Biophysical studies of the translation-regulating *add* adenine riboswitch from *Vibrio vulnificus*

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie
der Goethe-Universität
in Frankfurt am Main

von

Sven Warhaut

aus Darmstadt

Frankfurt am Main

2017

(D30)

Vom Fachbereich Biochemie, Chemie und Pharmazie der

Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Michael Karas

Gutachter: Prof. Dr. Harald Schwalbe

Prof. Dr. Mike Heilemann

Datum der Disputation: 8. Juni 2017

Table of content

Summary

Bacterial gene expression can be regulated at mRNA level by *cis*-acting mRNA elements termed riboswitches. Riboswitches adjust the expression level of their associated genes to the intracellular concentration of a specific ligand, typically a low-molecular weight metabolite. Riboswitches operate by conformational switching between a ligand-free and a ligand-bound state with different structures that either activate or inhibit gene expression. Since natural bacterial riboswitches regulate fundamental physiological pathways, they are potential drug targets for treating bacterial infections. As genetic switch modules, riboswitches further have potential applications in synthetic biology or molecular diagnostics.

This PhD thesis contributes to the molecular level understanding of full-length purine riboswitches. It presents biophysical investigations on the ligand-dependent folding of the full-length translation-regulating *add* adenine riboswitch from the gram-negative human pathogenic marine bacterium *Vibrio vulnificus* (Asw). Asw has the typical bipartite riboswitch architecture with a 5' ligand-sensing aptamer domain and a 3' regulatory domain termed expression platform. According to the working hypothesis, Asw employs a unique thermodynamically-controlled 3-state conformational switching mechanism between an apoB, an apoA and a holo conformation to regulate translation initiation in a temperature-compensated manner. The two apo conformations are the putative translation-OFF states and the holo conformation is the putative translation-ON state of Asw.

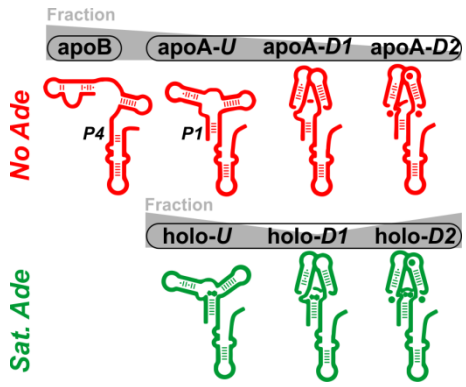

Figure 1: Conformational ensemble of ligand-free and ligand-bound ^{112}Asw

In the main project of this PhD thesis, which is presented in **Chapter 3**, single-molecule FRET (smFRET) experiments were established in collaboration with Martin Hengesbach, Philipp Höllthaler and Prof. Mike Heilemann, and an integrated nuclear magnetic resonance (NMR) and smFRET spectroscopic study of the full-length 112-nucleotide Asw (^{112}Asw) was performed. The adenine-dependent folding of ^{112}Asw was monitored at the level of base pairing interactions by NMR of the RNA imino protons, and at the level of three long-range intramolecular distances by smFRET of immobilized

molecules. This smFRET investigation of ^{112}Asw represents the first multivector fluorescence analysis of a purine riboswitch in the full-length sequence. The integrated NMR and smFRET spectroscopic study of ^{112}Asw yielded two major findings. First, NMR and smFRET both revealed that adenine binding to ^{112}Asw impedes apoB formation by stabilizing the apoA secondary structure in the holo conformation without modulating tertiary structural interactions between the two riboswitch domains. This finding highlights the central role of competitive P1 and P4 helix formation at the interface of the aptamer and the expression platform for switching the accessibility of the ribosome binding site of ^{112}Asw. Moreover, it strongly corroborates the hypothesis that purine riboswitches in general operate according to the key principle of a spatially decoupled secondary structural allosteric switch that proceeds without ligand-induced tertiary structural interactions between the aptamer domain and the expression platform. Second, it was uncovered by smFRET that the apoA and the holo conformation of ^{112}Asw do not adopt a single folding state at near-physiological Mg^{2+} concentration. Instead, apoA and holo exhibit a persistent dynamic equilibrium between substates with an undocked (U), a short-lived docked (D1; ~s) and a Mg^{2+}-bound long-lived docked (D2; ~10 s) aptamer kissing loop motif. In the holo conformation, the fractional population of the long-lived docked substate is ~2-fold increased compared to the apoA conformation, but undocked and docked substates are still comparably stable. The here described multiple folding states of the apoA and the holo conformation might have regulatory properties that are in between the apoB translation-OFF state and the holo-D2 translation-ON state.

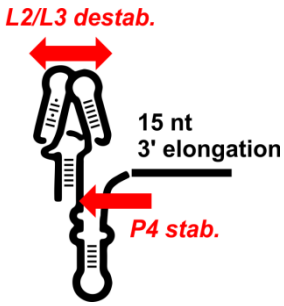

Figure 2: ^{127}Asw in the apoA conformation

In **Chapter 4**, an integrated NMR and smFRET analysis of 127-nucleotide Asw (^{127}Asw) is presented. Compared to ^{112}Asw, ^{127}Asw is 3'-elongated by 15 nucleotides of the adenosine deaminase encoding sequence of the *add* gene from *Vibrio vulnificus*. ^{127}Asw was chosen as mRNA template for future investigations of the interaction between Asw and the 30S ribosomal subunit. The NMR spectra of ^{127}Asw demonstrated that ^{127}Asw has the same overall secondary structure as ^{112}Asw. Like for ^{112}Asw, the combined NMR and smFRET analysis of ^{127}Asw showed that adenine binding impedes apoB formation and stabilizes a long-lived docked aptamer kissing loop fold. However, compared to ^{112}Asw, ^{127}Asw has a destabilized aptamer kissing loop motif

and a stabilized P4 helix in the expression platform. ^{127}Asw does not adopt the long-lived docked apoA-D2 substate at near-physiological Mg^{2+} concentration. And, unexpectedly, ^{127}Asw does not exhibit adenine-induced melting of the P4 helix, suggesting that its ribosome binding site is still masked in the holo conformation. It was thus hypothesized that the adenine-induced regulatory switch of Asw in a full-length protein coding mRNA chain occurs only after initial standby-binding of the 30S ribosome 3' to the riboswitch expression platform has eliminated effects of the protein coding sequence on the riboswitch fold.

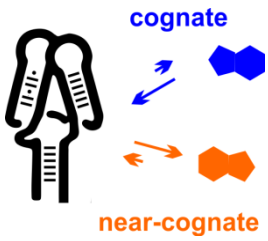

Figure 3: Cognate and near-cognate ligand binding to the purine riboswitch aptamer

In **Chapter 5**, ligand-observed studies of the transient encounter complex between Asw and the near-cognate ligand hypoxanthine are described. By competition binding WaterLOGSY NMR experiments with hypoxanthine and the adenine analogue 2,6-diaminopurine, it could be shown that hypoxanthine binds to the same binding site of ^{112}Asw as the cognate ligand adenine. The hypoxanthine binding constant measured with the WaterLOGSY method is in the low mM range (1.8 mM) and substantially exceeds the physiological hypoxanthine concentration in *E. coli* (~0.3 mM), thus ruling out that hypoxanthine binding can significantly impact the translational regulation of Asw *in vivo*. Additionally, Chapter 5 presents results obtained in collaboration with Albrecht Völklein, Henrik Gustmann and Prof. Wachtveitl. In this collaboration, the feasibility of Fourier transform infrared (FTIR) difference spectroscopy between isotope labelled and unlabelled ligand was tested for characterizing transient purine riboswitch aptamer complexes with near-cognate ligands and stable purine riboswitch aptamer complexes with cognate ligands. Preliminary FTIR difference spectra of ^{13}C,^{15}N-labelled and unlabelled hypoxanthine in complex with the *pbuE* adenine riboswitch aptamer and the *xpt* guanine riboswitch aptamer showed a pattern of multiple IR bands that appeared to be characteristic for the respective complex. However, due to spectral overlap, the FTIR difference spectra could not be properly evaluated.

Zusammenfassung in deutscher Sprache

Riboswitche sind Ligand bindende regulatorische mRNA Elemente die vorwiegend in Bakterien vorkommen. Sie sind meist in der 5’ untranslatierten Region bestimmter Gene kodiert und kontrollieren deren Expression *cis*-wirkend durch Ligand induzierte Faltung. Die spezifische Bindung eines Metaboliten ruft eine Strukturänderung des Riboswitches hervor, welche die Genexpression aktiviert oder inhibiert. Typischerweise kontrollieren Riboswitche entweder via Transkription oder via Translation die Expression von Proteinen welche an Stoffwechselwegen oder am Transport des Riboswitchliganden beteiligt sind. Das Expressionslevel dieser Proteine wird somit in Feedback zur zellulären Metabolit Konzentration reguliert. Riboswitche kontrollieren fundamentale Stoffwechselwege in Bakterien, kommen jedoch nicht in Säugetieren vor. Daher sind Riboswitche als potentielle Drugtargets für die Entwicklung neuer Antibiotika interessant. Darüber hinaus können Riboswitche als genetische Schalter in der Synthetischen Biologie oder als Biosensoren in der molekularen Diagnostik Anwendung finden.

In der vorliegenden Arbeit wurden biophysikalische Studien zur Ligand abhängigen Faltung des translationalen *add* Adenin-Riboswitchs (Asw) voller Länge aus dem humanpathogenen marinen Bakterium *Vibrio vulnificus* durchgeführt, um den regulatorischen Mechanismus dieses Riboswitchs auf molekularer Ebene zu charakterisieren.

Asw gehört zur Klasse der Purin-Riboswitche und besteht aus einer 5‘ Ligand bindenden Aptamerdomäne und einer 3‘ regulatorischen Expressionsplattform. Asw ist ein translationaler ON-switch, welcher durch Adenin Bindung die Expression von Adenosin-Desaminase aktiviert. Die Struktur und Ligand Bindung der Asw Aptamerdomäne ist in der Literatur bereits umfassend charakterisiert. Die Asw Aptamerdomäne besitzt die für Purin-Riboswitche typische Sekundärstrukur einer Drei-Wege-Kreuzung. Die Aptamer Tertiärstruktur ist im apo Zustand intrinsisch ungeordnet, und im holo Zustand hochgradig geordnet, wobei das gebundene Adenin Molekül im Kern des Aptamers eingekapselt ist und die apikalen Haarnadelschlaufen des Aptamers einen Schlaufenkontakt (Kissing-Loop Motiv) ausbilden, welcher durch Koordination von Mg^{2+}-Ionen stabilisiert ist. Kürzlich wurde der Asw voller Länge bestehend aus Aptamer und Expressionsplattform mittels Kernspinresonanzspektroskopie (NMR-Spektroskopie) charakterisiert und eine Arbeitshypothese für den regulatorischen Mechanismus dieses

Riboswitchs formuliert. Demnach ist Asw ein paradigmatischer 3-Zustand-Riboswich und reguliert die Translation des Adenosin-Desaminase Gens durch ein konformationelles Gleichgewicht zweier apo Konformationen, ApoB und ApoA, und einer Holo Konformation. ApoB besitzt keine Drei-Wege-Kreuzung Sekundärstruktur in der Aptamerdomäne und weist eine lange P4B Helix in der Expressionsplattform auf, welche die Ribosom Bindestelle inklusive AUG Start-Codon maskiert. ApoB ist daher ein bindungsinkompetenter Translation-OFF Zustand. ApoA zeichnet sich durch eine Drei-Wege-Kreuzung mit intrinsisch ungeordneter Tertiärstruktur ohne Kissing-Loop Motiv aus und besitzt eine kurze P4 Helix in der Expressionsplattform, welche die Shine-Dalgarno-Sequenz der Ribosom Bindestelle maskiert. Somit wurde ApoA als Adenin bindender Translation-OFF Zustand eingestuft. Die Holo Konformation ist der mutmaßliche Translation-ON Zustand des Riboswitchs. Sie weist eine Adenin gebundene Aptamerdomäne mit hochgradig geordneter Tertiärstruktur auf sowie eine Expressionsplattform mit geschmolzener P4 Helix und besitzt dadurch eine zugänglicher Ribosom-Bindestelle. Im Gegensatz zu einem 2-Zustandsmechanismus ermöglicht der 3-Zustandsmechanismus von Asw eine temperaturkompensierte Regulation der Translation.

In dieser Arbeit wurde der Asw voller Länge erstmals mit Einzelmolekül-Förster-Resonanzenergietransfer (smFRET) Spektroskopie charakterisiert. **Kapitel 3** beschreibt eine integrierte NMR und smFRET spektroskopische Untersuchung der Ligand-abhängigen Faltung des 112-Nukleotid langen Asw (^{112}Asw). Dabei wurden NMR spektroskopische Daten bezüglich der Basenpaarung der drei Konformationen von ^{112}Asw mit tertiärstrukturellen Informationen über intramolekulare Distanzen aus smFRET Experimenten komplementiert. Die vorhandene Zuordnung der NMR-Resonanzen der Imino-Gruppen von ^{112}Asw wurde genutzt, um die Adenin induzierte Umfaltung des Riboswitchs auf Imino-Reportersignalen für 25 Basenpaare auszuwerten. Im Gegensatz zu vorherigen Studien bei 10 °C wies die ApoA Konformation bei 25 °C in Gegenwart von 5 mM Mg^{2+} ein vorgefaltetes Aptamer Kissing-Loop Motiv auf und zeigte keine ApoA-spezifischen Imino-Signale für Basenpaare der Expressionsplattform. Die beobachtete Adenin induzierte Aptamer Faltung von ApoA zu Holo war eine vorwiegend lokale Faltung der Adenin Bindetasche. Die Ausbildung der Holo Konformation resultierte in einer deutlichen Depopulation der ApoB Konformation (von ~80% zu ~20% ApoB durch 1.1 eq Adenin) und im Aufschmelzen der P4 Helix in der Expressionsplattform. Ferner wurde gezeigt, dass sich neben der Adenin-Konzentration,

der Mg^{2+}-Ionen-Konzentration und der Temperatur, auch die Verschleppung von Li^{+}-Ionen während der Probenpräparation auf die Faltung von ^{112}Asw auswirken kann. Eine Li^{+}-Ionen-Konzentration um ~40 mM (bei einer Mg^{2+}-Ionenkonzentration von 5 mM) destabilisiert die Faltung der ApoA Aptamer Kissing-Loop Interaktion und stabilisiert Basenpaarung der Helix P4 in der Expressionsplattform. Die Adenin-Bindungsinkompetenz der ApoB Konformation wurde mittels einer ApoB-stabilisierten Mutante bestätigt. Die Mutante zeigte zudem, dass ApoB lediglich eine kurze P4 Helix aufweist und dass das AUG Start Codon im Gegensatz zum ursprünglich Modell keine stabile Basenpaarung ausbildet. smFRET Experimente an immobilisierten Molekülen wurden in Kollaboration mit Martin Hengesbach, Philipp Höllthaler und Prof. Mike Heilemann etabliert und an ^{112}Asw mit drei Fluorophor-Markierungsschemata durchgeführt: L2/P5, L3/P5 und L2/L3. Die Konstrukte L2/P5 und L3/P5 trugen Fluorophor Modifikationen an Aptamer und Expressionsplattform. Sie zeigten dass die beiden Riboswitch Domänen nicht über tertiärstrukturelle Wechselwirkungen miteinander interagieren und einen Ligand-unabhängigen Abstand zueinander aufweisen. Dies bestätigte einen zentralen Aspekt des auf NMR Basis aufgestellten Modells zur molekularen Funktionsweise von Asw, nämlich dass Asw die Zugänglichkeit der Ribosom Bindestelle auf Sekundärstrukturebene über eine kompetitive Ausbildung der Helices P1 und P4 moduliert. Mittels des L2/L3-markierten Konstrukts wurde die Faltung der Aptamer Kissing-Loop Interaktion in ^{112}Asw charakterisiert. Die beobachteten zwei FRET Populationen wurden über den Vergleich des Wildtyps mit Mutanten, in denen die ApoB bzw. die ApoA Konformation stabilisiert ist, zugeordnet. Saturierende Adenin-Konzentration verhinderte die Ausbildung der ApoB Konformation im Wildtyp, sodass das Kissing-Loop Faltungsgleichgewicht des Wildtyps dem der ApoA-stabilisierten Mutante glich. Die ApoA-stabilisierte Mutante wies sowohl ohne Adenin als auch in Gegenwart saturierender Adenin-Konzentration ein dynamisches Gleichgewicht zwischen Zuständen mit ungefalteter und gefalteter Aptamer Kissing-Loop Interaktion auf. Die zugrundeliegende Faltungsdynamik von ApoA und Holo wurde mittels Transition Occupancy Density Plots und Dwell-time Histogrammen analysiert. Hierzu wurden eigens verfasste MATLAB Skripte und die open-source Hidden Markov Modelling Software HaMMy verwendet. Auf Basis dieser Daten konnte ein erweitertes Modell für den molekularen Mechanismus von ^{112}Asw aufgestellt werden. Demnach populieren die ApoA und die Holo Konformation je 3 Faltungszustände: ApoA-U, ApoA-D1 und ApoA-D2, sowie Holo-U, Holo-D1 und Holo-D2. Im Zustand U ist das Aptamer Kissing-Loop

Motiv ungefaltet, im Zustand D1 ist es gefaltet und kurzlebig (~s) und im Zustand D2 gefaltet, Mg^{2+}-gebunden und langlebig (~10 s). Bei physiologischer Mg^{2+}-Ionen-Konzentration ist die Population der Zustände mit gefaltetem Kissing-Loop Motiv in Holo um Faktor 2 erhöht gegenüber ApoA. Dennoch weisen die Holo Zustände mit gefaltetem und ungefaltetem Kissing-Loop Motiv eine ähnliche Stabilität auf. Diese ausgeprägte strukturelle Heterogenität von ApoA und Holo ist aus NMR Spektren der Imino-Gruppen von Asw nicht ersichtlich und daher zuvor unbeachtet gewesen. Bei den verschiedenen Faltungszuständen von ApoA und Holo könnte es sich um intermediäre regulatorische Zustände zwischen dem ApoB Translation-OFF Zustand und dem Holo-D2 Translation-ON Zustand handeln.

In **Kapitel 4** wurde die Ligand abhängige Faltung von Asw in einer 127 Nukleotide langen Sequenz (^{127}Asw) mit NMR und smFRET Spektroskopie untersucht. ^{127}Asw ist gegenüber ^{112}Asw am 3‘ Ende um 15 kodierende Nukleotide des *add* Gens aus *Vibrio vulnificus* verlängert und soll als mRNA Templat für zukünftige Interaktionsstudien zwischen Asw und der 30S ribosomalen Untereinheit dienen. Die NMR Daten von ^{127}Asw beweisen, dass ^{127}Asw aufgrund der 3‘-Verlängerung keine neuen Sekundärstrukturelemente ausbildet und ein ähnliches Populationsverhältnis zwischen ApoB und ApoA wie ^{112}Asw aufweist. Im Vergleich zu ^{112}Asw zeigte ^{127}Asw eine ausgeprägtere Adenin Sensitivität der Imino-Signal Intensitäten in der Aptamerdomäne und kein Adenin induziertes Aufschmelzen der P4 Helix relativ zur P5 Helix. Die Aptamer Faltung von ^{127}Asw wirkt sich somit nicht signifikant auf die Zugänglichkeit der Shine-Dalgarno-Sequenz in der Expressionsplattform aus. In smFRET Experimenten mit L2/L3-markiertem ^{127}Asw war eine Destabilisierung des Aptamer Kissing-Loop Motivs in gegenüber ^{112}Asw beobachtbar. In Abwesenheit von Adenin lag ^{127}Asw bei physiologischer Mg^{2+}-Ionen-Konzentration nahezu vollständig in Faltungszuständen ohne Kissing-Loop Interaktion vor, und bei saturierender Adenin-Konzentration bildeten lediglich halb so viele Moleküle wie in ^{112}Asw das Kissing-Loop Motiv aus. Dass Adenin Bindung an ^{127}Asw nicht zum Aufschmelzen der P4 Helix führt, suggeriert, dass ^{127}Asw im Gegensatz zu ^{112}Asw kein regulatorisch repräsentatives RNA Konstrukt des *add* Adenin Riboswitchs darstellt. Dies ist mit der „Stand-by binding“ Hypothese zur Initiierung der Translation vereinbar. Demzufolge bindet das 30S Ribosom die mRNA des *add* Gens zunächst unspezifischen 3‘ zum AUG Start Codon (Stand-by binding event). In diesem anfänglichen mRNA-Ribosom-Komplex wären stabilisierende Effekte

der kodierenden mRNA Sequenz auf die Asw Expressionsplattform eliminiert, sodass der Riboswitch durch eine Adenin abhängige Ausbildung der P4 Helix die spezifische Bindung des 30S Ribosoms an die Shine-Dalgarno-Sequenz modulieren kann.

In **Kapitel 5** wurden Studien zum transienten Encounter-Komplex zwischen dem Adenin Riboswitch und dem schwach bindenden Liganden Hypoxanthin durchgeführt. Mittels WaterLOGSY NMR konnte gezeigt werden, dass Hypoxanthin mit 2,6-Diaminopurin um die Adenin-Bindestelle von ^{112}Asw kompetitiert. Ferner wurde die Affinität des Hypoxanthin-Encounter-Komplexes von ^{112}Asw mit der WaterLOGSY Methode bestimmt. Diese liegt im millimolaren Bereich (K_D = 1.8 mM) und ist weitaus höher als die zelluläre Konzentration von Hypoxanthin in *E. coli* (~0.3 mM). Somit kann davon ausgegangen werden, dass kompetitive Bindung von Hypoxanthin die Adenin-abhängige Regulation von ^{112}Asw *in vivo* nicht signifikant beeinflusst. Außerdem präsentiert Kapitel 5 Resultate einer Kollaboration mit Albrecht Völklein, Henrik Gustmann und Prof. Wachtveitl, in der getestet wurde, inwieweit schwache und hoch-affine Purin-Riboswitch-Ligand-Komplexe mittels Fourier-Transformations-Infrarot (FTIR) Differenzspektroskopie basierend auf unterschiedlicher Isotopenmarkierung des Liganden charakterisiert werden können. Die FTIR Differenzspektren von ^{13}C,^{15}N-markiertem Hypoxanthin und unmarkiertem Hypoxanthin für den schwachen Hypoxanthin-Encounter-Komplex eines Adenin bindenden Aptamers und den hoch-affinen Hypoxanthin-Komplex eines Guanin bindenden Aptamers zeigten ein charakteristisches Muster an IR Banden für den jeweiligen Komplex. Jedoch konnten die FTIR Differenzspektren aufgrund spektralen Überlapps nicht ausgewertet werden.

Chapter 1: General introduction

1.1. RNA structure and dynamics

RNA consists of polycondensed nucleotide monomers, adenosine-, cytidine-, guanosine- and uridine-5'-phosphate, that are 3'-5' connected via phosphodiester bonds (Figure 4A). Under physiological solution conditions, the RNA phosphodiester-backbone is deprotonated and coordinated by monovalent or divalent ions (Kirmizialtin *et al.*, 2012). RNA can fold into well-defined 3-dimensional structures that are comparable to the highly-organized structures of proteins. RNA folding is determined by the nucleotide sequence or primary structure of an RNA, which is conventionally represented from 5'- to 3'-terminus in the one-letter code of the genetic alphabet. The major thermodynamic driving forces of RNA folding are stacking of the nucleobases, hydrogen bonding and binding of metal ions (Burkard, Turner and Tinoco, 1999). Hydrogen bonding interactions in RNA mediate the base pairing of specific nucleotides, most prominently, the Watson Crick base pairs between adenosine and uridine (AU) and guanosine and cytidine (GC) (Figure 4B) (Leontis and Westhof, 1998). In the compact A-form double helix that is formed between two complementary RNA strands (Figure 4A) these base pairs are stabilized by pi stacking.

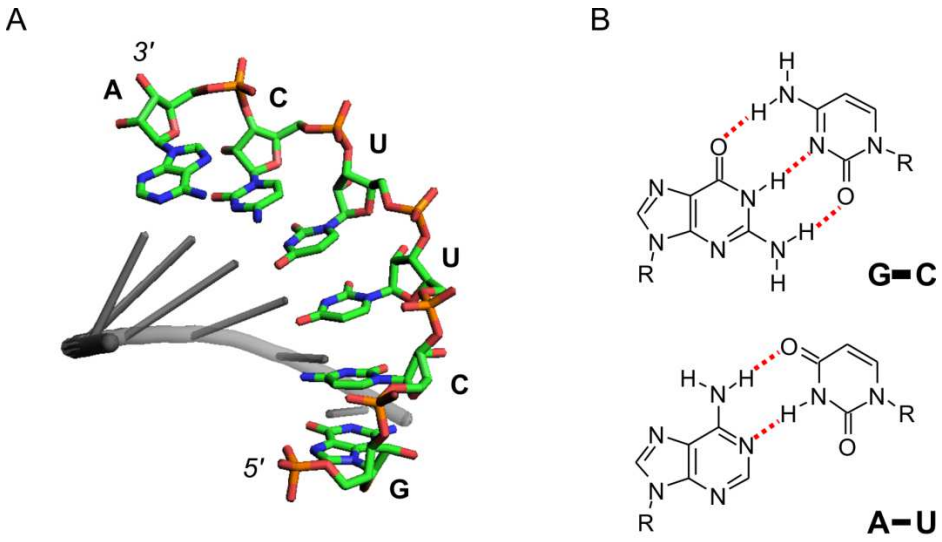

Figure 4: Structural features of RNA. (A) Three-dimensional structure of the A-form double helix in the P1 stem of the *add* adenine riboswitch aptamer (PDB ID: 1Y26) (Serganov *et al.*, 2004). The atoms of the 5'-3' strand are shown in stick representation. The complementary strand is depicted as cartoon. **(B)** Chemical formulas of guanosine (G), cytidine (C), adenosine (A) and uridine (U) in the two Watson-Crick base pairs GC and AU. R denotes the ribose moiety. Hydrogen bonds are indicated in red dashes.

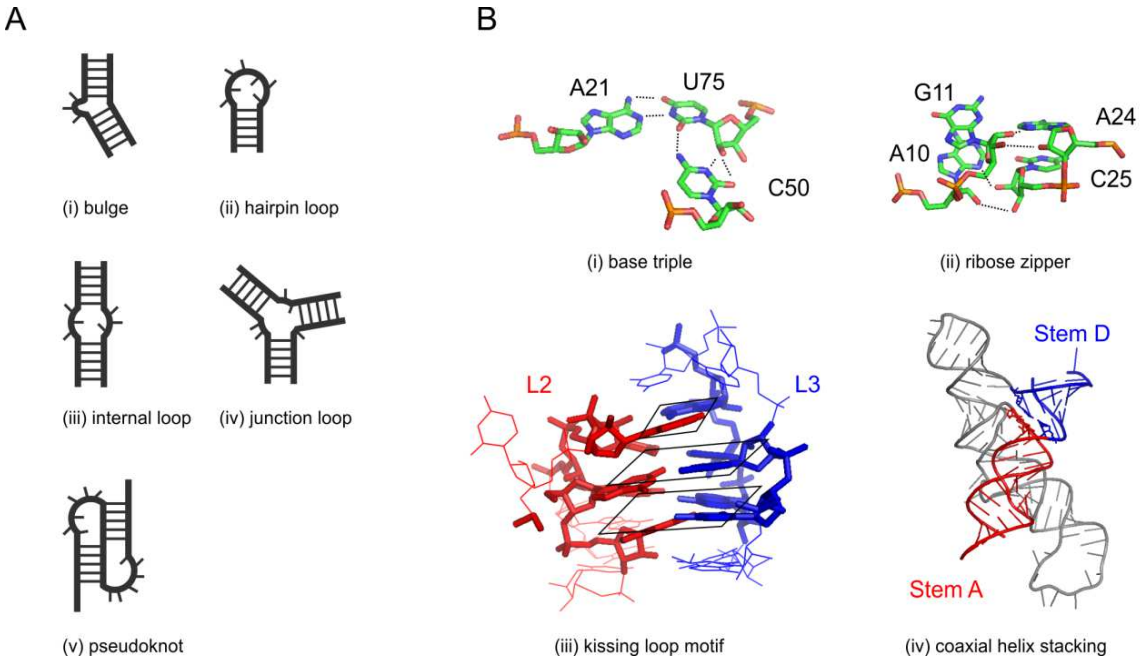

Figure 5: Motifs of RNA secondary and tertiary structure. (A) Schematic representation of typical RNA secondary structure motifs: (i) *bulge*, (ii) *hairpin loop*, (iii) *internal loop*, (iv) (3-way) *junction loop*, and (v) *pseudoknot*. Unpaired nucleotides are depicted as short dashes and base paired nucleotides as long dashes. **(B)** Examples for frequent RNA tertiary structure motifs. (i) *Base triple* in the adenine binding pocket of the adenine riboswitch aptamer (PDB ID: 1Y26) (Serganov *et al.*, 2004). C50 interacts with the minor groove side of the A21-U75 Watson Crick base pair. Hydrogen bonds are indicated in dashed lines. (ii) *Ribose zipper* between stems A and B of the hairpin ribozyme (PDB ID: 1M5K) (Rupert and Ferré-D'Amaré, 2001). The ribose 2'-OH of A10 is a hydrogen donor to the ribose 3'-phosphate of C25 and the ribose 2'-OH of C25 is a hydrogen donor to the adenine N3 of A10. Likewise, the ribose 2'-OH of A24 is a hydrogen donor to the ribose 3'-phosphate of G11 and the ribose 2'-OH of G11 is a hydrogen donor to the adenine N3 of A24. Hydrogen bonds are indicated in dashed lines. (iii) *Kissing loop motif* between hairpin loops L2 and L3 in the adenine riboswitch aptamer (PDB ID: 1Y26) (Serganov *et al.*, 2004). The two base tetrads and the terminal base pair that form the interactions between L2 and L3 are highlighted with black rhomboids. (iv) Coaxial helix stacking of stem A (red) and stem D (blue) in the hairpin ribozyme (PDB ID: 1M5K) (Rupert and Ferré-D'Amaré, 2001). The structure of the ribozyme is shown as cartoon and the stacking base pairs at the interface of stem A and stem D are shown in stick representation.

As opposed to DNA, RNA generally lacks full sequence complementarity under physiological conditions and folds in an intramolecular manner into a sequel of helical and non-helical segments. The distribution of helical structure over an RNA sequence is generally defined as secondary structure and the overall spatial arrangement of the RNA as tertiary structure (Brion and Westhof, 1997; Hendrix, Brenner and Holbrook, 2005). Typical RNA secondary structure elements are bulges, hairpin loops, internal loops, junction loops and pseudoknots (Figure 5A). Frequent RNA tertiary structure motifs are base triples, ribose zippers, kissing-loops or stacked helices (Figure 5B) (Hendrix, Brenner and Holbrook, 2005). However, the demarcation of RNA secondary and tertiary structure is not stringent. Current computational tools for RNA secondary structure

prediction cannot predict pseudoknots and consequently exclude pseudoknots from RNA secondary structure (Cragnolini, Derreumaux and Pasquali, 2015; Lorenz *et al.*, 2016). Additionally, in the field of computational RNA biochemistry, an entirely different definition of RNA secondary structure was suggested, which is defining the secondary structure of an RNA as a list of its AU, GC and GU base pairs (Zuker, 2000).

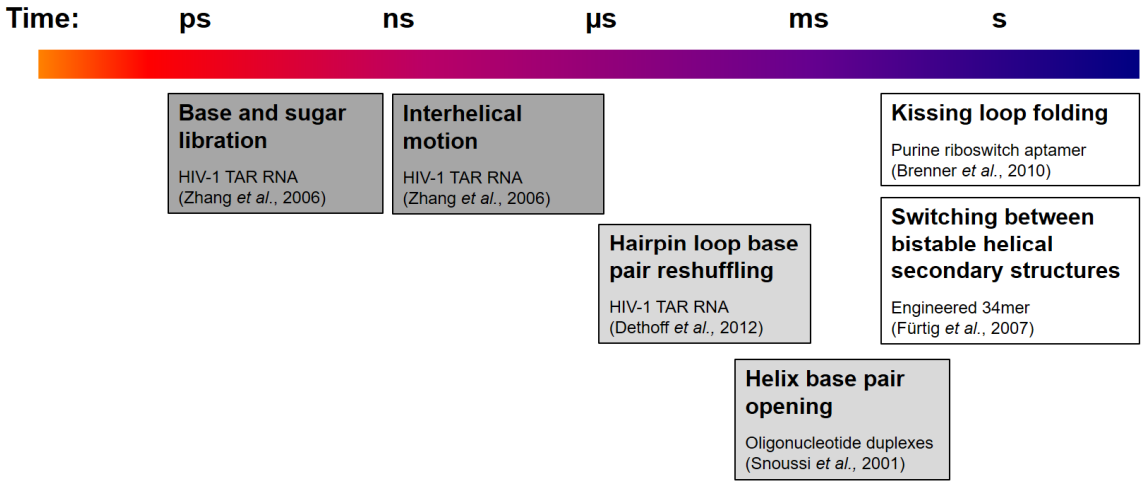

Figure 6: Structural RNA dynamics from picoseconds to seconds. Dark grey: Motional dynamics; Light grey: Local base pairing dynamics; White: Global folding dynamics. In all cases, one literature example is indicated (Snoussi and Leroy, 2001; Zhang *et al.*, 2006; Fürtig *et al.*, 2007; Brenner *et al.*, 2010; Dethoff *et al.*, 2012).

The function of various biologically relevant RNAs is related to RNA structural dynamics. In ribozymes, RNA structural dynamics mediate catalysis at the active site (Guo, Gooding and Cech, 2004; Klein, Been and Ferré-D'Amaré, 2007; Scott, Horan and Martick, 2013; Wilson *et al.*, 2016), riboswitches regulate gene expression by switching between alternative RNA structures (see section 1.3), and large macromolecular machines like the ribosome (Steitz, 2008) or the spliceosome (Sperling, Azubel and Sperling, 2008) are driven by dynamic RNA-RNA, RNA-protein and protein-protein interactions. Structural RNA dynamics, from local dynamics of specific nucleotides to global folding dynamics of secondary and or tertiary structure, span timescales from picoseconds to seconds (Figure 6) (Rinnenthal *et al.*, 2011; Mustoe, Brooks and Al-Hashimi, 2014). An ultimate goal of biophysical RNA research is to provide a molecular level understanding of how local and global RNA dynamics are coupled and mediate biological function.

1.2. Initiation of mRNA translation in bacteria

In prokaryotes, DNA transcription and mRNA translation are spatially coupled processes. Prokaryotic polysomes, complexes between mRNA and multiple actively translating 70S ribosomes are already formed at an mRNA chain that is being transcribed in the nucleoid. A polysome of a nascent mRNA chain from *E. coli* can contain up to forty 70S ribosomes, (Miller, Hamkalo and Thomas, 1970) the first of which is spatially close to the active RNA polymerase and prevents spontaneous backtracking of the mRNA producing enzyme (Nudler, 2012). As a consequence, the rate of transcription (12-42 nucleotides /s) correlates with the rate of translation (4-14 amino acids /s) in *E. coli* under various growth conditions (Proshkin *et al.*, 2010). The mean copy number of mRNAs per *E. coli* cell (~1000) (Taniguchi *et al.*, 2010) is of the same order of magnitude as the mean copy numbers of RNA polymerases (~5000) (Bakshi *et al.*, 2012) and low compared to the mean copy number of ribosomes (~60000) (Bakshi *et al.*, 2012). It is currently discussed whether mRNA translation occurs predominantly on polysomes in the nucleoid or on polysomes diffused into the segregated ribosome rich regions of the prokaryotic cell (Bakshi, Choi and Weisshaar, 2015).

Bacterial translation initiation is a multi-step process that proceeds via formation of a 30S initiation complex (30S IC) and a 70S initiation complex (70S IC) as critical intermediates. RNA-based regulation of bacterial translation initiation typically targets formation of the 30S IC (Duval *et al.*, 2015). The 30S IC consists of an mRNA, the 30S ribosomal subunit, three initiation factor proteins (IF1-IF3) and a formyl-methionine-loaded initiator tRNA (fMet-tRNA^{fMet}). 30S IC formation involves two stages: Component recruitment to a 30S pre-initiation complex (30S PIC) and rearrangement of the 30S PIC to the 30S IC (Figure 7) (Milón *et al.*, 2012). In the 30S IC, the translation start site of the mRNA is fixed in the decoding channel of the 30S ribosome via the codon-anticodon interaction between the mRNA start codon (83% AUG, 14% GUC, 3% UUG in *E. coli* (Blattner *et al.*, 1997)) and the complementary anticodon stem loop of fMet-tRNA^{fMet} (Ogle *et al.*, 2001; Simonetti *et al.*, 2008). mRNA recruitment of the 30S ribosome occurs in the absence of initiation factors and fMet-tRNA^{fMet}, but can be accelerated by the initiation factors and fMet-tRNA^{fMet} (Milón *et al.*, 2012). Specific binding of the 30S ribosomal subunit to the mRNA translation initiation site is mediated by base pairing between the 3'-tail of the ribosomal RNA (rRNA) and the partially complementary Shine Dalgarno sequence 5' adjacent to the mRNA start codon (Shine

and Dalgarno, 1974; Steitz and Jakes, 1975; Kaminishi *et al.*, 2007). The anti-Shine Dalgarno Sequence of the *E. coli* rRNA is a 13-nucleotide 5'-GAUCACCUCCUUA-3' sequence, which is highly conserved among prokaryotes (Nakagawa *et al.*, 2010). mRNA recruitment of the 30S ribosome is stimulated by the ribosomal protein S1 (Subramanian, 1983). Protein S1 was shown to be vital for *in vitro* translation of mRNAs with a weak Shine Dalgarno sequence (calculated free energy of base pairing < ~100 kJ/mol) (Farwell, Roberts and Rabinowitz, 1992). For mRNAs in which the translation initiation site is masked by helical secondary structure, it was suggested that the 30S ribosome first binds to a proximal single-stranded stand-by site, and then shifts into place as the secondary structure opens (De Smit and Van Duin, 2003). The three initiation factors can associate with the 30S ribosome before, concomitantly with, or after the mRNA (Milón *et al.*, 2012). They promote binding of fMet-tRNA^{fMet} and the adjustment of the codon-anticodon interaction (Laursen *et al.*, 2005). After assembly of the complete 30S IC with fixed codon-anticodon interaction, the 50S subunit joins and the 70S IC is formed, which progresses into the elongation phase under hydrolysis of an IF2-bound GTP and dissociation of the three IFs (Marshall, Aitken and Puglisi, 2009).

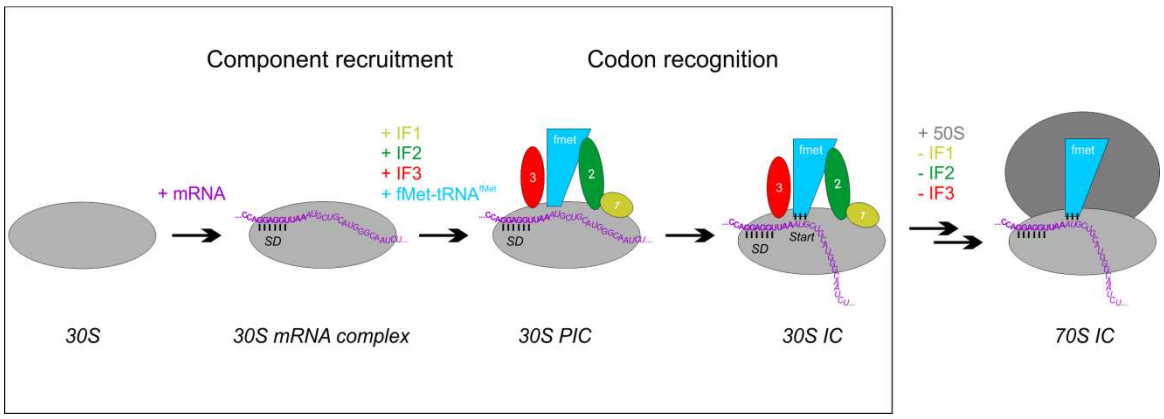

Figure 7: Schematic representation of bacterial translation initiation. Bacterial translation initiation is divided into formation of a 30S initiation complex (30S IC) and a 70S initiation complex (70S IC). RNA-based regulation of translation initiation typically targets formation of the 30S IC (boxed). 30S IC formation involves two stages. In the first stage (component recruitment), the 30S ribosome, the mRNA, three initiation factor proteins (IF1, IF2, IF3) and the initiator tRNA fmet-tRNA^{fmet} assemble to a 30S pre-initiation complex (30S PIC). The actual component recruitment order depends on the nature of the mRNA. The depicted recruitment order (mRNA prior to IFs and fmet-tRNA^{fmet}) was chosen to illustrate that the initiation factors are not necessary for mRNA binding. The 30S ribosome specifically binds to the Shine Dalgarno sequence (SD) in the 5'-untranslated region (5'-UTR; depicted in bold letters) of the mRNA. In the second stage (codon recognition), the 30S PIC undergoes conformational rearrangements to form the codon anticodon interaction between the mRNA start codon and the fmet-tRNA^{fmet}. The active 30S IC assembles with the 50S ribosomal subunit and forms the 70S IC under dissociation of the initiation factors. This Figure recapitulates a model from Milón et al (Milón *et al.*, 2012).

At mRNA level, 30S IC formation can be regulated by interfering with mRNA binding of the 30S ribosome or by interfering with the conformational rearrangement for codon recognition of the 30S PIC. A prevalent regulatory mechanism is controlling the specific binding of the 30S ribosome to the Shine Dalgarno sequence (SD) of the mRNA by masking or unmasking the SD (Duval *et al.*, 2015). Masking of the SD can be performed by RNA binding proteins that compete with the 30S ribosomal subunit for binding (Merianos, Wang and Moore, 2004; Schubert *et al.*, 2007). Likewise, bacterial small RNAs (sRNAs) can bind at and mask the SD, (Mitarai *et al.*, 2009) or unmask the SD by binding to an mRNA segment that initially masked it (Majdalani, Hernandez and Gottesman, 2002). Lastly, the mRNA itself can control the accessibility of the SD in *cis* via mRNA elements like riboswitches (see section 1.3) or RNA thermometers (Kortmann and Narberhaus, 2012) that undergo effector-dependent intramolecular switching between conformations with a masked and a liberated SD.

1.3. Riboswitches

Riboswitches are metabolite-responsive RNA conformational switch modules that are predominantly found in the 5'-untranslated region of bacterial genes. These *cis*-acting mRNA elements regulate the expression of downstream genes by a structural switch in response to binding of a specific metabolite. The ligand-induced structural switch of a riboswitch mRNA can control gene expression at the level of transcription, splicing, translation and or mRNA stability. The majority of riboswitches regulate transcription or translation, and control genes that encode proteins involved in the metabolic pathway or in the transport of their cognate ligand (Serganov and Nudler, 2013). Notably, *in vitro* selected RNA aptamers had been contrived as synthetic riboswitches for molecular biology applications to control the expression of specific genes by ligand addition already in 1998, before the first natural-occurring riboswitches were discovered (Werstuck and Green, 1998). The discovery of natural riboswitches is ascribed to Ronald Breaker, who reported in 2002 that the 5'-untranslated mRNA of the *btuB* gene in *E. coli* couples selective binding of vitamin B_{12} to a structural transition that blocks ribosome binding and inhibits translation of a transport protein for cobalamin compounds (Nahvi *et al.*, 2002). Since then, a plethora of natural-occurring riboswitches have been described that respond to biochemically diverse ligands ranging from ions (Mg^{2+} (Dann *et al.*, 2007), F^-

(Ren, Rajashankar and Patel, 2012)), nucleobases (adenine (Mandal and Breaker, 2004), guanine (Mandal *et al.*, 2003), deoxyguanosine (Kim, Roth and Breaker, 2007)), second messenger molecules (c-di-AMP (Gao and Serganov, 2014), c-di-GMP (Smith *et al.*, 2009)), amino acids (glycine (Mandal *et al.*, 2004), glutamine (Ren *et al.*, 2015), lysine (Sudarsan *et al.*, 2003)) and cofactors (SAM (Price, Grigg and Ke, 2014), FMN (Winkler, Cohen-Chalamish and Breaker, 2002), TPP (Thore, Leibundgut and Ban, 2006)) to biopolymers (tRNA (Zhang and Ferré-D'Amaré, 2015)). While most of the natural riboswitches are bacterial riboswitches, also eukaryotic riboswitches were found in fungi and in plants. Eukaryotic riboswitches typically control alternative splicing in response to TPP (Barrick and Breaker, 2007). Since riboswitches are prevalent in bacteria but not in eukaryotes, they have been envisaged as a potential drug target for novel antibiotics (Blount *et al.*, 2007; Kim *et al.*, 2009; Howe *et al.*, 2015). Moreover, riboswitches can find potential applications in synthetic biology (Sinha, Reyes and Gallivan, 2010; Miyamoto *et al.*, 2013) or in molecular diagnostics (Chappell *et al.*, 2015).

The majority of riboswitches consists of two domains, a 5' aptamer and a 3' expression platform. The aptamer is the ligand-sensing domain of a riboswitch. Riboswitch aptamers generally bind their ligand tightly (K_D ~nM) with high specificity over chemically related metabolites, and undergo ligand-induced folding from an intrinsically disordered to a highly-ordered structure. The expression platform is the regulatory domain. Ligand binding to the riboswitch aptamer changes the folding landscape of the riboswitch such that the expression platform switches from a gene-OFF state to a gene-ON state or vice versa. For many riboswitches, the isolated aptamer domain has been characterized in the ligand-bound state by X-ray crystallography. On the basis of these aptamer structures, more than 20 classes of riboswitches have been distinguished (Serganov and Nudler, 2013). For SAM-binding riboswitches alone, three riboswitch classes with up to three different subclasses have been described (Price, Grigg and Ke, 2014). This illustrates the astonishing capability of RNA to recognize the same ligand in diverse structural architectures. While the ligand-dependent structure of isolated riboswitch aptamers is well characterized, considerably less is known about the ligand-induced structural switch of full-length riboswitches. To explain the unique regulatory properties of different riboswitches, one needs to investigate how ligand-induced aptamer folding is coupled to refolding of the allosteric expression platform.

In the following, the distinct mode of action of transcriptional and translational riboswitches is exemplified by two riboswitches that have been characterized in the full-length sequence: A transcriptional ON-switch, the *pbuE* adenine riboswitch from *Bacillus subtilis*, and a translational OFF-switch, the $preQ_1$ riboswitch from *Thermoanaerobacter tengcongensis (Tte)*.

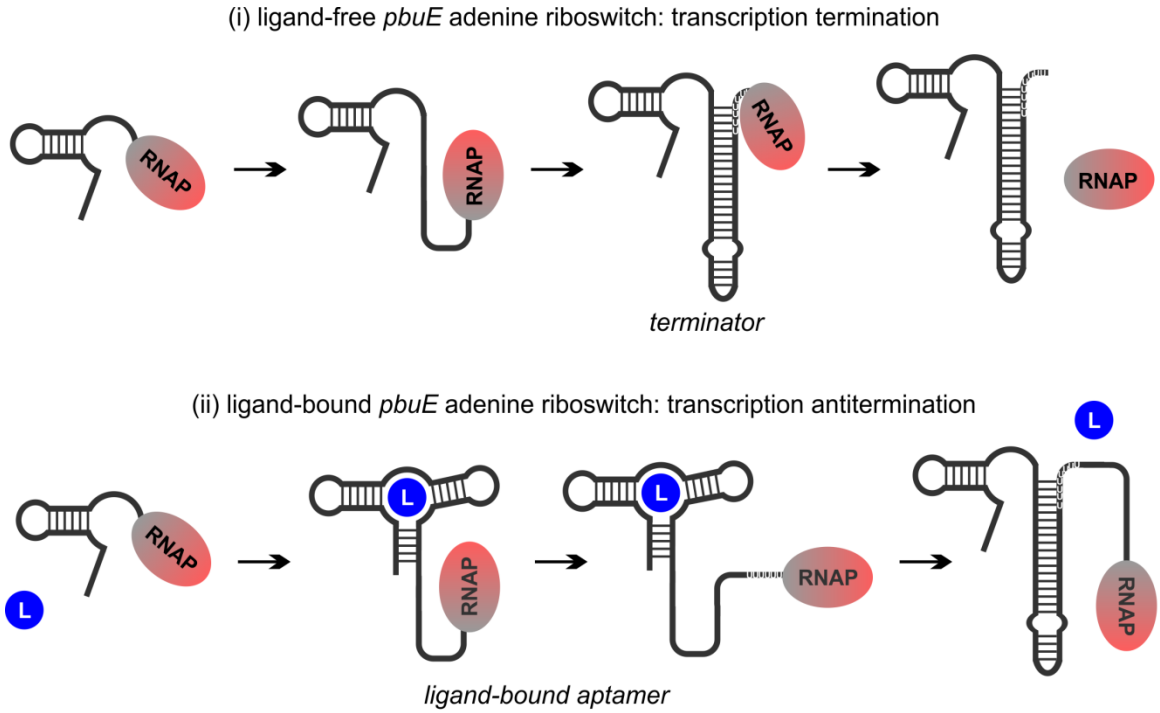

Figure 8: Schematic representation of transcriptional regulation by the *pbuE* adenine riboswitch from *Bacillus subtilis*. For clarity, only the nascent mRNA and the RNA polymerase (RNAP) are shown, and the DNA template is omitted. The depicted regulatory model is based on studies by Frieda *et al.* (Frieda and Block, 2012) and Lemay *et al.* (Lemay *et al.*, 2011). The *pbuE* adenine riboswitch is a transcriptional ON-switch. (i) At low adenine concentration, the riboswitch remains in a ligand-free state, and cotranscriptionally folds into a conformation with a terminator stem. The terminator stem pauses RNAP during the synthesis of a poly-U stretch. This triggers the release of the mRNA from the transcription complex and prevents transcription of the downstream gene. (ii) At high adenine concentration, a highly-structured adenine-bound aptamer is formed before RNAP has synthesized the full-length riboswitch. The aptamer and the terminator stem are mutually exclusive structures. With a pre-formed adenine-bound aptamer, formation of the thermodynamically favoured terminator stem is delayed until RNAP has passed the poly-U stretch. The downstream gene is transcribed.

