Domain Orientation in the RNA Helicase YxiN and the Role of Conformational Changes for RNA Unwinding

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Anne R. Karow

aus Gotha, Deutschland

Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Dagmar Klostermeier
Prof. Dr. Stephan Grzesiek

Basel, den 27.04.2010   Prof. Dr. Eberhard Parlow
                                                        Dekan

Abstract

The RNA helicase YxiN from *Bacillus subtilis* is a member of the family of DEAD box proteins. YxiN is able to unwind RNA double strands in an ATP-dependent manner. The ability to catalyse RNA rearrangement is *in vivo* presumably necessary for the bacterial ribosome biogenesis.

YxiN comprises a two-domain helicase core region and a C-terminal RNA binding domain. While crystal structures of the C-terminal core domain and the RNA binding domain separately have been determined before, the structure of full-length YxiN is not known. In the current project the orientation of these three domains to each other was determined employing fluorescence resonance energy transfer (FRET) experiments at the single-molecule level. Therefore the approximate architecture of the full-length enzyme in solution can now be described. The two core domains exhibit a conformation similar to the crystal structure of the DEAD box protein MjDeaD. The RNA binding domain is adjacent to the C-terminal core domain. Presumably the central β-sheet of the RNA binding domain faces towards a patch of the core domain that is formed by loops.

During catalysis YxiN undergoes a conformational change. The conformation of the core domains mentioned above is adopted in the absence of substrates and in the presence of RNA, ADP, ATP or ADPNP. In the presence of both RNA and ATP (or ADPNP) the core domains approach each other constituting a closed conformation. During the catalytic cycle this conformational change takes place initially after binding of RNA and ATP. The conformational change is necessary for RNA unwinding. But it is not sufficient since the YxiN_K52Q mutant adopts the closed conformation upon binding of RNA and ATP (or ADPNP) but is unwinding deficient.

Transitions between the open and the closed conformation could only rarely been detected in the FRET experiments on a confocal microscope due to a limited observation time. To be able to monitor the conformation of YxiN on longer timescale the protein was engineered for FRET experiments on a total internal reflection microscope. A protocol was developed that comprises fluorophore double labelling of YxiN and the attachment of a biotin at the protein's C-terminus. The biotinylation procedure is based on the reaction type of expressed protein ligation. The labelled and biotinylated YxiN construct could be specifically immobilized on a streptavidin coated surface for total internal reflection microscopy. Subsequently, YxiN could be monitored for up to a few seconds.

Expressed protein ligation was furthermore employed to develop a specific fluorophore double labelling strategy for FRET experiments. Employing this strategy a YxiN construct could be generated that carries one certain fluorophore exclusively at one position in the protein. A different fluorophore can attach to the same position or to one further site within the protein. The procedure was therefore termed semi-site-specific double labelling. In comparison with statistic labelling procedures the semi-site-specific double labelling allows for decreasing the sample heterogeneity in FRET experiments.

Taken together, this study revealed the conformation of the three-domain RNA helicase YxiN, its conformational change during catalysis which is essential for the activity of the helicase and the study established protein preparation techniques that provide the basis for further studies on the helicase mechanism.

Table of Contents

1 Introduction

1.1 DEAD Box Proteins Catalyse the Structural Rearrangement of RNA

1.1.1 RNA Folding

RNA is a biomolecule of enormous functional diversity. In all cells RNA molecules transport and transduce genetic information during transcription and translation. Several viruses employ RNA for the storage of genetic information. Moreover, small RNA molecules, such as small interfering RNAs or micro RNAs take part in cellular regulation processes. RNA molecules also can provide scaffolding function. For instance, the ribosome comprises 60 % RNA by weight (Moore & Steitz, 2002). This ribosomal RNA largely scaffolds the structure of the ribosome. Notably, also the ribosome's catalytic site is entirely composed of RNA. Nucleobases of the large ribosomal subunit catalyse the peptide bond formation. Furthermore, certain RNAs act as enzymes in RNA processing events and in RNA splicing.

Its variety in function is based on the ability of RNA to fold into a great plenty of architectures. As for proteins, one can distinguish interactions between the building blocks of RNA on the level of the primary sequence, on the secondary structure and on the level of tertiary interactions. Two or more RNAs might furthermore interact and form complexes of quaternary structure as in the ribosome.

On the level of secondary structure the nucleobases mostly interact via base pairing employing hydrogen bonds. The interactions are strong and secondary structure elements are cemented by base stacking between the nucleobases. In RNA structures the following basic elements can be distinguished: helix forming double stranded segments, single stranded loop regions and bulges of two RNA strands adjacent to helices as well as junctions of the mentioned elements. On the next level long-range tertiary contacts are established that fold the RNA molecule in a three-dimensional structure.

In comparison with the tertiary interactions the secondary structure elements are highly stable. The secondary structure elements are nearly independent of the tertiary structure. In most cases, isolated stretches adopt "their" secondary structure without the context of the whole molecule. If misfolding occurs on the level of secondary structure, the corresponding non-native structure is often long-lived and not corrected by the influence of tertiary contacts. This property is described as the kinetic folding problem of RNA (Herschlag, 1995). In addition, a thermodynamic problem arises from the propensity of RNA to be able to form a variety of tertiary contacts that each have a low and comparable contribution to the overall folding energy. Herschlag delineated that RNA is intrinsically prone to misfolding (Herschlag, 1995).

RNA misfolding was already observed in the 1960s (Gartland & Sueoka, 1966; Lindahl *et al.* 1966). In studies on a tRNA molecule two alternative conformers were detected. One of the structures could not be charged by the cognate aminoacyl-tRNA synthetase. This biologically inactive structure seemed to be kinetically trapped since it could be reverted to the active form via heating. In proceeding studies the inactive conformer was characterised as a slightly extended form of the tRNA that misses a few stacking interactions compared to the native molecule (Madore *et al.*, 1999).

Following up Gartland, Lindahl and colleagues, kinetically trapped conformers were also detected for the 5 S ribosomal RNA (Aubert *et al.*, 1968), for viroids (Riesner *et al.*, 1979) and for RNase P RNA (Altman & Guerrier-Takada, 1986). Nowadays the folding of intron RNA and the assembly of the ribosomal subunits is of major interest to the community.

In 1975 Karpel and colleagues reported that the protein UP1 is able to renature the metastable tRNA conformer discussed above (Karpel *et al.*, 1975, Karpel & Burchard, 1980). UP1 exhibits a high affinity to single stranded DNA and RNA and could therefore favour the dissociation of double strands. In case a non-native double strand had formed during RNA folding, UP1 can promote the dissociation of the strands. Thereby the inactive RNA conformer is released from its kinetic trap. Based on this observations a general role of RNA renaturing proteins in RNA physiology was suggested. Later on more proteins that accelerate rearrangement of RNA structures were discovered. Based on their mode of action they can be subdivided into two major groups. Including UP1, the members of the first group exhibit high affinity to RNA single strands, yet bind RNA with low or no specificity. They constitute usually positively charged proteins and often do not dissociate from the natively folded RNA. As a result they become part of RNA-protein complexes such as the proteins of ribonucleoprotein particles. Also some ribosomal proteins fall into this group. Notably, these ribosomal proteins are even able to assist the folding of non-ribosomal RNA *in vitro* (Coetzee *et al.*, 1994).

The members of the second group, in contrast, usually dissociate from their RNA substrate and, more strikingly, function in an ATP-dependent manner. They constitute the so-called family of DExD/H box proteins. These proteins share a common core domain structure and various characteristic sequence motifs such as the name giving sequence D-E-x-D/H. They have been shown to catalyse a variety of RNA structural transitions including the displacement of proteins from RNA strands (Jankowsky *et al.*, 2005; Yang & Jankowsky, 2005; Linder, 2006; Halls *et al.*, 2007). A subset of these RNA dependent ATPases is the family of DEAD box proteins that were analogously named after a common sequence motif. Characteristically, these proteins are able to catalyse the rearrangement of only short RNA stretches. For some cases specific substrates were identified.

While prokaryotes possess only a handful of different DEAD box proteins (e.g. 4 in *Bacillus subtilis*) the number markedly increases for eukaryotes (e.g. 25 in *Saccharomyces cerevisiae*; both data from the DExH/D database 2.0). In eukaryotes the presence of cellular compartments and an increased repertoire in regulation mechanisms require plenty of RNA reorganisations. After transcription every mRNA is capped and a polyA tail is attached. Eukaryotic mRNAs are usually spliced and some are even edited via deamination. Subsequently the mRNAs need to be transported out of the nucleus. All of these steps involve structural rearrangements of RNA that are mostly accompanied by DEAD box proteins. In addition, these proteins are involved in eukaryotic translation and its regulation.

The structure and mechanism of DEAD box proteins in general and in particular of the *Bacillus subtilis* DEAD box protein YxiN will be the subject of the following chapters.

1.1.2 Article:

The Mechanism of ATP-dependent RNA Unwinding by DEAD Box Proteins

Review

The mechanism of ATP-dependent RNA unwinding by DEAD box proteins

Manuel Hilbert, Anne R. Karow and Dagmar Klostermeier*

Biozentrum, Biophysical Chemistry, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

* Corresponding author
 e-mail: dagmar.klostermeier@unibas.ch

Abstract

DEAD box proteins catalyze the ATP-dependent unwinding of double-stranded RNA (dsRNA). In addition, they facilitate protein displacement and remodeling of RNA or RNA/protein complexes. Their hallmark feature is local destabilization of RNA duplexes. Here, we summarize current data on the DEAD box protein mechanism and present a model for RNA unwinding that integrates recent data on the effect of ATP analogs and mutations on DEAD box protein activity. DEAD box proteins share a conserved helicase core with two flexibly linked RecA-like domains that contain all helicase signature motifs. Variable flanking regions contribute to substrate binding and modulate activity. In the presence of ATP and RNA, the helicase core adopts a compact, closed conformation with extensive interdomain contacts and high affinity for RNA. In the closed conformation, the RecA-like domains form a catalytic site for ATP hydrolysis and a continuous RNA binding site. A kink in the backbone of the bound RNA locally destabilizes the duplex. Rearrangement of this initial complex generates a hydrolysis- and unwinding-competent state. From this complex, the first RNA strand can dissociate. After ATP hydrolysis and phosphate release, the DEAD box protein returns to a low-affinity state for RNA. Dissociation of the second RNA strand and reopening of the cleft in the helicase core allow for further catalytic cycles.

Keywords: ATP hydrolysis; conformational changes; coupling; helicase; protein dynamics; RNA duplex destabilization.

Introduction

Helicases unwind double-stranded nucleic acids in an ATP-dependent manner (Pyle, 2008). DNA helicases unwind double-stranded DNA regions and are involved in DNA replication, recombination and repair and overall genome stability (Singleton et al., 2007). RNA helicases, in contrast, remodel RNA or RNA/protein complex structures in all processes that involve RNA (Cordin et al., 2006). According to the occurrence of conserved sequence motifs, helicases are grouped into up to five superfamilies (Gorbalenya and Koonin, 1993). RNA helicases are part of the helicase superfamily 2 (SF2), which comprises the DExD/H proteins, the RecQ family, and members of the SWI/SNF family. DExD/H proteins are further divided into the DEAD, DEAH, and DExH family, according to the sequence in the corresponding signature motif. DExD/H proteins share a common domain organization but differ with respect to function and mechanism. In this review, we will focus on DEAD box proteins, which constitute the largest subfamily of SF2 helicases. Members of this family share a ~400 amino acid core region comprising highly conserved helicase signature motifs (Figure 1). Among them is the name giving DEAD box, a conserved motif with the amino acid sequence DEAD. The frequently used designation 'DEAD box helicases' refers to the ability of many but not all members of the DEAD box family to unwind RNA duplexes at the expense of ATP hydrolysis. Although for some time RNA unwinding has been believed to be the hallmark reaction catalyzed by all DEAD box proteins, unwinding activity was only demonstrated for some representatives, and in fact other DEAD box proteins have been implicated in a variety of functions such as remodeling structures of RNA or RNA/protein complexes (Cheng et al., 2005; Linder, 2006), dissociating RNA/protein complexes (Fairman et al., 2004) or RNA annealing (Yang and Jankowsky, 2005; Halls et al., 2007; Rajkowitsch et al., 2007). In these processes, they facilitate the (local) disruption of secondary and tertiary structure and of RNA-protein interactions. As such they are involved in all facets of RNA metabolism, from transcription, mRNA splicing and translation, RNA modification, and transport, ribosome biogenesis and RNA/protein complex assembly in general, to RNA degradation (Cordin et al., 2006).

DEAD box proteins share a common architecture. Their 'helicase core' consists of two independent RecA-like domains, connected by a short, flexible linker (Figures 1 and 2). In all DEAD box protein structures determined to date, the conserved helicase motifs are distributed in a similar three-dimensional arrangement on the RecA-like domains. The N-terminal RecA-like domain provides the structural scaffold for the motifs I–III, and the C-terminal RecA-like domain harbors motifs IV, V, and VI. Mutations in these helicase motifs affect nucleotide binding, ATP hydrolysis, RNA binding, or unwinding. A few mutations have been reported that uncouple ATP hydrolysis from RNA binding and unwinding without affecting binding of ATP or RNA (Pause and Sonenberg, 1992; Cheng et al., 2005; Sengoku et al., 2006; Karow and Klostermeier, 2009). From extensive mutational studies on several

Figure 1   Conserved motifs in the DEAD box helicase core and their contributions to nucleotide binding and hydrolysis, RNA binding, and interdomain contacts.

DEAD box proteins share a conserved core that consists of two RecA-like domains (gray) and can be flanked by variable N- and C-terminal extensions. The helicase core contains all helicase signature motifs. Consensus sequences are derived from an alignment of 920 DEAD box proteins from reviewed sequences in the UniProt database (UniProt-Consortium, 2008), using the presence of a DEAD box plus a Q-motif as a selection criterion. A single amino acid is given in the consensus if this amino acid was found in >65% of the sequences. The two most abundant amino acids are given if their combined occurrence is >70%. Positions with less conservation are marked with x.
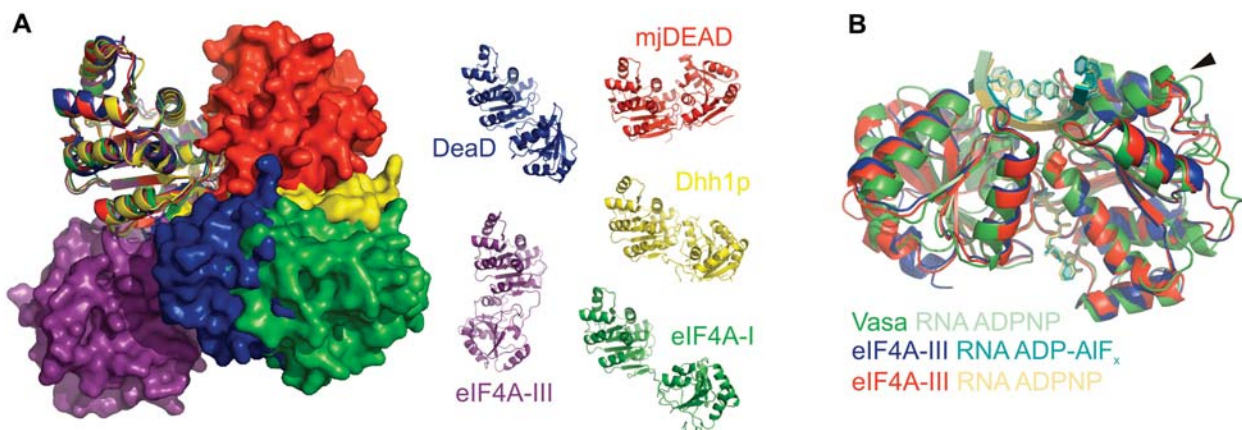
Figure 2   Open and closed conformations of DEAD box proteins.
(A) Different conformations of the DEAD box helicase core in the absence of ligands. The crystal structures of mjDeaD (*Methanococcus jannaschii*, red, PDB-ID 1hv8), Dhh1p (*Saccharomyces cerevisiae*, yellow, PDB-ID 1s2m), eIF4A-I (*S. cerevisiae*, green, PDB-ID 1fuu), DeaD (*Sulfolobus tokodaii*, blue, PDB-ID 2z0m), and eIF4A-III (human, purple, PDB-ID 2hxy) were superimposed on the N-terminal RecA-like domain, and the C-terminal domain is shown in surface representation. (B) Similar conformations of the helicase core of eIF4A-III and Vasa in the presence of ssRNA and ADPNP or ADP-AlF$_x$. The structure of Vasa (*Drosophila melanogaster*, PDB-ID 2db3, green) in complex with RNA and ADPNP (light green) is superimposed with eIF4A-III (human) in complex with ADPNP and RNA (blue/cyan) or RNA and ADP-AlF$_x$ (red/yellow). The overall structures of these three complexes are highly similar, including the positions and conformations of bound nucleotide and RNA. The triangle marks helix α7 that causes a kink in the bound RNA (see close-up in panel B).

DEAD box proteins, a consensus on the individual functions of the conserved helicase motifs has been obtained (Figure 1). In general, motifs I and II are implicated in

ATP binding and hydrolysis, with contributions from motif VI. Motif III is believed to couple ATP hydrolysis to RNA unwinding. Motifs IV, V, and VI contribute to RNA

binding. Details on the specific function of conserved motifs will be discussed in the following sections.

Notably, mutations in conserved motifs can have different effects in different DEAD box proteins, indicating that the specific structural context influences the significance of a motif. Structural studies have provided a first glimpse on the individual interactions between the conserved motifs in the absence or presence of ATP and ssRNA substrate (Caruthers et al., 2000; Story et al., 2001; Cheng et al., 2005; Andersen et al., 2006; Bono et al., 2006; Sengoku et al., 2006; Nielsen et al., 2008; Collins et al., 2009; von Moeller et al., 2009). When both substrates are bound, the helicase motifs engage in a complex network with a multitude of cooperative interactions, complicating the assignment of functional contributions for each individual motif in mutagenesis studies (Banroques et al., 2008). The complete interaction network and its role for DEAD box protein activity are thus not yet fully understood.

In the absence of nucleotide or RNA substrate, the two RecA-like domains in the helicase core do not interact. Crystal structures of various DEAD box proteins without ligands show a large variety of relative orientations of the two RecA-like domains (Figure 2A; Caruthers et al., 2000; Story et al., 2001; Cheng et al., 2005; Andersen et al., 2006) and point to a significant flexibility that is presumably provided by the linker region. Solution studies of the *Bacillus subtilis* DEAD box protein YxiN are in agreement with such a high flexibility. Small angle X-ray scattering (Wang et al., 2007) and single molecule FRET experiments (Theissen et al., 2008) show an extended conformation in solution, consistent with an open conformation of the helicase core. No nucleotide binding is observed when the two RecA-like domains are mixed without a covalent linkage (Karow et al., 2007), confirming that interactions between these domains are weak. In contrast to the open conformation in the absence of ligands, the helicase core of DEAD box proteins adopts a compact, closed conformation in the presence of ssRNA and nucleotide, as exemplified by the crystal structures of the DEAD box proteins eIF4A-III, Vasa, and Ddx19 in complex with ssRNA and the non-hydrolyzable ATP-analog ADPNP (Andersen et al., 2006; Bono et al., 2006; Sengoku et al., 2006; Collins et al., 2009; von Moeller et al., 2009; Figure 2B). In these structures, most of the conserved motifs face the interdomain cleft between the RecA-like domains and are engaged in a complex hydrogen bond network. The nucleotide is buried at the bottom of the interdomain cleft and interacts with residues from both domains. Similarly, the ssRNA contacts both RecA-like domains and stabilizes the closed conformation. The transition from the flexible open conformation to the more rigid closed conformation is only observed when both ATP and RNA are present (Theissen et al., 2008; Karow and Klostermeier, 2009). Thus, it appears to be an important checkpoint in the catalytic cycle of DEAD box proteins. The similar architecture of the closed conformations between DEAD box proteins from different organisms points to a common mechanism for the helicase core. In contrast to the similar arrangement of the two domains, the connecting linker region adopts different conformations in the closed

structures of eIF4AIII, Vasa, and Ddx19. Stepwise conversion of the eIF4A linker to the corresponding Vasa sequence resulted in enhanced ATPase activity, revealing the linker as a regulatory element (Low et al., 2007).

In all crystal structures, the phosphate backbone of the bound ssRNA shows a sharp bend in the same position that prevents canonic A-form RNA geometry. It was postulated that the formation of the closed conformer is coupled to this distortion of the RNA and thereby facilitates double-strand destabilization (Andersen et al., 2006; Bono et al., 2006; Sengoku et al., 2006). However, further experimental evidence for the distortion is lacking and the kink is observed in complex with ADPNP and ADP-AlF$_x$, nucleotides that do not allow for unwinding.

The helicase core on its own is an RNA-stimulated ATPase and has the ability to unwind short RNA duplexes (Rogers et al., 1999). Its affinities for RNA are regulated by the nucleotide state and it thus functions as a nucleotide-dependent molecular switch (Lorsch and Herschlag 1998a,b). However, only few examples of DEAD box proteins consist of an isolated helicase core, among them the translation initiation factor eIF4A, or the *Methanococcus jannaschii* DeaD protein. In numerous DEAD box proteins, the helicase core provides the basic DEAD box protein functions, and large N- and/or C-terminal extensions (Figure 1) modulate the activity of the helicase core by conferring substrate specificity or by mediating contacts with interacting proteins (Schmid and Linder 1992; see 'Modulation of the helicase core activity by interacting partners and flanking domains'). In addition, the helicase core appears as a module in several enzymes involved in nucleic acid processing, such as DNA topoisomerases (Confalonieri et al., 1993; Rodriguez and Stock, 2002), restriction enzymes (Gorbalenya and Koonin, 1991; Szczelkun, 2000), chromatin remodeling enzymes (Flaus and Owen-Hughes, 2001), or Dicer (Ma et al., 2008).

This year marks the 20th anniversary of the 'Birth of the DEAD box' (Linder et al., 1989). Despite a wealth of information gained during two decades, the catalytic mechanism of DEAD box proteins is not yet fully understood. Open questions include: What is the role of the conformational change in the helicase core for strand displacement? What is the conformation of the RNA substrate bound to the helicase core? How is ATP hydrolysis coupled to RNA unwinding, and at which stage does ATP hydrolysis occur? Which step of the nucleotide cycle triggers RNA release? What are the functions of additional domains and interacting proteins? In this review, we summarize the current knowledge on mechanistic aspects of DEAD box protein activity.

## Interaction of DEAD box proteins with adenine nucleotides and RNA

The interactions of DEAD box protein with adenine nucleotide and RNA substrates have been characterized using a variety of equilibrium and non-equilibrium methods. Nucleotide and RNA binding have been widely studied in crosslinking and filter binding experiments (Pause and Sonenberg, 1992; Pause et al., 1993; Lorsch and Hers-

chlag, 1998a; Tanner et al., 2003; Sengoku et al., 2006). These experiments are not performed under equilibrium conditions, but nevertheless they have been very valuable for comparing relative affinities of one DEAD box protein for different nucleotides, or for comparing nucleotide or RNA affinities of different DEAD box protein mutants. Nucleotide binding constants have also been derived from 'equilibrium filtration' experiments (Polach and Uhlenbeck, 2002), and RNA binding constants have been determined in electrophoretic mobility shift assays (Pause et al., 1993; Henn et al., 2001; Polach and Uhlenbeck, 2002; Cordin et al., 2004; Karginov et al., 2005; Talavera et al., 2006; Banroques et al., 2008; Liu et al., 2008). In these experiments, the equilibrium is only minimally perturbed. Equilibrium dissociation constants can be obtained directly in fluorescence titrations of fluorescently labeled nucleotides or RNA. Nucleotide binding can be investigated via fluorescence of mant nucleotides (Henn et al., 2002, 2008; Talavera and De La Cruz, 2005; Karow et al., 2007; Theissen et al., 2008; Karow and Klostermeier, 2009), and RNA binding can be monitored via fluorescence anisotropy using fluorescently labeled RNAs (Karow et al., 2007; Marintchev et al., 2009). Fluorescence spectroscopy has also been employed to directly determine rate constants for nucleotide binding and dissociation (Henn et al., 2002, 2008) and for RNA unwinding (Karow et al., 2007). The fact that DEAD box proteins are RNA-stimulated ATPases provides a further means to characterize their interactions with RNA and ATP (Lorsch and Herschlag, 1998a,b; Tsu and Uhlenbeck, 1998; Kossen and Uhlenbeck, 1999; Tsu et al., 2001; Polach and Uhlenbeck, 2002; Karow et al., 2007; Henn et al., 2008; Theissen et al., 2008; Karow and Klostermeier, 2009). Michaelis-Menten constants for ATP in the absence and presence of RNA substrates can provide a measure for the corresponding dissociation constants. Likewise, a 'Michaelis-Menten' constant for RNA in the presence of ATP can be determined that represents an apparent dissociation constant of DEAD box protein/ RNA complexes and, although not a genuine dissociation constant, allows for rapid comparison of RNA binding to different mutants and different DEAD box proteins. The interpretation of steady-state ATPase data is complicated by the fact that the nucleotide state under the experimental conditions depends on the nature of the rate-limiting step in the nucleotide cycle. Although all cited methods provide valuable insights on enzyme-substrate interaction, results obtained by different methods have to be compared with care, and in some cases comparison might be difficult or even impossible.

In contrast to the moderate RNA affinity of the DEAD box helicase core, flanking domains can bind RNA with high affinities. When overall RNA affinities are determined, nucleotide-dependent substrate binding properties of the core will thus be masked by the nucleotide-independent high RNA affinity of the flanking domain, and it is crucial to dissect the individual contributions. Recent evidence suggests that additional domains can also influence nucleotide binding (Collins et al., 2009; Fan et al., 2009; Napetschnig et al., 2009; von Moeller et al., 2009), see also the section on 'modulation of the helicase core activity by interacting partners and flanking domains'.
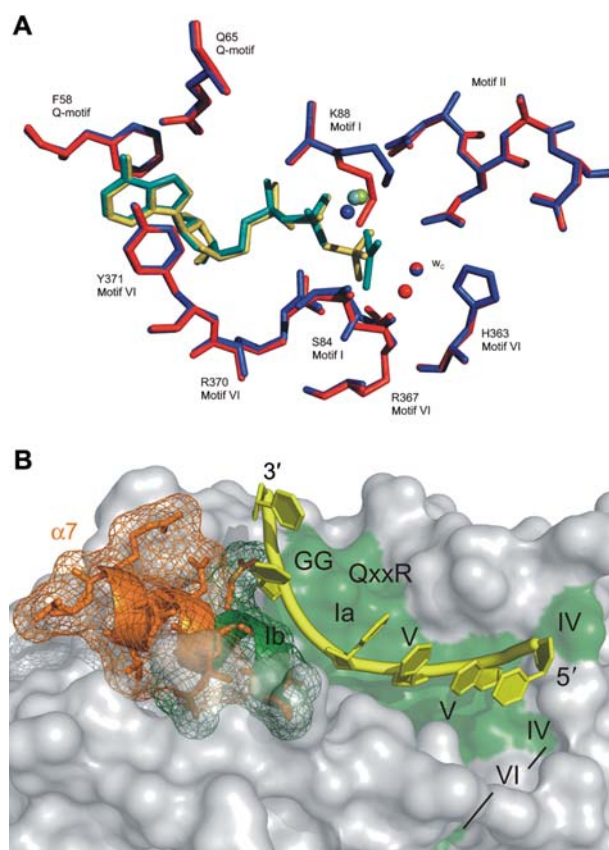
Figure 3 Nucleotide and RNA binding.
(A) Nucleotide binding pocket. Superposition of the structures of eIF4A-III in complex with ADPNP and ssRNA (red, nucleotide/ Mg^{2+} in yellow), and in complex with ADP-AlF$_x$ and ssRNA (blue, nucleotide/Mg^{2+} in cyan). Water molecules that contact the nucleotide in these structures are depicted as red and blue sphere, respectively. The side chain conformations of amino acids contributing to nucleotide binding and hydrolysis are similar in both structures, with the exception of S84 and K88 (motif I). The catalytic water molecule (w_c, red/blue sphere) is in nearly the same location in both structures. The side chain of D342 (motif VI) contacts the 2'-OH but has been omitted for clarity. (B) RNA binding site: the ssRNA (yellow) bound to Vasa (gray surface) is kinked. The RNA distortion is not compatible with regular A-RNA conformation and can thus destabilize RNA duplexes bound to DEAD box helicases. Steric hindrance owing to the position of helix α7 (orange) causes the kink after nucleotide U5. Conserved motifs involved in RNA binding are highlighted in green.

Nucleotide binding

The nucleotide binding pocket in DEAD box proteins is formed by a Walker A motif (motif I), a Walker B motif (DEAD box, motif II), and the Q-motif (Figure 3A). The Walker A motif contains the characteristic P-loop (Saraste et al., 1990; Leipe et al., 2003) with a conserved lysine that contacts the β- and γ-phosphates. Residues from motifs I and II contact the triphosphate part of the nucleotide directly and through water and Mg^{2+} (Andersen et al., 2006; Bono et al., 2006; Sengoku et al., 2006; von Moeller et al., 2009). In addition, motif VI from the C-terminal RecA-like domain interacts with the triphosphate in the closed conformation (Bono et al., 2006; Sengoku et al., 2006; von Moeller et al., 2009). Whereas motifs I and II are present in various other ATPases (Wal-

ker et al., 1982), the Q-motif upstream of motif I is unique to DEAD box proteins (Tanner, 2003; Tanner et al., 2003). Its highly conserved glutamine interacts with the adenine base and provides specificity for adenine nucleotides. Other nucleotides might compete for the binding site but are not efficiently hydrolyzed and do not stimulate RNA unwinding by DEAD box proteins (Du et al., 2002; Franca et al., 2007; Garcia and Uhlenbeck, 2008). It has therefore been suggested that the Q-motif also contributes to positioning of the nucleotide for hydrolysis (Tanner, 2003; Tanner et al., 2003). The phenylalanine of the Q-motif stacks with the adenine base that is also contacted by an aromatic or hydrophobic side chain from motif VI (Bono et al., 2006). Only the 3′-OH group of the ribose is contacted, rationalizing the weak discrimination between ATP and dATP (Du et al., 2002). In addition to direct and indirect contacts to the nucleotides, the conserved motifs from both RecA-like domains form an intricate interaction network around the ATPase site. In some cases, the isolated N-terminal RecA-domain interacts (weakly) with nucleotides (Rudolph et al., 2006; Fan et al., 2009; Napetschnig et al., 2009), whereas no interaction with nucleotides was detected for the YxiN N-terminal RecA-like domain (Karow et al., 2007). The interaction of the nucleotide with both domains provides the link to its influence on RNA binding and remodeling.

Early experiments to understand the nucleotide cycle of DEAD box proteins addressed binding of ADP and ATP. DEAD box proteins bind ADP one order of magnitude more tightly than ATP (Lorsch and Herschlag, 1998a; Talavera and De La Cruz, 2005; Karow et al., 2007). The higher affinity for ADP compared to ATP explains the observed product inhibition of the ATPase activity (Lorsch and Herschlag, 1998a). In addition, the lack of contributions of the γ-phosphate to nucleotide affinity is an indication for ATP-driven conformational changes. Limited proteolysis experiments provide further evidence for ATP-induced conformational changes (Lorsch and Herschlag, 1998b; Henn et al., 2002; Cheng et al., 2005; Low et al., 2007).

Tremendous insight into the mechanism of DEAD box protein activity has been gained from studies using ATP analogs such as ATPγS, ADPNP, ADP-BeF$_x$, and ADP-AlF$_x$. The role of nucleotides for RNA unwinding will be discussed in the section 'the mechanism of DEAD box helicases: ATP hydrolysis and its coupling to RNA unwinding' below.

RNA binding

DEAD box proteins functionally interact with RNA via the helicase core. Flanking domains can contribute high affinity and/or high specificity RNA binding. For example, the C-terminal domains of the splicing helicases CYT-19 and Mss116 mediates interactions with structured RNA (Grohman et al., 2007; Mohr et al., 2008). In contrast, the C-terminal domains of DbpA and YxiN specifically interact with a hairpin in ribosomal RNA (Diges and Uhlenbeck, 2001; Tsu et al., 2001; Kossen et al., 2002; Karginov et al., 2005; Wang et al., 2006), and the C-terminal domain of Hera provides high affinity for ribosomal RNA and RNase P RNA, among others (Morlang et al., 1999; Linden et al., 2008). The influence of domains

flanking the helicase core on RNA binding and on DEAD box mechanism in general will be discussed in detail in 'Modulation of the helicase core activity by interacting partners and flanking domains'. Here we focus on the RNA binding properties of the core domain.

The translation initiation factor eIF4A consists of the helicase core without any additional domains and thus represents a minimal DEAD box protein. As a consequence, the RNA binding properties of eIF4A reflect RNA binding to the isolated helicase core. eIF4A possesses a moderate RNA affinity without sequence specificity (Rogers et al., 1999, 2002). The K_d values for single-stranded RNA are in the low micromolar range, and no or weak binding of dsRNA has been detected (Lorsch and Herschlag, 1998a). In agreement with common properties of the helicase core, the truncation of flanking domains in DEAD box helicases reduces their RNA affinity to the level observed for eIF4A (Karginov et al., 2005; Grohman et al., 2007; Mohr et al., 2008).

The RNA binding site of the helicase core is formed by both RecA-like domains, and binding involves contacts to motifs Ia, GG, and Ib in the N-terminal domain, and to motifs IV, QxxR, V, and VI in the C-terminal domain (Pause et al., 1993; Andersen et al., 2006; Bono et al., 2006; Sengoku et al., 2006; Nielsen et al., 2008; Collins et al., 2009; von Moeller et al., 2009) (Figures 1 and 3B). In all available structures, the N-terminal RecA-like domain binds the 3′-region of the ssRNA substrate, and the C-terminal domain binds the 5′-region. The continuous bipartite RNA binding site is stabilized by a network of interdomain interactions. Motifs Ia from the N-terminal and motif QxxR from the C-terminal RecA-like domain contact the same phosphate and might therefore contribute to the stabilization of the closed conformation (Bono et al., 2006; Sengoku et al., 2006). Whereas the ribose phosphate backbone of the RNA is extensively contacted, no interactions occur with the nucleobases of the RNA, explaining the lack of sequence specificity. Direct contacts to 2′-OH groups of the ribose moieties lead to a discrimination against DNA as a substrate (Peck and Herschlag, 1999; Rogers et al., 2001a; Bono et al., 2006; Sengoku et al., 2006). RNA binding to the open conformation in the absence of ATP only occurs with very low affinity as the bipartite binding site is not aligned.

Most nucleotides of the ssRNA bound to the closed helicase core are stacked, but stacking is interrupted between nucleotides 5 and 6 (numbering according to the Vasa structure; Sengoku et al., 2006). This leads to the striking feature of a kink in the phosphate-ribose backbone between these nucleotides (Figure 3B). This kink has consistently been observed in all structures of closed DEAD box helicase cores known to date and is caused by a steric hindrance with helix α7 (Figure 3B) that counteracts a straight binding of RNA. As this kink is not compatible with canonical A-form geometry and thus with dsRNA, it has been proposed to constitute a first step towards RNA unwinding (Sengoku et al., 2006). Helix α7 equivalents in other SF2 families adopt different conformations that would not impose a kink on the bound RNA. Hence, the RNA distortion might be a special feature of DEAD box proteins and point towards a unique mechanism of unwinding.

Although all crystals of DEAD box proteins with RNA have been obtained with longer single-stranded oligo-uridine, only six or seven nucleotides of the RNA substrate are visible in the electron densities. However, no structural data on relevant physiological unwinding substrates of DEAD box proteins are available, limiting possible functional conclusions. First efforts to define the size of the RNA binding site investigated the interaction of eIF4A with RNA substrates of different lengths and revealed that ssRNAs of 11–18 nucleotides interact with eIF4A and stimulate its intrinsic ATPase activity (Abramson et al., 1987; Goss et al., 1987). A 20mer was used to describe the minimal kinetic and thermodynamic framework of the nucleotide cycle in the presence of RNA for eIF4A and demonstrated coupling between nucleotide state and RNA affinity (Lorsch and Herschlag, 1998a). Since then, functional interaction with unstructured RNAs of ∼15 nucleotides was shown for several DEAD box helicases (Peck and Herschlag, 1999; Bizebard et al., 2004; Garcia and Uhlenbeck, 2008). RNA protection experiments showed that ∼10 nucleotides are protected by binding to eIF4A (Le Hir et al., 2000; Ballut et al., 2005; Rozovsky et al., 2008). However, already 2–4 RNA nucleotides within a DNA molecule lead to a stimulation of ATPase activity (Peck and Herschlag, 1999). The necessity of much larger RNA fragments to fully stimulate ATPase activity of some DEAD box proteins most likely is not an intrinsic core characteristic but can be ascribed to interaction of a structured substrate with additional domains (Garcia and Uhlenbeck, 2008). Interacting proteins might also increase the size of the RNA binding site. The RNA binding site of eIF4A increases from ∼17 to ∼30 nucleotides in the presence of its interacting partners eIF4H or eIF4B (Rozovsky et al., 2008), and to >60 bp in the eIF4F complex (Kaye et al., 2009). Altogether, it appears that binding already occurs with rather short RNA molecules, but full activation of the DEAD box protein might require larger nucleic acids interacting with the core.

Cooperativity of RNA and nucleotide binding

In principle, binding of two different ligands can occur in a random or ordered manner. The binding sites can be independent, or coupled to each other via positive or negative cooperativity. Early experiments suggested sequential binding of ATP and RNA (Abramson et al., 1987; Pause et al., 1993). This conclusion resulted from the low RNA affinity of the helicase core in the absence of nucleotides, making RNA binding difficult to detect. For eIF4A, cooperative binding of ssRNA and ATP was demonstrated, but no evidence of sequential binding was observed (Lorsch and Herschlag, 1998a,b). Cooperative binding of RNA and ATP has been demonstrated for various other DEAD box proteins (Mohr et al., 2002; Polach and Uhlenbeck, 2002; Cordin et al., 2004; Elles and Uhlenbeck, 2008; Theissen et al., 2008) and has been established as a general feature, although kinetic data do not support this cooperativity for DbpA (Henn et al., 2008). In the other case where no apparent cooperativity has been observed (Grohman et al., 2007; Mohr et al., 2008), C-terminal domains contribute the largest

part of the RNA affinity and mask nucleotide-dependent RNA binding properties of the core. eIF4A appears to bind ADP and ssRNA with negative cooperativity (Lorsch and Herschlag, 1998a), whereas other DEAD box proteins investigated show little or no cooperativity in binding of ADP and single-stranded or structured RNA (Cordin et al., 2004; Henn et al., 2008; Nielsen et al., 2008; Theissen et al., 2008). A detailed analysis of the nucleotide cycle of DbpA revealed the post-hydrolysis ADP-P_i state rather than the initial ATP complex or the ADP state as a high-affinity state with regard to RNA (Henn et al., 2008). These results imply that phosphate release resets the helicase from the ADP-P_i state with high affinity for RNA to the ADP state with low RNA affinity. Consistent with this notion, duplex unwinding in the presence of ADP-BeF_x, but not ADP-AlF_x, was demonstrated (Liu et al., 2008), suggesting that unwinding occurs before ATP hydrolysis and hydrolysis is required for resetting the enzyme for further catalytic cycles (see the following section).

The flexible connection of the two helicase domains renders it likely that rearrangements of the domains occur during the catalytic cycle. Indirect evidence for such a conformational reorganization is available from different proteolysis patterns in the presence or absence of substrates (Lorsch and Herschlag, 1998b; Henn et al., 2002; Cheng et al., 2005; Low et al., 2007). Nucleotide binding kinetics support nucleotide-driven conformational rearrangements: ATP and ADP binding follows a single exponential in the absence of RNA, but a second slow phase appears when RNA is present. The slow step might reflect a (local or global) conformational reorganization (Henn et al., 2008). The domain reorientation upon cooperative binding of ADPNP and RNA binding has been demonstrated directly in single molecule FRET experiments (Theissen et al., 2008), and the first structure of the closed DEAD box protein core (Sengoku et al., 2006) has rationalized the observed communication between nucleotide and RNA binding sites. Ddx19 and eIF4A-III are the only DEAD box proteins for which structures have been determined both in open and closed conformations (Andersen et al., 2006; Collins et al., 2009; von Moeller et al., 2009). In the closed conformation, ATP and RNA interact with both RecA-like domains of the core. Thus, both ATP and RNA binding stabilize the closed conformation, explaining the observed cooperativity.

The mechanism of DEAD box helicases: ATP hydrolysis and its coupling to RNA unwinding

Characteristics of RNA helicase substrates

The unwinding mechanism of DEAD box proteins is principally different from DNA helicases. Highly processive DNA helicases are motor proteins that translocate with a defined directionality along DNA stretches and unwind thousands of base pairs without dissociation from their DNA substrates (Singleton et al., 2007). Remodeling of RNA structures by DEAD box proteins does not require unwinding of long stretches of dsRNA, and DEAD box proteins achieve duplex separation through local base

pair disruption. The closely related DExH RNA helicases, however, exhibit intermediate processivity (Jankowsky et al., 2000). Many of these enzymes are viral proteins and might require processivity for functions during viral replication.

DEAD box proteins bind tightly to single-stranded RNA (ssRNA) but do not or only weakly interact with dsRNA (Lorsch and Herschlag, 1998a). The presence of single-stranded regions facilitates loading of DEAD box protein onto their RNA substrate (Yang and Jankowsky, 2006; Halls et al., 2007). Interestingly, a continuous backbone of the ssRNA and the unwinding target is not required, but a close proximity is sufficient (Yang and Jankowsky, 2006). In addition, unwinding does not necessarily start at the end of a duplex but can also be facilitated internally and therefore does not depend on single-strand/double-strand junctions (Yang and Jankowsky, 2006; Halls et al., 2007). As an alternative to ssRNA loading, many DEAD box proteins are brought into proximity of their unwinding target by ancillary domains that mediate RNA binding and tether the DEAD box helicase core on large RNA substrates (see section on 'modulation of the helicase core activity by interacting partners and flanking domains'). As a consequence, duplex disruption or RNA/RNP remodeling by DEAD box proteins is an inherently local activity. Although different unwinding activities of DEAD box proteins for duplexes with 5′- or 3′-single stranded overhangs have been reported, this 'directionality' possibly reflects differences in unwinding that depend on the relative orientation of the DEAD box protein with regard to the target region. For example, the C-terminal domain of DbpA recognizes a hairpin in ribosomal 23S rRNA and the helicase core unwinds an adjacent helix (Fuller-Pace et al., 1993; Diges and Uhlenbeck, 2001; Tsu et al., 2001). Unwinding efficiencies depend on the position of the helix relative to the hairpin (Diges and Uhlenbeck, 2005). If the hairpin is located on the 3′-side of the helical region, the helix is efficiently unwound. When the hairpin is moved to the 5′-side, no duplex unwinding is observed. However, an increased linker length between the hairpin on the 5′-side and the helix restores unwinding efficiency, suggesting that the flexibility of the longer linker allows for proper positioning of the helix to be unwound for the helicase core (Diges and Uhlenbeck, 2005).

Mechanism of ATP hydrolysis

The catalytic site for ATP hydrolysis of DEAD box proteins is assembled at the interface of the two RecA-like domains in the helicase core by closure of the interdomain cleft in response to RNA and ATP binding. The mechanism of ATP hydrolysis by DEAD box proteins has become evident from mutagenesis studies and from structures of the closed helicase core in complex with ADPNP and ssRNA (Andersen et al., 2006; Bono et al., 2006; Sengoku et al., 2006; Collins et al., 2009; von Moeller et al., 2009), and with the transition state analog ADP-AlF$_x$ and ssRNA (Nielsen et al., 2008; Figure 3A). Mutations in motifs I and II (Walker A and B motifs) render DEAD box proteins ATPase-deficient. A conserved lysine in motif I, the glutamate in motif II, and an arginine in motif VI are essential for catalysis (Rozen et al., 1989;

Elles and Uhlenbeck, 2008). The motif I lysine bridges the β- and γ-phosphates and presumably coordinates the γ-phosphate while it is transferred from the β-phosphate to the catalytic water (Bono et al., 2006; Nielsen et al., 2008; Figure 3A). A water molecule embedded in interactions with side chains from motifs II, V, and VI close to the γ-phosphate is in appropriate geometry for nucleophilic attack on the scissile bond (Andersen et al., 2006; Sengoku et al., 2006). Similar to mutations in motifs I and II, mutation of a conserved arginine in motif VI abolishes the ATPase activity (Elles and Uhlenbeck, 2008). This arginine contacts the γ-phosphate and might stabilize the transition state (Sengoku et al., 2006; Elles and Uhlenbeck, 2008; Nielsen et al., 2008). Its role would be reminiscent of the arginine finger in GTPase activating proteins (Ahmadian et al., 1997). The structure of eIF4A-III in the exon junction complex with ADP-AlF$_x$, a transition state analog, is indicative of a dissociative mechanism for phosphoryl transfer (Nielsen et al., 2008). eIF4A-III adopts the closed conformation in complex with ADP-AlF$_x$. Interestingly, the overall structure is nearly identical to the ADPNP state of other DEAD box proteins, with small differences limited to the immediate environment of the γ-phosphate (Nielsen et al., 2008; Figure 3A).

Coupling of ATP hydrolysis and duplex separation

Strand separation is a thermodynamically unfavorable reaction. DEAD box protein mediated duplex unwinding requires ATP, and it has been inferred that the energy from ATP hydrolysis is required for base pair disruption. By contrast, many DEAD box proteins already promote limited strand separation in the absence of ATP, which has been interpreted as nonspecific duplex destabilization via an ssRNA capture mechanism (Del Campo et al., 2007; Chen et al., 2008). For a few DEAD box proteins, a strand annealing activity has been demonstrated (Yang and Jankowsky, 2005; Halls et al., 2007). In contrast to unwinding, strand annealing is an ATP-independent process (Halls et al., 2007) and is even inhibited by ATP (Yang and Jankowsky, 2005).

Unraveling the molecular mechanism that leads to coupling of ATP hydrolysis and duplex unwinding requires a thorough dissection of individual steps in the nucleotide cycle and assessment of their effects on RNA binding and unwinding. The nucleotide cycle regulates the transition of DEAD box proteins through different nucleotide states with different RNA affinities. The ADP and nucleotide-free states exhibit low RNA affinity (Lorsch and Herschlag, 1998a; Henn et al., 2008; Nielsen et al., 2008). In contrast, the ADP-P$_i$ state exhibits high RNA affinity (Henn et al., 2008). Non-hydrolyzable ATP analogs, such as ADPNP, ADP-BeF$_x$, and ADP-AlF$_x$, promote high affinity RNA binding (Ballut et al., 2005; Liu et al., 2008; Nielsen et al., 2008; Rozovsky et al., 2008; Theissen et al., 2008), but different results have been obtained for ATP itself: high RNA affinity was reported for the ATP state of DbpA (Polach and Uhlenbeck, 2002; Elles and Uhlenbeck, 2008), whereas the ATP states of Dbp5 and eIF4A exhibit low RNA affinity (Nielsen et al., 2008; Rozovsky et al., 2008).

The potential of ATP analogs (or lack thereof) to support unwinding supports the notion that ATP hydrolysis

is required for base-pair disruption. ADPNP is not hydro-lyzed by DEAD box proteins and does not support RNA unwinding (Liu et al., 2008; Theissen et al., 2008). In contrast, ATPγS, which is frequently used as a 'non-hydrolyzable' nucleotide analog for many ATPases, is efficiently hydrolyzed by eIF4A and supports RNA unwinding (Peck and Herschlag, 2003). Although ATPγS is hydrolyzed 10-fold more slowly than ATP, the unwind-ing rates with ATP- and ATPγS are similar, suggesting that the chemical step of ATP hydrolysis is not rate lim-iting for RNA unwinding (Peck and Herschlag, 2003). Dis-section of individual steps in the nucleotide cycle of DbpA has provided a kinetic and thermodynamic frame-work for the nucleotide cycle and its links to RNA binding and unwinding (Henn et al., 2008). Phosphate release was identified as a rate-limiting step both in the absence and in the presence of RNA (Henn et al., 2008), in line with similar rates of ATP- and ATPγS-supported RNA unwinding (Peck and Herschlag, 2003). Recently, the effect of the ATP analogs ADP-BeF$_x$ and ADP-AlF$_x$ on RNA unwinding has been examined (Liu et al., 2008). ADP-BeF$_x$ is a ground state ATP analog, whereas ADP-AlF$_x$ mimics the transition state. The structure of eIF4A-III in complex with ADP-AlF$_x$ confirms that it acts as a transition state analog for DEAD box proteins (Nielsen et al., 2008). ADP-BeF$_x$ and ADP-AlF$_x$ favor high-affinity ssRNA binding (Liu et al., 2008) similar to ADPNP. Inter-estingly, ADP-BeF$_x$ facilitates RNA unwinding by Ded1p under single turnover conditions. In contrast, ADP-AlF$_x$ and ADP-P$_i$ do not support RNA unwinding, suggesting that strand separation occurs prior to ATP hydrolysis (Liu et al., 2008). Duplex unwinding can be achieved by hydrolysis of one ATP molecule, and under certain conditions less than one ATP is hydrolyzed per duplex separation event (Chen et al., 2008). In fact, ATP bound to the helicase core was only hydrolyzed in half of the strand separation events (Chen et al., 2008), further sup-porting that strand separation can precede hydrolysis. Strikingly, ADP-BeF$_x$ does not support unwinding by the DExH box protein NPH-II (Liu et al., 2008). Thus, the mechanism of coupling the nucleotide cycle to duplex unwinding appears to be specific for DEAD box proteins.

Under multiple turnover conditions, where unwinding is efficient in the presence of ATP, ADP-BeF$_x$ is not suf-ficient to promote RNA unwinding. Consequently, ATP hydrolysis must be required to reset the enzyme and to enable multiple turnovers (Liu et al., 2008). ADP-BeF$_x$ appears to faithfully mimic the ATP conformation of DEAD box proteins, whereas it has been suggested that ADPNP traps a non-productive DEAD box protein/RNA complex (Henn et al., 2008) and does not elicit the same conformation of DEAD box proteins as ATP (Chen et al., 2008; Liu et al., 2008). Although ADPNP does not sup-port RNA unwinding, its binding to the helicase core allows for the closure of the interdomain cleft in response to RNA binding (Andersen et al., 2006; Bono et al., 2006; Theissen et al., 2008; Karow and Klostermeier, 2009; von Moeller et al., 2009). The closed conformation of YxiN in complex with ADPNP resembles the closed conformation with ATP (Theissen et al., 2008; Karow and Klostermeier, 2009), suggesting that global conformations are similar, and the difference between the ATP and ADPNP confor-

mations must be subtle. Consistent with this notion, the global structure of eIF4A-III/ADP-AlF$_x$ with ssRNA bound is virtually identical to the ADPNP complex, with small local differences around the γ-phosphate (Nielsen et al., 2008; Figure 3A). The structures of Mss116 in complex with ssRNA and ADPNP, ADP-BeF$_x$ and ADP-AlF$_x$ were recently determined (Del Campo and Lambowitz, 2009). All three structures do not reveal any significant differ-ences outside the ATP binding site, suggesting that there is no apparent structural reason for different capacities to support unwinding.

A unifying mechanism for DEAD box protein activity?

The observation that duplex separation might precede ATP hydrolysis is not unexpected. The kink in the back-bone of ssRNA bound to the closed conformation of the helicase core interferes with base pairing (Figure 3B). The introduction of this kink would thereby lead to the sep-aration of the first few base pairs of the RNA substrate, and it has been interpreted as a possible first step towards unwinding (Sengoku et al., 2006; Karow and Klostermeier, 2009). The low affinity of DEAD box pro-teins for ATP compared to ADP might not only reflect the energetic cost for ATP-induced conformational changes of the protein but also for distortion of the RNA. Kinking the substrate is a consequence of the closure of the interdomain cleft in the helicase core but does not require ATP hydrolysis (Karow and Klostermeier, 2009). Conse-quently, all nucleotides that favor the closed conforma-tion of the helicase core should lead to a kink in the RNA substrate. If the remaining double helical region is unsta-ble, disruption of these few base pairs would already result in dissociation of the first RNA strand and in unwinding (Rogers et al., 1999). However, ADPNP and ADP-AlF$_x$ allow for the distortion of the ssRNA and yet do not support RNA unwinding. In this respect, it is inter-esting to note that a rearrangement of the eIF4A/ATP complex towards a hydrolysis-competent conformation was proposed earlier (Peck and Herschlag, 1999). The recent observation of two phases for nucleotide binding in the presence of structured RNA also supports the notion of two ATP complexes (Henn et al., 2008). These data can be integrated in a model in which initial binding of ATP and dsRNA to DEAD box proteins is followed by a rearrangement to a hydrolysis- and unwinding-competent complex (Figure 4). The different capacities of ATP analogs to support unwinding suggest that this rearrangement can also take place in the presence of ADP-BeF$_x$. In contrast, the ADPNP and ADP-AlF$_x$ com-plexes do not undergo this activation, and thus ADPNP and ADP-AlF$_x$ cannot support RNA unwinding. From the activated complex, the first strand of the bound RNA can dissociate, resulting in RNA unwinding. Liu and col-leagues have speculated that the flexibility of the single-stranded RNA product is required for local conforma-tional adjustments that lead to ATP hydrolysis (Liu et al., 2008). In this scenario, ATP hydrolysis would be a con-sequence of (and subsequent to) unwinding rather than the cause, and the two processes would be tightly linked. If the first RNA strand dissociates from the helicase before ATP hydrolysis, as suggested by Liu and cowork-

ers, unwinding has already been favored by ATP binding and RNA distortion (Figure 4), rationalizing RNA unwinding without ATP hydrolysis (Chen et al., 2008). Kinetic partitioning between RNA dissociation and ATP hydrolysis could explain the influence of duplex length on ATP consumption (Figure 4): whereas one round of ATP hydrolysis leads to a large fraction of, or even complete, unwinding for short helices, more ATP is required for unwinding of longer helices (Rogers et al., 1999; Chen et al., 2008). The probability of a kink and unwinding of a few base pairs to cause complete strand separation is high for short duplexes, whereas for longer helices one round of nucleotide-regulated conformational changes results in a lower probability of strand separation. Thus, for longer duplexes, ATP is hydrolyzed while both RNA strands are still bound, generating the ADP-P_i state (Figure 4). Possibly, the increased flexibility of the complex owing to the local distortion of the RNA is sufficient to allow for ATP hydrolysis.

It is currently unclear which step in the nucleotide cycle triggers release of the second RNA strand. Its dissociation should be linked to the reopening of the interdomain cleft in the helicase core. The ADP-P_i state is difficult to populate and has not been accessible in single molecule FRET experiments (Theissen et al., 2008). However, this state shows tight coupling with RNA binding (Henn et al., 2008), suggesting that it must be in the closed conformation. By contrast, the helicase core is mainly in the open conformation in the presence of ADP and RNA (Theissen et al., 2008), suggesting that phosphate release from the ADP-P_i state should trigger reopening of the interdomain cleft. Consistent with these observations, kinetic evidence suggests that phosphate release might be linked to RNA release (Henn et al., 2008). Both findings are consistent with the reduced ssRNA affinity of DEAD box proteins in complex with ADP (Lorsch and Herschlag, 1998a; Iost et al., 1999; Cordin et al., 2004).

In the current scenario of the DEAD box protein mechanism, RNA unwinding is initiated and is sometimes even completed by distorting the bound RNA in the ATP-bound state. This state can be mimicked by ADP-BeF$_x$ but not by ADPNP. ATP hydrolysis is required to reset the enzyme to a low-affinity state for RNA and to allow for multiple rounds of unwinding (Chen et al., 2008). For DEAD box proteins with additional RNA binding domains, the helicase core might remain tethered to the RNA substrate, allowing for multiple catalytic cycles (Diges and Uhlenbeck, 2001; Tijerina et al., 2006). Remodeling of RNP complexes can be rationalized by the same mechanism and would be a consequence of local distortions of the RNA that interfere with protein binding.

Modulation of the helicase core activity by interacting partners and flanking domains

The helicase core provides all basic properties required for a functional DEAD box protein. Nevertheless, the properties of the helicase core are modulated in virtually all DEAD box proteins. Regulation can be either *in cis* or *in trans*. For example, few DEAD box proteins contain

insertions into the conserved core scaffold. Many DEAD box proteins contain additional flanking domains that affect the helicase core *in cis*. The prototypic DEAD box protein eIF4A that consists only of a helicase core is part of a large interaction network and is extensively regulated by interaction partners *in trans*.

Modulation by insertions and flanking domains

The human Ded1p homolog Ddx3x/y helicase core contains a unique 10 aa insertion between the P-loop and motif Ia. This insertion folds into a helix and forms a continuous positive patch together with a loop enriched
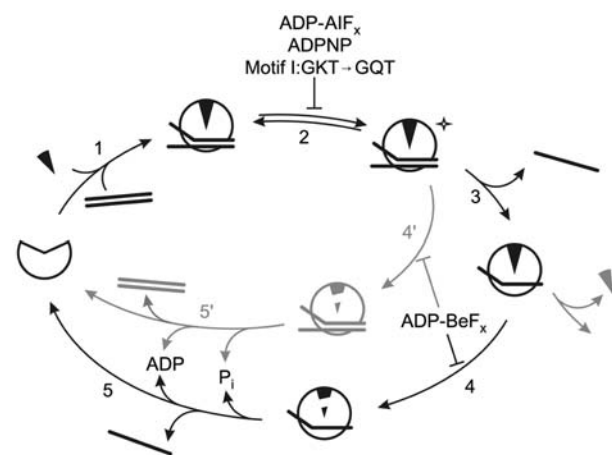
Figure 4 Possible catalytic cycle of RNA unwinding by DEAD box proteins.

