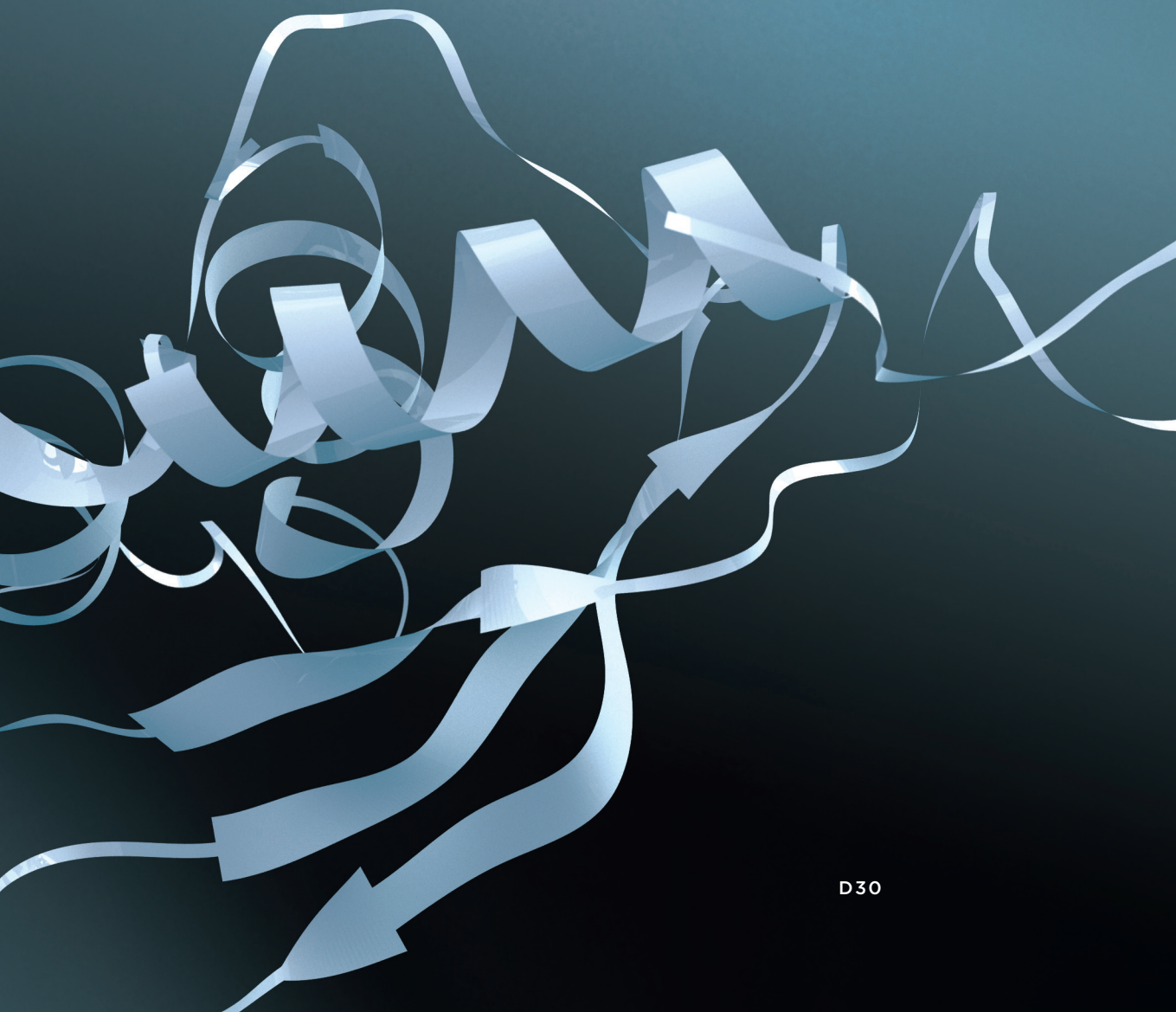
STEFFEN KASPAR GRIMM

Studies on protein:RNA complexes

A role for ribosomal binding factor A



D30

Studies on protein:RNA complexes

A role for ribosomal binding factor A

Dissertationsschrift
zur Erlangung des Doktorgrades
der Naturwissenschaften

— ❧ —

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

— ❧ —

von
Steffen Kaspar Grimm
aus Offenbach am Main

Frankfurt am Main
2007

Vom Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Harald Schwalbe

Gutachter: Prof. Dr. Jens Wöhnert und Prof. Dr. Harald Schwalbe

Datum der Dissputation: .....

Das ist wie bei jeder Wissenschaft,
am Schluss stellt sich dann heraus,
dass alles ganz anders war.

Karl Valentin

Contents

Figures V
Tables VII
Abbreviations VIII

Chapter 1 – Introduction 1
Abstract 2
1.1 Cold Shock Proteins 2
1.2 The Ribosome 4
1.3 The 16S rRNA 11
1.4 Accessory Factors to Ribosomal Assembly 17
Overview of this Thesis 22

Chapter 2 – Experimental Procedures 23
2.1 Cloning and Expression of RbfA in *E. coli* 24
2.2 Protein Purification 26
2.3 NMR Spectroscopic Techniques 28
2.4 Structure Determination and Refinement 33
2.5 CD Spectroscopy 35
2.6 Steady state Fluorescence Measurement 36
2.7 Calculation of Relative Contact Order 37
2.8 RNA 37
2.9 BioMagRes Database Accession Number 40

Chapter 3 – Results and Discussion 41
3.1 Choice of Model Organism 42
3.2 The Protein RbfA from *Thermotoga maritima* 45
3.3 The Protein RbfA from *Helicobacter pylori* 73
3.4 Comparing the RbfA Proteins 85
3.5 RNA Binding of RbfA 97

Conclusions 126
Appendix 130
References 159
Zusammenfassung 181
Danksagung 185
List of Publications 188
Curriculum Vitæ 189

Figures

1.1	The *E. coli* ribosome	5
1.2	Hybrid model of the ribosome elongation phase	7
1.3	The 30S subunit in vitro assembly map	8
1.4	Proposed Structure of the 16S rRNA precursor sequence	11
1.5	Secondary structure of the 16S rRNA from *E. coli*	13
1.6	Structure of RNA bases	14
1.7	Mutations in the helix I region	16
1.8	C23U mutation in the helix I region	20
3.1	Sequence alignment of 10 RbfA family members	42
3.2	Sequence alignment of *T. maritima* and *H. pylori* RbfA	43
3.3	SDS-PAGE for expression levels of *T. maritima* RbfA	45
3.4	Chromatogram of *T. maritima* RbfA on a cation exchange column	46
3.5	Chromatogram of *T. maritima* RbfA on a size exclusion column	47
3.6	^{1}H-^{15}N-HSQC spectrum of *T. maritima* RbfA	49
3.7	Strips from HNCA spectrum of *T. maritima*	51
3.8	Strips from HBHACONH spectrum of *T. maritima*	52
3.9	Strips from HCCH-TOCSY spectrum of *T. maritima*	53
3.10	Chemical shift indices of Cα and Hα atoms	55
3.11	Hydrogen/deuterium experiment	56
3.12	Strips from ^{15}N NOESY-HSQC spectrum of *T. maritima*	57
3.13	Plot of NOE contacts in the *T. maritima* RbfA β-sheets	59
3.14	^{3}J (HNHA) coupling constants of *T. maritima*	61
3.15	Hα-Cα residual dipolar couplings	62
3.16	Plot of the RDCs onto the final solution structure	63
3.17	Plot of the average backbone RMSD	64
3.18	Structural ensemble of the 10 lowest energy structures	65
3.19	Representation of the solution structure of *T. maritima* RbfA	68
3.20	Relaxation parameters of *T. maritima* RbfA	70
3.21	KH domain folding topologies	72
3.22	Chromatogram of *H. pylori* RbfA on a cation exchange column	74
3.23	^{1}H-^{15}N-HSQC spectrum of *H. pylori* RbfA	75
3.24	Chemical shift indices of Cα and Hα atoms	76
3.25	^{3}J (HNHA) coupling constants of *H. pylori*	77
3.26	Relaxation parameters of *H. pylori*	78
3.27	Comparison of chemical shift indices	80
3.28	Comparison of relaxation parameters	82
3.29	Overlay of solution structures from RbfA proteins	86

3.30	Charge distribution on RbfA proteins	87
3.31	Thermal unfolding CD spectra of *T. maritima* and *H. pylori* RbfA	89
3.32	Sequence alignment of RbfA proteins charged residues	91
3.33	Unfolding of *H. pylori* RbfA at various ion concentrations	92
3.34	Plot of relative contact order for RbfA proteins	94
3.35	Secondary structure of RNA constructs I–III	98
3.36	CD-spectra of Mg^{2+}-titrations of RNA I	99
3.37	HSQC spectra of *H. pylori* RbfA upon titration with RNA I	100
3.38	Plot of the composite chemical shift changes	101
3.39	Cartoon of the composite chemical shift changes	101
3.40	Fluorescence changes upon titration of *H. pylori* RbfA	102
3.41	Curve fitting of fluorescence data for titrations with RNA I–III	103
3.42	Curve fitting of fluorescence data for titrations with RNA IV	104
3.43	NOESY spectrum for RNA IV	106
3.44	Imino proton spectra of RNA IV upon titration with RbfA	107
3.45	Secondary structures of RNA V–VI	109
3.46	Curve fitting of fluorescence data for RNA V and VI	109
3.47	Curve fitting of fluorescence data for RNA VII–IX	110
3.48	Curve fitting of fluorescence data for titrations with RNA X	110
3.49	NOESY spectrum for RNA X	112
3.50	Imino proton spectra of RNA X upon titration with RbfA	113
3.51	Fold of the pre-16S rRNA sequences	114
3.52	Secondary structures of pre 16S rRNA constructs	115
3.53	Native gel analysis of interaction of RbfA and RNA XI	115
3.54	Curve fitting of fluorescence data for RNA XI and XII	116
3.55	Secondary structures of RNA XI–XIV	117
3.56	Curve fitting of fluorescence data for RNA XIII–XV	117
3.57	Curve fitting of fluorescence data for RNA XVI	118
3.58	Imino proton spectra of RNA XVI upon titration with Mg^{2+}	119
3.59	Imino proton spectra of RNA XVI upon titration with RbfA	120
3.60	HSQC spectra of *H. pylori* RbfA upon titration with RNA XVI	121
4.1	Folding of the precursor 16S rRNA	127

Tables

1.1	Modifications in the 16S rRNA from *E. coli*	15
1.2	Mutations in the 16S rRNA helix I region	17
2.1	Gradient protocol for purification of RbfA	26
2.2	Molecular weights of labeled *T. maritima* RbfA	27
2.3	Molecular weights of labeled *H. pylori* RbfA	27
2.4	NMR buffer for the RbfA proteins	27
2.5	Extinction coefficients of the RbfA protein	27
2.6	Extinction coefficients of the RNA constructs	38
3.1	Parameters and restraints used in the structural calculation	66
3.2	Change of CD-signal at various ion concentrations	93
3.3	Relative contact orders	95
3.4	Binding affinities of the RNA constructs	123
A.1	Acquisition and processing parameters for *T. maritima* RbfA	130
A.2	Acquisition and processing parameters for *H. pylori* RbfA	132
A.3	Acquisition and processing parameters for RNA constructs	133
A.4	Chemical shift indices for *H. pylori* RbfA	134
A.5	Chemical shift indices for *T. maritima* RbfA	138
A.6	Chemical shift indices for RNA IV and RNA VI	146
A.7	Experimental restraints for *T. maritima* RbfA	147
A.8	Relaxation parameter for *T. maritima* RbfA	151
A.9	Experimental restraints for *H. pylori* RbfA	155

Abbreviations

AA	Amino acid
Å	Angstroem [1 Å = 10^{-10} m]
ATP	adenine triphosphate
a. u.	arbitrary units
CD	circular dichroism
CO	contact order
COSY	correlated spectroscopy
CSD	cold shock domain
CSP	cold shock protein
CT	constant time
$\Delta\sigma$	composite chemical shift difference
DEAD box	amino acid sequence (Asp-Glu-Ala-Asp)
DMSO	dimethyl-sulfoide
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
ssDNA	single stranded DNA
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GDPNP	non-hydrolysable GTP analog
GTP	guanine triphosphate
HSQC	heteronuclear single quantum coherence
Hz	Hertz [1/s]
INEPT	insensitive nuclei enhanced by polarization transfer
IPAP	in-phase anti-phase
IPTG	Isopropyl-β-D-1-thiogalactopyranosid
kDa	kilo Dalton [1 Da = 1.66×10^{-27} kg]
LB	Luria Bertani media
MALDI	matrix assisted laser desorption/ionization
MS	mass spectoscopy
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser and exchange spectroscopy
OB	oligonucleotide/oligosaccharide binding fold
OD	optical density
PAGE	polyacrylamide gelelectrophoresis
PCR	polymerase chain reaction
ppm	parts per million
RDC	residual dipolar coupling

RMSD	root mean square deviation
RNA	ribonucleic acid
dsRNA	double stranded RNA
mRNA	messenger RNA
rRNA	ribosomal RNA
ssRNA	single stranded RNA
tRNA	transfer RNA
SDS	sodium dodecylsulfate
TOCSY	total correlation spectroscopy
UV	ultra violet
WATERGATE	Water suppression by gradient-tailored excitation
vdW	van der Waals

CHAPTER 1

Introduction

Abstract

Ribosome binding factor A (RbfA) is a bacterial cold-shock protein that associates with 30S ribosomal subunits but not with 70S ribosomes. It is required for the efficient processing of 16S ribosomal RNA. Here, we describe the three-dimensional structure of *Thermotoga maritima* RbfA, a 120 residue construct with 11 residues removed from the carboxyl terminus of full-length RbfA. The solution structure was determined at pH 4.5 by heteronuclear NMR methods. Using the same NMR methods, the backbone fold of the full length RbfA protein from *Helicobacter pylori* was determined. The structure revealed a type II KH-domain fold, which is related to conserved KH sequence family proteins whose βααβ subunits are characterized by a helix-turn-helix motif with the sequence signature GXXG at the turn. In RbfA, this βααβ subunit is characterized by a helix-kink-helix motif in which the GXXG sequence is replaced by a conserved AXG sequence. Dynamic hot spots of RbfA have been identified by R_1, R_2 relaxation rates and heteronuclear ^{15}N-{^{1}H}-Het-NOE measurements. The analysis of the distributions of charged amino acids and the internal dynamics together with the fold topology and patterns of sequence conservation suggested that RbfA is a nucleic acid-binding protein.

Previous genetic evidence pointed towards an interaction of RbfA with helix I of 16S rRNA. In this thesis, we demonstrate that the RbfA protein indeed is an RNA-binding protein and map the RNA-binding site of RbfA to a conserved surface of the protein by using chemical shift perturbation NMR experiments. Fluorescence titration experiments using a number of RNA substrates mimicking either helix I, the central pseudoknot region of 16S rRNA or regions of the pre-16S rRNA showed that RbfA preferentially binds to pre-16S rRNAs.

1.1 Cold shock proteins

All living organisms must adapt to changes in the environment. These adaptation processes are essential for their survival as dramatic changes such as heat shock, cold shock, acid shock, pressure and osmotic stress are lethal to most organisms (Thieringer et al., 1998). Therefore, organisms have developed various mechanisms to overcome such drastic changes in their environment. In the case of cold shock, many bacteria respond with the

reorganization of their entire protein expression pattern. In general, the protein synthesis of bulk proteins becomes inhibited whereas the expression of some proteins increases rapidly (Thieringer et al., 1998). Proteins with an increased expression level upon a cold shock were called cold shock proteins (CSPs). In bacteria, a cold shock is defined as a decrease in the surrounding temperature of 10–15° C relative to the optimal growth temperature. A cold shock causes two major problems: the fluidity of the cell membranes is strongly decreased (Fulco, 1974; Wunderlich et al., 1973) and the folding of DNA and RNA molecules is altered and non-native secondary structures are stabilized (Lopez et al., 2001; Jiang et al., 1996; Xia et al., 2002). To counteract these obstacles, the cell uses different mechanisms. The fluidity of the membrane is maintained by an increase in the fraction of unsaturated fatty acids. This is achieved through the action of a group of proteins that are called fatty acid desaturases (Fukushima et al., 1976; Kasai et al., 1976; Martin et al., 1976). To ensure the correct folding of DNA and RNA molecules, a number of cold shock proteins are overexpressed (Giuliodori et al., 2004). The cold shock proteins in *E. coli* are divided into two classes. Class I proteins are expressed in very low quantities at 37° C but are overexpressed at least ten fold when the cell is exposed to a cold shock. Class II proteins are already abundant at a certain level at the optimal growth temperature. When a decrease in temperature occurs, the concentration of these proteins in the cell is increased. To facilitate protein expression at low temperatures, some of the cold shock proteins possess a so-called cold box sequence, which is located in the 5' untranslated region of the mRNA (Jiang et al., 1996; Xia et al., 2002). This cold box sequence, for example from the cold shock protein A (CspA), is highly unstable at 37° C. After the temperature is lowered to 20° C, the secondary structure of the mRNA is stabilized by a factor of 100 (Xia et al., 2002). This RNA has a stem-loop structure followed by an AU rich sequence (Xia et al., 2002). It is not only highly stable under cold shock conditions; it also displays a high affinity to the ribosome. This leads to actual binding by the ribosome and initiation of the protein synthesis, despite the fact that the initiation codon as well as the downstream box is masked by non-native secondary structures after the cold shock (Xia et al., 2002). All proteins of the CSP class I family have this cold box sequence.

Many of the primary cold shock proteins are believed to be DNA and RNA binding proteins that help in folding of DNA and RNA molecules. This led to the idea that these cold shock proteins were actual DNA and RNA chaperones (Jiang et al., 1997; Brandi et al., 1999; Gualerzi et al., 2003). For example, the CspA protein seems to have RNA helicase activity

(Jiang et al., 1997), while the CspB protein binds to specific DNA sequences; a pentamer sequence (CCAAT) with an increased affinity for ssDNA (Lopez et al., 2001). CspB functions as a transcription activator for secondary cold shock proteins that have an ATTGG-box sequence in the promoter region of the cold shock genes (Lopez et al., 2001). Other cold shock proteins are known or suspected to bind rRNA. They seem to interact with the rRNA and ensure the proper functioning of the ribosomal machinery under cold shock conditions.

1.2 The Ribosome

The ribosome plays an essential role in every living organism or cell. It is the site of protein biosynthesis (translation). The ribosome itself is a large ribonucleoprotein complex consisting of two different subunits. Although the size and the individual components of a fully functional ribosome differ from organism to organism, the key structural features and functions are conserved throughout all living organisms (Woese et al., 1983; Hill et al., 1990). Part of the fully functional ribosome are the translatable mRNA molecules, tRNA molecules charged with their cognate amino acid and a number of initiation, elongation and termination factors.

In general, it is believed that the appearance of the ribosome is an essential turning point in early evolution (Ramakrishna and White, 1998). The prokaryote ribosome, namely that from *Escherichia coli,* is one of the best studied systems to date. The *E. coli* ribosome consists of two subunits named 30S and 50S based on their sedimentation coefficients. The 30S subunit consists of the 16S rRNA, 1,542 nucleotides in length, and 21 ribosomal proteins (S1–S21), while the 50S subunit consists of the 5S and 23S rRNAs, 120 (or 121) and 2,904 nucleotides in length respectively, as well as 33 ribosomal-proteins (L1–L6, L7/L12, L9–L11, L13–L25, L27–L36) (Zengel and Lindahl, 1994) With the exception of the L7/L12 heterodimer, the ribosome contains only single copies of each of the proteins. The 16S and 23S rRNAs and several of the ribosomal proteins are chemically modified by special enzymes. Together the 30S and 50S subunits form, after initiation of translation, the 70S ribosome or 70S polysomes (multiple ribosomes attached to a single mRNA).

Over the past decades huge efforts were made to fully understand the complex structure and function of the ribosome. These efforts included the

Figure 1.1
The *E. coli* ribosome. The 50S subunit is shown left (pdb file 1PNY). The 30S subunit is shown on the right (pdb file 1PNX). The ribosomal RNAs, 23S rRNA in the large ribosomal subunit and the 16S rRNA in the small ribosomal subunit, respectively, are colored orange, while the ribosomal proteins are colored brown.

sequencing of the involved rRNAs and ribosomal proteins (Brosius et al., 1978); the *in vitro* reconstitution of ribosomal subunits and the ribosome (e.g. Nomura, 1973; Held et al., 1974; Rohl and Nierhaus, 1982); models of the secondary structure of the rRNAs (Fox and Woese, 1975); determination of protein-protein (Yeh et al., 1986), protein-RNA and RNA-RNA interactions by bifunctional crosslinking (Brimacombe, 1995); the localization of proteins by immuno electron microscopy (Stoffler and Stoffler-Meilicke, 1987), neutron scattering (Moore et al., 1975; Langer et al., 1978), chemical modification (Noller, 1991) and the identification and characterization of the different functional phase cycles and groups involved in protein synthesis (Brimacombe, 1995; Noller, 1991; Stark et al., 1995). In recent years, several methods to elucidate the structure of the ribosome have been employed. Cryo electron microscopy and X-ray crystallography have been used to identify the structure of ribosomal components, whole subunits or complete 70S ribosomes. The highest resolution has been achieved by X-ray crystallography and a structural model of the ribosome was published, when high resolution X-ray structures of the two subunits and the complete ribosome were solved (Wimberly et al., 2000; Schluenzen et al., 2000; Cate, 1999; Agalamrov, 2000; Nikulin, 2000; Noller et al., 2001; Yusupov, 2001). However, before the complete structure was solved, a number of structures from individual

ribosomal components such as ribosomal proteins, rRNA fragments and protein factors were solved by X-ray crystallography and high resolution nuclear magnetic resonance (NMR) techniques. The first structure to be solved—a fragment of the L7/L12 heterodimer—was published in 1980 (Leijonmarck et al., 1980), the second one six years later (Wilson et al., 1986). Since then a large number of substructures have been published. These structures of ribosomal components were instrumental in assigning the initial electron density maps of the ribosomal subunits and the complete ribosome.

The structural data available for the ribosome and its subunits and individual components led to a better understanding of the functional features of the ribosome. This included the central feature of the ribosome, the protein synthesis, during which mRNAs possessing genetic information are translated into new polypeptides.

The translation process in *E. coli* cells can be divided into three parts: initiation, elongation and termination.

Initiation. During initiation of translation, the 30S subunit, initiator tRNA and mRNA form a complex with the help of several initiation factors (IF 1, IF 2, IF 3). The binding of the mRNA to the 30S ribosomal subunit is mediated by the ribosomal protein S1 and the anti-Shine-Dalgarno sequence on the 16S rRNA. The anti-Shine-Dalgarno sequence then forms basepairs with the Shine-Dalgarno sequence on the mRNA, which is located at the 5' end preceding the coding region (Shine and Dalgarno, 1974: Boni et al., 1991). The initiator tRNA (formylmethionine) binds to the peptidyl-site (P-site) of the 30S subunit and interacts via codon-anticodon basepairing with the start codon of the mRNA. IF 2 facilitates the binding of the initiator tRNA to the P-site and blocks the aminoacyl-site (A-site) (Carter et al., 2001). IF3 keeps the two ribosomal subunits apart and also inhibits initiation with non-initiator tRNAs (Hartz et al., 1990). The function of IF1 is less well characterized but seems to improve the function of the other initiation factors. After the primary initiation complex is formed, the 50S subunit binds to the initiation complex and the elongation can proceed (Gualerzi and Pon, 1990).

Elongation. During the elongation phase, ternary complexes consisting of elongation factor Tu (EF-Tu), aminoacylated tRNA and GTP, bind to the A-site so that the codon of the mRNA and the anticodon of the tRNA can interact. If this codon anticodon interaction is correct, the 30S subunit goes through a conformational change in which GTP is hydrolyzed and the EF-Tu factor is released (see Figure 1.2 B). After this a proofreading step occurs

and improper tRNAs are rejected. If the tRNA is not rejected the aminoacyl group of the tRNA is accommodated in the peptidyl transferase center of the 50S ribosomal subunit (see Figure 1.2 C).

At this point, the peptide on the P-site is transferred to the aminoacyl group on the A-site and the amino acid chain is elongated by one amino acid. The A-site tRNA will end up in a hybrid state, between the A-site of the 30S subunit and in the P-site of the 50S subunit (the A/P state). The P-site tRNA ends up in the P-site of the 30S subunit and the exit site (E) of the 50S subunit (the P/E state) (Moazed and Noller, 1989) (see Figure 1.2 D). After peptidyl transfer, the elongation factor G (EF-G) binds to, and induces a conformational change of the ribosome (Frank and Agrawal, 2000). EF-G hydrolyzes one GTP and translocation takes place; the deacylated tRNA in the P/E site shifts to the E-site and the peptidyl-tRNA in the A/P site shifts to the P-site (Rodnina et al., 1997). During the translocation, the mRNA is shifted by one codon and the next codon in the A-site will be available for interaction with the next tRNA (see Figure 1.2 E).

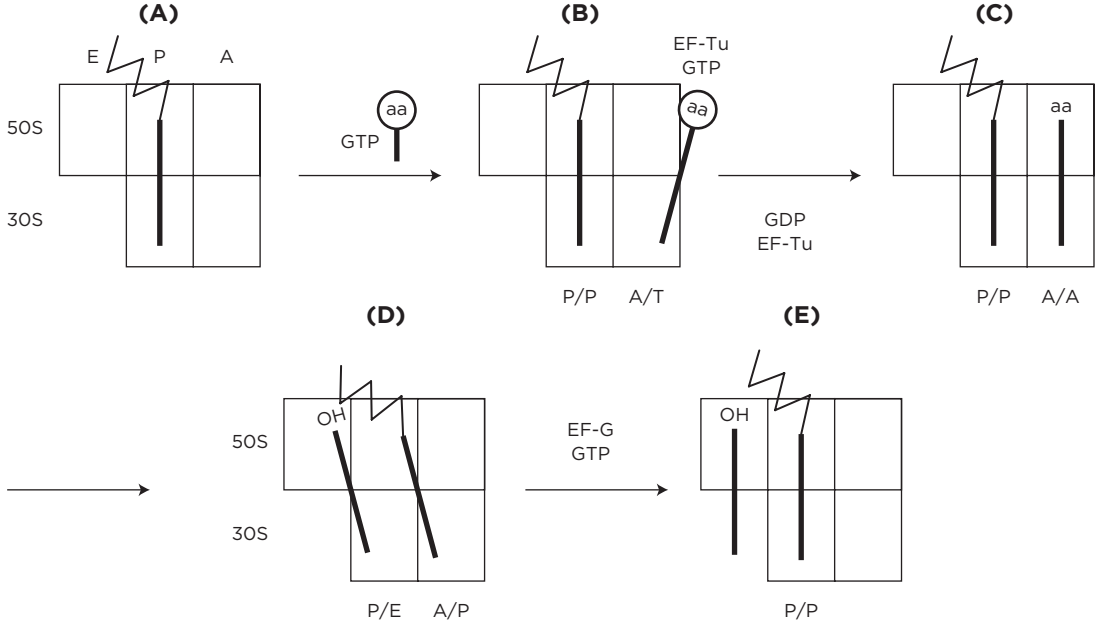


Figure 1.2
Hybrid states model of the elongation phase in the *E. coli* ribosome.

Termination. The termination of translation occurs when the ribosome encounters a stop codon, which is recognized by release factor 1 (RF1) or release factor 2 (RF2), where RF1 recognizes UAA and UAG stop codons and RF2 recognizes UAA and UGA stop codons. After that the peptide is

hydrolyzed from the P-site tRNA. RF3 then binds to the ribosome inducing the release of RF1 and RF2 after which RF3 hydrolyzes GTP to release itself (Kisselev and Buckingham, 2000).

The structural and dynamical organization of the 22 components (1 rRNA, 21 proteins) into a functional 30S subunit was the focus of much work. The recent high resolution analysis of the 70S ribosome, 30S subunit and 30S subdomains have expanded the understanding of the overall fold of the fully assembled ribosome. The discovery of catalytic RNAs (Cech, 1986) also changed the view of the ribosome being an actual ribozyme. The structure of the ribosome revealed that the ribosomal proteins are mere accessory factors

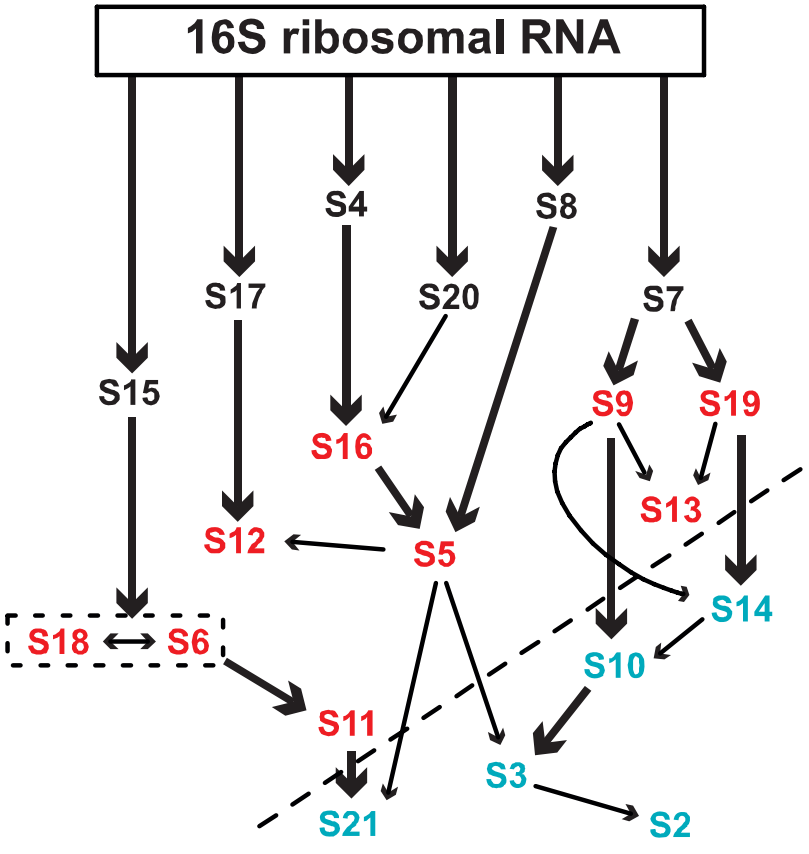


Figure 1.3
The 30S subunit *in vitro* assembly map with traditional protein designations. The 16S rRNA is represented as a rectangle. Arrows indicate interactions between components. The primary, secondary, and tertiary binding proteins are colored black, pink, and blue, respectively. The S6 and S18 are shown in a dashed line box to indicate that they bind as a heterodimer.

that aid or stabilize rRNA folding into functional conformations (Culver, 2003).

