation when bound to U2AF65 (Kellenberger et al., 2002).

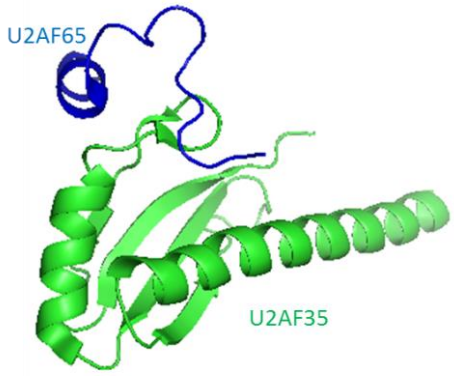


Figure 5.5 Interaction of U2AF35 with the N-terminus of U2AF65. Structure of U2AF35 (green) bound to the proline-rich N-terminus of U2AF65 (blue). PDB: 1JMT.

To understand the molecular mechanism of the function of U2AF65 RRM1,2, we applied spFRET experiments to monitor its conformational changes upon binding to various Py-tracts with varying strength and to U2AF35. Using the full information contained in PIE-MFD experiments of the protein in solution, we could show that the conformational change was initiated by intrinsic conformational dynamics. We could further elucidate the effect of RNA binding on the single U2AF65 subunit and on a complex of U2AF65 and U2AF35.

5.2 Results

5.2.1 Domain constructs of U2AF

To understand the conformational changes of U2AF65(RRM1,2), we employed U2AF constructs with various domain combinations. We used RRM1,2 to analyze the minimal RNA-binding tandem domain (Figure 5.6), because extensive structural studies were performed on this construct (Mackereth et al., 2011; Sickmier et al., 2006). A linker truncation of the inner linker region between amino acids 233 and 252 allowed an investigation of the linker function on the conformational changes. URRM1,2, the minimal RRM1,2 elongated by the ULM region is needed for recognition by U2AF35. The minimal U2AF35(UHM) used in this work consisted of the RRM domain and was deprived of its zinc finger motifs and the arginine-serine rich domain (Figure 5.6). This minimal heterodimeric construct of URRM1,2/U2AF35(UHM) was used in previous studies (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999) and its effect on conformational changes was probed here.

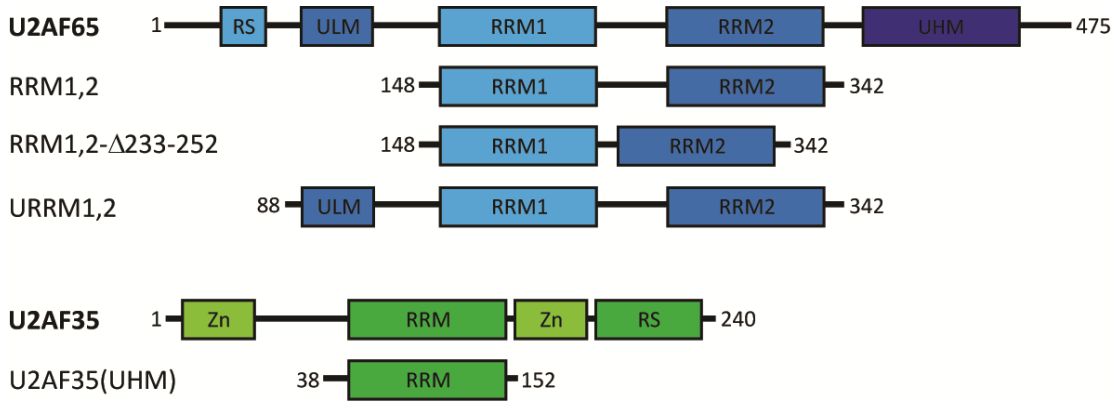


Figure 5.6 Domain composition of U2AF65 (blue) and U2AF35 (green). U2AF65 is composed of two RRMs framed by an RS domain, a ULM, and a UHM domain. The minimal RNA-recognizing construct, RRM1,2, consists of both RRM domains. A construct minimizing the linker between both domains was named RRM1,2-Δ233-252. URRM1,2 was used for dimerization with U2AF35. U2AF35 contains an RRM bordered by two zinc-finger domains and an RS domain. U2AF35(UHM) used in this work carried the RRM only.

To perform spFRET experiments, fluorophores were attached to the cysteines specifically engineered in both RRM domains of all U2AF constructs. Two constructs of U2AF65 were investigated with labeling positions C187/C318 or C187/C326 (Figure 5.7). The dye pair of donor and acceptor fluorophore was attached to these positions stochastically. During the investigations, we tested a multitude of different donor/acceptor dye pairs to exclude any dye-specific artifacts. This chapter will mainly focus on the RRM1,2-C187/C326 construct with Atto532 and Alexa647 as a FRET pair, except noted otherwise. Additional results with other RRM1,2 constructs and fluorophores can be found in Chapter 5.4 and in the Appendix.

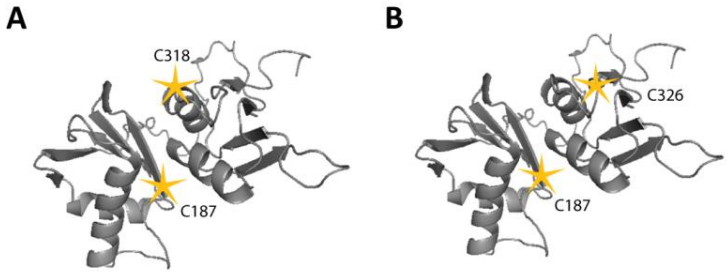


Figure 5.7 RRM1,2 constructs with two different labeling positions. (**A**) RRM1,2 is shown with labeling positions C187 and C318 for the donor-acceptor dye pair. (**B**) RRM1,2 where the fluorophores were bound to residues C187 and C326. PDB: 2YH0.

To determine the fluorophore positions and actual distance values, accessible volume calculations of the fluorophores attached to the RRM1,2 constructs were performed (Kalinin et al., 2012). By simulating the accessible volume of the fluorophores when attached to the structures of RRM1,2, the average distances of the closed and open conformation were estimated as 35-42 Å and 66-70 Å, respectively.

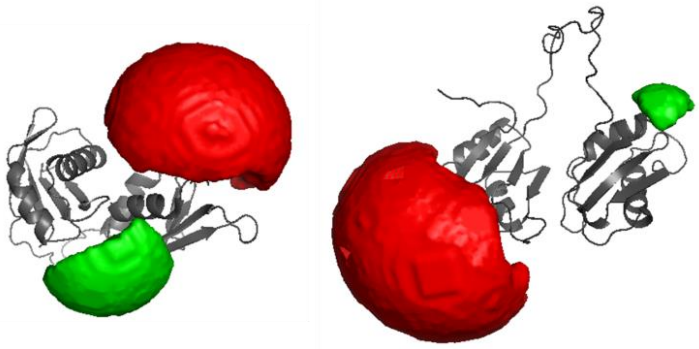


Figure 5.8 Accessible volumes of fluorophores bound to RRM1,2. RRM1,2 structure in the closed (left) and open (right) conformation with accessible volumes of the fluorophores at positions C187 and C326 displayed in green and red, respectively. PDB: 2YH0, 2YH1.

5.2.2 Conformational states and dynamics of RRM1,2

RRM1,2 was previously observed by NMR in a closed state in the absence of RNA and in an open conformation when bound to RNA (Mackereth et al., 2011). The conformational transition between these states was investigated in detail using spFRET. By the use of spFRET, conformational changes induced by the binding of RNA can be distinguished from intrinsic conformational dynamics of RRM1,2. The dual-labeled RRM1,2 constructs were measured on a confocal microscope with PIE-MFD.

The FRET efficiency of the free RRM1,2-C187/C326 was high with a distribution around a mean value of 0.78 (Figure 5.9A). Including the lifetime information obtained by PIE-MFD experiments, a two-dimensional histogram of FRET efficiency in dependence of donor lifetime was obtained. A static FRET line connects all static FRET species, while molecules undergoing a change in FRET efficiency during their transit time in the focal volume deviate to the right of this line due to intensity-weighted calculations of the donor lifetime (Chapter 2.10.2) (Kalinin et al., 2010; Kudryavtsev et al., 2012). The deviation from the static FRET

line in Figure 5.9B clearly shows that the free RRM1,2 molecules underwent conformational dynamics on the timescale of micro- to milliseconds. The obtained FRET efficiency is thus an average value of the conformational states between which the transitions occur.

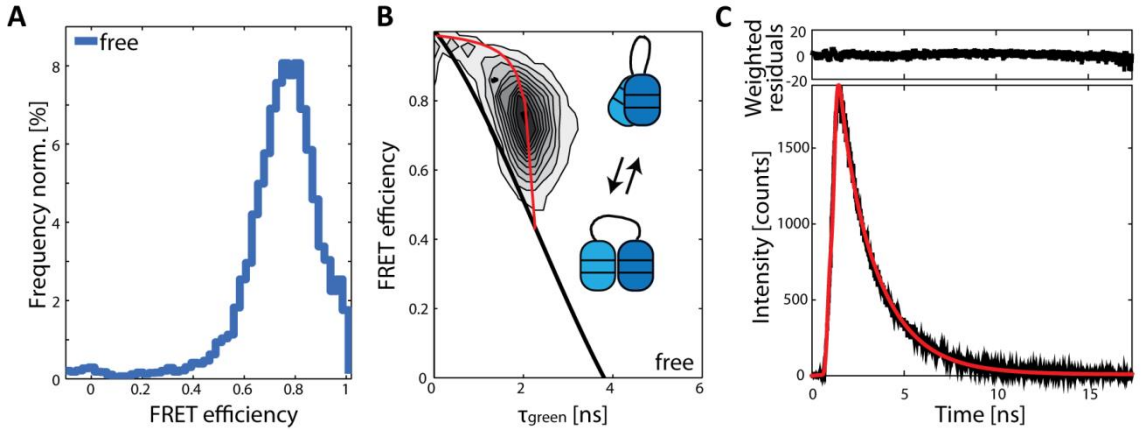


Figure 5.9 FRET efficiency and lifetime information for RRM1,2 obtained from MFD-PIE experiments in solution. (**A**) FRET efficiency histogram of RRM1,2 fluorescently labeled at positions C187 and C326. (**B**) Two-dimensional histogram of FRET efficiency versus donor lifetime. The population of RRM1,2 molecules deviates from the static FRET line (black) and can be described by a dynamic FRET curve (red). Insets show schematic representations of RRM1,2 in the closed (upper icon) and open (lower icon) state. (**C**) Microtime histogram of the donor lifetimes of RRM1,2 fit by a biexponential decay function to determine the lifetimes of the closed and open conformation.

To determine the donor lifetimes of the open and closed conformation, we used the combined lifetime information of all molecules in a microtime histogram. To obtain enough statistics to extract fluorescence lifetimes using biexponential fit functions these microtime histograms were used. A fit of the histogram, deconvoluted by the instrument response function, to a biexponential function is displayed in Figure 5.9C. Two donor lifetimes of 0.14 ns and 2.15 ns could be determined from this fit. The corresponding FRET efficiencies of the closed and open conformation could therefore be calculated as 0.96 and 0.43, respectively (Table 5.1). These FRET efficiencies display the fully closed and open conformational states between which the transitions occur. The corresponding distances (36 Å in the closed and 65 Å in the open state) for both FRET efficiencies agree well with the distances obtained from accessible volume calculations (35-42 Å and 66-70 Å) (Chapter 5.2.1). To illustrate the closed conformation of the complex, the FRET efficiency histogram of RRM1,2 molecules was filtered by the donor lifetime (τ_D < 0.1 ns) (Figure 5.10A). U13ACAGG-bound RRM1,2 with τ_D > 0.1 ns was used to display the open conformation. The equilibrium of RRM1,2 in the open conformation versus the closed state could be estimated by linearly combining the FRET efficiencies obtained for the fully open and closed state to fit the value of the obtained mean FRET efficiency.

The free form of RRM1,2 spent on average 34% of its time in the open state, thus mainly populating the closed conformation with high FRET efficiency.

Table 5.1 Overview of multiple parameters determined from spFRET experiments using PIE-MFD of different RRM1,2 constructs in the absence or presence of substrate.

Protein	RNA	⟨E⟩	τ_1 [ns] [a]	τ_2 [ns] [b]	E_1 [c]	E_2 [d]	D_1 [Å] [e]	D_2 [Å] [f]	A(τ_1)
RRM1,2 C187-C326	-	0.783 ± 0.004	2.15	0.14	0.43	0.96	64.8	36	0.34 ± 0.02
	U9	0.50 ± 0.02	2.11	0.10	0.44	0.97	64.3	34	0.89 ± 0.06
	U4A8U4	0.632 ± 0.008	2.28	0.13	0.40	0.97	66.3	35	0.59 ± 0.04
	U13ACAGG	0.460 ± 0.008	2.14	-	0.44	-	64.7	-	0.96 ± 0.06
	U4A8U5ACAGG	0.49 ± 0.02	2.26	0.15	0.41	0.96	66.1	36	0.85 ± 0.06
	U4A8U4ACAGG	0.51 ± 0.01	2.21	-	0.42	-	65.5	-	0.84 ± 0.05
URRM1,2 C187-C326	-	0.725 ± 0.005	2.11	0.16	0.45	0.96	64.3	37	0.46 ± 0.03
	U9	0.48 ± 0.02	2.12	0.18	0.44	0.95	64.4	37	0.92 ± 0.06
	U4A8U4	0.55 ± 0.01	2.35	0.21	0.38	0.95	67.2	39	0.69 ± 0.05
	U13ACAGG	0.431 ± 0.007	2.17	0.17	0.43	0.96	65.0	37	1.00 ± 0.06
	U4A8U5ACAGG	0.439 ± 0.009	2.28	0.15	0.40	0.96	66.3	36	0.93 ± 0.06
	U4A8U4ACAGG	0.44 ± 0.02	2.29	0.27	0.40	0.93	66.5	40	0.91 ± 0.06
URRM1,2/ U2AF35(UHM) C187-C326	-	0.642 ± 0.006	2.47	0.22	0.35	0.94	68.7	39	0.51 ± 0.04
	U9	0.432 ± 0.005	2.17	0.17	0.43	0.95	65.1	37	0.99 ± 0.06
	U4A8U4	0.456 ± 0.007	2.17	0.14	0.43	0.96	67.7	38	0.85 ± 0.06
	U13ACAGG	0.417 ± 0.005	2.39	0.21	0.37	0.95	67.7	38	1.02 ± 0.06
	U4A8U5ACAGG	0.405 ± 0.004	2.31	0.20	0.39	0.95	66.7	38	0.97 ± 0.06
	U4A8U4ACAGG	0.412 ± 0.005	2.37	-	0.38	-	67.4	-	0.94 ± 0.06
RRM1,2-Δ233-252 C187-C318	-	0.895 ± 0.004	1.44	0.11	0.62	0.97	58.9	36	0.22 ± 0.01
	U9	0.71 ± 0.02	1.44	0.19	0.62	0.95	59.0	39	0.72 ± 0.04
	U4A8U4	0.770 ± 0.006	1.47	0.10	0.61	0.97	59.2	35	0.57 ± 0.03

The mean FRET efficiency ⟨E⟩ was determined by a Gaussian fit to the FRET efficiency histogram. As an error, the 95% confidence intervall is given. The donor lifetimes (τ_1, τ_2) were obtained by biexponential fits to the microtime histogram (Figure 5.9C). These were used to calculate the FRET efficiencies E_1, E_2, and distances D_1, D_2 of the respective closed and open conformations. The Förster radius of the fluorophores attached to RRM1,2 was determined to be R_0=62 Å and the lifetime of the donor-only population was given as $\tau_{D(0)}$=3.8 ns. Standard deviations of three independent experiments with RRM1,2-Δ233-252 in the presence of U9 were given as errors a=0.08 ns, b=0.08 ns, c=0.02, d=0.02, e=0.8 Å, f=4 Å.The equilibrium between the closed and open conformation was calculated as A(τ_1). The value was calculated as the linear combination of the FRET efficiencies E_1, E_2 ($A(\tau_1)=\frac{\langle E\rangle-E_2}{E_1-E_2}$). Error propagation of errors of ⟨E⟩, E_1, and E_2 was used. Populations with fluorescence lifetimes too small for correct determination are shown as a dash.

We further observed the conformational dynamics of RRM1,2 upon the addition of Py-tracts of different strengths. The mean FRET efficiency shifted to lower values with increasing Py-tract strength (Figure 5.10A). For U4A8U4-bound constructs, the mean FRET efficiency value was 0.63, while it reached 0.5 for constructs bound to U9-RNA. The two-dimensional histograms including the donor lifetime information showed little dynamics, i.e. a population on the static FRET line, in the presence of U9 (Figure 5.10B).

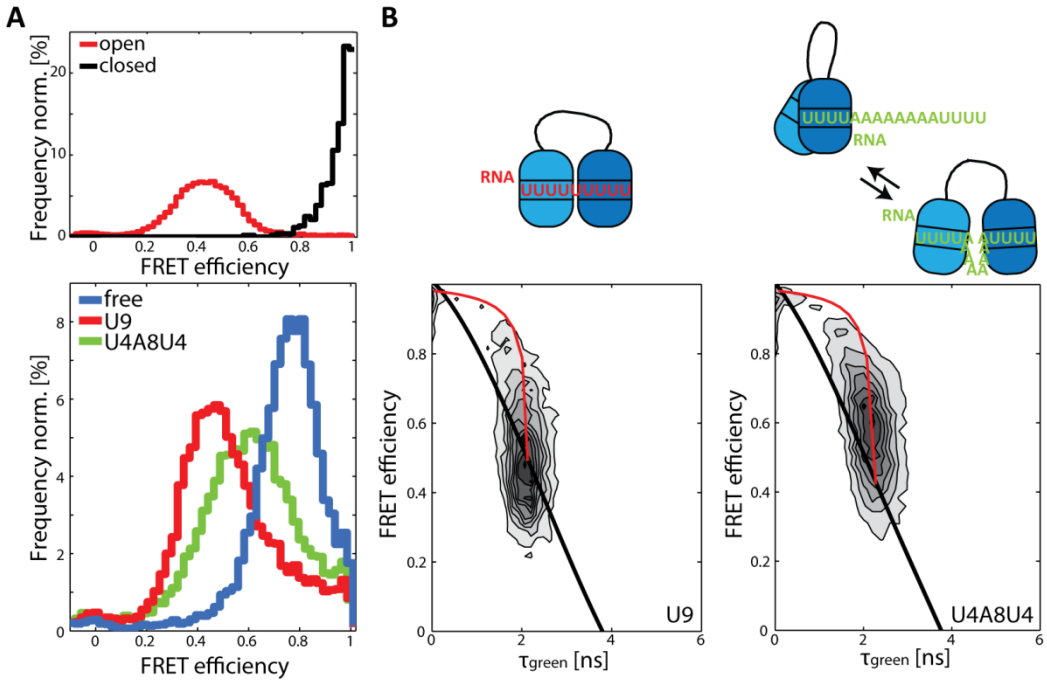


Figure 5.10 FRET efficiency and donor lifetime information of RRM1,2 in the absence or presence of Py-tract RNA measured with PIE-MFD. (**A**) Upper panel: FRET efficiency histogram of the closed (black) and open (red) conformation of RRM1,2 illustrated by using spFRET experiments of RRM1,2 molecules with $\tau_D < 0.1$ ns for the closed conformation and U13ACAGG-bound RRM1,2 with $\tau_D > 0.1$ ns for the open conformation. Lower panel: FRET efficiency histogram of RRM1,2 in the absence (blue) and presence of Py-tracts with different strengths (U9: red, U4A8U4: green). (**B**) Two-dimensional histograms of donor lifetime versus FRET efficiency of the U9-bound RRM1,2 (left) and of the U4A8U4-bound protein (right). The static FRET line is shown in black and a dynamic description of the data is displayed in red. Schematic respresentations of the conformational states are shown as blue icons.

Using the lifetime information, we determined the FRET efficiencies of the closed and open conformation in the presence of U9 as 0.97 and 0.44, respectively (Table 5.1). These values correspond to the ones obtained in the absence of RNA. Interestingly, the conformational states between which the dynamics were measured, were the same independent of the presence of RNA. However, the mean FRET efficiency changed. This means, that the equilibrium between the open and closed conformation shifted in the presence of RNA. For a Py-tract with weak affinity, U4A8U4, RRM1,2 spent on average 59% of its time in the open conformation. A strong Py-tract, U9, shifted the equilibrium, so that the molecules were found in the open state 89% of their time. This corresponds to a stabilization of the open conformation observed for RRM1,2 bound to U9 in the two-dimensional donor lifetime versus FRET efficiency histograms (Figure 5.10B).

We could exclude that the conformational dynamics observed were induced by a complete association or dissociation of the RNA. First of all, they were visible in the absence of RNA. Secondly, the association and dissocation rates determined from TIRF experiments of immobilized Py-tract RNA with RRM1,2 constructs associated to it were too slow. The half-life of the complexes was shown to be around 1 s (Table 5.2, Figure 5.11A), a time which does not correspond to the conformational dynamics observed on the timescale of micro- to milliseconds.

Table 5.2 Affinity, dissociation, and association rates of U2AF constructs with U9.

Protein	K_D [µM]	Number of molecules	t/2 [s]	k_{off} [1/s]	k_{on} [1/(s·µM)]
RRM1,2	1.34 ± 0.03	190/171	0.68 ± 0.05	1.02 ± 0.05	0.76 ± 0.06
URRM1,2	0.22 ± 0.01	140/211	0.8 ± 0.1	0.9 ± 0.1	4.1 ± 0.6
URRM1,2/U2AF35(UHM)	0.15 ± 0.01	251	0.76 ± 0.06	0.92 ± 0.06	6.1 ± 0.6

The affinity constants K_D of RRM1,2, URRM1,2, and URRM1,2/U2AF35(UHM) were determined from isothermal titration experiments by Hyun-Seo Kang (Prof. Michael Sattler, Technische-Universität München). The half-life (t/2), and dissociation rate (k_{off}) was obtained from experiments on the TIRF microscope. Biotinylated U9 elevated by a double-stranded RNA stretch was immobilized on the surface of the prism. The dissociation of the protein constructs were determined by exponentially fitting the lifetime of the given number of molecules on the prism surface. Due to use of low laser intensity, the measurement was unaffected by photobleaching. The values for the affinity constants and the dissociation rates were used to determine the association rate k_{on}. The errors given are standard deviations of two independent measurements or determined by error propagation for the heterodimer.

This provides additional evidence that the dynamics observed are intrinsic to the protein construct and not induced by the interaction with RNA. To distinguish, whether the motion in the presence of RNA might be induced by binding and release of one of the RRM subunits, most likely RRM1 due to its lower affinity, we investigated RRM1,2 in the presence of fluorescently labeled Py-tract RNA. Due to the low affinity of the protein to the RNA, however, the RNA concentrations used on the confocal microscope were not sufficient to investigate these triple labeled complexes. While concentrations above around 1 µM of RNA could induce the full conformational transition to the open conformation of RRM1,2, for lower concentrations only a fraction of molecules (500 nM U13) or none of the molecules (100 pM U13) shifted to the open state (Figure 5.11B, C). For the detection of single triple-labeled complexes, an RNA concentration in the picomolar range would be necessary. To obtain results with triple-labeled complexes in the here observed affinity range of a few micromolar, the complexes could be encapsulated in lipid vesicles or investigated using microfluidics.

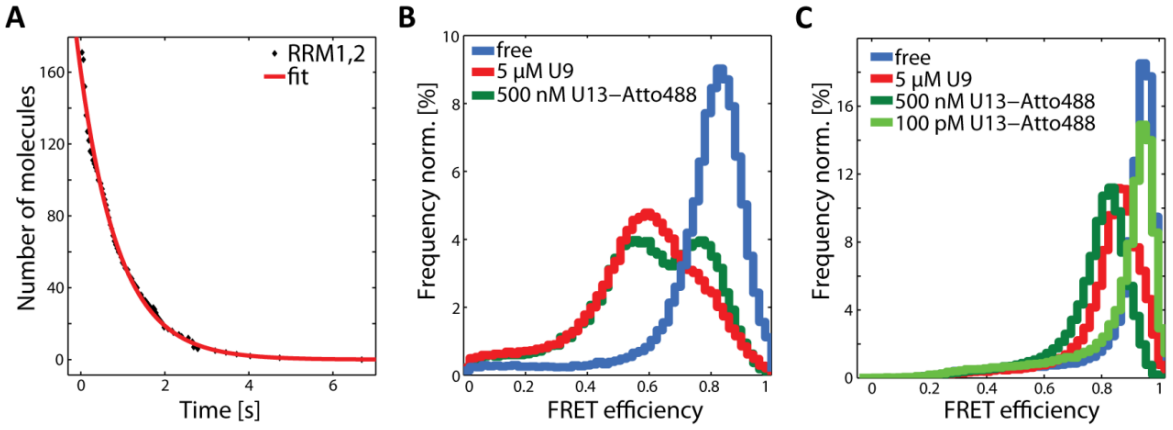


Figure 5.11 Investigations on RRM1,2 binding affinity to RNA. (**A**) Dissociation of RRM1,2-C187/C326 molecules (number of molecules as black diamonds) from U9-RNA attached to the prism surface in a TIRF experiment over time. The half-life and dissociation rate were determined by an exponential fit to the data (red). (**B**) FRET efficiency histograms of RRM1,2-C187/C318 labeled with Atto565/Alexa647 in the absence of RNA (blue), in the presence of 5 µM U9 (red) and in the presence of 500 nM U13-Atto488 (dark green). (**C**) FRET efficiency histograms of RRM1,2-Δ233-252-C187/C318 with Atto565/Alexa647 fluorophores in the absence of RNA (blue), in the presence of 5 µM U9 (red), 500 nM U13-Atto488 (dark green) or 100 pM U13-Atto488 (light green).

5.2.3 Effects of the interdomain linker on the conformational dynamics

RRM1,2 showed conformational switching between a closed and an open state. The conformational states observed by NMR displayed a slight opening and rotating motion of one of the domains around the other stressing the importance of a flexible segment between the domains (Figure 5.4) (Mackereth et al., 2011). The RRM domains are connected by a linker region spanning from amino acid 227 to 261 (Figure 5.6). To further study the significance of this linker region between both RRM domains for the conformational dynamics, a linker truncation mutant, RRM1,2-Δ233-252-C187/C318, was measured. The mutant protein still switched between a closed and an open conformation independent of the presence of substrate. The FRET efficiencies were 0.97 and 0.62 for the closed and open state, respectively, as determined from global lifetime information (Table 5.1). The FRET efficiency of the closed state was the same as observed for the full-length RRM1,2. The open conformation, however, adopted a state with higher FRET efficiency than for RRM1,2 (0.43). The linker length is therefore important for the extension of RRM1,2 in the open state. With missing residues in the linker region, the full extension of the open conformation cannot be reached.

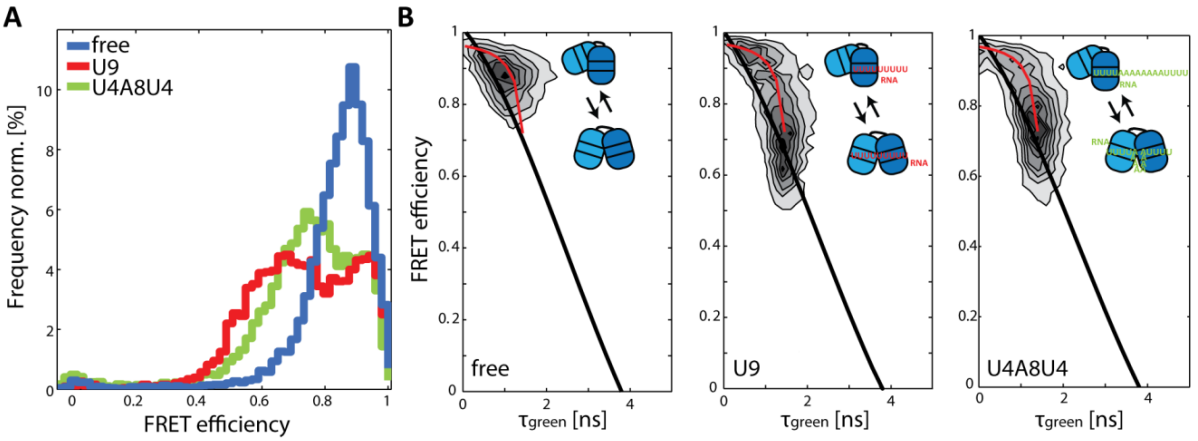


Figure 5.12 FRET efficiency and donor lifetime information of RRM1,2-Δ233-252 obtained from PIE-MFD experiments in solution. (**A**) FRET efficiency histogram of RRM1,2-Δ233-252 in the absence (blue) and presence of the strong Py-tract U9 (red) or the weak Py-tract U4A8U4 (green). (**B**) Two-dimensional histograms of donor lifetime versus FRET efficiency of RRM1,2-Δ233-252 in the absence of RNA (left), bound to U9 (middle) or bound to U4A8U4 (right). The static FRET line is shown in black and the dynamic FRET curve describing the data is displayed in red. Schematic respresentations of the conformational states are shown as blue icons.

Even with a truncation of the linker region, RRM1,2-Δ233-252 was still highly dynamic with molecules spending 22% of their time in the open state (Figure 5.12). Upon addition of Py-tract RNA, the equilibrium shifted towards the open conformation. In the presence of the strong Py-tract U9, RRM1,2-Δ233-252 spent 72% of its time in the open state, while this was populated by around 57% in the presence of U4A8U4. A small part of the population did not shift to the open conformation even in the presence of RNA (Figure 5.12). In these experiments, we cannot distinguish, whether RNA was bound to these RRM1,2-Δ233-252 molecules. Mackereth et al. (2011) suggested a mechanism in which RRM1 could perform a searching motion along the RNA once RRM2 has already bound. Due to the shortened and thus less flexible linker region, this searching motion might be impaired, leading to a lower

probability of stable RNA binding and thus open complex formation, although the dynamic conformational switching was otherwise not affected by the truncation of the linker.

5.2.4 Influence of extended RNA on the conformational dynamics of different U2AF constructs

In a succeeding step, we investigated the influence of RNA substrates elongated by several nucleotides, including the conserved AG dinucleotide, on U2AF constructs. As previously observed, the FRET efficiency of the closed and the open conformation stayed the same for RRM1,2 in the absence and presence of the different Py-tract RNAs. In the presence of these extended Py-tracts, a FRET efficiency of around 0.96 was observed for the closed state and around 0.4 for the open conformation (Table 5.1). The fully closed and open conformation were thus unaffected by the extended RNA. It did, however, shift the equilibrium between the closed and open conformation of RRM1,2. For strong Py-tracts, a stabilization of the complex in the open state on the static FRET line was observed (Figure 5.13B). In the presence of the strong U13ACAGG, the protein was in the open conformation 95% of the time, while U4A8U5ACAGG was in the open conformation 85% of the time (Figure 5.13, Table 5.1)

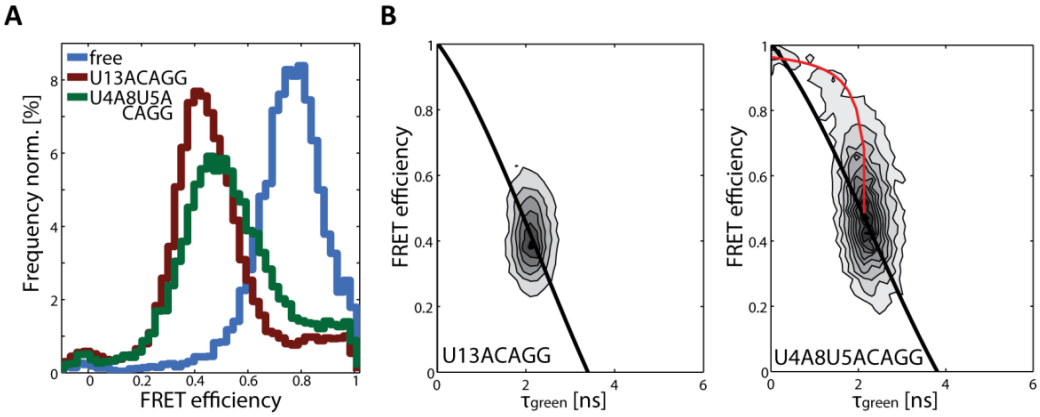


Figure 5.13 Parameters characterizing RRM1,2 in the presence of the long Py-tract RNAs U13ACAGG and U4A8U5ACAGG obtained from PIE-MFD experiments. (**A**) FRET efficiency distributions of RRM1,2 in the absence of RNA (blue), bound to the strong Py-tract U13ACAGG (dark red), and to the weak Py-tract U4A8U5ACAGG (dark green). (**B**) Two-dimensional histograms of donor lifetime versus FRET efficiency of RRM1,2 bound to U13ACAGG (left) or to U4A8U5ACAGG (right). Static populations lie on the static FRET line (black) and molecules undergoing dynamic conformational changes are described by the red curve.

A comparison to the construct URRM1,2, including the U2AF35 binding region (Figure 5.6), showed a similar trend in the presence of extended RNA with an overall slightly shifted equilibrium to the open conformation as compared to RRM1,2 (5-10% higher average fraction of molecules in the open conformation) (Figure 5.14). Also for this construct, the fully closed and open conformations had a FRET efficiency of around 0.96 and 0.43, respectively (Table 5.1).

In its free form, URRM1,2 spent 46% of its time in the open state. The time in this conformation increased from 69% for the U4A8U4-bound construct, to ~91-93% in the presence of U4A8U4ACAGG, U4A8U5ACAGG, and U9 to 100% for U13ACAGG-bound URRM1,2. The strong Py-tracts U9 and U13ACAGG clearly stabilized the open conformation. The extended weak Py-tracts, however, also shifted the equilibrium of

URRM1,2 to the open conformation in a similar fashion as U9. This stabilization is visualized in the histograms of FRET efficiency versus donor lifetime, in which the URRM1,2 populations bound to the extended RNA sequences clearly fall onto the static FRET line (Figure 5.14).

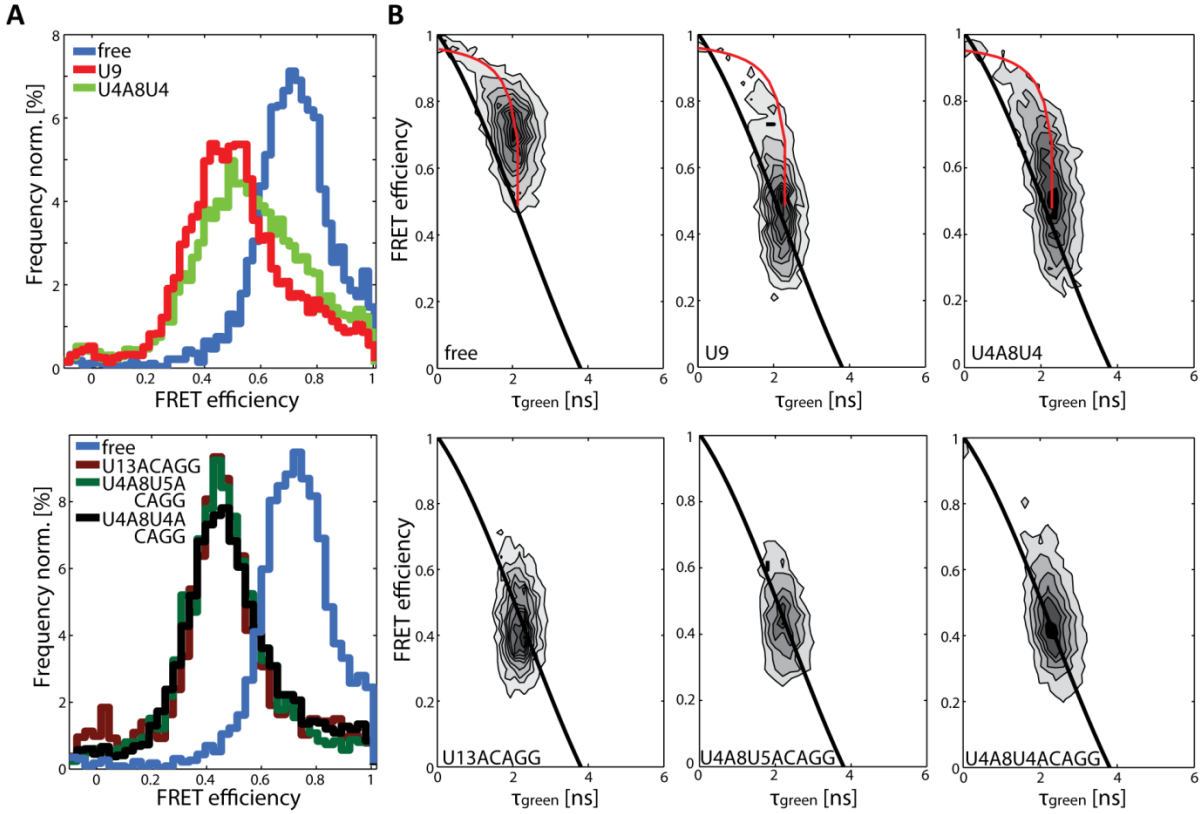


Figure 5.14 Characterization of URRM1,2 in the absence and presence of various Py-tract RNAs by spFRET with PIE-MFD. (**A**) FRET efficiency distributions of URRM1,2 in its free form (blue), bound to U9 (light red), or to U4A8U4 (light green) (upper graph) or in the presence of U13ACAGG (dark red), U4A8U5ACAGG (dark green), or U4A8U4ACAGG (black) (lower graph). (**B**) Histograms of donor lifetime versus FRET efficiency for the free form of URRM1,2 and in the presence of the different Py-tract RNAs U9, U4A8U4, U13ACAGG, U4A8U5ACAGG, and U4A8U4ACAGG. Static populations are found on the black static FRET curve, while dynamic molecules deviate from this line (red curve).

5.2.5 Effect of the additional U2AF35 subunit on U2AF65 dynamics

U2AF35 was added to the URRM1,2 construct containing ULM as the binding interface for the small subunit (Figure 5.6). The FRET efficiencies of the closed (0.94-0.96) and open (0.35-0.43) conformation obtained from the global lifetime information had similar values as for the aforementioned constructs (Table 5.1).

Overall, the presence of the small subunit even more significantly increased the shift of the equilibrium to the open conformation, in the absence and presence of Py-tract RNA (Figure 5.15). Already in its free form, the heterodimer spent 51% of the time in the open conformation (as compared to 34% for RRM1,2 and 46% for URRM1,2) (Table 5.1).

Upon the addition of a strong Py-tract (U9, U13ACAGG), the molecules were stabilized in the open form without ever switching to the closed state. The addition of U4A8U4 led to 85% of heterodimers found in the open conformation. The elongated Py-tract U4A8U5ACAGG, including the AG dinucleotide as a binding region for U2AF35, induced an almost fully open

heterodimer which was found to exist in the open state for 97% of the time. The conformational dynamics, observed in two-dimensional FRET efficiency versus donor lifetime histograms as deviations from the static FRET line, for free and U4A8U4-bound URRM1,2/U2AF35(UHM) diminished in the presence of strong Py-tracts or weak Py-tracts extended by the additional nucleotide sequence ACAGG (Figure 5.15).

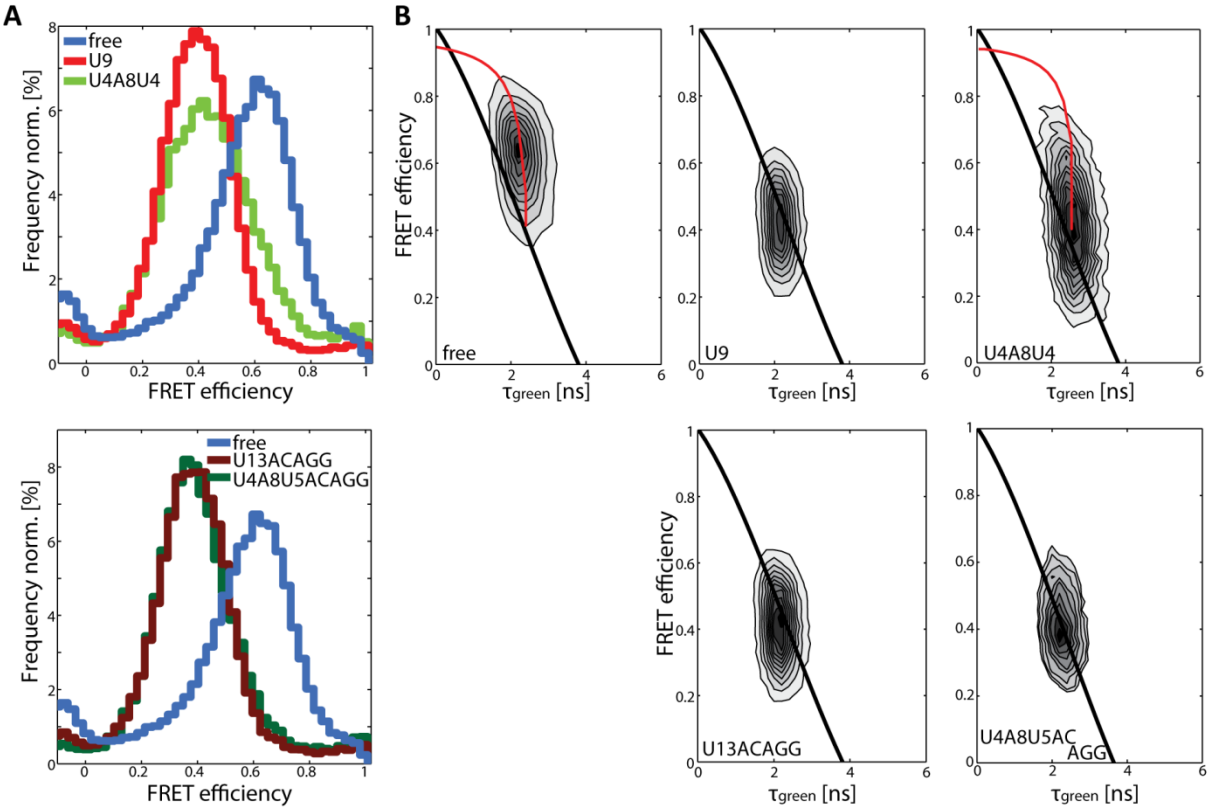


Figure 5.15 Characteristics of URRM1,2/U2AF35(UHM) in the absence and presence of Py-tract RNA investigated using spFRET with PIE-MFD. (**A**) FRET efficiency histograms of the free form of URRM1,2/U2AF35(UHM) (blue) and the protein bound to the strong Py-tract U9 (light red), or the weak Py-tract U4A8U4 (light green) (upper graph). In the lower graph FRET efficiency distributions of URRM1,2/U2AF35(UHM) bound to a strong (dark red) and weak (dark green) Py-tract sequence including the AG dinucleotide for U2AF35 recognition are displayed. (**B**) Two-dimensional histograms of donor lifetime versus FRET efficiency for the URRM1,2/U2AF35(UHM) heterodimer in the absence and presence of a variety of RNA substrates. The static FRET line is shown in black and a dynamic curve is displayed in red.

Comparing the equilibrium state between the closed and open conformation of the investigated constructs showed a dependency of the time spent in the open conformation on the size of the binding interface of the protein complex. This is most clearly visible for the free form of the protein. An elongation of RRM1,2 by the additional N-terminal region including ULM (URRM1,2) already slightly shifted the equilibrium towards the open conformation, while the addition of the small subunit U2AF35 enhanced this effect even further (Figure 5.14, Figure 5.15). A direct comparison of the FRET efficiency histograms of RRM1,2 and URRM1,2/U2AF35(UHM) as displayed in Figure 5.16 shows the shift of the equilibrium towards the open conformation as a shift of the mean FRET efficiency to lower values in the presence of additional U2AF35. A similar trend in the increase in binding affinity was observed for the different constructs in isothermal titration experiments (Table 5.2).

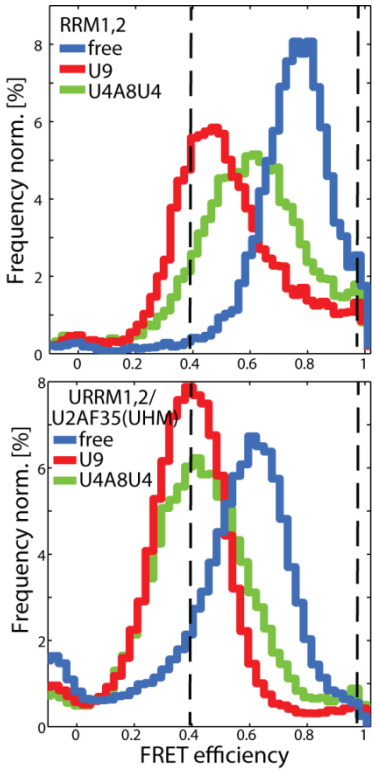


Figure 5.16 Comparison of the dynamic equilibrium between the open and closed conformation of RRM1,2 and URRM1,2/U2AF35(UHM). FRET efficiency histograms of RRM1,2 (upper panel) and URRM1,2/ U2AF35(UHM) (lower panel) in the absence (blue) and presence of the Py-tract RNAs U9 (red) and U4A8U4 (green). The dashed black lines show the mean FRET efficiency of the closed and open conformation determined from fitting the lifetime of the donor.

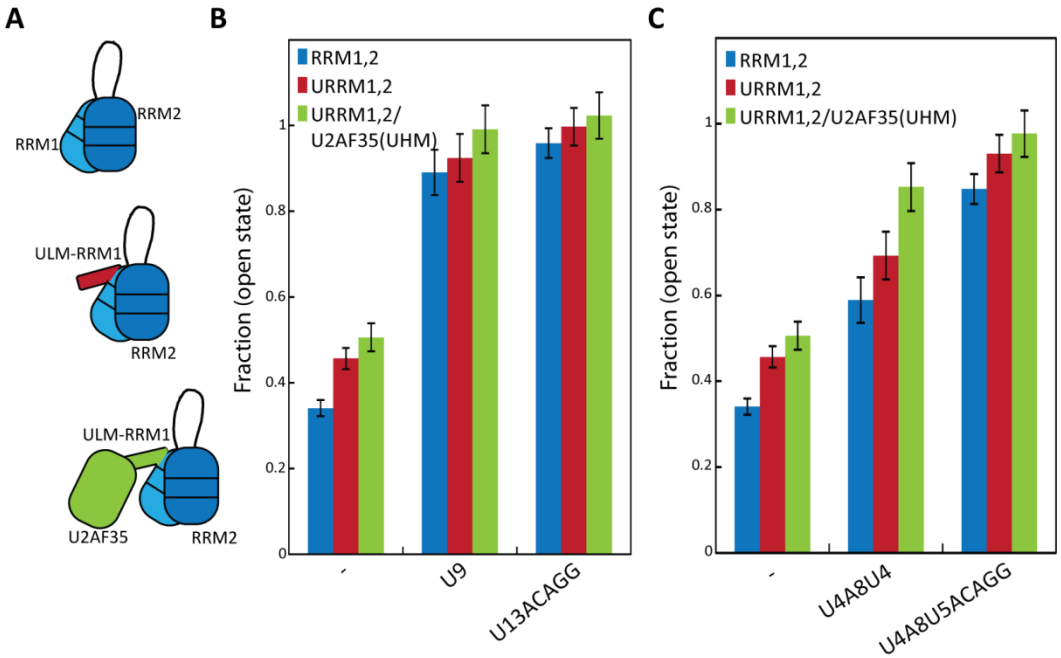


Figure 5.17 Equilibrium between the closed and open conformations displayed as the average fraction of molecules in the open state for different U2AF constructs. (**A**) Schematic representations of the three constructs shown: RRM1,2 (blue), URRM1,2 (red), URRM1,2/U2AF35(UHM) (green). (**B**) Fraction of the open state of the three constructs in the absence of RNA and in the presence of the strong Py-tracts U9 and U13ACAGG. (**C**) Fraction of time, which the U2AF constructs spent in the open state in the absence of RNA and in the presence of the weak Py-tracts U4A8U4 and U4A8U5ACAGG.