DEAD box proteins bind dsRNA (line pair) and ATP (black triangle, step 1). In complex with ATP and RNA, the DEAD box protein adopts a closed conformation. One strand of the bound RNA is kinked, leading to the local destabilization of (terminal or internal) base pairs. The initial complex undergoes a rearrangement towards an activated, hydrolysis- and unwinding-competent conformation (star, step 2; Peck and Herschlag, 1999; Henn et al., 2008). The structural difference between these two complexes is currently unclear. From the activated complex, the first RNA strand can dissociate (step 3) before ATP is hydrolyzed (step 4). Product release is coupled to the dissociation of the second RNA strand (step 5), thereby resetting the enzyme for further cycles. Possibly, ATP can dissociate before being hydrolyzed (gray, right arm), allowing for the completion of one cycle of unwinding in the absence of ATP hydrolysis (Chen et al., 2008). For longer RNA duplexes, destabilization by the kink does not lead to the same probability for complete strand separation, and dissociation of the first RNA strand from the activated complex. Possibly, the increased flexibility of the complex owing to transient local unwinding of the RNA allows for ATP hydrolysis to occur. In this case, ATP is hydrolyzed before the first RNA strand dissociates (step 4'). Product release and dissociation of the RNA (step 5') reset the enzyme for further cycles. The partitioning between dissociation of the first strand and ATP hydrolysis rationalizes the higher amount of ATP required for unwinding longer duplexes (Rogers et al., 1999; Chen et al., 2008). The nucleotide analogs ADPNP and ADP-AlF$_x$ do not promote the rearrangement towards the activated (hydrolysis- and unwinding-competent) complex (step 2) and thus do not support RNA unwinding. The unwinding deficiency of motif I mutants can be rationalized similarly. In the presence of ADP-BeF$_x$, the activation takes place and RNA unwinding becomes possible. However, in the absence of hydrolysis, the enzyme is not released from the second RNA strand and cannot undergo multiple catalytic cycles.

in positively charged side chains that is also unique to Ddx3x/y. As this patch is localized close to the expected position of the 3′-end of bound RNA, it has been suggested that the insertion increases the RNA binding surface and contributes to RNA binding (Hogbom et al., 2007). A striking example for a SF2-like module with an insertion is the helicase-like domain in reverse gyrase which provides the nucleotide-dependent DNA binding to this topoisomerase (Del Toro Duany et al., 2008). Here, a small domain of ~75 aa that is homologous to the transcription termination factor Rho, the so-called latch domain, is inserted into the C-terminal RecA-like domain (Rodriguez and Stock, 2002). The latch domain interacts with the topoisomerase domain and has been implicated in communication between the two domains during positive DNA supercoiling by reverse gyrase (Rodriguez, 2002, 2003).

The helicase core of Dbp5/Ddx19 is flanked by an N-terminal extension that appears to contribute to nucleotide binding (Collins et al., 2009; Fan et al., 2009; Napetschnig et al., 2009; von Moeller et al., 2009). A homologous region is also found in Ddx25. In Ddx19, the α-helix N-terminal to the helicase core is inserted in the cleft of the helicase core in the presence of ADP (Collins et al., 2009). When ADPNP and ssRNA are bound, the Ddx19 core adopts the closed conformation, and the N-terminal helix is packed onto the surface of the C-terminal RecA-like domain, whereas it interacts with the N-terminal RecA-like domain in the structure of Ddx19 with its interaction partner Nup214 (Napetschnig et al., 2009). Deletion experiments point towards a regulatory role of the Ddx19 N-terminal extension on Ddx19 activity (Collins et al., 2009).

In various DEAD box proteins, domains flanking the helicase core mediate interactions with RNA substrates or possibly also with other proteins. RNA binding domains can provide high affinity, high specificity, or both. The DEAD box proteins DbpA (*Escherichia coli*) and its *B. subtilis* homolog YxiN bind to hairpin 92 in the ribosomal 23S rRNA with high affinity and specificity and have been implicated in ribosome biogenesis (Fuller-Pace et al., 1993; Nicol and Fuller-Pace, 1995; Kossen and Uhlenbeck, 1999; Kossen et al., 2002). The crystal structure of the C-terminal RNA-binding domain (RBD) of YxiN adopts a classical RNA-recognition motif (RRM) fold (Wang et al., 2006). However, aromatic residues on the face of the β-sheet are not required for RNA binding, suggesting an unusual RNA binding mode. Instead, conserved lysines in two loops might be involved in RNA binding. The isolated RBD binds hairpin 92 with similar affinity as the full-length YxiN protein (Karginov et al., 2005; Wang et al., 2007). YxiN unwinds a short double helix adjacent to hairpin 92, and it has been suggested that the RBD functions as an anchor to direct the unwinding activity of the helicase core to this region. However, localization of the RBD with respect to the helicase core in the structure is currently unknown.

A similar RRM-fold domain, the so-called GUCT domain, is present in Ddx21 and Ddx50 (RNA helicase II-Gu isoforms α and β, respectively) (Ohnishi et al., 2008). Its negative surface potential, however, is not consistent with RNA binding and might point towards a different function.

The DEAD box protein Hera from *Thermus thermophilus* contains a flanking region C-terminal to the helicase core. As a unique feature so far, this C-terminal domain is bipartite and consists of a dimerization motif (Klostermeier and Rudolph, 2009; Rudolph et al., 2009), and an RBD that mediates high affinity interaction with ribosomal RNA fragments and RNase P RNA (Linden et al., 2008). Despite the lack of sequence similarity with YxiN, the Hera RBD folds into a modified RRM (Rudolph and Klostermeier, 2009). Hera is the first example where the orientation of the RNA binding domain with respect to the helicase core has been determined (Klostermeier and Rudolph, 2009; Rudolph et al., 2009; Rudolph and Klostermeier, 2009).

Some DEAD box proteins contain basic C-terminal extensions rich in arginines, serines and glycines. Examples are the splicing helicases Cyt-19 and Mss116 (Mohr et al., 2002; Huang et al., 2005; Halls et al., 2007) that are general RNA chaperones, and Ded1p, a helicase involved in translation (Iost et al., 1999). These proteins bind unspecifically to structured RNA via their basic C-terminal extensions, and their helicase cores then unwind loosely attached neighboring double helical regions (Grohman et al., 2007; Mohr et al., 2008). In Cyt-19 and Mss116, a α-helical region is located between the helicase core and the basic tail. A high RNA-stimulated ATPase activity is maintained when the basic tail is deleted, but mutations in the α-helical region strongly inhibit RNA-dependent ATPase activity, pointing to different functions of these two regions in modulating activity of the helicase core (Mohr et al., 2008). A recently reported structure of Mss116 that includes the α-helical region reveals a second kink in the bound RNA caused by the presence of the α-helical region (Del Campo and Lambowitz, 2009), indicating that these extensions might also contribute to RNA unwinding.

The substrate specificity of DEAD box proteins can be transferred by fusing the RNA binding domain to a different helicase core (Kossen et al., 2002). Similarly, the RNA specificity of the C-terminal domains can be exploited to direct a DEAD box protein to a specific unwinding substrate (Chen et al., 2008). These findings underscore the modular character of DEAD box proteins. In general, tethering of DEAD box proteins to their RNA substrate by high-affinity RNA binding of their C-terminal domains is not compatible with a translocation mechanism of processive helicases and local action is a unique feature of DEAD box helicases.

Regulation by other protein factors

Nucleotide-regulated transitions between open and closed conformations of the helicase core are at the heart of DEAD box protein activity. Stabilization of either of these conformations or affecting the kinetics of their interconversion constitutes an attractive and efficient way for the regulation of DEAD box protein activity. A prominent example for such a regulation mechanism is eIF4A. Its intrinsically very weak helicase activity is stimulated by various other translation factors (Rogers et al., 2001a; Korneeva et al., 2005; Oberer et al., 2005; Schutz et al., 2008; Marintchev et al., 2009). A well-studied example is the functional interaction of eIF4A with eIF4G. eIF4G

contacts both RecA-like domains of eIF4A with a conserved HEAT sequence (Oberer et al., 2005; Bellsolell et al., 2006; Schutz et al., 2008). As a consequence of this bipartite interaction, the two RecA-like domains in the eIF4A helicase core are fixed in a partly open conformation (Caruthers et al., 2000; Schutz et al., 2008). As the interaction with eIF4G stimulates the eIF4A ATPase and helicase activities, transitions of eIF4A between open and closed conformations have to be possible in the complex. Binding of eIF4G to the C-terminal RecA domain of eIF4A covers a large surface area and involves numerous interactions, whereas few contacts are formed with the N-terminal RecA domain (Schutz et al., 2008). The presence of two interaction sites, a primary, high-affinity site combined with a secondary low-affinity interaction site, could allow for the required flexibility of the eIF4A/eIF4G complex. It can be envisaged that the low-affinity interface transiently opens during the catalytic cycles of eIF4A, consistent with a previously proposed 'soft-clamp' mechanism (Oberer et al., 2005). Regulation of eIF4A activity by other translation factors might follow a similar regulation principle. For eIF4B and eIF4H, stable interactions with the N-terminal domain of eIF4A in the presence of ADPNP and ssRNA have been demonstrated, and there is evidence for additional interactions with the C-terminal domains (Rozovsky et al., 2008).

Based on a similar regulation principle, inhibition of DEAD box proteins can be achieved by interaction partners that stabilize a single conformation and interfere with shuttling between open and closed conformations. One example is the isoform III of eIF4A which is part of the exon junction complex. In the exon junction complex structure, eIF4A-III bound to ADPNP and ssRNA adopts a closed conformation (Andersen et al., 2006; Bono et al., 2006). Both domains of eIF4A are contacted by MLN51 (Brz). The bipartite interaction of MLN51 with eIF4A is stabilized by the other components of the complex, MAGO and Y-14 (Andersen et al., 2006; Bono et al., 2006). Although ATP hydrolysis is still possible, ATP turnover is inhibited by preventing product dissociation (Nielsen et al., 2008). As a consequence, eIF4A is trapped in the ADP-P_i form, tightly bound to its RNA substrate, consistent with stable RNA binding as its main function (Nielsen et al., 2008).

A similar model might be more common as a regulatory principle for helicase domains that are part of larger functional units where the helicase domain is just one among many other functional domains, such as reverse gyrase or Dicer. The variety of modulation principles indicates a large repertoire of mechanisms in the modulation of DEAD box protein activities and of tailoring the helicase core properties to a specific DEAD box protein function.

Conclusions

The DEAD box protein core is a nucleotide-dependent switch that adopts an open conformation in the absence of ligands and in the ADP state, and in a closed conformation in the ATP (and ADP-P_i) states. The closure of the interdomain cleft in the helicase core is coupled to high-affinity RNA binding and to a distortion of the bound RNA. This distortion is a critical feature of DEAD box protein activity. However, formation of a complex with a kinked RNA substrate does not necessarily lead to unwinding, suggesting that an activation step, a rearrangement of the initial complex to a hydrolysis- and unwinding-competent state, has to occur. Such an activation step can be inferred from biochemical data, but its nature is currently unknown. From this activated complex, one RNA strand can dissociate from the destabilized duplex and strand separation can precede ATP hydrolysis. Dissociation of the second RNA strand is most likely coupled to phosphate release. ATP hydrolysis is not required for strand separation but for release of the second RNA strand and for resetting the enzyme for multiple turnovers. The same model can also account for other activities of DEAD box proteins, such as protein displacement and structural remodeling. Variations of the common DEAD box protein scheme can be achieved by flanking domains that confer high RNA affinity or specificity, contribute to nucleotide binding, or mediate interactions with other proteins. Future experiments should address the existence and nature of the two ATP complexes, and the role of phosphate release for the mechanism of DEAD box proteins.

Acknowledgments

References

Abramson, R.D., Dever, T.E., Lawson, T.G., Ray, B.K., Thach, R.E., and Merrick, W.C. (1987). The ATP-dependent interaction of eukaryotic initiation factors with mRNA. J. Biol. Chem. *262*, 3826–3832.

Ahmadian, M.R., Stege, P., Scheffzek, K., and Wittinghofer, A. (1997). Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. Nat. Struct. Biol. *4*, 686–689.

Andersen, C.B., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L., Pedersen, J.S., Seraphin, B., Le Hir, H., and Andersen, G.R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. Science *313*, 1968–1972.

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B., and Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat. Struct. Mol. Biol. *12*, 861–869.

Banroques, J., Cordin, O., Doere, M., Linder, P., and Tanner, N.K. (2008). A conserved phenylalanine of motif IV in superfamily 2 helicases is required for cooperative, ATP-dependent binding of RNA substrates in DEAD-box proteins. Mol. Cell. Biol. *28*, 3359–3371.

Bellsolell, L., Cho-Park, P.F., Poulin, F., Sonenberg, N., and Burley, S.K. (2006). Two structurally atypical HEAT domains in the C-terminal portion of human eIF4G support binding to eIF4A and Mnk1. Structure *14*, 913–923.

Bizebard, T., Ferlenghi, I., Iost, I., and Dreyfus, M. (2004). Studies on three *E. coli* DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. Biochemistry *43*, 7857–7866.

Bono, F., Ebert, J., Lorentzen, E., and Conti, E. (2006). The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. Cell *126*, 713–725.

Caruthers, J.M., Johnson, E.R., and McKay, D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. Proc. Natl. Acad. Sci. USA *97*, 13080–13085.

Chen, Y., Potratz, J.P., Tijerina, P., Del Campo, M., Lambowitz, A.M., and Russell, R. (2008). DEAD-box proteins can completely separate an RNA duplex using a single ATP. Proc. Natl. Acad. Sci. USA *105*, 20203–20208.

Cheng, Z., Coller, J., Parker, R., and Song, H. (2005). Crystal structure and functional analysis of DEAD-box protein Dhh1p. RNA *11*, 1258–1270.

Collins, R., Karlberg, T., Lehtio, L., Schutz, P., van den Berg, S., Dahlgren, L.G., Hammarstrom, M., Weigelt, J., and Schuler, H. (2009). The DExD/H-box RNA helicase DDX19 is regulated by an a-helical switch. J. Biol. Chem. *284*, 10296–10300.

Confalonieri, F., Elie, C., Nadal, M., de La Tour, C., Forterre, P., and Duguet, M. (1993). Reverse gyrase: a helicase-like domain and a type I topoisomerase in the same polypeptide. Proc. Natl. Acad. Sci. USA *90*, 4753–4757.

Cordin, O., Tanner, N.K., Doere, M., Linder, P., and Banroques, J. (2004). The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. EMBO J. *23*, 2478–2487.

Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. Gene *367*, 17–37.

Del Campo, M. and Lambowitz, A.M. (2009). Structure of the yeast DEAD-box protein Mss116p reveals two wedges that crimp RNA. Mol. Cell *35*, 598–609.

Del Campo, M., Tijerina, P., Bhaskaran, H., Mohr, S., Yang, Q., Jankowsky, E., Russell, R., and Lambowitz, A.M. (2007). Do DEAD-box proteins promote group II Intron splicing without unwinding RNA? Mol. Cell *28*, 159–166.

Del Toro Duany, Y., Jungblut, S.P., Schmidt, A.S., and Klostermeier, D. (2008). The reverse gyrase helicase-like domain is a nucleotide-dependent switch that is attenuated by the topoisomerase domain. Nucleic Acids Res. *36*, 5882–5895.

Diges, C.M. and Uhlenbeck, O.C. (2001). *Escherichia coli* DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA. EMBO J. *20*, 5503–5512.

Diges, C.M. and Uhlenbeck, O.C. (2005). *Escherichia coli* DbpA is a 3′→5′ RNA helicase. Biochemistry *44*, 7903–7911.

Du, M.X., Johnson, R.B., Sun, X.L., Staschke, K.A., Colacino, J., and Wang, Q.M. (2002). Comparative characterization of two DEAD-box RNA helicases in superfamily II: human translation-initiation factor 4A and hepatitis C virus non-structural protein 3 (NS3) helicase. Biochem. J. *363*, 147–155.

Elles, L.M. and Uhlenbeck, O.C. (2008). Mutation of the arginine finger in the active site of *Escherichia coli* DbpA abolishes ATPase and helicase activity and confers a dominant slow growth phenotype. Nucleic Acids Res. *36*, 41–50.

Fairman, M.E., Maroney, P.A., Wang, W., Bowers, H.A., Gollnick, P., Nilsen, T.W., and Jankowsky, E. (2004). Protein displacement by DExH/D RNA helicases without duplex unwinding. Science *304*, 730–734.

Fan, J.S., Cheng, Z., Zhang, J., Noble, C., Zhou, Z., Song, H., and Yang, D. (2009). Solution and crystal structures of mRNA exporter Dbp5p and its interaction with nucleotides. J. Mol. Biol. *388*, 1–10.

Flaus, A. and Owen-Hughes, T. (2001). Mechanisms for ATP-dependent chromatin remodelling. Curr. Opin. Genet. Dev. *11*, 148–154.

Franca, R., Belfiore, A., Spadari, S., and Maga, G. (2007). Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity. Proteins *67*, 1128–1137.

Fuller-Pace, F.V., Nicol, S.M., Reid, A.D., and Lane, D.P. (1993). DbpA: a DEAD box protein specifically activated by 23s rRNA. EMBO J. *12*, 3619–3626.

Garcia, I. and Uhlenbeck, O.C. (2008). Differential RNA-dependent ATPase activities of four rRNA processing yeast DEAD-box proteins. Biochemistry *47*, 12562–12573.

Gorbalenya, A.E. and Koonin, E.V. (1991). Endonuclease (R) subunits of type-I and type-III restriction-modification enzymes contain a helicase-like domain. FEBS Lett. *291*, 277–281.

Gorbalenya, A.E. and Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. Curr. Opin. Struct. Biol. *3*, 419–429.

Goss, D.J., Woodley, C.L., and Wahba, A.J. (1987). A fluorescence study of the binding of eucaryotic initiation factors to messenger RNA and messenger RNA analogues. Biochemistry *26*, 1551–1556.

Grohman, J.K., Campo, M.D., Bhaskaran, H., Tijerina, P., Lambowitz, A.M., and Russell, R. (2007). Probing the mechanisms of DEAD-box proteins as general RNA chaperones: the C-terminal domain of CYT-19 mediates general recognition of RNA. Biochemistry *46*, 3013–3022.

Halls, C., Mohr, S., Del Campo, M., Yang, Q., Jankowsky, E., and Lambowitz, A.M. (2007). Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. J. Mol. Biol. *365*, 835–855.

Henn, A., Medalia, O., Shi, S.P., Steinberg, M., Franceschi, F., and Sagi, I. (2001). Visualization of unwinding activity of duplex RNA by DbpA, a DEAD box helicase, at single-molecule resolution by atomic force microscopy. Proc. Natl. Acad. Sci. USA *98*, 5007–5012.

Henn, A., Shi, S.P., Zarivach, R., Ben-Zeev, E., and Sagi, I. (2002). The RNA helicase DbpA exhibits a markedly different conformation in the ADP-bound state when compared with the ATP- or RNA-bound states. J. Biol. Chem. *277*, 46559–46565.

Henn, A., Cao, W., Hackney, D.D., and De La Cruz, E.M. (2008). The ATPase cycle mechanism of the DEAD-box rRNA helicase, DbpA. J. Mol. Biol. *377*, 193–205.

Hogbom, M., Collins, R., van den Berg, S., Jenvert, R.M., Karlberg, T., Kotenyova, T., Flores, A., Hedestam, G.B., and Schiavone, L.H. (2007). Crystal structure of conserved domains 1 and 2 of the human DEAD-box helicase DDX3X in complex with the mononucleotide AMP. J. Mol. Biol. *372*, 150–159.

Huang, H.R., Rowe, C.E., Mohr, S., Jiang, Y., Lambowitz, A.M., and Perlman, P.S. (2005). The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. Proc. Natl. Acad. Sci. USA *102*, 163–168.

Iost, I., Dreyfus, M., and Linder, P. (1999). Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. J. Biol. Chem. *274*, 17677–17683.

Jankowsky, E., Gross, C.H., Shuman, S., and Pyle, A.M. (2000). The DExH protein NPH-II is a processive and directional motor for unwinding RNA. Nature *403*, 447–451.

Karginov, F.V., Caruthers, J.M., Hu, Y., McKay, D.B., and Uhlenbeck, O.C. (2005). YxiN is a modular protein combining a DEx(D/H) core and a specific RNA-binding domain. J. Biol. Chem. *280*, 35499–35505.

Karow, A.R. and Klostermeier, D. (2009). A conformational change in the helicase core is necessary but not sufficient for RNA unwinding by the DEAD box helicase YxiN. Nucleic Acids Res. *37*, 4464–4471.

Karow, A.R., Theissen, B., and Klostermeier, D. (2007). Authentic interdomain communication in an RNA helicase reconstituted by expressed protein ligation of two helicase domains. FEBS J. *274*, 463–473.

Kaye, N.M., Emmett, K.J., Merrick, W.C., and Jankowsky, E. (2009). Intrinsic RNA binding by the eukaryotic initiation factor 4F depends on a minimal RNA length, but not on the m7G cap. J. Biol. Chem. *284*, 17742–17750.

Klostermeier, D. and Rudolph, M.G. (2009). A novel dimerization motif in the C-terminal domain of the *Thermus thermophilus* DEAD box helicase Hera confers substantial flexibility. Nucleic Acids Res. *37*, 421–430.

Korneeva, N.L., First, E.A., Benoit, C.A., and Rhoads, R.E. (2005). Interaction between the NH$_2$-terminal domain of eIF4A and the central domain of eIF4G modulates RNA-stimulated ATPase activity. J. Biol. Chem. *280*, 1872–1881.

Kossen, K. and Uhlenbeck, O.C. (1999). Cloning and biochemical characterization of *Bacillus subtilis* YxiN, a DEAD protein specifically activated by 23S rRNA: delineation of a novel sub-family of bacterial DEAD proteins. Nucleic Acids Res. *27*, 3811–3820.

Kossen, K., Karginov, F.V., and Uhlenbeck, O.C. (2002). The carboxy-terminal domain of the DExDH protein YxiN is sufficient to confer specificity for 23S rRNA. J. Mol. Biol. *324*, 625–636.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. *19*, 6860–6869.

Leipe, D.D., Koonin, E.V., and Aravind, L. (2003). Evolution and classification of P-loop kinases and related proteins. J. Mol. Biol. *333*, 781–815.

Linden, M.H., Hartmann, R.K., and Klostermeier, D. (2008). The putative RNase P motif in the DEAD box helicase Hera is dispensable for efficient interaction with RNA and helicase activity. Nucleic Acids Res. *36*, 5800–5811.

Linder, P. (2006). Dead-box proteins: a family affair – active and passive players in RNP-remodeling. Nucleic Acids Res. *34*, 4168–4180.

Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J., and Slonimski, P.P. (1989). Birth of the D-E-A-D box. Nature *337*, 121–122. [Erratum published in Nature *340* (1989), 246].

Liu, F., Putnam, A., and Jankowsky, E. (2008). ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. Proc. Natl. Acad. Sci. USA *105*, 20209–20214.

Lorsch, J.R. and Herschlag, D. (1998a). The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. Biochemistry *37*, 2180–2193.

Lorsch, J.R. and Herschlag, D. (1998b). The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. Biochemistry *37*, 2194–2206.

Low, W.K., Dang, Y., Bhat, S., Romo, D., and Liu, J.O. (2007). Substrate-dependent targeting of eukaryotic translation initiation factor 4A by pateamine A: negation of domain-linker regulation of activity. Chem. Biol. *14*, 715–727.

Ma, E., MacRae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of human dicer by its internal helicase domain. J. Mol. Biol. *380*, 237–243.

Marintchev, A., Edmonds, K.A., Marintcheva, B., Hendrickson, E., Oberer, M., Suzuki, C., Herdy, B., Sonenberg, N., and Wagner, G. (2009). Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. Cell *136*, 447–460.

Mohr, S., Stryker, J.M., and Lambowitz, A.M. (2002). A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. Cell *109*, 769–779.

Mohr, G., Del Campo, M., Mohr, S., Yang, Q., Jia, H., Jankowsky, E., and Lambowitz, A.M. (2008). Function of the C-terminal domain of the DEAD-box protein Mss116p analyzed *in vivo* and *in vitro*. J. Mol. Biol. *375*, 1344–1364.

Morlang, S., Weglohner, W., and Franceschi, F. (1999). Hera from *Thermus thermophilus*: the first thermostable DEAD-box helicase with an RNase P protein motif. J. Mol. Biol. *294*, 795–805.

Napetschnig, J., Kassube, S.A., Debler, E.W., Wong, R.W., Blobel, G., and Hoelz, A. (2009). Structural and functional analysis of the interaction between the nucleoporin Nup214 and the DEAD-box helicase Ddx19. Proc. Natl. Acad. Sci. USA *106*, 3089–3094.

Nicol, S.M. and Fuller-Pace, F.V. (1995). The DEAD box protein DbpA interacts specifically with the peptidyltransferase center in 23S rRNA. Proc. Natl. Acad. Sci. USA *92*, 11681–11685.

Nielsen, K.H., Chamieh, H., Andersen, C.B., Fredslund, F., Hamborg, K., Le Hir, H., and Andersen, G.R. (2008). Mechanism of ATP turnover inhibition in the EJC. RNA *15*, 67–75.

Oberer, M., Marintchev, A., and Wagner, G. (2005). Structural basis for the enhancement of eIF4A helicase activity by eIF4G. Genes Dev. *19*, 2212–2223.

Ohnishi, S., Paakkonen, K., Koshiba, S., Tochio, N., Sato, M., Kobayashi, N., Harada, T., Watanabe, S., Muto, Y., Guntert, P., et al. (2008). Solution structure of the GUCT domain from human RNA helicase II/Gub reveals the RRM fold, but implausible RNA interactions. Proteins *74*, 133–144.

Pause, A. and Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. EMBO J. *11*, 2643–2654.

Pause, A., Methot, N., and Sonenberg, N. (1993). The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. Mol. Cell. Biol. *13*, 6789–6798.

Peck, M.L. and Herschlag, D. (1999). Effects of oligonucleotide length and atomic composition on stimulation of the ATPase activity of translation initiation factor eIF4A. RNA *5*, 1210–1221.

Peck, M.L. and Herschlag, D. (2003). Adenosine 5′-O-(3-thio)triphosphate (ATPgS) is a substrate for the nucleotide hydrolysis and RNA unwinding activities of eukaryotic translation initiation factor eIF4A. RNA *9*, 1180–1187.

Polach, K.J. and Uhlenbeck, O.C. (2002). Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. Biochemistry *41*, 3693–3702.

Pyle, A.M. (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. Annu. Rev. Biophys. *37*, 317–336.

Rajkowitsch, L., Chen, D., Stampfl, S., Semrad, K., Waldsich, C., Mayer, O., Jantsch, M.F., Konrat, R., Blasi, U., and Schroeder, R. (2007). RNA chaperones, RNA annealers and RNA helicases. RNA Biol. *4*, 118–130.

Rodriguez, A.C. (2002). Studies of a positive supercoiling machine. Nucleotide hydrolysis and a multifunctional latch in the mechanism of reverse gyrase. J. Biol. Chem. *277*, 29865–29873.

Rodriguez, A.C. (2003). Investigating the role of the latch in the positive supercoiling mechanism of reverse gyrase. Biochemistry *42*, 5993–6004.

Rodriguez, A.C. and Stock, D. (2002). Crystal structure of reverse gyrase: insights into the positive supercoiling of DNA. EMBO J. *21*, 418–426.

Rogers, G.W. Jr., Richter, N.J., and Merrick, W.C. (1999). Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. J. Biol. Chem. *274*, 12236–12244.

Rogers, G.W. Jr., Lima, W.F., and Merrick, W.C. (2001a). Further characterization of the helicase activity of eIF4A. Substrate specificity. J. Biol. Chem. *276*, 12598–12608.

Rogers, G.W. Jr., Richter, N.J., Lima, W.F., and Merrick, W.C. (2001b). Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. J. Biol. Chem. *276*, 30914–30922.

Rogers, G.W. Jr., Komar, A.A., and Merrick, W.C. (2002). eIF4A: the godfather of the DEAD box helicases. Prog. Nucleic Acid Res. Mol. Biol. *72*, 307–331.

Rozen, F., Pelletier, J., Trachsel, H., and Sonenberg, N. (1989). A lysine substitution in the ATP-binding site of eucaryotic initiation factor 4A abrogates nucleotide-binding activity. Mol. Cell. Biol. *9*, 4061–4063.

Rozovsky, N., Butterworth, A.C., and Moore, M.J. (2008). Interactions between eIF4AI and its accessory factors eIF4B and eIF4H. RNA *14*, 2136–2148.

Rudolph, M.G., Heissmann, R., Wittmann, J.G., and Klostermeier, D. (2006). Crystal structure and nucleotide binding of the *Thermus thermophilus* RNA helicase Hera N-terminal domain. J. Mol. Biol. *361*, 731–743.

Rudolph, M.G. and Klostermeier, D. (2009). The Thermus thermophilus DEAD box helicase Hera contains a modified RNA recognition motif loosely connected to the helicase core. RNA. Doi: 10.1261/rna.1820009.

Rudolph, M.G., Wittmann, J.G., and Klostermeier, D. (2009). Crystallization and preliminary characterization of the *Thermus thermophilus* RNA helicase Hera C-terminal domain. Acta Crystallogr. Sect. F *65*, 248–252.

Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop – a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. *15*, 430–434.

Schmid, S.R. and Linder, P. (1992). D-E-A-D protein family of putative RNA helicases. Mol. Microbiol. *6*, 283–291.

Schutz, P., Bumann, M., Oberholzer, A.E., Bieniossek, C., Trachsel, H., Altmann, M., and Baumann, U. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. Proc. Natl. Acad. Sci. USA *105*, 9564–9569.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein *Drosophila* Vasa. Cell *125*, 287–300.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structures and mechanism of helicases and nucleic acid translocases. Annu. Rev. Biochem. *76*, 23–50.

Story, R.M., Li, H., and Abelson, J.N. (2001). Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. Proc. Natl. Acad. Sci. USA *98*, 1465–1470.

Szczelkun, M.D. (2000). How do proteins move along DNA? Lessons from type-I and type-III restriction endonucleases. Essays Biochem. *35*, 131–143.

Talavera, M.A. and De La Cruz, E.M. (2005). Equilibrium and kinetic analysis of nucleotide binding to the DEAD-box RNA helicase DbpA. Biochemistry *44*, 959–970.

Talavera, M.A., Matthews, E.E., Eliason, W.K., Sagi, I., Wang, J., Henn, A., and De La Cruz, E.M. (2006). Hydrodynamic characterization of the DEAD-box RNA helicase DbpA. J. Mol. Biol. *355*, 697–707.

Tanner, N.K. (2003). The newly identified Q motif of DEAD box helicases is involved in adenine recognition. Cell Cycle *2*, 18–19.

Tanner, N.K., Cordin, O., Banroques, J., Doere, M., and Linder, P. (2003). The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. Mol. Cell *11*, 127–138.

Theissen, B., Karow, A.R., Kohler, J., Gubaev, A., and Klostermeier, D. (2008). Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. Proc. Natl. Acad. Sci. USA *105*, 548–553.

Tijerina, P., Bhaskaran, H., and Russell, R. (2006). Nonspecific binding to structured RNA and preferential unwinding of an exposed helix by the CYT-19 protein, a DEAD-box RNA chaperone. Proc. Natl. Acad. Sci. USA *103*, 16698–16703.

Tsu, C.A. and Uhlenbeck, O.C. (1998). Kinetic analysis of the RNA-dependent adenosinetriphosphatase activity of DbpA, an *Escherichia coli* DEAD protein specific for 23S ribosomal RNA. Biochemistry *37*, 16989–16996.

Tsu, C.A., Kossen, K., and Uhlenbeck, O.C. (2001). The *Escherichia coli* DEAD protein DbpA recognizes a small RNA hairpin in 23S rRNA. RNA *7*, 702–709.

UniProt-Consortium (2008). The universal protein resource (UniProt). Nucleic Acids Res. *36*, D190–D195.

von Moeller, H., Basquin, C., and Conti, E. (2009). The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. Nat. Struct. Mol. Biol. *16*, 247–254.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the a- and b-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. *1*, 945–951.

Wang, S., Hu, Y., Overgaard, M.T., Karginov, F.V., Uhlenbeck, O.C., and McKay, D.B. (2006). The domain of the *Bacillus subtilis* DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold. RNA *12*, 959–967.

Wang, S., Overgaard, M.T., Hu, Y., and McKay, D.B. (2007). The *Bacillus subtilis* RNA helicase YxiN is distended in solution. Biophys. J. *94*, L01–L03.

Yang, Q. and Jankowsky, E. (2005). ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. Biochemistry *44*, 13591–13601.

Yang, Q. and Jankowsky, E. (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. Nat. Struct. Mol. Biol. *13*, 981–986.

1.1.3 The DEAD Box Protein YxiN

YxiN is a 479 amino acid DEAD box protein from *Bacillus subtilis*. It exhibits RNA stimulated ATPase activity that goes along with a high affinity and specificity for a fragment of the 23 S ribosomal RNA (Kossen & Uhlenbeck, 1999). Hairpin 92 of this ribosomal RNA was found to be specifically recognized by the enzyme. The change of only two nucleobases in the RNA sequence abolished the high affinity binding (Kossen *et al.*, 2002). C-terminal to the generic DEAD box protein core region, YxiN features an accessory domain. This domain provides the whole enzyme with the RNA substrate specificity (Kossen *et al.*, 2002). When C-terminally fused to an unspecific DEAD box protein, the domain could even confer RNA specificity to the chimeric protein. In isolation, this RNA binding domain (RBD) also binds the RNA substrate with high affinity and specificity (Karginov *et al.*, 2005).

The crystal structure of the RBD was published in 2006 (Wang *et al.*, 2006). The 76-residue domain adopts a RNA recognition motif (RRM) fold that is characterised by a four-stranded anti-parallel β-sheet which on one side packs onto two α-helices (figure 1.1B). The RRM fold is rarely found in prokaryotes but highly abundant in eukaryotic proteins (Maris *et al.*, 2005). In available structures of RRM domains bound to RNA, the substrate is directed across the β-sheet. Arginines and lysines contact the phosphate backbone and certain nucleobases. Hydrophobic residues are in contact with the ribose and aromatic side chains and histidines undergo stacking interactions with the nucleobases. Especially these planar amino acids are conserved in various RRM domains.
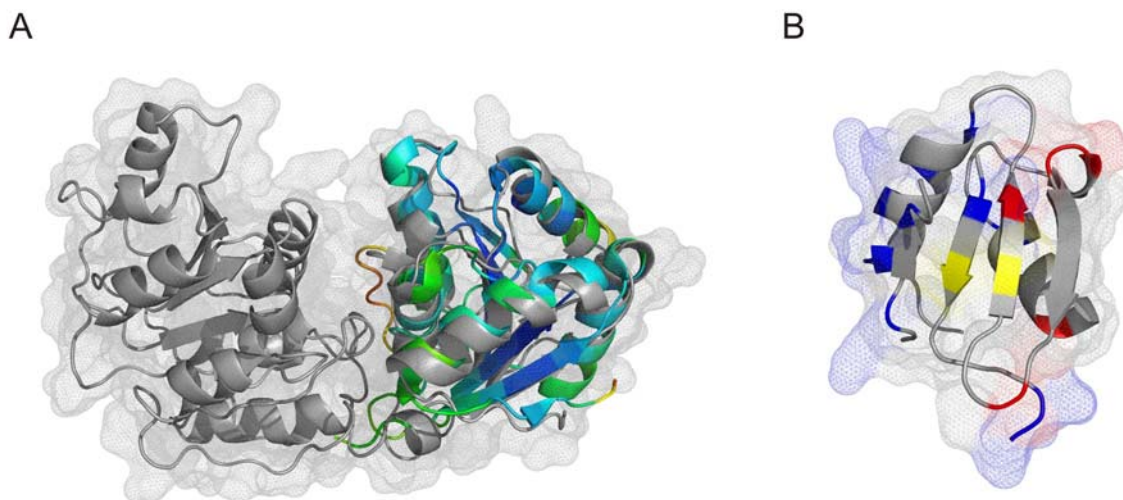
Figure 1.1
The three-dimensional structure of YxiN A) The helicase core was modelled using the structure of MjDeaD (pdbID 1hv8) and is coloured in grey. The structure of the C-terminal domain of YxiN has been determined (pdbID 2HJV) and is shown as a superimposition on the homology model. The colouring encodes the deviations of the backbone coordinates for both structures (blue: nearly identical –green–yellow– orange: large deviation). B) Structure of the YxiN RBD (pdbID 2G0C). Blue and red colouring denotes positively and negatively charged residues respectively. Tyrosines are coloured in yellow. Both tyrosines protrude from the central β-sheet to the solution and were therefore supposed to stack with nucleobases upon RNA binding. The orientation of the RBD towards the helicase core is not known.
The figures were created with PyMOL (DeLano, W.L. The PyMOL Molecular Graphics System, 2002).

Consistently, the YxiN RBD exposes two tyrosines from the surface of the β-sheet. To investigate the role of these tyrosines for RNA binding, Wang and colleagues determined the RNA affinity of a Y407C, a Y407A and a Y447C mutant (Wang *et al.*, 2006). The dissociation constants of all mutants and the wild-type were similar. Both tyrosines are thus dispensable for the high affinity RNA binding by the YxiN RBD. It was therefore proposed that the binding mode observed for many eukaryotic RRMs does not apply to YxiN. Rather the loop regions that connect the β-strands (on the basis in figure 1.1B) were suspected to recognize the hairpin 92 of the 23 S ribosomal RNA.

Structural information is also available for the C-terminal core domain of YxiN. The structure was published in 2006 (Caruthers *et al.*, 2006) and revealed a RecA-like topology highly similar to the structures of other DEAD box proteins. As discussed in chapter 1.1.2 also the N-terminal core domains of DEAD box proteins display a RecA-like fold and largely superimpose. Therefore this domain is likely to adopt the same fold in YxiN and can be modelled. Figure 1.1A (grey) shows a homology model of the YxiN core domain based on the structure of *Methanococcus jannaschii* DeaD (MjDeaD; Story *et al.*, 2001). When the crystal structure of the YxiN C-terminal core domain is superimposed, only small deviations can be observed (the extent of deviation is colour coded). In both proteins this domain displays the same topology and the locations of the secondary structure elements are similar. Deviations can be observed in particular for loop regions.

While the structure of the isolated YxiN domains is known or can be modelled, little is known about the domain orientations in the full-length enzyme. From small angle x-ray scattering data it was inferred that YxiN displays a distended shape (Wang *et al.*, 2008).

Due to its highly specific, tight binding to the hairpin 92 of ribosomal RNA it has been believed that *in vivo* YxiN plays a role in ribosome biogenesis. Up to now no experiments addressing the cellular function of YxiN have been published. However, the biological function of the *E. coli* homolog DbpA has been studied recently. In one study several *dbpA* mutants were overexpressed in *E. coli* (Sharpe Elles & Uhlenbeck, 2008). One mutant (R331A mutation in motif VI) showed a small colony phenotype in presence and the absence of endogenous wild-type DbpA. When the protein was characterised *in vitro*, the R331A mutant showed a drastically reduced ATPase activity, still bound the RNA substrate but was unable to perform unwinding.

In a subsequent study ribosome profiles of *E. coli* overexpressing the *dbpA*_R331A mutant were analysed (Sharpe Elles *et al.*, 2009). Compared to the *dbpA* wild-type overexpressing cells a smaller fraction of 70 S particles were detected in the R331A strain. Further analysis revealed that the 50 S ribosomal subunit is not properly assembled in the mutant strain. Instead, particles of 45 S accumulate that contain incompletely processed 23 S RNA and that miss some ribosomal proteins. Notably, no ribosome misassembly was observed in a Δ*dbpA* strain matching the finding that *dbpA* is a non-essential gene (Baba *et al.*, 2006; Iost & Dreyfus, 2006). In summary, it was concluded that DbpA assists the maturation of 50 S ribosomal subunits but is dispensable under the tested growth conditions. In the R331A overexpressing strain the inactive protein might have bound to the 23 S RNA in the immature subunit. But the protein was unable to catalyse RNA rearrangements and could not dissociate. As a result further assembly of ribosomes was blocked and the growth of the bacterium was affected (Sharpe Elles *et al.*, 2009).

1.2 Expressed Protein Ligation as a Bioconjugation Tool

The techniques for recombinant protein production nowadays allow for routine isolation and investigation of proteins *in vitro*. Eukaryotic proteins, however, may lack posttranslational modifications after heterologous production. Besides, biochemical studies may require the incorporation of unnatural amino acids, biophysical probes or stable isotopes into the protein of interest. This demand for modified proteins is approached by protein engineering techniques. Expressed protein ligation (EPL) constitutes one method for protein modification that, since its introduction (Muir *et al*., 1998; Evans *et al*., 1999), was employed for various applications. EPL is a semi-synthetic technique that allows for the ligation of two unprotected protein fragments or of a protein and a peptide under aqueous conditions. The following sections will describe the mechanism of EPL in detail and discuss its potential for various applications.

1.2.1 Origin and Mechanism of EPL

EPL is a regioselective and chemoselective method for the ligation of two proteins or a protein and a peptide. A native peptide bond is formed between the reactants. The ligation technique was independently developed by two groups (Muir *et al*., 1998; Evans *et al*., 1999). The development was on the one hand inspired by the catalytic mechanism of inteins (see below). On the other hand, it also bases on a ligation technique in peptide synthesis chemistry. In fact, the underlying mechanism of EPL (figure 1.2) likewise applies for native chemical ligation (NCL). NCL was introduced for the ligation of unprotected, synthetic peptides (Dawson *et al*., 1994; Tam *et al*., 1995) based on the peptide bond formation strategy of Wieland and colleagues (Wieland *et al*., 1953). As shown in figure 1.2, in NCL and EPL a protein or peptide with a N-terminal cysteine reacts with a second protein/peptide that is activated by a C-terminal thioester. A native peptide bond will be formed via the following mechanism. Initially, a reversible transesterification to the thiol of the N-terminal cysteine takes place. It is followed by a S-N acyl shift to the N-terminal amino group. The S-N acyl shift is kinetically favoured by the intramolecular nature of the reaction and the occurrence of a five-membered ring structure. The formation of an energetically favourable amide from a thioester constitutes the driving force of the overall reaction.

As already mentioned, fully unprotected proteins/peptides can be used for the ligation. If internal cysteines are present they may reversibly react with the activated carbonyl of the N-terminal fragment. But the irreversible S-N acyl shift will not take place due to an unfavourable geometry of the intermediate. As a further advantage the reaction proceeds in aqueous buffers at neutral pH. Although ligation reactions in peptide synthesis are frequently performed in the presence of 6 M guanidinium chloride the unfolding of the chains is not strictly required. Denaturation might increase the reaction yield but ligation reactions of folded protein domains have also been successful (Xu *et al*., 1999; Karow *et al.*, 2007).
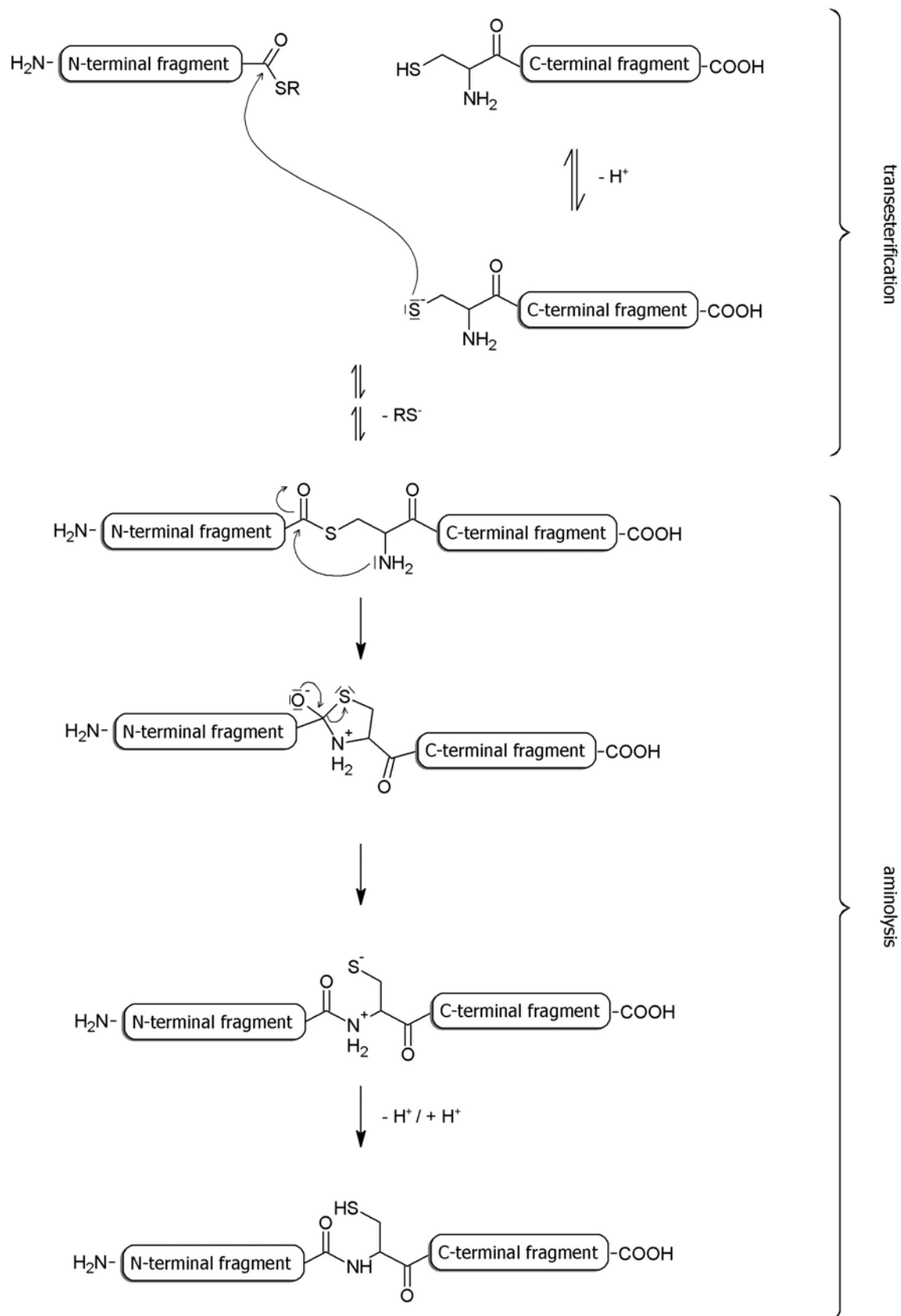
Figure 1.2
Mechanism of native chemical ligation (NCL) and expressed protein ligation (EPL). Without the need for protection groups a protein or peptide with a C-terminal thioester can be coupled to a protein or peptide that harbours a N-terminal cysteine. The initial transesterification is a reversible reaction that can also occur with internal cysteines. The proceeding N-S acyl shift is irreversible and favoured by the five-membered ring intermediate.

NCL and EPL reactions are essentially equal concerning the ligation mechanism. EPL rather constitutes an extension of the NCL method. The term NCL was coined for the ligation of fully synthetic peptides within the scope of generating long peptides (Dawson *et al*., 1994). Solid-phase peptide synthesis allows for the synthesis of peptides comprising no more than approximately 50 residues. Therefore, longer chains can be obtained via NCL of multiple segments. In EPL at least one fragment is derived via recombinant protein production. This provides several advantages. On the one hand the chain length of the reactants is not limited by the capability of solid-phase peptide synthesis. On the other hand, recombinant production allows for facile generation of the C-terminal thioester (explained below). In contrast, in NCL the synthesis of the thioester has always been a bottleneck (David *et al*., 2004). The expanded capabilities of EPL, therefore, allowed for a broad applicability of the ligation principle.

Generation of a Protein Thioester for EPL

A C-terminal thioester on the protein of interest can be generated employing engineered inteins. Inteins are protein domains that autocatalytically excise from a precursor protein after its ribosomal synthesis (for a detailed review Evans & Xu, 2002). Upon excision (also termed splicing) the flanking protein sequences (exteins) become covalently linked. The self-cleavage activity of inteins via a (thio)ester intermediate can be employed for the generation of thioester-tagged proteins.

For this purpose, the protein of interest is fused at the N-terminus of an engineered intein (figure 1.3A). At the junction the intein displays a serine or cysteine. The intein will catalyse the N-O(S) acyl shift of the preceding amide carbonyl onto the hydroxyl (thiol) of these amino acids. The intein is engineered in such a way that the breakdown of the resulting branched intermediate is prevented. Upon treatment with a thiol in solution the bond is cleaved and a thioester with the thiol component is formed at the target protein.
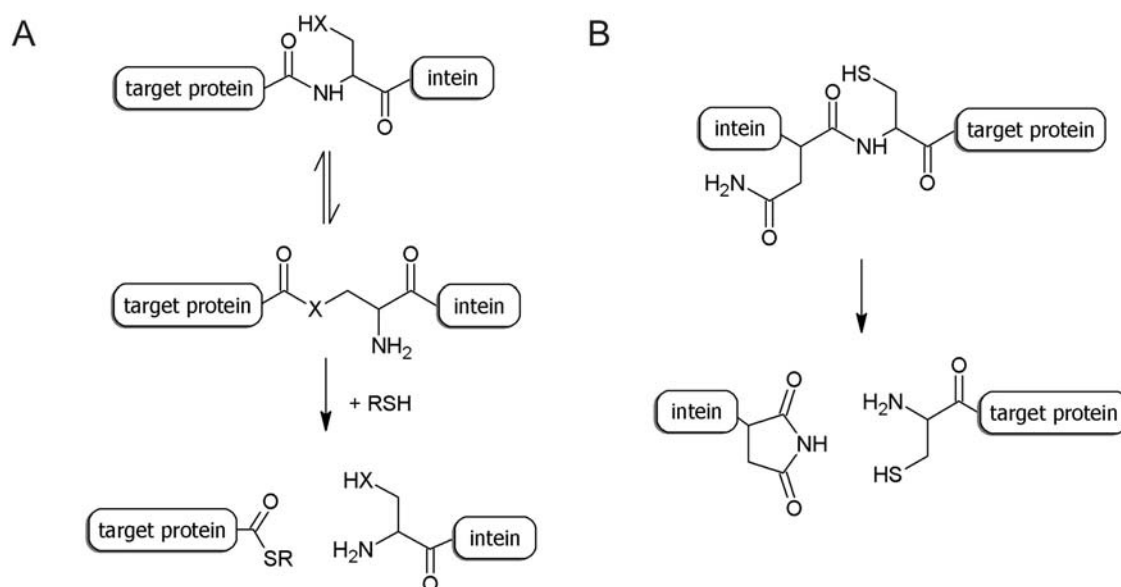
Figure 1.3
Intein-mediated generation of protein reactants for EPL. A) X = O or S. The target protein is fused to a mono-functional intein that can catalyse a reversible N-S(O) acyl shift of the carbonyl at the junction to the terminal cysteine or serine of the intein. Upon addition of a thiol in excess transesterification takes place and the target protein is released. B) The target protein, starting with any amino acid, is fused to the C-terminus of an engineered intein. Upon asparagine-cyclization the target protein is released and displays the desired N-terminal residue (a cysteine here). The cleavage reaction can be triggered via pH and temperature change.
Both strategies are often performed using affinity tagged inteins. This allows for purification of the fusion protein via affinity chromatography and facile elution of the target protein after the cleavage reaction.

Generation of a N-terminal Cysteine in a Recombinant Protein

During translation a N-terminal methionine is incorporated into any protein. Proteins starting with a different amino acid can be derived via proteolysis of a precursor protein (Erlanson *et al.*, 1996), via *in vivo* processing (Iwai & Plückthun, 1999) or employing engineered inteins (Evans *et al.*, 1999). For the latter strategy the protein of interest bearing the desired terminal residue is fused to the C-terminus of the intein (figure 1.3B). The intein will catalyse its cleavage from the fusion protein releasing the target protein. The cleavage reaction is very slow at 4 °C and slightly basic pH. It can therefore be triggered via a change to pH 7. Incubation at 25 °C further accelerates the reaction but is not strictly required (Xu & Evans, 2001).

For both, the generation of protein thioesters and proteins with a N-terminal cysteine, various systems are also commercially available that provide vectors with suitable intein constructs (Xu & Evans, 2001).

1.2.2 Applications of EPL – Variations of a Common Theme

Within the last decade the utilisation of EPL assisted various studies on protein structure and function. Employing EPL, it has been possible to incorporate spectroscopic probes and to introduce posttranslational modifications such as phosphorylations, glycosylations and lipid moieties (Cotton & Muir, 2000; Mukhopadhyay *et al.*, 2001; Huse *et al.*, 2001; Macmillan & Bertozzi, 2000; Alexandrov *et al.*, 2002). Cyclic proteins and proteins containing unnatural amino acids or even entire backbone mimetics have been generated (Iwai & Plückthun, 1999; Valiyaveetil *et al.*, 2004; Arnold *et al.*, 2002). In addition, the selective isotopic labelling of multi-domain proteins for NMR studies was achieved (Xu *et al.*, 1999; Camarero *et al.*, 2002). In the following, a few examples will be discussed in detail. In explaining the respective underlying methodology, the versatile nature of the EPL approach is emphasised. The choice of the starting reagents renders the basic reaction of a thioester and a cysteine-containing moiety to a multi-faceted methodology.

In a study on the mechanism of the potassium channel KcsA EPL was employed to introduce a D-amino acid at a position inside the selectivity filter of the channel (Valiyaveetil *et al.*, 2004). The selectivity filter contains a conserved glycine (G77) that is in an unusual left-handed helical conformation. A mutation of this gylcine to alanine inactivates KcsA. It was concluded that a glycine at this position is essential for the filtering function. In addition, it had been speculated that the conformation required for selective ion passage can be solely adopted by a glycine or a D-amino acid. Indeed, when D-alanine was incorporated on position 77 in a semi-synthetic approach the channel retained its function. In this remarkable study on a structure-function relationship a N-terminal fragment of the KcsA protein (residues 1 to 68) was derived via bacterial protein production as an intein fusion that allowed for the attachment of a C-terminal thioester. The C-terminal fragment (residues 69 to 122) was synthesized on solid-phase incorporating the D-alanine on position 77. Hence, in the subsequent EPL reaction a synthetic peptide was coupled to a recombinantly produced protein fragment.

It is not required to react a protein or peptide to the thioester molecule. As employed by the Ebright group, a cysteine only can serve as the C-terminal fragment (Mukhopadhyay *et al.*, 2001; Mekler *et al.*, 2002). In studies on the mechanism of transcription initiation, RNA polymerase was labelled with a cysteine-fluorophore moiety at the C-terminus. Employing an intein, the RNA polymerase core had been derived as a C-terminal thioester and was subsequently labelled with a cysteine-fluorescein moiety in an EPL reaction. Later on, the fluorescein served as a donor fluorophore for FRET experiments in which the σ^{70} factor or DNA were labelled with acceptor fluorophores. From the data the group was able to describe the structural organisation of the polymerase-

promoter DNA open complex and to detect the association of the σ^{70} factor with the polymerase during translocation.

Approaches ligating a synthetic N-terminal fragment to a protein of recombinant origin have also been successful. In an attempt to generate a homogenously phosphorylated protein Huse and colleagues performed the semi-synthesis of the type I transforming growth factor β receptor (Huse *et al*., 2001). A tetraphosphorylated synthetic peptide with a C-terminal thioester was coupled to a fragment of the receptor that had been derived from a baculoviral expression system. The phosphorylation displays a posttranslational modification that was thought to activate the receptor *in vivo*. In general, proteins containing defined, homogenous posttranslational modifications are hardly obtainable from heterologous expression systems. Employing the EPL approach the molecular mechanism of the receptor activation could be studied *in vitro* on a defined system.

In the next example two entire domains were ligated in an EPL reaction (Xu *et al*., 1999). The respective study dealt with the regulatory element of the Abl protein tyrosine kinase. This element is formed by a SH3 domain and a proceeding SH2 domain. The SH2 domain was separately produced by bacteria in ^{15}N media. In contrast, the SH3 domain was produced in bacteria without isotope labelling and was purified as a C-terminal thioester. Upon EPL of SH3 and SH2 the two-domain regulatory apparatus of the kinase could be reconstituted. The ligation product allowed for NMR experiments on the SH2 domain in its native context of the SH3 domain. This pioneering work suggested an approach to overcome the practical size limit for protein structure determination by NMR. For large proteins the application of segmental isotopic labelling results in a decreased spectral complexity facilitating the data analysis.

The generation of a N-terminal cysteine and a C-terminal thioester on the same protein allows for the backbone cyclisation of a target protein. In general, cyclic proteins are of interest for medical applications. Cyclisation is thought to increase the stability of pharmaceutically relevant proteins. Furthermore, proteins can be generated that are locked in an active conformer. As an example, a cyclic β-lactamase could be derived via intramolecular EPL of a precursor protein (Iwai & Plückthun, 1999). This engineered β-lactamase exhibited increased stability against exoproteolysis and to heat-precipitation.

All in all, the exemplarily discussed studies demonstrate the versatility of EPL in protein studies. Since EPL was introduced in the late 1990s, the investigation of various biological issues has been assisted by this protein engineering technique.

1.3 Aims of Research

The aim of this work is to characterise the conformation of the *B. subtilis* DEAD box protein YxiN with respect to the orientation of its three protein domains. Moreover, changes of the enzyme conformation during the catalytic cycle should be detected and linked to the enzyme function. The results should reveal the mechanism of the RNA helicase YxiN. For this purpose a combination of biochemical and biophysical methods should be employed. In particular, fluorescence resonance energy transfer (FRET) experiments at the single molecule level should be utilised to probe intra protein distances and thus the protein conformation.

It is believed that DEAD box proteins undergo conformational changes during their catalytic cycle. Results from proteolysis experiments with YxiN's *E. coli* homolog DbpA already indicated the occurrence of different conformers (Henn *et al.*, 2002). The conformation (and the putative conformational change) of the YxiN helicase core during the catalytic cycle should be investigated directly using FRET experiments. The characterisation of YxiN should furthermore cover the protein's ATPase activity, RNA unwinding activity and their coupling.

It had been shown that the mutation of conserved motifs can reveal insights into the role of these motifs for the ATPase and the unwinding activity in DEAD box proteins (Pause & Sonenberg, 1992; Pause *et al.*, 1993). In the current project, motif mutants shall also be employed to answer the following questions. How do the mutations influence the coupling of ATPase and RNA unwinding activity? Is there an impact on the protein conformation? The results should allow for a deeper understanding of the DEAD box protein mechanism in general.

The structure of the helicase core is known for many DEAD box proteins. But little information is available regarding the orientation of accessory domains to the core. The current project should also approach the orientation of the YxiN RNA binding domain to the core. The isolated structures of the C-terminal core domain and of the RNA binding domain have been published (Caruthers *et al.*, 2006 and Wang *et al.*, 2006 respectively). Employing single-molecule FRET experiments distances between the domains shall be measured and the orientation of the domains should be mapped. The information provides a basis for future studies that address the action of the RNA binding domain during RNA unwinding.

For single-molecule FRET experiments protein samples need to be labelled with two different fluorophores. Expressed protein ligation (EPL) should be used to develop a protocol that allows for site-specific double labelling of the protein of interest. In addition, a sample preparation technique for single-molecule FRET experiments on a TIR-microscope (total internal reflection microscope) shall be set up. Here, the molecule of interest needs to be immobilized on a surface. Making use of EPL a procedure should be developed that allows for attachment of the target protein without affecting its activity. Single-molecule FRET experiments on a TIR-microscope will be useful to study the protein conformation on seconds time scale thereby deepening the understanding of how the DEAD box protein catalyses structural conversions of RNA.