For transcriptional riboswitches, the ligand-induced switch occurs cotranscriptionally while the mRNA is being elongated by a host-specific RNA polymerase (RNAP). Transcriptional riboswitches typically operate under kinetic control. Cotranscriptional folding of the *pbuE* adenine riboswitch from the family of purine riboswitches (see section 1.4) was monitored *in vitro* by single-molecule force spectroscopy using optical

tweezers (Frieda and Block, 2012). In the absence of adenine (Figure 8, i), the riboswitch readily formed a terminator hairpin during transcription. This triggered RNAP dissociation from the template in ~90% of the assayed transcripts. Cotranscriptional folding of the expected adenine-sensing 3-way junction aptamer in the ligand-free state was not observed in this study, presumably because the apo aptamer secondary structure was too labile to affect transcript extension. In presence of saturating adenine concentration (Figure 8, ii), the *pbuE* adenine riboswitch was kinetically trapped in an adenine-bound aptamer that acted as an antiterminator and delayed formation of the thermodynamically favoured terminator hairpin. With adenine, transcription was terminated only in ~50% of the transcripts. In line with the hypothesis that the ligand-sensing aptamer domain of the *pbuE* adenine riboswitch is only transiently formed during transcription, it was found in a different study that the full-length *pbuE* adenine riboswitch is ligand binding incompetent (Lemay *et al.*, 2011).

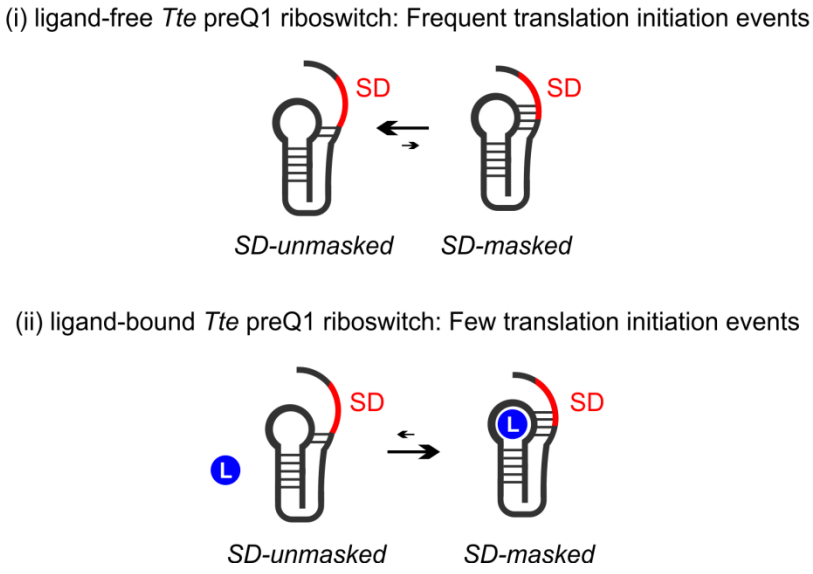

Figure 9: Schematic representation of translational regulation by the preQ$_1$ riboswitch from *Thermoanaerobacter tengcongensis (Tte)*. For clarity, only the full-length riboswitch is shown. The downstream coding mRNA chain and the 30S ribosome are omitted. The depicted regulatory model is based on studies by Suddala *et al.* (Suddala *et al.*, 2013) and Rinaldi *et al.* (Rinaldi *et al.*, 2016). The *Tte* preQ1 riboswitch is a translational OFF-switch. (i) At low preQ$_1$ concentration, the riboswitch remains in a ligand-free state and predominantly adopts a conformation in which the Shine Dalgarno sequence (SD) is accessible for binding of the 30S ribosome. (ii) At high preQ$_1$ concentration, ligand binding shifts the conformational equilibrium to a folding sate with a partially masked SD. Binding of the 30S ribosome to the SD is impeded.

In contrast to transcriptional riboswitches, several translational riboswitches were proposed to be thermodynamically controlled, since they can regulate translation via a ligand-dependent conformational equilibrium of their full-length sequence (Rieder *et al.*, 2007; Holmstrom *et al.*, 2014; Gong *et al.*, 2016; Rinaldi *et al.*, 2016). Translational riboswitches generally operate by conformational switching between states with a masked and a liberated Shine Dalgarno sequence (SD). Masking of the SD inhibits 30S ribosomal translation initiation while liberation of the SD activates it (see section 1.2). In a single-molecule FRET (smFRET) study, the small pseudoknot *Tte* $preQ_1$ riboswitch was shown to adopt two coexisting conformational states, one with a flexible and one with a partially base-paired and structurally restrained 3'-tail that contains the SD (Figure 9) (Suddala *et al.*, 2013). Without $preQ_1$, the majority of molecules adopted a state with flexible 3'-tail. At saturating $preQ_1$ concentration, the equilibrium was shifted to a conformation with restrained 3'-tail, in which the SD is partially occluded. This conformational switching behaviour is in agreement with *in vitro* translation assays that revealed 40% decreased translation in presence of saturating $preQ_1$ concentration (Rinaldi *et al.*, 2016).

1.4. Purine Riboswitches

Among the different classes of riboswitches, purine riboswitches are one of the best characterized and have emerged as a model system to study the structure, the dynamics and the function of riboswitch RNAs (Porter, Marcano-Velázquez and Batey, 2014). Purine riboswitches occur in Firmicutes, Fusobacteria and Proteobacteria (Barrick and Breaker, 2007). They commonly regulate the expression of purine salvage pathway enzymes, de-novo purine biosynthesis enzymes or purine transporter proteins (Singh and Sengupta, 2012). The vast majority of purine riboswitches either respond to guanine (and hypoxanthine) or to adenine. They were hence classified as guanine- and adenine-sensing riboswitches. A third class of purine riboswitches are deoxyguanosine-sensing riboswitches. Deoxyguanosine riboswitches, however, are rare and were only found in *Mesoplasma florum* (Kim, Roth and Breaker, 2007).

Common to purine riboswitches is the highly-conserved structure of their aptamer domain. The purine riboswitch aptamers adopt a characteristic three-way junction secondary structure (Figure 10, (i)) consisting of a P1 stem, a P2/L2 hairpin and a P3/L3 hairpin that are connected via three junctional segments J_{1-2}, J_{2-3} and J_{3-1}.

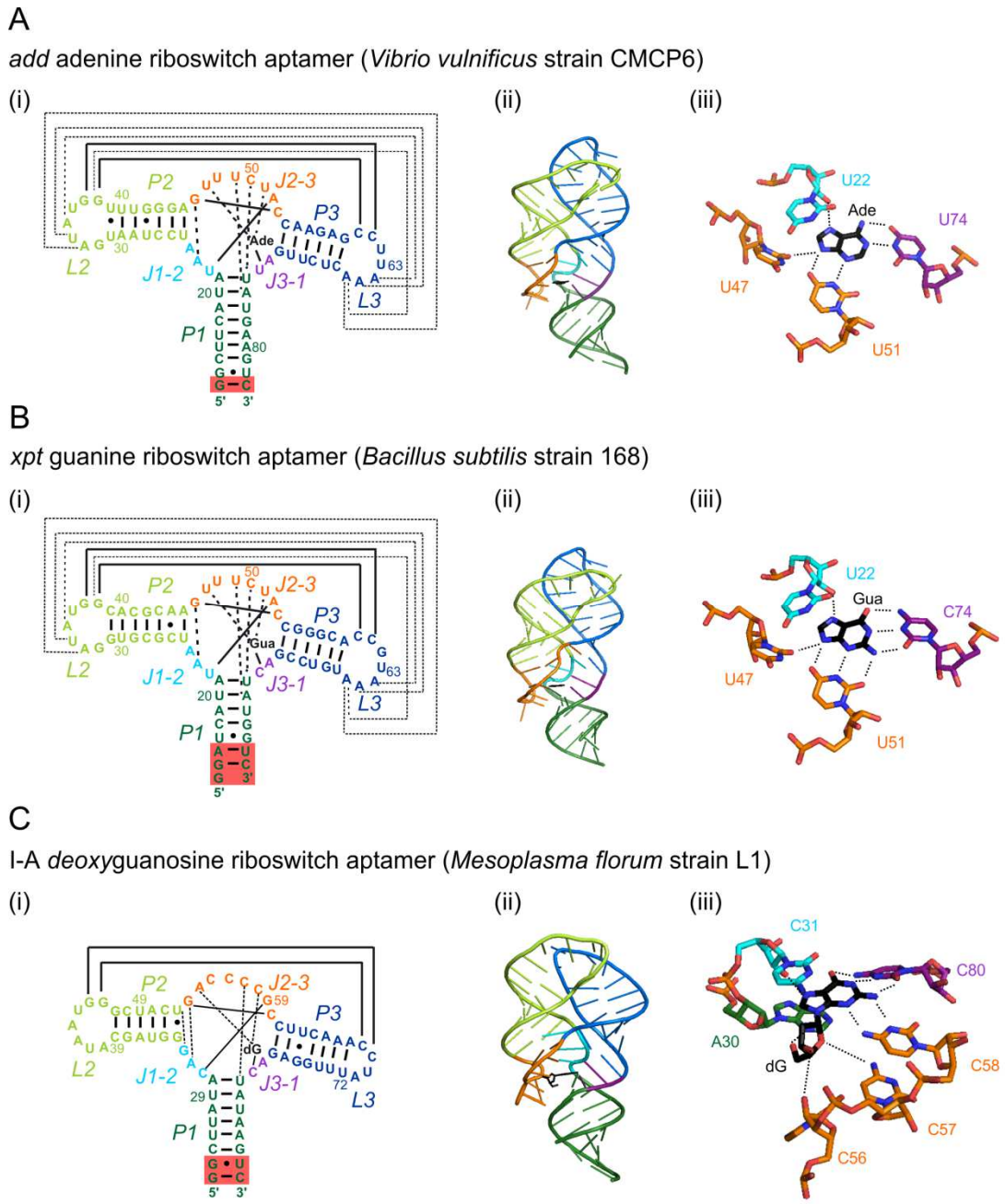

Figure 10: X-ray diffraction structure of the ligand-bound aptamer domain of three representative purine riboswitches. (A) Adenine-sensing riboswitch (PDB ID: 1Y26) (Serganov *et al.*, 2004). **(B)** guanine-sensing riboswitch (PDB ID: 1Y27) (Serganov *et al.*, 2004). **(C)** Deoxyguanosine-sensing riboswitch (PDB ID: 3SKI) (Pikovskaya *et al.*, 2011). (i) Secondary structure and long-range tertiary structural base pairing interactions. Watson Crick base pairs are represented as solid lines, GU wobble base pairs as filled circles and other non-canonical base pairing interactions as dashed lines. The primary structure of the three aptamers corresponds to the native mRNA sequence of the indicated strain but contains an artificial P1-stem. P1-modifications are highlighted in red. (ii) Tertiary structure in cartoon representation. (iii) Mode of ligand recognition. Possible hydrogen bonds are indicated with dashed lines.

X-ray diffraction (XRD) structures have been solved for the *add* adenine riboswitch from *Vibrio vulnificus*, (Serganov *et al.*, 2004) the *xpt* guanine riboswitch from *Bacillus subtilis*, (Serganov *et al.*, 2004; Gilbert *et al.*, 2009) the *pbuE* adenine riboswitch from

Bacillus subtilis, (Delfosse *et al.*, 2010) and the I-A deoxyguanosine riboswitch from *Mesoplasma florum* (Pikovskaya *et al.*, 2011). The XRD structures of the ligand-bound purine aptamers (Figure 10, (ii)) revealed their highly-organized tertiary structure which resembles a 'tuning fork' (Serganov *et al.*, 2004). The P1 helix forms the handle of the tuning fork, and the P2/L2 and the P3/L3 hairpin form the prongs that are held together by kissing loop interactions. The purine ligand is encapsulated in the aptamer core amidst $J_{1\text{-}2}$, $J_{2\text{-}3}$ and $J_{3\text{-}1}$. In adenine and guanine riboswitches, the ligand is bound by hydrogen bonding with U22 ($J_{1\text{-}2}$) at the N7-C6 (Hoogsteen) face, hydrogen bonding with U47 and U51 ($J_{2\text{-}3}$) at the N3-N9 (sugar) face and Watson Crick base pairing with Y74 ($J_{3\text{-}1}$) at the C6-C2 (Watson Crick) face (Figure 10, (iii)). The ligand specificity is mediated by Y74 which is a uridine in the adenine riboswitch and a cytidine in the guanine riboswitch. In the deoxyguanosine riboswitch, the ligand is recognized in a similar manner, by hydrogen bonding interactions between C31 ($J_{1\text{-}2}$), C58 ($J_{2\text{-}3}$) and C80 ($J_{3\text{-}1}$) and the nucleobase and between A30 (P1), C56 and C57 ($J_{2\text{-}3}$) and the deoxyribose moiety. The organization of the junctional core structure in these ligand-bound purine riboswitch aptamers stabilizes the global arrangement of their three helices. The $J_{3\text{-}1}$ segment mediates stacking of the P3 helix upon the P1 helix. The $J_{2\text{-}3}$ segment is engaged in hydrogen bonding interactions with the P1 helix and with the $J_{1\text{-}2}$ segment. This compact purine riboswitch aptamer tertiary structure is further stabilized by specifically bound Mg^{2+} ions at the aptamer core and at the peripheral kissing loop motif (Noeske, Schwalbe and Wöhnert, 2007; Pikovskaya *et al.*, 2011).

The ligand-free purine riboswitch aptamers adopt the 3-way junction secondary structure also in solution, but exhibit an intrinsically disordered tertiary structure, especially at the junction connecting segments of the aptamer core. In the ligand-free *xpt* guanine and *pbuE* adenine riboswitch aptamer, the L2/L3 kissing loop interaction is preformed already without Mg^{2+} (Noeske *et al.*, 2007; Noeske, Schwalbe and Wöhnert, 2007; Nozinovic *et al.*, 2014), albeit to a different extent. The I-A deoxyguanosine riboswitch aptamer requires Mg^{2+} for formation of the kissing loop interaction in the apo state (Wacker *et al.*, 2011). At saturating ligand conditions, all three P1-stabilized purine riboswitch aptamers adopt a homogenous solution structure that is in good agreement with the XRD structure of their ligand-bound state (Noeske *et al.*, 2007; Noeske, Schwalbe and Wöhnert, 2007; Wacker *et al.*, 2011). Time-resolved NMR studies of the ligand-induced folding of the *xpt* guanine riboswitch aptamer (Buck *et al.*, 2007) and the *add* adenine riboswitch aptamer

(Lee *et al.*, 2010) have demonstrated that ligand binding to the aptamer core first induces folding of the binding pocket and subsequently leads to helical tightening, in particular of P1, along with stabilization of the L2/L3 kissing loop interaction.

The high degree of structural conservation in the purine riboswitch family is limited to the aptamer domain. The *add* adenine riboswitch from *Vibrio vulnificus*, the *pbuE* adenine and the *xpt* guanine riboswitch from *Bacillus subtilis* and the I-A deoxyguanosine riboswitch from *Mesoplasma florum* each have an individual expression platform (Serganov *et al.*, 2004; Helmling *et al.*, 2017). The different expression platforms of these riboswitches have in common that they can form two mutually exclusive secondary structures, one with base pairing confined to the expression platform and one with base pairing between aptamer and expression platform. The first leaves the 3-way junction secondary structure of the aptamer intact and the latter disrupts it. According to the working hypothesis of purine riboswitch function, purine riboswitches operate via the general switching principle, that ligand binding to the aptamer domain stabilizes a conformation without base pairing between aptamer and expression platform. In the *pbuE* adenine riboswitch, adenine binding to the 3-way junction aptamer prevents formation of a transcription terminator that involves the nucleobases of the P1 and the P3 helix (see section 1.3, Figure 8). In the *add* adenine riboswitch adenine binding prevents formation of the P4 helix at the expense of the P1 stem and thereby liberates the ribosome binding site (see section 1.5 Figure 11).

1.5. The full-length *add* adenine riboswitch from *Vibrio vulnificus*

The *add* adenine riboswitch from the human pathogenic marine bacterium *Vibrio vulnificus* (strain CMCP6) is the sole translation regulating riboswitch of the structurally characterized purine riboswitches (see section 1.4). It regulates the expression of the *add* gene, which encodes the purine salvage pathway enzyme adenosine deaminase. The full-length *add* Asw encompassing the adenine-sensing aptamer domain and the expression platform up to the AUG start codon of the *add* gene is a 109-nucleotide long translational ON-switch. Its *in vivo* activity was demonstrated in *E. coli* in two different studies. Lemay *et al.* performed *in vivo* β-Galactosidase activity assays with translational *add* Asw *lacZ* fusions. They observed a ~3-fold increase in β-Galactosidase activity 2 h after supplementing the growth medium with 500 µM adenine (Lemay *et al.*, 2011). Dixon *et*

al. performed *in vivo* fluorescent protein expression assays with the *add* Asw fused to an eGFP gene, and measured a ~10-fold induction of eGFP expression in response to the adenine analogue 2-aminopurine at a concentration of 250 µM (Dixon *et al.*, 2012).

Several biophysical studies of the full-length *add* Asw provided insight into its molecular mechanism. Rieder *et al.* performed stopped flow fluorescence spectroscopic studies on fluorescent U48AP variants (AP = 2-aminopurine) of an isolated 71-nucleotide *add* Asw aptamer domain and a 111-nucleotide full-length *add* Asw (Rieder *et al.*, 2007). Upon folding of the *add* Asw aptamer tertiary structure, the AP at nucleotide position 48 becomes solvent exposed resulting in a substantial increase in fluorescence. The full-length U48AP *add* Asw exhibited a comparable adenine affinity to the isolated U48AP aptamer domain (K_D(111mer) = 2.3 µM, K_D(71mer) = 0.7 µM at 25 °C and 2 mM Mg^{2+}) and bound adenine with almost the same rate constant (k_{ON}(111mer) = 3.73*10^4 $M^{-1}s^{-1}$, k_{ON}(71mer) = 3.47*10^4 $M^{-1}s^{-1}$ at 20 °C and 2 mM Mg^{2+}). Since the adenine binding properties of the isolated aptamer domain were retained in the full-length riboswitch, it was proposed that the *add* Asw operates via a thermodynamically controlled ligand-dependent conformational switch of its full-length sequence. This hypothesis was corroborated by Lemay et al (Lemay *et al.*, 2011). Lemay *et al.* investigated the ligand-dependent folding of a 119-nucleotide full-length *add* Asw by SHAPE (selective 2'-hydroxyl acylation analysed by primer extension) and RNAse T1 partial cleavage. Their data demonstrated that aptamer folding induced by Mg^{2+} and or adenine increases the accessibility of the Shine Dalgarno sequence (SD) and the start codon (AUG) in the expression platform, in line with a conformational switch from a translational OFF-state to a translational ON-state.

More recently, Reining *et al.* revealed that the *add* Asw does not only switch between two but between three distinct long-lived conformational states (Figure 11) (Reining *et al.*, 2013). In their NMR spectroscopic study, Reining *et al.* assigned the imino NH resonances of adenine-free and adenine-bound 112-nucleotide *add* Asw at 10 °C and 5 mM Mg^{2+} via a divide-and-conquer approach using a P1/P2, a P3 and a P4/P5 fragment as well as the isolated aptamer domain as structural reference modules. They discovered secondary structure heterogeneity in the ligand-free riboswitch, between a structurally binding-incompetent apoB conformation and the anticipated adenine-sensing apoA conformation (Figure 11). According to the NMR derived base pairing pattern, apoB and apoA are functional OFF states with a masked translation initiation site, and the holo

conformation is a functional ON state with solvent exposed SD and AUG. The conformational pre-equilibrium between the apoB and the apoA conformation in the ligand-free *add* Asw proved to be temperature-dependent. With increasing temperature from 10 °C to 30 °C, the fractional population of the major apoB conformation decreased from 90% to 60%. This temperature dependence is reversed compared to the intrinsic temperature dependence of ligand binding. It was thus hypothesized that the apoB/apoA secondary structure bistability of the *add* Asw has an RNA thermometer function and compensates the temperature dependence of riboswitch ligand binding. Simulations of riboswitch ligand response profiles as a function of temperature showed that the 3-state *add* adenine riboswitch can maintain a largely constant switching efficiency between 5 °C and 30 °C, while the switching efficiency of a hypothetical 2-state *add* Asw would be substantially compromised between 5 °C and 15 °C (Reining *et al.*, 2013; Fürtig *et al.*, 2015). Due to its temperature compensation, the 3-state *add* Asw was termed a riboswitch thermostat (Reining *et al.*, 2013). Riboswitch thermostats might be of general importance for organisms that are exposed to varying temperatures, such as *Vibrio vulnificus*, which can be present in the marine environment or in a mammalian host.

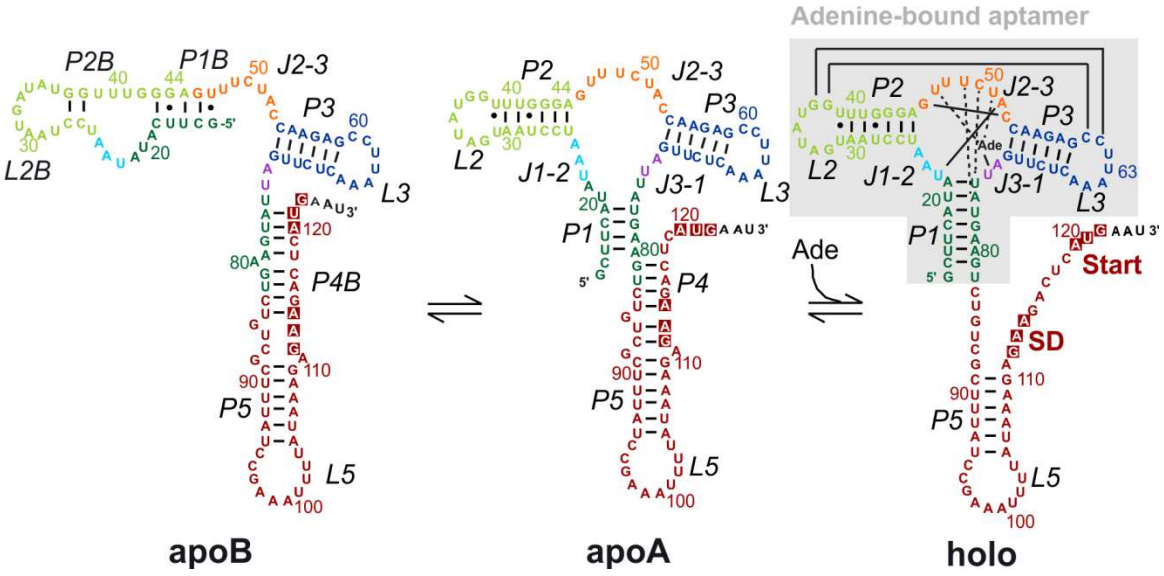

Figure 11: Three-state model of the full-length translation-regulating *add* adenine riboswitch from *Vibrio vulnificus*. NMR-derived 3-state solution structural model of the full-length *add* adenine riboswitch in a 112-nucleotid wildtype mRNA sequence according to Reining *et al.* (Reining *et al.*, 2013). The distinct base pairing patterns of the apoB, the apoA and the holo conformation are shown. ApoB lacks the adenine-sensing three-way junctional secondary structure of the aptamer domain and has an extended P4B helix in the expression platform that masks the AUG start codon. ApoA adopts the junctional purine riboswitch aptamer fold with a disordered aptamer tertiary structure and has shorter P4 helix that masks the Shine Dalgarno sequence (SD). Holo forms a compact aptamer tertiary structure and a disordered expression platform with solvent exposed SD and AUG.

Chapter 2: Material and Methods

2.1. General methods

2.1.1. Concentration measurements by UV/vis spectroscopy

All concentration measurements by UV/vis spectroscopy were performed on a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific). The concentrations were calculated from the measured absorbance at λ_{max} via Beer's Law using the extinction coefficients ε given in Table 1.

Table 1: UV/vis extinction coefficients used for concentration determination

Plasmids and oligonucleotides [a]	**λ_{max} [nm]**	**ε [µl*ng^{-1}*cm^{-1}]**
ssDNA	260	1/33
dsDNA	260	1/50
RNA	260	1/40
RNA constructs [b]	**λ_{max} [nm]**	**ε [mM^{-1}*cm^{-1}]**
112nt Asw	260	1050
127nt Asw	260	1230
66nt *pbuE* adenine aptamer	260	630
80nt *xpt* guanine aptamer	260	760
purine bases [c]	**λ_{max} [nm]**	**ε [mM^{-1}*cm^{-1}]**
adenine	260	13.5
2-aminopurine	305	6.0
2,6-diaminopurine	279	8.9
hypoxanthine	249	10.5
Cyanine dyes [d]	**λ_{max} [nm]**	**ε [mM^{-1}*cm^{-1}]**
Cy3	550	150
Cy5	650	250

[a] For plasmids and oligonucleotides, the default nucleic acids extinction coefficients of the NanoDrop were employed.
[b] For RNA constructs, the extinction coefficient was determined by the IDT Oligo Analyser 3.1 that is provided on the IDT website.
[c] For purine bases, the extinction coefficients reported by Mason *et al*. were used (Mason, 1954)
[d] For the cyanine dyes, the extinction coefficients provided in the manufacturer's manual of the Cy3 and Cy5 monoreactive dye pack (GE Healthcare) were used.

2.1.2. Analytical agarose gel electrophoresis

Analytical scale agarose gel electrophoresis was performed using a Sub-Cell® horizontal electrophoresis system (Bio-Rad). Gels were casted with 30 mL 1% w/w agarose (for analysis of plasmid DNA) or 2% w/w agarose (for analysis of PCR products) in TAE

buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0). Typically, 10 µL samples containing ~1 µg DNA were mixed with 2 µL 6x gel loading dye blue (New England Biolabs) and loaded adjacent to 5 µL 2-log DNA ladder (1 µg/µL; 0.1-10kB; New England Biolabs). The gels were run in TAE buffer at 120 V for 40 – 60 min. The bands were stained in GelRed™ (Biotium) and visualized on the UV-transilluminator of a Gel iX20 Imager (Intas).

2.1.3. Analytical denaturing polyacrylamide gel electrophoresis

Analytical scale denaturing polyacrylamide gel electrophoresis was performed using a Multigel system (Biometra). Gels were prepared with 10 mL 10% v/v acrylamid bis-acrylamide (29:1) in 2x TBE buffer (1x: 90 mM Tris, 90 mM boric acid, 2 mM EDTA, pH 8.0) with 7-8 M urea and polymerized by addition of 0.1% w/v ammonium persulfate and 0.1% v/v TEMED. Typically, 8 µL samples containing ~1 µg RNA were used and mixed with 8 µL formamide buffer (90% v/v formamide, 90 mM Tris, 90 mM boric acid, 2 mM EDTA, 0.01% w/v bromophenol blue and 0.01% w/v xylene cyanol) prior to loading. For fluorophore-labelled RNA, bromophenol blue and xylene cyanol were omitted in the formamide buffer. The gels were run in TBE buffer at 200 V for 40 min. For unlabelled or isotope-labelled RNA, the bands were stained in GelRed™ (Biotium) and visualized on the UV-transilluminator of a Gel iX20 Imager (Intas). For fluorophore-labelled RNA, the bands were visualized via fluorescence detection on a Tayphoon scanner (GE Healthcare).

2.1.4. Analytical native polyacrylamide gel electrophoresis

Analytical scale native polyacrylamide gel electrophoresis was performed using a Multigel system (Biometra). Gels were prepared with 10 mL 12% v/v acrylamide bis-acrylamide (29:1) in Tris acetate buffer (100 mM Tris, 100 mM sodium acetate, pH 8.0) and polymerized by addition of 0.1% w/v ammonium persulfate and 0.1% v/v TEMED. Typically, 5 µL samples containing ~0.5 µg RNA or ~20 ng fluorophore-labelled RNA were mixed with 5 µL 70 % v/v glycerol for loading. The gels were run in Tris acetate buffer at 40 V for 4 - 6 h in a 4°C cold-room. For unlabelled or isotope-labelled RNA, the bands were stained in GelRed™ (Biotium) and visualized on the UV-transilluminator of a

Gel iX20 Imager (Intas). For fluorophore-labelled RNA, the bands were visualized via fluorescence detection on a Tayphoon scanner (GE Healthcare).

2.2. Preparation of $^{13}C,^{15}N$-labeled adenine and hypoxanthine

The preparation of $^{13}C,^{15}N$-labelled adenine and hypoxanthine was performed by Albrecht Völklein as part of his Master thesis. $^{13}C,^{15}N$-labeled adenine was prepared via the acidic hydrolysis of $^{13}C,^{15}N$-labelled ATP (Silantes) as described (Noeske *et al.*, 2005). $^{13}C,^{15}N$-labeled hypoxanthine was prepared via oxidation of $^{13}C,^{15}N$-labelled adenine with sodium nitrite as reported (Noeske *et al.*, 2007).

2.3. Preparation of riboswitch RNAs by *in vitro* transcription

For this PhD thesis, five different riboswitch RNAs were prepared by *in vitro* transcription.

- 112-nucleotide wt *add* adenine riboswitch from *Vibrio vulnificus* CMCP6 (^{112}Asw) was prepared by *in vitro* transcription from a PCR product and purified via DEAE anion exchange chromatography and HPLC
- ^{112}Asw mutant apoB_{STAB} was prepared by *in vitro* transcription from a linearized plasmid and purified via DEAE anion exchange chromatography and HPLC
- 127-nucleotide wt Asw (^{127}Asw) from *Vibrio vulnificus* CMCP6 was prepared by *in vitro* transcription from a PCR product and purified via DEAE anion exchange chromatography and preparative denaturing polyacrylamide gel electrophoresis
- A 66-nucleotide *pbuE* adenine aptamer from *Bacillus subtilis* was prepared by Albrecht Völklein as part of his Master thesis by *in vitro* transcription from a PCR product and purified via DEAE anion exchange chromatography and HPLC
- An 80-nucleotide *xpt* guanine aptamer variant from *Bacillus subtilis* was prepared by Albrecht Völklein as part of his Master thesis by *in vitro* transcription from a linearized plasmid and purified via DEAE anion exchange chromatography and HPLC

The preparation procedure is described in detail in the following chapter.

2.3.1. DNA templates

The DNA template sequences used for *in vitro* transcription are given in Table 2.

Table 2: DNA sequences used for *in vitro* transcription of riboswitch RNAs (written in 5' to 3' direction)

^{112}Asw wt DNA [a]	GAATTCTAATACGACTCACTATAGGGAGATGAAGCCTGATG AGAGCGAAAGCTCGAAACAGCTGTGAAGCTGTCGCTTCATA TAATCCTAATGATATGGTTTGGGAGTTTCTACCAAGAGCCTT AAACTCTTGATTATGAAGTCTGTCGCTTTATCCGAAATTTTAT AAAGAGAAGACTCATGAATATTGGATCC
Forward Primer	TAATACGACTCACTATAGG
Reverse Primer	ATTCATGAGTCTTCTCTTTAT

^{112}Asw apoB$_{STAB}$ DNA [b]	GAATTCTAATACGACTCACTATAGGGAGAATGGCGCCTGAT GAGAGCGAAAGCTCGAAACAGCTGTGAAGCTGTCGCGCCAT ATAATCCTAATGATATGGTTTGGGCGTTTCTACCAAGAGCCT TAAACTCTTGATTATGAAGTCTGTCGCTTTATCCGAAATTTTA TAAAGAGAAGACTCATGAATATTGGATCC

^{127}Asw wt DNA [c]	GAATTCTAATACGACTCACTATAGGGAGATGAAGCCTGATG AGAGCGAAAGCTCGAAACAGCTGTGAAGCTGTCGCTTCATA TAATCCTAATGATATGGTTTGGGAGTTTCTACCAAGAGCCTT AAACTCTTGATTATGAAGTCTGTCGCTTTATCCGAAATTTTAT AAAGAGAAGACTCATGAATATTGGATCC
Forward Primer	TAATACGACTCACTATAGG
Reverse Primer	CGGCAGGTCAAAGTAATTCATGAGTCTTCTCTTTAT

pbuE adenine aptamer DNA [d]	AACTTGTAAAACGAATTCTAATACGACTCACTATAGGCTTGT ATAACCTCAATAATATGGTTTGAGGGTGTCTACCAGGAACCG TAAAATCCTGATTACAA
Forward Primer	TAATACGACTCACTATAGG
Reverse Primer	TTGTAATCAGGATTTTACGG

xpt guanine aptamer DNA [e]	GAATTCTAATACGACTCACTATAGGAACACTCATATAACTGC GTGGATATGGCACGCAGGTTTCTACCGGGCACCGTAAATGTC CGACTATGGGTGAGCAATGGATCC

[a] The DNA was provided by Anke Reining as an insert in the pUC57 plasmid. The colour coding is as follows: EcoRI restriction site (brown), T7 promotor (purple), hammerhead ribozyme (grey), ^{112}Asw wt (black), 3'-extension of SspI restriction site (red), BamH1 restriction site (orange). ^{112}Asw wt was prepared by *in vitro* transcription of a PCR product from the plasmid. The employed PCR primers are indicated and were ordered at MWG Eurofins (HPLC purified oligonucleotides).

[b] The DNA was purchased at GenScript as insert in the pUC57 plasmid. The colour coding is as described in *a*. ^{112}Asw apoB$_{STAB}$ was prepared by *in vitro* transcription of the linearized plasmid obtained by SspI restriction digest.

[c] ^{127}Asw wt was prepared by *in vitro* transcription of a PCR product from the pUC57 plasmid of ^{112}Asw. The employed PCR primers are indicated and were ordered at MWG Eurofins (HPLC purified oligonucleotides). The 5' extension of the reverse primer that yields the 3' extension of ^{127}Asw is underlined.

[d] The dsDNA template was provided by Senada Nozinovic. The colour coding is as follows: 5' overhang (grey), T7 promotor (purple), 66-nucleotide *pbuE* adenine aptamer (black). The *pbuE* adenine aptamer was prepared by *in vitro* transcription of a PCR product of the dsDNA. The employed PCR primers are indicated and were ordered at MWG Eurofins (HPLC purified oligonucleotides).

[e] The DNA template was provided by Hannah Steinert as an insert in the pUC19 plasmid. The colour coding is as follows: EcoRI restriction site (brown), T7 promotor (purple), 80-nucleotide *xpt* guanine aptamer (black), 3'-extension of BamH1 restriction site (orange). The *xpt* guanine aptamer was prepared by *in vitro* transcription of the linearized plasmid obtained by BsrDI restriction digest.

2.3.2. Plasmid DNA Mega prep

The pUC57 or pUC19 DNA plasmid with riboswitch sequence insert (see Table …) was amplified in DH5α competent *E. coli* cells (Thermo Fisher Scientific) and isolated using the Nucleobond® PC 2000 plasmid DNA purification kit (Macherey Nagel). For transformation, 20 µL cells were added to 0.5 µL plasmid DNA (c ~300 ng/µL) in a pre-chilled 1.5 µL Eppendorf tube. The mixture was flicked gently, incubated on ice for 30 min, subjected to a heat shock in a water bath at 42 °C for 30 s and again cooled on ice for 5 min. Subsequently, 980 µL LB medium were added and the cells were incubated in a thermoshaker (37°C, 150 rpm, 1 h). 100 µL of the cell suspension were then spread on an LB/ampicillin agar plate and incubated at 37 °C over night. For a starter culture, 5 mL LB medium containing 100 mg/L ampicillin were inoculated with a single colony of the LB/ampicillin agar plate and incubated in a thermoshaker (37°C, 150 rpm, 8 h). The starter culture was used to inoculate a 1 L LB/ampicillin main culture, which was grown over night in a thermoshaker (37°C, 150 rpm). The cells were harvested by centrifugation (4000 x g, 4 °C, 15 min). Plasmid isolation was then performed according to the manufacturer's manual of the Nucleobond® PC 2000 plasmid DNA purification kit. Briefly, the cells were resuspended in 10 mL Buffer S1 per gram cell pellet, lysed by addition of 1 vol Buffer S2 following 5 min incubation at room temperature, diluted to 1.5 vol with ice-cold Buffer S3 and chilled on ice for 5 min. After centrifugation (38000 x g, 30 min) the lysate was cleared by vacuum filtration over a 0.4 micron membrane filter. The AX-2000 gravity flow column was equilibrated with 20 mL buffer N2 prior to loading the lysate. It was then washed with 70 mL buffer N3 and eluted with 25 mL buffer N5. The plasmid DNA was precipitated by addition of 18 mL isopropanol, pelleted by centrifugation (10000 x g, 4°C, 1 h) and reconstituted in 1 mL water. The DNA concentration was determined by UV/vis spectroscopy (see section 2.1.1). The sequence identity was confirmed by Sanger Sequencing with the M13 uni primer (Eurofins Genomics).

2.3.3. Plasmid linearization by enzymatic restriction digestion

Linearization of the plasmid DNA templates by enzymatic restriction at the 3' end of the riboswitch coding sequence was performed with Ssp1 (New England Biolabs) for apoB_{STAB} ^{112}Asw in pUC57 and with BsrDI (New England Biolabs) for the *xpt* guanine

aptamer in pUC19. For the restriction digest with SspI, 2 mg DNA were incubated with 1000 U Ssp1 in 4 mL Ssp1 restriction buffer at 37 °C for 16 h. For the restriction digest with BsrDI, 1 mg DNA was incubated with 250 U BsrDI and 0.25 mg BSA in 2.5 mL NEBuffer 2 at 65 °C for 8 h following heating at 80°C for 20 min for enzyme inactivation. Completion of the restriction digest was verified by agarose gel electrophoresis (see section 2.1.2). The linearized plasmid was then isolated from the enzymatic reaction mixture by phenol chloroform extraction. The mixture was extracted 3 times with 1 vol Roti®-Phenol-Chloroform-Isoamyl alcohol (25:24:1, pH 7,5-8,0; Carl Roth) and 3 times with 1 vol of chloroform. In each extraction step rapid phase separation was achieved by centrifugation (10000 x g, 4 °C, 5 min). The DNA was then precipitated from the aqueous phase by addition of 0.1 vol 3 M NaOAc pH 5.2 and 2.5 vol ice-cold abs. ethanol and subsequent cooling at -80 °C for 10 min. The DNA was pelleted by centrifugation (10000 x g, 4 °C, 30 min), washed with ~5 mL 70% ethanol, centrifuged a second time (10000 x g, 4 °C, 15 min), air-dried for 15 min and reconstituted in 1 mL water. The DNA concentration was determined by UV/vis spectroscopy (see section 2.1.1).

2.3.4. PCR amplification of DNA templates

PCR amplification of DNA templates with a T7 promoter forward primer and a specifically designed reverse primer (see Table …) was performed as described by Helmling *et al* (Helmling *et al.*, 2015). To produce sufficient DNA for a preparative scale *in vitro* transcription, a 4 mL batch containing ~2 ng/µL pUC57 plasmid DNA template, 200 µM dATP, dCTP, dGTP and dTTP, 5 µM forward primer and 5 µM reverse primer in Phusion® HF buffer (New England Biolabs) was supplemented with 20 U/mL Phusion® High-Fidelity DNA polymerase (New England Biolabs), distributed to a 96-well plate in 80 µL aliquots and subjected to the thermocycling protocol indicated in Table 3.

Table 3: PCR cycling protocol for DNA template amplification

Step	T [°C]	t [s]	Repetitions
Initial denaturation	98 °C	120	1
Annealing	49°C	20	
Extension	72 °C	15	50
Denaturation	98°C	10	
Final hold	4°C		1

In case of the dsDNA template for the *pbuE* adenine aptamer, the PCR was performed at 2.5 mL scale with an initial template concentration of 0.02 ng/µL. The PCR product was analyzed by agarose gel electrophoresis (see section 2.1.2). 10 µL of the PCR product were purified using the QIAquick® PCR purification kit (Qiagen). The sequence identity of the purified PCR product was confirmed by Sanger Sequencing with T7 promoter forward primer (Eurofins Genomics).

2.3.5. *In vitro* transcription

In vitro transcriptions from linearized plasmids or from crude PCR products were performed with in-house prepared T7 RNA polymerase mutant P266L.

Table 4: Optimized conditions for *in vitro* transcriptions from linearized plasmids

RNA construct	^{112}Asw apoB$_{STAB}$	*xpt* guanine aptamer
Tris glutamate pH 8.1	200 mM	200 mM
Magnesium acetate	35 mM	20 mM
rNTP	4 mM	12 mM
Spermidine	2 mM	2 mM
Lin. Plasmid	50 ng/µL	50 ng/µL
DTT	20 mM	20 mM
T7 polymerase P266L	0.01 mg/mL	0.01 mg/mL
YIPP	-	0.3 U/mL
Batch scale	*30 mL*	*25 mL*
Incubation time	*O/N*	*4 h*

Table 5: Optimized conditions for *in vitro* transcriptions from crude PCR products

RNA construct	^{112}Asw	^{127}Asw	*pbuE* adenine aptamer
Tris glutamate pH 8.1	200 mM	200 mM	200 mM
Magnesium acetate	30 mM	35 mM	20 mM
rNTP	4 mM	6 mM	10 mM
Spermidine	2 mM	2 mM	2 mM
DMSO	20% v/v	20% v/v	20% v/v
Triton-X	0.01% v/v	0.01% v/v	0.01% v/v
Crude PCR product	15% v/v	15% v/v	10% v/v
DTT	20 mM	20 mM	20 mM
T7 polymerase P266L	0.01 mg/mL	0.01 mg/mL	0.01 mg/mL
Batch scale	*25 mL*	*25 mL*	*25 mL*
Incubation time	*O/N*	*4h*	*O/N*

The transcriptions were carried out with an rNTP mix whose composition matched the base composition of the synthesized transcript. All riboswitch RNAs for NMR experiments were prepared with ^{15}N-labelled GTP and UTP (Silantes). The transcription conditions were optimized with respect to the Mg^{2+}, the rNTP and the DNA concentration in small scale test transcriptions (25 µL). Test transcriptions were analysed by denaturing polyacrylamide gel electrophoresis (see section 2.1.3) and submitted to Elke Stirnal for quantification of the reaction yield by HPLC-UV/vis. The optimal conditions in terms of product homogeneity and yield were used for preparative transcription and are indicated in Table 4 for transcriptions from linearized plasmids and in Table 5 for transcriptions from crude PCR products. For preparative transcriptions, all components except DTT and T7 polymerase were mixed in a 50 mL Falcon tube and pre-incubated in a thermoshaker for temperature equilibration (37°C, 150 rpm, 10 min). The mixture was then supplemented with DTT and T7 polymerase and distributed equally over five 50 mL Falcon tubes to achieve efficient homogenization during shaking. The tubes were incubated in a thermoshaker for the indicated incubation times (37°C, 150 rpm). In case of the *xpt* guanine aptamer, yeast inorganic pyrophosphatase (YIPP; New England Biolabs) was added after 2 h in order to increase the transcription yield by enzymatic hydrolysis of pyrophosphate. After incubation, the transcription reactions were stored at -20 °C until purification of the target RNA.

2.3.6. Initial RNA purification by DEAE anion exchange chromatography

In a first purification step, the transcription reaction mixture was subjected to diethyl-aminoethyl (DEAE) anion exchange chromatography in order to remove the enzymes and non-converted rNTPs. A ~10 mL gravity-flow DEAE-sepharose column was washed with ~100 mL water, incubated with 0.1 % DEPC over night to inactivate RNases, washed with ~200 mL hot water (~60 °C) to remove residual DEPC and equilibrated with 70 mL 0.1 M sodium acetate, pH 5.5. The transcription reaction mixture was centrifuged (8000 x g, 4 °C, 30 min) to remove precipitated salt. The salt pellet was washed with 10 mL water and re-pelleted by centrifugation (8000 x g, 4 °C, 20 min). The combined supernatants were loaded to the column. After washing with 30 mL 0.1 M sodium acetate pH 5.5, the RNA was eluted by developing the column with 0.6 M (100 mL), 1 M (70 mL) and 2 M (50 mL) sodium acetate, pH 5.5. The eluate was collected in fractions of ~9 mL. For all fractions, the UV absorbance at 260 nm (A_{260}) was measured on a

NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific). Fractions with a significant absorbance were analysed by denaturing polyacrylamide gel electrophoresis (see section 2.1.3). All target RNA containing fractions were diluted with water to a salt concentration of 0.6 M sodium acetate, diluted with abs. ethanol to an ethanol concentration of 70% v/v and cooled at -20°C over night. The precipitated RNA was pelleted by centrifugation (8500 x g, 4 °C, 30 min), air-dried at room temperature for 15 min and reconstituted in water at a concentration of ~100 A_{260} units for purification via HPLC or at a concentration of ~1000 A_{260} units for purification via denaturing polyacrylamide gel electrophoresis.

2.3.7. RNA purification by HPLC

After DEAE anion exchange chromatography, reversed-phase HPLC purification was performed by Elke Stirnal on a Hitachi D-7000 HPLC system with a UV/vis detector using a Perfectsil RP 18 column (10 x 250 mm) and the following two solvents: (A) 50 mM KH_2PO_4/K_2HPO_4, 2 mM tetrabutylammonium bisulfate, pH 5.9, (B) Solvent A with 60% v/v acetonitrile. In case of wt and apoB_{STAB} ^{112}Asw, a separation of the riboswitch and the hammerhead RNA was achieved by a gradient from 42% B to 48% B over 25 min at a flow rate of 5 mL/min. For the *pbuE* adenine aptamer and the *xpt* guanine aptamer, a linear gradient from 42% B to 45% B over 25 min at a flow rate of 5 mL/min was used to separate the target RNA from transcription abortion and degradation fragments. Fractions were collected manually. All obtained HPLC fractions were analyzed by denaturing polyacrylamide gel electrophoresis. Fractions of comparable purity were lyophilized, reconstituted in water, combined, and desalted by solvent exchange into water via Vivaspin 20 centrifugal concentrators (5 kDa cutoff; Sartorius). All centrifugation steps were carried out at 4500 x g and 8 °C. The solvent exchange was performed until a constant UV-absorbance ratio A_{260}/A_{200} was reached. The RNA was subsequently concentrated to ~50 A_{260} units and precipitated via addition of 5 vol 2% lithium perchlorate in acetone and subsequent cooling at -20 °C over night. The RNA was pelleted by centrifugation (8500 x g, 4°C, 30 min), washed with 70% ethanol, centrifuged again (8500 x g, 4°C, 15 min), air-dried at room temperature for 15 min and reconstituted in water at ~300 A_{260} units for folding.

2.3.8. RNA purification by preparative polyacrylamide gel electrophoresis

As a substitute for HPLC purification, ^{127}Asw was purified via a 16.0 x 47.5 cm 8% denaturing polyacrylamide gel using a TVS1000 sequencer electrophoresis system (Biostep) connected to a high-voltage electrophoresis power supply (PowerPac 3000, Biorad). The gel was prepared with 250 mL 8% v/v acrylamide bis-acrylamide (29:1) and 7 M urea in 2x TBE buffer (1x: 90 mM Tris, 90 mM boric acid, 2 mM EDTA, pH 8.0) and polymerized by addition of 0.1% w/v ammonium persulfate and 0.1% v/v TEMED. The RNA sample containing ^{127}Asw and hammerhead RNA (0.7 mL, ~1000 A_{260} units) was mixed with an equal volume of formamide loading buffer (90% v/v formamide, 90 mM Tris, 90 mM boric acid, 2 mM EDTA) and loaded adjacent to a small reference lane containing formamide loading buffer with 0.01% w/v bromophenol blue and 0.01% w/v xylene cyanol. The gel was run in TBE buffer for 5 h at a constant power of 50 W under cooling with the integrated ventilator of the TVS1000 sequencer.

The bands were visualized at the long edge of the gel by UV-shadowing (254 nm). The target band was excised avoiding the illuminated region and cut into three slices that were each granulated by pushing the slice through a syringe and subsequently frozen at -80 °C for 15 min in order to facilitate the elution of the RNA. The granulated slices were eluted in separate falcon tubes, each containing 16.5 mL 0.3 M sodium acetate pH 5.5, by fixing the tubes on top of an Eppendorf thermoshaker (700 rpm, r.t., over night). The combined eluates were filtered through a sterile Corning® bottle top vacuum filer system with 0.2 µm pore size (Sigma-Aldrich), diluted with 2.3 vol abs. ethanol and cooled at -20 °C for 3 h to precipitate the RNA. The RNA was pelleted by centrifugation (8500 x g, 4°C, 30 min), air-dried at room temperature for 15 min and reconstituted in water at ~50 A_{260} units. The RNA was then precipitated by addition of 5 vol 2% lithium perchlorate in acetone and cooling at -20 °C for at least 4 h, pelleted by centrifugation (8500 x g, 4°C, 30 min), air-dried at room temperature for 15 min and reconstituted in water at ~50 A_{260} units. The lithium perchlorate precipitation was repeated two times to minimize contamination by low-molecular weight acrylamide. After the third lithium perchlorate precipitation, the RNA was reconstituted in water at ~300 A_{260} units for folding.

2.3.9. RNA folding and buffer exchange

After purification, all *in vitro* transcribed RNAs were folded at high concentration in water (~300 A_{260} units) by denaturation at 95 °C for 5 min following rapid dilution with 9 vol ice cold water and incubation in ice for 1 h. The folded RNAs were then buffer exchanged into 99.9% NMR buffer (25 mM K_2HPO_4/KH_2PO_4, 50 mM KCl, pH 6.2) by repeated dilution cycles in Vivaspin 2 centrifugal concentrators (5 kDa cutoff; Sartorius). All centrifugation steps were carried out at 4500 x g and 6 °C. The RNAs were then concentrated to concentrations between 0.3 and 0.6 mM and stored at 4 °C. The RNA concentrations were determined by UV/vis spectroscopy (see section 2.1.1). Homogenous folding of the RNAs was verified by native polyacrylamide gel electrophoresis (see section 2.1.4).

2.4. Preparation of fluorophore-labelled riboswitch RNAs

2.4.1. Construct design

For this PhD thesis, several Cy3/Cy5-labelled constructs of the 112-nucleotide and the 127-nucleotide *add* adenine riboswitch from *Vibrio vulnificus* CMCP6 were prepared. The constructs were assembled out of three oligonucleotide fragments by DNA-splinted enzymatic ligation (Figure 12A). The oligonucleotides were purchased from Dharmacon. Their sequences are shown in Table 6. Each oligonucleotide fragment contained a single 5-aminoallyl-modified uridine residue at L2 (U36), L3 (U62) and P5 (U92), respectively, and could be inserted in the unlabelled form or in a fluorophore-labelled form. For the fluorophore-labelled fragments, Cy3 or Cy5 was attached to the 5-aminoallyl-modified uridine by peptide coupling with an *N*-hydroxysuccinimide ester derivative of the dye (Figure 12B). Using this fragment-based approach, the following constructs were prepared. The detailed preparation is described in the next sections.