Although the equilibrium was generally strongly shifted to the open conformation in the presence of strong Py-tracts, we still observed small effects for increasing construct size (Figure 5.17B). In addition, an elongated strong Py-tract increased the time spent in the open state even more.

In the presence of the weak Py-tract U4A8U4, we found a dramatic increase of the open conformation for the heterodimeric constructs when compared to RRM1,2 or URRM1,2 (Figure 5.17C). This was shifted even further to the open state when the weak Py-tract contained an additional AG dinucleotide (U4A8U5ACAGG).

5.2.6 Timescales of the conformational dynamics

The described conformational transitions of U2AF constructs were observed by solution-based PIE-MFD experiments on a confocal microscope, meaning switching between different FRET states occurred on timescales of micro- to milliseconds. We used TIRF experiments to investigate whether additional dynamics were visible on longer timescales. For immobilization on the surface, the constructs were encapsulated in lipid vesicles (Figure 5.18A).

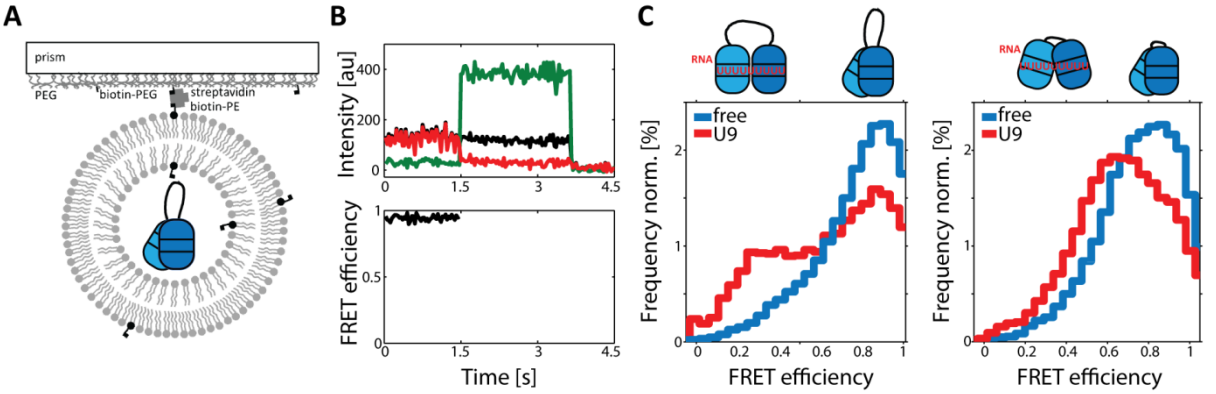


Figure 5.18 Surface-based TIRF experiments of RRM constructs. (**A**) Schematic representation of RRM1,2 encapsulated in lipid vesicles and attached to the prism surface (not to scale). (**B**) Exemplary intensity trace over time of the donor (green), acceptor (red), and total intensity (black) signal (upper panel) as well as the FRET efficiency (black) (lower panel). (**C**) Static framewise FRET efficiency distributions of free (blue) or U9-bound (red) RRM1,2 (left) or RRM1,2-Δ233-252 (right).

We did not observe any conformational switching above the temporal resolution limit of 30 ms. The static FRET efficiency histograms of the constructs RRM1,2-C187/C318 and RRM1,2-Δ233-252-C187/C318 resembled those from MFD experiments (Figure 5.18C, D). The amplitudes of the individual populations varied due to the selection criteria and timescales of the conformational dynamics.

To further investigate the timescales of the dynamics observed in spFRET experiments, we performed FRET-FCS and filtered FCS analysis. Unfortunately, these methods did not give any further distinct results on a single exact timescale of the conformational dynamics, but rather variations in the timescales observed for the different constructs and methods. Using these methods, the conformational dynamics showed variations from around 6 µs to a few milliseconds. An exemplary analysis of RRM1,2 in the absence of substrate is shown in Figure 5.19.

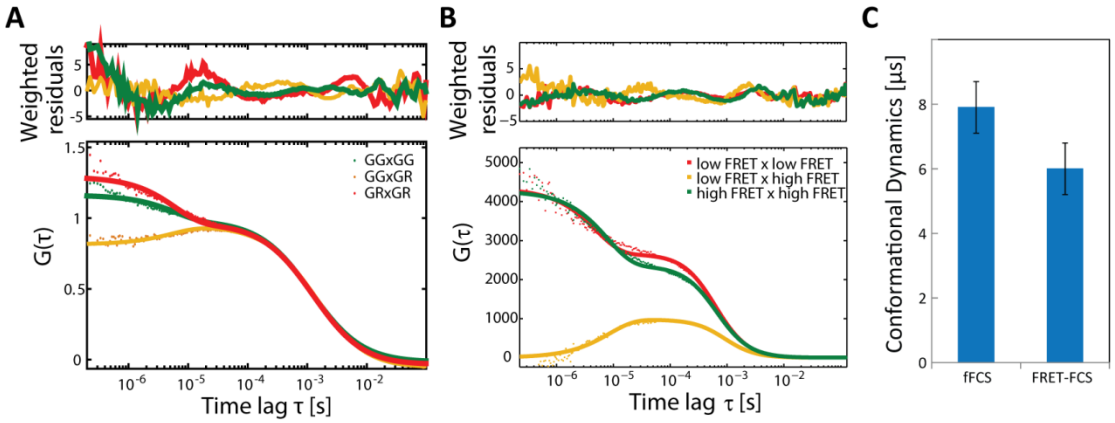


Figure 5.19 Exemplary analysis of RRM1,2 data with methods to evaluate the timescale of conformational dynamics observed in spFRET experiments with PIE-MFD. (**A**) FRET-FCS. Normalized correlation curves of the fluorophores after donor excitation. The autocorrelation curve of the detected donor signal is shown in green, the autocorrelation curve of the FRET signal is displayed in red and the cross-correlation function between both is shown in yellow. The anticorrelated signal of the cross-correlation at small timescales can be induced by conformational dynamics, while the molecule transverses through the confocal volume. (**B**) Filtered FCS analysis. The signals of the low FRET (red) and the high FRET (green) species were autocorrelated and cross-correlated (yellow). Anticorrelated signal can arise from conformational dynamics. (**C**) Timescales of conformational dynamics of RRM1,2 in the absence of RNA obtained from filtered FCS and FRET-FCS analysis.

Any fluctuations in the intensity can be extracted by these fluorescence fluctuation methods, the interpretation of the observed fluctuations has to be performed carefully keeping in mind additional effects which can lead to a change in the FRET efficiency, such as donor or acceptor photoblinking. From the experiments performed, however, it is clear that the opening-closing transition of RRM1,2 happens on the timescale of micro- to milliseconds. Possibly, the dynamic switching time of the molecules shows a distribution ranging from a few to several hundred microseconds depending on the construct.

5.3 Discussion

U2AF is highly important for the selective recognition of introns in pre-mRNA. Its large subunit, U2AF65, binds to the Py-tract region and U2AF35 contacts the conserved AG dinucleotide at the 3'splice site. The minimal RNA-binding construct of U2AF65, RRM1,2, is known to exist in a closed conformation in its free form, while it assumes an open state when bound to RNA (Mackereth et al., 2011). Using spFRET, we investigated these conformational states. We could show that RRM1,2 in its free form is highly dynamic, switching between an open and a closed conformation on the timescale of micro- to milliseconds. These conformations were the same for RRM1,2 in the presence of a variety of substrates. The equilibrium between these two states, however, changed depending on the substrate present (Figure 5.20). Py-tracts with a high affinity dramatically shifted the equilibrium to the open conformation. For weak Py-tracts, the complexes were still more dynamic, switching between the closed and open conformation, although spending already most of their time in the open state. This shift in the equilibrium presents the importance of RRM1,2 in the recognition of differences in the Py-tract sequences. A large variety of Py-tract sequences is known. The splicing efficiency from these sequences can vary greatly leading to the effect of alternative

splicing. This in turn is unequivocally important for the development of higher order organisms. Up to 90% of all human genes are spliced alternatively (Kim et al., 2006; Wang et al., 2008). The efficiency of splicing depends on the number of consecutive uridines (Coolidge et al., 1997; Reed, 1989). When the sequences are interrupted or the number of uridines is decreased, splicing from those sites occurs less frequently. We could link this effect of Py-tract affinity to the equilibrium of the conformational dynamics observed in RRM1,2. Therefore, the protein has an important role in the regulation of correct recognition of highly variable introns. Dependent on the strength of the Py-tract, RRM1,2 is stabilized to varying degrees in its open conformation. In agreement with this, a stabilizing effect of RNA binding on specific regions of the RRM structure has been observed previously (Thickman et al., 2007).

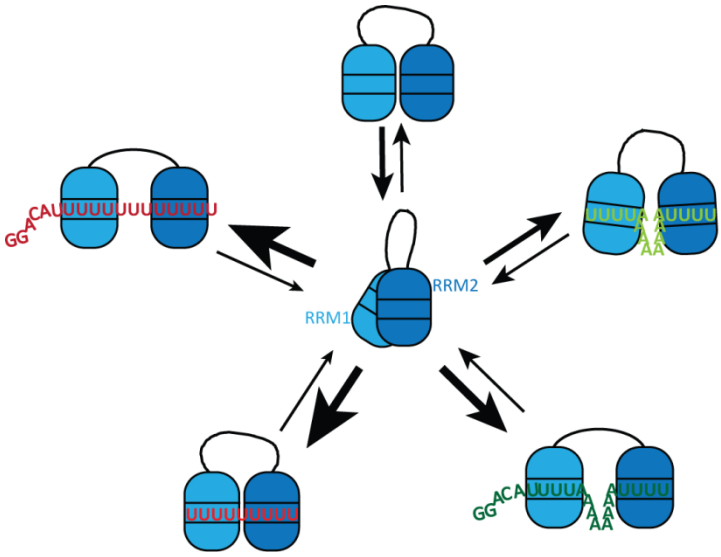


Figure 5.20 Equilibrium between the closed and open conformations of RRM1,2 in the absence and presence of a variety of Py-tracts. The free form of RRM1,2 undergoes conformational dynamics between a closed and an open conformation (blue). It spends less than half of its time in the open state. As soon as RNA binds to the protein, the time spent in the open conformation increases. This is most strongly pronounced for the strong Py-tracts U9 (light red) and U13ACAGG (dark red). For weak Py-tracts, the increase in affinity of gradually elongated RNA also shifts the equilibrium increasingly towards the open conformation. The arrows represent qualitative descriptions of the times spent in the open and closed conformation.

The shift in equilibrium towards the open state was not only introduced by the addition of RNA substrate, but was also achieved by increasing the affinity of U2AF65 by additional amino acids for U2AF35 binding. The elongation of RRM1,2 by the N-terminal region of U2AF65 including ULM induced an overall shift of the equilibrium towards the open conformation (Figure 5.17). Compared to RRM1,2, URRM1,2 showed a slight shift in the equilibrium under all conditions observed. Since the time spent in the open conformation was already increased for the proteins in the absence of RNA, the additional U2AF65 sequence stretch already provided a propensity for opening of the domains or posed a small hindrance to the closing transition. In agreement with this observation, the affinity of URRM1,2 for RNA was increased six-fold as compared to RRM1,2 (Table 5.2).

In particular, we observed a significant influence of the small subunit U2AF35 on the equilibrium between the closed and open state. The heterodimeric complex

URRM1,2/U2AF35(UHM) showed an even more distinct shift of the equilibrium towards the open state. Already in the free form, the heterodimer spent more than half of its time in the open conformation. As observed for URRM1,2, the affinity of the heterodimer was increased nine-fold in comparison to RRM1,2.

U2AF35 has been described as highly important for the splicing from weak Py-tract sequences (Guth et al., 1999; Pacheco et al., 2006). This agrees well with the here observed data, where the effect of the additional subunit is most pronounced for the stabilization of the open state in the presence of weak Py-tracts (Figure 5.17). Several studies showed the necessity of U2AF35 for splicing of AG-containing introns (Kanaar et al., 1993; Rudner et al., 1996; Zorio and Blumenthal, 1999), while intron sequences with a strong Py-tract can be spliced independently of U2AF35 (Guth et al., 1999; Pacheco et al., 2006; Reed, 1989). Our findings can provide the underlying molecular mechanisms explaining this observation. While U2AF65 on its own is well stabilized on strong Py-tracts, the additional small subunit U2AF35 conveys increased stability to its open state and thus binding to weak Py-tract sequences. A stabilization of the complex on the RNA might be necessary for the further recruitment of additional splicing factors.

While the RRM domain of U2AF35 was believed to drive the interaction of this subunit with the AG dinucleotide (Chusainow, 2005; Merendino et al., 1999; Tavanez et al., 2012; Wu et al., 1999; Zorio and Blumenthal, 1999), it was recently shown that this binding occurs via the zinc-finger motifs of the protein (Yoshida et al., 2015). Interestingly, the new results suggest only a protein stabilizing role of the RRM domain for U2AF35. The here observed increase of the time spent in the open conformation therefore cannot have been attributed to the additional sequence-specific interaction of the small subunit with the AG dinucleotide. This is in good agreement with the stabilization of the open state already observed in the absence of RNA or in the presence of RNA sequences, which are missing this dinucleotide binding region.

We detected a direct correlation of the equilibrium between both states with the affinity of the protein construct. The dynamic conformational changes and the shift in the equilibrium disclose the molecular explanation for the binding of U2AF to the RNA. The flexibility of RRM1,2 to occur in an open conformation even in the absence of RNA increases the probability of RNA binding as soon as it is added to the protein. A gradual increase in the time spent in the open conformation for U2AF constructs of increasing size therefore leads to an increase in affinity. Thus, we could explain the importance of U2AF35 for efficient splicing from weak Py-tracts and at the same time, by investigating U2AF65 dynamics provide a glimpse on the molecular mechanism behind alternative splicing.

5.4 Excursion: The effect of fluorophores on the conformational dynamics of U2AF65

SpFRET experiments require proteins fluorescently labeled with a donor and acceptor dye pair. The fluorophores are chosen according to their photophysical and biochemical properties. For the observation of single molecules, it is necessary to detect a sufficiently large number of photons from each fluorophore. Thus, the fluorophores have to provide these in the limited amount of time spent in the confocal volume before photodestruction. For *in vitro* experiments, organic fluorophores with a high photostability and large quantum yields are usually chosen. For FRET, the pair of fluorophores has to additionally show a certain spectral overlap to obtain reasonable Förster radii. Typically used donor dyes in the green wavelength range are Cy3, or Atto532, while Cy5, and Alexa647 are often used as acceptor fluorophores in the red spectral range. Di Fiori and Meller (2010) investigated the effects of different FRET pairs on nucleic acids and Chung et al. (2010) described the consequences of fluorophore photophysics on FRET efficiency distributions of immobilized molecules. We here investigated the effect of several combinations of fluorophores on the conformational dynamics of RRM1,2 constructs in solution. In addition, we combined spFRET and NMR experiments to show the effects of attaching a fluorescent moiety on protein structure.

5.4.1 Photophysical and biochemical properties of the fluorophores

In a first step, the fluorophores were characterized and compared regarding their photophysical properties. Both donor fluorophores, Atto532 and Atto565, are rhodamine-based dyes while the utilized acceptors are a carbopyronine (Atto647N) or carbocyanine dyes (Alexa647, Cy5). The typical Π-electron ring systems of the respective fluorophore families are visualized in Figure 5.21, where the chemical structures of the fluorophores under investigation are displayed. Atto532 shows a typical rhodamine-ring system, a derivate of the xanthene three-ring system. The system is slightly extended for Atto565. The carbopyronine Atto647N is structurally very similar to the rhodamine-based Atto532 and Atto565 fluorophores. Cyanine dyes consist of two heteroaromatic moieties, an electron-donor and electron-acceptor, connected by a polyene.

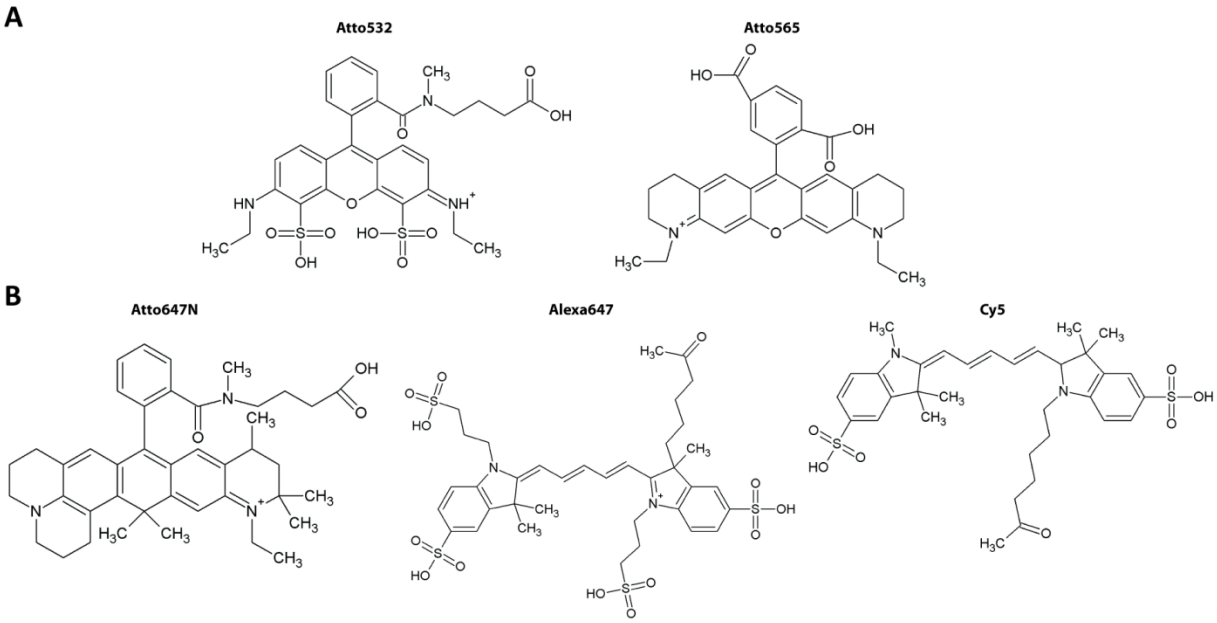


Figure 5.21 Chemical structures of the donor and acceptor fluorophores under investigation. (**A**) Donor fluorophores Atto532 (left) and Atto565 (right). (**B**) Acceptor fluorophores Atto647N (left), Alexa647 (middle) and Cy5 (right).

To ensure a high photon yield especially important for single molecule experiments, an investigation of the parameters extinction coefficient, quantum yield, and photostability is of special importance (Table 5.3). Rhodamine dyes typically show high photostability and quantum yields as well as little triplet formation (Kolmakov et al., 2010). Here, the quantum yields of both donor fluorophores were high with around 90%, while quantum yields of many red dyes are generally lower. Different red fluorescent dyes were tested and compared in Buschmann et al. (2003). The here tested acceptor dyes had quantum yiels of 30/33% for Cy5 and Alexa647 or 65% for Atto647N (Table 5.3). Atto647N additionally comprises a high photostability, while carbocyanine dyes often undergo photoblinking. From its photophysical properties, Atto647N would be the most favorable acceptor fluorophore in this spectral range. Because of its high photostability, it was used in a variety of single molecule experiments including live-cell STED microscopy (Hovius et al., 2006; Westphal et al., 2008). In addition, both red cyanine dyes have significantly shorter fluorescence lifetimes (Table 5.3).

Table 5.3 Physical and biochemical properties of donor and acceptor fluorophores.

Fluorophore	Fluorescence lifetime [ns]	Fluorescence quantum yield [%]	Extinction coefficient [1/(M·cm)]	Charge	Properties
Atto532	3.8	90	115000	-1	
Atto565	3.9	90	120000	+1	hydrophobic
Atto647N	3.5	65	150000	+1	hydrophobic
Alexa647	1.0	33	239000	-3	
Cy5	0.9	30	250000	+1	

Biochemically, however, Atto647N poses difficulties due to its hydrophobic nature, which might affect protein structure or function. Both cyanine-based acceptor fluorophores, as well as Atto532, were well dissolvable in aqueous solutions. Moreover, additional biochemical properties, such as molecular size, charge, and pH sensitivity have to be considered. All

organic fluorophores used in this study had a molecular size between 0.7 and 1.3 kDa. The overall charge of the dyes varied from a single positive charge for Atto565, Atto647N, and Cy5, via a single negative charge for Atto532 to a charge of -3 for Alexa647. Due to these charges, the fluorophores can have the propensity to interact unspecifically with the molecule of interest, induce structural changes or misfolding, and affect proper molecular functioning.

From the above characterized fluorophores, we chose several combinations of FRET pairs for our experiments. The FRET pairs under investigation had Förster radii ranging from 57 Å for Atto532/Atto647N to 68 Å for Atto565/Alexa647 (Table 5.4).

Table 5.4 Förster radii of the different combinations of FRET dye pairs.

Förster radius [Å]	Atto647N	Alexa647	Cy5
Atto532	57.2	61.8	63.2
Atto565	62.8	68.3	69.2

The local environment of fluorophores can affect their photophysical properties. The binding to biomolecules can, for example, quench the fluorophores. SpFRET experiments with PIE-MFD can reveal changes of the photophysical properties of fluorophores by investigating their lifetime. Analysis of the FRET efficiency histogram of RRM1,2-C187/C318 labeled with Atto532 and Cy5 showed an additional population besides the dynamic high FRET population observed for all combinations of FRET pairs (Figure 5.22A). This population appeared at low FRET efficiencies. A more detailed investigation on a two-dimensional histogram of FRET efficiency versus stoichiometry shows, that the additionally observed FRET state was induced by trailing of the populations between the donor-only species and the double-labeled or the acceptor-only population (Figure 5.22B). Usually, trailing between the double-labeled and the donor-only species can be associated with photobleaching of the acceptor or events during which several proteins travel through the focal volume simultaneously. Filtering the data using the information of the mean macroscopic arrival times, T_{DD+DA}-T_{AA}, removes these events from further analysis. Here, however, the trailing was not sufficiently removed, suggesting photoconversions of the fluorophores. Cy5 is known to undergo an isomerization between a cis- and a trans-configuration (Widengren and Schwille, 2000). This switching leads to the existence of a bright and a dark state of the fluorophore. Whenever the acceptor undergoes this transition while it is transversing through the focal volume, molecules will appear in the region between the donor-only species and the double-labeled population. The isomerization of Cy5 is thus directly visible in the two-dimensional histograms obtained in MFD experiments.

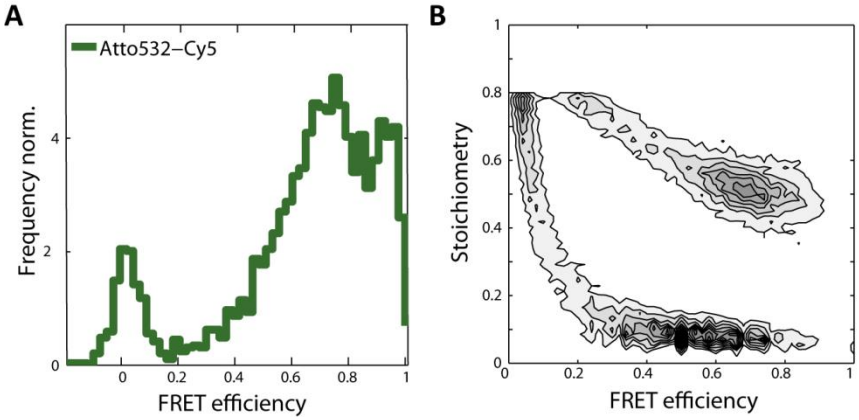


Figure 5.22 Isomerization of Cy5 observed in spFRET experiments with PIE-MFD. (**A**) FRET efficiency distribution of RRM1,2 labeled with Atto532/Cy5. Panel (**B**) shows the histogram of FRET efficiency versus stoichiometry. Trailing between the donor-only and the double-labeled population or the acceptor-only population was caused by the isomerization of Cy5

.

Furthermore, in spFRET experiments with PIE-MFD, one can gain information about additional parameters. We used the lifetime information of the acceptor fluorophore to observe its photophysical characteristics. Unlike the lifetime of the donor, the acceptor lifetime should be unaffected by the FRET efficiency. Whenever the value of the acceptor lifetime decreases nevertheless, it means that an additional radiationless decay of the fluorophore to the ground state has to be considered. The acceptor fluorophore is quenched by a surrounding moiety. We observed quenching of Alexa647 in experiments with RRM1,2 in the presence of U9. The molecules found in populations with different proximity ratios also showed differences in their acceptor lifetimes (Figure 5.23). The lifetime of the acceptor decreased from 2 ns to 1.6 ns for decreasing FRET efficiency. The acceptor was thus quenched due to the binding of the RNA or structural changes by opening of the domains. The differences in acceptor lifetime can be accounted for by applying a different correction factor γ to each of the populations. The populations corrected by this means then show a static value for the stoichiometry (Figure 5.23A, C). The use of PIE-MFD thus allows for the observation and correction of changes in the photophysical properties of the fluorophores.

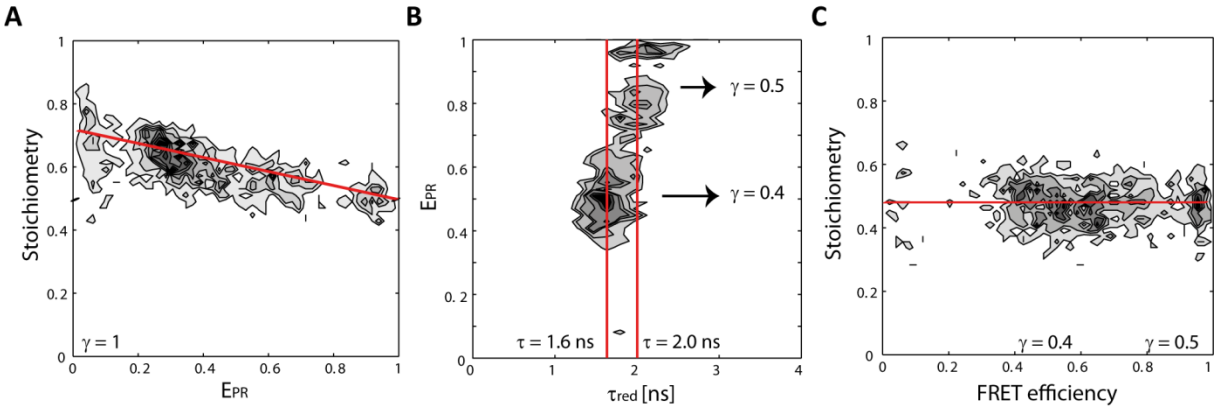


Figure 5.23 Quenching of the Alexa647 fluorophore bound to RRM1,2-C187/C318 in different conformational states. (**A**) Histogram of proximity ratio versus stoichiometry for RRM1,2 molecules. (**B**) A histogram of acceptor lifetime versus proximity ratio shows different acceptor lifetimes (τ) for the different RRM1,2 populations observed. These populations can be corrected individually by their respective γ-factor to obtain a corrected histogram of FRET efficiency versus stoichiometry (**C**). The red lines are provided as a guide to the eye.

5.4.2 Effects of the fluorophores on RRM1,2 structure

Besides the fluorescent properties of the individual fluorophores, we compared their effect on the structure of RRM1,2. To this end, we used a combined approach of spFRET and NMR experiments. The NMR experiments were performed in a collaborative work with Carolina Sanchéz Rico from the group of Professor Sattler (Technische Universität München, Germany) and will therefore be mentioned here in less detail.

Structures of RRM1,2 determined from NMR experiments with the addition of fluorescent labels to the positions C187 and C318 were prepared by Lisa Warner (Professor Michael Sattler, Technische Universität München) and are shown in Figure 5.24A. The proteins were stochastically labeled, meaning that each of the fluorophores could be positioned on one of the two cysteines. We analyzed RRM1,2 with different FRET pairs using PIE-MFD. The FRET efficiency histograms obtained for the different donor fluorophores Atto532 and Atto565 in combination with the acceptor Alexa647 showed similar distributions and complete dynamic opening of the domains in the presence of U9 (Figure 5.24B).

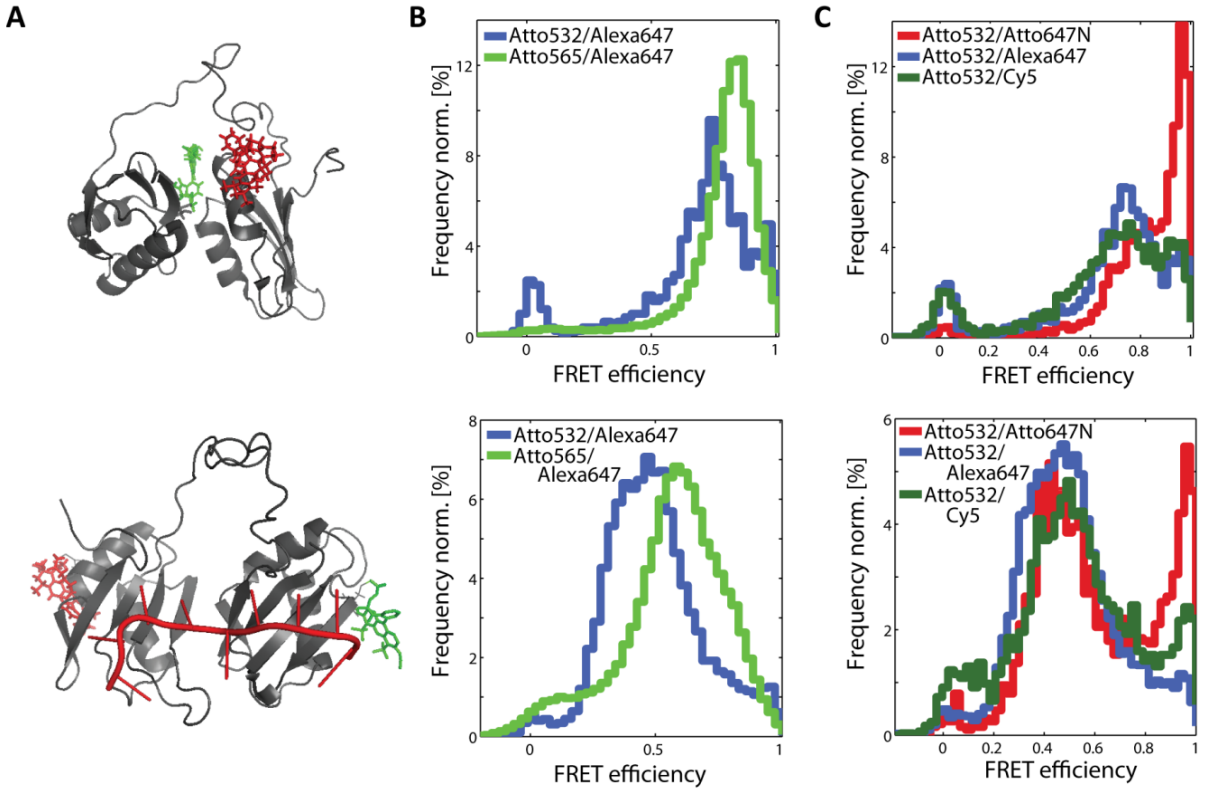


Figure 5.24 Characterization of the RRM1,2-C187/C318 structure with different FRET fluorophore pairs using spFRET experiments with PIE-MFD. (**A**) The structure of RRM1,2 in the free form (upper structure) or bound to RNA (lower structure) with fluorophores Atto532 and Atto647N displayed in green and red, respectively. (**B**) FRET efficiency distributions of RRM1,2 fluorescently labeled with the acceptor Alexa647 and the donor fluorophores Atto532 (blue) and Atto565 (light green) in the free form (upper graph) and in the U9-bound conformation (lower graph). (**C**) FRET efficiency histograms of RRM1,2 labeled with the donor fluorophore Atto532 and the acceptors Atto647N (red), Alexa647 (blue), or Cy5 (dark green) in the absence of substrate (upper graph) and in the presence of U9 (lower graph).

The overall shifts in the FRET efficiency distributions are due to differences in the Förster radii. When different acceptor fluorophores were compared, slight differences of the FRET efficiency distributions were visible while the overall trend was similar. In the presence of

Atto647N, a very strong distinct peak occurred at high FRET efficiencies of around 0.95 (Figure 5.24C). Even in the presence of U9, this population was still present. Attachment of Atto647N thus introduced a fraction of molecules with a non-physiological high FRET efficiency. For Cy5-labeled molecules, a very small population of these RRM1,2 molecules with a high FRET efficiency not observed for other fluorophores was visible.

By comparing NMR data of non-fluorescently labeled RRM to RRM labeled at each of the cysteines, we found Atto647N to introduce structural changes to the protein. Especially the RRM2 domain was affected. Using the chemical shifts of the amino acids of RRM1,2 labeled with the different dyes at both cysteine positions, we could observe that Atto532 and Alexa647 were the donor and acceptor fluorophores introducing the least pertubations to RRM1,2 structure. This agrees well with the low hydrophobicity of these fluorophores. Moreover, both dyes carry a negative charge which seems to introduce less structural changes to RRM1,2 than the positively-charged fluorophores.

In a following step, the position of one fluorophore was moved from a site in an α-helical region to an unstructured region next to it (C318 to C326). This positional rearrangement, however, still did not fully eliminate the inactive high FRET population in the presence of Atto647N although clearly diminishing it (Figure 5.25).

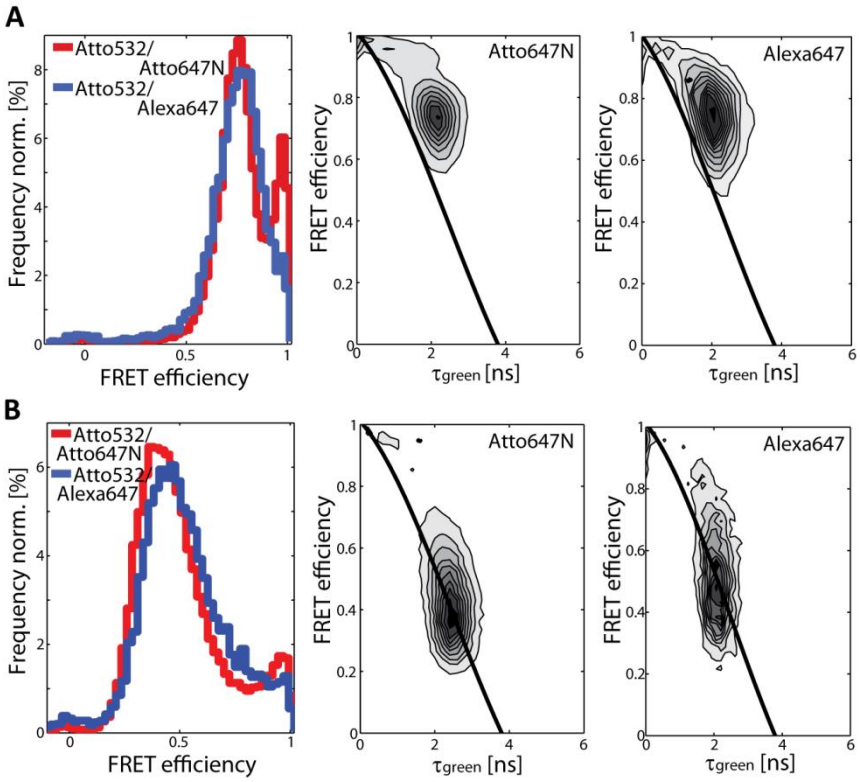


Figure 5.25 Comparison of different acceptor fluorophores attached to RRM1,2-C187/C326 using spFRET with PIE-MFD. FRET efficiency histograms and two-dimensional histograms including the donor lifetime information of RRM1,2 labeled with Atto532 and the acceptor dyes Atto647N (red) and Alexa647 (blue) in the absence (**A**) and in the presence of U9-RNA (**B**).

5.4.3 Conformational dynamics of RRM1,2 affected by fluorophores

The effects of the different acceptor fluorophores on the conformational dynamics were investigated using two-dimensional histograms of donor lifetime versus FRET efficiency. While RRM1,2 labeled with Alexa647 or Cy5 was clearly dynamic in the absence of RNA, the protein labeled with Atto647N appeared less dynamic due to the large fraction of inactive molecules with high FRET efficiency (Figure 5.26A). In the presence of U9, both, RRM1,2-C187/C318-Atto532/Alexa647 and RRM1,2-C187/C318-Atto532/Cy5, in their open conformation clearly fall onto the static FRET curve showing little conformational switching on the timescale of micro- to milliseconds. Again, differences in the behavior of U9-bound RRM1,2 labeled with Atto647N were observed (Figure 5.26B). Molecules were switching between a closed and an open conformation in the presence of U9 only when RRM1,2 was labeled with Atto647N. The fluorophore thus affected the dynamics of the conformational switching possibly shifting it to slower timescales. Atto647N may attach unspecifically to additional positions within the protein inhibiting the free movement of the domains.

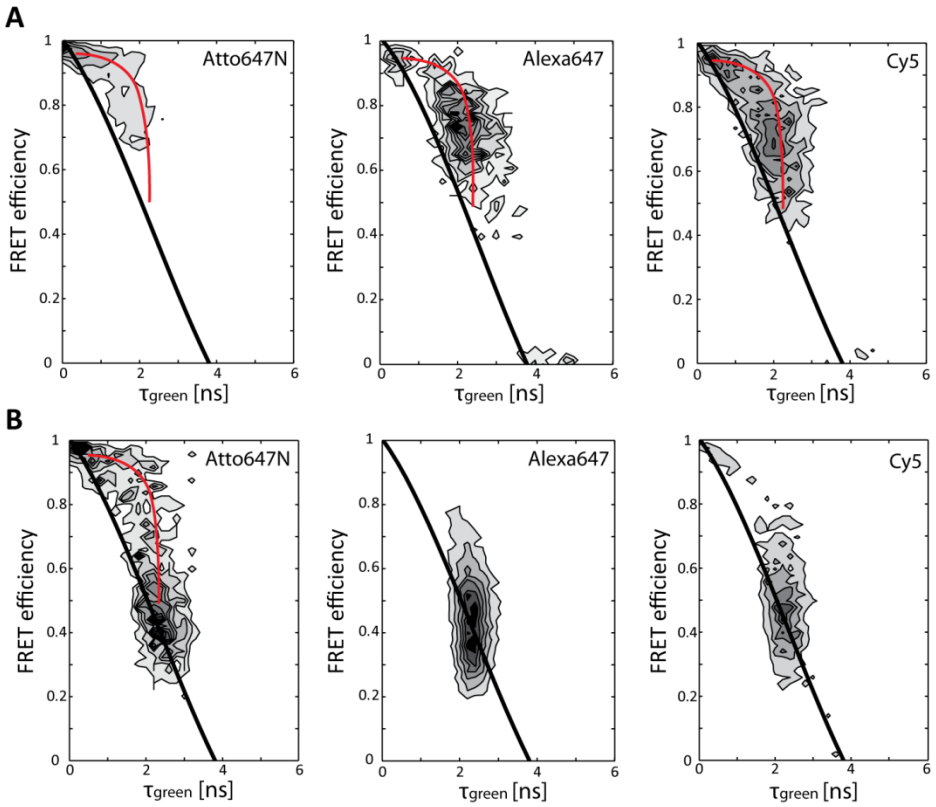


Figure 5.26 Effect of the fluorophores attached to RRM1,2-C187/C318 on the conformational dynamics. Two-dimensional histograms of donor lifetime versus FRET efficiency. RRM1,2 labeled with Atto532 and the acceptor fluorophores Atto647N (left), Alexa647 (middle), and Cy5 (right) in the absence of substrate (**A**) or bound to U9-RNA (**B**). Static populations fall onto the static FRET line (black) and dynamics were described by dynamic FRET curves (red).

Furthermore, conformational dynamics on the timescale of seconds were observed for RRM1,2-C187/C318 labeled with Atto532/Atto647N encapsulated in lipid vesicles, which did not occur when RRM1,2 was labeled with a different acceptor fluorophore (1-5% dynamic molecules) (Table 5.5). Both, in the absence and in the presence of RNA, 21-29% of Atto532/Atto647N-labeled RRM1,2 molecules showed conformational dynamics on a TIRF

microscope. An exemplary intensity trace of a dynamic molecule is displayed in Figure 5.27A. The framewise FRET efficiency histogram shows a low and a high FRET state of RRM1,2 in the absence and presence of RNA (Figure 5.27B). The switching dynamics between these states was extracted by an HMM analysis, which yielded transition density plots of the protein showing populations in at least two different states (Figure 5.27C). The transition rate from one state to another was determined by fitting the survival probability of the molecules in one state to a single exponential decay function (Figure 5.27D, E). Transition rates of RRM1,2 in the absence and presence of RNA were between 1.5-2 s for both the transition from the low FRET to the high FRET state and vice versa. The linker truncation mutant RRM1,2-Δ233-252-C187/C318 labeled with Atto532/Atto647N showed similar results. Transitions from the low FRET to the high FRET state were well fit by a single exponential, while the opposite transition was less well described for all four combinations of protein constructs and RNA, which suggests that this transitions might involve multiple states not distinguishable by the FRET efficiency.

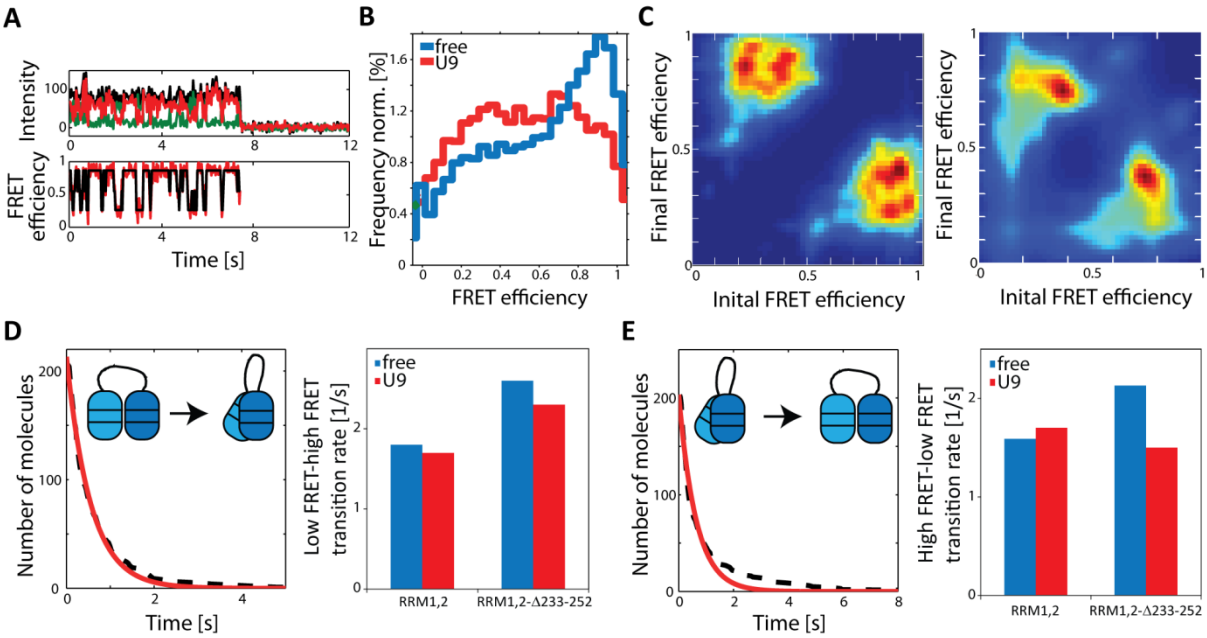


Figure 5.27 Conformational dynamics of RRM1,2-C187/C318-Atto532/Atto647N obtained by TIRF microscopy. (**A**) Exemplary intensity time trace of a dynamic RRM1,2 molecule. The donor (green) and acceptor (red) signal (upper graph) are anticorrelated while the total intensity (black) stays constant. The FRET efficiency (lower graph) shifts between two states. The FRET efficiency trace was fit by an HMM (black). (**B**) FRET efficiency histogram of dynamic RRM1,2 molecules in the absence (blue) and presence (red) of U9-RNA. (**C**) Transition density plots obtained by HMM analysis of the free RRM1,2 (left) and of the U9-bound RRM1,2 (right). (**D**) RRM1,2 and RRM1,2-Δ233-252 molecules shifting from the low FRET state to the high FRET state. Left: Exemplary survival probability of the free form of RRM1,2 (dashed black line) with an exponential fit (red curve). Right: Transition rates for the transition from the low FRET state to the high FRET state of free (blue) or U9-bound (red) RRM1,2 and RRM1,2-Δ233-252. (**E**) RRM1,2 constructs transiting from the high FRET to the low FRET state. Left: Survival probability of free RRM1,2 molecules (black) with exponential fit to the data (red). Right: Transition rates for the transition from the high FRET state to the low FRET state of free (blue) and U9-bound (red) RRM1,2 and RRM1,2-Δ233-252.

Table 5.5 Fraction of dynamic molecules observed using TIRF microscopy.

Protein	RNA	Acceptor fluorophore	Number of static molecules	Number of dynamic molecules	Dynamic molecules [%]
RRM1,2-C187/C318	-	Atto647N	313	129	29
	U9		135	35	21
	-	Alexa647	234	3	1
	U9		90	5	5

The observation of conformational dynamics on the timescale of seconds whenever Atto647N was bound to RRM1,2 in combination with the shifted equilibrium observed in MFD experiments (Figure 5.24) led us to conclude that Atto647N influenced the timescale of the conformational transitions of RRM1,2. The fluorophore appeared to interact with RRM1,2 in such a way that the switching rate between the closed and open conformation was decreased. In addition, the hydrophobic Atto647N might lead to an interaction of RRM1,2 with the lipids of the surrounding vesicle.

5.4.4 Discussion

This short excursion on the influence of fluorescent moieties on RRM1,2 constructs shows the importance of the careful selection of fluorescent dyes used in spFRET studies. In single molecule experiments, the detection of a large number of photons is of special importance. Thus, fluorophores with a small amount of photoblinking and high photostability and quantum yield are clearly favored. Different photophysical characteristics of the fluorophores become directly apparent in the multiple parameters observed by spFRET with PIE-MFD. In some cases, however, the improved photophysical properties come with biochemical characteristics of the fluorescent dyes which are poorly suited for biological molecules. Zanetti-Domingues et al. (2013) described the effect of hydrophobic fluorophores on single molecule tracking experiments. Especially in the red spectral range, organic fluorophores with high quantum yields and photostability are particularly scarce. Compared to other fluorophores in this range, Atto647N has strongly improved photophysical characteristics for single molecule experiments. In this chapter, however, we could show biochemical characteristics of the dye affecting RRM1,2 function, in particular the influence of Atto647N on protein structure and dynamics. We observed that each of the acceptor fluorophores analyzed came with its own characteristics when analyzing FRET efficiency histograms. Acceptor quenching of Alexa647 and photoisomerization of Cy5 could be directly observed and accounted for. Buschmann et al. (2003) observed quenching of different red fluorophores by the covalent attachment of other small moieties such as biotin or streptavidin. In agreement with this, the protein and the RNA offer an even larger interacting surface possibly inducing the quenching of Alexa647. Atto647N, on the other hand, introduced a high FRET population, which was largely inactive even in the presence of substrate. Effects of fluorophores on spFRET efficiency distribution histograms have been described before. While Di Fiori and Meller (2010) investigated the influence of fluorophores on nucleic acids, Chung et al. (2010) described the effects of fluorophore photophysics of immobilized proteins. Using spFRET

with PIE-MFD of proteins in solution, we were able to observe the effects of fluorophores on the dynamics of the conformational changes. For Atto647N-labeled proteins, we observed changes in the timescale of the conformational dynamics towards slower motion, which could even be observed in TIRF experiments. In histograms of donor lifetime and FRET efficiency of PIE-MFD experiments, Atto647N-labeled constructs differed from the constructs labeled with other acceptors.
The fluorophore itself was responsible for shifting the equilibrium to a non-physiological high FRET state, while the dynamics observed on the TIRF microscope might have been additionally influenced by the interaction of the hydrophobic dye with the lipid environment of the vesicles. Since the transition rates neither differed largely between constructs nor in the presence or absence of substrate, transitions describing the interaction with the surrounding lipids are likely. This idea is supported by the notion of an increased lipophilicity for negatively charged molecules with aromatic rings (Ogawa et al., 2009). Hughes et al. (2014) investigated the ratio of fluorophores retained in liposomes for a large number of dyes. While Atto647N showed a strong retension, Atto532 and Alexa647 went into the aqueous solution in neutral and negatively charged liposomes. The interaction of the Atto647N fluorophore with the lipids may therefore cause the dynamic behavior on the TIRF microscope. The FRET efficiencies of the two FRET states observed by TIRFM, on the other hand, correspond well to the closed and open FRET conformations observed in MFD experiments, indicating a shifting of the dynamic timescales due to Atto647N even in the absence of lipids.