2 Materials and Methods

2.1 Reagents and Enzymes

Aldrich (Steinheim, D) $MgSO_4$, 99 %; triisopropylsilane, 99 %

AppliChem (Darmstadt, D) glycerol, anhydrous, p.A.; L-glutathione, reduced; $MnCl_2x4\ H_2O$

Bio-Rad Laboratories (Hercules, USA) Avidin-HRP; 4-chloro-1-naphthol (HRP color development reagent)

Carl Roth (Karlsruhe, D) NaCl > 99.8 %; peptone; yeast extract, pulv.

EPICENTRE Biotechnologies (Madison, USA) T7 R&DNA Polymerase

Fermentas (St. Leon-Rot, D) PageRuler Plus prestained protein ladder; RNA loading dye, 2x

Finnzymes Oy (Espoo, FIN) Phusion High-Fidelity DNA Polymerase (2000 U/ml), 5x Phusion HF Buffer

Fluka (Buchs, CH) activated charcoal, purum; DCM; diethyl ether, p.A.; DMF, absolute; formic acid; methanol, p.A.; 4-methylmorpholine, purum; Silica Gel 60; TFA; thiophenol, purum

GERBU Biochemicals (Gaiberg, D) ampicillin sodium salt

Invitrogen (Eugene, USA) Alexa Fluor 488 C5-maleimide; Alexa Fluor 546 C5-maleimide; Alexa Fluor 555 C2-maleimide; streptavidin; TCEP; TMR C5-maleimide; TEV protease

Jena Bioscience (Jena, D) mant-ADP, > 95 %; mant-ADPNP, > 90 %

Laysan Bio (Arab, USA) Biotin-PEG-SVA-5000, mPEG-Succinimidyl Valerate, MW 5,000

Merck (Darmstadt, D) N-α-t-Boc-S-trityl-L-cysteine, $\geq$ 98 %; ethanol, absolute; MESNa, for synthesis; potassium acetate, puriss.; 2-propanol, p.A.; TBTU, $\geq$ 98 %; HOBt

New England BioLabs (Ipswich, USA) BSA (10 mg/ml); DNA ladder, 1 kb, 100 bp; *Dpn*I (20,000 U/ml); *Nde*I (20,000 U/ml); NEBuffer 4; *Sap*I (2000 U/ml); T4 DNA Ligase (400,000 cohesive end U/ml); T4 DNA Ligase Reaction Buffer; Thermostable Inorganic Pyrophosphatase (2,000 U/ml)

Pharma Waldhof (Düsseldorf, D) ATP and ADP disodium salt

Promega (Madison, USA) GoTaq DNA polymerase (5 U/μl); *Pfu* DNA polymerase (2-3 U/μl); *Pfu* DNA polymerase 10x reaction buffer; SAP (1 U/μl); 10x SAP buffer; Wizard SV Gel and PCR Clean-Up System

Qiagen (Hilden, D) QIAprep Spin Miniprep Kit (250); QIAquick Nucleotide Removal Kit (50)

Radiant Dyes Laser Acc. (Wermelskirchen, D) *Uranin* (fluorescein)

Reuss-Chemie (Tägerig, CH) 2-propanol, tech.

Roche Diagnostics (Mannheim, D) Complete EDTA-free Protease Inhibitor Cocktail Tablets; hexokinase, from yeast "overproducer"; lactate dehydrogenase from rabbit muscle; NADH disodium salt; PEP; proteinase K, recombinant, PCR grade; pyruvate kinase from rabbit muscle

Sigma Aldrich (St. Louis, USA) ADPNP tetralithium salt hydrate, ~ 95% (HPLC); Albumin, biotin labelled bovine, lyophilized powder; ATP disodium salt, Grade I, $\geq$ 99 %; $CoCl_2x\ 6\ H_2O$; $CuCl_2x\ 2\ H_2O$; CTP disodium salt, $\geq$ 95 %; deoxynucleotide Set (dATP, dCTP, dGTP, dTTP sodium salts); DMSO $\geq$ 99.9 %; ethanol, purum; $FeCl_3x\ 6\ H_2O$; GTP sodium salt hydrate, $\geq$ 95 % (HPLC); $Na_2MoO_4x\ 2\ H_2O$, spermidine, min. 98 % (GC); N-[3-(Trimethoxysilyl)propyl]ethylenediamine, 97 %; UTP trisodium salt hydrate, Type IV, 90-95 %

Thermo Fisher Scientific (Rockford, USA) 5-(biotinamido)pentylamine; Biotin Quantitation Kit

VWR (Briare, F) glycerol ~ 87 %

Chemicals not listed above, were purchased in p.a. or puriss. quality from Carl Roth (Karlsruhe, D).

Aqueous solutions were made using ultrapure water, purified by the Nanopure Cartridge System (Skan AG, Basel, CH). Stock solutions were filtered with sterilizing membrane filters (Durapore 0.22 μm Millipore, Billerica, USA).

2.2 Oligonucleotides

DNA oligonucleotides were purchased from PURIMEX (Grebenstein, D), MWG Biotech (Ebersberg, D), Microsynth (Balgach, CH) or Sigma Aldrich (St. Louis, USA). RNA oligonucleotides were synthesized by Dharmacon (Lafayette, USA), Microsynth (Balgach, CH) or PURIMEX (Grebenstein, D). The sequences are noted in 5' to 3' direction.

Whenever suitable sequencing primers could not be provided by the sequencing company, following DNA primers were added. Mxe-intein_rev was also used in PCR screens.

Ssp-DnaB-intein_for: ACT GGG ACT CCA TCG TTT CT

Mxe-intein_rev: GGC ACG ATG TCG GCG ATG

M70_rev: TCA GCG GTG GCA GCA GCC AAC TCA

YxiN and *yxiN*_C61A_C267A_A115C_D262C were cloned into pTWIN1 using

YxiN_1-211_for: GGT GGT CAT ATG AGT CAT TTT AAA AAC TAT C (forward primer)

YxiN_GS_Sap: GGT GGT TGC TCT TCC GCA GGT GCC GCC GCT GCC GCC GCC GCT GCC GCC GCC TTT ATT CGC TTT ATT CAC C (reverse primer)

Following DNA primers were used in mutagenesis PCR and/or in PCR screens.

YxiN_Q29A_for: GAA CCG ACC AAG GTGGCG CAG AGT GTG ATC C

YxiN_Q29A_rev: GGA TCA CAC TCT GCG CCA CCT TGG TCG GTT C

YxiN_Q47C_for: CTT GTC GTC AAA TCA TGA ACA GGA AGC GGG

YxiN_Q47C_rev: CCC GCT TCC TGT GCA TGA TTT GAC GAC AAG

YxiN_K52Q_for: GAC AGG AAG CGG GCA AAC GGC TTC GTT CGG G

YxiN_K52Q_rev: CCC GAA CGA AGC CGT TTG CCC GCT TCC TGT C

YxiN_C61A_for: CGG GAT TCC TCT CGC GGA GCT GGC GAA TTG GG

YxiN_C61A_rev: CCC AAT TCG CCA GCT CCG CGA GAG GAA TCC CG

YxiN_K70C_for: GGG ATG AAA ACT GTC CGC AGG CGC TTA TTT TAA CAC C

YxiN_S108C_for: CA GCC GTA TTT GGA AAA TGC TCC TTT GAT AAA CAA A

YxiN_S108C_rev: TTT GTT TAT CAA AGG AGC ATT TTC CAA ATA CGG CTG

Yxin_N158C_for: GGA TGA GAT GCT GTG TAT GGG CTT CAT TGA GC

Yxin_N158C_rev: GCT CAA TGA AGC CCA TAC ACA GCA TCT CAT CC

YxiN_S182A/T184A_for: CGA TGC TGT TTG CCG CGG CGC TTC CGC AGG

YxiN_S182A/T184A_rev: CCT GCG GAA GCG CCG CGG CAA ACA GCA TCG

YxiN_Q187C_for: GCG ACG CTT CCG TGC GAT ATC GAG AAG C

YxiN_Q187C_rev: GCT TCT CGA TAT CGC ACG GAA GCG TCG C

YxiN_T212C_rev: CCG CAT GTT CAA TAT TTC TAC AGG TTA GGC CGG CCG C

YxiN_E224C_for: GTG ATT CAA GTA AGA TGC GAG AAT AAG TTT TC

YxiN_S229C_for: GAA GAG AAT AAG TTT TGC TTG CTG AAA GAT GTG

YxiN_S229C_rev: CAC ATC TTT CAG CAA GCA AAA CTT ATT CTC TTC

YxiN_T237C_for: GAT GTG CTG ATG TGC GAG AAT CCC GAC

YxiN_D262C_for: CGA TGA ATT GGA TTG CTT GGG ATA TCC

YxiN_D262C_rev: GGA TAT CCC AAG CAA TCC AAT TCA TCG

YxiN_C267A_for: GGG ATA TCC AGC GGA TAA AAT TCA CGG CGG

YxiN_C267A_rev: CCG CCG TGA ATT TTA TCC GCT GGA TAT CCC

YxiN_R288C_rev: CGG TAC TCG CCA CAT TTA AAT TCA TTC ATG

YxiN_I275C_for: CAC GGC GGA ATG TGT CAG GAA GAC CG

YxiN_I275C_rev: CGG TCT TCC TGA CAC ATT CCG CCG TG

YxiN_E290C_for: GAA TGA ATT TAA ACG TGG CTG TTA CCG TTA CTT AGT GG

YxiN_E290C_rev: CCA CTA AGT AAC GGT AAC AGC CAC GTT TAA ATT CAT TC

YxiN_G303A_for: CGC CGC GCG CGC GAT TGA TAT TGA AAA TAT C

YxiN_G303A_rev: GATATTTTCAATATCAATCGCGCGCGCGGCG

YxiN_K337C_for: GGG ACG CGC AGG GAA CTG TGG AAA GGC CAT TTC G

YxiN_K337C_rev: CGA AAT GGC CTT TCC ACA GTT CCC TGC GCG TCC C

YxiNRBD_405K_for: CAG GGC GCC ATG AAG CTG TAT TTT AAC GGC

YxiNRBD_405K_rev: GCC GTT AAA ATA CAG CTT CAT GGC GCC CTG

YxiN_D429C_for: GGA ACG ATT GCC AAA ATT TGC GGT GTG TCA GCT GAC G

YxiN_D429C_rev: CGT CAG CTG ACA CAC CGC AAA TTT GGC AAT CGT TTC C

YxiN_T440C_for: CTG ACG ACA TCG GCA TCA TCT GCA TCA TGG ATA ACG CCT C

YxiN_T440C_rev: GAG GCG TTA TCC ATG ATG CAG ATG ATG CCG ATG TCG TCA G

YxiN_N444C_for: GCA TCA TCA CGA TCA TGG ATT GCG CCT CAT ACG TTG AG

YxiN_N444C_rev: CTC AAC GTA TGA GGC GCA ATC CAT GAT CGT GAT GAT GC

YxiN_N452C_for: CGT TGA GAT TTT ATG TGG TAA AGG CCC TCA TGT TC

YxiN_N452C_rev: GAA CAT GAG GGC CTT TAC CAC ATA AAA TCT CAA CG

YxiN_N464C_for: CAT GTT CTC AAA GTG ATG AAG TGC ACA ACC GTC AAA GG

YxiN_N464C_rev: CCT TTG ACG GTT GTG CAC TTC ATC ACT TTG AGA ACA TG

YxiN_L472C_for: CCG TCA AAG GGA AAC AGT GCA AGG TGA ATA AAG C

YxiN_L472C_rev: GCT TTA TTC ACC TTG CAC TGT TTC CCT TTG ACG G

The DNA primers listed below were used in PCR screens.

pGEX_for: GGG CTG GCA AGC CAC GTT TGG TG

T7_promotor: TAA TAC GAC TCA CTA TAG GG

T7_terminator: GCT AGT TAT TGC TCA GCG G

The DNA template used in the *in-vitro* transcription reaction was generated using the following primers.

IVT_for: TAA TAC GAC TCA CTA TAG GCT TAT CTC CCC CAA G

IVT_rev: GAA CTG TCT CAC GAC GTT C

The following RNA oligonucleotides were used in unwinding assays.

32mer: CGA GGU CCC AAG GGU UGG GCU GUU CGC CCA UU

9mer: UUG GGA CCU

9mer-F: UUG GGA CCU(-F)   F denotes a fluorescein label at the 3' end of the strand.

2.3 Consumables

Amersham Biosciences Europe (Freiburg, D) SEC columns NAP-10 and NAP-5 (Sephadex G-25)

BD (Franklin Lakes, USA) 50 ml conical tubes, PP; Microlance-3 needles, 0.90 x 40 mm; serological pipettes 1 ml, 25 ml

Bio-Rad Laboratories (Hercules, USA) chitin resin beads in 20 % ethanol; Econo-Pac Columns, Micro Bio-Spin 30 Columns, RNase-free; Micro Bio-Spin 6 Columns

Brand (Wertheim, D) UV-cuvettes PLASTIBRAND micro, z = 8.5 mm, 2 x 3.5 mm

CODAN Medical (Rødby, DK) single use syringes 2 ml, 10 ml

Eppendorf (Hamburg, D) Safe-Lock tubes ambra 1,5 ml; Safe-Lock tubes 2 ml

Erie Electroverre SA (Romont, CH) Micro slides 76 x 26 mm; 0.8-1 mm, white precleaned

Fluka (Buchs, CH) TLC cards with fluorescent indicator

G. Kisker (Steinfurt, D) Quali-'Low Retention' tubes, 1.7 ml

Life Systems Design (Merenschwand, CH) Gel-Saver II pipette tips 1-200 μl

Merck (Darmstadt, D) Neutralit pH-indicator strips pH 5.0-10.0

Paul Marienfeld (Lauda-Königshofen, D) Cover slips 24 mm x 32 mm (1.5)

Sarstedt (Nürnbrecht, D) 15 ml conical tubes, PP; cuvettes 10 x 4 x 45 mm; Filtropur S 0.2 μm; Micro Tube 1.5 ml, PP; Multiply-Pro cup, 0.2 ml, PP

Sartorius Stedim Biotech (Aubagne, F) filter Midisart 2000; Vivaspin 20 concentrator 10,000 MWCO PES; Vivaspin 6 concentrator 10,000 MWCO PES; Vivaspin 500 concentrator 10,000 MWCO PES; Vivaspin 15R concentrator 2,000 MWCO

Schleicher&Schuell (Dassel, D) ELUTRAP BT-1 membranes; ELUTRAP BT-2 membranes; nitrocellulose transfer membrane Protran BA 85

Spectrum Laboratories (Rancho Dominguez, USA) Spectra/Por dialysis membrane 12-14,000 MWCO and 3,500 MWCO

Starlab (Ahrensburg, D) TipOne Tips 10 μl, 200 μl, 1000 μl

Thermo Fisher Scientific (Rockford, USA) Lab-Tek chamber slides, 8 chambers, glass slide

Treff AG (Degersheim, CH) Microtubes ClickFit, 1.5 ml

2.4 Instrumentation

Centrifugation Heraeus Biofuge pico with rotor 33258; Sorvall RC 5C plus with rotors SLA3000 and SS34 (both Kendro Laboratory Products, Langenselbold, D); centrifuge 5402 with rotor F-45-18-11 (Eppendorf, Hamburg, D)

Electrophoresis and Western blotting Hoefer multiple gel caster; gel electrophoresis unit Mighty Small II (8 x 7 cm); power supply EPS 300, 301 and 1001 (all of them Amersham Biosciences Europe, Freiburg, D); gel electrophoresis unit NA-L 200 (LTF Labortechnik, Wasserburg, D); protein blotting cell Mini Trans-Blot (Bio-Rad Laboratories, Hercules, USA); gel documentation system EDAS 290 (Kodak, Stuttgart, D) equipped with Transilluminator FT-20/254/365 (Herolab, Wiesloch, D)

Fluorescence measurements fluorimeter FluoroMax-3 (HORIBA JOBIN YVON, Stanmore, UK) and precision cells made of Quartz SUPRASIL for fluorescence measurements no. 105.250, path 10 x 2 mm (Hellma, Müllheim, D)

Absorption measurements Ultrospec 2100 pro (Amersham Biosciences Europe, Freiburg, D) and precision cells made of Quartz SUPRASIL for absorption measurements no. 105.202, path 10 mm (Hellma, Müllheim, D); UV/Vis photometer BioPhotometer (Eppendorf, Hamburg, D)

LC-ESI-MS HP 1100 HPLC, (Agilent, Santa Clara, USA) equipped with RP columns (BioSuite C18 PA-B, Waters, Milford, USA; GraceVydac Protein C4 350 µl, Grace, Deerfield, USA) and connected to a microTOF (Bruker Daltonik, Bremen, D)

Analytical HPLC purifier (GE Healthcare, Chalfont St. Giles, UK) equipped with a C18-RP column (Prontosil 120-5-C18, Bischoff Chromatography, Leonberg, D)

Preparative HPLC LC system ÄKTAprime (GE Healthcare, Chalfont St. Giles, UK) equipped with SEC columns HiLoadTM 16/60 Superdex 200 pg and Superdex 75 pg, resins Glutathione Sepharose 4 FF and Chelating Sepharose FF, LC column XK 16, Superloop 50 ml (all Amersham Biosciences Europe, Freiburg, D)

Preparative RP-HPLC LC-20A prominence (Shimadzu, Duisburg, D) equipped with a semipreparative C18 RP column (Chromolith Performance RP-18 endcapped 100-4.6; Merck, Darmstadt, D)

MALDI-MS analysis Voyager-DE PRO (Applied Biosystems, Foster City, USA)

Electroelution system Elutrap (Schleicher&Schuell, Dassel, D)

Membrane pHmeter HI 8314 (Hanna instruments, Kehl, D)

Microfluidizer M-110L (Microfluidics, Newton, USA)

PCR machine Mastercycler gradient (Eppendorf, Hamburg, D)

Shaker Minitron (Infors, Bottmingen, CH)

Thermostat ThermoStat plus (Eppendorf, Hamburg, D)

Confocal Fluorescence Microscope for SmFRET Measurements

FRET events from single molecules freely diffusing in solution were detected on a home-built confocal microscope. The excitation light was generated by a modelocked, solid-state Ti:sapphire laser (Tsunami, Spectra-Physics, Mountain View, USA) that was pumped by a continuous wave Nd:VO$_4$ diode laser (Millennia pro, Spectra-Physics). The emitted light of 951 nm passed a frequency doubler (Model 3980, Spectra-Physics) and a frequency doubler/tripler (GWU, Spectra-Physics) resulting in light of 475 nm. Passing a $\lambda/2$ plate, an adjustable neutral density filter and a notch filter the beam was directed into an inverted microscope (IX71, Olympus, Tokyo, J). The beam was then focussed into the sample by a water immersion objective (UPlanApo, N_A = 1.2, Olympus). The fluorescence was collected by the same objective and transmitted to the detection part via a dichroic beam splitter (DM505; Olympus). A lens focused the light to a pinhole (100 µm) selecting the in-focus light only. After passing a second lens the light was split via a dichroic beam splitter to an APD (SPCM-AQR-14, PerkinElmer, Dumberry, CAN) detecting donor photons and a second APD detecting acceptor photons. The signal from the APDs was sent via a router to a single photon counting card (SPC-630, Becker & Hickl, Berlin, D) and processed.

TIR-microscope for SmFRET Measurements

Fluorescence from single molecules immobilized on a surface was recorded using a home-built TIR-microscope. The excitation light was generated by a solid-state cyan laser (Cyan CDRH, Spectra-Physics, Mountain View, USA) emitting continuous wave light of 488 nm. To fill the back aperture of the objective, the beam was

expanded via a telescope. The light was then transmitted through an adjustable neutral density filter, a λ/4 plate and a focussing lens. Then the light entered an inverted microscope (BXFM components, Olympus, Tokyo, J) equipped with an oil immersion objective (UIS2 PLAPON 60XOTIRFM, N_A = 1.45, Olympus) under TIR conditions ($\theta > 62°$). Emission light was collected by the same objective and transmitted to the detection part via a dichroic beam splitter (505DCXR, Chroma, Bellows Falls, USA). The light passed a tube lens, a notch filter (NFU-488.0-25, MK Photonics, Albuquerque, USA) and was splitted via a dichroic beam splitter (565DCXR, Chroma) into an acceptor and a donor path. On each path the light passed filters (donor: D535/40m; acceptor: 570LP and 610/75m, all Chroma) and a lens that focused the beam on a CCD camera (sensicam qe 670KS, PCO, Kelheim, D). Donor and acceptor beam were imaged side by side onto one CCD camera via a dichroic beam splitter (565DCXR).

2.5 Plasmids and Bacterial Strains

Ahead of this project, nucleotides 2481-2634 of *B. subtilis* 23S rRNA had been cloned into the *Bam*HI/*Hin*dIII restriction sites of pUC18 (Fermentas, St. Leon-Rot, D). This construct was used for *in vitro* transcription of a 153mer RNA.

YxiN and *yxiN*_C61A_C267A_A115C_D262C were cloned into the *Nde*I/*Sap*I sites of pTWIN1 (New England Biolabs, Ipswich, USA) as described in section 2.7.

Before the project started, *yxiN*1-211 and *yxiN*C212-479 had been cloned into the pTWIN1 vector (*Nde*I/*Sap*I and *Sap*I/*Pst*I sites, respectively). For the production of wild type and mutant YxiN, YxiN1-386 and YxiN404-479, constructs cloned into the restriction sites *Nco*I/*Xho*I of pETM30 (G. Stier, EMBL, Heidelberg, D) were used.

In *yxiN*404-479 amino acid 405 was changed due to the restriction sites for cloning. Subsequently it was mutated to the native lysine as indicated by the name of the construct *yxiN*404-479_405K.

In *yxiN* and *yxiN*1-386 amino acid 2 was changed due to the restriction sites for cloning. The constructs harbour a S2G mutation.

All plasmids were propagated in *E. coli* XL1-Blue (Stratagene, La Jolla, USA).
recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'*proAB lacI*q*ZDM15 Tn10* (Tet^r)]

The proteins YxiN404-479_405K and YxiN_C61A_C267A_A115C_S229C_K52Q were produced in strain *E. coli* Rosetta 2 (DE3) (Merck, Darmstadt, D)
F^- *ompT hsd*S_B(r_B^- m_B^-) *gal dcm*(DE3) pRARE2 (*argU, argW, argX, glyT, ileX, leuW, metT, proL ,thrT, thrU, tyrU;* Cam^r).

All other proteins were purified from strain *E. coli* Rosetta (DE3) (Merck, Darmstadt, D)
F^- *ompT hsd*S_B(r_B^- m_B^-) *gal dcm*(DE3) pRARE (*argU, argW, glyT, ileX, leuW, metT, proL ,thrT, thrU, tyrU;* Cam^r).

Culture Media

Bacteria were cultivated in LB medium (10 g peptone, 5 g yeast extract, 10 g NaCl per 1 l medium) or on LB plates (1 l medium + 19 g agar). Depending on the construct (see section 2.9.1) auto induction medium (Studier, 2005) or LB medium was used for protein production. All media were supplemented with appropriate antibiotics (20 μg/ml kanamycin, 100 μg/ml ampicillin and/or 5 μg/ml chloramphenicol).

2.6 General Methods

2.6.1 Agarose Gel Electrophoresis

DNA fragments up to 1 kb or DNA larger than 1 kb were separated in 2 % or 0.8 % agarose gels respectively. The samples were supplemented with 6x loading buffer (10 mM Tris/HCl, pH 7.6, 60 mM EDTA, 60 % (v/v) glycerol, 0.03 % (w/v) bromophenol blue, 0.03 % (w/v) xylencyanol) before the electrophoresis was performed for 20 min at 8 V/cm in 0.5x TBE). Finally the gel was stained in ethidium bromide solution and the DNA was visualised under UV light.

2.6.2 SDS-PAGE (discontinuous)

Stacking gel	125 mM Tris/HCl, pH 6.8 (25 °C); 0.06 % (w/v) SDS; 4.5 % acrylamide; 0.1 % (v/v) *N,N'*-methylenebisacrylamide; polymerized by addition of 0.1 % (w/v) APS and 0.1 % (v/v) TEMED
Resolving gel	375 mM Tris/HCl, pH 8.8 (25 °C); 0.1 % (w/v) SDS; 15 % (v/v) acrylamide; 0.4 % (v/v) *N,N'*-methylenebisacrylamid; polymerized by addition of 0.086 % (w/v) APS and 0.057 % (v/v) TEMED
Running buffer	24 mM Tris base; 0.1 % (w/v) SDS; 0.2 M glycin
4x loading buffer	130 mM Tris/HCl, pH 6.8 (25 °C); 200 mM DTT; 20 % (v/v) glycerol; 4 % (w/v) SDS; 0.01 % (w/v) bromophenol blue
Staining solution	50 % (v/v) methanol; 10 % (v/v) acetic acid; 0.1 % (w/v) Brillant Blue R-250
Destaining solution	20 % (v/v) ethanol; 10 % (v/v) acetic acid

Protein samples were analysed on denaturing polyacrylamide gels consisting of a large-pored stacking gel and a resolving gel. This method, published by Laemmli (Laemmli, 1970), allows for increased resolution. The samples are first applied to the stacking gel, where the band is focused. After entering the resolving gel, proteins of different size are separated.

Samples were supplemented with loading buffer, incubated at 95 °C for approx. 5 min and loaded to the gel. The electrophoresis was performed at 30 mA until the front arrived at the lower part of the gel. If fluorescently labelled samples were analysed, the gel was photographed under UV illumination. In the case of non-fluorescent samples or subsequent to UV documentation the gel was stained in staining solution for at least 30 min, then incubated in destaining solution for 10 to 20 min and kept in water for further destaining. The gel was photographed for documentation.

2.6.3 Transformation of *E. coli* Cells

PCR samples or 50 to 100 ng plasmid DNA were added to $CaCl_2$ competent *E. coli* cells. The mixture was incubated on ice for 5 min. After 90 s at 42 °C, the cells were again incubated for 5 min on ice. Then 700 μl of LB medium were added and the suspension was incubated for 1 h at 37 °C. The cells were collected by centrifugation and then plated on antibiotic-containing LB plates to select for transformants.

2.6.4 Site-directed Mutagenesis

Mutants were constructed using the QuikChange site-directed mutagenesis method (Stratagene, La Jolla, USA). The plasmid containing the insert of interest is amplified in a PCR. In this reaction, primers are used that contain the desired mutation and anneal to opposite strands of the plasmid. The primers are incorporated into the amplicons generating mutated plasmids.

The reactions contained

1 ng/μl template DNA

0.2 μM forward primer

0.2 μM reverse primer

0.2 mM dATP, dCTP, dGTP, dTTP each

0.04 to 0.06 U/μl *Pfu* DNA polymerase

in 1x *Pfu* buffer.

PCR cycle:

1	5 min 95 °C	
2	30 s 95 °C	
3	60 s variable temperature, depending on the primer	
4	14.5 min 68 °C (12.5 min for YxiN404-479, 15 min for YxiN1-211 or YxiNC212-479)	

The series of steps 2 to 4 was repeated 24 times.

Mutagenesis of G303A was done with Phusion DNA polymerase, in reactions containing

1 ng/μl template DNA

0.5 μM forward primer

0.5 μM reverse primer

0.2 mM dATP, dCTP, dGTP, dTTP each

0.02 U/μl Phusion DNA polymerase

in 1x HF buffer + 3 % DMSO.

PCR cycle:

1	30 s 98 °C	
2	10 s 98 °C	
3	25 s 59 °C	
4	2.5 min 72 °C	The series of steps 2 to 4 was repeated 24 times.
5	5 min 72 °C	

Then 800 U/ml *Dpn*I endonuclease were added to the reaction mixture to digest the bacterial template DNA. After incubation at 37 °C for 3 h, the mixture was transformed into *E. coli* XL1-Blue (section 2.6.3) and transformants were selected on LB plates containing antibiotic.

In some cases insertions are created at the mutagenesis site in the PCR. To exclude the corresponding clones, a screening reaction was performed. In this screening PCR, a fragment of 200 to 300 bp containing the mutagenesis site is amplified from the bacterial DNA. The primers were chosen in such a way that the forward primer anneals about 100 bp upstream from the mutagenesis site, the reverse primer about 100 bp downstream. The length of the PCR product is then compared to a control.

The reactions contained:

bacterial cells from a colony

0.5 µM forward primer

0.5 µM reverse primer

0.5 mM dATP, dCTP, dGTP, dTTP each

5 U/l GoTaq DNA polymerase

in buffer (20 mM Tris/HCl pH 8.55, 16 mM $(NH_4)_2SO_4$, 0.01 % (v/v) Tween 20, 2 mM $MgCl_2$).

PCR cycle:

1   5 min 95 °C

2   45 s 95 °C

3   30 s 52 °C

4   30 s 72 °C   the series of steps 2 to 4 was repeated 24 times

The PCR product was then analysed on a 2 % agarose gel. Finally, overnight cultures of *E. coli* clones with insertion-free plasmids were grown. The plasmids were isolated (QIAprepSpin Miniprep Kit) according to the manual (but elution of the DNA with 20 % EB) and sequenced (Microsynth, Balgach, CH).

2.6.5 Absorption Measurements

To determine the concentration of DNA, RNA, proteins or labelled proteins in solution the absorption was measured at single wavelengths (BioPhotometer) or an absorption spectrum was taken (Ultrospec 2100 pro). Absorption values were corrected for buffer contributions and the concentration was calculated using the Lambert-Beer law (eq. 2.1).

$$A_x = c \cdot \varepsilon_x \cdot d$$

(eq. 2.1)

A_x – absorption at x nm, c – concentration, ε_x – extinction coefficient at x nm, d – path length, 1 cm

The following relations and extinction coefficients were used

for DNA 1 A_{260} = 50 ng/ml

153mer RNA ε_{260} = 1425000 $M^{-1}cm^{-1}$

32mer RNA ε_{260} = 211200 $M^{-1}cm^{-1}$

9mer RNA ε_{260} = 95200 $M^{-1}cm^{-1}$

YxiN, YxiN_GS and mutants ε_{280} = 23540 $M^{-1}cm^{-1}$ (Theissen, 2006)

YxiN1-368 ε_{280} = 20400 $M^{-1}cm^{-1}$ (Gasteiger *et al*., 2005)

YxiN404-479_405K ε_{280} = 2980 $M^{-1}cm^{-1}$ (Gasteiger *et al*., 2005)

YxiN1-211 and mutants ε_{280} = 11460 $M^{-1}cm^{-1}$ (Gasteiger *et al*., 2005)

YxiNC212-479 and mutants ε_{280} = 11920 $M^{-1}cm^{-1}$ (Gasteiger *et al*., 2005)

TMR ε_{550} = 95000 $M^{-1}cm^{-1}$ (ε_{280} = 18531 $M^{-1}cm^{-1}$ considered when calculating the protein concentration of the same sample, ε_{493} = 11259 $M^{-1}cm^{-1}$ considered when calculating the A488 concentration of the same sample)

A488 ε_{493} = 72000 $M^{-1}cm^{-1}$ (ε_{280} = 7922 $M^{-1}cm^{-1}$ considered when calculating the protein concentration of the same sample)

A546 ε_{554} = 93000 (ε_{280} = 10538 $M^{-1}cm^{-1}$ considered when calculating the protein concentration of the same sample, ε_{493} = 7025 $M^{-1}cm^{-1}$ considered when calculating the A488 concentration of the same sample)

A555 $\varepsilon_{556} = 158000$ ($\varepsilon_{280} = 11250$ M^{-1}cm^{-1} considered when calculating the protein concentration of the same sample, $\varepsilon_{493} = 27622$ M^{-1}cm^{-1} considered when calculating the A488 concentration of the same sample)

2.7 Cloning of Full-length YxiN Constructs in pTWIN1

Ahead of this project the *Sap*I restriction site in the *yxiN* gene had been removed by site-directed mutagenesis without changing the amino acid sequence. The corresponding constructs of *yxiN* or *yxiN*_C61A_C267A_A115C_D262C in pETM30 were amplified in a PCR and then ligated into the multiple cloning site of pTWIN1 creating a fusion to the intein sequence in the vector. In the resulting fusion protein the intein will be located at the C-terminal region.

At first, pETM30_*yxiN* or pETM30_*yxiN*_C61A_C267A_A115C_D262C was linearized by incubation with 80 U/ml *Sap*I in NEBuffer 4 for 2,5 h at 37 °C, providing the template for the PCR. The forward and the reverse primer for the PCR contained the sequence for a *Nde*I and *Sap*I restriction site respectively. The reverse primer also contained the sequence for a linker (Gly$_3$-Ser-Gly$_3$-Ser-Gly$_2$-Thr) downstream to the protein C-terminus.

The reactions contained

1 ng/μl template

0.2 μM forward primer

0.2 μM reverse primer

0.4 mM dATP, dCTP, dGTP, dTTP each

2 mM MgSO$_4$

0.04 to 0.06 U/μl *Pfu* DNA polymerase

in 1x *Pfu* buffer.

PCR cycle:

1   2 min 95 °C

2   30 s 95 °C

3   30 s 55 °C

4   3 min 72 °C   the series of steps 2 to 4 was repeated 29 times

5   5 min 72 °C

The PCR product was purified (Promega PCR Clean-Up System) and incubated with 80 U/ml *Sap*I and 800 U/ml *Nde*I in NEBuffer 4 for 3 h at 37 °C followed by heat inactivation of the enzymes for 10 min at 65 °C. 80 ng pTWIN1 were treated analogously and then de-phosphorylated for 30 min at 37 °C using 0.03 U/μl SAP in 1x SAP buffer. Insert and vector were purified (QIAquick Nucleotide Removal Kit, 5 volumes PN buffer, elution with EB) and mixed in a ligation reaction with 20 U/μl T4 ligase in ligase buffer (16 °C, 16 h). The reaction mixture was transformed into *E. coli* XL1-Blue (section 2.6.3) and transformants were selected on LB plates containing ampicillin. Transformants with the empty vector only could be excluded in a PCR screen (section 2.6.4, forward primer annealed within the insert, reverse primer annealed within the *mxe* gene). Finally, an overnight culture of a positive clone was grown, the plasmid was isolated (QIAprepSpin Miniprep Kit, elution with 20 % EB) and the sequence was verified (Microsynth, Balgach, CH).

2.8 *In-vitro* Transcription of the 153mer RNA Substrate

IVT-buffer	80 mM Tris/HCl, pH 7.5(25 °C); 22 mM MgCl2; 1 mM spermidine; 5 mM DTT; 0.12 mg/ml BSA
TBE buffer	89 mM Tris, 89 mM boric acid, 10.7 mM EDTA
8 M urea 10 % PA gel	8 M urea; 89 mM Tris, 89 mM boric acid, 10.7 mM EDTA; 10 % (v/v) acrylamide; 0.267 % (v/v) *N,N'*-methylenebisacrylamid; polymerized by addition of 0.045 % (w/v) APS and 0.113 % (v/v) TEMED; prerun for some minutes before use

A 153mer RNA fragment from *B. subtilis* 23 S rRNA (nucleotides 2483-2635) was prepared by run-off *in-vitro* transcription as described (Karow *et al*., 2007). The RNA was transcribed from a DNA template that had been amplified via PCR and purified by ethanol precipitation before.

The reaction mixture contained 90 ng/ml DNA template, 20 mM ATP, CTP, GTP and UTP each, 16 U/ml Thermostable Inorganic Pyrophosphatase and T7 RNA Polymerase in IVT-buffer and was incubated at 37 °C for 4 h. The reaction was stopped by adding 1/10 volume 0.5 M EDTA (pH 8), 1/10 volume 3 M sodium acetate and 3 volumes ethanol. The nucleic acids were precipitated at -20 °C for 20 min and centrifuged (F-45-18-11, 13000 rpm, 4 °C, 30 min). The pellet was washed with 70 % ethanol and dissolved in water and 2x Fermentas RNA loading dye. After incubation at 65 °C for 5 min the sample was loaded on an 8 M urea 10 % PA gel and ran for 1.5 h, at 25 W in TBE buffer. The bands were visualised under UV-illumination ($\lambda = 254$ nm) and the product-containing band was excised from the gel and kept in TBE buffer. The RNA was eluted in TBE at 4 °C, 250 V, 100 mA for 4 h and precipitated by addition of 1/10 volume 3 M sodium acetate, 2.75 volumes ethanol and over-night incubation at -20 °C. The precipitate was collected by centrifugation (F-45-18-11, 13000 rpm, 4 °C, 1 h), the pellet washed with 70 % ethanol and let dry for some hours. The RNA was dissolved in 20 mM HEPES/NaOH pH 7.2. The concentration was determined via absorption spectroscopy.

2.9 Protein Production and Purification

2.9.1 Protein Production

YxiN constructs in pTWIN1 and the *yxiN*404-479_405K gene in pETM30 were expressed in LB medium containing chloramphenicol and ampicillin (pTWIN1) or chloramphenicol and kanamycin (pETM30). First, an overnight culture was inoculated with a single *E. coli* clone and grown at 37 °C. Then medium was inoculated with 1/100 volume of overnight culture and kept at 37 °C until reaching an OD_{600} of 0.5. Then, protein production was induced by addition of 0.15 mM IPTG and the culture was kept at 30 °C for 4 h. The cells were collected by centrifugation (10 min, 4 °C, 5000 rpm, rotor SLA3000) and resuspended in the buffers mentioned in the sections describing the purification. Cells harbouring pTWIN1 were again centrifuged and the pellet was resuspended in fresh buffer. This washing step ensures maximal stability of the intein fusions that are able to undergo pH-dependent self-cleavage.

Resuspended cells were stored at - 70 °C.

All other *YxiN* constructs in pETM30 were expressed in auto induction medium (Studier, 2005) containing kanamycin and chloramphenicol.

First, an LB overnight culture was inoculated with a single *E. coli* clone and grown at 37 °C. The auto induction medium was then inoculated with 1/100 volume of overnight culture, incubated for approx. 3 h at 37 °C and then approx. 24 h at 30 °C. The cells were collected by centrifugation (10 min, 4 °C, 5000 rpm, rotor SLA3000) and

stored at -70 °C or resupended in approx. 40 ml buffer (50 mM Tris/HCl pH 8.5(4 °C), 0.5 M NaCl) and stored at -70 °C.

2.9.2 Purification of YxiN1-368, YxiN and Related Mutants

The proteins carrying a glutathione *S*-transferase and a hexahistidine tag were produced as described in section 2.9.1. The cells were disrupted by the Microfluidizer and the protein was purified by glutathione-affinity chromatography and cleaved from the tags. Then immobilized metal affinity chromatography, ammonium sulphate precipitation and SEC followed. Experimental details are presented in Theissen (2006) and Karow *et al.* (2007). After shock freezing the proteins were stored in SEC-buffer at -80 °C.

In the case of mutants prone to aggregation the standard protocol was modified as follows. Before the ammonium sulphate precipitation, the protein samples of YxiN_C61A_C267A_A115C_S229C_S182A_T184A and YxiN_C61A_C267A_A115C_S229C_Q29A were concentrated to about half of their volume.

2.9.3 Purification of YxiN404-479_405K

Buffer A   50 mM Tris/HCl, pH 8.5(4 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol
Buffer B   50 mM Tris/HCl, pH 8.5(4 °C); 1 M NaCl; 2 mM 2-Mercaptoethanol
Buffer C   50 mM Tris/HCl, pH 8.5(4 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol; 20 mM glutathione
Buffer D   50 mM Tris/HCl, pH 8.5(4 °C); 500 mM NaCl

The protein was produced as described in section 2.9.1. The cells were resuspended in buffer A containing half a protease inhibitor tablet and disrupted in the Microfluidizer. The suspension was clarified by centrifugation (SS34, 13000 rpm, 4 °C, 30 min) and loaded onto a glutathione sepharose column (4 °C, 0.3 ml/min). The column was washed with buffer B (1 ml/min) and subsequently the protein was eluted by flushing with buffer C. TEV protease was added and the eluate was dialysed (MWCO 3500) against 2x 1 l buffer A at room temperature for 1 day. The sample was supplemented with 20 mM imidazole and loaded onto a Ni^{2+} chelating sepharose column (4 °C, 1 ml/min). The flow-through was collected and concentrated in a Vivaspin MWCO 2000. The sample was centrifuged (F-45-18-11, 13000 rpm, 4 °C, 2 min) and purified on a Superdex 75 SEC column (Buffer D, 1 ml/min, 4 °C). Fractions containing the protein in high purity were pooled and concentrated (Vivaspin MWCO 2000). The concentration was determined via absorption spectroscopy and after shock freezing the protein was stored at -80 °C.

2.9.4 Purification of YxiN_GS and YxiN_C61A_C267A_A115C_D262C_GS

Buffer A   20 mM Tris/HCl, pH 7(25 °C); 500 mM NaCl; 1 mM EDTA
Buffer B   20 mM Tris/HCl, pH 8(4 °C); 500 mM NaCl; 1 mM EDTA; 0.1 M MESNa
Buffer C   50 mM Tris/HCl, pH 7(25 °C); 500 mM NaCl

The proteins were produced as intein fusions from pTWIN1 as described in section 2.9.1. The cells were resuspended in buffer A containing half a protease inhibitor tablet and disrupted in the Microfluidizer. The suspension was clarified by centrifugation (SS34, 13000 rpm, 4 °C, 30 min) and applied to chitin resin equilibrated with buffer A at 4 °C. About 5 ml chitin resin was used to purify the protein from a 0.5 l culture.

The column was washed with 15 column volumes of buffer A before the cleavage was initiated by flushing with 1.5 column volumes buffer B. After overnight incubation the protein, now harbouring a C-terminal thioester, was eluted using buffer B. Fractions with high amount of cleavage product were pooled and concentrated. The buffer was exchanged to buffer C using a NAP-10 SEC column followed by further concentrating the protein.

The concentration was determined via absorption spectroscopy and after shock freezing the protein was stored at -80 °C.

2.9.5 Purification of YxiN1-211 and YxiNC212-479 and Related Mutants

The proteins, fused to an intein and a chitin-binding domain, were produced as described in section 2.9.1. The cells were disrupted by the Microfluidizer and the protein was purified via chitin affinity chromatography followed by cleavage from the intein as described (Karow *et al.*, 2007).

The protein concentration was determined via absorption spectroscopy and after shock freezing the protein was stored at -80 °C.

A modified purification protocol was applied for YxiN1-211_C61A_S108C. The bacterial cells were resuspended in 20 mM Tris/HCl, pH 7(25 °C); 500 mM NaCl; 1 mM EDTA; 2 M urea, disrupted and the crude extract was applied to a chitin column equilibrated with the same buffer. During the washing procedure the buffer was exchanged to 20 mM Tris/HCl, pH 7(25 °C); 500 mM NaCl; 1 mM EDTA and the following steps equalled the standard protocol.

2.10 Labelling of YxiN Constructs with Fluorophores

2.10.1 Statistic Labelling and Removal of Free Fluorophores

Buffer A   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 0.5 mM TCEP
Buffer B   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol
Fluorophore stock solution 20 mM in DMSO

Before the labelling reaction, the buffer of the protein solution was exchanged to buffer A via Micro Bio-Spin 30 SEC columns and a typically 30 to 50 μM protein solution was prepared. For donor-acceptor labelling maleimide-functionalized donor and acceptor dye were added simultaneously. The relative concentration of the dyes was optimized depending on the specific mutant. Typically a 3-fold excess of donor and a 4-fold excess of acceptor dye or, in case of labelling with one dye, a 5-fold excess relative to the protein concentration was used. The mixture was incubated in low-retention tubes at 25 °C, for 20 to 50 min in the dark. In the case of mutants prone to aggregation the temperature was decreased to 20 °C and the incubation time shortened to 15 min. After incubation, unreacted dye was removed by passing the solution at least two times over a Micro Bio-Spin 30 SEC column equilibrated with buffer B. In case of single-labelled proteins used for the determination of FRET correction parameters or the Förster distance, unreacted dye was always removed in three runs.

A sample of the labelled protein was supplied to SDS-PAGE for the amount of unreacted dye. The concentration and labelling degrees were determined via absorption spectroscopy correcting A_{280} for dye contributions and A_{493} for acceptor contribution (section 2.6.5).

2.10.2 Fluorophore Labelling of YxiN1-211_C61A_S108C and YxiNC212-479_C267A_S229C and Subsequent EPL

Buffer A   50 mM Tris/HCl, pH 7.4(25 °C); 500 mM NaCl; 1 mM TCEP
Buffer B   50 mM Tris/HCl, pH 7.4(25 °C); 500 mM NaCl; 0.4 mM TCEP; 3.64 M urea
Fluorophore stock solution 10 or 20 mM in DMSO
All reactions were incubated in low-retention tubes.

YxiNC212-479_C267A_ S229C was dialysed against buffer A, the protein concentration was determined and a variable amount of maleimide-functionalized A488 was added. The protein-A488 ratio was optimized with

respect to high yield in the subsequent EPL reaction at a sufficient labelling degree. The reaction contained about 200 μM protein and was incubated at 25 °C for 20 min in the dark. Then, the labelling reaction was stopped by the addition of 3.75 % (v/v) thiophenol.

To 92 μM YxiN1-211_C61A_S108C (harbouring a C-terminal thioester) in buffer B a 4-fold excess of maleimide-functionalized A546 was added and incubated at 25 °C for 5 min. The labelling reaction was stopped by the addition of 2.6 % (v/v) thiophenol.

For the EPL reaction the labelled proteins were mixed, the pH was adjusted by addition of 25 mM Tris/HCl, pH 8(25 °C) and the mixture incubated for 16 h at 15 °C in the dark. In total, the EPL samples contained 50 μM YxiN1-211_C61A_S108C, about 70 μM YxiNC212-479_C267A_ S229C in buffer comprising 2 M urea and 3 % (v/v) thiophenol. After centrifugation at 13000 rpm (rotor 33258), 25 °C the pellet and the supernatant were supplemented with 4x SDS-PAGE loading buffer (section 2.6.2) and analysed via SDS-PAGE.

2.10.3 Donor Fluorophore Labelling of YxiNC212-479_C267A_D429C or YxiNC212-479_C267A_N444C, Subsequent EPL to YxiN1-211_C61A and Purification of the Full-length Protein

Buffer A   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 0.5 or 1 mM TCEP
Buffer B   50 mM Tris/HCl, pH 8(25 °C); 500 mM NaCl
Buffer C   50 mM Tris/HCl, pH 7.5(4 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol
Fluorophore stock solution 20 mM in DMSO
All reactions were incubated in low-retention tubes.

YxiNC212-479_C267A_D429C (or ~N444C) was dialysed against buffer A or the storage buffer was exchanged to buffer A using a Micro Bio-Spin 6 SEC column. After determination of the protein concentration, a variable amount of maleimide-functionalized A488 was added. During the optimization process, the protein-A488 ratio was varied to achieve a high yield in the subsequent EPL reaction at a sufficient labelling degree. Based on the result of the latter, a protein-A488 ratio of 1 was chosen in preparative reactions. The reactions contained about 400 to 600 μM protein, were incubated at 25 °C for 30 min in the dark and stopped by the addition of approx. 6.5 % (v/v) thiophenol.

For the EPL reaction, YxiN1-211_C61A (harbouring a C-terminal thioester) and buffer B were added to the labelled protein. The reaction contained 100 to 200 μM YxiN1-211_C61A, 100 to 200 μM YxiNC212-479_C267A_D429C (or ~N444C), 2.5 % (v/v) thiophenol and the volume was adjusted by addition of buffer B. The mixture was incubated for 16 h in the dark, at 15 °C or 25 °C (as indicated). After centrifugation at 13000 rpm (rotor 33258), 25 °C thiophenol and unreacted fluorophore were removed by SEC (NAP-5 column or Micro Bio-Spin 6 column). The EPL product was purified on a Superdex 75 SEC column (1 ml/min, 4 °C, buffer C). Fractions containing a high amount of full-length protein were pooled and concentrated (Vivaspin 500, 4 °C). The protein concentration and the labelling degree were determined via absorption spectroscopy correcting A_{280} for A_{488} contribution (section 2.6.5).

2.10.4 Acceptor Fluorophore Labelling of the EPL Product and Removal of Free Fluorophore

Buffer A   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 0.5 mM TCEP
Buffer B   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol
Fluorophore stock solution 20 mM in DMSO

The buffer of the A488-labelled EPL product YxiN_C61A_T212C_C267A_N444C was exchanged to buffer A via a Micro Bio-Spin 30 SEC column. A 4-fold excess of A546 was added and the mixture was incubated at 25 °C for 1 h in low-retention tubes, in the dark. After incubation, unreacted A546 was removed by passing the solution two times over a Micro Bio-Spin 30 SEC column equilibrated with buffer B. A sample of the labelled protein was checked by SDS-PAGE for the amount of unreacted dye. The concentration and labelling degrees were determined via absorption spectroscopy correcting A_{280} for dye contributions and A_{493} for A546 contribution (section 2.6.5).

2.11 C-terminal Biotinylation of YxiN

2.11.1 Synthesis of Cys-Biotin

The protocol is a modified version of the synthesis by Lesaicherre and colleagues (Lesaicherre *et al.,* 2002).
0.15 mmol *N*-α-*t*-Boc-*S*-trityl-L-cysteine, 0.18 mmol TBTU and 0.225 mmol HOBt were dissolved in 1 ml DMF (dried over molecular sieves). The mixture was stirred under argon for 30 min at room temperature. 0.45 mmol 4-methylmorpholine and 0.15 mmol 5-(biotinamido)pentylamine were added and the mixture was stirred under argon for 4.5 h at room temperature. After evaporation of the solvent, the product was dissolved in 11 ml DCM and extracted 3 times with 11 ml H_2O. The organic phase was dried with $MgSO_4$ and evaporated. The crude product was further purified by chromatography (Silica Gel 60, 8 % methanol in DCM), followed by evaporation of the solvent. Product formation and purification were followed by TLC (10 % methanol in DCM) and LC-ESI-MS (BioSuite C18 PA-B column).

The compound was deprotected by addition of 3 ml TFA, 94 ml H_2O and 0.07 g triisopropylsilane and stirring for 30 min at room temperature. Then the crude product was precipitated in 40 ml cold diethyl ether, lyophilised and dissolved in 5 ml H_2O. The aqueous phase was extracted 3 times with DCM and contained the deprotected compound as verified by LC-ESI-MS. After lyophilisation the product was further purified via HPLC with a C18 RP column (ACN/H2O gradient + 0.1 % TFA), followed by MALDI-MS analysis. TFA was removed in a second RP-HPLC run using 0.1 % formic acid as ion pairing agent. After neutralization by the addition of NH_3 the final product was lyophilised yielding 9 mg (0.021 mmol) of white powder (MALDI-MS: 432.3 = $[M+H]^+$). Quantification via the Biotin Quantitation Kit confirmed a yield of 0.0268 mmol.

2.11.2 Western or Dot Blotting and Detection of Biotinylated Protein

Transfer buffer   25 mM Tris base; 150 mM glycin; 20 % EtOH
TBS   20 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl
TBS-T   20 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 0.2 % (v/v) Tween-20

a) For Western Blotting, the samples from biotinylation reactions were subjected to SDS-PAGE. After the run the gel was covered with a nitrocellulose membrane and sandwiched between filter paper and sponges, all soaked with transfer buffer. Then, the proteins were transferred in a wet blot cell (150 mA, 1 h, 4 °C).

b) For Dot Blotting, a nitrocellulose membrane was wet in TBS and dried for 5 min. The protein samples were spotted onto the membrane and let dry completely.

After Western Transfer or Dot Blotting the membrane was washed 2 times in TBS-T for 5 min. The membrane was incubated in TBS-T containing 0.1 % (v/v) Avidin-HRP solution for 1 h. Then the membrane was washed 2 times in TBS-T for 5 min and 2 times in TBS for 5 min.

A solution of 4-chloro-1-naphthol in cold MeOH was prepared and added to 5 volumes 0.018 % (v/v) H_2O_2 in TBS (HRP color development solution). The membrane was incubated in the HRP color development solution until stained dots became visible (typically 5 to 15 min). The color development was stopped by washing the membrane in water. The membrane was photographed for documentation.

2.11.3 Biotinylation of YxiN_GS

Buffer A   50 mM Tris/HCl, pH 7(25 °C); 500 mM NaCl; 0.32 mM TCEP
Buffer B   50 mM Tris/HCl, pH 8(25 °C); 500 mM NaCl

To optimize the C-terminal biotinylation of YxiN_GS, the reaction was done in analytical scale (10 µl) at varying conditions. Each reaction contained 70 µM YxiN_GS (harbouring a C-terminal thioester) in buffer A, a variable concentration of MESNa solution (pH 8) and a variable concentration of Cys-Biotin solution (in buffer B). Variable volumes of MESNa and Cys-Biotin solutions in one series were adjusted by addition of buffer B and the reaction was incubated at 25 °C. At defined time points a 0.714 µl sample was added to 25 µl 4-fold SDS-PAGE loading buffer (section 2.6.2) and shock frozen in liquid nitrogen. For analysis the samples were thawed, subjected to SDS-PAGE (application of 5 µl sample corresponding to 10 pmol protein) followed by western blotting and the biotinylated protein was detected as described in section 2.11.2.

2.11.4 Labelling of YxiN_C61A_C267A_A115C_D262C_GS with Fluorophores and Subsequent Biotinylation

Buffer A   50 mM Tris/HCl, pH 7 (25 °C); 500 mM NaCl; 0.4 mM TCEP
Buffer B   50 mM Tris/HCl, pH 8(25 °C); 500 mM NaCl
Buffer C   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol
Fluorophore stock solution 20 mM in DMSO

About 60 µM YxiN_C61A_C267A_A115C_D262C_GS (harbouring a C-terminal thioester) in buffer A were reacted with a mixture of about 0.18 mM donor and 0.24 mM acceptor dye in low-retention tubes. After a 10 min incubation at 25 °C in the dark, the reaction was stopped by adding 5.7 mM MESNa, pH 8. After 5 min at 25 °C, 3 mM Cys-Biotin solution (in buffer B) were added and the biotinylation at a final MESNa concentration of 5 mM proceeded for 3 to 4 hours at 25 °C in the dark. Unreacted dye and free Cys-Biotin were removed by passing the solution at three times over a Micro Bio-Spin 30 SEC column equilibrated with buffer C.

2.12 Steady-state ATPase Activity Assay

Assay buffer
50 mM HEPES/KOH, pH 7,5; 175 mM KCl; 10 mM $MgCl_2$; 100 µM DTT; 200 µM NADH; 1 mM PEP; 13 µg/ml lactate dehydrogenase; 23 µg/ml pyruvate kinase

The steady-state ATPase activity of YxiN was assayed in a coupled spectroscopic test (Adam, 1962), in which the formed ADP is regenerated to ATP by pyruvate kinase consuming PEP. The reaction product pyruvate is subsequently reduced to lactate by lactate dehydrogenase consuming NADH. The resulting decrease in NADH concentration can be followed via a decreasing A_{340} over time.

The reactions were performed as described (Kossen & Uhlenbeck, 1999; Karow *et al.*, 2007) in assay buffer, at 37 °C, with varying concentration of 153mer RNA (in presence of 5 mM ATP) or varying concentration of ATP (in presence of 153mer as indicated). The reaction was initiated by addition of YxiN at mutant specific concentration (10 nM to up to 1 µM) as indicated. For testing the ATPase activity of the helicase domain

YxiN1-368 and RBD YxiN404-479_405K *in trans* a mixture of both proteins (10 nM to up to 1 μM each) was added to start the reaction.

Data evaluation

The decrease in NADH absorption over time can be converted into the decrease in NADH concentration using the Lambert-Beer law and the extinction coefficient of $\varepsilon_{340} = 6300$ $M^{-1}cm^{-1}$ (Bergmeyer, 1975). Since the consumption of one NADH molecule is directly linked to the consumption of one ATP molecule, the time derivative of c_{NADH} gives the velocity of ATP hydrolysis.

When assaying the RNA or ATP dependent activity of YxiN, the basal rate in the absence of RNA or ATP respectively was subtracted. ATP dependencies were fitted using the Michaelis-Menten formalism (eq. 2.2) and calculating the maximal velocity of the reaction v_{max} and an apparent K_M for ATP. The turnover number k_{cat} was calculated using equation 2.3. As the RNA substrate concentration is in the range of the YxiN concentration, the data from RNA dependent measurements were analyzed using the explicit binding equation (eq. 2.4, Segel, 1993, Kossen & Uhlenbeck, 1999).

$$v = \frac{v_{max}[S]_0}{K_M + [S]_0}$$
(eq. 2.2)

$$k_{cat} = \frac{v_{max}}{[YxiN]}$$
(eq. 2.3)

$$k = \frac{k_{cat}}{[YxiN]}\left[\frac{[S]_0 + [YxiN] + K_{app,RNA}}{2} - \sqrt{\left(\frac{[S]_0 + [YxiN] + K_{app,RNA}}{2}\right)^2 - [S]_0[YxiN]}\right]$$
(eq. 2.4)

v – velocity of the reaction, v_{max} – maximal velocity, $[S]_0$ – total substrate concentration,

K_M – Michaelis-Menten constant, k_{cat} – turnover number, $K_{app,RNA}$ – apparent dissociation constant for RNA

2.13 RNA Unwinding Activity Assay

5-fold assay buffer	250 mM HEPES/KOH, pH 7.2; 750 mM KCl; 25 % (v/v) glycerol; 0.5 mM DTT; 0.5 mg/ml BSA; 25 mM $MgCl_2$
Storage buffer	50 mM Tris/HCl, pH 8.5(4 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol
2-fold quenching buffer	2 % (w/v) SDS; 100 mM DTT; 10 % (v/v) glycerol; 0.1 % (w/v) bromophenol blue
Native PA-gel	15 % (v/v) acrylamide; 0.4 % (v/v) *N,N'*-methylenebisacrylamid; 100 mM Tris, 100 mM boric acid, 1 mM EDTA; 10 mM $MgCl_2$; polymerized by addition of 0.1 % (w/v) APS and 0.067 % (v/v) TEMED
GE running buffer	100 mM Tris, 100 mM boric acid, 1 mM EDTA

The RNA double strand unwinding activity of YxiN and related mutants was assayed as described (Karow *et al.*, 2007). In detail, a minimal RNA substrate for YxiN was generated by incubating 25 μM 9mer-F and 50 μM 32mer RNA at 25 °C for 30 min. The annealed 9/32mer was kept on ice until added to the reaction mixture. The 10 μl reactions contained a variable volume of H_2O, 2 μl 5-fold assay buffer, 2 μl of the annealed 9/32mer, 50 μM 9mer (unlabelled), 2 μl of 50 μM protein in storage buffer (unless indicated differently) and 10 μM ATP. Control reactions without protein or nucleotide were instead supplemented with the same volume of storage buffer or H_2O respectively.

The unwinding reaction was initiated by addition of the nucleotide and incubated at 25 °C. After incubation for the indicated time, it was stopped by addition of 10 μl quenching buffer and the mixture was applied to a native

15 % PA-gel. Duplex and single-stranded RNA were separated at 10 V/cm, at 4 °C. Fluorescent RNA bands were visualised under UV light and the intensities were quantified using the KODAK 1D Image Analysis Software.