- Four single-fluorophore-labelled ^{112}Asw wt constructs
    - L2-Cy3
    - L3-Cy5
    - P5-Cy3
    - P5-Cy5
- Three dual-fluorophore-labelled ^{112}Asw wt FRET constructs
    - L2-Cy3 / L3-Cy5

- L2-Cy3 / P5-Cy5
- L3-Cy5 / P5-Cy3

- One dual-fluorophore-labelled ^{112}Asw mutant apoB$_{STAB}$ FRET construct
  - L2-Cy3 / L3-Cy5
- One dual-fluorophore-labelled ^{112}Asw mutant apoA$_{STAB}$ FRET construct
  - L2-Cy3 / L3-Cy5
- Two dual-fluorophore-labelled ^{127}Asw wt FRET constructs
  - L2-Cy3 / L3-Cy5
  - L3-Cy5 / P5-Cy3
- Two dual-fluorophore-labelled ^{127}Asw apoB$_{STAB}$ FRET constructs
  - L2-Cy3 / L3-Cy5
  - L3-Cy5 / P5-Cy3
- Two dual-fluorophore-labelled ^{127}Asw apoA$_{STAB}$ FRET constructs
  - L2-Cy3 / L3-Cy5
  - L3-Cy5 / P5-Cy3

Table 6: Oligonucleotide sequences (written in 5' to 3' direction) used for the DNA splinted ligations [a]

Fragment 1	
^{112}Asw, ^{127}Asw wt	GCUUCAUAUAAUCCUAAUGAUA(5-N-U)GGUUUGGGAGUUUCUAC
^{112}Asw, ^{127}Asw apoB$_{STAB}$	GCGCCAUAUAAUCCUAAUGAUA(5-N-U)GGUUUGGGCGUUUCUAC
^{112}Asw, ^{127}Asw apoA$_{STAB}$	GCUUCAUAUAAUCCUCGUGAUA(5-N-U)GGUCGGGGAGUUUCUAC

Fragment 2	
^{112}Asw, ^{127}Asw all variants	p-CAAGAGCC(5-N-U)UAAACUCUUGAUUAUGAAGUCUGUC

Fragment 3	
^{112}Asw all variants	p-GCUU(5-N-U)AUCCGAAAUUUUAUAAAGAGAAGACUCAUGAAU-bi
^{127}Asw, all variants	p-GCUU(5-N-U)AUCCGAAAUUUUAUAAAGAGAAGACUCAUGAAUUAC UUUGACCUGCCG-bi

DNA splint [b]
ATTCATGAGTCTTCTCTTTATAAAATTTCGGATAAAGCGACAGACTTCATAATCAAGAGTT TAAGGCTCTTGGTAGAAACTCCCAAACCATATCATTAGGATTATATGAAGC

[a] Mutated residues are underlined, dye-modification sites are highlighted in yellow; abbreviations: 5-aminoallyl-modified uridine (5-N-U), 5'-phosphate (p), 3'-biotin linker (bi)
[b] The DNA splint with the ^{112}Asw wt reverse-complement sequence was used for all construct preparations

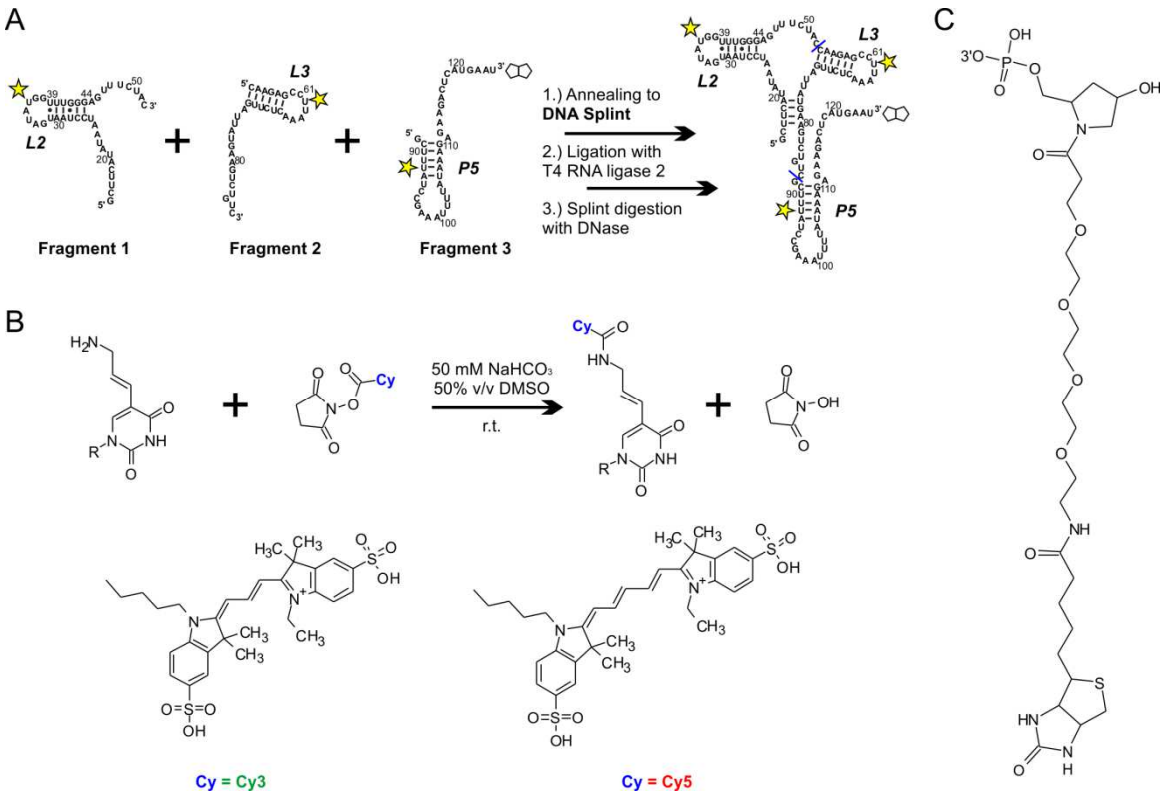

Figure 12: (A) Reaction scheme for the assembly of the full-length *add* adenine riboswitch (here: ^{112}Asw wt) from 3 oligonucleotide fragments by DNA-splinted enzymatic ligation. The ligation sites in the full-length ligation product are indicated in blue. 5-aminoallyl-modified uridine residues for dye-attachment are indicated with yellow stars. The 3'-biotin linker for surface immobilization is represented as a polygon. **(B)** Reaction scheme for the cyanine dye-coupling to 5-aminoallyl-modified uridines. **(C)** Chemical formula of the 3'-biotin modification.

2.4.2. Dye coupling and deprotection of RNA oligonucleotides

The 2'-ACE protected RNA oligonucleotides, the *add* Asw fragments 1, 2 and 3 (Table …), were obtained as solid pellets from Dharmacon and each reconstituted at a concentration of 1 mM in water. For dye coupling and deprotection, 30 µL oligonucleotide solution were first subjected to an ethanol precipitation by adding 0.1 vol 3 M sodium acetate pH 5.2 and 2.3 vol abs. ethanol, and cooling at -80 °C for 10 min. The RNA was pelleted by centrifugation in a benchtop centrifuge (13000 rpm, -4 °C, 15 min), dried by heating at 95 °C for 30 s and reconstituted in 20 µL 100 mM sodium bicarbonate for dye coupling, or in 200 µL deprotection buffer (100 mM acetic acid adjusted to pH 3.8 with TEMED; Dharmacon) for deprotection.

For dye coupling, the dye of a Cy3 or Cy5 Mono-Reactive dye pack vial (Amersham) was dissolved in 20 µL DMSO and mixed with the oligonucleotide in 20 µL 100 mM sodium bicarbonate. The mixture was incubated under exclusion of light at room temperature for

90 min. The RNA was then precipitated by addition of 0.1 vol 3 M sodium acetate pH 5.2 and 2.3 vol abs. ethanol and cooling at -80 °C for 10 min. After centrifugation in a benchtop centrifuge (13000 rpm, -4 °C, 20 min) the pellet was air-dried for 5 min and reconstituted in 200 µL deprotection buffer.

For deprotection, dye-labelled or unlabeled RNA oligonucleotides in 200 µL deprotection buffer were heated at 60 °C for 30 min. For the 3'-biotinylated *add* Asw fragments 3, the incubation time was increased to 2 h. Dye-labelled oligonucleotides were incubated under exclusion of light. After incubation, the RNA was precipitated by addition of 0.1 vol 3 M sodium acetate pH 5.2 and 2.3 vol abs. ethanol, and cooling at - 80 °C for 10 min. The precipitated RNA was pelleted by centrifugation in a benchtop centrifuge (13000 rpm, -4 °C, 20 min) and air-dried for 5 min. Unlabeled RNA oligonucleotides were reconstituted in 200 µL water, subjected to UV/vis spectroscopy for concentration measurement (see section 2.1.1) and stored at -20 °C. Dye-labelled RNA oligonucleotide samples were reconstituted in 60 µL TEAA buffer (100 mM triethylamine adjusted to pH 7.0 with acetic acid) for HPLC purification.

2.4.3. HPLC purification of dye-labelled RNA oligonucleotides

HPLC purification of dye-labelled RNA oligonucleotides was carried out on an Äkta purifier using a Kromasil C8 HPLC column (250 x 4.6 mm) and the following two solvents: (A) 100 mM triethylamine adjusted to pH 7.0 with acetic acid, (B) acetonitrile. Prior to loading, the oligonucleotide samples were heated at 60 °C for 3 min. Separation of the unlabeled and dye-labelled RNA was achieved by employing a linear gradient from 0% B to 100% B over 160 min at a flow rate of 0.5 mL/min. The elution was monitored by simultaneous detection of the UV/vis absorption at the RNA and the fluorophore absorption maximum. Fractions of dye-labelled RNA were collected manually. The fractions were subjected to centrifugation in an Eppendorf Vacufuge plus concentrator (10 min, 30 °C, vacuum setting V_{aq}). The RNA was then precipitated by addition of 0.1 vol 3 M sodium acetate pH 5.2 and 2.3 vol abs. ethanol and cooling at -80 °C for 10 min. After centrifugation in a benchtop centrifuge (13000 rpm, -4 °C, 15 min) the RNA pellets of each fraction were air-dried for 5 min and reconstituted in water at ~30 A_{260} units. The fractions were analyzed by denaturing polyacrylamide gel electrophoresis (see section 2.1.3) and by UV/vis spectroscopy. Fractions that showed a single band on

the denaturing gel and a comparable UV/vis absorption ratio at the fluorophore over the RNA absorption maximum were combined and stored under exclusion of light at -20 °C.

2.4.4. DNA splinted enzymatic RNA ligation

The fluorophore-labelled full-length *add* Asw constructs (see section 2.4.1) were assembled from the three oligonucleotide fragments by DNA splinted ligations with T4 RNA ligase 2 (New England Biolabs). Ligation reactions were performed under exclusion of light on 450 pmol scale. The three oligonucleotides and the DNA splint were annealed at 2 µM concentration in T4 RNA ligase 2 buffer (New England Biolabs) containing 20% v/v PEG 8000 by heating at 75 °C for 3 min and cooling at room temperature for 10 min. The mixture was incubated with 200 U/mL T4 RNA ligase 2 (New England Biolabs) at 37 °C for 1 h. Then 70 U/mL T4 RNA ligase 2 were added and the incubation was continued for 30 min, before 70 U/mL Turbo DNase (Ambion) were added for another 30 min incubation time. Finally, the mixture was extracted three times with 1 vol Roti®-Aqua-P/C/I (Carl Roth) and twice with chloroform. In each extraction step rapid phase separation was achieved by centrifugation in a benchtop centrifuge (2000 rpm, 4 °C, 3 min). The RNA was then precipitated from the aqueous phase by addition of 0.1 vol 5 M ammonium acetate and 2.5 vol abs. ethanol, and cooling at -80 °C for 10 min. The precipitated RNA was pelleted by centrifugation in a benchtop centrifuge (13000 rpm, -4 °C, 30 min), dried in an Eppendorf Vacufuge plus concentrator (2 min, 30 °C, vacuum setting V_{al}) and reconstituted in a mixture of 12 µL water and 12 µL formamide buffer (90% v/v formamide, 90 mM Tris, 90 mM boric acid, 2 mM EDTA) for purification by denaturing polyacrylamide gel electrophoresis.

2.4.5. RNA purification

The full-length *add* Asw RNA was purified from ligation by-products by denaturing polyacrylamide gel electrophoresis using a Multigel system (Biometra). Gels were prepared with 10 mL 10% v/v acrylamide bis-acrylamide (29:1) in 2x TBE buffer (1x: 90 mM Tris, 90 mM boric acid, 2 mM EDTA, pH 8.0) with 8 M urea and polymerized by addition of 0.1% w/v ammonium persulfate and 0.1% v/v TEMED. The gel was pre-run in TBE buffer at 225 V for 10 min. The crude ligation product was loaded adjacent to a

reference lane containing 3% of the ligation product. The gel was then run under exclusion of light at 225 V for 30 min. The target band was identified by eye, excised and frozen at -80 °C for ≥10 min. Successful excision of the desired band was subsequently verified by imaging the gel via fluorescence detection on a Tayphoon scanner (GE Healthcare). The RNA was eluted from the gel slice in 400 µL 0.5 M ammonium by shaking in an Eppendorf thermoshaker (450 rpm, r.t., over night) and precipitated by adding 2.5 vol abs. ethanol and cooling at -80 °C for 1.5 h. The precipitated RNA was pelleted by centrifugation in a benchtop centrifuge (13000 rpm, -4 °C, 60 min), dried in an Eppendorf Vacufuge plus concentrator (5 min, 30 °C, vacuum setting V_{al}) and reconstituted in 17 µL water. The concentration was determined by UV/vis spectroscopy via the absorbance of the RNA (see section 2.1.1). The sample was stored under exclusion of light at -20 °C.

2.4.6. Validation of RNA folding

Prior to biophysical analysis, small aliquots of the fluorophore-labelled *add* Asw constructs were folded at low nM concentration in smFRET immobilization buffer (25 mM potassium phosphate pH 7.0, 50 mM potassium chloride) by heating at 85 °C for 3 min and subsequent cooling at room temperature for 15 min. For each construct, it was verified in a separate experiment that the RNA exhibited a homogenous fold after this folding protocol. For this, 0.5 pmol RNA were folded in 10 µL smFRET buffer and analysed via native polyacrylamide gel electrophoresis (see section 2.1.4).

2.5. NMR spectroscopy

All NMR samples were prepared in NMR buffer (25 mM K_2HPO_4/KH_2PO_4, 50 mM KCl, pH 6.2) with 90% H_2O / 10% D_2O and 100 µM DSS as internal ^{1}H chemical shift standard. ^{15}N chemical shifts were indirectly referenced with the Ξ ratio 0.101329118 (Wishart *et al.*, 1995). All spectra were acquired at Bruker Avance spectrometers and processed in Topspin 3.1.

$^{1}H,^{15}N$-BEST-TROSY spectra of RNA imino protons were acquired on an 800 MHz spectrometer with a 5 mm TXI cryo ^{1}H, ^{15}N, ^{13}C Z-GRD probe using a pulse program with the modifications proposed by Brutscher *et al.* (Favier and Brutscher, 2011; Solyom

et al., 2013). In the direct dimension, 2048 data points were recorded at a spectral width of 25 ppm with the ^{1}H transmitter frequency offset at the water resonance frequency. For the indirect dimension, 256 points were collected at a spectral width of 28 ppm with the ^{15}N transmitter frequency offset at 153 ppm. The INEPT delay was set to 2.7 ms. The lengths of the BEST-TROSY shaped pulses were calculated such that optimal excitation of the imino proton chemical shift region of 12.4 ± 3.0 ppm was achieved. The relaxation delay was set to 0.5 s. For processing, 1024*256 data points were Fourier transformed employing a shifted quadratic sine bell window function (sine bell shift 2). A phase correction and an automated baseline correction in the ^{1}H dimension of the spectrum between 5 and 16 ppm was performed. The chemical shifts were corrected for the TROSY shift by subtracting 47 Hz from the ^{1}H spectrum reference frequency and adding 47 Hz to the ^{15}N spectrum reference frequency. The spectra were further evaluated in Sparky 3.114 (Goddard and Kneller, 2008) to extract the imino signal intensities as peak heights.

^{1}H,^{15}N-HSQC spectra of RNA imino protons were recorded on a 600 MHz spectrometer with a 5 mm TCI cryo ^{1}H, ^{15}N, ^{13}C Z-GRD probe using the FHSQC pulse sequence with a binomial WATERGATE for water suppression (Mori *et al.*, 1995). In the direct dimension, 2048 data points were acquired at a spectral width of 25 ppm with the ^{1}H transmitter frequency offset at the water resonance frequency. For the indirect dimension, 200 points were collected at a spectral width of 28 ppm with the ^{15}N transmitter frequency offset at 153 ppm. The INEPT delay was set to 2.3 ms. The delay for binomial water suppression was 66.7 µs (2*DW). Heteronuclear decoupling during ^{15}N chemical shift evolution was achieved via a 180° hard pulse on ^{1}H. For composite pulse decoupling during acquisition, the GARP scheme was used (Shaka and Keeler, 1987). The relaxation delay was set to 1.5 s. For processing, 1024*200 data points were Fourier transformed employing a shifted quadratic sine bell window function (sine bell shift 2). A phase correction and an automated baseline correction in the ^{1}H dimension of the spectrum between 5 and 15 ppm was performed. The spectra were further evaluated in Sparky 3.114 (Goddard and Kneller, 2008) to extract the imino signal intensities as peak heights.

^{1}H,^{1}H-NOESY spectra of RNA imino protons were recorded on a 900 MHz spectrometer with a TXI cryo ^{1}H,^{15}N,^{13}C Z-GRD probe using a pulse program with jump-return water suppression (Sklenář and Bax, 1987) and a transmitter frequency offset switch between the direct and the indirect dimension. In the direct dimension, 2048 data points were

recorded at a spectral width of 23.6 ppm with the ^{1}H transmitter frequency offset at the water resonance frequency. For the indirect dimension, 768 points were collected at a spectral width of 16 ppm with the ^{1}H transmitter frequency offset at 7.3 ppm. The mixing time was 150 ms. The jump-return echo delay was 27.5 µs. Heteronuclear decoupling during chemical shift evolution in the indirect dimension was achieved via a 180° hard pulse on ^{15}N. The ^{15}N transmitter frequency offset was set to 153 ppm. For composite pulse decoupling during acquisition, the GARP scheme was used (Shaka and Keeler, 1987). The relaxation delay was 1.5 s. For processing, the FIDs were zero-filled to 4096*2048 data points and Fourier transformed employing a shifted quadratic sine bell window function (sine bell shift 2). A phase correction and an automated baseline correction in the direct dimension of the spectrum between 5 and 15 ppm was performed. Assignment of the spectra was performed in Sparky 3.114 (Goddard and Kneller, 2008).

^{1}H-WaterLOGSY experiments were performed on a 600 MHz spectrometer equipped with a 5 mm TXI ^{1}H, ^{15}N, ^{13}C Z-GRD probe using the pulse sequence developed by Dalvit *et al.* (Dalvit *et al.*, 2000, 2001). 16384 data points were recorded at a spectral width of 14 ppm with the ^{1}H transmitter frequency offset at the water resonance frequency. To suppress RNA resonances, a Carr-Purcell-Meiboom-Gill (CPMG) pulse train was employed after NOE mixing (Van, Chmurny and Veenstra, 2003). Water suppression was achieved by excitation sculpting. The NOE mixing time and relaxation delay were 2 s and 2.5 s, respectively. For processing, the FIDs were zero-filled to 32768 data points and Fourier transformed employing an e-function (line broadening factor 4 Hz). The spectra were phase corrected and manually baseline corrected between 7 and 9 ppm. Relative signal intensities of a series of waterLOGSY spectra were obtained by integration via the ERETIC tool.

^{7}Li 1D experiments were performed on a 300 MHz spectrometer with a 5 mm BBFO probe. 16384 data points were recorded at a spectral width of 32 ppm with the transmitter frequency offset at 0 ppm. The relaxation delay was set to 120 s. For processing, the FIDs were zero-filled to 32768 data points and Fourier transformed employing an e-function (line broadening factor 1 Hz). The spectra were phase corrected and baseline corrected.

^{7}Li inversion recovery experiments of Li^{+} were performed on a 300 MHz spectrometer with a 5 mm BBFO probe using the Bruker t1ir pulse program. In the direct dimension, 37280 data points were recorded at a spectral width of 32 ppm with the transmitter

frequency offset at 0 ppm. For the indirect dimension of the pseudo-2D experiment, a variable delay list was specified with the inversion recovery delays 180 s, 60 s, 40 s, 30 s, 20 s, 15 s, 10 s, 5 s, 3 s, and 1 s. The relaxation delay was set to 180 s. For processing, the FIDs were stripped to 32768 data points and Fourier transformed employing an e-function (line broadening factor 1 Hz). The spectra were phase corrected and baseline corrected. The relative intensities of the Li^+ resonance were determined by integration with the ERETIC tool.

2.6. Fluorescence binding assays with 2-aminopurine

Equilibrium binding of the fluorescent adenine analogue 2-aminopurine to *add* Asw RNA was monitored by measuring the fluorescence of 2-aminopurine in presence of an increasing excess of *add* Asw. 200 µL samples containing 50 nM 2-aminopurine, 2 mM $MgCl_2$ and 0 - 2.5 µM *in vitro* transcribed ^{112}Asw wt or 0 - 10 µM *in vitro* transcribed ^{112}Asw apoB_{STAB} were prepared in NMR buffer (25 mM K_2HPO_4/KH_2PO_4, 50 mM KCl, pH 6.2) and kept on ice until measurement. Blank samples contained 2 mM $MgCl_2$ in NMR buffer. Fluorescence spectroscopy was performed on a Fluorolog-3 spectrophotometer (Horiba Scientific) at 25 °C. Emission spectra were recorded after a temperature equilibration time of 5 min by accumulating 2 scans at a bandwidth of 9 nm and a response time of 200 ms. The excitation wavelength was 300 nm and the monitored emission range was 330-550 nm at a data interval of 1 nm. Data analysis was performed in Origin 9G. The blank spectra were subtracted from the sample spectra and the fluorescence intensities were picked at 370 nm to calculate the quantity

$$\frac{dF}{F} = \frac{F(0) - F(c)}{F(0)}$$

where F(0) denotes the 2-aminopurine fluorescence intensity in absence of RNA and F(c) the 2-aminopurine fluorescence intensity at the RNA concentration c. Under the approximation that in presence of RNA, the total RNA concentration equals the free RNA concentration, since the RNA is in large excess over 2-aminopurine, and that binding occurs with 1:1 stoichiometry, dF/F as a function of c was fitted by

$$\frac{dF}{F} = (1 - a) * \frac{c}{K_{D,app} + c}$$

to determine the 2-aminopurine dissociation constant ($K_{D,app}$) and the ratio of the fluorescence response coefficients of 2-aminopurine in the RNA bound state over the free state (a).

2.7. Bulk fluorescence anisotropy measurements on Cy3 and Cy5

Bulk fluorescence anisotropy measurements were performed at 20 °C on a FluoroMax-4 spectrophotometer (Horiba Scientific) in L-format geometry. Samples containing 8 nM Cy3, Cy5, Cy3-coupled to ^{112}Asw or Cy5-coupled to ^{112}Asw were prepared in 500 µL smFRET immobilization buffer (25 mM potassium phosphate pH 7.0, 50 mM potassium chloride). For the free dyes, the sample concentrations were adjusted via the fluorophore extinction coefficients (ε_{650}(Cy5) = 250 $mM^{-1}cm^{-1}$, ε_{550}(Cy3) = 150 $mM^{-1}cm^{-1}$), and for the RNA-coupled dyes, the sample concentrations were adjusted via the RNA extinction coefficient (ε_{260}(^{112}Asw) = 1052 $mM^{-1}cm^{-1}$). The samples were heated at 85 °C for 3 min, subsequently cooled at room temperature for 10 min, diluted with an equal volume of 4 mM $MgCl_2$ in smFRET immobilization buffer and then kept on ice until measurement. After a temperature equilibration time of 2 min, the fluorescence anisotropies were measured in accumulations of 20 scans at constant excitation / emission wavelengths 525 nm / 565 nm for Cy3 and 625 nm / 665 nm for Cy5. The integration time was 100 ms and the excitation and emission bandwidth were 7 nm. The fluorescence anisotropy r was calculated as $r = (I_{VV} - G*I_{VH}) / (I_{VV} + 2G*I_{VH})$ where I_{VV} denotes the intensity of vertically polarized emission and I_{VH} the intensity of horizontally polarized emission detected at excitation with vertically polarized light. The correction factor $G = I_{HV}/I_{HH}$ (G = 0.71 for Cy3 and G = 0.51 for Cy5) that corrects for the different detection efficiency of vertically and horizontally polarized light was determined from the intensity I_{HV} of vertically polarized emission and the intensity I_{HH} of horizontally polarized emission measured at excitation with horizontally polarized light. The absence of significant background fluorescence intensity in the measurement of I_{VV}, I_{VH}, I_{HV} and I_{HH} was confirmed by measuring blank samples of 2 mM Mg^{2+} in smFRET immobilization buffer.

2.8. Single-molecule FRET spectroscopy

2.8.1. Sample preparation

For single-molecule fluorescence measurements, 0.5 pmol FRET construct were folded at 1 nM concentration in smFRET immobilization buffer (25 mM K_2HPO_4/KH_2PO_4, 50 mM KCl, pH 7.0) by 3 min denaturation at 85 °C and 15 min cooling at room temperature. The folded FRET construct was diluted to a series of 100 pM RNA samples in smFRET immobilization buffer that were supplemented with different concentrations of $MgCl_2$ and adenine or PEG 6000. The samples were kept on ice until measurement within 24 h.

For all but PEG 6000 containing samples, measurement slides were assembled from standard glass microscope slides and 24 x 60 mm glass coverslips (Carl Roth). The glass microscope slides were washed in 1 M KOH for 15 min, in water for 15 min and dried with compressed air. The glass coverslips were cleaned in N_2 plasma for 10 min (Diener Electronic). Small stripes of parafilm were vertically aligned along the long edge of the glass slide with a ~3 mm spacing to confine measurement channels with ~10 µL volume. The cleaned surface of the coverslip was gently pressed onto the parafilm stripes and the assembly was fixed by placing the slide onto an 85 °C heatblock for 1 min. Shortly before measurement, the channels were flushed with 30 µL 1 mg/mL biotinylated BSA (Sigma Aldrich) in buffer T50 (10 mM Tris-HCl, pH 8.0, 50 mM NaCl) and incubated for 10 min. After washing with 100 µL buffer T50, the channels were flushed with 30 µL 0.2 mg/mL streptavidin (Molecular Probes, Thermo Fisher Scientific) in buffer T50 and incubated for 10 min. The channels were then washed with 100 µL smFRET immobilization buffer. The BSA-biotin-streptavidin coated slides were used within 12 h.

Each 100 pM RNA sample was immobilized directly before measurement. For this, a channel was first flushed with 200 µL smFRET immobilization buffer containing equivalent concentrations of $MgCl_2$ and adenine to the sample. Then, 50 µL of the RNA sample were flushed through the channel. After 1min incubation, the channel was rinsed with 125 µL imaging buffer. The imaging buffer included all components of the immobilisation buffer, the respective concentrations of $MgCl_2$ and adenine, an enzymatic oxygen scavenging system (10% glucose, 14 U/mL glucose oxidase, 1000 U/mL catalase) and saturating amounts of trolox. The immobilization procedure typically resulted in an optimal surface density for data acquisition (~300 molecules per field of view; i.e. ~0.1 molecules per µm²). If this was not the case, the RNA sample was re-immobilized in a

fresh channel by diluting the RNA concentration of the sample or by flushing the channel with multiples of 50 µL of the RNA sample.

smFRET measurements of viscous RNA samples with PEG 6000 were performed on flexiPERM slides (Sarstedt). The flexiPERM silicon attachment was placed on a plasma-cleaned 24 x 60 mm glass coverslip (Carl Roth) for compartmentalization of 8 measurement chambers. To immobilize the FRET construct, the chambers were filled with 150 µL 1 mg/mL biotinylated BSA (Sigma Aldrich) in buffer T50 and incubated for 10 min. After washing twice with 150 µL buffer T50, the chambers were incubated with 150 µL 0.2 mg/mL streptavidin (Molecular Probes, Thermo Fisher Scientific) in buffer T50 for 10 min. Next, the chambers were washed twice with 150 µL smFRET immobilization buffer. Immediately before measurement of an RNA sample, a chamber was washed twice with 150 µL smFRET immobilization buffer containing adjusted $MgCl_2$ and PEG concentrations. The chamber was then filled with 150 µL of the RNA sample, and, after 1 min, washed with 150 µL smFRET immobilization buffer containing adjusted $MgCl_2$ and PEG. Then, 200 µL imaging buffer were applied. In addition to the components of the smFRET immobilization buffer, $MgCl_2$ and PEG, the imaging buffer included an enzymatic oxygen scavenging system (10% glucose, 14 U/mL glucose oxidase, 1000 U/mL catalase) and saturating amounts of trolox.

2.8.2. smFRET measurement and analysis

smFRET experiments were performed at room temperature (22 °C) on a custom objective type total internal reflection microscopy setup with 532 nm laser excitation (Figure 13). The microscope (IX71, Olympus) was equipped with a green (532 nm, 150 mW) and a red (637 nm, 140 mW) diode laser (OBIS, Coherent), an acousto-optical tunable filter (AOTFnC-400.650-TN, AA Opto-Electronic), a NA 1.4 100x oil immersion objective (UPLSAPO100XO, Olympus), a dual view emission splitting system (DV2, Photometrics) and an electron multiplying charge coupled device (EMCCD) camera (iXon 897, Andor Technology) that was computer-operated by Andor Solis 4.2. The ENCCD camera was synchronized with the AOTF via a PCI delay generator (BME_SG08p, Bergmann Messgeräte Entwicklung). Before measurement of smFRET data, a 20-frame calibration *.tif movie of a 1:500 dilution of 0.1 µm TetraSpeck Microspheres (Thermo Fisher Scientific) in phosphate buffered saline was acquired. All

immobilized RNA samples were imaged ~3 min after buffer exchange into imaging buffer. The fluorescence emission of Cy3 and Cy5 was recorded at an integration time of 100 ms with maximum EM gain. For smFRET histogram data, ~20 20-frame *.tif movies were acquired per sample. For smFRET time traces, ~10 1500-frame *.tif movies were recorded per sample. Prior to data analysis, the *.tif movies were rotated in the windows explorer such that the acceptor channel was shown on top of the donor channel. The movies were then converted to 16 bit integer *.raw format using the batch file conversion tool in Andor Solis 4.2.

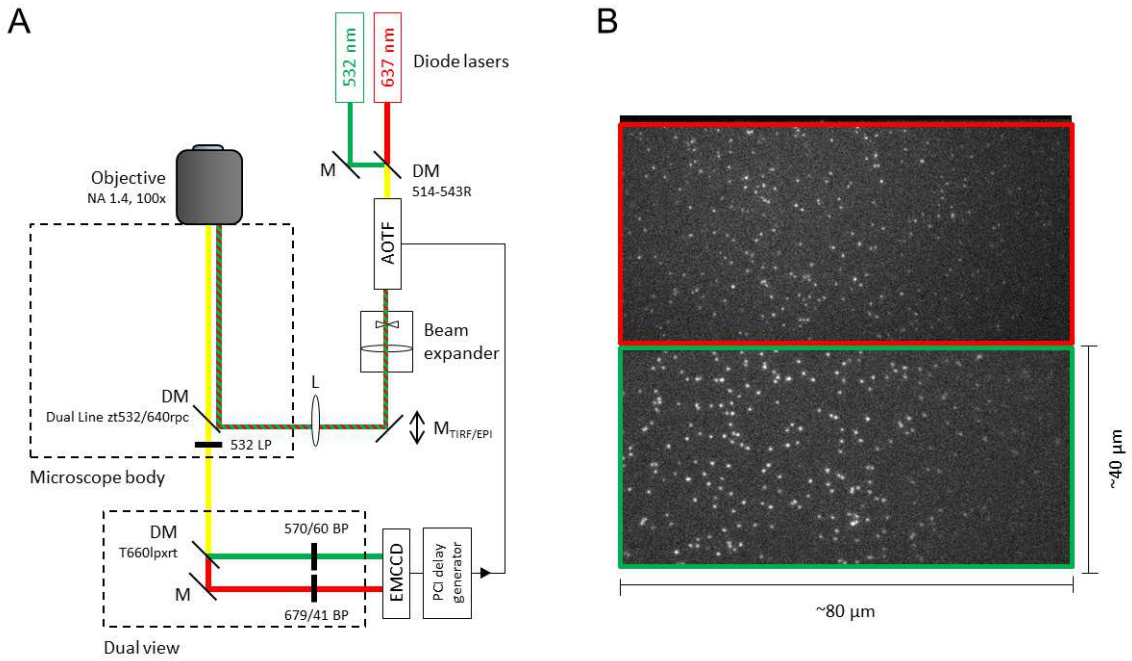

Figure 13: (A) Schematic representation of the employed objective-type total internal reflection fluorescence microscopy setup. The red and the green laser beam are aligned with a dichroic mirror (DM) and pass an acousto-optical tunable filter (AOTF), which is synchronized with an electron multiplying charge coupled device (EMCCD) camera via a PCI delay generator. The AOTF allows running the setup either at continuous 532 nm laser excitation for conventional smFRET experiments or at alternating 532 nm / 637 nm laser excitation for smALEX experiments. The excitation beam is widened with a beam expander. A movable mirror ($M_{TIRF/EPI}$) reflects the beam onto a lense (L) that focuses the beam at the back focal plane of the objective. By transversely moving the position of the mirror one can switch between TIRF and EPI illumination. The Dual Line DM in the microscope body reflects excitation light at 532 nm and at 634 - 640 nm and transmits fluorescence emission at 550 - 615 nm and at 660 - 750 nm. The emission light is further cleaned by a 532 nm long pass (LP) filter and split into the FRET donor and the FRET acceptor channel by a DM and two band pass filters (BP) in the Dual view. The EMCCD camera simultaneously images the donor and the acceptor fluorescence emission in 512 x 512 pixels on an 8192 x 8192 µm chip. **(B)** Sample EMCCD camera image (integration time: 100 ms) of immobilized L2(Cy3)/L3(Cy5)-labelled ^{112}Asw molecules with continuous fluorescence excitation at 532 nm in TIRF mode. The dimensions of the imaged field are indicated, considering the 100x magnification of the microscope. The donor and the acceptor channel are framed in green and red, respectively.

For data analysis, a custom program written in IDL (Exelis) and several Matlab scripts (MathWorks) were used. The IDL program was written by Hazen Babcock and Greg Bokinsky, (Bokinsky *et al.*, 2003) and was kindly provided by Martin Hengesbach. It consists of 6 scripts. Two scripts (*beads_co2*, *beads_map*) were used to create a transformation map for donor and acceptor channel alignment from the calibration movie file. Four scripts (*dat_all2*, *ffp_dat2*, *ap_dat2*, *ave_tr_dat*) were used to extract background corrected single-molecule FRET donor and acceptor fluorescence intensities from the smFRET movie files by Gaussian point spread function fitting. The program exported the FRET donor and acceptor fluorescence intensity trajectories of all molecules per movie in a single binary *.traces file. The STD threshold for the peak finding algorithm in *ffp_dat2* was set to 2. Further analysis of the *.traces files was performed in Matlab.

For analysis of smFRET histogram data, a Matlab script was used that was kindly provided by Martin Hengesbach. The script is attached in Appendix 2. It averaged the donor and acceptor fluorescence intensities, I_D and I_A, of each molecule over the first three frames (i.e. 300 ms), corrected the acceptor fluorescence intensities for ~10% leakage of the donor fluorescence into the acceptor channel and calculated the apparent FRET efficiency $E = I_A/(I_A+I_D)$. For all molecules that satisfied the fluorescence intensity threshold $I_A+I_D \geq I_{sum}^{min}$ it exported the FRET efficiency values binned in intervals of 0.025. I_{sum}^{min} was typically set to 4000 or 8000 a.u. Plotting and Gaussian distribution fitting of the FRET histograms was performed in Origin 9G. The FRET peak at E = 0, which arises due to donor-only molecules, was removed from the histograms by subtraction of the corresponding Gaussian. For the two-state FRET histograms of L2/L3-labeled *add* Asw, the fraction of molecules in the docked (high-FRET) state was determined from the Gaussian peak areas of the docked (high-FRET) and the undocked (low-FRET) state. The error was calculated from the fitting errors by Gaussian error propagation.

smFRET trace data of L2/L3-labelled *add* Asw were evaluated according to a newly established analysis routine that makes use of three Matlab scripts (attached in Appendix 3) and the open source Hidden Markov modelling software Hammy (McKinney, Joo and Ha, 2006). A detailed description of the analysis routine is presented in section 3.2.7. Briefly, after correcting the acceptor fluorescence intensities for ~10% donor leakage, fluorescence trajectories of all molecules that satisfied the fluorescence intensity

threshold $I_A+I_D \geq 4000$ a.u. were exported for Hidden Markov modelling of the FRET trajectory (E) alongside the normalized donor (I_D/I_D^{max}) and the normalized acceptor fluorescence intensity trajectory (I_A/I_A^{max}). Hidden Markov modelling was performed with an initial guess of 4 states. The Hidden Markov models of E, I_D/I_D^{max} and I_A/I_A^{max} were compared to unambiguously identify anticorrelated fluorescence intensity transitions that reflect a significant change in FRET efficiency ($\Delta E \geq 0.1$). Based on the identified transitions, a transition occupancy density plot was created and dwell-times were extracted. The dwell-times were sorted in Microsoft Excel 2010 and subjected to thresholding. For ^{112}Asw, all docked dwell-times with $E_{before} > 0.6$ and $E_{after} < 0.6$ and all undocked dwell-times with $E_{before} < 0.6$ and $E_{after} > 0.6$ were considered. For ^{127}Asw the thresholds were $E_{before} > 0.8$ and $E_{after} < 0.6$ for the docked state, and $E_{before} < 0.6$ and $E_{after} > 0.8$ for the undocked state. Using these thresholds, ~80% of the extracted dwell-times were included in the dwell-time analysis. Dwell-time histograms were created in Origin 9.0G with a bin width of 0.5 s and fitted with single-exponential decay functions to extract the rate constants for undocking and docking. The dwell-time analysis was performed in duplicate over smFRET trace data from two independent experiments.

2.8.3. smALEX measurement and analysis

smALEX experiments were performed under the same conditions as the smFRET experiments (see section 2.8.2), except that for smALEX histograms ~20 40-frame *.tif movies were recorded with alternating laser excitation. Prior to data analysis, the *.tif movies were rotated in the windows explorer such that the acceptor channel was shown on top of the donor channel. Data analysis of the *.tif movies was performed with the open source software iSMS (Preus *et al.*, 2015). Two-dimensional smALEX histograms that include donor-only and acceptor-only molecules were created by plotting all 10 coordinate pairs (FRET efficiency E / stoichiometry factor S) measured over 20 frames in a two dimensional grid. One-dimensional histogram slices of E and S were created with a bin width of 0.05.

2.9. FTIR spectroscopy

Fourier transform infrared spectroscopy (FTIR) experiments were performed by Albrecht Völklein and Henrik Gustmann. Briefly, aqueous solutions of hypoxanthine and ^{13}C,^{15}N-

labelled hypoxanthine with and without riboswitch RNA were prepared in NMR buffer (25 mM potassium phosphate pH 6.2, 50 mM potassium chloride) with 16 mM $MgCl_2$ and exchanged into 99.9% D_2O (Deutero) by two cycles of lyophilization. For measurement, 5 µL samples were sealed in sandwich cuvettes assembled of CaF_2 windows (Crystal GmbH) and a 50 µm PTFE spacer (Späh GmbH). The FTIR data were acquired on a Bruker Vertex 80 spectrometer using a Mercury-Cadmium-Telluride (MCT) detector. 128 scans were accumulated over a spectral range of 900 – 4000 cm^{-1} at a resolution of 1 cm^{-1} with an aperture of 2 mm and a mirror speed of 20 kHz. The high-pass filter was open and the low-pass filter was set to 10 kHz. Fourier transformation was performed with a phase resolution of 8 points following a Blackman Harris 3-term apodization with a zero filling factor of 4. Processing and creation of FTIR difference spectra was performed in Opus 7.2. All spectra were H_2O vapour corrected. Moreover, HDO and buffer related peaks were removed by subtraction of the respective spectra.

Chapter 3: Integrated NMR and smFRET analysis of ^{112}Asw

3.1. Introduction

Based on NMR spectroscopic investigations of the full-length 112-nucleotide *add* adenine riboswitch from *Vibrio vulnificus* (^{112}Asw), Reining *et al.* proposed a 3-state model for its regulatory mechanism (see section 1.5). This model assumed a single folding state for the apoB the apoA and the holo conformation. However, previous single-molecule FRET studies of P1 stabilized variants of the *add* Asw aptamer domain have shown that the isolated apoA-form aptamer can adopt two distinct folding states with undocked and docked aptamer kissing loop motif (Lemay *et al.*, 2006; Dalgarno *et al.*, 2013; Liu *et al.*, 2015). A 3-state model thus might not be sufficient to describe the regulatory mechanism of the *add* adenine riboswitch. Moreover, the available NMR data of the full-length *add* Asw contained only limited information on the tertiary structure of ^{112}Asw. Specifically, it remains unclear, whether long-range tertiary structural interactions between aptamer and expression platform in the apoB, in the apoA or in the holo conformation contribute to the relative stability of these conformers.

Besides NMR spectroscopy, routinely employed methods to characterize the solution structure of riboswitches are structural probing techniques, small-angle-scattering, single-molecule FRET and single-molecule force spectroscopy (Liberman and Wedekind, 2012). Of these methods, NMR and single-molecule FRET (smFRET) exhibit a particularly striking complementarity. First, NMR and smFRET reveal complementary RNA structural information. NMR of the RNA imino protons yields information on base pairing interactions (Fürtig *et al.*, 2003) and base pair stabilities (Steinert, Rinnenthal and Schwalbe, 2012) at nucleotide resolution. By using time-resolved NMR methods such as N_{ZZ} exchange or rapid mixing, RNA refolding kinetics can be characterized at timescales ranging from sub-seconds to hours (Rinnenthal *et al.*, 2011). smFRET is a powerful method to uncover heterogeneity of long-range intramolecular distances in the RNA conformational ensemble (Xie *et al.*, 2004). When conducted on immobilized molecules, smFRET can reveal real-time RNA folding dynamics in the range of sub-seconds to minutes, a similar timescale to the one accessible by NMR (Al-Hashimi and Walter, 2008). Moreover, NMR and smFRET complement one another regarding experimental conditions. NMR requires half-millimolar RNA concentration and ^{13}C/^{15}N isotope-labelled samples. smFRET operates at sub-nanomolar RNA concentration and uses dual-

fluorophore labelled RNA constructs. The low RNA concentration employed by smFRET is closer to the physiological RNA concentration (~1 nM) (Church *et al.*, 2000) and allows to titrate riboswitch ligand and metal ion cofactors in large excess over the RNA.

The major project of this PhD thesis, which is presented in this chapter, was an integrated NMR and smFRET spectroscopic investigation of the ligand-induced conformational switch of ^{112}Asw at ambient temperature. This is the first integrated NMR and smFRET spectroscopic study of the very same sequence of a purine riboswitch which allowed direct comparison and integration of the results of these complementary biophysical techniques. The first aim of this investigation was to obtain an enhanced model for the regulatory mechanism of the *add* Asw. Specifically, this meant to deduce the actual number of folding states of ^{112}Asw, to characterize both their secondary and tertiary structure and to establish the pathways of their interconversion by integrating complementary solution structural information from NMR and smFRET. The second aim was to characterize by smFRET how varying concentrations of adenine and Mg^{2+} in large excess over the RNA affect the conformational equilibrium of ^{112}Asw. The smFRET experiments shed light on how ligand and cofactor modulate the folding of ^{112}Asw at physiologically representative concentrations of all interaction partners.

The here presented NMR experiments are a direct follow-up on the NMR spectroscopic study of ^{112}Asw performed by Reining *et al.* (Reining *et al.*, 2013). In section 3.2.1, the adenine-induced switch of the global base pairing structure of ^{112}Asw is characterized. This characterization was achieved by evaluating the adenine-dependent intensity of selected imino reporter signals in two-dimensional ^{1}H,^{15}N-correlation spectra. In section 3.2.2, it is demonstrated that Li^{+} carryover from the lithium perchlorate precipitation performed during RNA sample preparation can critically affect the results of this evaluation. In section 3.2.3, an apoB-stabilized mutant of ^{112}Asw is described. This mutant was devised to verify the proposed secondary structure as well as the hypothesized adenine binding incompetency of the apoB conformation.

The here presented smFRET investigation of ^{112}Asw was the first smFRET analysis of RNA performed in the Schwalbe and Heilemann labs. The smFRET experiments were established in collaboration with Martin Hengesbach and Philipp Höllthaler. In this study, three Cy3/Cy5-FRET labelling schemes were employed. An aptamer loop labelling scheme (L2/L3) was adopted from previous studies of the isolated P1-stabilized *add* Asw

aptamer (Lemay *et al.*, 2006; Dalgarno *et al.*, 2013) and reports on the folding of the aptamer kissing loop motif in the full-length riboswitch. Additionally, two aptamer / expression platform labelling schemes (L2/P5 and L3/P5) were devised that report on distance changes across the two riboswitch domains. A quality validation of the three labelling schemes based on bulk fluorescence anisotropy measurements is presented in section 3.2.4. In section 3.2.5, the global adenine-induced switch of ^{112}Asw is described as monitored by static smFRET in the three labelling schemes. In section 3.2.6, a detailed analysis of the aptamer kissing-loop docking equilibrium of ^{112}Asw in the L2/L3-labelling scheme by static smFRET is presented. To discern FRET signatures of the apoB and the apoA conformation, the wild-type was compared with an apoA-stabilized mutant, (Reining *et al.*, 2013) and the apoB-stabilized mutant introduced in section 3.2.3. The effects of varying adenine concentration, varying Mg^{2+} concentration and varying concentration of the macromolecular crowding agent PEG 6000 was investigated. In section 3.2.7, a software routine is described that was established in this thesis in order to analyse the aptamer folding dynamics of ^{112}Asw. The adenine- and Mg^{2+}- dependent aptamer docking dynamics of L2/L3-labelled ^{112}Asw are presented in section 3.2.8.

In section 3.3, the two perspectives of the conformational equilibrium of the *add* Asw obtained by NMR and smFRET are discussed, and new implications for the molecular mechanism of this riboswitch are highlighted.

3.2. Results

3.2.1. NMR analysis of the adenine-dependent base pairing

The adenine-induced switch in the base pairing of ^{112}Asw was investigated by evaluating the adenine-dependent signal intensity of guanosine and uridine base pair reporter imino resonances. The imino signal intensities were measured in ^{1}H,^{15}N-BEST-TROSY spectra of ^{15}N G,U-labelled ^{112}Asw. These spectra were recorded at 298 K to assure comparability with the single-molecule FRET data presented in sections 3.2.5 - 3.2.8. The 298 K spectra exhibited a highly similar peak pattern to the previously published and assigned 283 K spectra (Figure 14) (Reining *et al.*, 2013), but they showed two significant differences. First, the imino resonances U31, G38, U40, U41, G44, G59 and U77 were observed already in the absence of adenine, demonstrating pre-formation of the aptamer L2/L3 kissing-loop motif in the apoA conformation. Second, the absence of the

imino signals assigned to G81b, U82b and G115b indicated that the P4 helix adopted only a single conformation. Importantly, the apoB-stabilized mutant also showed a single set of P4 imino signals (see section 3.2.3 and Appendix 1), which matched the set of peaks that was originally assigned to the apoA conformation (Reining *et al.*, 2013). This suggests that the P4 imino signals of ^{112}Asw in fact belong to apoB. However, it cannot be fully excluded, that also apoA shows P4 imino signals with equivalent ^{1}H,^{15}N chemical shifts to the P4 signals of apoB.

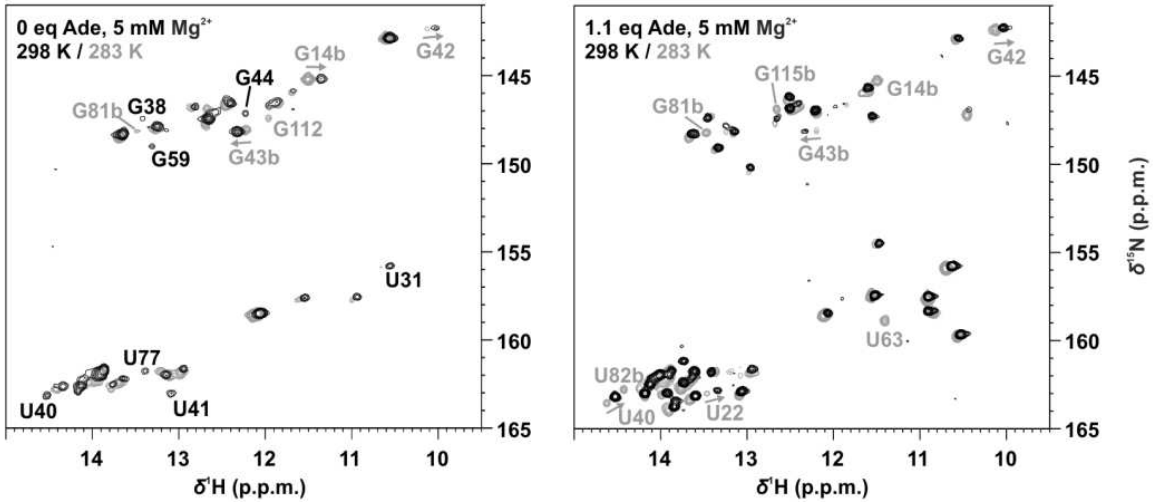

Figure 14: Transferability of the imino NH resonance assignment of ^{112}Asw from 283 K to 298 K. Overlays of the imino region of the ^{1}H,^{15}N-BEST-TROSY spectra measured at 298 K (black) with the assigned ^{15}N-HSQC spectra measured at 283 K (grey) (Reining *et al.*, 2013) without and with adenine in presence of 5 mM Mg^{2+}. Assigned signals that are missing or shifted beyond linewidth at 298 K are annotated in grey. Assigned signals that appear at 298 K are annotated in black.