Depending on the nature of the molecule of interest, it is highly important to not only consider photophysical but also biochemical characteristics of the fluorescent labels used. Besides positioning the fluorophores to obtain the best possible separation for FRET, the positions should be chosen such that the activity and structure of the protein are unaffected. An overview of some considerations necessary for the appropiate choice of the fluorophore as well as the whole single molecule experiment is given in Gust et al. (2014) and Roy et al. (2008). To facilitate the choice of the most optimal fluorophore, recently a variety of new organic dyes, especially in the red wavelength range, are being developed (Brabetz and Weber; Cooper et al., 2004; Kolmakov et al., 2010; Wurm et al., 2012). Additionally, known fluorophores are stabilized by the addition of small moieties whose stabilizing mechanism is under investigation (Altman et al., 2011, 2012; Tinnefeld and Cordes, 2012). With the increasing number of fluorescent dyes available, the gap between optimal biochemical properties and good photophysical characteristics might be closed, leading to even more exact single molecule studies.

6 Comparative analysis of the allosteric effects and coordinated motion in Hsp70 chaperones from different organelles

Already during protein translation, chaperones bind to the newly synthesized protein and act co-translationally to reduce misfolding and aggregation. Heat shock proteins 70 (Hsp70) are one of the first chaperones to bind to these nascent polypeptide chains, recognizing stretches of seven hydrophobic amino acids. To perform their function, these chaperones undergo a conformational cycle based on their nucleotide-bound state. In the ADP-bound form, the affinity for substrate peptides is high and upon binding of ATP it decreases and therefore the substrate can leave its binding pocket. The substrate can undergo several binding and release cycles of Hsp70 as long as hydrophobic amino acids are still exposed. Hsp70 chaperones thus prevent aggregation and misfolding. The conformational cycle of different Hsp70s has been observed in detail previously using spFRET. The investigation of Hsp70 chaperones was extended here by the use of three-color FRET. By the simultaneous detection of three distances, we could directly monitor the coordinated behavior of these chaperones. The interconversion between the conformational states of the endoplasmic reticulum-associated BiP, the mitochondrial Ssc1, and the prokaryotic DnaK were clearly coordinated in their interdomain and lid conformation when bound to ATP. ADP-bound Ssc1 showed higher conformational flexibility of its lid for a distinct large interdomain distance. Differences of these Hsp70 chaperones in their peptide and nucleotide exchange factor binding and in their selectivity for nucleotides were revealed. And the effect of Hsp70 inhibitors on these chaperones was shown by investigating the induced conformational changes. Knowledge about the specificity of these inhibitors is of great importance to understand the molecular mechanism of the Hsp70 chaperones and considering their effectiveness as antitumoral and antibacterial drugs.

Our collaboration partners Dr. Dejana Mokranjac (Ludwig-Maximilians-Universität München, Germany), Professor Johannes Buchner, and Dr. Mathias Rosam (Technische Universität München, Germany) provided plasmids for the Hsp70 chaperones. Dr. Edward Lemke, Swati Tyagi, and Christine Köhler (European Molecular Biology Laboratory, Heidelberg, Germany) introduced us to the expression of the proteins using non-natural amino acids. Professor Matthias Mayer (Ruprecht-Karls-Universität Heidelberg, Germany) provided information on DnaK and ADP.

6.1 Introduction

Mature mRNA is transported into the cytosol and translated into protein by ribosomes at the endoplasmic reticulum (ER). The primary amino acid sequence contains information regarding the tertiary folded structure of the newly formed protein (Anfinsen's dogma). Single domain proteins therefore often refold spontaneously *in vitro*. *In vivo*, however, they usually only fold posttranslationally once the whole sequence is present, since their folding relies on the interaction of many weak and non-covalent bonds which can be many amino acids apart

and the exit channel of the ribosome is only 20 Å in width (Hartl et al., 2011). To prevent partial misfolding, molecular chaperones, such as trigger factor 1, bind to the protein cotranslationally. Multi-domain proteins, on the other hand, often start the folding of individual domains cotranslationally since protein synthesis is slow (~4-20 amino acids/s). Protein folding in the cytosol does not occur for all polypeptides. Rather, many of them have to be transported into organelles such as mitochondria or chloroplasts or into the ER for secretion. Molecular chaperones are therefore also involved in protein transport and translocation and are ubiquitously found in all organelles. Around 20-30% of all mammalian proteins are intrinsically disordered and only fold upon binding to their interaction partners (Hartl et al., 2011).
Molecular chaperones are necessary for *de novo* protein folding and to prevent misfolding and aggregation of proteins under cellular stress conditions. Due to their increased expression levels as a response to heat stress or the accumulation of misfolded proteins, they are often called heat shock proteins. Some chaperones are constitutively expressed while others appear only as a response to stress. Chaperones are found in all cellular compartments and organelles and make up for 1-2% of all proteins in the cell. Six major classes of chaperones are known, which are classified according to their molecular weight (Hsp40, Hsp60, Hsp70, Hsp90, Hsp100, and small Hsps). Chaperones are very diverse in structure and function. Although all of them participate in the process of protein folding and conformational change, some additionally partake in other functions such as the assembly and disassembly of protein complexes, switching between inactive and active conformations, quality control, intracellular transport, or degradation. The focus of this chapter will be on the Hsp70 chaperone family, which is one of the most abundant chaperone families and found in all domains of life.

6.1.1 Hsp70 chaperones

Hsp70 chaperones are involved in protein folding in the cytosol or different organelles, prevent misfolding and aggregation, stabilize proteins for membrane transport and participate in complex assembly and disassembly. Hsp70 proteins are the most highly conserved proteins known. Prokaryotes usually carry one or few Hsp70 genes, while eukaryotes contain several Hsp70-encoding genes. Eukaryotic Hsp70s are grouped into four families: cytosolic, endoplasmic reticulum-associated, mitochondrial, and plastid Hsp70 chaperones. We here investigated the Hsp70s DnaK from prokaryotes, mitochondrial Ssc1, and BiP from the endoplasmic reticulum. Interestingly, the different families are highly conserved throughout different phyla, which means that the major eukaryotic Hsp70 groups arose prior to the divergence of the earliest eukaryotes around 2 billion years ago. Additionally, duplication events occured more recently to give rise to more specialized Hsp70s.

Hsp70 chaperones consist of an N-terminal 44 kDa nucleotide-binding domain, connected to a 28 kDa substrate-binding domain via a flexible linker (Figure 6.1). The substrate-binding domain contains a β-sheet with a hydrophobic cleft for substrate recognition and an α-helical lid domain for stabilization. Hsp70 proteins act early during protein translation and bind to a

stretch of seven residues enriched in hydrophobic amino acids. These stretches occur on average every 50 to 100 amino acids in a protein (Hartl et al., 2011).

6.1.2 Conformational cycle of Hsp70s

When performing their function, Hsp70s undergo conformational changes from an open to a closed conformation (Figure 6.1). The conformational changes are induced by the binding of a nucleotide to the NBD or by substrate binding to the SBD. In the ATP-bound form, the interdomain distance between the NBD and the SBD is small and the α-helical lid occurs in an open state (Kityk et al., 2012). In this conformation, the association of substrate and the rate of substrate release are both very high, and thus proteins quickly associate and dissociate from the chaperone. Stable peptide binding to the Hsp70 occurs after hydrolysis of ATP. Both, the binding and the release rate decrease by two and three orders of magnitude, respectively. The overall affinity is therefore strongly increased by ten- to 50-fold for the ADP-bound Hsp70 (Mayer et al., 2000; McCarty et al., 1995; Schmid et al., 1994). In the ADP-bound conformation, the α-helical lid closes upon the β-sheet of the SBD and the interdomain distance is increased.

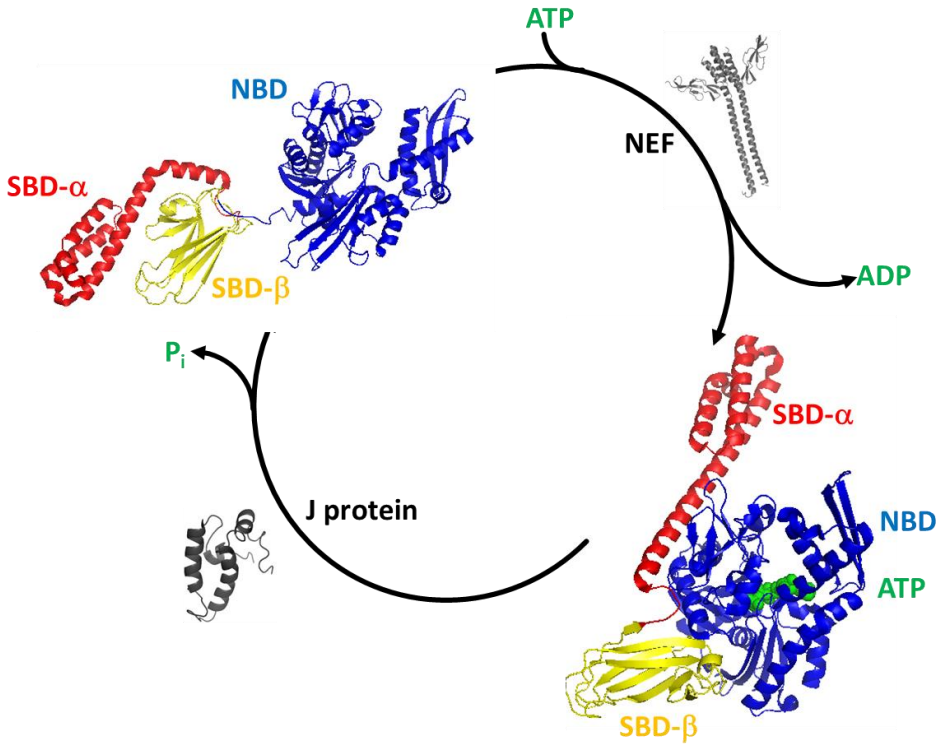


Figure 6.1 Conformational cycle of Hsp70 chaperones depicted exemplarily for DnaK. In the ADP-bound form (upper left), Hsp70s adopt a conformation where the nucleotide binding domain (NBD, blue) and the substrate binding domain (yellow, red) are far apart and the α-helical lid (red) closes the binding pocket of the SBD. ATP-binding to the NBD leads to docking of the NBD and SBD and opening of the lid (lower right). The intrinsic ATPase activity of Hsp70s is low. Nucleotide exchange factors (NEF) increase the release rate of ADP from the NBD and J-domain proteins (J protein) act to increase the ATP hydrolysis rate. PDB: 4B9Q, 3A6M, 2KHO, 1HDJ.

The intrinsic ATPase activity of Hsp70 chaperones is low with a turnover rate of around one ATP per 30 minutes. In the presence of peptide substrate, the intrinsic ATPase activity is increased. Cochaperones are required to further enhance the rate of ATP hydrolysis.

Nucleotide exchange factors aid in the release of ADP from the Hsp70 chaperone and subsequent rebinding of ATP to the nucleotide binding domain. Hsp40 chaperones (or J proteins) bind to the NBD via their 70 amino acids large J-domain and enhance the rate of ATP hydrolysis (Frydman, 2001; Genevaux et al., 2007; Mayer and Bukau, 2005). The binding of substrate to the SBD in synergism with Hsp40 binding induces a conformational change in the ATPase domain stimulating the rate of ATP hydrolysis up to 5000-fold (Karzai and McMacken, 1996; Laufen et al., 1999).

6.1.3 DnaK, a prokaryotic Hsp70

The most abundant prokaryotic Hsp70 is DnaK. It is highly conserved throughout the whole domain and was first described by Bardwell and Craig (1984). DnaK is involved in *de novo* protein folding, protein assembly, and stress response. Its conformation in the ADP- and ATP-bound states was recently solved by crystallization (Bertelsen et al., 2009; Kityk et al., 2012) (Figure 6.1). The protein was even crystallized in the presence of its nucleotide exchange factor GrpE (Wu et al., 2012b). GrpE binds to DnaK in the ADP-bound conformation (Zylicz et al., 1987). It was found to accelerate the ADP/ATP exchange of DnaK by 5000-fold even though its binding site does not include the nucleotide binding domain (Harrison, 2003; Mally and Witt, 2001; Packschies et al., 1997). The Hsp40 DnaJ alone, on the other hand, increases the intrinsic ATPase activity of $5 \cdot 10^{-4}$/s by sevenfold. Several studies proposed a direct binding of DnaJ to the substrate protein and targeting of the substrate to DnaK (Gamer et al., 1996; Schröder et al., 1993; Szabo et al., 1994). To convey correct folding, DnaK requires several posttranslational cycles of binding and release of its substrate (Agashe et al., 2004; Pierpaoli et al., 1997). Substrate release from DnaK is achieved by binding of ATP to the Hsp70 (McCarty et al., 1995; Palleros et al., 1993).

6.1.4 Eukaryotic Hsp70s

Eukaryotes carry at least five to six Hsp70 encoding genes. Their Hsp70 proteins are found in the cytosol and all organelles. Cytosolic Hsp70s are responsible for cotranslational binding of nascent proteins, protein folding, targeting of proteins to downstream chaperonins, or transport to different organelles. The Hsp70 cofactor Bag1 in the eukaryotic cytosol is the functional analog of GrpE.

Ssc1

Mitochondria and plastids contain Hsp70 proteins and nucleotide exchange factors, which are close homologs of DnaK and GrpE (Genevaux et al., 2007; Naylor et al., 1996; Schroda et al., 2001). All higher eukaryotes carry a single mitochondrial Hsp70, mtHsp70 (Schilke et al., 2006). It resides in the mitochondrial matrix. Due to its multiple functions, it is essential for cell survival (Craig et al., 1987; Yoneda et al., 2004). In combination with the Hsp40 Mdj1 and the nucleotide exchange factor Mge1, mtHsp70 is responsible for folding and prevents aggregation of proteins in mitochondria (Horst et al., 1997; Kang et al., 1990; Rowley et al., 1994; Westermann et al., 1996). When acting jointly with Jac1 and Mge1, mtHsp70 plays a role in the biogenesis of iron-sulfur clusters (Craig and Marszalek, 2002). Upon interaction

with Tim44 and Pam16/18, mtHsp70 works as a molecular motor to transport proteins across the inner mitochondrial membrane (Endo, 2003; Koehler, 2004; Neupert and Brunner, 2002; Neupert and Herrmann, 2007; Rehling et al., 2004). Tim44 is necessary for the recruitment of mtHsp70 to the translocon complex of the inner mitochondrial membrane with the channel protein Tim23. Pam16 and Pam18 act as J-domain containing proteins, thus stimulating the Hsp70s ATPase rate (Bolender et al., 2008; Craig et al., 2006; Mokranjac et al., 2006).
In yeast, three mitochondrial Hsp70s are known (Voos and Röttgers, 2002). The major mtHsp70, Ssc1, is responsible for folding and membrane transport. Ssq1 is involved in iron-sulfur cluster biogenesis and the function of Ecm10 is still under debate (Baumann et al., 2000).

BiP

Protein folding, assembly, and control of membrane-bound or secreted proteins take place in the endoplasmic reticulum. Around one third of all proteins are folded in the ER (Chen et al., 2005). Due to the high concentration of luminal proteins (~100 mg/mL) and Ca^{2+} concentrations of 5-10 mM, protein folding in the gel-like environment of the endoplasmic reticulum is especially challenging. Chaperones of the ER aid in the folding process and the folding is tightly regulated by the ER quality control system. The compartment contains a single conventional Hsp70, BiP (immunoglobulin heavy-chain binding protein) (Munro and Pelham, 1986). The human Hsp70 BiP interacts with around 20% of all human proteins. It is important for protein folding, assembly, translocation, degradation, as a stress sensor (Bertolotti et al., 2000; Shen et al., 2002), and for the maintenance of calcium stores (Hendershot, 2004). Similar to Ssc1, BiP fulfills its functions by interacting with multiple different J-proteins. Seven ERdj proteins are known. For its role in protein translocation through the ER membrane, BiP interacts with the membrane-anchored ERdj2. BiP directly interacts with the translocon and therefore aids in the contranslational transport across the membrane while, at the same time, inhibiting aggregation. Protein folding occurs in the ER lumen upon interaction with ERdj3 and ERdj6 (Guo and Snapp, 2013) and for endoplasmic reticulum associated degradation (ERAD) ERdj4 and ERdj5 are required (Behnke et al., 2015; Ellgaard et al., 1999; Hammond and Helenius, 1995). The roles of ERdj1 and ERdj7 are still under debate (Behnke et al., 2015). As for the other Hsp70s, the intrinsic ATP hydrolysis rate is low with around 0.37/min (Kassenbrock and Kelly, 1989; Mayer et al., 2003; Welsh et al., 2011). Interestingly, the conditions in the ER (neutral pH and high Ca^{2+} concentrations) inhibit BiP's ATPase activity. Ca^{2+} increases its affinity for ADP around 1000-fold making nucleotide exchange rather than ATP hydrolysis the rate-limiting step in BiP's chaperone cycle (Lamb et al., 2006). To accelerate the nucleotide exchange, two nucleotide exchange factors are known, Grp170 and BAP/Sil1. BAP selectively binds to BiP in the ADP-state and increases the ATPase turnover rate by a factor of 1.5-2 (Chung, 2002; Weitzmann et al., 2007). Grp170's task does not solely lie in its function as a nucleotide exchange factor, it rather has chaperone functions on its own (Behnke et al., 2015). The overall action of BiP can be regulated by modifications, such as the phosphorylation of serine and threonine residues or ADP ribosylation (Carlsson and Lazarides, 1983; Hendershot et al., 1988).

With the investigation of the three Hsp70 chaperones DnaK, Ssc1, and BiP, the function of chaperones in different eukaryotic compartments can be compared to their prokaryotic homolog. All three chaperones share a high sequence similarity and many perform overlapping functions. Besides their role in protein folding, Ssc1 and BiP, for example, are involved in protein translocation. A functional substitution of DnaK for Ssc1 and vice versa, however, is not possible (Deloche et al., 1997; Moro et al., 2002). A detailed investigation of the underlying molecular mechanism for different functionalities of the Hsp70s is therefore of high interest.

6.1.5 Conformational cycle of the Hsp70 BiP observed by spFRET

In several previous studies, the conformational cycle of different Hsp70 chaperones was observed using spFRET (Mapa et al., 2010; Marcinowski et al., 2011; Sikor et al., 2013). The Hsp70 chaperones were either fluorescently labeled on both of the domains to observe the interdomain distance or on the β-sheet of the SBD and its α-helical lid to analyze the conformational change of the lid. An example of the FRET efficiencies observed for the conformational cycle of BiP is displayed in Figure 6.2. In the nucleotide free form and in the ADP-bound state, the lid adopts a closed conformation with a FRET efficiency of 0.8 (Figure 6.2A), while the interdomain distance is large with a FRET efficiency of 0.4 (Figure 6.2B). Upon ATP or AMP-PNP binding, the lid opens (E=0.1) and a small interdomain distance (E=0.7) is observed.

These data agree well with observations of the conformations of BiP in different nucleotide states by Marcinowski et al. (2011). In this publication commercially available ADP was used, which contains around 5% ATP. The high affinity of BiP for ATP led to a second population in the presence of commercially available ADP. In the here performed experiments, ADP was purified by anion exchange chromatography and thus a clear major population was observed for ADP-bound BiP. Data of the same protein in the presence of commercially available ADP can be found in the Appendix.

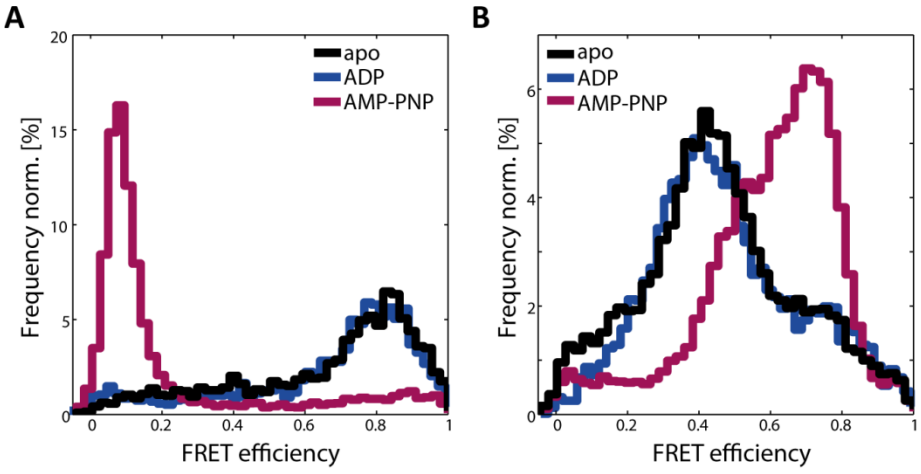


Figure 6.2 Conformational cycle of the Hsp70 chaperone BiP observed by spFRET using PIE-MFD. FRET efficiency histograms are shown for BiP alone (black), ADP-bound BiP (blue), and the AMP-PNP-bound protein (magenta). (**A**) The conformation of the lid was monitored using BiP-C519-C638-Atto532/Atto647N. (**B**) BiP-C167-C519-Atto532/Atto647N was used to analyze the interdomain distance.

SpFRET is ideally suited to not only investigate the conformational changes, but also their dynamics. Conformational dynamics on the timescale of micro- to milliseconds were not observed for BiP (Marcinowski et al., 2011). In the case of Ssc1, spFRET experiments using PIE-MFD showed conformational switching for the interdomain distance of the ADP- and ATP-bound form (Mapa et al., 2010; Sikor, 2011). Interestingly, Ssc1 displayed dynamic conformational changes of the lid and between the SBD and the NBD on the timescale of seconds observed by TIRF microscopy. The lid dynamics were observed in the ADP-bound form on timescales of around 4 s, while the domains underwent dynamics on timescales of 1 s when ADP was bound and changed to shorter switching times in the presence of ATP (Sikor et al., 2013). Here, we set out to investigate these conformational dynamics in detail especially regarding the importance of allosteric communication between the domains of the Hsp70s.

6.2 Results

Conformational dynamics of Hsp70 chaperones were analyzed in a comparative approach with Hsp70 chaperones from three different organelles and organisms. Prokaryotic DnaK was obtained from *E. coli*, mitochondrial Ssc1 originated from *S. cerevisiae*, and the ER-associated BiP stemmed from *M. musculus*. To get an idea of the sequence conservation, we searched for amino acid sequence similarities between the three proteins. The Hsp70 chaperones are highly conserved with an amino acid sequence identity of 58% between DnaK and Ssc1, 48% between DnaK and BiP, and 51% between Ssc1 and BiP. For the alignment of the primary sequences see the Appendix. Generally, the conservation of Hsp70s within a single organelle throughout all eukaryotic species is higher than the similarities of Hsp70 chaperones within one species. Hsp70 chaperones therefore existed in the α-proteobacteria and cyanobacteria before the first and second endosymbiotic event. The comparison of the three chosen Hsp70 chaperones, even though from different species, therefore allows the study of differences in conformational motion and functions of Hsp70s from different organelles.

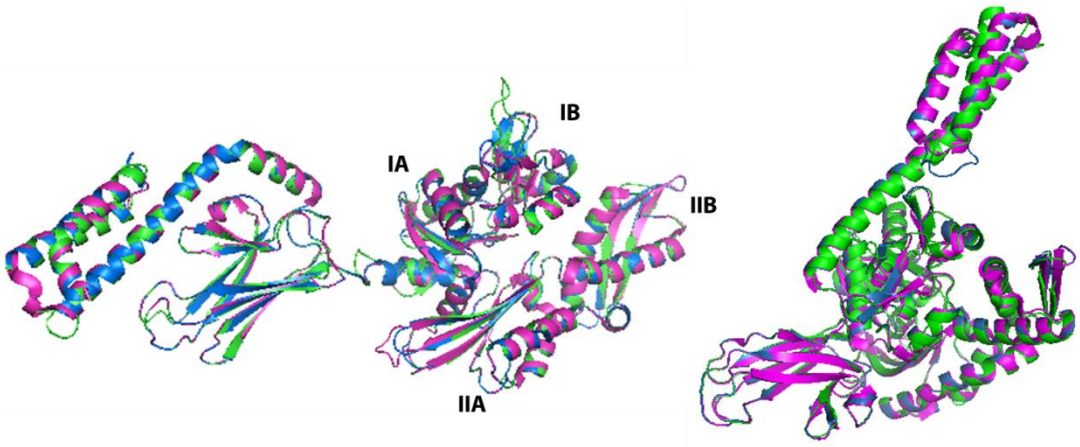


Figure 6.3 Structural models of Ssc1 and BiP based on the known crystal structure of DnaK (PDB: 2KHO, 4B9Q) in the ADP-bound form (left) and in the ATP-bound conformation (right). DnaK: magenta, Ssc1: blue, BiP: green. IA, IIA, IB, and IIB indicate different subdomains on the nucleotide binding domain.

To investigate the structural similarity, the primary sequences of Ssc1 and BiP were modelled onto the known ADP- and ATP-bound crystal structures of DnaK (Figure 6.3). The structure of all three proteins could be modelled to highly similar three-dimensional structures for both nucleotide states. In addition, the recently solved crystal structure of BiP in the ATP-bound form shows a high similarity to the ATP-bound form of DnaK (Yang et al., 2015).

6.2.1 Functional integrity of the Hsp70s after attachment of fluorescent moieties

The allosteric effects of conformational dynamics were investigated by the attachment of three fluorophores to simultaneously measure the interdomain distance as well as the lid conformation. To this end, the chaperones were recombinantely expressed carrying a non-natural amino acid and two cysteines. The positions for the three fluorescent labels were placed at the same sites as for previous Hsp70 constructs, with which spFRET experiments have been performed (Kudryavtsev et al., 2012; Mapa et al., 2010; Marcinowski et al., 2011). DnaK thus contained the non-natural amino acid PrK at position 318 and the cysteines at C425 and C563. Similarly, Ssc1-PrK341-C448-C590 and BiP-PrK167-C519-C638 were used. The attachment points of the fluorophores for all three chaperones are displayed in Figure 6.4. With the use of one non-natural amino acid, one fluorescent label (Atto488) was attached site-specifically and both cysteines were labeled stochastically with the fluorophores Atto565 and Atto647N. The lid conformation can therefore be extracted directly by illumination with the green laser. The quality of the purification and the fluorescent labeling of the proteins was monitored by UV-illumination and subsequent Coomassie staining of the proteins separated by SDS-PAGE.

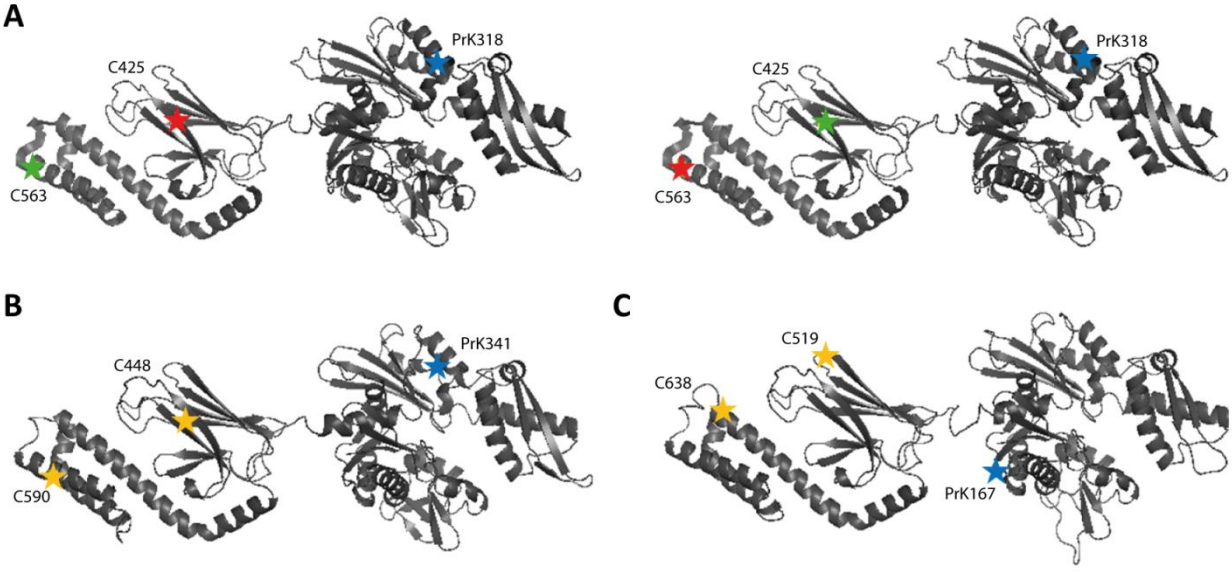


Figure 6.4 Site-specific attachment of fluorophores to the Hsp70s DnaK (**A**), Ssc1 (**B**), and BiP (**C**). Atto488 (blue star) was attached site-specifically to the non-natural amino acid PrK. Atto565 and Atto647N were attached to two cysteine residues. In (A) both possible combinations of Atto565 (green star) and Atto647N (red star) fluorophore attachment to the cysteines are displayed exemplarily for DnaK. PrK was placed at position 318 for DnaK (A), 341 for Ssc1 (B), and 167 for BiP (C). The cysteine residues were site-specifically integrated at positions 425 and 563 for DnaK (A), 448 and 590 for Ssc1 (B), and 519 and 638 for BiP (C).

The functional activity of the fluorescently labeled proteins was tested by analysis of their ATPase activity. To achieve this, an ATP-NADH coupled ATPase assay was used. The ATPase activity rate of BiP without a non-natural amino acid (BiP-C519-C638) was 0.18/min. The insertion of the non-natural amino acid did not affect the activity (ATP hydrolysis rate: 0.22/min), neither did the attachment of three fluorophores lower the activity significantly (0.18/min for BiP-PrK166-C519-C638-Atto488/565/647N). The ATP hydrolysis rates found here are in agreement with previously reported values (Bonomo et al., 2010; Welsh et al., 2011). The relative values of different BiP constructs compared to unlabeled BiP-C519-C638 are displayed in Figure 6.5. The overall ATPase activity of triple-labeled DnaK and Ssc1 under the same conditions was slightly higher with around 1.7/min. and 1.46/min. Previously reported values were in the range of 0.4-0.7/min (Bonomo et al., 2010; Welsh et al., 2011). Considering, that the values strongly depend on the experimental conditions, such as temperature, pH, and ion concentration, and that different approaches of determining the amount of ATP can have different sensitivities, these values are in good agreement with previous reports.

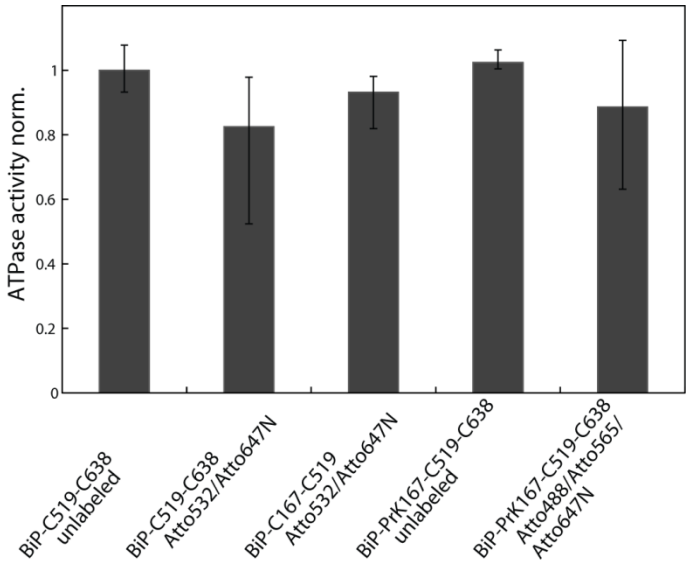


Figure 6.5 ATPase activity of BiP constructs in comparison. The ATPase activity of BiP-C519-C638 was normalized to a value of 1. The ATPase activity of the other constructs is shown as a relative value. Unlabeled BiP with two modified cysteines or the additional non-natural amino acid N-propargyl-lysine is displayed in comparison to the corresponding fluorescently labeled BiP constructs.

To investigate, whether one labeling position was preferred for one of the fluorophores due to the protein environment, mass spectrometry and HPLC were used. Different mass spectrometric approaches and digestion into peptides by ArgC with the separation of the peptides by PAGE or HPLC could not answer the question, which fluorophore was attached to which of the cysteines. Mass spectrometry could, however, confirm the purity of the protein of interest.

To get an estimate of the effects of the stochastic labeling at both cysteines on the FRET efficiency distributions E_{BG} and E_{BR}, we purified BiP constructs containing one non-natural amino acid and a single cysteine. These two positions were then site-specifically labeled with Atto488 bound to PrK and either Atto565 or Atto647N bound to the cysteine. SpFRET

experiments of all combinations of constructs in the presence of different nucleotides were performed in solution using PIE-MFD (Figure 6.6). The FRET efficiency distributions of BiP-PrK167-C638 and BiP-PrK167-C519 investigating the interdomain distance showed strong similarities for the FRET efficiencies E_{BG} and E_{BR}. In the nucleotide-free state, BiP mainly adopted a conformation with a mean FRET efficiency E_{BG} of 0.3 for BiP-PrK167-C638 or for BiP-PrK167-C519, both labeled with Atto488/Atto565 (FRET efficiencies E_{BR} of 0.2 and 0.1 for both constructs labeled with Atto488/Atto647N, respectively). Small impurities in the spFRET experiment of BiP-PrK167-C519-Atto488/Att565 broadened the FRET efficiency histogram. In this conformation, the position of the fluorophore (C519 versus C638) cannot be distinguished by the FRET efficiency. The mean FRET efficiencies were similar also for the ATP- and ADP-bound BiP in most cases. A deviation of the mean FRET efficiency E_{BR} for the ATP-bound BiP-PrK167-C638 compared to BiP-PrK167-C519 might be the result of the larger number of corrections necessary for the FRET efficiency value E_{BR}.

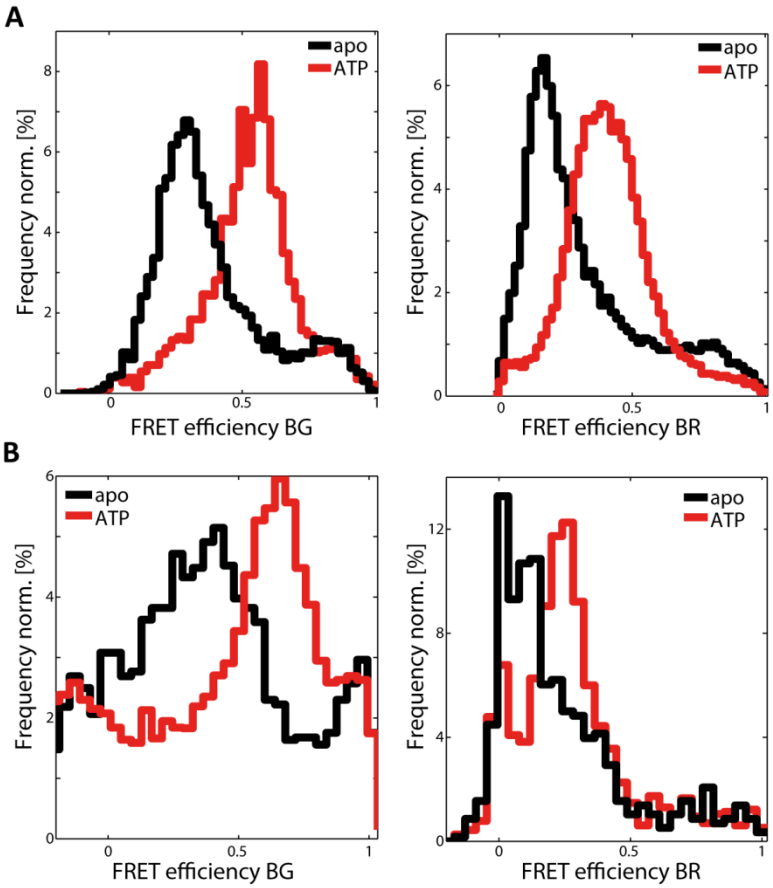


Figure 6.6 SpFRET of dual-labeled BiP by PIE-MFD. BiP was measured in the absence of nucleotides (black) and in the presence of ATP (red). (**A**) BiP-PrK167-C638 labeled with Atto488/Atto565 (left) and labeled with Atto488/Atto647N (right). (**B**) BiP-PrK167-C519 labeled with Atto488/Atto565 (left) and Atto488/Atto647N (right).

The interdomain distance can therefore be determined from a mixture of the FRET efficiencies from both positions, C519 and C638, for all nucleotide states. A distinction between both interdomain constructs is thus not necessary for a correct interpretation of the data. All three combinations of one-dimensional FRET efficiency histograms, E_{GR}, E_{BG}, and E_{BR}, will be presented from here onwards. For the two-dimensional analysis, however, the

focus will be on the lid conformation (FRET efficiency E_{GR}) compared to the interdomain distance (as a mixture of NBD-β-sheet SBD and NBD-lid) (FRET efficiency E_{BG}).

6.2.2 Conformational changes of Hsp70s observed by three-color FRET

Active, triple-labeled BiP, Ssc1, and DnaK were used to study coordinated motion along their conformational cycle. They were studied in solution by three-color FRET using PIE-MFD. The FRET efficiency histograms of triple-labeled Hsp70s simultaneously contain the distance information of the interdomain space and the lid conformation.

BiP

The FRET efficiency distributions of triple-labeled BiP are displayed in Figure 6.7. The FRET efficiency E_{GR} describes the conformation of the lid, while the FRET efficiencies E_{BG} and E_{BR} offer information about the interdomain distance. The mean FRET efficiencies E_{GR} observed for the major population of the lid conformation were 0.87 for the apo state and 0.2 for the ATP-bound conformation of BiP. BiP bound to ADP occurred in a similar conformation as observed for the apo state with a mean FRET efficiency E_{GR} of 0.89. Additionally, the interdomain distance was probed by FRET between the fluorophore on the nucleotide binding domain and both positions on the substrate binding domain. We observed mean FRET efficiencies E_{BG} of 0.12 for the apo state and 0.56 for the ATP-bound conformation. In the ADP-bound form, the protein showed a large interdomain distance with a FRET efficiency E_{BG} of 0.15 for its major population. The FRET efficiency E_{BG} showed a broad distribution. Its shoulder was described by a minor population with a mean FRET efficiency E_{BG} of 0.44. Even though the FRET efficiencies from the nucleotide-binding domain to both positions on the substrate-binding domain were very similar (Figure 6.6), the large width of the FRET efficiency distribution E_{BG} may be attributed to two different distances, which result from the stochastic labeling of both cysteines on the substrate binding domain with the Atto565 fluorophore.

The conformations obtained from three-color FRET data of triple-labeled BiP nicely correspond to the previous results of both dual-labeled BiP constructs (BiP-C519-C638, BiP-C167-C519) (Figure 6.2).

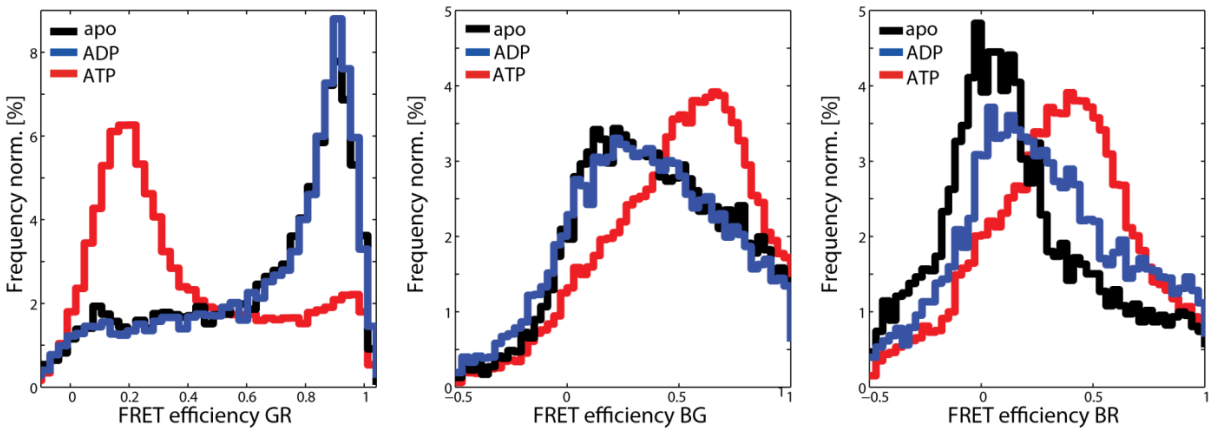


Figure 6.7 Single-molecule three-color FRET of BiP by PIE-MFD. FRET efficiency histograms E_{GR} (left) sampling the lid conformation, E_{BG} (middle), and E_{BR} (right) investigating the interdomain distance. BiP was measured in the absence of nucleotides (black) and when bound to ADP (blue), or ATP (red).

The careful analysis of the one-dimensional FRET efficiency histograms already suggests an anticorrelated behavior of the lid conformation compared to the interdomain distance. To investigate the allosteric effects of the different domains in detail, we exploited the full information contained in three-color experiments and analyzed two-dimensional FRET efficiency histograms (Figure 6.8). Coordinated motion becomes apparent when analyzing the trend in which the populations appear in the two-dimensional histograms. When comparing BiP in all nucleotide states, a population with a high FRET efficiency E_{GR} always corresponded to a population with low FRET efficiency E_{BG}. This is clearly visible for the major populations of BiP in the apo and ADP-bound or ATP-bound state (Figure 6.8A-C). BiP therefore either existed with a short interdomain distance and an open lid or with a large interdomain distance and a closed lid.

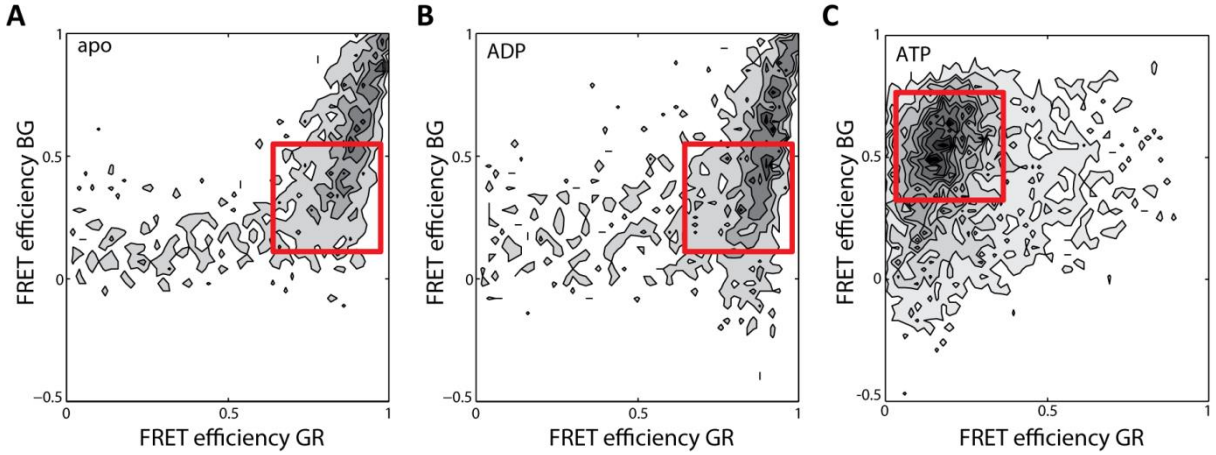


Figure 6.8 Two-dimensional FRET efficiency histograms of three-color FRET data of BiP obtained using PIE-MFD experiments. The FRET efficiencies E_{GR}, containing information about the lid conformation, and E_{BG}, probing the interdomain distance, are displayed for the nucleotide-free form of BiP (**A**), ADP-bound (**B**), and ATP-bound (**C**) BiP.

In general, the correct analysis of these two-dimensional FRET efficiency histograms is difficult to perform due to the large effects of the correction factors. Correlations are inherent to the proximity ratios in GR and BG, because a higher FRET efficiency from blue to green

results in a higher red intensity, if FRET from green to red occurs. Therefore, Anders Barth implemented a three-color PDA analysis, in which the original proximity ratios are used to determine absolute distances (manuscript in preparation). An exemplary three-dimensional PDA fit of BiP in both nucleotide-bound states is displayed in Figure 6.9. The data is fit to three populations, for each of which three distances and widths of the distance distributions for all combinations of fluorophores are obtained.

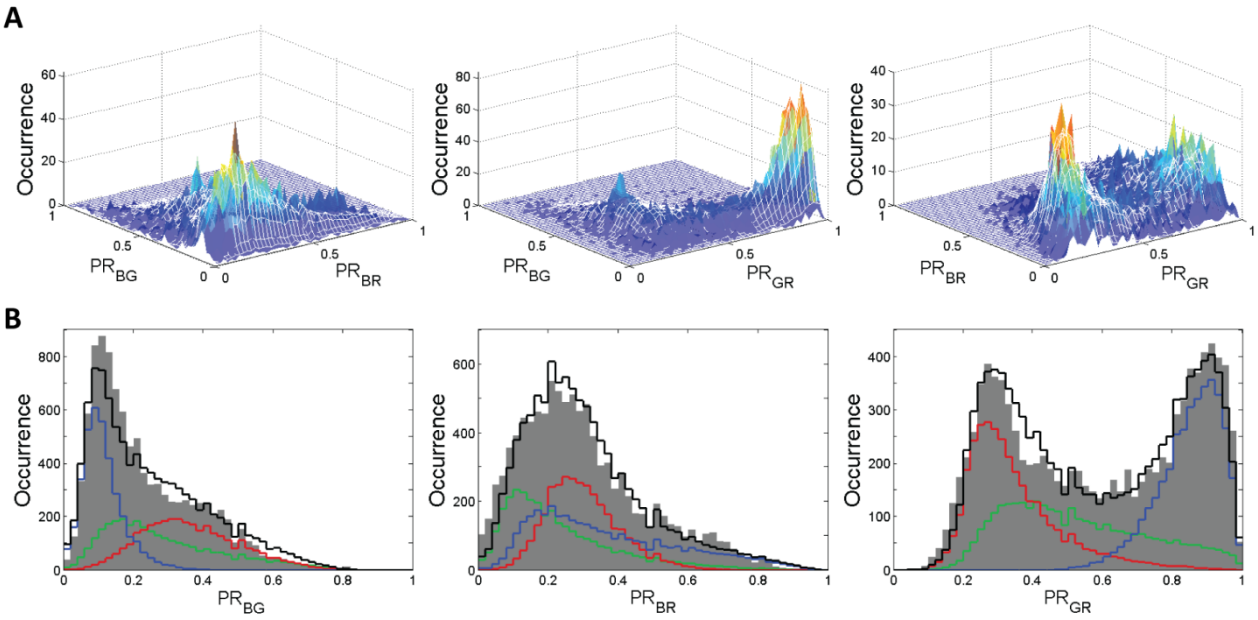


Figure 6.9 Three-color PDA fit shown exemplarily for BiP in the presence of commercially available ADP. (**A**) Three-dimensional histograms of the data (color) with their PDA fit (white lines). (**B**) One-dimensional projections of the histograms with the overall fit (black) and the individual fits for each of the three subpopulations (blue, green, red) of BiP.