2.14 Mant-nucleotide Titration

Buffer A 50 mM Tris/HCl, pH 7.5(25 °C); 150 mM NaCl; 5 mM MgCl₂; 1 mM 2-Mercaptoethanol

Nucleotide binding to YxiN mutants was assayed in mant-ADP or mant-ADPNP titrations (Karow *et al.*, 2007). These nucleotide analogues display a fluorescence signal sensitive to their environment (Hiratsuka, 1983). Therefore, the signal can serve as a probe for binding to a protein i.e. the transfer from aqueous solution to a less hydrophilic environment.

Assay conditions were as described (Karow *et al.*, 2007): 1 µM mant-ADP (or mant-ADPNP) in buffer A was titrated with YxiN at 20 °C. After 2 min of equilibration mant fluorescence was excited at 365 nm (excitation slit 2 nm) and collected for 20 s at 430 nm (emission slit 3 nm).

2.15 SmFRET Techniques

2.15.1 Fluorescence Resonance Energy Transfer (FRET)

Fluorescence Resonance Energy Transfer (FRET) is a spectroscopic technique that has found wide use in biochemistry and biophysics. FRET has already been discussed by Förster in 1946 and was applied as "A Spectroscopic Ruler" in 1967 (Stryer & Haugland, 1967). It is based on non-radiative energy transfer from a donor fluorophore to an acceptor fluorophore via dipole-dipole interactions. The energy transfer is highly distance dependent therefore the distance (r) between the fluorophores can be calculated from the transfer efficiency (E) (eq. 2.5).

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

(eq. 2.5)

The so called Förster distance, R_0 depends on the spectral properties of the fluorophores, the relative orientation of donor and acceptor transition dipoles and the refractive index of the solvent. The Förster distance has to be determined for a specific dye pair under the conditions chosen for the FRET experiment (i.e. in the bioconjugated form and in assay buffer). The procedure is described in section 2.15.5. Values for R_0 are typically in the low nanometer range making FRET well suited to study inter- and intramolecular phenomena of biomacromolecules. As the FRET efficiency scales with r^{-6} (eq. 2.5) the measurements will be most sensitive to distances around R_0. It is generally not recommended to measure distances outside the range of $0.5R_0$ to $2R_0$ (Lakowicz, 2006).

2.15.2 Measurements at the Single Molecule Level

Enzyme kinetics, biochemical binding events and spectroscopic properties are widely determined from ensembles of molecules. Here, the probed signal always represents the average signal from all molecules. Usually different states cannot be distinguished and the need for synchronization techniques arises. As a matter of principle the detection of single molecules does not have these disadvantages. As each single molecule is

studied separately, a distribution of states, including rare and transient ones, can be uncovered. Changes over time can be probed without synchronisation of an ensemble.

In the last few years FRET at the single molecule level has proven to be a powerful technique. Firstly, fluorescence itself is well suited for single molecule detection (SMD) since the signal to background ratio is rather large. Secondly, as FRET is a ratiometric technique, part of the high noise one is faced with in SMD is eliminated (Dahan *et al.*, 1999).

2.15.3 Confocal and Total Internal Reflection Microscopy as Optical Techniques for Single Molecule FRET (smFRET)

The detection of single molecules (not only for FRET experiments) can be achieved by minimising the observation volume using confocal or Total Internal Reflection (TIR) microscopy (fig. 2.1). In a confocal microscope the donor fluorophores are excited with light from a parallel laser beam. The beam is focused in the sample via an objective with high numerical aperture. The emitted fluorescence is collected and directed to a pinhole at the detection path. The pinhole will block out-of-focus light. Thus, only the fluorescence from the focal plane passes and is finally detected. The volume from which photons are collected is typically in the order of 10^{-15} l. By choosing nanomolar concentrations of the sample there is virtually never more than a single molecule in the confocal volume. The molecules are freely diffusing in solution so each molecule can only be observed while crossing the confocal volume. The dwell time for proteins is typically several ms and processes on larger time-scales cannot be studied. TIR microscopy of immobilized molecules can serve as a complementary technique in a seconds time-scale. Here, the molecules need to be fixed on a surface and thereby will not leave the observation volume. In TIR microscopy the excitation beam is directed in such a way that it is totally reflected at the interface of a glass substrate and buffer solution. Under these conditions an evanescent field is created in the solution and is able to excite fluorophores close to the interface (typically in the range of 10^{-7} m). The fluorescence emitted by the immobilized molecules within the illuminated area is usually detected by CCD cameras.
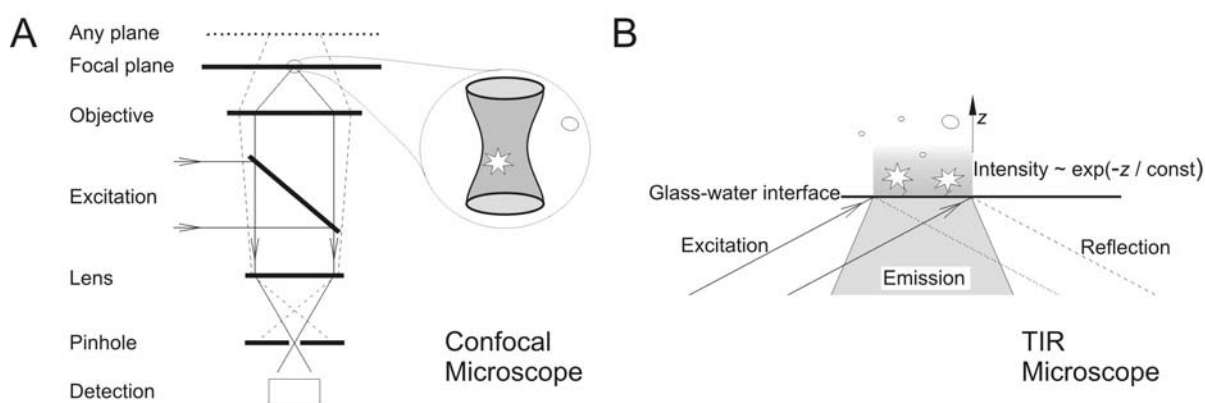
Figure 2.1
Optical techniques for smFRET experiments. A) In a confocal microscope the excitation light is focused via a high aperture objective to the sample. Emission light is collected by the same objective and directed to the detection path where out-of-focus light is blocked by a pinhole. Only photons from an observation volume of femtoliter size are detected. B) In a TIR-microscope the excitation light is totally reflected by the sample-glass interface. An evanescent wave is created in the sample that can only excite molecules close to the surface. The emission light is collected from the illuminated area and detected spatially resolved with respect to the x- and y-direction.

2.15.4 SmFRET Experiments on a Confocal Microscope

Buffer A   50 mM Tris/HCl, pH 7.5(25 °C); 150 mM NaCl; 5 mM MgCl$_2$
                (active charcoal treated to reduce the amount of fluorescent impurities)
Buffer B   50 mM Tris/HCl, pH 7.5(25 °C); 500 mM NaCl; 2 mM 2-Mercaptoethanol

SmFRET was measured using a home-built confocal microscope (instrumentation section 2.4) in cuvettes (Lab-Tek chamber slides) that had been saturated with 10 μM YxiN1-368 in buffer A for approx. 1 h. In the cuvette, the fluorescently labelled sample was diluted to 40 pM (concentration of the donor fluorophore, in buffer A) and substrates as indicated were added. The donor was excited at 475 nm, 75 μW and the sm fluorescence was recorded for 20 to 30 min.

Data evaluation

Whenever a molecule crosses the confocal volume, it causes a burst of photons. Not only the fluorophore labelled proteins but also small fluorescent impurities provoke photon counts that need to be distinguished. Therefore an algorithm calculating the noise and detecting bursts (Spreitler, 2006) was applied to the raw data. Then, bursts of more than 50 (100 or 200 as indicated) photons were selected.

The FRET efficiency for each burst can be calculated using eq. 2.6. Here one assumes equal quantum efficiencies and detection efficiencies for donor and acceptor as well as a perfect splitting of donor and acceptor photons to donor and acceptor APD. These conditions are not fulfilled in real life. To account for crosstalk i.e. donor photons being detected by the acceptor APD and acceptor photons being detected by the donor APD the correction parameters α and β respectively are implemented. In addition, the parameters γ, accounting for the different quantum yields and detection efficiencies of donor and acceptor, and δ, considering the direct excitation of the acceptor at 475 nm are needed. Therefore, an expanded version of eq. 2.6 was used to calculate the FRET efficiency for each burst (eq. 2.7). The values for all events in one experiment were then plotted in a histogram.

$$E = \frac{I_A}{I_D + I_A} \tag{eq. 2.6}$$

$$E = \frac{(1 + \beta\gamma\delta) \cdot (I_A - \frac{\alpha + \gamma\delta}{1 + \beta\gamma\delta} I_D)}{(1 + \beta\gamma\delta) \cdot (I_A - \frac{\alpha + \gamma\delta}{1 + \beta\gamma\delta} I_D) + (\gamma + \gamma\delta)(I_D - \beta I_A)} \tag{eq. 2.7}$$

E – FRET efficiency, I_A – counting rate in acceptor channel,

I_D – counting rate in donor channel, $\alpha, \beta, \gamma, \delta$ – correction parameters

Using the labelling procedure described in section 2.10.1, a statistic labelling of the two cysteines in a protein with donor and acceptor dye is performed. The set of correction parameters for a double-cysteine mutant is determined assuming donor on position X and acceptor on Y as well as vice versa. When a FRET histogram is calculated, the mean value of the parameters from both situations is taken or the histograms are calculated for each situation.

Likewise, to compute distance histograms the mutant specific Förster distances and correction parameters were needed. For distance calculations of a statistically labelled mutant XY the distances assuming donor on position X and acceptor on Y as well as *vice versa* were calculated or the average α, β, γ and R_0 were used.

2.15.5 Determination of the Correction Parameters for the SmFRET Histogram Calculations

Single-cysteine mutants were labelled with donor or acceptor. After removal of free dye as described in section 2.10.1 the absorbance of the protein coupled dyes at 475 nm was determined (25 °C, buffer B). Then the protein was digested by incubation with 0.2 mg/ml proteinase K for 30 min at 37 °C. An absorption spectrum of the solution was taken, corrected for buffer and proteinase K contribution and used to calculate the dye concentration. Knowing the dye concentration and the A_{475}, the ε_{475} of donor or acceptor dye coupled to the specific mutant can be derived.

The parameter δ for a specific double-cysteine mutant XY is defined by the ratio of ε_{475}acceptor(X) to ε_{475}donor(Y).

Furthermore the fluorescence of the labelled single mutants (50 to 100 nM solution in buffer A with respect to the dye) and the corresponding buffer in saturated cuvettes, at 75 µW, 25 °C is measured by the confocal smFRET setup (neutral density filter between objective and APDs added). The counting rates are normalized to the fluorophore concentration and α, β and γ are calculated using equations 2.8 to 2.10.

$$\alpha = \frac{I_A^{Donor} - I_A^{Buffer}}{I_D^{Donor} - I_D^{Buffer}} \qquad \text{(eq. 2.8)}$$

$$\beta = \frac{I_D^{Acceptor} - I_D^{Buffer}}{I_A^{Acceptor} - I_A^{Buffer}} \qquad \text{(eq. 2.9)}$$

$$\gamma = \frac{I_A^{Acceptor} - I_A^{Buffer}}{I_D^{Donor} - I_D^{Buffer}} \delta^{-1} \qquad \text{(eq. 2.10)}$$

I_D – counting rate in the donor channel, I_A – counting rate in the acceptor channel

2.15.6 Determination of the Förster Distances of Protein Coupled Donor-acceptor Fluorophores

To convert a FRET efficiency into a distance between the protein-coupled fluorophores a mutant specific Förster distance (R_0) is needed. The R_0 for a double-cysteine mutant XY can be determined assuming donor on position X and acceptor on Y as well as *vice versa*.

$$R_0 = \sqrt[6]{\frac{9000 \cdot \ln 10}{128 \cdot \pi^5 N_A} \phi_D \kappa^2 n^{-4} J} \qquad \text{(eq. 2.11)}$$

R_0 – Förster distance, N_A – Avogadro number, ϕ_D – donor quantum yield,
κ^2 – orientation factor, n – index of refraction, J – overlap integral

In the calculation of R_0 according to equation 2.11 the orientation factor κ^2 was assumed to be 2/3, a value characteristic for unrestrained rotational flexibility (for some positions rotational averaging had been confirmed via anisotropy measurements, chapter 3.1.1). For the index of refraction $n_{water} = 1.33$ was used since the samples diffuse in aqueous solution. The mutant specific donor quantum yield ϕ_D and the overlap integral of donor emission and acceptor absorption spectra J were determined as follows. The absorption spectrum of the corresponding acceptor labelled single-cysteine mutant was recorded and then the protein was digested (section 2.15.5) for determination of the dye concentration. From these data a normalized absorption spectrum was calculated. For the complementary donor labelled single-cysteine mutant A_{460} and the dye concentration was

determined. Then the fluorescence spectrum of a 10 nM solution was recorded (λ_{exc}= 460 nm, excitation slit 3 nm, emission slit 5 nm, 0.5 s integration time, buffer A) and normalized to have an integrated area of 1. The overlap integral J was then calculated with equation 2.12 using the normalized acceptor absorption and donor emission spectra. Equation 2.12 is an adapted version of the general definition of J presented elsewhere (Lakowicz, 2006) and can be used for A488/A546 as a donor/acceptor pair.

$$J = \sum_{i=475}^{700} F_D(\lambda_i) \cdot \varepsilon_A(\lambda_i) \cdot \lambda_i^4 \qquad \text{(eq. 2.12)}$$

J – overlap integral, $F_D(\lambda_i)$ – donor fluorescence at i nm from the normalized fluorescence spectrum, $\varepsilon_A(\lambda_i)$ – acceptor extinction coefficient at i nm, λ_i – wavelength in nm

The concentration, A_{460} and the fluorescence spectrum (λ_{exc}= 460 nm, excitation slit 3 nm, emission slit 5 nm, 0.5 s integration time) of fluorescein in 0.1 M NaOH were determined and the donor quantum yield ϕ_D was determined relative to fluorescein as described by Parker and Rees (Parker & Rees, 1960) with $\phi_{Fluorescein}$ = 0.92 (Magde *et al.*, 2002).

2.15.7 Glass Chamber Preparation for TIR-FRET Measurements

Glass chambers were assembled from micro slides and coverslips after rigorous cleaning. After cleaning with Alconox detergent (Alconox, White Plains, USA) in a first step, they were sonicated for 2 times 15 min in 0.1 M KOH and 15 min in EtOH. After each step the glasses were thoroughly rinsed with H_2O.

The slides were incubated in methanol containing 2 % (v/v) N-[3-(Trimethoxysilyl)propyl]ethylenediamine, 135 mM acetic acid and 4 % (v/v) H_2O. The mixture was stirred at 25 °C and incubated for 3 h. The glasses were rinsed with H_2O and dried under a nitrogen stream.

Then the surface was functionalized with a mixture of PEG and biotinylated PEG. A solution of 40 mM mPEG-Succinimidyl valerate and 0.4 mM Biotin-PEG-SVA-5000 in 50 mM MOPS (pH 7) was prepared and centrifuged to remove particles. The solution was applied to the glass and incubated for 3 h in the dark. After rinsing with H_2O the surface was dried under a nitrogen stream and chambers were assembled from micro slides and coverslips as shown in Roy *et al.* (2008). The chambers were stored desiccated in the dark.

2.15.8 Immobilization of Fluorophore Labelled Protein for TIR-FRET Measurements

To minimize the unspecific binding of labelled and biotinylated protein (YxiN_C61A_C267A_A115C_D262C_GS), several buffers (all treated with active charcoal to minimize the amount of fluorescent impurities) were tested. Their composition will, related to the results, be mentioned in chapter 3.3.2. The chambers were completely filled with buffer which was then exchanged to a solution of 0.2 mg/ml streptavidin in 10 mM Tris/HCl (pH 7.5), 50 mM NaCl. After incubation for 10 min, the unbound streptavidin was removed by flushing with buffer. The labelled protein in buffer (variable concentration during the optimization) was flushed through, allowed to bind and unbound protein was removed by flushing with buffer. To check for unspecific binding, the same procedure was done without applying streptavidin.

The molecules were visualized using the TIR-microscope (instrumentation section 2.4) and time traces of their fluorescence signals were recorded.

2.16 Quantitation of the ATP Content in ADP Solutions and Reduction of the ATP Concentration via a Hexokinase/Glucose System

Due to its fabrication process, commercially available ADP always contains a reasonable amount of ATP. Since the conformational changes of YxiN were also assayed in the presence of RNA and ADP, it turned out to be relevant to determine the relative amounts of ATP and ADP under the smFRET assay conditions. Additionally, a removal of ATP was favourable to study the effects of ADP only.

To quantify the relative amounts of the nucleotides, they were separated on a C-18 RP-column (1 ml/min; 100 mM sodium phosphate buffer, pH 6.5) recording the A_{260} in the chromatogram. The chromatogram of a reference solution containing ATP, ADP and AMP was used to identify the nucleotides in the samples via their retention volume. From the integrated areas under each absorption peak the relative amount of each species in one sample was calculated.

In smFRET experiments the reduction of the ATP content was achieved by incubation of 5 mM ADP, 4.7 to 6.5 U/ml hexokinase and 25 mM to 37 mM glucose in smFRET assay buffer (buffer A section 2.15.4) for 1 or 2 h at 25 °C or 37 °C (as indicated). Then the fluorescently labelled protein and RNA were added and the smFRET measurement was started.

The activity of the hexokinase was checked by incubation of ADP with hexokinase and glucose under the same conditions and quantitation of the ATP/ADP content as described above. Before the solution was applied to the RP-column hexokinase was precipitated by addition of 91 mM $HClO_4$, followed by neutralization with 1.65 M KOAc and removal of the precipitate via centrifugation. The supernatant was diluted 1:20 or 1:50 with 100 mM sodium phosphate buffer, pH 6.5 and analyzed.

3 Results

3.1 The Mechanism of DEAD Box Protein YxiN

3.1.1 Article:

Cooperative Binding of ATP and RNA Induces a Closed Conformation in a DEAD Box RNA Helicase

Our understanding of enzyme activities can be widely broadened via monitoring the kinetics of the catalysed reaction and studying substrates, cofactors and inhibitors of the reaction. On the next level, the determination of the three dimensional structure using X-ray crystallography or, for comparatively small proteins, NMR studies can unravel enzyme-substrate interactions and often allows for mechanistic conclusions. However, a X-ray structure is a rather static entity. One can only speculate about dynamics and simulate transition states. But enzyme catalysis is a highly dynamic process ranging from local fluctuations and rearrangements up to large conformational changes. FRET is a distance dependent spectroscopic technique that is thus ideally suited to probe relative reorientations within an enzyme. Exploiting FRET in solution, the enzymes can freely diffuse and access their conformational space.

FRET experiments at the single molecule level were established for studying the DEAD box protein YxiN[1]. DEAD box proteins all comprise a core structure with two domains connected by a flexible linker. Therefore they were thought to undergo a large conformational change i.e. a relative rearrangement of these domains during catalysis. This conformational change should be probed employing FRET. To provide YxiN constructs suitable for FRET studies, cysteines accessible for fluorophores had to be removed in a first step. Subsequently, cysteine pairs were introduced on positions that were expected to change their distance during catalysis. The resulting mutants were checked for ATPase and unwinding activity. Finally their conformation in the absence of substrate and in the presence of ADP, ADPNP or RNA was determined in smFRET experiments on a confocal microscope. A conformational change from the open to a closed form was observed in the presence of both ADPNP and RNA. In the ADPNP/RNA ligated state the two core domains approach each other, completing the catalytic centre for ATP hydrolysis. Transitions between the open and the closed form have only rarely been seen. Since lacking a statistic significant number of events, quantitation of the transition is so far not possible.

[1] A.R.K. was involved in discussion of results, in the generation of figures and performed FRET experiments in the presence of ADP/P_i that were asked for during the revision process.

Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase

Bettina Theissen[†], Anne R. Karow[†], Jürgen Köhler[‡], Airat Gubaev[†], and Dagmar Klostermeier[†¶]

[†]Department of Biophysical Chemistry, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland; and [‡]Department of Experimental Physics IV, University of Bayreuth, Universitätsstrasse 30, 95440 Bayreuth, Germany

RNA helicases couple the energy from ATP hydrolysis with structural changes of their RNA substrates. DEAD box helicases form the largest class of RNA helicases and share a helicase core comprising two RecA-like domains. An opening and closing of the interdomain cleft during RNA unwinding has been postulated but not shown experimentally. Single-molecule FRET experiments with the *Bacillus subtilis* DEAD box helicase YxiN carrying donor and acceptor fluorophores on different sides of the interdomain cleft reveal an open helicase conformation in the absence of nucleotides, or in the presence of ATP, or ADP, or RNA. In the presence of ADP and RNA, the open conformation is retained. By contrast, cooperative binding of ATP and RNA leads to a compact helicase structure, proving that the ATP- and ADP-bound states of RNA helicases display substantially different structures only when the RNA substrate is bound. These results establish a closure of the interdomain cleft in the helicase core at the beginning of the unwinding reaction, and suggest a conserved mechanism of energy conversion among DEAD box helicases across kingdoms.

ATP-induced conformational changes | cooperativity | single-molecule FRET | RNA unwinding | YxiN

RNA helicases are ubiquitous enzymes that convert the energy of ATP hydrolysis into structural changes of RNA or RNA/protein complexes. They participate in virtually all processes involving RNA, from transcription, RNA editing, splicing, and translation to RNA decay (for a recent review see ref. 1). DEAD box helicases constitute the largest family of RNA helicases and derive their name from the conserved DEAD sequence in the Walker B motif involved in ATP hydrolysis. In contrast to the highly processive DNA helicases, which unwind thousands of base pairs before dissociating from their DNA substrate, DEAD box helicases unwind only 5–6 nucleobases (2). Recently, an unwinding mode distinct from translocating DNA helicases was suggested for DEAD box helicases (3).

A minimal DEAD box helicase consists of two RecA-like domains, the so-called helicase core, which contains nine conserved motifs required for ATP binding and hydrolysis, RNA binding, and helicase activity. In addition, many helicases contain flanking sequences of varying length that mediate additional functions such as substrate specificity. Although various structures of isolated RecA-like domains of helicases have been determined (4–9), structural information of complete DEAD box helicases (or helicase cores) is limited to eIF4A from *Saccharomyces cerevisiae* (PDB ID code 1fuu) (7), DeaD from *Methanococcus jannaschii* (mjDeaD, PDB ID code 1hv8) (10), the human splicing helicase UAP56, with (PDB ID code 1xtj) and without ADP (PDB ID code 1xti) (11), *S. cerevisiae* Dhh1p (PDB ID code 1s2m) (12), the *Drosophila melanogaster* helicase Vasa in complex with ADPNP and ssRNA (PDB ID code 2db3) (13), and human eIF4A-III as part of the exon junction complex, with ADPNP and ssRNA bound (PDB ID code 2hyi) (14, 15). Most of these structures capture the DEAD box helicases in a surprisingly wide range of open conformations in which the two RecA-like domains of the helicase core are separated and

juxtaposed in different relative orientations, and contacts between the domains are largely absent. In contrast, Vasa and eIF4A-III, both in complex with ADPNP and ssRNA, display closed conformations with extensive contacts between the two RecA-like domains. Based on these structures, the formation of a closed state upon simultaneous binding of ATP and RNA was proposed (13–15). Conformational changes in the catalytic cycle of DEAD box helicases have also been predicted from biochemical data, and a number of models for RNA unwinding have been suggested that require conformational changes (reviewed in ref. 1). Hitherto, structures of open and closed conformations for one particular DEAD box helicase have not been determined. Single-molecule FRET (smFRET) experiments are ideally suited to investigate conformational changes in the catalytic cycle of RNA helicases in solution. Here, we identify different conformational states of the DEAD box RNA helicase YxiN from *Bacillus subtilis* that depend on the ligation state of the helicase. YxiN is involved in ribosome biogenesis (19). Specific binding of YxiN to hairpin 92 of the 23S rRNA is mediated by a C-terminal extension of the helicase core (20), which contains a classical RNA recognition motif (21).

YxiN adopts an open conformation in the absence of ligand, or in the presence of nucleotides or RNA. The open conformation is retained upon simultaneous binding of ADP and RNA. In contrast, cooperative binding of RNA and ATP promotes a closure of the cleft between the RecA-like domains in the helicase core, leading to a compact, closed conformation. Upon ATP hydrolysis, YxiN returns to the open conformation. We observe direct switching between these states in smFRET experiments. Altogether, these data can be integrated into a "switch-kink" model for YxiN-mediated RNA unwinding and for RNA unwinding by DEAD-box helicases in general.

Results

Double-Cysteine Mutants for Donor/Acceptor Labeling of YxiN Are Functional Helicases. To investigate conformational changes in the catalytic cycle of the RNA helicase YxiN, we constructed double-cysteine mutants with one cysteine for fluorescent labeling on each side of the interdomain cleft. YxiN contains four intrinsic cysteines at positions 61, 243, 247, and 267. According to modification reactions of individual cysteine to serine mutants with Ellman's reagent (22), or Alexa Fluor488− (A488[§]−) or tetramethylrhodamine-maleimide, two of these cysteines (C243, C247) are not solvent-accessible and do not interfere with labeling reactions at site-specifically introduced cysteines.

Therefore, the YxiN C61S/C267S mutant, YxiN*, was generated. Positions 108 and 115 in the N-terminal and positions 224, 229, and 262 in the C-terminal domain of the helicase core were selected for labeling (Fig. 1A), and double-cysteine mutants (S108C/E224C, S108C/S229C, A115C/E224C, A115C/S229C, and A115C/D262C) were generated. All mutants are wild type-like in terms of secondary structure, as demonstrated by their similar far-UV CD spectra (data not shown). They exhibit wild type-like RNA-stimulated ATPase activity, with $K_{M,app}$ values for RNA from 114 to 496 nM (wild type: 156 nM), and k_{cat} values of 0.64–1.41 s^{-1} (wild type: 1.48 s^{-1}, Table 1). Most importantly, all mutants unwind a model RNA substrate in an ATP-dependent manner (Fig. 1B) with yields comparable with wild-type YxiN, confirming that, despite the modifications, all constructs are functional RNA helicases.

YxiN Conformation in the Absence of Ligands. A limited number of crystal structures of DEAD box RNA helicases have been determined, most of these in the absence of ligands (7, 10–12). In all cases, the structures of the individual N- and C-terminal RecA-like domains of the helicase core are highly similar. The domains are splayed apart, with no contacts between them, and their relative orientation varies tremendously between the different structures.

To define the global conformation of YxiN in solution, smFRET experiments were performed on freely diffusing helicase molecules in a confocal microscope. YxiN mutants with one cysteine on each side of the interdomain cleft were labeled with A488 (donor) and Alexa Fluor 546 (A546, acceptor). The donor quantum yields and Förster distances are summarized in supporting information (SI) Tables 2 and 3. FRET histograms for all constructs (Fig. 1C) show a unimodal distribution of FRET efficiencies (E_{FRET}), demonstrating that the helicase in solution adopts a single conformation already in the absence of ligands, in stark contrast to the different conformations observed in the crystal structures. The interdye distances calculated from E_{FRET} distributions, R_{calc}, (Fig. 1C) range from 4.82 to 5.65 nm and are in agreement with an open conformation of the interdomain cleft. Because of the low sequence homology between DEAD box helicases outside the conserved motifs, homology models of YxiN were generated with Geno3D by using the structures of eIF4A (PDB ID code 1fuu), mjDeaD (PDB ID code 1hv8), Dhh1p (PDB ID code 1s2m), Vasa (PDB ID code 2db3), and eIF4A-III (PDB ID code 2hyi) as templates (23) (see SI Methods) to facilitate the comparison of our experimental data with the structures of different helicases. The Cβ–Cβ distances for the cysteines used for dye coupling according to these models (SI Table 4) are virtually identical to those determined directly from the crystal structures. Strikingly, all experimental distances are shorter than expected from these models, indicating that the helicase core conformation in solution is much more compact. Most likely, the crystal structures represent an extreme opening of the cleft due to stabilization of artificial open conformations by crystal contacts. Close inspection of the low-FRET region shows that YxiN does not adopt a wide-open conformation similar to the eIF4A structure to any significant extent (SI Fig. 6 and SI Table 5).

The best agreement of the experimental distances is found with the structure of the *M. jannaschii* DeaD helicase (Fig. 1 A and C), with a mean distance deviation of 0.55 nm, and individual deviations ranging from 0.21 to 0.92 nm. Taking into account the length of the linkers between the cysteines and the attached dyes (0.5–1.0 nm), deviations in this range are expected.

The full width at half maximum (FWHM) of the E_{FRET} distributions is 0.3–0.4, corresponding to an average interdye distance variation of 1.4 nm (SI Table 6). Although the width of the distribution is influenced by the segmental flexibility of the
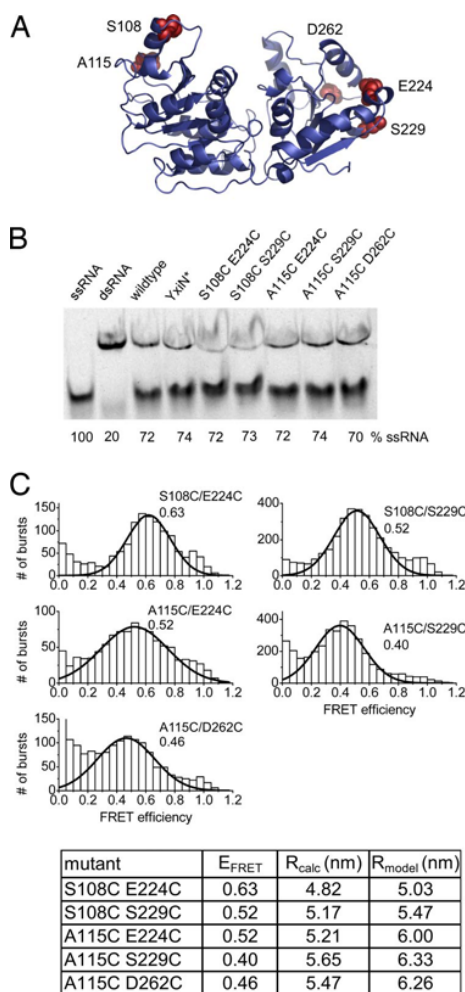
Fig. 1. Conformation of YxiN in the absence of ligands. (*A*) Homology model of YxiN (residues 3–367) according to the structure of *M. jannaschii* DeaD (mjDeaD, PDB ID code 1hv8). The two RecA-like domains of the helicase core are separated by a large interdomain cleft. The positions used for fluorophore attachment, 108 and 115 in the N-terminal, and 224, 229, and 262 in the C-terminal RecA-like domain, are highlighted. (*B*) RNA unwinding by wild-type YxiN and mutants. All constructs show similar unwinding yields as wild-type YxiN after 30 min of incubation with RNA substrate. (*C*) FRET histograms (*Upper*) for donor–acceptor-labeled YxiN constructs. All histograms display a unimodal distribution of FRET efficiencies, E_{FRET}, consistent with a single conformation of YxiN in the absence of ligand. Lines show Gaussian distributions fitted to the data (see SI Table 6). The numbers denote the mean FRET efficiency. The table (*Lower*) summarizes the values for E_{FRET}, the experimental distances calculated from E_{FRET}, R_{calc}, in nm, and the distances calculated from the homology model in A (Cβ–Cβ), R_{model}.

mutant	E_{FRET}	R_{calc} (nm)	R_{model} (nm)
S108C E224C	0.63	4.82	5.03
S108C S229C	0.52	5.17	5.47
A115C E224C	0.52	5.21	6.00
A115C S229C	0.40	5.65	6.33
A115C D262C	0.46	5.47	6.26

dyes attached to the protein, this broad range of experimental FRET efficiencies points toward a significant flexibility of the interdomain linker and thus of the open conformation of YxiN.

No Influence of Nucleotide and RNA Substrate on the Global Helicase Conformation. To probe possible nucleotide-induced conformational changes of YxiN, we measured the FRET efficiency of all double-labeled constructs in the presence of saturating amounts of ADP, ATP, or the nonhydrolyzable ATP analog ADPNP (Fig. 2A). According to the K_D values of the YxiN/nucleotide complexes (ref. 24 and Fig. 3C), >94% saturation is achieved in these experiments. The FRET histograms are similar to the histograms for YxiN in the absence of nucleotide for each construct, irrespective of the nucleotide present. These results clearly show

Table 1. RNA-stimulated ATPase of YxiN constructs

| Mutant	k_{cat}, s^{-1}	$K_{M,app}$ (RNA), nM
YxiN (wild type)	1.48 ($\pm$0.22)	156 ($\pm$74)
YxiN*	0.98 ($\pm$0.10)	170 ($\pm$61)
YxiN* S108C/E224C	0.85 ($\pm$0.07)	458 ($\pm$76)
YxiN* S108C/S229C	0.64 ($\pm$0.07)	114 ($\pm$47)
YxiN* A115C/E224C	1.40 ($\pm$0.26)	496 ($\pm$180)
YxiN* A115C/S229C	1.30 ($\pm$0.25)	324 ($\pm$149)
YxiN* A115C/D262C	1.41 ($\pm$0.21)	422 ($\pm$130)

that the ATP- and ADP-bound conformations of YxiN in solution are virtually identical and resemble the open conformation in the absence of ligand. Hence, nucleotide binding does not induce a significant global conformational change. Similarly, FRET histograms of YxiN in the presence of RNA substrate, but without nucleotide, show no indication for a change of the donor–acceptor distances upon RNA binding (Fig. 2*B*). Based on the $K_{M,app}$ values for the RNA substrate (Table 1), the saturation of YxiN with RNA is 30–65%, and a second species with a different FRET efficiency would be clearly visible in the histogram. The results are consistent within all constructs and establish that YxiN retains the open conformation of the helicase core upon RNA binding. Neither nucleotide nor RNA alone are sufficient to convert YxiN into a compact, catalytically competent helicase.

Cooperative Binding of RNA and ATP Induces a Closure of the Interdomain Cleft. In contrast, smFRET experiments in the presence of both ADPNP/ATP and RNA substrate, show the appearance of a high-FRET species for all YxiN constructs. The data illustrating the closure of the interdomain cleft in response to ADPNP and RNA binding for the YxiN* S108C/S229C and A115/S229C mutants are shown in Fig. 3, and the distances determined for the closed conformation for all constructs are summarized in Fig. 4. In contrast, no such conformational change occurs in the presence of ADP and RNA (Fig. 3), but the open conformation is retained.

The structures of the DEAD box RNA helicases Vasa and
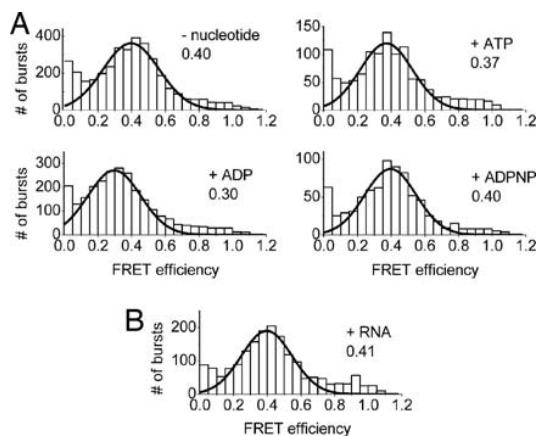
Fig. 2. Nucleotide or RNA binding does not influence the conformation of YxiN. (*A*) FRET histograms for YxiN* A115C/S229C (A488/A546), in the absence of ligands, and in the presence of ADP, ATP, and the nonhydrolyzable ATP analog ADPNP. (*B*) FRET histogram for YxiN* A115C/S229C (A488/A546) in the presence of 200 nM RNA substrate. RNA binding does not induce a conformational change of YxiN. Nucleotides or RNA do not affect the FRET efficiencies and thus do not induce a global conformational change. Lines show Gaussian distributions fitted to the data (SI Table 6). The numbers denote the mean FRET efficiency.
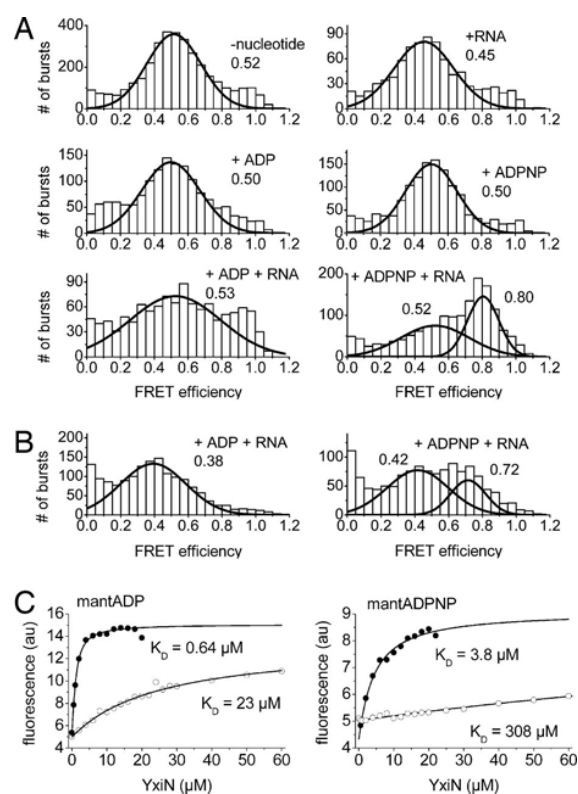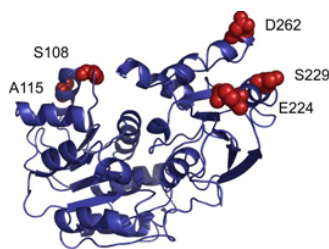

Fig. 3. Cooperative binding of RNA and ADPNP induces a closure of the interdomain cleft. Neither ADP or ADPNP nor RNA binding show any effect on the helicase conformation, whereas binding of RNA and ADPNP decreases the FRET efficiency from $\approx$0.5 in the apoprotein to $\approx$0.8, corresponding to a change in interdye distance from 5.16 to 4.13 nm. No such effect is detected upon binding of ADP and RNA. Consequently, the structures of the helicase in the ATP and the ADP form are significantly different when RNA is bound. (*A*) FRET histograms for YxiN* S108C/S229C in the absence of ligands, in the presence of ADP or ADPNP, of RNA, or of ADP and RNA or ADPNP and RNA.(*B*) FRET histograms for YxiN* A115C/S229C in the presence of ADP and RNA or ADPNP and RNA. (*C*) Cooperative binding of ADPNP and RNA. mantADP (1 μM) (*Left*) or mantADPNP (1 μM) (*Right*) were titrated with YxiN (wild type) in the presence (filled symbols) and absence (open symbols) of 154-mer RNA. The K_D values determined are 23 ($\pm$5.9) μM (mantADP/YxiN) and 308 ($\pm$16) μM (ADPNP) in the absence, and 0.64 ($\pm$0.05) μM (mantADP) and 3.8 ($\pm$0.6) μM (ADPNP) in the presence of RNA, corresponding to a 35-fold increase in mantADP affinity and a 80-fold increase in mantADPNP affinity when RNA is bound.

eIF4A-III in complex with RNA and ADPNP (refs. 13–15 and Fig. 4) show a closed conformation of the helicase core, with an extensive interface between the two RecA-like domains. The donor–acceptor distances calculated from the FRET histograms of double-labeled YxiN in the presence of ADPNP and RNA agree with the values expected from Vasa and eIF4A-III for all constructs studied here. The average differences between experimental and expected distances are low, with 0.56 nm (eIF4A-III, range 0.18–0.87 nm) or 0.55 nm (Vasa, range 0.11–0.71 nm), which confirms that YxiN adopts a similar compact conformation as Vasa and eIF4A-III in the presence of ADPNP and RNA substrate. Additionally, the FWHM of the E_{FRET} distribution is reduced to 0.1–0.2 in the closed conformation, corresponding to an average interdye distance variation of 0.87 nm (SI Table 6), consistent with an increased rigidity compared with the open conformation.

For a few helicases, cooperative binding of ATP or ADPNP and RNA substrate, but not of ADP and RNA substrate, has

53

Fig. 4. The closed conformation of the RNA helicase YxiN. Homology model of YxiN (1–354) based on the crystal structure of the DEAD box helicase Vasa in the presence of RNA and ADPNP (PDB ID code 2db3). The table below summarizes FRET efficiencies of the closed conformation from smFRET experiments for all constructs, the interdye distance calculated from E_{FRET}, R_{calc}, and the distance expected from the homology model (Cβ–Cβ), R_{model}, in nm.

mutant	E_{FRET}	R_{calc} (nm)	R_{model} (nm)
S108C E224C	0.70	4.54	3.71
S108C S229C	0.80	4.13	4.03
A115C E224C	0.81	4.13	4.84
A115C S229C	0.72	4.52	5.11
A115C D262C	0.88	3.81	4.30

been demonstrated (3, 16, 18). We have shown before that mantADP is a suitable fluorescent probe for nucleotide binding to YxiN (24) and we have now investigated the effect of RNA substrate on the interaction of YxiN with nucleotides in fluorescence equilibrium titrations (Fig. 3C). For the mantADPNP/YxiN complex in the absence of RNA, the K_D value is 308 ($\pm$16) μM. In the presence of RNA, this value decreases 80-fold to 3.8 ($\pm$0.6) μM, revealing a clear coupling between mantADPNP and RNA binding. Although mantADP binding is also affected by RNA, the increase in affinity is only 35-fold, with a K_D value of 23 ($\pm$6) μM in the absence and 0.64 ($\pm$0.05) μM in the presence of RNA. As a consequence of these different extents of cooperativity, the affinities for ATP and ADP are different in the absence of RNA [346 ($\pm$29) and 55 ($\pm$4) μM, respectively (24)] but very similar in the presence of RNA ($\approx$4.3 and $\approx$1.6 μM). In turn, these results prove that RNA binds 2- to 3-fold more tightly to the ATP form of YxiN than to the ADP form. A similar factor is estimated from comparison of $K_{M,app}$ values for RNA from ATPase assays at saturating and subsaturating ATP concentrations (data not shown).

These results demonstrate that the ATP- and ADP-bound states of RNA helicases display substantially different structures only in the presence of RNA substrate. Although the cooperative binding of ATP and RNA promotes the closure of the interdomain cleft, binding of ADP and RNA is not sufficient for this conformational change.

Direct Transitions Between Open and Closed Conformations. A strict coupling of the simultaneous binding of ATP and RNA substrate with the conformational change of the YxiN helicase ensures a tight regulation of ATP hydrolysis and RNA helicase activity. Consequently, a coordination of the switching between the two conformations with ATP hydrolysis is expected. The rate constants for ATP hydrolysis by the YxiN constructs at saturating RNA concentrations are 0.64–1.41 s^{-1} (Table 1), translating into one hydrolysis event every 0.71–1.56 s on average. The observation time in the confocal microscope is typically <5 ms, with only a few events >10 ms. Therefore, the probability to observe such a switching event on this time scale is <1%. Nevertheless we recorded FRET traces with a jump from low to high FRET, corresponding to a closure of the interdomain cleft or with a sudden drop from high to low FRET, reflecting an opening of the interdomain cleft (SI Fig. 7). Although confocal microscopy

does not allow kinetic experiments on this time scale, our experiments open avenues to investigate the conformational changes of RNA helicases during their catalytic cycle in real time by increasing the observation time, e.g., by immobilizing the double-labeled helicases on a surface.

In summary, we have shown that ATP and RNA bind cooperatively to the DEAD box helicase YxiN and promote a closure of the interdomain cleft in the helicase core. ATP hydrolysis induces a reopening of the interdomain cleft and reduces the affinity for RNA. The similar closed conformations of YxiN, Vasa, and eIF4A-III suggest a conserved mechanism of energy conversion among DEAD box helicases.

Discussion

We have shown here that the helicase core of the RNA helicase YxiN adopts a flexible, open conformation in solution, with a wide cleft between the two RecA-like domains. In the absence of RNA, this open conformation is retained irrespective of nucleotides bound. Only the cooperative binding of ATP and RNA induces a concerted conformational change, leading to a well defined closed conformation of the helicase core. In single-molecule experiments, the direct switching of the helicase between open and closed conformations during the unwinding cycle is observed.

Flexible proteins such as DEAD box helicases are notoriously difficult to crystallize and are prone to artifacts, which may explain why only a limited number of full-length helicase structures have been determined. SmFRET is an ideal technique to characterize the conformation of such a flexible enzyme in solution. The measured E_{FRET} distributions are rather broad, characteristic of high interdomain flexibility, but do not cover such a broad range as the crystal structures suggest. All interdye distances from FRET experiments are shorter than expected from the structures of open helicase conformations, demonstrating that the crystal structures do not reflect the open conformation of DEAD box helicases in solution. Rather, the opening of the interdomain cleft is exaggerated, probably because of capture of rare conformations that maximize crystal contacts.

In 1998, a nucleotide- and RNA-induced conformational cycle was suggested for the DEAD box helicase eIF4A based on different limited proteolysis patterns in the absence and presence of ADP, ADPNP, or RNA (16, 17). Despite numerous efforts (11, 12, 25, 26), the nature of these conformational changes and the trigger for their interconversion has remained elusive. The proposed closure of the interdomain cleft in helicases in response to ATP binding has been difficult to prove because no ATP-bound DEAD box helicase crystal structure has been reported, and there is no DEAD box helicase for which structures of two (or more) different conformations have been determined. Our single-molecule studies reveal open and closed conformations of the same helicase and demonstrate that neither ADP nor ATP binding alone induce a conformational change in the helicase core. Similarly, the helicase does not respond to RNA binding alone but retains its open conformation. Only when ATP/ADPNP and RNA are bound simultaneously does the closed conformation become populated. The closed conformation is well defined and in agreement with the helicase core conformation in the structures of Vasa (13) and eIF4A-III (14, 15). The structural conservation of the DEAD box helicase core and the identical arrangement of all nine conserved helicase motifs on this scaffold suggest that the closed conformation is conserved among DEAD box helicases across the kingdoms of life.

In crystal structures of open helicase conformations, residues from both RecA-like domains that are involved in ATP hydrolysis are far apart. The ATPase site becomes fully assembled only when the closed conformation is induced (13–15). The switching from an inactive open conformation to the active closed con-

formation only when both ATP and RNA are present ensures a tight coupling between ATP hydrolysis and structural changes in the RNA substrate. Motif IV is the only motif that interacts with both ADPNP and RNA in the closed conformation (13), which renders this motif a likely candidate to act as a sensor for the two ligand-binding sites. In addition, various "uncoupled" mutants that retain ATPase and RNA binding but lack RNA unwinding activity have been described. These mutants may be deficient in undergoing the conformational change in the presence of RNA and ATP, thus preventing helicase activity.

The structures of Vasa and eIF4A-III (13–15) define the bipartite RNA-binding site constituted by residues on the surface of both RecA-like domains. Tight RNA binding therefore requires contacts with both domains, which is impossible in the open conformation. Chemical and enzymatic footprinting of RNA bound to DbpA revealed that the conformation of helicase-bound RNA is not affected by ADP binding but changes significantly when ADPNP is bound (27). Consistent with this change in footprinting patterns in the presence of ADPNP, the RNA bound to eIF4A-III and Vasa exhibits a kinked backbone, and it was suggested that the kink is stabilized only when ATP binding favors closure of the interdomain cleft that leads to formation of the extended RNA-binding surface (13–15). The kink in the RNA backbone is not compatible with double-stranded RNA (13), and switching between the two helicase conformations thus leads to local RNA unwinding. This distortion of the local RNA structure also explains other structural changes in the RNA substrate promoted by DEAD box helicases, such as displacement of proteins (28).

A destabilization model has been proposed for the mechanism of RNA helicases (1, 29), according to which either ATP binding or ATP hydrolysis induce a closure of the interdomain cleft and thereby change the affinity of the helicase for its RNA substrate. The smFRET experiments allowed us to capture the conformations of the DEAD box helicase YxiN before ATP hydrolysis (in the presence of ADPNP and RNA) and after hydrolysis (in the presence of ADP and RNA), and the open and closed conformations are now unambiguously assigned to the individual ligation states: The interdomain cleft is closed in the ATP state and open in the ADP state.

These results can be reconciled in a modified destabilization or "switch-kink" model for DEAD box helicase-mediated RNA remodeling (Fig. 5). In the absence of ligand, or in the presence of either ATP or RNA, the helicase adopts an inactive, open conformation. In the presence of RNA, the helicase acts as a nucleotide-dependent switch: ATP, but not ADP, binding leads to a closure of the interdomain cleft. In the closed conformer, the ATPase site and the bipartite RNA binding site are assembled. At the same time, the unwinding reaction is initiated by kinking the RNA backbone, which separates the adjacent base pairs. When ATP hydrolysis generates the ADP state, the helicase returns to the open form. As a consequence, the ATPase- and RNA-binding sites are disrupted, and the affinity for the RNA is reduced. RNA and ADP release allow for further catalytic cycles.

This "switch-kink" model rationalizes the RNA-stimulated ATPase activity and the cooperative binding of RNA and ATP or ADPNP but not (to the same extent) of RNA and ADP (16, 18, 30, 31) (Fig. 3C). In addition, it illustrates why DEAD box helicases do not act as highly processive motor proteins but as nonprocessive switches. The RNA kink is observed with the nonhydrolyzable ADPNP, suggesting that ATP binding is sufficient to induce this structural change in the RNA. However, even in large excess over RNA substrate, YxiN does not efficiently catalyze RNA unwinding in the presence of ADPNP (ref. 20 and SI Fig. 8). Hence, an additional "power stroke" upon ATP hydrolysis appears to be required to complete the unwinding reaction.
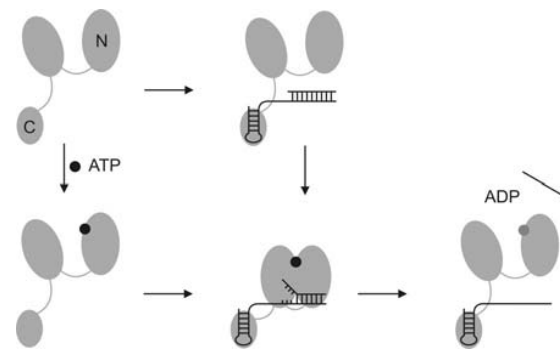
Fig. 5. "Switch-kink model" for DEAD box helicase-mediated RNA unwinding. The two RecA-like domains of the helicase core are depicted as large gray ovals (N, N-terminal domain), and the C-terminal domain mediating RNA binding is depicted as a small oval (C, C-terminal RNA-binding domain). The hairpin of the RNA substrate binds to the RNA recognition motif in the C-terminal domain, and ATP binds to the N-terminal RecA domain in the helicase core. Upon binding of RNA and ATP, YxiN adopts a closed conformation. The conformational change of the helicase core assembles the ATPase site and the RNA-binding site, establishing contacts with the adjacent double-helical region of the RNA. Introduction of a kink into the RNA backbone leads to local unwinding ("switch-kink" model). Upon ATP hydrolysis, YxiN returns to the open conformation. The affinity for RNA substrate is reduced, and the RNA is released. Kinking the RNA is not sufficient for unwinding, but an additional "power stroke" upon ATP hydrolysis is required to complete the catalytic cycle.

ATP binding and ADP release are rapid compared with the steady-state ATPase rate of YxiN and DbpA (data not shown and ref. 32), leaving either ATP hydrolysis, phosphate release, or an associated conformational change as the rate-limiting step. It is likely that the large conformational change from the open to the closed helicase constitutes this rate-limiting step in the overall helicase cycle. Our experiments open avenues to track this conformational change in the catalytic cycle in real time and to correlate its kinetics with unwinding activity to answer this question.

Materials and Methods

Cloning, Mutagenesis, Protein Production, and Purification. YxiN wild type was purified as described (24). The YxiN mutant C61S/C267S (YxiN*) was generated via stepwise site-directed mutagenesis (QuikChange; Stratagene) and used as a template for generating double-cysteine mutants for donor–acceptor labeling (S108C/E224C, S108C/S229C, A115C/E224C, A115C/S229C, and A115C/D262C). Sequences were confirmed, and proteins were produced and purified as wild-type YxiN.

Fluorescence Measurements. Fluorescence measurements were performed at 20°C by using a Fluoromax-3 fluorimeter. mantADP (1 μM) or mantADPNP (1 μM) in 50 mM Tris·HCl (pH 7.5), 150 mM NaCl, 5 mM MgCl$_2$, 2 mM 2-mercaptoethanol were titrated with YxiN in the absence and presence of saturating concentrations of 154-mer RNA substrate (200 nM), and K_D values were determined as described (24). Amplitudes in the absence and presence of RNA were fixed during the fit.

Steady-State ATPase Assay. ATP hydrolysis was monitored in a coupled enzymatic assay at 37°C via the decrease in A$_{340}$ because of oxidation of NADH to NAD$^+$ (33) as described (24).

RNA Substrates and Unwinding Assay. A 154-mer comprising nucleotides 2,481–2,634 of the *B. subtilis* 23S rRNA was generated by T7 polymerase *in vitro* transcription as described (24). In all experiments, a C2538G/C2606G mutant (lacking an internal T7 promotor sequence) was used.

For RNA-unwinding assays, a minimal helicase substrate consisting of a synthetic, fluorescently labeled 9-mer and a 32-mer was used. Unwinding reactions were performed as described (24).

Fluorescent Labeling. Proteins were labeled with a mixture of A488- (donor) and A546-maleimide (acceptor), respectively, for 1 h at 25°C in the presence of 1 mM Tris(2-carboxyethyl)phosphine. The ratio of the dyes was varied from 1:0.5:5 (protein/donor/acceptor) to 1:5:10 to maximize the formation of donor/acceptor-labeled protein. Labeling efficiencies were determined from absorbance ratios at 495 nm (A488, corrected for A546 contributions) or 555 nm (A546) and 280 nm (protein, corrected for dye contributions).

Determination of Quantum Yields and Förster Distances. Quantum yields of the donor A488 attached to positions 108, 115, 224, 229, or 262 of YxiN were determined as described (34) relative to fluorescein in 0.1 M NaOH (ϕ = 0.92) (35).

Förster distances were calculated according to Eq. **1** (36) from normalized absorbance spectra of acceptor-only labeled protein, $\varepsilon_A(\lambda)$, normalized fluorescence spectra of the donor-only labeled protein, $F_D(\lambda)$, and the quantum yield of the donor, ϕ_D. The refractive index n of water is 1.33, and N is the Avogadro constant. Rapid rotational averaging of the dyes attached to the protein was confirmed in anisotropy decays of singly labeled YxiN (see *SI Methods*), and the orientation factor κ^2 was therefore set as 2/3.

$$R_0^6 = \frac{9{,}000 \cdot \ln 10 \cdot \phi_D \cdot \kappa^2}{128 \cdot \pi^5 \cdot N \cdot n^4} \cdot \int_0^\infty F_D(\lambda) \cdot \varepsilon_A(\lambda) \cdot \lambda^4 d\lambda. \quad \textbf{[1]}$$

Single-Molecule FRET Experiments. SmFRET experiments were performed by using a home-built confocal microscope (*SI Methods*). Only fluorescence bursts above a threshold of 50 photons in total were considered in the analysis. Measured background-corrected fluorescence intensities (I_D, I_A) were corrected for cross-talk (α, donor cross-talk in acceptor channel, β, acceptor cross-talk in donor channel), different quantum yields and detection efficiencies of donor and acceptor fluorescence (γ), and direct excitation of the acceptor (δ) and converted into FRET efficiencies according to Eq. **2**.

$$E_{\mathrm{FRET}}$$
$$= \frac{(1 + \beta\gamma\delta) \cdot \left(I_A - \dfrac{\alpha + \gamma\delta}{1 + \beta\gamma\delta} \cdot I_D \right)}{(1 + \beta\gamma\delta) \cdot \left(I_A - \dfrac{\alpha + \gamma\delta}{1 + \beta\gamma\delta} \cdot I_D \right) + (\gamma + \gamma\delta) \cdot (I_D - \beta I_A)}. \quad \textbf{[2]}$$

Measurements were performed at room temperature (25°C) in 50 mM Tris·HCl (pH 7.5), 150 mM NaCl, 5 mM MgCl$_2$ with 40 pM YxiN (concentration of donor fluorophore), 5 mM nucleotide, and 200 nM 154-mer RNA.

1. Cordin O, Banroques J, Tanner NK, Linder P (2006) *Gene* 367:17–37.
2. Rogers GW, Jr, Richter NJ, Merrick WC (1999) *J Biol Chem* 274:12236–12244.
3. Yang Q, Jankowsky E (2006) *Nat Struct Mol Biol* 13:981–986.
4. Rudolph MG, Heissmann R, Wittmann JG, Klostermeier D (2006) *J Mol Biol* 361:731–743.
5. Caruthers JM, Hu Y, McKay DB (2006) *Acta Crystallogr F* 62:1191–1195.
6. Kurimoto K, Muto Y, Obayashi N, Terada T, Shirouzu M, Yabuki T, Aoki M, Seki E, Matsuda T, Kigawa T, *et al.* (2005) *J Struct Biol* 150:58–68.
7. Caruthers JM, Johnson ER, McKay DB (2000) *Proc Natl Acad Sci USA* 97:13080–13085.
8. Benz J, Trachsel H, Baumann U (1999) *Structure (London)* 7:671–679.
9. Johnson ER, McKay DB (1999) *RNA* 5:1526–1534.
10. Story RM, Li H, Abelson JN (2001) *Proc Natl Acad Sci USA* 98:1465–1470.
11. Shi H, Cordin O, Minder CM, Linder P, Xu RM (2004) *Proc Natl Acad Sci USA* 101:17628–17633.
12. Cheng Z, Coller J, Parker R, Song H (2005) *RNA* 11:1258–1270.
13. Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S (2006) *Cell* 125:287–300.
14. Andersen CB, Ballut L, Johansen JS, Chamieh H, Nielsen KH, Oliveira CL, Pedersen JS, Seraphin B, Le Hir H, Andersen GR (2006) *Science* 313:1968–1972.
15. Bono F, Ebert J, Lorentzen E, Conti E (2006) *Cell* 126:713–725.
16. Lorsch JR, Herschlag D (1998) *Biochemistry* 37:2180–2193.
17. Lorsch JR, Herschlag D (1998) *Biochemistry* 37:2194–2206.
18. Polach KJ, Uhlenbeck OC (2002) *Biochemistry* 41:3693–3702.
19. Kossen K, Uhlenbeck OC (1999) *Nucleic Acids Res* 27:3811–3820.
20. Kossen K, Karginov FV, Uhlenbeck OC (2002) *J Mol Biol* 324:625–636.
21. Wang S, Hu Y, Overgaard MT, Karginov FV, Uhlenbeck OC, McKay DB (2006) *RNA.*
22. Ellman GL (1959) *Arch Biochem Biophys* 82:70–77.
23. Combet C, Jambon M, Deleage G, Geourjon C (2002) *Bioinformatics* 18:213–214.
24. Karow AR, Theissen B, Klostermeier D (2007) *FEBS J* 274:463–473.
25. Talavera MA, Matthews EE, Eliason WK, Sagi I, Wang J, Henn, A., De La Cruz EM (2006) *J Mol Biol* 355:697–707.
26. Henn A, Shi SP, Zarivach R, Ben-Zeev E, Sagi I (2002) *J Biol Chem* 277:46559–46565.
27. Karginov FV, Uhlenbeck OC (2004) *Nucleic Acids Res* 32:3028–3032.
28. Fairman ME, Maroney PA, Wang W, Bowers HA, Gollnick P, Nilsen TW, Jankowsky E (2004) *Science* 304:730–734.
29. Rogers GW, Jr, Lima WF, Merrick WC (2001) *J Biol Chem* 276:12598–12608.
30. Tsu CA, Uhlenbeck OC (1998) *Biochemistry* 37:16989–16996.
31. Iost I, Dreyfus M, Linder P (1999) *J Biol Chem* 274:17677–17683.
32. Talavera MA, De La Cruz EM (2005) *Biochemistry* 44:959–970.
33. Adam H (1962) in *Methoden der Enzymatischen Analyse* (Bergmeyer, Weinheim, Germany), pp 573–577.
34. Parker CA, Rees WT (1960) *Analyst* 85:587–600.
35. Magde D, Wong R, Seybold PG (2002) *Photochem Photobiol* 75:327–334.
36. Förster T (1959) *Disc Faraday Soc* 27:7–17.