Having transferred the imino NH resonance assignment of ^{112}Asw to 298 K, the adenine-dependent imino signal intensities were determined for 25 imino reporter signals (Figure 15A, B). The selected reporter signals represent key structural elements of the apoB-form aptamer (helix P1b), the apoA-form aptamer (helices P1, P2, loop L2 and junctions J1-2, J2-3) and the expression platform (helices P4, P5). The P3 helix was omitted in this analysis. The P3 imino resonances cannot be specifically ascribed to the apoB-form or to the apoA-form aptamer. Figure 15C shows the adenine-induced switch of the imino reporter signal intensities.

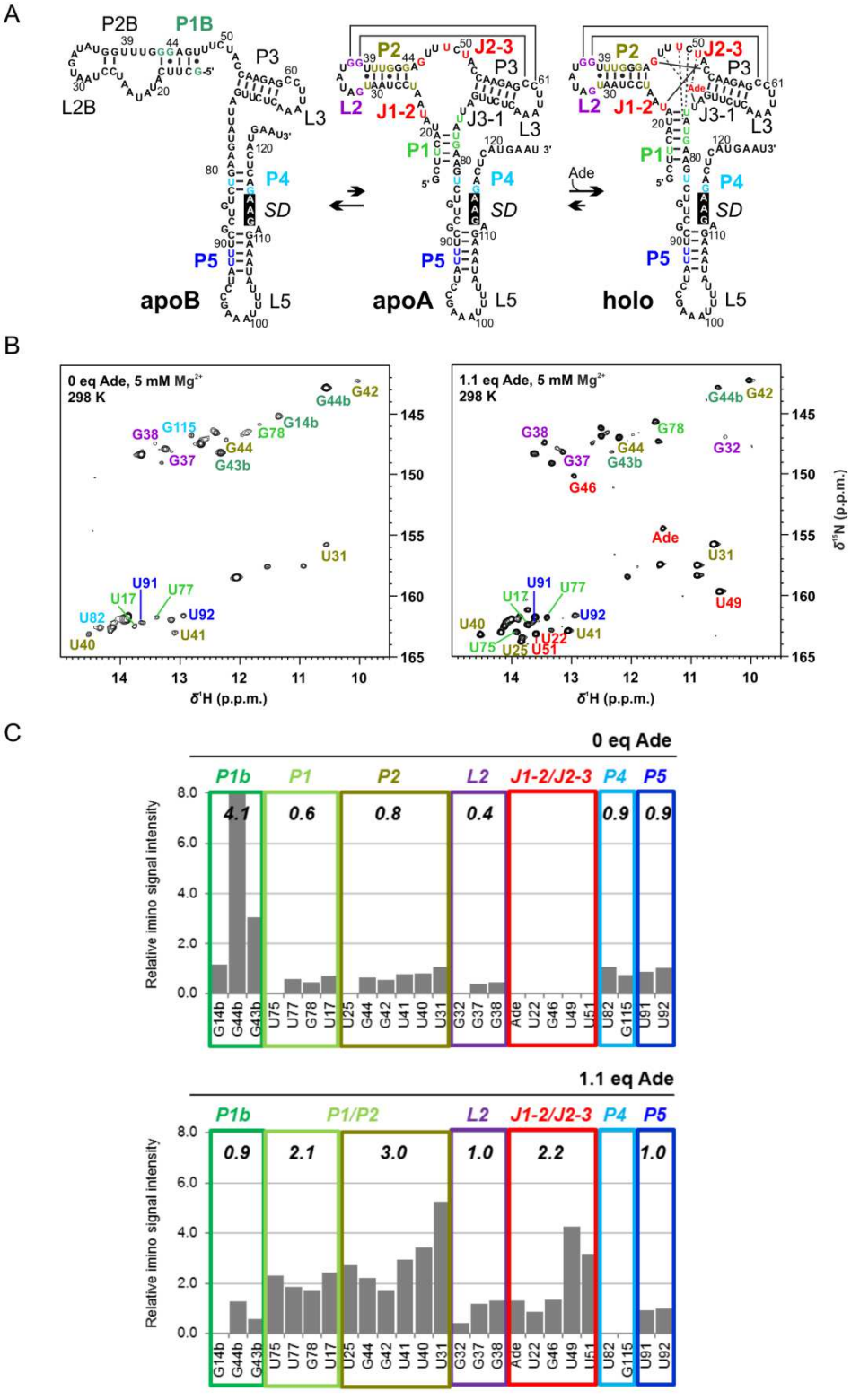

Figure 15: Adenine-induced switch in ^{112}Asw base pairing monitored by NMR. (A) Model of the 3-state conformational equilibrium with colour-coded G/U imino reporter residues and highlighted Shine Dalgarno sequence (SD). The scheme was modified after Reining *et al.* (Reining *et al.*, 2013). **(B)** Imino region of the ^{1}H,^{15}N-BEST-TROSY spectra of the *add* Asw (0.3 mM) in absence and presence of 1.1 eq adenine at 5 mM Mg^{2+} and 298 K. The imino reporter signals are annotated. **(C)** Relative intensities of the imino reporter signals (normalized to U92) without adenine (top) and with 1.1 eq adenine (bottom). The average imino reporter signal intensity of the structural modules is indicated in black.

Without adenine, ^{112}Asw exhibited dominating imino signals of the apoB-specific helix P1b, which were ~6-fold stronger than the imino signals of the P1/P2 module of the apoA aptamer and ~4-fold stronger than those of the expression platform helices P4 and P5. Given that the helices P1b, P4 and P5 are all part of the apoB conformation (see section 3.2.3), the substantially stronger imino signal intensities of P1b are striking. Apparently, the P1b base pairs were more stable against solvent exchange than the P4 and P5 base pairs. Specific P4 and 5'-terminal P1 imino signals of the apoA-form secondary structure were not observed. This suggests that base pairing at the P1/P4 interface was transient.

With 1.1 eq adenine, the imino signals of the junctional elements J1-2 and J2-3 as well as U75 in P1 and U25 in P2 appeared, indicating folding of the holo aptamer core. The average imino reporter signal intensity was ~4-fold decreased for the P1b helix and ~4-fold increased for the P1/P2 module. This anti-correlated intensity change for apoB-specific and apoA/holo-specific signals suggests a structural switch from the apoB-form to the apoA-form secondary structure. Adenine did not induce an increase of the relative signal intensity of L2 over P2. This observation suggests a similar stability of the aptamer kissing loop interaction in the apoA and in the holo conformation. The imino signals of P4 disappeared with adenine.

In summary, the adenine-induced changes in the relative imino signal intensities of ^{112}Asw demonstrate that adenine binding to apoA induced folding of the holo aptamer core, and, thereby shifted the secondary structure equilibrium from the apoB-form to the apoA-form. These conformational transitions were associated with a residual melting of the P4 helix.

3.2.2. Effect of Li^+ carryover during NMR sample preparation

The NMR data presented in Figure 15 were collected from an *in vitro* transcribed RNA sample that was purified according to the standard protocol used in the Schwalbe lab (see sections 2.3.5 - 2.3.9). In this protocol, the RNA is precipitated with a five-fold volume of 2% $LiClO_4$ in acetone, folded by thermal denaturation in water, and buffer exchanged into >99% NMR buffer (25 mM potassium phosphate pH 6.2, 50 mM potassium chloride) by repeated dilution cycles in Vivaspin centrifugal concentrators. In a modified version of this protocol, the buffer exchange was omitted, and the folded RNA was instead rapidly

buffered by the addition of 0.1 vol 10x concentrated NMR buffer. Both procedures yielded the desired pH of 6.2. However, in comparison with the buffer-exchanged NMR sample here denoted ^{112}Asw-*x*, the NMR sample with omitted buffer exchange, ^{112}Asw-*nx*, showed changes in the imino region of the ^{1}H,^{15}N-BEST-TROSY spectra. In ^{112}Asw-*nx* without adenine (Figure 16A), the imino signals G38, G59 and U41, which indicate prefolding of the apoA aptamer tertiary structure, were below noise level. Additionally, the apoA signals G42 and G44 as well as the apoB signal G43b were shifted beyond linewidth, and the expression platform signal G112 appeared. In ^{112}Asw-*nx* with adenine (Figure 16B), G14b, the apo-specific additional signals of G72 and U71, as well as the expression platform signals U82, G112 and G115 were still visible. Moreover, the apoA signals U22 and U77 and the apoB signal G43b were shifted beyond linewidth.

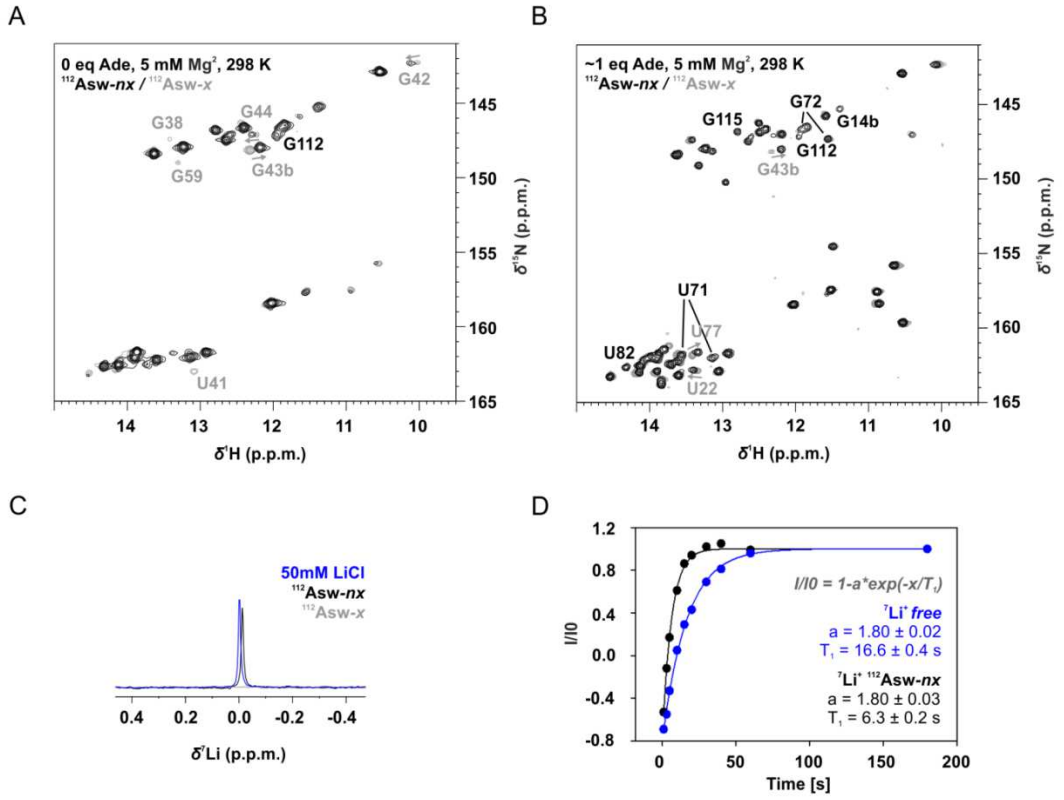

Figure 16: Comparison of NMR data obtained for ^{112}Asw-*x* prepared with buffer exchange and ^{112}Asw-*nx* prepared without buffer exchange. (A) Overlay of the imino region of the ^{1}H,^{15}N-BEST-TROSY spectra of ^{112}Asw-*nx* (0.5 mM; black) and ^{112}Asw-*x* (0.3 mM; grey) in absence of adenine at 298 K and 5 mM Mg^{2+}. Assigned signals that are missing or shifted beyond linewidth in ^{112}Asw-*nx* are annotated in grey. Assigned signals that appear in ^{112}Asw-*nx* are annotated in black. **(B)** Overlay of the imino region of the ^{1}H,^{15}N-BEST-TROSY spectra of ^{112}Asw-*nx* (0.4 mM; black) and ^{112}Asw-*x* (0.3 mM; grey) with ~1 eq adenine at 298 K and 5 mM Mg^{2+}. Assigned signals that are missing or shifted beyond linewidth in ^{112}Asw-*nx* are annotated in grey. Assigned signals that appear in ^{112}Asw-*nx* are annotated in black. **(C)** ^{7}Li NMR spectra of ^{112}Asw-*nx* and ^{112}Asw-*x* at the conditions specified in (B) and of a reference sample containing 50 mM LiCl with 0.4 mM adenine and 5 mM Mg^{2+}. **(D)** ^{7}Li T_1 measurements via inversion-recovery experiments at 300 MHz and 298 K for ^{112}Asw-*nx* (black) and for the 50 mM LiCl reference sample (blue).

Previously reported thermal denaturation experiments have shown that amongst the group I metal ions Li^+ can most potently stabilize tertiary structure of the P1-stabilized isolated adenine riboswitch aptamer domain. Specifically, in buffer containing 140 mM Li^+, the diaminopurine-bound aptamer tertiary structure was about 2 kcal/mol more stable than in buffer containing 140 mM K^+ (Lambert *et al.*, 2009). Since Li^+ can impact the folding of the adenine riboswitch aptamer, it was tested whether omitting the Vivaspin buffer exchange procedure for ^{112}Asw-*nx* led to a significant carryover of Li^+ ions from the $LiClO_4$ precipitation. ^{7}Li NMR measurements of ^{112}Asw-*x* and ^{112}Asw-*nx* revealed that ^{112}Asw-*nx* contained 40 mM Li^+ ions (Figure 16C), whereas no residual Li^+ was detected in ^{112}Asw-*x* (c << 1mM). The ^{7}Li resonance of Li^+ exhibited a 0.01 ppm upfield shift in ^{112}Asw-*nx* compared to a 50 mM LiCl reference sample (Figure 16C) and the ^{7}Li T_1 relaxation time in ^{112}Asw-*nx* was decreased from 16.6 ± 0.4 s to 6.3 ± 0.2 s (Figure 16D). This decrease of the ^{7}Li T_1 of Li^+ in presence of ~1.1 nucleotide equivalents of the 36 kDa *add* Asw was comparable to the decrease of the ^{7}Li T_1 of Li^+ from 17.5 s to 5.8 s in presence of ~0.8 nucleotide equivalents of ~1000 kDa calf thymus DNA, which was measured by Hald and Jacobsen (at 250 MHz and 300 K) (Hald and Jacobsen, 1991). In ^{112}Asw-*nx*, Li^+ was thus in fast exchange between the free form and RNA bound forms that exhibit faster T_1 relaxation in the anionic environment of the slowly tumbling macromolecule. Given that these data clearly indicate Li^+ binding to the *add* Asw in ^{112}Asw-*nx* and since Li^+ was present in ~8-fold excess over Mg^{2+}, it was concluded that Li^+ coordination led to the observed changes between ^{112}Asw-*nx* and ^{112}Asw-*x*.

The adenine induced switch in the intensity of the selected imino reporter signals was then evaluated for the 40 mM Li^+ containing NMR sample ^{112}Asw-*nx* (Figure 17) in an analogous manner to the NMR sample ^{112}Asw-*x* which contained no Li^+ (see section 3.2.1 Figure 15C). In ^{112}Asw-*nx* without adenine, the relative imino signal intensities of P4 and P5 were ~5-fold increased, compared to adenine-free ^{112}Asw-*x*. This observation indicates that the expression platform base pairs were stabilized by the Li^+ carryover. The addition of ~1 eq adenine to ^{112}Asw-*nx* clearly induced formation of the holo conformation, as evidenced by the appearing of the $J_{1\text{-}2}$ and $J_{2\text{-}3}$ signals with similar intensities to P1 and P2. Holo formation in ^{112}Asw-*nx* was associated with an ~8-fold increase of the P1/P2 signals, which is comparable to the ~4-fold increase that was observed for ^{112}Asw-*x*. However, in ^{112}Asw-*nx* only a minute adenine-induced decrease of the P1b and the P4 signals was observed. Thus, from the adenine-induced changes in the imino signal

intensities of ^{112}Asw-*nx* one can infer adenine-induced folding of the apoA aptamer but no pronounced equilibrium shift from apoB to apoA and no substantial melting of the P4 helix.

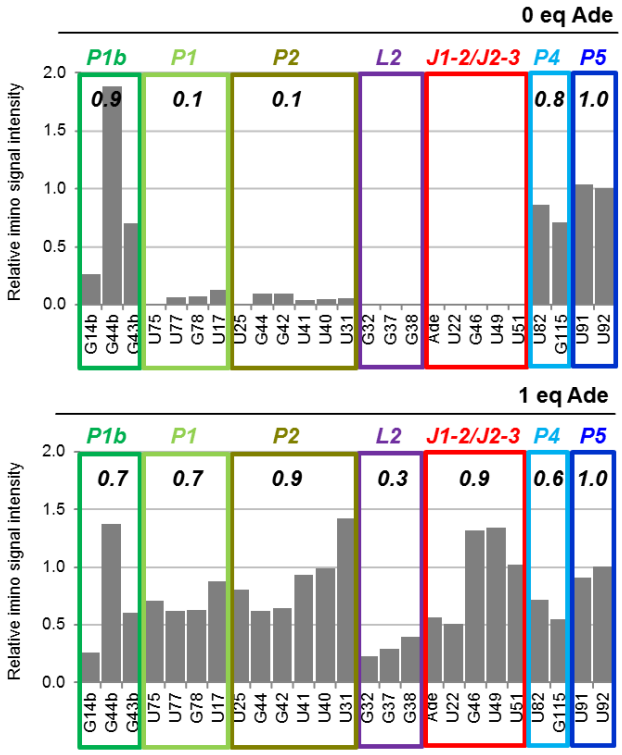

Figure 17: Adenine-induced switch in ^{112}Asw base pairing monitored by NMR of the sample ^{112}Asw-*nx* that contained a carryover of 40 mM Li^+. Relative signal intensity of the imino reporter signals (compare section 3.2.1) measured without adenine (top) and with 1 eq adenine (bottom) in ^{1}H,^{15}N-BEST-TROSY spectra at 5 mM Mg^{2+} and 298 K. The average imino reporter signal intensity of the structural modules is indicated in black.

3.2.3. NMR characterization of the apoB-stabilized mutant apoB$_{STAB}$

As the full-length *add* Asw exhibits a bistable secondary structure with the apoA-form and the apoB-form in dynamic equilibrium, its macroscopic structural and functional properties are ensemble averages over the distinct properties of apoA and apoB. To individually examine apoA and apoB, their secondary structure needs to be stabilized by mutations. Reining *et al.* stabilized apoA in ^{112}Asw by mutating two P2 base pairs, which completely suppressed apoB formation, as judged by ^{1}H,^{15}N-HSQC NMR data (Reining *et al.*, 2013). Here, a specifically devised apoB-stabilized mutant of ^{112}Asw with two mutated base pairs in the P1b helix (U16-A45 to G16-C45 and U17-G44 to C17-G44; Figure 18C) was characterized in terms of secondary structure and adenine-binding competency.

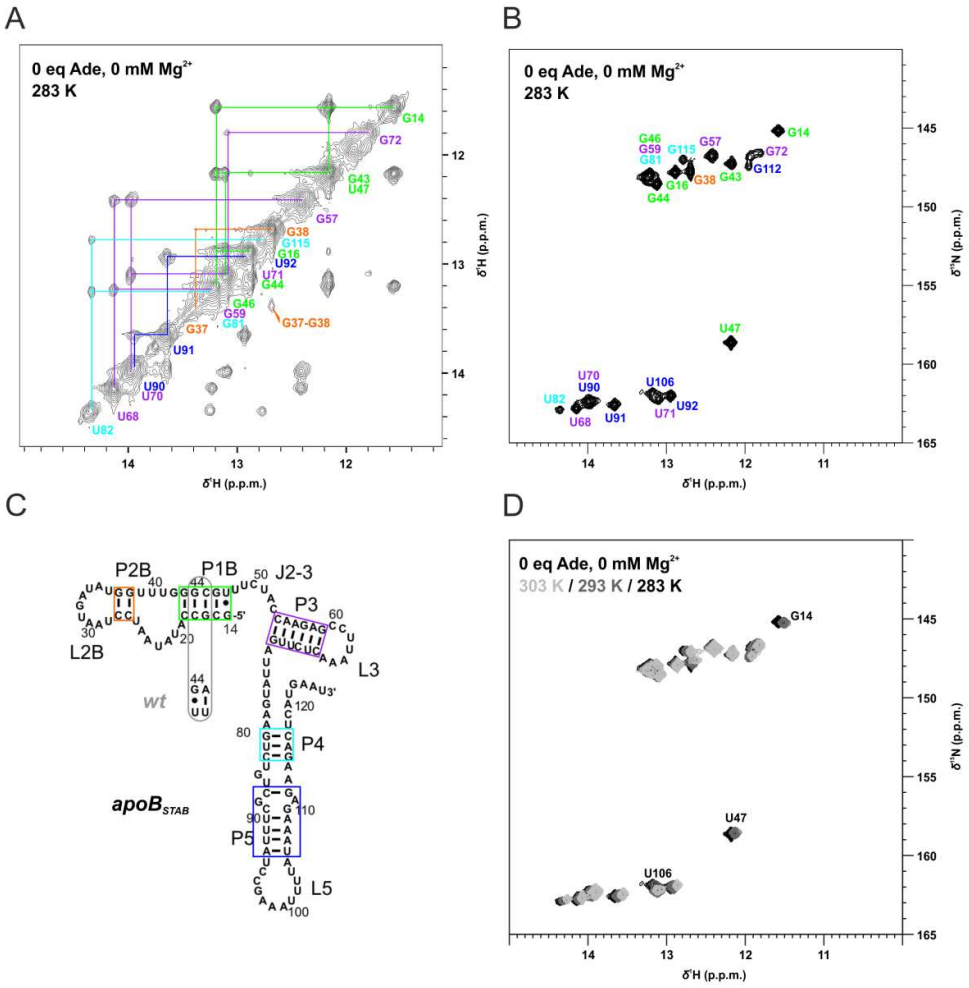

Figure 18: Assignment of the imino NH resonances of the ^{112}Asw mutant apoB$_{STAB}$. (A) Imino region of the ^{1}H,^{1}H-NOESY spectrum of apoB$_{STAB}$ (0.5 mM) at 283 K without adenine and Mg^{2+}. The imino residues are coloured according to structural elements. **(B)** ^{1}H,^{15}N-HSQC spectrum of apoB$_{STAB}$ at the conditions specified in (A). **(C)** Secondary structure model of apoB$_{STAB}$. Mutated base pairs are highlighted in comparison to the wildtype. **(D)** Overlay of the imino region of the ^{1}H,^{15}N-HSQC spectra of apoB$_{STAB}$ measured at 283 K (black), 293 K (dark grey) and 303 K (light grey) without adenine and Mg^{2+}. Signals that disappear with higher temperature are annotated in black.

The secondary structure of the apoB-stabilized mutant (apoB$_{STAB}$) was validated by ^{1}H,^{15}N-HSQC and ^{1}H,^{1}H-NOESY NMR experiments. At 283 K, the ^{1}H,^{15}N-HSQC spectrum of apoB$_{STAB}$ showed 19 signals (Figure 18B), which could be assigned to 21 guanosine and uridine residues (see Appendix 1) by means of the ^{1}H,^{1}H-NOESY spectrum (Figure 18A) and via chemical shift comparison to the wild-type. Importantly, apoA-specific signals were completely absent both in the ^{1}H,^{15}N-HSQC spectrum and in the ^{1}H,^{1}H-NOESY spectrum of apoB$_{STAB}$. The ^{1}H,^{1}H-NOESY spectrum confirmed that the mutated P1b helix of apoB$_{STAB}$ formed all six expected base pairs. It further demonstrated that the P4 helix of apoB$_{STAB}$ formed only three stable base pairs that

involve the imino protons of G81, U82 and G115. An extended P4b helix, as suggested by Reining et al. (Reining *et al.*, 2013), was not observed. On the contrary, the chemical shifts of G81 and G115 in apoB_{STAB} matched with the chemical shifts of the G81 and the G115 signal that Reining *et al.* ascribed to a P4 helix of apoA ($\Delta\delta^{1}H \leq 0.05$ ppm and $\Delta\delta^{15}N \leq 0.2$ ppm). These observations suggest that there exists only a single form of the P4 helix of ^{112}Asw, which might be adopted exclusively by apoB or by both the apoB and the apoA conformation. Additionally, the temperature dependence of the secondary structure of apoB_{STAB} was investigated in a $^{1}H,^{15}N$-HSQC temperature series (Figure 18D). These spectra revealed only two minor changes between 283 K and 303 K, which are the disappearing of the imino signals for the P1b terminal base pair (G14 and U47) and the P5 terminal base pair (U106).

The adenine binding competency of apoB_{STAB} was characterized in three different experiments. First, $^{1}H,^{15}N$-HSQC NMR was used to investigate whether adenine affects the base pairing structure of apoB_{STAB} in presence of 5 mM Mg^{2+} at 283 K (Figure 19A, B). The $^{1}H,^{15}N$-correlation spectra of apoB_{STAB} were found unchanged by the addition of adenine up to 10 eq over the RNA (Figure 19A), except for a small ^{1}H chemical shift change of G14 from 11.55 ppm without adenine to 11.49 ppm with 10 eq adenine. Also, the signal intensities of selected imino reporter signals of apoB_{STAB} for the structural modules P1b, P2b, P3, P4 and P5 did not change with adenine, except for a small (~20%) decrease in the relative imino signal intensities of the helices P1b and P2b with 10 eq adenine. These data clearly show that adenine did not alter the folding of apoB_{STAB}. The minor chemical shift change of G14 and the minute intensity decrease of P1b and P2b might indicate weak interactions of adenine with the P1b/P2b module.

Second, it was tested whether ligand-observed binding can be detected for apoB_{STAB} in fluorescence quenching experiments with the fluorescent adenine analogue 2-aminopurine (Figure 19C). The fluorescence quenching assay was performed in comparison to wild-type ^{112}Asw, which served as a positive control for 2-aminopurine binding. As opposed to the wild-type, apoB_{STAB} did not quench 2-aminopurine fluorescence up to a concentration of 10 μM. This suggests that the adenine analogue, if at all, is bound only weakly by apoB_{STAB} with a K_D in the high micromolar range. The 2-aminopurine binding affinity obtained for the wild-type (0.27 μM) is in good agreement with the previously reported 2-aminopurine binding affinity of ^{112}Asw that was calculated

from the observable rate constant in stopped-flow 2-aminopurine fluorescence quenching experiments under comparable solution conditions (0.55 µM) (Reining *et al.*, 2013).

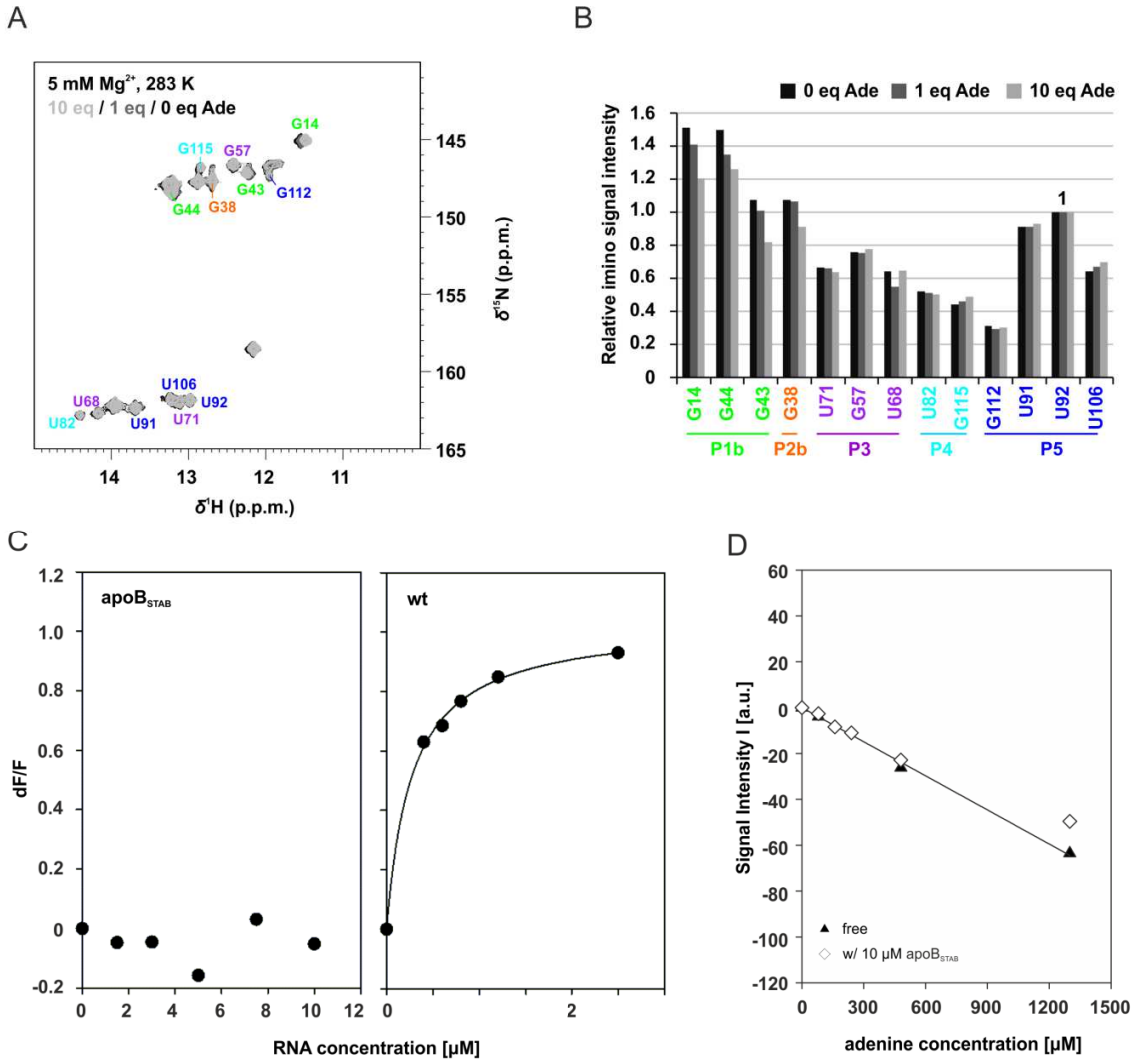

Figure 19: Adenine binding studies of the ^{112}Asw mutant apoB$_{STAB}$. (A) Overlay of the imino region of the ^{1}H,^{15}N-HSQC spectra of apoB$_{STAB}$ (~0.3 mM) measured at 283 K in presence of 5 mM Mg^{2+} with 0 eq (black), 1 eq (dark grey) and 10 eq adenine (light grey). Selected imino reporter signals are annotated. **(B)** Relative intensities of the imino reporter signals (normalized to U92) determined from the spectra shown in (A). **(C)** 2-aminopurine fluorescence quenching as a function of the RNA concentration for apoB$_{STAB}$ (left) and wt ^{112}Asw (right). The data were acquired at 298 K using a 2-aminopurine concentration of 50 nM in NMR buffer containing 2 mM Mg^{2+}. For wt, dF/F was fitted using the equation $dF/F = (1-a)*x/(x+K_D)$ which yielded $a = -0.03 \pm 0.02$ and $K_D = 0.27 \pm 0.02$ µM. **(D)** WaterLOGSY signal intensity of adenine (H2 + H8) as a function of the adenine concentration measured in the absence of apoB$_{STAB}$ (filled triangles) and with 10 µM apoB$_{STAB}$ (empty diamonds). The signal intensities in absence of apoB$_{STAB}$ were fitted linearly (slope = 0.050 ± 0.001). The WaterLOGSY spectra were recorded with an NOE mixing time of 2 s at 298 K in presence of 2 mM Mg^{2+}. The parameters of the CPMG pulse train for suppression of the RNA resonances were n = 70 and τ = 0.2 ms.

Third, ligand observed adenine binding of apoB$_{STAB}$ was assayed with the WaterLOGSY NMR method (Figure 19D). The adenine signal intensity in WaterLOGSY spectra of adenine measured in the absence and in the presence of apoB$_{STAB}$ did not differ significantly up to an adenine concentration of 0.5 mM. At 1.3 mM adenine, the WaterLOGSY signal intensity of adenine was decreased in presence of apoB$_{STAB}$. These observations indicate that adenine did not interact with apoB$_{STAB}$ at sub-millimolar concentrations but that low affinity binding set on at adenine concentrations above 1 mM. The WaterLOGSY data thus confirm that the small changes observed in the ^{1}H,^{15}N-HSQC spectra of apoB$_{STAB}$ with 10 eq adenine were due to weak interactions between adenine and the RNA. Given that the bacterial adenine concentration is far below ~1 mM (1 µM – 102 µM in *E. coli*) (Ishii *et al.*, 2007; Bennett *et al.*, 2009), such weak adenine binding of the apoB conformation can clearly be considered insignificant.

Taken together, the apoB conformation of ^{112}Asw was successfully stabilized in the mutant apoB$_{STAB}$ and the ligand binding experiments conducted with apoB$_{STAB}$ proved that the adenine binding competency of the wild-type can be entirely ascribed to the apoA conformation.

3.2.4. Bulk fluorescence anisotropies of single-fluorophore-labelled ^{112}Asw

To investigate the degree of interaction between the FRET dyes Cy3 and Cy5 and surrounding RNA nucleotides in the FRET constructs of ^{112}Asw, bulk fluorescence anisotropy measurements were performed. For these measurements, single-fluorophore labelled ^{112}Asw constructs were prepared. Since the dual-fluorophore labelling schemes of the three ^{112}Asw FRET constructs used in this study were L2(Cy3)/L3(Cy5), L2(Cy3)/P5(Cy5) and L3(Cy5)/P5(Cy3), the investigated single-fluorophore-labelled ^{112}Asw constructs were L2(Cy3), P5(Cy3), L3(Cy5) and P5(Cy5) (Figure 20).

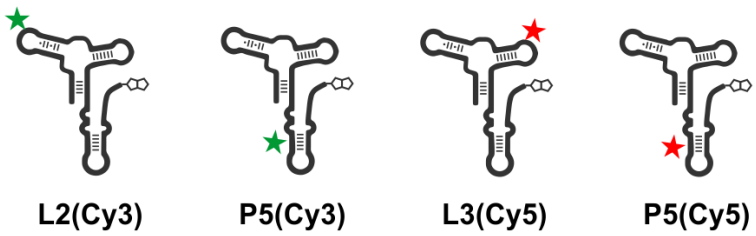

Figure 20: Schematic representation of the four single-fluorophore labelled ^{112}Asw constructs in the apoA conformation. Cy3 (green star), Cy5 (red star) and the 3' biotin-immobilization site (white polygon) are highlighted.

The fluorescence anisotropies of both cyanine dyes were significantly increased in the RNA-coupled forms compared to the free form (Table 7). In case of Cy3, an increase from r = 0.244 in the free form to r = 0.285 at the L2-labelling site and to r = 0.296 at the P5-labelling site was observed. In case of Cy5, the increase was from r = 0.127 in the free form to r = 0.238 at L3 and r = 0.281 at P5. These data indicate that orientational averaging of the RNA-coupled dyes on the nanosecond timescale was perturbed in the RNA constructs. However, such increase in the fluorescence anisotropy of RNA-coupled dyes is common in smFRET studies of RNA. The anisotropies of Cy3 and Cy5 at the labelling sites of ^{112}Asw were comparable to the anisotropies of Cy3 and Cy5 in a FRET construct that was recently used to study the folding of the diels alderase (DA) ribozyme (r(Cy3-DA) = 0.26 and r(Cy5-DA) = 0.25) (Kobitski *et al.*, 2007). It was thus concluded that the increase in the fluorescence anisotropy of Cy3 and Cy5 in the FRET constructs of ^{112}Asw did not reflect specific interactions between the fluorophores and their environment in ^{112}Asw.

Table 7: Bulk fluorescence anisotropies of Cy3 (Ex/Em = 525 nm / 565 nm) and Cy5 (Ex/Em = 625 nm / 665 nm) in the free form and in RNA-coupled forms in single-fluorophore-labelled ^{112}Asw constructs [a]

Construct	r
Cy3 free	0.244 ± 0.006
L2(Cy3)-^{112}Asw	0.285 ± 0.004
P5(Cy3)-^{112}Asw	0.296 ± 0.004
Cy5 free	0.127 ± 0.004
L3(Cy5)-^{112}Asw	0.238 ± 0.009
P5(Cy5)-^{112}Asw	0.281 ± 0.008

[a] The data were measured at a concentration of ~4 nM in smFRET immobilization buffer (25 mM potassium phosphate pH 7.0, 50 mM potassium chloride) with 2 mM Mg^{2+} at 20 °C. The errors are standard errors obtained over 20 replicates.

3.2.5. FRET histogram analysis of the adenine-dependent global folding

The adenine-induced global tertiary structure folding of ^{112}Asw was analysed by collecting single-molecule FRET histograms of L2/L3-, L2/P5- and L3/P5-labelled ^{112}Asw in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. The 100 µM adenine concentration is a ~100-fold excess over the reported adenine dissociation constant of ^{112}Asw (K_D ~ 0.5 µM at 2 mM Mg^{2+} and 25 °C) (Reining *et al.*, 2013). The FRET histograms of L2/L3-labelled ^{112}Asw (Figure 21, left panel) revealed persistent conformational heterogeneity of the L2/L3 distance, with a low-FRET state at E~0.2 and a high-FRET state at E~0.9. Two equivalent FRET states were previously reported for the

P1-stabilized L2/L3-labeled aptamer domain of the *add* Asw (Lemay *et al.*, 2006). Accordingly, the low-FRET and the high-FRET state of ^{112}Asw were assigned to folding states with undocked (*U*) and docked aptamer kissing loop motif (*D*). While the P1-stabilized L2/L3-labeled *add* Asw aptamer domain was predominantly in the docked state even without adenine at a comparable Mg^{2+} concentration (4 mM), (Dalgarno *et al.*, 2013) the fractional population of the docked state in the native full-length ^{112}Asw was considerably lower, with ~20% *D* in absence and ~50% *D* in presence of saturating amounts of adenine.

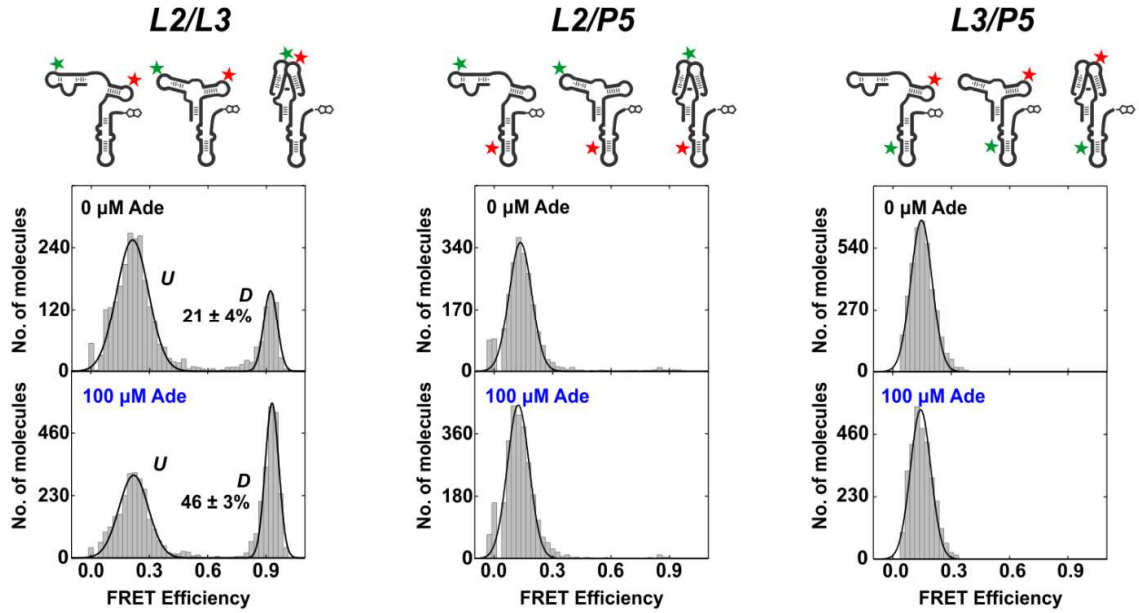

Figure 21: Adenine-induced switch in the global tertiary structure of ^{112}Asw monitored by smFRET. FRET histograms of L2/L3-, L2/P5- and L3/P5-labelled ^{112}Asw in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. A schematic representation of each FRET construct in the apo conformations is depicted above. For the FRET histograms of the L2/L3-labelled construct, the fractional population of the docked state (*D*) is indicated in percent.

The FRET histograms of L2/P5- and L3/P5-labelled ^{112}Asw (Figure 21, middle and right panels) showed a single low-FRET state at E~0.15 independent of adenine. The adenine induced conformational switch of ^{112}Asw thus did not lead to a detectable change in the L2/P5 and in the L3/P5 distance. However, in conventional smFRET experiments one cannot detect conformational states with E~0. The population of molecules that does not show FRET is excluded from analysis, since it contains donor-only molecules that do not possess a photoactive acceptor fluorophore due to incomplete labelling or due to acceptor

photo-bleaching. Therefore, the FRET histogram analysis was additionally performed using single-molecule alternating laser excitation (smALEX) (Kapanidis *et al.*, 2004).

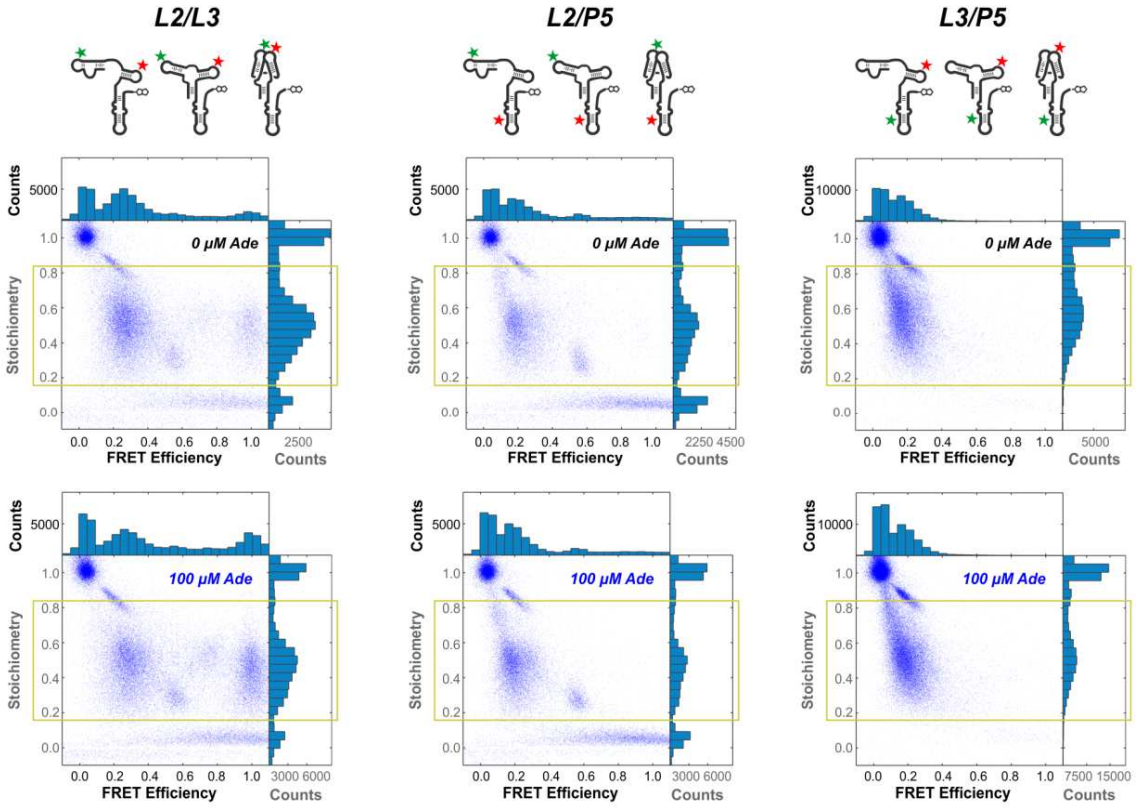

Figure 22: Adenine-induced switch in the global tertiary structure of ^{112}Asw monitored by smALEX. Two-dimensional FRET efficiency / stoichiometry histograms of L2/L3-, L2/P5- and L3/P5-labelled ^{112}Asw in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. These histograms show both the dual- and the single-fluorophore-labelled molecules of the samples. The stoichiometry axis sorts the FRET counts of the molecules according to a stoichiometry factor S. Donor-only molecules appear at S~1, dual-fluorophore-labelled molecules at S~0.5 and acceptor-only molecules at S~0 (Kapanidis *et al.*, 2004). The FRET counts associated with dual-fluorophore labelled molecules are boxed. A schematic representation of each FRET construct in the apo conformations is depicted above.

The smALEX histograms (Figure 22) showed that the dual-fluorophore-labelled molecules in the three FRET constructs did not populate an additional FRET state with E~0. This validates the above presented smFRET data. L2/P5- and L3/P5-labelled ^{112}Asw thus exhibited a single and a similar distance between the two fluorophores on the nm scale. As the NMR data have indicated that the P5 helix is not involved in the functional dynamics of ^{112}Asw, it is reasonable to assume that dye-labelling of P5 did not perturb the adenine-induced conformational transitions of the RNA. The single distance between the aptamer domain and the expression platform inferred from the two constructs hence demonstrates that single-stranded regions of the *add* Asw expression platform did not

interact with upstream sequence elements of the aptamer domain in the ligand-free and in the ligand-bound state. Furthermore, the fact that aptamer docking was undetectable in these constructs suggests a perpendicular arrangement of the aptamer and the expression platform.

3.2.6. FRET histogram analysis of aptamer folding

To decipher the aptamer conformational heterogeneity that gives rise to the two FRET states of L2/L3-labelled ^{112}Asw, the ligand-dependent folding of L2/L3-labelled ^{112}Asw was investigated in a detailed smFRET histogram analysis.

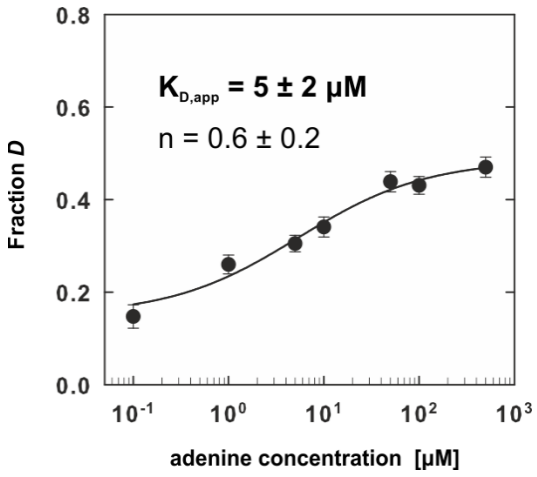

Figure 23: Adenine-dependent aptamer kissing-loop docking of ^{112}Asw monitored by smFRET. Plot of the fractional population of the docked state determined from FRET histograms of L2/L3-labelled ^{112}Asw with different adenine concentrations at 2 mM Mg^{2+}. The data have been fitted using the Hill equation. The half-saturating adenine concentration for aptamer docking ($K_{D,app}$) and the Hill coefficient (n) are indicated.

First, the adenine-dependent folding of the wild-type L2/L3-labelled ^{112}Asw was analysed. FRET histograms were collected at adenine concentrations between 0.1 and 500 µM in presence of 2 mM Mg^{2+} and the fractional population of the docked state was plotted against the adenine concentration (Figure 23). These data were fitted using the Hill equation (Prinz, 2010) to extract the half-saturating adenine concentration for aptamer docking, which is an apparent adenine binding constant $K_{D,app}$. The half-maximal adenine concentration for aptamer docking of L2/L3-labelled ^{112}Asw (5 ± 2 µM) was comparable to the value determined by smFRET of a P1-stabilized *add* Asw aptamer in a similar labelling-scheme (3 ± 1 µM), (Liu *et al.*, 2015) but 10-fold higher compared to the

ligand-observed binding constant determined by fluorescence quenching assays with the fluorescent adenine analogue 2-aminopurine (~0.5 µM) (Reining *et al.*, 2013). The binding curve corroborated the finding that L2/L3-labelled ^{112}Asw populated an equal fraction of molecules with undocked and docked aptamer kissing loop motif at saturating adenine conditions (c ≥ 50 µM), and indicated tertiary structure heterogeneity of the ligand-saturated state.

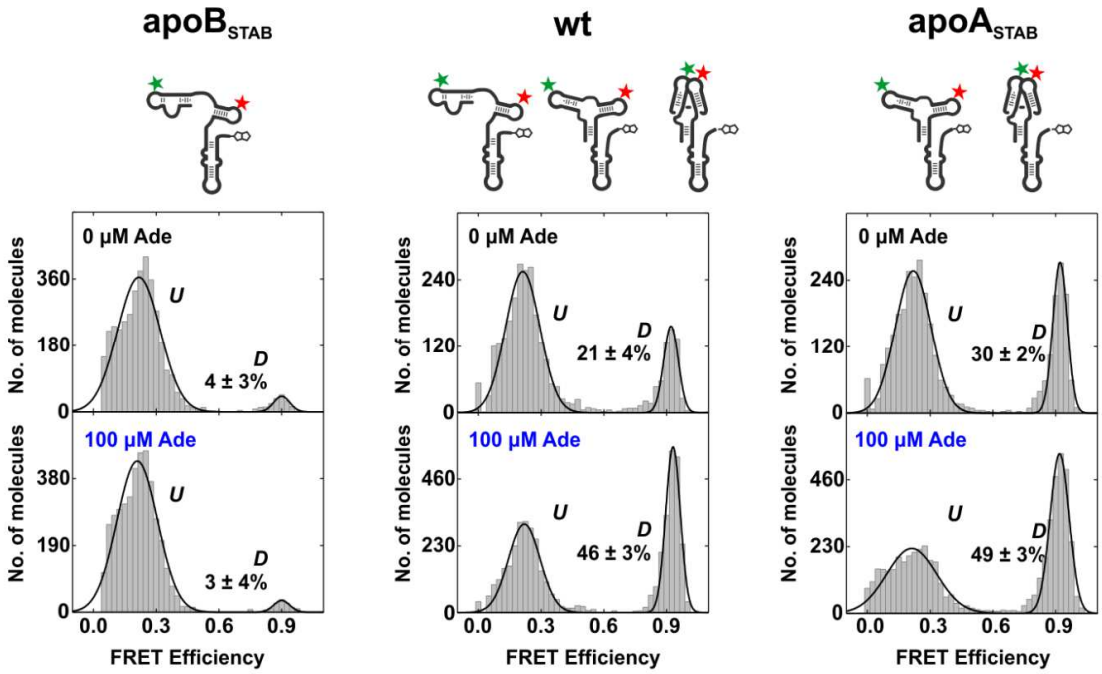

Figure 24: Mutational smFRET analysis of aptamer docking in ^{112}Asw. FRET histograms of L2/L3-labeled apoB$_{STAB}$, wt and apoA$_{STAB}$ in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. A schematic representation of each variant in the apo conformations is depicted above. The fractional population of the docked state (*D*) is indicated in percent.