As mentioned above, one of the advances in the field of ribosome research was the reconstitution of the 30S ribosomal subunit. Most striking was the discovery that a functional 30S subunit could be reconstituted *in vitro* by using 16S rRNA and a mixture of the 30S ribosomal proteins (Traub and Nomura, 1968; Krzyzosiak et al., 1987; Culver and Noller, 1999). These experiments revealed that all the necessary information for appropriate 30S subunit assembly is contained within the molecular components of the subunit itself.

Furthermore, there were several combinatorial approaches, coupled with sequential addition of the proteins to the 16S rRNA under reconstitution conditions, which gave insight into how the 30S subunits assemble. This work facilitated the division of the small subunit ribosomal proteins into the three categories of primary, secondary and tertiary binding proteins (Mizushima et al., 1970; Held et al., 1973; Schaup et al., 1971; Held et al., 1974). The assembly process starts at the 5' end of the 16S rRNA and progresses towards the 3' end (Powers et al., 1993). The ribosomal protein binds to the nascent particle in an ordered fashion. While some ribosomal proteins can bind independently to the rRNA, others require certain other proteins to bind first.

The set of primary proteins (S4, S7, S8, S15, S17 and S20) can bind directly and independently to the 16S rRNA. The secondary binding proteins (S5, S6, S9, S11, S12, S13, S16, S18 and S19) require at least one of the primary proteins to be bound already to the 16S rRNA. Assembly of the tertiary binding proteins (S2, S3, S10, S14 and S21) requires at least one protein from both of the previous groups to be bound, to fully assemble a functional 30S subunit. These studies led to the determination of an *in vitro* 30S subunit assembly map (Held et al., 1974; Nierhaus, 1991; Grondek and Culver, 2004) (see Figure 1.3).

For reasons that will be discussed later I would now like to concentrate on the ribosomal protein S5. In the assembly map, the S5 protein is defined as a secondary protein (see Figure 1.3). The kinetic assembly studies (Powers et al., 1993) group the protein with the mid-late assembly proteins. The S5 protein is strongly dependent on the S8 protein for assembly and is also dependent on the proteins S4 and S16. In a study of a set of *E. coli* cold sensitive strains, which are defective in ribosome assembly (Guthrie et al., 1969), one cold sensitive strain accumulated an assembly intermediate containing 16S rRNA, which sediments at 21S. Responsible was a single mutation and this mutation was mapped to the coding region for ribosomal protein S5

(Nashimoto et al., 1971). The study of Nashimoto and colleagues suggested that S5 performs a critical role during ribosome assembly, perhaps participating in structural rearrangements of the assembling ribosome. In addition, the S5 protein has been implicated in various functions of the ribosome, most notably aminoacyl-tRNA selection (translational fidelity) and translocation (tRNA movement). Isolation of ribosomal ambiguity mutations (ram) in S5 demonstrated the influence of this protein on the accuracy of tRNA selection (Rosset and Gorini, 1969; Piepersberg et al., 1975a). Specific mutations in S5 also suppress streptomycin dependence (Stoffler et al., 1971; Kreider and Brownstein, 1972). Other mutations in S5 confer resistance to the translocation inhibitor spectinomycin (Bollen et al., 1969; Dekio and Takata, 1969; Bollen and Herzog, 1970; Piepersberg et al., 1975b). The presence of S5 on the 30S subunit is required for ribosome-dependent EF-G-catalyzed GTPase activity, again suggesting a link between S5 and translocation (Marsh and Parmeggiani, 1973). From its effects on tRNA selection and movement, one might expect to find S5 at the subunit interface, where the tRNAs are bound (Agrawal et al., 1996). However, contrary to this suggestion, both neutron diffraction and immunoelectron microscopy studies, in conjunction with the determination of the 30S subunit structure, placed the S5 protein on the solvent surface of the 30S subunit in the helix I region of the 16S rRNA (Leonard and Lake, 1979; Stoffler and Stoffler Meilicke, 1984; Oakes et al., 1986; Capel et al., 1987; Wimberly et al., 2000).

Although there is considerable data concerning the *in vitro* assembly of the ribosome, much less is known about the ribosome assembly *in vivo*. The process *in vitro* is slow and requires specific conditions. Additionally there is pronounced temperature dependence for the *in vitro* reconstitution of functional 30S subunits. The *in vivo* process is quite complex due to the fact that rRNA transcription, processing and modification as well as ribosomal protein binding and folding into the complex have to be well coordinated. In addition, a number of non-ribosomal proteins that seem to assist in the subunit formation have been identified (see chapter 1.4).

Precursor particles that are intermediates in the 30S and 50S subunit production have been identified (Mangiarotti et al., 1968, Hayes and Hayes, 1971). These particles are similar but not identical to the *in vitro* assembly intermediates. Formation of the 30S subunit proceeds via a 21S particle to an immature 30S particle and eventually into the mature 30S subunit (Lindahl, 1975). It has been suggested the 21S precursor particle contains nine ribosomal proteins (S1, S4, S5, S8, S13, S15, S16, S17 and S20) (Nierhaus et al., 1973).

1.3 The 16S rRNA

In bacteria, archaea, mitochondria, and chloroplasts the small ribosomal subunit contains the 16S rRNA. The bacterial 16S rRNA genes (from seven operons all containing 16S, 23S and 5S rRNA as well as one or more tRNAs) are typically organized as a co-transcribed operon. There may be one or more copies of the operon dispersed in the genome. All the rRNA genes are

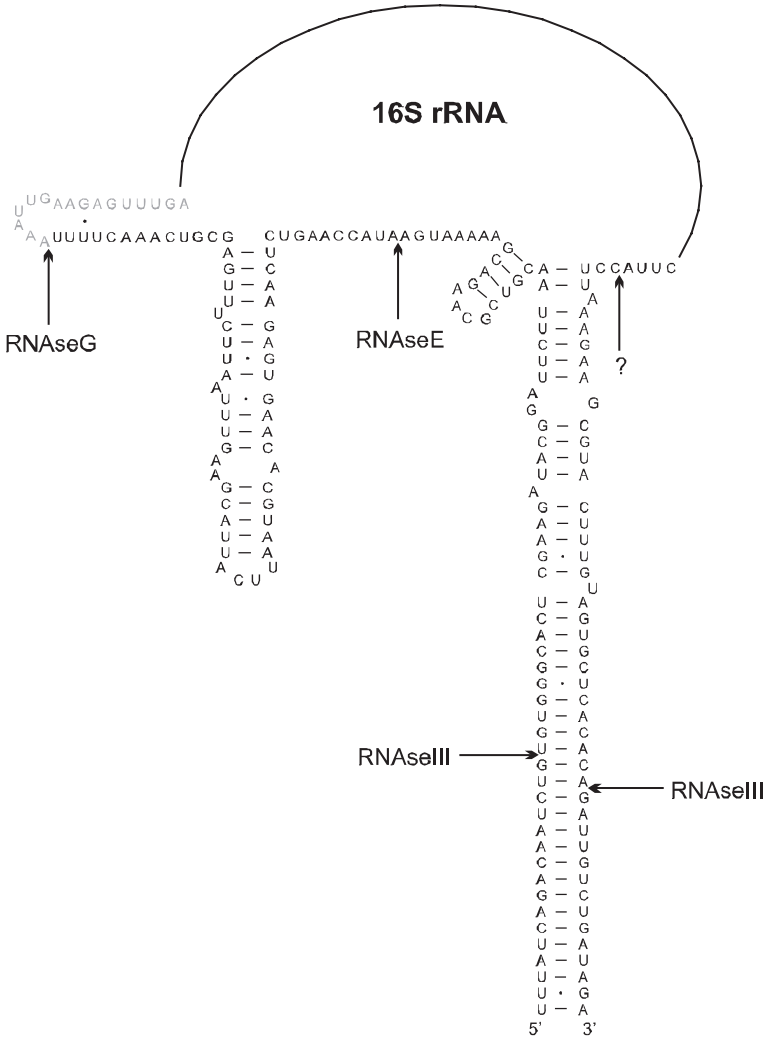


Figure 1.4
Proposed structure of the 16S rRNA precursor from the rrnB operon. Residues retained in the mature rRNA are shown in gray. However, most of the mature rRNA is shown schematically as a black line. The different RNAse cleavage sites are marked by an arrow. A question mark indicates the cleavage site for an unknown RNAse.

transcribed into one long transcript (the 30S rRNA), which is processed by different RNAses to produce the mature product. First RNAse III separates the precursors for 5S, 16S and 23S rRNA by cleaving at certain dsRNA structures (Nikolaev, 1973; Ginsburg and Steitz, 1975). For the 16S rRNA the cleavage at the 5' end requires that the 3' end has been transcribed (King and Schlessinger, 1983). The 16S rRNA precursor produced by RNAseIII is the 17S RNA, which has 115 extra nucleotides in the 5' end and 33 additional nucleotides in the 3' end (Young and Steitz, 1978). Ribosomes produced *in vitro* that contain the 17S precursor are deficient in translation (Wireman and Sypherd, 1974). After the RNaseIII cleavage, the RNA is further processed at the 5' end by RNaseE to a precursor, which is 66 nucleotides longer than the mature 16S rRNA (Dahlberg et al., 1978; Li et al., 1999). After this step, there is an immediate cleavage by RNAseG to yield the mature 5' end (Li et al., 1999b). However, both RNAseE and RNAseG can form mature 5' ends in absence of the other (Li et al., 1999).

The RNAse involved in the processing of the 3' end of the precursor 16S rRNA has not been identified, however, this processing step is less efficient if the 5' end has not been cleaved by RNAseE (Li et al., 1999). The cleavage of the 5' end is a late step in the assembly of the 30S subunit and 17S rRNA is found in particles co-sedimenting with mature 30S subunits. It has been suggested that the conversion of 17S to 16S rRNA is dependent on the functionality of the ribosomal subunit and that this processing step occurs when the newly formed subunit initiates translation (Mangiarotti et al., 1974).

The mature 16S rRNA makes up the bulk of the 30S subunit. It is important for subunit association and translational accuracy. It consists of 1,542 bases and contains the substrate binding A-, P-, and E-sites. The P-site is occupied by peptidyl-tRNA and is located in a major groove of the rRNA. The A-site is the attachment site for an incoming aminoacyl-tRNA, and is wide and shallow, which gives it a lower affinity for tRNA facilitating later translocation to the P-site. The E-site, which is occupied by deacylated tRNAs when they exit, is associated with more ribosomal proteins than the A- or P-site.

This might facilitate the dissociation of the codon-anticodon pair during translation (Wimberly et al., 2000). The primary structure of 16S rRNA is highly conserved. The arrangement of the 16S rRNA creates a 5‘ domain, a central domain, a 3‘ major domain and a 3‘ minor domain (Brimacombe et al., 1988; Carter et al., 2000; Wimberly et al., 2000). The 5‘ domain consists of 19 helices that make up the bulk of the body. It diagonally traverses the ribosome, when viewed from the 50S ribosomal subunit (Wimberly et al.,

2000). The central domain of the rRNA generates the platform and is an elongated, curved structure of nine helices, with the junction of helices 20, 21, and 22 located at the heart of it. The 3‘ major domain contains 15 helical elements and forms the head. The 3‘ minor domain contains 2 helices and projects from the subunit to interact with the 50S subunit.

The *E. coli* rRNAs have several modified nucleotides, 11 in the 16S rRNA (see Table 1.1), 25 in the 23S rRNA, but none in the 5S RNA.

The enzymes responsible for the different modifications are only known in half of the cases. The modifications are clustered in the functionally important regions (Brimacombe et al., 1988; Smith et al., 1992).

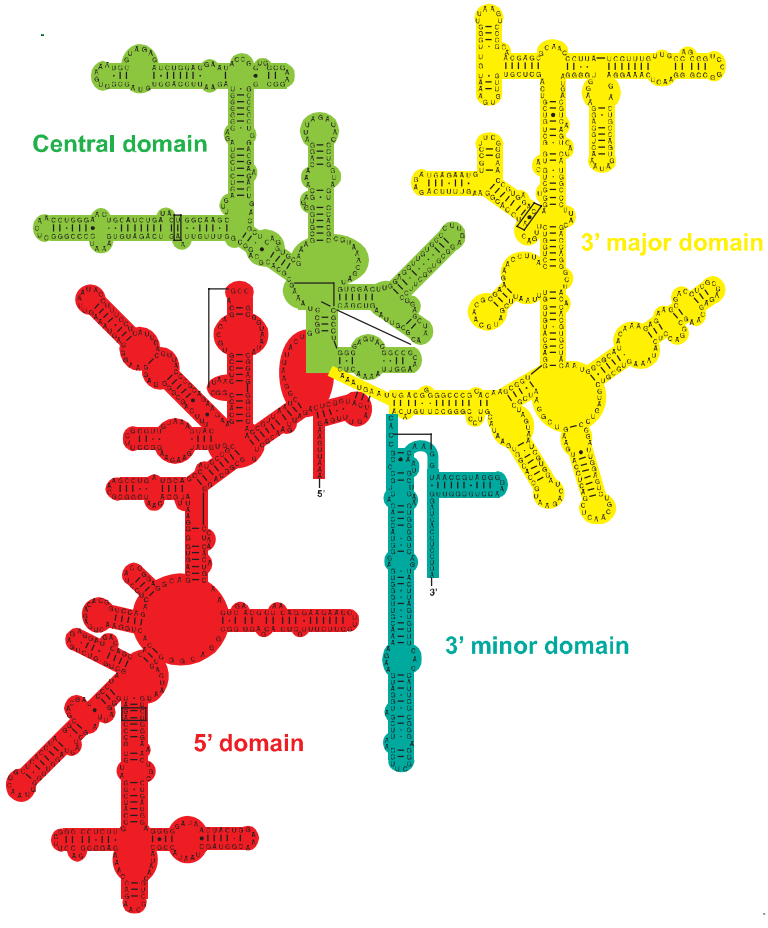


Figure 1.5
Secondary structure of the 16S rRNA from *Escherichia coli;* with its 5’, central, 3’ major and 3’ minor domain colored red, green, yellow and blue respectively.

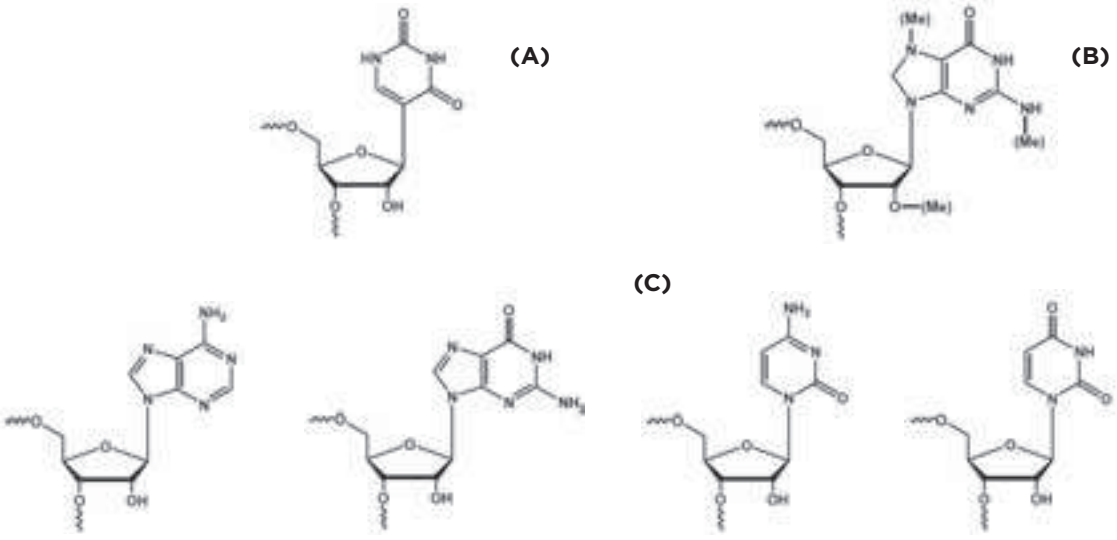


Figure 1.6
(A) Structure of pseudouridine and an example of a methylated guanine base **(B)**. **(C)** Structure of the unmodified RNA bases, from left to right: Adenine, guanine, cytosine and uracil.

Pseudouridines. Pseudouridines are synthesized by isomerization of uridines so that the uracil base is attached to the ribose by the C5 instead of the N1. There is one pseudouridine (Ψ) in the 16S rRNA (Del Campo et al., 2001). The pseudouridine synthetase responsible for the modification in the 16S rRNA has been identified (Del Campo et al., 2001), whereas pseudouridine synthetases responsible for isomerizations in the 23S rRNA remain to be identified.

Methylations. Methyl groups can be added to nucleotides both on the 2'-oxygen of the ribose and at several positions on the base. Only some of the genes for the methyltransferases have been identified; RsmB, which methylates C967 to m^5C (Tscherne et al., 1999), RmsC, which methylates G1207 to m^2G (Tscherne et al., 1999) and RsmA, which methylates both A1518 and A1519 to m^6A (van Buul and van Knippenberg, 1985). The modifications are made during the assembly and maturation of the ribosomal particles in such a way that all positions that are accessible to the modification enzymes are fully modified. The enzymes that can bind to free rRNA as a substrate are likely to be involved in early assembly steps, while others are involved in later steps. The leader sequence preceding and the internal spacers between the mature 16S rRNA sequence are important for ribosomal assembly.

Table 1.1
Modifications in the 16S rRNA of *E. coli*.

Modified nucleotide	Position	Gene for modifying enzyme	Reference
Ψ	516	rsuA	Wrzesinski et al., 1995
m^7G	527		
m^2G	966		
m^5C	967	rmsB	Tscherne et al., 1999a
m^2G	1207	rmsC	Tscherne et al., 1999b
m^4Cm	1402		
m^5C	1407		
m^3U	1498		
m^2G	1516		
$m^{6_2}A$	1518	rmsA	Helser et al., 1971; van Buul and van Knippenberg, 1985
$m^{6_2}A$	1519	rmsA	Helser et al., 1971; van Buul and van Knippenberg, 1985

Ψ pseudouridine, m^xN nucleotide N methylated at position x of the base,
Nm nucleotide N methylated at the 2'-O of the ribose.

Mutations in the 16S rRNA leader sequence that are present in the 17S precursor rRNA but not in mature 16S rRNA affect the production of functional 30S ribosomal subunits (see Table 1.2).

The 30S subunits produced with 16S rRNA that have altered leader sequences are largely non-functional in translation (Mori et al., 1990; Theissen et al., 1993). This is probably due to different folding of the rRNA (Balzer and Wagner, 1998; Besancon and Wagner, 1999). One feature of the 16S rRNA secondary structure first proposed in 1985 is the central pseudoknot. This pseudoknot structure involves interactions of nucleotides 17–19 and 916–918 (*E. coli* numbering) (Pleij et al., 1985). These base pairs connect nucleotides in the loop of the first stem-loop structure and nucleotides 916–918 that connect the central domain and the 3' major domain. The first stem-loop structure (helix I) in the 16S rRNA itself is formed by nucleotides 9–13 and 21–25 (*E. coli* numbering), at the 5' end of the mature 16S rRNA. The helix I region forms part of the binding site of ribosomal protein S5 in the mid-late

assembly step. It is located in the central part of the small ribosomal subunit. The functional importance of the central pseudoknot has been investigated by mutational analysis.

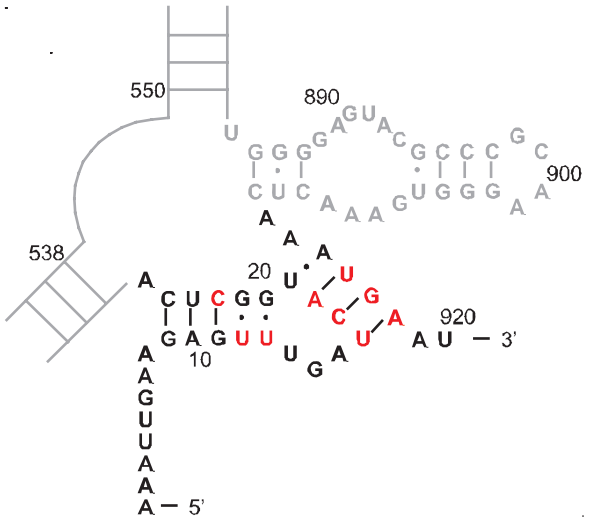


Figure 1.7
Secondary structure of *E. coli* 16S rRNA helix I region. Nucleotides in red denote nucleotides that upon mutation lead to defective ribosomes upon mutation.

Brink et al. (1993) mutated C18 to A, G, or U and observed inhibition of *in vivo* translation activity. The mutant 30S subunits were not able to form 70S ribosomal complexes or polysomes, so a defect in translation initiation was suspected (Brink et al., 1993). The mutation of the adjacent AU base pairs (U17C and A19C) also led to inhibition of translational activity (Poot et al., 1998).

The functional defect in the subunits with the disrupted central pseudoknot base pairs was thought to be in subunit stability, because those subunits were deficient in proteins S1, S2, S18, and S21 (Poot et al., 1998). Data from the mutational analysis led to the conclusion that the base complementarity rather than sequence or thermodynamic stability was important for a functional helix II (Poot et al., 1998). Further studies tested several other mutations in this central pseudoknot region (see Table 1.2 and Figure 1.7).

Table 1.2 Mutations in the 16S rRNA within the helix I region.

Mutation	Effect	Reference
G –59 A	Temperature sensitive	Mori et al., 1990
G –30 A	Temperature sensitive	Mori et al., 1990
G –13 A	Temperature sensitive	Mori et al., 1990
U 12 C	Suppressor of A900G mutation	Belanger et al., 2005
Δ U12	Suppressor of A900G mutation	Belanger et al., 2005
U 13 A	Suppressor of A900G mutation	Belanger et al., 2005
U 17 C	Inhibition of translation	Poot et al., 1996
C 18 A	Inhibition of translation	Brink et al., 1993
C 18 G	Inhibition of translation	Brink et al., 1993
C 18 U	Inhibition of translation	Brink et al., 1993
A 19 C	Inhibition of translation	Poot et al., 1996
C 23 U	Temperature sensitive	Dammel and Noller, 1995

1.4 Accessory factors to ribosomal assembly

As discussed above, purified bacterial ribosomal components are capable of assembling into functional 30S subunits *in vitro* in the absence of other factors. These observations suggested that this process is uncatalyzed and does not require any additional cellular factors or metabolic energy *in vivo* (Culver, 2003). A similar view was also widely held for protein folding until the discovery of molecular chaperones. Given some of the requirements of the *in vitro* reconstitution reactions, it seems that extra ribosomal or non-ribosomal auxiliary factors actively participate in ribosome assembly *in vivo*. Studies of eukaryotic ribosome assembly support this hypothesis (Steitz and Tycowski, 1995; Venema and Tollervey, 1999). Although, there is no direct evidence for such factors in *E. coli*, a few cellular non-ribosomal proteins have been implicated in the assembly process (Culver, 2003). Such auxiliary factors would assist in folding and assembly of the ribosomal subunits.

One group of proteins associated with ribosomal subunit assembly is comprised of the DEAD box RNA helicases. Three of this helicases SrmB, DbpA and CsdA have thus far been identified in *E. coli*. Deletion of one of these proteins gives a cold sensitive phenotype (Nashimoto, 1993; Jones et al., 1996; Charollais, 2003). The DEAD box helicases unwind double stranded RNA structures in an ATP dependent manner (Tanner and Linder, 2001). These proteins are believed to assist in ribosomal assembly by resolving secondary structures, which otherwise would inhibit proper folding of the rRNA. The CsdA protein was identified as a high copy suppressor of temperature sensitive mutations in the ribosomal protein S2 (Toone et al., 1991). CsdA is a cold shock protein that unwinds double stranded RNA in an ATP dependent manner (Jones et al., 1996). Several functions have been suggested, all of which are consistent with an RNA helicase activity: stabilization of mRNAs (Iost and Dreyfus, 1994; Brandi et al., 1999), degradation of mRNAs (Jones et al., 1996; Yamanaka and Inouye, 2001), enhancement of translational initiation and ribosomal maturation (Lu and Inouye, 1999).

Protein chaperones are also thought to be involved in the assembly of the ribosomal subunits since both DnaK and GroEL seem to participate in this process. Temperature sensitive DnaK mutants have decreased levels of 70S monosomes and accumulate ribosomal precursor particles at higher temperatures (Alix and Guerin, 1993). This suggests that the chaperones are more important at higher temperatures. The precursor particles are able to form mature particles and disappear when the temperature is decreased. The 21S precursor contains 17S rRNA but lacks ribosomal proteins S3, S10, S14 and S21 and contains substochiometric amounts of S1, S2 and S5 (El Hage and Alix, 2004). DnaK has a strong affinity for the ribosomal proteins S4, S12, S17 and S21 as well as weaker affinity for S3, S5, S8, S16 and S19 (Maki et al., 2002). The temperature sensitive DnaK mutation can be suppressed by overproduction of ribosomal protein S4 (Maki et al., 2004). Overexpression of GroEL can partially compensate for the lack of DnaK at higher temperatures (El Hage, 2001). Temperature sensitive GroEL mutants have reduced levels of mature 70S particles and accumulate a 45S precursor of the 50S particle (El Hage, 2001). GroEL has also been implicated in the RNAseE dependent processing of the pre-5S rRNA (Sohlberg et al., 1993). The ribosomal proteins S2 and L9 as well as the L7/L12 heterodimer immuno-precipitate together with GroEL (Houry et al., 1999).

In addition to the proteins mentioned above, several other individual proteins have been suggested to be involved in ribosomal maturation. The Era protein is an essential GTPase (March et al., 1988). Its gene is co-transcribed with the gene

for RNAse III. The Era protein consists of two domains, the N-terminal GTPase domain and a C-terminal KH domain (Chen et al., 1999). The K-homology domain of Era mediates the binding to the cytoplasmic membrane (March et al., 1988; Lin et al., 1994; Hang and Zhao, 2003) and to 16S rRNA (Sayed et al., 1999; Meier et al., 2000; Hang and Zhao, 2003). Several different functions have been suggested for the Era protein, such as roles in DNA replication and cell division (Gollop and March, 1991; Britton et al., 1997; Britton et al., 1998), coordination between nitrogen and carbon uptake (Powell et al., 1995), expression of heat shock proteins (Lerner and Inouye, 1991), as well as ribosomal maturation. Overexpression of Era suppresses an rbfA (see below) mutation (Inoue et al., 2003). The amount of ribosomal protein S6 increases if cells are depleted for Era (Lerner and Inouye, 1991). In addition to its ability to bind 16S rRNA, Era can also bind to 30S subunits (Sayed et al., 1999). Moreover, the GTPase activity of Era is stimulated by 16S rRNA (Meier et al., 1999; Meier et al., 2000). Some mutations in the Era protein increase the amount of 17S rRNA (Nashimoto et al., 1985; Inoue et al., 2003).

The RimM protein has also been identified as an auxiliary protein for ribosomal assembly. The gene for the RimM protein is located in an operon, which also contains the genes for ribosomal proteins S16 and L19 (Bystroem et al., 1983). A rimM deletion mutant has a low growth rate at 37° C and this difference is accentuated at temperatures of 42° C or 21° C (Persson et al., 1995; Bylund et al., 1997). The rimM deletion mutant is also deficient in translation initiation and shows a reduced translational elongation rate (Bylund et al., 1997). This mutant accumulates the 17S rRNA precursor to the 16S rRNA (Bylund et al., 1998). The slow growth and the translation deficiencies can be suppressed by alterations in the C-terminal part of ribosomal protein S13 and by overexpression of the RbfA protein (Bylund et al., 1997; Bylund et al., 1998). The RimM protein binds to 30S subunits but not to 70S subunits and the amount of the RimM protein is highest in the slowest sedimenting part of the 30S peak after sucrose gradient centrifugation of cellular extracts. This indicates that the RimM protein binds to an immature subpopulation of the 30S particle (Bylund et al., 1997). The C-terminal half of the RimM protein is proposed to fold into a PRC β-barrel structure (Anantharaman and Aravind, 2002).

The YrdC protein was identified as a suppressor of a temperature sensitive mutation in the gene for the RF1 protein. The yrdC mutant shows increased levels of the 17S precursor to 16S rRNA and of free ribosomal subunits, resulting in fewer polysomes (Kaczanowska and Rydén-Aulin, 2004). The YrdC protein binds to double stranded RNA *in vitro* (Teplova et al., 2000).

The RsgA protein is a GTPase that is activated by ribosomes, especially by 30S subunits but also by 70S ribosomes and to a lesser extent by 50S subunits (Daigle et al., 2002; Daigle and Brown, 2004). This protein has an OB β-barrel fold that mediates binding to the ribosome, which is strongest in the presence of the non-hydrolyzable GTP analogue GDPNP (Daigle et al., 2002; Daigle and Brown, 2004). The binding of the RsgA protein to the ribosome is abolished in the presence of aminoglycoside antibiotic binding at the A-site (Himeno et al., 2004). Deletion of the RsgA protein results in a reduction in growth rate and reduced amounts of 70S particles compared to 50S and 30S particles and also in an increased amount of 17S rRNA (Himeno et al., 2004). Free 30S subunits from the deletion mutants cannot activate the GTPase activity of RsgA *in vitro* but free 30S subunits from the wild type strain or from 70S particles can (Himeno et al., 2004).

RbfA, ribosomal binding factor A, was identified as a suppressor of a cold-sensitive mutation (C23U) in 16S rRNA (see Figure 1.8) (Dammel and Noller, 1995). The C23U mutation weakens the base pairing of this base with G11 and destabilizes the 5' helix (helix I) of the 16S rRNA (Dammel and Noller, 1995). Helix I of 16S rRNA is part of the so-called 'central pseudoknot' that links the 5'-domain and the central domain in the 16S rRNA secondary structure (see above). Cold shock conditions lead to an enhanced expression of RbfA in *E. coli* whereas RbfA-deletion mutants display a cold-sensitive phenotype quite similar to that caused by the mutation in the helix I of 16S rRNA (Dammel and Noller, 1995; Bylund et al., 1997).

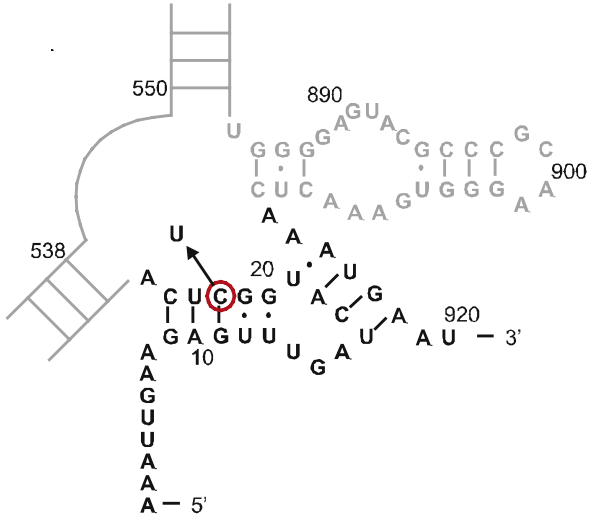


Figure 1.8 Secondary structure of *E. coli* 16S rRNA helix I region. The red circle denotes the cold sensitive C23U mutation.