The obtained distance distributions of BiP are displayed as two-dimensional distance distribution histograms in Figure 6.10. Fitting the data with a photon distribution for three different populations clearly shows a shift of the populations between the different nucleotide-bound states. Whenever the lid was in a closed conformation in the apo or ADP-bound state (distance R_{GR}=50 Å), the interdomain distance was large (R_{BG}=68 Å) (Figure 6.10A, B). While on the other hand, in the ATP-bound state the interdomain distance was small (R_{BG}=57 Å) and the lid appeared in an open conformation (R_{GR}=87 Å) (Figure 6.10C).
The subpopulations did not only differ in their absolute distances, but also the widths of the distributions changed. For a highly distinct lid conformation (R_{GR}=49 Å, σ=6 Å), the interdomain distance showed a large variation (R_{BG}=70 Å, σ=19 Å) (Figure 6.10A, B), while a distinct low interdomain distance (R_{BG}=60 Å, σ=10 Å) appeared in combination with larger variations in the lid conformation (R_{GR}=87 Å, σ=15 Å) (Figure 6.10C). BiP in the apo or ADP-bound state was therefore more flexible in its interdomain distance, while ATP-bound BiP showed a higher flexibility of the lid.

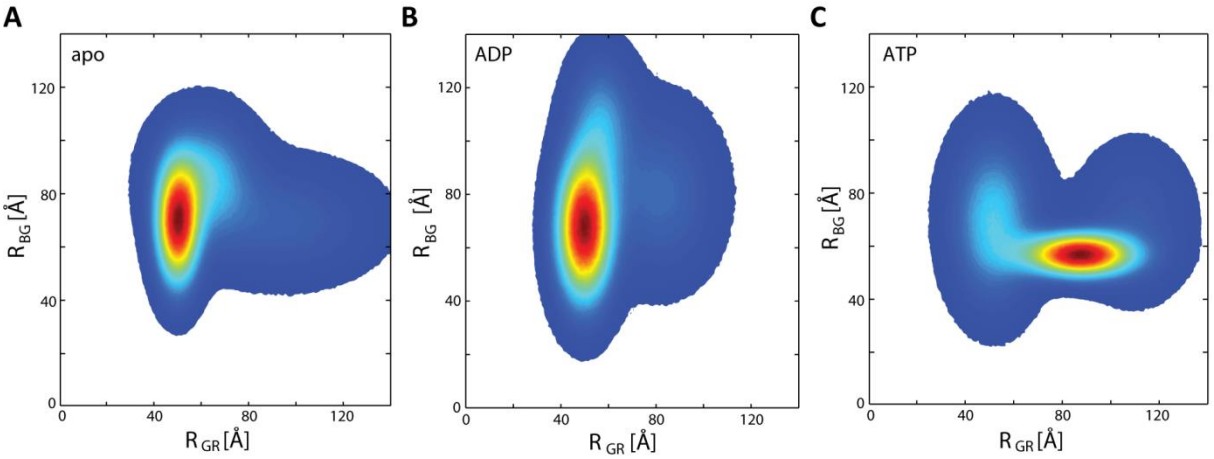


Figure 6.10 Two-dimensional distance distributions of the lid conformation and the interdomain distance of BiP obtained by the three-color PDA analysis. The distance distributions were obtained by fitting the proximity ratios to three subpopulations. The two-dimensional distance distributions are displayed for BiP in the nucleotide-free state (**A**), ADP-bound BiP (**B**), and ATP-bound BiP (**C**).

Ssc1

The mitochondrial Hsp70 Ssc1 was investigated by the same means. Ssc1 needs the presence of nucleotides for its stability. Purification of the protein was thus performed in the presence of ATP, which in the Ssc1-bound form was hydrolyzed to ADP. Its conformation was therefore only investigated in the presence of ADP and ATP. Three-color FRET experiments of triple-labeled Ssc1 were performed using PIE-MFD. The interdomain distance was monitored by the FRET efficiency E_{BG} and the lid conformation was observed by analyzing the FRET efficiency E_{GR}. The major population of ADP-bound Ssc1 occurred with a high FRET efficiency E_{GR}=0.93 and a low FRET efficiency E_{BG}=0.04 (Figure 6.11). ADP-bound Ssc1 therefore adopted a large interdomain distance with a closed lid. The major population of ATP-bound Ssc1, on the other hand, was found with a small distance between its domains (E_{BG}=0.83) and an open lid (E_{GR}=0.15). A minor population with FRET efficiencies E_{GR}=0.68 and E_{BG}=0.42 was additionally observed. This population may correspond to Ssc1 molecules carrying a hydrolyzed ATP.

The here observed three-color FRET data agree well with spFRET data by Martin Sikor, which showed the conformational change upon ADP- and ATP-binding to the Ssc1 constructs Ssc1-C341-C448 probing the interdomain distance and Ssc1-C448-C590 investigating the lid conformation (Mapa et al., 2010). As for BiP, a fraction of ATP-bound Ssc1 in the ADP-bound state for the spFRET experiments has to be taken into consideration when comparing the results. Martin Sikor described conformational dynamics of Ssc1 in the ATP-bound as well as in the ADP-bound form, which can explain the broad distributions of Ssc1's FRET efficiency histogram (Figure 6.11). Furthermore the stochastic labeling with Atto565 and Atto647N on both cysteines in the SBD may lead to broadening of the FRET efficiency E_{BG}.

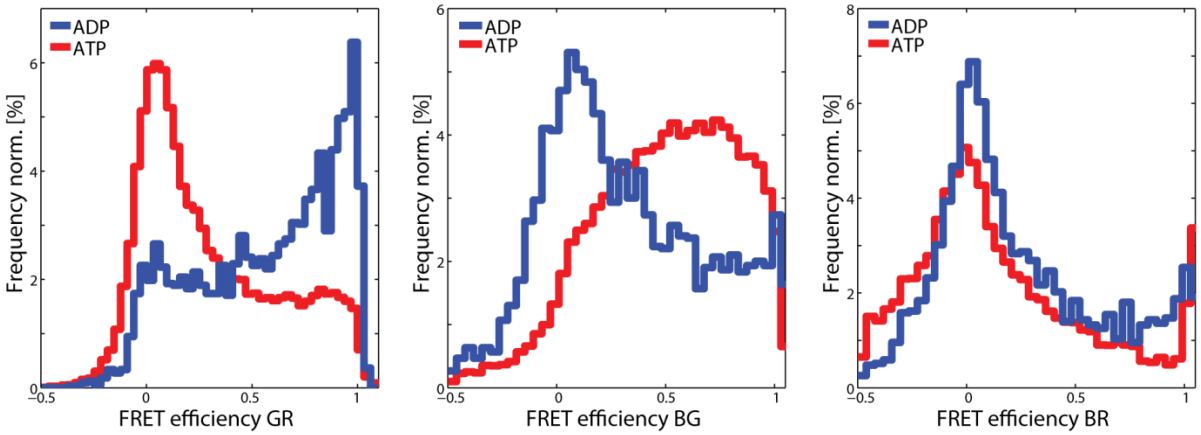


Figure 6.11 Single-molecule three-color FRET of Ssc1 by PIE-MFD. FRET efficiency histograms E_{GR} (left) sampling the lid conformation, and E_{BG} (middle), and E_{BR} (right) investigating the interdomain distance. Ssc1 was measured in the presence of 1 mM ADP (blue) or 1 mM ATP (red).

As for BiP, a three-color PDA analysis was performed for the ADP- and ATP-bound conformation of Ssc1. Although the distance distribution histograms are easier to interpret than the original FRET efficiency data, an analysis has to be performed carefully, as the fitting of the data can already bias the obtained results. The number of populations used to fit the data can determine, whether a coordinated behavior is observed in the two-dimensional distance distributions. Furthermore, fitting to multiple populations can be ambiguous and can also affect the distance distributions. Even though, the data of two-dimensional FRET efficiency histograms is more difficult to describe, we therefore chose to display these histograms in addition to the distance distributions (Figure 6.12A). The two-dimensional FRET efficiency histograms of Ssc1 in its ADP- and ATP-bound conformation show an anticorrelated behavior of its domains with either a high FRET efficiency E_{GR} and a low FRET efficiency E_{BG} or vice versa for the major population of molecules. In addition, a small population of molecules in the ADP-bound state with both, a low FRET efficiency E_{GR} and E_{BG}, existed. For a constant interdomain distance, the lid of ADP-bound Ssc1 therefore shows some flexibility in its position. The obtained two-dimensional distance distributions are displayed in Figure 6.12B. The major population of ADP-bound Ssc1 showed a large interdomain distance (R_{BG1}=67 Å, σ=25 Å) and a closed lid (R_{GR1}=46 Å, σ=5 Å). The histogram was described by two additional minor populations with interdomain distances R_{BG2}=66 Å (σ=23 Å) and R_{BG3}=70 Å (σ=26 Å) and lid conformations R_{GR2}=66 Å (σ=10 Å), and R_{GR3}=110 Å (σ=14 Å), respectively. For ATP-bound Ssc1, the interdomain distance decreased (R_{BG1}=60 Å (σ=18 Å), R_{BG2}=59 Å (σ=16 Å), R_{BG3}=59 Å (σ=14 Å)), while the lid adopted an open conformation (R_{GR1}=100 Å (σ=20 Å), R_{GR2}=71 Å (σ=7 Å), R_{GR3}=53 Å (σ=7 Å)) (Figure 6.12B). For ADP-bound Ssc1, several conformations of the lid with different lid to β-sheet separations were observed, while for all these states the interdomain separation stayed constant with a distance of around 60 Å. Therefore, in the ADP-bound state, the lid seems to move independently of the interdomain configuration. In the ATP-bound state, minor populations with a much lower fraction of molecules as compared for the ADP-state with a similar behavior were observed. Some flexibility of the lid was thus found even

for the ATP-bound Ssc1. The flexibility was, however, more strongly pronounced in the ADP-bound conformation, where the interdomain distance is large.

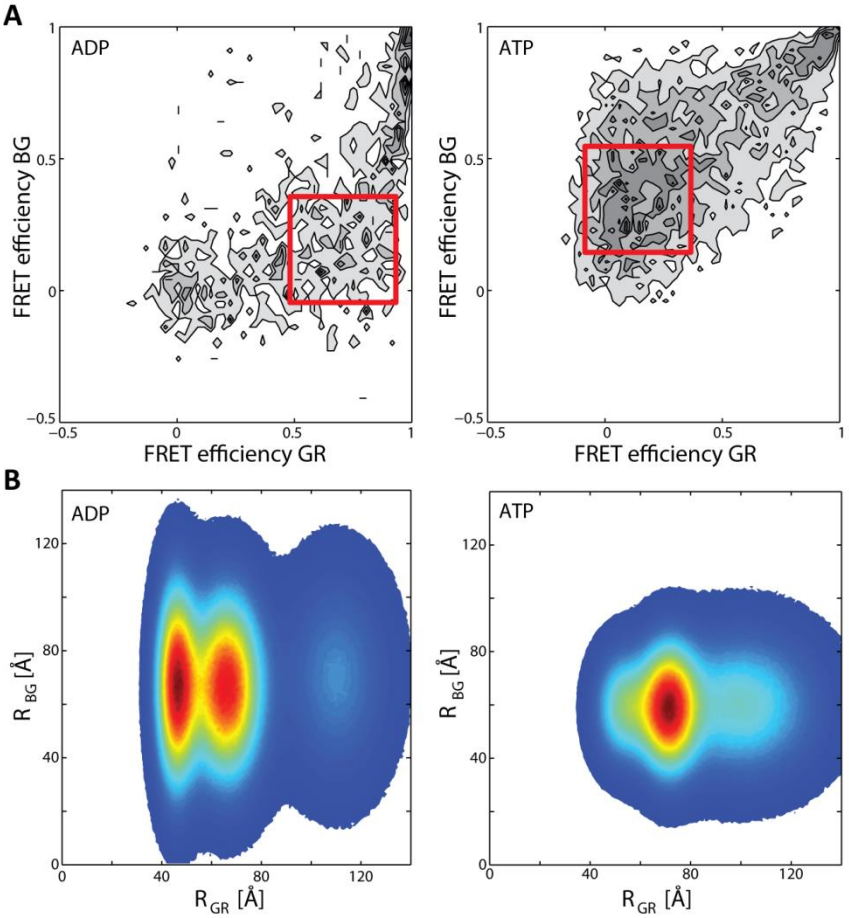


Figure 6.12 Two-dimensional FRET efficiency and distance distributions of the lid conformation (E_{GR}/R_{GR}) and the interdomain distance (E_{BG}/R_{BG}) of Ssc1. (**A**) Two-dimensional FRET efficiency histograms of three-color FRET data of Ssc1 in the presence of ADP (left) or ATP (right) obtained by PIE-MFD experiments. (**B**) Distance distribution histograms by fitting the proximity ratios PR_{GR} and PR_{BG} to three subpopulations with a three-color PDA analysis. The two-dimensional distance distributions are shown for ADP-bound (left), and ATP-bound Ssc1 (right).

DnaK

The conformational cycle of the prokaryotic DnaK was investigated by single molecule three-color FRET experiments as described for the other Hsp70 chaperones. DnaK was followed in its ADP-bound form, where it showed a closed lid conformation (E_{GR}=0.89) and a large interdomain distance (E_{BG}=0.13) for its major subpopulation. The minor subpopulation of DnaK exhibited the FRET efficiencies E_{GR}=0.40 and E_{BG}=0.59. In the presence of ATP, DnaK adopted two conformations. The major population showed an open lid configuration (E_{GR}=0.24) and a small interdomain distance (E_{BG}= 0.78). The minor subpopulation adopted a conformation similar to the one observed for ADP-bound DnaK with FRET efficiencies E_{GR}=0.9 and E_{BG}=0.28. This fraction might represent DnaK molecules containing a hydrolyzed ATP. The effect of ATP hydrolysis should be more pronounced for DnaK than for BiP due to its slightly higher intrinsic ATPase activity.

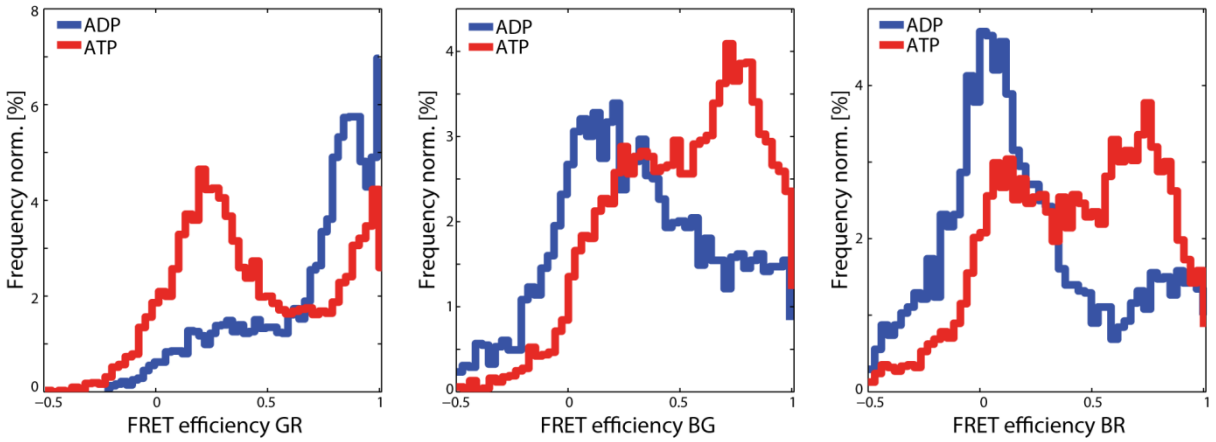


Figure 6.13 Single-molecule three-color FRET of DnaK by PIE-MFD. FRET efficiency histograms E_{GR} (left) sampling the lid conformation, and E_{BG} (middle), and E_{BR} (right) investigating the interdomain distance. DnaK was measured in the presence of 1 mM ADP (blue), or 1 mM ATP (red).

To investigate the allosteric coordination between the domains, two-dimensional FRET efficiency histograms are shown in Figure 6.14A. When comparing these histograms for the ADP- and ATP-bound form of DnaK, a clear difference in both FRET efficiencies E_{BG} and E_{GR} becomes apparent. The majority population of ADP-bound DnaK was found with a high FRET efficiency E_{GR} and a lower FRET efficiency E_{BG}, while ATP-bound DnaK appeared with a low FRET efficiency E_{GR} and a high FRET efficiency E_{BG}. Interestingly, a second population with the same low E_{GR} but also low E_{BG} was visible.

As for the other Hsp70 constructs, we investigated the coordinated behavior of DnaK by performing a three-color PDA analysis of the data including three subpopulations. The results are displayed as two-dimensional distance distribution histograms in Figure 6.14B. The major population of ADP-bound DnaK appeared with a large interdomain distance of R_{BG1}=82 Å (σ=4 Å) and a closed lid (R_{GR1}=52 Å, σ=4 Å). The minor subpopulations were described by the interdomain distances R_{BG2}=87 Å (σ=2 Å) and R_{BG3}=82 Å (σ=2 Å) and the lid-SBD distances R_{GR2}=37 Å (σ=5 Å) and R_{GR3}=63 Å (σ=9 Å). When DnaK was bound to ATP, the distance effect was reversed and the chaperone was mainly found with a short interdomain distance R_{BG1}=69 Å (σ=11 Å) and an open lid (R_{GR1}=85 Å, σ=13 Å). Two minor populations with the interdomain distances R_{BG2}=58 Å (σ=7 Å) and R_{BG3}=54 Å (σ=12 Å) and the lid-SBD distances R_{GR2}=96 Å (σ=15 Å) and R_{GR3}=61 Å (σ=11 Å) led to the broad appearance of the distance distribution histogram.

In the ADP- as well as in the ATP-bound state, DnaK adopted a clear major population with either a large interdomain separation and a closed lid (for the ADP-state) or a short interdomain distance and an open lid (for the ATP-state).

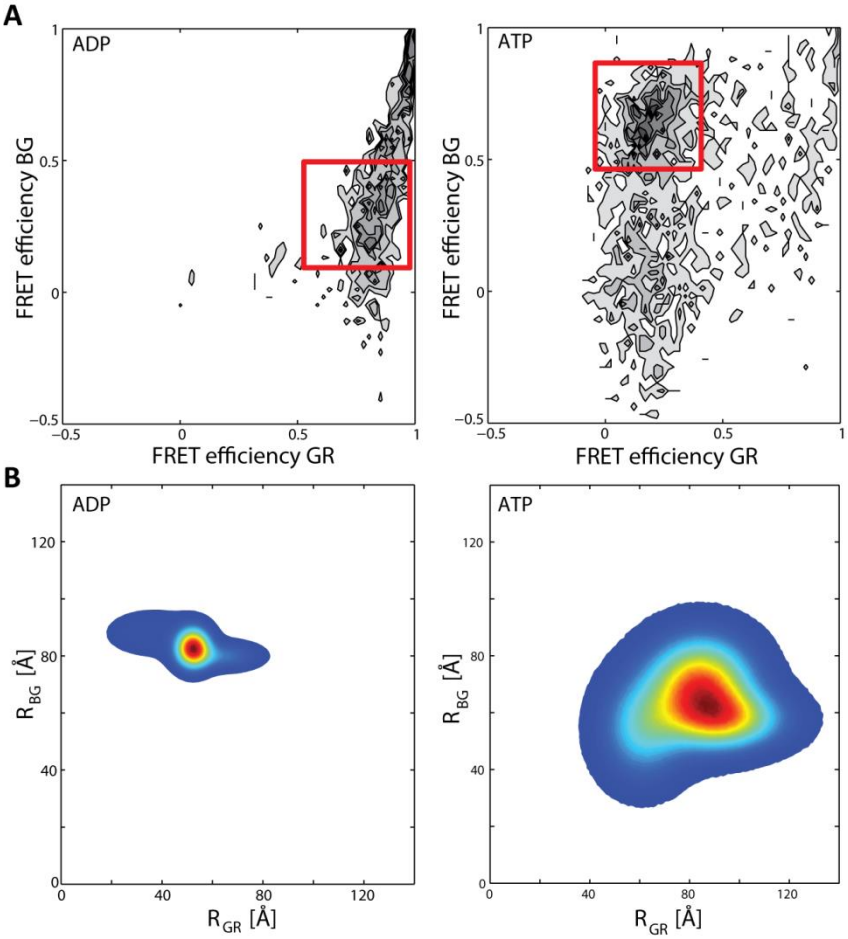


Figure 6.14 Two-dimensional FRET efficiency and distance distributions of the lid conformation (E_{GR}/R_{GR}) and the interdomain distance (E_{BG}/R_{BG}) of DnaK. (**A**) Two-dimensional FRET efficiency histograms of three-color FRET data of DnaK in the presence of ADP (left) or ATP (right) obtained by PIE-MFD experiments. (**B**) Distance distribution histograms by fitting the proximity ratios PR_{GR} and PR_{BG} to three subpopulations with a three-color PDA analysis. The two-dimensional distance distributions are shown for ADP-bound (left), and ATP-bound DnaK (right).

6.2.3 Comparative analysis of the allosteric effects in Hsp70s

Previously, differences in the allosteric communication of Hsp70 chaperones from different organelles have been reported (Behnke et al., 2015; Yang et al., 2015). Especially, the interdomain communication for the ADP-bound state of the Hsp70s was shown to differ. While for DnaK only transient contacts were observed, strong contacts between the domains were described for other Hsp70s (Behnke et al., 2015). The results obtained from PDA fits of the three-color FRET data of DnaK, Ssc1, and BiP were used to study differences in their coordinated behavior. A clearly anticorrelated behavior of the interdomain distance and the lid conformation was observed for the three chaperones when comparing their ADP- and ATP-bound states (Figure 6.8, Figure 6.10, Figure 6.12, Figure 6.14). Hsp70s in the ADP-bound form were primarily found with a large interdomain distance and closed lid, while ATP-bound Hsp70s appeared with a short interdomain distance and open lid.

When comparing the behavior of the subpopulations within one nucleotide-bound state, differences between the Hsp70s become apparent. While each nucleotide state of BiP and DnaK could be described by a single major population (Figure 6.8, Figure 6.10), several

subpopulations were necessary to explain the data for Ssc. Besides the major population with a large interdomain distance and a closed lid, additional populations of similar amplitudes with the same interdomain distance but different lid conformations were observed (Figure 6.12). This shows a higher flexibility of Ssc1 in the ADP-bound conformation with a constant large interdomain distance when compared to BiP or DnaK. This agrees with dynamic conformational changes observed by Martin Sikor (Mapa et al., 2010). Interestingly, DnaK rather showed a flexibility in its interdomain distance in the ATP-bound state, while the lid conformation stayed constant (Figure 6.14). The allosteric communication between both domains seems to have different strengths for Hsp70 chaperones from different organelles as already suggested by Behnke et al. (2015).

6.2.4 Behavior of the Hsp70 chaperones in the presence of substrate peptide and cofactors

The conformational cycle of the Hsp70 chaperones including ATP binding, hydrolysis, and ADP release is coupled to the binding of substrate to the SBD. In the ATP-bound conformation, the substrate binding as well as the release rate are high, leading to a fast substrate exchange. Upon ATP hydrolysis, the release rate decreases by three orders of magnitude, thus increasing the affinity of the Hsp70 for the substrate. Stable substrate binding therefore occurs in the ADP-bound Hsp70 conformation. Substrate binding to the Hsp70 chaperone was shown to increase its ATP hydrolysis rate. In addition to substrate binding, the cochaperones Hsp40 and nucleotide exchange factors can increase the rate of ATP hydrolysis. To fully understand the conformational cycle of the investigated chaperones, the conformation and allosteric coordination of the domains were therefore analyzed in the presence of substrate and nucleotide exchange factors.

BiP

Naturally, BiP binds to secreted proteins, especially antibody heavy chains, which are intrinsically disordered before binding to their respective light chains (Haas and Wabl, 1983; Hendershot, 1990). Marcinowski et al. (2013) investigated short peptide substrates derived from antibody heavy chains. They showed that binding of a peptide is not only sequence- but also substrate conformation-dependent. The described peptide HTFPAVL is used here to analyze the effect of substrate binding on BiP's conformation and coordinated motion.

For BiP in the presence of HTFPAVL, we obtained similar FRET efficiency distributions as observed for BiP alone. The mean FRET efficiencies E_{GR} observed for the lid conformation of BiP in the presence of the peptide were 0.88 for the nucleotide-free form and 0.16 for the major ATP-bound conformation. In the ADP-state, BiP's major population had a FRET efficiency E_{GR}=0.92. The interdomain distance showed mean FRET efficiency values E_{BG1}=0.24 and E_{BG2}=0.79 for the nucleotide-free BiP in the presence of peptide and E_{BG}=0.57 for ATP-bound protein. The mean FRET efficiencies E_{BG} in the ADP-bound conformation were E_{BG1}=0.07 and E_{BG2}=0.43. The addition of the substrate peptide HTFPAVL therefore did not introduce any observable conformational change, when compared to BiP in the absence of

substrate (Figure 6.7). This agrees with previously published spFRET data of ADP-bound BiP in the presence of the same substrate (Marcinowski et al., 2011).

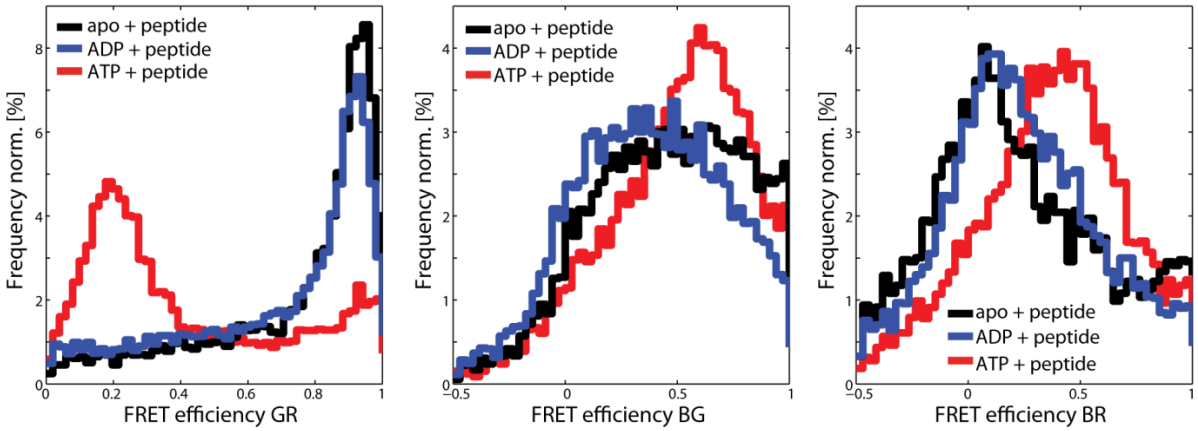


Figure 6.15 Single-molecule three-color FRET of BiP in the presence of substrate peptide HTFPAVL by PIE-MFD. FRET efficiency histograms E_{GR} (left) sampling the lid conformation, and E_{BG} (middle), and E_{BR} (right) investigating the interdomain distance. BiP was measured in the presence of 100 µM peptide (black) and 1 mM ADP (blue), or 1 mM ATP (red).

Ssc1

Mitochondrial Ssc1 is responsible for the transport of mitochondrial proteins into the matrix where it then aids in their folding. It interacts with stretches of around seven hydrophobic residues. Addition of the substrate peptide CALLLSAPRR to Ssc1 in the ADP-bound form led to a clearly discernable population with a large interdomain distance with FRET efficiency E_{BG}=0.15 and a closed lid (E_{GR}=0.82) for the majority of molecules (Figure 6.16A). In the presence of ATP-bound Ssc1, the binding of peptide shifted the FRET efficiencies E_{GR} to higher and E_{BG} to lower values (Figure 6.16B). The FRET efficiencies E_{BG} and E_{GR} of the major subpopulation were E_{BG}=0.13 and E_{GR}=0.69, respectively. We could thus visualize that peptide binding to Ssc1 occurred in the presence of both nucleotides. Whether peptide-bound Ssc1 directly hydrolyzed its ATP and formed a stable Ssc1-ADP-peptide complex could not be discerned from these histograms. Lid closing upon the substrate and a large interdomain distance were, however, clearly observed for Ssc1 in the presence of its artificial substrate CALLLSAPRR.

Ssc1 bound to its nucleotide exchange factor Mge1 was found in two populations. One subpopulation adopted a conformation similar to the one observed for ATP-bound Ssc1. The lid assumed a state with the same FRET efficiency E_{GR} as for the ATP-bound Ssc1. The interdomain distance adopted a value similar to the one obtained for ATP-bound Ssc1 in the presence of peptide. Due to the increased ATP hydrolysis rate in the presence of peptide and Mge1, a second population similar to the ADP-bound Ssc1 was additionally observed. The presence of peptide in combination with Mge1 induced both populations. An overview of the mean FRET efficiencies E_{GR} and E_{BG} for the major and minor populations of Ssc1 in the presence of the different factors under investigation is given in Table 6.1.

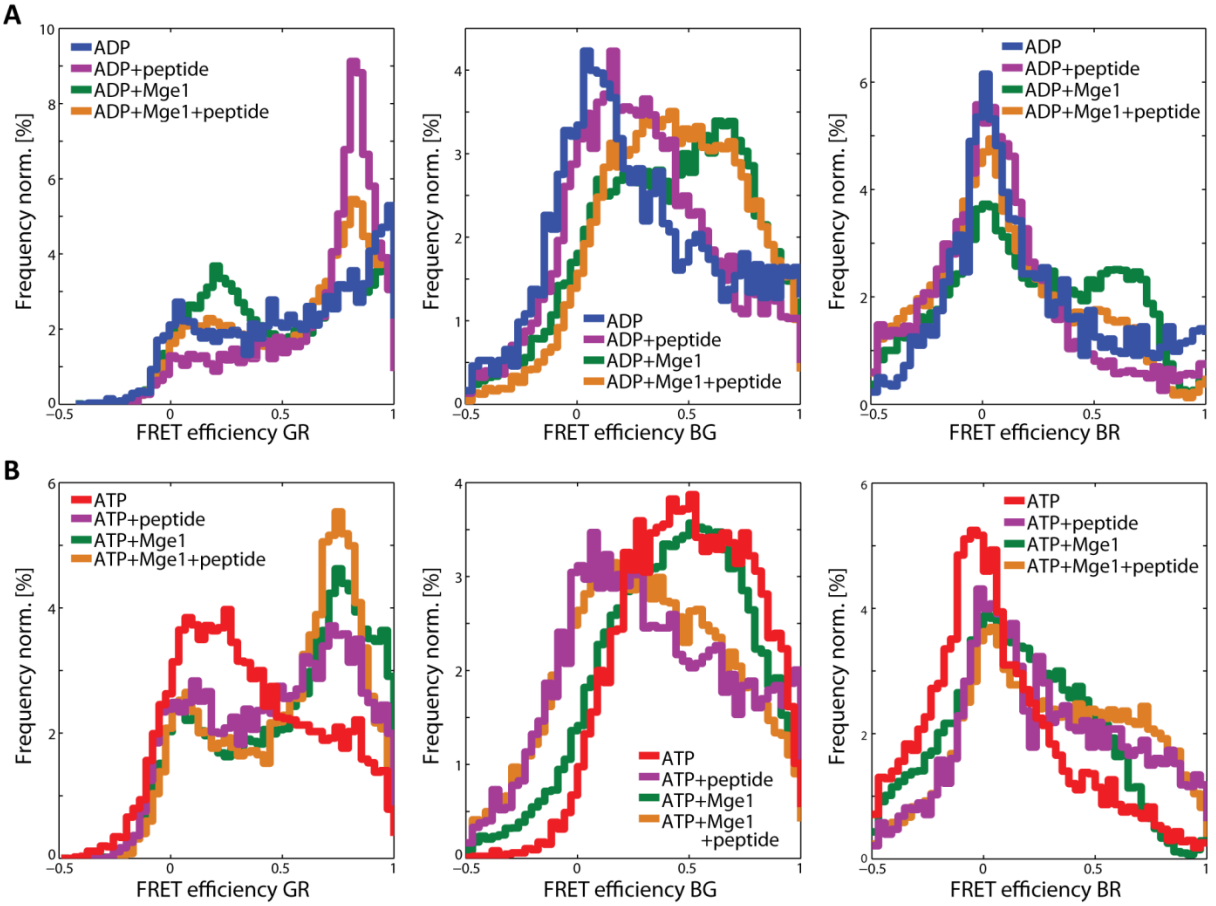


Figure 6.16 Single-molecule three-color FRET of Ssc1 in the presence of substrate peptide CALLLSAPRR and its nucleotide exchange factor Mge1 by spFRET with PIE-MFD. FRET efficiency histograms E_{GR} (left) sampling the lid conformation, and E_{BG} (middle), and E_{BR} (right) investigating the interdomain distance. (**A**) Ssc1 was measured in the presence of 1 mM ADP (blue) and 100 µM peptide (magenta), or 1 mM ADP, 5 µM Mge1 (green), and peptide (yellow). (**B**) Ssc1 in the presence of 1 mM ATP (red), and 100 µM peptide (violet), or 1 mM ATP, 1 µM Mge1 (green), and peptide (yellow).

Table 6.1 Mean FRET efficiencies of the major and minor populations of ADP- or ATP-bound Ssc1 in the presence of Mge1 or the peptide CALLLSAPRR.

Ssc1	E_{GR} (major)	E_{GR} (minor)	E_{BG} (major)	E_{BG} (minor)
+ADP	0.87	0.39	0.04	0.47
+ADP+peptide	0.82	0.67	0.15	0.49
+ADP+Mge1	0.85	0.19	0.20	0.69
+ADP+Mge1+peptide	0.81	0.21	0.27	0.71
+ATP	0.15	0.68	0.83	0.42
+ATP+peptide	0.69	0.10	0.13	0.77
+ATP+Mge1	0.77	0.12	0.70	0.31
+ATP+Mge1+peptide	0.73	0.16	0.16	0.71

DnaK

The prokaryotic DnaK interacts with all bacterial proteins, that require the aid of chaperones during protein folding. To this end, it binds to stretches of seven amino acids containing a high content of hydrophobic residues similar to the other Hsp70s. Substrate peptide binding to DnaK is known to introduce an increase in the ATP hydrolysis rate (Slepenkov and Witt, 2003). The conformations of DnaK were probed here in the presence of the substrate peptide CALLLSAPRR. A comparison of DnaK in the respective nucleotide states in the absence and presence of the substrate is displayed in Figure 6.17. The ADP-bound conformation of DnaK was largely unaffected by the peptide (Figure 6.17A). The mean FRET efficiencies E_{GR} and E_{BG} of the major subpopulation were 0.87 and 0.13, respectively (Table 6.2). The addition of substrate peptide to the ATP-bound DnaK induced FRET efficiency distributions which were highly similar to the ones observed for DnaK in the ADP-bound state (Figure 6.17B). The major subpopulation shifted from a mean FRET efficiency E_{GR} of 0.25 for the ATP-bound DnaK to a value of 0.88 in the presence of peptide. And the FRET efficiency distributions E_{BG} and E_{BR} sampling the interdomain distance showed an increase in their low FRET population (Figure 6.17A, B). The mean distances for the major and minor populations of DnaK in the presence of peptide are found in Table 6.2.

The presence of the substrate peptide therefore induced a DnaK conformation with FRET efficiencies similar to the ones observed for the ADP-bound protein. The conformational transition of the ATP-bound DnaK to a state similar to the ADP-bound form in the presence of peptide can simply be the result of peptide binding. In addition, the binding of substrate peptide increases the ATP hydrolysis rate, thus inducing stable peptide binding. Therefore the effect could be additionally induced by a larger fraction of DnaK actually bound to ADP and peptide.

In a next step, the conformation of DnaK was investigated in the presence of its nucleotide exchange factor GrpE. The effect of GrpE on the conformation of ADP-bound DnaK visible by the FRET efficiencies E_{GR}, E_{BG}, and E_{BR}, was minimal (Figure 6.17A). The addition of GrpE to ATP-bound DnaK, on the other hand, increased the fraction of molecules found in the conformation observed for ADP-bound DnaK (Figure 6.17B). This is in agreement with previous descriptions, where GrpE introduced a conformation similar to the ADP-bound state of DnaK. The crystal structures of *Geobacillus kaustophilus* DnaK in complex with GrpE shows DnaK in a conformation with a closed lid and large interdomain distance (Wu et al., 2012b). The here observed high FRET efficiency E_{GR} and low FRET efficiency E_{BG} agree very well with the previously described crystal structure. Moreover, the nucleotide exchange factor led to an increased ATP turnover and therefore the equilibrium may have shifted towards the ADP state. In the presence of substrate peptide, the effect was more strongly pronounced, especially visible for the FRET efficiency E_{GR}, because not only the nucleotide exchange but also the ATP hydrolysis was increased (Figure 6.17B). The mean FRET efficiencies of ADP- and ATP-bound DnaK in the presence of GrpE or in the combined presence of GrpE and peptide are summarized in Table 6.2.

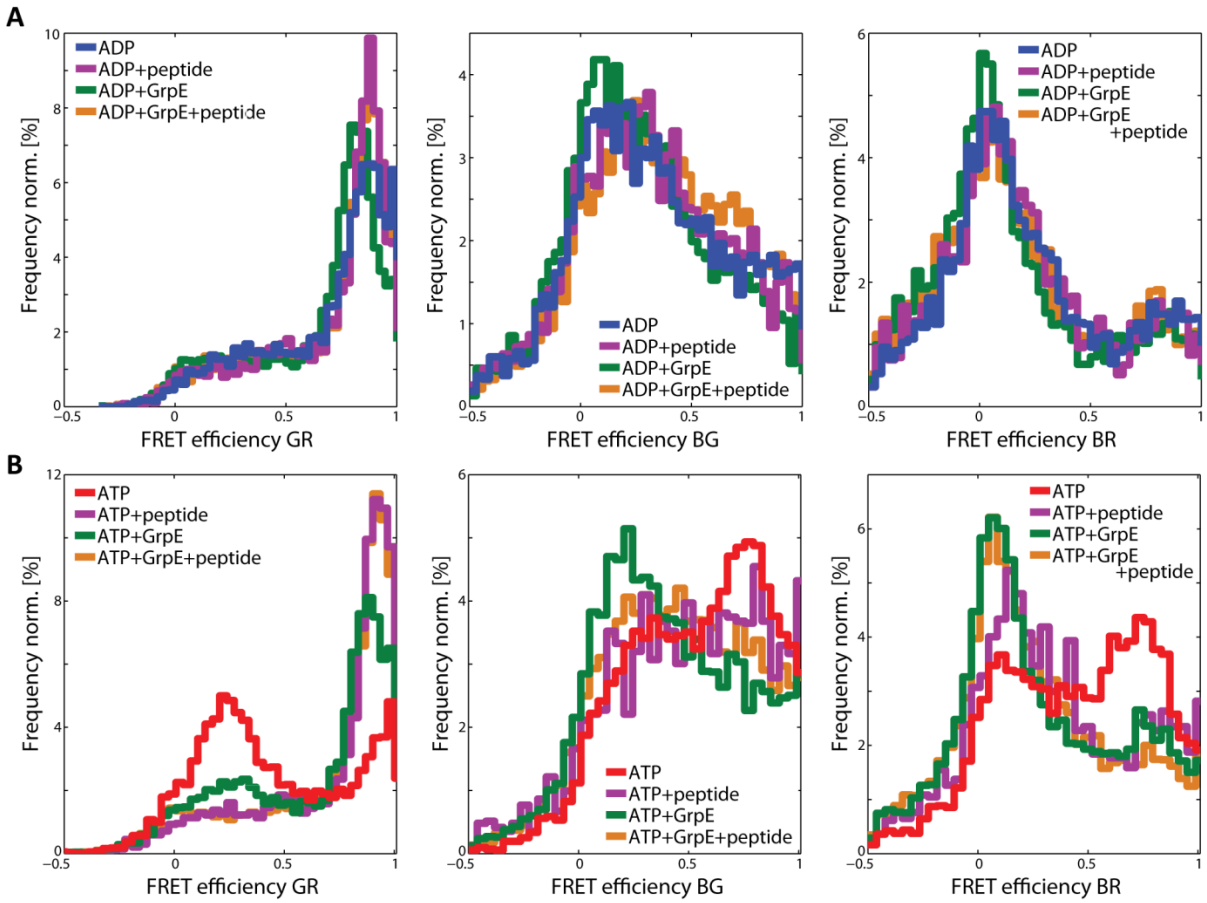


Figure 6.17 Single-molecule three-color FRET of DnaK in the presence of substrate peptide CALLLSAPRR and its nucleotide exchange factor GrpE by spFRET with PIE-MFD. FRET efficiency histograms E_{GR} (left) sampling the lid conformation, and E_{BG} (middle), and E_{BR} (right) investigating the interdomain distance. (**A**) DnaK was measured in the presence of 1 mM ADP (blue), and 100 µM peptide (magenta), or 1 mM ADP, 1 µM GrpE (green), and peptide (yellow). (**B**) DnaK in the presence of 1 mM ATP (red), and 100 µM peptide (magenta), or 1 mM ATP, 1 µM GrpE (green), and peptide (yellow).

Table 6.2 Mean FRET efficiencies of the major and minor populations of ADP- and ATP-bound DnaK in the presence of the peptide CALLLSAPRR or GrpE.

DnaK	E_{GR} (major)	E_{GR} (minor)	E_{BG} (major)	E_{BG} (minor)
+ADP	0.89	0.40	0.13	0.59
+ADP+peptide	0.87	0.79	0.22	0.81
+ADP+GrpE	0.81	1.00	0.08	0.35
+ADP+GrpE+peptide	0.87	0.45	0.14	0.47
+ATP	0.25	0.90	0.79	0.30
+ATP+peptide	0.88	0.29	0.22	0.78
+ATP+GrpE	0.91	0.48	0.37	0.86
+ATP+GrpE+peptide	0.91	0.51	0.33	0.85

6.2.5 Effects of individual nucleotide moieties on the allosteric communication

The investigation of the conformational states of Hsp70 chaperones in the presence of nucleotides offered to study the importance of individual groups of atoms of these nucleotides on the Hsp70 conformation. We concentrated here on the differences in nucleotide binding of prokaryotic DnaK and eukaryotic BiP. To assess the effect of the atoms, nucleotide analogs were used in addition to the naturally occurring ADP and ATP. The non-hydrolyzable ATP derivatives used for this investigation are visualized in Figure 6.18. Interestingly, the nucleotide analogs affected the interdomain distance as well as the lid conformation usually in a similar manner than either of the natural substrates ADP or ATP. The effect on the Hsp70 conformation was therefore mainly monitored using the information from the FRET efficiency E_{GR}, because here the difference in the nucleotide state was most pronounced. In cases where the interdomain distance was affected differently from the lid conformation, both FRET efficiencies E_{GR} and E_{BG} are presented.

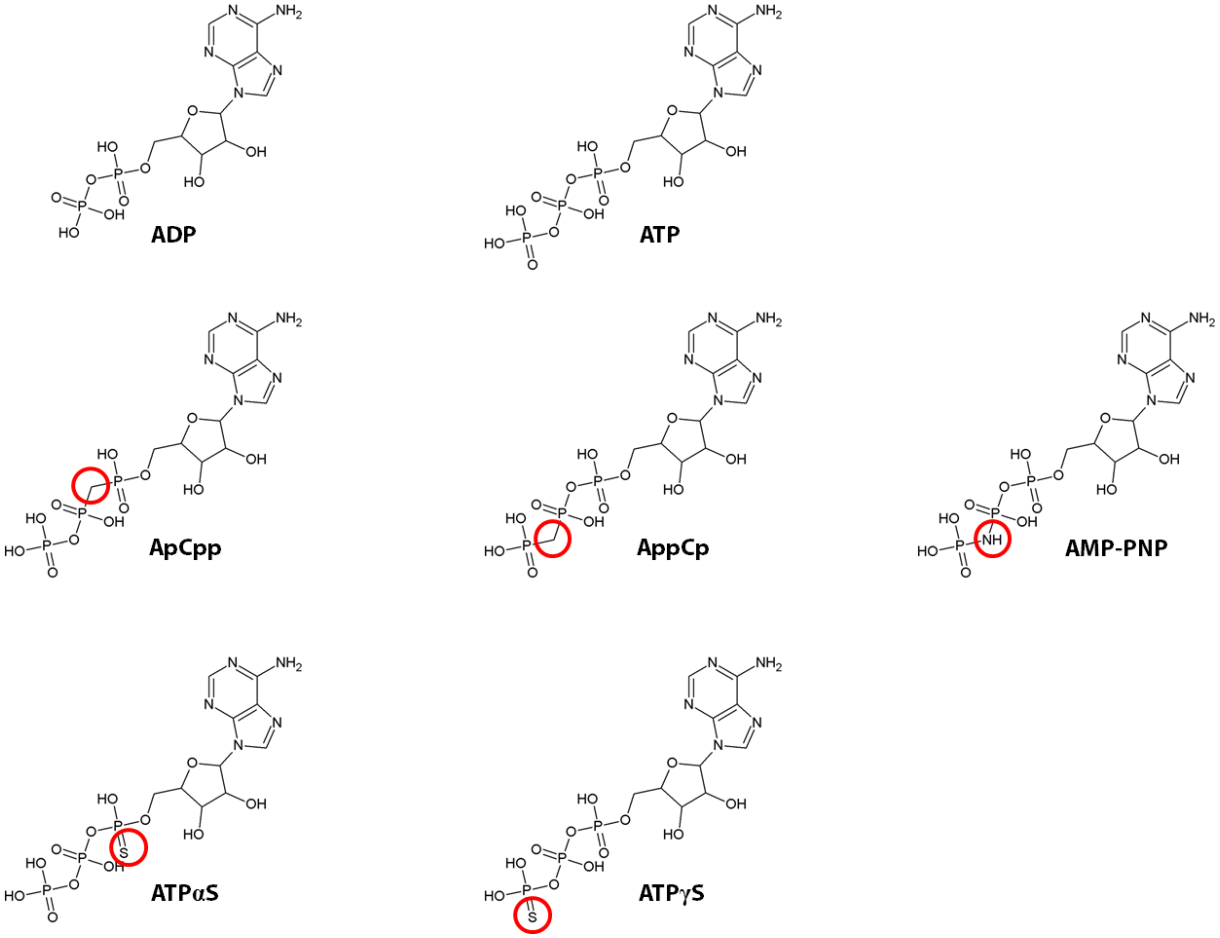


Figure 6.18 Chemical structures of the nucleotides and ATP analogs used to investigate Hsp70 conformations. The differences of the nucleotide analogs compared to ATP are marked by a red circle.

BiP

The conformational state of molecules upon hydrolysis of ATP is often probed using ADP+PO_4 or ADP+AlF_x. Interestingly, BiP was found in its ADP-bound conformation in the presence of these nucleotides (Figure 6.19A-B). BiP thus already adopts its ADP

conformation directly upon hydrolysis, when the inorganic phosphate may still be bound. A high phosphate concentration in the ER might additionally lead to an inactivation of the chaperone in the presence of kinases. BiP is inactivated due to phosphorylation of serine and threonine residues (Hendershot et al., 1988).

The nucleotide analog AppCp introduced a conformation similar to the apo state for the lid conformation (Figure 6.19A), while BiP in the presence of AMP-PNP, ATPαS, or ATPγS assumed the conformation observed for the ATP-bound state (Figure 6.19C). For BiP in the presence of ApCpp two subpopulations, sampling the ADP- as well as the ATP-conformation were found. This effect may be caused by a lower binding affinity of BiP for ApCpp which introduced a population as observed for the apo state, while ApCpp actually bound to BiP induced the ATP-bound conformation. This hypothesis could be verified by analyzing BiP in the presence of different ApCpp concentrations.