Second, a comparative FRET histogram analysis was performed between the wild-type (wt) and the apoA-stabilized mutant (apoA$_{STAB}$) as well as the apoB-stabilized mutant (apoB$_{STAB}$) of ^{112}Asw. ApoA$_{STAB}$ was previously shown to only adopt the apoA and the holo conformation and to bind adenine with ~5-fold higher affinity than wt (Reining *et al.*, 2013). The NMR characterization of apoB$_{STAB}$ presented in section 3.2.3 demonstrated that apoB$_{STAB}$ exclusively forms the apoB conformation. The FRET histograms of apoB$_{STAB}$ and apoA$_{STAB}$ in absence and in presence of 100 µM adenine at 2 mM Mg^{2+} showed equivalent FRET states to the wt (Figure 24). As expected from the NMR data, apoB$_{STAB}$ was almost exclusively in the undocked state (~ 95%) and did not exhibit an adenine-dependent docking equilibrium. Without adenine, wt and apoA$_{STAB}$

showed a slightly different aptamer docking equilibrium with ~20% docked molecules in wt and ~30% docked molecules in apo$A_{STAB.}$ Remarkably, at saturating adenine conditions, wt and apoA_{STAB} populated an equal fraction of docked molecules (~50%). Hence, the holo state of wt could be reproduced with the apoA-stabilized mutant. The finding that the decreased fractional population of the docked state in adenine-free wt compared to adenine-free apoA_{STAB} was overcome by the addition of adenine is consistent with an adenine-induced switch from the apoB-form to the apoA-form, as observed by NMR spectroscopy (see section 3.2.1). The relative populations of *U* and *D* measured for wt and apoA_{STAB} were converted to free docking enthalpies (Table 8). From apoA_{STAB} it can be deduced that adenine stabilized the docked apoA-form aptamer over the undocked apoA-form aptamer about ~0.5 kcal/mol in ^{112}Asw. This value is strikingly similar to the guanine-induced stabilization of the docked state in the related *xpt* guanine riboswitch aptamer at sub-saturating Mg^{2+} (~0.5 kcal/mol) (Brenner *et al.*, 2010). The free adenine binding enthalpy of ^{112}Asw determined by isothermal titration calorimetry is ~10-fold higher (~5 kcal/mol for wt at 2 mM Mg^{2+} and 30 °C) (Reining *et al.*, 2013), which illustrates that the weak adenine-induced stabilization of the L2/L3 kissing-loop motif contributes little to the thermodynamics of adenine binding. The difference in the free docking enthalpy between adenine-free wt and adenine-free apoA_{STAB} was ~0.3 kcal/mol. This value is of the same order of magnitude as the free enthalpy difference between the apoA and the apoB conformation previously determined by ^{1}H,^{15}N-HSQC N_{ZZ}-exchange NMR spectroscopy (~0.6 kcal/mol) (Reining *et al.*, 2013). This additionally supports the hypothesis that the difference in the fractional population of the docked state in L2/L3-labelled wt and apoA_{STAB} without adenine was associated with the apoB/apoA conformational equilibrium of wt.

Table 8: Adenine-dependent free enthalpy for aptamer docking [a] of wt and apoA_{STAB} at 2 mM Mg^{2+}

Variant	Ade [µM]	ΔG(*UD*) [kcal/mol]
Wt	0	0.8 ± 0.1
	100	0.1 ± 0.1
apoA_{STAB}	0	0.5 ± 0.1
	100	0.0 ± 0.1

[a] The free docking enthalpies were calculated from the FRET distribution observed in the FRET histograms shown in Figure 24 via ΔG(*UD*) = - RT*ln(*D*/*U*). The reported error was calculated from the fitting errors of the FRET populations by Gaussian error propagation.

Next, the Mg^{2+}-dependence of the aptamer docking equilibrium was characterized for all three variants. For $apoB_{STAB}$, the great majority of molecules remained undocked (>90%) at Mg^{2+} concentrations up to 40 mM in absence and presence of 100 µM adenine (Figure 25). Mg^{2+}, though, had an effect on the broadness of the undocked FRET state, which had an FWHM of ~0.1 in the absence and of ~0.2 in the presence of Mg^{2+}. This indicates that Mg^{2+} increases the conformational flexibility of $apoB_{STAB}$.

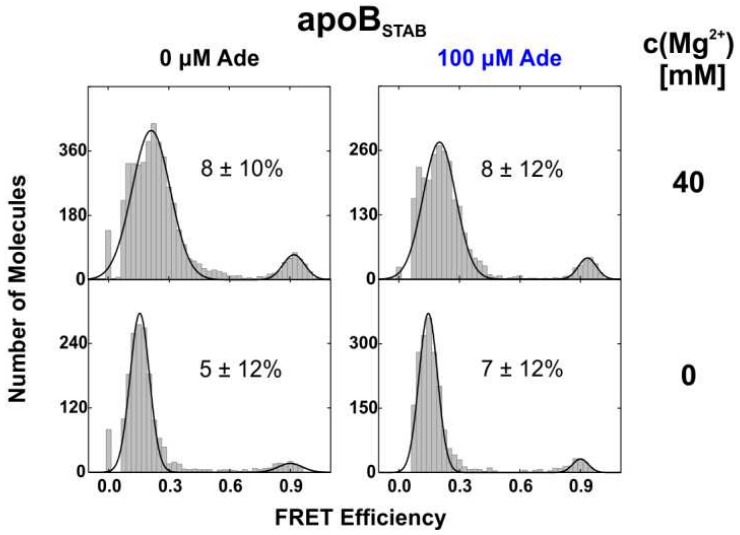

Figure 25: Mg^{2+}-dependent aptamer kissing-loop docking of $apoB_{STAB}$ ^{112}Asw monitored by smFRET. FRET histograms of L2/L3-labeled $apoB_{STAB}$ at 0 and 40 mM Mg^{2+} without adenine (left panel) and with 100 µM adenine (right panel). The fractional population of the docked state (*D*) is indicated in percent.

In contrast, wt and $apoA_{STAB}$ showed a strongly Mg^{2+}-dependent aptamer docking equilibrium. For these two variants, full Mg^{2+} titrations were performed with Mg^{2+} concentrations from 0.05 mM to 40 mM in absence and presence of adenine (100 µM). In all cases, an apparent two-state folding transition was observed and fitted using the Hill equation (Prinz, 2010) (Figure 26). The midpoint of the folding transition $[Mg^{2+}]_{1/2}$ was decreased with adenine from ~4 mM to ~1 mM for wt and from ~1.3 mM to ~0.3 mM for $apoA_{STAB}$. This indicates positive cooperativity between ligand and cofactor binding. The Hill coefficients n were between ~1.9 and ~1.2 and were slightly decreased (within error limits) by the addition of adenine and by the mutational stabilization of apoA. Notably, mutational stabilization of the apoA conformation and the addition of 100 µM adenine had a similar effect on the aptamer docking equilibrium of ^{112}Asw: Wt had a ~3-fold higher $[Mg^{2+}]_{1/2}$ than $apoA_{STAB}$ independent of adenine, and adenine led to a ~4-fold

reduction of $[Mg^{2+}]_{1/2}$ in both wt and $apoA_{STAB}$. The $[Mg^{2+}]_{1/2}$ for L2/L3-docking of $apoA_{STAB}$ without adenine was well comparable to the $[Mg^{2+}]_{1/2}$ for L2/L3-docking of the isolated *pbuE* adenine riboswitch aptamer determined by NMR spectroscopy (1.5 ± 0.4 mM at 10°C w/o adenine) (Noeske, Schwalbe and Wöhnert, 2007). However, the $[Mg^{2+}]_{1/2}$ reported for L2/L3-docking of the isolated P1-stabilized *add* Asw aptamer domain (14 bp P1 with 7 GC) is ~100-fold lower (22 µM) (Lemay *et al.*, 2006), suggesting that P1-stabilization strongly favors the conformational transition into a Mg^{2+} bound L2/L3-docked state.

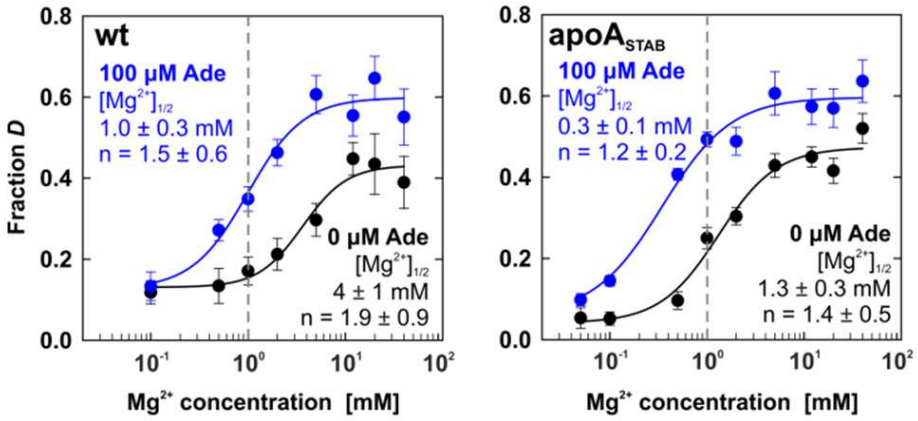

Figure 26: Mg^{2+}-dependent aptamer kissing-loop docking of wt and $apoA_{STAB}$ ^{112}Asw monitored by smFRET. Fractional population of the docked state *D* in smFRET histograms of L2/L3-labeled wt and $apoA_{STAB}$ as a function of the Mg^{2+} concentration without adenine (black) and with 100 µM adenine (blue). The data have been fitted using the Hill equation to obtain the Mg^{2+} concentration for half-maximal docking $[Mg^{2+}]_{1/2}$ and the Hill coefficient n. The dashed line indicates near-physiological free Mg^{2+} concentration (~1 mM).

In terms of populations, wt and $apoA_{STAB}$ ^{112}Asw were predominantly undocked (~90%) at low Mg^{2+} concentrations (<0.1 mM), regardless of the presence of adenine. At near physiological Mg^{2+} concentration (~1 mM) (Tyrrell *et al.*, 2013), wt populated less docked molecules than $apoA_{STAB}$ without and with adenine and showed greatest sensitivity towards adenine in switching from 20% to 40% docked molecules. At saturating Mg^{2+} (>10 mM) the distribution of undocked and docked molecules was again equal in wt and $apoA_{STAB}$ (~40% docked w/o and 60% docked w/ adenine). These observations suggest that not only adenine, but also Mg^{2+} can equalize the aptamer docking equilibria of wt and $apoA_{STAB}$ by eliminating apoB folding in wt. The fact that it was at near-physiological Mg^{2+} concentration where the aptamer docking equilibrium of wt and $apoA_{STAB}$ showed greatest sensitivity towards adenine, and where the difference

between wt and apoA$_{STAB}$ was most pronounced, indicated that the conformational dynamics of the *add* Asw are adapted to physiological levels of Mg^{2+}.

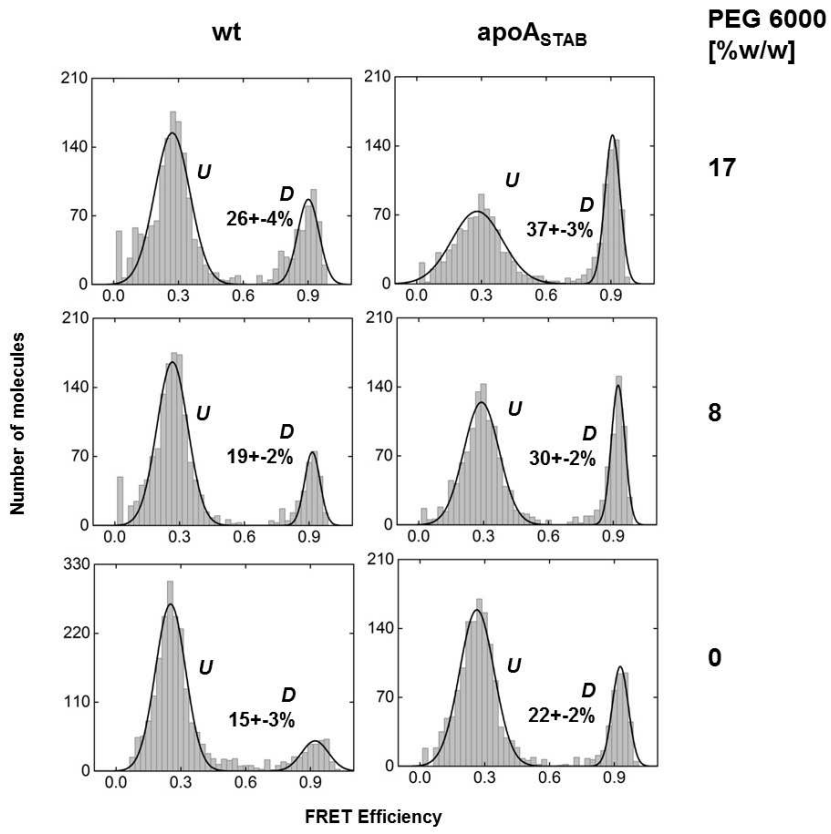

Figure 27: Polyethylene glycol (PEG) dependent aptamer kissing-loop docking of wt and apoA$_{STAB}$ ^{112}Asw monitored by smFRET. FRET histograms of L2/L3-labeled wt and apoA$_{STAB}$ with increasing amounts of PEG 6000 at 2 mM Mg^{2+}. The fractional population of the docked state *D* is indicated in percent.

Last, the effect of macromolecular crowding on the aptamer docking equilibrium of L2/L3-labeled wt and apoA$_{STAB}$ was investigated. FRET histograms were collected at different concentrations of the crowding agent polyethylene glycol (PEG) 6000 at 2 mM Mg^{2+} without adenine (Figure 27). Increasing amounts of PEG 6000 led to an increased fractional population of the docked state *D* in wt and in apoA$_{STAB}$. However, as opposed to Mg^{2+} and adenine, the addition of PEG 6000 did not counterbalance the difference in the fractional population of *D* between wt and apoA$_{STAB}$. This might signify that macromolecular crowding shifted the docking equilibrium of ^{112}Asw by destabilizing the apoA-*U* state towards both apoB and apoA-*D*.

3.2.7. Establishment of an automated analysis routine for smFRET traces

Extracting kinetic information on RNA conformational dynamics from time-resolved smFRET data requires the analysis of a large set of fluorescence time trajectories of individual molecules (~1000). The general approach of this analysis depends on the investigated construct. For the L2/L3-labelled ^{112}Asw, two well-separated FRET states could clearly be observed in the FRET histogram. In such systems, kinetic information can be obtained from dwell-time analysis. However, upon initial inspection of the collected smFRET traces it was realized that the L2/L3-labelled ^{112}Asw exhibited a pronounced heterogeneity between molecules that transitioned between the two FRET states and molecules that did not. In such a case, it is advisable to visualize the distribution of dynamics over the ensemble by creating transition occupancy density plots (TODPs) (Blanco, 2010). Furthermore, some molecules showed artificial changes in the FRET trajectory due to changes in fluorescence intensity of only the acceptor (I_A) or only the donor (I_D). Such uncorrelated fluorescence intensity changes can arise due to fluctuations in the local environment of either fluorophore as a consequence of nucleobase dynamics or cofactor binding (Blanco, 2010). The resulting artificial FRET changes must be sorted out and not considered in the analysis of the true FRET changes, which are characterized by an anticorrelated fluorescence intensity change of I_A and I_D. As a practical solution to this problem Blanco *et al.* suggested to perform Hidden Markov modelling of the FRET trajectory (E), of the normalized acceptor fluorescence intensity trajectory (I_A/I_A^{max}) and of the normalized donor fluorescence intensity trajectory (I_D/I_D^{max}), in order to subsequently compare the three idealized trajectories and identify the true transitions in E with an anticorrelated intensity change in donor and acceptor fluorescence (Blanco, 2010). This strategy was implemented in a here established software-based analysis routine for the smFRET traces of L2/L3-labelled ^{112}Asw, which is described in the following. The analysis routine can in principle be employed for any system with two or three well separated FRET states ($\Delta E>0.1$).

The analysis routine (Figure 28) consists of three MATLAB scripts (Appendix 3) and the open source Hidden Markov modelling software HaMMy (McKinney, Joo and Ha, 2006). It uses as input binary *.traces files that contain the background corrected donor and acceptor fluorescence intensities of all molecules observed in a movie. These *.traces files are created from Andor Solis *.raw movie files by a custom program written in IDL, that was kindly provided by Martin Hengesbach (see section 2.8.2). The MATLAB script

export_fluo_trajectories.m processes the data and saves the fluorescence intensities $I_D(t)$ and $I_A(t)$ of individual molecules from the *.traces files in *.dat files for Hidden Markov modelling (HMM) in HaMMy. It performs three processing steps. First, the script corrects the acceptor intensities for 10% donor leakage. Second, it scales back negative fluorescence intensities, which can arise due to noise in the data, to positive values such that the calculated FRET efficiency falls into the interval between 0 and 1 through the entire trace. This is a pre-requisite for the trace to be accepted by HaMMy, and, this correction is relevant for the smFRET traces of L2/L3-labelled *add* Asw due to the high-FRET state at E~0.9 that occasionally fluctuates to E>1. The intensity scaling is implemented in the script by simply replacing $I_A<0$ or $I_D<0$ with an insignificant intensity value (=10). Third, the script removes donor-only trajectories and terminates each donor-acceptor trajectory at the first photobleaching event.

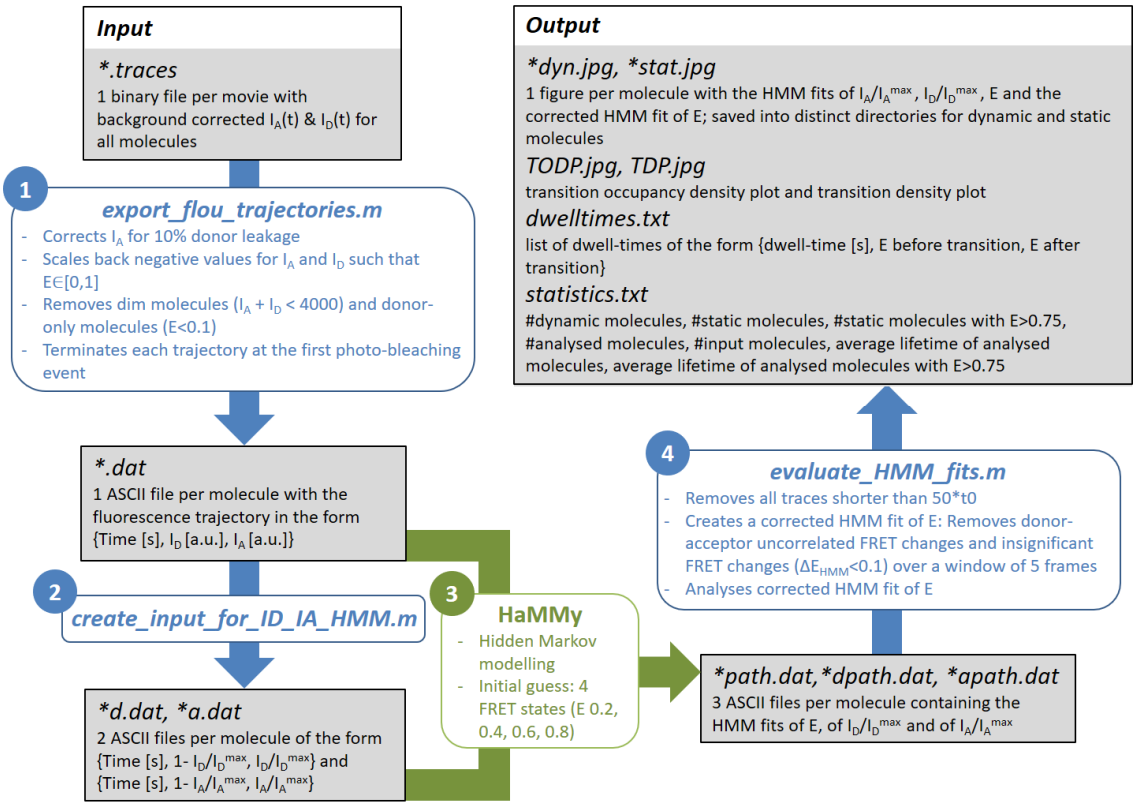

Figure 28: Software-based analysis routine for the smFRET traces of the *add* Asw. Input / output files are presented as grey boxes. The MATLAB scripts *export_flou_trajectories.m* and *create_input_for_ID_IA_HMM* were used to pre-process the data for Hidden Markov modelling (HMM). HMM fitting was performed with the open source software HaMMy (McKinney, Joo and Ha, 2006). The HaMMy output was analysed with the MATLAB script *evaluate_HMM_fits.m.* The analysis routine was established on the basis of a strategy outlined by Blanco *et al.* (Blanco, 2010). The MATLAB script code is attached in Appendix 3.

This is achieved by employing a threshold on the sum of I_A and I_D ($\geq$4000) and on the FRET efficiency ($\geq$0.1) over the median-filtered fluorescence trace (averaging window 3). The output of *export_fluo_trajectories.m* is one *.dat file per molecule. It is noteworthy that the output filenames contain an extension with the number of the respective molecule (*_m1.dat, *_m2.dat, *_m3.dat). The numbering is not continuous, since donor-only molecules and dim molecules have been removed by the script, but it is retained throughout the following analysis, which is necessary to compare the three HMM fits (for E, I_D/I_D^{max} and I_A/I_A^{max}) of each individual molecule at a later stage. The MATLAB script *create_input_for_ID_IA_HMM.m* uses the *.dat output of *export_fluo_trajectories.m* and creates the additional HaMMy input files *d.dat and *a.dat for Hidden Markov modelling of I_D/I_D^{max} and I_A/I_A^{max}, respectively. HaMMy expects ASCII input files of the form {Time [s], x(t), y(t)} and then fits a Hidden Markov model to z(t) = y(t) / (x(t) + y(t)). Therefore, the data for Hidden Markov modelling of I_D/I_D^{max} and I_A/I_A^{max} must be of the form {Time [s], 1- I_D/I_D^{max}, I_D/I_D^{max}} and {Time [s], 1- I_A/I_A^{max}, I_A/I_A^{max}}, respectively (Blanco, 2010).

For a statistically relevant population of ~1000 molecules one needs to process ~3000 input files with HaMMy. When running HaMMy one needs to consider that the software can only process a limited number of files at a time (~200). For the initial guess of FRET states, it is important to select at least two FRET states more than one can actually infer from the FRET histogram (McKinney, Joo and Ha, 2006). HaMMy creates 3 output files per input file (*dwell.dat, *report.dat, *path.dat). Of the HaMMy output, the three *path.dat files per molecule (*path.dat, *d.path.dat and *a.path.dat) need to be copied in a single directory for further analysis in MATLAB.

The MATLAB script *evaluate_HMM_fits.m* analyses the HaMMy *path.dat output for all molecules with a trajectory of at least 50 time points. It compares the HMM fits of E, I_D/I_D^{max} and I_A/I_A^{max} and creates a corrected HMM fit of E, in which FRET transition artifacts due to intensity fluctuations of a single fluorophore are skipped. This is achieved by comparing the transition points of the three HMM fits. The transition points of the HMM fits are identified from their corresponding difference vectors z(t+t0)-z(t) which equal zero at all time points except for the transition points (t0 denotes the time resolution of the data array). A transition in the HMM fit of E (E_{HMM}) is then accepted as true if (i.) the HMM fits of I_D/I_D^{max} and I_A/I_A^{max} contain an anticorrelated intensity change of at least

10% amplitude in a window of ±2 time points and (ii.) the change in E_{HMM} is significant ($\Delta E_{HMM} \geq 0.1$).

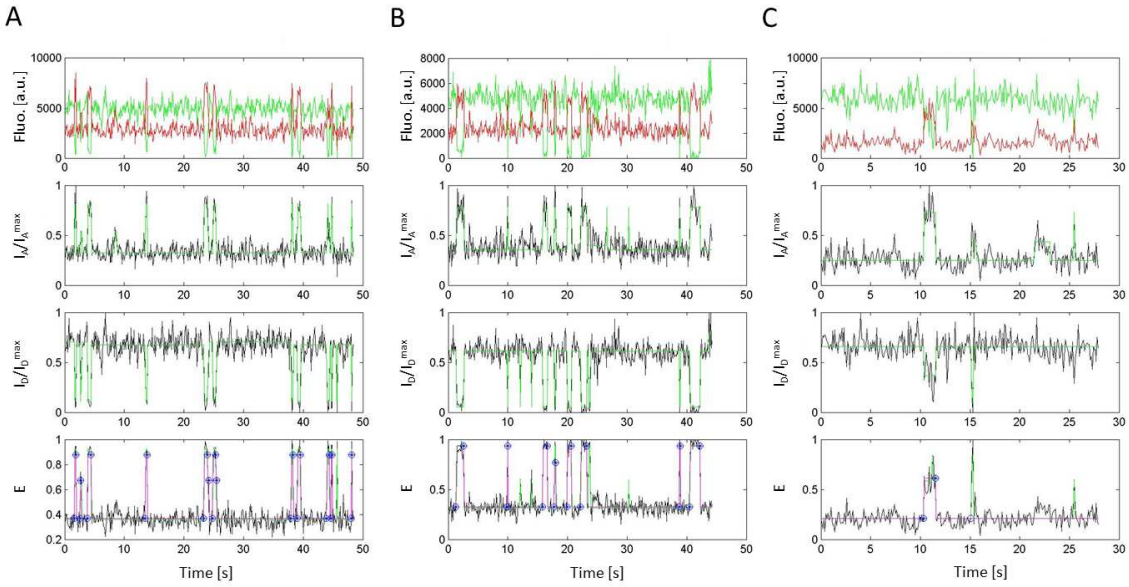

Figure 29: Three example *dyn.jpg output files of *evaluate_HMM_fits.m* for smFRET traces of ^{112}Asw (2 mM Mg^{2+} w/o adenine). The HMM fits performed by HaMMy are depicted in green. The corrected HMM fit of the MATLAB variable true_steps is depicted in purple. Transition points in the MATLAB variable trace_steps are depicted as circles and transition points in true_steps are depicted as stars.

Based on the accepted transition points, the script separates the trace into sequential steps and saves the duration and the mean values of I_D, I_A, E and E_{HMM} for all intervals between the trace-start, the transitions points and the trace-end in the variable trace_steps. The variable trace_steps hence constitutes a representation of the trace in the form of dwell-times. Ideally, the consecutive mean values of E_{HMM} in trace_steps would already adequately model the FRET trajectory for all smFRET traces. However, due to noise in the fluorescence traces, the employed checking routine of transitions occasionally accepts a transition into a state but not a subsequent transition out of that state (or vice versa). This issue is dealt with in the script by additionally performing the following corrections of trace_steps in the variable true_steps. First, all mean values in trace_steps for E_{HMM} with $\Delta E_{HMM} < 0.1$ are replaced by their average. Then, sequential steps with equal mean values for E_{HMM} are concatenated. Based on true_steps, the script analyses the corrected HMM fit of E and creates a transition occupancy density plot (TODP), a transition density plot (TDP) as well as a list of dwell-times. Additionally, it reports a statistics file stating the number of molecules with different properties and their average lifetime.

Importantly, to allow for a visual inspection of the corrected HMM fit of E in true_steps, the script exports a *.jpg figure for each analysed molecule with the scaled fluorescence trajectory, the HMM fits of I_A/I_A^{max} , I_D/I_D^{max} and E, as well as the corrected HMM fit of E (Figure 29). These *.jpg files are named with the ending *dyn.jpg for dynamic molecules and *stat.jpg for static molecules.

3.2.8. Single-molecule aptamer folding dynamics

The aptamer docking dynamics of L2/L3-labelled wt and apoA_{STAB} ^{112}Asw were analysed at equilibrium by acquiring time-resolved smFRET data. The smFRET time traces were collected at three different solution conditions: At 2 mM Mg^{2+}, at 20 mM Mg^{2+} and at 2 mM Mg^{2+} with 100 µM adenine. These three conditions were selected to investigate the effect of the change from sub-saturating to saturating Mg^{2+} conditions and the effect of saturating amounts of adenine at near-physiological Mg^{2+} concentration on the aptamer docking dynamics of the two variants.

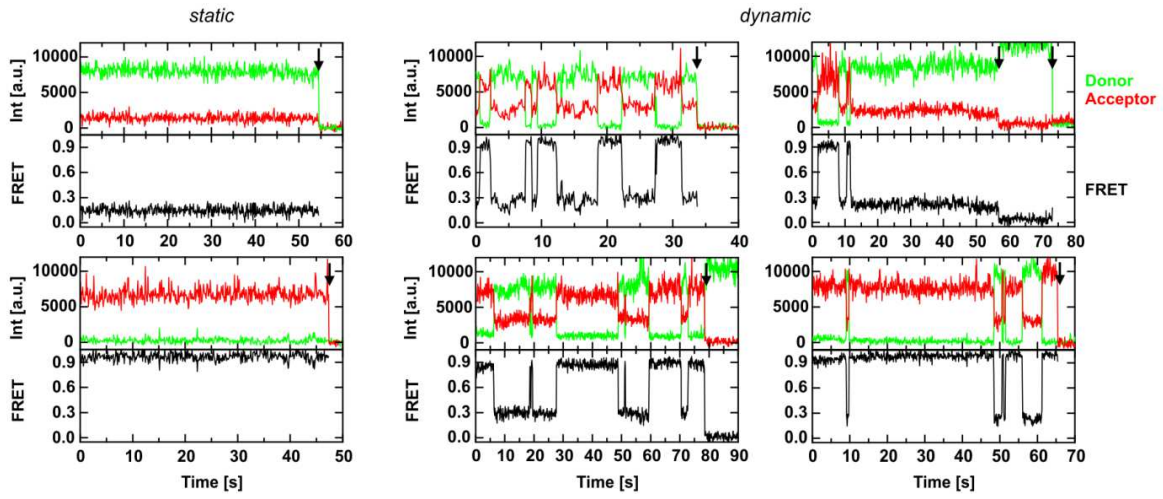

Figure 30: Representative smFRET time traces of L2/L3-labeled ^{112}Asw. The data were collected for wt at 2 mM Mg^{2+} without adenine. The aptamer domain of single riboswitches either statically remained in a long-lived undocked or docked state (left panel), or exhibited dynamics between short-lived and long-lived undocked and docked states (middle and right panel). Photobleaching events are indicated by a black arrow.

Under all three conditions, the smFRET time traces of wt and apoA_{STAB} exhibited heterogeneous dynamics. The traces were grouped into static and dynamic traces. Static traces showed molecules that permanently resided in a long-lived undocked or docked state (Figure 30, left panel). Dynamic traces exhibited at least one transition and showed

molecules that populated long-lived and short-lived undocked and docked states (Figure 30, middle and right panel). The dwell-times of the long-lived undocked and docked states were of the same order of magnitude as the average observation window until photo-bleaching of Cy3 or Cy5 (~30 s). The dwell-times of the short-lived states were on the sub-second to second timescale. The traces were analysed by creating transition occupancy density plots (TODPs) and dwell-time histograms. TODPs are heat maps that give an overview of the distribution of N observed molecules between molecules that showed static behaviour and molecules that exhibited dynamics. In TODPs, every static molecule is plotted as a diagonal peak at its respective FRET efficiency in form of a two-dimensional Gauss function (amplitude 1/N, full width at haf maximum 0.1) and dynamic molecules are plotted as cross-peaks at each pair of FRET efficiencies ($E_{initial}$, E_{final}) for which at least one transition was observed until photo-bleaching (Blanco, 2010).

The TODPs for wt and apoA_{STAB} ^{112}Asw (Figure 31) showed that Mg^{2+} and adenine had a similar effect on the proportioning of static and dynamic molecules: The FRET distribution of the static molecules was shifted from the undocked to the docked state (diagonal of the plots) and the fraction of molecules that showed transitions was decreased (off-diagonal of the plots). Notably, ~30% of the molecules in wt and apoA_{STAB} still showed dynamics at adenine saturation. This strongly suggests that the holo conformation of ^{112}Asw exhibited dynamic tertiary structure heterogeneity between states with undocked and docked aptamer kissing loop motif.

The rate constants for docking and undocking of wt and apoA_{STAB} were obtained from single-exponential fits of dwell-time histograms of the undocked and the docked state, respectively. This analysis was performed in duplicate over 2 datasets of smFRET time traces that were collected in independent experiments. Representative dwell-time histograms are shown in Figure 31. The rate constants obtained from the in duplicate analysis are listed in Table 9. At 2 mM Mg^{2+}, the docking rate constant of wt ^{112}Asw ($k_{dock} = 0.7 \pm 0.1\ s^{-1}$) was equal to the docking rate constant reported for the isolated P1-stabilized *add* Asw aptamer domain ($k_{dock} = 0.68 \pm 0.03\ s^{-1}$), (Dalgarno *et al.*, 2013) and the undocking rate constant of wt ^{112}Asw ($k_{undock} = 0.98 \pm 0.03\ s^{-1}$) was increased ~2-fold compared to the aptamer variant ($k_{undock} = 0.37 \pm 0.04\ s^{-1}$) (Dalgarno *et al.*, 2013). For apoA_{STAB}, k_{dock} and k_{undock} were slightly decreased compared to wt, presumably due to the internal mutations in the P2 helix.

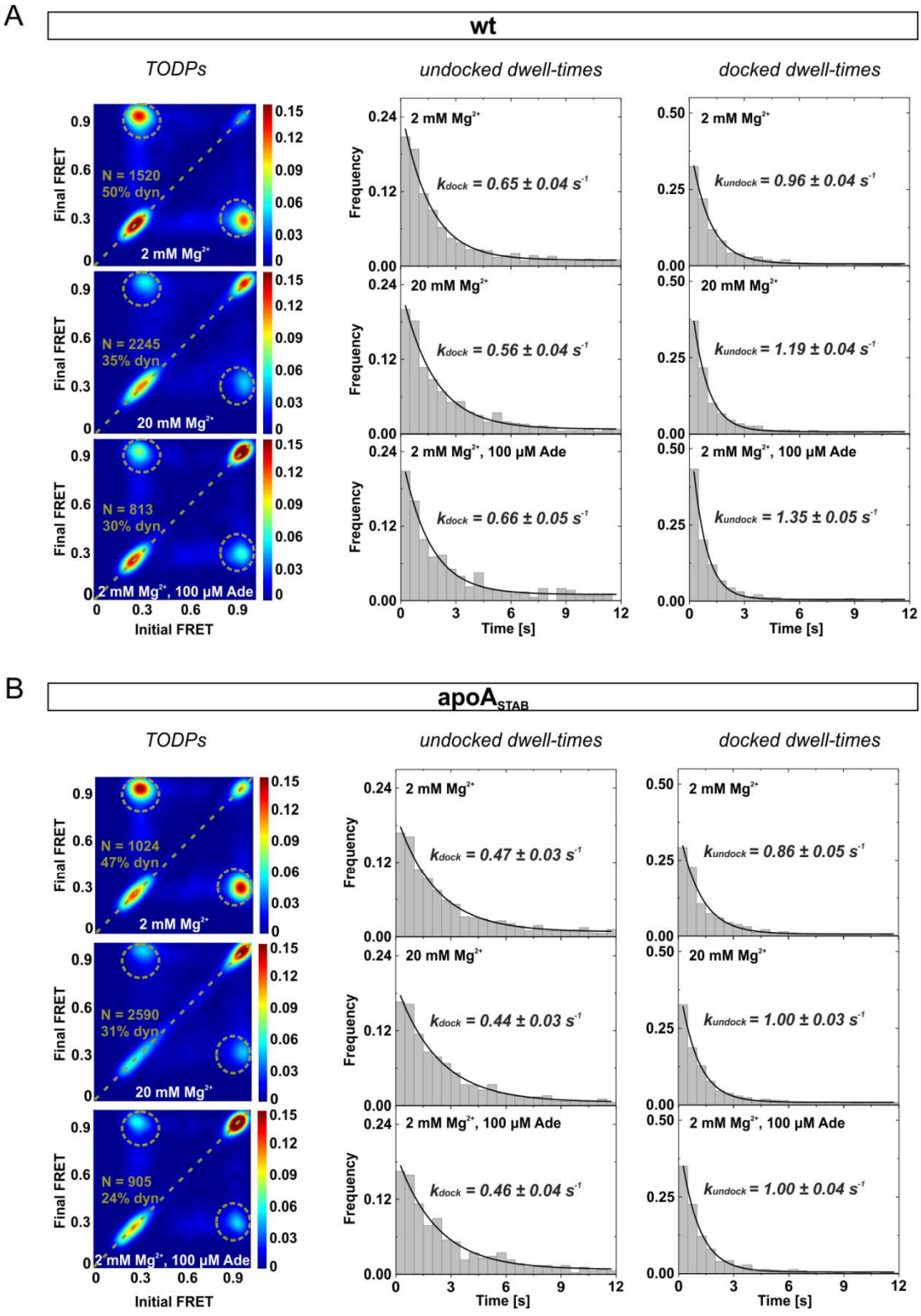

Figure 31: Mg^{2+}- and adenine-dependent equilibrium docking dynamics of the aptamer domain in wt and apoA_{STAB} ^{112}Asw from smFRET time traces. Transition occupancy density plots (TODPs) and dwell-time histograms of the undocked and the docked state obtained from the analysis of smFRET time traces of L2/L3-labelled wt (A) and L2/L3-labelled apoA_{STAB} (B) at 2 mM Mg^{2+}, at 20 mM Mg^{2+} and at 2 mM Mg^{2+} with 100 µM adenine. The number of analysed traces N and the proportion of molecules that exhibited dynamics (*dyn*) are indicated in the TODPs. The dwell-time histograms created from the dynamic sub-population were fitted using single-exponential decay functions which yielded the indicated docking and undocking rate constants.

Table 9: Summarized folding parameters for aptamer docking of wt and apoA$_{STAB}$ ^{112}Asw obtained by smFRET.

Variant	Mg [mM]	Ade [μM]	ΔG(*UD*) *Hist*[a] [kcal/mol]	k_{dock}[b] [s^{-1}]	k_{undock}[b] [s^{-1}]	ΔG_{dock} *Kin*[c] [kcal/mol]
Wt	2	0	0.8 ± 0.1	0.7 ± 0.1	0.98 ± 0.03	0.2 ± 0.1
	20	0	0.2 ± 0.1	0.7 ± 0.1	1.18 ± 0.02	0.3 ± 0.1
	2	100	0.1 ± 0.1	0.7 ± 0.1	1.5 ± 0.2	0.4 ± 0.1
apoA$_{STAB}$	2	0	0.5 ± 0.1	0.5 ± 0.1	0.9 ± 0.1	0.3 ± 0.1
	20	0	0.2 ± 0.1	0.5 ± 0.1	1.1 ± 0.1	0.5 ± 0.1
	2	100	0.0 ± 0.1	0.49 ± 0.04	1.2 ± 0.3	0.5 ± 0.1

[a] Free enthalpy of the docked state relative to the undocked state determined from the FRET distribution observed in the FRET histogram (see section 3.2.6) as $\Delta G(UD) = -RT*\ln(D/U)$. The reported error was calculated from the fitting errors of the FRET populations by Gaussian error propagation.

[b] The reported rate constants are mean values with standard deviations from single-exponentially fitted dwell-time histograms of two independent experiments.

[c] Free docking enthalpy determined from the observable kinetics as $\Delta G_{dock} = -RT*\ln(k_{dock}/k_{undock})$. The reported error was calculated from the standard deviations of the rate constants from two independent experiments by Gaussian error propagation.

Remarkably the Mg^{2+}- and adenine-induced changes in the docking equilibrium did not manifest themselves in the observed rate constants. The docking rate constant of wt and apoA$_{STAB}$ was completely independent of the concentrations of Mg^{2+} and adenine. The undocking rate constant was slightly increased with Mg^{2+} and adenine. The adenine-induced increase of the undocking rate constant translated to a weak ~0.2 kcal/mol increase of the free docking enthalpy calculated from the ratio of k_{dock} over k_{undock}. A comparison between the free docking enthalpies obtained from the observed kinetics and from the relative population of the undocked and the docked state measured in the FRET histogram is included in Table 9. This comparison illustrates that neither the ligand-induced stabilization of the docked state in wt and apoA$_{STAB}$ nor the stabilization of the docked state in apoA$_{STAB}$ towards wt was manifested in the observed kinetics. Single-molecule dwell-time histograms only include a statistically significant number of dwell-times for dwell-times that are substantially shorter than the observation time of the molecules. The TODPs of wt and apoA$_{STAB}$ (Figure 31) have revealed a major proportion of static traces at all investigated conditions (≥ 50%) that were in a long-lived undocked or docked state with a comparable lifetime to the observation time (~30 s). It was thus concluded that the kinetics that accounted for the observed differences in the fractional population of *D* remained hidden in the long-lived states. Given the pronounced similarity of the observed kinetics between wt and apoA$_{STAB}$ in the ligand-free and in the ligand-

saturated form, it was reasoned that the observed kinetics were specifically associated with similar substates of apoA and holo, as discussed further in the section 3.3.

3.3. Discussion

Complementary NMR and smFRET spectroscopic investigation of ^{112}Asw

The ligand-induced conformational switch of ^{112}Asw was characterized at ambient temperature by NMR and by smFRET spectroscopy. The aim of this investigation was to elucidate the number of states, their structural features and their folding pathways in the equilibrium conformational ensemble of ^{112}Asw by integrating complementary data from NMR and smFRET. Figure 32 illustrates the folding of ^{112}Asw as observed by the two solution structural spectroscopies in absence and in presence of adenine at near-physiological Mg^{2+} concentration.

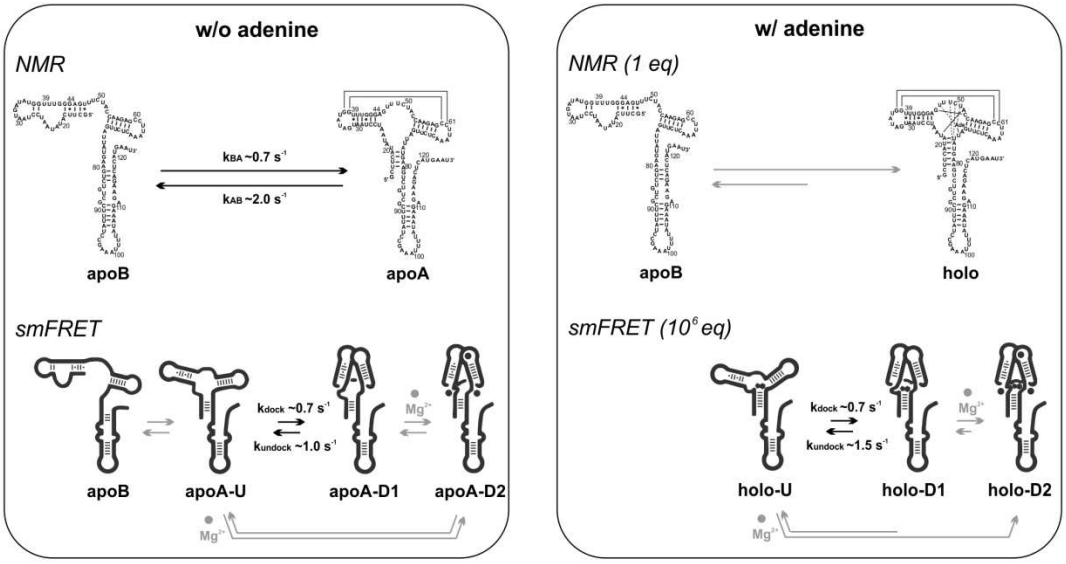

Figure 32: Adenine-dependent equilibrium folding of ^{112}Asw at ambient temperature and near-physiological Mg^{2+} concentration observed by NMR and by smFRET. Without adenine, NMR showed the apoB and an apoA conformation with docked aptamer. The indicated rate constants were determined in a previous study using N_{ZZ}-exchange experiments (Reining *et al.*, 2013). smFRET revealed apoB and 3 substates of apoA: ApoA-U with an undocked aptamer, apoA-D1 with a docked aptamer and apoA-D2 with a Mg^{2+}-bound docked aptamer. For illustrative purposes, the three Mg^{2+} binding sites identified for the isolated aptamer domain were depicted in apoA-D2 (Noeske, Schwalbe and Wöhnert, 2007). The indicted rate constants were Mg^{2+}-independent and hence specifically associated with the folding between apoA-U and apoA-D1. With adenine, NMR exposed an equilibrium shift from the apoB-form to the apoA-form which was captured in the holo conformation. smFRET showed equivalent 3 substates to apoA in the holo conformation. The substate equilibrium was shifted towards holo-D2 and undocking of holo-D1 was modestly accelerated.

NMR and smFRET revealed consistent results regarding the secondary structure bistability of ^{112}Asw between the apoB- and the apoA-form. By NMR, the apoB and the apoA conformation were directly observed. By the comparative smFRET analysis of L2/L3-labelled wild-type (wt), apoB-stabilized mutant ($apoB_{STAB}$) and apoA-stabilized mutant ($apoA_{STAB}$), apoB was indirectly observed as an additional fraction of molecules with undocked aptamer in wt over $apoA_{STAB}$. Both NMR and smFRET revealed an adenine-induced switch from the apoB-form to the apoA-form secondary structure. In NMR, the addition of ~1.1eq adenine over RNA induced a ~4-fold decrease of apoB- and a ~4-fold increase of apoA-specific signals, which indicated a population shift from 80% apoB / 20% apoA to 20% apoB / 80% apoA. In smFRET, the fractional population of apoB could not be quantified, since the apoB and the undocked apoA conformation had a similar L2/L3 distance. But the observation that saturating adenine concentrations equalized the kissing loop docking equilibria of wt and $apoA_{STAB}$ demonstrated that adenine binding eliminated apoB-folding in wt. Given that the different RNA concentrations employed by NMR (0.3 mM) and smFRET (0.1 nM) translated to a six-orders of magnitude difference in the equivalents of Mg^{2+} over RNA, this consistency of the NMR and the smFRET results is striking.

Regarding the allosteric switch of ^{112}Asw, NMR and smFRET both indicated a spatially decoupled switching mechanism that involves no long-range tertiary structural interactions between the aptamer domain and the expression platform. In this study and in the recent study by Reining *et al.* (Reining *et al.*, 2013), no NMR imino resonances were observed that would report on tertiary structural base pairing interactions between single stranded segments of the two riboswitch domains. The smFRET data of L2/P5- and L3/P5-labelled ^{112}Asw confirmed this result. Further, the smALEX data indicated that a perpendicular orientation between the aptamer and the expression platform with a similar L2/P5- and L3/P5-distance was maintained in the distinct conformers of ^{112}Asw. This strongly suggests that the structural basis of the adenine-induced allosteric switch of the *add* adenine riboswitch is base pair formation of the P1 helix coupled to base pair disruption of the P4 helix.

A major difference between the NMR and the smFRET results of ^{112}Asw was that the two methods revealed a different aptamer kissing loop fold of the apoA-form secondary structure. NMR indicated that the apoA and the holo conformation predominantly adopted a folding state with docked aptamer kissing loops at 5 mM Mg^{2+}. The ^{1}H,^{15}N-

BEST-TROSY spectra acquired in absence and in presence of adenine (1.1 eq) showed a single set of imino resonances for the P2/L2 hairpin. Adenine did not significantly affect the imino signal intensity ratio of the L2 base pair reporters over the P2 base pair reporters. smFRET exposed a persistent equilibrium between comparably stable folding states with undocked and docked aptamer kissing loop motif in the apoA and in the holo conformation. The corresponding low-FRET and high-FRET state and transitions between the two FRET states were detected in L2/L3-labelled wt and $apoA_{STAB}$ at all investigated ligand conditions up to saturating concentrations of adenine (100 µM) and Mg^{2+} (20 mM). The analysis of time-resolved smFRET data of wt and $apoA_{STAB}$ at three specific conditions (2 mM Mg^{2+}, 20 mM Mg^{2+} and 2 mM Mg^{2+} with 100 µM adenine) shed new light on the kissing loop docking mechanism of the apoA-form aptamer in the full-length 112-nucelotide sequence. The data showed that Mg^{2+} and adenine did not substantially affect the observable docking and undocking kinetics (~s timescale) but induced formation of a long-lived docked state that exhibited more than 10-fold slower undocking (~30 s timescale). The minimum model that can account for these observations is a three-substate aptamer kissing loop docking mechanism of the apoA-form secondary structure between one undocked (U) and two docked substates (D1 and D2; Figure 32). In absence of adenine, the Mg^{2+}-independent docking equilibrium between apoA-U and apoA-D1 was coupled to Mg^{2+}-dependent folding into apoA-D2. Folding into apoA-D2 could have occurred via the conformational capture pathway from apoA-U over apoA-D1 to apoA-D2, or via the induced fit pathway from apoA-U to apoA-D2. The single-exponential docking rate constant was consistent with the conformational capture pathway. However, a slower induced fit pathway beyond the accessible timescale of the single-molecule dwell-time histograms (~15s) cannot be ruled out. At saturating adenine conditions, equivalent three substates were observed for the holo conformation. Compared to adenine-free conditions, the equilibrium between holo-U, holo-D1 and holo-D2 was shifted towards holo-D2 and the undocking rate constant of holo-D1 was slightly accelerated. Remarkably, the kissing-loop folding kinetics between the two substates U and D1 of the apoA and the holo conformation were on the same timescale as the secondary structure switching kinetics between apoB and apoA that were previously measured by N_{ZZ}-exchange NMR spectroscopy (Reining *et al.*, 2013). Moreover, it is noteworthy that the persistent heterogeneity between undocked and docked substates of the apoA and the holo conformation of ^{112}Asw at near-physiological Mg^{2+} concentration was not observed in the P1 stabilized *add* Asw aptamer (14 bp P1 with 7 GC) (Lemay *et*

al., 2006; Dalgarno *et al.*, 2013). The latter adopted almost exclusively the docked state under comparable solution conditions. This comparison indicates a positive correlation between the stability of the P1 helix and the stability of the kissing loop motif in the *add* Asw aptamer. With this regard, the *add* Asw aptamer differs from the related *pbuE* adenine riboswitch aptamer, for which P1 stabilization resulted in a destabilization of the aptamer kissing loop interaction (Nozinovic *et al.*, 2014).