56

3.1.2 Article:

A Conformational Change in the Helicase Core is Necessary but not Sufficient for RNA Unwinding by the DEAD box Helicase YxiN

As the previous article depicts, the DEAD box protein YxiN adopts at least two different conformations during catalysis. The closed conformation is populated when ADPNP and the RNA substrate are cooperatively bound. Although the YxiN constructs used were active RNA helicases, their ATPase activity was impaired to some extent. The k_{cat} and the apparent RNA affinity were lower than for the wild-type enzyme. This may have influence on the kinetics of transitions. To improve the system, constructs that more closely display wild-type like behaviour are required. In the preceding work, YxiN mutants were used in which two native cysteines had been replaced by serines. Replacing the cysteines by alanines led to a significant improvement to wild-type like ATPase kinetic parameters. In addition, using these constructs it was now possible to verify the main conclusions from the previous article and to even populate the ATP/RNA ligated state of the enzyme.

It has been a matter of debate whether ADPNP is a suitable ATP analogue and the ADPNP/RNA ligated form resembles the prehydolysis state. Hence, an alternative to ADPNP usage in FRET experiments was established. DEAD box proteins carrying a mutation in the Walker A motif (motif I) have been reported to still bind nucleotide but to be ATPase and unwinding deficient. Therefore a hydrolysis deficient YxiN construct was created and its conformation in the absence and presence of substrates was studied. Instead of providing the non-hydrolysable analogue ADPNP in FRET experiments, ATP can be used and should stall the enzyme mutant in the prehydrolysis state.

As it was discussed in the previous article the study of so called uncoupled mutants might broaden our understanding of the DEAD box protein mechanism. Mutants of the helicase motif III and motif V have been described to retain ATPase activity while being unwinding deficient (Pause & Sonenberg, 1992; Rocak *et al.*, 2005; Cheng *et al.*, 2005). What happens to the conformational change in these mutants? Answering this question can provide a causal link between the conformational change and the activities. Moreover, it can help to understand how ATPase activity is actually coupled to unwinding. In the following article the conformations of YxiN uncoupling mutants are presented and related to ATPase and helicase kinetic data.

A conformational change in the helicase core is necessary but not sufficient for RNA unwinding by the DEAD box helicase YxiN

Anne R. Karow and Dagmar Klostermeier*

University of Basel, Biozentrum, Biophysical Chemistry, Klingelbergstrasse 70, 4056 Basel, Switzerland

ABSTRACT

Cooperative binding of ATP and RNA to DEAD-box helicases induces the closed conformation of their helicase core, with extensive interactions across the domain interface. The bound RNA is bent, and its distortion may constitute the first step towards RNA unwinding. To dissect the role of the conformational change in the helicase core for RNA unwinding, we characterized the RNA-stimulated ATPase activity, RNA unwinding and the propensity to form the closed conformer for mutants of the DEAD box helicase YxiN. The ATPase-deficient K52Q mutant forms a closed conformer upon binding of ATP and RNA, but is deficient in RNA unwinding. A mutation in motif III slows down the catalytic cycle, but neither affects the propensity for the closed conformer nor its global conformation. Hence, the closure of the cleft in the helicase core is necessary but not sufficient for RNA unwinding. In contrast, the G303A mutation in motif V prevents a complete closure of the inter-domain cleft, affecting ATP binding and hydrolysis and is detrimental to unwinding. Possibly, the K52Q and motif III mutants still introduce a kink into the backbone of bound RNA, whereas G303A fails to kink the RNA substrate.
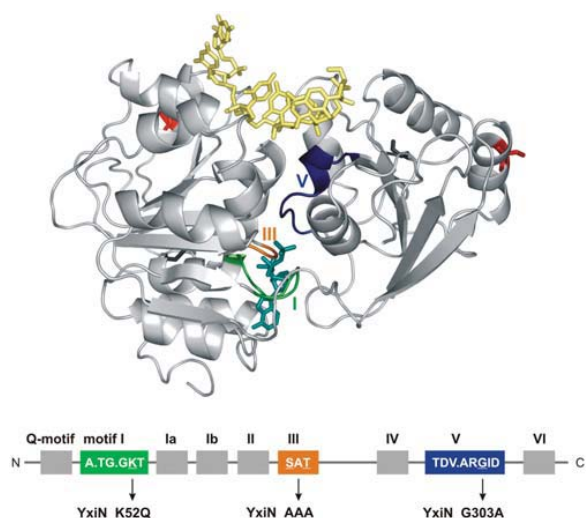
INTRODUCTION

Members of the DEAD-box family of RNA helicases participate in all processes involving structural changes of RNA. Their common structure is a helicase core composed of two flexibly linked RecA domains (Figure 1). In many helicases, the core is flanked by additional domains that mediate interactions with other proteins, or provide high affinity- and/or high specificity binding sites for RNA substrates. Thus, these flanking domains define the specific function of a particular helicase,

whereas the helicase core is a degenerate nucleotide-dependent switch with affinities for nucleic acids regulated by the nucleotide state. The core region carries all conserved signature motifs in a similar 3D arrangement in different helicases (1), pointing to similar contributions of the conserved motifs to DEAD box helicase activities. A conserved phenylalanine, the Q-motif, motifs I, Ia, Ib, II and III are located in the N-terminal RecA-like domain, and motifs IV, V and VI are in the C-terminal domain (1). In extensive mutational analyses of a number of DEAD box helicases, and in structural studies (2–6), ATP binding and hydrolysis were assigned to motifs I (Walker A motif, AxTGxGKT) and II (Walker B motif, DEAD), with contributions from motif VI (HRIGRTGR). The conserved glutamine in the Q-motif determines the specificity for adenine nucleotides (7). The motifs Ia (PTRELA) and Ib (TPGR), and IV, V (ARGID) and VI contribute to RNA binding (1,2,5,6). Most importantly, coupling of ATPase and unwinding activities has been ascribed to motif III [SAT (8,9)]. Altogether, these studies could confirm similar roles for the conserved signature motifs in many DEAD box helicases, but revealed complex contributions of various motifs to ATP binding and hydrolysis, to RNA binding and to unwinding. Structural studies demonstrated that extensive interactions between conserved motifs within the N- and C-terminal RecA-like domain most likely relay the nucleotide state to the RNA-binding site and thus contribute to RNA unwinding. The crystal structures of the DEAD box helicases Vasa [Drosophila (2)], eIF4A-III [human (3,4)] and Ddx19 ([human (5)] or Dbp5 [yeast (6)] in complex with ADPNP and a single-stranded RNA capture the closed conformation of DEAD box helicases and reveal an extensive interaction network between the conserved motifs from both RecA-like domains across the domain interface. ADPNP is bound in the cleft between the domains, and RNA is bound to a bipartite RNA-binding site formed by both RecA-like domains (Figure 1). The simultaneous assembly of the ATPase site and the RNA-binding site in the closed conformer rationalizes the RNA-stimulated ATPase activity and the cooperativity between ATP and RNA binding.

*To whom correspondence should be addressed. Tel: +41 61 267 2381; Fax: +41 61 267 2189; Email: Dagmar.Klostermeier@unibas.ch

Figure 1. DEAD box helicase motifs and position of mutations in motifs I, III and V. The homology model for the YxiN helicase core was constructed with Geno3D (23) using Vasa (PDB-ID 2db3) as a template. Mutations were introduced into motifs I (green, K52Q), motif III (orange, S182A/T184A) and motif V (blue, G303A). ADPNP (cyan) and RNA (yellow) are depicted in stick representation as they are bound to Vasa. The lower panel illustrates the context of the conserved helicase motifs. The positions for fluorophore attachment are indicated in red.

The single-stranded RNA bound to Vasa and Ddx19 is kinked, and it has been suggested that this kink destabilizes an RNA duplex as a first step towards unwinding (2,5).

YxiN is a DEAD box helicase from *Bacillus subtilis* that specifically binds to hairpin 92 in the 23S ribosomal RNA via its C-terminal domain. RNA binding stimulates the intrinsic ATPase activity by three orders of magnitude (10–12). YxiN is in an open conformation with a large inter-domain cleft in the absence of RNA or nucleotide (12,13). In response to cooperative binding of the non-hydrolyzable ATP-analog ADPNP and RNA, YxiN undergoes a conformational change in its helicase core, leading to a closed conformation similar to Vasa and eIF4A-III (12). To investigate the relation of this conformational change to RNA unwinding, we used YxiN variants carrying a mutation in the Walker A motif (motif I, YxiN_K52Q, Figure 1), a mutation in the SAT motif (motif III, S182A/T184A, YxiN_AAA), or a mutation in motif V (YxiN_G303A). The lysine in motif I contacts the β- and γ-phosphates of ATP, and mutations have been reported to diminish ATPase activity in various DEAD box helicases (8,14). The role of ATP binding for DEAD box helicases has previously been addressed using the non-hydrolyzable ATP analog ADPNP, but the concern has been raised that the ADPNP bound state is an off-pathway intermediate (15). The K52Q mutant is thus ideally suited to study the ATP state of DEAD box helicases in more detail without relying on non-hydrolyzable ATP analogs. The two other mutants are examples of so-called 'uncoupled' mutants: motif III has been implicated in coupling ATP hydrolysis to RNA unwinding. Mutations do not interfere with ATP hydrolysis or RNA binding, but abolish or reduce helicase activity (8,16). Motif III is the only conserved motif that does not interact with ADPNP or RNA in the Vasa structure (2). In contrast, motif V is the only motif that interacts with both ADPNP and RNA (2). Similar to mutations in motif III, mutation of the conserved glycine in motif V leads to a DEAD box protein that hydrolyzes ATP and binds RNA but does not show unwinding activity (17). Thus, these mutants should be ideally suited to study coupling effects. To dissect the relation between the conformational change in the helicase core and RNA unwinding, we compared the RNA-dependent ATPase activity, the propensity to undergo the conformational change, and RNA unwinding activity for the three YxiN constructs with mutations in motif I, III, or V that are impaired at different points of the catalytic cycle.

MATERIAL AND METHODS

Construction and purification of YxiN mutants

Site-directed mutagenesis was performed according to the Quikchange protocol (Stratagene). YxiN wild-type and mutants were purified as described (11,12).

Steady-state ATPase activity

Adenine nucleotides were purchased from Pharma Waldhof or JenaBioscience. Steady-state ATP hydrolysis was monitored in a coupled enzymatic assay at 37°C *via* the decrease in A_{340} due to oxidation of NADH to NAD^+ (18) as described (10–12). Assay conditions were 50 mM HEPES/KOH, pH 7.5, 175 mM KCl, 10 mM $MgCl_2$, 100 μM DTT, 200 μM NADH, 1 mM PEP, 13 μg ml^{-1} LDH, 23 μg ml^{-1} PK and 10 nM YxiN wild-type, YxiN′, 30 nM YxiN_AAA, YxiN′_AAA, 50 nM YxiN_G303A, YxiN′_G303A, 1 μM YxiN_K52Q or 200 nM YxiN′_K52Q. Initial reaction velocities were calculated from the absorbance change $\Delta A_{340}/\Delta t$ using the extinction coefficient $\varepsilon_{340,\ NADH} = 6300\,M^{-1}\,cm^{-1}$. Data from ATP-dependent experiments were analyzed using the standard Michaelis–Menten model. Data from RNA-dependent experiments were analyzed using the explicit binding equation

$$k_{\mathrm{obs}} = \frac{k_{\mathrm{cat}}}{[\mathrm{YxiN}]}\left[\frac{[\mathrm{RNA}]+[\mathrm{YxiN}]+K_{\mathrm{app,RNA}}}{2}\right.$$
$$\left.-\sqrt{\left(\frac{[\mathrm{RNA}]+[\mathrm{YxiN}]+K_{\mathrm{app,RNA}}}{2}\right)^2-[\mathrm{RNA}][\mathrm{YxiN}]}\right]$$

$$\mathbf{1}$$

k_{obs} is the observed rate constant at a certain RNA concentration, k_{cat} is the turnover number and $K_{\mathrm{app,RNA}}$ is the apparent dissociation constant for RNA. For all constructs except YxiN_AAA, $K_{\mathrm{app,RNA}}$ values from the standard Michaelis–Menten or the explicit binding model were in good agreement. For YxiN_AAA and YxiN′_AAA, data were not well-described by the explicit

binding model, and thus were analyzed using the standard Michaelis-Menten model (Figure 2a, Table 1). All experiments were repeated at least three times.

To determine coupling energies for RNA and ATP binding, $K_{M,app,ATP}$ values were determined in ATP-dependent experiments at four different RNA concentrations. $K_{M,app,ATP}$ values were plotted as a function of the RNA concentration, and described using

$$K_{M,app,ATP}$$
$$= K_{M,ATP} + \frac{\Delta K_{M,app,ATP}}{[YxiN]} \left[\frac{[RNA]+[YxiN]+K_{app,RNA}}{2} \right.$$
$$\left. - \sqrt{\left(\frac{[RNA]+[YxiN]+K_{app,RNA}}{2} \right)^2 - [RNA][YxiN]]} \right] \quad \mathbf{2}$$

$K_{app,RNA}$ is the apparent dissociation constant for RNA. $K_{app,RNA}$ values from the fit were in reasonable agreement with the values determined directly (Figure 2a). $K_{M,ATP}$ denotes the K_M value in the absence of RNA, and the $K_{M,app,ATP}$ at saturating RNA concentration is $K_{M,ATP} + \Delta K_{M,app,ATP}$. These K_M values were converted to ΔG values; their difference is the coupling energy.

RNA substrates

The 153mer RNA substrate comprising nucleotides 2483–2635 of the *B. subtilis* 23S rRNA was generated by T7 polymerase *in vitro* transcription as described (11).

The 32/9mer minimal RNA substrate comprising hairpin 92 of the 23S rRNA was constructed by annealing a synthetic 32mer and a synthetic 9mer as described (11,12). Unwinding assays were performed with $5\,\mu M$ RNA and $10\,\mu M$ YxiN in $50\,mM$ HEPES, pH 7.2, $150\,mM$ KCl, 5% glycerol, $0.1\,mM$ DTT, $0.1\,mg/ml$ BSA, $5\,mM$ MgCl$_2$ at $25°C$ and products were analyzed by native polyacrylamide gel electrophoresis as described (11,12). All experiments were reproduced at least twice.

Fluorescent labeling and smFRET experiments

Fluorescent labeling of cysteines was performed in $50\,mM$ Tris–HCl, pH 7.5, $500\,mM$ NaCl, $0.5\,mM$ TCEP at a protein concentration of $\sim 50\,\mu M$ with a 3-fold molar excess of Alexa488-maleimide (A488, donor) and a 4-fold molar excess of Alexa546-maleimide (A546, acceptor) for 1 h at $25°C$. The reaction was stopped by adding $1\,mM$ BME, and free dye was removed by size exclusion chromatography on Bio-Rad Micro Bio-Spin Columns. Labeling efficiencies were determined from absorbance ratios at 493 nm (A488, corrected for A546 contributions) or 554 nm (A546) and 280 nm (protein, corrected for dye contributions).

Single molecule FRET experiments were performed using a home-built confocal microscope as described (12). Only fluorescence bursts above a threshold of 100 photons were considered in the analysis. Measured background-corrected fluorescence intensities were corrected for crosstalk (α donor crosstalk in acceptor channel, β acceptor crosstalk in donor channel), different quantum yields and detection efficiencies of donor and acceptor fluorescence (γ), and direct excitation of the acceptor (δ), and

converted into FRET efficiencies as described (12). Measurements were performed at room temperature ($25°C$) in $50\,mM$ Tris–HCl, pH 7.5, $150\,mM$ NaCl, $5\,mM$ MgCl$_2$ with $40\,pM$ fluorescently labeled protein (concentration of donor fluorophore), $5\,mM$ nucleotide and $200\,nM$ 153mer RNA. For conversion of FRET efficiencies into distances, previously determined Förster distances were used (12).

RESULTS AND DISCUSSION

Motif mutations affect RNA stimulated ATPase activity and RNA unwinding differently

To test the effect of mutations in the conserved motifs on the RNA-stimulated ATP hydrolysis, the steady-state ATPase activity was determined for wild-type YxiN and the mutants in the presence of a 153mer derived from 23S ribosomal RNA (Figure 2a, Table 1). YxiN is a Michaelis–Menten enzyme with a k_{cat} of $2.4\,s^{-1}$ and the apparent $K_{app,RNA}$ value for this RNA of 7 nM. The $K_{M,app,ATP}$ value for ATP in the presence of RNA is 1.0 mM (Figure 2b, Table 2). ATP hydrolysis by the motif I mutant YxiN_K52Q was close to the background hydrolysis under the experimental conditions, confirming that it is ATPase-deficient. The motif III double mutant YxiN_AAA exhibits a reduced k_{cat} of $1.4\,s^{-1}$, demonstrating that the rate-limiting step in ATP hydrolysis is decelerated. For the *Escherichia coli* homolog of YxiN, DbpA, it has been shown previously that ATP hydrolysis and phosphate release are rate-limiting in the RNA-stimulated ATPase reaction (15). Consequently, the apparent $K_{app,RNA}$ for RNA reflects binding to the nucleotide-states populated under steady-state conditions, which for DbpA are the ATP- and ADP-P$_i$ states. The $K_{app,RNA}$ value of YxiN_AAA is slightly increased compared to wild-type (15 nM). Assuming similar steady-state populations for YxiN as for its *E. coli* homolog, this suggests that the ATP and/or ADP-P$_i$ state bind RNA with similar affinities compared to wild-type YxiN. The $K_{M,app,ATP}$ of YxiN_AAA for ATP is slightly reduced compared to wild-type to 0.6 mM, indicative of a minor increase in ATP affinity. The mutation G303A in motif V drastically reduced the ATPase rate to $k_{cat} = 0.55\,s^{-1}$ and the $K_{M,app,ATP}$ value is increased ~ 2-fold compared to wild-type YxiN. Thus, ATP binding and the ATP hydrolysis rate are affected when the conserved glycine is replaced by alanine. In contrast, the $K_{app,RNA}$ value (9 nM) is similar to wild-type YxiN.

A hallmark feature of DEAD box helicases is the cooperative binding of ATP and RNA substrate (19). Experimentally, cooperativity can be assessed by determining nucleotide affinities in the absence and presence of RNA substrate (12), or by determining RNA affinities in the absence and presence of nucleotide. However, RNA binding to YxiN is dominated by the high-affinity interaction of hairpin 92 in ribosomal RNA with the RNA-recognition motif in the C-terminal domain, and this interaction will mask nucleotide-dependent RNA affinity changes of the helicase core in titration experiments. In the reciprocal titration experiment of nucleotide
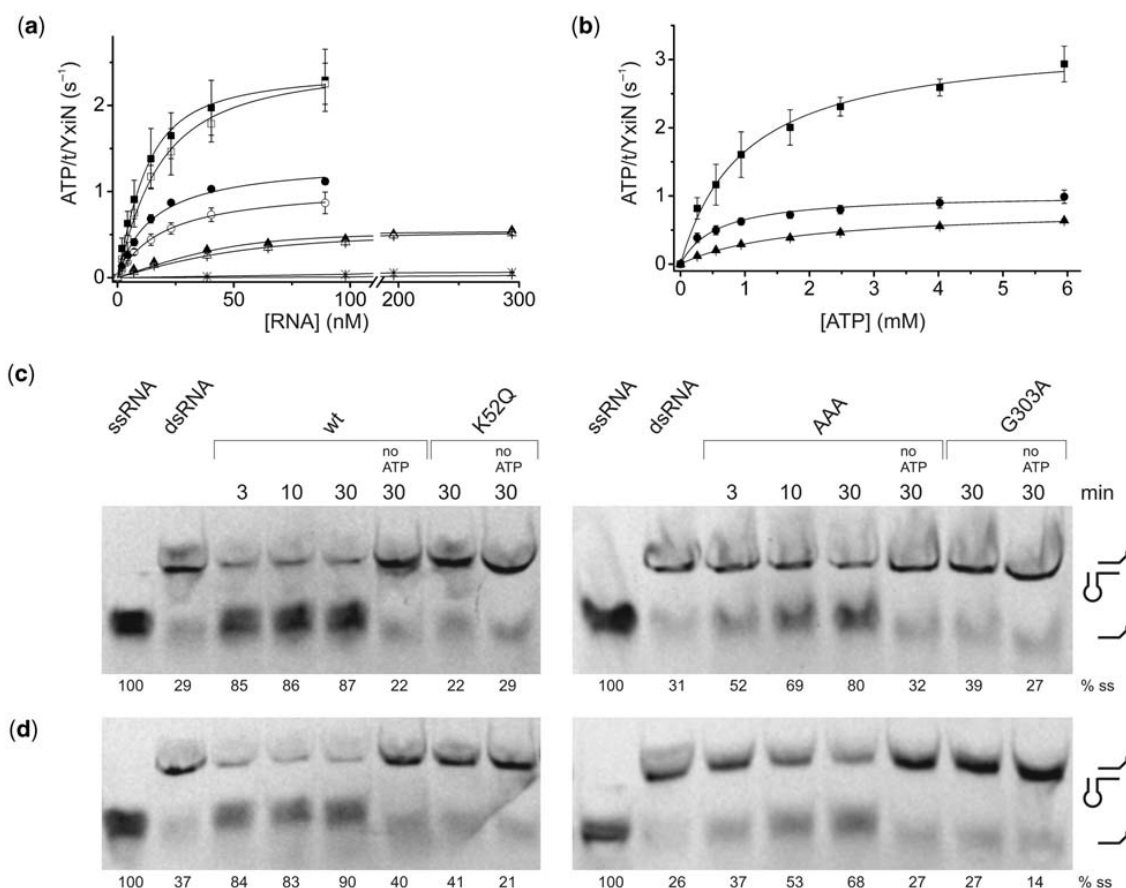
Figure 2. Steady-state ATPase activity and RNA unwinding. (**a**) ATP hydrolysis rates as a function of RNA concentration in the presence of 5 mM ATP. Wild-type YxiN (squares) is a Michaelis–Menten enzyme with a k_{cat} of 2.4 s^{-1} and a $K_{app,RNA}$ of 7 nM. YxiN′ (open squares) has wild-type like properties. YxiN_AAA (circles), YxiN′_AAA (open circles), YxiN_G303A (triangles), YxiN′_G303A (open triangles). YxiN_K52Q hydrolysis was within background (stars). Data for YxiN_AAA and YxiN′_AAA could not be described using the explicit binding model and were thus described using the standard Michaelis–Menten model (see Methods section). All experiments were performed at least in triplicate. $K_{app,RNA}$ and k_{cat} values are summarized in Table 1. (**b**) ATP hydrolysis rates as a function of ATP concentration for YxiN_wild-type (squares), YxiN_AAA (circles) and YxiN_G303A (triangles) in the presence of RNA (260 nM, 55 nM, and 603 nM, respectively). Experiments were performed in triplicate. Michaelis–Menten parameters are summarized in Table 2. (**c**) Unwinding of the 32/9mer minimal RNA substrate by mutants in wildtype background. YxiN_AAA unwinds RNA to the same extent as wild-type YxiN, but much more slowly. YxiN_K52Q and YxiN_G303A are unwinding-deficient. None of the proteins unwinds RNA in the absence of ATP. Numbers below the lanes indicate the fraction of single strand. (**d**) RNA unwinding by mutants in YxiN′ background. The effect on RNA unwinding is identical to the results in (c) where the mutations were introduced into wild-type YxiN. All depicted data are representative results and have been reproduced at least twice.

Table 1. Parameters for steady-state ATP hydrolysis by YxiN wild-type and mutants

	k_{cat} (s^{-1})	$K_{app,RNA}$ (nM)
YxiN_wild-type	2.4 (±0.3)	7 (±2)
YxiN_K52Q	n.d.	n.d.
YxiN_AAA	1.36 (±0.03)[a]	15 (±1)[a]
YxiN_G303A	0.55 (±0.03)	9 (±4)
YxiN′	2.6 (±0.1)	13 (±4)
YxiN′_K52Q	n.d.	n.d.
YxiN′_AAA	1.1 (±0.2)[a]	20 (±3)[a]
YxiN′_G303A	0.55 (±0.02)	18 (±4)

YxiN′ denotes the FRET construct YxiN_C61/267A_A115/S229C. Standard deviations were calculated from individual parameters of at least three independent experiments (Figure 2a).
[a]Parameters were obtained using the standard Michaelis–Menten model to describe the data (see 'Methods' section).

with YxiN, low nucleotide affinities require high protein concentrations exceeding 100 µM. Solubility problems of the mutants used here precluded such a determination of K_d values, but when the analysis was restricted to data for protein concentrations below 80 µM, relative affinities could be deduced from initial slopes of the titration curves of mantADPNP with YxiN (data not shown). These data point towards reduced ADPNP affinities of YxiN_G303A and YxiN_K52Q compared to wild-type. YxiN_AAA appears to bind ADPNP with similar affinity as wild-type, but the analysis had to be restricted to protein concentrations below 10 µM. In addition, we therefore determined apparent $K_{M,app,ATP}$ values for ATP at different RNA concentrations in steady-state ATPase assays as a measure for ATP affinity (Table 2, data not

Table 2. Michaelis–Menten parameters for steady-state ATP hydrolysis by YxiN wild-type and mutants: $K_{M,app,ATP}$ values for ATP in the presence of 153mer RNA (see Figure 2b)

	$K_{M,app,ATP}$ (mM)
YxiN_wild-type	1.0 (±0.4)
YxiN_AAA	0.6 (±0.1)
YxiN_G303A	1.7 (±0.2)

Standard deviations were calculated from individual $K_{M,app,ATP}$ values determined in at least three independent experiments (Figure 2b).

shown). For wild-type YxiN, the $K_{M,app,ATP}$ value for ATP is markedly reduced with increasing RNA concentrations as a result of thermodynamic coupling. From the extrapolated $K_{M,app,ATP}$ at RNA saturation (0.5 mM) and the extrapolated $K_{M,ATP}$ in the absence of RNA (9 mM, data not shown), a coupling energy of 7.5 kJ/mol (1.8 kcal/mol) can be calculated (see 'Methods' section), consistent with the coupling energy for the same RNA substrate in the *E. coli* homolog DbpA [1.3 kcal/mol, (19)]. YxiN_G303A and YxiN_AAA show less cooperativity between ATP and RNA binding. The coupling energies are reduced to ~3-4 kJ/mol (~0.7–1.0 kcal/mol), indicating a substantial loss in thermodynamic coupling in both mutants.

To address the effect of the mutations on coupling of the ATPase activity to RNA unwinding, helicase assays were performed. As unwinding of the helix adjacent to hairpin 92 is reversible in the 153mer RNA substrate and cannot be detected experimentally, a bipartite 32/9mer subfragment of the 153mer was used as the minimal YxiN substrate (Figure 2c). Due to the limited thermodynamic stability of the 9 bp double helix, these experiments have to be performed at 25°C. It should be noted that unwinding assays are performed under single turnover conditions with respect to RNA, but provide multiple turnover conditions with respect to ATP hydrolysis. As a consequence, it cannot be excluded that YxiN undergoes multiple ATPase cycles while it remains bound to the RNA *via* its C-terminal domain.

Unwinding of this 32/9mer by wild-type YxiN is complete within three minutes and requires ATP (11,12) (Figure 2c). YxiN_AAA unwinds the RNA substrate to a similar extent as the wild-type protein, but displaces the 9mer more slowly. This reduced unwinding velocity might be a consequence of the reduced k_{cat} for ATP hydrolysis, or of the less efficient coupling between RNA and ATP binding. In contrast, neither for YxiN_K52Q nor for YxiN_G303A was helicase activity observed, indicating that both mutants cannot complete the catalytic cycle. This is expected for YxiN_K52Q due to its deficiency in ATP hydrolysis. YxiN_G303A, however, still shows RNA-stimulated ATP hydrolysis, suggesting that the functional interaction between motifs from both RecA-like domains is maintained. The lack of unwinding activity indicates that the catalytic cycle is impaired at a different stage for YxiN_G303A.

Closure of the inter-domain cleft in the helicase core

We next probed whether the YxiN mutants were still able to undergo the conformational change in the helicase core upon binding of RNA and ATP in single molecule FRET (smFRET) experiments. For these measurements, YxiN mutants were used in which the two solvent-accessible cysteines C61 and C267 were replaced by alanines, and two cysteines were introduced at positions 115 and 229 in each RecA domain in the helicase core (YxiN_C61/267A_A115/S229C, abbreviated as YxiN') for fluorescent labeling. In a previous smFRET study we have used a different construct with the intrinsic cysteines replaced by serines (12), which displayed markedly reduced RNA-stimulated ATPase activities compared to wild-type. In contrast, YxiN' exhibits similar RNA-stimulated ATPase and RNA unwinding activities as wild-type YxiN, and thus constitutes an improved YxiN construct for smFRET experiments (Supplementary Figure 1). The mutations in the conserved motifs were therefore introduced into YxiN'. These mutants displayed similar k_{cat} and $K_{app,RNA}$ values in steady-state ATPase assays (Figure 2a, Table 1) and identical RNA unwinding activities as the corresponding motif mutants in wild-type background (Figure 2d).

SmFRET experiments were performed on freely diffusing donor/acceptor labeled YxiN' in a confocal microscope. FRET histograms (Figure 3) show a FRET efficiency of ~0.38 for all constructs in the absence of ligands, suggesting that the conformations of the mutants are similar. Upon addition of 153mer RNA substrate and the non-hydrolyzable ATP analog ADPNP, the FRET efficiency in the wild-type protein shifted to a value of ~0.7. Using the combined results from FRET experiments with five constructs carrying donor and acceptor fluorophores in different positions on each side of the inter-domain cleft, we have previously established that the increase in FRET efficiency between dyes at positions 115 and 229 reflects a closure of the inter-domain cleft in response to RNA and ATP binding (12). The FRET efficiency remained at 0.38, and the protein retained the open conformation when only RNA or nucleotide was present (data not shown), in agreement with our earlier observations (12). In the presence of ATP and RNA, two populations were observed in the FRET histogram (Figure 3), with FRET efficiencies of 0.38 (open conformation) and 0.70 (closed conformation). Under these conditions, YxiN can hydrolyze ATP and complete catalytic cycles while in the confocal volume, and is thus captured in different conformations at different points in its catalytic cycle. In contrast, it is trapped in the closed pre-hydrolysis conformation in the presence of the non-hydrolyzable ADPNP.

The motif I mutant YxiN_K52Q is in the open conformation in the absence of substrate as well as in the presence of ATP, ADPNP or RNA alone. In the presence of RNA and ADPNP, or RNA and ATP, the FRET efficiency increased to 0.70 (Figure 3), characteristic of the closed conformation. The FRET histograms for the ADPNP- and ATP-states of YxiN_K52Q are similar, and closely resemble the histogram for wild-type YxiN in the
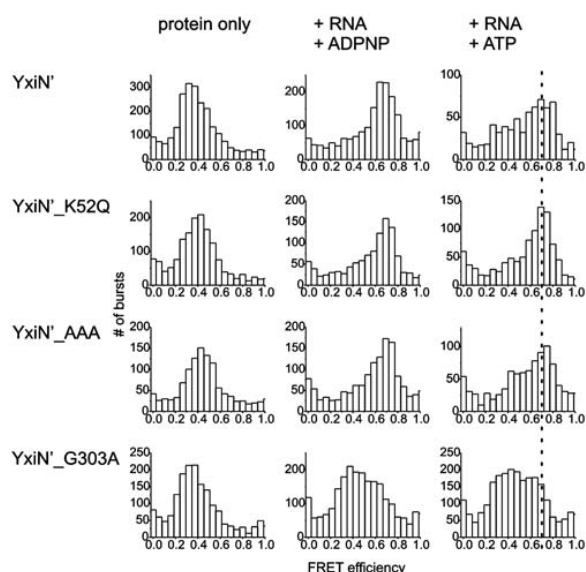
Figure 3. Conformational changes in the helicase core: smFRET experiments. FRET histograms of YxiN′, YxiN′_K52Q, YxiN′_AAA and YxiN′_G303A in the absence of ligands, and in the presence of 153mer RNA and ADPNP or ATP. The broken line indicates the FRET efficiency for the (completely) closed conformer of YxiN. YxiN′_K52Q and YxiN′_AAA have a similar propensity to form the closed conformer as wild-type YxiN, and the FRET efficiencies for the closed conformer are identical. YxiN′_G303A does not populate the closed conformer to the same extent, and the slightly reduced FRET efficiency of the closed conformer indicates that the cleft is wider compared to the closed conformation of wild-type YxiN.

RNA- and ADPNP-bound form. These data confirm that the global conformation of YxiN in the ADPNP-bound state is indeed similar to the ATP-state, in contrast to previous suggestions (15,20,21). Importantly, they provide strong evidence that ATP binding, not a subsequent step, populates the closed helicase conformer with high RNA affinity. Furthermore, the observation that YxiN_K52Q adopts a closed conformation in the RNA- and ATP-bound state, but is unwinding-deficient, demonstrates that the closure of the inter-domain cleft in the helicase core is not sufficient to promote RNA unwinding. This conclusion is consistent with the observation that ADPNP promotes closure of the inter-domain cleft in the helicase core, but does not support RNA unwinding in YxiN (12,22) and other DEAD-box helicases (21).

The motif III mutant YxiN_AAA showed similar FRET histograms to wild-type YxiN in the absence of ligands, and in the presence of RNA and ATP, or RNA and ADPNP (Figure 3), indicating that the SAT motif is not required for the conformational change to occur. Strikingly, the histograms for wild-type YxiN and YxiN_AAA in the presence of RNA and ATP are similar, indicating similar equilibrium populations of open and closed states despite the different rates of ATP hydrolysis. These similar populations suggest that opening and closing rates are equally affected in the YxiN_AAA mutant. While the global conformation of YxiN_AAA appears

to be similar to wild-type YxiN, the reduced k_{cat} points to local structural differences that lead to non-optimal alignment of residues involved in ATP hydrolysis, and the reduced thermodynamic coupling indicates differences in inter-domain communication.

Finally, the motif V mutant G303A showed a unique behavior in the presence of RNA and ADPNP, or RNA and ATP, with a very broad distribution of FRET efficiencies (Figure 3). The distribution can be described by two underlying populations, one with a mean FRET efficiency of around ∼0.37 for the open state, and a second state with a mean FRET efficiency of ∼0.65 resembling the closed conformation. The difference between FRET efficiencies of 0.65 for YxiN_G303A and 0.70 as observed for the closed conformations of wild-type YxiN, YxiN_K52Q and YxiN_AAA is significant and was reproduced in multiple independent experiments. The observed difference in the FRET efficiencies translates into a difference in donor/acceptor distance of 0.2 nm, and points towards a slightly wider cleft between the RecA domains in the closed conformation than in the other constructs. The incomplete closure of the cleft is not unexpected as the replacement of the highly conserved glycine by an alanine leads to steric hindrance at the inter-domain interface, as is evident from the Vasa, eIF4A-III and Ddx19 structures (2–6). Such an incomplete cleft-closure would lead to a misassembled catalytic site for ATP hydrolysis, consistent with the significantly reduced ATPase rate of YxiN_G303A. Similarly, the slightly increased $K_{app,RNA}$ value for RNA suggests that the two RNA-interaction sites on the RecA domains in YxiN_G303A may not be optimally aligned to form the bipartite RNA-binding site, and the partial loss in thermodynamic coupling of ATP and RNA binding indicates differences in inter-domain communication, again consistent with a different, more open conformation of the closed helicase core.

The model for RNA unwinding that has been proposed based on the Vasa structure predicts that uncoupled mutants still adopt a closed conformation of the helicase core, but do not close completely, and thus may not introduce a kink into the RNA substrate (2). Assuming that the formation of a completely closed conformer is linked to this RNA distortion, our results suggest that YxiN_K52Q and YxiN_AAA that form similar closed conformers as wild-type YxiN should kink the RNA (Figure 4), whereas YxiN_G303A does not. Similar to wild-type YxiN in the presence of ADPNP, YxiN_K52Q is unwinding-deficient because a critical step subsequent to ATP binding and formation of the closed conformer is blocked. YxiN_AAA unwinds RNA more slowly than wild-type because ATP hydrolysis (or product release) is decelerated. YxiN_G303A, in contrast, binds ATP and RNA, but does not adopt a completely closed conformation (Figure 4), and hence might fail to distort the bound RNA. Such a lack of initial distortion would thus prevent the efficient coupling of ATP hydrolysis to RNA unwinding. In this scenario, YxiN_G303A constitutes an uncoupled mutant as proposed by the Vasa unwinding model. Future experiments that directly address the conformation of RNA, such as the structure determination of
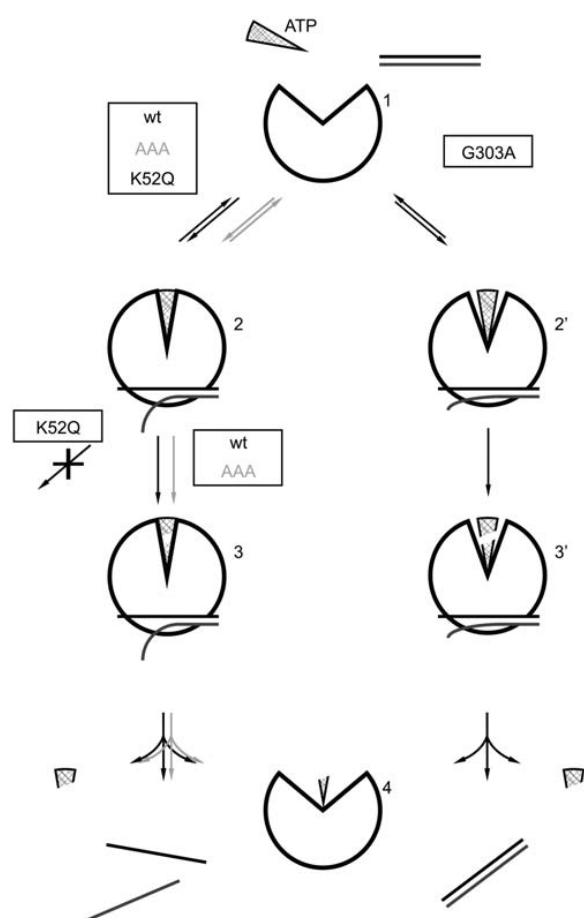
Figure 4. Effect of mutations on the RNA helicase catalytic cycle. Possible model for RNA unwinding consistent with the data presented in this work. The left branch represents a productive unwinding cycle (interrupted for YxiN_K52Q), the right branch depicts the uncoupled cycle for YxiN_G303A. YxiN is in an open conformation in the absence of ligands (1). Binding of ATP (triangle) and RNA (line-pair) induces a closure of the inter-domain cleft in the helicase core, which introduces a kink into the bound RNA and unwinds the terminal base-pairs (2). YxiN_G303A does not close completely, and possibly does not bend the bound RNA to the same extent (2'). ATP is hydrolyzed, and product release is coupled to dissociation of the unwound RNA (3,4). YxiN_K52Q cannot undergo hydrolysis, and the initial kink is not sufficient for RNA unwinding, rendering YxiN_K52Q unwinding-deficient. The RNA bound to YxiN_G303A dissociates in a double-stranded form upon ADP or phosphate release (3'–4'). The catalytic cycle for YxiN_AAA is similar to wild-type YxiN, but individual steps and thus the overall unwinding reaction are decelerated. Future experiments need to focus on the conformation of bound RNA, and on the identity of the second critical step for RNA unwinding after formation of the closed conformer.

double-stranded RNA bound to the closed helicase, could shed light on RNA distortions and their origin.

CONCLUSIONS

We have shown here that three mutants of the DEAD box helicase YxiN with mutations in motifs I, III and V that are unwinding-deficient or show reduced unwinding activity still adopt a closed conformation in the presence of RNA and ATP. The conformation of the ATPase-deficient YxiN_K52Q/ATP/RNA complex resembles the corresponding ADPNP state of wild-type YxiN, demonstrating that binding of nucleotide and RNA is sufficient to stabilize the closed helicase conformer. Furthermore, our previous assignment of the closed conformation as the pre-hydrolysis state is independently confirmed. Mutation of the SAT motif affects RNA unwinding by YxiN, but does not completely decouple ATP hydrolysis and RNA unwinding. YxiN_G303A adopts a less compact closed conformation, which is most likely caused by steric hindrance due to the alanine side-chain. Our results show that mutations can affect k_{cat}, either by removal of the catalytic amino acid (YxiN_K52Q), or by affecting the alignment of catalytic residues at the interface of the RecA domains (YxiN_AAA and YxiN_G303A). As the ATPase activity provides a timing function for the catalytic cycle, RNA unwinding is similarly affected. Additionally, mutations can increase $K_{app,RNA}$ and $K_{M,app,ATP}$, indicative of weaker interactions with RNA and ATP due to a misalignment of the bipartite RNA-binding site and the ATPase site, and decrease the thermodynamic coupling between interactions with ATP and RNA. Consequently, subtle changes in the inter-domain communication can affect RNA helicase activity in a complex manner. The two mutants studied here represent different degrees of misalignment and of decoupling: In YxiN_AAA, the global conformation of the helicase core and thus RNA binding are not affected, but local conformational differences lead to a decreased k_{cat}. Altogether, this mutant can still unwind RNA, but the reaction occurs more slowly. In YxiN_G303A, the global conformation is affected, leading to a decrease in ATPase activity and a severe drop in ATP and RNA affinities. The combination of these effects is detrimental for unwinding activity. It has previously been suggested that the introduction of a kink into the RNA upon formation of the closed conformer initiates unwinding (2), and it may be speculated that YxiN_K52Q and YxiN_AAA still kink their RNA substrate, whereas YxiN_G303A may not kink the RNA (or not to the same extent). Consistent with the observation that ADPNP does not support unwinding (12,22), we postulate that ATP hydrolysis or product release is coupled to a subsequent final step of unwinding. This critical step is absent in YxiN_K52Q, and significantly slowed down in YxiN_AAA. To identify the final step that leads to RNA unwinding, future studies will have to address the conformation of RNA bound to the helicase directly.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Ines Hertel for excellent technical assistance.

FUNDING

*Conflict of interest statement*. None declared.

REFERENCES

1. Cordin,O., Banroques,J., Tanner,N.K. and Linder,P. (2006) The DEAD-box protein family of RNA helicases. *Gene*, **367**, 17–37.
2. Sengoku,T., Nureki,O., Nakamura,A., Kobayashi,S. and Yokoyama,S. (2006) Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell*, **125**, 287–300.
3. Andersen,C.B., Ballut,L., Johansen,J.S., Chamieh,H., Nielsen,K.H., Oliveira,C.L., Pedersen,J.S., Seraphin,B., Le Hir,H. and Andersen,G.R. (2006) Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science*, **313**, 1968–1972.
4. Bono,F., Ebert,J., Lorentzen,E. and Conti,E. (2006) The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell*, **126**, 713–725.
5. Collins,R., Karlberg,T., Lehtio,L., Schutz,P., van den Berg,S., Dahlgren,L.G., Hammarstrom,M., Weigelt,J. and Schuler,H. (2009) The DExD/H-box RNA helicase DDX19 is regulated by an alpha-helical switch. *J. Biol. Chem*, **284**, 10296–10300.
6. von Moeller,H., Basquin,C. and Conti,E. (2009) The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat. Struct. Mol. Biol.*, **16**, 247–254.
7. Tanner,N.K., Cordin,O., Banroques,J., Doere,M. and Linder,P. (2003) The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell*, **11**, 127–138.
8. Pause,A. and Sonenberg,N. (1992) Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J.*, **11**, 2643–2654.
9. Pause,A., Methot,N. and Sonenberg,N. (1993) The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. *Mol. Cell Biol.*, **13**, 6789–6798.
10. Kossen,K. and Uhlenbeck,O.C. (1999) Cloning and biochemical characterization of *Bacillus subtilis* YxiN, a DEAD protein

11. Karow,A.R., Theissen,B. and Klostermeier,D. (2007) Authentic interdomain communication in an RNA helicase reconstituted by expressed protein ligation of two helicase domains. *FEBS J.*, **274**, 463–473.
12. Theissen,B., Karow,A.R., Kohler,J., Gubaev,A. and Klostermeier,D. (2008) Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. *Proc. Natl Acad. Sci. USA*, **105**, 548–553.
13. Wang,S., Overgaard,M.T., Hu,Y. and McKay,D.B. (2007) The *Bacillus subtilis* RNA Helicase YxiN is Distended in Solution. *Biophys. J.*, **94**, L01–L03.
14. Rozen,F., Pelletier,J., Trachsel,H. and Sonenberg,N. (1989) A lysine substitution in the ATP-binding site of eucaryotic initiation factor 4A abrogates nucleotide-binding activity. *Mol. Cell Biol.*, **9**, 4061–4063.
15. Henn,A., Cao,W., Hackney,D.D. and De La Cruz,E.M. (2008) The ATPase cycle mechanism of the DEAD-box rRNA helicase, DbpA. *J. Mol. Biol.*, **377**, 193–205.
16. Rocak,S., Emery,B., Tanner,N.K. and Linder,P. (2005) Characterization of the ATPase and unwinding activities of the yeast DEAD-box protein Has1p and the analysis of the roles of the conserved motifs. *Nucleic Acids Res.*, **33**, 999–1009.
17. Cheng,Z., Coller,J., Parker,R. and Song,H. (2005) Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA*, **11**, 1258–1270.
18. Adam,H. (1962) *Methoden der enzymatischen Analyse*. Bergmeyer, H.U.(Hrsg.), Verlag Chemie, Weinheim, pp. 573–577.
19. Polach,K.J. and Uhlenbeck,O.C. (2002) Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. *Biochemistry*, **41**, 3693–3702.
20. Chen,Y., Potratz,J.P., Tijerina,P., Del Campo,M., Lambowitz,A.M. and Russell,R. (2008) DEAD-box proteins can completely separate an RNA duplex using a single ATP. *Proc. Natl Acad. Sci. USA*, **105**, 20203–20208.
21. Liu,F., Putnam,A. and Jankowsky,E. (2008) ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc. Natl Acad. Sci. USA*, **51**, 20209–20214.
22. Kossen,K., Karginov,F.V. and Uhlenbeck,O.C. (2002) The carboxy-terminal domain of the DExDH protein YxiN is sufficient to confer specificity for 23S rRNA. *J. Mol. Biol.*, **324**, 625–636.
23. Combet,C., Jambon,M., Deleage,G. and Geourjon,C. (2002) Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics*, **18**, 213–214.

specifically activated by 23S rRNA: delineation of a novel subfamily of bacterial DEAD proteins. *Nucleic Acids Res.*, **27**, 3811–3820.

3.1.3 Helicase Activity of Fluorophore Labelled Motif Mutants

In the previous chapters smFRET experiments with the DEAD box protein YxiN were presented. From the observed conformational change depending on the ligation state of the enzyme mechanistic conclusions were drawn. For FRET experiments, donor and acceptor fluorophores need to be conjugated to cysteines in the protein. The labels may have an influence on protein stability and enzymatic activity. To be able to assess the influence of the fluorophores, the unwinding activity of labelled YxiN constructs was assayed. The corresponding unwinding gels are shown in figure 3.1.

The unwinding results presented in all other chapters of this work were obtained under quasi single-turnover conditions (protein-RNA ratio 2 to 1 in the standard protocol). The experiments presented in the current chapter, however, were performed using a protein-RNA ratio of about 2 to 5, leaving the single-turnover regime of the standard protocol. Because of the labelling procedure, the labelled proteins have a stock concentration of around 10 to 20 μM. Employing the standard protocol for the unwinding assay, the concentration is decreased to a final assay concentration of around 2 to 4 μM (at 5 μM dsRNA). To achieve single turnover conditions with the labelled proteins of low concentration, on the one hand the RNA concentration could be decreased. However, this is not beneficial for the quantification of the RNA. On the other hand, one could increase the volume of added protein stock solution. The addition of more than 2 μl protein to the reaction (10 μl total volume) was tested. When replacing the water, normally added to adjust the reaction volume, by the protein stock solution, markedly impaired unwinding activity was observed. This was true for the unlabelled and the labelled protein. Most probable this is due to the increased ionic strength in the sample caused by the addition of protein in storage buffer (containing 500 mM NaCl). Therefore it is not advantageous to add more than 2 μl protein stock solution to the reaction. The unwinding assay with labelled protein could, thus, only be performed employing a protein-RNA ratio of about 2 to 5.
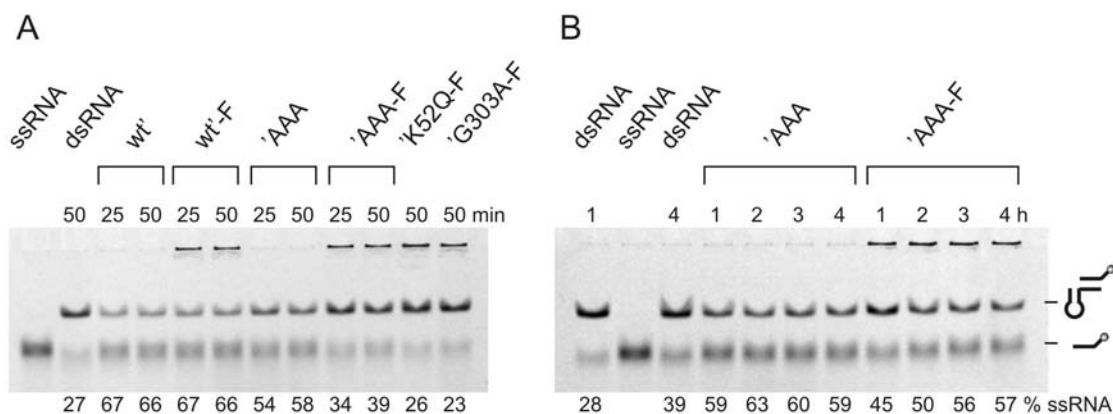
Figure 3.1
RNA unwinding activity of unlabelled and labelled YxiN constructs. The 32/9mer minimal RNA substrate was incubated with protein for the indicated time span, subsequently ds and ssRNA were separated on a native polyacrylamide-gel and the fluorescein labelled 9mer was detected. The activity of wild-type and motif mutants all in YxiN_C61A_C267A_A115C_S229C (= YxiN') background was assayed and equal concentrations of unlabelled and the corresponding labelled protein were used (2.4 μM YxiN`, 2 μM YxiN'_AAA, 3.3 μM YxiN'_K52Q, 1.8 μM YxiN'_G303A)

The labelled mutant YxiN_C61A_C267A_A115C_S229C (= YxiN') exhibits the same unwinding activity as the unlabelled mutant. As it was shown in chapter 3.1.2, mutants of helicase motif I (YxiN_K52Q) and V (YxiN_G303A) are unwinding deficient. The same behaviour is observed for the labelled YxiN' motif I and V mutants. The motif III mutant YxiN'_AAA yet shows a different behaviour. Within 50 min of incubation, the fluorophore labelled construct unwinds to a lesser extent than the unlabelled protein (fig. 3.1A). For this mutant also longer incubation times were tested (fig. 3.1B). Within 1 h of incubation approx. 60 % of single stranded RNA have been released by the unlabelled mutant. The same amount was unwound by the labelled one after 4 h. YxiN'_AAA indeed seems to be impaired by the labels regarding the velocity of the unwinding reaction. It still constitutes an active helicase and achieves the same endpoint of unwinding.

One may argue that unwinding is catalysed by the fraction of unlabelled species in the solution of labelled protein. However, the labelling degree of YxiN' was 87 % donor 112 % acceptor (relative to the protein concentration) here. Thus, 1 % single labelled protein and practically no unlabelled species were present in the reaction (in case of equal reactivity of the 2 labelled cysteines). Assuming only these being active in the reaction, a drastic decrease in velocity for labelled YxiN' would be expected but is not observed (fig. 3.1A).

The labelling degree of YxiN'_AAA used for the assay was 65 % donor, 92 % acceptor. In that case approx. 5 % unlabelled protein were present in the solution. About 34 % carried only one label. Since a massive decrease in unwinding velocity was observed, the reaction indeed could have been catalysed by the unlabelled and single labelled proteins. YxiN'_AAA might in fact be impaired when labelled with two fluorophores and FRET results from this mutant should therefore be interpreted with care.

3.1.4 SmFRET Experiments with a Q-motif Mutant of YxiN

In chapter 3.1.2 the biochemical characterization of YxiN constructs carrying mutations in the DEAD box motif I, III or V was presented. ATPase und unwinding activities were related to the results from smFRET studies probing the conformational change of the mutants. Combining all findings, mechanistic conclusions could be drawn. The current chapter summarises the smFRET study of YxiN carrying a mutation in the Q-motif.

The Q-motif is involved in nucleotide binding, in particular in the recognition of the adenine moiety (chapter 1.1.2). The motif was speculated to relay nucleotide binding to other motifs thereby regulating the ATP hydrolysis reaction (Cordin *et al.*, 2006). The lack of regulation in a Q-motif mutant might have an influence on the conformational changes of YxiN during catalysis. Therefore a Q29A mutation was introduced in the YxiN_C61A_C267A_A115C_S229C (=YxiN') background. The protein was fluorophore labelled and assayed in smFRET experiments. The corresponding histograms are presented in figure 3.2. YxiN'_Q29A is in the open conformation in the absence of ligands similar to YxiN'. No change is observed after addition of 200 nM 153mer RNA substrate or 5 mM ADPNP or 5 mM ATP. In the presence of both 200 nM RNA and 5 mM ADPNP the histogram becomes slightly broadened which could either be due to increasing noise or the appearance of a second population of higher FRET efficiency. The data however suffer from insufficient statistics caused by the increased noise from the substrate solutions. A measurement in the presence of 200 nM RNA and 5 mM ATP also showed a slightly broadened histogram. In an attempt to populate the presumably second conformation to a higher extent by adding 200 nM RNA and 6 mM ADPNP, no change was observed. Again, the histogram is slightly broadened in comparison to the one for protein only. The usage of an even higher substrate concentration was not experimentally accessible at that point because of the fluorescent impurities in the substrate solutions.

The question arises whether the Q-motif mutant is able to bind ATP (or ADPNP) and RNA under the experimental conditions. Since the mutation affects the nucleotide binding site, the affinity might be altered and the ATP (or ADPNP)+RNA ligated state could not be populated. To test the nucleotide affinity of YxiN'_Q29A a mant-nucleotide titration was performed. Mant-ADP as a probe was chosen since, in the absence of RNA, the wild-type has a higher affinity for it than for mant-ADPNP (chapter 3.1.1). When 1 μM nucleotide was titrated with YxiN'_Q29A no change in mant-fluorescence could be observed up to a protein concentration of 10 μM. The further increase of the protein concentration led to precipitation of the protein. In titrations with different YxiN constructs a signal change had been observed at protein concentrations below 10 μM before (Karow *et al.*, 2007; chapter 3.1.1). Thus, the Q29A mutation seemed to cause a drastic decrease in affinity for mant-ADP (estimated $K_D > 0.5$ mM) and presumably for other adenine nucleotides. The hydrogen bonds between the side chain of the native glutamine 29 and the adenine N7 and N6 appear to contribute substantially to the binding energy. Probably the lack of these hydrogen bonds also decreases the affinity of Q29A for adenine nucleotides in the presence of RNA. Therefore, in smFRET measurements the Q29A construct was possibly not saturated with substrate.

Regardless of the presence of substrates YxiN'_Q29A could only be captured in the open conformation in smFRET experiments. Either the closed conformation could not be populated due to the mutant's massively decreased affinity for adenine nucleotides or the mutation directly impairs the conformational change.
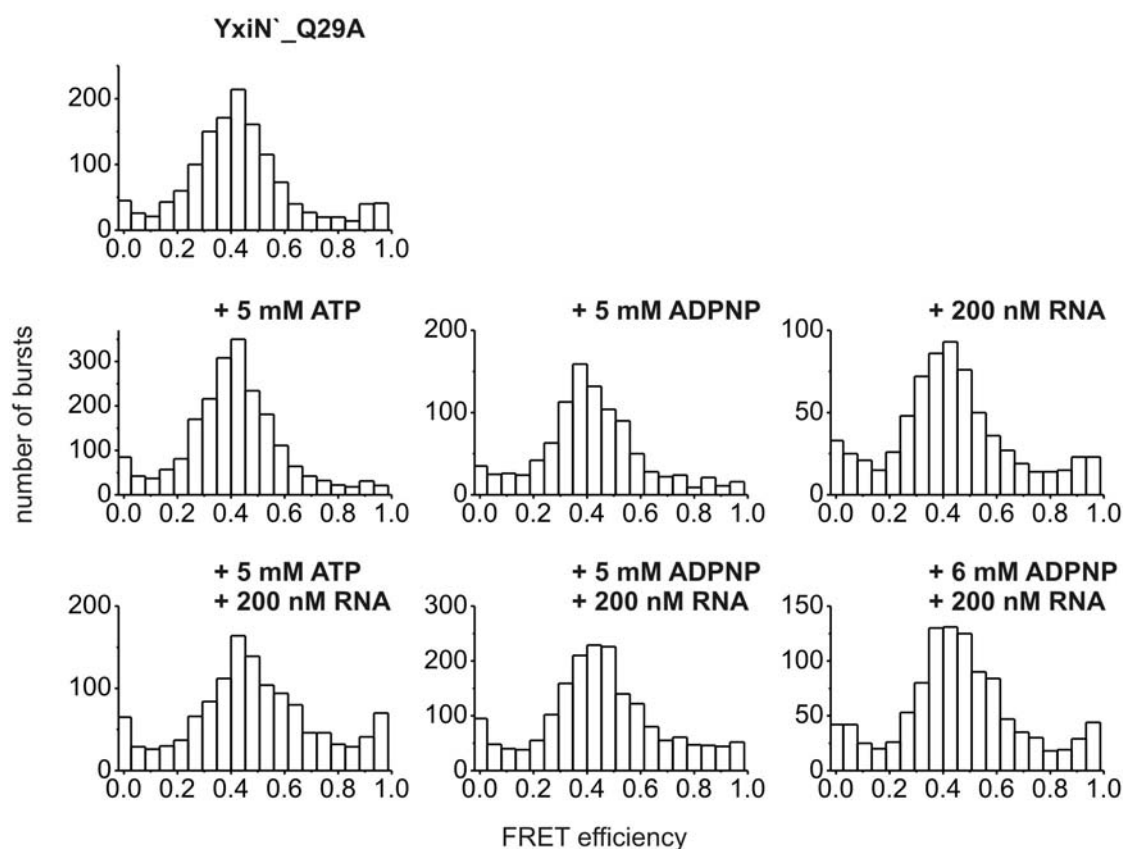
Figure 3.2
Histograms from smFRET experiments with YxiN'_Q29A. The protein only and the protein in the presence of substrates ATP, ADPNP, 153mer RNA as indicated were studied. Only bursts above a threshold of 100 photons were considered in the calculation of the histograms.