In both cases, sucrose gradient sedimentation analysis of polysomes revealed that RbfA is found in the pre-30S and 30S subunit fractions, but not associated with 50S subunits, 70S monosomes, or polysomes (Dammel and Noller, 1995; Jones and Inouye, 1996). These 30S ribosomal subunits contained an incompletely processed pre-16S rRNA. Therefore, RbfA seems to enhance the efficiency of the final processing steps for pre-16S rRNA that take place within the 30S ribosomal subunits (Bylund et al., 1998). Interestingly, phenotypes of the *rbfA* knockout strain appear to be very similar to those observed in the S5 mutant strain mentioned above. The 16S rRNA processing deficiencies and the inability to form polysomes becomes increasingly more severe with decreasing temperature (Xia et al., 2003). The expression rate of RbfA increases during a cold shock and in RbfA deletion mutant the cold shock state becomes permanent after a temperature downshift (Jones and Inouye, 1996)

The RbfA protein is a 12–15 kDa protein found in the genomes of almost all eubacteria sequenced to date. Structural studies of an N-terminal RbfA fragment from *E. coli* by NMR-spectroscopy revealed a KH-domain fold (Huang et al., 2003), which is common for many RNA-binding proteins. However, instead of the usual GXXG motif, an AXG motif is conserved in all RbfA sequences known to date (Huang et al., 2003). In addition, the charge distribution at the protein surface and the conservation of many basic amino acids throughout the RbfA-sequences suggested that RbfA is probably an RNA-binding protein (Huang et al., 2003). If the 25 C-terminal residues of *E. coli* RbfA are removed, the resulting protein, which does not bind to 30S subunits, can trans-complement the cold shock and 16S processing defects but not completely restore the low growth rate of the rbfa deletion mutant (Xia et al., 2003). In addition, this C-terminal truncation cannot suppress the phenotype of the C23U mutation in the 16S rRNA. These findings appear to support the earlier suggestion based on the suppression of the phenotype of the C23U mutation in 16S rRNA that RbfA might bind directly to the helix I of 16S rRNA during subunit assembly and stabilize RNA secondary and/or tertiary structure in this region.

Alternatively, RbfA could favor the formation of helix I of 16S rRNA by interacting with sequences of the pre-16S rRNA that are capable of forming an alternative secondary structure with nucleotides 1–16 of the mature 16S rRNA. Furthermore, it has been shown for other RNA-binding motifs that they can be utilized in protein-protein interactions. Recent studies suggest that RbfA, along with RimM are essential for efficient processing of pre-16S rRNA (Bylund et al., 1998).

Overview of this thesis

In the first chapter of this thesis a general introduction of the assembly and function of the ribosome is given. Chapter two outlines the experimental procedures performed during this thesis. In chapter 3.2, we describe how the three dimensional solution structure of the cold shock protein ribosomal binding factor A (RbfA) from *Thermotoga maritima* was solved using multi dimensional, heteronucleic nuclear magnetic resonance (NMR) spectroscopy. In chapter 3.3, the backbone fold of the RbfA protein from *Helicobacter pylori* is described. Several NMR experiments are used to compare the secondary structure as well as the dynamical behavior of the two proteins. This leads to the general question of the differences between thermophile and mesophile systems. This question is further investigated in chapter 3.4, where other members of the RbfA family are compared to the proteins used in this study.

It has been suggested that the RbfA protein interacts with the helix I region from the 16S rRNA. Chapter 3.5 describes efforts to test if the RbfA protein is an RNA binding protein and to define the binding mode between the protein and the 16S rRNA. Taking advantage of the tryptophan residue in the *Helicobacter pylori* RbfA sequence, fluorescence quenching experiments were used to extract apparent binding constants of the RNA ligand to the protein. NMR spectroscopic techniques are used to give insights into the binding mode of the RNA ligand and identify the binding surface on the protein. The fourth chapter then presents the conclusions derived from our experimental data.

CHAPTER 2

Experimental Procedures

2.1 Cloning and Expression of RbfA in E. coli

Genomic DNA was obtained from American Type Culture Collection for *T. maritima* (ATCC: 43589/MSB8) and *H. pylori* (ATCC: 700392/26695). These DNAs were extracted by standard protocols (Sambrook and Pollack, 1974) and used as a template for a polymerase chain reaction to amplify the *rbfa* gene. The DNA fragments containing the respective *rbfa* genes were used in the PCR reaction with two synthetic DNA-Primers. The primer *rbfaNdeF* (tmrbfa: 5'-CGG AAT TCC ATA TGA ACC CAG CCT ATA GA-3'; hprbfa: 5'-CGG AAT TCC ATA TGT TAG TCA TTA TCT TT-3') is complementary to the first 15 nucleotides of the *rbfa* gene including the methionine start codon and contains an NdeI and EcoRI restriction site. The primer *rbfaBamR* (tmrbfa: 5'-GAT CCT CTG AAG GAC TAA GGA TTC GC-3'; hprbfa: 5'-TTT ATG AGC GTT CAT TAA GGA TTC GC-3') is complementary to the last 15 nucleotides of the *rbfa* gene and includes the stop codon and a BamHI restriction site. The following PCR protocol was used:

40 µL H_2O	1. 95° C 1:30 min (initial denaturizing)
0.5 µL DNA	2. 95° C 1:00 min (denaturizing)
0.5 µL each Primer	3. 60° C 2:00 min (Primer annealing)
0.5 µL dNTPs [50 mM]	4. 72° C 1:00 min (polymerization)
5 µL PCR buffer (–$MgCl_2$)	5. 72° C 20:00 min (final elongation)
1.5 µL $MgCl_2$ [50 mM]	
1 µL Taq DNA Polymerase	repeated steps: 2–4 25×

The PCR product was purified using the QIAquick PCR purification kit (Qiagen) and was analyzed on a 1.5 % agarose gel. The purified PCR product was cloned into the plasmid vector pET11a (Novagen) using the NdeI and BamHI restriction sites. pET11a is a protein overexpression vector with a T7-RNA-Polymerase promoter and a gene encoding ampicillin resistance. Insertion of the gene encoding *rbfa* was confirmed by DNA sequence analysis. The *rbfa:*pET11a constructs were subsequently transformed in host strain *E. coli* BL21(DE3) competent cells and were selected on an ampicillin (0.1 mg/mL) agar plate. Expression cultures were grown in LB medium (Fisher BioRea-

gents at a concentration of 25 g/L) at 37° C. In 50 mL LB media the cells were grown until OD_{600} of 0.5–0.7 was reached, centrifuged at 5,000 g for 10 min. and washed with fresh LB medium. The cells were resuspended in LB medium (~25 mL) and the main culture (1L) was inoculated with those cells. This culture was grown to an OD_{600} of 0.8 and overexpression was induced with IPTG (250 mg/L). After 3.5 h the cells were harvested by centrifugation at 5,000 g for 20 min. and kept in the freezer at –80° C. Samples were taken before the induction and every hour until the overexpression was stopped by harvesting the cells. The level of protein overexpression in these samples was examined by gel electrophoresis using 12 % SDS-polyacrylamide gels. For isotope labeling, M9 media was used with ^{15}N labeled ammonium chloride and ^{13}C labeled glucose as the sole nitrogen and carbon source, respectively.

M9 media	
Disodium hydrogen phosphate	42 mM
Potassium dihydrogen phosphate	22 mM
Sodium chloride	8.5 mM
[^{15}N] – Ammonium chloride	18.5 mM
Calcium chloride	0.1 mM
Magnesium sulfate	2 mM
Glucose or [^{13}C] glucose	50 mM/12.5 mM
Fe^{3+}	30 μM
Thiamine hydrochloride	5 mg/L
Zn^{2+}	1 μM

For isotopic labeling cells were grown in a 50 mL LB media culture until the cells reached an OD_{600} of 0.5–0.7, centrifuged at 5,000 g for 10 min. and washed with fresh M9 medium. The cells were then resuspended in M9 medium (~25 mL) and the main culture was inoculated with those cells. This culture was grown to an OD_{600} of 0.8 and overexpression was induced with IPTG (250 mg/L). After 3–3.5 h the cells were harvested by centrifugation at 5,000 g for 20 min. and kept in the freezer at –80° C. Samples for SDS-PAGE were collected before induction and then every hour to examine the overexpression levels. In all media the antibiotic ampicillin was used as a selection agent with a concentration of 100 μg/mL.

2.2 Protein Purification

For purification of the thermophile RbfA from *T. maritima* the *E. coli* cells containing the overexpressed protein were resuspended and homogenized in heat shock buffer (50 mM Tris-HCl (pH 7.5), 100 mM KCl and 2 mM EDTA), and sonicated using a sonoplus sonicater (Bandelin). Cell debris was removed by centrifugation for 30 min. at 10,000 g. The soluble fraction was heated to 75° C for 15 min. and centrifuged again at 10,000 g for 30 min. The soluble fraction was dialyzed against Buffer A (25 mM KH_2PO_4 (pH 7.2), 25 mM KCl and 2 mM EDTA). Further purification of the proteins was achieved using cation-exchange-chromatography with a 5 mL HiTrap™ SP-sepharose column (Amersham) on a fast protein liquid chromatography (FPLC) system. The column was run at room temperature using a flow rate of 1 mL/min. The proteins were eluted using a salt gradient (Buffer B: 25 mM KH_2PO_4 (pH 7.2), 500 mM KCl and 2 mM EDTA).

Table 2.1
Gradient protocol for purification of RbfA using a SP-Sepharose-column

Volume (in column volumes)	Buffer A	Buffer B
2	100 %	0 %
8	100 %–20 %	0 %–80 %
2	0 %	100 %
2	100 %	0 %

Fractions containing protein were identified by their absorbance at 280 nm and by gel electrophoresis using 12 % SDS-polyacrylamide gels (SDS-PAGE). These fractions were combined and concentrated to a final volume of 5 mL. The protein solution was further purified by size-exclusion-chromatography on a HiPrep™ 16/60 Sephacryl™ S200 gel filtration column (Amersham) with gel filtration buffer (25 mM KH_2PO_4 (pH 7.2), 200 mM KCl and 2 mM EDTA). The gel filtration was run at room temperature with a flow rate of 0.5 mL/min. Protein was again detected by absorbance at 280 nm and SDS-PAGE. The protein containing fractions were combined and stored at 4° C. The purity of the protein was verified by SDS-PAGE and MALDI-MS. The molecular weight of purified, labeled *T. maritima* RbfA is given in the table below.

Table 2.2
Molecular weights of labeled *T. maritima* RbfA

***T. maritima* RbfA**	**Unlabeled protein**	**^{15}N labeled**	**^{15}N, ^{13}C labeled**
Molecular weight	14133 Da	14311 Da	14954 Da

Purification of the *H. pylori* RbfA was achieved by resuspending, homogenizing and sonicating the *E. coli* cells containing the protein in Buffer A. Protein resuspended in buffer A was purified using a cation-exchange and gel filtration step, using the same buffers and conditions as described above for *T. maritima* RbfA. The protein samples were stored at 4° C and the purity of the protein was verified by SDS-PAGE and MALDI-MS. The expected molecular weight of purified, labeled *H. pylori* RbfA is given in the table below.

Table 2.3
Molecular weights of labeled *H. pylori* RbfA

***H. pylori* RbfA**	**Unlabeled protein**	**^{15}N labeled**	**^{15}N, ^{13}C labeled**
Molecular weight	12487 Da	12645 Da	13200 Da

For the NMR-experiments, the proteins were dialyzed into NMR buffer (see Table 2.4) and concentrated by centrifugation at 5,000 g to their respective concentration used in the NMR experiments.

Table 2.4
NMR buffers for the RbfA proteins

***T. maritima* RbfA**	***H. pylori* RbfA**
20 mM Acetate (pH 4.5)	25 mM KH_2PO_4 (pH 6.5)
50 mM KCl	50 mM KCl

To determine the protein concentration UV measurements were used. Samples were 100× diluted and measured in a 1 cm path length cuvette. The extinction coefficients for the RbfA proteins were:

Table 2.5
Extinction coefficients for the RbfA proteins

***T. maritima* RbfA**	**276 nm**	**278 nm**	**280 nm**
Extinction coefficient	8,700 M^{-1} cm^{-1}	8,400 M^{-1} cm^{-1}	7,680 M^{-1} cm^{-1}

***H. pylori* RbfA**	**276 nm**	**278 nm**	**280 nm**
Extinction coefficient	6,995 M^{-1} cm^{-1}	7,127 M^{-1} cm^{-1}	7,090 M^{-1} cm^{-1}

2.3 NMR Spectroscopic Techniques

Assignment and NOE distance restraints

For the *T. maritima* RbfA, NMR spectra were recorded at 42° C, the *H. pylori* RbfA NMR spectra were recorded at 25° C, using 600 MHz DMX, 700 MHz AV Ultrashield, 800 MHz AV and 900 MHz AV Bruker spectrometers equipped with triple-resonance cryogenic probes and z-axis pulsed field gradients. NMR samples contained 90 % H_2O/10 % D_2O with a protein concentration of 1 mM for *T. maritima* RbfA and 0.7 mM for *H. pylori* RbfA. For broadband decoupling of $^{13}C/^{15}N$ nuclei, GARP (Shaka et al., 1987) sequences were typically used. For proton broadband decoupling, the DIPSI-2 (Shaka et al., 1988) sequence was used. In the experiments described, the WATERGATE-technique (Piotto et al., 1992; Sklenar et al., 1994) was used for water suppression. The acquisition and processing parameters for all NMR experiments described below are given in Tables A.1 and A.2 in the appendix.

First, heteronuclear 2D single bond correlation spectra (HSQC) were recorded (Bodenhausen and Ruben, 1980). The ^{15}N-edited HSQC (Mori et al., 1995) was used to identify the number of signals at the given sample conditions and so was used to test for buffer and temperature conditions. In the case of the ^{13}C-edited HSQC (Palmer III et al., 1991; Schleucher et al., 1994; Kay et al., 1992) spectra the aliphatic and aromatic ^{13}C chemical shift regions were recorded in two separate spectra. Backbone resonance assignments were obtained using standard three-dimensional triple resonance HNCA (Grzesiek and Bax, 1992b; Schleucher et al., 1993; Kay et al., 1994), HNCACB (Muhandiram and Kay, 1994; Wittekind and Mueller, 1993), HNCO (Grzesiek and Bax, 1992b; Schleucher et al., 1993; Kay et al., 1994), HNCACO (Clubb et al., 1992; Kay et al., 1994) HBHA(CBCACO)NH (Grzesiek and Bax, 1993; Muhandiram et al., 1994) and CBCA(CO)NH spectra (Grzesiek and Bax, 1992a; 1993). Taken together these spectra correlate the backbone H_N protons to the N, Cα, Cβ, and C` signals of the same and adjacent amino acid residues. Side-chain resonance assignments for the *T. maritima* RbfA were obtained by analyzing three-dimensional ^{13}C-resolved HCCH-COCSY (Kay et al., 1993) and ^{13}C-resolved HCCH-TOCSY (Kay et al., 1993) spectra with samples in 100 % ^2H_2O. In order to exchange the protein into ^2H_2O buffer the protein samples were frozen at −80° C and

lyophilized. The freeze-dried protein was then dissolved in $^{2}H_2O$ buffer. For the TOCSY-spinlock period the DIPSI-3 (Shaka et al., 1988) sequence was used with a magnetization transfer delay of 15 ms.

NOE cross-peaks were detected using three-dimensional ^{1}H-^{1}H-^{15}N NOESY-HSQC (Davis et al., 1992), and three-dimensional ^{1}H-^{1}H-^{13}C NOESY-HSQC (Davis et al., 1992; Palmer III et al., 1991; Schleucher et al., 1994; Kay et al., 1992) spectra. For the ^{13}C-edited NOESY experiments, samples were dissolved in 100 % $^{2}H_2O$ buffer. The NOE mixing time was 100 ms for spectra used to derive distance constraints from the ^{15}N-edited NOESY-HSQC spectrum. For the ^{13}C-edited NOESY-HSQC spectrum analyzing the aliphatic side chain data, a mixing time of 100 ms was used, whereas the ^{13}C-edited NOESY-HSQC for analyzing the aromatic side chains used a mixing time of 80 ms. In the ^{15}N-edited NOESY-HSQC experiments ^{15}N broad band decoupling was achieved using the GARP (Shaka and Pines, 1987) sequence. In the ^{13}C-edited NOESY-HSQC experiments the GARP sequence was used for broadband ^{13}C decoupling.

NMR-data were processed using XWINNMR 3.5 and TOPSPIN 1.3 (Bruker). ^{1}H, ^{15}N, and ^{13}C chemical shifts were referenced as recommended (Wishart et al., 1995), with proton chemical shifts relative to internal 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 0 ppm. The 0 ppm ^{13}C and ^{15}N reference frequencies were determined by multiplying the 0 ppm ^{1}H reference frequency by the gyromagnetic ratios for ^{13}C (0.251 449 530) and ^{15}N (0.101 329 118). During processing of the spectra the NMR data were optimized for digital resolution by zero filling (SI) of the data matrix. For optimization of sensitivity and/or resolution, prior to the Fourier transformation the QSINE window function (square sine bell window function) was typically applied with a shifting of the sine bell of $\pi/2 = 90°$ (SSB = 2). For automatic baseline correction a polynomial of degree 5 was subtracted from the fid (qpol). After initial processing of the data zero and first order phase corrections were applied. (For details of the processing parameters see Appendix Tables A.1 and A.2)

Measurement of scalar coupling constants

In order to determine the homonuclear three bond coupling between the backbone Hα and H_N ^{3}J (HNHA), a 3D HNHA experiment was used (Vuitser and Bax, 1993; 1994; Kuboniwa et al., 1994). The experiment was carried out at 40° C (*Tm.* RbfA) and 25° C (*Hp.* RbfA) on a 600 MHz Bruker

spectrometer equipped with a cryogenic probe head. The coupling constants were determined by the ratio between the cross (Ic) and diagonal (Id) signal intensity;

Equation 2.1: $^{3}J = (\arctan - (Ic/Id))^{1/2}/(\pi\,(\delta + \delta'))$

where $(\delta + \delta')$ describe the delay times within the pulse sequence during which the H_N-Hα coupling is active (3D HNHA: $(\delta + \delta') = 25$ ms). The pulse program used an acquisition time in the t_2 dimension, which is much shorter than the transverse relaxation time of either H_N or Hα. As a result the line shape in the F2 dimension is determined primarily by the apodization function. Therefore, the line shapes of the cross and diagonal peak are identical and the intensity ratio provides a direct measure of the coupling constant.

Measurement of residual dipolar couplings

Weak alignment of proteins prevents complete averaging of dipolar interactions. This weak alignment is induced by an alignment medium. For this study, 8 % polyacrylamide gels were used. This system for inducing alignment makes use of anisotropic compression or stretching of polyacrylamide gels. The polyacrylamide alignment medium is inert and stable over a wide range of temperature (at least 5° to 45° C), pH values (at least 2.0 to 8.5) and ionic strength (0 mM to at least 200 mM) (Ishii et al., 2001). The gel was cast using an acrylamide/bisacrylamide molar ratio of 83:1 and 0.08 % ammonium persulfate (final concentration). Stretching of the gel was achieved using the gel sample kit from New Era spectro, where a larger diameter gel polymerized in the provided chamber is pressed through a beveled exit into a regular NMR tube. The alignment in electrically neutral polyacrylamide gels is dominated by steric interactions. Impurities in the polyacrylamide gels might disturb the ^{1}H-^{15}N spectrum of the macromolecule studied by introducing several peaks between 114 to 116 ppm ^{15}N-chemical shift and 7 to 7.8 ppm ^{1}H-chemical shift. This was overcome by washing the gel after polymerization in Millipore water for 24 h and than washing the gel in the respective buffer for another 24 h before the actual sample was allowed to diffuse into the gel.

The H_N-N residual dipolar couplings were extracted from an IPAP-HSQC spectrum (Wang et al., 1998). The IPAP-HSQC spectra were recorded with an isotropic sample, and with the sample in alignment media using the conventional IPAP-HSQC technique (Ottiger et al., 1998). The IPAP-HSQC

contains three INEPT-type polarization transfer periods of length (^{1}JNH) before the t_1 period, the third one present to generate the anti-phase doublet spectrum and absent to generate the in-phase doublet spectrum. The resulting two interleaved spectra are processed using a macro, which reads out two separate spectra. By subtraction and addition of the two separate spectra the H_N resonances separated by the H_N-N coupling constants were extracted and analyzed by comparing the resonance positions in the respective spectra. The difference of the isotropic coupling constants and the coupling constants from the anisotropic sample gave the residual dipolar couplings. Hα-Cα residual dipolar couplings were extracted from a HNCA-experiment without Hα-decoupling during carbon evolution. The experiment was recorded with the isotropic sample to evaluate acquisition parameters and then repeated with the anisotropic sample. The active Hα-Cα coupling during the t_1 time then leads to splittings in the F3 dimension of the HNCA spectrum. By comparison of the two isotropic and anisotropic experiments the Hα-Cα residual dipolar coupling could be extracted. Only peaks that could be tracked reliably and positions that could be determined unambiguously were analyzed. Data for residues with ^{1}H-^{15}N heteronuclear NOE values smaller than 0.65 were excluded from the structure calculations.

Measurements of ^{15}N nuclear relaxation parameters

^{15}N relaxation parameters (longitudinal relaxation rates *(R_1),* transversal relaxation rates *(R_2),* and heteronuclear ^{15}N-{^{1}H}-Het-NOE) were measured for *T. maritima* RbfA at 30° C and 40° C on a Bruker 600 MHz spectrometer. The relaxation parameters for the *H. pylori* RbfA protein was measured at 25° C, on a Bruker 600 MHz Spectrometer equipped with a triple resonance cryogenic probe. For the relaxation experiments, both proteins were dissolved at a concentration of 0.5 mM in 90 % H_2O/10 % ^2H_2O, in their respective buffer. Gradient-enhanced pulse sequences were used to minimize water saturation. Twelve experiments were performed for R_1 measurements, using different values of the relaxation delay 10, 50, 100, 200, 300, 400, 600, 900, 1,200 and 1,500 ms and two repeated delays (100 ms, 400 ms). R_2 data sets were obtained employing the following relaxation delays: 18, 36, 54, 90, 126, 198, 270, 306 and 405 ms with two repeated delays (36 ms, 90 ms). During the relaxation period τ of the R_1 measurement, proton decoupling with DIPSI-2 (Shaka et al., 1988) was employed to eliminate effects of cross-correlation between dipolar and chemical shift anisotropic (CSA) relaxation

mechanisms. For R_2 measurements, ^{15}N spin-locking was carried out using the Carr-Purcell-Meiboom-Gill (CPMG) spin-echo sequence during the transverse relaxation period T. The recycle delay between transients was set to 1.5 s in the pulse sequences for R_1 and R_2 measurements. Two identical pairs of ^{15}N-{^1H}-Het-NOE experiments were recorded at each temperature. In one experiment of each pair, protons were saturated for 3 s during the 5 s recycle delay; in the other experiment, a 5 s recycle delay was used without proton saturation.

Relaxation Data Processing and Analysis

The data for the measurement of relaxation parameters were Fourier transformed after application of a cosine-squared apodization function to yield a matrix of 2,048 × 512 data points. Relaxation rate constants and ^{15}N-{^1H}-Het-NOEs were calculated from cross-peak heights in the ^1H-^{15}N correlation spectra. The longitudinal relaxation times, T_1, were obtained by a two-parameter (I_∞, T_1) non-linear least squares fit of the following equation to the experimental data:

Equation 2.2: $I(t) = I_\infty - I_\infty [1 - 2\exp(-t/T_1)]$

Similarly, transverse relaxation times, T_2 were obtained by a two-parameter (I_∞, T_2) fit of the following equation:

Equation 2.3: $I(t) = I_\infty \exp(-t/T_2)$

The non-linear least-squares-fit was performed using the Levenberg-Marquardt algorithm, using Sigma Plot 9.0 (Systat Software Inc.).

Measurement of Diffusion constants

Most biochemical interactions involve translational or rotational diffusion of molecules. The translational diffusion is described by the diffusion coefficient D_s. Self diffusion is defined as translational motion reflecting the random motions of a molecule in the absence of a concentration gradient (Altieri et al., 1995). Measurement of the D_s coefficient was done by using pulsed field gradient (PFG) NMR methods (Stejskal and Tanner, 1965). The pulse se-

quence includes two PFG pulses, which create an echo. By increasing the PFG pulses in successive experiments the echo amplitude is differentially attenuated in each 1D spectrum due to the translational diffusion. The attenuation for 40 signals out of the recorded pseudo 2D experiment is then fitted with the formula:

Equation 2.4: $A = A(0) \exp [-(\gamma \delta G)^2(\Delta - \delta/3)D_s]$

where γ is the ^1H gyromagnetic ratio, δ is the PFG duration [s], G is the gradient strength [G/cm] and Δ is the time between PFG pulses [s], which is the time that allows for diffusion. Since baseline correction is crucial for the read out of the decaying signals, resonances from the 1D methyl region (2–0 ppm) and from the amide region (6–10 ppm) were used. The baseline correction was done for both regions separately.

The overall rotational dynamics of a quasi rigid structure can be expressed in terms of the correlation time, which is derivable from the ratio of longitudinal and transversal relaxation times. For non-spherical proteins, rotational dynamics is governed by a tensorial quantity, the rotational diffusion tensor D_r. The rotational correlation time can then be extracted from this tensor. For theoretical calculations of the rotational correlation time the program HydroNMR (Garcia de la Torre et al., 2000) was used, which requires the (pdb) file of a given structure as input data. The program builds an appropriate hydrodynamic model of the protein and computes the fully anisotropic rotational diffusion tensor. The correlation time was then obtained from this tensor. The rotational correlation time was also estimated experimentally from the R_2/R_1 ratio from the relaxation data.

2.4 Structure Determination and Refinement

NOE signals identified in NOESY spectra were classified as strong, medium and weak based on cross peak intensity and the corresponding restraints were assigned upper limits of 2.7 Å, 3.5 Å and 5.5 Å respectively. The intensities were referenced to NOE intensities from fixed distances of intraresidual amino acid topology (Hβ1–Hβ2) or fixed distances in secondary structural elements ($H_{N(i)} - H_{N(i+1)}$). NOEs from methyl groups were added an additional upper limit of 0.2 Å to give a total of 2.9 Å, 3.7 Å and 5.7 Å upper limit

restraints accounting for the pseudoatom assignments. For the exchangeable H_N protons a fourth class of upper limit restraints for H_N-H_N NOEs was used, with an upper limit of 4.4 Å. Hydrogen bond restraints were added in the late stages of refinement and were identified on the basis of amide proton exchange data and/or were identified if the structural calculations consistently suggested a hydrogen bond between the respective amino acids. For the distance between the H_N and O' an upper limit for N-H···O=C hydrogen bonds of 2.4 Å was used.

Peak tables generated from the NOESY data sets were used as input for the structural calculation program ARIA 1.2 (Linge et al., 2001.). Initial strucures were calculated from unambiguous NOESY peak assignments, and these structures were used to filter possible assignments based on distances in subsequent rounds of refinement. All calculations were performed with CNS 1.1 (Brunger et al., 1998), using the ARIA 1.2 (Linge et al., 2001) setup and protocols. The protein allhdg 5.3 force field (Linge and Nilges, 1999) was used. After each of the first 8 iterations (0–7) in which 50 structures were calculated, the NOE distance restraints were recalibrated by ARIA based on the 10 lowest energy structures. The violation tolerance was progressively reduced to 0.1 Å in the last iteration (iteration 8) in which 200 structures were calculated. For the structure calculations, a four stage simulated annealing (SA) protocol was used using torsion angles dynamics (TAD). The high temperature stage consisted of 10,000 steps at 10,000 K. This was followed by three cooling stages, 8,000 steps to 2,000 K, 5,000 steps to 1,000 K and 10,000 steps to 50 K. During the SA protocol the force constant for the NOE restraints was set to 0, 10, 20 and 50 kcal mol^{-1} $Å^{-2}$. The final 20 lowest energy structures were further analyzed with procheck 3.5.4 (Laskowski et al., 1993; Laskowski et al., 1996) and refined in explicit water (Linge et al., 2003).

As experimental restraints NOE based distance restraints, ϕ dihedral angle restraints and residual dipolar couplings were used. ^3J (HNHA) coupling constants with tolerances of ±2 Hz were entered for restraints of the backbone dihedral angle ϕ. Internal parameterization constants for the Karplus equation [$J(\phi) = A \cos^2(\phi - 60) + B \cos(\phi - 60) + C$] within the program ARIA were used to calculate the dihedral angles. The residual dipolar coupling restraints as extracted from the experiment were entered with tolerances of ±1.7 Hz. If an interatomic distance in a calculated structure after a round of refinement exceeded the upper limit restraint, the restraint was loosened or removed from the restraint list. The ARIA program initially starts calculations from a randomized extended structure. The amino acids arginine and lysine are calculated as being positively charged, aspartic acid and glutamic

acid as negatively charged and histidine in a neutral form. For the methyl groups and the aromatic protons (Hδ, Hε and Hζ) of phenylalanine and tyrosine (Hδ and Hε), pseudoatoms were defined with their respective upper limit restraints. ARIA estimates NOE distance restraints for the next round of structural calculations and thereby corrects for spin diffusion effects. In order to use this algorithm, the spectrometer frequencies, rotation correlation time from the relaxation dynamics analysis, and the mixing time from the NOESY experiments had to be provided. The resulting set of structures was analyzed for the structural quality by the provided final structural parameters from the calculation and by visualizing and analyzing the structures using MOLMOL (Koradi et al., 1996) and PyMOL (DeLano, 2002). The provided unambiguous assignments list was then manually checked with the NOESY spectra. Furthermore, the ambiguous assignment list provided by ARIA was also checked manually with the NOESY spectra. All NOE restraints, which were violated consistently in the structural calculations, were also checked manually with the NOESY spectra. After cross checking the assignment lists, the new unambiguous assignment lists were used to start a next cycle of structural calculations. A short molecular dynamics trajectory in a thin layer of explicit solvent (water) refines the final structural ensemble.

2.5 CD Spectroscopy

CD spectra were recorded on a Jasco J-810 instrument. For melting curve measurements the temperature in the range from 15° C to 97° C was controlled by a Jasco PTC 423S Peltier type temperature controller. The protein sample was diluted in buffer (20 mM acetate, pH 4.5, 50 mM KCl (*T. maritima* RbfA) or 25 mM KH_2PO_4 pH 6.5, 50 mM KCl (*H. pylori* RbfA) to a final concentration of 40 µM. CD spectra were obtained at a wavelength of 222 nm using a 1 mm path length cuvette. The spectra were recorded with 5 nm/min scanning speed, with a data pitch of 1 nm, a response time of 2 s and a band width of 1 nm. For the *T. maritima* RbfA, which showed a melting temperature >100° C, increasing concentrations of guanidine hydrochloride (0 M, 0.2 M, 0.4 M) were added to the buffer and the measurements were repeated as described above. To test the influence of the ionic strength of the solvent on the proteins thermal stability, melting curve measurements were conducted using different salt concentrations in the sample buffer. The potassium chloride concentrations were successively increased with final

concentrations of 0 mM KCl, 20 mM KCl, 50 mM KCl and 100 mM KCl. Both proteins were measured in 25 mM KH_2PO_4, pH 6.5 buffer with a final protein concentration of 10 μM. The temperature range for the measurements was 15° C to 97° C. The CD signal at 222 nm was followed using a 1 mm path length cuvette. The temperature scanning speed was 5° C/min with a 2 s response time and a band width of 1 nm.