When comparing BiP's conformation in the presence of different nucleotide analogs, groups of atoms necessary for the induction of each of the well described conformations and allosteric coupling could be determined. The binding of AppCp, ApCpp, or ADP+AlF_x was not sufficient to induce the ATP conformation of BiP. Using the information of these measurements, we could show that, besides the γ-phosphate, the oxygen atoms between the α- and β-phosphate, as well as between the β- and γ-phosphate could not be replaced by carbon moieties for adaptation of the ATP conformation and for the allosteric communication of the information to the lid of the SBD. The replacement of the oxygen between β- and γ-phosphate by a nitrogen atom (AMP-PNP) or of the oxygen side groups in the α- and γ-phosphate (ATPαS, ATPγS), on the other hand, did not inhibit the formation of the ATP-conformation of BiP. Variable side chains in these positions can thus still induce BiP molecules in the ATP state.

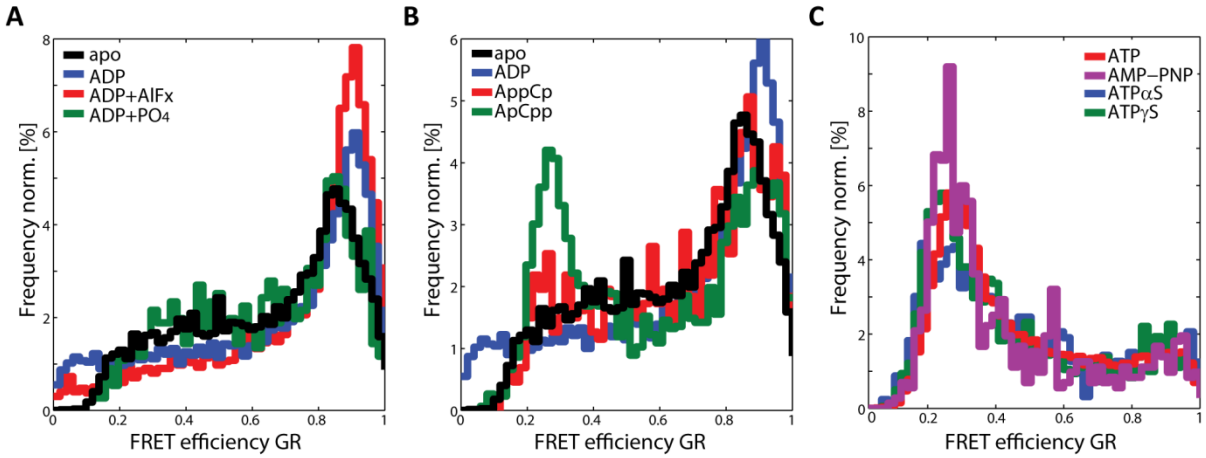


Figure 6.19 FRET efficiency histograms of triple-labeled BiP in the presence of a variety of nucleotide analogs obtained by three color PIE-MFD experiments. (**A**) BiP in its nucleotide-free form (black) and in the presence of ADP (blue), ADP+AlFx (red), or ADP+PO_4 (green). (**B**) BiP in the absence of nucleotide (black) or bound to ADP (blue), or to the ATP analogs AppCp (red) or ApCpp (green). (**C**) BiP bound to ATP (red), or to the ATP analogs AMP-PNP (magenta), ATPαS (blue), or ATPγS, in the presence of which it adopted the ATP-conformation.

The interdomain conformation of BiP as observed from the FRET efficiencies E_{BG} and E_{BR} displayed the same effect as found for the FRET efficiency E_{GR} of the lid conformation for

BiP in the presence of ADP+PO_4, ADP+AlF_x, AMP-PNP, ATPαS, and ATPγS. ApCpp and AppCp, on the other hand, no longer showed the clear correlation of the domains. ApCpp introduced an interdomain distance comparable to the nucleotide-free state of BiP, while the lid appeared as if partially found in the ADP-state (Figure 6.19B, Figure 6.20A). The FRET efficiency histogram of the lid conformation of BiP bound to AppCp was comparable to the one of nucleotide-free BiP, while the interdomain distance appeared similar to the one of ATP-bound BiP (Figure 6.19A, Figure 6.20B). This effect is especially notable when investigating the mechanism of the allosteric communication between the domains. In the two-dimensional FRET efficiency histograms displaying the interdomain distance with respect to the lid conformation, the anticorrelated behavior of the domains observed for the ADP-bound conformations of BiP was lost. Rather, an increased interdomain distance corresponded to an increasingly open lid state. The replaced moieties of ApCpp and AppCp are therefore highly important for conferring the allosteric communication between the domains.

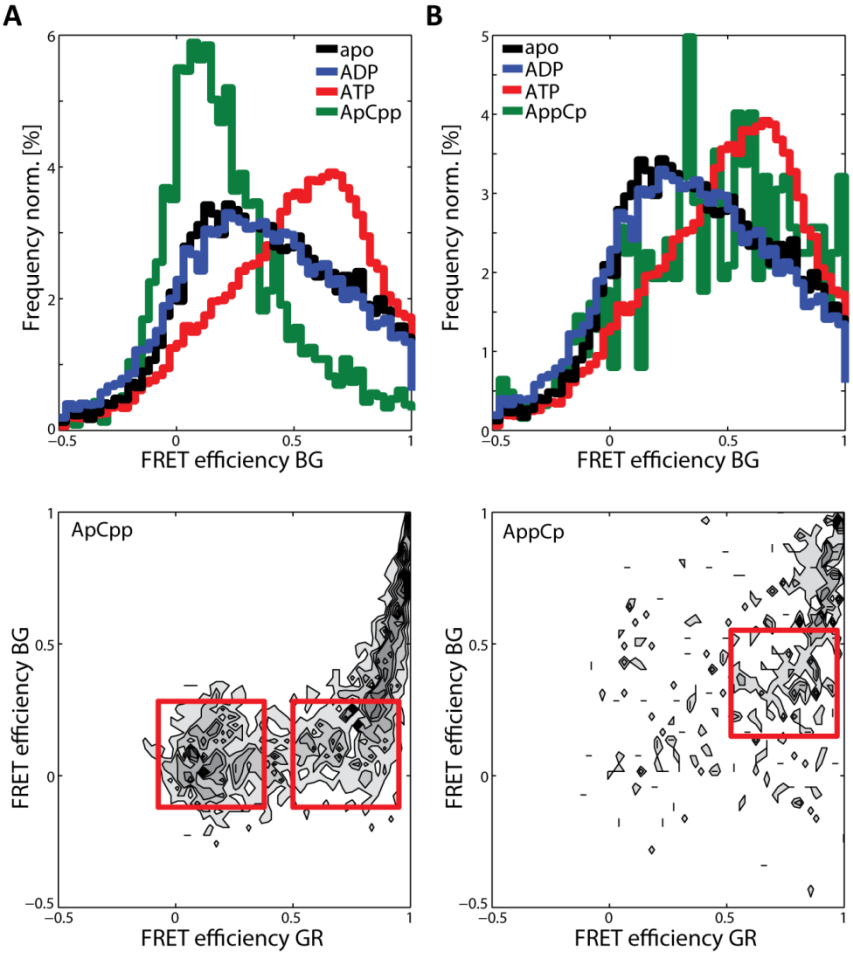


Figure 6.20 Interdomain distance of BiP investigated using the FRET efficiency histograms E_{BG} of triple-labeled protein (upper row) and two-dimensional FRET efficiency histograms E_{GR} versus E_{BG} (lower row). The one-dimensional E_{BG} histograms show the comparison of the FRET efficiency distributions of ApCpp- (**A**) and AppCp-bound (**B**) BiP to BiP in its apo, ADP-, or ATP-bound state. The two-dimensional FRET efficiency histograms display ApCpp-bound BiP (A) or AppCp-bound BiP (B).

DnaK

The same investigation was performed analyzing the conformations of DnaK in the presence of the nucleotide analogs. Interestingly, the observed pattern here was different from the one determined for BiP. DnaK was found in a conformation with the same FRET efficiencies E_{GR} and E_{BG} as for the ADP-bound form in the presence of the ATP analogs AMP-PNP, ATPγS, ApCpp, and AppCp, as well as for ADP+AlF_x and ADP+PO_4 (Figure 6.21A-B). Only ATPαS was able to induce a lid conformation as observed for ATP-bound DnaK (Figure 6.21C). For the interdomain distance, again ATPαS was the only nucleotide analog, in whose presence DnaK resembled the ATP-bound conformation slightly shifted to lower FRET efficiency values (Figure 6.21C).

To adopt the ATP-bound state, DnaK has a high specificity for the individual moieties of the nucleotide. The selectivity is induced by binding of the NBD to the triphosphate of the nucleotide. Modifications of the atoms between the phosphate groups of the triphosphate or an alteration at the γ-phosphate inhibited a conformation as observed for ATP-bound DnaK. Only for a modification of the α-phosphate, an ATP-like conformation could be recovered, suggesting little interaction of this phosphate moiety with the NBD of DnaK.

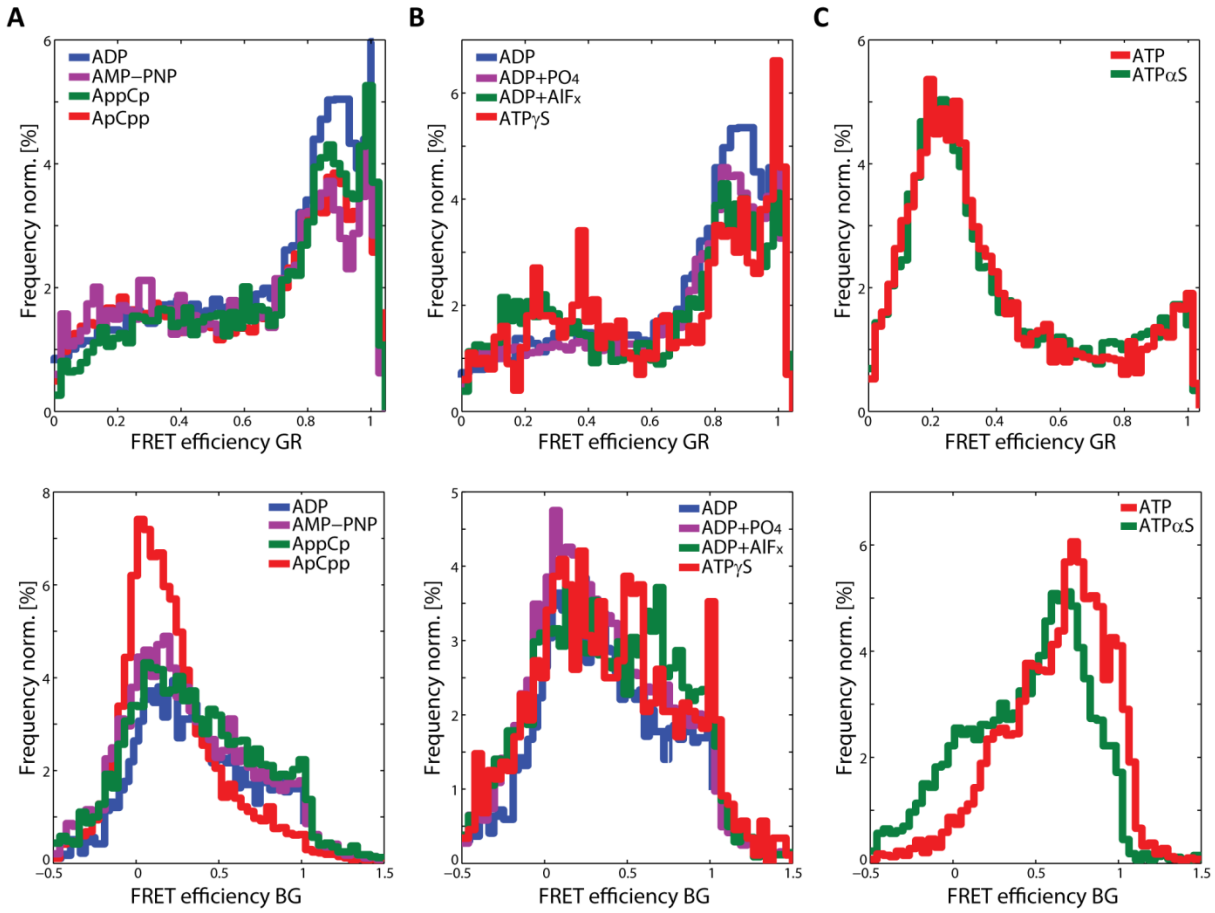


Figure 6.21 FRET efficiency distributions E_{GR} (upper row) and E_{BG} (lower row) of triple-labeled DnaK in the presence of nucleotide analogs compared to the native nucleotides ADP and ATP. (**A**) ADP-bound DnaK (blue) and DnaK in the presence of AMP-PNP (magenta), AppCp (green), or ApCpp (red). (**B**) ADP-bound DnaK (blue) and DnaK in the presence of ADP+PO_4 (magenta), ADP+AlF_x (green), or ATPγS (red). (**C**) DnaK bound to ATP (red), or ATPαS (green).

DnaK and BiP discriminate their conformation dependent on different atoms in the triphosphate of the nucleotide. DnaK showed a higher selectivity for the nucleotides than BiP. While BiP adopted a conformation similar to the ATP-bound state in the presence of the ATP analogs AMP-PNP, ATPαS, and ATPγS, DnaK only assumed this state when bound to ATPαS. The side-groups of the γ-phosphate seem therefore more important for DnaK to recognize its highly energetic state, than for BiP, where the oxygen connecting the β- and γ-phosphate could be replaced by a nitrogen and the double-bonded oxygen at the γ-phosphate could be a sulfur atom without losing its ATP-bound conformation. These experiments contribute a small hint at the recognition mechanism of nucleotides for the Hsp70 chaperones from different organelles. Although their nucleotide binding pocket is highly conserved, small differences in the nucleotides can still lead to variable conformational states for different Hsp70s. We observed that slight deviations of this recognition mechanism could be associated to the chaperones from the ER and from the prokaryotic cystosol. Variable recognition sites allow fine-tuning of the conformational cycle and can differ for the Hsp70 proteins in prokaryotes and eukaryotes. Whether these differences in the mechanism of recognition at different nucleotide positions plays an important role *in vivo*, where only ADP and ATP are present, has to be investigated further. With respect to the moieties important for the allosteric communication between the domains, it is highly interesting to note that some of the nucleotide analogs induced a different conformational state for the lid than they did for the interdomain distance independent of their functional role *in vivo*.

6.2.6 Mechanistic function of Hsp70 inhibitors

Due to their fundamental role in protein folding and transport in all cellular compartments, a malfunction of Hsp70 chaperones can lead to severe diseases. Hsp70 is associated with protein folding disorders such as Corea Huntington or Alzheimer's disease as well as many types of cancer (Broadley and Hartl, 2009; Murphy, 2013). In many cancers, Hsp70 chaperones prevent intrinsic and extrinsic apoptosis and cell senescence and are involved in proliferation, invasion, and differentiation of cancer cells (Alexiou et al., 2013, 2014; Lianos et al., 2015; Murphy, 2013; Yoshidomi et al., 2014). Eight human Hsp70 members are known to engage in cancer initiation and progression (Lianos et al., 2015). Many of these are targeted by Hsp70 inhibitors, over a dozen of which are known. It is still under debate whether targeting of the heat-inducible isoform of Hsp70 is sufficient for anti-tumoral activity, because in several studies dual-targeting of Hsc70 and Hsp70 was necessary (Schlecht et al., 2013). Therefore, it is important to investigate the specificity of the inhibitors, since Hsp70 chaperones are highly conserved and an inhibition of constitutively expressed Hsp70 should be avoided. The highest conservation is found for the nucleotide binding domain of these chaperones, which is targeted by some of the known inhibitors, such as VER155008 or MKT-007 (Murphy, 2013). Others bind to the less conserved substrate binding domain and inhibit the attachment of hydrophobic peptides. The here investigated Hsp70 inhibitors include two SBD-binding small molecules, geranylgeranylacetone and gentamicin sulfate, and the NBD-binding VER155008. Geranylgeranylacetone and gentamicin inhibit the binding of substrate peptides to the Hsp70s in a competitive manner. Thus they hinder the chaperone activity in

protein folding and the prevention of aggregation (Otaka et al., 2007; Yamamoto et al., 2010). Since gentamicin is not only found in the cytosol upon administration, but also resides in the nucleus, mitochondria, and the endoplasmic reticulum, an analysis of the effect of these inhibitors on chaperones from different compartments is important. VER155008 was shown to inhibit tumor cell growth by interacting with the nucleotide binding pocket of Hsp70 in competition with ATP and ADP (Williamson et al., 2009). It arrests the NBD in a half-open conformation and hinders the formation of the ATP-bound low affinity conformation of Hsp70. Using three-color FRET, this allosteric inhibition of the Hsp70 chaperone by VER155008 can be studied and can give a detailed mechanistic understanding of the functional mechanism of this inhibitor, which decreases the substrate binding and dissociation rates (Schlecht et al., 2013; Williamson et al., 2009).
In addition to these cancer cell-targeting Hsp70 inhibitors, inhibitors of the prokaryotic DnaK are used as antibacterial drugs (Liebscher et al., 2007). Hydrophobic straight carbon chains, similar to the ones for geranylgeranylacetone, are used in these benzamido antibacterial molecules to target the substrate binding pocket of the Hsp70.

Using three-color FRET, the binding of three Hsp70 inhibitors to BiP and DnaK and the conformational change induced in these chaperones was monitored. All three inhibitors were found to bind to the ER chaperone BiP (Figure 6.22). They are thus not selective for the cytosolic Hsc70 and Hsp70. Although gentamicin and geranylgeranylacetone are known to bind to the SBD, we observed different conformational states of the interdomain distance as well as the lid. When bound to gentamicin, a small interdomain distance and a closed lid were observed at the same time (Figure 6.22A). The lid conformation resembled the one for nucleotide-free BiP, while the fluorophores monitoring the interdomain distance were in even closer proximity to the conformation observed for ATP-bound BiP. A competition of ATP with gentamicin could shift the conformations towards the ATP-bound state. For geranlygeranylacetone, the opposite effect was observed. The interdomain as well as the lid conformation showed large distances (Figure 6.22B). In the presence of geranylgeranylacetone, the interdomain conformation observed by FRET rather resembled the nucleotide-free state, while the lid adopted a conformation similar to ATP-bound BiP. In both cases, the proximity of the domains did not correspond to the lid conformation expected from the natural ADP/ATP conformational cycle of BiP. The inhibitors therefore not only act as competitors for substrate binding to BiP, but may also uncouple the interdomain conformation from the SBD state. BiP in the presence of VER155008 adopted a conformation undistinguishable from ATP-bound BiP. The open lid was clearly visible by the FRET efficiency histogram E_{GR} of BiP in the presence of VER155008 (Figure 6.22C). It is known that VER155008 binds to the NBD and thus competes with ADP and ATP attachment. We could show that even though VER155008 binds to the NBD, it transfers the conformational information of this domain to the SBD in a similar fashion than ATP. By inhibiting ATP and ADP binding and thus ATP hydrolysis, the Hsp70 is therefore kept in a low substrate affinity

state.

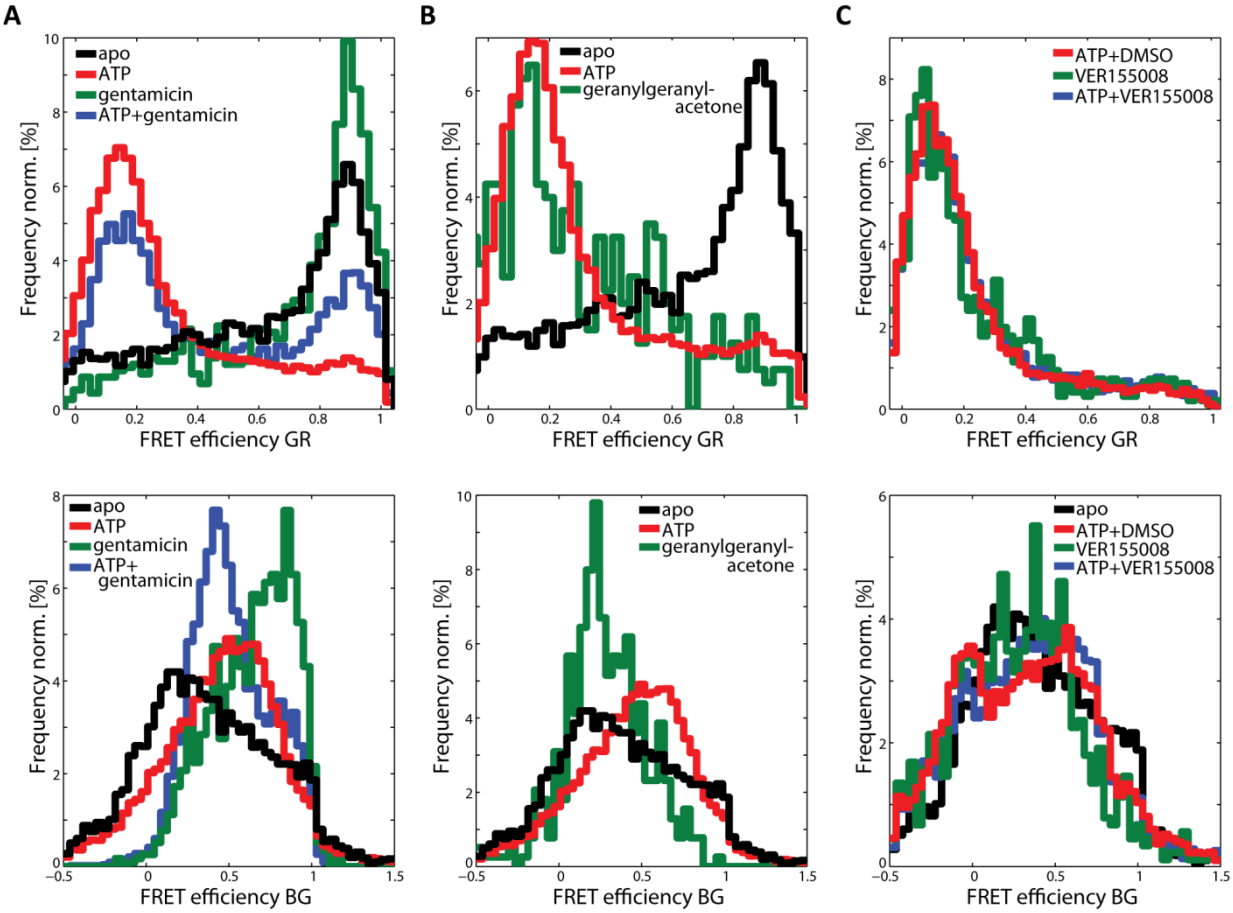


Figure 6.22 FRET efficiency histograms E_{GR} (upper row) and E_{BG} (lower row) of BiP in the presence of Hsp70 inhibitors gentamicin sulfate at a concentration of 3 mM (**A**), 300 µM geranylgeranylacetone (**B**), or 100 µM VER155008 (**C**) (green). The competition of 10 µM VER155008 and 1 mM ATP or 3 mM gentamicin sulfate and 1 mM ATP (blue) was tested. The inhibitor-bound BiP was compared to BiP in its apo state (black) or in the presence of ATP (red).

Even though DnaK is of prokaryotic origin and does therefore not act as a potential cancer target, the effect of the same inhibitors as for BiP were analyzed here. As general Hsp70 inhibitors, they might play a role as antibacterial molecules.

DnaK in the presence of gentamicin adopted a conformation similar to the ADP-bound state of DnaK with a closed lid and two interdomain conformations with a large and short distance (Figure 6.23A). DnaK in the presence of gentamicin showed an overall similar behavior to gentamicin-bound BiP. Similarly to gentamicin, geranylgeranylacetone acts as a competitive inhibitor of substrate binding by adopting a conformation resembling the ADP-bound state of DnaK (Figure 6.23B). The NBD-binding VER155008 induced a lid conformation as observed for ATP-bound DnaK and the interdomain distance was large (Figure 6.23C). The binding to the NBD thus induced a DnaK conformation, for which the coupling of the two domains was different from the natural conformational cycle of DnaK. Even with a large interdomain distance, an open lid conformation was adopted. Due to this stabilization of the Hsp70 conformation in an ATP-like state, VER155008 inhibits the exchange for the nucleotides ADP and ATP and possibly substrate binding to the SBD. To determine its effect on substrate

binding by DnaK, the affinity for substrate peptides could be measured in the presence of different VER155008 concentrations.

Binding of all three investigated human Hsp70 inhibitors to the prokaryotic DnaK was observed here. These inhibitors could therefore be used as antibacterial therapeutics. Due to the high sequence conservation, however, side effects on human Hsp70s may occur.

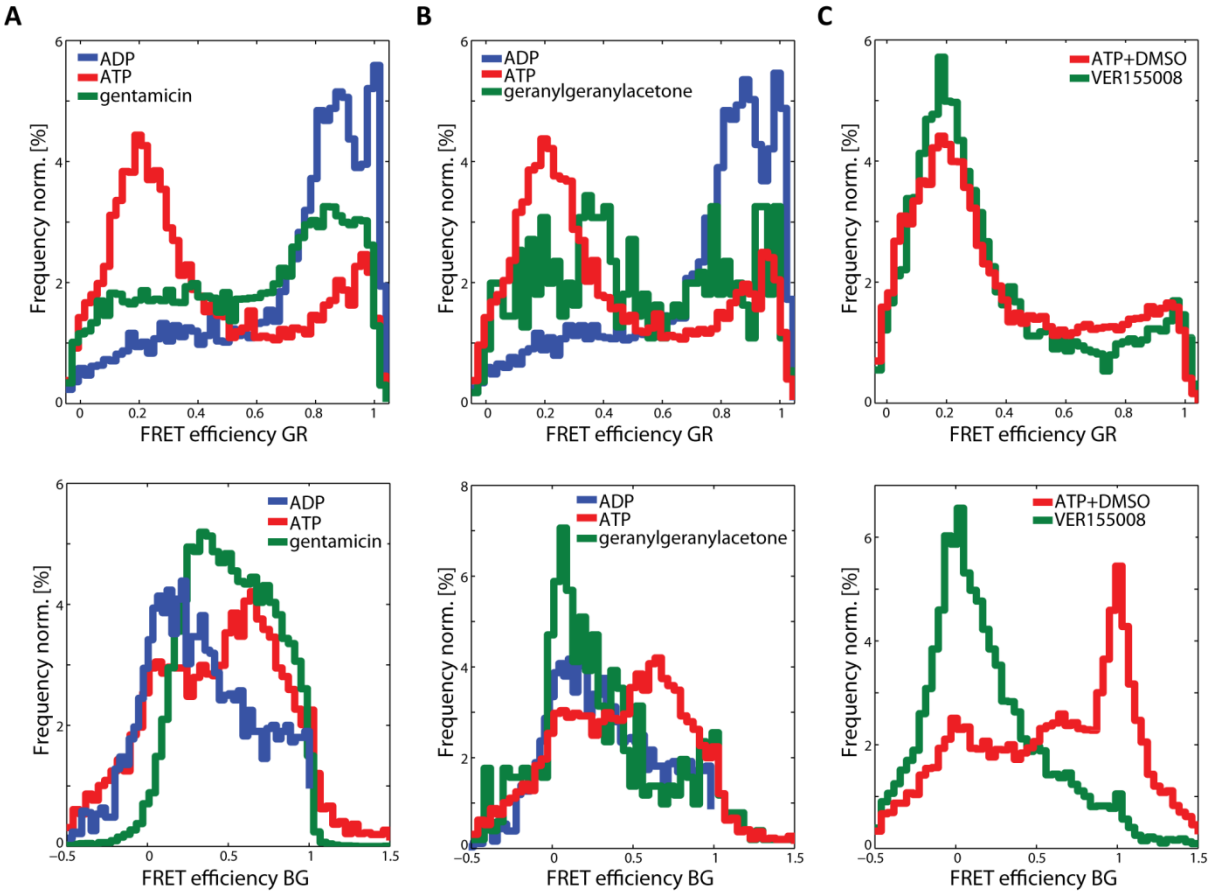


Figure 6.23 FRET efficiency distributions E_{GR} (upper row) and E_{BG} (lower row) of DnaK in the presence of Hsp70 inhibitors. The presence of 3 mM gentamicin sulfate (**A**), 300 µM geranylgeranylacetone (**B**), or 100 µM VER155008 (**C**) (green) was compared to DnaK bound to ADP (blue), or ATP (red).

Interestingly, the conformations obtained for the different Hsp70 inhibitors were similar for DnaK and BiP, while they largely differed between the inhibitors themselves. The SBD-binding inhibitors gentamicin and geranylgeranylacetone rather induced Hsp70 conformations with a large domain separation and a closed lid, while VER155008 binding to the NBD conferred a large domain distance and an open lid. For all inhibitors, the Hsp70 domain separation was rather large, suggesting an uncoupling of the domains in the presence of these inhibitors.

These experiments provide a first glimpse on the mechanistic function of the inhibitors regarding the chaperone's conformation. Further detailed investigations are needed to analyze competitive effects with nucleotides or substrate peptide. Furthermore, other Hsp70 inhibitors investigated in clinical trials, such as the mitochondrial Hsp70 targeting MKT-077, would be interesting to examine. Using three-color FRET, one could directly observe the allosteric communication between the domains for a variety of Hsp70 chaperone inhibitors.

6.3 Discussion

Hsp70 chaperones are known to undergo a conformational cycle, which is nucleotide dependent. Binding of a nucleotide to the NBD hereby induces a conformational change not only in the NBD itself, but the binding is allosterically communicated via a linker to the SBD. There it introduces large scale conformational rearrangements in the configuration of the lid with respect to the β-sheet of the SBD. It was previously shown that, in the ADP-bound conformation, the domains move more independently from each other with a larger interdomain separation. The domains only transiently contact each other in this state. The intensities of these contacts, however, seem to depend on the specific Hsp70. For DnaK, only transient contacts were described (Bertelsen et al., 2009; Chang et al., 2008; Swain et al., 2007), while in Hsc70 from *Bos primigenius taurus* both domains contact each other to a high degree (Behnke et al., 2015; Jiang et al., 2005). A similar observation was made by Sikor et al. (2013), who found independent motion of the interdomain conformation compared to the SBD state of the mitochondrial Ssc1 in the presence of ADP. Here, we used three-color FRET of single Hsp70 molecules to investigate the coordination of motion between their domains and therefore further elucidate the mechanism of allostery in these chaperones. To easily compare to previous results, the Hsp70 chaperones and their labeling positions were chosen as the ones used for spFRET experiments (Kudryavtsev et al., 2012; Mapa et al., 2010; Marcinowski et al., 2011). Because of their diverse functions in different organelles, we chose to compare Hsp70s from two eukaryotic organelles to the prokaryotic major Hsp70 DnaK. This approach is of evolutionary interest because the eukaryotic organelles stem from early prokaryotes. By an endosymbiotic event of an early eukaryotic cell taking up an α-proteobacteria, mitochondria evolved 1.5 billion years ago (Embley and Martin, 2006; Gray et al., 2001). Similarily, this ancestor then took up a cyanobacterium which became the chloroplast (McFadden, 2001; Raven and Allen, 2003). Hsp70 chaperones show a higher sequence similarity between all different organisms within one organelle than for the different Hsp70s within one species. The use of Hsp70s from three different compartments, even though from different organisms, therefore reveals information about the general functional mechanism of the Hsp70s in one specific organelle.

The three investigated chaperones, the prokaryotic DnaK, mitochondrial Ssc1, and BiP from the endoplasmic reticulum, share certain similarities in their conformational cycle. The conformational states of the Hsp70 chaperones depend on their bound nucleotides and other interaction partners. The majority population of all three Hsp70s, DnaK, Ssc1, and BiP, in their ATP-bound form had a short interdomain distance and an open lid. This ATP-bound conformation with a high communication between the NBD and the SBD and the thus induced open α-helical lid is therefore highly conserved not only throughout different phyla, but also between the Hsp70s of different cellular compartments. The conservation of this conformation points to the immense importance of the interdomain communication in this state for the functional mechanism of these chaperones. In addition, their function is strongly influenced by their interaction with cochaperones. Interestingly, the crystal structure of ATP-bound BiP was recently solved (Yang et al., 2015). It showed an overall highly similar

structure to ATP-bound DnaK (Kityk et al., 2012). Some differences, important for the allosteric communication between the domains, were however observed. First of all, BiP contains a larger hydrophobic interface between its lid and the NBD. Secondly, some residues were found to be conserved in cytosolic and ER-associated Hsp70s only, such as a conserved arginine (Arg532), which is necessary for the allosteric coupling in the ATP-bound state. Yang et al. (2015) discovered a high similarity between prokaryotic and mitochondrial or cytosolic and ER-associated Hsp70s with a lower conservation between these two groups of chaperones. These minor differences were not observed in the here performed three-color FRET studies on ATP-bound Hsp70s. We did find differences between the investigated chaperones, however, when comparing their conformational behavior in the presence of ADP. While all three chaperones generally adopted a conformation with a fully closed lid and a large interdomain distance in the presence of ADP, a second population with a large interdomain separation as well as open lid conformation was found for ADP-bound Ssc1. These results further expand on previous observations, in which only transient domain contacts were described for some ADP-bound Hsp70s such as DnaK (Bertelsen et al., 2009; Chang et al., 2008; Swain et al., 2007), while others interacted more strongly between their domains. The results from one-dimensional FRET efficiency histograms already displayed differences in the ADP-bound conformation of the different Hsp70s; their coordinated behavior was further investigated in more detail using two-dimensional FRET efficiency and distance distribution histograms. BiP and DnaK showed a more coordinated, anticorrelated behavior of the interdomain and lid conformational state in the ADP-bound form than Ssc1. As observed by Yang et al. (2015), mitochondrial and prokaryotic Hsp70s are more closely related to eachother than to ER-associated BiP. Similar to DnaK, for which only little interdomain dockings were described in the ADP-bound state, Ssc1 might undergo very few docking events between its domains in this state. The lid can thus act independently from the NBD in this conformation. This independent motion of the individual Ssc1 domains was visualized in spFRET experiments of surface attached protein by Sikor et al. (2013). The mitochondrial Ssc1 is most closely related to plastid Hsp70s (Appendix). In a next step, we will investigate the interdomain communication of one of the plastid Hsp70 isoforms in a collaborative project with Dr. Serena Schwenkert (Ludwig-Maximilians-Universität München, Germany).

The coordination of the domains is necessary for the allosteric regulation of the Hsp70 protein by the nucleotide. Extensive studies have been performed investigating the allosteric coupling of *E. coli* DnaK using crystal structures, NMR, and molecular dynamics simulations (Bhattacharya et al., 2009; Nicolaï et al., 2013; Zhuravleva and Gierasch, 2011, 2015; Zhuravleva et al., 2012). Several main features for the coupling behavior were observed: The SBD and NBD domains collide repeatedly in the ADP-bound form, but do not dock (Bertelsen et al., 2009). The binding of ATP leads to stable docking. This docking is induced by rotations of the subdomains IIB and IIA with respect to IA in the NBD (Bhattacharya et al., 2009). The β1-β2 turn in subdomain IIA recognizes the γ-phosphate of ATP and subsequent rotation of the domain IIA and IA lead to the exposure of a hydrophobic surface,

which in turn acts as a binding surface for the interdomain linker (Bhattacharya et al., 2009; Zhuravleva and Gierasch, 2011). This stabilizing connection in the ATP-bound state is further supported by an interaction of helix B of the α-helical lid with the NBD (Zhuravleva et al., 2012). Furthermore, increased conformational dynamics of the β-sheet of the SBD were observed in the presence of ATP (Zhuravleva and Gierasch, 2015).
It is known that ATP is contacted by all four subunits of the NBD and that the γ-phosphate is bound by the domain IIA. Here, we used different non-hydrolyzable ATP analogs to investigate the effect of different nucleotide moieties on the allosteric conformational change of the prokaryotic DnaK compared to the eukaryotic BiP, which according to Yang et al. (2015) represent examples for each of the large groups of Hsp70s. We found DnaK to be highly selective in its conformational change upon nucleotide binding. The only ATP analog inducing a conformation similar to the ATP-bound state, was ATPαS, which differs from its natural partner in the replacement of a single oxygen at the α-phosphate by a sulfur atom. This atom is either not recognized by the NBD of DnaK or its contact site is not important for the induction of the ATP active state. BiP shows a lower selectivity for the formation of the ATP-state. The recognition of different, and possibly less moieties, of the nucleotide by the NBD of BiP is therefore hypothesized. Interestingly, the binding of some ATP analogs induced an allosteric decoupling of the domains, which was observed for BiP bound to ApCpp or AppCp. The replaced moieties in the nucleotides therefore seem of high importance for the induction of the allosteric coupling by rotation of the NBD subdomains and further interdomain communication by the linker and the helical lid.

Besides these differences in the selectivity and induction of its conformational states, biochemical studies showed differences in the substrate binding kinetics between prokaryotic and eukaryotic Hsp70s (Mapa et al., 2010; Marcinowski et al., 2013). We investigated DnaK, Ssc1, and BiP in the presence of artificial substrate peptide and their nucleotide exchange factors. The binding of peptide introduced conformational changes in DnaK and Ssc1 to a conformation with a large interdomain distance and closed lid. It is interesting to note that substrate binding can also induce conformational changes in the interdomain communication. This agrees with previous descriptions of bidirectional allosteric signal propagation in these chaperones (Zhuravleva and Gierasch, 2011). The FRET efficiency distributions of BiP in the presence of peptide showed little conformational change, although a slight stabilization of the interdomain distance was observed in two-dimensional distance distributions. The here observed differences in the conformational states of peptide-bound Hsp70s might describe the underlying conformational effects to the differences of the observed substrate binding kinetics. A necessary conformational change of the Hsp70 can strongly influence the association or dissociation behavior of the substrate.
Interestingly, the binding of Mge1 to Ssc1 introduced a conformation similar to the ATP-bound state of Ssc1, while GrpE rather introduced the ADP-bound state with a large interdomain distance and a closed lid in DnaK. The effect observed for GrpE-bound DnaK agrees with previous reports of DnaK-GrpE structures with a closed lid and large domain separation (Bhattacharya et al., 2009; Wu et al., 2012a). Mge1, on the other hand, uses a

different mode to increase the exchange of ADP for ATP. It seems to induce the ATP conformation of Ssc1 to increase the rate of ATP binding to the chaperone.

In a last step, we investigated the effect of Hsp70 inhibitors on the two Hsp70s DnaK and BiP from both major classes of Hsp70s. We found the inhibitors gentamicin, geranylgeranylacetone, and VER155008 to not only target cytosolic Hsp70s, but also the prokaryotic DnaK and the ER-associated BiP. These inhibitors therefore do not act specifically on the stress induced Hsc70 and constitutively expressed Hsp70 from the cytosol. This observation is interesting for anticancer treatment, because the unspecific interaction with Hsp70s from different compartments may cause side effects, when the inhibitor is targeted also to these compartments. Furthermore, the interaction with DnaK offers the possibility to use these Hsp70 inhibitors as antibacterial agents keeping in mind their possible side effects on eukaryotic Hsp70s.
The binding of inhibitors to the Hsp70 chaperones is additionally of interest, when considering their allosteric coupling behavior. Some of the nucleotide inhibitors induce conformational states in which the interdomain separation is uncoupled from the lid conformation, as was observed for example for the VER155008-bound DnaK. Not only the competitive binding of the inhibitors to either the substrate binding domain or to the nucleotide binding domain is therefore important for its inhibitory effect, but also the functional conformational cycle of the Hsp70s can be disrupted by their binding.

The application of single molecule three-color FRET revealed high similarities between Hsp70 chaperones from different cellular compartments in their ATP-bound conformation. This active state is highly conserved from prokaryotes to humans. Differences in the coupling of the domains in the low energetic state, in their conformation upon peptide binding, and in their selectivity for nucleotides can be important for the different functionalities of the Hsp70s in their compartments. Interestingly, the sphere of action of Hsp70s is largely driven by their J-domain containing interaction partners. In many organelles, only a single Hsp70 is present, which can interact with multiple different Hsp40s. Mitochondria of higher eukaryotes, for example, contain a single Hsp70 and four different Hsp40s. Similarly, a single Hsp70 in the endoplasmic reticulum is known, which can interact with six different Hsp40 homologs. Further analyses of the Hsp70 chaperones could therefore include the comparison of their conformational states and dynamics in the presence of different Hsp40s.

7 First glimpse on the conformational states of σ^{32} and their changes upon Hsp70 chaperone binding

Gene expression is highly regulated in any eukaryotic or prokaryotic cell and adapts in response to different cellular conditions. The cell adjusts to elevated temperatures by the upregulation of certain heat shock induced genes. σ^{32} is one of these heat shock proteins, which is upregulated in response to an increase in temperature. σ^{32} in itself is a transcription factor involved in the regulation of protein expression under stress. This 32 kDa sized protein is found in the prokaryotic cytosol, where DnaK and RNA polymerase can bind to it competitively depending on the cellular conditions. The conformational state of σ^{32} in the absence and presence of its binding partners was investigated using spFRET. The protein was found in a condensed conformation in its free form, whereas the binding of DnaK led to an increase in the distance between both termini of the protein. The conformation of σ^{32} is, most likely, highly important for the initiation of gene transcription.
The project was performed in collaboration with Professor Matthias Mayer, Roman Kityk (Ruprecht-Karls-Universität Heidelberg, Germany), and Ganesh Agam (Ludwig-Maximilians-Universität München, Germany).

7.1 Introduction

The expression level of a variety of proteins largely changes in response to cellular stress such as a shift in temperature. Large scale changes in the expression level are often induced by a regulation of transcription. RNA polymerase, responsible for gene transcription, interacts with multiple σ factors, which confer specificity for certain groups of promoters. In *E. coli* seven different σ factors, σ^{70}, σ^{54}, σ^{32}, σ^{F}, σ^{S}, σ^{E}, and σ^{fecI} are known (Zhao et al., 2005). Some of these σ factors are responsible for the upregulation of certain genes upon heat shock. σ^{70} is the major factor responsible for gene expression under normal conditions, while σ^{32} causes the transcription of certain genes in response to elevated temperatures or other stress conditions.

7.1.1 The transcription factor σ^{32}

To initiate transcription of genes upon heat shock, σ^{32} forms a complex with RNA polymerase. Under normal conditions, this complex formation is hindered by the binding of DnaK/DnaJ to native σ^{32}. In this case, the transcription factor has a rapid turnover and short half-life of below 1 min. The degradation of σ^{32} is promoted by the binding of DnaK/DnaJ to the N-terminal domain, inducing a conformational change in σ^{32} by which the ATP-dependent protease FtsH can start degradation from the N-terminus (Straus et al., 1990). σ^{32} is, however, stabilized at elevated temperatures, when the Hsp70 chaperone system is involved in protein folding and releases the transcription factor. It is then free to bind to RNA polymerase and DNA. The binding of RNA polymerase and DnaK occurs competitively with a K_D of 12 nM for the σ^{32}-RNA polymerase complex, 5 μM for σ^{32}-DnaK, and 19 nM for σ^{32}-DnaJ (Gamer et al., 1996; McCarty et al., 1996; Rodriguez et al., 2008). For a long time, the binding of DnaK to σ^{32} was believed to occur with a molar ratio of 1:1 (Gamer et al., 1996; Liberek and

Georgopoulos, 1993; Liberek et al., 1992, 1995; Straus et al., 1990). The binding sites of RNA polymerase and DnaK on the homology model of σ^{32} are displayed in Figure 7.1. The shown structures are energy minimized homology models of *E. coli* σ^{32} on the structures of anti-σ factor bound FliA from *Aquifex aeolicus* and σ^{A} from *Thermus thermophilus* bound to RNA polymerase (Rodriguez et al., 2008; Sorenson and Darst, 2006; Sorenson et al., 2004; Vassylyev et al., 2002). Very recently, Noguchi et al. (2014) discovered DnaK-σ^{32} complexes of different molecular mass, suggesting the formation of complexes with up to three DnaK molecules. ADP-bound DnaK was shown to bind to the Leu-containing binding site of σ^{32} independent of DnaJ. Binding to the other two sites only occurs in the presence of DnaJ and ATP. We used our expertise on spFRET to investigate the conformational changes of σ^{32} in the absence and presence of the chaperone DnaK.

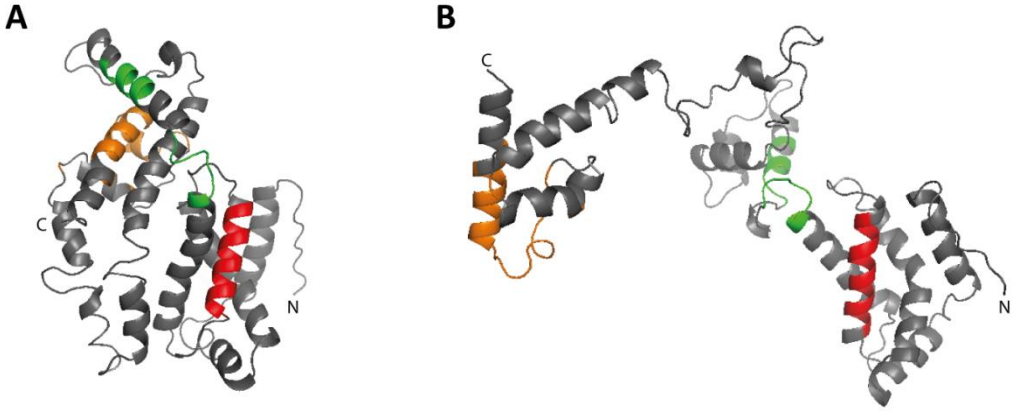


Figure 7.1 Energy-minimized homology models of σ^{32} obtained by modelling of the amino acid sequence of *E. coli* σ^{32} to the anti-σ factor bound FliA of *Aquifex aeolicus* (PDB: 1RP3) (**A**) or to the RNA polymerase-bound structure of σ^{A} from *Thermus thermophilus* (PDB: 1IW7) (**B**). The RNA polymerase-binding domain is shown in red, DNA-binding regions are displayed in orange, and the DnaK-binding domain is visualized in green.

7.2 Results

Homology modeling of σ^{32} on the structures of FliA bound to anti-σ factor and σ^{A} bound to RNA polymerase (Rodriguez et al., 2008; Sorenson and Darst, 2006; Sorenson et al., 2004; Vassylyev et al., 2002) suggests that σ^{32} populates different conformational states (Figure 7.1). The different conformations may well be associated with the activity state of the protein, since it is inactive and readily degradable when bound to DnaK, while it actively promotes transcription from certain promoters when bound to RNA polymerase and DNA. Therefore DnaK acts as the anti-σ factor of σ^{32}. To investigate potential conformational transitions between different states, four cysteine mutants of σ^{32} were designed. Large scale conformational rearrangements of the whole protein could be observed by placing cysteines close to both termini of σ^{32}. In mutant σ^{32} I, cysteines were site-specifically introduced at positions C9 and the C-terminus C284 (Figure 7.2A). For σ^{32} II, the sites were placed at C30 and C284 (Figure 7.2B). The difference between these two mutants thus carries information about the most N-terminal α-helix. To probe conformational rearrangements of individual domains, two additional mutants were analyzed containing cysteines at one of the termini and a position next to the center of σ^{32}, in each case respectively. σ^{32} III and IV were fluorescently

labeled at the cysteines C30 and C165 (Figure 7.2C) or C165 and C284 (Figure 7.2D) to probe the conformation of the N-terminal or C-terminal domain, respectively.
The four σ^{32} mutants were used to investigate the conformation of the protein in its free form or when bound to the Hsp70 chaperone DnaK. The binding of DnaK to σ^{32} was examined by FCS. SpFRET experiments of σ^{32} in solution were performed on a confocal microscope using PIE-MFD to gain detailed insights into the conformational changes of σ^{32}.