3.1.5 The Conformation of ADP/RNA Ligated YxiN

The conformation of YxiN is regulated by its ligation state. As it was shown in chapters 3.1.1 and 3.1.2, a transition from the open to the closed conformation occurs after binding of RNA and ATP (or ADPNP). In contrast, in the presence of ADP and RNA YxiN was found in the open conformation (construct YxiN_C61S_C267S_A115C_S229C) or the smFRET histogram showed a very broad distribution with no distinct maximum (construct YxiN_C61S_C267S_S108C_S229C). In chapter 3.1.1 it was therefore concluded that YxiN adopts the open conformation when RNA and ADP are bound.

The alanine variant YxiN_C61A_C267A_A115C_S229C (= YxiN') was introduced as an improved construct for smFRET experiments in chapter 3.1.2. Since it closely resembles the wild-type enzyme in terms of turnover number and RNA affinity, smFRET experiments in the presence of ADP+RNA were also performed using this construct. The corresponding histogram is shown in figure 3.3. Additionally, the smFRET histograms for the motif mutants in the presence of ADP+RNA are included in this figure. The histogram for YxiN' in the presence of ADP+RNA displays a bimodal distribution. The low-FRET state, characteristic for the open conformation, and the high-FRET state, characteristic for the closed conformation, are populated. The same result was obtained for the corresponding motif III mutant YxiN'_AAA. The population of two states was also found for the motif I mutant YxiN'_K52Q and the motif V mutant YxiN'_G303A. In comparison with YxiN' however, for YxiN'_K52Q a higher fraction of high-FRET events and for YxiN'_G303A of low-FRET events were observed in the presence of RNA+ADP.

Either YxiN' adopts the closed conformation upon binding of RNA+ADP or, since the population of the closed conformation had only been seen in the presence of RNA+ATP(ADPNP) before, the presence of RNA and traces of ATP or another impurity effected the closure of the inter domain cleft.

To localize a potential impurity, the experiment was repeated with different batches of ADP and 153mer RNA, and with freshly prepared buffer stock solutions. Cuvettes for smFRET experiments were saturated with BSA instead of YxiN1-368 that might have been copurified with bound nucleotides. In addition, the smFRET experiments were repeated with newly-labelled YxiN'. In all cases the smFRET histograms for YxiN' in the presence of ADP+RNA showed a bimodal distribution. Lowering the concentration of ADP, however, decreased the fraction of closed conformer (figure 3.3B). Altogether these observations indicate that ADP itself or an impurity in the ADP stock solution provoked the high-FRET state of the enzyme in the presence of RNA.

The ADP solution was thus analysed via RP-HPLC. The commercially available ADP always contained about 2 % of ATP. In a smFRET experiment, typically 5 mM of ADP were used corresponding to the presence of about 100 μM ATP. To check whether this "impurity" causes the population of the closed state, experiments in the presence of 50, 150, 500 μM and 3 mM ATP (+ 200 nM RNA) were performed (figure 3.4). At 50 and 150 μM ATP the high-FRET state was not populated to the same extent as in the presence of 5 mM ADP (with approx. 100 μM ATP). Only at 500 μM ATP the relative portion of low to high-FRET events was similar to the experiment at 5 mM ADP. For YxiN'_K52Q an even higher ATP concentration was necessary to populate the closed conformation to the same extent. A similar trend was observed for YxiN'_AAA and YxiN'_G303A.
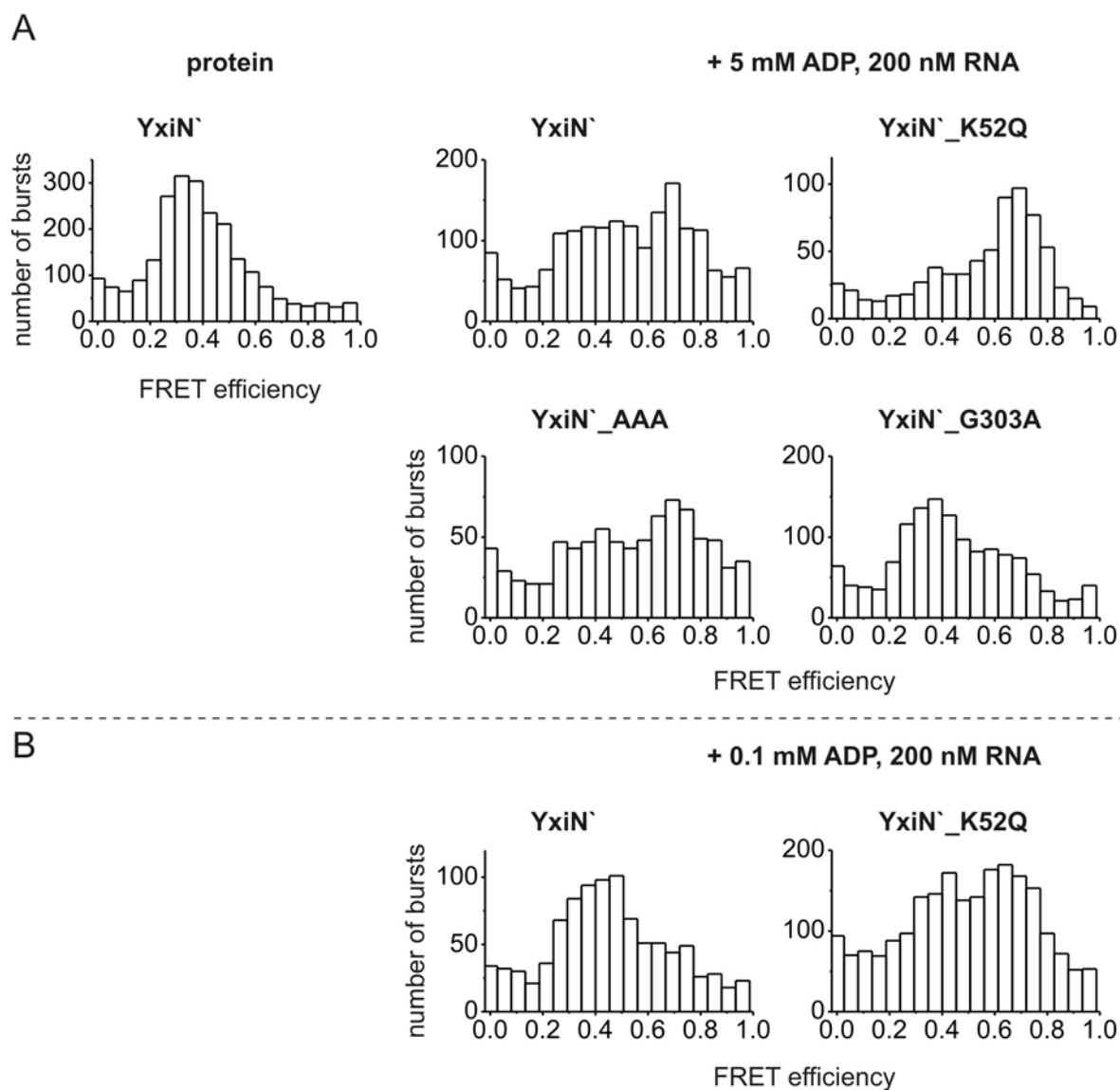
Figure 3.3
Histograms from smFRET measurements in the presence of 200 nM 153mer RNA and 5 mM ADP. In the absence of substrates YxiN' adopts the open conformation. In the presence of both RNA and ADP YxiN' and corresponding motif mutants also populate the closed conformation. Only bursts above a threshold of 100 photons were considered in the calculation of the histograms.

Since a certain ATP concentration was not sufficient to populate the closed conformation to the same extent as ADP containing the equal amount of ATP, the fraction of ATP in ADP might change under the experimental conditions. Possibly YxiN' or a copurified contaminant increased the concentration of ATP in the assay solution via a disproportionation reaction (2 ADP to ATP + AMP). To test for such a reaction, the amount of ATP under smFRET assay-like conditions (5 mM ADP, 60 pM unlabelled YxiN') and after incubation for 25 min was determined via RP-HPLC. It turned out that the fraction of ATP does not increase during the experiment.
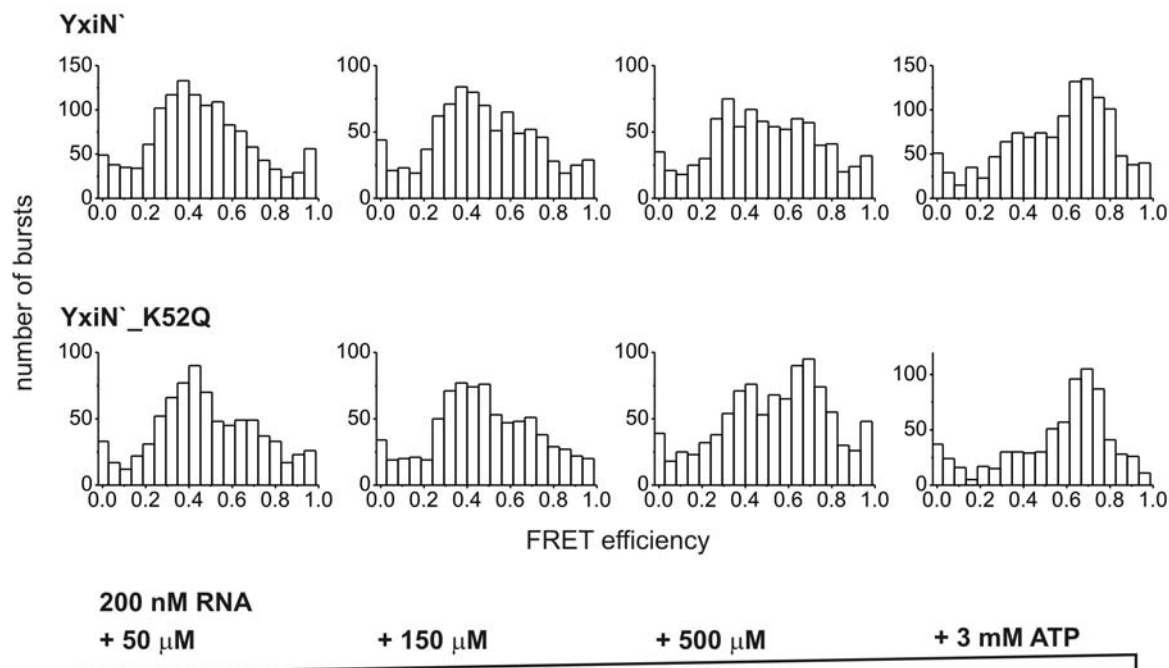
Figure 3.4
SmFRET histograms of YxiN' and YxiN'_K52Q in the presence of 200 nM 153mer RNA and varying concentration of ATP. Only bursts above a threshold of 100 photons were considered in the calculation of the histograms.

As a next step, the ATP in ADP ought to be removed or reduced. Hexokinase catalyses the phosphorylation of glucose consuming ATP (Kunst *et al.*, 1984). Therefore, hexokinase and glucose in combination can be used for *in-situ* depletion of ATP in ADP solutions.

To minimise the amount of ATP in smFRET experiments with RNA+ADP, 5 mM ADP in assay buffer were incubated with glucose/hexokinase for up to 2 h at 25 °C or 37 °C. Subsequently RNA and labelled YxiN' were added and the smFRET measurement was performed. To check whether the hexokinase/glucose system was effectively depleting ATP under the experimental conditions, the amount of ATP was determined via RP-HPLC (all analyses in triplicate). When 5 mM ADP was incubated in smFRET assay buffer for 2 h at 25 °C, about 1 % of ATP was detected. After incubation in the presence of hexokinase/glucose, ATP could not be detected by the UV absorption of the HPLC system. The hexokinase/glucose system indeed reduced the ATP content considerably.

Histograms from measurements in the presence of RNA+ADP (hexokinase/glucose treated) are shown in figure 3.5. Different incubation protocols were tested and in all cases the histograms show a bimodal distribution. In an attempt to quantify the fraction of closed to open population, the histograms were fitted with two Gaussian distributions. The fitting procedure might give an estimate of the ratio between low- and high-FRET species. According to this quantification and to an inspection of the unfitted histograms, the fraction of closed conformer was not decreasing in the presence of hexokinase/glucose. Since the RP-HPLC analysis revealed the ATP concentration in ADP is indeed minimised by hexokinase/glucose, the smFRET results strongly indicate that YxiN populates the closed conformation in the presence of ADP+RNA. The effect of ADP+RNA might have been missed before, because the measurements in chapter 3.1.1 were performed on YxiN constructs displaying a lower substrate affinity than YxiN'.
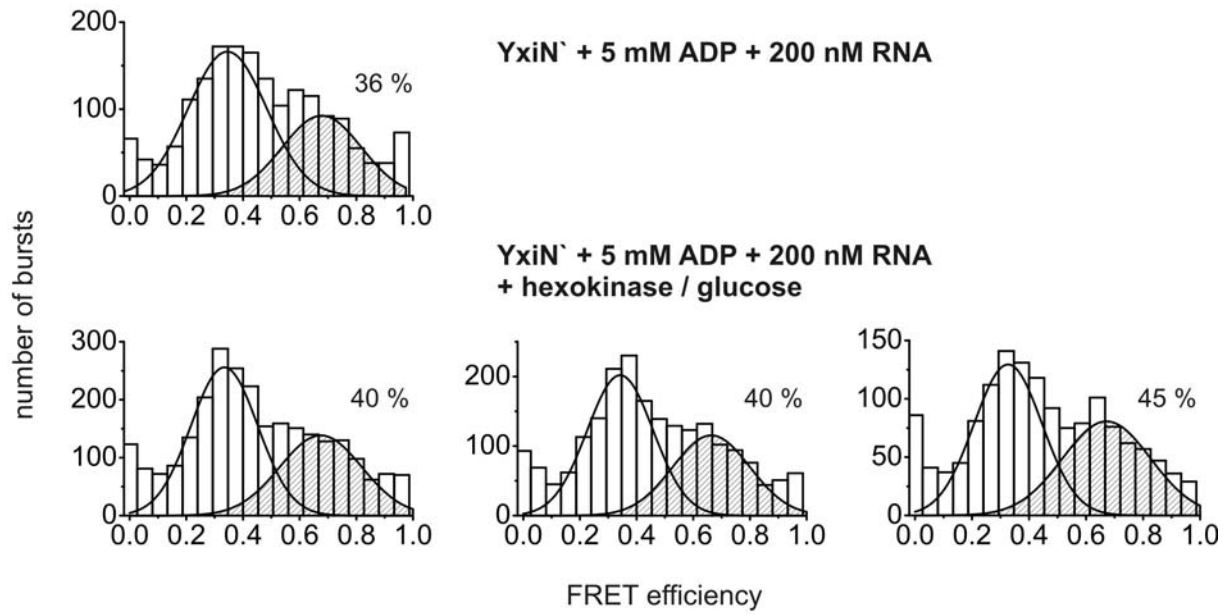
Figure 3.5
Histograms from smFRET measurements of YxiN' in the presence of 200 nM 153mer RNA and 5 mM ADP (upper row: untreated, lower row hexokinase/glucose treated).

Lower row from left to right: incubation with 4.7 U/ml hexokinase, 25 mM glucose, 1 h, 25 °C; 6.5 U/ml hexokinase, 37 mM glucose, 1 h, 37 °C; 6.5 U/ml hexokinase, 37 mM glucose, 2 h, 25 °C

Only bursts above a threshold of 100 photons were considered in the calculation of the histograms. The distributions were fitted with 2 Gaussians fixing the offset to 0. All other parameters were not restrained.
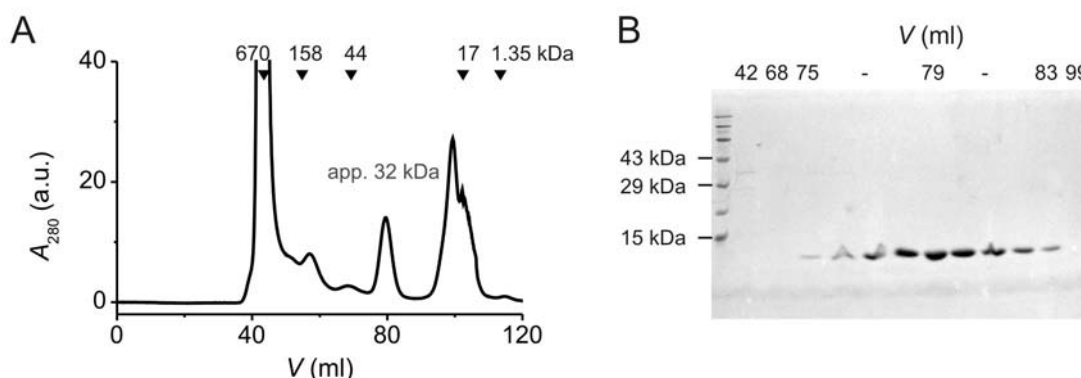
3.2 The RNA Binding Domain of YxiN

3.2.1 Purification of the RNA Binding Domain

The DEAD box protein YxiN consists of the typical helicase core domains and an additional C-terminal domain that was shown to provide specificity towards the RNA substrate of YxiN (Kossen *et al.*, 2002). To study this so called RNA binding domain (RBD, ranging from amino acid 404 to 479), the region coding for the corresponding gene was cloned into pETM30. Due to the cloning procedure (the 5' *Nco*I restriction site) the second amino acid of the RBD (lysine 405) was replaced by a glycine. Consequently, a positive charge on the protein would be removed. But positively charged amino acids play a role for RNA binding and a change in surface charge might influence the binding properties. Therefore, the native lysine 405 was restored after the cloning via site-directed mutagenesis.

The gene was expressed in *E. coli*. The corresponding GST-RBD fusion protein could be purified using the protocol for full-length YxiN with some modifications. When immobilized on the glutathione-affinity column, the column was washed with buffer containing 1 M NaCl to remove bound nucleic acids. TEV protease did not efficiently cleave the fusion protein. Therefore the cleavage reaction was allowed to proceed for 1 day at room temperature. The RBD did not precipitate in the presence of ammonium sulphate (80 % saturation) and therefore this purification step was omitted.

Notably, when purified on a size exclusion column in the last purification step, the RBD (8.4 kDa monomeric mass) did not elute as an apparent monomer. The corresponding chromatogram and a SDS-gel analysis of selected fractions are shown in figure 3.6. Using elution volumes of reference proteins, the column was calibrated and an apparent mass of 32 kDa could be calculated for the RBD. Thus the protein exhibited an apparent mass of a RBD tetramer under the experimental conditions of high ionic strength (500 mM NaCl). But the SEC data solely do not allow for conclusions on the oligomeric state of the RBD since the SEC elution volume is sensitive towards the hydrodynamic volume which is also determined by the shape of the species.

Full-length YxiN, however, is a monomer as judged by SEC and an extensive study on the *E. coli* homologue DbpA also revealed a monomeric state of the enzyme (Talavera *et al.*, 2006). Therefore it is doubtful that the potential RBD tetramer is of biological relevance.

Figure 3.6
The RNA binding domain (RBD, YxiN404-479_405K) elutes as an apparent tetramer on a SEC-column.
A) Elution profile of the SEC during purification of the RBD. The triangles mark the elution volumes of reference proteins of the indicated mass. B) Fractions were analysed on a denaturing gel and the RBD (8.4 kDa) could be identified eluting from 75 to 83 ml. According to the calibration of the SEC-column, an elution volume of 79 ml corresponds to an apparent mass of 32 kDa.

3.2.2 The Action of the RNA Binding Domain *in trans*

The YxiN helicase core binds RNA unspecifically and with low affinity (Karginov *et al.*, 2005). RNA specificity and binding with a K_D in the nanomolar range is provided by the YxiN RBD. In an ATPase assay in the presence of the RNA 32mer, it had been shown before that the RBD cannot fulfil its function *in trans* (Karginov *et al.*, 2005). In this study an ATPase assay was performed with a mixture of the isolated core and the RBD of YxiN (protein concentration 1 μM each). The ATPase activity was not stimulated by the 32mer RNA as it was with the full-length enzyme.

In the same study crosslinking experiments revealed that YxiN contacts the 153mer RNA substrate at various positions and not only at the region mimicked by the 32mer (Karginov *et al.*, 2005). Thus, it might be possible to facilitate an interaction of the RBD and the catalytic core by the presence of the 153mer substrate. The 153mer RNA might be able to bridge helicase core and RBD and could thereby provide a linkage between the protein fragments.

Therefore, in the current study the ATPase stimulation of the core by the RBD *in trans* was assayed in the presence of the 153mer RNA substrate. If the RBD was able to provide high affinity RNA binding to the core *in trans*, the ATPase activity should be stimulated by RNA concentrations similar to those stimulating the full-length enzyme. The helicase core (YxiN1-368) only, the RBD (YxiN404-479_405K) only and a 1:1 mixture of both were added to the assay solution and the decrease of A_{340} (reflecting the decrease in ATP concentration, methods section 2.12) over time was measured. The experiment was performed in the absence and presence of RNA and at protein concentrations of 10 nM, 200 nM and 1014 nM. In every set of measurements a sample containing no protein and no RNA was included. The ΔA_{340} of this reference sample was subtracted from the ΔA_{340} of the other samples.

The data from two independent measurements at 1014 nM helicase core, RBD or both are plotted in figure 3.7. When the helicase core, the RBD or the mixture of both were assayed, the ΔA_{340} fluctuated around the background value and ATPase activity was not detectable. Moreover, no stimulation was observed in the presence of 196 nM 153mer RNA. The RNA concentration was chosen such that it saturates the RBD (K_D of approx. 5 nM for full-length YxiN or RBD, Karginov *et al.*, 2005). The result of the measurements at 10 and 200 nM protein concentration was similar and the data are therefore not included in figure 3.7.

Possibly, a potential action of the RBD *in trans* could not be detected at the used protein concentrations. But higher concentrations were not experimentally accessible at that point. If a higher volume of protein stock solution (in 500 mM NaCl) was used, the increased ionic strength in the assay would influence substrate binding. Figure 3.7 additionally includes data from an ATPase measurement with full-length YxiN (10 nM protein concentration). In the absence of RNA the ΔA_{340} was close to the background value. But in the presence of 196 nM 153mer RNA the signal increased approximately 40fold. The ATPase activity of the full-length enzyme was significantly stimulated by the RNA substrate.

Taken together, the ATPase activity of YxiN is stimulated by its RNA substrate only when the RBD is covalently bound to the helicase core. The addition of the isolated RBD to the helicase core in the presence or absence of the 153mer RNA does not influence the ATPase activity.
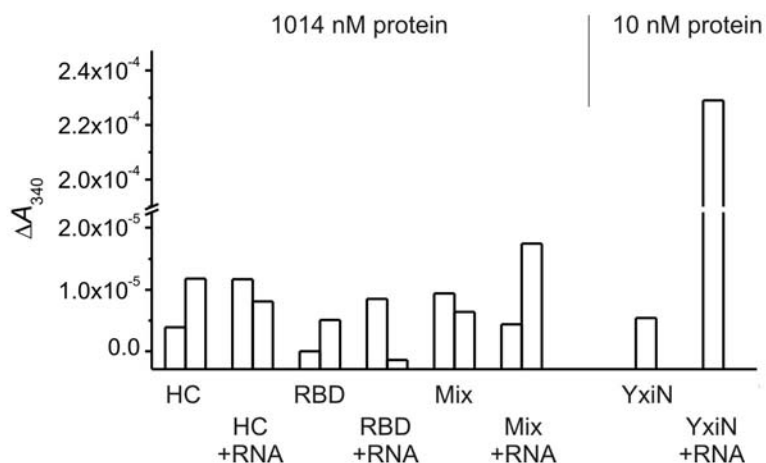
Figure 3.7
Steady-state ATPase activity of the YxiN helicase core (HC), the YxiN RBD or the HC/RBD mix in comparison with full-length YxiN. The measurements were performed in the absence and presence of 196 nM 153mer RNA. The plotted decrease of A_{340} over time reflects the consumption of ATP. The observed slope was always corrected by the value of a sample without RNA and protein. The measurements were performed at 1014 nM HC, 1014 nM RBD, 1014 nM HC+1014 nM RBD (all in duplicate) or 10 nM full-length YxiN. The ATPase activity of YxiN is markedly stimulated by the presence of the RNA substrate. No stimulation is observed for the HC and could also not be facilitated by the RBD *in trans*.

3.2.3 Orientation of the RNA Binding Domain to the Helicase Core

All DEAD box proteins show a high conservation of the helicase core region with respect to amino acid sequence and three-dimensional structure (chapter 1.1.2). Consistently, the crystal structure of the second domain of the YxiN core (Caruthers *et al.*, 2006) exhibits RecA like topology similar to other DEAD box proteins. The structure of full-length YxiN is not known but the core region can be approximated by the MjDeaD structure (chapter 3.1.1). Furthermore, the isolated RBD was crystallised (Wang *et al.*, 2006). From the crystal structure and biochemical data a possible mode of RNA binding by the RBD was suggested. The data however, did not allow for conclusions on how the RNA is directed to the core, whether the RBD might assist the unwinding reaction and how the RBD acts during substrate release.

To be able to answer these questions and gain a comprehensive understanding of the mechanism of YxiN it is required to know the position of the RBD in context of the full-length protein. This chapter presents a strategy for mapping the orientation of the RBD relative to the helicase core using various distance information that were obtained in smFRET experiments.

Selecting Positions for Fluorophore Labelling of the RBD

FRET experiments require fluorophore attachment to the positions of interest. For reaction with maleimide coupled fluorophores, cysteines have to be introduced at these positions. Corresponding mutations were introduced in YxiN_C61A_C267A (=YxiN#), a construct without solvent accessible cysteines having wild-type like activity (chapter 3.1.2). Cysteines were introduced on position D429, T440, N444, N452, N464 or L472. The set of positions covers the RBD quite uniformly and the positions exhibit no marked conservation in an alignment of homologous proteins (Karginov *et al.*, 2005).

Unwinding Activity of RBD Cysteine Mutants

Although non-conserved residues were mutated, the replacement by a cysteine might influence the folding of the RBD. All cysteine single mutants could be purified using the standard protocol and constituted soluble proteins. The ability to bind RNA substrate was indirectly tested via performing a RNA unwinding assay with the

mutants. The corresponding unwinding gels are shown in figure 3.8. In the presence of ATP all constructs unwound the 32/9mer to the same extent as the wild-type enzyme. The unwinding reaction was complete within 3 min of incubation. In the control samples lacking ATP no unwinding was observed. All constructs are thus able to bind the minimal RNA substrate and to subsequently unwind the double stranded region. This result indirectly provides evidence for the wild-type like folding of the mutants which is a prerequisite for assigning the upcoming distance information to the wild-type enzyme.
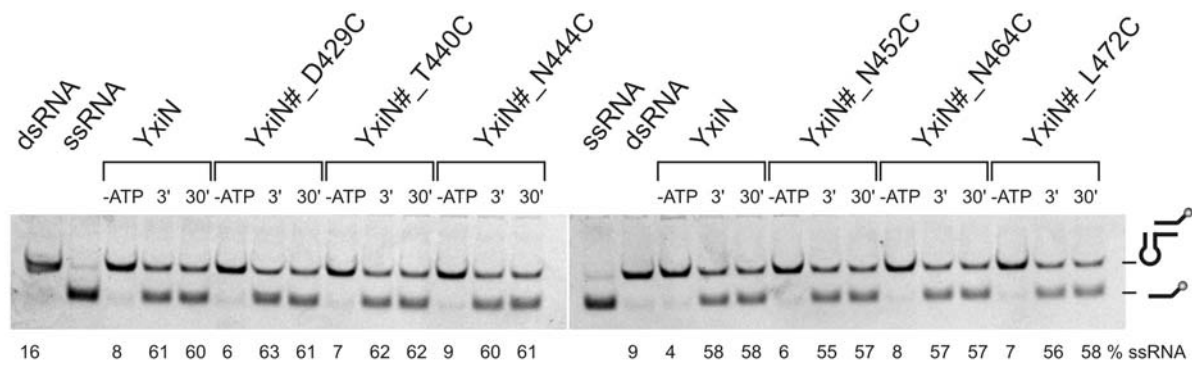
Figure 3.8
RNA unwinding activity of RBD cysteine mutants. The 32/9mer minimal RNA substrate was incubated with protein in the presence of ATP for the indicated time span or in the absence of ATP for 30 min. The activity of wild-type YxiN and RBD mutants in YxiN_C61A_C267A (=YxiN#) background was assayed.

SmFRET Experiments on Double Cysteine Mutants

In a next step an additional cysteine was introduced. Since the structure of the C-terminal helicase domain is known, a position within this domain was chosen. The mutation D262C was introduced because a cysteine on this position can be labelled with fluorophores and the mutation did not inactivate YxiN (chapter 3.1.1). The corresponding double cysteine mutants were purified using the standard protocol. The proteins were soluble and could be purified in high yield, except for YxiN#_D262C_N452C which was only obtained in low amounts.

The proteins were labelled with donor and acceptor fluorophores for smFRET measurements. Again, YxiN#_D262C_N452C exhibited a special behaviour. The mutant was prone to aggregation and could be labelled with low efficiency only (25 % A488, 39 % A546 in contrast to labelling efficiencies of >50 % A488, >50 % A546 for the other mutants). SmFRET experiments were performed at least in duplicate. The corresponding histograms (figure 3.9) were at that point calculated using FRET correction parameters (section 2.15.4) for non-conjugated A488/A546. All histograms peak at a FRET efficiency of around 0.85 to 0.9. No significant difference in FRET efficiency for the different mutants could be observed. In general, FRET efficiencies close to 0 and 1 are outside the dynamic range of FRET. Ideally, the FRET efficiencies should adopt values close to 0.5 allowing for distance determination within the most sensitive range (section 2.15.1). Therefore, in the performed experiments the distances were probed with suboptimal precision.

In an attempt to shift the dynamic range via influencing the Förster distance of the conjugated fluorophore pair, smFRET measurements in the presence of KI were performed. KI acts as a collisional quencher and reduces the quantum yield of fluorophores. According to equation 2.11, a decreased quantum yield leads to a lower Förster distance. Applied to the FRET measurements between D262C and positions within RBD, a lower FRET efficiency than 0.9 would be expected. Indeed, when 50 or 100 mM KI were added to the assay buffer, the mean FRET efficiencies decreased to around 0.75 or 0.68 respectively (exemplary values from measurements with

YxiN#_D262C_N464C). Unfortunately the presence of the quencher also influenced the overall photon counts. The background increased severely and the statistics of the measurements worsened. Probably KI quenches both, the donor and the acceptor fluorophores causing a drop in the overall signal. Additionally, the substance seemed to contain fluorescent impurities that could not be removed by the usual active charcoal treatment. Due to these unfavourable side effects measurements in the presence of quencher were not pursued further.
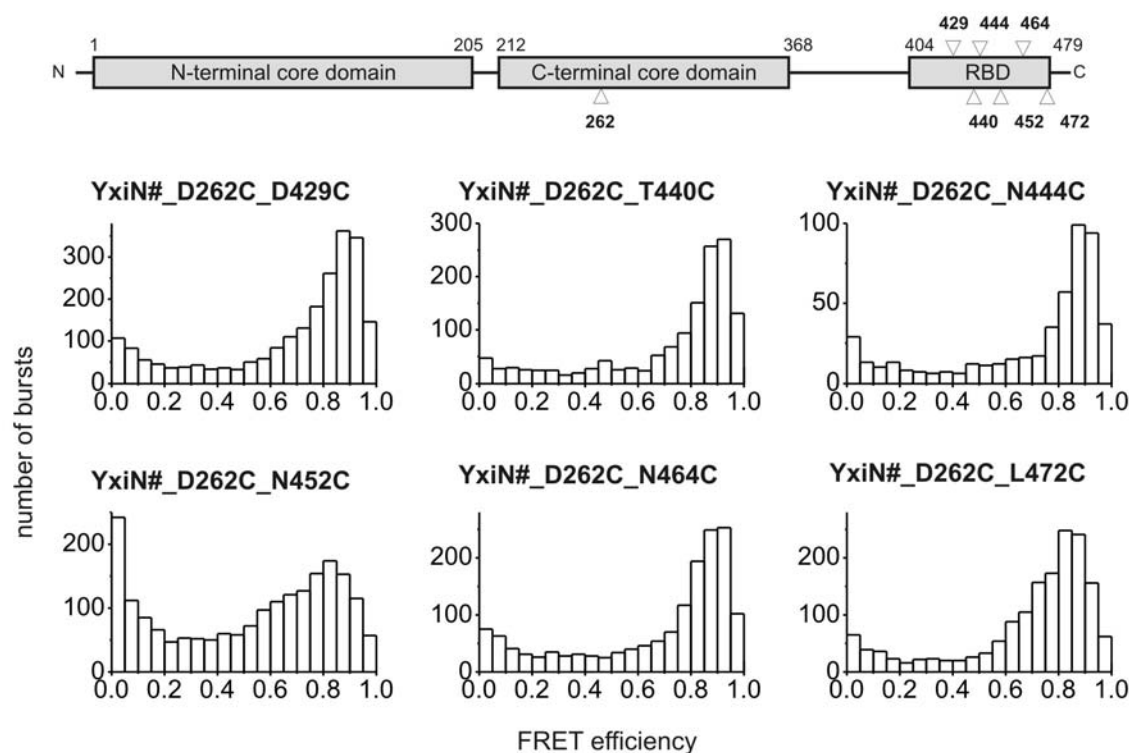
Figure 3.9
SmFRET histograms for YxiN constructs having fluorophores attached to the C-terminal core domain and the RBD. The scheme at the top assigns the labelled residues to the domains of YxiN. YxiN# denotes the YxiN_C61A_C267A background of the double cysteine mutants. The histograms were calculated using correction parameters of non-conjugated A488/546 and considering only bursts above a threshold of 100 photons.

Instead, positions other than 262 were chosen for fluorophore attachment. These positions should have a larger distance to the RBD than 262 leading to FRET efficiencies within the dynamic range of FRET measurements. Assuming the YxiN RBD might have a similar position as the RBD of *Thermus thermophilus* DEAD box helicase Hera (Rudolph & Klostermeier, 2009) the C-terminal core domain of YxiN was superimposed with the structure of Hera_208-510. According to this alignment the distance between position 262 and the RBD is approximately 3 to 4 nm. A larger distance can, among others, be found between position 275 and the RBD (5 to 6.5 nm) as well as between position 290 and the RBD (4.6 to 6.5 nm). Therefore, the constructs YxiN#_I275C_D429C, ~N464C, YxiN#_E290C_D429C and ~N464C were generated. The proteins could be purified with high yield, were fluorophore labelled and smFRET experiments were performed. A mean FRET efficiency of 0.85 to 0.95 was determined for these constructs (table 3.1 summarizes all results). Again, the probed distances were outside the dynamic range and a calculation of distances from the FRET events would be error-prone. Nevertheless, the results indicate that the RBD of YxiN occupies a position different from the Hera RBD. The measured FRET efficiencies of around 0.9 translate to a distance of around 4 nm (in an approximation

using the Förster distance of the non-conjugated fluorophores). In contrast, the distances between Hera RBD and the corresponding positions are 5 to 6 nm, as mentioned above.

In a further attempt to increase the distance between donor and acceptor fluorophores, a position within the N-terminal core domain was chosen. Although the structure of this domain has not been determined, it most probably resembles the fold known from other DEAD box proteins. All determined structures possess a RecA fold and largely superimpose (chapter 1.1.2). In addition, the arrangement of the N-terminal and the C-terminal core domains can be approximated using the MjDeaD structure (chapter 3.1.1). The mutation S108C was introduced because a cysteine on this position can be labelled with fluorophores and the mutation did not inactivate YxiN (chapter 3.1.1). The constructs YxiN#_S108C_D429C, ~N444C, ~N464C, and ~L472C constituted soluble proteins and could be purified with high yield. In contrast, YxiN#_S108C_T440C and ~N452C appeared in the insoluble fraction of the bacterial crude extract and could not be purified quantitatively.
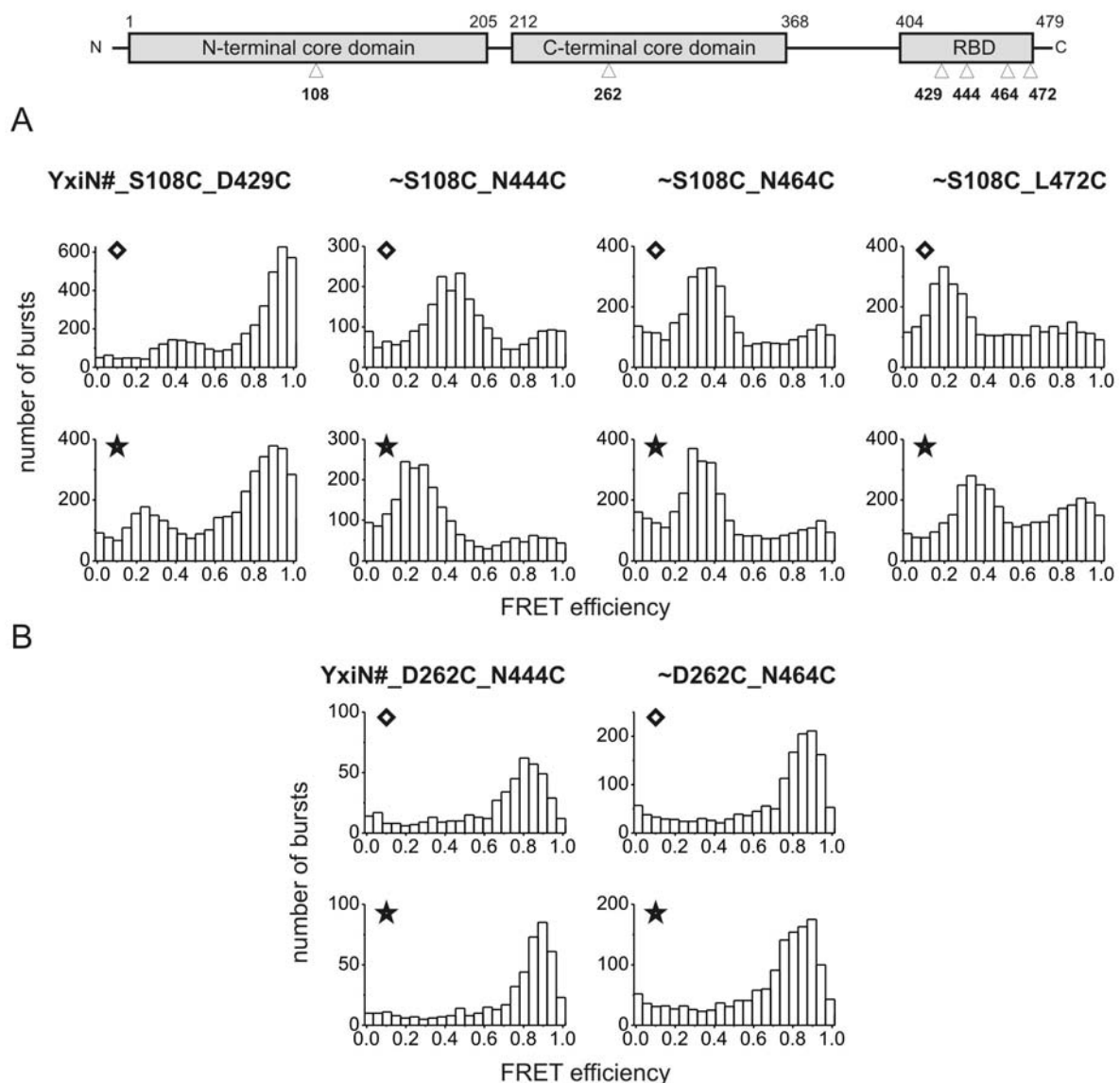

Figure 3.10
SmFRET histograms for YxiN constructs having fluorophores attached to the N- or C-terminal core domain and the RBD. The scheme at the top assigns the labelled residues to the domains of YxiN. YxiN# denotes the YxiN_C61A_C267A background of the double cysteine mutants. The histograms were calculated using correction parameters assuming A546 conjugated to the RBD (squares) or A488 conjugated to the RBD (stars). Only bursts above a threshold of 100 photons were considered for the calculation of the histograms.

The soluble constructs were fluorophore labelled and smFRET experiments were performed at least in duplicate. For each protein a representative histogram is shown in figure 3.10A. Mutant specific correction parameters for the conjugated fluorophores were determined and considered for the calculation of histograms. Since it is not known whether the fluorophores attach with equal probability to the cysteines in the double mutant, the boundary conditions were considered i.e. histograms were calculated for the RBD labelled with acceptor, S108C labelled with donor fluorophore and *vice versa*. The same procedure was employed when calculating the distance histograms. The correction parameters and the Förster distance for both extrema were considered. The mean distance from each histogram sets an upper and a lower limit for the distance between the fluorophores in a randomly double-labelled mutant. All results are summarized in table 3.1.

Table 3.1

Results of smFRET measurements for YxiN constructs labelled at a position within the RBD and a position within the helicase core. The mean FRET efficiency E_{FRET} was determined from FRET histograms calculated using correction parameters of the non-conjugated fluorophores. The experimentally determined mean distance between the positions r was determined from a distance histogram that was generated by summarizing the distance histograms from 2 to 4 independent measurements. The resulting histogram was fitted with Gaussians to obtain the centres of the peaks. The calculations were done using the correction parameters for two cases: the acceptor fluorophore is at the RBD and the donor is at the helicase core (RBD-A546) or the donor fluorophore is at the RBD and the acceptor is at the helicase core (RBD-A488).

*An unusually high γ was determined for the pair S108C-A546/N444C-A488. Therefore the calculated distance should be considered with care.

construct (YxiN# background)		E_{FRET}	$r_{RBD-A546}$	$r_{RBD-A488}$	remarks
D262C	D429C	0.88			
D262C	T440C	0.9			
D262C	N444C	0.89	4.5 nm	3.7 nm	
D262C	N452C	0.83			low yield, low labelling efficiency
D262C	N464C	0.88	4.2 nm	4.1 nm	
D262C	L472C	0.85			
I275C	D429C	0.9			
I275C	N464C	0.85			
E290C	D429C	0.94			
E290C	N464C	0.95			
S108C	D429C		5.9 nm	6.4 nm	predominant high-FRET species
S108C	T440C				no quantitative purification
S108C	N444C		5.6 nm	6.0 nm*	
S108C	N452C				no quantitative purification
S108C	N464C		6.0 nm	6.3 nm	
S108C	L472C		6.6 nm	6.1 nm	

The histograms for the RBD mutants in YxiN#_S108C background peak at values around 0.2-0.4, meaning the probed distances are within the dynamic range of the FRET experiments. However, in all cases a population of high-FRET value was also observed. In the case of YxiN#_S108C_D429C it even is the dominating species. For YxiN#_S108C_N444C and ~N464C the fraction of the high-FRET species could be reduced by centrifugation of the labelled protein before the smFRET experiments or by performing the preceding labelling reaction at lower temperature. This indicates that the high-FRET species may originate from small aggregates of labelled protein. In the case of YxiN#_S108C_D429C the fraction of high-FRET events could not be reduced. Nevertheless, taking the distance information from the other constructs into account, it is geometrically impossible that the high-FRET species probes the S108C-D429C distance. The origin of the high-FRET species remains elusive and the 0.3 to 0.4 state was therefore considered for positioning the RBD.

To gain additional information, FRET correction parameters and the Förster distances were determined for YxiN#_D262C_N444C and ~N464C. The recalculated FRET histograms are shown in figure 3.10B and the mean value from distance histograms can be found in table 3.1. Although the measurements probed distances outside the dynamic range as mentioned before, they can be used to approximate the distance between D262C and the RBD.

Taking all distance information into account, the position of the RBD relative to the helicase core could be mapped manually. Two possible orientations that satisfy the distance constraints are shown in figure 3.11. It is emphasized that the three-dimensional arrangements presented in the figure were created assuming the helicase core resembles the X-ray structure of MjDeaD. This assumption is supported by the nearly identical structures of the C-terminal core domains of YxiN and DeaD (chapter 1.1.2) and distance information between positions in the N- and C-terminal core domains of YxiN (chapter 3.1.1).

The depicted orientations are different from the RBD orientation in the structure of the DEAD box protein Hera (Rudolph & Klostermeier, 2009) as the results from the distance measurements with the E290C and I275C constructs already indicated. Both orientations are feasible regarding the structurally unknown linker region. The linking sequence between the C-terminus of the core and the N-terminus of the RBD is comprised of 36 amino acids. In the most extended conformation the linker can span a distance of up to 8.8 nm. In both proposed orientations the distance between the termini does not exceed this value.

There is evidence arguing in favour of orientation A. Taking a closer look at the structure of the isolated C-terminal core domain of YxiN, a shallow cavity can be found on its surface. No rigid secondary elements such as α-helices or β-sheets are present there. Instead, the surface is mainly formed by flexible loops possibly allowing for an adaptable interface. The RBD is located close to this cavity in orientation A.

When residues close to or at this cavity were mutated in a previous study, a drop in protein stability or activity had been observed (Theissen, 2006). In detail, mutation T237C, E238C, N239C and R288C caused a decreased ATPase activity and/or a lower melting temperature in unfolding experiments and/or the protein mutants could be purified with low yield only. Possibly, the overall folding stability of the proteins was influenced by affecting the interface between the RBD and the C-terminal core domain.

In the current study the mutation of two RBD residues that are located at the domain interface in orientation A, also affected protein stability. In orientation A, N452 and T440 are at the interface to the C-terminal core domain of YxiN. In the corresponding single cysteine mutants the overall protein stability was not affected at first sight. But slight instabilities might multiply when further mutations are present. Indeed, the double cysteine mutants YxiN#_D262C_N452C, YxiN#_S108C_N452C and YxiN#_S108C_T440C were prone to aggregation or
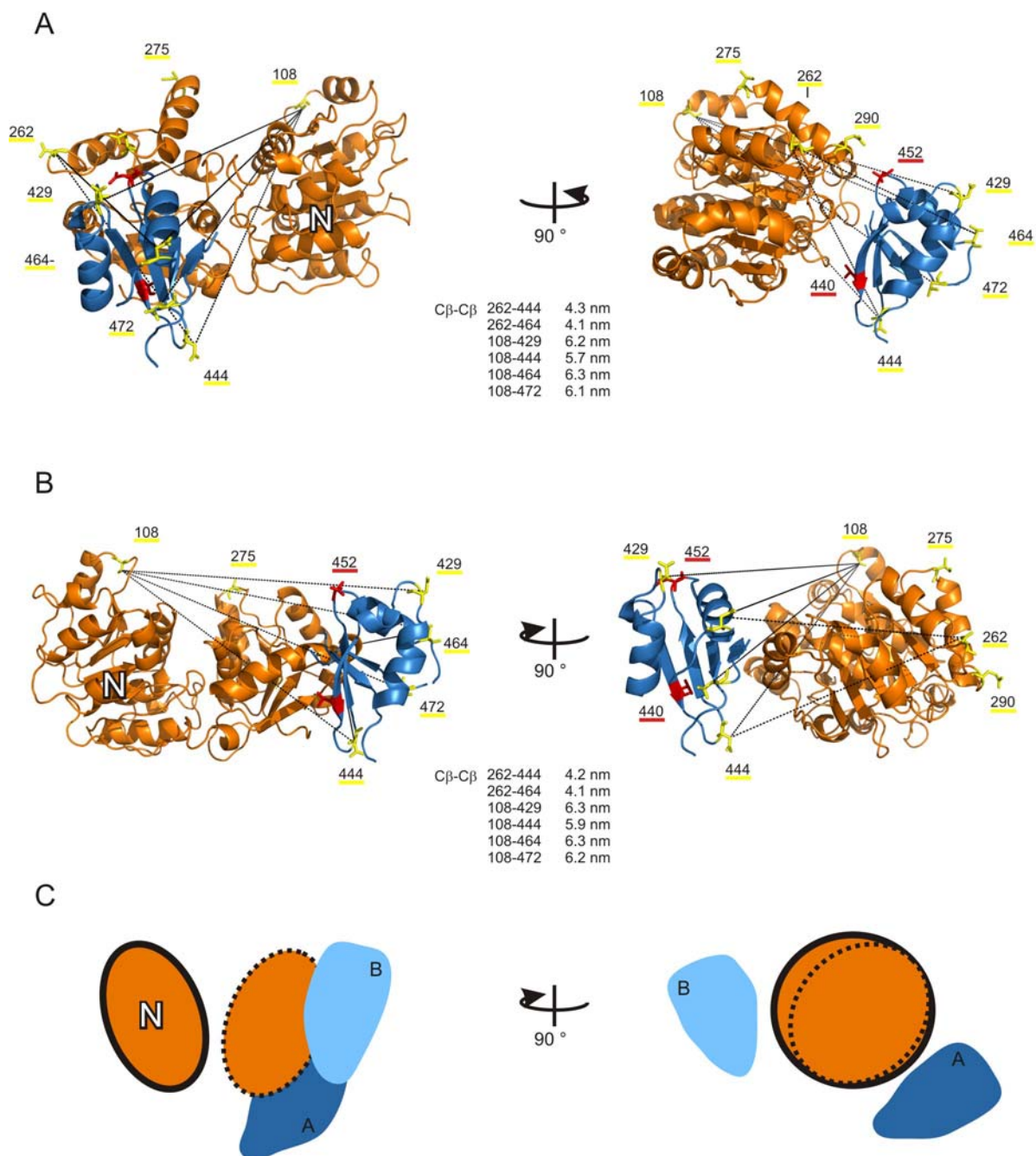
Figure 3.11

Orientation of the RBD (blue) of YxiN relative to the helicase core (orange) based on distance information from smFRET experiments. "N" labels the N-terminal core domain of YxiN.

A+B) The residues in stick representation were mutated to cysteines for fluorophore labelling for smFRET experiments. Distance constraints from these experiments (table 3.1) were used to position the X-ray structure of the isolated RBD (pdbID 2G0C) relative to a homology model of the YxiN helicase core. The orientations A and B satisfy the distance constraints as indicated by the Cβ-Cβ distances between residues in the models. The mutation of residue 440 or 452 (in red) caused a lower solubility of YxiN#_S108C constructs.

C) Schematic comparison of the orientations in A and B.

constituted insoluble proteins indicating a misfolding tendency. In addition, YxiN#_D262C_N452C could be fluorophore labelled with only low efficiency indicating that this position might be shielded from the solvent. Consistently, N452C is at the domain interface in orientation A and might be barely accessible for fluorophore labelling. These conclusions and related results from studies in literature will be discussed comprehensively in chapter 4.1.2.

In summary, it was possible to locate the RBD of YxiN relative to the helicase core using distance information gained from smFRET experiments. Two orientations fulfil the experimental distance constraints. Biochemical data favour orientation A in figure 3.11. Here the RBD is located in such a way that its central β-sheet faces a patch of the C-terminal core domain that is formed by various loops.

3.3 Development of EPL Based Techniques for SmFRET Sample Preparation

The development of biophysical techniques has often been accompanied by new sample preparation techniques. As an example, expressed protein ligation (EPL, chapter 1.2) was combined with various bioconjugation techniques to generate labelled or functionalized biomolecules. Chapter 3.3 describes the development of two techniques involving an EPL step. At first, a strategy for site-specific labelling with donor and acceptor fluorophores is described. This strategy constitutes an alternative to statistic double labelling of proteins. Secondly, an EPL-based technique is presented which functionalizes proteins in such a way that they can be immobilized on a surface. This allows for studying the proteins in TIR-FRET experiments.
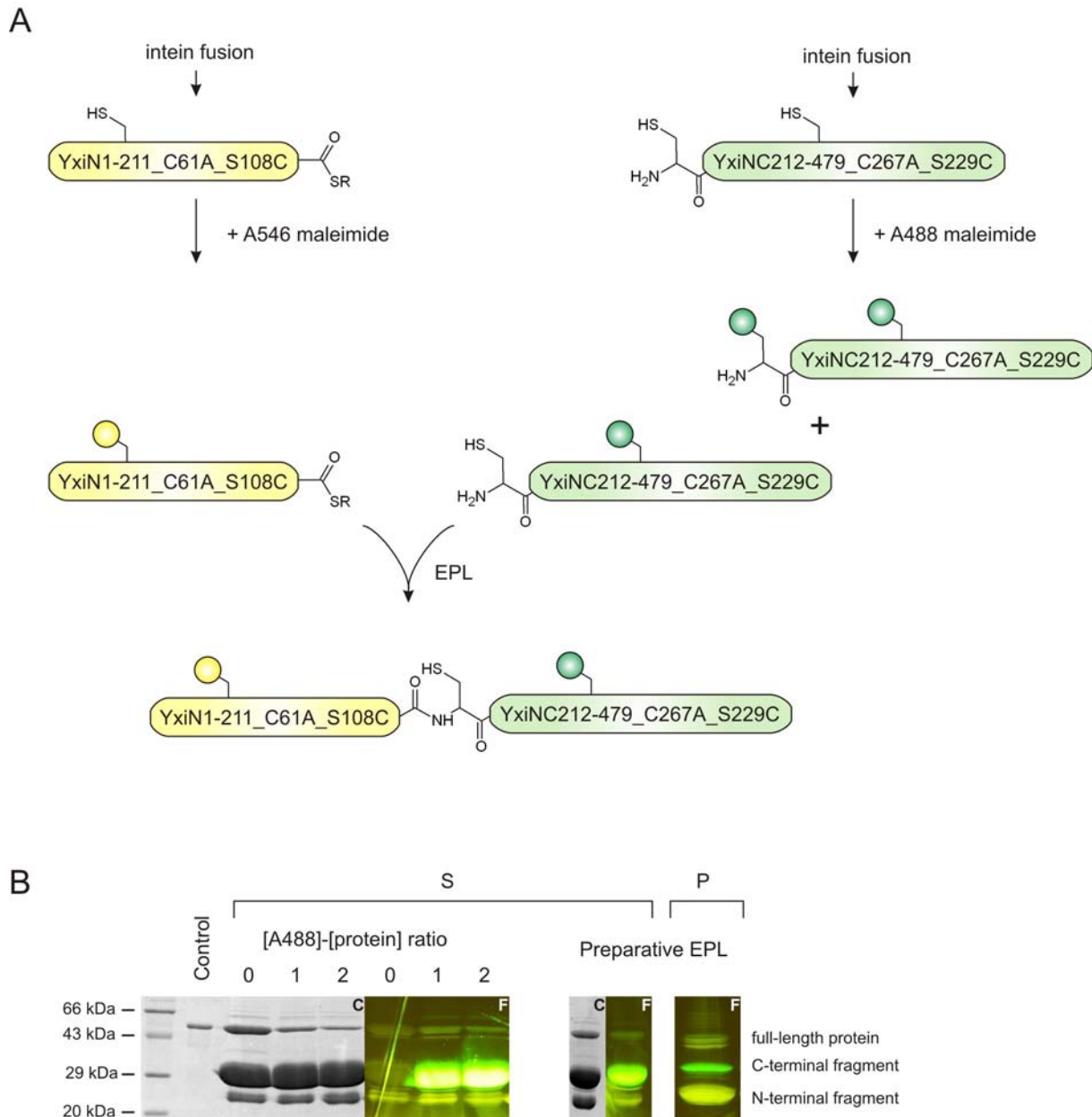
3.3.1 Site-Specific Double Labelling of YxiN

For smFRET experiments the molecules of interest need to be labelled with one donor and one acceptor fluorophore. Most prominently, proteins are labelled via the conjugation of a maleimide functionalized fluorophore to cysteines. For labelling with two fluorophores, two cysteines have to be accessible for the reaction. Often, these cysteines display similar reactivity and each cysteine can react with both donor or acceptor fluorophore. When a mixture of the fluorophores is added to the protein of interest, the protein will be labelled randomly. For a cysteine mutant XY the following species will be generated: X+Y-acceptor, X+Y-donor, X-acceptor+Y-donor, X-donor+Y-acceptor. In real life, the reaction will have a yield below 100 %. Therefore, unlabelled proteins and proteins carrying one fluorophore only will also be present. As a consequence, the labelling procedure causes a large heterogeneity within the sample which can affect the signal to noise ratio or can cause broadening of the smFRET histograms. To circumvent these effects, a procedure for site-specific labelling of two cysteines within one protein should be developed.

A Strategy for Site-specific Double Labelling

At first, the strategy visualized in figure 3.12A was employed on the DEAD box helicase YxiN. In principle, the procedure can be adapted for any protein. The three-domain protein YxiN was genetically split between domain 1 and 2 and the fragments were fused to inteins. The fusion proteins were designed as introduced in 1.2.1 and during purification the inteins catalyse the cleavage from the YxiN fragments. A C-terminal thioester is generated at domain 1 and the second fragment exhibits a N-terminal cysteine. It has been shown before that the full-length protein can in principal be generated from the isolated fragments using EPL (Karow *et al.,* 2007). This study furthermore revealed that the reconstituted YxiN construct displays wild-type like activity.