The fact that NMR and smFRET displayed a different apoA-form aptamer kissing loop fold of ^{112}Asw highlights the necessity to employ complementary solution structural techniques to unravel the folding of large junctional riboswitches. The source of this difference between NMR and smFRET could be related to the following two points:

i. The substate dynamics of the apoA-form aptamer did not lead to chemical shift changes of observable imino NH resonances
ii. The smFRET specific experimental conditions (RNA labelling, RNA immobilization, enzymatic oxygen scavenging system) destabilized the aptamer kissing loop motif

As for point i., it is a known drawback that NMR imino resonances can only be observed for RNA nucleotides in stable base pairs which protect the imino proton from solvent exchange (Fürtig *et al.*, 2016). Characteristic imino signals for the undocked apoA-form aptamer would correspond to the solvent-exposed loop-terminal base pairs of the P2/L2 and the P3/L3 hairpin. Thus, solvent exchange might explain why such signals were not observed for ^{112}Asw. The lack of specific reporter signals for the L2/L3-undocked folding state of the *add* Asw was also evident in a previous NMR spectroscopic study of the isolated *add* Asw aptamer domain (Lee *et al.*, 2010). In future studies, one could attempt to detect the undocked substates of the apoA-form aptamer by NMR with recently reported ^{13}C,^{15}N-SQC experiments (Fürtig *et al.*, 2016). Using such experiments on ^{112}Asw with position specific ^{13}C/^{15}N-labelled G32, G37 or G38 in L2, it may be possible to observe CN correlations of the imino nitrogen associated with an L2/L3 base paired and an L2/L3 unpaired conformation. As for point ii., a destabilizing effect of the dye labels on the aptamer kissing loop interaction of ^{112}Asw is unlikely, since the labelled uridine residues U36 and U62 were solvent exposed in the X-ray structure of the *add* Asw aptamer (Serganov *et al.*, 2004; Lemay *et al.*, 2006). This was further corroborated by bulk fluorescence anisotropy data (section 3.2.4), which contained no indication of a

specific interaction between the fluorophore labels and their environment in the L2/L3 labelling scheme. An effect of 3'-immobilization on the aptamer fold of ^{112}Asw cannot be excluded, but is improbable, because there was a spacing of 8 single-stranded nucleotides (~0.5 nm per nucleotide) and the chemical biotin linker (~2 nm) in between the P4 helix and the streptavidin-biotin-BSA coated coverslip surface. This should have allowed for a global structural arrangement of the riboswitch that is comparable to its free form in solution. The employed glucose oxidase / catalase oxygen system is state of the art in single-molecule fluorescence experiments to inhibit oxygen based photo-bleaching reactions. However, it requires supplementing half-molar D-glucose to the measurement buffer (Swoboda *et al.*, 2012). Recently, it was shown that a comparable concentration of ethylene glycol changed the SHAPE reactivity profile of the *add* Asw aptamer domain (Tyrrell, Weeks and Pielak, 2015). Therefore, besides surface immobilisation, also the smFRET specific buffer conditions could have had an impact on the folding of ^{112}Asw in the smFRET experiment. In conclusion, NMR of the imino protons is intrinsically biased towards L2/L3-docked states of ^{112}Asw. Additionally, the specific experimental conditions of smFRET could have affected the folding equilibrium of ^{112}Asw, potentially favouring L2/L3-undocked states.

Conformational dynamics of ^{112}Asw at near physiological concentrations of RNA, Mg^{2+} and adenine

The smFRET experiments of ^{112}Asw provided insight into the conformational dynamics of ^{112}Asw at near-physiological concentrations of RNA, Mg^{2+} and adenine. The bacterial RNA concentration (~1nM) (Church *et al.*, 2000) is close to the RNA concentration employed by smFRET (0.1 nM). Metabolomic studies of *E. coli* have suggested that the bacterial adenine concentration changes in the range between 1 µM and 100 µM (Ishii *et al.*, 2007; Bennett *et al.*, 2009). The Mg^{2+} concentration in *E. coli* was recently determined ~1 mM (Tyrrell *et al.*, 2013). The here observed adenine dependence of the aptamer kissing loop docking equilibrium of ^{112}Asw at near-physiological Mg^{2+} concentration (2 mM) demonstrated that a switch from the lowest (1 µM) to the highest (100 µM) adenine concentration measured in *E.coli* would saturate the riboswitch with ligand, eliminate apoB folding and induced a ~2-fold increase in the population of a Mg^{2+}-bound long-lived docked substate of the apoA-form aptamer. At a 10-fold lower

Mg^{2+} concentration, no adenine induced aptamer folding could be observed. At a 10-fold higher and saturating Mg^{2+} concentration, the aptamer kissing-loop folding equilibrium of adenine-free ^{112}Asw was strikingly similar to adenine-saturated ^{112}Asw: Saturating Mg^{2+} conditions also impaired apoB folding and induced a ~2-fold increase in the population of a long-lived docked apoA-form aptamer. The Mg^{2+} titration data hence indicated that the *add* adenine riboswitch is specifically adapted to provide greatest ligand sensitivity of the aptamer kissing loop folding equilibrium at physiological Mg^{2+} concentration. Previously reported smFRET data have revealed a fundamentally distinct Mg^{2+}-dependence for aptamer kissing loop folding of the structurally related, transcriptional-operating *xpt* guanine riboswitch (Brenner *et al.*, 2010). The *xpt* guanine aptamer showed greatest guanine sensitivity of the kissing loop folding equilibrium at 0.1 mM Mg^{2+}, which is 10-fold lower than the physiological Mg^{2+} concentration. The different Mg^{2+}-sensitivity of aptamer kissing loop folding in the *add* adenine and the *xpt* guanine aptamer is in line with a fundamentally different role of aptamer kissing loop folding for translational- and transcriptional-operating purine riboswitches. In case of the translational-operating *add* adenine riboswitch, the ligand-dependent kissing loop docking equilibrium can contribute to the relative stability of conformational states whose equilibrium interconversion directly modulates translational regulation. In case of the transcriptional *xpt* guanine riboswitch, the enhanced stability of the L2/L3-docked state at physiological Mg^{2+} concentration might be required to achieve efficient ligand binding during transcription which is necessary for ligand-modulated co-transcriptional riboswitch folding.

The regulatory conformational switch of ^{112}Asw

The NMR characterization of the apoB-stabilized mutant of ^{112}Asw strongly suggested that apoB adopts a short P4 helix with only three stable base pairs and that the AUG start codon does not form persistent base pairing interactions. The NMR data of wt ^{112}Asw demonstrated that adenine binding decreases the fractional population of apoB and melts the P4 helix. These observations are in excellent agreement with the hypothesis that the *add* Asw regulates the specific binding of the 30S ribosome to the Shine Dalgarno sequence by switching between P4-closed (gene OFF) and P4-open (gene ON) states. However, it remains unknown, which specific conformational transitions correlate with P4 melting. P4 melting could be induced already by conformational switching between

apoB and apoA-U (as depicted in Figure 32). This assumption is consistent with two previous data points of the apoA-stabilized mutant. The $^{1}H,^{15}N$-HSQC spectrum of the apoA-stabilized mutant without Mg^{2+} and adenine lacked the reporter signals of the P4 base pairs (Reining *et al.*, 2013). And, in transcription-translation coupled luciferase *in vitro* expression assays with a luciferase reporter gene under riboswitch control the apoA-stabilized mutant proved to be an ON-state mutant (Reining *et al.*, 2013). Still, it could be possible that folding of the aptamer tertiary structure modulates P4 stability, and that apoA-U, apoA-D1, apoA-D2, holo-U, holo-D1 and holo-D2 all have a differentially stable P4 helix. Further studies should aim at elucidating the structural differences between these states with respect to P4 in order to map their regulatory property. Recently published simulations of 3-state riboswitch regulation profiles have shown that it has a pronounced effect on the temperature dependence of the switching efficiency of the *add* Asw, whether apoA is to be considered an OFF-state or an ON-state (Fürtig *et al.*, 2015). In these simulations, the scenario of two functional OFF states (apoB and apoA) and one functional ON state (holo) resulted in a largely constant switching efficiency between 4 °C and 30 °C (≥70%). The scenario of one functional OFF state (apoB) and two functional ON states (apoA and holo) resulted in a decrease of the switching efficiency between 25 °C and 30 °C (from 70% to 50%). Thus, not only the riboswitch function, but also the thermostat function of the *add* Asw critically depends on the structural and functional properties of the apoA conformation.

Chapter 4: Integrated NMR and smFRET analysis of ^{127}Asw

4.1. Introduction

The mRNA channel of the bacterial ribosome can incorporate approximately 30 mRNA nucleotides that are wrapped around the neck of the 30S ribosomal subunit (Yusupova *et al.*, 2001). Since the Shine Dalgarno sequence and the AUG start codon of the *add* adenine riboswitch are located within the last 3'-terminal 15 nucleotides of ^{112}Asw, this means that a larger mRNA construct than ^{112}Asw is required to study the formation of a biologically representative 30S translation initiation complex between the riboswitch and the 30S ribosomal subunit. In the Schwalbe lab, ribosome interaction studies have therefore been initiated with a 127-nucleotide construct of the native full-length *add* Asw, ^{127}Asw, which contains a 3'-extension of 15 nucleotides that match the adenosine deaminase encoding sequence from *Vibrio vulnificus* CMCP6.

In this study, ^{127}Asw was investigated by NMR and by smFRET spectroscopy to elucidate whether the 3'-extension of ^{112}Asw to ^{127}Asw affects the folding and or the relative stability of the apoB, the apoA and the holo conformation. In section 4.2.1, the adenine-induced switch of the global base pairing structure of ^{127}Asw is characterized by NMR of the imino protons. Sections 4.2.2 and 4.2.3 present an smFRET analysis of the ligand-dependent aptamer kissing loop folding equilibrium and the single-molecule aptamer folding dynamics of L2/L3-labelled ^{127}Asw. The NMR and smFRET results suggested that the the 3'-extension of the *add* Asw stabilized the P4 helix and destabilized the aptamer kissing loop interaction. Consequently, in section 4.2.4, it was tested whether smFRET of P5/L3-labelled ^{127}Asw could shed light into a possible interaction between the 3'-extension and the *add* Asw aptamer domain. The structural effects of the 3'-extension of ^{112}Asw to ^{127}Asw and their implication for the regulatory mechanism of the *add* Asw are discussed in section 4.3.

4.2. Results

4.2.1. NMR analysis of the adenine-dependent base pairing

To investigate whether the base pairing structure of ^{127}Asw differs from ^{112}Asw, the adenine-induced changes in the imino signal intensities of ^{127}Asw were evaluated in an analogous manner to ^{112}Asw (see section 3.2.1).

The $^{1}H,^{15}N$-BEST-TROSY spectra of ^{127}Asw were acquired at 298 K and 5 mM Mg^{2+} in absence and in presence of 1.1 eq adenine (Figure 33A,B). They showed the same peak pattern that was observed for ^{112}Asw, albeit with a few differences. In ^{127}Asw without adenine, G43b was shifted beyond linewidth, G44 split into a double peak and G112 appeared. Furthermore, four new and unassigned isolated signals of ^{127}Asw were observed: G?1 and U?1 in the ^{1}H chemical shift region about 11 ppm, which is characteristic of non-canonical base pairs, and G?2 and U?2 with more downfield ^{1}H chemical shifts typical of canonical base pairs. The four new signals had a comparable intensity to the apoA-specific resonances of the aptamer domain. However, these signals could not be assigned due to a lack of cross peaks in $^{1}H,^{1}H$-NOESY spectra of ^{127}Asw at 298 K (data not shown). In ^{127}Asw with adenine, the expression platform signals U82, G112 and G115 remained visible. G43b was shifted, and G14b as well as the apo-specific additional P3 signals of G72 and U71 were still observed. The P1 residues G78 and U77 showed double peaks consisting of a major peak with a minor peak shoulder (~15% intensity of the major peak). The imino signal G?2 was still detected but G?1, U?1 and U?2 were barely visible at contour levels above noise.

The adenine-dependent relative intensities of the selected imino reporter signals of the *add* Asw in the 127-nucleotide long construct are shown in Figure 33C. Without adenine, ^{127}Asw exhibited a ~5-fold higher intensity of P1b over P1/P2. This indicated that the population ratio of apoB to apoA was similar to ^{112}Asw, for which a ~6-fold higher intensity of P1b over P1/P2 had been measured (see section 3.2.1). However, the relative intensities of the P4 and P5 imino reporter signals of ^{127}Asw were ~5-fold increased compared to ^{112}Asw. With this regard, ^{127}Asw resembled ^{112}Asw-*nx* that contained 40 mM Li^{+} (see section 3.2.2). In contrast to ^{112}Asw and ^{112}Asw-*nx*, the adenine-free ^{127}Asw also exhibited a ~2-fold higher intensity of the P1 reporter signals over P2. Upon addition of 1.1 eq adenine, ^{127}Asw formed a holo conformation as evidenced by the appearance of the J1-2/J2-3 resonances with comparable intensities to the P1 resonances. Adenine binding induced a ~2-fold decrease of the apoB-specific P1b signals and a very pronounced, ~30-fold increase of the signals of the P1/P2 module. These observations suggest that adenine did not only shift the secondary structure equilibrium from the apoB-form to the apoA-form, but also induced considerable global structural changes of the apoA-form aptamer that stabilized the P1/P2 imino protons against solvent exchange.

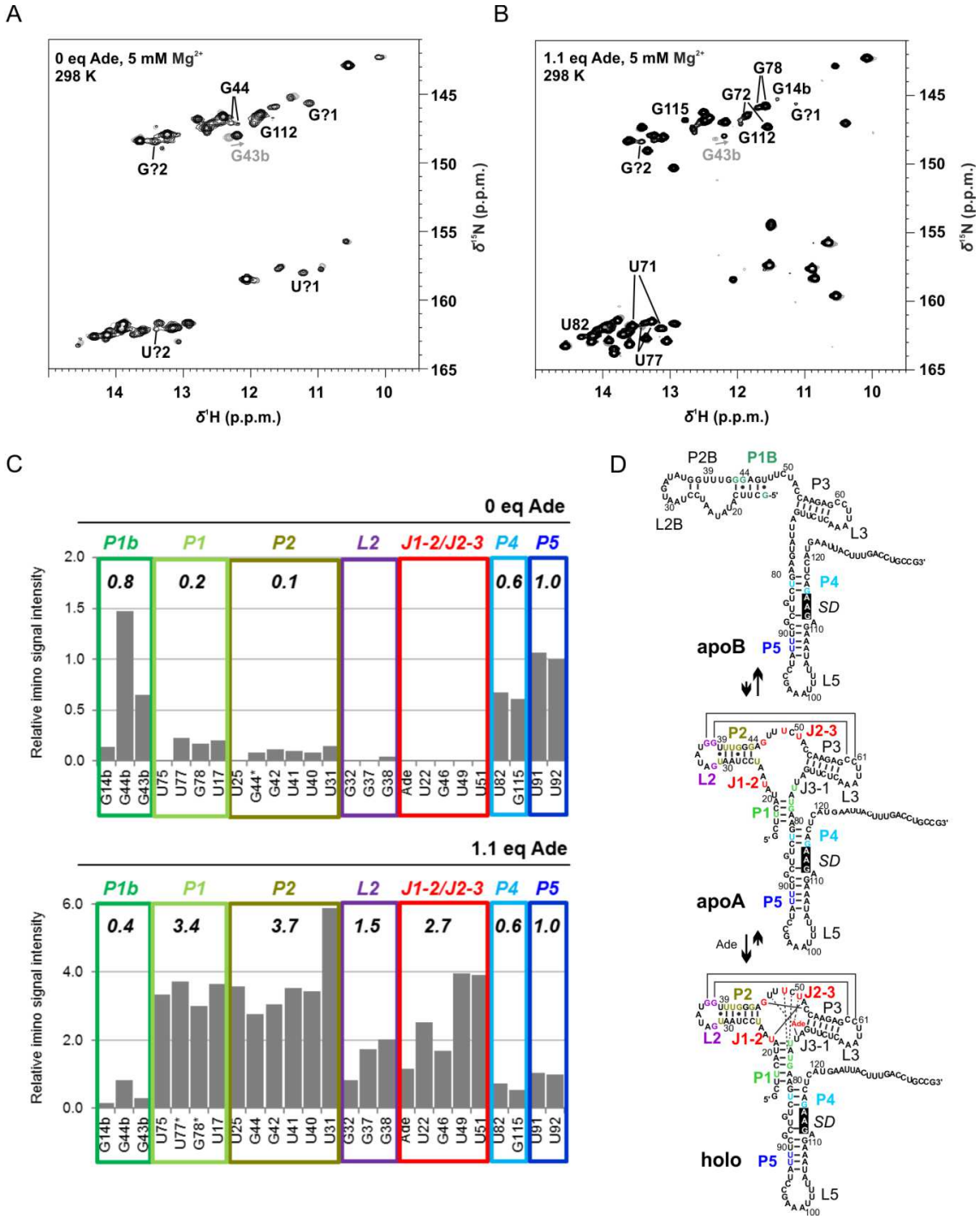

Figure 33: Adenine-induced switch in ^{127}Asw base pairing monitored by NMR. (A), (B) Overlay of the imino region of the ^{1}H,^{15}N-BEST-TROSY spectra of ^{127}Asw (0.7 mM; black) and ^{112}Asw (0.3 mM; grey) at 298 K and 5 mM Mg^{2+} in absence (A) and in presence (B) of 1.1 eq adenine. Assigned signals that are missing or shifted beyond linewidth in ^{127}Asw are annotated in grey. Assigned signals that appear in ^{127}Asw are annotated in black. **(C)** Relative intensities of the imino reporter signals (normalized to U92) without adenine (top) and with 1.1 eq adenine (bottom). The average imino reporter signal intensity of the structural modules is indicated in black. **(D)** Model of the 3-state conformational equilibrium of ^{127}Asw with colour-coded G/U imino reporter residues and highlighted Shine Dalgarno sequence (SD). Residues denoted * exhibited double peaks and the peak height of the major peak was used as a measure of the signal intensity.

Intriguingly, adenine did not affect the relative imino signal intensities of P4 and P5 in ^{127}Asw.

Taken together, the NMR data of ^{127}Asw showed that the additional 15 3'-terminal nucleotides of ^{127}Asw compared to ^{112}Asw stabilized the expression platform base pairs without changing the chemical shifts of the corresponding imino resonances. This stabilization had the consequence that adenine-induced formation of the holo conformation did not induce an allosteric melting of P4 compared to P5. The structural basis of the adenine induced allosteric switch of the *add* Asw is that base pair formation of the P1 helix is coupled to base pair disruption of the P4 helix (see section 3.3). With a three-nucleotide overlap only between P1 and P4, it is highly unlikely that adenine-induced P1-stabilization could have destabilized both the P4 and the P5 helix of ^{127}Asw in an equal manner. The NMR data of ^{127}Asw therefore suggest that the 3-state conformational equilibrium of ^{127}Asw needs to be described by a model in which apoB, apoA and holo all exhibit a P4-closed expression platform (Figure 33D).

4.2.2. FRET histogram analysis of aptamer folding

Like for ^{112}Asw, the aptamer kissing loop docking equilibrium of ^{127}Asw was analysed via FRET histograms of L2/L3 labelled constructs (compare section 3.2.6). First, adenine-dependent kissing-loop docking of the wild-type, the apoB-stabilized mutant (apoB$_{STAB}$) and the apoA-stabilized mutant (apoA$_{STAB}$) was monitored in their respective 127-nucleotide long sequences at 2 mM Mg^{2+}. Second, the Mg^{2+}-dependence was analysed for wt and the apoA-stabilized mutant by comparing three different Mg^{2+} concentrations (0, 2 and 20 mM) in absence and in presence of adenine.

Figure 34 shows the adenine-induced switch in the aptamer kissing loop docking equilibrium of L2/L3-labeled apoB$_{STAB}$, wt and apoA$_{STAB}$ ^{127}Asw at 2 mM Mg^{2+}. The 127mer variants exhibited identical two FRET states to the 112mer variants (see section 3.2.6): A low-FRET state at E~0.2 that corresponds to L2/L3-undocked molecules (*U*) and a high-FRET state at E~0.9 that corresponds to L2/L3-docked molecules (*D*). However, apoB$_{STAB}$, wt and apoA$_{STAB}$ ^{127}Asw without adenine were all almost exclusively in the undocked state (>95%), while wt and apoA$_{STAB}$ of ^{112}Asw adopted 20% and 30% docked molecules, respectively (see section 3.2.6). ApoB$_{STAB}$ ^{127}Asw was unresponsive to adenine. For wt and apoA$_{STAB}$ ^{127}Asw, the addition of 100 μM adenine

induced a shift of the conformer distribution to ~20% docked state, which is ~2-fold less docked molecules than observed in ^{112}Asw (see section 3.2.6). Since 2 mM Mg^{2+} could not induce significant folding of the docked aptamer in ^{127}Asw, wt and apoA_{STAB} were indistinguishable both in absence and in presence of adenine. The secondary structure bistability of wt thus remained entirely hidden in these data. The sensitivity of the docking equilibrium of ^{127}Asw towards adenine was higher than in ^{112}Asw, in that adenine induced a ~7-fold increase in the fractional population of *D*, while this increase was ~2-fold for ^{112}Asw (see section 3.2.6).

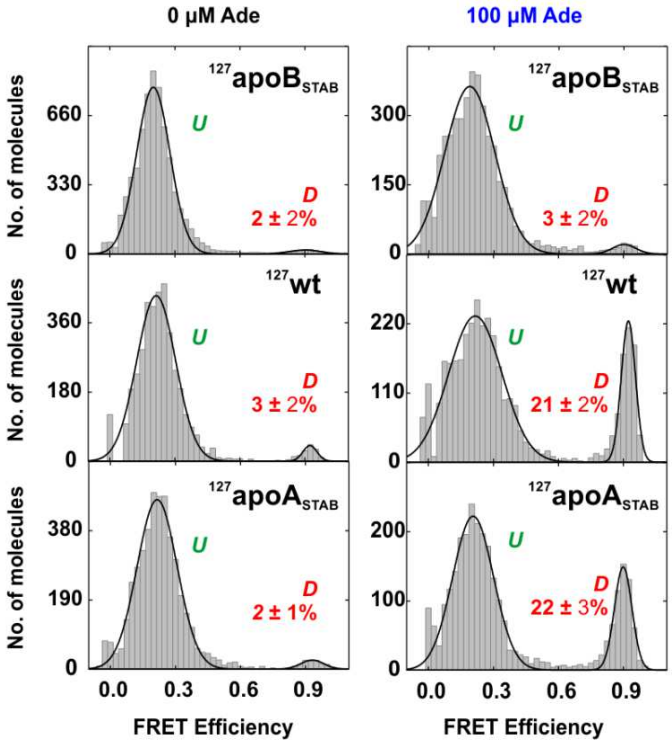

Figure 34: FRET histogram analysis of adenine-dependent aptamer docking of ^{127}Asw. smFRET histograms of L2/L3-labeled 127-nucleotide apoB_{STAB}, wt and apoA_{STAB} in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. The fractional population of the docked state *D* is indicated in percent.

The Mg^{2+}-dependence of aptamer docking of wt and apoA_{STAB} ^{127}Asw was investigated by collecting FRET histograms at 0 mM, 2 mM and 20 mM Mg^{2+} in absence and in presence of 100 µM adenine, and plotting the fractional population of the docked state in a bar diagram (Figure 35). As expected, ^{127}Asw showed no adenine-dependent aptamer docking in the absence of Mg^{2+}. The docked state of ^{127}Asw could be stabilized by the addition of Mg^{2+} in the absence of adenine. But the aptamer kissing loop docking equilibrium of ^{127}Asw without adenine was less Mg^{2+}-responsive than for ^{112}Asw. The

addition of 20 mM Mg^{2+} in the absence of adenine induced only ~15% docked molecules in wt and in apoA_{STAB} ^{127}Asw compared to ~40% in ^{112}Asw (see section 3.2.6). Notably also the addition of both 100 µM adenine and 20 mM Mg^{2+} did not compensate the difference in the fractional population of the docked state between ^{127}Asw and ^{112}Asw. At all investigated solution conditions, wt and apoA_{STAB} ^{127}Asw exhibited an equal fractional population of *D*.

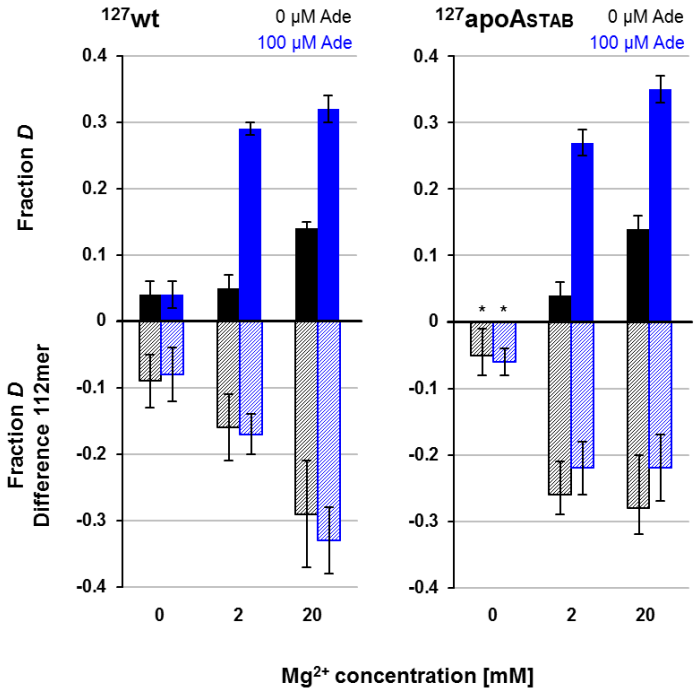

Figure 35: FRET histogram analysis of Mg^{2+}-dependent aptamer docking of ^{127}Asw. Fractional population of the docked state *D* in smFRET histograms of L2/L3-labeled 127-nucleotide wt and apoA_{STAB} at 0, 2 and 20 mM Mg^{2+} without adenine (black bars) and with 100 µM adenine (blue bars) and corresponding difference to the fractional population of *D* in the 112nt constructs. *docked state not detectable

4.2.3. Single-molecule aptamer folding dynamics

The aptamer kissing loop equilibrium docking dynamics of L2/L3-labelled wt ^{127}Asw without and with 100 µM adenine at 2 mM Mg^{2+} were analysed by collecting smFRET time traces. These data were acquired by Klara R. Mertinkus as part of a supervised Bachelor thesis. As expected, the traces of ^{127}Asw showed static molecules that permanently resided in a long-lived undocked or docked state and dynamic molecules that exhibited transitions between long-lived and short-lived undocked and docked states (data

not shown). The traces were evaluated by creating TODPs and dwell-time histograms (Figure 36).

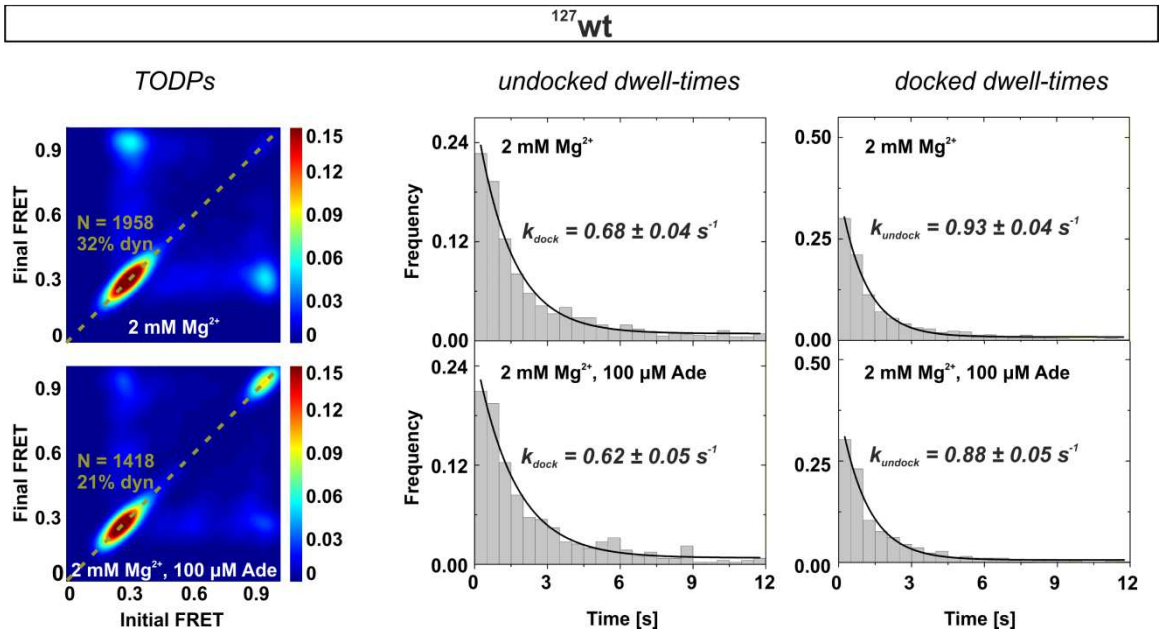

Figure 36: Adenine-dependent aptamer docking dynamics of wt ^{127}Asw. Transition occupancy density plots (TODPs) and dwell-time histograms of the undocked and the docked state obtained from the analysis of smFRET time traces of L2/L3-labelled wt ^{127}Asw in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. The number of analysed traces N and the proportion of molecules that exhibited dynamics (*dyn*) are indicated in the TODPs. The dwell-time histograms created from the dynamic sub-population were fitted using single-exponential decay functions which yielded the indicated docking and undocking rate constants.

As evident from the TODPs (Figure 36, left panel), the addition of adenine to ^{127}Asw decreased the fraction of molecules that showed transitions and increased the fractional population of static molecules that resided in a long-lived docked state. Without adenine, the number of dynamic molecules was reduced in ^{127}Asw (~30% dyn at 2 mM Mg^{2+}) compared to ^{112}Asw (~50% dyn at 2 mM Mg^{2+}; see section 3.2.8). With adenine, ^{127}Asw exhibited only a very weak TODP crosspeak for molecules that transitioned between E~0.3 and E~0.9. Thus, for this condition, the ~20% smFRET traces that were identified as dynamic traces by the automated analysis routine were additionally inspected manually. This revealed that ~10% of all traces actually showed transitions between E~0.3 and E~0.9, while 4% showed transitions between E~0.3 and E~0.6 and 6% corresponded to erroneous smFRET traces that exhibited multiple donor or multiple acceptor bleaching events. Representative smFRET traces of ^{127}Asw that transiently sampled a FRET state with E~0.6 are shown in Figure 37. The comparison of transition density plots (TDPs) created for the dynamic sub-population of ^{127}Asw and ^{112}Asw in

absence and in presence of 100 µM adenine at 2 mM Mg^{2+} (Figure 38) illustrated that only adenine-bound ^{127}Asw exhibited appreciable transition density between E~0.3 and E~0.6. Since the FRET state with E~0.6 was only detected in 4% of the molecules for ^{127}Asw at one specific solution condition, it was not considered significant.

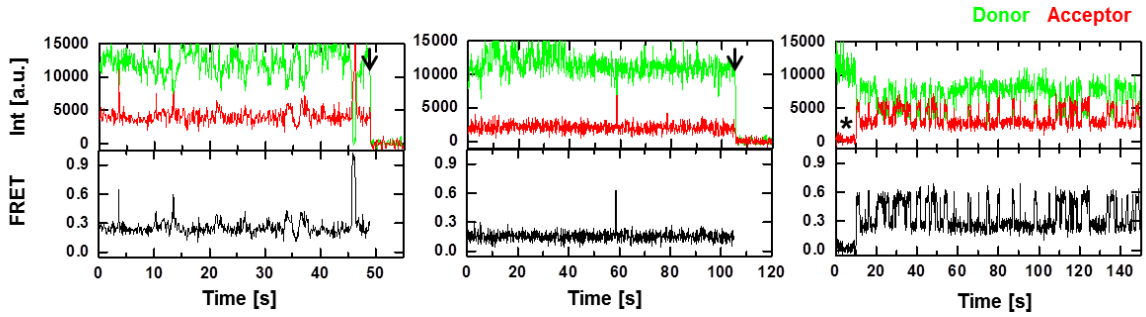

Figure 37: smFRET time traces of L2/L3-labeled ^{127}Asw that showed transitions between states with low FRET efficiency (E~0.3) and intermediate FRET efficiency (E~0.6). The data was collected at 2 mM Mg^{2+} with 100 µM adenine. Photobleaching events are indicated by a black arrow. The black star indicates an acceptor blinking event.

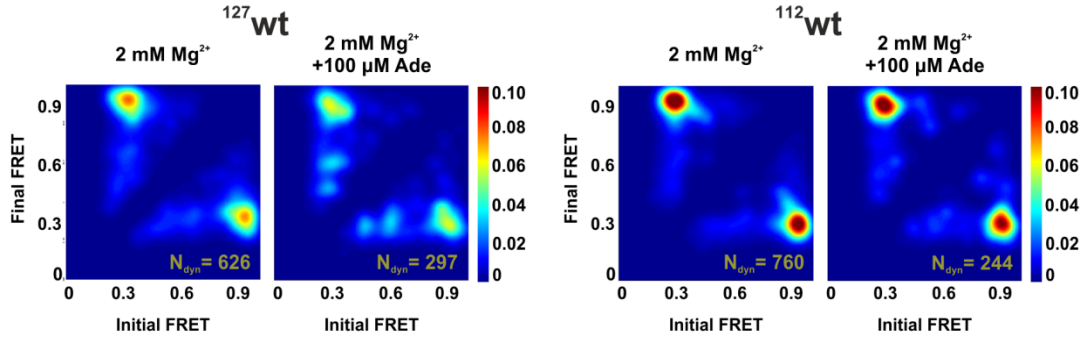

Figure 38: Transition density plots (TDPs) of L2/L3-labeled ^{127}wt and ^{112}wt. The TDPs visualize as a heat map the relative number of transitions between different pairs of FRET states in the subpopulation of dynamic molecules (N_{dyn}) and were constructed from the same datasets that was used for the TODPs shown in Figure 36 for ^{127}wt and in Figure 31for ^{112}wt.

The dwell-time data of ^{127}Asw (Figure 36 middle and right panel) show that the addition of adenine did not affect the observable rate constants for docking (k_{dock}) and undocking (k_{undock}). Intriguingly, k_{dock} and k_{undock} of ^{127}Asw without adenine were equal (within error limits) to k_{dock} and k_{undock} of ^{112}Asw (see Table 10). The rate constants obtained in presence of 100 µM adenine were also similar to ^{112}Asw, but the subtle increase of k_{undock} that was observed for ^{112}Asw with adenine could not be detected for ^{127}Asw. While the free enthalpies for aptamer kissing loop docking calculated from the FRET histograms were pronouncedly increased in ^{127}Asw compared to ^{112}Asw (+1.2 kcal/mol in absence and +0.7 kcal/mol in presence of adenine) the free enthalpies for aptamer kissing loop

docking calculated from the observable kinetics were equal (within experimental error) for ^{127}Asw and ^{112}Asw (see Table 10).

Overall, the TODPs and the dwell-time histograms showed that the effect of adenine on the aptamer docking dynamics of ^{127}Asw was similar to ^{112}Asw. The difference between ^{112}Asw and ^{127}Asw observed in the FRET histograms was thus due an additional undocked state that did not exhibit docking at the accessible timescale of the dwell-time data (~12 s).

Table 10: Summarized folding parameters for aptamer docking of ^{127}Asw in comparison to ^{112}Asw. The data for ^{112}Asw are taken from section 3.2.8.

Variant	Mg [mM]	Ade [µM]	$\Delta G(UD)$ *Hist*[a] [kcal/mol]	k_{dock}[b] [s^{-1}]	k_{undock}[b] [s^{-1}]	ΔG_{dock} *Kin*[c] [kcal/mol]
^{112}Asw	2	0	0.8 ± 0.1	0.7 ± 0.1	0.98 ± 0.03	0.2 ± 0.1
	2	100	0.1 ± 0.1	0.7 ± 0.1	1.5 ± 0.2	0.4 ± 0.1
^{127}Asw	2	0	2.0 ± 0.2	0.64 ± 0.06	0.89 ± 0.06	0.2 ± 0.1
	2	100	0.78 ± 0.04	0.56 ± 0.08	0.8 ± 0.1	0.2 ± 0.1

[a] Free enthalpy of the docked state relative to the undocked state determined from the FRET distribution observed in the FRET histogram (see section 4.2.2) as $\Delta G(UD) = -RT*\ln(D/U)$. The reported error was calculated from the fitting errors of the FRET populations by Gaussian error propagation.

[b] The reported rate constants are mean values with standard deviations from single-exponentially fitted dwell-time histograms of two independent experiments.

[c] Free docking enthalpy determined from the observable kinetics as $\Delta G_{dock} = -RT*\ln(k_{dock}/k_{undock})$. The reported error was calculated from the standard deviations of the rate constants from two independent experiments by Gaussian error propagation.

4.2.4. FRET histogram analysis of the L3/P5-labelled construct

Since the population of an additional undocked state in ^{127}Asw was inferred from the smFRET data of the L2/L3-labelled construct, FRET histograms of ^{127}Asw were also collected in the L3/P5-labelling scheme. It was reasoned that the L3/P5-labelled construct might reveal an interaction of the 3'-extension with the aptamer domain, since such interaction could affect the orientation of the P3 helix. Figure 39 shows the FRET histograms of P5/L3-labeled apoB_{STAB}, wt and apo$_{ASTAB}$ ^{127}Asw at 2 mM Mg^{2+} in absence and in presence of 100 µM adenine. These data were acquired by Klara R. Mertinkus as part of a supervised Bachelor thesis. The FRET histograms all show a similar FRET distribution with a single Gaussian FRET state centred at E~0.15. The FRET efficiency of P5/L3-labelled ^{127}Asw was thus equal to the FRET efficiency of P5/L3-labelled ^{112}Asw (E~0.15; see section 3.2.5). A difference in the relative orientation of the P3 helix and the

expression platform between ^{127}Asw and ^{112}Asw could not be observed in this labelling scheme.

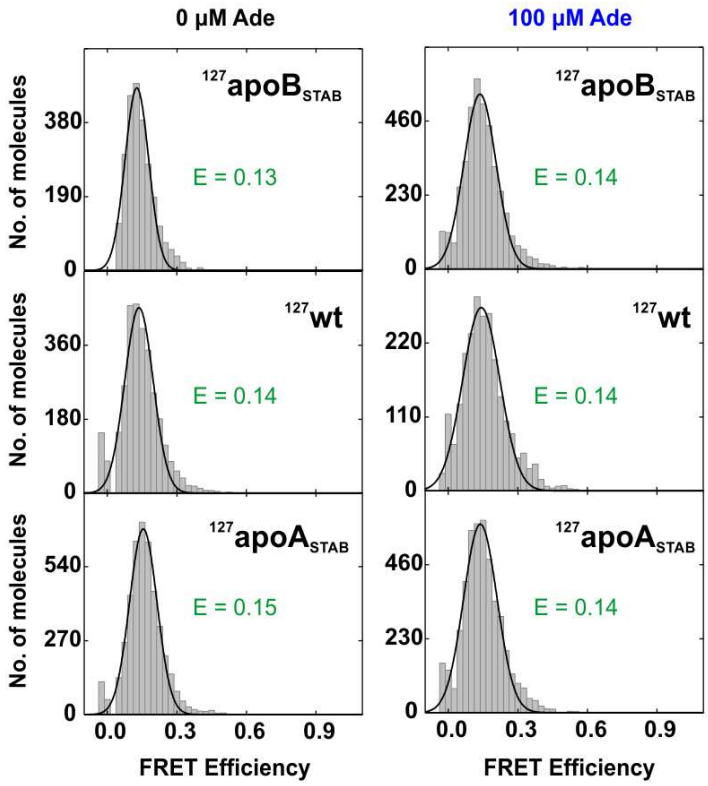

Figure 39: FRET histogram analysis of L3/P5-labelled ^{127}Asw. smFRET histograms of P5/L3-labeled apoB-stabilized mutant (apoB$_{STAB}$), wild-type (wt) and apoA-stabilized mutant (apoA$_{STAB}$) ^{127}Asw in absence and presence of 100 µM adenine at 2 mM Mg^{2+}. The observed low-FRET state was fitted with a Gaussian with the indicated peak centre FRET efficiency E.

4.3. Discussion

In this chapter, ^{127}Asw was characterized by NMR and by smFRET using the L2/L3 and the P5/L3 labelling scheme. While ^{127}Asw formed no other long-lived secondary structure elements than ^{112}Asw, the 15-nucleotide 3'-terminal extension of the full-length riboswitch led to an unexpected destabilization of the aptamer kissing loop motif and appeared to decouple aptamer folding from P4 base pairing in the expression platform.

^{127}Asw without adenine exhibited a similar population ratio between apoB and apoA to ^{112}Asw, as judged by the relative NMR signal intensities of imino base pair reporter signals. However, the apoA aptamer of ^{127}Asw was less structurally pre-organized. The FRET histograms acquired in the L2/L3-labelling scheme showed that near-physiological

Mg^{2+} concentration (2 mM) did not induce folding of the aptamer kissing-loop motif in the wild-type and in the apoA-stabilized mutant of ^{127}Asw (fractional population of *D* < 5%). The destabilized aptamer kissing-loop interaction correlated with an increased stability of the P4 and the P5 base pairs against solvent exchange. With adenine, ^{127}Asw clearly adopted the compact aptamer tertiary structure of the holo conformation but exhibited no melting of the P4 helix. It was therefore suggested that the apoB, the apoA and the holo conformation of ^{127}Asw all had a P4-closed expression platform. Also, the folding of the holo aptamer of ^{127}Asw differed from ^{112}Asw. The FRET histograms of the wild-type and the apoA-stabilized mutant of ^{127}Asw with 100 µM adenine and 20 mM Mg^{2+} showed a ~2-fold reduced fractional population of the docked state compared to ^{112}Asw, indicating that the kissing-loop motif was less stable in ^{127}Asw even under ligand-saturation. The NMR spectra in presence of adenine showed P1 and P3 related double peaks. That neither the NMR data of ^{127}Asw nor the smFRET data of P5/L3-labelled ^{127}Asw provided evidence for an interaction between the 3'-extension of the expression platform and upstream sequence elements of the aptamer domain points to an unspecific effect of the 3'-extension on P4-stability. Taken together, the NMR and smFRET results are consistent with the hypothesis that a stabilized P4 helix translates to a destabilized P1 helix, which in turn translates to a destabilized aptamer kissing loop interaction in the *add* Asw (see section 3.3).

The finding that ^{112}Asw but not ^{127}Asw exhibited an adenine-induced melting of the P4 helix suggests that only ^{112}Asw can modulate the accessibility of the Shine Dalgarno sequence as a function of the adenine concentration. At first glance, this seems puzzling. The *add* Asw is considered a thermodynamically-controlled riboswitch that regulates ribosome binding to the fully transcribed mRNA (riboswitch + protein coding sequence) via an adenine-dependent equilibrium between states with a liberated (P4-open) and a masked (P4-closed) Shine Dalgarno sequence (see section 1.5). Accordingly, its regulatory properties must not change with an elongated mRNA sequence. It was recently proposed that stand-by binding of the 30S ribosomal subunit to single-stranded regions upstream or downstream of a masked Shine Dalgarno sequence is a general feature of translation initiation for mRNAs with helical secondary structure at the translation initiation site (see section 1.2) (De Smit and Van Duin, 2003). An initial stand-by binding event of the 30S ribosome would occur 3' to the closed P4 helix of the *add* Asw and could eliminate stabilizing effects of the protein coding mRNA sequence on the P4 base

pairs. It is tempting to speculate that the adenine-induced switch of the *add* Asw might thus modulate the relocation of 30S ribosomes from a stand-by site to the Shine Dalgarno sequence.

Chapter 5: Characterization of the Asw encounter complex with hypoxanthine

5.1. Introduction

Purine riboswitches are known to bind their respective ligand with high specificity over chemically related metabolites. The *pbuE* adenine riboswitch aptamer binds its cognate ligand adenine (K_D = 0.3 µM) with at least 30-fold higher affinity over guanine (K_D >10 µM) and 1000-fold higher affinity over hypoxanthine and xanthine (K_D > 300 µM), as was determined by in-line probing assays (Mandal and Breaker, 2004). Yet, NMR data suggest that the *pbuE* adenine riboswitch weakly binds hypoxanthine and that the *xpt* guanine riboswitch weakly binds adenine (Buck *et al.*, 2007). This weak binding of near-cognate ligands by purine riboswitch aptamers has been interpreted as encounter complex formation (Buck *et al.*, 2007). An encounter complex describes an initial docking event between a ligand and the riboswitch binding pocket. Encounter complex formation between a riboswitch aptamer and its cognate ligand triggers induced fit folding of the binding pocket such that the cognate ligand becomes strongly bound by the aptamer. Encounter complex formation between an aptamer and a near-cognate ligand is unproductive. It does not induce a specific folding transition of the riboswitch RNA and the weakly bound ligand dissociates.

One goal of biophysical studies of purine riboswitch aptamer ligand encounter complexes is to shed light on the initial ligand binding site and the molecular mechanism of riboswitch ligand recognition. Specifically, it remains unknown whether initial ligand binding occurs at the junction $J_{2\text{-}3}$ and initiates folding of the junction $J_{3\text{-}1}$, or vice versa (Porter, Marcano-Velázquez and Batey, 2014). A second goal is to investigate whether the weak binding of near-cognate ligands to purine riboswitches could contribute to riboswitch regulation under physiological conditions. The measured concentrations of purines in *E.coli* are 1-100 µM for adenine (Ishii *et al.*, 2007; Bennett *et al.*, 2009), 3-200 µM for guanine (Ishii *et al.*, 2007) and 277 µM for hypoxanthine (Ishii *et al.*, 2007). It is thus possible that near-cognate purine ligands could be present in large excess (~100) over the cognate ligand and affect the regulatory response of a purine riboswitch.

In section 5.2.1, the existence of a site-specific encounter complex of low mM affinity between hypoxanthine and the adenine binding pocket of the full-length *add* Asw was unambiguously shown using WaterLOGSY NMR. Further, the affinity of the *add* Asw hypoxanthine encounter complex was determined by WaterLOGSY NMR. Section 5.2.2

summarizes results that were obtained in the Master thesis of Albrecht Völklein that was co-supervised together with Henrik Gustmann and Prof. Wachtveitl. This Master thesis tested the feasibility of FTIR difference spectroscopy for the analysis of purine riboswitch encounter complexes. Since near-cognate ligand binding to purine riboswitches is fast on the NMR timescale but slow on the IR timescale, it was reasoned that encounter complexes might be directly observed through FTIR. FTIR difference spectra of ^{13}C,^{15}N-labelled hypoxanthine and unlabelled hypoxanthine in complex with the *pbuE* adenine aptamer and the *xpt* guanine aptamer are presented. The data suggest that binding of hypoxanthine in a weak riboswitch encounter complex and in a strong riboswitch ligand complex both has a pronounced effect on the IR bands of hypoxanthine. However, the FTIR difference spectra of the RNA-hypoxanthine complexes with the different isotope labelling of the ligand could not be interpreted due to spectral overlap. In section 5.3, the implications of these results are discussed briefly.

5.2. Results

5.2.1. Hypoxanthine binding of ^{112}Asw analysed by WaterLOGSY NMR

The site specificity and the strength of the hypoxanthine binding interaction of ^{112}Asw were characterized using the WaterLOGSY method. WaterLOGSY is a one-dimensional ^{1}H-NMR experiment that can be used to screen samples of ligands in large excess over a macromolecule for weak binders (Dalvit *et al.*, 2000, 2001). It employs selective water excitation followed by NOE mixing to distinguish weak binders from non-binders and strong binders. Weak binders exchange rapidly between a hydrated free and a hydrated bound state, such that their NMR signal is affected by the positive water-ligand NOE of the bound state, while non-binders and strong binders are almost exclusively in the hydrated free state, which has a negative water-ligand NOE.

Competition binding WaterLOGSY experiments were conducted to investigate whether low-affinity hypoxanthine binding of the *add* Asw occurs specifically at the adenine binding site. WaterLOGSY spectra of hypoxanthine were acquired in buffer, with ^{112}Asw and with both ^{112}Asw and 2,6-diaminopurine. 2,6-diaminopurine is an adenine analogue (Figure 40A), which binds to the adenine binding site of the *add* Asw with ~30-fold higher affinity than adenine (Mandal and Breaker, 2004).