2.6 Steady state Fluorescence Measurement

The fluorescence of *H. pylori* RbfA samples with a protein concentration of 1 μM (200 μL) was measured at 20° C on a Varian Eclipse Fluorescence spectrometer using a Varian single cell peltier accessory as temperature control. Emission spectra were recorded over the wavelength range of 310–390 nm with an excitation wavelength of 295 nm. The spectral bandpass was 5 nm for all spectra. In order to determine the equilibrium binding constant samples with a fixed protein concentration were titrated with increasing amounts of RNA. The dissociation constant (K_d) was extracted from the titrations by following the decrease in fluorescence at 350 nm. Assuming single-site binding of the RNA the obtained titration curves were fitted to equation:

Equation 2.5:

$$F = \{(F_0 - F_f)/2[HpRbfA]_{tot}\} \times \{b - (b^2 - 4[RNA]_{tot}[HpRbfA]_{tot})^{1/2}\} + F_0$$

$$b = K_d + [RNA]_{tot} + [HpRbfA]_{tot}$$

---

where F_0 and F_f are the initial and final fluorescence intensities respectively, $[HpRbfA]_{tot}$ is the total protein concentration and $[RNA]_{tot}$ is the total concentration of the RNA. Control experiments were done to test if the fluorescence quenching was due to binding of the RNA to the protein. For these experiments a solution of 1 μM tryptophane and a solution with 1 μM concentration of the peptide LLRWFEQNLEKMLPQPPK were used.

2.7 Calculation of relative Contact order

The calculation of the relative contact order was performed using the web based program supplied by the Baker laboratory (http://depts.washington.edu/bakerpg/contact_order/) using the parameters as reviewed in (Plaxco et al., 1998). The relative contact order is the average sequence distance between tertiary contacts in the native protein structure normalized to the number of residues.

Equation 2.6: Relative CO = $(1 / N_c L) \sum^{Nc} \Delta L_{i,j}$

Where N_c is the total number of contacts within a contact threshold of 6 Å. $\Delta L_{i,j}$ is the number of residues separating the contacting residues i and j.

2.8 RNA

RNAse free work

In all biochemical procedures the contamination of RNA samples with RNAses has to be avoided. All buffers and solutions used during this work with RNA were prepared using Millipore water, which was autoclaved before usage. All glassware was heated to 200° C for at least 4 h. Everything that cannot be heated to such high temperatures was washed with 0.1 % Diethyl-pyrocarbonate-solution (DEPC) at 37° C and after that autoclaved, washed with autoclaved Millipore water and dried at 80° C. Reduced volume NMR-tubes (Shigemi) were also washed with DEPC solution and freeze-dried in vacuum after they were washed with autoclaved Millipore water.

RNA preparation

All RNAs used in this study were chemically synthesized and purchased from Dharmacon (Boulder, CO). They were deprotected and lyophilized as described by the manufacturer's protocol. Purity was checked using native polyacrylamide gels and HPLC. All RNA constructs were folded into a

homogeneous, monomeric form by heating them to 95° C for 10 min. at a concentration of ~3 mM and fast-cooling on ice. The RNA samples were then exchanged into buffer (25 mM KH_2PO_4 pH 6.5, 50 mM KCl) for further use in the experiments.

Table 2.6
Extinction coefficients for the RNA constructs

Name	Sequence	Extinction coefficient [L/mole cm]	Molecular weight [g/mole]
RNA I	GAGUUUGAUCCUGGCUCA	175,800	5,709.4
RNA II	UUUAUGGAGAGUUUGAUCCUGGCUCA	260,400	8,282.9
RNA III	UUUAUGGAGAGUUUGAUCCUGGUCUCA	269,900	8,582.2
RNA IV	GGAGUCUGAUCCUGGCUCC	176,000	6,029.6
RNA V	AGAGUUUGAUCCUGGCUCAUUUUUUAAAGGAUA	341,800	10,518.3
RNA VI	UUUAUGGAGAGUUUGAUCCUGGCUCAUUUUUUAAAGGAUA	445,000	14,068.4
RNA VII	AGAGUUUGAUCCUGGCUCAUUUUUUUUUAAAGGAU	356,200	11,107.6
RNA VIII	GGAGUUUGAUCCUGGCUCCUUUUUUUAAAGGAU	327,000	10,487.2
RNA IX	AGAGUUUGACCCUGGCUCAUUUUUUUAAAGGGAU	331,000	10,510.3
RNA X	GGAGUCUGAUCCUGGCUCCAUUUUUUUAAAGGAU	337,300	10,815.4
RNA XI	AGCUUUAGGACAAACACUUUUAUGGAGAGUUUG	343,200	10,580.4
RNA XII	AGCUUUAGGACAAACACUUUUAUGGAGAGUUUGAUCCUGGCUCA	445,000	14,068.4
RNA XIII	GAUCCCGAAGCAAACACUUUUAUGGAGAGUUUG	337,100	10,578.4
RNA XIV	UUUUUUUUUUCAAACACUUUUAUGGAGAGUUUG	334,700	10,395.1
RNA XV	AGCUUUAGGAC	113,000	3,490.1
RNA XVI	CAAACACUUAUAUGGAGAGUUUG	241,300	7,356.5
Poly U	UUUUUUUUU	87,500	2,693.5
ssRNA helix I	UUUAUGGA	86,500	2,511.5

The sequences for the RNA constructs were obtained from DNA sequences of the *Helicobacter pylori* 16S rRNA gene (ATCC 700392/26695) starting at nucleotide 1209081. To determine the RNA constructs concentration UV measurements were used. Samples were diluted by a factor of 1:1,000 and

measurements in a 1 cm path length cuvette were done. Spectra were recoded from 300 nm to 200 nm. The extinction coefficients at 260 nm for the RNA constructs are given in Table 2.6.

RNA protein gel mobility shift assay

To identify the binding of the proteins to the RNA constructs gel mobility shift assays were conducted. For the gel mobility shift assay RNA samples with a concentration of 5 μM and various increasing concentrations of *H. pylori* RbfA samples were used. The samples were incubated for 5 min. in buffer (25 mM KH_2PO_4 pH 6.5, 50 mM KCl) prior to addition of loading buffer (50 mM Tris/Acetate pH 7.0, 0.1 % brome phenol blue, 40 % glycerol). Samples were analyzed in non-denaturing 12 % polyacrylamide gels (50 mM Tris-Acetate (pH 7.0); 19:1 acrylamide/bisacrylamide (w/v) and run at 10–15° C in 50 mM Tris/Acetate running buffer pH 7.0. The resulting gels were stained in ethidium bromide solution (0.5 μg/mL) and documented under UV light (320 nm) using a camera with an UV filter.

RNA oligonucleotide assignment

NMR experiments for determining the RNA fold were recorded at 10° C on a 600 MHz Bruker spectrometer. NMR samples typically contained 90 % H_2O/10 % ^2H_2O. Restricted volume NMR tubes (Shigemi) were used for all experiments. ^{1}D-^{1}H spectra (Sklenar and Bax, 1987) were recorded using WATERGATE pulses for suppression of the water signal. For optimal water suppression the phases of the WATERGATE pulses have to be corrected manually. Spectra were processed using XWINNMR 3.5. The baseline correction for the imino proton region (9–15 ppm) was done using manual baseline correction in XWINNMR 3.5. For the assignment of the imino proton resonances standard 2D ^{1}H-NOESY (Davis et al., 1992) experiments were used. These NOESY experiments used a mixing time of $\tau_m = 80$ ms. Water suppression was achieved by using the WATERGATE technique. The offset for the indirect dimension is set to the imino proton region (9 ppm) and is changed back to the water frequency for the WATERGATE pulses and the recording of the direct dimension. The spectra were processed using XWINMR 3.5, internal baseline correction was used.

2.9 BioMagRes Database accession number

The chemical shift list corresponding to the *T. maritima* structure determination has been deposited in the BioMagRes Database under identifier code BMRB-6314.

CHAPTER 3

Results and Discussion

3.1 Choice of model organism

All previous *in vivo* studies regarding the functional and genetic characterization of RbfA have been carried out with *E. coli* as the model organism. However, the full-length *E. coli* RbfA is unstable *in vitro* and slowly aggregates (Swapna et al., 2001; Huang et al., 2003). Moreover, truncations of the last 25 C-terminus amino acids in the *E. coli* protein that made it amenable to biophysical characterization interfere with some of the biological functions of the protein (Huang et al., 2003).

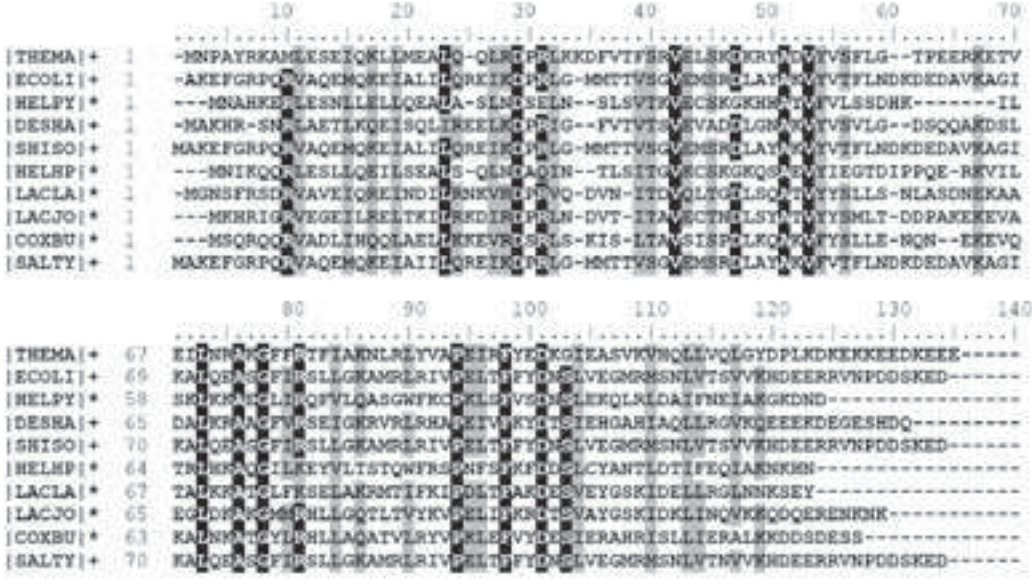


Figure 3.1.
Multiple sequence alignment of 10 bacterial RbfA family members with sequence identities >20 % obtained by database searches using Blossum 62 (Henikoff and Henikoff, 1992) as a matrix. Sequences denoted with a (+) have more than 120 residues, sequences denoted with a (*) have less than 120 amino acid residues. Amino acid residues identical in over 50 % of the sequences are shaded black, residues which were similar in over 50 % are shaded gray.

For this study, two model organisms were chosen which provide the chance of getting further insights into what role the RbfA protein plays in ribosomal assembly. The model organisms were chosen for their RbfA protein characteristics. One organism *(Thermotoga maritima)* synthesizes an RbfA protein with almost the same length as the *E. coli* RbfA and possesses a high simi-

larity in the protein sequence. In addition, its thermophilic character provides an excellent reference for studies regarding the structural differences in thermophile and mesophile systems. Therefore, the *Thermotoga maritima* RbfA was chosen for structural characterization. The full length construct of the *Thermotoga maritima* protein aggregates in solution. Deletion of the highly charged last 11 amino acids in the C-terminus (KEKKEEDKEEE) results in a protein that was stable in solution at concentrations up to 1 mM, which is highly favorable for NMR experiments.

The second organism *(Helicobacter pylori)* synthesizes a 111 amino acid long RbfA protein (full length), which was stable in solution at concentrations favorable for NMR-experiments. The slightly shorter *H. pylori* RbfA protein has also a high sequence similarity to the *E. coli* protein. In addition, *H. pylori* RbfA possesses a single tryptophan residue (W77) close to the putative RNA-binding site rendering it amenable to fluorescence quenching based ligand binding studies. Therefore, *H. pylori* RbfA is a good mimic for the more completely characterized *E. coli* protein in biophysical studies regarding the function of RbfA.

The RbfA family includes proteins with a single polypeptide chain of molecular masses ranging from 11 kDa to 15 kDa. Multiple sequence alignment of bacterial members of the RbfA family, obtained by database searches using Blossum 62 (Henikoff and Henikoff, 1992) as a matrix, showed several strongly conserved regions of the protein sequence.

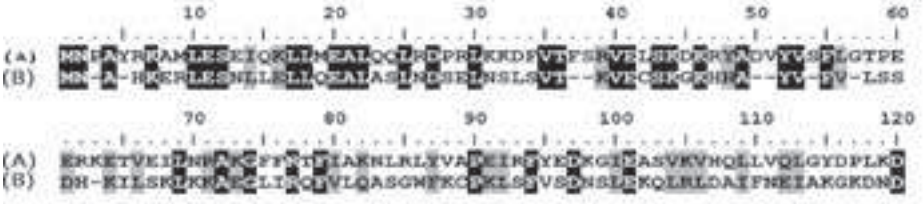


Figure 3.2.
Sequence alignment of the ribosomal binding factor A sequences from **(A)** *T. maritima* and **(B)** *H. pylori.* The alignment was obtained by database searches using Blossum 62 (Henikoff and Henikoff, 1992) as a matrix. Amino acid residues identical in the two proteins are shaded black, similar amino acids are shaded gray.

Noteworthy is the highly conserved AXG motif in the RbfA family which replaces the conserved GXXG motif formed in the helix-turn-helix core of many other KH domain binding domains (Huang et al., 2003; Grishin et al., 2001). The majority of sequence variations across the RbfA family occur around residues numbering 55–65 and the N- and C-terminal regions.

The sequence alignment of the RbfA family showed 30 % identity and 52 % similarity between the *E. coli* and the *T. maritima* homolog. For the *E. coli* and *H. pylori* homolog the sequence alignment showed 23 % identity and 46 % similarity. The two protein constructs used in this study from *T. maritima* (Δ11) and *H. pylori* were aligned by database searches using Blossum 62 (Henikoff and Henikoff, 1992) as a matrix with default parameters. The alignment of the two RbfA protein constructs (see Fig. 3.2) showed that the sequences have 28 % identity and 46 % similarity.

The alignment shown in Figure 3.2 showed identical residues in the regions of helix α1, the β1- and β2-strands and residues surrounding the kink region between helices α2 and α3. The beginning of helix α2 and the end of helix α3 showed the highest sequence variations. The C-terminal region also showed significant variations.

3.2 The protein ribosomal binding factor A (RbfA) from Thermotoga maritima

Purification of the T. maritima RbfA protein.

The gene encoding RbfA from *Thermotoga maritima* was cloned by PCR from genomic DNA. The PCR construct was inserted into a pET11a overexpression vector using Nde1 and BamH1 restriction sites. Sequencing of the resulting plasmid pET11a_Tmrbfa confirmed the insertion of the correct sequence coding for the *T. maritima* RbfA fragment (1–120). The protein was overexpressed successfully in either unlabeled, ^{15}N or $^{15}N/^{13}C$-labeled form in *Escherichia coli* BL21(DE3) cells as described in experimental procedures (see chapter 2.1). For isotopic labeling the cells were grown in M9 minimal medium.

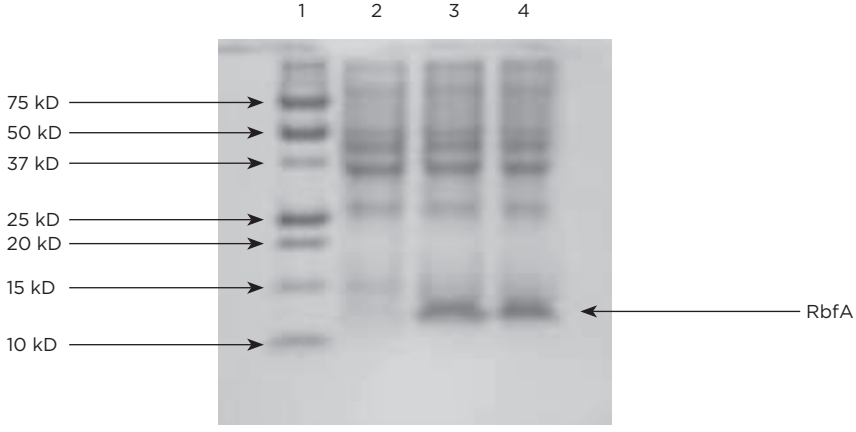


Figure 3.3
Coomassie Brilliant blue stained SDS PAGE (12%) of *T. maritima* RbfA to test for overexpression level following induction with IPTG. Each lane contains cells from a 1 mL sample of the expression media, resuspended in 50 µL buffer (25 mM KH_2PO_4 pH 7.0, 50 mM KCl) and 50 µL SDS sample buffer. Lane 1: Molecular weight marker (Biorad, Precision Plus Protein™), lane 2: before induction, lane 3: 2 h after induction, lane 4: 4 h after induction.

The overexpression of *T. maritima* RbfA is indicated by the appearance of a novel protein band corresponding to a molecular weight of 14 kDa in SDS-PAGE (see Figure 3.3) upon induction with IPTG. The protein expression

rate was tested at certain intervals after the IPTG induction, by SDS-PAGE analysis (see Figure 3.3). The polyacrylamide gels showed that the protein band corresponding to a molecular weight of 14.2 kDa becomes more intense with time after induction. The distinct difference with respect to basal protein expression levels indicated overexpression of the cloned RbfA. An optimal growth time of 3.5 h at 37° C after induction was determined. Densitometric analysis showed that the overexpressed protein accounted for ~20 % of the total cellular protein at this time. Overexpression yields are similar for cells grown on unlabeled LB-medium or labeled minimal medium.

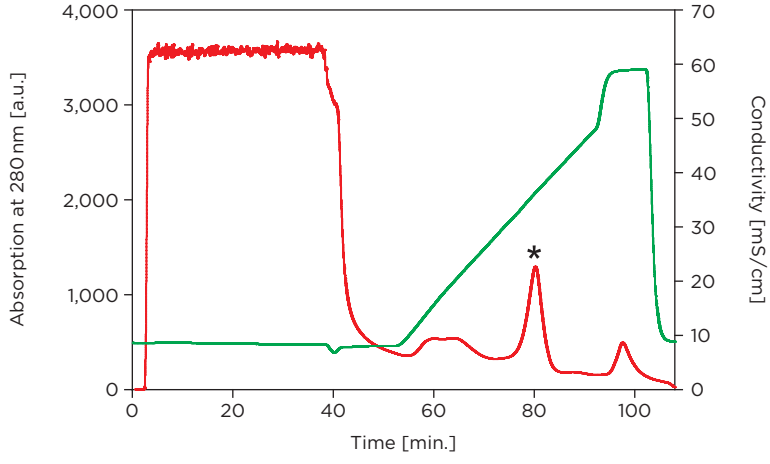


Figure 3.4
Chromatogram of *T. maritima* RbfA purification on a cation exchange column (Amersham) with SP-Sepharose resin. The column was loaded with cell extract after heat shock with a flow rate of 1 mL/min flow rate. Red trace: relative absorption units at 280 nm; green trace: conductivity. The bulk absorbance seen immediately after the sample was loaded are DNA and RNA molecules of the cell extract as well as all negatively charged or non-charged proteins. Application of a salt gradient results in the elution of positively charged proteins from the column. The asteriks marks the eluted fractions of *T. maritima* RbfA at ~80 min.

A purification protocol was established for isolation of *T. maritima* RbfA from *E. coli* as described above in experimental procedures.

After the crude cell extract of the harvested cells was resuspended in heat shock buffer (50 mM Tris/HCl pH 7.5, 100 mM KCl, 2 mM EDTA) and sonicated for cell disruption, centrifugation was used as described to separate the fraction of the soluble cytosolic protein from insoluble protein aggregates and membrane fragments. The *T. maritima* RbfA protein was found in

the soluble fraction after centrifugation. This fraction was treated with a heat shock at 70° C for 10 min. leading to aggregation of the major portion of the *E. coli* proteins.

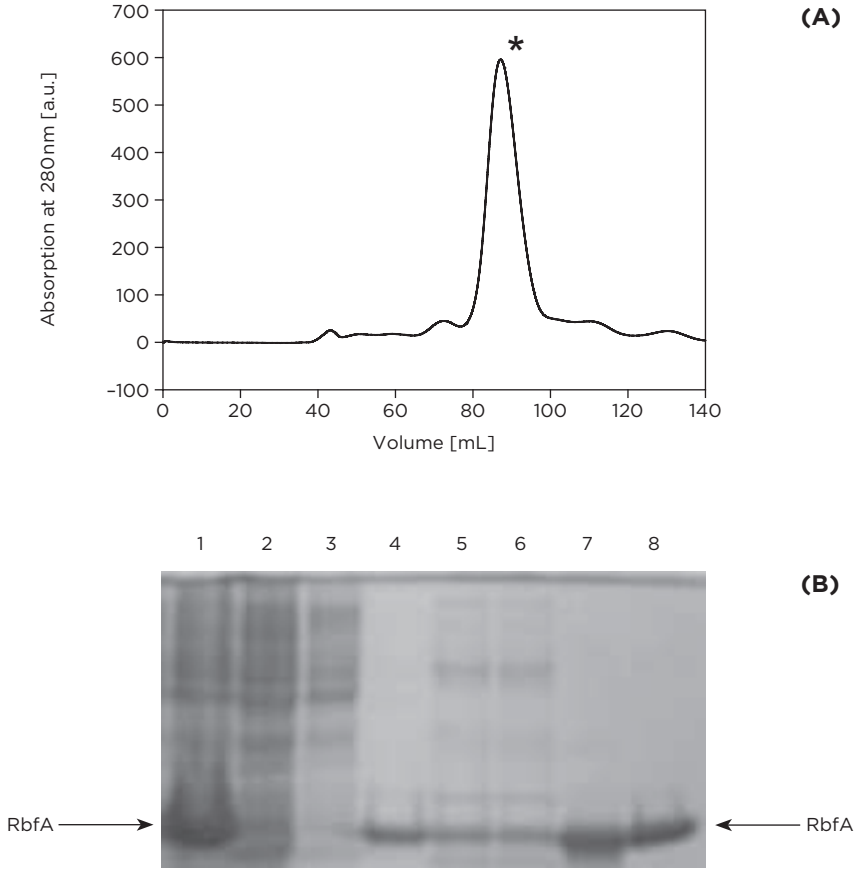


Figure 3.5
(A) Chromatogram of *T. maritima* RbfA using a Sephacryl S-100 gel filtration column (Amersham). The column was loaded with the combined fractions from the cation exchange column concentrated to a volume of 5 mL. Black trace: relative absorption at 280 nm. The star marks the fractions containing *T. maritima* RbfA at ~85 mL.
(B) Coomassie stained SDS PAGE gel (12 %) containing samples of every step of the purification procedure. Samples taken during the purification procedure: lane 1: crude cell extract after overexpression, lane 2: cell pellet after sonication, lane 3: pellet after heat shock, lane 4: *T. maritima* Rbfa as a reference, lane 5: and 6: supernatant following heat shock, lane 7: Rbfa containing fractions from the Sp-Sepharose-column, lane 8: Rbfa containing fractions from the gel filtration column.

After centrifugation, the RbfA protein was found highly enriched and in soluble form in the supernatant. The RbfA protein was separated from the remaining heat stable *E. coli* proteins using a cation exchange chromatography step on a HiTrap™ SP-sepharose-column (Amersham). For separation a KCl gradient from 25 mM to 500 mM was used. The RbfA protein elutes at a KCl concentration of ~200 mM (see Figure 3.4). An additional size exclusion chromatography step on a Sephacryl S-100 gel filtration column (Amersham) was used to further purify the RbfA protein. The RbfA protein elutes after ~85 mL, which corresponds to a molecular weight of ~14 kDa (see Figure 3.5 A).

Protein purified by this procedure was soluble up to concentrations of 1 mM, >95 % pure and showed a molecular weight of 14.1 kDa as verified by SDS-PAGE (Figure 3.5 B) and MALDI-MS. The final yield was ~4 mg protein per liter of *E. coli* culture when isotopically labeled minimal media was used. Samples of soluble protein for NMR-spectroscopy were prepared by exchanging the protein into NMR buffer (20 mM acetic buffer, pH 4.5, 50 mM KCl and 10 % $^{2}H_2O$) and could be concentrated to a final protein concentration of roughly 1.0 mM.

Assignment of T. maritima RbfA backbone resonances

Suitable buffer conditions for the assignment and structure determination of the *T. maritima* RbfA were determined using the ^{15}N-^{1}H single bond correlation experiment (^{15}N-^{1}H-HSQC). After testing several pH conditions ranging from pH 7 to pH 4.5 and several temperatures ranging from 20° C to 40° C, a buffer consisting of 20 mM acetic acid, pH 4.5, 50 mM KCl and 10 % $^{2}H_2O$ was found to give the needed dispersion of the amide proton resonances at 40° C. Of particular importance was the dispersion of the ^{1}H and ^{15}N resonances, since the dispersion of these signals is used in the 3D NMR experiments to enhance the resolution and overcome signal overlap. In addition, a 3D HNHA experiment was conducted in order to test the dispersion of the Hα resonances. The ^{15}N-^{1}H-HSQC spectrum shown in Figure 3.6 and the 3D HNHA experiment show a single set of signals with sufficient signal half width, reflecting a homogeneously folded monomeric protein. In addition, the spectra showed sufficient amide ^{1}H and ^{15}N dispersions, except for some degenerate resonances, which were later determined to be in the flexible C-terminus of the protein. Additionally, the resonances of the arginine ε-amino groups were well resolved. For the 120 residue fragment of *T. maritima* RbfA 114 backbone amide group resonances were expected.

This excludes the resonance of the N-terminal residues, which typically can not be detected due to fast exchange with the solvent and the 5 proline residues that lack an amide group. From the 114 expected signals 108 backbone resonances were resolved in the ^{15}N-^{1}H-HSQC spectrum. Additionally, signals from side chain amino groups of 5 glutamines, 3 asparagines and 10 arginine ε-amino groups were expected. All of the glutamine and asparagine side chain signals were identified, as well as 9 out of 10 arginine ε-amino group side chain signals were detected (see Figure 3.6). Triple resonance NMR experiments were used for the sequential assignment process, which do not necessitate any knowledge about the identity of the spin systems. Instead, the sequential connectivitys via ^{1}J and ^{2}J couplings are used to establish correlations between amino acids.

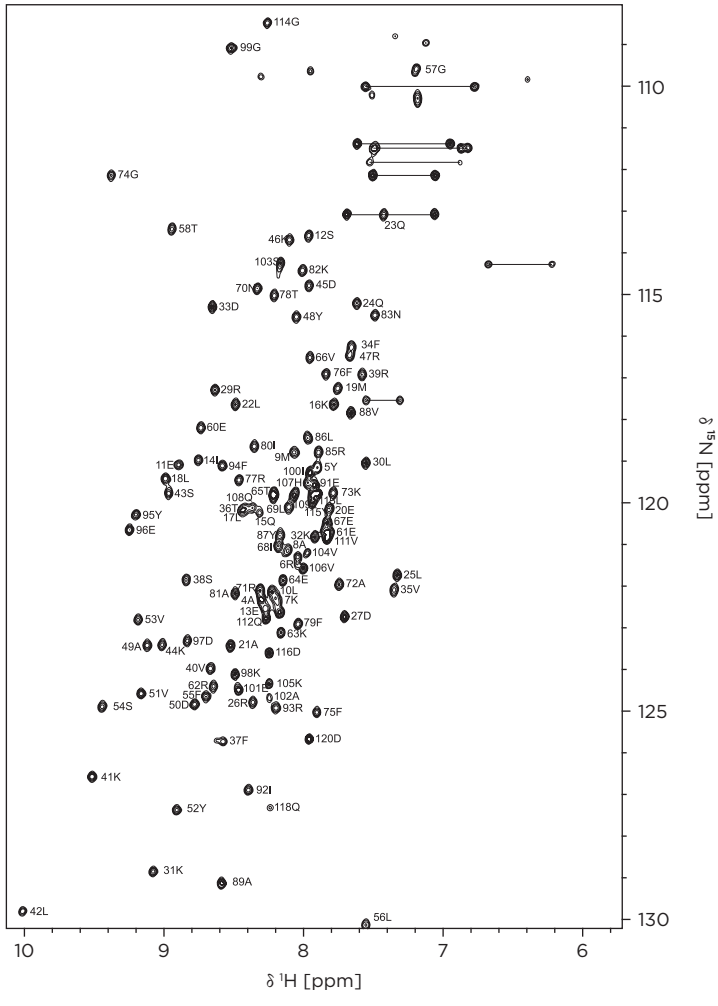


Figure 3.6
(^{15}N-^{1}H) HSQC spectra of *T. maritima* RbfA from a ^{15}N labeled sample in NMR buffer (20 mM acidic acid at pH 4.5, 50 mM KCl) at 40° C on a 700 MHz Bruker spectrometer. The side chain amino groups of 3 asparagines and 5 glutamines are indicated by horizontal lines. The peak labels show the assignment as extracted from the assignment experiments.

In the following section, a brief outline of the general assignment strategy is given. Assignment of the backbone residues started from the HNCA (see Figure 3.7 A) and HNCACB (see Figure 3.7 B) data. The HNCA and HNCACB spectra have three frequency axis [ω3 (^{1}H), ω2 (^{13}C), ω1 (^{15}N)]. The spectra were used to correlate both the intraresidue Cα (i) and interresidue Cα (i–1) in the case of HNCA and the intraresidue Cα (i), Cβ (i) and interresidue Cα (i–1) and Cβ (i–1) in the case of the HNCACB. The projection of the [ω3 (^{1}H), ω1 (^{15}N)] plane shows resonances as seen in the ^{15}N edited HSQC spectra. Each signal represents a single amino acid. At the frequency of each amide proton there are two cross signals in the Cα dimension at the respective nitrogen frequencies shown in the [ω2 (^{13}C), ω3 (N_H)] strips (Figure 3.7 A). Using these cross signals of different intensities, which are due to the difference in coupling constants (intraresidue N-Cα 11 Hz; interresidue N-Cα 7 Hz), a chain of correlations through the whole amino acid sequence could be established. However, *T. maritima* RbfA has 5 proline residues at positions P3, P28, P59, P90 and P117. These proline residues interrupt that chain due to the absence of an amide proton in this amino acid. The HNCA spectrum alone was not adequate for extracting all backbone connectivities due to chemical shift degeneracies in the Cα resonances. Therefore, the HNCACB was used in addition and provided an independent route for assignments through the Cβ resonances. Thereby the ambiguities from the HNCA-spectrum were almost completely overcome. Neighboring residues were identified readily on the basis of the common Cα and Cβ shifts observed (see Figure 3.7 B).