Since the protein fragments can be obtained separately, they can be labelled individually. This constitutes the methodical basis for the site-specific double labelling of the protein. After donor-labelling of one fragment and acceptor-labelling of the corresponding fragment the protein is reconstituted via EPL (as shown in figure 3.12A). Both EPL and the labelling reaction rely on proteinogenic cysteines. The C-terminal protein fragment was designed in such a way that it harbours an internal cysteine for labelling and a N-terminal cysteine for EPL. But upon addition of the maleimide coupled fluorophore to the C-terminal fragment the fluorophore can conjugate to both cysteines. If the N-terminal cysteine is labelled the reactivity of the thiol for EPL will be abolished. The EPL reaction categorically depends on the N-terminal cysteine (preceding EPL experiments with C-terminal fragments carrying a N-terminal serine or threonine did not show a significant formation of ligation product). To still allow for the ligation of the YxiN fragments after labelling, one has to make sure that a certain fraction of

C-terminal fragment is not labelled at the N-terminal cysteine. Therefore, quantitative labelling of the C-terminal fragment has to be avoided.

Figure 3.12
Site-specific double labelling of YxiN. A) The underlying strategy is depicted schematically. YxiN fragments, carrying cysteines for labelling, were expressed as intein fusions. The fusions were designed such that the N-terminal fragment is functionalized with a C-terminal thioester and that a N-terminal cysteine is generated at the C-terminal fragment. The fragments are individually labelled with maleimide-functionalized fluorophores. Subsequently, the fragments are mixed for an EPL reaction. A native peptide bond is formed between the fragments thereby reconstituting the full-length protein. Only C-terminal fragments that are not fluorophore labelled at the N-terminal cysteine will react with the N-terminal fragment. Further unlabelled or single labelled species will also form but are omitted for clarity. B) The reaction can be analysed by SDS-PAGE detecting protein bands via the fluorescence of the labelled species (F) or via coomassie staining (C). The reactions shown here were performed as follows: 92 μM YxiN1-211_C61A_S108C labelled with 367 μM A546 in buffer containing 3.67 M urea; 200 μM YxiNC212-479_C267A_S229C labelled with 200 or 400 μM A488; labelling reaction stopped by addition of 3 % (v/v) thiophenol; mixing the labelled fragments for EPL and incubation at 15 °C for 16 h. The EPL was also performed in preparative scale (same conditions as mentioned before but 165 μM C-terminal fragment were labelled with 152 μM A488). In all samples only a low amount of N-terminal fragment is in the supernatant (S). This fragment, in particular when labelled, is found in the precipitate (P).

The YxiN N-terminal domain comprises amino acids 1 to 211. The internal cysteine 61 was replaced by an alanine in this construct and a cysteine at position 108 was introduced for labelling. The fragment was purified using an intein system (Karow *et al*., 2007). In contrast to the wild-type fragment YxiN1-211, the double mutant was prone to aggregation. During the initial purification steps, it was necessary to solubilise the protein in buffer containing 2 M urea. The buffer was then exchanged to buffer without urea while the fusion protein was immobilized to an affinity column. The C-terminal fragment YxiNC212-479 carrying mutations C267A and S229C could be purified in urea-free buffer. The protein was stable and obtained with high yield.

The isolated fragments could be labelled separately and were subsequently mixed for EPL. The C-terminal fragment was labelled adding varying fluorophore concentrations to the protein. The concentration was optimized with respect to a high yield in the EPL reaction (i.e. low amount of modified N-terminal cysteine) at high labelling efficiency (C229 labelled preferentially). As expected, the formation of ligation product decreased upon addition of a higher fluorophore concentration (SDS-PAGE analysis figure 3.12B). The overall yield, however, is low. YxiN1-211_C61A_S108C tends to precipitate during the EPL and a large amount is therefore not effectively available for the reaction. When the incubation temperature was lowered to 15 °C and the reaction was performed in the presence of 2 M urea, precipitation was only slightly decreased.

Nevertheless, the EPL reaction could be performed in preparative scale (300 μl reaction mixture). It was possible to reconstitute a full-length YxiN construct that is labelled with acceptor fluorophore only at position S108C and donor fluorophore only at position S229C. However, the subsequent purification of the double-labelled protein via SEC was not successful due to the low yield of the reaction.

Semi-site-specific Double Labelling

Since the yield of the EPL reaction was predominantly limited by the low stability of YxiN1-211_C61A_S108C, an alternative strategy was developed that does not include the labelling of this fragment. Instead of S108C the cysteine at the ligation site (here T212C) should be labelled with the second fluorophore. The labelling of T212C shall be performed after the EPL reaction.

The entire procedure is depicted in figure 3.13. The N-terminal YxiN fragment YxiN1-211_C61A was expressed as an intein fusion, cleaved from the intein and functionalized with a C-terminal thioester. The C-terminal YxiN fragment was also purified from an intein fusion. It harbours an internal cysteine for labelling (the experiments were performed with a D429C or N444C construct) and a N-terminal cysteine. Both fragments were stable and could be purified using urea-free buffers.

The C-terminal fragment was fluorophore labelled in a next step. As mentioned before, the labelling was optimized with respect to a high yield in the EPL reaction at a sufficient labelling degree. The EPL reaction provides a selection step: C-terminal fragments that are labelled at the N-terminal cysteine do not react with the N-terminal fragment. The full-length protein consequently comprises the N-terminal fragment and an unlabelled C-terminal fragment or a C-terminal fragment that is labelled at the internal cysteine only.

As shown in figure 3.14A the formation of the EPL product thus decreased with increasing fluorophore concentration. For preparative EPL reactions a fluorophore-protein ratio of 1 was chosen. The reaction of YxiN1-211_C61A and labelled YxiNC212-479_C267A_D429C was successful in preparative scale. The product, however, could not be purified to acceptable purity. A side product of unknown origin was present in the samples. When YxiN1-211_C61A was ligated with labelled YxiNC212-479_N444C a higher yield of full-length protein was obtained. The product was purified via SEC to acceptable purity (figure 3.14B). A labelling efficiency of approximately 25 % A488 was determined neglecting the contribution of a small amount of

impurities. It is emphasized that this fluorophore is exclusively conjugated to N444C in this full-length YxiN. Subsequently the product was incubated with A546 that reacts with the unmodified cysteines (C212 at the ligation site and unlabelled N444C). The following double labelled species are generated in this reaction: one A546 molecule each is bound to C212 and N444C; one A546 at C212 and one A488 at N444C. Additionally unlabelled or single labelled species were present in the sample. In contrast to double labelling in a random fashion proteins carrying two A488 molecules are absent.

Since the attachment of one fluorophore (here A488) could be directed in a double cysteine mutant, the procedure displays a way for semi-site-specific labelling of a protein. The approach can basically be adapted to any protein.
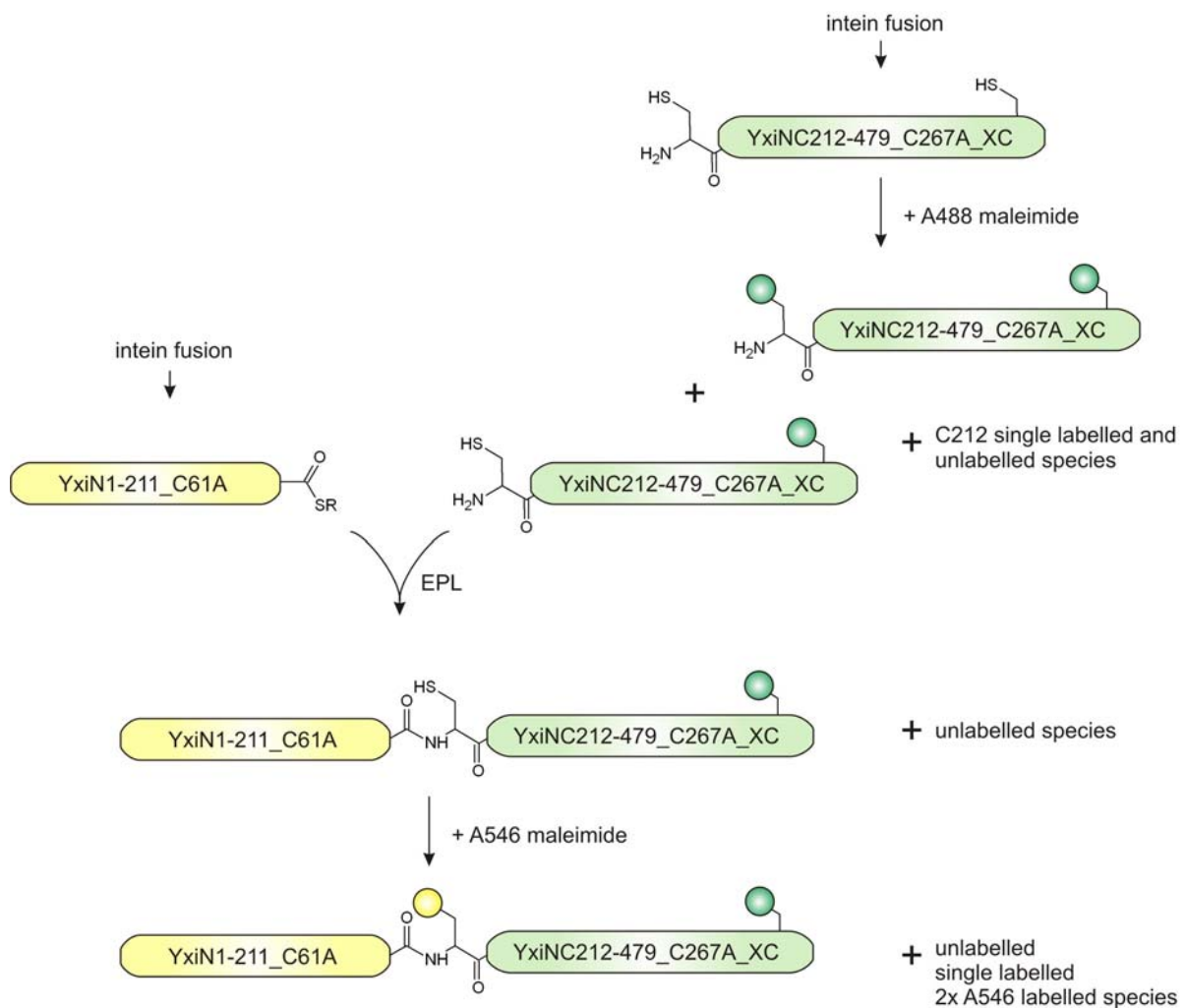
Figure 3.13
Strategy for semi-site-specific double labelling of YxiN constructs. YxiN fragments were expressed as intein fusions. The fusions were designed such that the N-terminal fragment is functionalized with a C-terminal thioester and that a N-terminal cysteine is generated at the C-terminal fragment. In addition the C-terminal fragment carries an internal cysteine for labelling. It is individually labelled and subsequently mixed with the N-terminal fragment for EPL. Only species that possess an unmodified N-terminal cysteine will undergo the ligation reaction. The full-length protein is finally labelled with a second fluorophore that will mainly attach to the cysteine at the junction.
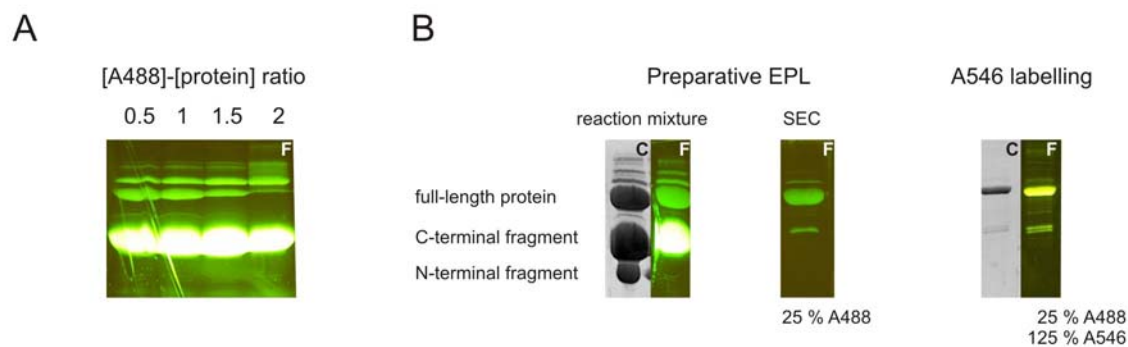
Figure 3.14

SDS-PAGE analysis of YxiN that was double labelled using the semi-site-specific strategy. Protein bands were detected via the fluorescence of the labelled species (F) or using coomassie staining (C).

A) As the C-terminal fragment is labelled with increasing concentration of fluorophore, the yield in EPL decreases due to increased modification of the essential N-terminal cysteine. 510 μM YxiNC212-479_C267A_D429C were labelled with a 0.5fold, 1fold, 1.5fold or 2fold excess of A488 maleimide, the reaction was quenched by adding 2.5 % (v/v) thiophenol. Upon addition of 110 μM YxiN1-211_C61A the mixture was incubated for 16 h at 15 °C for EPL.

B) In preparative scale 493 μM YxiNC212-479_C267A_N444C were labelled with 493 μM A488. The reaction was quenched with 2.5 % (v/v) thiophenol and 100 μM YxiN1-211_C61A were added. After incubation for 16 h at 15 °C the mixture was applied to a SEC column and fractions containing a high amount of full-length protein were pooled. The product is exclusively labelled at position N444C with 25 % efficiency. Subsequently the protein was A546 labelled with 125 % efficiency. A546 was conjugated to C212 and N444C.

SmFRET Measurements

In the following, smFRET experiments were performed on the double labelled YxiN construct derived via the presented EPL strategy. The corresponding histograms were compared with histograms of a statistically labelled YxiN_C61A_C267A_T212C_N444C (figure 3.15). The latter construct was expressed and purified as a full-length protein and was labelled by adding a mixture of A488 and A546 to the protein. The labelling degree was 25 % A488, 125 % A546 for both the EPL derived sample and the randomly labelled sample. The histograms for both constructs peak at a FRET efficiency of 0.78 meaning both exhibit the same global conformation. Thus, the EPL derived sample seems to be suitable for FRET experiments.

Broadening of a smFRET histogram can be caused by differences in donor quantum yield depending on the site of attachment. In an extreme case the quantum yield and therefore the Förster distance is significantly different between the donor at position X (here A488 at N444C) and donor at position Y (here A488 at T212C). For each case, the probed distance between X and Y would display a different FRET efficiency. In the smFRET histogram of a randomly labelled species, presumably two overlapping distributions appear that show up as one broad distribution i.e. a distribution with a large width. Via fitting the histograms with Gaussian distributions, the width can be determined. For the EPL derived sample a width of 0.15 and from a second experiment a width of 0.173 was determined. In two experiments on the randomly labelled construct, histograms displaying a width of 0.169 and of 0.167 were calculated. Thus, no significant difference in histogram width between the randomly and the semi-site-specific labelled YxiN construct was observed. For the examples presented here, the width of the histogram is presumably determined by experimental noise and segmental flexibilities. The width is not severely influenced by the donor quantum yield depending on the attachment to N444C or T212C in the corresponding mutant.

Semi-site specific labelling might be an alternative for double cysteine mutants of YxiN (and principally other proteins) that do display a very broad smFRET histogram due to a different donor quantum yield depending on

the attachment site. Employing this labelling procedure the attachment of the donor can be directed to one cysteine at which the donor displays one corresponding quantum yield.

The statistically labelled construct has a labelling degree of 25 % A488 (as the EPL sample). Due to the labelling procedure, A488-A488 (donor-donor) labelled species are present in the sample and, together with species carrying one donor, cause a distinct distribution at -0.1 FRET efficiency (negative FRET efficiencies are not of physical meaning but result from the histogram calculation procedure). In contrast, no donor-donor labelled proteins are present in the EPL sample as mentioned before. Consistently, no peak at negative FRET efficiencies is observed for the corresponding sample. This result exemplifies that a decrease in sample heterogeneity via semi-site specific labelling is beneficial for the signal to noise ratio in smFRET experiments.
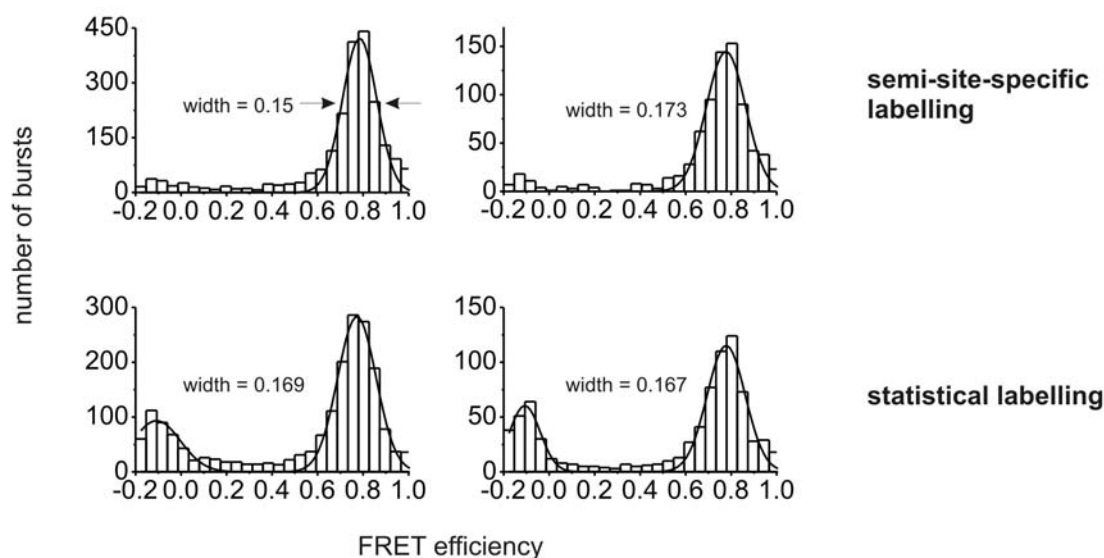
Figure 3.15
SmFRET histograms for YxiN_C61A_C267A_T212C_N444C. The donor-acceptor labelled protein was generated using an EPL based technique that allows for semi-site-specific labelling (upper row). Only protein species having the donor fluorophore attached to N444C were present in the sample. For comparison, YxiN_C61A_C267A_T212C_N444C was also purified as full-length protein and subsequently labelled in a random fashion (smFRET histograms, lower row). The histograms of two independent measurements each are shown here. For the calculation, only bursts above a threshold of 200 photons were considered and correction parameters of the non-conjugated fluorophores were used. The shape of the histograms and the width of the distributions, as determined via fitting with Gaussians, are similar. In the statistically labelled sample donor-donor labelled species are present that are registered as events of negative FRET efficiency. In the EPL product these species are absent.

3.3.2 C-terminal Biotinylation of Double Labelled YxiN for Immobilization in TIR-FRET Experiments

The smFRET experiments presented so far were performed on a confocal microscope with double labelled YxiN constructs freely diffusing in solution. Only a handful of switching events between two FRET states have been observed for YxiN until now (chapter 3.1.1). These events might be detectable in significant numbers if YxiN was observed on longer timescales. However, employing smFRET on a confocal microscope the time of observation is usually limited to several milliseconds. As already mentioned in section 2.15.3, TIR-FRET constitutes a complementary technique that provides for single molecule experiments on the seconds time scale. Thus, the increased observation time may allow for the observation of YxiN switching between two FRET states. Information from dwell time histograms would finally elucidate the kinetics of the conformational change discussed in chapter 3.1. For example, it would be possible to analyze the influence of motif mutations (chapter 3.1.2) on the transition kinetics between two conformations.

To increase the observation time, TIR-FRET experiments shall be set up for the DEAD box helicase YxiN. As TIR-FRET probes molecules close to the glass-buffer interface, YxiN needs to be immobilized on a functionalized glass surface. For this purpose, the approach depicted in figure 3.16 was designed. The protein of interest is produced as an intein fusion and functionalized with a C-terminal thioester during the purification procedure. After donor-acceptor labelling of the protein, an EPL-type reaction is performed. A cysteine-biotin moiety is added, that is covalently bound to the protein C-terminus via the mechanism described in chapter 1.2.1.
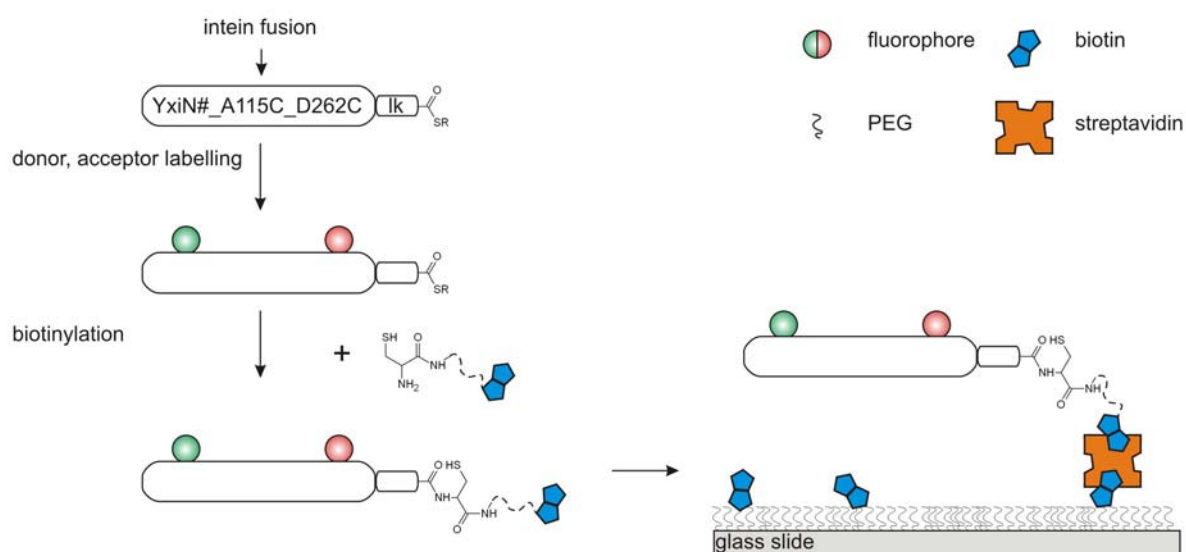
Figure 3.16
Immobilization of double labelled YxiN for TIR-FRET experiments. YxiN is purified and functionalized with a C-terminal thioester using an intein system. It is then labelled with donor-acceptor fluorophores at defined internal cysteines. After quenching of the reaction, a cysteine that is C-terminally linked to a biotin moiety (Cys-Biotin) is added. In an EPL reaction the molecule will attach to the C-terminus of YxiN.
The glass slide used for TIR-FRET experiments is coated with a PEG layer incorporating a few biotinylated PEG molecules. Streptavidin is applied that binds to the biotinylated PEG. Upon flushing the surface with biotinylated YxiN, streptavidin will also bind the latter biotin thereby immobilizing the conjugated protein. The streptavidin tetramer harbours 4 biotin binding sites that each display a high binding affinity (K_D 10^{-15} to 10^{-14} M).
In principle, the overall strategy can be adapted to any protein. It can be varied with respect to the spacer between the cysteine and the biotin in the Cys-Biotin moiety. PEG molecules of different chain length or branching characteristics can be used. A linker sequence (lk) can be introduced C-terminal to the protein of interest to provide a flexible attachment and sufficient spacing between the protein and the streptavidin molecule.

The subsequent immobilization procedure is widely used (Rasnik *et al.*, 2005). It makes use of the high affinity binding of streptavidin and biotin. A glass slide is functionalized with PEG molecules incorporating approximately 1 % biotinylated PEG molecules. When streptavidin is flushed over the surface, it will bind to these biotinylated PEG molecules. Since one streptavidin displays four biotin binding sides, the immobilized streptavidin can subsequently bind the biotinylated protein of interest.

Purification of Full-length YxiN Harbouring a C-terminal Thioester

For fluorophore labelling, the mutations A115C and D262C were introduced in YxiN_C61A_C267A (=YxiN#). The corresponding gene was amplified via PCR using a reverse primer that adds a linker sequence (GGGSGGGSGGT, abbreviated as GS) at the C-terminus. The linker sequence should provide flexibility and a sufficient distance between the immobilized streptavidin and YxiN. As the C-terminal domain of YxiN binds the RNA substrate, this domain must not be shielded by the surface or the streptavidin.
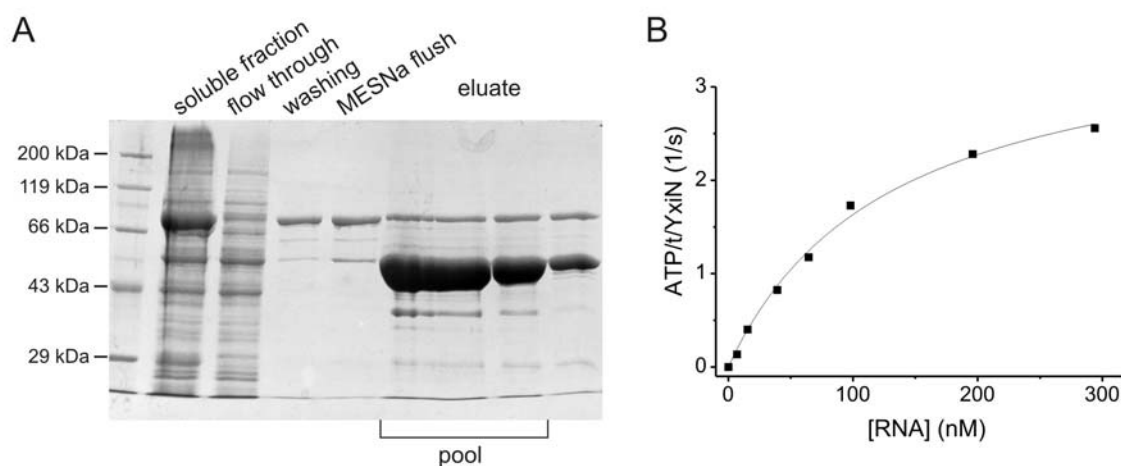
Figure 3.17
A) SDS-PAGE analysis of samples from the purification of YxiN#_A115C_D262C_GS. The protein was produced as an intein fusion additionally harbouring a chitin binding domain (total mass 83 kDa). The crude extract was applied to a chitin affinity column and non-specifically bound proteins were removed by washing. Cleavage from the intein was initiated by flushing with MESNa solution. After over-night incubation fractions containing a large amount of YxiN#_A115C_D262C_GS were eluted and pooled (experimental details in method section 2.9.4). B) The wild-type YxiN version carrying the C-terminal linker sequence (YxiN_GS) was purified analogously. The construct was assayed for 153mer RNA stimulated steady-state ATPase activity ($k_{cat} = 3.7$ s^{-1}, $K_{app,RNA} = 129$ nM). Exclusively for this assay, the protein was further purified on a SEC column to exclude that contaminant proteins catalyze the ATP hydrolysis. For labelling and biotinylation reactions the protein eluate from the chitin affinity column was used.

The construct was cloned into pTWIN1 as a N-terminal fusion to the *mxe* GyrA intein (experimental details in method section 2.7). Using the same procedure wild-type YxiN was also cloned into pTWIN1. The fusion constructs were produced and constituted soluble proteins. During the purification YxiN_GS was cleaved from the intein upon addition of MESNa. This thiol component forms a thioester with the terminal carboxyl group of the protein. Figure 3.17A shows the SDS-PAGE analysis of samples from the purification (method section 2.9.4). The protein was obtained in high quantity (16 mg from 1 l bacterial culture) and in acceptable purity. To check whether the attachment of the GS linker affects the protein activity, YxiN_GS was assayed in a steady-state ATPase assay. For YxiN_GS a k_{cat} of 3.7 s^{-1} and a $K_{app,RNA}$ of 129 nM were determined (figure 3.17B). With the same batch of RNA the wild-type enzyme kinetic parameters vary from a k_{cat} of 1.5 s^{-1} and $K_{app,RNA}$ of 156 nM (chapter 3.1.1) to a k_{cat} of 4.1 s^{-1} and $K_{app,RNA}$ of 33 nM (Karow *et al.*, 2007) depending on the protein

batch. The values for YxiN_GS are in the same range as these wild-type values. Thus, the GS linker does not impair the RNA stimulated ATPase activity of the protein.

After the elution from the chitin column, the protein can be labelled with fluorophores. For the moment, however, the optimization of the biotin attachment to the protein will be discussed.

Optimization of the Biotinylation Reaction

Figure 3.18A depicts the structure of the cysteine-biotin moiety (Cys-Biotin) that was attached to the C-terminus of YxiN_GS. Cys-Biotin was generated via coupling of 5-(biotinamido)pentylamine to a cysteine using common peptide synthesis protocols. If desired slightly altered reactants could also be chosen. Only the biotin and the cysteine parts are essential for the presented approach. For example, one could use a biotin moiety with a PEG chain between the biotin and the terminal amino group. After attachment of the corresponding modified Cys-Biotin to the target protein the PEG part might increase the solubility of this protein.
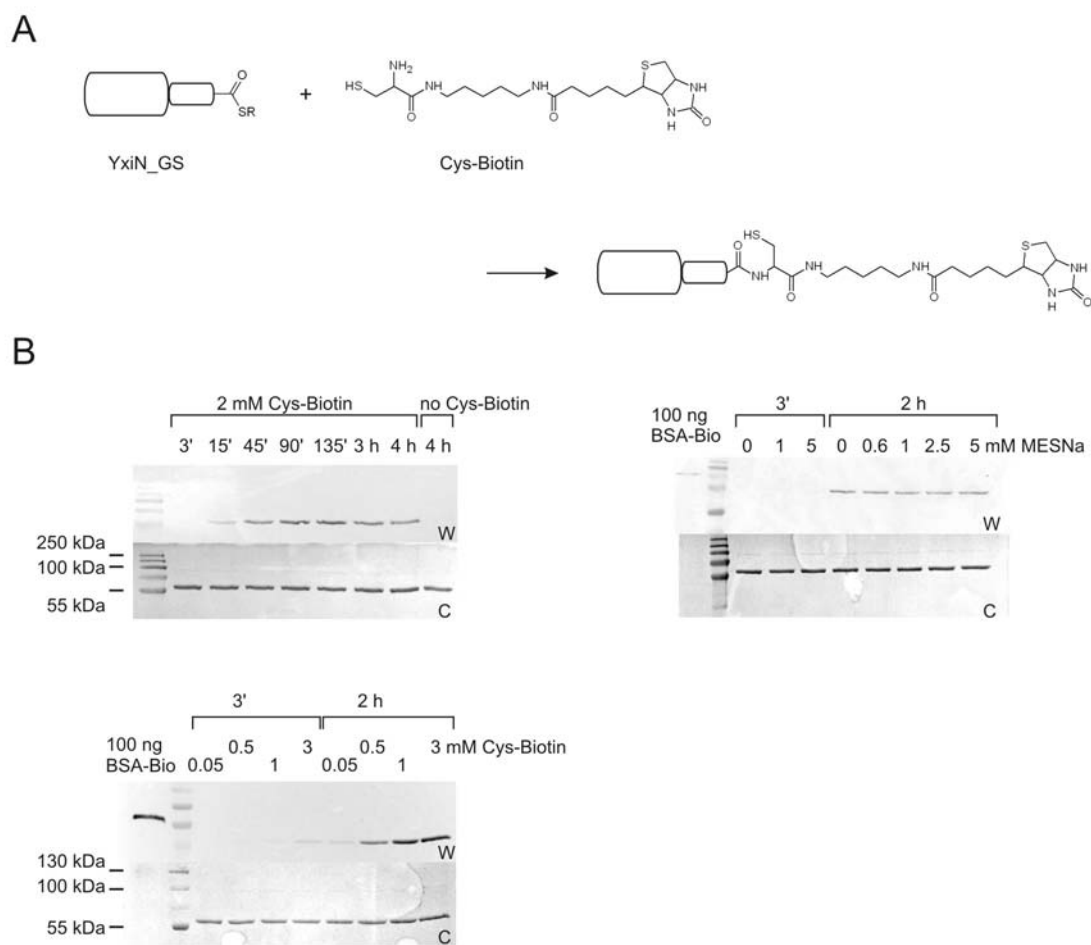
Figure 3.18
C-terminal biotinylation of YxiN_GS. A) In an EPL reaction, a cysteine-biotin moiety (Cys-Biotin) is coupled to the protein that is C-terminally activated by a thioester. B) The reaction was optimized with respect to incubation time, concentration of MESNa thiol cofactor and Cys-Biotin starting concentration. Unless indicated differently, 70 μM YxiN_GS were incubated with 2 mM Cys-Biotin and 5 mM MESNa at 25 °C for the indicated time. For analysis, the protein samples were subjected to SDS-PAGE/Western blotting followed by staining via avidin conjugated horse radish peroxidase (W). Bands on SDS-PAGE gels were detected via coomassie staining (C). Biotinylated BSA (BSA-Bio) served as a positive control. In the absence of Cys-Biotin no band is detected on the blot (9th lane, upper left panel, negative control).

The coupling to the protein occurs upon mixing of Cys-Biotin with the protein harbouring a C-terminal thioester at aqueous conditions, without the need of protection groups. As in any EPL like reaction, the reaction velocity can be enhanced by the addition of thiol cofactors (Dawson *et al.*, 1997; Johnson & Kent, 2006). To minimise health risks, a non-volatile thiol cofactor such as MESNa should be used whenever possible.

The reaction conditions were optimized with respect to MESNa concentration, Cys-Biotin concentration and incubation time. For this purpose protein samples were biotinylated in various series under different conditions and analysed.

The attempt to determine the biotinylation degree of these samples using a commercially available Quantitation Kit (Pierce) did not yield consistent results. As an alternative, SDS-PAGE/Western blot analysis was set up for biotin detection. This Western blot analysis provides a rather qualitative analysis. However, it is suitable to compare samples that contain the same amount of protein starting material but were incubated under different reaction conditions.

The corresponding blots are summarized in figure 3.18B. In the presence of a higher concentration of Cys-Biotin reactant the amount of biotinylated protein increases. Based on the result of this series a concentration of 3 mM Cys-Biotin was used in preparative biotinylation reactions. The amount of biotinylated protein levels off after about 2 h of incubation. Therefore, preparative biotinylations were incubated for 3 to 4 h. Unexpectedly, the reaction is not enhanced by the addition of MESNa. Nevertheless, 5 mM MESNa were added in preparative biotinylation reactions. The preparative biotinylation reaction follows close upon the fluorophore labelling of the protein (figure 3.16, method section 2.11.4). The remaining fluorophores must not attach to the thiol of the Cys-Biotin moiety. Therefore, the thiol component MESNa was added in huge excess to quench the labelling reaction. After incubation for 5 min, Cys-Biotin was added. Employing this procedure, it becomes very improbable that a fraction of fluorophores remains reactive for reaction with thiols and can conjugate to the C-terminal cysteine in the biotinylated protein. Therefore, fluorophores will only be present at the desired internal cysteines of YxiN (here A115C D262C).

Taken together, the preparation of fluorophore labelled and biotinylated YxiN comprises the following steps:

1) intein-mediated purification of the protein with a C-terminal thioester

2) protein labelling with donor and acceptor fluorophores

3) quenching of the labelling reaction via addition of MESNa

4) addition of Cys-Biotin for C-terminal biotinylation under the optimized conditions mentioned above

5) purification of the protein via SEC.

SmFRET on a Confocal Microscope

To check whether the fluorophore labelling and biotinylation procedure affects the activity of the protein, the propensity of YxiN#_A115C_D262C_GS to undergo a conformational change upon binding of RNA and ADPNP was assayed using the confocal smFRET setup. Histograms for the biotinylated protein only and in the presence of ADPNP and 153mer RNA are shown in figure 3.19. The experiments were done on samples that were labelled with A488/TMR, A488/A546 or A488/A555 and subsequently biotinylated. In all cases the characteristic conformational change (chapter 3.1) occurs upon binding of ADPNP and RNA. Concerning their activity, the constructs constitute suitable samples for TIR-FRET analysis. The C-terminal GS linker, the labelling and biotinylation procedure do not impair the enzyme.

As mentioned above, the biotinylation degree could not be assessed so far. Therefore the influence of the attached Cys-Biotin cannot be determined from the confocal smFRET experiments. One may argue that only the

non-biotinylated proteins are able to populate the closed conformation. However, the GS linker is present in any molecule. It does not impair the enzyme. Therefore, an influence of the comparatively short Cys-Biotin appears unlikely.
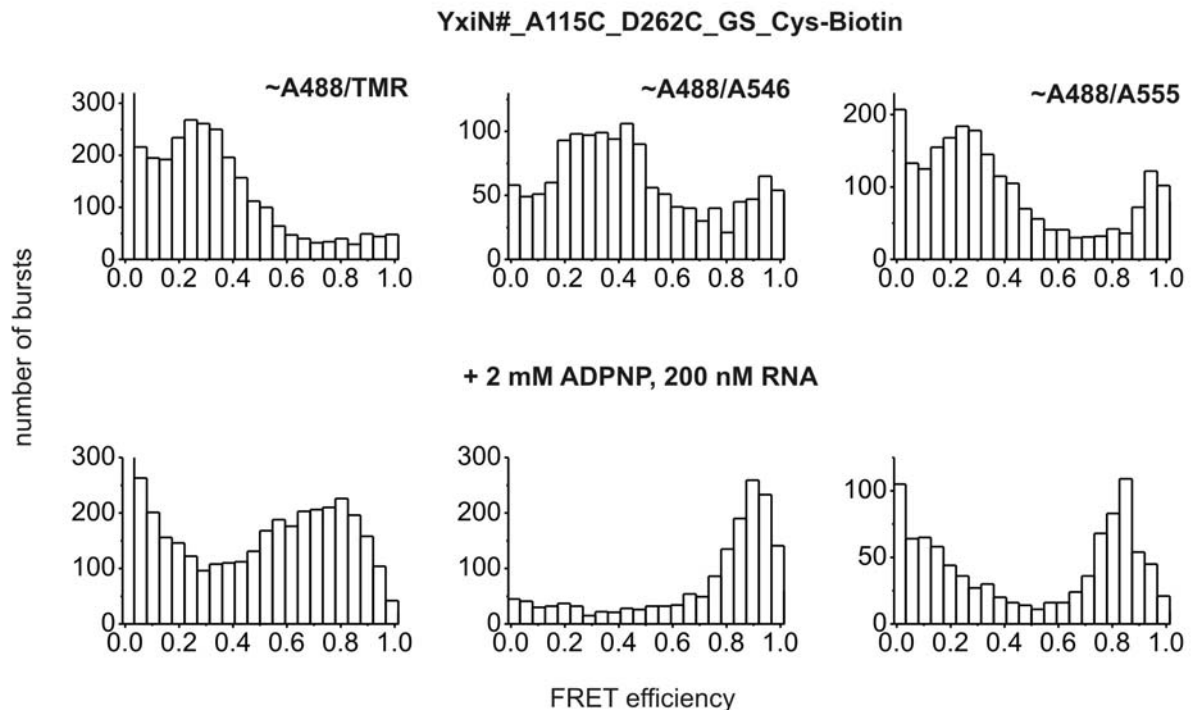
Figure 3.19
SmFRET histograms of biotinylated YxiN#_A115C_D262C. The protein was labelled with A488 donor fluorophore and different acceptor fluorophores (TMR, A546 or A555) and subsequently biotinylated. SmFRET measurements on a confocal microscope were performed on the protein only and in the presence of 2 mM ADPNP and 200 nM 153mer RNA substrate. In all cases, a conformational change occurs in the presence of these substrates. The protein is still active after the labelling and biotinylation procedure.
Only bursts above a threshold of 50 photons were considered for the calculation of the histograms

TIR Microscopy on Fluorophore Labelled and Biotinylated YxiN

To check whether the biotinylated YxiN construct can be immobilized specifically, TIR-microscopy was performed. Glass-slides that were functionalized with PEG/PEG-biotin were initially imaged on a TIR-microscope to assess the fluorescence background from the slides and the buffer used in the experiment. Subsequently the surface was flushed with labelled and biotinylated YxiN (in the absence of streptavidin). The surface was imaged to check for potential unspecific binding. Since streptavidin had not been applied yet, the biotinylated YxiN should not attach to the surface unless it adsorbs unspecifically. In contrast, after flushing the PEGylated surface with streptavidin and subsequently with the YxiN sample, bright fluorescent spots should appear on the TIR-image due to specifically immobilized YxiN molecules. The experiment was performed varying the buffer composition, the YxiN concentration and the excitation power. Using the following buffers little or no unspecific binding was observed:

50 mM Tris/HCl pH 7.5(25 °C); 1 M NaCl; 5 mM Mg_2Cl
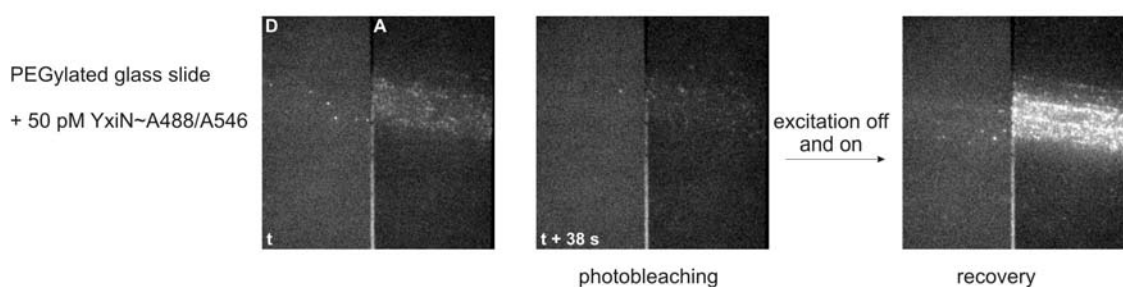
50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl

50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl; 4 % (v/v) glycerol

50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl; 10 % (v/v) glycerol.

In all experiments the preparation of the slides including cleaning and PEGylation performance turned out to be critical. A few times high background fluorescence was already observed before flushing with protein solution. In these cases the signal from the immobilized proteins could not be distinguished from the fluorescent impurities precluding the analysis of the data.

The integrity of the PEG layer is also crucial. PEG passivates the glass surface thereby minimizing the unspecific adsorption of sample molecules. At defects of the PEG coating ("scratches"), protein adsorbs non-specifically on the surface as illustrated in picture 3.20. After flushing with protein, without removal of unbound protein by a buffer flush, the area was imaged until all fluorescence was bleached. Subsequently the shutter was closed for some seconds. When the same area was imaged again, fluorescence had recovered indicating molecules from the solution still adsorb to the presumable layer defect. This behaviour was not observed for the surrounding regions.

Figure 3.20
Fluorophore labelled protein adsorbs at "scratches" of the PEG layer. 50 pM biotinylated, A488/A546 labelled YxiN#_A115C_D262C_GS in 50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl; 4 % (v/v) glycerol were flushed over a PEGylated glass slide (no preceding streptavidin immobilization!). The area was imaged on a TIR-microscope at 50 mW excitation power and 200 ms exposure time. D and A indicate the image from donor and acceptor channel.

Figure 3.21 shows exemplary series of TIR-images taken under different experimental conditions. In series A (excitation power of 50 mW) a rather high background fluorescence was observed. Nevertheless this series already indicates that high intensity spots appear when both streptavidin and subsequently YxiN were applied to the surface. In an attempt to monitor the surface for a longer time span before all fluorescence was bleached, the excitation power was decreased to 20 mW (3.12B+C). More obvious than in 3.21A YxiN does not bind non-specifically to the surface. The fluorophores still bleach within several 100 ms. To be able to monitor the protein of interest for a sufficient time span, a suitable oxygen scavenger system that decelerates bleaching will be necessary. The specific immobilization was successful with A488/TMR, A488/A546 and A488/A555 labelled YxiN.

In summary, this chapter describes a strategy for site-specific biotinylation of fluorophore labelled proteins for TIR-FRET experiments. The procedure was successfully employed on the DEAD box helicase YxiN but can, in principle, be adapted to any protein. Biotinylated YxiN could directly be detected on Western blots. The protein sample binds specifically to a surface that had been coated with PEG/biotinPEG and streptavidin which serves as an indirect evidence for the biotinylation of the YxiN construct. Employing the presented sample preparation, YxiN can now be immobilized and studied in TIR-FRET experiments.
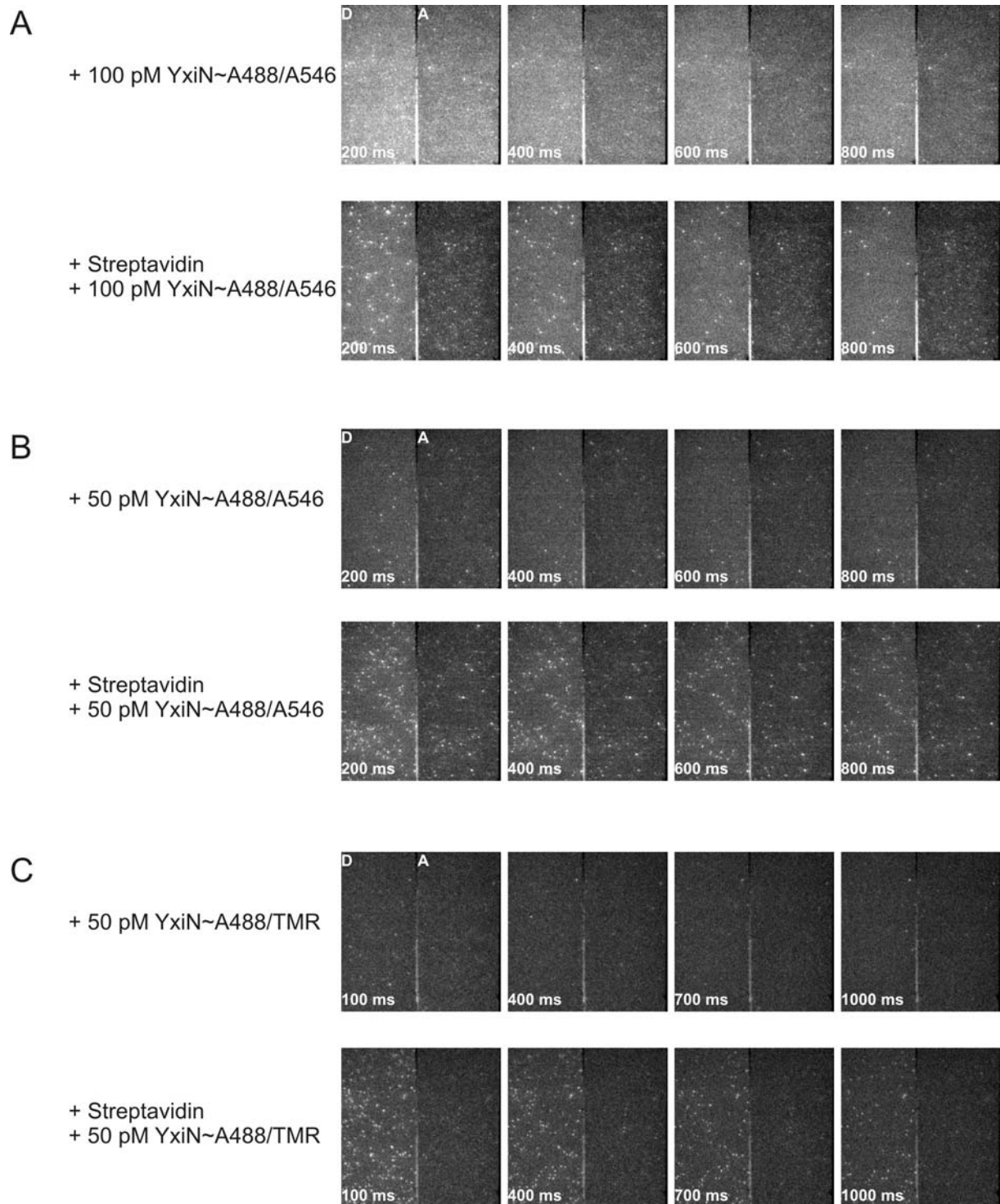
Figure 3.21
TIR-images of immobilized YxiN. The images from donor (D) and acceptor (A) channel at certain time points are depicted. Biotinylated, fluorophore labelled YxiN#_A115C_D262C_GS was flushed over glass slides that had been PEGylated before. This experiment serves as a control to assess the unspecific binding of the protein to the surface. In contrast, the protein should be immobilized when streptavidin was bound to the surface before. The experiment was performed under various conditions. A) 50 mW excitation power (determined at the light path before the microscope, the intensity at the glass slide is about 3fold lower); 200 ms exposure time; 50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl; 10 % (v/v) glycerol B) 20 mW excitation power (about 3fold lower at the glass slide); 200 ms exposure time; 50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl C) 20 mW excitation power; 100 ms exposure time; 50 mM Tris/HCl pH 7.5(25 °C); 150 mM NaCl; 5 mM Mg_2Cl; 4 % (v/v) glycerol

4 Discussion and Outlook

4.1 SmFRET Experiments Provide Information about Domain Orientations

4.1.1 The Structure of the YxiN Helicase Core

As introduced in section 1.1.2, any DEAD box protein comprises a two-domain helicase core. In all crystal structures of this core region or of the isolated domains a RecA-like topology can be observed for each domain. The isolated domains of different DEAD box proteins largely superimpose as it was also discussed in the introduction. The three-dimensional structure of the domains was additionally examined employing NMR experiments. Namely, NMR solution structures of DEAD box protein core domains have been published for Dbp5p (Fan *et al.*, 2009). In this study a crystal structure of Dbp5p was compared with the NMR structures of the isolated core domains. The crystal structure and the NMR solution structure of each domain are indeed very similar for Dbp5p. A solution structure of the full-length helicase core is not available yet.

In crystal structures of the whole core, the relative orientation of the domains in the absence of ligands is quite different (section 1.1.2 figure 2). This observation points towards a considerable flexibility of the linker between the core domains. In addition, there seems to be no or just a weak interaction between the domains.

The structure of the YxiN C-terminal core domain has been determined (Caruthers *et al.*, 2006) and was introduced in section 1.1.3. The YxiN N-terminal core domain most probably resembles the structure of this domain in other DEAD box proteins and can therefore be modelled. However, the relative orientation of the YxiN core domains is not known. To define the global conformation of the core, smFRET experiments were employed in the current project. It turned out that the YxiN core domains are in a less extended conformation than the core domains in known crystal structures. The experimental distances agreed best with the distances derived from the MjDeaD structure (Story *et al.*, 2001). Therefore the homology model of YxiN presented in section 1.1.3 constitutes a reasonable model for the YxiN helicase core.

The observation that the crystal structures display a much more extended conformation than observed for YxiN in smFRET experiments, might be due to the influence of crystal contacts on the core structure. In contrast to crystal structure determinations, in smFRET experiments the protein of interest is monitored freely diffusing. One is able to probe the solution structure of the protein.

Wang and colleagues also studied the conformation of YxiN in solution (Wang *et al.*, 2008). They set up small angle x-ray scattering experiments on the YxiN helicase core diffusing in solution. The data indicated that the core domains are in an orientation in which the two domains are distinguishable. They do not associate to one compact entity. This result is consistent with the core structure that, in the current project, was derived from homology modelling and smFRET experiments. In the derived YxiN homology model the domains are also separated. From the scattering data a centroid-to-centroid distance of 4.54 nm between the YxiN core domains was calculated (Wang *et al.*, 2008). In the structure of eIF4A (Caruthers *et al.*, 2000; depicted in section 1.1.2 figure 2) the corresponding distance is 4.86 nm (Wang *et al.*, 2008). Therefore the authors argued that the YxiN structure displays an extended shape similar to the eIF4A crystal structure. The smFRET data of the current

project, in contrast, indicate a more compact structure (centroid-to-centroid distance in MjDeaD homology model estimated 3.6 nm).

The small angle x-ray scattering experiments did not allow for conclusions on the relative rotation and position of the core domains (Wang *et al.*, 2008). From these experiments only an envelope of the molecule could be computed. Wang and colleagues only manually placed the domains into the envelope which is a subjective procedure. In contrast, the smFRET data and homology modelling in the current project directly provide experimental information about the domain orientation.

In the current study positions in the N-terminal core domain were chosen that are located rather close to the inter-domain cleft (compare section 3.1.1 figure 1A). Therefore no information from the rather distal regions of the N-terminal domain was obtained. As a result the tilting of this domain relative to the inter-domain cleft might have been probed with low precision. In future experiments one could choose further positions for fluorophore attachment that are equally distributed all over this domain. This would allow for a refinement of the smFRET results.

4.1.2 The Orientation of the YxiN RNA Binding Domain to the Helicase Core

The helicase core of YxiN is C-terminally flanked by an accessory domain. This domain provides high RNA affinity and specificity for the RNA substrate of YxiN (Kossen *et al.*, 2002; Karginov *et al.*, 2005). The crystal structure of this RNA binding domain (RBD) has been published (Wang *et al.*, 2006). However, the orientation relative to the helicase core is not known. This information would be helpful to obtain a comprehensive picture of the mechanism of YxiN with respect to the action of the RBD during RNA unwinding.

Therefore, the position of the RBD was determined in the current project employing smFRET experiments. The cysteines that were introduced for fluorophore labelling did not affect the RNA binding capabilities and the folding of the RBD as judged from unwinding assays.

Initially, distances between positions within the RBD and positions on the C-terminal core domain of YxiN were probed. FRET efficiencies close to 1 were determined meaning the probed distances were outside the dynamic range of the experiments. Distances between the RBD and the N-terminal core domain of YxiN were probed next. For various constructs FRET efficiencies within the dynamic range were determined. From the experimental data distances between S108C in the N-terminal core domain and various positions in the RBD were calculated. Taking these distance constraints into account the orientation of the RBD relative to the core could be mapped manually. Two orientations satisfy the distance constraints. However, further biochemical data favours one orientation (figure 3.11A, figure 4.1A). Namely, when residues that are located at the RBD-core interface according to this orientation were mutated the stability of the protein decreased. (table 3.1; Theissen, 2006). These residues at position 440 and 452 (RBD) as well as 237, 238, 239 and 288 (C-terminal core domain) are coloured in red in figure 4.1A. Possibly, affecting the interface between RBD and C-terminal core impairs the overall stability of YxiN.

In addition, cysteines at the positions 237, 238, 239, 288 and 337 could not be labelled with TMR-maleimide (Theissen, 2006) indicating that they are shielded from the solvent. In contrast, cysteines at position 224, 229, 262, 275 and 290 readily react with fluorophores (Theissen, 2006 and section 3.2.3). They are coloured in green in figure 4.1A and appear to be solvent accessible in this orientation.
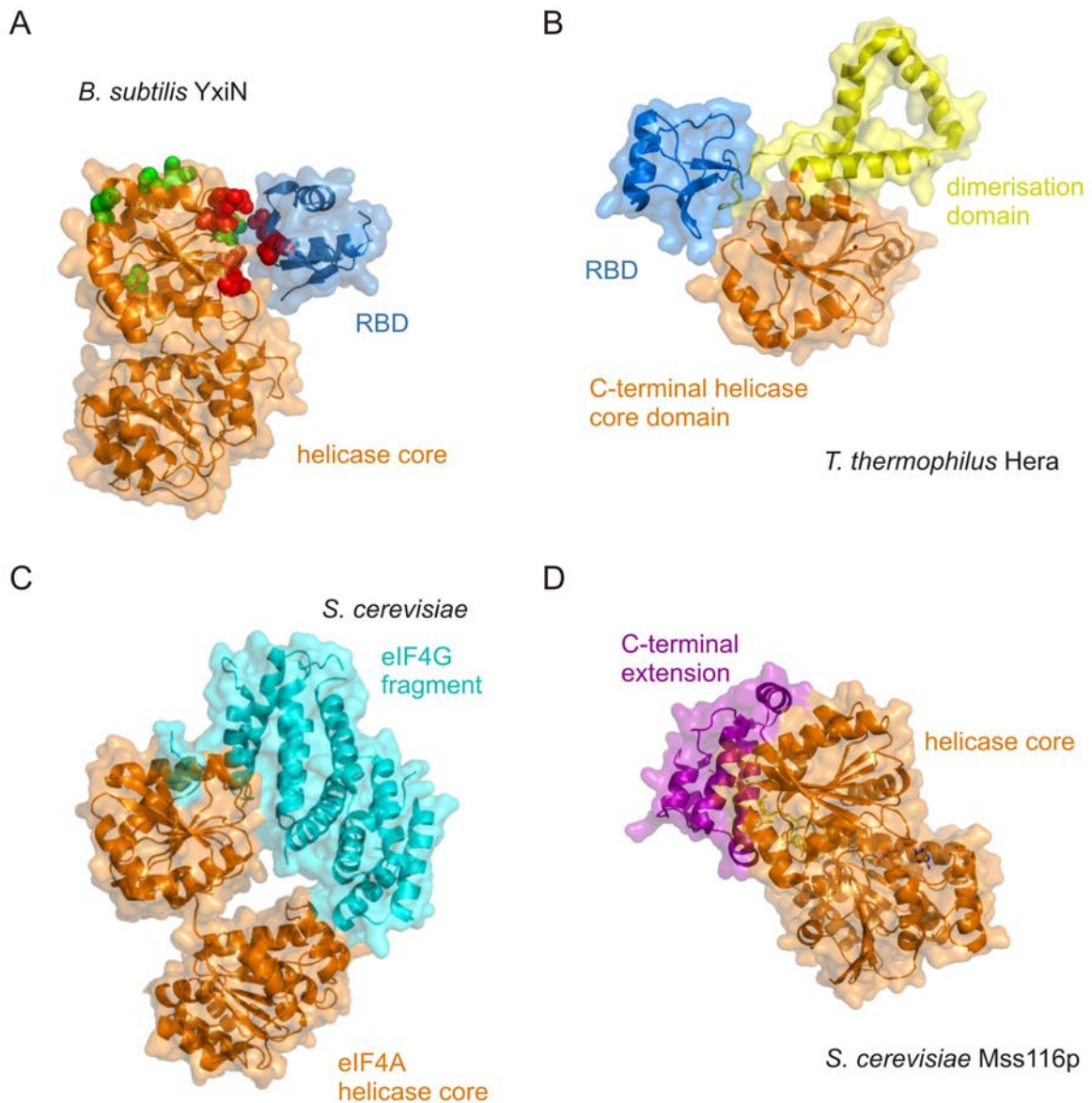
Figure 4.1
Structures of the core region of DEAD box proteins in context with accessory domains or interacting proteins. All structures are rotated in such a way that the C-terminal core domain is in a similar orientation in A, B, C, D. A) Structure of the core domain of YxiN (orange, homology model using the structure of MjDeaD, pdbID 1HV8 as a template) and the YxiN RBD (blue, pdbID 2G0C). The orientation of the RBD is proposed from smFRET data and biochemical data in the current project. In red: residues that upon mutation to cysteines decreased the protein stability and/or were inaccessible for labelling with TMR-maleimide. In green: cysteines at these positions can be labelled with fluorophores. B) Structure of the C-terminal core domain of Hera (orange), the flanking dimerisation domain (yellow) and the Hera RBD (blue), pdbID 3I32. C) Structure of eIF4A (orange) with a fragment of eIF4G (cyan), pdbID 2VSO. D) Structure of the Mss116p helicase core (orange) flanked by the α-helical C-terminal extension of the protein (purple). The protein was crystallised in the presence of ADPNP (stick representation) and RNA (yellow, stick representation), pdbID 3I5X.