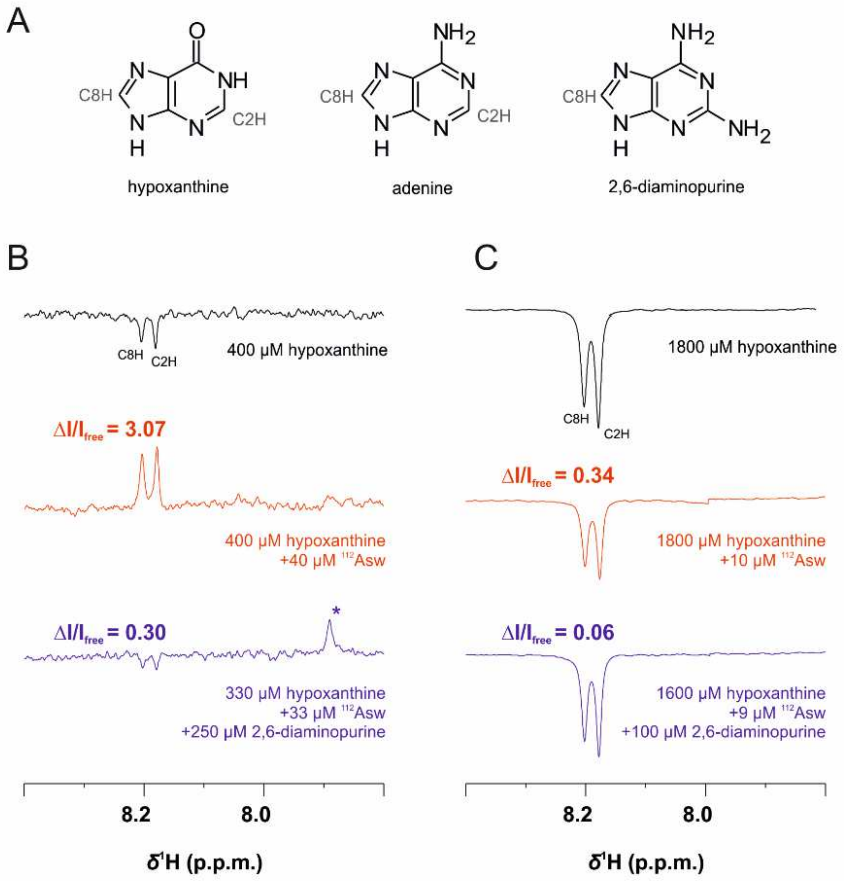

Figure 40: Competition binding WaterLOGSY experiments of ^{112}Asw with hypoxanthine and 2,6-diaminopurine. (A) Structural formulas of the purine ligands hypoxanthine, adenine and 2,6-diaminopurine with the aromatic CH groups indicated. **(B)** One-dimensional WaterLOGSY spectra of free hypoxanthine (top), hypoxanthine with ^{112}Asw (middle) and hypoxanthine with both ^{112}Asw and 2,6-diaminopurine (bottom). For the spectra with ^{112}Asw, the relative hypoxanthine signal intensity difference to the signal intensity of free hypoxanthine $\Delta I/I_{free}$ is indicated. The signal denoted * stems from the C8H proton of 2,6-diaminopurine. All spectra were recorded with an NOE mixing time of 2 s at 298 K in presence of 5 mM Mg^{2+}. The parameters of the CPMG pulse train for suppression of the RNA resonances were n = 14 and τ = 1 ms. **(C)** One-dimensional WaterLOGSY spectra of hypoxanthine as in (B) but at different RNA/ligand concentrations. All spectra were recorded with an NOE mixing time of 2 s at 298 K in presence of 2 mM Mg^{2+}. The parameters of the CPMG pulse train for suppression of the RNA resonances were n = 70 and τ = 0.2 ms.

Here, 2,6-diaminopurine was used as a replacement for adenine, because the ^{1}H-NMR resonances of adenine overlap with the hypoxanthine resonances. A first series of WaterLOGSY experiments was performed using 400 µM hypoxanthine and 40 µM RNA (Figure 40B). The WaterLOGSY NMR signal of hypoxanthine clearly changed sign with the addition of 0.1 eq ^{112}Asw (Figure 4036B, black and red spectrum). This effect was reversed by the addition of ~8 eq 2,6-diaminopurine over the RNA (Figure 40B, purple spectrum), thus proving that hypoxanthine and 2,6-diaminopurine competed for the same binding site.

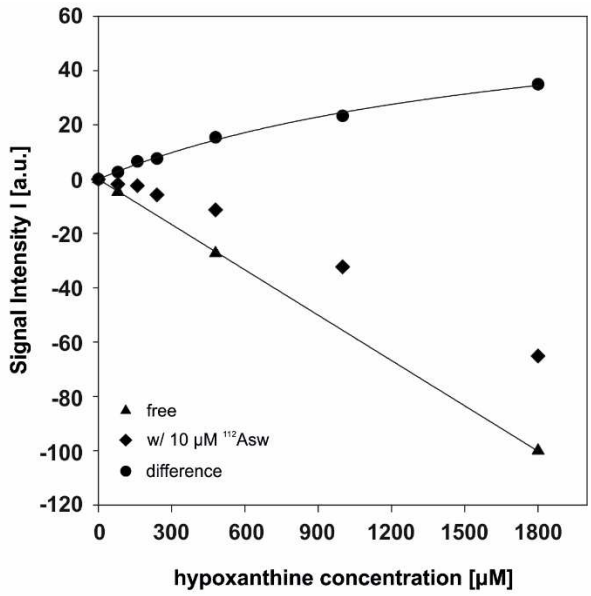

Figure 41: Determination of the hypoxanthine binding affinity of ^{112}Asw with the WaterLOGSY method. WaterLOGSY signal intensity of hypoxanthine as a function of the hypoxanthine concentration measured without ^{112}Asw (triangles) and with 10 µM ^{112}Asw (diamonds). The signal intensities without ^{112}Asw were fitted linearly (slope = 0.0556 ± 0.0002) to calculate the signal intensity difference of hypoxanthine without and with RNA (circles). The signal intensity difference was fitted using the equation $I = -I_{max}/(1+x/K_D)+I_{max}$ (Dalvit *et al.*, 2001) which yielded $I_{max} = 70 \pm 7$ and $K_D = 1800 \pm 300$ µM. The WaterLOGSY data were recorded with an NOE mixing time of 2 s at 298 K in presence of 2 mM Mg^{2+}. The parameters of the CPMG pulse train for suppression of the RNA resonances were n = 70 and $\tau = 0.2$ ms.

However, the hypoxanthine signal intensity in the spectrum of hypoxanthine with ^{112}Asw and 2,6-diaminopurine was still 30% reduced (ΔI/Ifree) compared to the spectrum without RNA (Figure 40B, purple and black spectrum). Additionally, a positive signal of 2,6-diaminopurine was detected, although the signal of 2,6-diaminopurine should be negative, since 2,6-diaminopurine is a strong binder. These observations indicated that both purine ligands bound also unspecifically to ^{112}Asw at the employed ligand and RNA concentrations. It was then tested whether a lower RNA concentration (10 µM) would allow monitor hypoxanthine binding to ^{112}Asw up to mM ligand concentrations without significant contribution of unspecific binding (Figure 40C). A positive NOE contribution to the hypoxanthine signal was observed for 1800 µM hypoxanthine in presence of ^{112}Asw in form of a ~34% reduction of the signal intensity (ΔI/Ifree) compared to the spectrum without ^{112}Asw (Figure 40C, red and black spectrum). At this RNA concentration, the signal intensity of 1600 µM hypoxanthine with ^{112}Asw and ~10 eq 2,6-diaminopurine over the RNA did only marginally deviate from the intensity of free hypoxanthine (~6%), and no positive 2,6-diaminopurine signal was detected. These

conditions were thus adjudged suitable to determine the hypoxanthine affinity of ^{112}Asw from a concentration series of WaterLOGSY spectra.

To determine the K_D of the *add* Asw hypoxanthine encounter complex, the hypoxanthine signal intensity (C2H+C8H) was extracted from WaterLOGSY spectra that were measured in absence and in presence of 10 µM ^{112}Asw at different hypoxanthine concentrations. The intensity difference between the spectra with and without RNA lied on a conventional dose-response curve (Figure 41). From this curve, the ligand dissociation constant was determined as described by Dalvit *et al.* (Dalvit *et al.*, 2001). The hypoxanthine dissociation constant (K_D) of ^{112}Asw extracted from these data was 1800 ± 300 µM. This value matched the highest employed ligand concentration, indicating that more data points would be required for a reliable measure of K_D. However, at hypoxanthine concentrations higher than 1800 µM unspecific binding occurred (data not shown). Therefore, higher ligand concentrations could not be used to improve the accuracy of the hypoxanthine dissociation constant determined from the WaterLOGSY titration. To conclude, the WaterLOGSY method provided an estimate for the K_D of the *add* Asw hypoxanthine encounter complex (1800 µM) that substantially exceeded the physiological hypoxanthine concentration in *E. coli* (277 µM) (Ishii *et al.*, 2007). This suggests that physiologically relevant concentrations of hypoxanthine do not significantly impact the adenine-dependent gene regulation of the *add* Asw by competitive binding.

5.2.2. FTIR of hypoxanthine in purine riboswitch aptamer complexes

In a co-supervised master thesis, it was tested whether FTIR difference spectroscopy between isotope labelled and unlabelled ligand could shed light into the different binding modes of hypoxanthine in a weak riboswitch ligand encounter complex and in a high-affinity riboswitch ligand complex. The idea behind this approach is that FTIR spectra of two identical samples with different isotope labelling of the ligand are subtracted from each other, to yield a difference spectrum that exclusively shows the IR bands of the ligand. In this study, hypoxanthine and ^{13}C,^{15}N-labelled hypoxanthine was used, and riboswitch complexes were prepared using a 66-nucleotide *pbuE* adenine riboswitch aptamer for the encounter complex and an 80-nucleotide *xpt* guanine riboswitch aptamer for the high-affinity complex. The FTIR experiments were performed by Albrecht Völklein and Henrik Gustmann.

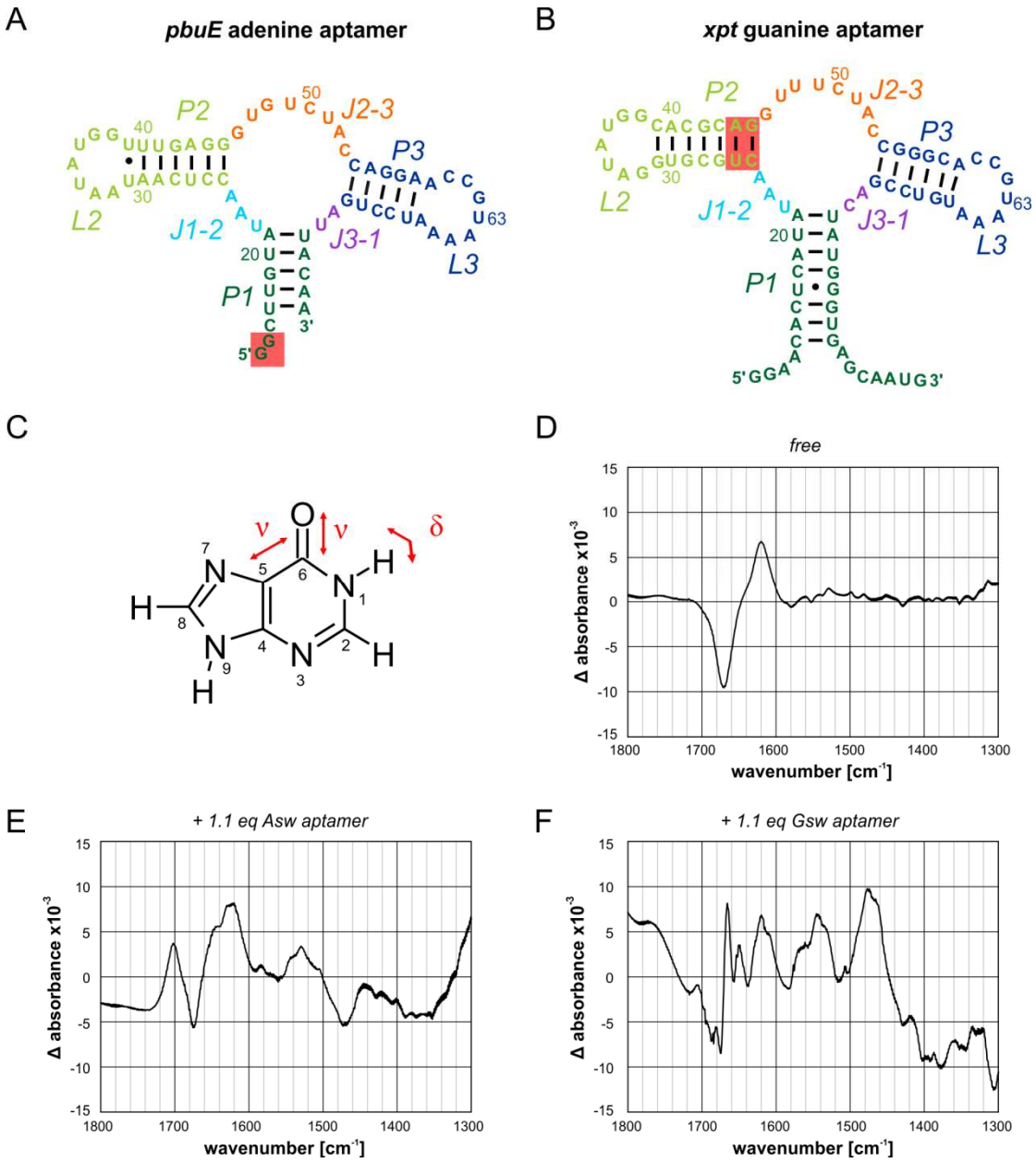

Figure 42: FTIR difference spectra for ^{13}C,^{15}N-labelled hypoxanthine subtracted by hypoxanthine in the free form and in complex with an adenine and a guanine riboswitch aptamer. (A) Secondaary structure of the 66-nucleotide *pbuE* adenine riboswitch aptamer. Modifications to the wildtype sequence from *Bacillus subtilis* are highlighted in red. **(B)** Secondary structure of the 80-nucleotide *xpt* guanine riboswitch aptamer. Modifications to the wildtype sequence from *Bacillus subtilis* are highlighted in red. **(C)** Structural formula of hypoxanthine. The molecular vibrations (ν stretching, δ bending) that contribute to the observed IR band in (D) according to a theoretical study by Fernandez Quejo *et al.* are indicated (Fernández-Quejo, de la Fuente and Navarro, 2005). **(D)** Hypoxanthine FTIR difference spectrum at a concentration of 1.5 mM in buffered D_2O (25 mM potassium phosphate pH 6.2, 50 mM potassium chloride, 16 mM magnesium chloride) at 293 K. **(E)** Hypoxanthine FTIR difference spectrum at the same conditions as in (D) but with 1.1 eq 66-nucleotide *pbuE* adenine riboswitch aptamer. **(F)** Hypoxanthine FTIR difference spectrum at the same conditions as in (D) but with 1.1 eq 80-nucleotide *xpt* guanine riboswitch aptamer. Acquisition and processing of the spectra was performed by Albrecht Völklein.

The FTIR difference spectra of $^{13}C,^{15}N$-labelled hypoxanthine and hypoxanthine that were obtained in the free form, in complex with the adenine aptamer and in complex with the guanine aptamer are shown in Figure 42. The difference spectrum without RNA (Figure 42D) showed a single IR band for $^{13}C,^{15}N$-labelled hypoxanthine (1620 cm^{-1}; positive band) and hypoxanthine (1670 cm^{-1}; negative band). According to theoretical calculations, this IR band corresponds to vibrational modes involving stretching of the C6-O bond and the C6-C5 bond as well as bending of the N1-H bond (Figure 42C) (Fernández-Quejo, de la Fuente and Navarro, 2005). The corresponding band in polyinosinic acid (1678 cm^{-1}) has been shown to undergo a 19 cm^{-1} shift towards higher frequency upon duplex formation with polycytidylic acid (Miles, 1959). This IR band hence is a promising reporter for Watson Crick type base pairing. The FTIR difference spectrum of $^{13}C,^{15}N$-labelled hypoxanthine and hypoxanthine in presence of the adenine aptamer (Figure 42E) was obtained at RNA and ligand concentrations in the range of the hypoxanthine dissociation constant of the related *add* adenine riboswitch (K_D ~1.8 mM; see section 5.2.1). One would thus expect IR bands for hypoxanthine in the free form and in the RNA bound form. Indeed, the FTIR difference spectrum obtained in presence of the adenine aptamer was changed compared to the FTIR difference spectrum without RNA, suggesting that an encounter complex was formed between hypoxanthine and the adenine aptamer. However, it showed no clear pattern of positive and negative peaks. The FTIR difference spectrum of $^{13}C,^{15}N$-labelled hypoxanthine and hypoxanthine in presence of the guanine aptamer (Figure 42F) was even more complex and showed a multitude of overlapping bands. However, at the here employed guanine riboswitch aptamer and ligand concentrations, hypoxanthine should be exclusively in the bound state, because the hypoxanthine dissociation constant of the *xpt* guanine aptamer is in the low micromolar range (Batey, Gilbert and Montange, 2004). Notably, a similarly crowded FTIR difference spectrum was obtained for $^{13}C,^{15}N$-labelled adenine and adenine in the cognate complex with the adenine aptamer (data not shown). It might thus be a general feature of purine riboswitches, that formation of the cognate aptamer ligand complex induces pronounced vibrational changes of the purine ligand. To conclude, since the IR bands of $^{13}C,^{15}N$-labelled hypoxanthine and hypoxanthine could not be unambiguously deduced from the FTIR difference spectra obtained in presence of the adenine aptamer and the guanine aptamer, it was not possible to evaluate these data.

5.3. Discussion

In this chapter, WaterLOGSY NMR spectroscopy and FTIR difference spectroscopy were employed to characterize the encounter complex of hypoxanthine with the adenine riboswitch RNA.

The WaterLOGSY competition binding experiments with 2,6-diaminopurine proved that hypoxanthine is recognized by the binding pocket of the full-length 112-nucleotide *add* Asw. By means of WaterLOGSY titrations, an estimate of the hypoxanthine binding constant of the *add* Asw was obtained, which was in the low millimolar range (K_D ~1.8 mM) and ~5-fold higher than the physiological hypoxanthine concentration in *E. coli* (~0.3mM) (Ishii *et al.*, 2007). The purine nucleobases that are involved in the adenosine salvage pathway, which is regulated by the *add* adenine riboswitch in *Vibrio vulnificus*, are adenine, hypoxanthine and xanthine. Of these three metabolites, adenine and hypoxanthine are structurally most closely related. The fact that the here determined hypoxanthine binding constant of the *add* Asw clearly exceeds the bacterial hypoxanthine concentration provides increasing evidence that the *add* adenine riboswitch responds only to a single metabolite of the adenosine salvage pathway.

The FTIR difference spectrum of $^{13}C,^{15}N$-labelled hypoxanthine and hypoxanthine showed changes upon formation of an encounter complex with the *pbuE* adenine riboswitch aptamer and upon formation of a high-affinity complex with the *xpt* guanine riboswitch aptamer. However, these changes could not be evaluated, since no clear pattern of positive and negative IR bands was observed in the FTIR difference spectra with the aptamers. Still, it was striking that the complexity of the FTIR difference spectrum of hypoxanthine increased from the free ligand to the encounter complex and from the encounter complex to the stable riboswitch ligand complex. One might follow up on this preliminary data and perform titrations with varying RNA and Mg^{2+} concentrations to gradually monitor the formation of these purine riboswitch ligand complexes and unravel the accompanying changes of the FTIR difference spectrum. However, this would not help resolve the complex spectral pattern of the FTIR difference spectrum obtained for the holo state of the *xpt* guanine riboswitch with hypoxanthine at high Mg^{2+} concentration (16 mM). Since the *xpt* guanine riboswitch aptamer did not show global tertiary structure heterogeneity in smFRET experiments at Mg^{2+} concentrations exceeding 5 mM (Brenner *et al.*, 2010), the multitude of IR bands in the FTIR difference

spectrum of the complex between hypoxanthine and the *xpt* guanine riboswitch likely corresponds to hypoxanthine encapsulated in a homogenously folded aptamer with 'tuning fork' tertiary structure (see section 1.4). Thus, the IR spectra of the purine ligands in the purine riboswitch aptamers might simply be too complex to be accessible by FTIR difference spectroscopy between uniformly $^{13}C,^{15}N$-labelled and unlabelled ligands. To address this issues, future FTIR difference spectroscopic studies could be performed with position specific $^{13}C,^{15}N$-labelled nucleobases (Abad, Gaffney and Jones, 1999; Shallop, Gaffney and Jones, 2003).

Chapter 6: Concluding remarks

NMR, smFRET and FTIR as tools for biophysical studies of riboswitches

In this PhD thesis, three different biophysical techniques were employed to investigate the molecular mechanism of purine binding riboswitches: Nuclear magnetic resonance (NMR), single-molecule FRET (smFRET) and Fourier transform infrared (FTIR) spectroscopy.

NMR and smFRET were used to study the adenine-induced conformational switch of the full-length translation-regulating *add* adenine riboswitch from *Vibrio vulnificus* (Asw). NMR is a well-established and powerful method for secondary structure elucidation of RNA. In case of Asw, the secondary structure bistability in the apo state, which is at the heart of the biological riboswitch thermostat function of Asw, was only uncovered by NMR (Reining *et al.*, 2013) and remained hidden in structural probing data (Lemay *et al.*, 2011) and in single-molecule force experiments (Neupane *et al.*, 2011). smFRET has proven to be a valuable tool for the investigation of the ligand-dependent tertiary structure of purine riboswitch aptamer domains (Brenner *et al.*, 2010; Dalgarno *et al.*, 2013). In this thesis, NMR and smFRET experiments were for the first time performed on the same purine riboswitch sequence such that the NMR and the smFRET data could be directly compared. The main methodological conclusion of the combined NMR and smFRET study of Asw is that routinely employed NMR and smFRET experiments as stand-alone biophysical techniques could only capture a subset of the long-lived conformational states of Asw. By NMR of the base pair reporter imino groups, folding states with undocked aptamer kissing loop motif remained hidden, although they are long-lived (lifetime ~s) and stable (fractional population ~50%) as revealed by smFRET. By smFRET in three different labelling schemes (L2/L3, L2/P5, L3/P5), one could not obtain a specific FRET signature for the apoB conformation, which is long-lived (lifetime ~s) and stable (fractional population ~50%) as demonstrated by NMR. Only the combination of NMR and smFRET allowed deducing a comprehensive model of the multi-state conformational equilibrium of Asw that accounts for both the secondary structure heterogeneity of the apo state and the aptamer tertiary structure heterogeneity of the holo state. Since increasing evidence accumulates that full-length riboswitches can occupy multiple-long-lived conformational states (Haller *et al.*, 2011; Liberman and Wedekind, 2012; Reining *et al.*, 2013; Helmling *et al.*, 2017), the integrated NMR and smFRET analysis of

riboswitches represents a powerful means to uncover the complex conformational space of these intriguing regulatory RNA elements at the level of both secondary and tertiary structure.

FTIR spectroscopy is an attractive method to investigate fast dynamics that average out on the timescale of NMR and smFRET. In case of purine riboswitches, FTIR could in principle be used to compare the binding mode of near-cognate ligands in transient encounter complexes to the binding mode of cognate ligands in a stable complex. Here, FTIR difference spectroscopy between $^{13}C,^{15}N$-labelled and unlabelled hypoxanthine was employed to observe exclusively the IR bands of hypoxanthine in a stable complex with a guanine riboswitch aptamer and in a transient encounter complex with an adenine riboswitch aptamer. Specific FTIR difference spectra could be obtained as fingerprints for the stable ligand complex and the transient ligand encounter complex. However, in contrast to the FTIR difference spectrum of free hypoxanthine in solution, the spectra of hypoxanthine in complex with the guanine and the adenine riboswitch aptamer could not be analysed since they showed multiple overlapping bands. Unexpectedly, the FTIR difference spectrum for hypoxanthine in complex with the guanine riboswitch aptamer was most crowded, although it corresponded to an aptamer complex with homogenous RNA tertiary structure. Since riboswitches usually undergo large conformational re-arrangements in response to concentration changes of ligand and Mg^{2+}, other FTIR difference spectroscopy techniques like concentration jump difference experiments can be expected to result in even more complex spectra. The issue of overlapping bands is the major limitation of marker-free infrared spectroscopy of biomolecules (Barth, 2007). The fact that the specific IR bands of hypoxanthine could not be deciphered for the complex of hypoxanthine with the *xpt* guanine riboswitch aptamer by FTIR difference spectroscopy on the basis of readily available isotope-labelling schemes of the ligand, suggests that FTIR difference spectroscopy has no potential to become a broadly applicable technique in the field of riboswitches.

Significance of the new insight into the functional dynamics of Asw

The here performed studies of Asw revealed two mechanistic features that are of general significance for full-length purine riboswitches.

First, the multivector smFRET investigation of Asw with dye labels across aptamer and expression platform demonstrated that Asw operates via a spatially decoupled allosteric switch that involves no long-range tertiary structural interaction between the ligand-bound aptamer domain and the expression platform. Hence, the allosteric switch is a ligand-dependent stabilization of the 3-way junction aptamer secondary structure towards alternative secondary structures with a helix formed between complementary sequence elements of the aptamer and the expression platform. This mechanistic principle of a spatially decoupled allosteric switch was hypothesized to be the general switching principle of purine riboswitches and could now be corroborated experimentally.

Second, the smFRET study of the full-length ^{112}Asw in the aptamer loop labelling scheme uncovered tertiary structure heterogeneity of the holo conformation between substates with undocked and docked aptamer kissing loop motif at near physiological Mg^{2+} concentration, while the isolated P1-stabilized aptamer domain of this riboswitch is homogenously folded and has a docked aptamer kissing loop motif at near physiological Mg^{2+} concentration (Dalgarno *et al.*, 2013). This finding demonstrates that the extension of purine riboswitch aptamers by their native expression platform does not only expand the secondary structure space of the riboswitch, but also alters the tertiary structure folding landscape of the 3-way junction aptamer domain. In the native sequence of purine riboswitches, tertiary structure folding of the 3-way junction aptamer might be tuned to fulfil specific requirements related to the riboswitch mode of operation (transcriptional vs translational) or to achieve a specific induction factor in gene expression. Notably, the dynamic interplay between folding of the P4 helix, the P1 helix and the aptamer kissing loop motif of the full-length ^{112}Asw changed further upon elongation of ^{112}Asw to ^{127}Asw by 15 nucleotides of the protein coding sequence of the *add* gene. This highlights the necessity to study translation-regulating riboswitches also within their respective protein-coding mRNA transcript.

With increasing available structural information on transcriptional and translational full-length purine riboswitches (Helmling *et al.*, 2017; Warhaut *et al.*, 2017) the new frontier in the mechanistic study of purine riboswitches is shifted to the investigation of these riboswitches in the context of the macromolecular machinery of transcription and translation. Given the multitude of conformational states observed for the full-length purine riboswitches, the major challenge of these studies is to figure out which specific

conformational transition is central to the regulation of transcription termination or translation initiation.

References

Abad, J. L., Gaffney, B. L. and Jones, R. A. (1999) '15N-Multilabeled adenine and guanine nucleosides. Syntheses of [1,3,NH2-15N3]- and [2-13C-1,3,NH2-15N3]-labeled adenosine, guanosine, 2'-deoxyadenosine, and 2'-deoxyguanosine', *Journal of Organic Chemistry*, 64(18), pp. 6575–6582. doi: 10.1021/jo982372k.

Al-Hashimi, H. M. and Walter, N. G. (2008) 'RNA dynamics: it is about time.', *Current opinion in structural biology*, 18(3), pp. 321–9. doi: 10.1016/j.sbi.2008.04.004.

Bakshi, S., Choi, H. and Weisshaar, J. C. (2015) 'The spatial biology of transcription and translation in rapidly growing Escherichia coli', *Frontiers in Microbiology*, 6, p. 636. doi: 10.3389/fmicb.2015.00636.

Bakshi, S., Siryaporn, A., Goulian, M. and Weisshaar, J. C. (2012) 'Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells', *Molecular Microbiology*, 85(1), pp. 21–38. doi: 10.1111/j.1365-2958.2012.08081.x.

Barrick, J. E. and Breaker, R. R. (2007) 'The distributions, mechanisms, and structures of metabolite-binding riboswitches.', *Genome biology*, 8(11), p. R239. doi: 10.1186/gb-2007-8-11-r239.

Barth, A. (2007) 'Infrared spectroscopy of proteins', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1767(9), pp. 1073–1101. doi: 10.1016/j.bbabio.2007.06.004.

Batey, R. T., Gilbert, S. D. and Montange, R. K. (2004) 'Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine.', *Nature*, 432(7015), pp. 411–5. doi: 10.1038/nature03037.

Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J. and Rabinowitz, J. D. (2009) 'Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli', *Nature Chemical Biology*, 5(8), pp. 593–599. doi: 10.1038/nchembio.186.

Blanco, M. (2010) 'Analysis of Complex Single-Molecule FRET Time Trajectories', *Methods in Enzymology*, 472, pp. 153–178. doi: 10.1016/S0076-6879(10)72011-5.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. and Shao, Y. (1997) 'The complete genome sequence of Escherichia coli K-12.', *Science*, 277(5331), pp. 1453–62. doi: 10.1126/science.277.5331.1453.

Blount, K. F., Wang, J. X., Lim, J., Sudarsan, N. and Breaker, R. R. (2007) 'Antibacterial lysine analogs that target lysine riboswitches', *Nature Chemical Biology*, 3(1), pp. 44–49. doi: 10.1038/nchembio842.

Bokinsky, G., Rueda, D., Misra, V. K., Rhodes, M. M., Gordus, A., Babcock, H. P., Walter, N. G. and Zhuang, X. (2003) 'Single-molecule transition-state analysis of RNA folding.', *Proceedings of the National Academy of Sciences of the United States of America*, 100(16), pp. 9302–7. doi: 10.1073/pnas.1133280100.

Brenner, M. D., Scanlan, M. S., Nahas, M. K., Ha, T. and Silverman, S. K. (2010) 'Multivector Fluorescence Analysis of the xpt Guanine Riboswitch Aptamer Domain and

the Conformational Role of Guanine', *Biochemistry*, 49(8), pp. 1596–1605. doi: 10.1021/bi9019912.

Brion, P. and Westhof, E. (1997) 'HIERARCHY AND DYNAMICS OF RNA FOLDING', *Annual Review of Biophysics and Biomolecular Structure*, 26(1), pp. 113–137. doi: 10.1146/annurev.biophys.26.1.113.

Buck, J., Fürtig, B., Noeske, J., Wöhnert, J. and Schwalbe, H. (2007) 'Time-resolved NMR methods resolving ligand-induced RNA folding at atomic resolution.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(40), pp. 15699–704. doi: 10.1073/pnas.0703182104.

Burkard, M., Turner, D. and Tinoco, I. J. (1999) 'The Interactions That Shape RNA Structure.', in *Cold Spring Harbor Monograph Archive, North America, 37*, pp. 233–264.

Chappell, J., Watters, K. E., Takahashi, M. K. and Lucks, J. B. (2015) 'A renaissance in RNA synthetic biology: New mechanisms, applications and tools for the future', *Current Opinion in Chemical Biology*, pp. 47–56. doi: 10.1016/j.cbpa.2015.05.018.

Church, G. M., Selinger, D. W., Cheung, K. J., Mei, R., Johansson, E. M., Richmond, C. S., Blattner, F. R. and Lockhart, D. J. (2000) 'RNA expression analysis using a 30 base pair resolution Escherichia coli genome array.', *Nature Biotechnology*, 18(12), pp. 1262–1268. doi: 10.1038/82367.

Cragnolini, T., Derreumaux, P. and Pasquali, S. (2015) 'Ab initio RNA folding.', *Journal of physics. Condensed matter: an Institute of Physics journal*, 27(23), p. 233102. doi: 10.1088/0953-8984/27/23/233102.

Dalgarno, P. A., Bordello, J., Morris, R., St-Pierre, P., Dube, A., Samuel, I. D. W., Lafontaine, D. A. and Penedo, J. C. (2013) 'Single-molecule chemical denaturation of riboswitches', *Nucleic Acids Research*, 41(7), pp. 4253–4265. doi: 10.1093/nar/gkt128.

Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M. and Stockman, B. (2001) 'WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability.', *Journal of biomolecular NMR*, 21(4), pp. 349–59. doi: 10.1023/A:1013302231549.

Dalvit, C., Pevarello, P., Tatò, M., Veronesi, M., Vulpetti, A. and Sundström, M. (2000) 'Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water.', *Journal of biomolecular NMR*, 18(1), pp. 65–8. doi: 10.1023/A:1008354229396.

Dann, C. E., Wakeman, C. A., Sieling, C. L., Baker, S. C., Irnov, I. and Winkler, W. C. (2007) 'Structure and Mechanism of a Metal-Sensing Regulatory RNA', *Cell*, 130(5), pp. 878–892. doi: 10.1016/j.cell.2007.06.051.

Delfosse, V., Bouchard, P., Bonneau, E., Dagenais, P., Lemay, J. F., Lafontaine, D. A. and Legault, P. (2010) 'Riboswitch structure: an internal residue mimicking the purine ligand', *Nucleic Acids Research*, 38(6), pp. 2057–2068. doi: 10.1093/nar/gkp1080.

Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. and Al-Hashimi, H. M. (2012) 'Visualizing transient low-populated structures of RNA', *Nature*, 491(7426), p. 724. doi: 10.1038/nature11498.

Dixon, N., Robinson, C. J., Geerlings, T., Duncan, J. N., Drummond, S. P. and Micklefield, J. (2012) 'Orthogonal Riboswitches for Tuneable Coexpression in Bacteria',

Angewandte Chemie International Edition, 51(15), pp. 3620–3624. doi: 10.1002/anie.201109106.

Duval, M., Simonetti, A., Caldelari, I. and Marzi, S. (2015) 'Multiple ways to regulate translation initiation in bacteria: Mechanisms, regulatory circuits, dynamics', *Biochimie*, 114, pp. 18–29. doi: 10.1016/j.biochi.2015.03.007.

Farwell, M. A., Roberts, M. W. and Rabinowitz, J. C. (1992) 'The effect of ribosomal protein S1 from Escherichia coli and Micrococcus luteus on protein synthesis in vitro by E. coli and Bacillus subtilis.', *Molecular microbiology*, 6(22), pp. 3375–83. doi: 10.1111/j.1365-2958.1992.tb02205.x.

Favier, A. and Brutscher, B. (2011) 'Recovering lost magnetization: polarization enhancement in biomolecular NMR.', *Journal of biomolecular NMR*, 49(1), pp. 9–15. doi: 10.1007/s10858-010-9461-5.

Fernández-Quejo, M., de la Fuente, M. and Navarro, R. (2005) 'Theoretical calculations and vibrational study of hypoxanthine in aqueous solution', *Journal of Molecular Structure*, 744-747, pp. 749–757. doi: 10.1016/j.molstruc.2004.10.092.

Frieda, K. L. and Block, S. M. (2012) 'Direct observation of cotranscriptional folding in an adenine riboswitch.', *Science*, 338(6105), pp. 397–400. doi: 10.1126/science.1225722.

Fürtig, B., Nozinovic, S., Reining, A. and Schwalbe, H. (2015) 'Multiple conformational states of riboswitches fine-tune gene regulation', *Current Opinion in Structural Biology*, 30, pp. 112–124. doi: 10.1016/j.sbi.2015.02.007.

Fürtig, B., Richter, C., Wöhnert, J. and Schwalbe, H. (2003) 'NMR Spectroscopy of RNA', *ChemBioChem*, 4(10), pp. 936–962. doi: 10.1002/cbic.200300700.

Fürtig, B., Schnieders, R., Richter, C., Zetzsche, H., Keyhani, S., Helmling, C., Kovacs, H. and Schwalbe, H. (2016) 'Direct 13C-detected NMR experiments for mapping and characterization of hydrogen bonds in RNA', *Journal of Biomolecular NMR*, 64(3), pp. 207–221. doi: 10.1007/s10858-016-0021-5.

Fürtig, B., Wenter, P., Reymond, L., Richter, C., Pitsch, S. and Schwalbe, H. (2007) 'Conformational dynamics of bistable RNAs studied by time-resolved NMR spectroscopy', *Journal of the American Chemical Society*, 129(51), pp. 16222–16229. doi: 10.1021/ja076739r.

Gao, A. and Serganov, A. (2014) 'Structural insights into recognition of c-di-AMP by the ydaO riboswitch', *Nature Chemical Biology*, 10(9), pp. 787–792. doi: 10.1038/nchembio.1607.

Gilbert, S. D., Reyes, F. E., Edwards, A. L. and Batey, R. T. (2009) 'Adaptive Ligand Binding by the Purine Riboswitch in the Recognition of Guanine and Adenine Analogs', *Structure*, 17(6), pp. 857–868. doi: 10.1016/j.str.2009.04.009.

Goddard, T. D. and Kneller, D. G. (2008) 'Sparky 3', *University of California, San Francisco*.

Gong, S., Wang, Y., Wang, Z., Wang, Y. and Zhang, W. (2016) 'Reversible-Switch Mechanism of the SAM-III Riboswitch.', *The journal of physical chemistry. B*, 120(48), pp. 12305–12311. doi: 10.1021/acs.jpcb.6b09698.

Guo, F., Gooding, A. R. and Cech, T. R. (2004) 'Structure of the Tetrahymena ribozyme:

base triple sandwich and metal ion at the active site.', *Molecular cell*, 16(3), pp. 351–62. doi: 10.1016/j.molcel.2004.10.003.

Hald, M. and Jacobsen, J. P. (1991) 'Sodium-23 and lithium-7 NMR spin-lattice relaxation measurements in the study intercalation in DNA.', *Biophysical chemistry*, 41(2), pp. 113–24. doi: 10.1016/0301-4622(91)80011-F.

Haller, A., Rieder, U., Aigner, M., Blanchard, S. C. and Micura, R. (2011) 'Conformational capture of the SAM-II riboswitch', *Nature Chemical Biology*, 7(6), pp. 393–400. doi: 10.1038/nchembio.562.

Helmling, C., Keyhani, S., Sochor, F., Fürtig, B., Hengesbach, M. and Schwalbe, H. (2015) 'Rapid NMR screening of RNA secondary structure and binding', *Journal of Biomolecular NMR*, 63(1), pp. 67–76. doi: 10.1007/s10858-015-9967-y.

Helmling, C., Wacker, A., Wolfinger, M. T., Hofacker, I. L., Hengesbach, M., Fuertig, B. and Schwalbe, H. (2017) 'NMR structural profiling of transcriptional intermediates reveals riboswitch regulation by metastable RNA conformations', *Journal of the American Chemical Society*, 139(7), p. jacs.6b10429. doi: 10.1021/jacs.6b10429.

Hendrix, D. K., Brenner, S. E. and Holbrook, S. R. (2005) 'RNA structural motifs: building blocks of a modular biomolecule', *Quarterly Reviews of Biophysics*, 38(03), p. 221. doi: 10.1017/S0033583506004215.

Holmstrom, E. D., Polaski, J. T., Batey, R. T. and Nesbitt, D. J. (2014) 'Single-molecule conformational dynamics of a biologically functional hydroxocobalamin riboswitch.', *Journal of the American Chemical Society*, 136(48), pp. 16832–43. doi: 10.1021/ja5076184.

Howe, J. A., Wang, H., Fischmann, T. O., Balibar, C. J., Xiao, L., Galgoci, A. M., Malinverni, J. C., Mayhood, T., Villafania, A., Nahvi, A., Murgolo, N., Barbieri, C. M., Mann, P. A., Carr, D., Xia, E., Zuck, P., Riley, D., Painter, R. E., Walker, S. S., Sherborne, B., de Jesus, R., Pan, W., Plotkin, M. A., Wu, J., Rindgen, D., Cummings, J., Garlisi, C. G., Zhang, R., Sheth, P. R., Gill, C. J., Tang, H. and Roemer, T. (2015) 'Selective small-molecule inhibition of an RNA structural element', *Nature*, 526(7575), pp. 672–677. doi: 10.1038/nature15542.

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., Hoque, A., Ho, P. Y., Kakazu, Y., Sugawara, K., Igarashi, S., Harada, S., Masuda, T., Sugiyama, N., Togashi, T., Hasegawa, M., Takai, Y., Yugi, K., Arakawa, K., Iwata, N., Toya, Y., Nakayama, Y., Nishioka, T., Shimizu, K., Mori, H. and Tomita, M. (2007) 'Multiple High-Throughput Analyses Monitor the Response of E. coli to Perturbations', *Science*, 316(5824), pp. 593–597. doi: 10.1126/science.1132067.

Kaminishi, T., Wilson, D. N., Takemoto, C., Harms, J. M., Kawazoe, M., Schluenzen, F., Hanawa-Suetsugu, K., Shirouzu, M., Fucini, P. and Yokoyama, S. (2007) 'A Snapshot of the 30S Ribosomal Subunit Capturing mRNA via the Shine-Dalgarno Interaction', *Structure*, 15(3), pp. 289–297. doi: 10.1016/j.str.2006.12.008.

Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E. and Weiss, S. (2004) 'Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(24), pp. 8936–41. doi: 10.1073/pnas.0401690101.

Kim, J. N., Blount, K. F., Puskarz, I., Lim, J., Link, K. H. and Breaker, R. R. (2009) 'Design and Antimicrobial Action of Purine Analogues That Bind Guanine Riboswitches', *ACS Chemical Biology*, 4(11), pp. 915–927. doi: 10.1021/cb900146k.

Kim, J. N., Roth, A. and Breaker, R. R. (2007) 'Guanine riboswitch variants from Mesoplasma florum selectively recognize 2'-deoxyguanosine.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(41), pp. 16092–7. doi: 10.1073/pnas.0705884104.

Kirmizialtin, S., Pabit, S. A., Meisburger, S. P., Pollack, L. and Elber, R. (2012) 'RNA and its ionic cloud: solution scattering experiments and atomically detailed simulations.', *Biophysical journal*, 102(4), pp. 819–28. doi: 10.1016/j.bpj.2012.01.013.

Klein, D. J., Been, M. D. and Ferré-D'Amaré, A. R. (2007) 'Essential role of an active-site guanine in glmS ribozyme catalysis', *Journal of the American Chemical Society*, 129(48), pp. 14858–14859. doi: 10.1021/ja0768441.

Kobitski, A. Y., Nierth, A., Helm, M., Jaschke, A. and Nienhaus, G. U. (2007) 'Mg2+-dependent folding of a Diels-Alderase ribozyme probed by single-molecule FRET analysis', *Nucleic Acids Research*, 35(6), pp. 2047–2059. doi: 10.1093/nar/gkm072.

Kortmann, J. and Narberhaus, F. (2012) 'Bacterial RNA thermometers: molecular zippers and switches.', *Nature reviews. Microbiology*, 10(4), pp. 255–65. doi: 10.1038/nrmicro2730.

Lambert, D., Leipply, D., Shiman, R. and Draper, D. E. (2009) 'The Influence of Monovalent Cation Size on the Stability of RNA Tertiary Structures', *Journal of Molecular Biology*, 390(4), pp. 791–804. doi: 10.1016/j.jmb.2009.04.083.

Laursen, B. S., Sorensen, H. P., Mortensen, K. K. and Sperling-Petersen, H. U. (2005) 'Initiation of Protein Synthesis in Bacteria', *Microbiology and Molecular Biology Reviews*, 69(1), pp. 101–123. doi: 10.1128/MMBR.69.1.101-123.2005.

Lee, M.-K., Gal, M., Frydman, L. and Varani, G. (2010) 'Real-time multidimensional NMR follows RNA folding with second resolution.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(20), pp. 9192–7. doi: 10.1073/pnas.1001195107.

Lemay, J. F., Desnoyers, G., Blouin, S., Heppell, B., Bastet, L., St-Pierre, P., Massé, E. and Lafontaine, D. A. (2011) 'Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms', *PLoS Genetics*, 7(1), p. e1001278. doi: 10.1371/journal.pgen.1001278.

Lemay, J.-F., Penedo, J. C., Tremblay, R., Lilley, D. M. J. and Lafontaine, D. A. (2006) 'Folding of the Adenine Riboswitch', *Chemistry & Biology*, 13(8), pp. 857–868. doi: 10.1016/j.chembiol.2006.06.010.

Leontis, N. B. and Westhof, E. (1998) 'Conserved geometrical base-pairing patterns in RNA.', *Quarterly reviews of biophysics*, 31(4), pp. 399–455. doi: 10.1017/S0033583599003479.

Liberman, J. A. and Wedekind, J. E. (2012) 'Riboswitch structure in the ligand-free state.', *Wiley interdisciplinary reviews. RNA*, 3(3), pp. 369–84. doi: 10.1002/wrna.114.

Liu, Y., Holmstrom, E., Zhang, J., Yu, P., Wang, J., Dyba, M. A., Chen, D., Ying, J.,

Lockett, S., Nesbitt, D. J., Ferré-D'Amaré, A. R., Sousa, R., Stagno, J. R. and Wang, Y.-X. (2015) 'Synthesis and applications of RNAs with position-selective labelling and mosaic composition.', *Nature*, 522(7556), pp. 368–72. doi: 10.1038/nature14352.

Lorenz, R., Wolfinger, M. T., Tanzer, A. and Hofacker, I. L. (2016) 'Predicting RNA secondary structures from sequence and probing data', *Methods*, 103, pp. 86–98. doi: 10.1016/j.ymeth.2016.04.004.

Majdalani, N., Hernandez, D. and Gottesman, S. (2002) 'Regulation and mode of action of the second small RNA activator of RpoS translation, RprA', *Molecular Microbiology*, 46(3), pp. 813–826. doi: 10.1046/j.1365-2958.2002.03203.x.

Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. and Breaker, R. R. (2003) 'Riboswitches Control Fundamental Biochemical Pathways in Bacillus subtilis and Other Bacteria', *Cell*, 113(5), pp. 577–586. doi: 10.1016/S0092-8674(03)00391-X.

Mandal, M. and Breaker, R. R. (2004) 'Adenine riboswitches and gene activation by disruption of a transcription terminator', *Nature Structural & Molecular Biology*, 11(1), pp. 29–35. doi: 10.1038/nsmb710.

Mandal, M., Lee, M., Barrick, J. E., Weinberg, Z., Emilsson, G. M., Ruzzo, W. L. and Breaker, R. R. (2004) 'A Glycine-Dependent Riboswitch That Uses Cooperative Binding to Control Gene Expression', *Science*, 306(5694), pp. 275–279. doi: 10.1126/science.1100829.

Marshall, R. A., Aitken, C. E. and Puglisi, J. D. (2009) 'GTP Hydrolysis by IF2 Guides Progression of the Ribosome into Elongation', *Molecular Cell*, 35(1), pp. 37–47. doi: 10.1016/j.molcel.2009.06.008.

Mason, S. F. (1954) 'Purine studies. Part II. The ultra-violet absorption spectra of some mono- and poly-substituted purines', *Journal of the Chemical Society (Resumed)*, (0), pp. 2071–2081. doi: 10.1039/jr9540002071.

McKinney, S. A., Joo, C. and Ha, T. (2006) 'Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling', *Biophysical Journal*, 91(5), pp. 1941–1951. doi: 10.1529/biophysj.106.082487.

Merianos, H. J., Wang, J. and Moore, P. B. (2004) 'The structure of a ribosomal protein S8/spc operon mRNA complex.', *RNA*, 10(6), pp. 954–964. doi: 10.1261/rna.7030704.

Miles, H. T. (1959) 'A Proposed Interpretation of Infra-red Spectral Changes occurring upon the Interaction of Polynucleotides', *Nature*, 183(4678), pp. 1814–1814. doi: 10.1038/1831814a0.

Miller, O. L., Hamkalo, B. A. and Thomas, C. A. (1970) 'Visualization of Bacterial Genes in Action', *Science*, 169(3943), p. 392 LP – 395. doi: 10.1126/science.169.3943.392.

Milón, P., Maracci, C., Filonava, L., Gualerzi, C. O. and Rodnina, M. V (2012) 'Real-time assembly landscape of bacterial 30S translation initiation complex', *Nature Structural & Molecular Biology*, 19(6), pp. 609–615. doi: 10.1038/nsmb.2285.

Mitarai, N., Benjamin, J.-A. M., Krishna, S., Semsey, S., Csiszovszki, Z., Masse, E. and Sneppen, K. (2009) 'Dynamic features of gene expression control by small regulatory RNAs', *Proceedings of the National Academy of Sciences*, 106(26), pp. 10655–10659. doi: 10.1073/pnas.0901466106.

Miyamoto, T., Razavi, S., DeRose, R. and Inoue, T. (2013) 'Synthesizing Biomolecule-Based Boolean Logic Gates', *ACS Synthetic Biology*, 2(2), pp. 72–82. doi: 10.1021/sb3001112.

Mori, S., Abeygunawardana, C., Johnson, M. O. and van Zijl, P. C. (1995) 'Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation.', *Journal of magnetic resonance. Series B*, 108(1), pp. 94–98. doi: 10.1006/jmrb.1995.1109.

Mustoe, A. M., Brooks, C. L. and Al-Hashimi, H. M. (2014) 'Hierarchy of RNA functional dynamics.', *Annual review of biochemistry*, 83(1), pp. 441–66. doi: 10.1146/annurev-biochem-060713-035524.

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L. and Breaker, R. R. (2002) 'Genetic control by a metabolite binding mRNA', *Chemistry and Biology*, 9(9), pp. 1043–1049. doi: 10.1016/S1074-5521(02)00224-7.

Nakagawa, S., Niimura, Y., Miura, K. and Gojobori, T. (2010) 'Dynamic evolution of translation initiation mechanisms in prokaryotes.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), pp. 6382–7. doi: 10.1073/pnas.1002036107.

Neupane, K., Yu, H., Foster, D. A. N., Wang, F. and Woodside, M. T. (2011) 'Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism', *Nucleic Acids Research*, 39(17), pp. 7677–7687. doi: 10.1093/nar/gkr305.

Noeske, J., Buck, J., Fürtig, B., Nasiri, H. R., Schwalbe, H. and Wöhnert, J. (2007) 'Interplay of "induced fit" and preorganization in the ligand induced folding of the aptamer domain of the guanine binding riboswitch', *Nucleic Acids Research*, 35(2), pp. 572–583. doi: 10.1093/nar/gkl1094.

Noeske, J., Richter, C., Grundl, M. A., Nasiri, H. R., Schwalbe, H. and Wöhnert, J. (2005) 'An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(5), pp. 1372–7. doi: 10.1073/pnas.0406347102.

Noeske, J., Schwalbe, H. and Wöhnert, J. (2007) 'Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain', *Nucleic Acids Research*, 35(15), pp. 5262–5273. doi: 10.1093/nar/gkm565.

Nozinovic, S., Reining, A., Kim, Y.-B., Noeske, J., Schlepckow, K., Wöhnert, J. and Schwalbe, H. (2014) 'The importance of helix P1 stability for structural pre-organization and ligand binding affinity of the adenine riboswitch aptamer domain', *RNA Biology*, 11(5), pp. 83–82. doi: 10.4161/rna.29439.

Nudler, E. (2012) 'RNA Polymerase Backtracking in Gene Regulation and Genome Instability', *Cell*, 149(7), pp. 1438–1445. doi: 10.1016/j.cell.2012.06.003.

Ogle, J. M., Brodersen, D. E., Clemons, W. M., Tarry, M. J., Carter, A. P. and Ramakrishnan, V. (2001) 'Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit', *Science*, 292(5518), pp. 897–902. doi: 10.1126/science.1060612.