The sequential assignment was further simplified when the intra- and interresidue cross signals could be distinguished not only by their intensity. To differentiate between the intra- and interresidual cross signals the data from a CBCACONH experiment was used. In a CBCACONH spectrum only the interresidue Cα (i–1) and Cβ (i–1) resonances are correlated. The magnetization is transferred via the CO atom, thus only the interresidual cross signals can be observed. Therefore, the CBCACONH data was then used to confirm the sequential assignment obtained from the HNCA and HNCACB spectra. In the process of assigning the resonances to their respective residues, the unique chemical shifts observed for Gly, Ser, Thr and Ala residues provided useful checkpoints. To complete the backbone assignment the HNCO experiment was used to provide the resonances of the backbone carbonyl carbons. Following the described assignment strategy nine sequentially connected sections were identified. The correlations were interrupted as expected by the proline residues due to the lack of the H_N-proton and due

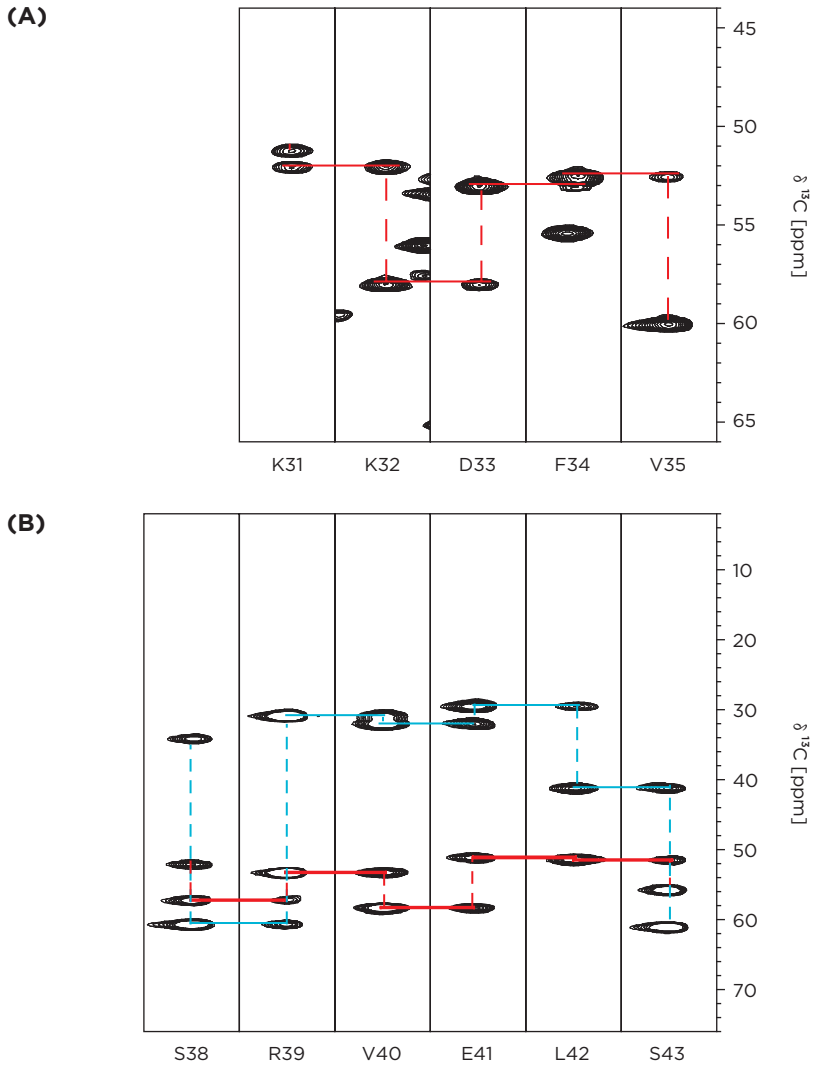


Figure 3.7
(A) 3D HNCA spectrum strip plot for residues K31 to V35 of *T. maritima* RbfA at 40° C on a 700 MHz Bruker spectrometer. The [ω2 (^{13}C), ω3 ($^{15}N_H$)] strips were extracted from their respective N_H and H_N chemical shifts and plotted in their sequence order. The red solid lines mark the connection between Cα (i–1) and Cα (1) between two strips. The dashed lines connect the intraresidue Ca (i) and Ca (i–1) residues within the same strip.
(B) 3D HNCACB spectrum strip plot for residues S38 to S43 at 40° C. The strips were extracted from the spectrum as described above. The solid red lines indicate the Cα (i–1) and Cα (1) connectivities between the strips; the blue lines indicate the Cβ (i–1) and Cβ (1) connectivities between the strips. The broken lines indicate the connection of residues within the same strip.

to signal degeneracy for residues L17/L18 and Q23/Q24. No assignment was found for residue N2 which lies in the N-terminal sequence M1-N2-P3 and no connections were identified for residues T65, V66 and R93. The assignments reached the following extent: 92 % H_N, 88 % Cα, 84 % Cβ and 84 % CO.

In order to provide an additional unique pathway for resonance assignments, as well as to obtain the assignments for the Hα and Hβ resonances necessary for the side chain assignment process, the HNHA and HBHA(CBCA)CONH experiments were used (see Figure 3.8). The HNHA experiment correlates the intraresidue Hα resonances with the N_H and thus provided a unique check on the assignments made above. Together with the Hα and Hβ resonance assignment from the HBHA(CBCA)CONH experiment, which correlates the interresidue Hα (i–1) and Hβ (i–1), the assignment of the backbone protons was accomplished to the following extent: 97 % H_N, 97 % N_H, 95 % Hα and 33 % Hβ.

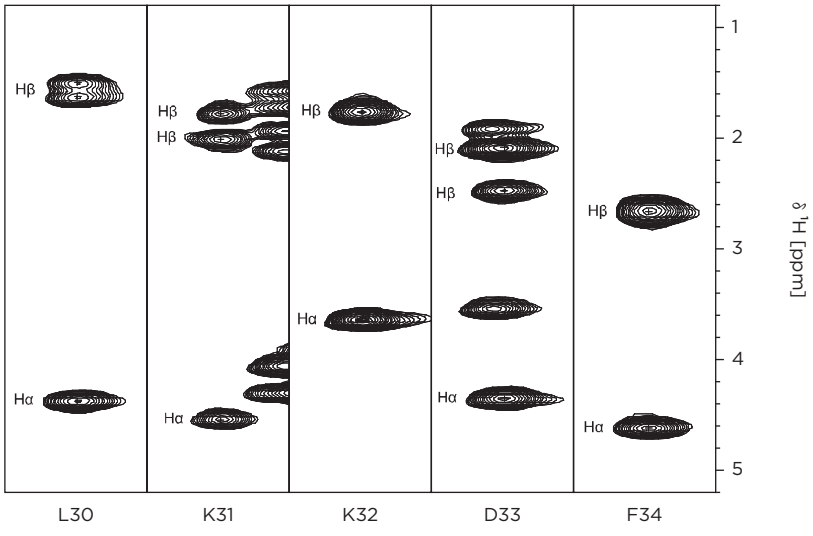


Figure 3.8
3D HBHA(CBCACO)-NH spectrum strip plot for residues L30 to F34 of *T. maritima* RbfA at 40° C. The [ω2 (^{1}H), ω3 ($^{15}N_H$)] strips were extracted from their respective N_H+1 and H_N+1 chemical shifts and plotted according to their sequence. For the assignment of degenerate signals of methylene groups, a pseudoatom was assigned, as seen for residues K32 and F34.

For the cases where Hβ protons can not be differentiated in the spectra, e. g. alanine methyl group protons, a pseudoatom was assigned. Using both experiments the assignments could be validated and a set of starting points for the side chain assignment was established.

From the 3D NMR experiments mentioned above the backbone assignment for the 120 amino acid *T. maritima* RbfA was essentially complete, with the exception of residues M1 and N2 and all 5 proline residues. In addition, there were no assignments for the residues T65, V66 and R93.

Assignment of T. maritima RbfA side chain resonances

Based on the unique Hα and Cα resonance assignments described above, the Hα/Cα region of the ^{13}C-^{1}H(CT-HSQC) spectrum was assigned. These Hα/Cα peaks were then used as the anchor points for the assignments of the aliphatic side chains. Additionally, the chemical shift data for the Hβ and Cβ was also used whenever it was possible. The assignment for the aliphatic side chain resonances was obtained using 3D CC(CO)NH, HCCH TOCSY and HCCH COSY experiments. The CC(CO)NH experiment first gives assignments for the aliphatic side chain ^{13}C resonances, which were used together with the Cα and Cβ assignments as the starting points for further assignment using the 3D TOCSY and COSY spectra. The HCCH TOCSY data was recorded with a mixing time of 15 ms yielding side chain correlations through multiple C-C bonds. The delay time was chosen to allow medium range correlations, so the side chain signals from Cα up to Cδ were identified. The side chain signals of Lysine Cε and Arginine Cδ could not be assigned for all cases using the given delay time. The HCCH COSY data was employed to obtain definitive identification of neighboring carbons and hydrogens in the aliphatic side chains. During the side chain assignment particular effort was made to assign the resonances of the methyl groups, since those are part of the hydrophobic core of the protein and give rise to readily assignable NOE signals.

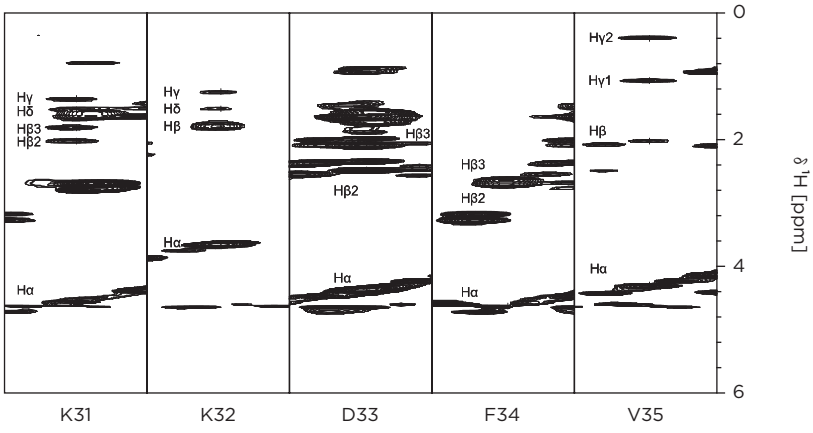


Figure 3.9
3D HCCH TOCSY spectrum strip plot for residues K31 to V35 of *T. maritima* RbfA at 40° C. The [ω1 (^{1}H), ω3 (^{1}H)] strips were extracted from their respective ^{13}C aliphatic side chain atom chemical shifts and plotted according to their sequence. For the assignment of degenerate signals for the two protos of methylene groups, a pseudoatom was assigned, as seen for residues K32 and F34.

The limited bandwidth of the DIPSI-3 mixing sequence in the HCCH-TOCSY experiment does not allow obtaining correlations from the aromatic side chains. In order to complete the assignment of the aromatic resonances, a combination of an aromatic ^{13}C-^{1}H HSQC and ^{1}H-^{1}H-^{13}C NOESY data was used. 85 % Cδ/Hδ, 85 % Cε/Hε and 14 % Cζ/Hζ of the phenylalanine side chain atoms could be assigned with this strategy. No assignments were obtained for the side chain of residue F76 due to signal degeneracy. 83 % Cδ/Hδ and 50 % Cε/Hε of the tyrosine residue side chains were assigned, whereas no assignment could be made for the side chain of residue Y115. The side chain of residue H107 was fully assigned using the ^{13}C-^{1}H HSQC and ^{1}H-^{1}H-^{13}C NOESY data.

In summary, backbone and non-aromatic side chain assignments from the triple resonance experiments were made to the following extents: 97 % H_N, 97 % N_H, 97 % Cα, Hα; 95 % Cβ, 93 % Hβ; 33 % Cγ and 66 % Hγ; 40 % Cδ and 59 % Hδ, and 97 % CO.

The assigned ^{1}H,^{15}N,^{13}C chemical shifts of the *Thermotoga maritima* RbfA have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession number BMRB-6314. A list of the resonance assignment is provided in the appendix (see Table 5.5).

Secondary structure of the RbfA protein from T. maritima

The chemical shifts of the backbone atoms Hα, Cα, Cβ, and C' were identified to have a strong correlation with the secondary structure of the residue involved (Spera and Bax, 1991; Wishart and Sykes, 1994). Wishart and Sykes developed a chemical shift index (CSI) using the chemical shifts of all four nuclei that predicts the secondary structure with relatively high accuracy.

Therefore, the assignments from the backbone residues of *T. maritima* RbfA were used to assess the secondary structure of the protein. Figure 3.10 shows the plot of the Cα and Hα chemical shift indices of *T. maritima* RbfA. Analysis of the secondary structure of RbfA showed it to be a mixed α/β protein. The secondary structure of RbfA is comprised of three α-helices and three β-strands. The protein exhibits secondary structural elements ordered α1-β1-β2-α2-α3-β3 along the sequence. A notable feature of the RbfA family is the kink region between helices α2 and α3 with the highly conserved AXG motif. This corresponds to residues A72-K73-G74 in *T. maritima* RbfA. In the polypeptide segment E60-L84 residue A72 is the only residue, which showed non-helical Cα chemical shift values.

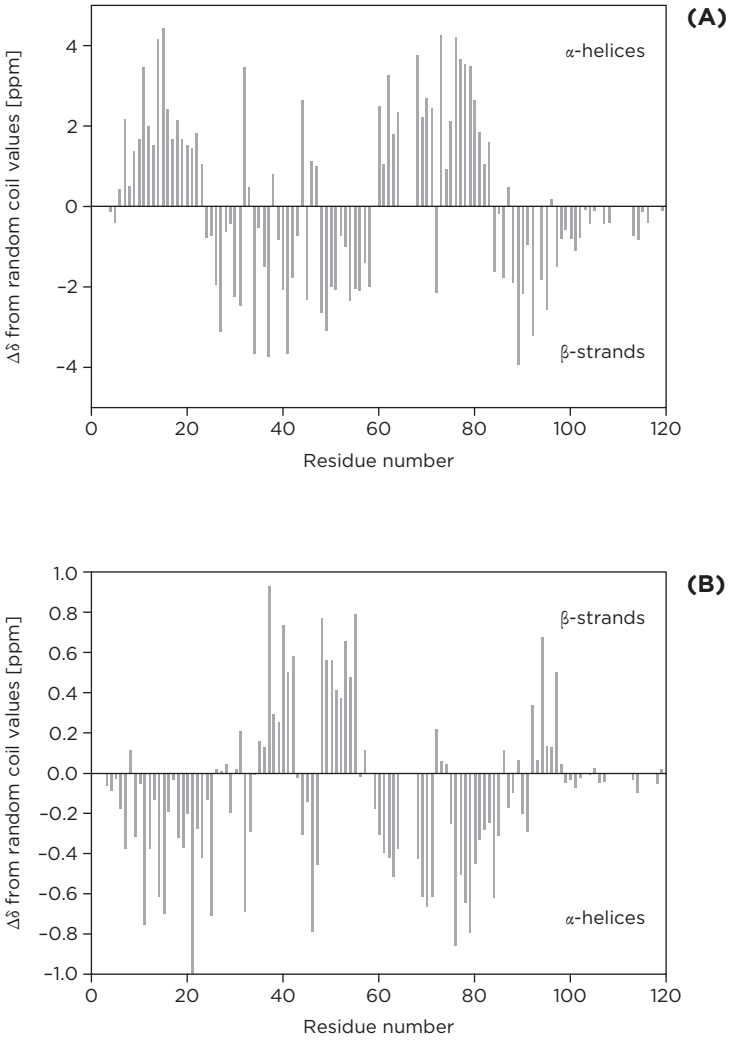


Figure 3.10
The chemical shift differences to random coil values for the Cα **(A)** and Hα **(B)** atoms plotted against the residue number of *T. maritima* RbfA. The chemical shift values were calculated by subtracting the assigned Cα and Hα values from chemical shift values of residues in random coil conformation.

The hydrogen bonded amide protons of α-helices and β-sheets displayed slow exchange with the bulk solvent upon dissolution in $^{2}H_2O$ buffer. This provided additional means of identifying the residues involved in regular secondary structures.

Therefore, samples of the *T. maritima* protein were freeze dried and resolved in 100 % $^{2}H_2O$. ^{15}N-^{1}H HSQC spectra were recorded at certain time spans and the observable signals were identified according to the backbone assignment. In this manner amide groups, which are involved in hydrogen bonds, within protein secondary structures were identified. For *T. maritima* RbfA, residues L18, A21 and L22 in helix α1, residues T36, V40 and E41 in β-strand β1, residues A49, V51, Y52, V53 and S54 in β-strand β2 residues E61, I68, L69, N70, R71 and A72 in helix α2, residue I80 in helix α3 and residues R93 and Y95 in β-strand β3 were protected and thus identified as having backbone amide group protons involved in hydrogen bond interactions. From the experiments described above the corresponding secondary structure elements of *T. maritima* RbfA were identified: Glu11–Leu25 (α1), Val35–Glu41 (β1), Ala49–Phe55 (β2), Glu60–Arg71 (α2), Phe75–Leu84 (α3) and Glu91–Asp97 (β3).

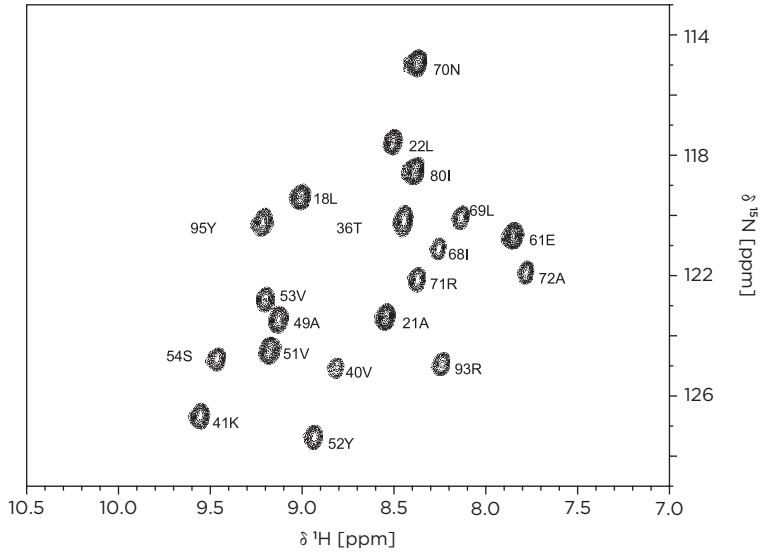


Figure 3.11
A portion of the ^{15}N edited HSQC spectrum of *T. maritima* at 40° C. The spectrum was recorded 24 h after the protein sample was resuspended in 100 % $^{2}H_2O$ buffer. Only backbone N_H resonances that participate in hydrogen bond interactions can be observed.

Structural restraints for T. maritima RbfA

In the following sections the strategy of the gradual assignment of ^{1}H-^{1}H NOE correlations will be described, which were later used as distance restraints in the structural calculations.

Using 3D ^{15}N edited NOE spectra (^{1}H-^{1}H-^{15}N-NOESY-HSQC) in 90 % H_2O/10 % $^{2}H_2O$, with a mixing time of 100 ms, the signals of 1887 NOE correlations were initially identified. Using a 3D ^{13}C edited NOE spectra (^{1}H-^{1}H-^{13}C-NOESY-HSQC) in 100 % $^{2}H_2O$, with 100 ms mixing time, 2,248 correlations were initially identified. In addition, 173 NOE correlations could be extracted from a 3D aromatic side chain specific ^{13}C edited NOE experiment (^{1}H-^{1}H-^{13}C-NOESY-HSQC).

The residue specific assignments from imported peak lists of the assignment spectra served as starting points for further identification of NOEs. Identification of NOE signals started with the sequential walk from the diagonal resonances to the respective neighboring residues. In a next step, NOEs involving the signals of the well resolved methyl group protons and Hα resonances were assigned.

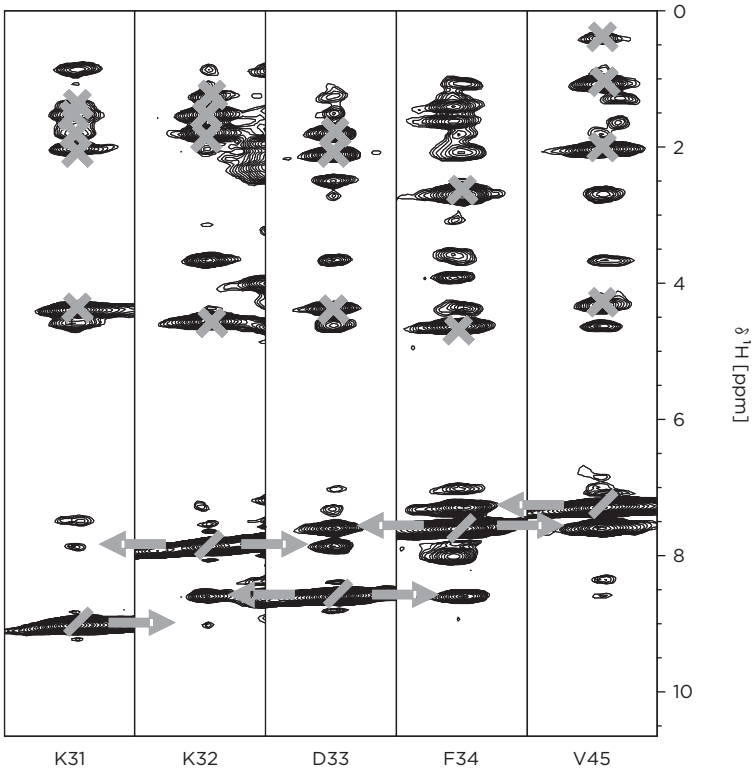


Figure 3.12
[ω3 (^{1}H), ω1 (^{1}H)] strips from ^{15}N edited NOESY-HSQC for residues K31 to V35. The diagonal peaks are marked by a diagonal slash. Intraresidual peaks are marked by a cross. Interresidual peaks to the neighboring amino acid residues are marked by arrows.

The correlations were picked manually and the final assignment was entered when the equivalent signal of the cross peak was found in the diagonal peak of the [ω3 (^{1}H), ω1 (^{1}H)] strips. After the assignment of the well resolved resonances, further efforts concentrated on the defined secondary structures. Regular secondary structure elements in proteins entail characteristic NOEs (Wüthrich et al., 1986). These NOEs can be used to identify the boundaries of the secondary structural elements. In particular a dense array of sequential and medium range NOEs are often observed in α-helices. These included $N_H(i) - N_H(i+1)$, $N_H(i) - N_H(i+2)$, $H\alpha(i) - N_H(i+3)$, $H\alpha(i) - N_H(i+4)$, and $H\alpha(i) - H\beta(i+3)$ NOEs. In the case of β-strands, strong sequential $H\alpha(i) - N_H(i+1)$ NOEs are observed, as well as $H\alpha - N_H$ and $H\alpha - H\alpha$ NOEs across the strands of a β-sheet. These interstrand NOEs were used for determination of the topology of the β-sheet as well as to define the termini of the β-sheet strands.

The NOE data from *T. maritima* RbfA showed strong characteristic α-helix NOEs for residues Glu11–Leu25, which constitute α-helix α1. The ladder of characteristic helical NOEs is disrupted at residues L10 and R26, which were then determined as the N-terminal and C-terminal ends of helix α1 respectively. Following this approach the boundaries of all other helical secondary structures were determined. Helix α2 extends from residues Glu60–Arg71 and residues Phe75–Leu84 constitute helix α3. Within the kink region of *T. maritima* RbfA comprising residues A72–G74 the characteristic helical NOE pattern was disrupted. This was in perfect agreement with the chemical shift data described above.

The β-strands were identified by a combination of strong $H\alpha$-H_N NOEs, slowly exchanging amide protons and their characteristic interstrand NOEs. The first β-strand β1 extends from residue Thr36–Lys41. This beta strand showed strong intrastrand $H\alpha$-H_N NOE correlations, and the chemical shift index also identified these residues as a β-strand. In addition, the slowly exchanging amide protons of residues T36, Arg39, Val40 and Glu41 were indicative of a β-strand. The characteristic NOE pattern was disrupted at residue S38, although the signal line width did not show any distortions for this residue. The second β-strand β2 extends from residues Ala49–Phe55, showing the characteristic NOE pattern and chemical shift indices for a β-strand. Additionally, residues Ala49, Val51, Tyr52, Val53 and Ser54 were identified to have slowly exchanging amide protons. The interstrand NOE correlations between β-strands β1 and β2 suggested that the two β-strands were oriented antiparallel to each other. The third β-strand β3 extends from residue Glu91–Lys98. Within this β-strand, residues Arg93 and Tyr95

showed slowly exchanging amide protons. From the interstrand NOE correlations it was concluded that this third β-strand is parallel with respect to beta strand β2. For this β-strand, as well as in the others, the chemical shift index data is in good agreement with both the NOE data and the data from the deuterium exchange experiments.

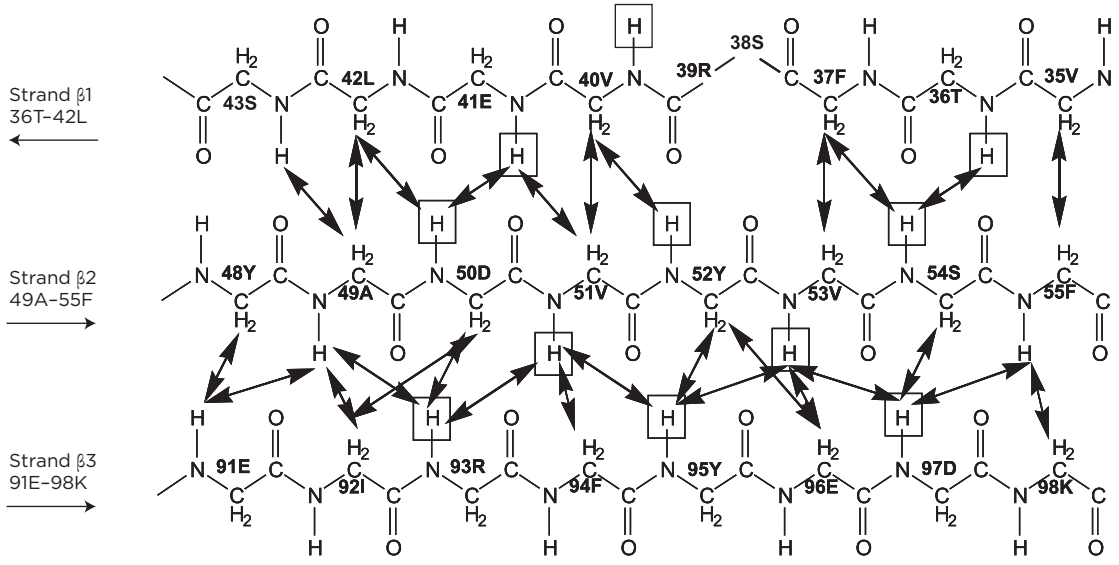


Figure 3.13
Plot of NOE contacts in the β-sheets of *T. maritima* RbfA. The arrows mark NOE correlations as identified in the recorded NOESY experiments. Boxed protons are protected against H/D exchange.

After the manual assignment of the well resolved NOE signals and the correlation of the secondary structural NOEs, the first structural calculations were done using DYANA (Guntert at al., 1997). From this initial structure calculation 20 of the 80 calculated structures, which showed the lowest target function, were chosen to audit the exported peak list for wrong and/or redundant assignments. NOE restraints that were violated in all of the 20 structures were checked again for their assignments. From this point on further NOE cross peak correlations were assigned in a semi-automatic fashion, where recommended correlations from the software (XEASY (Bartels et al., 1995), SPARKY (Goddard and Kneller, 2004) were verified manually and finally assigned. From this semi-automated assignment of NOE restraints, the final peak list used in the first structure calculation cycles included 693 restraints from the ^{15}N edited NOESY and 989 restraints from the ^{13}C edited NOESY experiment. Of those NOE distance restraints already assigned, 408 were interresidual (i–j = 0), 341 sequential (i–j = 1), 45 medium range (i–j < 5) and 85 long range (i–j > 5).

These NOE correlations were converted into distance restraints based on their signal intensities. Known $^{1}H–^{1}H$ distances from secondary structural elements were used to calibrate the upper limit distance restraints. Within the program CYANA (Herrmann et al., 2002) redundant upper limit restraints were eliminated, and at the same time pseudo atom corrections (for e. g. methyl groups) were implemented. After that, the next round of structure calculations was started. These calculation cycles included NOE distance restraints from which 788 were interresidual (i–j = 0), 455 sequential (i–j = 1), 264 medium range (i–j < 5) and 175 long range (i–j > 5). Also included was the complete peak list of all unassigned NOE signals that were used by the CYANA implemented program CANDID (Herrmann et al., 2002) to automatically suggest additional NOE correlations. These proposed correlations were checked manually again after the structural calculation. The NOE restraints lists were cross checked for the assignments of peaks that were violated in more than half of the structures with the lowest target function. In later structural calculation cycles the program ARIA 1.2 (Ambiguous Restraints for Iterative Assignment, Linge et al., 2003) was used to include solvent refinement.

Structural restraints: coupling constants

For further structure refinement ϕ angle restraints were determined from a 3D HNHA experiment. Given the obtained spectral quality, 76 HNHA coupling constants could be determined from the ratio of the respective cross and diagonal peak intensities. No coupling constants could be measured for the first 5 residues in *T. maritima* RbfA and the C-terminal residues I100 to D120 due to fast proton exchange with the solvent and signal overlap. Several residues—particularly those in the loop regions connecting the secondary structural elements—showed H_N-Hα correlations, which were very weak or absent.

The ϕ angles were predicted from the measured coupling constants using the Karplus equation:

$$J(\phi) = A \cos^2(\phi - 60) + B \cos(\phi - 60) + C$$

The J values were then introduced in the structural calculation program ARIA 1.2 which used implemented parameterization factors (A = 6.98, B = –1.38, C = 1.72).

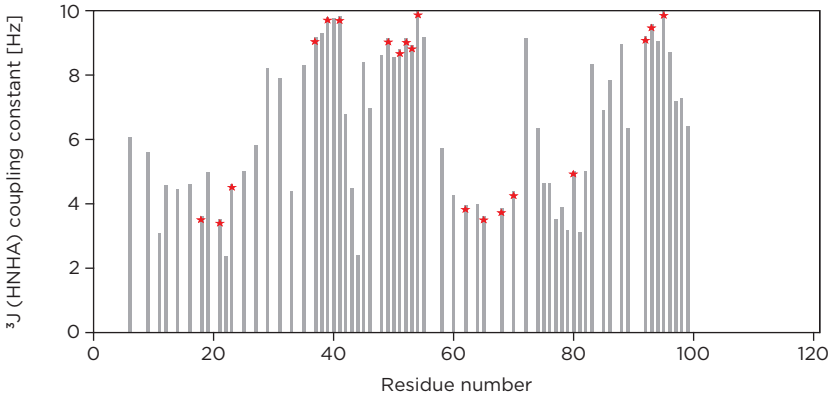


Figure 3.14
Plot of experimentally determined ^{3}J (HNHA) coupling constants against the sequence. The values were extracted from a 3D HNHA experiment recorded at 25° C on a 700 µM protein sample in acetic buffer at pH 4.5. The stars mark the residues that are protected against H/D exchange.