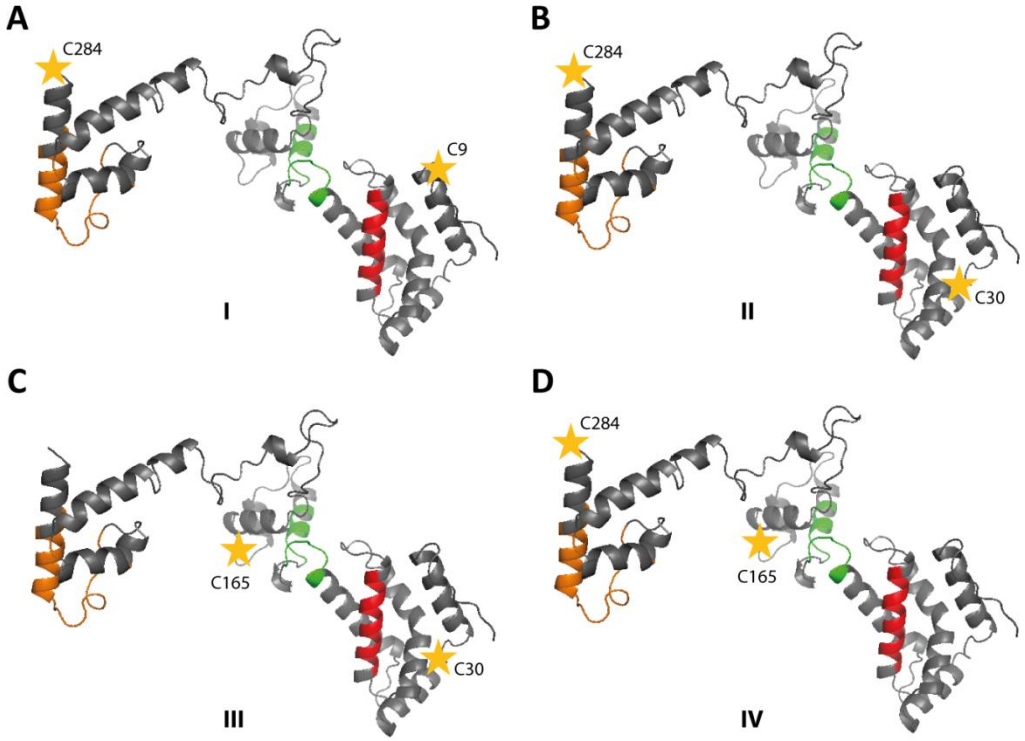


Figure 7.2 Sites of fluorophore attachment on the investigated σ^{32} mutants. (**A**) σ^{32} I contains cysteines at positions C9 and C284, which were labeled fluorescently with Atto532 and Alexa647. (**B**) σ^{32} II was labeled at positions C30 and C284. (**C**) For σ^{32} III the labeling sites were moved to C30 and C165 and (**D**) σ^{32} IV was fluorescently labeled at C165 and C284.

7.2.1 Binding of DnaK to σ^{32} observed by FCS

FCS was used to monitor the binding of DnaK to σ^{32}. The obtained diffusion coefficients agree with the expected values for a protein of this size. The binding of the 70 kDa large DnaK to σ^{32} increased the molecular weight of the complex by a factor of around 3. The expected changes in the diffusion time are small with a difference of around 1.4, but still detectable by FCS. An increase in the diffusion time with factors of 1.1-1.4 in the presence of DnaK for the four σ^{32} mutants was observed here (Figure 7.3). The increase occurred in the presence of both nucleotides, ADP and ATP, suggesting binding of DnaK to σ^{32} in both nucleotide states. We observed slight shifts in the diffusion times for the different constructs and nucleotide states. Multiple reasons for the differences are possible. A second population of free σ^{32} or different conformational states leading to different diffusion times could introduce a second term in the autocorrelation function. Furthermore, Noguchi et al. (2014) described three potential binding sites for DnaK. Differences in the stoichiometry of the complexes could affect the diffusion coeffient as well. Although these slight differences in the diffusion coefficient of the different complexes were observed, DnaK clearly binds to all four

complexes in the presence of both nucleotides. The observation of conformational changes of σ^{32} in the presence of DnaK can additionally support these results.

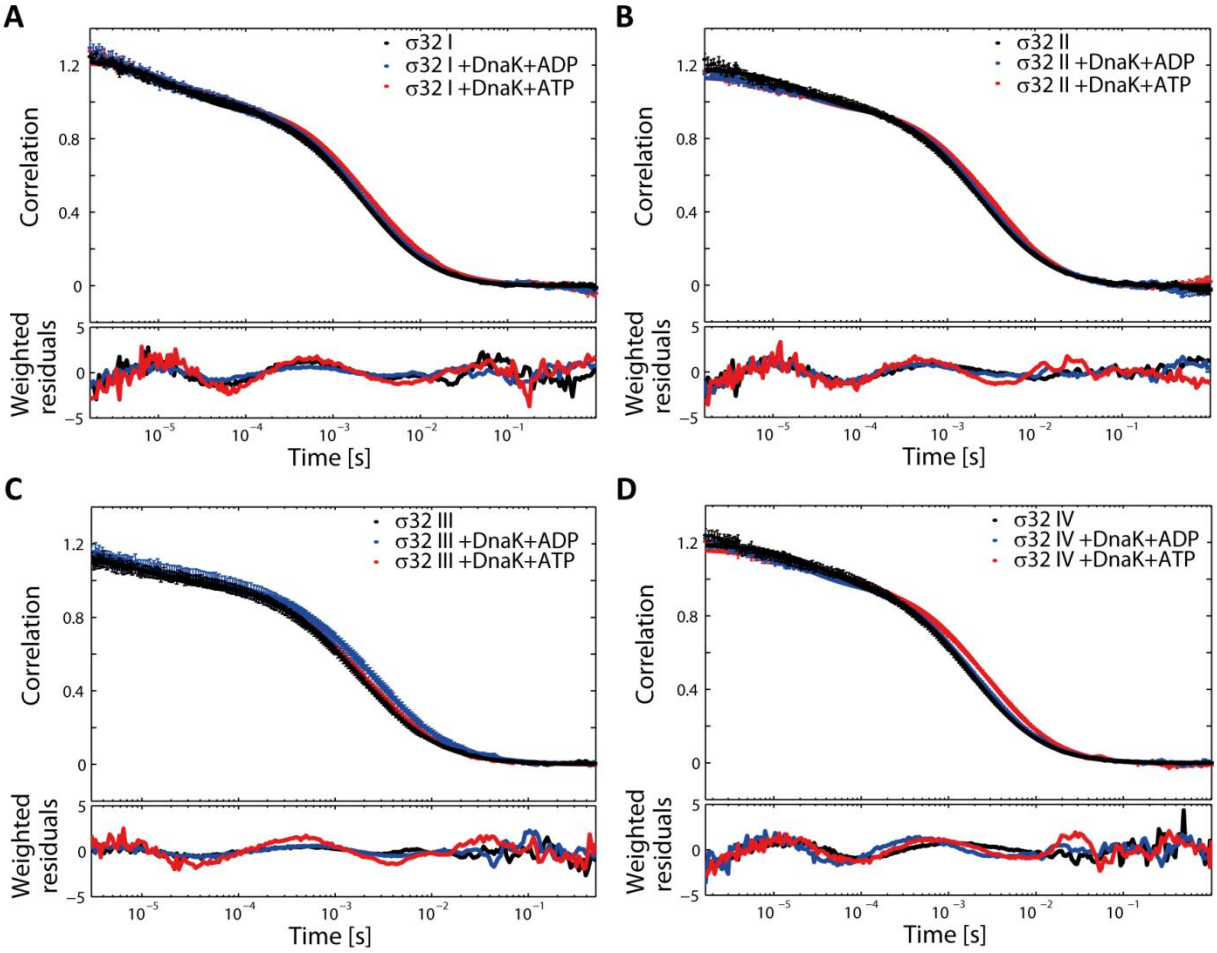


Figure 7.3 Fluorescence correlation spectroscopy of σ^{32} in the absence (black) and presence of 20 µM DnaK and 40 µM ADP (blue) or ATP (red). The correlation functions including a fit to a single species and the weighted residuals of the fits are shown for the mutants σ^{32} I (**A**), II (**B**), III (**C**), and IV (**D**).

7.2.2 Conformational states of σ^{32} and their changes upon binding of DnaK

Once the binding of DnaK to σ^{32} was confirmed, we investigated the effect of DnaK binding on the conformation of σ^{32}. Using the FRET efficiency information obtained by spFRET experiments with PIE-MFD, σ^{32} was found to exist in a compact conformation in its free form. The mean FRET efficiencies were 0.88 for σ^{32} I, 0.67 for σ^{32} II, and 0.92 and 0.87 for σ^{32} mutants III and IV (Table 7.1). All attachments sites of the fluorophores were thus found to be in close proximity with an average distance of around 46-48 Å for σ^{32} I, III, and IV or 55 Å for the major population of σ^{32} II. Interestingly, σ^{32} II had a slightly lower FRET efficiency, suggesting a closer proximity of the most N-terminal end of the first α-helix to the C-terminal domain than for position C30.

A conformational change of σ^{32} was observed for mutants I and II in the presence of DnaK-ADP or DnaK-ATP. At least two subpopulations of molecules were observed for σ^{32} in the presence of the chaperone. One of the populations corresponded well to the FRET efficiency obtained for free σ^{32} and may thus account for the unbound protein. σ^{32} I-DnaK complexes had a mean FRET efficiency of 0.15-0.19 (Figure 7.4A) and σ^{32} II-DnaK complexes showed a

FRET efficiency value of 0.06-0.07 (Figure 7.4B). These FRET efficiencies correspond to approximate distances of 65-85 Å for σ^{32} I or 72-79 Å for σ^{32} II (Table 7.1). For σ^{32} III and IV, the obtained FRET efficiencies did not describe any large scale conformational rearrangements between the attachment positions of the fluorophores when bound to DnaK (Figure 7.4C, D). Since binding of the chaperone to these mutants was ensured by FCS (Chapter 7.2.1), a failure of DnaK binding can be excluded. Therefore, the protein regions under investigation did not undergo any, by the FRET efficiency measurable, conformational changes. The binding of DnaK left these regions unaffected, suggesting a bending of the protein between its N- and C-terminal domains in the region of amino acid C165.

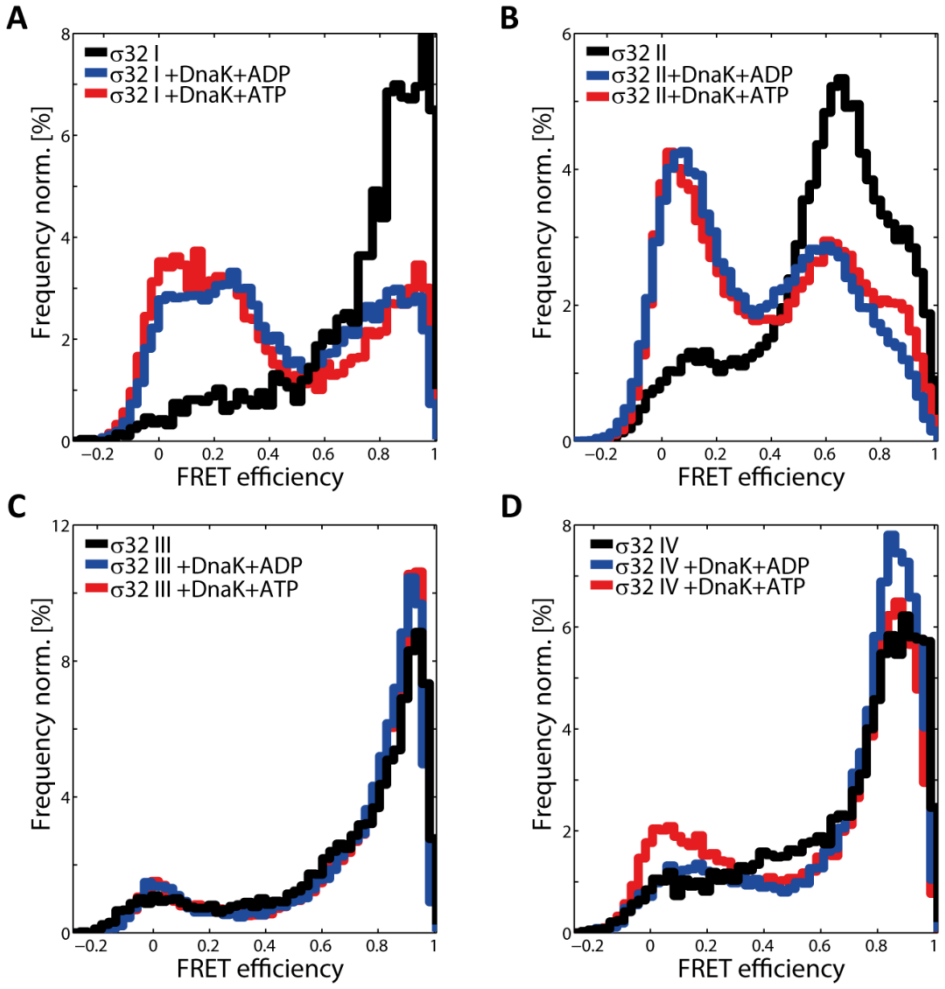


Figure 7.4 FRET efficiency distributions of σ^{32} obtained by PIE-MFD experiments. σ^{32} was measured in its free form (black), in the presence of 20 µM DnaK and 40 µM ADP (blue), or in the presence of 20 µM DnaK and 40 µM ATP (red). The graphs show σ^{32} mutants I (**A**), II (**B**), III (**C**), and IV (**D**).

We could show here, that large scale conformational rearrangements between the N- and C-terminal domain, as observed for σ^{32} I and II, take place upon binding of DnaK. The observation that for σ^{32} I and II conformational changes were visible upon DnaK binding, while they were not observed for σ^{32} IV agrees well with previous results of hydrogen-deuterium exchange experiments, which showed that residues 31-49 of σ^{32} are less compact in the presence of DnaK (Rodriguez et al., 2008). An observation of these effects would also be

expected for σ^{32} III. Small distance changes might, however, not be visible as a difference for the less sensitive high FRET efficiency state.

Table 7.1 Distances obtained by PDA fits of the spFRET experiments of σ^{32} constructs in the absence and presence of DnaK.

σ^{32}	$D_1(\sigma_1)$ [Å]	A_1 [%]	$D_2(\sigma_2)$ [Å]	A_2 [%]	$D_3(\sigma_3)$ [Å]	A_3 [%]
I	48(15)	87			112(4)	13
I+DnaK+ADP	47(6)	30	67(11)	60	115(7)	10
I+DnaK+ATP	47(6)	36	65(8)	30	85	33
II	55(5)	61	69(22)	20	44(5)	19
II+DnaK+ADP	55(8)	45	72(15)	40	128(14)	15
II+DnaK+ATP	55(9)	61	79(7)	18	119(7)	18
III	46(7)	79	68(7)	13	133(1)	8
III+DnaK+ADP	46(5)	85	67(1)	7	100(1)	8
III+DnaK+ATP	46(6)	87	67(1)	6	100(1)	7
IV	46(7)	60	64(8)	33	133(1)	7
IV+DnaK+ADP	48(5)	73	67(20)	27		
IV+DnaK+ATP	48(5)	65	70(21)	35		

We analyzed the conformational dynamics of σ^{32} in the absence and presence of DnaK using histograms of donor lifetime versus FRET efficiency (Figure 7.5). The high FRET populations of σ^{32} mutants III, and IV in their free form showed conformational dynamics on the timescale of micro- to milliseconds, while the population of σ^{32} II molecules found on the static FRET line suggests static σ^{32}. Interestingly, the conformational dynamics which occur within one of the domains were less well observed when analyzing distance changes of the whole protein. A comparison of σ^{32} I with σ^{32} II could suggest that the most N-terminal α-helix of the protein may be dynamic, the dynamics observed for mutants III and IV, however, cannot be explained by motion of this helix. Conformational dynamics occurring within either the N-terminal or C-terminal domain may not be observed with the single linear distance information obtained by spFRET between both termini when looking at mutants I and II.
When bound to DnaK, the low FRET populations of mutants I and II showed a static behavior, while the subpopulation of molecules with a high FRET efficiency was dynamic. DnaK thus seems to induce the conformational change to the low FRET conformation in which σ^{32} is then stabilized when bound to its anti-σ-factor. In the presence of DnaK, σ^{32} mutants III and IV still showed conformational dynamics in a histogram of donor lifetime versus FRET efficiency. A stabilization of these σ^{32} mutants by the binding of DnaK could not be observed here.

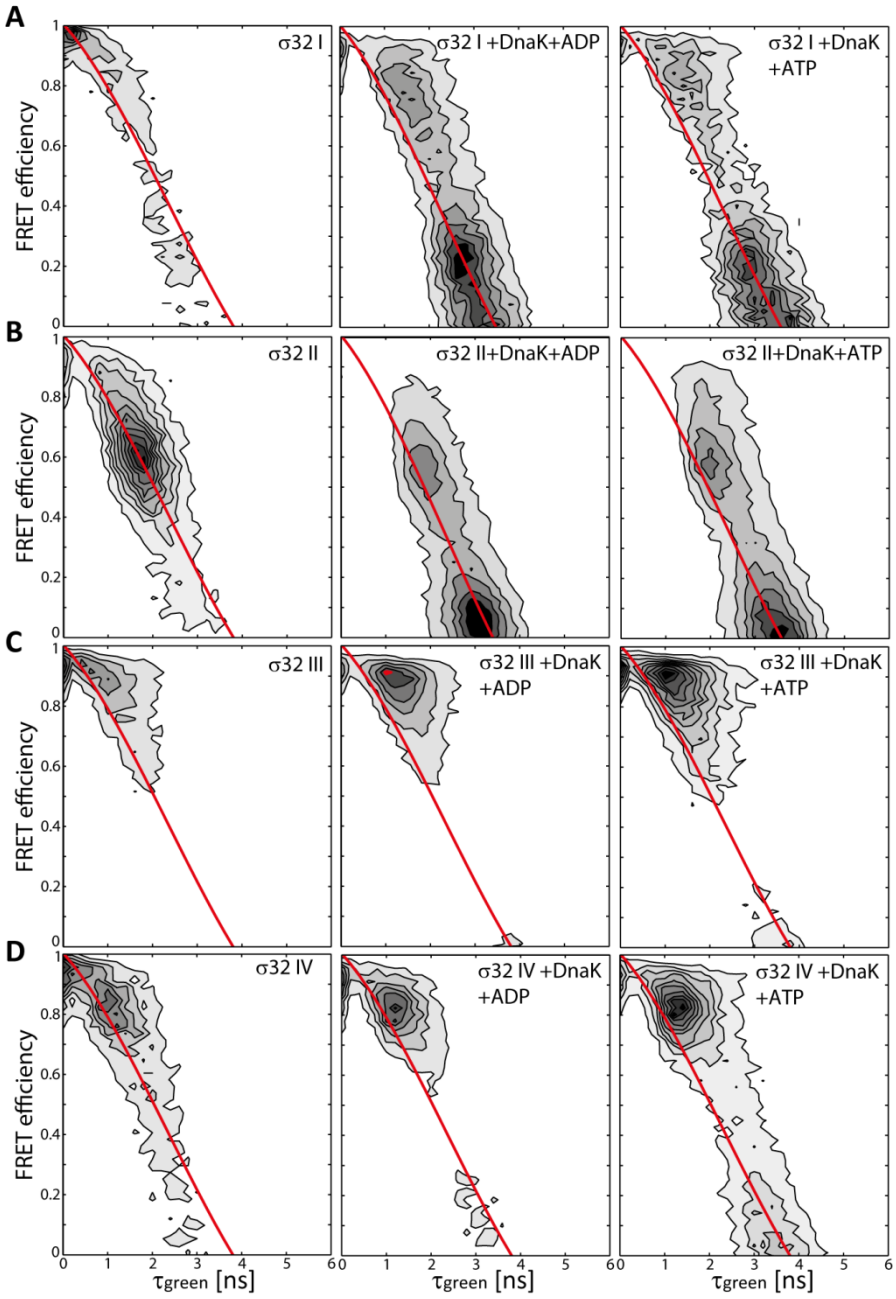


Figure 7.5 Conformational dynamics of σ^{32}. Histograms of donor lifetime versus FRET efficiency obtained from spFRET experiments with PIE-MFD of σ^{32} mutants I (**A**), II (**B**), III (**C**), and IV (**D**) in the absence (left) and presence of DnaK-ADP (middle) or DnaK-ATP (right).

7.3 Discussion and prospects

Previous studies on the crystal structures of σ^{32} homologs from different prokaryotes suggested a conformational change of σ^{32} dependent on binding of its interaction partner (Sorenson and Darst, 2006; Sorenson et al., 2004). When bound to the DnaK/DnaJ/GrpE chaperone system, σ^{32} is kept in an inactive form readily degradable by FtsH protease. The degradation from N- to C-terminus of σ^{32} is believed to occur due to an unravelling of the N-terminal residues 31-49 in the presence of DnaK (Rodriguez et al., 2008).

First spFRET experiments of σ^{32} in the absence and presence of DnaK revealed a large conformational change of σ^{32} between its N-terminal and C-terminal domain. In its free form,

σ^{32} appeared in a compact, dynamic conformation with an approximate distance between its termini of 48 Å and, in complex with DnaK, the distance between both termini was increased to around 67 Å. It is interesting to note that conformational changes were observed upon binding of DnaK to σ^{32} I and II, for which the distances of the whole protein were monitored. σ^{32} III and IV, which carried information on the distances within either the N- or C-terminal domain respectively, showed a conformation which was unchanged between the free and the DnaK-bound form when considering the FRET efficiency distribution. The conformational change induced by DnaK binding of σ^{32} therefore most likely occurs in the region between its N-terminal and C-terminal domain around the residue C165. Interestingly, the low FRET state induced by the binding of DnaK to σ^{32} I and II had a static appearance, while the high FRET states observed for all four mutants showed conformational dynamics in histograms of donor lifetime versus FRET efficiency. σ^{32} in its free forms was thus slightly flexible, while DnaK binding induced a static conformation with a large distance between the N- and C-terminus of the protein. In a next step, the conformational dynamics can be further analyzed using dynamic PDA, FRET-FCS, or filtered FCS as described for U2AF65 (Chapter 5.2.6).

We observed the binding of DnaK and the conformational change leading to an increase in the distance between both termini of the protein in the presence of both nucleotides, ADP and ATP. It is of high interest to observe the effect of nucleotide exchange factors to investigate the binding of DnaK to σ^{32} in both nucleotide states in more detail by keeping DnaK in an ATP-bound conformation even when hydrolysis occurs. Future measurements performed by Ganesh Agam will thus include experiments in the presence of DnaK's nucleotide exchange factor GrpE and the additional J-domain containing cofactor of DnaK, DnaJ. Often J-proteins play a role in the binding of Hsp70 substrates. The conformation of σ^{32} in the presence of DnaJ alone as well as in combination with DnaK can thus be investigated. For the addition of σ^{32} to DnaK and DnaJ, a high association of stable complexes was observed in the presence of ATP (Gamer et al., 1996; Liberek and Georgopoulos, 1993; Liberek et al., 1995; Noguchi et al., 2014) while binding of DnaK to σ^{32} alone has only been observed in the presence of ADP (Liberek et al., 1995; Noguchi et al., 2014; Rodriguez et al., 2008). σ^{32}-DnaK complexes formed in the absence of DnaJ and ATP were more stable than σ^{32}-DnaK-DnaJ complexes (Gamer et al., 1996). Using different concentrations of the cofactors and the nucleotides, the affinity of the formation of the complexes as well as their stability can be investigated.

In a subsequent step, it will be interesting to investigate the conformational change observed in the presence of RNA polymerase. RNA polymerase binding should render σ^{32} in an active conformational state, ready to induce transcription from promoters of heat shock induced genes. The possible occurrence of RNA polymerase-bound σ^{32} in a compact conformation may raise the question whether free σ^{32} is actively able to bind to DNA even in the absence of RNA polymerase. This hypothesis could be tested using FCS experiments of σ^{32} in the presence of σ^{32} promoter-containing DNA.

The here observed large scale conformational changes between the N-terminal and C-terminal domains of σ^{32} in the presence of DnaK, open up a whole field of investigations on this protein using spFRET experiments in solution or on the surface to obtain information about the actual conformational transitions. The active and inactive conformational states of σ^{32} can be unravelled by studying the protein in the presence of different interaction partners. This can lead to the discovery of the underlying conformational states necessary for the functional mechanism of the protein.

8 Conclusions and Summary

The tight regulation of all cellular processes is highly important for cell maintenance, differentiation, or apoptosis. The regulation of gene expression takes place at several steps on the way to a fully functional protein. Since the energy cost for protein production is high, most long-term regulations are performed on the level of transcription initiation. The expression level can be further regulated by adjusting the mRNA for translation. Short term regulations can take place by limiting the amount of functional protein through inhibition and degradation. The chapters of this thesis comprise one of these aspects of the regulation of expression levels each, starting from transcription initiation.

The transcription factor HMGA1 and its Epstein Barr virus mimick EBNA1 were shown to bind to palindromic AATATT-DNA sequences. The DNA-protein complex did not assume a simple 1:1 stoichiometry, but rather multiple HMGA1 or EBNA1 molecules interacted with the DNA forming higher order complexes by electrostatic interaction. A more detailed analysis is necessary to investigate the effect of HMGA1 binding on the DNA conformation as was performed for TBP.

The transcription initiation factor TBP is the first protein binding to TATA-box containing promoters. It regulates the association of additional subunits of the early RNA polymerase complex on the DNA. The binding of TBP to the DNA is strongly regulated by a variety of transcription cofactors. The effects of Mot1, NC2, and PC4 on the dissociation behavior of TBP from the DNA were investigated. Mot1 induced dissociation of the protein when present in the ternary complex as well as free in solution. A two step dissociation mechanism was proposed in which the stably bound Mot1 induced a conformational change of the complex, which could be observed by spFRET. It's ATPase activity was not sufficient to dissociate the complex, but rather primed the complex for dissociation. Mot1 in solution was necessary for the dissocation of TBP from these weakened complexes.
The bending dynamics of DNA in the presence of TBP (Heiss, 2011) were further investigated using fluorescently labeled TBP. The experiments gave a hint to the orientation-independent binding of TBP. Measurements of the complex in the presence of TFIIA, which promoted the orientation-specificity of TBP on the DNA supported these observations.

The regulation of protein expression on the mRNA level was investigated by analyzing a splicing factor involved in alternative splicing. U2AF takes part in the formation of the early pre-splicing complex. It binds to the Py-tract with its large subunit (65 kDa) and to the conserved AG-dinucleotide via its small subunit (35 kDa). The conformation of the minimal RNA-binding large subunit, RRM1,2, was shown to dynamically transit between a closed and an open conformation. The binding of Py-tract RNA of increasing strength gradually shifted the equilibrium of RRM1,2 towards the open conformation. An elongated U2AF65 with the binding domain for U2AF35 similarly led to a shift of the equilibrium to the open conformation, while the additional U2AF35 subunit even further strengthened that effect. A

gradually increasing affinity was observed for these constructs. This is of particular interest because splicing from high affinity Py-tracts was shown to occur independently of U2AF35 while this small subunit is necessary for splicing from weak Py-tracts. The occurrence of Py-tracts with varying strengths opens the possibility for alternative splicing. Alternative splicing takes place in 90% of all human genes and determines the ratio of one gene product over a second longer or shorter version.

Moving downstream the road of protein expression after successful translation, the proteins need to adopt their native tertiary structure. Many newly translated proteins require the aid of at least one molecular chaperone to obtain their final structure. Hsp70 chaperones bind to nascent proteins and enable them to fold, prevent aggregation, or aid in protein transport. Using three-color FRET, the coordinated motion of these chaperones in their conformational cycle was analyzed. A comparison between chaperones from different organelles showed a strongly coordinated behavior of ADP-bound BiP and DnaK while the probed distance distributions of Ssc1's domains appeared to move slightly more independently of each other. The binding of a short substrate peptide affected the conformation of Ssc1, and DnaK, while it left BiP's conformation unchanged. The nucleotide exchange factors Mge1 and GrpE induced changes in Ssc1's and DnaK's conformational states similar to the peptide-bound conformations. The chaperones also differed in their selectivity towards nucleotides. DnaK was highly selective, adopting its ATP-bound conformation only in the presence of a single ATP analog, while BiP was more likely to be found in the ATP-bound state with different nucleotide analogs. Several inhibitors for the cytosolic human Hsp70s were shown to interact with BiP and DnaK and can therefore not only function as anti-cancer drugs, but also as antibacterials.

σ^{32}, as a native substrate of DnaK, was investigated regarding its conformation in the presence of different binding partners. σ^{32} is active in the transcription from heat shock induced promoters when bound to RNA polymerase, while it is kept in an inactive, readily degradable form when bound to DnaK. Using spFRET, the conformation of σ^{32} was shown to differ between its free form and the DnaK-bound protein. σ^{32} adopted a condensed closed structure when free in solution and the distance between its N-terminal and C-terminal domains was extended when bound to DnaK. This extension of the protein offers the N-terminus as a target for degradation by FtsH protease and its occurrence is thus regulated by the binding to DnaK.

The transition between different conformational states is important for the function of many proteins on the path to protein expression. Using single molecule fluorescence methods, this conformational state and its dynamics become accessible. The used techniques spFRET, single molecule three-color FRET, and FCS, offer the possibility to obtain information about the interaction, distances within or between molecules, the timescale of their dynamics, and the coordination of motion. All this information was used to link the underlying conformational mechanism of the described proteins to their function. The ability to look at

these aspects of the functional mechanism of proteins promises the elucidation of a variety of cellular processes in the future.

Bibliography

Agashe, V.R., Guha, S., Chang, H.-C., Genevaux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., and Barral, J.M. (2004). Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. Cell *117*, 199–209.

Alberts, B. (2015). Molecular biology of the cell (New York, NY: Garland Science, Taylor and Francis Group).

Alexiou, G.A., Vartholomatos, G., Stefanaki, K., Patereli, A., Dova, L., Karamoutsios, A., Lallas, G., Sfakianos, G., Moschovi, M., and Prodromou, N. (2013). Expression of heat shock proteins in medulloblastoma. J Neurosurg Pediatr *12*, 452–457.

Alexiou, G.A., Karamoutsios, A., Lallas, G., Ragos, V., Goussia, A., Kyritsis, A.P., Voulgaris, S., and Vartholomatos, G. (2014). Expression of heat shock proteins in brain tumors. Turk Neurosurg *24*, 745–749.

Altman, R.B., Terry, D.S., Zhou, Z., Zheng, Q., Geggier, P., Kolster, R.A., Zhao, Y., Javitch, J.A., Warren, J.D., and Blanchard, S.C. (2011). Cyanine fluorophore derivatives with enhanced photostability. Nature Methods *9*, 68–71.

Altman, R.B., Zheng, Q., Zhou, Z., Terry, D.S., Warren, J.D., and Blanchard, S.C. (2012). Enhanced photostability of cyanine fluorophores across the visible spectrum. Nature Methods *9*, 428–429.

Andrau, J.-C., Van Oevelen, C.J., Van Teeffelen, H.A., Weil, P.A., Holstege, F.C., and Timmers, H.T. (2002). Mot1p is essential for TBP recruitment to selected promoters during in vivo gene activation. The EMBO Journal *21*, 5173–5183.

Andrec, M., Levy, R.M., and Talaga, D.S. (2003). Direct Determination of Kinetic Rates from Single-Molecule Photon Arrival Trajectories Using Hidden Markov Models. J Phys Chem A *107*, 7454–7464.

Antonik, M., Felekyan, S., Gaiduk, A., and Seidel, C.A. (2006). Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. The Journal of Physical Chemistry B *110*, 6970–6978.

Aravind, L., and Landsman, D. (1998). AT-hook motifs identified in a wide variety of DNA-binding proteins. Nucleic Acids Res. *26*, 4413–4421.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics *22*, 195–201.

BIBLIOGRAPHY

Auble, D.T. (2009). The dynamic personality of TATA-binding protein. Trends in Biochemical Sciences *34*, 49–52.

Auble, D.T., and Hahn, S. (1993). An ATP-dependent inhibitor of TBP binding to DNA. Genes & Development *7*, 844–856.

Auble, D.T., Hansen, K.E., Mueller, C.G., Lane, W.S., Thorner, J., and Hahn, S. (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. Genes & Development *8*, 1920–1934.

Bacia, K., and Schwille, P. (2007). Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. Nature Protocols *2*, 2842–2856.

Banerjee, H. (2004). The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF65) is essential in vivo but dispensable for activity in vitro. RNA *10*, 240–253.

Bardwell, J.C., and Craig, E.A. (1984). Major heat shock gene of Drosophila and the Escherichia coli heat-inducible dnaK gene are homologous. Proc. Natl. Acad. Sci. U.S.A. *81*, 848–852.

Baum, L.E., and Petrie, T. (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. The Annals of Mathematical Statistics *37*, 1554–1563.

Baum, L.E., Petrie, T., Soules, G., and Weiss, N. (1970). A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. The Annals of Mathematical Statistics *41*, 164–171.

Baumann, F., Milisav, I., Neupert, W., and Herrmann, J.M. (2000). Ecm10, a novel hsp70 homolog in the mitochondrial matrix of the yeast Saccharomyces cerevisiae. FEBS Lett. *487*, 307–312.

Behnke, J., Feige, M.J., and Hendershot, L.M. (2015). BiP and its nucleotide exchange factors Grp170 and Sil1: mechanisms of action and biological functions. J. Mol. Biol. *427*, 1589–1608.

Benecke, A., Eilebrecht, S., Benecke, A., and Eilebrecht, S. (2015). RNA-Mediated Regulation of HMGA1 Function. Biomolecules *5*, 943–957.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407–412.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5'terminus of adenovirus 2 late mRNA. Proceedings of the National Academy of Sciences *74*, 3171–3175.

Berk, A.J., and Sharp, P.A. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell *12*, 721–732.

Bertelsen, E.B., Chang, L., Gestwicki, J.E., and Zuiderweg, E.R. (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proceedings of the National Academy of Sciences *106*, 8471–8476.

Bertolotti, A., Zhang, Y., Hendershot, L.M., Harding, H.P., and Ron, D. (2000). Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nat. Cell Biol. *2*, 326–332.

Bewley, C.A., Gronenborn, A.M., and Clore, G.M. (1998). Minor groove-binding architectural proteins: structure, function, and DNA recognition. Annu Rev Biophys Biomol Struct *27*, 105–131.

Bhattacharya, A., Kurochkin, A.V., Yip, G.N.B., Zhang, Y., Bertelsen, E.B., and Zuiderweg, E.R.P. (2009). Allostery in Hsp70 Chaperones Is Transduced by Subdomain Rotations. Journal of Molecular Biology *388*, 475–490.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research *42*, W252–W258.

Bielinsky, A.K., and Gerbi, S.A. (2001). Where it all starts: eukaryotic origins of DNA replication. J. Cell. Sci. *114*, 643–651.

Binnig, null, Quate, null, and Gerber, null (1986). Atomic force microscope. Phys. Rev. Lett. *56*, 930–933.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. *72*, 291–336.

Blair, R.H., Goodrich, J.A., and Kugel, J.F. (2012). Single-Molecule Fluorescence Resonance Energy Transfer Shows Uniformity in TATA Binding Protein-Induced DNA Bending and Heterogeneity in Bending Kinetics. Biochemistry *51*, 7444–7455.

Blencowe, B.J. (2006). Alternative splicing: new insights from global analyses. Cell *126*, 37–47.

Bochkarev, A., Barwell, J.A., Pfuetzner, R.A., Bochkareva, E., Frappier, L., and Edwards, A.M. (1996). Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein, EBNA1, bound to DNA. Cell *84*, 791–800.

Bochkarev, A., Bochkareva, E., Frappier, L., and Edwards, A.M. (1998). The 2.2 Ang structure of a permanganate-sensitive DNA site bound by the Epstein-Barr virus origin binding protein, EBNA1. Journal of Molecular Biology *284*, 1273–1278.

BIBLIOGRAPHY

Böhmer, M., Wahl, M., Rahn, H.-J., Erdmann, R., and Enderlein, J. (2002). Time-resolved fluorescence correlation spectroscopy. Chemical Physics Letters *353*, 439–445.

Bolender, N., Sickmann, A., Wagner, R., Meisinger, C., and Pfanner, N. (2008). Multiple pathways for sorting mitochondrial precursor proteins. EMBO Rep. *9*, 42–49.

Bonomo, J., Welsh, J.P., Manthiram, K., and Swartz, J.R. (2010). Comparing the functional properties of the Hsp70 chaperones, DnaK and BiP. Biophysical Chemistry *149*, 58–66.

Brabetz, W., and Weber, C. Novel combination of fluorescent dyes for the detection of nucleic acids.

Broadley, S.A., and Hartl, F.U. (2009). The role of molecular chaperones in human misfolding diseases. FEBS Lett. *583*, 2647–2653.

Brooks Shera, E., Seitzinger, N.K., Davis, L.M., Keller, R.A., and Soper, S.A. (1990). Detection of single fluorescent molecules. Chemical Physics Letters *174*, 553–557.

Buschmann, V., Weston, K.D., and Sauer, M. (2003). Spectroscopic study and evaluation of red-absorbing fluorescent dyes. Bioconjug. Chem. *14*, 195–204.

Butryn, A., Schuller, J.M., Stoehr, G., Runge-Wollmann, P., Förster, F., Auble, D.T., and Hopfner, K.-P. (2015). Structural basis for recognition and remodeling of the TBP: DNA: NC2 complex by Mot1. Elife *4*, e07432.

Carlsson, L., and Lazarides, E. (1983). ADP-ribosylation of the Mr 83,000 stress-inducible and glucose-regulated protein in avian and mammalian cells: modulation by heat shock and glucose starvation. Proceedings of the National Academy of Sciences *80*, 4664–4668.

Cech, T.R. (1986). The generality of self-splicing RNA: relationship to nuclear mRNA splicing. Cell *44*, 207–210.

Chang, Y.-W., Sun, Y.-J., Wang, C., and Hsiao, C.-D. (2008). Crystal Structures of the 70-kDa Heat Shock Proteins in Domain Disjoining Conformation. Journal of Biological Chemistry *283*, 15502–15511.

Chase, M.B., Haga, S.B., Hankins, W.D., Williams, D.M., Bi, Z., Strovel, J.W., Obriecht, C., and Berg, P.E. (1999). Binding of HMG-I(Y) elicits structural changes in a silencer of the human beta-globin gene. Am. J. Hematol. *60*, 27–35.

Chatterjee, A., Sun, S.B., Furman, J.L., Xiao, H., and Schultz, P.G. (2013). A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in *Escherichia coli*. Biochemistry *52*, 1828–1837.

Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. Nature Reviews Genetics *8*, 93–103.

Chen, Y., Zhang, Y., Yin, Y., Gao, G., Li, S., Jiang, Y., Gu, X., and Luo, J. (2005). SPD-a web-based secreted protein database. Nucleic Acids Res. *33*, D169–D173.

Chow, L.T., Roberts, J.M., Lewis, J.B., and Broker, T.R. (1977). A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. Cell *11*, 819–836.

Chung, K.T. (2002). BAP, a Mammalian BiP-associated Protein, Is a Nucleotide Exchange Factor That Regulates the ATPase Activity of BiP. Journal of Biological Chemistry *277*, 47557–47563.

Chung, H.S., Louis, J.M., and Eaton, W.A. (2010). Distinguishing between protein dynamics and dye photophysics in single-molecule FRET experiments. Biophys. J. *98*, 696–706.

Chusainow, J. (2005). FRET analyses of the U2AF complex localize the U2AF35/U2AF65 interaction in vivo and reveal a novel self-interaction of U2AF35. RNA *11*, 1201–1214.

Clamme, J.-P., and Deniz, A.A. (2005). Three-Color Single-Molecule Fluorescence Resonance Energy Transfer. ChemPhysChem *6*, 74–77.

Collart, M.A. (1996). The NOT, SPT3, and MOT1 genes functionally interact to regulate transcription at core promoters. Molecular and Cellular Biology *16*, 6668–6676.

Coolidge, C.J., Seely, R.J., and Patton, J.G. (1997). Functional analysis of the polypyrimidine tract in pre-mRNA splicing. Nucleic Acids Research *25*, 888–896.

Cooper, M., Ebner, A., Briggs, M., Burrows, M., Gardner, N., Richardson, R., and West, R. (2004). Cy3B: improving the performance of cyanine dyes. J Fluoresc *14*, 145–150.

Corsini, L., Bonna, S., Basquin, J., Hothorn, M., Scheffzek, K., Valcárcel, J., and Sattler, M. (2007). U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. Nature Structural & Molecular Biology *14*, 620–629.

Cox, J.M., Hayward, M.M., Sanchez, J.F., Gegnas, L.D., van der Zee, S., Dennis, J.H., Sigler, P.B., and Schepartz, A. (1997). Bidirectional binding of the TATA box binding protein to the TATA box. Proc. Natl. Acad. Sci. U.S.A. *94*, 13475–13480.

Craig, E.A., and Marszalek, J. (2002). A specialized mitochondrial molecular chaperone system: a role in formation of Fe/S centers. Cell. Mol. Life Sci. *59*, 1658–1665.

Craig, E.A., Kramer, J., and Kosic-Smithers, J. (1987). SSC1, a member of the 70-kDa heat shock protein multigene family of Saccharomyces cerevisiae, is essential for growth. Proceedings of the National Academy of Sciences *84*, 4156–4160.

Craig, E.A., Huang, P., Aron, R., and Andrew, A. (2006). The diverse roles of J-proteins, the obligate Hsp70 co-chaperone. Rev. Physiol. Biochem. Pharmacol. *156*, 1–21.

Creton, S. (2002). The NC2 alpha and beta subunits play different roles in vivo. Genes & Development *16*, 3265–3276.

Curtis, D., Lehmann, R., and Zamore, P.D. (1995). Translational regulation in development. Cell *81*, 171–178.

Dahan, M., Deniz, A.A., Ha, T., Chemla, D.S., Schultz, P.G., and Weiss, S. (1999). Ratiometric measurement and identification of single diffusing molecules. Chemical Physics *247*, 85–106.

Darst, R.P. (2003). Mot1 Regulates the DNA Binding Activity of Free TATA-binding Protein in an ATP-dependent Manner. Journal of Biological Chemistry *278*, 13216–13226.

Darst, R.P., Wang, D., and Auble, D.T. (2001). MOT1-catalyzed TBP–DNA disruption: uncoupling DNA conformational change and role of upstream DNA. The EMBO Journal *20*, 2028–2040.

Dasgupta, A., Juedes, S.A., Sprouse, R.O., and Auble, D.T. (2005). Mot1-mediated control of transcription complex assembly and activity. The EMBO Journal *24*, 1717–1729.

Davis, J.L., Kunisawa, R., and Thorner, J. (1992). A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast Saccharomyces cerevisiae. Molecular and Cellular Biology *12*, 1879–1892.

Deloche, O., Kelley, W.L., and Georgopoulos, C. (1997). Structure-function analyses of the Ssc1p, Mdj1p, and Mge1p Saccharomyces cerevisiae mitochondrial proteins in Escherichia coli. J. Bacteriol. *179*, 6066–6075.

Deniz, A.A., Dahan, M., Grunwell, J.R., Ha, T., Faulhaber, A.E., Chemla, D.S., Weiss, S., and Schultz, P.G. (1999). Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. Proc. Natl. Acad. Sci. U.S.A. *96*, 3670–3675.

Deniz, A.A., Mukhopadhyay, S., and Lemke, E.A. (2008). Single-molecule biophysics: at the interface of biology, physics and chemistry. J R Soc Interface *5*, 15–45.

DePamphilis, M.L. (1999). Replication origins in metazoan chromosomes: fact or fiction? Bioessays *21*, 5–16.

Di Fiori, N., and Meller, A. (2010). The Effect of Dye-Dye Interactions on the Spatial Resolution of Single-Molecule FRET Measurements in Nucleic Acids. Biophysical Journal *98*, 2265–2272.

Disney, J.E., Johnson, K.R., Magnuson, N.S., Sylvester, S.R., and Reeves, R. (1989). High-mobility group protein HMG-I localizes to G/Q- and C-bands of human and mouse chromosomes. J. Cell Biol. *109*, 1975–1982.

Dragan, A.I., Liggins, J.R., Crane-Robinson, C., and Privalov, P.L. (2003). The energetics of specific binding of AT-hooks from HMGA1 to target DNA. J. Mol. Biol. *327*, 393–411.

Durbin, R. (1998). Biological sequence analysis: probabalistic models of proteins and nucleic acids. Cambridge University Press, Cambridge, UK : New York.

Eggeling, C., Berger, S., Brand, L., Fries, J.R., Schaffer, J., Volkmer, A., and Seidel, C.A.M. (2001). Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. Journal of Biotechnology *86*, 163–180.

Ellgaard, L., Molinari, M., and Helenius, A. (1999). Setting the standards: quality control in the secretory pathway. Science *286*, 1882–1888.

Elson, E.L. (2011). Fluorescence Correlation Spectroscopy: Past, Present, Future. Biophysical Journal *101*, 2855–2870.

Elson, E.L., and Magde, D. (1974). Fluorescence correlation spectroscopy. I. Conceptual basis and theory. Biopolymers *13*, 1–27.

Embley, T.M., and Martin, W. (2006). Eukaryotic evolution, changes and challenges. Nature *440*, 623–630.

Endo, T. (2003). Functional cooperation and separation of translocators in protein import into mitochondria, the double-membrane bounded organelles. Journal of Cell Science *116*, 3259–3267.

Epstein, W., and Beckwith, J.R. (1968). Regulation of gene expression. Annual Review of Biochemistry *37*, 411–436.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. Mol. Cell *36*, 593–608.

Falvo, J.V., Thanos, D., and Maniatis, T. (1995). Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). Cell *83*, 1101–1111.

Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A., and Seidel, C.A.M. (2012). Filtered FCS: Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules. ChemPhysChem *13*, 1036–1053.

Fonfría-Subirós, E., Acosta-Reyes, F., Saperas, N., Pous, J., Subirana, J.A., and Campos, J.L. (2012). Crystal Structure of a Complex of DNA with One AT-Hook of HMGA1. PLoS ONE *7*, e37120.

Foo, Y.H., Naredi-Rainer, N., Lamb, D.C., Ahmed, S., and Wohland, T. (2012). Factors Affecting the Quantification of Biomolecular Interactions by Fluorescence Cross-Correlation Spectroscopy. Biophysical Journal *102*, 1174–1183.

Frappier, L., and O'Donnell, M. (1991). Epstein-Barr nuclear antigen 1 mediates a DNA loop within the latent replication origin of Epstein-Barr virus. Proc. Natl. Acad. Sci. U.S.A. *88*, 10875–10879.

Frydman, J. (2001). Folding of newly translated proteins in vivo: the role of molecular chaperones. Annu. Rev. Biochem. *70*, 603–647.

Gambin, Y., and Deniz, A.A. (2010). Multicolor single-molecule FRET to explore protein folding and binding. Molecular BioSystems *6*, 1540.

Gamer, J., Multhaup, G., Tomoyasu, T., McCarty, J.S., Rüdiger, S., Schönfeld, H.J., Schirra, C., Bujard, H., and Bukau, B. (1996). A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the Escherichia coli heat shock transcription factor sigma32. EMBO J. *15*, 607–617.

Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Tóth, K., Langowski, J., and Seidel, C.A. (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET. Proceedings of the National Academy of Sciences *106*, 15308–15313.

Geierstanger, B.H., Volkman, B.F., Kremer, W., and Wemmer, D.E. (1994). Short peptide fragments derived from HMG-I/Y proteins bind specifically to the minor groove of DNA. Biochemistry *33*, 5347–5355.

Geisberg, J.V., Holstege, F.C., Young, R.A., and Struhl, K. (2001). Yeast NC2 Associates with the RNA Polymerase II Preinitiation Complex and Selectively Affects Transcription In Vivo. Molecular and Cellular Biology *21*, 2736–2742.

Geisberg, J.V., Moqtaderi, Z., Kuras, L., and Struhl, K. (2002). Mot1 Associates with Transcriptionally Active Promoters and Inhibits Association of NC2 in Saccharomyces cerevisiae. Molecular and Cellular Biology *22*, 8122–8134.