Pikovskaya, O., Polonskaia, A., Patel, D. J. and Serganov, A. (2011) 'Structural principles of nucleoside selectivity in a 2′-deoxyguanosine riboswitch', *Nature Chemical*

Biology, 7(10), pp. 748–755. doi: 10.1038/nchembio.631.

Porter, E. B., Marcano-Velázquez, J. G. and Batey, R. T. (2014) 'The purine riboswitch as a model system for exploring RNA biology and chemistry.', *Biochimica et biophysica acta*, 1839(10), pp. 919–930. doi: 10.1016/j.bbagrm.2014.02.014.

Preus, S., Noer, S. L., Hildebrandt, L. L., Gudnason, D. and Birkedal, V. (2015) 'iSMS: single-molecule FRET microscopy software', *Nature Methods*, 12(7), pp. 593–594. doi: 10.1038/nmeth.3435.

Price, I. R., Grigg, J. C. and Ke, A. (2014) 'Common themes and differences in SAM recognition among SAM riboswitches.', *Biochimica et biophysica acta*, 1839(10), pp. 931–938. doi: 10.1016/j.bbagrm.2014.05.013.

Prinz, H. (2010) 'Hill coefficients, dose-response curves and allosteric mechanisms.', *Journal of chemical biology*. Springer, 3(1), pp. 37–44. doi: 10.1007/s12154-009-0029-3.

Proshkin, S., Rahmouni, A. R., Mironov, A. and Nudler, E. (2010) 'Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation', *Science*, 328(5977), pp. 504–508. doi: 10.1126/science.1184939.

Reining, A., Nozinovic, S., Schlepckow, K., Buhr, F., Fürtig, B. and Schwalbe, H. (2013) 'Three-state mechanism couples ligand and temperature sensing in riboswitches', *Nature*, 499(7458), pp. 355–359. doi: 10.1038/nature12378.

Ren, A., Rajashankar, K. R. and Patel, D. J. (2012) 'Fluoride ion encapsulation by Mg2+ ions and phosphates in a fluoride riboswitch', *Nature*, 486(7401), pp. 85–9. doi: 10.1038/nature11152.

Ren, A., Xue, Y., Peselis, A., Serganov, A., Al-Hashimi, H. M. and Patel, D. J. (2015) 'Structural and Dynamic Basis for Low-Affinity, High-Selectivity Binding of L-Glutamine by the Glutamine Riboswitch', *Cell Reports*, 13(9), pp. 1800–1813. doi: 10.1016/j.celrep.2015.10.062.

Rieder, R., Lang, K., Graber, D. and Micura, R. (2007) 'Ligand-Induced Folding of the Adenosine Deaminase A-Riboswitch and Implications on Riboswitch Translational Control', *ChemBioChem*, 8(8), pp. 896–902. doi: 10.1002/cbic.200700057.

Rinaldi, A. J., Lund, P. E., Blanco, M. R. and Walter, N. G. (2016) 'The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs show ligand-dependent accessibility bursts', *Nature Communications*, 7, pp. 1–10. doi: 10.1038/ncomms9976.

Rinnenthal, J., Buck, J., Ferner, J., Wacker, A., Fürtig, B. and Schwalbe, H. (2011) 'Mapping the landscape of RNA dynamics with NMR spectroscopy.', *Accounts of chemical research*, 44(12), pp. 1292–301. doi: 10.1021/ar200137d.

Rupert, P. B. and Ferré-D'Amaré, A. R. (2001) 'Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis.', *Nature*, 410(6830), pp. 780–786. doi: 10.1038/35071009.

Schubert, M., Lapouge, K., Duss, O., Oberstrass, F. C., Jelesarov, I., Haas, D. and Allain, F. H.-T. (2007) 'Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA', *Nature Structural & Molecular Biology*, 14(9), pp. 807–813. doi: 10.1038/nsmb1285.

Scott, W. G., Horan, L. H. and Martick, M. (2013) 'The Hammerhead Ribozyme', in

Progress in molecular biology and translational science, pp. 1–23. doi: 10.1016/B978-0-12-381286-5.00001-9.

Serganov, A. and Nudler, E. (2013) 'A Decade of Riboswitches', *Cell*, 152(1), pp. 17–24. doi: 10.1016/j.cell.2012.12.024.

Serganov, A., Yuan, Y.-R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., Hobartner, C., Micura, R., Breaker, R. R. and Patel, D. J. (2004) 'Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-Sensing mRNAs', *Chemistry & Biology*, 11(12), pp. 1729–1741. doi: 10.1016/j.chembiol.2004.11.018.

Shaka, A. J. and Keeler, J. (1987) 'Broadband spin decoupling in isotropic-liquids', *Progress in Nuclear Magnetic Resonance Spectroscopy*, 19(1), pp. 47–129. doi: 10.1016/0079-6565(87)80008-0.

Shallop, A. J., Gaffney, B. L. and Jones, R. A. (2003) 'Use of 13C as an Indirect Tag in 15N Specifically Labeled Nucleosides. Syntheses of [8-13C-1,7,NH2-15N3]Adenosine, -Guanosine, and Their Deoxy Analogues', *Journal of Organic Chemistry*, 68(22), pp. 8657–8661. doi: 10.1021/jo0345446.

Shine, J. and Dalgarno, L. (1974) 'The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites.', *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), pp. 1342–6. doi: 10.1073/pnas.71.4.1342.

Simonetti, A., Marzi, S., Myasnikov, A. G., Fabbretti, A., Yusupov, M., Gualerzi, C. O. and Klaholz, B. P. (2008) 'Structure of the 30S translation initiation complex', *Nature*, 455(7211), pp. 416–420. doi: 10.1038/nature07192.

Singh, P. and Sengupta, S. (2012) 'Phylogenetic analysis and comparative genomics of purine riboswitch distribution in prokaryotes', *Evolutionary Bioinformatics*, 2012(8), pp. 589–609. doi: 10.4137/EBO.S10048.

Sinha, J., Reyes, S. J. and Gallivan, J. P. (2010) 'Reprogramming bacteria to seek and destroy an herbicide', *Nature Chemical Biology*, 6(6), pp. 464–470. doi: 10.1038/nchembio.369.

Sklenář, V. and Bax, A. (1987) 'Spin-echo water suppression for the generation of pure-phase two-dimensional NMR spectra', *Journal of Magnetic Resonance (1969)*, 74(3), pp. 469–479. doi: 10.1016/0022-2364(87)90269-1.

De Smit, M. H. and Van Duin, J. (2003) 'Translational standby sites: How ribosomes may deal with the rapid folding kinetics of mRNA', *Journal of Molecular Biology*, 331(4), pp. 737–743. doi: 10.1016/S0022-2836(03)00809-X.

Smith, K. D., Lipchock, S. V, Ames, T. D., Wang, J., Breaker, R. R. and Strobel, S. A. (2009) 'Structural basis of ligand binding by a c-di-GMP riboswitch.', *Nature structural & molecular biology*, 16(12), pp. 1218–23. doi: 10.1038/nsmb.1702.

Snoussi, K. and Leroy, J. L. (2001) 'Imino proton exchange and base-pair kinetics in RNA duplexes', *Biochemistry*, 40(30), pp. 8898–8904. doi: 10.1021/bi010385d.

Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D. and Brutscher, B. (2013) 'BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins', *Journal of Biomolecular NMR*, 55(4), pp. 311–321. doi:

10.1007/s10858-013-9715-0.

Sperling, J., Azubel, M. and Sperling, R. (2008) 'Structure and Function of the Pre-mRNA Splicing Machine', *Structure*, 16(11), pp. 1605–1615. doi: 10.1016/j.str.2008.08.011.

Steinert, H. S., Rinnenthal, J. and Schwalbe, H. (2012) 'Individual basepair stability of DNA and RNA studied by NMR-detected solvent exchange.', *Biophysical journal*, 102(11), pp. 2564–74. doi: 10.1016/j.bpj.2012.03.074.

Steitz, J. and Jakes, K. (1975) 'How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli', *PNAS*, 72(12), pp. 4734–4738. doi: 10.1073/pnas.72.12.4734.

Steitz, T. A. (2008) 'A structural understanding of the dynamic ribosome machine', *Nature Reviews Molecular Cell Biology*, 9(3), pp. 242–253. doi: 10.1038/nrm2352.

Subramanian, A.-R. (1983) 'Structure qnd Functions of Ribosomal Protein S1', *Progress in Nucleic Acid Research and Molecular Biology*, 28, pp. 101–142. doi: 10.1016/S0079-6603(08)60085-9.

Sudarsan, N., Wickiser, J. K., Nakamura, S., Ebert, M. S. and Breaker, R. R. (2003) 'An mRNA structure in bacteria that controls gene expression by binding lysine', *Genes & Development*, 17(21), pp. 2688–2697. doi: 10.1101/gad.1140003.

Suddala, K. C., Rinaldi, A. J., Feng, J., Mustoe, A. M., Eichhorn, C. D., Liberman, J. A., Wedekind, J. E., Al-Hashimi, H. M., Brooks, C. L. and Walter, N. G. (2013) 'Single transcriptional and translational preQ1 riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure', *Nucleic Acids Research*, 41(22), pp. 10462–10475. doi: 10.1093/nar/gkt798.

Swoboda, M., Henig, J., Cheng, H. M., Brugger, D., Haltrich, D., Plumeré, N. and Schlierf, M. (2012) 'Enzymatic oxygen scavenging for photostability without ph drop in single-molecule experiments', *ACS Nano*, 6(7), pp. 6364–6369. doi: 10.1021/nn301895c.

Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A. and Xie, X. S. (2010) 'Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells', *Science*, 329(5991), pp. 533–538. doi: 10.1126/science.1188308.

Thore, S., Leibundgut, M. and Ban, N. (2006) 'Structure of the Eukaryotic Thiamine Pyrophosphate Riboswitch with Its Regulatory Ligand', *Science*, 312(5777), pp. 1208–1211. doi: 10.1126/science.1128451.

Tyrrell, J., McGinnis, J. L., Weeks, K. M. and Pielak, G. J. (2013) 'The cellular environment stabilizes adenine riboswitch RNA structure.', *Biochemistry*, 52(48), pp. 8777–85. doi: 10.1021/bi401207q.

Tyrrell, J., Weeks, K. M. and Pielak, G. J. (2015) 'Challenge of Mimicking the Influences of the Cellular Environment on RNA Structure by PEG-Induced Macromolecular Crowding', *Biochemistry*, 54(42), pp. 6447–6453. doi: 10.1021/acs.biochem.5b00767.

Van, Q. N., Chmurny, G. N. and Veenstra, T. D. (2003) 'The depletion of protein signals in metabonomics analysis with the WET–CPMG pulse sequence', *Biochemical and Biophysical Research Communications*, 301(4), pp. 952–959. doi: 10.1016/S0006-

291X(03)00079-2.

Wacker, A., Buck, J., Mathieu, D., Richter, C., Wohnert, J. and Schwalbe, H. (2011) 'Structure and dynamics of the deoxyguanosine-sensing riboswitch studied by NMR-spectroscopy', *Nucleic Acids Research*, 39(15), pp. 6802–6812. doi: 10.1093/nar/gkr238.

Warhaut, S., Mertinkus, K. R., Höllthaler, P., Fürtig, B., Heilemann, M., Hengesbach, M. and Schwalbe, H. (2017) 'Ligand-modulated folding of the full-length adenine riboswitch probed by NMR and single-molecule FRET spectroscopy', *Nucleic Acids Research*, 54, pp. 6447–6453. doi: 10.1093/nar/gkx110.

Werstuck, G. and Green, M. R. (1998) 'Controlling Gene Expression in Living Cells Through Small Molecule-RNA Interactions', *Science*, 282(5387), p. 296 LP – 298. doi: 10.1126/science.282.5387.296.

Wilson, T. J., Liu, Y., Domnick, C., Kath-Schorr, S. and Lilley, D. M. J. (2016) 'The Novel Chemical Mechanism of the Twister Ribozyme', *Journal of the American Chemical Society*, 138(19), pp. 6151–6162. doi: 10.1021/jacs.5b11791.

Winkler, W. C., Cohen-Chalamish, S. and Breaker, R. R. (2002) 'An mRNA structure that controls gene expression by binding FMN', *Proceedings of the National Academy of Sciences*, 99(25), pp. 15908–15913. doi: 10.1073/pnas.212628899.

Wishart, D., Bigam, C., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. and Sykes, B. (1995) '1H, 13C and 15N chemical shift referencing in biomolecular NMR', *Journal of Biomolecular NMR*, 6(2), pp. 135–140. doi: 10.1007/BF00211777.

Xie, Z., Srividya, N., Sosnick, T. R., Pan, T. and Scherer, N. F. (2004) 'Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(2), pp. 534–9. doi: 10.1073/pnas.2636333100.

Yusupova, G. Z., Yusupov, M. M., Cate, J. H. D. and Noller, H. F. (2001) 'The path of messenger RNA through the ribosome', *Cell*, 106(2), pp. 233–241. doi: 10.1016/S0092-8674(01)00435-4.

Zhang, J. and Ferré-D'Amaré, A. R. (2015) 'Structure and mechanism of the T-box riboswitches.', *Wiley interdisciplinary reviews. RNA*, 6(4), pp. 419–33. doi: 10.1002/wrna.1285.

Zhang, Q., Sun, X., Watt, E. D. and Al-Hashimi, H. M. (2006) 'Resolving the Motional Modes That Code for RNA Adaptation', *Science*, 311(5761).

Zuker, M. (2000) 'Calculating nucleic acid secondary structure', *Current Opinion in Structural Biology*, pp. 303–310. doi: 10.1016/S0959-440X(00)00088-9.

List of abbreviations

% w/v	percent weight per volume
°C	degree Celsius
µm	micrometer
µs	microsecond
^{112}Asw	112-nucleotide Asw
^{112}Asw-*nx*	112-nucleotide Asw sample prepared without buffer exchange
^{112}Asw-*x*	112-nucleotide Asw (x highlights that the sample was buffer exchanged)
^{127}Asw	127-nucleotide Asw
5'-UTR	5' untranslated region
A	adenosine
a.u.	arbitrary units
A_{260}	absorbance at 260nm UV light
ACE	bis(acetoxyethoxy)-methyl ether
AOTF	acousto-optical tunable filter
AP	2-aminopurine
apoA_{STAB}	apoA-stabilized mutant
apoB_{STAB}	apoB-stabilized mutant
aqua-P/C/I	aqueous phenol chlorophorm isoamylalcohol
Asw	translational *add* adenine riboswitch from *Vibrio vulnificus*
BBFO	double-resonance broad band NMR probe featuring 19F observation
BEST-TROSY	Band□selective Excitation Short□Transient Transverse Relaxation Optimized Spectroscopy
BSA	bovine serum albumine
C	cytidine
c-di-AMP	cyclic diadenylate monophosphate
c-di-GMP	cyclic diguanylate monophosphate
cm	centimeter
CPMG	Carr-Purcell-Meiboom-Gill
cryo	cryogenic
Cy3	cyanine 3 dye
Cy5	cyanine 5 dye
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
DEAE	diethylaminoethyl
DEPC	diethyl pyrocarbonate
dGTP	deoxyguanosine triphosphate
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA

DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	dithiothreitol
dTTP	deoxythimidine triphosphate
DW	dwelltime
E	FRET efficiency
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescent protein
EHMM	HMM fit of E
Em	emission
EMCCD	electron multiplying charge coupled device
eq	equivalent
ERETIC	Electronic REference To access In vivo Concentrations
Ex	excitation
F	fluorescence intensity
FHSQC	Fast HSQC
FID	free induction decay
fMet	formyl methionine
FMN	flavin mononucleotide
FRET	Förster resonance energy transfer
FTIR	Fourier transform infrared
G	guanosine
g	relative centrifugal force
GARP	Globally Optimized Alternating Phase Rectangular Pulse decoupling scheme
h	hour
HMM	hidden Markov model
HPLC	high pressure liquid chromatography
HSQC	Heteronuclear Single Quantum Coherence
Hz	Hertz
I_A	FRET acceptor fluorescence intensity
IC	initiation complex
I_D	FRET donor fluorescence intensity
IDT	integrated DNA technologies
IF	initiation factor
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
K	Kelvin
k_{AB}	rate constant for the apoA to apoB transition
k_{BA}	rate constant for the apoB to apoA transition
KD	dissociation constant
$K_{D,app}$	apparent dissociation constant
kDa	kilo Dalton
k_{dock}	docking rate constant
kHz	kilo Hertz

k_{ON}	binding rate constant
k_{undock}	undocking rate constant
L	lysine
L	liters
LB	lysogeny broth
MCT	Mercury-Cadmium-Telluride
MHz	Megahertz
min	minutes
mL	milliliter
mM	millimolar
mm	millimeter
mRNA	messenger RNA
ms	millisecond
mW	milliwatt
n	Hill coefficient
nM	nanomolar
nM	nanomolar
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
NOESY	Nuclear Overhauser Effect Spectroscopy
ns	nanosecond
P	proline
PCI	periphal component interconnect
PCR	polymerase chain reaction
PDB ID	RCSB Protein data bank identification code
PEG	polyethylene glycol
pH	potential of hydrogen
PIC	pre initiation complex
pM	picomolar
pmol	picomol
ppm	parts per million
$preQ_1$	pre-queuosine$_1$
ps	picosecond
PTFE	polytetrafluoroethylene
r	fluorescence anisotropy
RNA	ribonucleic acid
rNTP	ribonucleotide triphosphate
RP	reversed phase
rpm	rounds per minute
S	Svedberg unit for sedimentation rate
S	stoichiometry factor (smALEX)
s	second
SAM	S-adenosyl methionine
SD	Shine Dalgarno sequence

SHAPE	selective 2'-hydroxyl acylation analysed by primer extension
smALEX	single-molecule alternating laser excitation
smFRET	single-molecule Förster resonance energy transfer
ssDNA	single-stranded DNA
T4	phage T4
T7	phage T7
TAE	tris acetate EDTA
TBE	tris borate EDTA
TCI	proton-optimized triple resonance NMR 'inverse' probe
TDP	transition density plot
TEAA	triethylamine acetate
TEMED	tetramethylethylenediamine
tif	tagged image file format
TODP	transition occupancy density plot
TPP	thiamine pyrophosphate
tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
Tte	*Thermoanaerobacter tengcongensis*
TXI	triple resonance X1+ X2 nucleus decoupling NMR 'inverse' probe
U	uridine
U	enzymatic activity units
UV	ultra violet light
V	Volt
vis	visible light
vol	volume
W	Watt
WATERGATE	WATER suppression by GrAdient Tailored Excitation
WaterLOGSY	water-Ligand Observed via Gradient SpectroscopY
wt	wildtype
XRD	x-ray diffraction
Y	Pyrimidine
YIPP	yeast inorganic pyrophosphatase
Z-GRD	z-gradient
δ	chemical shift
ε	extinction coefficient
λ_{max}	maximal absorbance wavelength
µL	microliter
µM	micromolar

Appendix

Appendix 1: Assignment of the imino NH resonances of mutant apoB_{STAB} ^{112}Asw without adenine but with 5 mM Mg^{2+} at 283 K

Mutated residues are highlighted in red. Residues that showed chemical shift changes ($\Delta\delta^{1}$H >0.05 or $\Delta\delta^{15}$N >0.2) compared to the wildtype are denoted *. These are the two guanosines adjacent to the mutated base pairs (G43 and G46) and two P4 helix residues (G81 and G115).

Residues	**^{1}H**	**^{15}N**
G14	11.55	145.1
G16	12.89	147.8
G38	12.70	147.8
G43*	12.22	147.2
G44	13.19	148.5
G46*	13.21	147.9
U47	12.17	158.6
G57	12.42	146.7
G59	13.21	147.9
U68	14.18	162.8
U70	13.96	162.3
U71	13.11	162.1
G72	11.81	146.6
G81*	13.21	147.9
U82	14.40	162.9
U90	13.96	162.3
U91	13.69	162.5
U92	12.98	161.9
U106	13.22	161.9
G112	11.94	147.5
G115*	12.85	146.8

Appendix 2: MATLAB script for the analysis of smFRET histogram data

This script was provided by Martin Hengesbach

```
%Program to read binary traces file output by Hazens IDL program written by
%Michael Stone 01/17/2004
clear all;
total_frets = [];
n_molecules = [];


s = pwd;
files = dir(s);

for i = 3:length(files)

fileName = files(i).name;
s = strcat(s,'\');
fullFileName = strcat(s,fileName);
FID = fopen(fullFileName);

% data = importdata(fullFileName);
accepted_traces = []; % an array to hold good molecule data for future analysis
fretthreshold = 4000; % this is a threshold value to prevent crazy fret traces

% %First read in file from directory
% [fileName,path1] = uigetfile('*.traces','Read Traces File');
% addpath(path1)  %add path to path list each time to ensure proper file access
% FID = fopen(fileName);

%find number of frames and peaks from traces header file
[n_fr,z1] = fread(FID,1,'int32');    %z1 is just the number of indicated data types read (should be 1)
[n_traces,z2] = fread(FID,1,'int16');

rate_fr = 10; %input('At what frame rate was the data collected?');
n_peaks = n_traces/2;
n_molecules = [n_molecules n_peaks];
exp_length =  n_fr/rate_fr; %(seconds)
time = (1/rate_fr:(1/rate_fr):exp_length);


%%Now read in the rest of the data from the traces file
[total_data,z3] = fread(FID,[n_traces+1,n_fr],'int32');
fclose=(FID);
total_data = total_data(2:n_traces,1:n_fr);   %% this gets rid of the frame number col, which is the first colum in a traces file


%RemoveBleachedMolecules

total_donor = [];
total_acceptor = [];
total_acceptor_corrected = [];

total_donor(1:n_peaks-1,1:3) = total_data(2*(0:n_peaks-2)+1,1:3);
%%remember matlab starts arrays at 1
total_acceptor(1:n_peaks-1,1:3) = total_data(2*(0:n_peaks-2)+2,1:3);


%correct the acceptor intensities for 10% leakage from the donor channel

total_acceptor_corrected(1:n_peaks-1,1:3)= total_acceptor(1:n_peaks-1,1:3)-0.1*total_donor(1:n_peaks-1,1:3);
%total_acceptor_corrected = total_acceptor;

for i = 1:n_peaks-1
   i
   dyesum = total_donor(i,:)+total_acceptor_corrected(i,:);
   FRET = total_acceptor_corrected(i,:)./dyesum;
   [rows] = find(dyesum <fretthreshold);
   FRET(rows) = [];
   total_frets = [total_frets mean(FRET)];
   %remove FRET below -0.1 and above 1.1:
   [rowsoutsidelow] = find(total_frets <-0.1);
   total_frets(rowsoutsidelow) = [];
```

```
   [rowsoutsidehigh] = find(total_frets >1.1);
   total_frets(rowsoutsidehigh) = [];
end

end

%this creates the vector for binning into 0.025 FRET steps
bins = [-0.1   -0.075  -0.05   -0.025  0   0.025   0.05    0.075   0.1 0.125   0.15    0.175   0.2 0.225   0.25    0.275   0.3 0.325   0.35
0.375   0.4 0.425   0.45    0.475   0.5 0.525   0.55    0.575   0.6 0.625   0.65    0.675   0.7 0.725   0.75    0.775   0.8 0.825   0.85    0.875
0.9 0.925   0.95    0.975   1   1.025   1.05    1.075   1.1];
bins = bins';
%and this makes the vector for exporting the histogram figures (to Origin)
binned_data = hist(total_frets,bins);
exp_hist = [bins, transpose(binned_data)];
%finally, this gives the raw data in a vector for export
export_frets = total_frets';

   figure
   hist(total_frets,bins)
   %plot A
   title(fileName)
   saveas(gcf,'histAVG3_STD2_b10_ft4000.jpg');



file2=strcat('histAVG3_STD2_b10_ft4000','.txt');

save(file2, 'exp_hist', '-ascii'); %this saves binned AVG3 FRET values into a txt file for origin!
```

Appendix 3: MATLAB scripts for the analysis of smFRET time trace data

Script 1: *export_flou_trajectories.m*

```
%Script works on a directory containing all .traces files of a specific measurement condition (and no other files)
%Needs frame rate in frames per second as an input

%The first part is an adaptation of Martins Script "readhisttracesandor3.m"
%reads in the background corrected donor and acceptor intensities
%corrects acceptor intensities for 10% donor leakage
%safes I(frame) as column vectors for all molecules in accepted_traces; (ID1,IA1, ID2, IA2 ... IDn, IAn)

clear all;
total_frets = [];
n_molecules = [];

alllifetimes=[]
list=[]
rate_fr = input('At what frame rate was the data collected?');

s = pwd;
files = dir(s);

for i = 3:length(files)

fileName = files(i).name;
s = strcat(s,'\');
fullFileName = strcat(s,fileName);
FID = fopen(fullFileName);

accepted_traces = []; % an array to hold good molecule data for future analysis
fretthreshold = 4000; % this is the threshold value for photobleaching used in the second part of the script
fretthreshold2 = 10; % this is the threshold value for intensity scaling used in the second part of the script

%find number of frames and peaks from traces header file
[n_fr,z1] = fread(FID,1,'int32');    %z1 is just the number of indicated data types read (should be 1)
[n_traces,z2] = fread(FID,1,'int16');

n_peaks = n_traces/2;
n_molecules = [n_molecules n_peaks];
exp_length =  n_fr/rate_fr; %(seconds)
time = (1/rate_fr:(1/rate_fr):exp_length);

% Now read in the rest of the data from the traces file
[total_data,z3] = fread(FID,[n_traces+1,n_fr],'int32');
fclose=(FID);
total_data = total_data(2:n_traces,1:n_fr);   % this gets rid of the frame number col, which is the first colum in a traces file

total_donor = [];
total_acceptor = [];
total_acceptor_corrected = [];

total_donor(1:n_peaks-1,1:n_fr) = total_data(2*(0:n_peaks-2)+1,1:n_fr);
total_acceptor(1:n_peaks-1,1:n_fr) = total_data(2*(0:n_peaks-2)+2,1:n_fr);

%correct the acceptor intensities for 10% leakage from the donor channel
total_acceptor_corrected(1:n_peaks-1,1:n_fr)= total_acceptor(1:n_peaks-1,1:n_fr)-0.1*total_donor(1:n_peaks-1,1:n_fr);


%safes all ID and all corrected IA in column pairs into the accepted_traces variable; 1 row per frame!
total_data_Acorr = reshape([total_donor(:) total_acceptor_corrected(:)]',2*size(total_donor,1), [])
accepted_traces = total_data_Acorr.'



%The second part pre-processes the intensity trajectories for Hidden Markov Modelling in HaMMy
%It scales the intensities such that E-FRET falls between 0 and 1
%It terminates the trajectory at a photobleaching event i.e. if dyesum<fretthreshold or E<0.1 in ave3 medfiltered data
%It removes all molecules that were terminated within the first three frames: dim molecules and D-only molecules
%It exports one ASCII file per molecule of the form "Time [s], ID [a.u.],IA [a.u.]"
%Additionally, for all saved molecule of each *.traces file, the script exports the lifetime [s] in a *.txt file (one list of lifetimes per
*.traces file)
```

```
%Scale the intensities so that E[0,1] and set trace_end right before avedyesum<fretthreshold or E<0.1 (3-frame average)

data = accepted_traces;
num_tr = size(data,1);
k=1

      while k < size(data,2)/2+1

      index = k*2;
      donor = data(:,index-1);
      accep = data(:,index);

      %SW: scale back negative intensities
      [rows] = find(donor<fretthreshold2); % creates a vector that contains indices of cells with FRET > 1 due to negative donor
intensities
      donor(rows) = 10;  % substitutes donorI<0 values in [rows] with I=10

      [rows] = find(accep<fretthreshold2); % creates a vector that contains indices of cells with FRET < 0 due to negative accep
intensities
      accep(rows) = 10;  % substitutes accepI<0 values in [rows] with I=10

      %SW: Create a median filtered trajectory to determine trace_end

      cutfilter = 3 %each point n will be replaced by average of the data interval [n-1:n+1]

      aveDonor = medfilt1(donor',cutfilter); %median averaging the donor intensity
      aveAccep = medfilt1(accep',cutfilter); %median averaging the acceptor intensity
      avedyesum = aveAccep+aveDonor;

      [rows] = find(avedyesum<fretthreshold)% creates a vector that contains indices of cells with dim intensities
      aveAccep(rows) = 0;  % substitutes IA values in [rows] with IA=0, so that they become E=0 and are filtered with the following E-
FRET criterion
      avefret = aveAccep./avedyesum

      [hammy_rows] = find(avefret < 0.1) %creates a vector that contains indices of cells with E<0.1

      if [hammy_rows] > 0
      trace_end = min(hammy_rows)-2 % means the end of the trace should be 2 frames before the avedyesum is smaller than the
fretthreshold

      else
      trace_end = num_tr
      end

% Then either export or skip the trace of the molecule
% Traces are skipped if the molecule bleaches during the first 3 frames (avedyesum<fretthreshold or avefret<0.1; s.a.)

      if  (trace_end < 3 )

            k=k+1

      else

            donor=donor'
            accep=accep'

            hammy_trace = [time(1:trace_end); donor(1:trace_end); accep(1:trace_end)];
            hammy_traceexport = hammy_trace';

            pattern = '.traces';
            replacement = '_m';
            fileName1 = regexprep(fileName,pattern,replacement)
            file1=strcat(num2str(fileName1),num2str(k),'.dat');
            save (file1, 'hammy_traceexport', '-ascii', '-tabs');

            %SW: Puts the molecule's lifetime into alllifetimes
            lifetime = trace_end/rate_fr

            list=vertcat(alllifetimes,lifetime)
            alllifetimes=list

            k=k+1

      end
```

```
    end

    %SW: Saves all lifetimes for that movie

    file2=strcat(num2str(fileName),'_lifetimes.txt');
    save (file2, 'alllifetimes', '-ascii')
    alllifetimes=[]


end
```

Script 2: *create_input_for_ID_IA_HMM.m*

This script was written by Klara R. Mertinkus as part of a supervised Bachelor thesis

```
%Script works on a directory containing the *.dat files with the "Time,IA,ID" trajectories of all molecules for a specific measurement condition (and no other files)
%It creates the corresponding trajectories for Hidden Markov modelling of the normalized donor intensity "Time, 1-ID/IDmax, ID/IDmax" in the files *d.dat
%and the corresponding trajectories for Hidden Markov modelling of the normalized acceptor intensity "Time, 1-IA/IAmax, IA/IAmax" in the files *a.dat

clear all

files=dir('*.dat');

for j=1:numel(files)
filename = files(j).name;

%load dat data: time, ID, IA
data=loadascii(filename,'%')

time=data(:,1)

donor=data(:,2)
normdonor=data(:,2)/max(donor)
subt=1-normdonor

data_d=[time subt normdonor]
hammy_traceexport = data_d

pattern = '.dat';
replacement = 'd';
filename1 = regexprep(filename,pattern,replacement)

file=strcat(num2str(filename1),'.dat');
save (file, 'hammy_traceexport', '-ascii', '-tabs');

accep=data(:,3)
normaccep=data(:,3)/max(accep)
subt2=1-normaccep

data_a=[time subt2 normaccep]
hammy_traceexport2 = data_a

pattern = '.dat';
replacement2 = 'a';
filename2 = regexprep(filename,pattern,replacement2)

file2=strcat(num2str(filename2),'.dat');
save (file2, 'hammy_traceexport2', '-ascii', '-tabs');

end
```

Script 3: *evaluate_HMM_fits.m*

This script was finalized together with Klara R. Mertinkus as part of a supervised Bachelor thesis

```
%Script works on a directory containing the 3 Hammy output files *path.dat, *apath.dat and *dpath.dat for each molecule of a specific measurement condition (and no other files)
%Needs the additional function loadascii.m from http://schwalbe.org.chemie.uni-frankfurt.de/research

clear all

rate_fr = input('At what frame rate was the data collected?');

alldwell =[]

TDPlist=[]

TODPlist = []
TODPdyn=0
TODPstat=0
TODPhistat=0

alllifetimes = []
histatlifetimes = []

mkdir('dyn_jpg')
mkdir('histat_jpg')
mkdir('lostat_jpg')

%read acceptor paths
files_a=dir('*apath.dat');

for j=1:numel(files_a)
filename_a=files_a(j).name;

data_a=loadascii(filename_a,'%'); %Format: Frame, 1-IA, IA, Anorm, Afit
trace_end=size(data_a,1)

if trace_end < 50  %skip if shorter than 50 frames

%then do nothing

else

%load FRET path and donor path and analyze the molecule

pattern='apath.dat';
replacement='';
moleculename=regexprep(filename_a,pattern,replacement);

files_e=strcat(num2str(moleculename),'path.dat');
filename_e=files_e; %only FRET name

files_d=strcat(num2str(moleculename),'dpath.dat');
filename_d=files_d; %only donor name

data_d=loadascii(filename_d,'%'); %Format: Frame, 1-ID, ID, Dnorm, Dfit
data_e=loadascii(filename_e,'%'); %Format: Frame, ID, IA, E, Efit

frame=data_e(1:end,1);

Efit=data_e(:,5);
deltaEfit=Efit(2:trace_end,1)-Efit(1:trace_end-1,1);
deltaEfit(trace_end,:)=0;

Dfit=data_d(:,5);
deltaDfit=Dfit(2:trace_end,1)-Dfit(1:trace_end-1,1);
deltaDfit(trace_end,:)=0;

Afit=data_a(:,5);
deltaAfit=Afit(2:trace_end,1)-Afit(1:trace_end-1,1);
```

```
deltaAfit(trace_end,:)=0;

slope_data=[frame deltaEfit deltaDfit deltaAfit];

%%export slope_data as optional control
%hammy_slopeexport=slope_data;
%file_slope=strcat(num2str(moleculename),'slope.dat');
%save(file_slope, 'hammy_slopeexport', '-ascii', '-tabs')

%compare slope data to find "real" transitions
%"Real" transitions have Dfit intensity change >10% in frame n
%Anticorrelated Afit intensity change >10% within frame n-2 and frame n+2
%And Efit change >0.1 within frame n-2 and frame n+2

arr=zeros(trace_end,1)
%intvlist=[]
   for n=3:trace_end-2
        if abs(slope_data(n,3))>0.1   %donor change

        if abs(slope_data(n,4))>0.1 & slope_data(n,3)*slope_data(n,4)<0  %anticorrelated acceptor change in same frame
           arr(n,1)=1;
           slope_data(n,3)=0;
           slope_data(n,4)=0;
           %intv=0;
        elseif abs(slope_data(n-1,4))>0.1 & slope_data(n,3)*slope_data(n-1,4)<0 || abs(slope_data(n+1,4))>0.1 &
slope_data(n,3)*slope_data(n+1,4)<0 % anticorrelated acceptor change in frame-1 or frame +1
           arr(n,1)=1;
           slope_data(n,3)=0;
           slope_data(n-1:n+1,4)=0;
           %intv=1;
        elseif abs(slope_data(n-2,4))>0.1 & slope_data(n,3)*slope_data(n-2,4)<0 || abs(slope_data(n+2,4))>0.1 &
slope_data(n,3)*slope_data(n+2,4)<0 % anticorrelated acceptor change in frame-2 or frame +2
           arr(n,1)=1;
           slope_data(n,3)=0;
           slope_data(n-2:n+2,4)=0;
           %intv=2;
        end
        %intvlist=vertcat(intvlist,intv)

        if arr(n,1)==1 && max(abs(slope_data((n-2):(n+2),2)))<0.1 %removes transitions with no FRET change > 0.1 within frame-2
and frame+2
        arr(n,1)=0
        end

        end
   end


%Create matrix Trace_steps according to the "real" transitions
%Trace_steps is the trace separated into steps between transitions; it contains
%as rows: "t0 to transition 1", "transition 1 to transition 2", ..., "last transition to Trace_end"
%as columns: steptime, mean(ID), mean(IA), mean(E), mean(Efit)
%Each steptime corresponds to a dwelltime, except the first steptime (t=t0 until transition 1) and the last steptime (last transition until
trace_end)
%Molecules with no Hammy transition will have 1 steptime (t=t0 until trace_end)
%The means are the respective means during the steptime

trans_rows=find(abs(arr)>0);
trace_trans=vertcat(trans_rows,trace_end);

Trace_steps=[];
k=1;
   for i=1:size(trace_trans)

   steptime=data_e(trace_trans(i),1)/rate_fr-data_e(k,1)/rate_fr+1/rate_fr

   if steptime == 1/rate_fr

   stepD=mean(data_e(k:trace_trans(i),2));
   stepA=mean(data_e(k:trace_trans(i),3));
   stepE=mean(data_e(k:trace_trans(i),4));
   stepEfit=mean(data_e(k:trace_trans(i),5));

   else
```

```
   stepD=mean(data_e(k:(trace_trans(i)-1),2));
   stepA=mean(data_e(k:(trace_trans(i)-1),3));
   stepE=mean(data_e(k:(trace_trans(i)-1),4));
   stepEfit=mean(data_e(k:(trace_trans(i)-1),5));

   end

   Trace_steps=vertcat(Trace_steps,[steptime,stepD,stepA,stepE,stepEfit]);

   k=trace_trans(i)+1;

   end


%Replace FRET-states in stepEfit with deltaEfit<0.1 by their average

FRETstates=unique(Trace_steps(:,5))

for s=1:size(FRETstates,1)

   deltaFRETstates=FRETstates-FRETstates(s)
   [sameFRETstate_rows]=find(abs(deltaFRETstates)<0.1)

   if size([sameFRETstate_rows],1) > 1

      FRETs1=FRETstates([sameFRETstate_rows])
      FRETs2=ones(size(FRETs1,1),1)*mean(FRETs1)

      FRETstates=changem(FRETstates,FRETs2,FRETs1)
      stepEfits = Trace_steps(:,5)
      stepEfits = changem(stepEfits,FRETs2,FRETs1)
      Trace_steps(:,5) =  stepEfits

   end

end

%Remove wrong steps from Trace_steps
%Wrong steps are sequential steps in Trace_steps with no change in stepEfit due to false positive "real" transitions
%Such false positives arise if a transitions out of a state was accepted but the transition into that state not (or vice versa)

True_steps=Trace_steps(1,:)
k=1

for s=2:size(Trace_steps,1)

   if  True_steps(k,5) == Trace_steps(s,5)
      True_steps(k,1)=True_steps(k,1)+Trace_steps(s,1)
      True_steps(k,2:5)=(True_steps(k,2:5)+Trace_steps(s,2:5))./2
   else
      True_steps=vertcat(True_steps,Trace_steps(s,:))
      k=k+1
   end

end

%Also calculate the transition time points true_trans for True_steps
%which is the equivalent of trace_trans for Trace_steps
true_trans=cumsum(True_steps(:,1))

%Now save the corrected HMM path according to True_steps in Efitcorr

s=[]
Efitcorr=[]
for s=1:size(True_steps,1)
edge=round(True_steps(s,1)*rate_fr)
segment=ones(edge,1).*True_steps(s,5)
Efitcorr=vertcat(Efitcorr,segment)
end


%Classify and analyze corrected HMM-paths

frame=data_e(:,1)+1;
frametime=frame/rate_fr;
```

```
if size(True_steps,1) > 1 % then the molecule is dynamic
TODPdyn=TODPdyn+1   %count it as dynamic

h(j)=figure;

subplot(24,1,2:6) % ID, IA
plot(frametime, data_e(:,2),'g', frametime, data_e(:,3),'r' )
ylabel('Intensity')
title(filename_e)

subplot(24,1,8:12) % IA/IAmax, Afit
plot(frametime, data_a(:,4),'k', frametime, data_a(:,5),'g' )
ylabel('Intensity')

subplot(24,1,14:18) % ID/IDmax, Dfit
plot(frametime, data_d(:,4),'k', frametime, data_d(:,5),'g' )
ylabel('Intensity')

subplot(24,1,20:24) % Efret, Efit, Efit_corrected with accepted transition points trace_trans and actual transition points true_trans
plot(frametime, data_e(:,4),'k', frametime, data_e(:,5),'g', frametime, Efitcorr,'m',  trace_trans(1:end-1)./10, Trace_steps(1:end-1,5),
'ob', true_trans(1:end-1), True_steps(1:end-1,5), '*b')
xlabel('Time/s')
ylabel('E-FRET')

set(gcf,'PaperPosition', [1 1 12.5 20]) % size of plot-output

saveas(h(j), strcat(filename_e, 'dyn.jpg'))

close(h(j))

if size(True_steps,1)>2 %the molecule has at least 2 transitions, and, hence actual dwelltimes

%Add the dwelltimes to the list alldwell of the format dwelltime, Efit, Efitafter
realdwelltimes=[True_steps(2:end-1,1),True_steps(2:end-1,5), True_steps(3:end,5)]
alldwell=vertcat(alldwell,realdwelltimes)

%Save the molecules TODP/TDP peak coordinates Einitial/Efinal
%To extract the TODP coordinates, you have to multipy by 1000, round, do the unique command, and then divide by 1000
%because somehow the unique command does not work on rows with decimal numbers
transitionpairs=[True_steps(1:end-1,5),True_steps(2:end,5)]
TDPlist=vertcat(TDPlist,transitionpairs)
transitionpairs=round([True_steps(1:end-1,5),True_steps(2:end,5)].*1000)
unique_transitionpairs=unique(transitionpairs, 'rows')./1000
TODPlist=vertcat(TODPlist,unique_transitionpairs)

else %the molecule has just 1 transition and no actual dwelltime

% Save the molecules TODP/TDP peak coordinate (s.a.)
transitionpairs=[True_steps(1:end-1,5),True_steps(2:end,5)]
TDPlist=vertcat(TDPlist,transitionpairs)
transitionpairs=round([True_steps(1:end-1,5),True_steps(2:end,5)].*1000)
unique_transitionpairs=unique(transitionpairs, 'rows')./1000
TODPlist=vertcat(TODPlist,unique_transitionpairs)

end

% Save lifetime in allliftimes
lifetime = trace_end/rate_fr
alllifetimes = vertcat(alllifetimes,lifetime)


else % it is static
TODPstat=TODPstat+1   %count it as static

h(j)=figure;

subplot(24,1,2:6) % ID, IA
plot(frametime, data_e(:,2),'g', frametime, data_e(:,3),'r' )
ylabel('Intensity')
title(filename_e)

subplot(24,1,8:12) % IA/IAmax, Afit
plot(frametime, data_a(:,4),'k', frametime, data_a(:,5),'g' )
ylabel('Intensity')
```

```
subplot(24,1,14:18) % ID/IDmax, Dfit
plot(frametime, data_d(:,4),'k', frametime, data_d(:,5),'g' )
ylabel('Intensity')

subplot(24,1,20:24) % Efret, Efit, Efit_corrected with accepted transition points trace_trans and actual transition points true_trans
plot(frametime, data_e(:,4),'k', frametime, data_e(:,5),'g', frametime, Efitcorr,'m', trace_trans(1:end-1)./10, Trace_steps(1:end-1,5), 'ob',
true_trans(1:end-1), True_steps(1:end-1,5), '*b')
xlabel('Time/s')
ylabel('E-FRET')

set(gcf,'PaperPosition', [1 1 12.5 20]) % size of plot-output

saveas(h(j), strcat(filename_e, 'stat.jpg'))

close(h(j))


% Save the molecules TODP peak coordinate: Include its E-FRET as diagonal peak in TODP

DPstat=[Trace_steps(1,5) Trace_steps(1,5)]
TODPlist = vertcat(TODPlist, DPstat)


% Save lifetime in allliftimes
lifetime = trace_end/rate_fr
alllifetimes= vertcat(alllifetimes,lifetime)


%this part is construct specific: here I separately count the number and
%lifetimes of static molecules in the high FRET state (E>0.75)

if Trace_steps(1,5) > 0.75 %then it is in the highFRET state
TODPhistat=TODPhistat+1 %count it as static and highFRET static
histatlifetimes = vertcat(histatlifetimes,lifetime)  %save lifetime in the separate list for highFRET static lifetimes

movefile(strcat(filename_e, 'stat.jpg'), 'histat_jpg')


end

end

end
end

%Make TODP plot:
%Plot a 2D Gauss function with amplitude 1 and sigma^2=0.002 at the coordinates (Efitbefore/Efitafter) for each transition FRET pair
Efitbefore/Efitafter found per molecule
%(Dynamic molecules: 1 Peak at E1/E2 or two peaks at E1/E2 and E2/E1; Static molecules: 1 Peak at E1/E1)
% Over all traces, this basically counts how many of the molecules exhibit a transition between a specific FRET pair or no transition
%(Example: In 3 molecules 2 exhibit transitions between E1/E2, 1 between E2/E1 and 1 is static and hence remains at E1/E1)

%This is as described for TDPs by Sean A. McKinney, Chirlmin Joo and Taekjip Ha in "Analysis of Single-molecule FRET
Trajectories Using Hidden Markov Modeling"
%Biophys J BioFAST, published on June 9, 2006 as doi:10.1529/biophysj.106.082487 Copyright 2006
%But each transition is counted once per trace only!

% Normalize by the number of molecules in order to "scale transitions based on the fraction of molecules that exhibit them at least
once"
% as described by Mario Blanco and Nils Walter in "Analysis of Complex Single Molecule FRET Time Trajectories"
% Methods Enzymol. 2010; 472: 153–178. doi:  10.1016/S0076-6879(10)72011-5


x = [0: 0.005: 1];
y = [0: 0.005: 1];
z = zeros(length(x), length(y)); % empty matrix init

[x_,y_] = meshgrid(x,y);

for k = 1:size(TODPlist,1)
z = z + exp((-(x_-TODPlist(k,1)).^2-(y_-TODPlist(k,2)).^2)./0.002) %each TODP entry is a Gaussian with amplitude 1 and
sigmasquare 0.002
end
```

```
Molecules=TODPdyn+TODPstat

z=z./Molecules %normalize amplitudes by the number of molecules

%now plot the TODP with the fine grid

h(j)=figure
surf(x,y,z,'LineStyle','none', 'FaceLighting','phong')
view(2)
colormap(jet) % heat map
%colorbar()
caxis([0 0.15])
colorbar('YTick',[0,0.03,0.06,0.09,0.12,0.15])
axis equal
axis([0 1 0 1])
xlabel('initial FRET')
ylabel('final FRET')
set(gca,'XTick', [0 0.2000 0.4000 0.6000 0.8000 1])
set(gca,'YTick', [0 0.2000 0.4000 0.6000 0.8000 1])
set(gca, 'FontWeight', 'bold', 'FontName', 'Arial', 'FontSize',14)
%h=get(gca)
saveas(gcf,'TODP.tif');
close(h(j))



%Make TDP plot:
%Plot a 2D Gauss function with amplitude 1 and sigma^2=0.002 at the coordinate (Efitbefore/Efitafter) for each transition FRET pair Efitbefore/Efitafter found over all traces
%Normalize by the overall number of transitions found

x = [0: 0.005: 1];
y = [0: 0.005: 1];
z = zeros(length(x), length(y)); % empty matrix init

[x_,y_] = meshgrid(x,y);

for k = 1:size(TDPlist,1)
z = z + exp((-(x_-TDPlist(k,1)).^2-(y_-TDPlist(k,2)).^2)./0.002) %each TDP entry is a Gaussian with amplitude 1 and sigmasquare 0.002
end

NoTransitions=size(TDPlist,1)

z=z./NoTransitions %normalize amplitudes by the number of transitions

%now plot the TDP with the fine grid

h(j)=figure
surf(x,y,z,'LineStyle','none', 'FaceLighting','phong')
view(2)
colormap(jet) % heat map
%colorbar()
caxis([0 0.1])
colorbar('YTick',[0,0.02,0.04,0.06,0.08,0.1])
axis equal
axis([0 1 0 1])
xlabel('initial FRET')
ylabel('final FRET')
set(gca,'XTick', [0 0.2000 0.4000 0.6000 0.8000 1])
set(gca,'YTick', [0 0.2000 0.4000 0.6000 0.8000 1])
set(gca, 'FontWeight', 'bold', 'FontName', 'Arial', 'FontSize',14)
%h=get(gca)
saveas(gcf,'TDP.tif');
close(h(j))


%calculate and export statistics file

aveliftime=mean(alllifetimes)
aveliftimehistat=mean(histatlifetimes)

header={'dyn_molecules', 'stat_molecules', 'histat_molecules', 'ave_lifetime','ave_histatlifetime',  'accepted_molecules', 'input_molecules'}
statistics = dataset({[TODPdyn,TODPstat,TODPhistat,aveliftime,aveliftimehistat, Molecules, numel(files_a)],header{:}})
```

```
export(statistics)

%EXPORT COLLECTED DWELLTIMES

file1='dwelltimes.txt';
save (file1, 'alldwell', '-ascii');


movefile('*dyn.jpg', 'dyn_jpg')
movefile('*stat.jpg', 'lostat_jpg')
```

Publications

S. Warhaut, K.R. Mertinkus, P. Höllthaler, B. Fürtig, M. Heilemann, M. Hengesbach and H. Schwalbe (2017) 'Ligand-modulated folding of the full-length adenine riboswitch probed by NMR and single-molecule FRET spectroscopy', *Nucleic Acids Res*, doi: 10.1093/nar/gkx110

Conference contributions

S. Warhaut, K.R. Mertinkus, P. Höllthaler, B. Fürtig, M. Heilemann, M. Hengesbach and H. Schwalbe, 'Ligand-Directed Conformational Dynamics of the Adenine-Sensing Riboswitch Thermostat', *Poster presentation*, 61st Annual Meeting of the Biophysical Society, New Orleans, USA, February 11-15, 2017, doi: 10.1016/j.bpj.2016.11.1996

S. Warhaut, 'Two perspectives of functional RNA dynamics: NMR and single-molecule FRET spectroscopy of the adenine-sensing riboswitch thermostat', *Oral presentation*, SFB 902 Conference RNA Structural Biology, Bad Homburg, Germany, November 28-29, 2016

S. Warhaut, 'NMR and Single-Molecule-FRET spectroscopy of the translational adenine riboswitch', *Oral presentation*, Workshop 8th BMRZ symposium, Bad Homburg, Germany, June 21, 2016

S. Warhaut, B. Fürtig, M. Hengesbach, P. Höllthaler, M. Heilemann and H. Schwalbe, 'Structural dynamics of a full-length adenine riboswitch', *Poster presentation*, GDCh FGMR 37th Annual Meeting, Darmstadt, Germany, September 7-10, 2015

S. Warhaut, B. Fürtig, M. Hengesbach, P. Höllthaler, M. Heilemann and H. Schwalbe, 'Structural dynamics of a full-length adenine riboswitch', *Poster presentation*, GDCh JCF 17th Frühjahrssymposium – spring symposium, Münster, Germany, March 25-28, 2015