Structure refinement: residual dipolar couplings

Recently, developed methods for the partial alignment of macromolecules provide additional information to overcome the limitations of classical structure determination. The classical structure determination relied mainly on the short ranged (<5 Å) proton-proton distances (as described above) or torsion angle restraints (as described above). Residual dipolar couplings (RDCs) represent a distinct class of experimental NMR data that provide long-range angular information that complements the Nuclear Overhauser Effect (NOE) distance and *J*-coupling torsion angle restraints. If a macromolecule experiences restricted orientation sampling, due to the presence of a liquid crystal or due to the paramagnetic properties of the molecule, strong first order interactions, such as chemical shift anisotropy or dipolar coupling, are no longer averaged to zero as is the case in isotropic solution (Saupe and Englert, 1963; Gayathri et al., 1982). The long-range structural information derived from these RDCs can be used to improve the accuracy and precision of NMR-based structure determinations. A set of ^{1}Hα-^{13}Cα and ^{1}H-^{15}N RDCs were measured on *T. maritima* RbfA samples. These RDCs were used as structural restraints in the refinement of the *T. maritima* RbfA structure. Partial alignment of the (^{15}N,^{13}C) or ^{15}N-labeled RbfA samples was achieved using 8 % polyacrylamide gels. Deuterium quadrupole splittings of ~10.0 Hz were measured for the ^2H_2O resonances in the (^{15}N, ^{13}C) and ^{15}N-labeled RbfA samples. One bond ^{1}Hα-^{13}Cα RDCs were measured for 91 residues and ranged from −36 Hz to 31 Hz. A total of 96 one bond ^{1}H-^{15}N imino RDCs were measured in the ^{15}N-labeled RbfA sample and ranged from −16 Hz to 16 Hz.

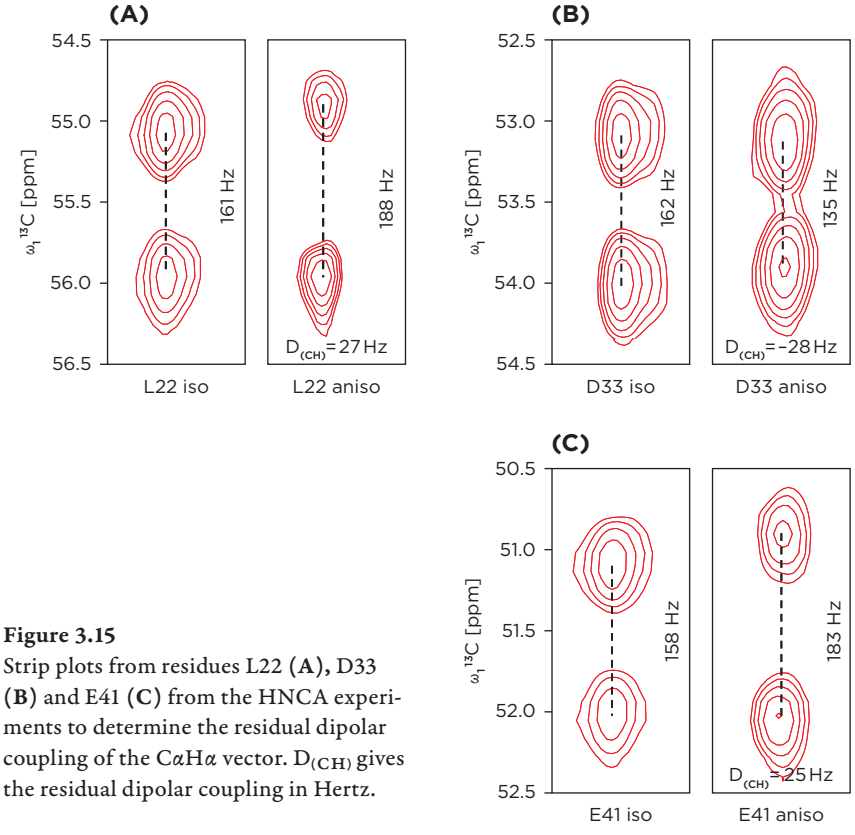


Figure 3.15
Strip plots from residues L22 **(A)**, D33 **(B)** and E41 **(C)** from the HNCA experiments to determine the residual dipolar coupling of the CαHα vector. $D_{(CH)}$ gives the residual dipolar coupling in Hertz.

A grid search procedure using the program Module (Dosset et al., 2001) that arrayed D_a (axial component of the alignment tensor) and R ($R = D_r/D_a$; D_r rhombic component of the alignment tensor) was used to determine the best fit of the alignment tensor for a set of RDCs and the experimentally determined structure of *T. maritima* RbfA derived from the NOE data. The significant differences between the experimental and predicted RDCs indicated that the local and/or global structure of the *T. maritima* RbfA was not represented well by the NOE-based structures. Therefore, the RDC lists derived from the two datasets were included in the ARIA structure calculation procedure.

The RDCs from the two datasets corresponding to flexible residues within the loop regions of the protein (as determined by R_1, R_2, and ^{15}N-{^1H}-Het-NOE data) were not included in the structural refinement. During the cycles of structural calculation the difference between the experimental and predicted RDCs reached a minimum and the plots shown in Figure 3.17 B and C have correlation factors of 0.987 for the H_N-N residual dipolar couplings and 0.988 for the Hα-Cα RDC, respectively.

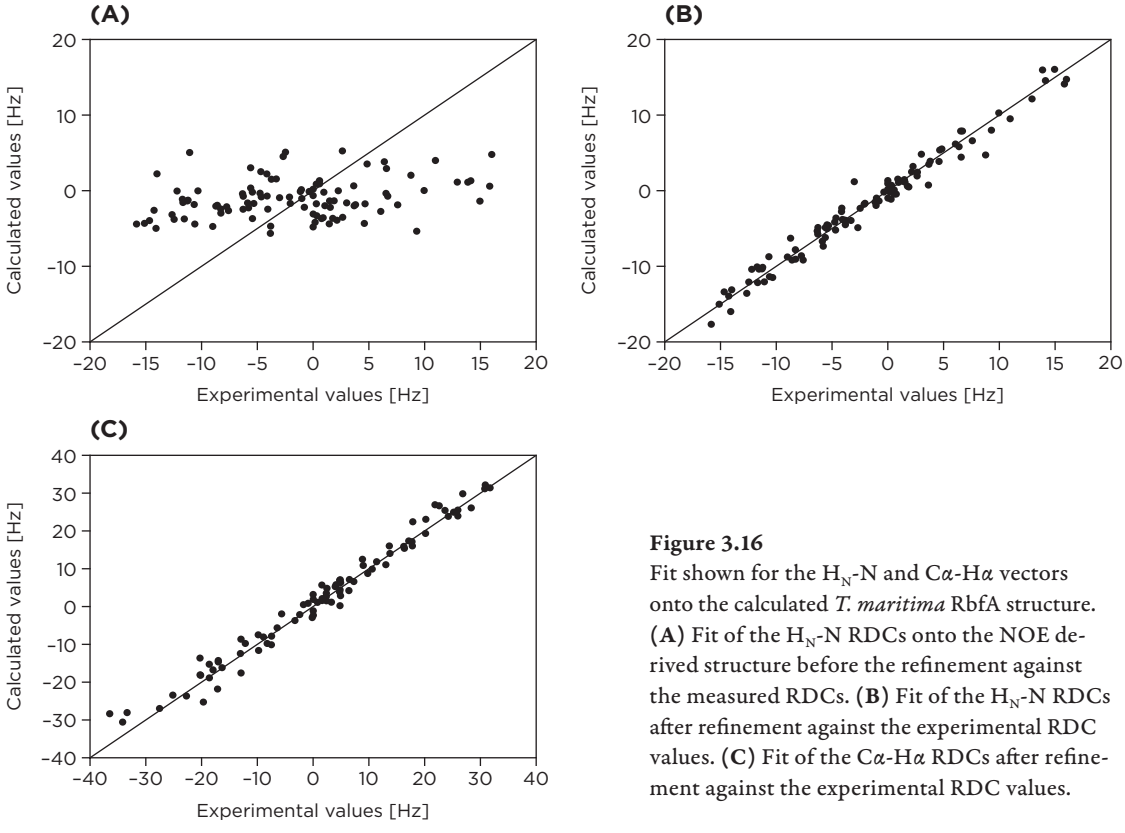


Figure 3.16
Fit shown for the H_N-N and Cα-Hα vectors onto the calculated *T. maritima* RbfA structure. **(A)** Fit of the H_N-N RDCs onto the NOE derived structure before the refinement against the measured RDCs. **(B)** Fit of the H_N-N RDCs after refinement against the experimental RDC values. **(C)** Fit of the Cα-Hα RDCs after refinement against the experimental RDC values.

Final structure calculation

The experimentally derived distance, torsion angles and dipolar coupling restraints described above were used for the final structure calculation using ARIA 1.2 (Linge et al., 2003). The method is called Ambiguous Restraints for Iterative Assignment (ARIA). The starting point for ARIA was an almost complete assignment list of the proton chemical shifts, and a list of assigned and/or unassigned NOEs, consisting of short range and long range NOEs. The restraint list was then augmented by automatically interpreting peak lists generated by the manually picked peaks. The central task of ARIA was the assignment of NOEs during the structure calculation using a combination of unambiguous and ambiguous distance restraints and an iterative assignment strategy. In addition, ARIA calibrates the NOE intensities to derive distance restraints, merges overlapping data sets to remove duplicate information, and uses empirical rules to identify erroneous peaks (Nilges et al., 1997). Many

NOE contacts were already assigned unambiguously in the course of resonance assignment, such as intra-residue and sequential NOEs and many easily identifiable long-range contacts. These NOEs were sufficient to define the overall fold of the *T. maritima* RbfA molecule. The main aim of the calculation strategy was to extract in a semi-automatic fashion the information from the available NOE spectra necessary to define a refined structure. In the first round of calculations (iteration zero), an initial ensemble of structures was calculated from an extended random coil starting structure based on the manually prepared list. Any ambiguities in this list were treated as ambiguous distance restraints. The lowest energy structures of this ensemble were selected for use in the iterative assignment. Each of the following iteration began with ordering the ensemble from the previous iterations with respect to total energy, and selecting the structures with lowest total energy as the basis for interpreting the spectra. All the restraints extracted from the spectra were analyzed for restraint violations in the chosen structures. Any restraint that was systematically violated was removed from the list. After that a new set of structures was calculated. The whole procedure was iterated until all restraints were accepted. Manual inspection of the lists of rejected restraints and peaks was done to locate errors or degenerate assignments. The final list of restraints was also inspected and modified by hand, and a set of final structures was calculated. A short molecular dynamics trajectory in a thin layer of solvent (water) refined the final structure ensemble (Linge et al., 2003).

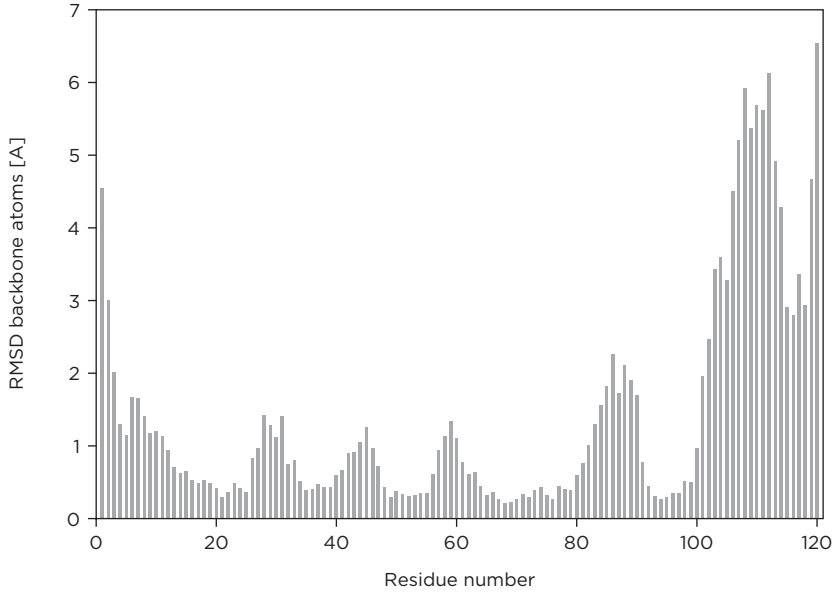


Figure 3.17
Plot of the average backbone RMSD for each residue from the mean. The values shown are the average RMSD calculated from the ensemble of the 20 lowest energy structures.

Hydrogen bonds as set parameters were included in the later structural refinement steps. The hydrogen bonds were identified by the data of the $^{2}H_2O$ exchange experiment and were cross checked by the structure calculation when consecutive calculations showed consistently a hydrogen bond at the respective position and these positions correlated with the earlier identified secondary structural elements. The final structure calculation cycles then included NOE distance restraints from which 731 were interresidual (i–j = 0), 675 sequential (i–j = 1), 438 medium range (i–j < 5) and 597 long range (i–j > 5); 66 HNHA coupling constants; 96 one bond ^{1}H-^{15}N RDC and 91 $^{1}H\alpha$-$^{13}C\alpha$ RDC; and 12 identified hydrogen bonds. This corresponds to an average of 26 restraints per residue.

Figure 3.18
NMR structural ensembles of the 10 lowest energy structures out of an ensemble of 80 final structures. This figure was generated using the program PyMOL.

The ensemble of calculated structures gave an overall backbone root mean square deviation of 0.77 Å. By plotting the individual RMSD values per residue (Figure 3.18), the well defined regions were found to include the secondary structural elements and parts of the loops connecting the α-helices and β-strands. Less well defined were the N and C-terminal regions. This was due to the fact that these portions of the protein seems to be highly flexible and only very limited NOE data could be analyzed for these regions. The high RMSD values for the loop region between β-strand β1 and β2 (residues 42–48) also was due to the lack of sufficient NOE data. In addition, the loop

region between a helix α3 and β-strand β3 (residues 85–90) is flexible (see *Backbone dynamics*) and the resulting signal broadening led to the identification of only a limited number of NOE correlations.

The quality of the structure is additionally judged by other structure related parameters. The Ramachandran plot analysis of the backbone dihedral angle distribution led to the following result; for the final structure ensemble 87 % of the residues are in most favoured or favoured regions and the remaining 13 % are in the generally allowed regions.

A summary of the experimentally derived restraints and the structure calculation statistics is given in Table 3.1.

Table 3.1
Structural parameters and restraints used in the structural calculation.

Number of experimental restraints	
Intra-residue unambiguous NOEs	731
Sequential unambiguous NOEs	675
Medium-range unambiguous NOEs	438
Long-range unambiguous NOEs	597
Total unambiguous NOEs	2,441
Total ambiguous NOEs	144
Dihedral angles (a)	66
Hydrogen bonds (b)	12
Residual dipolar couplings	
HN-N	96
Cα-Hα	91
RMSD [Å] from the mean	
All backbone atoms (1–101)	0.771
All heavy atoms (1–101)	1.333
Secondary structure backbone atoms (c)	0.404
Secondary structure heavy atoms (c)	0.999
Non bonded energy values after water refinement [kcal mol-1]	
E vdW	387.6 ± 34.2

RMSD from idealized covalent geometry	
Bonds [Å]	0.005 ± 0.0001
Angles [°]	0.807 ± 0.022
Impropers [°]	0.995 ± 0.043
RMSD from experimental data	
Distance [Å]	0.053 ± 0.012
Dihedral [°]	2.403 ± 0.541
Ramachandran analysis (d)	
Residues in most favored region [%]	59
Residues in additional allowed regions [%]	34
Residues in generously allowed regions [%]	5
Residues in disallowed regions [%]	2
G-factor overall	−0.401 ± 0.021

(a) Dihedral angle restraints are derived from ^{3}J (HNHA) measurements.
(b) The hydrogen bond restraints are derived from hydrogen-deuterium exchange experiments.
(c) Secondary structural elements comprise residues 11–25, 35–41, 49–55, 61–71, 75–84 and 92–98.
(d) The Ramachandran pot analysis was done for residues 1–102.

The final structure ensemble is shown in Figure 3.19. The structure of the *T. maritima* RbfA consists of an α + β two layer sandwich with an α – β plate topology structure, exhibiting regular secondary structure elements in the order α1-β1-β2-α2-α3-β3 along the sequence. All three helices are located on one side of the β-sheet. The helices α1 and α3 are organized with antiparallel helix packing. The strands of the β-sheet are ordered along the sequence, in which strands β1 and β2 are antiparallel and strands β2 and β3 are parallel. The β1-strand contains a β-bulge at position S38, this position is conserved in the RbfA family as either a serine or threonine residue. The strands β1 and β2 are linked by a seven amino acid long loop, which consists mostly of charged residues K44, D45, K46 and R47.

The aforementioned feature of the RbfA family, the kink region between helices α2 and α3, with the highly conserved AXG motif is formed by the residues A72-K73-G74. Residue A72 is the only residue in the polypeptide segment E61–L84, which shows non-helical Cα chemical shift values (see Figure 3.10) as well as a ^{3}J (HN-HA) coupling constants that correspond to a β-sheet torsion angle ϕ (Figure 3.15). The structure ensemble shows the kink between the two helices α2 and α3 which orients them at an angle of ~115° relative to each other.

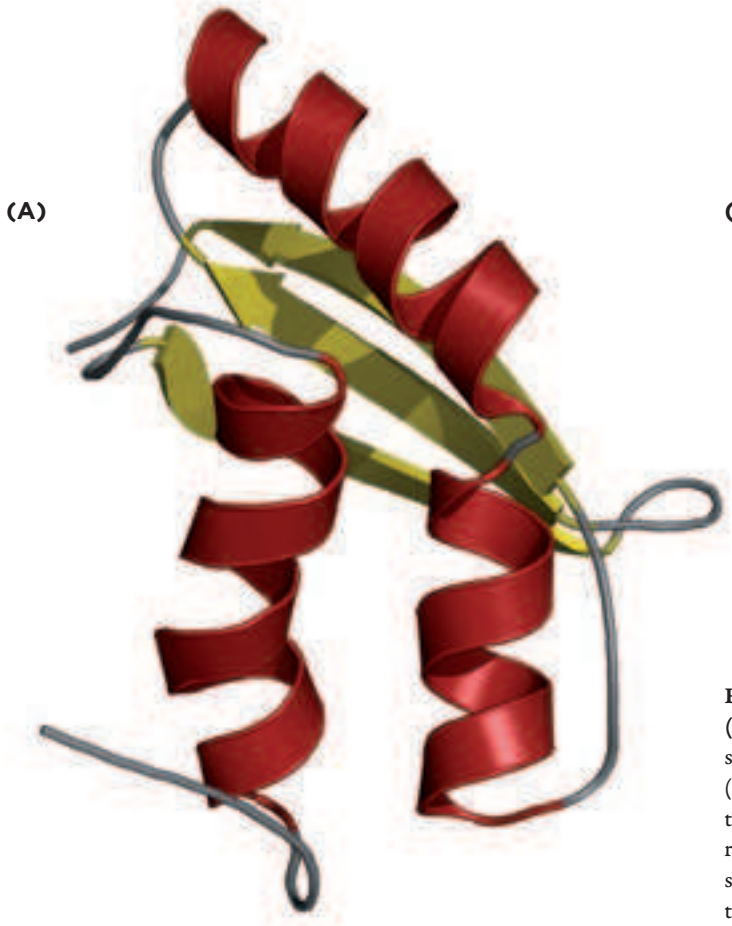


Figure 3.19
(A) Cartoon representation of the NMR solution structure of *T. maritima* RbfA (Δ11). The structure is composed of three helices (α1, α2 and α3 shown in red) and three β-strands (β1, β2 and β3 shown in yellow). **(B)** Schematic plot of the secondary structure.

Backbone dynamics

In order to characterize the apparent internal dynamics of the RbfA protein, the three relaxation parameters R_1, R_2, and ^{15}N-{^{1}H}-Het-NOE have been measured for *T. maritima* RbfA at temperatures of 30° C and 40° C. The correlation time for overall tumbling, τc (assuming isotropic tumbling) was estimated from the R_2/R_1 ratio to be 4.1 ns (± 0.2 ns) for *T. maritima* RbfA at 40° C and 4.2 ns (± 0.2 ns) at 30° C. Experimental values and uncertainties have been determined as described in experimental procedures. Graphs of the relaxation parameters versus residue numbers are shown in Figure (3.20). Relaxation parameters are generally constant along the secondary structure elements, as expected for a rigid structure, whereas at the N and C-terminal regions relaxation parameters well below the mean values can be identified. Furthermore, at the C-terminus, negative values of the ^{15}N-{^{1}H}-Het-NOE were found at higher temperature (Figure 20 C).

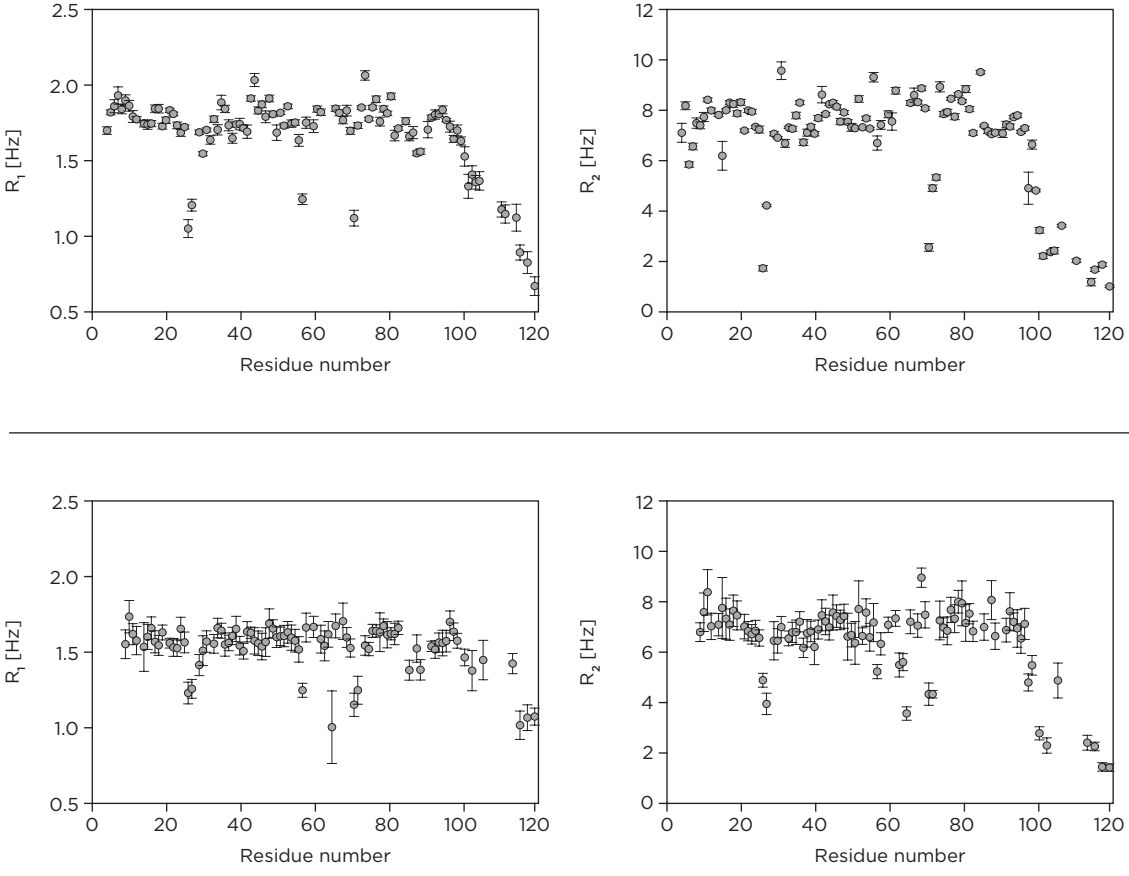
The resonances corresponding to residues M1–M9 of *T. maritima* RbfA, exhibit R_1 and R_2 values that are faster than the average protein backbone values. The ^{15}N-{^{1}H}-Het-NOE values in those regions are lower than the

average ^{15}N-{^{1}H}-Het-NOE values. Similarly, the carboxy termini of the protein showed lower than average ^{15}N-{^{1}H}-Het-NOE values, faster R_2 rates and faster R_1 rates. This indicated that the C- and N-termini in the truncated *T. maritima* RbfA are highly flexible. Within the core of the protein most residues exhibit an average ^{15}N-{^{1}H}-Het-NOE value of 0.72 for *T. maritima* RbfA at 40° C and 0.65 at 30° C. There are several residues that differ from this average value. Residues Arg 26 and Asp 27 at the C terminal end of α1 helix have lower ^{15}N-{^{1}H}-Het-NOE values, at both 30° C and 40° C. Residues G57, the kink region residues (Arg 71 to Lys 73) and the residues preceding β-strand β3 (Leu 86 to Ala 89) showed the same decrease in ^{15}N-{^{1}H}-Het-NOE values.

The ^{15}N-{^{1}H}-Het-NOE values at 30° C were lower than average for residues K63 to T65. The same residues could not be analyzed in the 40° C experiment due to signal overlap. The aforementioned residues also show faster than average R_1 relaxation and faster than average R_2 relaxation. Surprisingly, the loop connecting the β1- and β2-strands showed no indication of faster relaxation then the secondary structures surrounding it. Therefore, the previously observed high RMSD values for this region are not a result of high internal flexibility (see Figure 3.18). Overall the relaxation data for the *T. maritima* RbfA at the two temperatures of 30° C and 40° C show the same flexible features within the two proteins. Some of the highly dynamical residues seen at 30° C could not be analyzed at 40° C due to signal overlap. The relaxation rates at 30° C are slightly lower than the relaxation rates measured at 40° C.

Summary

The gene coding for the *T. maritima* RbfA (1–120) protein was successfully inserted in a vector that made it possible to overexpress the protein heterologuously in *E. coli* cells. The yield of the protein was sufficient to prepare the necessary samples for the various biophysical experiments. With the established purification protocol the RbfA protein could be purified to an extend that made it amenable for structural determination using NMR spectroscopic techniques. Using the empirically determined buffer conditions, all NMR experiments for the assignment of the backbone and side chain resonances could be performed and resulted in an almost complete assignment of the protein. Under the given buffer conditions is was also possible to derive structural restraints, which could be used to calculate the solution structure



of *T. maritima* RbfA. Hence the 3D solution structure of *T. maritima* RbfA (Δ11) is presented with an RMSD value from the mean of 0.4 Å, for the well defined core of the protein. The structure shows strong similarities with the type II K-homology domain folds. The RbfA protein exhibits the secondary structural arrangement αββαaβ and a three dimensional fold, which is similar to other KH domain proteins. The KH modules are widespread RNA-binding motifs that have been characterized largely by sequence similarity (Ashley et al., 1993; Siomi et al., 1993; Gibson et al. 1993). KH modules often occur as multiple tandemly repeated domains, or in combinations with other RNA-binding domains. Structural analyses of KH modules have revealed that they actually occur with two different topological folds: type I folds with secondary structural elements arranged as βααββα (see Figure 3.21 A) and type II folds with secondary structural elements arranged as α(α)ββαaβ (see

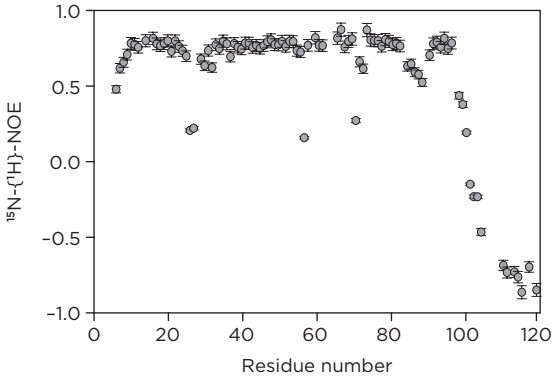
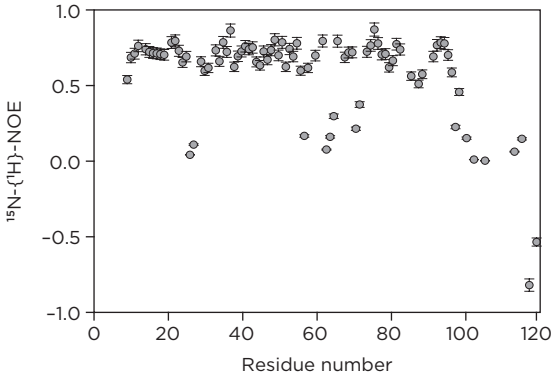


Figure 3.20
Parameters defining the backbone dynamics of the RbfA proteins from *T. maritima* (700 μM samples pH 4.5). The relaxation parameters are plotted as a function of the residue numbers.

(A) Plot of R_1, R_2 and ^{15}N-{^{1}H}-Het-NOE from *T. maritima* at 40° C.



(B) Plot of R_1, R_2 and ^{15}N-{^{1}H}-Het-NOE from *T. maritima* at 30° C.

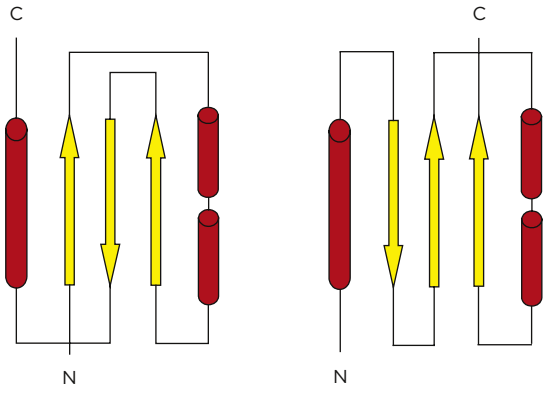
Figure 3.21 B) (Grishin, 2001). In both fold classes, the KH sequence motif shows the βααβ structural motif, at the core of which is a helix-turn-helix motif with a conserved GXXG sequence. The RbfA protein is a KH type II domain with the notable feature that the GXXG sequence is replaced by the conserved AXG sequence including the strongly conserved A72 residue. This A72 residue is forming an interhelical kink with an angle of about 115°, which is similar to the helix-turn-helix angle of other KH domains of about 100°–120° (Grishin, 2001).

The β-bulge, which is conserved throughout the RbfA family, is located at residue S38 in the β1-strand of the *T. maritima* RbfA. The chemical shift data and the HNHA coupling constants did not give a clear indication for the presence of this β-bulge in the *T. maritima* RbfA β1-strand. However, the NOE data and the experimentally derived residual dipolar couplings clearly

indicated that the β-bulge is formed by residue S38. This β-bulge in the β1-strand is also observed in β-strands of other structurally characterized type II KH domain folds. These include the KH domains of NusA, ribosomal protein S3 and the Era GTPase (Worbs et al., 2001; Wimberly et al., 2000; Chen at al., 1999).