Genevaux, P., Georgopoulos, C., and Kelley, W.L. (2007). The Hsp70 chaperone machines of Escherichia coli: a paradigm for the repartition of chaperone functions. Mol. Microbiol. *66*, 840–857.

Gietl, A., Holzmeister, P., Blombach, F., Schulz, S., von Voithenberg, L.V., Lamb, D.C., Werner, F., Tinnefeld, P., and Grohmann, D. (2014). Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. Nucleic Acids Research.

Gilbert, D.M. (2001). Making sense of eukaryotic DNA replication origins. Science *294*, 96–100.

Golub, T.R. (1999). Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science *286*, 531–537.

Gozani, O., Feld, R., and Reed, R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev. *10*, 233–243.

Gozani, O., Potashkin, J., and Reed, R. (1998). A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol. Cell. Biol. *18*, 4752–4760.

Graveley, B.R. (2001). Alternative splicing: increasing diversity in the proteomic world. Trends Genet. *17*, 100–107.

Gray, M.W., Burger, G., Lang, B.F., and others (2001). The origin and early evolution of mitochondria. Genome Biol *2*, 1018–1.

Guex, N., Peitsch, M.C., and Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis *30*, S162–S173.

Gumbs, O.H., Campbell, A.M., and Weil, P.A. (2003). High-affinity DNA binding by a Mot1p–TBP complex: implications for TAF-independent transcription. The EMBO Journal *22*, 3131–3141.

Guo, F., and Snapp, E.L. (2013). ERdj3 regulates BiP occupancy in living cells. Journal of Cell Science *126*, 1429–1439.

Gust, A., Zander, A., Gietl, A., Holzmeister, P., Schulz, S., Lalkens, B., Tinnefeld, P., and Grohmann, D. (2014). A Starting Point for Fluorescence-Based Single-Molecule Measurements in Biomolecular Research. Molecules *19*, 15824–15865.

Guth, S., Martínez, C., Gaur, R.K., and Valcárcel, J. (1999). Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). Mol. Cell. Biol. *19*, 8263–8271.

Ha, T., Enderle, T., Ogletree, D.F., Chemla, D.S., Selvin, P.R., and Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between

a single donor and a single acceptor. Proceedings of the National Academy of Sciences *93*, 6264–6268.

Haas, I.G., and Wabl, M. (1983). Immunoglobulin heavy chain binding protein. Nature *306*, 387–389.

Hammond, C., and Helenius, A. (1995). Quality control in the secretory pathway. Curr. Opin. Cell Biol. *7*, 523–529.

Harrison, C. (2003). GrpE, a nucleotide exchange factor for DnaK. Cell Stress & Chaperones *8*, 218.

Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. Nature *475*, 324–332.

Haustein, E., and Schwille, P. (2007). Fluorescence correlation spectroscopy: novel variations of an established technique. Annu Rev Biophys Biomol Struct *36*, 151–169.

Hearing, J., Mülhaupt, Y., and Harper, S. (1992). Interaction of Epstein-Barr virus nuclear antigen 1 with the viral latent origin of replication. J. Virol. *66*, 694–705.

Heiss, G. (2011). Single-Molecule Microscopy Study of Nano-Systems: From synthetic photo-switchable nano-devices to the dynamic of naturally occurring transcription factors. Dissertation Ludwig-Maximilians-Universität, München.

Hell, S.W., and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt Lett *19*, 780–782.

Hendershot, L.M. (1990). Immunoglobulin heavy chain and binding protein complexes are dissociated in vivo by light chain addition. J. Cell Biol. *111*, 829–837.

Hendershot, L.M. (2004). The ER function BiP is a master regulator of ER function. Mt. Sinai J. Med. *71*, 289–297.

Hendershot, L.M., Ting, J., and Lee, A.S. (1988). Identity of the immunoglobulin heavy-chain-binding protein with the 78,000-dalton glucose-regulated protein and the role of posttranslational modifications in its binding function. Molecular and Cellular Biology *8*, 4250–4256.

Hendrix, J., Schrimpf, W., Höller, M., and Lamb, D.C. (2013). Pulsed Interleaved Excitation Fluctuation Imaging. Biophysical Journal *105*, 848–861.

Herschel, J.F.W. (1845). Formula No. I. On a Case of Superficial Colour Presented by a Homogeneous Liquid Internally Colourless. Philosophical Transactions of the Royal Society of London *135*, 143–145.

Hieb, A.R., Gansen, A., Bohm, V., and Langowski, J. (2014). The conformational state of the nucleosome entry-exit site modulates TATA box-specific TBP binding. Nucleic Acids Research *42*, 7561–7576.

Hirschfeld, T. (1976). Optical microscopic observation of single small molecules. Appl Opt *15*, 2965–2966.

Horst, M., Oppliger, W., Rospert, S., Schönfeld, H.J., Schatz, G., and Azem, A. (1997). Sequential action of two hsp70 complexes during protein import into mitochondria. EMBO J. *16*, 1842–1849.

Hoskins, A.A., Friedman, L.J., Gallagher, S.S., Crawford, D.J., Anderson, E.G., Wombacher, R., Ramirez, N., Cornish, V.W., Gelles, J., and Moore, M.J. (2011). Ordered and Dynamic Assembly of Single Spliceosomes. Science *331*, 1289–1295.

Hovius, R., Guignet, E., Segura, J.-M., Piguet, J., Vogel, H., and Bäumle, M. (2006). Single Molecule Detection with Atto647N NTA. BioFiles *1.3*.

Hsieh, D.J., Camiolo, S.M., and Yates, J.L. (1993). Constitutive binding of EBNA1 protein to the Epstein-Barr virus replication origin, oriP, with distortion of DNA structure during latent infection. EMBO J. *12*, 4933–4944.

Hsu, J.-Y., Juven-Gershon, T., Marr, M.T., Wright, K.J., Tjian, R., and Kadonaga, J.T. (2008). TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription. Genes & Development *22*, 2353–2358.

Huang, J., Warner, L.R., Sanchez, C., Gabel, F., Madl, T., Mackereth, C.D., Sattler, M., and Blackledge, M. (2014). Transient Electrostatic Interactions Dominate the Conformational Equilibrium Sampled by Multidomain Splicing Factor U2AF65: A Combined NMR and SAXS Study. Journal of the American Chemical Society *136*, 7068–7076.

Hughes, L.D., Rawle, R.J., and Boxer, S.G. (2014). Choose Your Label Wisely: Water-Soluble Fluorophores Often Interact with Lipid Bilayers. PLoS ONE *9*, e87649.

Hung, S.C., Kang, M.-S., and Kieff, E. (2001). Maintenance of Epstein–Barr virus (EBV) oriP-based episomes requires EBV-encoded nuclear antigen-1 chromosome-binding domains, which can be replaced by high-mobility group-I or histone H1. Proceedings of the National Academy of Sciences *98*, 1865–1870.

Huth, J.R., Bewley, C.A., Nissen, M.S., Evans, J.N., Reeves, R., Gronenborn, A.M., and Clore, G.M. (1997). The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. Nat. Struct. Biol. *4*, 657–665.

Jacob, F., and Monod, J. (1961). Genetic regulatory Mechanisms in the Synthesis of Proteins. J. Mol. Biol. *2*, 318–356.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nature Genetics *33*, 245–254.

Jiang, J., Prasad, K., Lafer, E.M., and Sousa, R. (2005). Structural basis of interdomain communication in the Hsc70 chaperone. Mol. Cell *20*, 513–524.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., and Shoemaker, D.D. (2003a). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science *302*, 2141–2144.

Johnson, S.A.S., Dubeau, L., Kawalek, M., Dervan, A., Schonthal, A.H., Dang, C.V., and Johnson, D.L. (2003b). Increased Expression of TATA-Binding Protein, the Central Transcription Factor, Can Contribute to Oncogenesis. Molecular and Cellular Biology *23*, 3043–3051.

Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C., and Ha, T. (2008). Advances in Single-Molecule Fluorescence Methods for Molecular Biology. Annual Review of Biochemistry *77*, 51–76.

Kaiser, K., and Meisterernst, M. (1996). The human general co-factors. Trends Biochem. Sci. *21*, 342–345.

Kalinin, S., Valeri, A., Antonik, M., Felekyan, S., and Seidel, C.A.M. (2010). Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. The Journal of Physical Chemistry B *114*, 7983–7995.

Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and Seidel, C.A.M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nature Methods *9*, 1218–1225.

Kamada, K., Shu, F., Chen, H., Malik, S., Stelzer, G., Roeder, R.G., Meisterernst, M., and Burley, S.K. (2001). Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. Cell *106*, 71–81.

Kanaar, R., Roche, S.E., Beall, E.L., Green, M.R., and Rio, D.C. (1993). The conserved pre-mRNA splicing factor U2AF from Drosophila: requirement for viability. Science *262*, 569–573.

Kang, P.J., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A., and Pfanner, N. (1990). Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. Nature *348*, 137–143.

Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-

laser excitation of single molecules. Proceedings of the National Academy of Sciences of the United States of America *101*, 8936–8941.

Kapanidis, A.N., Laurence, T.A., Lee, N.K., Margeat, E., Kong, X., and Weiss, S. (2005). Alternating-Laser Excitation of Single Molecules. Accounts of Chemical Research *38*, 523–533.

Karzai, A.W., and McMacken, R. (1996). A bipartite signaling mechanism involved in DnaJ-mediated activation of the Escherichia coli DnaK protein. J. Biol. Chem. *271*, 11236–11246.

Kassenbrock, C.K., and Kelly, R.B. (1989). Interaction of heavy chain binding protein (BiP/GRP78) with adenine nucleotides. EMBO J. *8*, 1461–1467.

Kays, A.R., and Schepartz, A. (2000). Virtually unidirectional binding of TBP to the AdMLP TATA box within the quaternary complex with TFIIA and TFIIB. Chemistry & Biology *7*, 601–610.

Kearsey, S.E., and Cotterill, S. (2003). Enigmatic variations: divergent modes of regulating eukaryotic DNA replication. Mol. Cell *12*, 1067–1075.

Kellenberger, E., Stier, G., and Sattler, M. (2002). Induced folding of the U2AF35 RRM upon binding to U2AF65. FEBS Lett. *528*, 171–176.

Kelly, T.J., and Brown, G.W. (2000). Regulation of chromosome replication. Annu. Rev. Biochem. *69*, 829–880.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. Nucleic Acids Research *37*, D387–D392.

Kielkopf, C.L., Rodionova, N.A., Green, M.R., and Burley, S.K. (2001). A novel peptide recognition mode revealed by the X-ray structure of a core U2AF 35/U2AF 65 heterodimer. Cell *106*, 595–605.

Kim, J.L., and Burley, S.K. (1994). 1.9 A resolution refined structure of TBP recognizing the minor groove of TATAAAAG. Nat. Struct. Biol. *1*, 638–653.

Kim, E., Magen, A., and Ast, G. (2006). Different levels of alternative splicing among eukaryotes. Nucleic Acids Research *35*, 125–131.

Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B. (1993). Crystal structure of a yeast TBP/TATA-box complex. Nature *365*, 512–520.

Kityk, R., Kopp, J., Sinning, I., and Mayer, M.P. (2012). Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. Molecular Cell *48*, 863–874.

Koehler, C.M. (2004). New developments in mitochondrial assembly. Annu. Rev. Cell Dev. Biol. *20*, 309–335.

Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew. Chem. Int. Ed. Engl. *40*, 2004–2021.

Kolmakov, K., Belov, V., Bierwagen, J., Ringemann, C., Müller, V., Eggeling, C., and Hell, S. (2010). Red-Emitting Rhodamine Dyes for Fluorescence Microscopy and Nanoscopy. Chemistry - A European Journal *16*, 158–166.

Konarska, M.M., Vilardell, J., and Query, C.C. (2006). Repositioning of the reaction intermediate within the catalytic center of the spliceosome. Mol. Cell *21*, 543–553.

Koshioka, M., Sasaki, K., and Masuhara, H. (1995). Time-Dependent Fluorescence Depolarization Analysis in Three-Dimensional Microspectroscopy. Appl. Spectrosc. 224–228.

Krichevsky, O., and Bonnet, G. (2002). Fluorescence correlation spectroscopy: the technique and its applications. Reports on Progress in Physics *65*, 251.

Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C.A.M., and Lamb, D.C. (2012). Combining MFD and PIE for Accurate Single-Pair Förster Resonance Energy Transfer Measurements. ChemPhysChem *13*, 1060–1078.

Kügel, W. (2012). Development and application of advanced single molecule fluorescence methods using PIE-MFD.

Lamb, H.K., Mee, C., Xu, W., Liu, L., Blond, S., Cooper, A., Charles, I.G., and Hawkins, A.R. (2006). The affinity of a major Ca2+ binding site on GRP78 is differentially enhanced by ADP and ATP. J. Biol. Chem. *281*, 8796–8805.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947–2948.

Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J., and Bukau, B. (1999). Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. Proc. Natl. Acad. Sci. U.S.A. *96*, 5452–5457.

Lee, N.K., Kapanidis, A.N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R.H., and Weiss, S. (2005). Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. Biophysical Journal *88*, 2939–2953.

Lee, S., Lee, J., and Hohng, S. (2010). Single-Molecule Three-Color FRET with Both Negligible Spectral Overlap and Long Observation Time. PLoS ONE *5*, e12270.

Lianos, G.D., Alexiou, G.A., Mangano, A., Mangano, A., Rausei, S., Boni, L., Dionigi, G., and Roukos, D.H. (2015). The role of heat shock proteins in cancer. Cancer Letters *360*, 114–118.

Liao, M., Zhang, Y., Kang, J.-H., and Dufau, M.L. (2011). Coactivator function of positive cofactor 4 (PC4) in Sp1-directed luteinizing hormone receptor (LHR) gene transcription. J. Biol. Chem. *286*, 7681–7691.

Liberek, K., and Georgopoulos, C. (1993). Autoregulation of the Escherichia coli heat shock response by the DnaK and DnaJ heat shock proteins. Proc. Natl. Acad. Sci. U.S.A. *90*, 11019–11023.

Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C., and Zylicz, M. (1991). Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. Proc. Natl. Acad. Sci. U.S.A. *88*, 2874–2878.

Liberek, K., Galitski, T.P., Zylicz, M., and Georgopoulos, C. (1992). The DnaK chaperone modulates the heat shock response of Escherichia coli by binding to the sigma 32 transcription factor. Proceedings of the National Academy of Sciences *89*, 3516–3520.

Liberek, K., Wall, D., and Georgopoulos, C. (1995). The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the sigma 32 heat shock transcriptional regulator. Proceedings of the National Academy of Sciences *92*, 6224–6228.

Liebscher, M., Jahreis, G., Lücke, C., Grabley, S., Raina, S., and Schiene-Fischer, C. (2007). Fatty acyl benzamido antibacterials based on inhibition of DnaK-catalyzed protein folding. J. Biol. Chem. *282*, 4437–4446.

Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E.M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nature Biotechnology *25*, 117–124.

Lund, T., Holtlund, J., Fredriksen, M., and Laland, S.G. (1983). On the presence of two new high mobility group-like proteins in HeLa S3 cells. FEBS Lett. *152*, 163–167.

Mackereth, C.D., Madl, T., Bonnal, S., Simon, B., Zanier, K., Gasch, A., Rybin, V., Valcárcel, J., and Sattler, M. (2011). Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. Nature *475*, 408–411.

Madison, J.M., and Winston, F. (1997). Evidence that Spt3 functionally interacts with Mot1, TFIIA, and TATA-binding protein to confer promoter-specific transcriptional control in Saccharomyces cerevisiae. Mol. Cell. Biol. *17*, 287–295.

Malik, S., Guermah, M., and Roeder, R.G. (1998). A dynamic model for PC4 coactivator function in RNA polymerase II transcription. Proceedings of the National Academy of Sciences *95*, 2192–2197.

Mally, A., and Witt, S.N. (2001). GrpE accelerates peptide binding and release from the high affinity state of DnaK. Nat. Struct. Biol. *8*, 254–257.

Mapa, K., Sikor, M., Kudryavtsev, V., Waegemann, K., Kalinin, S., Seidel, C.A.M., Neupert, W., Lamb, D.C., and Mokranjac, D. (2010). The Conformational Dynamics of the Mitochondrial Hsp70 Chaperone. Molecular Cell *38*, 89–100.

Marcinowski, M., Höller, M., Feige, M.J., Baerend, D., Lamb, D.C., and Buchner, J. (2011). Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. Nature Structural & Molecular Biology *18*, 150–158.

Marcinowski, M., Rosam, M., Seitz, C., Elferich, J., Behnke, J., Bello, C., Feige, M.J., Becker, C.F.W., Antes, I., and Buchner, J. (2013). Conformational Selection in Substrate Recognition by Hsp70 Chaperones. Journal of Molecular Biology *425*, 466–474.

Masson, P., Leimgruber, E., Creton, S., and Collart, M.A. (2007). The dual control of TFIIB recruitment by NC2 is gene specific. Nucleic Acids Research *36*, 539–549.

Masters, K.M., Parkhurst, K.M., Daugherty, M.A., and Parkhurst, L.J. (2003). Native Human TATA-binding Protein Simultaneously Binds and Bends Promoter DNA without a Slow Isomerization Step or TFIIB Requirement. Journal of Biological Chemistry *278*, 31685–31690.

Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: Cellular functions and molecular mechanism. Cellular and Molecular Life Sciences *62*, 670–684.

Mayer, M., Reinstein, J., and Buchner, J. (2003). Modulation of the ATPase Cycle of BiP by Peptides and Proteins. Journal of Molecular Biology *330*, 137–144.

Mayer, M.P., Schröder, H., Rüdiger, S., Paal, K., Laufen, T., and Bukau, B. (2000). Multistep mechanism of substrate binding determines chaperone activity of Hsp70. Nature Structural & Molecular Biology *7*, 586–593.

McCarty, J.S., Buchberger, A., Reinstein, J., and Bukau, B. (1995). The role of ATP in the functional cycle of the DnaK chaperone system. J. Mol. Biol. *249*, 126–137.

McCarty, J.S., Rüdiger, S., Schönfeld, H.J., Schneider-Mergener, J., Nakahigashi, K., Yura, T., and Bukau, B. (1996). Regulatory region C of the E. coli heat shock transcription factor, sigma32, constitutes a DnaK binding site and is conserved among eubacteria. J. Mol. Biol. *256*, 829–837.

McFadden, G.I. (2001). Chloroplast origin and integration. Plant Physiology *125*, 50–53.

McKinney, S.A., Joo, C., and Ha, T. (2006). Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. Biophysical Journal *91*, 1941–1951.

Meisterernst, M., Roy, A.L., Lieu, H.M., and Roeder, R.G. (1991). Activation of class II gene transcription by regulatory factors is potentiated by a novel activity. Cell *66*, 981–993.

Merendino, L., Guth, S., Bilbao, D., Martínez, C., and Valcárcel, J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. Nature *402*, 838–841.

Milles, S., Tyagi, S., Banterle, N., Koehler, C., VanDelinder, V., Plass, T., Neal, A.P., and Lemke, E.A. (2012a). Click Strategies for Single-Molecule Protein Fluorescence. Journal of the American Chemical Society *134*, 5187–5195.

Milles, S., Koehler, C., Gambin, Y., Deniz, A.A., and Lemke, E.A. (2012b). Intramolecular three-colour single pair FRET of intrinsically disordered proteins with increased dynamic range. Molecular BioSystems *8*, 2531.

Mitchell, P., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371–378.

Mokranjac, D., Bourenkov, G., Hell, K., Neupert, W., and Groll, M. (2006). Structure and function of Tim14 and Tim16, the J and J-like components of the mitochondrial protein import motor. EMBO J. *25*, 4675–4685.

Mollet, I., Barbosa-Morais, N.L., Andrade, J., and Carmo-Fonseca, M. (2006). Diversity of human U2AF splicing factors.: Based on the EMBO Lecture delivered on 7 July 2005 at the 30th FEBS Congress in Budapest. FEBS Journal *273*, 4807–4816.

Moro, F., Okamoto, K., Donzeau, M., Neupert, W., and Brunner, M. (2002). Mitochondrial protein import: molecular basis of the ATP-dependent interaction of MtHsp70 with Tim44. J. Biol. Chem. *277*, 6874–6880.

Moyle-Heyrman, G., Viswanathan, R., Widom, J., and Auble, D.T. (2012). Two-step Mechanism for Modifier of Transcription 1 (Mot1) Enzyme-catalyzed Displacement of TATA-binding Protein (TBP) from DNA. Journal of Biological Chemistry *287*, 9002–9012.

Müller, B.K., Zaychikov, E., Bräuchle, C., and Lamb, D.C. (2005). Pulsed Interleaved Excitation. Biophysical Journal *89*, 3508–3522.

Müller, C.B., Loman, A., Pacheco, V., Koberling, F., Willbold, D., Richtering, W., and Enderlein, J. (2008). Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. EPL (Europhysics Letters) *83*, 46001.

Munro, S., and Pelham, H.R. (1986). An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell *46*, 291–300.

Murphy, K. (1998). Hidden Markov Model (HMM) Toolbox for Matlab.

Murphy, M.E. (2013). The HSP70 family and cancer. Carcinogenesis *34*, 1181–1188.

Naylor, D.J., Hoogenraad, N.J., and Hoj, P.B. (1996). Isolation and characterisation of a cDNA encoding rat mitochondrial GrpE, a stress-inducible nucleotide-exchange factor of ubiquitous appearance in mammalian organs. FEBS Lett. *396*, 181–188.

Neupert, W., and Brunner, M. (2002). The protein import motor of mitochondria. Nat. Rev. Mol. Cell Biol. *3*, 555–565.

Neupert, W., and Herrmann, J.M. (2007). Translocation of proteins into mitochondria. Annu. Rev. Biochem. *76*, 723–749.

Nicolaï, A., Delarue, P., and Senet, P. (2013). Decipher the Mechanisms of Protein Conformational Changes Induced by Nucleotide Binding through Free-Energy Landscape Analysis: ATP Binding to Hsp70. PLoS Computational Biology *9*, e1003379.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. Nature *377*, 119–128.

Nilsen, T.W. (2003). The spliceosome: the most complex macromolecular machine in the cell? BioEssays *25*, 1147–1149.

Nir, E., Michalet, X., Hamadani, K.M., Laurence, T.A., Neuhauser, D., Kovchegov, Y., and Weiss, S. (2006). Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. The Journal of Physical Chemistry B *110*, 22103–22124.

Noguchi, A., Ikeda, A., Mezaki, M., Fukumori, Y., and Kanemori, M. (2014). DnaJ-Promoted Binding of DnaK to Multiple Sites on 32 in the Presence of ATP. Journal of Bacteriology *196*, 1694–1703.

Norseen, J., Thomae, A., Sridharan, V., Aiyar, A., Schepers, A., and Lieberman, P.M. (2008). RNA-dependent recruitment of the origin recognition complex. The EMBO Journal *27*, 3024–3035.

Norseen, J., Johnson, F.B., and Lieberman, P.M. (2009). Role for G-Quadruplex RNA Binding by Epstein-Barr Virus Nuclear Antigen 1 in DNA Replication and Metaphase Chromosome Attachment. Journal of Virology *83*, 10336–10346.

Ogawa, M., Kosaka, N., Choyke, P.L., and Kobayashi, H. (2009). In vivo Molecular Imaging of Cancer with a Quenching Near-Infrared Fluorescent Probe Using Conjugates of Monoclonal Antibodies and Indocyanine Green. Cancer Research *69*, 1268–1272.

Otaka, M., Yamamoto, S., Ogasawara, K., Takaoka, Y., Noguchi, S., Miyazaki, T., Nakai, A., Odashima, M., Matsuhashi, T., Watanabe, S., et al. (2007). The induction mechanism of the molecular chaperone HSP70 in the gastric mucosa by Geranylgeranylacetone (HSP-inducer). Biochemical and Biophysical Research Communications *353*, 399–404.

Pacheco, T.R., Coelho, M.B., Desterro, J.M.P., Mollet, I., and Carmo-Fonseca, M. (2006). In Vivo Requirement of the Small Subunit of U2AF for Recognition of a Weak 3' Splice Site. Molecular and Cellular Biology *26*, 8183–8190.

Packschies, L., Theyssen, H., Buchberger, A., Bukau, B., Goody, R.S., and Reinstein, J. (1997). GrpE accelerates nucleotide exchange of the molecular chaperone DnaK with an associative displacement mechanism. Biochemistry *36*, 3417–3422.

Palleros, D.R., Reid, K.L., Shi, L., Welch, W.J., and Fink, A.L. (1993). ATP-induced protein-Hsp70 complex dissociation requires K+ but not ATP hydrolysis. Nature *365*, 664–666.

Parkhurst, K.M., Brenowitz, M., and Parkhurst, L.J. (1996). Simultaneous Binding and Bending of Promoter DNA by the TATA Binding Protein: Real Time Kinetic Measurements †. Biochemistry *35*, 7459–7465.

Parkhurst, K.M., Richards, R.M., Brenowitz, M., and Parkhurst, L.J. (1999). Intermediate species possessing bent DNA are present along the pathway to formation of a final TBP-TATA complex. J. Mol. Biol. *289*, 1327–1341.

Patel, A.A., and Steitz, J.A. (2003). Splicing double: insights from the second spliceosome. Nat. Rev. Mol. Cell Biol. *4*, 960–970.

Patikoglou, G.A., Kim, J.L., Sun, L., Yang, S.-H., Kodadek, T., and Burley, S.K. (1999). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. Genes & Development *13*, 3217–3230.

Pawley, J.B. (2006). Handbook of biological confocal microscopy. Springer *3rd ed*, New York, NY.

Perez-Howard, G.M., Weil, P.A., and Beechem, J.M. (1995). Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. Biochemistry *34*, 8005–8017.

Perkins, T.T., Smith, D.E., and Chu, S. (1994). Direct observation of tube-like motion of a single polymer chain. Science *264*, 819–822.

Pierpaoli, E.V., Sandmeier, E., Baici, A., Schönfeld, H.J., Gisler, S., and Christen, P. (1997). The power stroke of the DnaK/DnaJ/GrpE molecular chaperone system. J. Mol. Biol. *269*, 757–768.

Plass, T., Milles, S., Koehler, C., Schultz, C., and Lemke, E.A. (2011). Genetically Encoded Copper-Free Click Chemistry. Angewandte Chemie International Edition *50*, 3878–3881.

Poon, D., Campbell, A.M., Bai, Y., and Weil, P.A. (1994). Yeast Taf170 is encoded by MOT1 and exists in a TATA box-binding protein (TBP)-TBP-associated factor complex distinct from transcription factor IID. J. Biol. Chem. *269*, 23135–23140.

Price, E.S., Aleksiejew, M., and Johnson, C.K. (2011). FRET-FCS Detection of Intralobe Dynamics in Calmodulin. The Journal of Physical Chemistry B *115*, 9320–9326.

Qin, F., Auerbach, A., and Sachs, F. (2000). A Direct Optimization Approach to Hidden Markov Modeling for Single Channel Kinetics. Biophysical Journal *79*, 1915–1927.

Rabiner, L.., and Juang, B.. (1986). An introduction to hidden Markov models. IEEE ASSP Magazine *3*, 4–16.

Raven, J.A., and Allen, J.F. (2003). Genomics and chloroplast evolution: what did cyanobacteria do for plants. Genome Biol *4*, 209.

Reed, R. (1989). The organization of 3'splice-site sequences in mammalian introns. Genes & Development *3*, 2113–2123.

Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. Gene *277*, 63–81.

Reeves, R. (2010). Nuclear functions of the HMG proteins. Biochim. Biophys. Acta *1799*, 3–14.

Reeves, R., and Nissen, M.S. (1990). The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. J. Biol. Chem. *265*, 8573–8582.

Reeves, R., and Nissen, M.S. (1995). Cell cycle regulation and functions of HMG-I(Y). Prog Cell Cycle Res *1*, 339–349.

Rehling, P., Brandner, K., and Pfanner, N. (2004). Mitochondrial import and the twin-pore translocase. Nature Reviews Molecular Cell Biology *5*, 519–530.

Rigler, R., and Elson, E.S. (2001). Fluorescence Correlation Spectroscopy Theory and Applications (Berlin, Heidelberg: Springer Berlin Heidelberg).

Rodriguez, F., Arsène-Ploetze, F., Rist, W., Rüdiger, S., Schneider-Mergener, J., Mayer, M.P., and Bukau, B. (2008). Molecular Basis for Regulation of the Heat Shock Transcription Factor σ32 by the DnaK and DnaJ Chaperones. Molecular Cell *32*, 347–358.

Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. Sci. *21*, 327–335.

Roeder, R.G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Letters *579*, 909–915.

Rotman, B. (1961). Measurement of activity of single molecules of beta-D-galactosidase. Proc. Natl. Acad. Sci. U.S.A. *47*, 1981–1991.

Rowley, N., Prip-Buus, C., Westermann, B., Brown, C., Schwarz, E., Barrell, B., and Neupert, W. (1994). Mdj1p, a novel chaperone of the DnaJ family, is involved in mitochondrial biogenesis and protein folding. Cell *77*, 249–259.

Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. Nature Methods *5*, 507–516.

Rudner, D.Z., Kanaar, R., Breger, K.S., and Rio, D.C. (1996). Mutations in the small subunit of the Drosophila U2AF splicing factor cause lethality and developmental defects. Proceedings of the National Academy of Sciences *93*, 10333–10337.

Rudner, D.Z., Kanaar, R., Breger, K.S., and Rio, D.C. (1998). Interaction between subunits of heterodimeric splicing factor U2AF is essential in vivo. Molecular and Cellular Biology *18*, 1765–1773.

Rüttinger, S., Macdonald, R., Krämer, B., Koberling, F., Roos, M., and Hildt, E. (2006). Accurate single-pair Förster resonant energy transfer through combination of pulsed interleaved excitation, time correlated single-photon counting, and fluorescence correlation spectroscopy. Journal of Biomedical Optics *11*, 024012.

Sarkari, F., Sanchez-Alcaraz, T., Wang, S., Holowaty, M.N., Sheng, Y., and Frappier, L. (2009). EBNA1-Mediated Recruitment of a Histone H2B Deubiquitylating Complex to the Epstein-Barr Virus Latent Origin of DNA Replication. PLoS Pathogens *5*, e1000624.

Schaffer, J., Volkmer, A., Eggeling, C., Subramaniam, V., Striker, G., and Seidel, C.A.M. (1999). Identification of Single Molecules in Aqueous Solution by Time-Resolved Fluorescence Anisotropy. The Journal of Physical Chemistry A *103*, 331–336.

Schepers, A., Ritzi, M., Bousset, K., Kremmer, E., Yates, J.L., Harwood, J., Diffley, J.F., and Hammerschmidt, W. (2001). Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. EMBO J. *20*, 4588–4602.

Schilke, B., Williams, B., Knieszner, H., Pukszta, S., D'Silva, P., Craig, E.A., and Marszalek, J. (2006). Evolution of mitochondrial chaperones utilized in Fe-S cluster biogenesis. Curr. Biol. *16*, 1660–1665.

Schlecht, R., Scholz, S.R., Dahmen, H., Wegener, A., Sirrenberg, C., Musil, D., Bomke, J., Eggenweiler, H.-M., Mayer, M.P., and Bukau, B. (2013). Functional analysis of Hsp70 inhibitors. PloS One *8*, e78443.

Schluesche, P. (2007). Untersuchung der Funktion und Dynamik von DNA-Transkriptionsfaktoren mittels Einzelmolekül-Fluoreszenzmikroskopie. Dissertation Ludwig-Maximilians-Universität, München.

Schluesche, P., Stelzer, G., Piaia, E., Lamb, D.C., and Meisterernst, M. (2007). NC2 mobilizes TBP on core promoter TATA boxes. Nature Structural & Molecular Biology *14*, 1196–1201.

Schluesche, P., Heiss, G., Meisterernst, M., and Lamb, D.C. (2008). Dynamics of TBP binding to the TATA box. In Single Molecule Spectroscopy and Imaging, J. Enderlein, Z.K. Gryczynski, and R. Erdmann, eds. p. 68620E – 68620E – 8.

Schmid, D., Baici, A., Gehring, H., and Christen, P. (1994). Kinetics of molecular chaperone action. Science *263*, 971–973.

Schroda, M., Vallon, O., Whitelegge, J.P., Beck, C.F., and Wollman, F.A. (2001). The chloroplastic GrpE homolog of Chlamydomonas: two isoforms generated by differential splicing. Plant Cell *13*, 2823–2839.

Schröder, G.F., and Grubmüller, H. (2003). Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments. The Journal of Chemical Physics *119*, 9920.

Schröder, H., Langer, T., Hartl, F.U., and Bukau, B. (1993). DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J. *12*, 4137–4144.

Schwille, P., Meyer-Almes, F.-J., and Rigler, R. (1997). Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. Biophysical Journal *72*, 1878.

Sears, J., Ujihara, M., Wong, S., Ott, C., Middeldorp, J., and Aiyar, A. (2004). The amino terminus of Epstein-Barr Virus (EBV) nuclear antigen 1 contains AT hooks that facilitate the replication and partitioning of latent EBV genomes by tethering them to cellular chromosomes. J. Virol. *78*, 11487–11505.

Shen, J., Chen, X., Hendershot, L., and Prywes, R. (2002). ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals. Dev. Cell *3*, 99–111.

Sickmier, E.A., Frato, K.E., Shen, H., Paranawithana, S.R., Green, M.R., and Kielkopf, C.L. (2006). Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. Molecular Cell *23*, 49–59.

Sikor, M. (2011). Single-molecule fluorescence studies of protein folding and molecular chaperones. Dissertation Ludwig-Maximilians-Universität, München.

Sikor, M., Mapa, K., von Voithenberg, L.V., Mokranjac, D., and Lamb, D.C. (2013). Real-time observation of the conformational dynamics of mitochondrial Hsp70 by spFRET. The EMBO Journal *32*, 1639–1649.

Singh, R., Valcárcel, J., and Green, M.R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science *268*, 1173–1176.

Slepenkov, S.V., and Witt, S.N. (2003). Detection of a concerted conformational change in the ATPase domain of DnaK triggered by peptide binding. FEBS Letters *539*, 100–104.

Smith, C.W., and Valcárcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. Trends in Biochemical Sciences *25*, 381–388.

Smith, S., Finzi, L., and Bustamante, C. (1992). Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. Science *258*, 1122–1126.

Solomon, M.J., Strauss, F., and Varshavsky, A. (1986). A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. Proc. Natl. Acad. Sci. U.S.A. *83*, 1276–1280.

Sorenson, M.K., and Darst, S.A. (2006). Disulfide cross-linking indicates that FlgM-bound and free sigma28 adopt similar conformations. Proc. Natl. Acad. Sci. U.S.A. *103*, 16722–16727.

Sorenson, M.K., Ray, S.S., and Darst, S.A. (2004). Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. Mol. Cell *14*, 127–138.

Spedale, G., Meddens, C.A., Koster, M.J.E., Ko, C.W., van Hooff, S.R., Holstege, F.C.P., Timmers, H.T.M., and Pijnappel, W.W.M.P. (2012). Tight cooperation between Mot1p and NC2 in regulating genome-wide transcription, repression of transcription following heat shock induction and genetic interaction with SAGA. Nucleic Acids Research *40*, 996–1008.

Spencer, J.V., and Arndt, K.M. (2002). A TATA Binding Protein Mutant with Increased Affinity for DNA Directs Transcription from a Reversed TATA Sequence In Vivo. Molecular and Cellular Biology *22*, 8744–8755.

Sprouse, R.O., Brenowitz, M., and Auble, D.T. (2006). Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. The EMBO Journal *25*, 1492–1504.

Sprouse, R.O., Shcherbakova, I., Cheng, H., Jamison, E., Brenowitz, M., and Auble, D.T. (2008). Function and Structural Organization of Mot1 Bound to a Natural Target Promoter. Journal of Biological Chemistry *283*, 24935–24948.

Straus, D., Walter, W., and Gross, C.A. (1990). DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. Genes Dev. *4*, 2202–2209.

Strauss, F., and Varshavsky, A. (1984). A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell *37*, 889–901.

Stryer, L. (1978). Fluorescence energy transfer as a spectroscopic ruler. Annu. Rev. Biochem. *47*, 819–846.

Swain, J.F., Dinler, G., Sivendran, R., Montgomery, D.L., Stotz, M., and Gierasch, L.M. (2007). Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol. Cell *26*, 27–39.

Szabo, A., Langer, T., Schröder, H., Flanagan, J., Bukau, B., and Hartl, F.U. (1994). The ATP hydrolysis-dependent reaction cycle of the Escherichia coli Hsp70 system DnaK, DnaJ, and GrpE. Proc. Natl. Acad. Sci. U.S.A. *91*, 10345–10349.

Tavanez, J.P., Madl, T., Kooshapur, H., Sattler, M., and Valcárcel, J. (2012). hnRNP A1 Proofreads 3′ Splice Site Recognition by U2AF. Molecular Cell *45*, 314–329.

Thickman, K.R., Sickmier, E.A., and Kielkopf, C.L. (2007). Alternative Conformations at the RNA-binding Surface of the N-terminal U2AF 65 RNA Recognition Motif. Journal of Molecular Biology *366*, 703–710.

Thomae, A.W., Pich, D., Brocher, J., Spindler, M.-P., Berens, C., Hock, R., Hammerschmidt, W., and Schepers, A. (2008). Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins. Proceedings of the National Academy of Sciences *105*, 1692–1697.

Thomas, M.C., and Chiang, C.-M. (2006). The General Transcription Machinery and General Cofactors. Critical Reviews in Biochemistry and Molecular Biology *41*, 105–178.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research *22*, 4673–4680.

Tinnefeld, P., and Cordes, T. (2012). "Self-healing" dyes: intramolecular stabilization of organic fluorophores. Nat. Methods *9*, 426–427; author reply 427–428.

Tolić-Nørrelykke, S.F., Rasmussen, M.B., Pavone, F.S., Berg-Sørensen, K., and Oddershede, L.B. (2006). Stepwise Bending of DNA by a Single TATA-Box Binding Protein. Biophysical Journal *90*, 3694–3703.

Toor, N., Keating, K.S., Taylor, S.D., and Pyle, A.M. (2008). Crystal Structure of a Self-Spliced Group II Intron. Science *320*, 77–82.

Torres, T., and Levitus, M. (2007). Measuring Conformational Dynamics: A New FCS-FRET Approach. The Journal of Physical Chemistry B *111*, 7392–7400.

Tyagi, S., and Lemke, E.A. (2013). Genetically Encoded Click Chemistry for Single-Molecule FRET of Proteins. In Methods in Cell Biology, (Elsevier), pp. 169–187.

Tyagi, S., and Lemke, E.A. (2015). Single-molecule FRET and crosslinking studies in structural biology enabled by noncanonical amino acids. Current Opinion in Structural Biology *32*, 66–73.

Umen, J.G., and Guthrie, C. (1995). The second catalytic step of pre-mRNA splicing. RNA *1*, 869–885.

Valcárcel, J., Gaur, R.K., Singh, R., and Green, M.R. (1996). Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science *273*, 1706–1709.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. Nature *417*, 712–719.

Viswanathan, R., and Auble, D.T. (2011). One small step for Mot1; one giant leap for other Swi2/Snf2 enzymes? Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1809*, 488–496.

Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory *13*, 260–269.

Vogel, B., Loschberger, A., Sauer, M., and Hock, R. (2011). Cross-linking of DNA through HMGA1 suggests a DNA scaffold. Nucleic Acids Research *39*, 7124–7133.

Voos, W., and Röttgers, K. (2002). Molecular chaperones as essential mediators of mitochondrial biogenesis. Biochim. Biophys. Acta *1592*, 51–62.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell *136*, 701–718.

Walter, N.G., Huang, C.-Y., Manzo, A.J., and Sobhy, M.A. (2008). Do-it-yourself guide: how to use the modern single-molecule toolkit. Nature Methods *5*, 475–489.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Weiss, S. (1999). Fluorescence Spectroscopy of Single Biomolecules. Science *283*, 1676–1683.

Weiss, S. (2000). Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. Nature Structural Biology *7*, 724–729.

Weitzmann, A., Baldes, C., Dudek, J., and Zimmermann, R. (2007). The heat shock protein 70 molecular chaperone network in the pancreatic endoplasmic reticulum − a quantitative approach: Pancreatic endoplasmic reticulum chaperone network. FEBS Journal *274*, 5175–5187.

Welsh, J.P., Bonomo, J., and Swartz, J.R. (2011). Localization of BiP to translating ribosomes increases soluble accumulation of secreted eukaryotic proteins in an Escherichia coli cell-free system. Biotechnology and Bioengineering *108*, 1739–1748.

van Werven, F.J., van Bakel, H., van Teeffelen, H.A.A.M., Altelaar, A.F.M., Koerkamp, M.G., Heck, A.J.R., Holstege, F.C.P., and Timmers, H.T.M. (2008). Cooperative action of NC2 and Mot1p to regulate TATA-binding protein function across the genome. Genes & Development *22*, 2359–2369.

Westermann, B., Gaume, B., Herrmann, J.M., Neupert, W., and Schwarz, E. (1996). Role of the mitochondrial DnaJ homolog Mdj1p as a chaperone for mitochondrially synthesized and imported proteins. Molecular and Cellular Biology *16*, 7063–7071.

Westphal, V., Rizzoli, S.O., Lauterbach, M.A., Kamin, D., Jahn, R., and Hell, S.W. (2008). Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. Science *320*, 246–249.

Whittington, J.E., Delgadillo, R.F., Attebury, T.J., Parkhurst, L.K., Daugherty, M.A., and Parkhurst, L.J. (2008). TATA-Binding Protein Recognition and Bending of a Consensus Promoter Are Protein Species Dependent †. Biochemistry *47*, 7264–7273.

Widengren, J., and Schwille, P. (2000). Characterization of photoinduced isomerization and back-isomerization of the cyanine dye Cy5 by fluorescence correlation spectroscopy. The Journal of Physical Chemistry A *104*, 6416–6428.

Will, C.L., and Luhrmann, R. (2011). Spliceosome Structure and Function. Cold Spring Harbor Perspectives in Biology *3*, a003707–a003707.

Williamson, D.S., Borgognoni, J., Clay, A., Daniels, Z., Dokurno, P., Drysdale, M.J., Foloppe, N., Francis, G.L., Graham, C.J., Howes, R., et al. (2009). Novel adenosine-derived inhibitors of 70 kDa heat shock protein, discovered through structure-based design. J. Med. Chem. *52*, 1510–1513.

Wollmann, P., Cui, S., Viswanathan, R., Berninghausen, O., Wells, M.N., Moldt, M., Witte, G., Butryn, A., Wendler, P., Beckmann, R., et al. (2011). Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. Nature *475*, 403–407.

Wolner, B.S., and Gralla, J.D. (2001). TATA-flanking Sequences Influence the Rate and Stability of TATA-binding Protein and TFIIB Binding. Journal of Biological Chemistry *276*, 6260–6266.

Wu, J. (2001). DNA Sequence-dependent Differences in TATA-binding Protein-induced DNA Bending in Solution Are Highly Sensitive to Osmolytes. Journal of Biological Chemistry *276*, 14623–14627.

Wu, S.Y. (1998). TAFII-independent activation mediated by human TBP in the presence of the positive cofactor PC4. The EMBO Journal *17*, 4478–4490.

Wu, S.-Y., and Chiang, C.-M. (1998). Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro. Journal of Biological Chemistry *273*, 12492–12498.

Wu, B., Wang, Z., Huang, Y., and Liu, W.R. (2012a). Catalyst-Free and Site-Specific One-Pot Dual-Labeling of a Protein Directed by Two Genetically Incorporated Noncanonical Amino Acids. ChemBioChem *13*, 1405–1408.

Wu, C.-C., Naveen, V., Chien, C.-H., Chang, Y.-W., and Hsiao, C.-D. (2012b). Crystal Structure of DnaK Protein Complexed with Nucleotide Exchange Factor GrpE in DnaK Chaperone System: Insight into intermolecular communication. Journal of Biological Chemistry *287*, 21461–21470.

Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature *402*, 832–835.

Wurm, C.A., Kolmakov, K., Göttfert, F., Ta, H., Bossi, M., Schill, H., Berning, S., Jakobs, S., Donnert, G., Belov, V.N., et al. (2012). Novel red fluorophores with superior performance in STED microscopy. Optical Nanoscopy *1*, 1–7.

Yamamoto, S., Nakano, S., Owari, K., Fuziwara, K., Ogawa, N., Otaka, M., Tamaki, K., Watanabe, S., Komatsuda, A., Wakui, H., et al. (2010). Gentamicin inhibits HSP70-assisted protein folding by interfering with substrate recognition. FEBS Letters *584*, 645–651.

Yang, H. (2003). Protein Conformational Dynamics Probed by Single-Molecule Electron Transfer. Science *302*, 262–266.

Yang, J., Nune, M., Zong, Y., Zhou, L., and Liu, Q. (2015). Close and Allosteric Opening of the Polypeptide-Binding Site in a Human Hsp70 Chaperone BiP. Structure *23*, 2191–2203.

Yoneda, T., Benedetti, C., Urano, F., Clark, S.G., Harding, H.P., and Ron, D. (2004). Compartment-specific perturbation of protein handling activates genes encoding mitochondrial chaperones. J. Cell. Sci. *117*, 4055–4066.

Yoshida, H., Park, S.-Y., Oda, T., Akiyoshi, T., Sato, M., Shirouzu, M., Tsuda, K., Kuwasako, K., Unzai, S., Muto, Y., et al. (2015). A novel 3′ splice site recognition by the two zinc fingers in the U2AF small subunit. Genes & Development *29*, 1649–1660.

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. Nature *478*, 64–69.

Yoshidomi, K., Murakami, A., Yakabe, K., Sueoka, K., Nawata, S., and Sugino, N. (2014). Heat shock protein 70 is involved in malignant behaviors and chemosensitivities to cisplatin in cervical squamous cell carcinoma cells: Role of HSP70 in cervical SCC cells. Journal of Obstetrics and Gynaecology Research *40*, 1188–1196.

Young, T.S., Ahmad, I., Yin, J.A., and Schultz, P.G. (2010). An Enhanced System for Unnatural Amino Acid Mutagenesis in E. coli. Journal of Molecular Biology *395*, 361–374.

Zamore, P.D., Patton, J.G., and Green, M.R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. Nature *355*, 609–614.

Zanetti-Domingues, L.C., Tynan, C.J., Rolfe, D.J., Clarke, D.T., and Martin-Fernandez, M. (2013). Hydrophobic Fluorescent Probes Introduce Artifacts into Single Molecule Tracking Experiments Due to Non-Specific Binding. PLoS ONE *8*, e74200.

Zentner, G.E., and Henikoff, S. (2013). Mot1 Redistributes TBP from TATA-Containing to TATA-Less Promoters. Molecular and Cellular Biology *33*, 4996–5004.

Zhao, J., Jin, S.-B., Björkroth, B., Wieslander, L., and Daneholt, B. (2002). The mRNA export factor Dbp5 is associated with Balbiani ring mRNP from gene to cytoplasm. EMBO J. *21*, 1177–1187.

Zhao, K., Liu, M., and Burgess, R.R. (2005). The Global Transcriptional Response of Escherichia coli to Induced 32 Protein Involves 32 Regulon Activation Followed by Inactivation and Degradation of 32 in Vivo. Journal of Biological Chemistry *280*, 17758–17768.

Zhou, Z., Lin, I.-J., Darst, R.P., and Bungert, J. (2009). Maneuver at the transcription start site. Epigenetics *4*, 1–4.

Zhuravleva, A., and Gierasch, L.M. (2011). Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. Proc. Natl. Acad. Sci. U.S.A. *108*, 6987–6992.

Zhuravleva, A., and Gierasch, L.M. (2015). Substrate-binding domain conformational dynamics mediate Hsp70 allostery. Proceedings of the National Academy of Sciences 201506692.

Zhuravleva, A., Clerico, E.M., and Gierasch, L.M. (2012). An Interdomain Energetic Tug-of-War Creates the Allosterically Active State in Hsp70 Molecular Chaperones. Cell *151*, 1296–1307.