A strong argument for the proposed orientation also comes from the literature. Talavera and colleagues published a study on the hydrodynamic characterization of YxiN's *E.coli* homolog DbpA (Talavera *et al*., 2006). In this study sedimentation coefficients and diffusion coefficients of DbpA were determined via analytical ultracentrifugation. Furthermore various models for full-length DbpA were created. At that time the crystal structure of the RBD had not been published yet. Hence, the authors used a *de novo* generated RBD model. The fold of this model deviates from the RBD crystal structure. Therefore only the shape of the protein comprising the core and the RBD can be assessed. For the models hydrodynamic coefficients were calculated and compared with the experimental data. The values agreed for models in which the RBD of DbpA is located at nearly the same location to the C-terminal core domain as in figure 4.1A.

For the DEAD box protein Hera a completely different orientation is observed in the crystal structure (Klostermeier & Rudolph, 2009; figure 4.1B). This does not argue against the proposed RBD orientation for YxiN because both proteins differ in many aspects. Hera is a DEAD box protein from the thermophile *Thermus thermophilus*. In contrast, YxiN is a mesophilic enzyme from *Bacillus subtilis*. Secondly, the folding of the RBD in both DEAD box proteins is not identical (Rudolph & Klostermeier, 2009) and Hera has been shown to display moderate substrate specificity (Linden *et al*., 2008) whereas only one RNA substrate has been found to activate YxiN (Kossen & Uhlenbeck, 1999; Kossen *et al*., 2002). In contrast to YxiN, Hera forms a very stable dimer in solution. These differences may indicate a different mode of RNA binding in Hera and YxiN. Both proteins may exhibit a distinct manner of RNA binding which requires a certain orientation of the RBD.

In general, little information about the orientation of C-terminal extensions of DEAD box proteins is available. Besides the Hera structure only one further structure of a DEAD box protein inclusive its C-terminal region has been published. Namely, the Mss116 protein from yeast was crystallised in the presence of RNA and various nucleotides (Del Campo & Lambowitz, 2009). The core region of this protein is flanked by an α-helical C-terminal extension that precedes a basic tail (Mohr *et al*., 2008). In the structure the extension does not fold into an independent domain (figure 4.1D). It rather expands the C-terminal core domain and contacts the polyU RNA substrate (Del Campo & Lambowitz, 2009). The position of the extension is completely different from the position of the Hera RBD and the position of the YxiN RBD proposed in the current project. But the C-terminal extensions in Hera, YxiN and Mss116p are not related and exhibit a different folding. In addition, they display functional differences. The Hera and YxiN RBD provide high affinity RNA binding and moderate or high specificity respectively for their RNA substrate. In contrast, the extension of Mss116p is thought to bind RNA non-specifically (Halls *et al*., 2007, Mohr *et al*., 2008). Furthermore, in the crystal structure of Mss116p this extension is causing a kink in the RNA substrate (Del Campo & Lambowitz, 2009). Therefore it was proposed that the extension might also assist the unwinding reaction. For YxiN and Hera so far no indications for a direct role of the RBDs for unwinding activity have been found.

In the YxiN domain orientation that is proposed in figure 4.1A, the RBD contacts a patch of the C-terminal core domain that is formed by various loops. A similar interaction can be found in the crystal structure of the eIF4A-eIF4G complex (Schütz *et al*., 2008; figure 4.1C). In this structure eIF4G contacts both core domains of the DEAD box helicase eIF4A. The C-terminal domain of eIF4A is contacted at this patch formed by various loops. This interaction site contributes most to the overall interaction of the proteins (Schütz *et al*., 2008). In this area Schütz and colleagues detected 23 hydrogen bonds and 14 salt bridges between eIF4G and amino acids of eIF4A that are located in the mentioned loops.

The eIF4A-eIF4G intermolecular interaction was reported to be strong and enhances the ATPase activity of eIF4A (Oberer *et al.*, 2005; Schütz *et al.*, 2008). In contrast, when the isolated YxiN RBD was mixed with the YxiN helicase core in the presence of RNA no stimulation in activity was observed (section 3.2.2). *In trans* the RBD is not capable to provide high affinity RNA binding to the whole enzyme. This indicates that the interaction between the RBD and the C-terminal core domain is comparably weak. A covalent linkage of both domains is necessary. With respect to its function, yet, the interaction between eIF4A and eIF4G is different from the interaction between YxiN RBD and core domain. eIF4G fixes eIF4A in a certain orientation to prevent the formation of a non-productive conformation (Schütz *et al.*, 2008). In contrast, the YxiN RBD binds the RNA substrate with high affinity and therefore provides the catalytic core with RNA. The presence of RNA then stimulates the ATPase activity of the core.

Notably, the YxiN-RBD interaction and the principally different inter-protein interaction between eIF4A and eIF4G involve the "loop patch" of the C-terminal DEAD box core domain. Thus, this patch may constitute a variable surface module for DEAD box proteins. It is not constrained by rigid secondary structure elements but comprises rather flexible elements that are capable to adapt to various interacting surfaces.

In future work the smFRET analysis should be extended to validate the proposed orientation of the YxiN RBD. As mentioned in the results section 3.2.3 some of the obtained smFRET histograms showed two distributions. One distribution at 0.2 to 0.4 FRET efficiency and one high-FRET distribution were observed. It was argued that the high-FRET distribution originates from small aggregates of fluorophore labelled protein in the assay solution. Therefore it would be preferable to identify constructs that are more stable. For example the A115C mutation could be employed instead of the S108C mutation in the current project. In addition, probing the distance between the RBD and an alternative position within the N-terminal core domain would provide further distance constraints. Using these complementary constraints the current model could be refined.

Future studies could also focus on the proposed interface of RBD and C-terminal core domain. As mentioned above, a few residues that might be located at the interface have already been identified (red colouring in figure 4.1.A). Following up, one could introduce a bulky residue at the presumable interface. Subsequently, the activity and stability of the corresponding mutant should be assessed. In addition, the bulky residue could also be introduced into one of the smFRET constructs. SmFRET experiments on the resulting protein should be performed to detect whether the bulky residue causes a different location of the RBD. Analogously, one could introduce mutations that affect putative interactions between amino acids at the interface.

After careful validation of the RBD position one could examine how the RBD directs the RNA substrate to the core. Moreover, smFRET experiments could assay the position of the RBD during the catalytic cycle of YxiN to clarify whether the RBD contributes to RNA unwinding.

4.2 Probing Conformational Changes during the Catalytic Cycle Provides Valuable Insights into the Mechanism of DEAD Box Proteins

4.2.1 YxiN Adopts a Closed Conformation upon Binding of RNA and Nucleotide

In the available crystal structures of DEAD box proteins the helicase core domains are not in contact with each other in the absence of ligands. Proteolysis experiments indicated that the conformation of DEAD box proteins changes upon RNA and nucleotide binding (Lorsch & Herschlag, 1998b; Henn *et al*., 2002; Cheng *et al*., 2005). The publication of the first structures of DEAD box proteins bound to RNA and ADPNP (a non-hydrolysable ATP analogue) finally revealed that in the presence of both ligands the core domains approach each other (protein Vasa: Sengoku *et al*., 2006; protein eIF4A-III Andersen *et al*., 2006). As a result of this, an intricate network is formed between residues of both domains and the substrates. In addition, direct inter-domain contacts are established. In the current project it was intended to detect conformational changes for the DEAD box protein YxiN. The study should be performed on the freely-diffusing enzyme. This approach would therefore be complementary to crystal structure determination and would allow to investigate the dynamics of the conformational change during the catalytic cycle of the enzyme. For this purpose smFRET experiments on the YxiN helicase were designed. The corresponding constructs for smFRET experiment constituted active helicases (section 3.1.1). In addition, the fluorophore attachment for FRET experiments did not impair the enzyme activity as judged by an unwinding assay (section 3.1.2). Here, even the ability of the labelled protein to re-enter the catalytic cycle was probed since multiple-turnover conditions were applied in the assay. Thus, the fluorophore labelled YxiN constructs were suitable to study the conformation of the protein during catalysis.

As discussed in 4.1.1 the results of smFRET experiments suggested that the YxiN core structure in solution is similar to the MjDeaD crystal structure. Upon addition of 153mer RNA substrate, ADP, ATP or ADPNP the global conformation does not change. In contrast, in the presence of both RNA and ADPNP (section 3.1.1) or both RNA and ATP (section 3.1.2) the core domains come closer to each other. From the smFRET data distances between the sites of fluorophore attachment could be calculated. The distances resemble those between the respective sites in the closed structure of Vasa and eIF4A-III (Sengoku *et al*., 2006; Andersen *et al*., 2006). It was therefore concluded that the global conformation of YxiN bound to RNA and ADPNP (ATP) resembles the conformation in the crystal structures. In summary, employing smFRET experiments it was possible to detect two different conformations of the DEAD box protein YxiN in solution. A closure of the inter-domain cleft was observed upon binding of RNA and ADPNP(ATP).

In the presence of RNA and ADPNP YxiN appeared to be locked in the closed conformation such as Vasa or eIF4A-III in the crystal. In contrast, when ATP and RNA were present during a smFRET measurement the open and the closed conformation were populated by YxiN. As ATP is hydrolysed, YxiN proceeds through the catalytic cycle switching between both conformations. The open and the closed conformer are in a steady-state. As it will be discussed later in detail, the varying status of the bound nucleotide (prehydrolysis to posthydrolysis) effects the distribution of the conformers.

In the crystal structures of Vasa and eIF4A-III a kink in the single stranded RNA substrate was observed. It was speculated that this kink is incompatible with a RNA double helix and therefore a RNA double strand would be unwound (Sengoku *et al*., 2006; Andersen *et al*., 2006). Moreover, Sengoku and colleagues suggested that

protein mutants that fail to close completely do not bend the RNA efficiently. Combining both notions, the closed conformation would be considered as a prerequisite for RNA unwinding. In the current project evidence for this assumption was found. Namely, smFRET experiments and activity assays were performed on a YxiN construct that carries a mutation in the helicase motif V. Here, glycine 303 had been replaced by an alanine. The corresponding construct was unwinding deficient but was able to hydrolyse ATP. Results of smFRET experiments showed that this mutation impairs the propensity of the protein to undergo a conformational change. It further appeared that in the closed conformation the domains are a bit farther apart than in the wild-type construct. According to the crystal structures of the closed Vasa and eIF4A-III the glycine 303 that was mutated is in close proximity to the γ-phosphate of the bound nucleotide. The G303A mutation places a methyl group at this position. This group would collide with the nucleotide and therefore forces the inter-domain cleft to be slightly opened. Since the G303A mutant was unwinding deficient it can be concluded that a complete closure of the inter-domain cleft is a prerequisite for RNA unwinding.

Evidence for an altered conformation of this glycine mutant is also present in literature. The DEAD box protein Dhh1p was subjected to limited proteolysis in the presence of various substrates (Cheng *et al.*, 2005). The experiments were also performed with a motif V glycine mutant of Dhh1p. In the presence of RNA and ADPNP this mutant was less protease resistant than the wild-type under the same conditions. The authors concluded that this mutant fails to adopt a conformation that is as compact as in the wild-type. The results of the smFRET experiments in the current project now are in agreement with the conclusions of Cheng and colleagues. In the future the smFRET experiments might be refined to obtain a clearer picture. A motif V mutant might be created in which the critical glycine is replaced by a more bulky residue than an alanine. In smFRET experiments with the corresponding construct in the presence of RNA and ADPNP one might observe an even wider opening of the inter-domain cleft.

The experiments with Dhh1p (Cheng *et al.*, 2005) and YxiN (current project) only probed the protein conformation. The question arises how the conformation of the RNA changes if the protein adopts the closed conformation incompletely. Therefore future experiments should address the conformation of the RNA.

In the current study it was shown that binding of RNA and nucleotide take place in a cooperative fashion. Cooperativity in RNA and nucleotide binding has been observed for several DEAD box proteins (Lorsch & Herschlag, 1998a; Iost *et al.*, 1999; Polach & Uhlenbeck, 2002). In the current project, the K_D for the protein mant-ADP or mant-ADPNP complex decreased ~40- or 80-fold respectively when RNA substrate was added. Since YxiN adopts a closed conformation in the presence of RNA and ADPNP, the conformational change can be interpreted as the structural reason for the cooperativity. In the presence of both substrates the enzyme switches to a tight binding state i.e. the closed conformation. Thereby the ATPase site is assembled and ATP consumption is allowed. This provides a regulation mechanism coupling ATPase activity to the presence of RNA and thus preventing futile ATP hydrolysis in the absence of RNA.

While the ATP and RNA ligated state could be clearly assigned to the closed conformation, the results from smFRET experiments in the presence of RNA and ADP were not as clear. For smFRET constructs in which two internal cysteines had been replaced by serines no closure was observed upon binding of ADP and RNA. In contrast, a construct in which the two cysteines had been replaced by alanines populated the open and the closed conformation in the presence of ADP and RNA. In comparison with the cysteine-to-serine constructs, the cysteine-to-alanine construct more closely resembles the wild-type protein as judged by RNA stimulated ATPase activity. Presumably, the latter construct might also display a more wild-type like behaviour in the smFRET

experiments. The cysteine-to-serine construct might adopt the closed conformation to a lesser extent making the detection of the closed conformer more difficult.

For the cysteine-to-alanine construct a comprehensive set of control experiments was performed in the current project: different substrate batches were tested, the formation of ATP from ADP via disproportionation could be excluded and contaminant ATP from the ADP solution was depleted *in-situ* via hexokinase treatment. The results of all of these experiments argue for a population of the closed conformer in the presence of ADP and RNA. This observation is also supported by the observed cooperativity in RNA and ADP binding for YxiN. The cooperativity is yet lower than for binding of RNA and ATP. Consistently, the fraction of closed conformer in the presence of RNA and ADP is lower than in the presence of RNA and ATP.

Notably, regarding cooperativity in binding of RNA and ADP there is no uniform notion in the literature. Employing filter binding assays no cooperativity was observed for Ded1p and DbpA (Iost *et al.*, 1999; Polach & Uhlenbeck, 2002). In contrast, cooperativity was observed for eIF4A (Lorsch & Herschlag, 1998a), and a conformational change of eIF4A upon binding of RNA and ADP was delineated from a proteolysis study (Lorsch & Herschlag, 1998b). A recent study addressed the binding kinetics of the ligands (Henn *et al.*, 2008). The authors observed biphasic mant-ADP binding in the presence of RNA. In the absence of RNA only one phase was detected. The two phases in the presence of RNA were interpreted as reflecting the initial binding and an isomerisation step. The same group subsequently extended their analysis and delineated a pathway for the unwinding reaction of DbpA (Henn *et al.*, 2010). The authors state that the ADP ligated state exhibits a higher RNA affinity than the nucleotide-free state. It appears that the detection of cooperativity in binding of ADP and RNA depended on the employed assay. Nevertheless, indications for a conformational change were observed in at least four independent studies (Lorsch & Herschlag, 1998b; Henn *et al.*, 2002; Henn *et al.*, 2008; current project). The smFRET data in the current project indicate that the ADP/RNA ligated YxiN adopts a conformation that is similar to the global conformation of the ADPNP/RNA ligated protein. Though, in the presence of ADP and RNA a smaller fraction populated the closed conformer than in the presence of ADPNP and RNA.

In summary, it can be concluded that the DEAD box protein YxiN is in an open conformation in the absence of ligands and the presence of RNA substrate, ADP, ATP or ADPNP. Upon binding of both RNA and an adenine nucleotide YxiN is capable to adopt a closed conformation. The extent of populating this conformation depends on the nature of the nucleotide. The presence of a γ-phosphate seems to contribute to efficient closure of the inter-domain cleft. The nucleotide state thus regulates the conformation of the helicase core thereby constituting or reducing the RNA-binding surface. The complete closure might kink the RNA and promote the unwinding. One might speculate that the closed conformation needs to be populated for a certain time span to allow for unwinding. Possibly in the presence of ADP the closed conformation is only shortly populated. Consequently, ADP does not promote unwinding as it is also found experimentally. The kinetics of the transition between the conformations could not be investigated in the current project. The presented smFRET data were obtained employing a confocal microscope where the observation time is limited to several milliseconds. Conformational transitions of YxiN occur on a longer timescale, and switching between the open and the closed state could therefore only be detected in a handful of events. As introduced in the method section 2.15.3 TIR-microscopy allows for the study of enzymes at longer time-scales. In future work YxiN could be studied via TIR-FRET for example to compare conformational transitions of the ADP/RNA and the ATP/RNA ligated YxiN. For this

purpose a sample preparation technique has already been developed during the current project. It will be discussed in 4.3.

4.2.2 The Closure of the Inter-domain Cleft is not Sufficient for RNA Unwinding

To gain further insight into the role of the conformational change for the activity of DEAD box proteins, a mutational study was performed in the current project. As discussed in the introduction DEAD box proteins are characterized by several conserved motifs. Mutations in these motifs affect their activity differently. Therefore mutational studies allow for dissecting the ligand binding, ATPase and unwinding activity and allow for the study of the coupling of all of these processes. In the current project mutations in the motif I, III, V and the Q-motif of YxiN were introduced. These motifs are involved in ATP binding and hydrolysis (I), the formation of inter-domain contacts (III, V) and ATP binding via adenine recognition (Q-motif). The motif mutants were characterized biochemically and their propensity to undergo a conformational change was examined. The behaviour of the motif V mutant and the resulting conclusion were already discussed in the previous section.

Motif III has for a long time been proposed to be a "coupling motif" responsible for linking ATPase to unwinding activity. In a pioneering study a motif III mutant of eIF4A showed unaffected RNA-stimulated ATPase activity but was unwinding deficient (Pause & Sonenberg, 1992). The authors concluded that motif III couples ATPase to unwinding activity. No such clear statement could be drawn from a study on motif III mutants of the DEAD box protein Has1p (Rocak *et al.*, 2005). Depending on the mutated residue the mutants exhibited 35 to 78 % of the wild-type unwinding activity. In the current project the motif III mutant of YxiN (YxiN_AAA) exhibited a decreased k_{cat} in ATPase activity. In addition it was found that the coupling energy for the binding of RNA and ATP was lower than for the wild-type. Furthermore, the mutant exhibited a decreased unwinding velocity in comparison with the wild-type but was not unwinding deficient. Therefore mutating motif III does not strictly uncouple ATPase and unwinding activity. In other words motif III is not essential for their coupling.

In the crystal structure of the closed Vasa or eIF4A-III motif III forms inter-domain contacts with motif II and VI (in the N-terminal and C-terminal core domain respectively) (Sengoku *et al.*, 2006; Andersen *et al.*, 2006). Therefore one could suspect motif III is important for the formation of the closed conformer. In smFRET experiments the motif III mutant of YxiN was able to adopt the closed conformation in the presence of RNA and ADPNP(ATP). The corresponding histograms resembled the histograms of the wild-type like construct. However, in control experiments the fluorophore labelled protein used in the smFRET experiments exhibited an impaired unwinding activity in comparison with the unlabelled motif III mutant. The fluorophores seem to affect the activity of this mutant. Thus, the results of the smFRET experiments have to be interpreted with care. But clearly, the motif III does not seem to have an essential role for the conformational change to occur.

Combining the biochemical data and the structural information, it appears that motif III contributes to the alignment of the domain interface in the closed conformation. Only when all inter-domain contacts are formed the DEAD box protein can hydrolyse ATP efficiently and therefore catalyse RNA unwinding with maximal velocity. This conclusion is in line with the results of a very current study on the DEAD box protein Ded1p (Banroques *et al.*, 2010). The authors performed a comprehensive study on the role of each amino acid of motif III for the activity of Ded1p. They further extended their analysis to the *in vivo* phenotype of Ded1p motif III mutants. The *in vitro* experiments showed that, in comparison with the wild-type, Ded1p motif III mutants exhibit a lower k_{cat} in ATP hydrolysis, unwind RNA with reduced velocity and have an altered affinity

for single-stranded RNA. *In vivo* experiments revealed a "strong phenotype" for motif III mutants (Banroques *et al.*, 2010). These data for Ded1p also indicate a role of motif III for stabilising the closed conformer thereby aligning the ATPase site and forming an extended RNA binding surface.

In the current project furthermore a YxiN construct with a mutation in the Q-motif was studied. The Q-motif closely contacts the adenine of the bound nucleotide and was suspected to relay nucleotide binding to other conserved motifs (Cordin *et al.*, 2006). When the propensity of a Q-motif mutant of YxiN to undergo a conformational change was assayed, the population of the closed conformation in presence of RNA and ADPNP(ATP) could not be detected. But the corresponding construct was presumably not sufficiently ligated with the substrates in these experiments since it exhibited a massively decreased nucleotide affinity in comparison with the wild-type enzyme. The results for the Q-motif mutant of YxiN are in agreement with the proposed role for the Q-motif in nucleotide binding. But its role for the conformational change cannot be addressed in the absence of nucleotide binding and the smFRET experiments could not provide novel insights.

Motif I constitutes the Walker A motif of ATPases in general and is responsible for ATP binding and hydrolysis (introduction section 1.1.2). Upon mutation of motif I a loss of ATPase and unwinding activity of DEAD box proteins has been observed (Rozen *et al.*, 1989; Pause & Sonenberg, 1992; Cordin *et al.*, 2004; Rocak *et al.*, 2005; Solem *et al.*, 2006). In the current project the motif I mutant YxiN_K52Q was generated. It turned out to be ATPase and unwinding deficient. SmFRET experiments showed that this motif I mutant populated the closed conformation in the presence of RNA substrate and ADPNP (or ATP or ADP) like the wild-type YxiN. Therefore, affecting motif I does not inhibit the conformational change. Nevertheless, the mutant construct was unwinding deficient leading to the conclusion that the conformational change is not sufficient for RNA unwinding.

According to the smFRET histograms the YxiN_K52Q mutant populates the closed conformation in the presence of RNA and ATP to a larger extent than the wild-type. Since the K52Q mutant is ATPase deficient it cannot proceed through the catalytic cycle but is stabilized in the closed conformation initially after ligand binding. Recent studies on YxiN's analogue DbpA further dissected the event of ATP binding in the presence of RNA (Henn *et al.*, 2008; Henn *et al.*, 2010). In their kinetic analyses the authors distinguish an initial ATP/RNA complex with low RNA affinity from a subsequent complex of high RNA affinity. They suggest that the protein undergoes an "isomerisation" from the low affinity complex to the high affinity complex. The YxiN motif I mutant is probably captured at this high affinity state in the presence of ATP and RNA. A recent study addressed the conformation of YxiN at various stages of the catalytic cycle (Aregger & Klostermeier, 2009). The authors employed ATP analogues that should mimic different states of the nucleotide during its hydrolysis. YxiN was found to adopt the closed conformation in the presence of ADP·BeF$_x$+RNA and ADP·MgF$_x$+RNA. From these data and the observations for the YxiN_K52Q mutant in the current study one can conclude that YxiN adopts the closed conformation upon binding of ATP and RNA, remains closed during product formation and opens upon product release.

Future work could further dissect the catalytic cycle. It appears that after binding of RNA and ATP an activation step takes place that allows for RNA unwinding and ATP hydrolysis. In the motif I mutant this activation does not occur rationalising why this mutant is unwinding deficient. Since the global conformation of the wild-type and the motif I mutant are similar the activation step might constitute a local rearrangement which should be addressed by future studies. For this purpose, the substrate binding sites or the nucleotide (similarly as done before: Henn *et al.*, 2008) could be utilised as probes. One might for instance compare the corresponding signals

from experiments with the wild-type enzyme and the motif I mutant to reveal whether the mutant is impaired with respect to this local rearrangement.

It was mentioned in the previous section that the closed conformation may be regarded as the structural requirement for kinking and unwinding the RNA substrate. But a direct link between protein and RNA conformation cannot be drawn so far since no structural information about a bound RNA double strand is available. Therefore, future studies should also address the conformation of the RNA during the catalytic cycle.

A comprehensive dissection of the catalytic cycle including the information about the RNA conformation could then explain why the motif I mutant in the current study was not able unwind although it adopted the closed conformation efficiently.

4.3 YxiN can be Immobilized for smFRET Studies on a TIR-microscope

TIR-microscopy constitutes a complementary technique to confocal microscopy in smFRET studies as it extends the observation time to several seconds. Employing TIR-FRET might allow for the investigation of the transition between different conformers of the DEAD box protein YxiN. From TIR-FRET data dwell times for observed states and the kinetics of the opening and closing of YxiN could be derived.

In TIR-FRET only molecules close to the slide-buffer interface are detected. Therefore the molecules of interest usually need to be immobilized. The immobilization procedure should be specific and the molecules should not orient randomly. In addition, the molecules must not be impaired with respect to their activity and their binding characteristics to ligands or interaction partners. For example, an immobilized enzyme should still be able to bind its substrate.

In the current project a TIR-FRET sample preparation approach was designed that allows for an oriented immobilization of the DEAD box protein YxiN. The employed strategy was inspired by the work of Lesaicherre and colleagues (Lesaicherre *et al*., 2002). The authors designed an immobilization procedure for protein microarrays. Therefore they produced the target protein as an intein fusion (as in figure 1.3A). After binding the fusion protein to a chromatography column they incubated the protein with a cysteine-biotin moiety. Thereby the biotin attached to the protein C-terminus via the EPL mechanism. The biotinylated protein could be eluted and spotted onto an avidin-functionalized slide.

The approach in the current project also employed intein-mediated functionalization of the target protein YxiN. But before attaching the cysteine-biotin, the YxiN construct was isolated with a C-terminal thioester. It was subsequently labelled with donor and acceptor fluorophores. Afterwards a C-terminal biotin was attached in an EPL reaction. This chronological order ensures that the thiol of the cysteine-biotin moiety is not conjugated to a fluorophore. For immobilization, the construct was applied to a PEGylated glass slide that was functionalized with streptavidin. It was observed that the fluorophore labelled protein bound specifically onto the surface. This provides the basis for prospective TIR-FRET studies of donor-acceptor labelled and immobilized YxiN.

The correspondingly prepared YxiN construct was analysed for its activity i.e. the propensity to undergo a conformational change. SmFRET experiments on a confocal microscope showed that the protein adopts a closed conformation in the presence of ADPNP and RNA similar to the wild-type. But so far it cannot be judged whether the immobilization procedure affects the enzyme. This has to await a comprehensive set of TIR-FRET experiments with YxiN in the presence of its substrates. In case the immobilization procedure inactivates YxiN an alternative approach might be developed. Starting from the presented strategy one could vary the length or

constitution of the linker between protein and surface. Also a principally different approach might be useful. Various alternatives can be found in the literature (for example reviewed in Roy *et al*., 2008). For nucleic acid binding proteins it is often decided to immobilize their DNA or RNA substrate. Here, the nucleic acid needs to be biotinylated and is attached to a streptavidin-coated surface. The protein will come into view after binding to the DNA/RNA. Such an approach was for instance used in a study on the Rep DNA helicase that studied the translocation of this helicase on its substrate (Myong *et al*., 2005). In the assay the DNA was biotinylated and labelled with one fluorophore and the protein was labelled with a second fluorophore. This did not allow for the study of the protein conformation. But the principle can also be applied for the study of the conformational change of YxiN. One needs to immobilize its RNA substrate and will detect the labelled protein upon RNA binding.

The latter strategy does not only constitute an alternative it also is a complementary approach to the established immobilization of biotinylated YxiN. Ideally, one should study the behaviour of biotinylated YxiN that is directly immobilized and additionally of YxiN binding to the immobilized RNA substrate. The comparison of the obtained results could help to exclude artefacts caused by the respective immobilization procedure.

4.4 Site-specific Labelling of YxiN for FRET Experiments

FRET experiments require the attachment of donor and acceptor fluorophore at the desired positions in a biomolecule. In the case of proteins, usually maleimide-functionalized fluorophores are conjugated to cysteines (the least common amino acid in proteins). Preferably, one would generate homogenously labelled proteins with the donor and acceptor at defined positions. This can be achieved for small proteins that are obtained via chemical synthesis (e.g. Deniz *et al.*, 2000). Peptide synthesis, yet, has a size limitation and large proteins are hardly obtainable. Instead, larger proteins are widely produced in bacteria and are frequently labelled in a random fashion at the desired cysteines for FRET experiments. This procedure causes large sample heterogeneity. To decrease this heterogeneity the removal of unlabelled and single-labelled species was for instance attempted by hydrophobic interaction chromatography for the helicase YxiN (Theissen, 2006). An alternative approach suggests the purification of the donor-acceptor labelled species via affine RNA molecules (Hilbert, 2009). But in general the chromatographic separation turned out to be difficult and has to be optimized for the respective protein and the used fluorophores. Therefore an alternative to statistic labelling which initially causes less heterogeneity in the sample, would be beneficial.

The semi-synthetic approach for site-specific labelling of the DEAD box protein YxiN in the current project constitutes one alternative. Here, YxiN was purified in two separate fragments which each were labelled with donor or acceptor. The specifically labelled fragments were subsequently ligated in an EPL reaction. The approach based on a study which showed that the reconstitution of an active YxiN from two separate fragments is possible (Karow *et al.*, 2007). But in contrast to the ligation of the unlabelled YxiN fragments in the latter study, the reaction with the labelled fragments proceeded with low yield. The desired full-length product with donor and acceptor fluorophore at defined positions could not be generated in preparative scale. A major drawback in the experiments was the instability of the (labelled) N-terminal fragment of YxiN. However, the strategy might be successful for a different protein.

The approach for YxiN was modified. Namely, the donor labelled C-terminal fragment was ligated to an unlabelled N-terminal fragment. Subsequent to the ligation the full-length protein was labelled with acceptor dye. As a result a YxiN construct was generated which carried the donor fluorophore at a defined internal cysteine. The acceptor could potentially be attached to two different cysteines. The strategy was thus coined semi-site-specific labelling.

The obtained YxiN construct displayed the same global conformation as the statistically labelled YxiN. The smFRET histograms revealed that no donor-donor species were present in the YxiN sample obtained via semi-site-specific labelling. Therefore the heterogeneity in this sample was indeed reduced in comparison with the statistically labelled sample.

The EPL reaction requires the presence of an unmodified N-terminal cysteine at the C-terminal fragment. However, when the C-terminal fragment is labelled at an internal cysteine the N-terminal cysteine might also be modified. Such a double-labelled fragment will not be ligated in the EPL reaction. This side-reaction limited the yield of the employed strategy. In future experiments it might be overcome by reversible protection of the N-terminal cysteine during labelling. N-terminal cysteines can be specifically protected with ketones or aldehydes forming a thiazolidine structure. Upon treatment with methoxylamine and a reducing agent the native cysteine can be recovered (Villain *et al.*, 2001). Therefore, a reversible protection of the C-terminal EPL fragment might be useful for proteins that can withstand the procedure.

Alternatively, the EPL reaction could be replaced by ligation of the protein fragments via *trans*-splicing (figure 4.2B). Here, the ligation reaction is catalysed by an intein (ref. introduction 1.2.1). The reaction does not require the presence of a N-terminal cysteine at the C-terminal fragment, serines also work (Wu *et al.*, 1998; Southworth *et al.*, 1998). Figure 4.2 illustrates the principle of *trans*-splicing in comparison with EPL. The reactants can be labelled with fluorophores at internal cysteines before ligation. In *trans*-splicing theses cysteines exhibit orthogonal reactivity to the serine at the C-terminal fragment and the catalytic serine of the intein. Thus, the labelling reaction does not compete with the *trans*-splicing reaction as it happens in the EPL approach.

For *trans*-splicing the fragments of a target protein are fused to a split intein. The intein will only catalyse the splicing reaction upon mixing and reassembly of the intein halves. The fusion of YxiN fragments to a split *Ssp* DnaB intein was already performed in the author's group. However, one fusion protein exhibited very low solubility and had to be purified under denaturing conditions. So far, the refolding to an active intein was not successful. In future work the use of a different split intein might be worth considering. For example, the *Psp*-GDB Pol intein was employed in a recent study achieving site-specific labelling on a cysteine and subsequent *trans*-splicing of OmpF (Brenzel *et al.*, 2009).

In summary, this study presents an alternative to statistic labelling for large proteins. However in the developed strategy fluorophore labelling competes with ligation of the protein fragments thus limiting the yield of labelled full-length protein. The yield could be increased if a reversible protection step was embedded in the procedure. Alternatively, the ligation of labelled fragments of the target protein via *trans*-splicing could be tested.
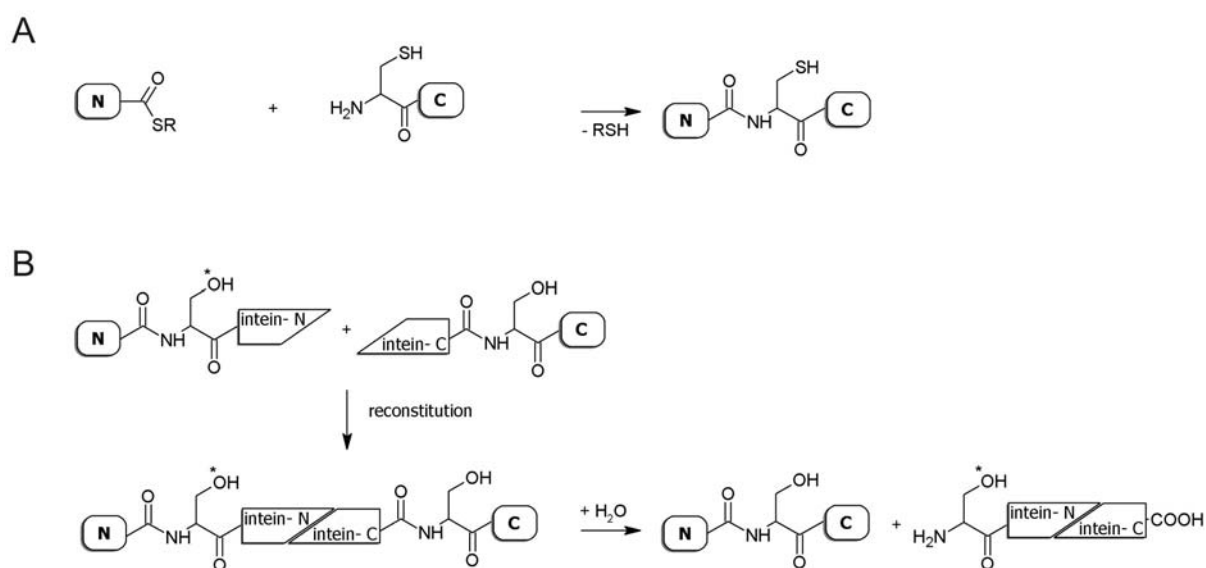
Figure 4.2
Protein ligation techniques. N and C denote the N and C-terminal fragment of the target protein respectively. Both fragments might be labelled before the ligation reaction e.g. for site-specific fluorophore labelling.
A) For expressed protein ligation the N-terminal fragment must harbour a C-terminal thioester and the C-terminal fragment must exhibit a N-terminal cysteine. For the ligation reaction to occur, the thiol of this N-terminal cysteine must not be modified e.g. by fluorophore attachment.
B) For *trans*-splicing the fragments of the target protein need to be fused to a split intein. Intein-N and intein-C denote the N-terminal and C-terminal fragment of the intein respectively. Few split inteins have been discovered that reconstitute to their active structure upon mixing. The reconstituted intein can then catalyse the splicing reaction i.e. ligation of the fragments of the target protein and excision of the intein. When *trans*-splicing shall be used in combination with fluorophore labelling at internal cysteines, one should choose an intein with a catalytic serine at the N-terminus (*). The C-terminal fragment should exhibit a N-terminal serine (threonine might also work). Upon labelling of internal cysteines with maleimide coupled fluorophores, the hydroxyl of serine (or threonine) is not modified. Therefore the fluorophore labelling of the target protein will in principle not compete with the splicing reaction.

5 References

Adam, H. (1962) Adenosin-5'Diphosphat und Adenosin-5'Monophosphat . In H. Bergmeyer (ed.), *Methoden der Enzymatischen Analyse*. Verlag Chemie, Weinheim, 573-577.

Alexandrov, K., Heinemann, I., Durek, T., Sidorovitch, V., Goody, R.S. & Waldmann, H. (2002) Intein-mediated synthesis of geranylgeranylated Rab7 protein *in vitro*. *J Am Chem Soc* **124** 5648-5649.

Altman, S. & Guerrier-Takada, C. (1986) M1 RNA, the RNA subunit of *Escherichia coli* ribonuclease P, can undergo a pH-sensitive conformational change. *Biochemistry* **25** 1205-1208.

Andersen, C.B.F., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L.P., Pedersen, J.S., Séraphin, B., Le Hir, H. & Andersen, G.R. (2006) Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* **313** 1968-1972.

Aregger, R. & Klostermeier, D. (2009) The DEAD box helicase YxiN maintains a closed conformation during ATP hydrolysis. *Biochemistry* **48** 10679-10681.

Arnold, U., Hinderaker, M.P., Nilsson, B.L., Huck, B.R., Gellman, S.H. & Raines, R.T. (2002) Protein prosthesis: a semisynthetic enzyme with a beta-peptide reverse turn. *J Am Chem Soc* **124** 8522-8523.

Aubert, M., Scott, J.F., Reynier, M. & Monier, R. (1968) Rearrangement of the conformation of *Escherichia coli* 5S RNA. *Proc Natl Acad Sci U S A* **61** 292-299.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. & Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2** 2006.0008.

Banroques, J., Doère, M., Dreyfus, M., Linder, P. & Tanner, N.K. (2010) Motif III in Superfamily 2 "Helicases" Helps Convert the Binding Energy of ATP into a High-Affinity RNA Binding Site in the Yeast DEAD-Box Protein Ded1. *J Mol Biol* **396** 949-966.

Bergmeyer, H.U. (1975) New values for the molar extinction coefficients of NADH and NADPH for the use in routine laboratories. *Z Klin Chem Klin Biochem* **13** 507-508.

Brenzel, S., Cebi, M., Reiss, P., Koert, U. & Mootz, H.D. (2009) Expanding the scope of protein *trans*-splicing to fragment ligation of an integral membrane protein: towards modulation of porin-based ion channels by chemical modification. *Chembiochem* **10** 983-986.

Camarero, J.A., Shekhtman, A., Campbell, E.A., Chlenov, M., Gruber, T.M., Bryant, D.A., Darst, S.A., Cowburn, D. & Muir, T.W. (2002) Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proc Natl Acad Sci U S A* **99** 8536-8541.

Caruthers, J.M., Hu, Y. & McKay, D.B. (2006) Structure of the second domain of the *Bacillus subtilis* DEAD-box RNA helicase YxiN. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62** 1191-1195.

Caruthers, J.M., Johnson, E.R. & McKay, D.B. (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc Natl Acad Sci U S A* **97** 13080-13085.

Cheng, Z., Coller, J., Parker, R. & Song, H. (2005) Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* **11** 1258-1270.

Coetzee, T., Herschlag, D. & Belfort, M. (1994) *Escherichia coli* proteins, including ribosomal protein S12, facilitate *in vitro* splicing of phage T4 introns by acting as RNA chaperones. *Genes Dev* **8** 1575-1588.

Cordin, O., Banroques, J., Tanner, N.K. & Linder, P. (2006) The DEAD-box protein family of RNA helicases. *Gene* **367** 17-37.

Cordin, O., Tanner, N.K., Doère, M., Linder, P. & Banroques, J. (2004) The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J* **23** 2478-2487.

Cotton, G.J. & Muir, T.W. (2000) Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed protein ligation. *Chem Biol* **7** 253-261.

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

David, R., Richter, M.P.O. & Beck-Sickinger, A.G. (2004) Expressed protein ligation. Method and applications. *Eur J Biochem* **271** 663-677.

Dawson, P.E., Churchill, M.J., Ghadiri, M.R. & Kent, S.B.H. (1997) Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives. *Journal of the American Chemical Society* **119** 4325-4329.

Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B. (1994) Synthesis of proteins by native chemical ligation. *Science* **266** 776-779.

Del Campo, M. & Lambowitz, A.M. (2009) Structure of the Yeast DEAD box protein Mss116p reveals two wedges that crimp RNA. *Mol Cell* **35** 598-609.

Deniz, A.A., Laurence, T.A., Beligere, G.S., Dahan, M., Martin, A.B., Chemla, D.S., Dawson, P.E., Schultz, P.G. & Weiss, S. (2000) Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc Natl Acad Sci U S A* **97** 5179-5184.

Erlanson, D.A., Chytil, M. & Verdine, G.L. (1996) The leucine zipper domain controls the orientation of AP-1 in the NFAT.AP-1.DNA complex. *Chem Biol* **3** 981-991.

Evans, T. & Xu, M. (2002) Mechanistic and kinetic considerations of protein splicing. *Chem Rev* **102** 4869-4883.

Evans, T.C.J., Benner, J. & Xu, M.Q. (1999) The *in vitro* ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *J Biol Chem* **274** 3923-3926.

Fan, J., Cheng, Z., Zhang, J., Noble, C., Zhou, Z., Song, H. & Yang, D. (2009) Solution and crystal structures of mRNA exporter Dbp5p and its interaction with nucleotides. *J Mol Biol* **388** 1-10.

Förster, T. (1946) Energiewanderung und Fluoreszenz. *Naturwissenschaften* **33** 166-175.

Gartland, W.J. & Sueoka, N. (1966) Two interconvertible forms of tryptophanyl sRNA in *E. coli*. *Proc Natl Acad Sci U S A* **55** 948-956.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. & Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook* 571-607.

Halls, C., Mohr, S., Del Campo, M., Yang, Q., Jankowsky, E. & Lambowitz, A.M. (2007) Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. *J Mol Biol* **365** 835-855.

Henn, A., Cao, W., Hackney, D.D. & De La Cruz, E.M. (2008) The ATPase cycle mechanism of the DEAD-box rRNA helicase, DbpA. *J Mol Biol* **377** 193-205.

Henn, A., Cao, W., Licciardello, N., Heitkamp, S.E., Hackney, D.D. & De La Cruz, E.M. (2010) Pathway of ATP utilization and duplex rRNA unwinding by the DEAD-box helicase, DbpA. *Proc Natl Acad Sci U S A* **107** 4046-4050.

Henn, A., Shi, S., Zarivach, R., Ben-Zeev, E. & Sagi, I. (2002) The RNA helicase DbpA exhibits a markedly different conformation in the ADP-bound state when compared with the ATP- or RNA-bound states. *J Biol Chem* **277** 46559-46565.

Herschlag, D. (1995) RNA chaperones and the RNA folding problem. *J Biol Chem* **270** 20871-20874.

Hilbert, M. (2009) Einfluss der elF4A bindenden Domäne aus elF4G auf Konformation und Aktivität der minimalen DEAD Box Helikase elF4A aus *Saccharomyces cerevisiae*. PhD thesis University of Basel

Hiratsuka, T. (1983) New ribose-modified fluorescent analogs of adenine and guanine nucleotides available as substrates for various enzymes. *Biochimica et Biophysica Acta* **742** 496-508.

Huse, M., Muir, T.W., Xu, L., Chen, Y.G., Kuriyan, J. & Massagué, J. (2001) The TGF beta receptor activation process: an inhibitor- to substrate-binding switch. *Mol Cell* **8** 671-682.

Iost, I. & Dreyfus, M. (2006) DEAD-box RNA helicases in *Escherichia coli*. *Nucleic Acids Res* **34** 4189-4197.

Iost, I., Dreyfus, M. & Linder, P. (1999) Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. *J Biol Chem* **274** 17677-17683.

Iwai, H. & Plückthun, A. (1999) Circular beta-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett* **459** 166-172.

Jankowsky, E., Fairman, M.E. & Yang, Q. (2005) RNA helicases: versatile ATP-driven nanomotors. *J Nanosci Nanotechnol* **5** 1983-1989.

Johnson, E.C.B. & Kent, S.B.H. (2006) Insights into the mechanism and catalysis of the native chemical ligation reaction. *J Am Chem Soc* **128** 6640-6646.

Karginov, F.V., Caruthers, J.M., Hu, Y., McKay, D.B. & Uhlenbeck, O.C. (2005) YxiN is a modular protein combining a DEx(D/H) core and a specific RNA-binding domain. *J Biol Chem* **280** 35499-35505.

Karow, A.R., Theissen, B. & Klostermeier, D. (2007) Authentic interdomain communication in an RNA helicase reconstituted by expressed protein ligation of two helicase domains. *FEBS J* **274** 463-473.

Karpel, R.L. & Burchard, A.C. (1980) Physical studies of the interaction of a calf thymus helix-destablizing protein with nucleic acids. *Biochemistry* **19** 4674-4682.

Karpel, R.L., Swistel, D.G., Miller, N.S., Geroch, M.E., Lu, C. & Fresco, J.R. (1975) Acceleration of RNA renaturation by nucleic acid unwinding proteins. *Brookhaven Symp Biol* 165-174.

Klostermeier, D. & Rudolph, M.G. (2009). A novel dimerization motif in the C-terminal domain of the *Thermus thermophilus* DEAD box helicase Hera confers substantial flexibility. *Nucleic Acids Res* **37** 421-430.

Kossen, K. & Uhlenbeck, O.C. (1999) Cloning and biochemical characterization of *Bacillus subtilis* YxiN, a DEAD protein specifically activated by 23S rRNA: delineation of a novel sub-family of bacterial DEAD proteins. *Nucleic Acids Res* **27** 3811-3820.

Kossen, K., Karginov, F.V. & Uhlenbeck, O.C. (2002) The carboxy-terminal domain of the DExDH protein YxiN is sufficient to confer specificity for 23S rRNA. *J Mol Biol* **324** 625-636.

Kunst, A., Draeger, B. & Ziegenhorn, J. (1984) D-Glucose. In H. Bergmeyer, J. Bergmeyer and U. Grassl (ed.), *Methods of Enzymatic Analysis*. Verlag Chemie Weinheim, 163-172.

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

Lakowicz, J. (2006) Principles of Fluorescence Spectroscopy. Springer

Lesaicherre, M., Lue, R.Y.P., Chen, G.Y.J., Zhu, Q. & Yao, S.Q. (2002) Intein-Mediated Biotinylation of Proteins and Its Application in a Protein Microarray. *Journal of the American Chemical Society* **124** 8768-8769.

Lindahl, T., Adams, A. & Fresco, J.R. (1966) Renaturation of transfer ribonucleic acids through site binding of magnesium. *Proc Natl Acad Sci U S A* **55** 941-948.

Linden, M.H., Hartmann, R.K. & Klostermeier, D. (2008) The putative RNase P motif in the DEAD box helicase Hera is dispensable for efficient interaction with RNA and helicase activity. *Nucleic Acids Res* **36** 5800-5811.

Linder, P. (2006) Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res* **34** 4168-4180.

Lorsch, J.R. & Herschlag, D. (1998a) The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry* **37** 2180-2193.

Lorsch, J.R. & Herschlag, D. (1998b) The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochemistry* **37** 2194-2206.

Macmillan, D. & Bertozzi, C. (2000) New directions in glycoprotein engineering. *Tetrahedron* **56** 9515-9525.

Madore, E., Florentz, C., Giegé, R. & Lapointe, J. (1999) Magnesium-dependent alternative foldings of active and inactive *Escherichia coli* tRNA(Glu) revealed by chemical probing. *Nucleic Acids Res* **27** 3583-3588.

Magde, D., Wong, R. & Seybold, P.G. (2002) Fluorescence quantum yields and their relation to lifetimes of rhodamine 6G and fluorescein in nine solvents: improved absolute standards for quantum yields. *Photochem Photobiol* **75** 327-334.

Maris, C., Dominguez, C. & Allain, F.H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272** 2118-2131.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R. & Ebright, R.H. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108** 599-614.

Mohr, G., Del Campo, M., Mohr, S., Yang, Q., Jia, H., Jankowsky, E. & Lambowitz, A.M. (2008) Function of the C-terminal domain of the DEAD-box protein Mss116p analyzed *in vivo* and *in vitro*. *J Mol Biol* **375** 1344-1364.

Moore, P.B. & Steitz, T.A. (2002) The involvement of RNA in ribosome function. *Nature* **418** 229-235.

Muir, T.W., Sondhi, D. & Cole, P.A. (1998) Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A* **95** 6705-6710.

Mukhopadhyay, J., Kapanidis, A.N., Mekler, V., Kortkhonjia, E., Ebright, Y.W. & Ebright, R.H. (2001) Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* **106** 453-463.

Myong, S., Rasnik, I., Joo, C., Lohman, T.M. & Ha, T. (2005) Repetitive shuttling of a motor protein on DNA. *Nature* **437** 1321-1325.

Oberer, M., Marintchev, A. & Wagner, G. (2005) Structural basis for the enhancement of eIF4A helicase activity by eIF4G. *Genes Dev* **19** 2212-2223.

Parker, C. & Rees, W. (1960) Correction of Fluorescence Spectra and Measurement of Fluorescence Quantum Efficiency. *Analyst* **85** 587-600.

Pause, A. & Sonenberg, N. (1992) Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J* **11** 2643-2654.

Pause, A., Méthot, N. & Sonenberg, N. (1993) The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. *Mol Cell Biol* **13** 6789-6798.

Polach, K.J. & Uhlenbeck, O.C. (2002) Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. *Biochemistry* **41** 3693-3702.

Rasnik, I., McKinney, S.A. & Ha, T. (2005) Surfaces and orientations: much to FRET about?. *Acc Chem Res* **38** 542-548.

Riesner, D., Henco, K., Rokohl, U., Klotz, G., Kleinschmidt, A.K., Domdey, H., Jank, P., Gross, H.J. & Sänger, H.L. (1979) Structure and structure formation of viroids. *J Mol Biol* **133** 85-115.

Rocak, S., Emery, B., Tanner, N.K. & Linder, P. (2005) Characterization of the ATPase and unwinding activities of the yeast DEAD-box protein Has1p and the analysis of the roles of the conserved motifs. *Nucleic Acids Res* **33** 999-1009.

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

Rozen, F., Pelletier, J., Trachsel, H. & Sonenberg, N. (1989) A lysine substitution in the ATP-binding site of eucaryotic initiation factor 4A abrogates nucleotide-binding activity. *Mol Cell Biol* **9** 4061-4063.

Rudolph, M.G. & Klostermeier, D. (2009) The *Thermus thermophilus* DEAD box helicase Hera contains a modified RNA recognition motif domain loosely connected to the helicase core. *RNA* **15** 1993-2001.

Schütz, P., Bumann, M., Oberholzer, A.E., Bieniossek, C., Trachsel, H., Altmann, M. & Baumann, U. (2008) Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci U S A* **105** 9564-9569.

Segel, I.H. (1993) *Enzyme Kinetics*. John Wiley & Sons, Inc.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S. & Yokoyama, S. (2006) Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell* **125** 287-300.

Sharpe Elles, L.M., Sykes, M.T., Williamson, J.R. & Uhlenbeck, O.C. (2009) A dominant negative mutant of the *E. coli* RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res* **37** 6503-6514.

Sharpe Elles, L.M. & Uhlenbeck, O.C. (2008) Mutation of the arginine finger in the active site of *Escherichia coli* DbpA abolishes ATPase and helicase activity and confers a dominant slow growth phenotype. *Nucleic Acids Res* **36** 41-50.

Solem, A., Zingler, N. & Pyle, A.M. (2006) A DEAD protein that activates intron self-splicing without unwinding RNA. *Mol Cell* **24** 611-617.

Southworth, M.W., Adam, E., Panne, D., Byer, R., Kautz, R. & Perler, F.B. (1998) Control of protein splicing by intein fragment reassembly. *EMBO J* **17** 918-926.

Spreitler, F. (2006). Einzelmolekül-FRET an Biomolekülen. Diplomarbeit Universität Bayreuth

Story, R.M., Li, H. & Abelson, J.N. (2001) Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. *Proc Natl Acad Sci U S A* **98** 1465-1470.

Stryer, L. & Haugland, R.P. (1967) Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci U S A* **58** 719-726.

Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41** 207-234.

Talavera, M.A., Matthews, E.E., Eliason, W.K., Sagi, I., Wang, J., Henn, A. & De La Cruz, E.M. (2006) Hydrodynamic characterization of the DEAD-box RNA helicase DbpA. *J Mol Biol* **355** 697-707.

Tam, J.P., Lu, Y.A., Liu, C.F. & Shao, J. (1995) Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc Natl Acad Sci U S A* **92** 12485-12489.

Theissen, B. (2006) Conformational changes in the catalytic cycle of the RNA-helicase YxiN - fluorescence resonance energy transfer in single molecules. Dissertation Universität Bayreuth

Valiyaveetil, F.I., Sekedat, M., Mackinnon, R. & Muir, T.W. (2004) Glycine as a D-amino acid surrogate in the K(+)-selectivity filter. *Proc Natl Acad Sci U S A* **101** 17045-17049.

Villain, M., Vizzavona, J. & Rose, K. (2001) Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides. *Chem Biol* **8** 673-679.

Wang, S., Hu, Y., Overgaard, M.T., Karginov, F.V., Uhlenbeck, O.C. & McKay, D.B. (2006) The domain of the *Bacillus subtilis* DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold. *RNA* **12** 959-967.

Wang, S., Overgaard, M.T., Hu, Y. & McKay, D.B. (2008) The *Bacillus subtilis* RNA helicase YxiN is distended in solution. *Biophys J* **94** L01-3.

Wieland, T., Bokelmann, E., Bauer, L., Lang, H. & Lau, H. (1953) Über Peptidsynthesen. 8. Bildung von S-haltigen Peptiden durch intramolekulare Wanderung von Aminoacylresten. *Ann Chem* **583** 129-149.

Wu, H., Xu, M.Q. & Liu, X.Q. (1998) Protein *trans*-splicing and functional mini-inteins of a cyanobacterial dnaB intein. *Biochim Biophys Acta* **1387** 422-432.

Xu, M.Q. & Evans, T.C.J. (2001) Intein-mediated ligation and cyclization of expressed proteins. *Methods* **24** 257-277.

Xu, R., Ayers, B., Cowburn, D. & Muir, T.W. (1999) Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc Natl Acad Sci U S A* **96** 388-393.

Yang, Q. & Jankowsky, E. (2005) ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry* **44** 13591-13601.

6 List of Abbreviations

Amino acids are abbreviated with the common three letter or single letter code.

A_x	absorption at $\lambda = $ x nm
A488	Alexa Fluor 488
A546	Alexa Fluor 546
A555	Alexa Fluor 555
ACN	acetonitrile
ADP	adenosine 5'-diphosphate
ADPNP	adenosine 5'-(β,γ-imido)triphosphate
APD	avalanche photodiode
APS	ammonium peroxodisulfate
ATP	adenosine 5'-triphosphate
bp	base pairs
B. subtilis	*Bacillus subtilis*
BSA	bovine serum albumine
CCD	charge-coupled device
CTP	cytidine 5'-triphosphate
dATP	2'-deoxyadenosine 5'-triphosphate
DCM	dichloromethane
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
DMF	*N,N*-dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ds	double stranded
DTT	1,4-dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
E. coli	*Escherichia coli*
EDTA	*N,N,N',N'*-ethylenediamine-tetraacetic disodium salt
EPL	expressed protein ligation
ESI	electrospray ionization
FRET	fluorescence resonance energy transfer
GTP	guanosine 5'-triphosphate
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HOBt	*N*-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilo base pairs
LB	Luria-Bertani broth
LC	liquid chromatography
MALDI	matrix assisted laser desorption/ionization
mant-	2'/3'-*O*-(*N*-methyl-anthraniloyl)-
MESNa	2-mercaptoethanesulfonic acid sodium salt
Mj	*Methanococcus jannaschii*
MOPS	3-(*N*-morpholino)propanesulfonic acid
MS	mass spectrometry
Mxe	*Mycobacterium xenopi*
NADH	nicotinamide adenine dinucleotide (reduced)
OD_{600}	optical density (measurement at $\lambda = $ 600 nm)
PAGE	polyacrylamide gel electrophoresis
PEP	phosphoenolpyruvate
PCR	polymerase chain reaction
Psp	*Pyrococcus species*
RBD	RNA binding domain
RNA	ribonucleic acid

RP	reversed phase
rpm	revolutions per minute
RRM	RNA recognition motif
rRNA	ribosomal ribonucleic acid
S. cerevisiae	*Saccharomyces cerevisiae*
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
sm	single molecule
SMD	single molecule detection
ss	single stranded
Ssp	*Synechocystis species*
T. thermophilus	*Thermus thermophilus*
TBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate
TCEP	tris-(2-carboxyethyl)phosphine hydrochloride
TEMED	*N,N,N',N'*-tetramethylethylenediamine
TFA	trifluoroacetic acid
TIR	total internal reflection
TLC	thin layer chromatography
TMR	tetramethylrhodamine
Tris	tris-(hydroxymethyl)aminomethane
UTP	uridine 5'-triphosphate
YxiN*	YxiN_C61S_C267S
YxiN#	YxiN_C61A_C267A
YxiN'	YxiN_C61A_C267A_A115C_S229C

7 Acknowledgement

This work was carried out in the laboratories of Prof. Dagmar Klostermeier at the Biozentrum of the University of Basel and financially supported by the VolkswagenStiftung and the SNF.

I would like to thank

most sincerely Prof. Dagmar Klostermeier for providing me with this colourful project and for sharing insights into some "bench-beyonds". Thank you for always being fair and for eloquently putting our results into words a lot of times. Thank you for your ideas but also the freedom and faith I experienced.

my family for their life-long support, for cheering me up many times and for sending numerous *Carepakete* in the last years.

the members of the Klostermeier group for their support, co-working and the (other) fun parts. In particular, thanks to Ines (Hermes) Hertel for a lot of help in the purification, mutagenesis and IVT business. And thank you for your friendship. Thank you Airat Gubaev and Manuel Hilbert for enormous help, constructive criticism and helpful suggestions. I very much appreciate the work of Bettina Theißen that provided a solid basis for this project. Thanks, Andreas Schmidt for donating the Super-HUBA ... water – every mutagenesis worked! Thank you, Martin Linden for switching on the rather positive blues in hours of darkness - thanks for your friendship. Merci, Bruno Stephanidis for proofreading and for glistening hours in the dark cold (cold dark?). Many thanks to Miriam Steiner for sharing your Postdoc wisdom and for one very nice hen's year.

all members of the biophysical chemistry department for the nice atmosphere here at the 6th floor.

Matthias Zwick for proofreading, reliable co-running of the Coffee Corner and for always reminding me to worry rather too late than too early. Hold on, Capun yeah!

Rita Müller, for the advice in the biotin detection business and for the co-western blotting. I enjoyed your company in and outside the lab a lot.

Andrè Ziegler for help with the synthesis of the Cys-Biotin.

Dirk de Bruyn for purifying the Cys-Biotin with me and thanks to Prof. Wolfgang Meier for providing the corresponding infrastructure.

Bea Lang for clearing up our mess and always spreading good cheer.

Leo Faletti and Simon Sarner for keeping everything running and for lending me essential tools, merci.

8 Curriculum Vitae

- not included in online version -