The type II KH folds seems to be used by several protein families, which possess similar biological functions involving cold-shock adaptation and/or RNA-binding. In this context, it is tempting to speculate that type I and type II KH fold topologies may have arisen from a common ancestral KH protein by structural rearrangements in the course of evolution (Grishin, 2001). Protein domains with type II KH folds are known to interact with other protein fold motifs, including OB-fold proteins and GTPase catalytic domains (Gopal et al., 2001; Wimberly et al., 2000). Thus the KH fold appears to be a protein scaffold capable of mediating both protein-RNA and protein-protein interactions.

Figure 3.21
KH domains have two folding topologies. **(A)** Type I fold βααββα. **(B)** Type II fold αββααβ. The KH sequence motif is the βααβ structural unit. At the core of this motif is a helix-turn-helix structure.



Structures of a large number of RNA binding proteins have been determined. These include several cold shock proteins, which are interacting with DNA or RNA ligands. The CspA and CspB proteins form a closed β-barrel consisting of five β-strands with a large number of positively charged residues. This fold, consisting of an antiparallel β-barrel, was initially identified in proteins that bind oligonucleotides or oligosaccharides, hence the name the OB fold (Murzin, 1993). A majority of OB fold proteins are known to bind nucleic acids such as several identified cold shock proteins (Schindelin et al., 1993). Another very similar fold also from a cold shock protein is the cold shock domain (CSD), which was identified in the cold shock protein CsdA. Other members of the cold shock protein family like the NusA protein have a KH domain fold. Therefore, it seems that the majority of cold shock proteins involved in DNA and RNA binding mainly use two different protein folds to interact with their respective DNA or RNA ligands. Some of the cold shock proteins like RbfA and NusA have a KH domain fold whereas many other cold shock proteins utilize the OB fold.

3.3 The protein ribosomal binding factor A from Helicobacter pylori

Purification of H. pylori RbfA

The RbfA protein from *H. pylori* was overexpressed in ^{15}N or $^{15}N/^{13}C$-labeled form in *Escherichia coli* BL21(DE3) cells. Isotopically labeled protein was produced in cells grown in M9 minimal medium as described above (experimental procedures 2.1).

A protocol was established for the purification of *H. pylori* RbfA (see experimental procedures 2.2). The protein expression rate was tested at certain times after IPTG induction by SDS-PAGE electrophoresis. An optimal overexpression time of 3.5 h at 37° C after induction was determined. After harvesting the cells were resuspended in buffer containing 25 mM KH_2PO_4 pH 6.5, 50 mM KCl and sonicated for cell disruption. The supernatant, which contained the soluble RbfA protein, was further purified by a cation exchange chromatography step on a SP-sepharose column (Amersham). To separate the RbfA protein from the *E. coli* proteins a salt gradient from 25 mM KCl to 500 mM KCl was used. The *H. pylori* RbfA elutes at a salt concentration of 250 mM KCl.

An additional step on a Sephacryl S-100 gel filtration column (Amersham) was used to separate the full length RbfA protein. Protein purified by this procedure was >95 % pure and showed an apparent molecular weight of 12.5 kDa on a SDS-PAGE and was verified by MALDI-MS. The overall yield of the labeled *H. pylori* RbfA protein was ~4 mg Protein/L [M9].

Assignment of H. pylori RbfA

The backbone assignment for the *H. pylori* RbfA was determined as described above for the *T. maritima* RbfA protein. The NMR sample was dialyzed into buffer containing 25 mM KH_2PO_4 pH 6.5, 50 mM KCl, with a final protein concentration of 0.7–0.8 mM. Higher concentration led to aggregation after several days of storage.

Under the chosen buffer conditions at 25° C the ^{1}H-^{15}N-HSQC (Figure 3.23) spectra showed the required signal dispersion. For this 111 amino acid protein, which contained one proline residue at position P81, 109 back-

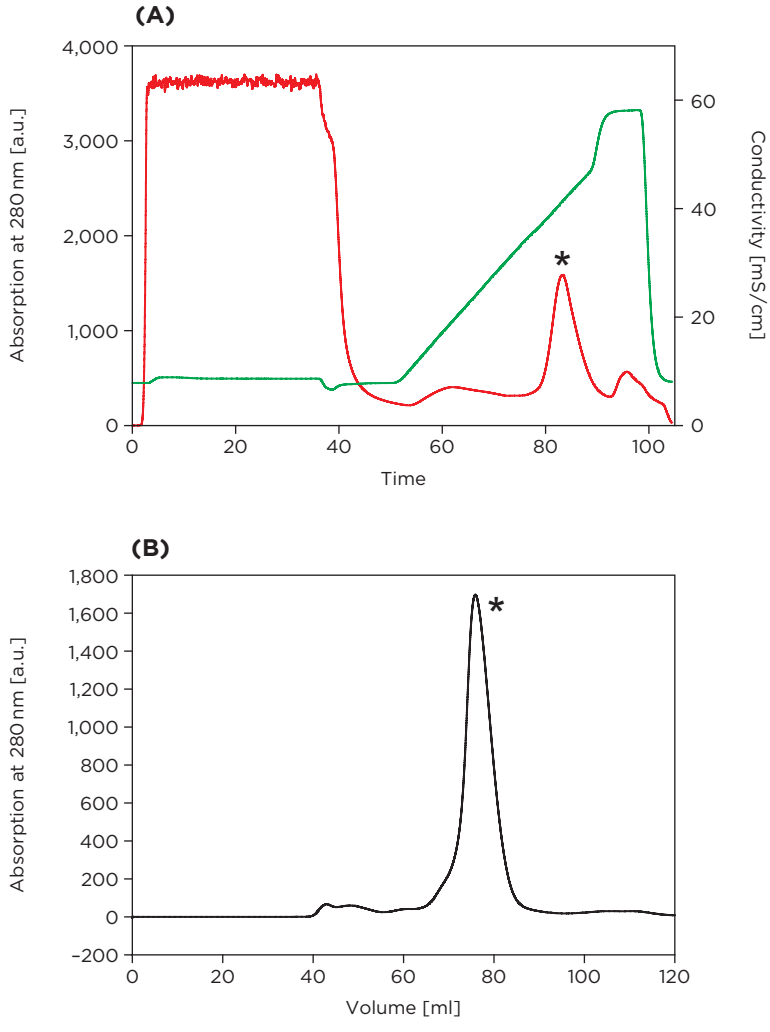


Figure 3.22
(A) Chromatogram of *H. pylori* RbfA using a cation exchange column (Amersham) with SP-Sepharose resin. The column was loaded with cell extract. Red trace: relative absorption at 280 nm; green trace: conductivity. Bulk absorbance seen immediately after the sample is loaded are DNA and RNA of the cell extract as well as all negatively charged or non-charged proteins. Positively charged proteins are eluted from the column by the application of a salt gradient from 25 mM KCl to 500 mM KCl. The asterisk marks protein containing fractions of *H. pylori* RbfA.
(B) Chromatogram of *H. pylori* RbfA using a gel filtration column (Amersham) with Sephacryl resin. The column was loaded with the fractions from the cation exchange column concentrated to a volume of 5 mL. Black trace: relative absorption at 280 nm. The asterisk marks the eluted fractions of *H. pylori* RbfA.

bone amide signals were expected from which 101 resonances were resolved. This value includes the side chain signal from a tryptophan residue at position W77. The expected resonances from the side chain amino groups of 7 glutamine and 4 asparagine residues were also detected.

The same protein backbone assignment strategy was employed for the *H. pylori* RbfA as described for the *T. maritima* RbfA. This comprised the three dimensional NMR experiments HNCA, HNCACB, HNCO and CBCA-CONH. Again, the HNHA and HBHA(CBCA)CONH experiments were used to identify the proton assignments.

From the conducted 3D NMR experiments the assignment was completed to the following extent: 91 % Cα; 90 % Hα; 88 % Cβ; 86 % Hβ; 93 % N_H, 93 % H_N and 93 % CO. No signals were found corresponding to residues N2, K42, H43, H44, Q73, F78, K79 and C80. From the assignment of the

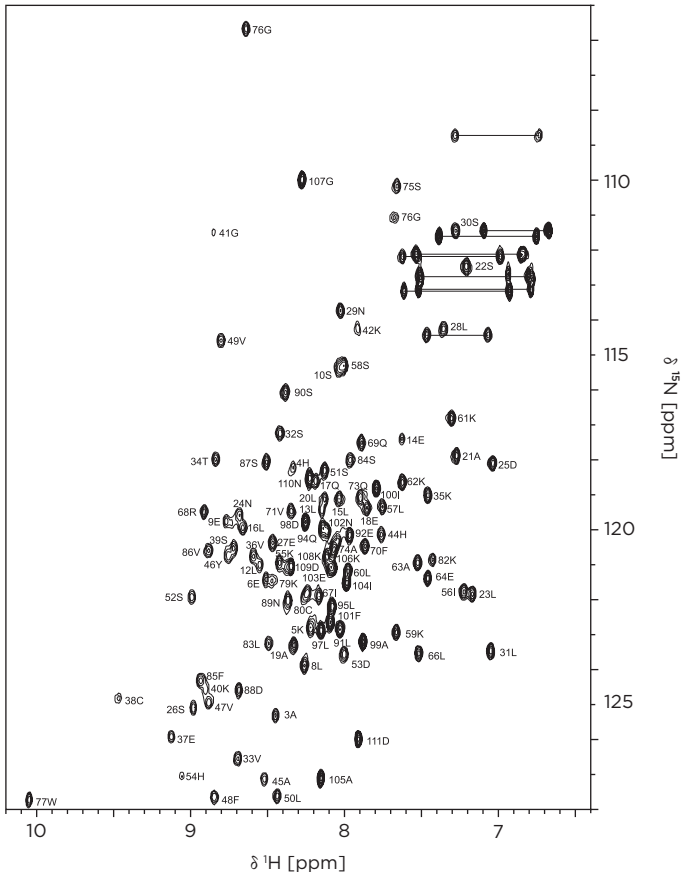


Figure 3.23 (^{1}H-^{15}N)-HSQC spectra of *H. pylori* RbfA from a ^{15}N labeled sample in NMR buffer (25 mM KH_2PO_4 at pH 6.5, 50 mM KCl) at 25° C on a 700 MHz Bruker spectrometer. The side chain amide groups of asparagine and glutamine are marked by lines. The peak labels correspond to the assignment as extracted from the assignment experiments.

backbone residues of the *H. pylori* RbfA protein, its secondary structure was predicted. A plot of the Cα and Hα chemical shift indices of *H. pylori* RbfA is shown in Figure 3.24. Analysis of the secondary structure indicated that the *H. pylori* RbfA is a mixed α/β protein.

The secondary structure is comprised of three α-helices and three β-strands. The secondary structure elements are ordered α1-β1-β2-α2-α3-β3 along the sequence. Residue A63 is the only residue in the polypeptide segment of helices α1 and α2, which shows non-helical chemical shift values. Residue T36 which is within the β1-strand region is the only residue in this

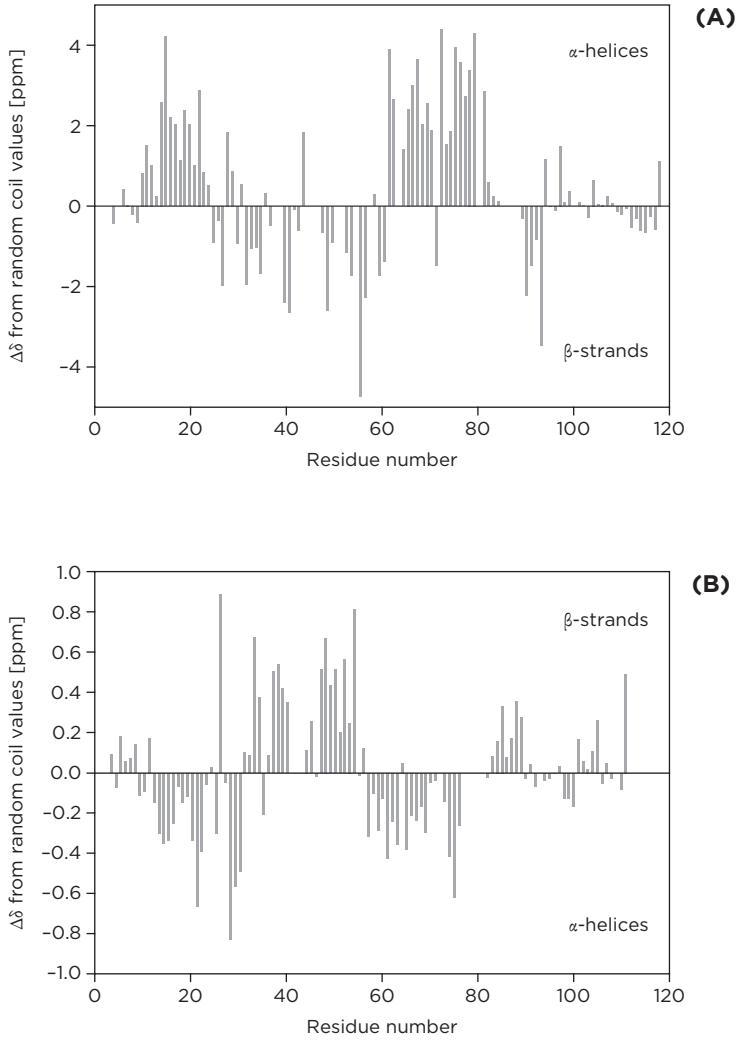


Figure 3.24
Sequence and Δσ chemical shift values of Cα **(A)** and Hα **(B).** The chemical shift values were calculated by subtracting the assigned Cα and Hα values from chemical shifts of residues in random coil conformation.

region that shows a non β-strand chemical shift value. A list containing the chemical shifts for the backbone resonances of *H. pylori* RbfA is placed in the appendix (see Table A.4).

In order to identify hydrogen bond interactions samples of the *H. pylori* protein were freeze dried and dissolved in buffer containing 100 % $^{2}H_2O$. ^{1}H-^{15}N-HSQC spectra were recorded at certain time intervals and the signals were assigned according to the backbone resonance assignment. From this experiment the following residues were identified with the corresponding backbone amide proton involved in hydrogen bond interactions: E14, L15, L16, Q17, A19, L20 and A21 in helix α1, residues T34, K35, E37 and S39 in β-strand β1, residues F48, V49 and L51 in β-strand β2, L60, K61 and K62 in helix α2, residue Q73 in helix α3 and residues S84 and V86 in β-strand β3. From the experiments described above the corresponding secondary structure elements of *H. pylori* Rbfa were extracted and placed in the polypeptide segments Glu9–Leu23 (α1), Leu31–Lys40 (β1), Val47–His54 (β2), Leu57–Leu62 (α2), Leu66–Phe78 (α3) and Leu83–Leu91 (β3).

Measurement of coupling constants for H. pylori RbfA

Using a 3D HNHA experiment, ^{3}J (HNHA) coupling constants were determined for the *H. pylori* RbfA. Given the obtained spectral quality, 68 HNHA coupling constants could be determined from the ratio of the respective cross and diagonal peak intensities. No coupling constants could be measured for the three N-terminal amino acids, several residues in helix α1, the loop between β-strand β1 and β2, the beginning of β-strand β2 itself and the loop region succeeding helix α3 due to fast proton exchange with the solvent and

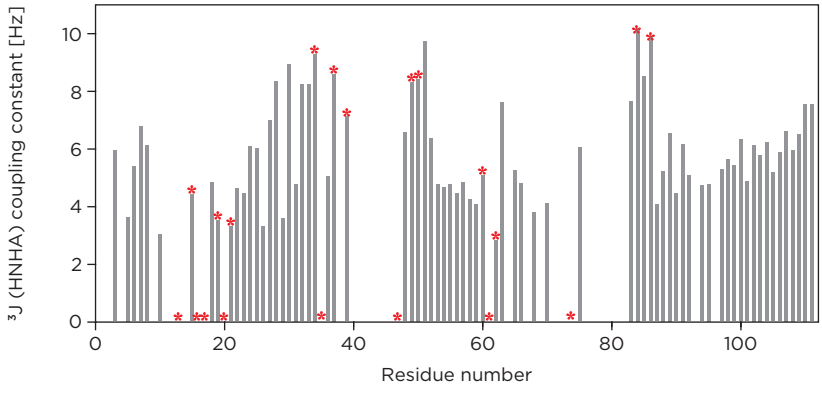


Figure 3.25
Plot of experimentally determined ^{3}J (HNHA) coupling constants against the sequence. The values were extracted from a 3D HNHA experiment recorded at 25° C on a 700 µM protein sample in phosphate buffer at pH 6.5. The asterisk mark the residues that are protected against H/D exchange.

signal overlap. In case of several residues, H_N-Hα correlations were too weak to be detected. This was especially true for signals of amino acids in the loop regions connecting the secondary structural elements.

Backbone dynamics

The three relaxation parameters R_1, R_2, and ^{15}N-{^{1}H}-Het-NOE have been measured for *H. pylori* RbfA at 25° C. The correlation time for overall tumbling, τ_C (assuming isotropic tumbling) of 4.6 ns (± 0.2 ns) at 25° C was estimated for *H. pylori* RbfA. Experimental values and uncertainties have been determined as described in Experimental procedures. Graphs of the relaxation parameters versus residue numbers are shown in Figure 3.26. Relaxation parameters are generally constant along the secondary structure elements, as expected for a rigid structure whereas in the N- and C-terminal regions relax-

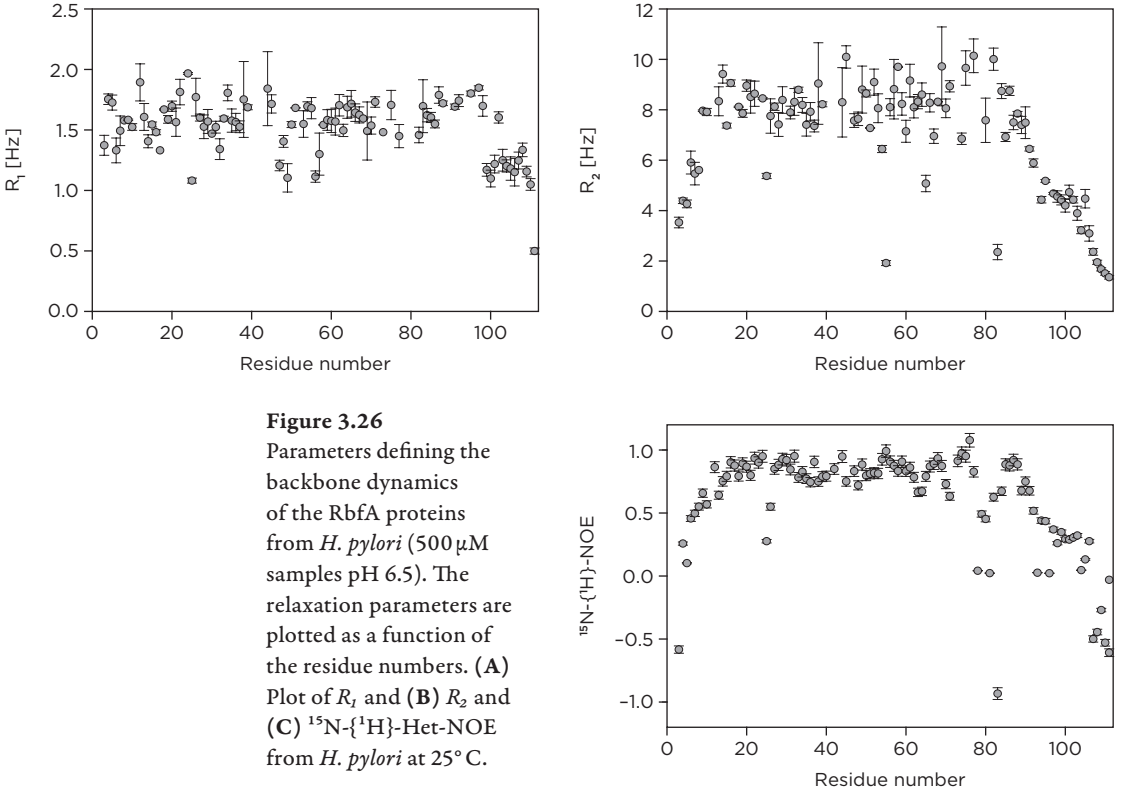


Figure 3.26
Parameters defining the backbone dynamics of the RbfA proteins from *H. pylori* (500 µM samples pH 6.5). The relaxation parameters are plotted as a function of the residue numbers. **(A)** Plot of R_1 and **(B)** R_2 and **(C)** ^{15}N-{^{1}H}-Het-NOE from *H. pylori* at 25° C.

ation parameters well below the mean values were found. The resonances of residues M1–S10 in *H. pylori* RbfA, exhibit R_1 and R_2 values, which are faster than the average protein backbone values. The ^{15}N-{^{1}H}-Het-NOE values in those regions are lower than the average ^{15}N-{^{1}H}-Het-NOE values. Similarly, the C-terminus showed lower than average ^{15}N-{^{1}H}-Het-NOE values and faster R_2 and R_1 rates.

This indicated that the N- and C-termini in the full length *H. pylori* RbfA are highly flexible.

Residues N24 and D25 of the *H. pylori* RbfA protein, at the end of helix α1, exhibit reduced values for the relaxation parameters. The kink region and the residues preceding the β3-strand showed flexibility in *H. pylori* RbfA. This was also found for the residues in the loop connecting the β-strands β1 and β2.

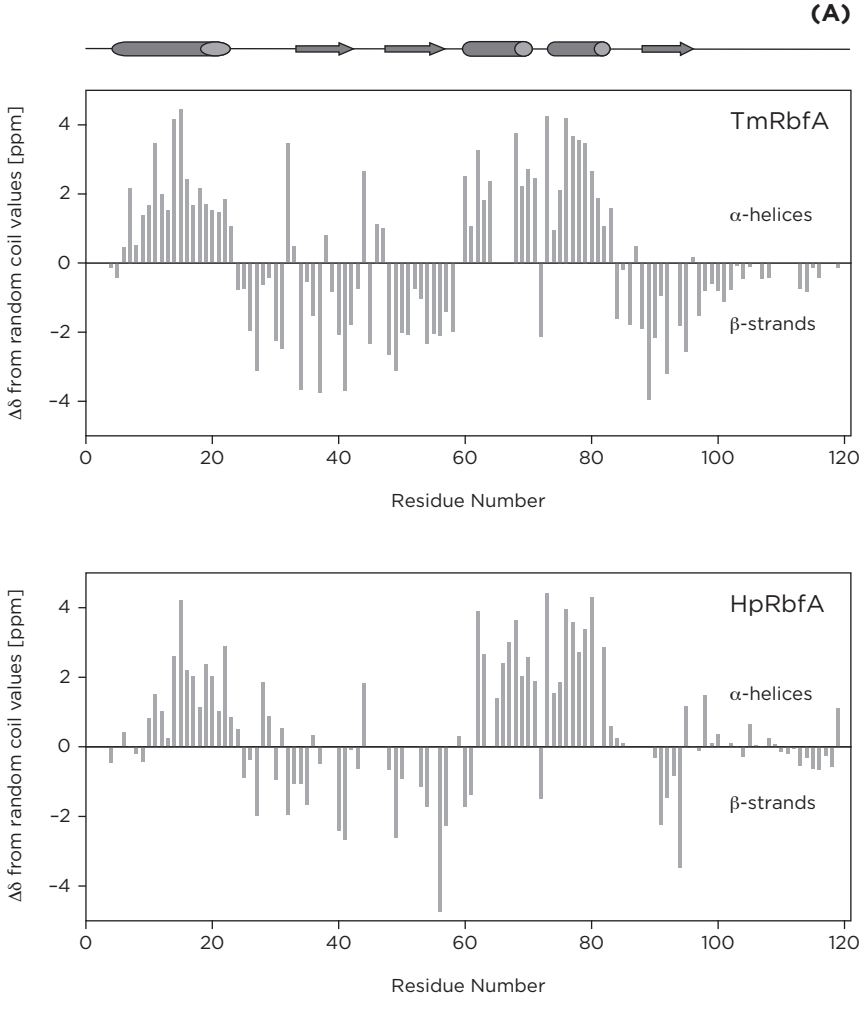
Comparison of T. maritima RbfA and H. pylori RbfA

Comparing the proteins

For both proteins the translational diffusion constants (D_s) were determined from pulsed field gradient (PFG) NMR methods (Stejskal and Tanner, 1965). The D_s was obtained by fitting the attenuation of the observed signals with increasing gradient strength. For *T. maritima* RbfA (14.1 kDa) a value of 1.29×10^{-6} cm^2/s was determined. For *H. pylori* RbfA (12.5 kDa) the experiment yielded a value of 1.39×10^{-6} cm^2/s. These results suggest that both proteins are in a monomeric, globular form in solution. The measured values are in agreement with other monomeric proteins (Ubiquitin (8.7 kDa) D_s = 1.49×10^{-6} cm^2/s, Lysozyme (14.5 kDa) $D_s = 1.08 \times 10^{-6}$ cm^2/s).

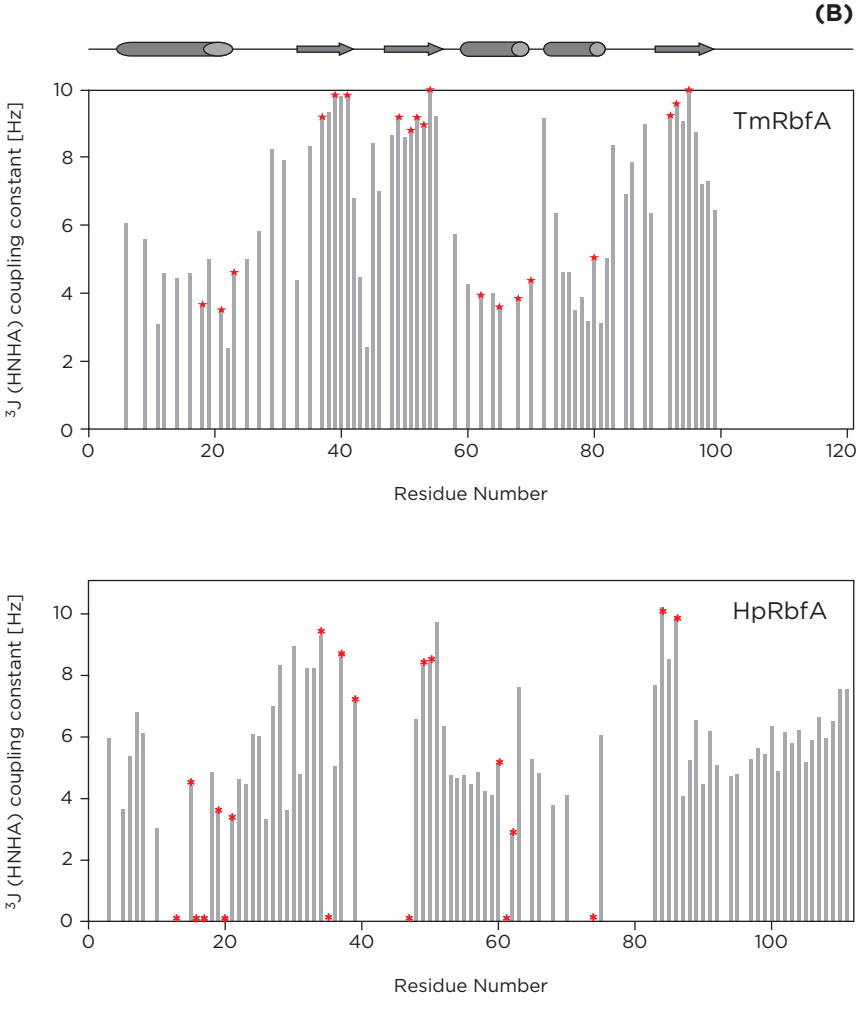
The backbone fold of the two RbfA proteins was determined using the chemical shift index analysis (CSI), ^{3}J (HNHA) coupling constants, extracted from 3D HNHA NMR experiments. ^{15}N edited NOESY spectra were used to identify the characteristic NOE signals of secondary structural elements. Figure 3.27 shows the comparison of the secondary structure of the two proteins. The loop regions are extended by one or two amino acids in the *T. maritima* RbfA. The β-strands are two amino acids longer in the *H. pylori* RbfA. The α1-helix has the same length in both proteins; in contrast the helix α2 is much longer in the *T. maritima* protein, whereas the helix α3 is longer in the *H. pylori* RbfA. Although the *H. pylori* RbfA sequence is shorter and the secondary structures differ in their length by a few amino acids, the arrangement of the secondary structure elements and the overall fold of the protein was

Figure 3.27
(A) Delta chemical shift values of Cα atoms from *H. pylori* (HpRfA) and *T. maritima* RbfA (TmRbfA). The position of the secondary structural elements in *T. maritima* RbfA is marked by the cartoon above. The chemical shift values were calculated by subtracting the assigned Cα values from random coil chemical shift values. The residue numbers are aligned to the *T. maritima* sequence.
(B) Plot of experimentally determined ^{3}J (HNHA) coupling constants versus the sequence for *H. pylori* (HpRfA) and *T. maritima* RbfA (TmRbfA). The residue numbers are aligned to the *T. maritima* sequence. An asterisk marks the residues that are protected against H/D exchange.



the same as in *T. maritima* RbfA. From these results it was concluded that the overall 3D structure of the *H. pylori* RbfA is similar to the *T. maritima* RbfA protein. Analysis of the ^{3}J (HNHA) scalar coupling constants revealed significant structural similarity even for subtle structural features such as the aforementioned β-bulge at the conserved residues S38 *(T. maritima)* and T36 *(H. pylori),* and the kink region between helices α2 and α3. This kink region is formed in the *H. pylori* RbfA by residues A63–G65 (*T. maritima* RbfA A72–G74) where residue A63 shows a non-helical coupling constant.

This suggests that the conformations of residues in or around the kink region consensus sequence are indeed very similar between the two structures.



Backbone dynamics

In order to characterize the apparent internal dynamics of the RbfA proteins, the three relaxation parameters R_1, R_2, and ^{15}N-{^{1}H}-Het-NOE have been measured for *H. pylori* RbfA at 25° C and *T. maritima* RbfA at temperatures of 30° C and 40° C. The correlation time for overall tumbling τ_C (assuming isotropic tumbling) was estimated from the R_2/R_1 ratio to be 4.1 ns (± 0.2 ns) for *T. maritima* RbfA at 40° C and 4.2 ns (± 0.2 ns) at 30° C. A τ_C value of 4.6 ns (± 0.2 ns) at 25° C was estimated for *H. pylori* RbfA. Relaxation parameters are generally constant along the secondary structure elements, as expected for a rigid structure. At the N- and C-terminal regions, relaxation parameters well below the mean values were identified. Furthermore, at the C-terminus negative values of the ^{15}N-{^{1}H}-Het-NOE were found at higher temperature (Figure 3.28).