Zorio, D.A., and Blumenthal, T. (1999). Both subunits of U2AF recognize the 3' splice site in Caenorhabditis elegans. Nature *402*, 835–838.

Zuo, P., and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes & Development *10*, 1356–1368.

Zylicz, M., Ang, D., and Georgopoulos, C. (1987). The grpE protein of Escherichia coli. Purification and properties. J. Biol. Chem. *262*, 17437–17442.

Appendix

Additional Materials: Chapter 2

Table 2.2 Names of plasmids including vector backbones, genes of interests and their donor organisms, and mutations used in this work for direct recombinant expression or further cloning.

Plasmid	Protein of interest	Donor Organism	Mutations	Original plasmid provided from
pET-15b TBP C61	TATA-binding protein (TBP)	*S. cerevisiae*	S61C, C78A, C164A	Prof. Meisterernst (WWU Münster)
pET-15b TBP no Cys	TATA-binding protein (TBP)	*S. cerevisiae*	C78A, C164A	Prof. Meisterernst (WWU Münster)
pET-15b TBP C128	TATA-binding protein (TBP)	*S. cerevisiae*	C78A, S128C, C164A	Prof. Meisterernst (WWU Münster)
pET-15b TBP C218	TATA-binding protein (TBP)	*S. cerevisiae*	C78A, C164A, K218C	Prof. Meisterernst (WWU Münster)
pProEX-BiP-C167-C519	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167C, T519C	Prof. Buchner (TUM München)
pProEX-BiP-C519-C638	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T519C, T638C	Prof. Buchner (TUM München)
pET-Duet-1-Ssc1-C448-C590-Hep1	Stress-seventy subfamily C (Ssc1)	*S. cerevisiae*	D341C, I448C	Dr. Mokranjac (LMU München)
pET-Duet-1-Ssc1-C341-C448-Hep1	Stress-seventy subfamily C (Ssc1)	*S. cerevisiae*	D341C, I448C	Dr. Mokranjac (LMU München)
pET-Duet-1-DnaK-C318-C425	Hsp70 chaperone DnaK	*E. coli*	C15A, E318C, V425C	Dr. Mokranjac (LMU München)
pET-Duet-1-DnaK-C458-C563	Hsp70 chaperone DnaK	*E. coli*	C15A, N458C, T563C	Dr. Mokranjac (LMU München)
pQU60-mGrpEΔ43	Mitochondrial GrpE domain chaperone (Mge1)	*S. cerevisiae*		Dr. Mokranjac (LMU München)
pEvol AcPhe	p-acetylphenyl-alanine tRNA, aminoacyl-tRNA synthetase	*M. jannaschii*	Young et al. (2010)	Dr. Lemke (EMBL Heidelberg)
pEvol PylRS WT	propargyllysine tRNA, aminoacyl-tRNA synthetase	*M. jannaschii*	Young et al. (2010)	Dr. Lemke (EMBL Heidelberg)
pEvol PylRS AF	Cyclooctyne-containing UAA tRNA, aminoacyl-tRNA synthetase	*M. jannaschii*	(Plass et al., 2011; Young et al., 2010)	Prof. Schultz (Scripps La Jolla)

Additional Materials: Chapter 2

Table 2.3 Names of plasmids including vectors, genes of interests with their corresponding donor organisms, and mutations produced during this research study.

Plasmid	Protein of interest	Donor Organism	Mutations	Original Plasmid
pProEX-BiP-C167-UAA519	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167C, T519UAA	pProEx-BiP-C167-C519
pProEX-BiP-C167-C519-C638	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167C, T519C,T638C	pProEx-BiP-C167-C519
pProEX-BiP-C167-UAA519-C638	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167C, T519UAA, T638C	pProEx-BiP-C167-C519
pProEX-BiP-UAA167-C519-C638	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167UAA, T519C, T638C	pProEx-BiP-C167-C519
pBAD-BiP-C167-UAA519-C638	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167UAA, T519C, T638C	pProEx-BiP-C167-C519
pBAD-BiP-C167-C519-C638	Binding immunoglobulin Protein (BiP)	*M. musculus*	C42S, C421S, T167C, T519C	pProEx-BiP-C167-C519
pET-Duet-1-Ssc1-C341-UAA448	Stress-seventy subfamily C (Ssc1)	*S. cerevisiae*	D341C, I448UAA	pET-Duet-1-Ssc1-C341-C448-Hep1
pET-Duet-1-Ssc1-C341-C448-C590	Stress-seventy subfamily C (Ssc1)	*S. cerevisiae*	D341C, I448C, D590C	pET-Duet-1-Ssc1-C341-C448-Hep1
pET-Duet-1-Ssc1-C341-UAA448-C590	Stress-seventy subfamily C (Ssc1)	*S. cerevisiae*	D341C, I448UAA, D590C	pET-Duet-1-Ssc1-C341-C448-Hep1
pET-Duet-1-Ssc1-UAA341-C448-C590	Stress-seventy subfamily C (Ssc1)	*S. cerevisiae*	D341UAA, I448C, D590C	pET-Duet-1-Ssc1-C341-C448-Hep1
pET-Duet-1-DnaK-C318-UAA425	Hsp70 chaperone DnaK	*E. coli*	C15A, E318C, V425UAA	pET-Duet-1-DnaK-C318-C425
pET-Duet-1-DnaK-C318-C425-C563	Hsp70 chaperone DnaK	*E. coli*	C15A, E318C, V425C, T563C	pET-Duet-1-DnaK-C318-C425
pET-Duet-1-DnaK-C318-UAA425-C563	Hsp70 chaperone DnaK	*E. coli*	C15A, E318C, V425UAA, T563C	pET-Duet-1-DnaK-C318-C425
pET-Duet-1-DnaK-UAA318-C425-C563	Hsp70 chaperone DnaK	*E. coli*	C15A, E318UAA, V425C, T563C	pET-Duet-1-DnaK-C318-C425
pBAD-DnaK-C318-UAA425-C563	Hsp70 chaperone DnaK	*E. coli*	C15A, E318C, V425UAA, T563C	pET-Duet-1-DnaK-C318-C425
pBAD-DnaK-UAA318-C425-C563	Hsp70 chaperone DnaK	*E. coli*	C15A, E318UAA, V425C, T563C	pET-Duet-1-DnaK-C318-C425

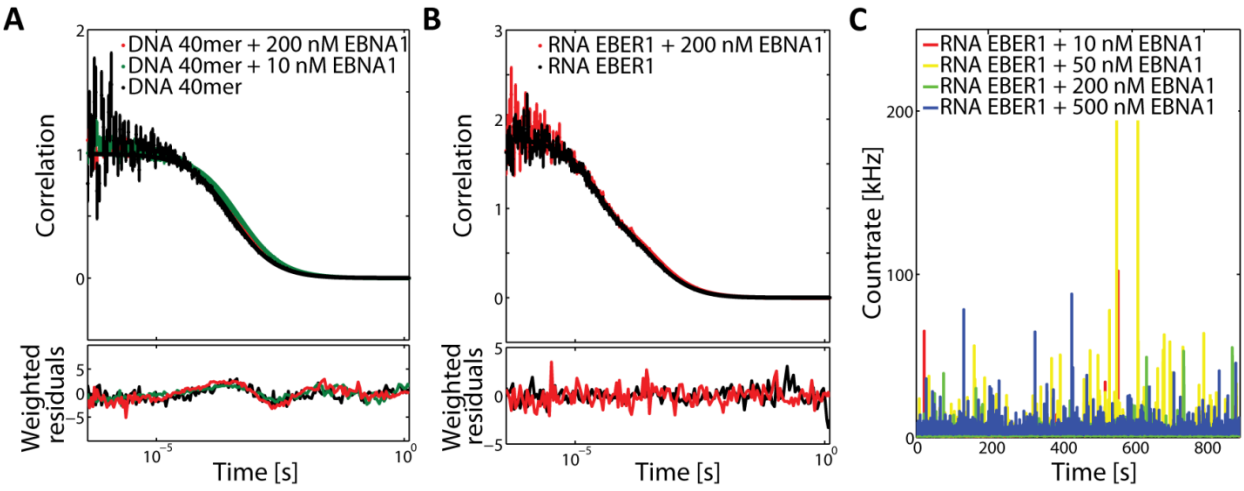
Additional Materials: Chapter 2

Table 2.4 Names and sequences of oligonucleotides (DNA and RNA) used in this work.

Oligonucleotide	**Sequence (5'→3')**	**Fluorophores/ *Additional Information***
TATA-DNA 25.3	Biotin-CTT CAC CTT ATT TGC ATA AGC GAT TCT ATA TAA AAG CGC CTC GCT CTA CCC TGC TCA CGC TGT TTT TCC TTT TCG TTG GC	T41-Atto647N, Complementary strand: T58-Atto532
TATA-DNA 18-Atto532	Biotin-CTT CAC CTT ATT TGC ATA AGC GAT TCT ATA TAA AAG CGC CTT GTC ATA CCC TGC TCA CGC TGT TTT TCC TTT TCG TTG GC	T41-Atto532
TATA-DNA 18-Atto647N	Biotin-CTT CAC CTT ATT TGC ATA AGC GAT TCT ATA TAA AAG CGC CTT GTC ATA CCC TGC TCA CGC TGT TTT TCC TTT TCG TTG GC	T41-Atto647N
TATA-DNA 18 (nick)-Cy5	Biotin-CTT CAC CTT ATT TGC ATA AGC GAT TCT ATA TAA AAG CGC CTT GTC ATA CCC TGC TCA CGC TGT TTT TCC TTT TCG TTG GC	T41-Cy5, complementary strand T58-nick
TATA-DNA18-2 (nick)-Cy5	Biotin-CTT CAC CTT ATT TGC ATA AGC GAT TCT ATA TAA AAG CGC CTT GTC ATA CCC TGC TCA CGC TGT TTT TCC TTT TCG TTG GC	T41-Cy5, complementary strand T58-nick
TATA-DNA 18'	Biotin-CTT CAC CTT ATT TGC ATA AGC GAT TCT ATA TAA AAG CGC CTT GTC ATA CCC TGC TCA CGC TGT TTT TCC TTT TCG TTG GC	-
TATA-DNA B11	Biotin-GTA CGG CCG GGC GCC CCG GAT GGG GGC CTA TAA AAG GGC TGG GCG	Complementary strand C21-Cy5-nick, *Provided by Prof. Auble*
HMGA1-DNA 30mer	GACCGCGAATATTCGCGCGTTATTTCGATAGC/ GCTATCGAAATAACGCGCGAATATTCGCGGTC	G1-Atto532, C30-Atto647N
HMGA1-DNA 40mer	TCCAGGACCGCGAATATTCGCGCGTTATTTCGATAGCAATA/ TATTGCTATCGAAATAACGCGCGAATATTCGCGGTCCTGGA	T1-Atto532, A40-Atto647N
EBER1	aggaccuacgcugcccuagagguuuugcuagggaggagacgugugguggcuguagccacccg ucccggguacaagucccgggguggugaggacggugucuguggguugucuucccagacucugcu uucugccgucuucggucaaguaccagcuggugguccgcauguuuu	*Provided by Dr. Schepers*
EBER2	aggacagccguugcccuaguggguuucggacacaccgccaacgcucagugcggcgcuaccgac ccgaggucaaguccccgggggaggagaagagaggcuucccgccuagagcauuugcaagucagg auucucuaaucccucugggagaaggguauucggcuuguccgcuguuuuu	*Provided by Dr. Schepers*
U9	uuuuuuuuu	
U4A8U4	uuuuaaaaaaaauuuu	
U13ACAGG	uuuuuuuuuuuuuacagg	
U4A8U4ACAGG	uuuuaaaaaaaauuuuacagg	
U4A8U5ACAGG	uuuuaaaaaaaauuuuuacagg	
random	cugcccuuggguca	
Bio-random-U13	Biotin-ugacccaagggcaguuuuuuuuuuu	
Bio-random-U9	Biotin-ugacccaagggcaguuuuuuuuu	
U13-Atto488	uuuuuuuuuuuuu	5'-Atto488
BiP-UAA519 fw	Phosphate-ATC ACA ATT ACC AAT GAC CAA AAT CGC C	
BiP-UAA519 rev	Phosphate-TTT GTT TTT GTT TCC CTA ACC TTT GTC TTC AG	
BiP-C638 fw	Phosphate-TCC CCC AAC TGG TGA AGA GGA TAC ATC	
BiP-C638 rev	Phosphate-GGG CCG CCA CAT CCA TAG AGT T	
BiP-UAA167 fw	Phosphate-TGG GAA AGA AGG TTT AGC ATG CAG TTG TTA	
BiP-UAA167 rev	Phosphate-AAT ATG CTT CAG CAG TTT CTT TCA TTT TAG TGA G	
BiP seq fw	GAC AAC CAT CTT CTG GGT ACA TT	
BiP seq rev	TTA CTA CAA CTC ATC TTT TTC TGA TGT ATC	
BiP seq167 fw	GTT GAT ATT GGA GGT GGG CAA	

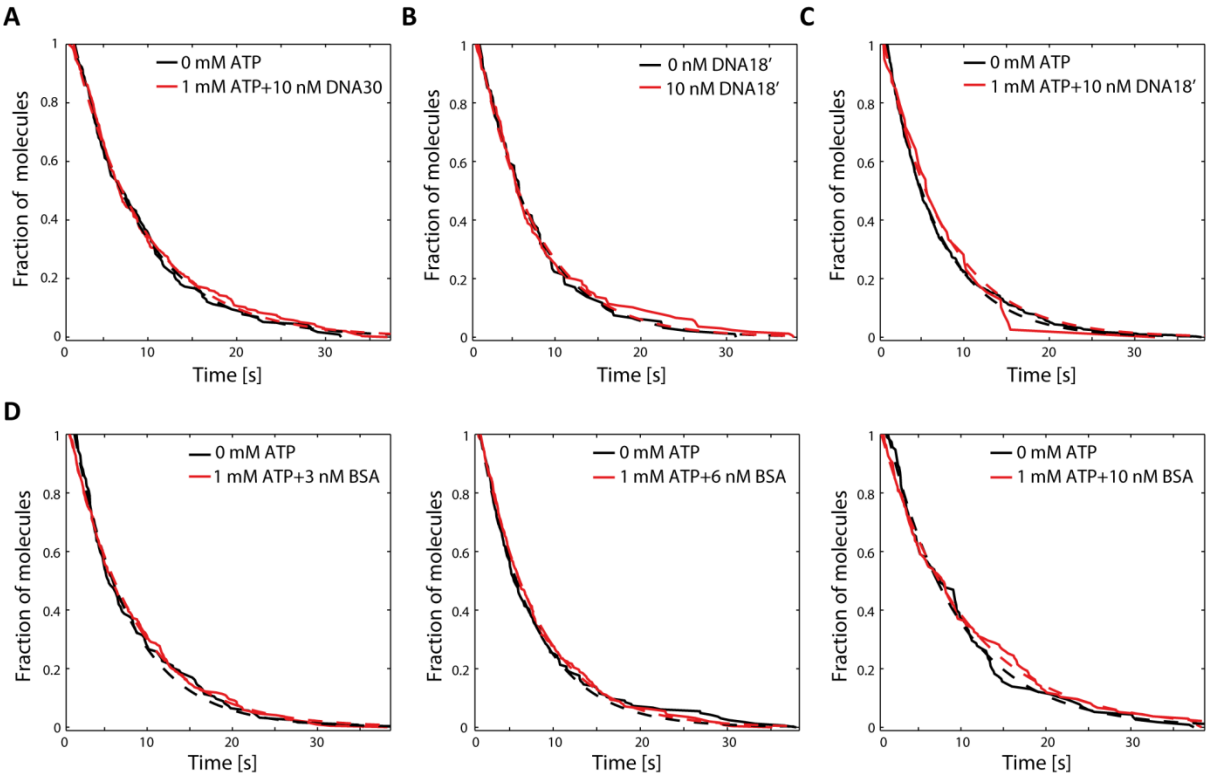
Oligonucleotide	**Sequence (5'→3')**	**Fluorophores/ *Additional Information***
BiP seq167 rev	CTT TCT GGA CTG GCT TCA TGG	
BiP 1-1 seq fw	AGC GGA TAA CAA TTT CAC ACA GG	
BiP 1-2 seq fw	GAA ACT GCT GAA GCA TAT TTG G	
BiP 1-3 seq fw	GAG ACC CTG ACT CGG GC	
Ssc1-UAA448 fw	Phosphate-AAG AAA TCT CAA TAG TTC TCC ACT GCC G	
Ssc1-UAA448 rev	Phosphate-TGT TGG AAT AGT AGT GTT TCT TGG AAT CAA TC	
Ssc1-C590 fw	Phosphate-CCA AAA GGT TAG GTG TCA AAT CAC TTC CT	
Ssc1-C590 rev	Phosphate-GCT TCA GCC TTG TCA ACC TTA CCT TC	
Ssc1-UAA341 fw	Phosphate-GTT AAG AGA ACT GTC TAG CCA GTC AAG AA	
Ssc1-UAA341 rev	Phosphate-TAG TGG GGC TGT CAA AGT CTC GAA TTG	
Ssc1 seq fw	ACT GAC GTC TTA TTA TTA GAT GTT ACC C	
Ssc1 seq rev	GTT TCA CCA GAT TCG GCA	
Ssc1 seq341 rev	TTC AGC GTC GTT AAC CAT TT	
Ssc1 1-1 seq fw	TCC CCT CTA GAA ATA ATT TTG	
Ssc1 1-2 seq fw	CGA AAT TGA ACA AAT GGT TAA C	
Ssc1 2-1 seq fw	TCC CCA TCT TAG TAT ATT AGT	
Ssc1 2-2 seq fw	GTC ACT GTC CCA GCT TAT T	
Ssc1 2-3 seq fw	CAC TGA AAT CAA CCT GCC	
DnaK-UAA425 fw	Phosphate-AAG CAC AGC CAG TAG TTC TCT ACC GCT	
DnaK-UAA425 rev	Phosphate-GGT CGG GAT AGT GGT GTT TTT CGC G	
DnaK-C563 fw	Phosphate-GCT GAC GAC AAA TGT GCT ATC GAG TCT	
DnaK-C563 rev	Phosphate-CGG CAG TTT GTC GCC TGC TTC TTC	
DnaK-UAA318 fw	Phosphate-ACC GTT CCA TTT AGC CGC TGA AAG TTG	
DnaK-UAA318 rev	Phosphate-TTA CCA GAT CTT CAA CCA GGC TTT CCA G	
DnaK-C425 fw	Phosphate-AAG CAC AGC CAG TGT TTC TCT ACC GCT	
DnaK seq fw	CCG CTG TCT CTG GGT ATC	
DnaK seq rev	GAT TTC CAT CAG TTT CTG GGA A	
DnaK mut seq fw	GAC CTG CGC AAC GAT CCG	
DnaK mut seq rev	ATC TAG GTT GAA CTG ACC CAG A	

Additional Materials: Chapter 3



FCS of nucleic acids in the presence of EBNA1. (**A**) Autocorrelation functions of DNA 40mer in the absence and presence of different EBNA1 concentrations and 200 mM NaCl. (**B**) Autocorrelation functions of EBER1-RNA in the absence and presence of 200 nM EBNA1. (**C**) Fluorescence intensity time trace of EBER1-RNA in the presence of different EBNA1 concentrations.

Additional Materials: Chapter 4



Control experiments for the dissociation of TBP-Atto647N molecules bound to DNA18-Atto532 on the prism surface of a TIRF microscope. (**A**) Survival time of DNA/TBP complexes as a sum of photobleaching and dissociation by the addition of ATP and DNA30 (red) compared to the survival time of molecules decreasing due to photobleaching (black). DNA30 did not contain a binding site for TBP. (**B**) Survival time of DNA/TBP complexes as a sum of photobleaching and dissociation by the addition of DNA18' containing a TATA-box sequence (red) compared to the survival time of molecules decreasing due to photobleaching (black). (**C**) Experiments as in (B), but for the addition of a combination of DNA18' and 1 mM ATP. (**D**) Survival time of DNA/TBP complexes as a sum of photobleaching and dissociation in the presence of ATP and different BSA concentrations (red) compared to the survival time due to photobleaching (black).

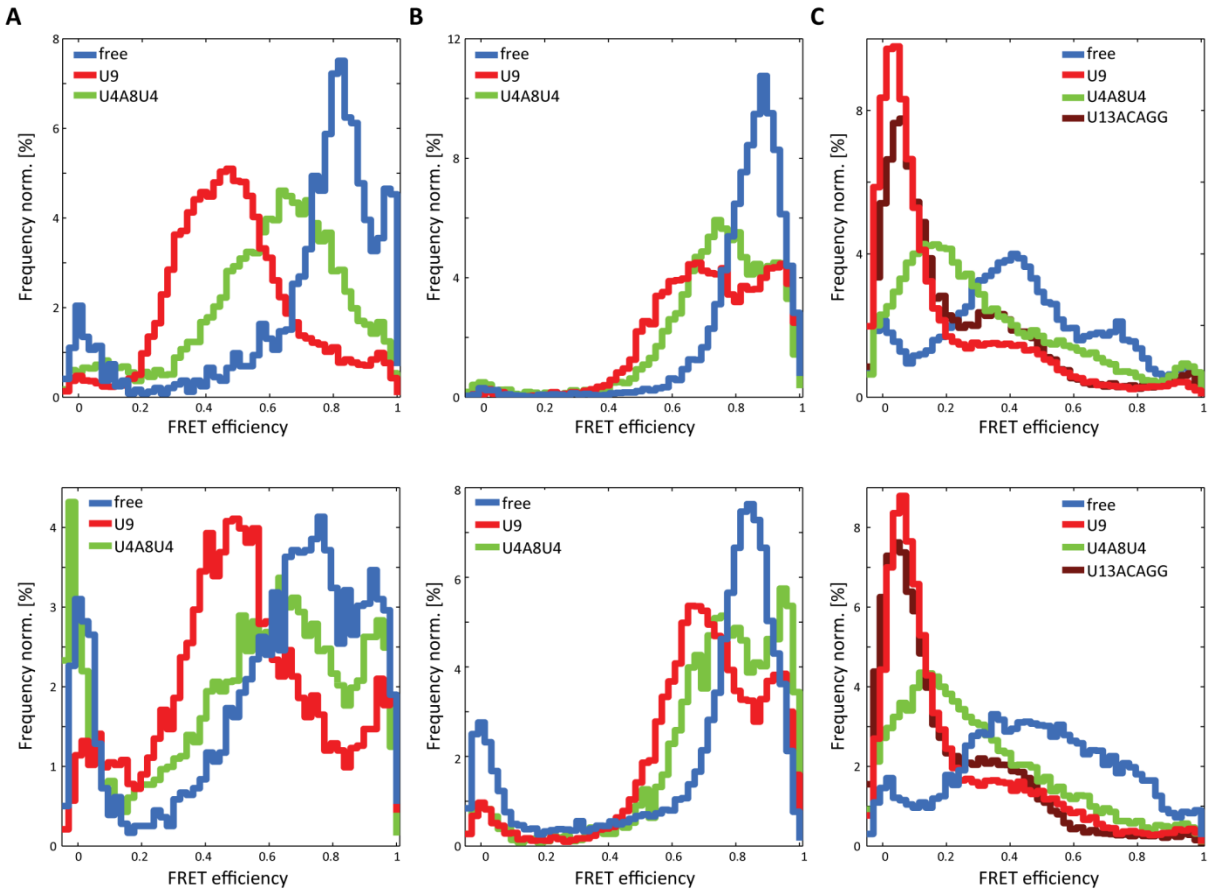
Additional Materials: Chapter 5

Static Photon Distribution Analysis of RRM1,2-Atto532-Alexa647 constructs.

Protein	RNA	A(D(0)) [%]	D1 [Å]	σ [Å]	A1 [%]	D2 [Å]	σ [Å]	A2 [%]
RRM1,2 C187-C318	-	4.4	54	5.8	82	37	4	13
	U9		70	6.9	89	49	7	11
	U4A8U4	7.6	59	9	92			
RRM1,2-Δ233-252 C187-C318	-		44	5.1	100			
	U9		55±2	4.6±0.7	64±9	41±2	5±2	37±9
	U4A8U4		53	5.2	80	41	4	20
URRM1,2/ U2AF35(UHM) C187-C318	-	5.2	63	10.2	93	39	5	2
	U9		86	30.9	91	69	4	9
	U4A8U4		71	19.1	100			
	U13ACAGG		86	32.8	80	69	5	20
RRM1,2 C187-C326	-		53	4.9	93	37	3	7
	U9	2	62	7.6	93	38	4	5
	U4A8U4		58	7.5	96	39	3	4
	U13ACAGG	2.3	67	4.4	80	50	9	18
URRM1,2 C187-C326	-		54	4.7	92	39	4	8
	U9		62	7.3	95	39	5	5
	U4A8U4		59	8.3	97	39	4	3
	U13ACAGG	8.6	64	7.0	89	38	0	2
URRM1,2/ U2AF35(UHM) C187-C326	-	4.5	59	6.0	96			
	U9	2.1	68	6.0	98			
	U4A8U4	2.0	65	8.5	98			
	U13ACAGG	1.9	69	6.1	98			

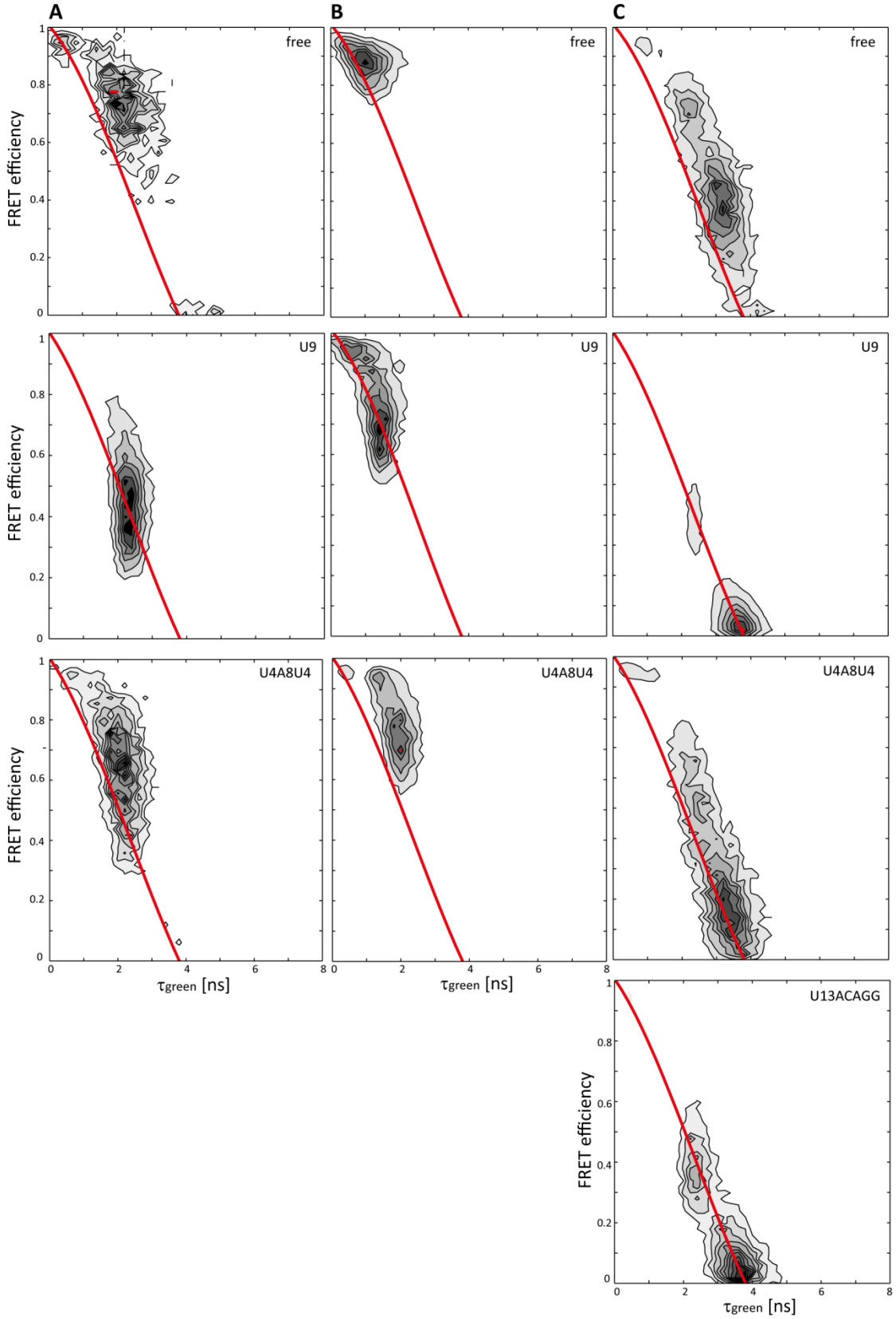
The U2AF65 constructs were measured using MFD-PIE in the absence and presence of different polypyrimidine tracts as indicated. Histograms were created using photons from individual bursts binned into time bins of 1 ms. The resulting histograms were described by a PDA fit consisting of one to three Gaussian distributions as static populations. For each population with amplitude A, the respective distance between the fluorophores (D), and the distribution width (σ) are given. A(D(0)) is a no-FRET population for which the acceptor photobleached during the transit of the molecule through the confocal volume. Standard deviations of the mean for each of the parameters from three independent measurements of RRM1,2-Δ233-252+U9 are given as a rough estimation of the error.

Additional Materials: Chapter 5



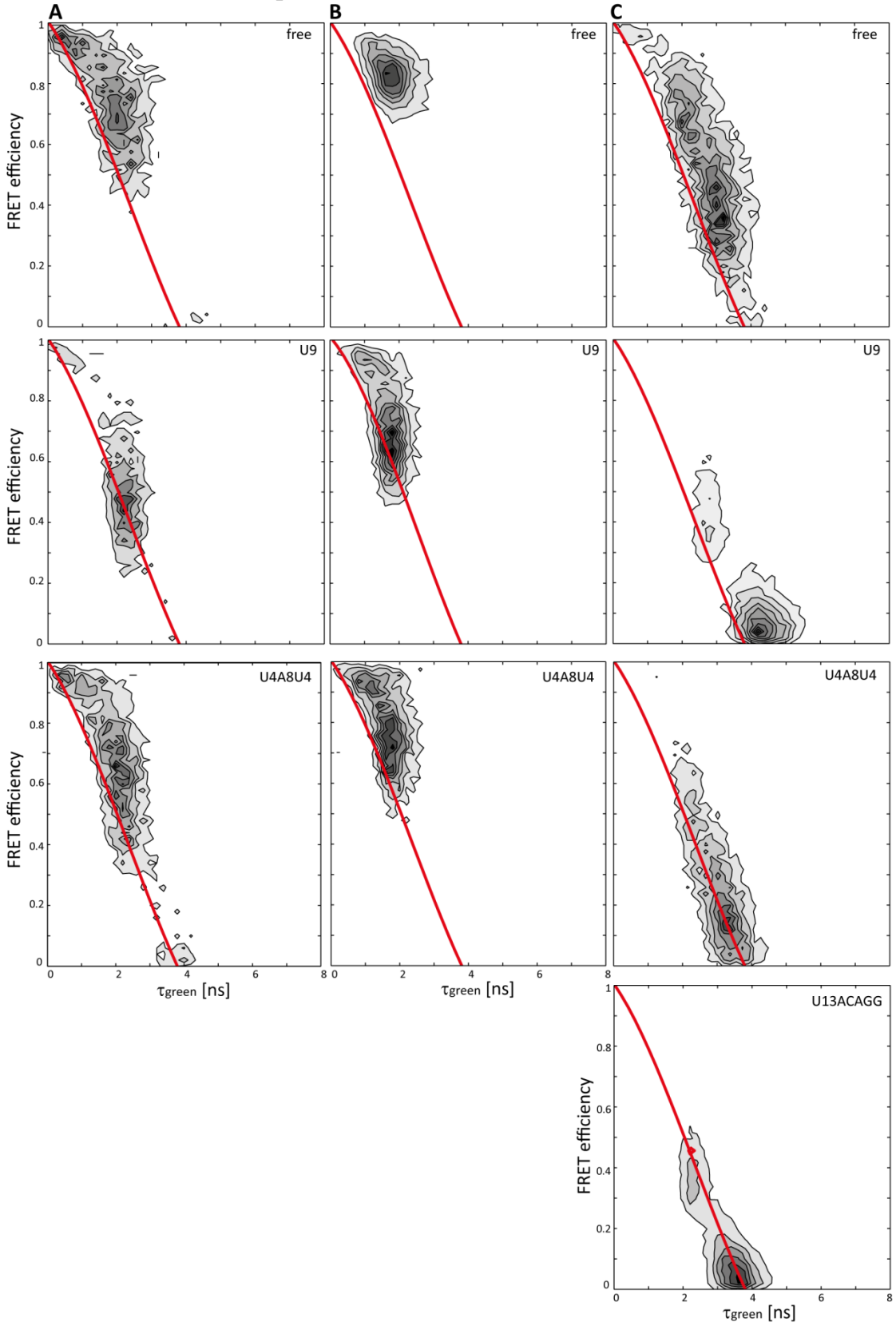
FRET efficiency histograms of RRM1,2-C187-C318 (**A**), RRM1,2-Δ233-252-C187-C318 (**B**), and URRM1,2-C187-C318/U2AF35(RRM) (**C**) labeled with Atto532 and Alexa647 (upper row) or Atto532 and Cy5 (lower row). The constructs were measured in the absence (blue) and in the presence of the RNAs U9 (red), U4A8U4 (green), and U13ACAGG (dark red).

Additional Materials: Chapter 5



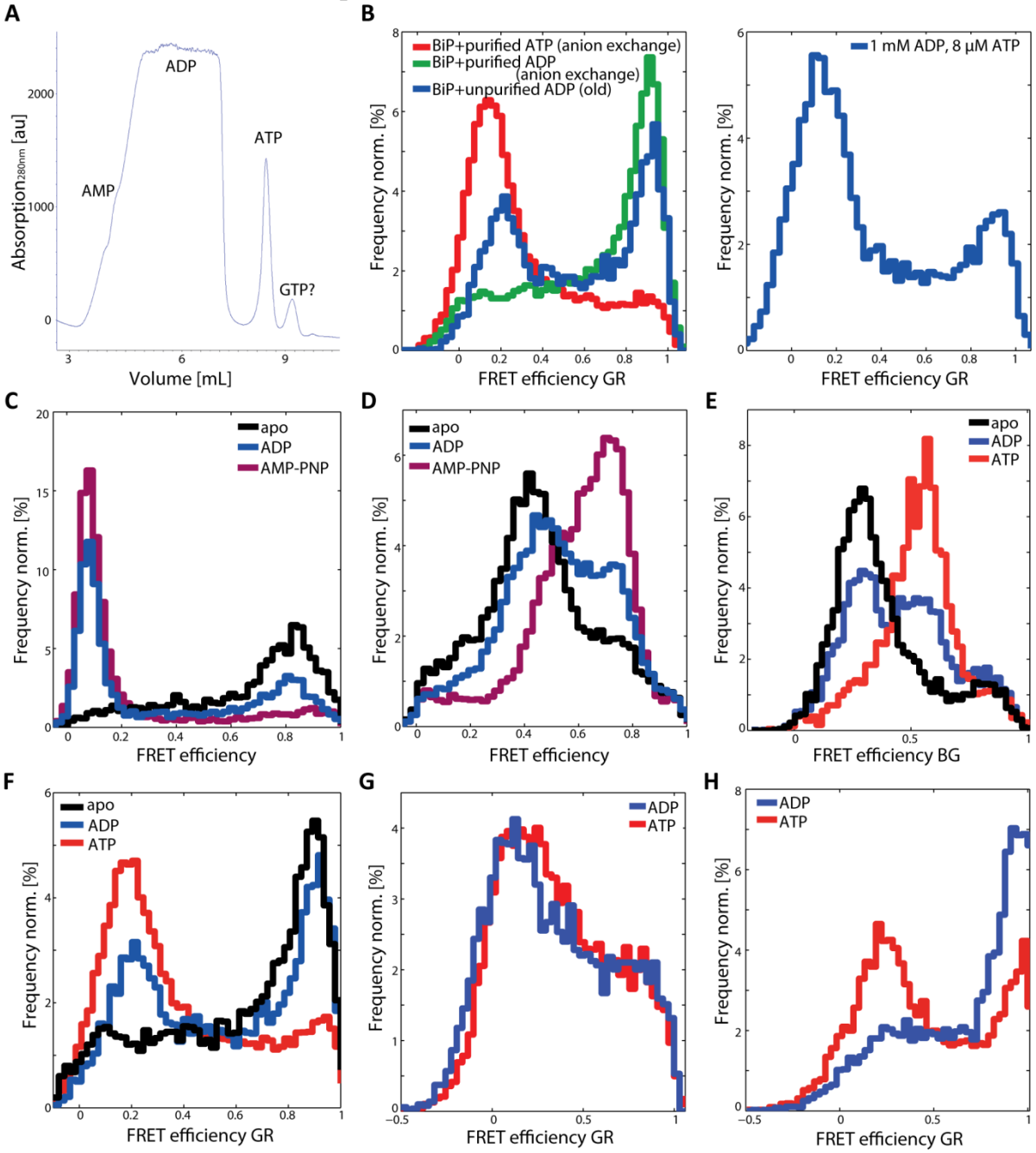
Donor lifetime versus FRET efficiency histograms for RRM1,2-C187-C318 (**A**), RRM1,2-Δ233-252-C187-C318 (**B**), and URRM1,2-C187-C318/U2AF35(RRM) (**C**) labeled with Atto532 and Alexa647 in the absence of RNA (top row) and in the presence of the RNAs U9 (second row), U4A8U4 (third row), and U13ACAGG (bottom row).

Additional Materials: Chapter 5



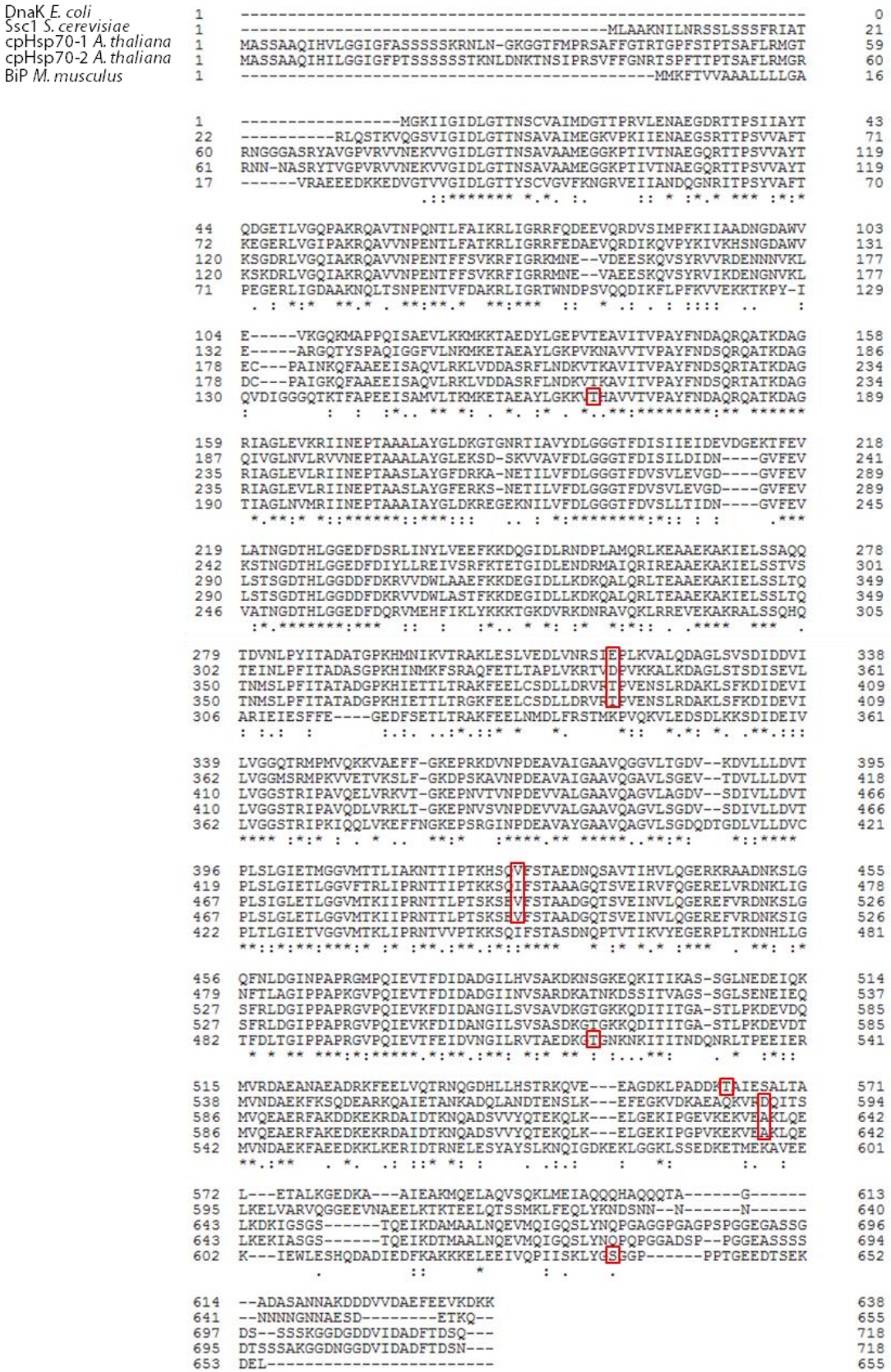
Donor lifetime versus FRET efficiency histograms for RRM1,2-C187-C318 (**A**), RRM1,2-Δ233-252-C187-C318 (**B**), and URRM1,2-C187-C318/U2AF35(RRM) (**C**) labeled with Atto532 and Cy5 in the absence of RNA (top row) and in the presence of the RNAs U9 (second row), U4A8U4 (third row), U13ACAGG (bottom row).

Additional Materials: Chapter 6



Purity of commercially available ADP. (**A**) Purification of ADP by anion exchange chromatography in Hepes buffer with salt concentrations ranging from 0 M to 1 M KCl. ADP and ATP eluted in different fractions. (**B**-**H**) FRET efficiency histograms of dual- and triple-labeled Hsp70 chaperones in the presence of unpurified ADP (blue), purified ADP (green), ATP (red), AMP-PNP (magenta) or in their apo state (black). (**B**) FRET efficiency histogram E_{GR} of BiP-PrK167-C519-C638. FRET efficiency histograms of dual-color BiP-C519-C638 (**C**) and BiP-C167-C519 (**D**). FRET efficiency histogram E_{BG} of BiP-PrK167-C638-Atto488/Atto565 (**E**). (**F**-**H**) FRET efficiency histograms E_{GR} of triple-labeled BiP-PrK167-C519-C639 (**F**), Ssc1-PrK341-C448-C590 (**G**), and DnaK-PrK318-C425-C563 (**H**).

Additional Materials: Chapter 6



Sequence alignment of the Hsp70 chaperones DnaK (*E. coli*), Ssc1 (*S. cerevisiae*), cpHsp70-1 (*A. thaliana*), cpHsp70-2 (*A. thaliana*), and BiP (*M. musculus*). The labeling positions are displayed in red rectangles.

Acknowledgements

My supervisor, Prof. Don C. Lamb, greatly influenced my scientific work during the last years. He offered me the opportunity to discover a completely new field of research. I thank him for expanding my scientific horizons, his advice, and his patience throughout the years.

Prof. Christoph Bräuchle allowed us to share his laboratories with their excellent infrastructure. Thanks for interesting and inspiring discussions in our combined seminars and for the support as a second advisor.

The investigations that were part of the many projects underlying this thesis were only possible with the help of our collaboration partners. Thanks to Dr. Aloys Schepers, Christoph-Erik Meyer and Julia Damaschke, who prepared the proteins HMGA1 and EBNA1 and shared their wisdom. Thanks to Aloys for reading and commenting on part of my thesis. Prof. David Auble, Prof. Michael Meisterernst, and Ramya Viswanathan prepared proteins and shared their protocols for the TBP project. Many thanks to Prof. Michael Sattler, Carolina Sanchéz Rico, Lisa Warner, and Hyun-Seo Kang for providing proteins and NMR data for the U2AF projects and for our countless discussions. My work on the Hsp70 chaperones was facilitated by Dr. Edward Lemke, Swati Tyagi, and Christine Köhler, who introduced me to expression with non-natural amino acids and supported me with helpful advice. My thanks also to Dr. Dejana Mokranjac, Prof. Johannes Buchner, and Dr. Mathias Rosam for providing plasmids for the three Hsp70 chaperones. With Prof. Matthias Mayer and Roman Kytik we collaborated on the Sigma32 project.

I was very lucky to spend these important years in such a stimulating and supportive work environment. With Ellen, I didn't only find a wise colleague but also a wonderful friend. I do not know, what I would have done without my great office mate Waldi with all his helpful advice, critical comments, discussions, and kind words. I will always remember our telephone conferences to align the Toptica laser. Many thanks to Gregor and Martin, who introduced me to TIRF experiments with TBP and Ssc1. Great thanks go to Niko, who helped me to discover the trickiness of HMGA1. Thanks for creating such a great, cooperative atmosphere in the lab. Thanks to Sushi for many great discussions about the TIRF microscope and three-color TIRF experiments, as well as Indian and German culture. Regarding three-color FRET experiments I am grateful to Anders, who reconstructed the confocal microscope and programmed the analysis software, for countless discussions about different analysis programs. Daniela, it was great to share all the ups and downs of the old MFD and a lot of organizational work with you. Fabian, thanks for being the good soul who keeps the lab and especially all the computers running. Thanks to Jens for his cooperativeness in all respects and for keeping pace with me on the way to lunch. Philipp, our boulder king, challenged me in many difficult routes. Ganesh, it's great to discuss with you your passion for science, Indian culture, and haircuts. Many thanks to all the postdocs in the lab, who shared their advice and wisdom. Aurélie, hiking to Lengrieser Hütte with Mona is an unforgotten event. Giulia, thanks for

sharing your wonderful spirit and smiles. Viola, thanks for always being a great source of biochemical knowledge. Evelyn, I'm glad that you are going to continue with some of my projects. Having in mind Jelle's strong will, I remember our first walk around Nymphenburg castle, where we decided, that he would become a great professor. Many thanks to Alex for his open ear, kindness, and always offering a helping hand. Sharing the office with Alvaro opened my eye to many aspects of scientific research and German bureaucracy. Over the last few years lots of new faces appeared in the group. Thanks to all the new graduate students, Maria, Kira, Ivo, Bässem, Nader, and Chen for continuing in the lab with the same friendly spirit.
Ellen, Vroni, Adriano, and Leonhard, the last survivors of Prof. Bräuchle's original group, created a wonderful work environment. Thanks to Prof. Michaelis' group (Julia, Moni, Thilo, and all the others) who created a wonderful working environment in the biolab during the early days of my PhD.
I cannot thank Dr. Moritz Ehrl and Silke Steger enough for taking care of all the organizational work, which made our studies here possible and also for Moritz's many amusing stories.
To the people underground (Michael Gayer, Heidi Buchholz and many more), the electronic and mechanical workshop, and the cleaning personnel, your efforts to ease our lives was not unnoticed and is very much appreciated.

My students Sigurd, Benjamin, Laura, and Franziska were a great motivation for me. It was fun to work together and I hope I contributed to your learning. The students I had the pleasure of teaching in courses helped create an atmosphere of excitement about science and discovery.

I have saved the biggest thanks of all to my wonderful and loving parents and my sister without whom none of this would have been possible. Their perpetual worried voices of advice "please, don't work so hard" have followed me throughout my whole education. Thanks for carrying me through the positive and the difficult times. Thank you, Britta, Jule, Daniel, Noggi, and Sebastian for filling my days with joy and happiness. Sebastian, thanks for sharing good and bad times, always teasing out a smile. Thanks to all the others in my family and my great friends from all over the world, especially my Canadian mom, for all their support and patience.