The resonances of residues M1-M9 in *T. maritima* RbfA and M1-S10 in *H. pylori* RbfA exhibit R_1 and R_2 values, which are faster than the average protein backbone values. The ^{15}N-{^{1}H}-Het-NOE values in those regions are lower than the average ^{15}N-{^{1}H}-Het-NOE values. Similarly, the C-termini of the two proteins show lower than average ^{15}N-{^{1}H}-Het-NOE values and faster R_2 and R_1 rates. This indicated that the N- and C-termini in the truncated *T. maritima* RbfA and in the full length *H. pylori* RbfA are highly flexible.

The trend of the relaxation parameters and of the heteronuclear ^{15}N-{^{1}H}-Het-NOE values of the two proteins at different temperatures essentially depends on the secondary structure architecture of the protein scaffold. The analysis of the R_1 and R_2 values of the two proteins, along the examined temperature range, revealed that they are rather similar.

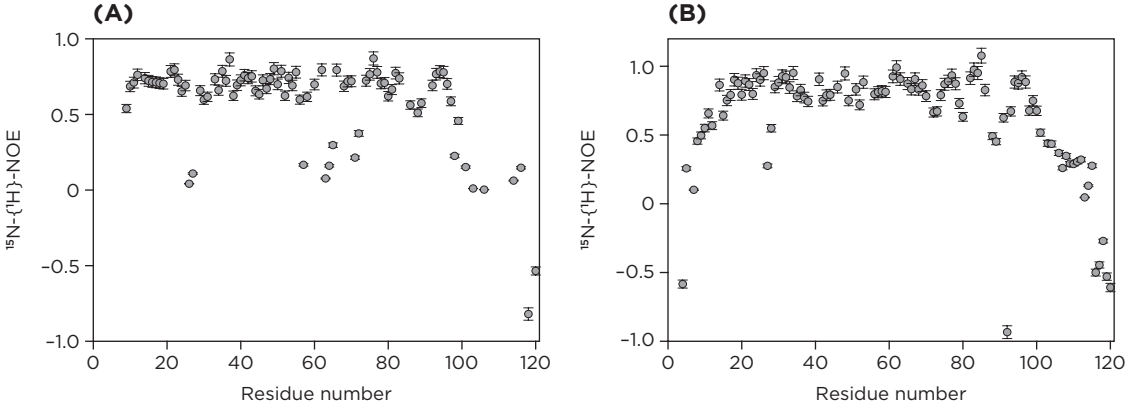


Figure 3.28
Parameters defining the backbone dynamics of the RbfA proteins from *T. maritima* (700 µM sample pH 4.5) and *H. pylori* (500 µM sample pH 6.5). The relaxation parameters are plotted as a function of the residue numbers.) **(A)** Plot of ^{15}N-{^{1}H}-Het-NOE from *T. maritima* at 30° C and **(B)** *H. pylori* at 25° C.

Summary

The construct for the *H. pylori* RbfA protein was successfully inserted in a vector that made it possible to overexpress the protein heterologuously in *E. coli* cells. The yield of the protein and its purification with the established purification protocol made it amenable for conducting high resolution NMR studies. Using the empirically determined buffer conditions, all NMR experiments for the assignment of the backbone could be performed and resulted in an almost complete assignment of the protein backbone. With the data derived from the conducted NMR experiments it was possible to compare the secondary structure, the overall fold and the dynamics of the two RbfA proteins from *H. pylori* and *T. maritima*. The comparison of the chemical shift data and coupling constants suggested that the overall secondary structural fold of the *H. pylori* protein is quite similar to the one of the *T. maritima* RbfA. Therefore, it was concluded that the three dimensional fold of the protein must be quite similar as well. This is supported by the notion that several structural features within the core of the protein are present in both proteins, like the β-bulge and the kink region between α-helices 2 and 3. Both proteins are in a monomeric and globular form in solution as was shown from their translational diffusion constants. The melting temperatures for both proteins verified their mesophilic *(H. pylori)* and thermophilic *(T. maritima)* character. The comparison of the dynamical behavior of the two proteins revealed even more similarities between the mesophile and thermophile RbfA protein. As mentioned above, the polypeptide segment that connects the α1-helix and the β1-strand constitutes part of a dynamical hot spot, where residues adjacent to the α1-helix show high flexibility in both proteins. This clustered hot spot is near a region with several charged residues. The corresponding polypeptide sequence of the KH domain of ribosomal protein S3 is involved in intramolecular protein-protein interactions with its own C-terminus, and the corresponding portion of the KH domain of the GTPase Era has protein-protein interactions with its N-terminal GTPase domain (Huang et al., 2003). Accordingly this region of the RbfA protein may be involved in RNA binding mechanism or in not yet defined protein-protein interactions, which might assist other proteins in forming functional complexes with RbfA and its targets. The region at the C-terminus of the α1 helix was previously identified as a dynamic hot spot in *E. coli* RbfA and could also be identified in the two RbfA proteins from *Thermotoga maritima* and *Helicobacter pylori.* The data from *E. coli* RbfA led to the conclusion that the region surrounding the

β-bulge belongs as well to this dynamic hotspot. The data for *T. maritima* RbfA and *H. pylori* RbfA do not support this notion. All of the resonances near the β-bulge had line widths typical for well ordered secondary structure elements and the relaxation parameters did not support the presence of a dynamical hot spot.

However, the kink region in both proteins and in RbfA from *E. coli* have a higher than average flexibility. In contrast, the region prior to α2-helix consisting of residues K63 to T65 (*T. maritima* RbfA numbering) was not shown to be highly flexible in the *E. coli* protein (Huang et al., 2003) Residues 90 to 93 in *E. coli* preceding the β3-strand showed the same high flexibility as seen for the *T. maritima* and *H. pylori* RbfA. These results suggested that the *E. coli* RbfA shares overall the same flexible regions within the protein core.

The data from *E. coli* RbfA led to the conclusion that the region surrounding the β-bulge belongs as well to this dynamic hotspot. The data for *T. maritima* RbfA and *H. pylori* RbfA do not support this notion. All of the resonances near the β-bulge had line widths typical for well ordered secondary structure elements and the relaxation parameters did not support the presence of a dynamical hot spot.

However, the kink region in both proteins and in RbfA from *E. coli* have a higher than average flexibility. In contrast, the region prior to α2-helix consisting of residues K63 to T65 (*T. maritima* RbfA numbering) was not shown to be highly flexible in the *E. coli* protein (Huang et al., 2003) Residues 90 to 93 in *E. coli* preceding the β3-strand showed the same high flexibility as seen for the *T. maritima* and *H. pylori* RbfA. These results suggested that the *E. coli* RbfA shares overall the same flexible regions within the protein core.

3.4 Comparing the RbfA proteins

Comparison of the structures of T. maritima, E. coli and M. pneumonia Rbfa

A comparison of the *T. maritima* RbfA (residues 1–101), the *E. coli* RbfA (residues 1–108) and *M. pneumonia* RbfA (residues 1–120) reveals structural similarities between the three proteins as well as some differences. Database searches using Blossum 62 (Henikoff and Henikoff, 1992) as a matrix with default parameters showed that the sequences of *T. maritima* RbfA and *E. coli* RbfA have 30 % identity and 52 % similarity. Comparison of two representative structures of each RbfA protein, using PyMOL the root mean square deviation (RMSD) of superimposed backbone Cα atoms was 3.7 Å for the RbfA proteins from *T. maritima* and *E. coli*. For the alignment of *T. maritima* and *M. pneumonia* the RMSD value was 3.8 Å. Figure 3.29 shows the structures of *T. maritima*, *E. coli* and *M. pneumonia* RbfA superimposed using PyMOL alignment.

The three proteins showed a high structural similarity of their overall fold. The structural alignment matches the important structural elements through the amino acid patterns that exist in the secondary structures. The proteins showed the same conformational kink region between helices α2 and α3. This kink region (*T. maritima* RbfA residues A72–G74, *E. coli* RbfA residues A75–G77, *M. pneumonia* residues A69–G71) is part of the highly conserved AXG motif as mentioned before. This AXG motif was also identified in *H. pylori* RbfA and other RbfA proteins (Rubin et al., 2003; Bonander et al., 2003) this confirmed that the AXG sequence is a conserved structural motif throughout the RbfA protein family. Another structural feature is the reported β-bulge formed in the β1-strand by residue S39 (*E. coli* RbfA) (Huang et al., 2003). This β-bulge was also found in the RbfA proteins from *T. maritima*, *M. pneumonia* (T36) and *H. pylori*. The residue is conserved throughout the RbfA family as either a serine or threonine. With the exception of the loop regions the length of the secondary structural elements was essentially the same. The loop region connecting the β2- and β3-strand is extended by two residues in the *T. maritima* RbfA protein as compared to *E. coli* RbfA, while all the other loop regions in *T. maritima* RbfA are shorter.

The three dimensional arrangement of the secondary structural elements in the 3D solution structure of *T. maritima* RbfA showed significant differences compared to the *E. coli* and *M. pneumonia* RbfA. The angle in which

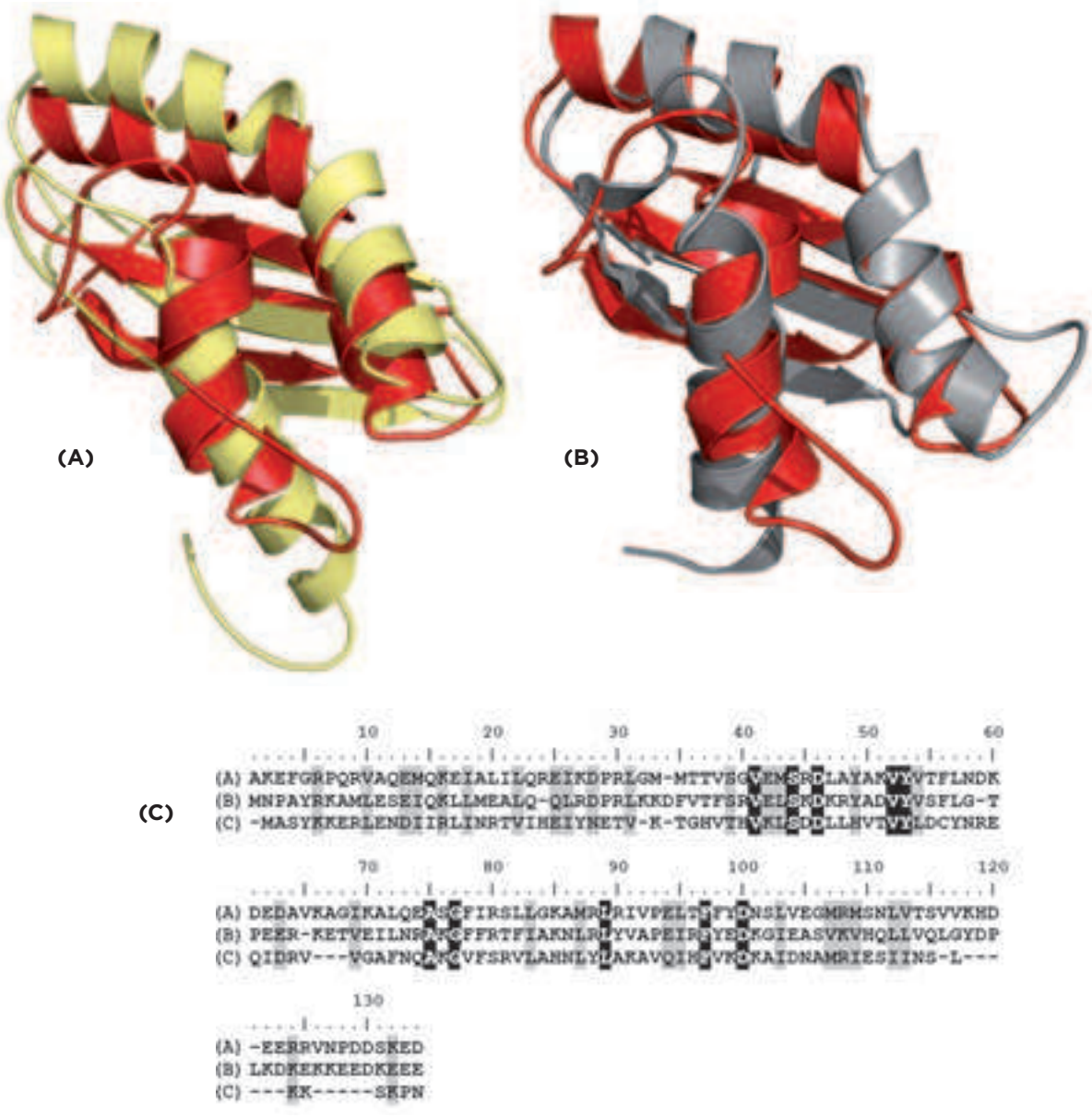


Figure 3.29
(A) Overlay of the NMR solution structures from *T. maritima* (red) and *E. coli* (brown) (Huang et al., 2003). An RMSD value of 3.74 Å was calculated by the program PyMOL. **(B)** Overlay of the solution structures of *T. maritima* (red) RbfA and *M. pneumonia* (grey) (Rubin et al., 2003) RbfA. An RMSD value of 3.79 Å was calculated using PyMOL. **(C)** Sequence alignment of the RbfA sequences from *E. coli* (A), *T. maritima* (B) and *M. pneumonia* (C). The alignment was obtained by database searches using Blossum 62 as a matrix. Amino acid residues identical in the two proteins are shaded in black, similar amino acid residues are shaded in grey.

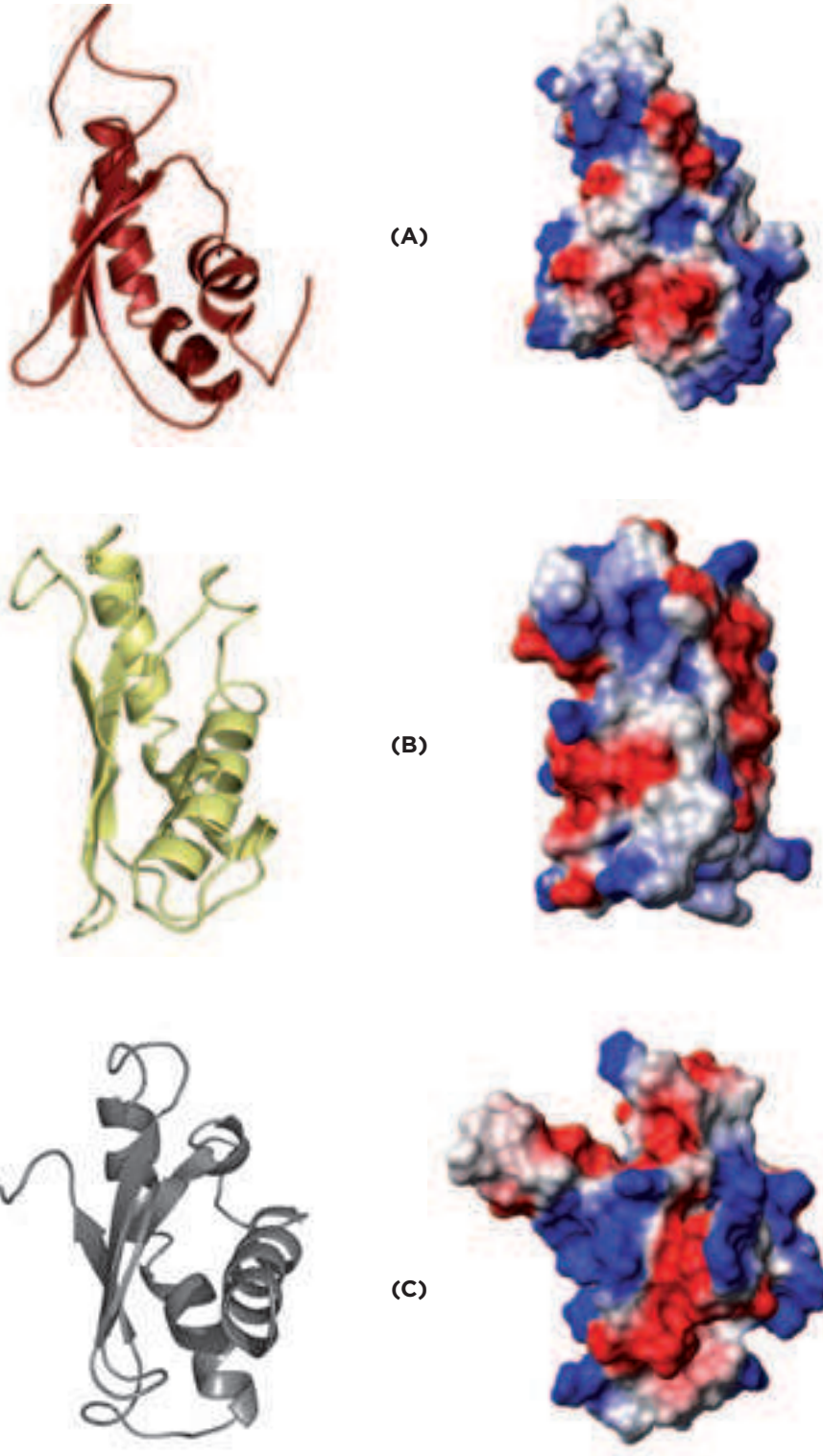


Figure 3.30
Structures and electrostatic surface potentials of **(A)** *T. maritima* RbfA, **(B)** *E. coli* RbfA and **(C)** *M. pneumonia* RbfA. The surface maps were created using the program MOLMOL.

the two helices (helix α2 and α3) are arranged to each other by the kink region varies between *E. coli* RbfA (about 130°), *M. pneumonia* RbfA (about 130°) and *T. maritima* RbfA (about 115°). The helices α1 and α3 are packed parallel to each other in the *T. maritima* RbfA structure as indicated by many long range NOEs observed between the residues of the two helices.

To further investigate the biophysical features of the three proteins, the electrostatic surface potential was plotted for the RbfA proteins (see Figure 3.30). *T. maritima* RbfA displays a bipolar distribution, which is characteristic for nucleic acid binding proteins. A strongly positively charged surface is created by a cluster of positively charged residues in the loop between α3-helix and β3-strand. Basic residues at the N-terminal end of helix α1 and a cluster of charged residues in the loop between β-strands β1 and β2 also contribute to this positively charged electrostatic surface. In addition, positively charged residues cluster around the kink region. These clusters involve several conserved residues such as K45 and K46 in the loop between the β1- and β2-strand. A negatively charged region is located at the N-terminal region of the α1-helix and at the N-terminal region of the α2-helix, which is the least conserved region in the RbfA family. Other regions of negatively charged residues are located in the β3-strand, which involves the conserved residue D97 and at the C-terminus.

The electrostatic potential for the RbfA proteins from *E. coli* and *M. pneumonia* share the same distribution of charged residues. It also shows the before mentioned bipolar distribution. In addition, the charged clusters are found in the same structural arrangement.

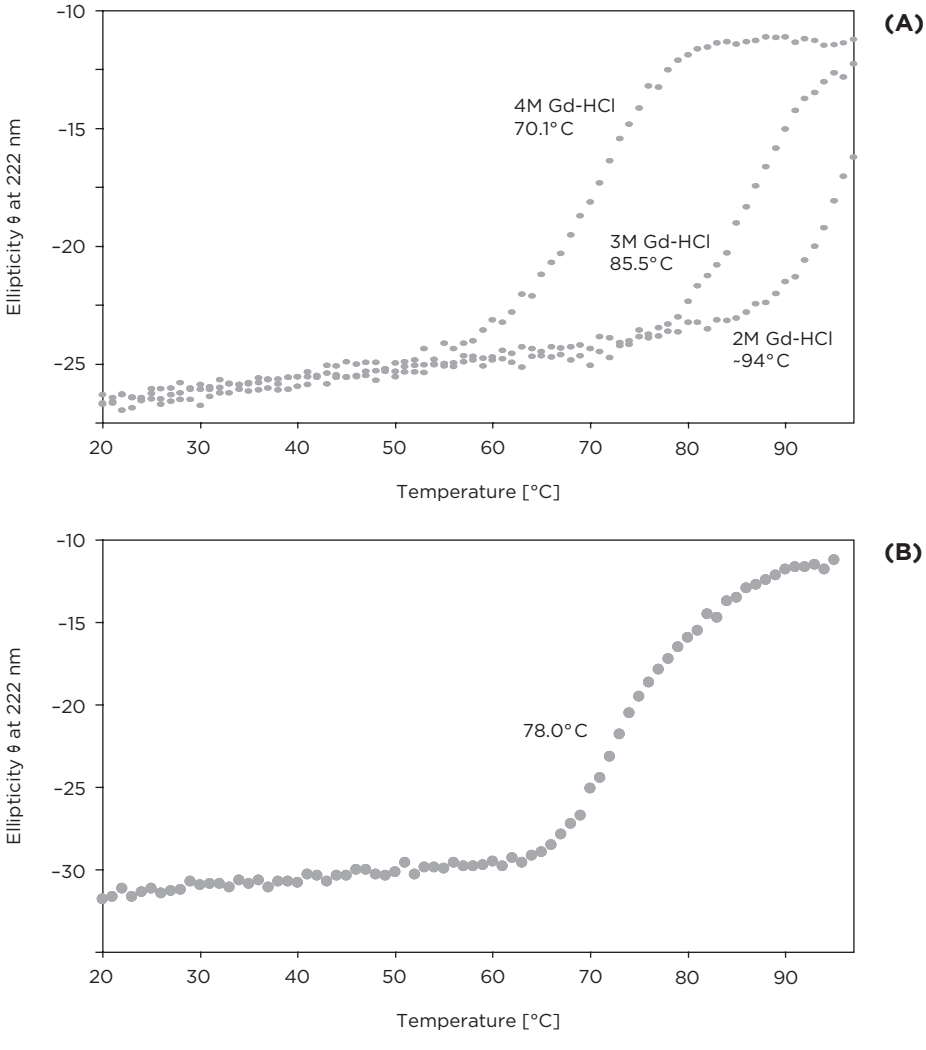
Thermostability

Thermal stabilities

Thermal stabilities of the RbfA proteins from *T. maritima* and *H. pylori* were assessed using CD spectroscopy. Changes in the spectrum were monitored with increasing temperature. The wavelength of 222 nm was chosen since this wavelength is indicative for α-helical folds. The adsorption at this wavelength was monitored continuously as the temperature was raised.

Both proteins unfold in cooperative transitions with t_m values of 78° C (*H. pylori* RbfA) and >100° C (*T. maritima* RbfA) in 25 mM phosphate buffer (pH 6.5) with 50 mM KCl. The melting temperature of *T. maritima* RbfA is decreased to 70° C in the presence of 4M guanidinium-hydrochloride. In Figure 3.29 (A) the circular dichroic melting curves are shown for the two

Figure 3.31
Thermal unfolding spectra of *T. maritima* RbfA at various Gd-HCl concentrations (**A**) and *H. pylori* RbfA (**B**). The spectra were recorded on 10 μM samples in phosphate buffer in the presence of 50 mM KCl. The transitions were monitored by the decrease of the CD signal at 222 nm in a 1 mm path length cuvette. Heating rates were 60° C h^{-1}. The melting temperature (t_m) was extracted by analyzing the first derivative of the melting curves.



proteins. In both cases the ellipticity only slightly increased in the premelting temperature range. This increase of intensity corresponds to only 9 % of the change observed at the melting temperature.

Adaptation to higher temperature as found in hyperthermophile organisms implies that molecular components of the cell, notably the proteins, are stable at temperatures above 80° C. Many structural features have been linked to protein thermostability, the simplest measure being protein length. However, in the case of RbfA the differences in sequence length are much more pronounced between proteins from mesophilic bacteria such as *E. coli* (133 AA) and *H. pylori* (111 AA) RbfA, then between the hyperthermophile *T. maritima* RbfA (131 AA) and *E. coli* RbfA. Longer secondary structural elements have also been suggested as a factor in temperature stability (Kumar et al., 2000), but in case of the *H. pylori* RbfA the β-strands are longer and the loop regions shorter than the corresponding structural elements in the *T. maritima* RbfA. Recent studies focused on compactness of the protein structure; assessing parameters such as accessible surface area (Kumar et al., 2000), cavity size as measured by total surface area of cavities (Szilagyi and Zavodisky, 2002) and solvent accessibility by distinguishing polar and nonpolar residues (Chakravarty and Varadarajan, 2002). None of these studies found significant differences between the mesophile and thermophile homologs, or the differences found were not the main feature leading to thermostability. Several groups have identified an increased number of ion pairs inside thermostable proteins (Perl et al., 2001). It was shown that overall coulombic interactions contribute favorably to the stability of the thermophilic protein Bc-Csp, whereas only two residues on the protein surface near the N- and C-terminus account for the additional stability of the thermophilic protein (Perl et al., 2001). In addition, it was found that the Bc-Csp protein uses destabilization of the unfolded state by unfavorable charge-charge interactions as a mechanism for increasing stability (Zhou et al., 2003). The effects of ionic strength and temperature are global in nature and are much less prone to errors in treating a particular local factor. Therefore, these global effects were used to yield insight into the contributions of electrostatic interactions to thermal stability. The thermal unfolding transitions of the RbfA proteins from *T. maritima* and *H. pylori* were monitored under different ion concentrations by following the change in circular dichroism at 222 nm.

To identify the amino acid residues that are responsible for the ionic interactions, the sequences of *T. maritima, H. pylori, M. pneumonia* and *E. coli* RbfA were aligned and analyzed for charged side chain residues. Fig 3.31 shows the alignment of the sequences where the positively charged amino acids (Lys and Arg) are color coded in blue, histidine residues are colored in light blue, and negatively charged residues (Asp and Glu) are color coded in red. *T. maritima* RbfA has 18 negatively charged residues and 22 positively

charged residues. The mesophile proteins *M. pneumonia*, *E. coli*, and *H. pylori* Rbfa have 13, 15 and again 15 negatively charged residues, respectively. *M. pneumonia* RbfA has 18, *E. coli* 17 and *H. pylori* 16 positively charged residues. Comparison of the aligned sequences show 15 charged residues in *T. maritima* Rbfa that are changed to uncharged residues in the mesophile proteins. On the other side 15 residues in the mesophile proteins are changed to uncharged residues in the thermophile protein. Despite these rearrangements of charged residues, the proteins share the same overall charge distribution in their structural fold. These findings made it too complicated to identify the particular amino acids that could be responsible for charge-charge energy contributions. In this case, the thermophilic protein differs from the mesophilic homologs at too many positions throughout the sequence.

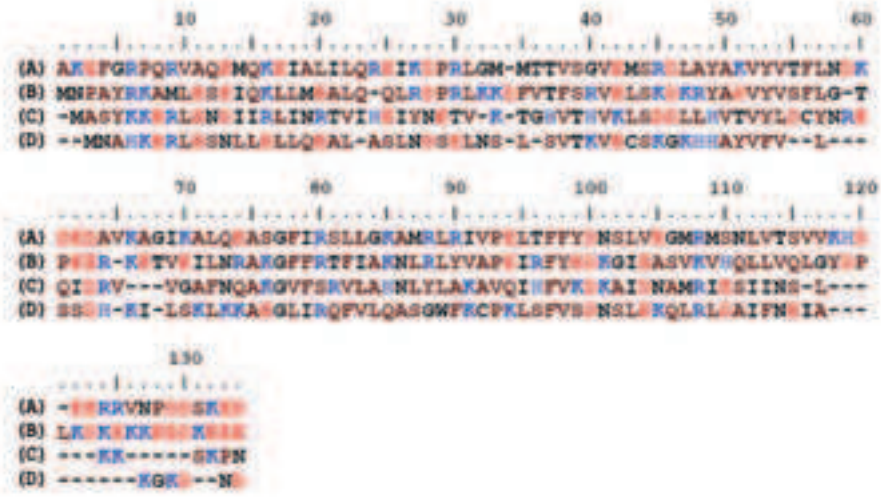


Figure 3.32
Sequence alignment of the ribosomal binding factor A sequences from *E. coli* **(A)** *T. maritima* **(B)**, *M. pneumonia* **(C)** and *H. pylori* **(D)**. The alignment was obtained by database searches using Blossum 62 as a matrix. Amino acid residues, which carry a positive charge, are colored in blue. Amino acid residues, which carry a negative charge, are colored in red.

To test the overall influence of charge-charge interactions the dependence of the midpoint of thermal unfolding transition on increasing salt concentrations was investigated. The thermal stability of proteins is usually salt dependent. Hofmeister effects, which depend on the nature of the added salt, affect stability even at high salt concentrations. For the screening of protein surface charge-charge interactions, potassium chloride was used. For *H. pylori* RbfA the midpoint of thermal unfolding is increased by ~10° C when the salt concentration was increased from 1 mM to 100 mM KCl (see Figure 3.32). From this result it was concluded that for the mesophilic *H. pylori* RbfA protein the ionic interactions are overall unfavorable. As a consequence, when these interactions are weakened by adding salt, the thermal unfolding transition is increased. In case of the thermophilic RbfA protein from *T. maritima* it

was shown that the thermal unfolding transition lies above 100° C. Therefore, the influence of salt concentration on protein stability was monitored at the premelting temperature of 60° C. For the mesophilic protein *H. pylori* RbfA the increase in salt concentration led to an increase in signal intensity at 222 nm.

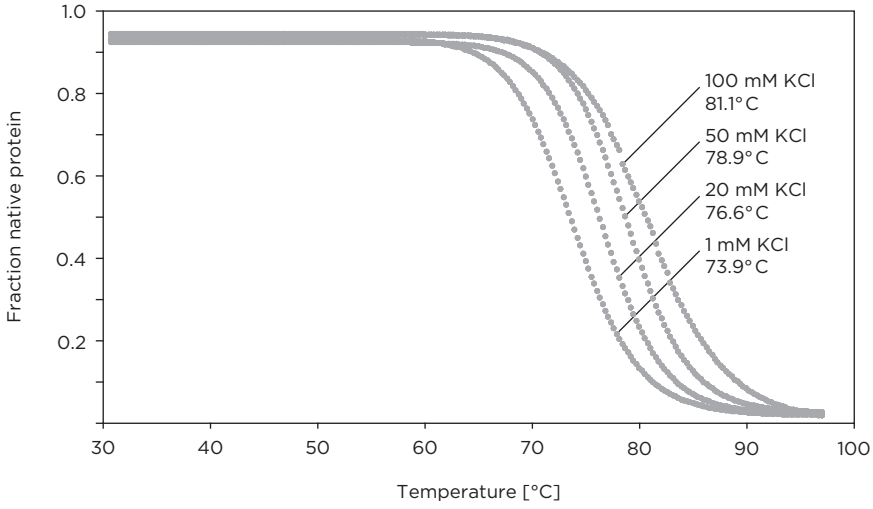


Figure 3.33
Thermal unfolding transitions of *H. pylori* RbfA at various ion concentrations. Measurements were performed in phosphate buffer pH 6.5. The transitions were monitored by the decrease of the CD signal at 222 nm in a 1 mm path length cuvette. Heating rates were 60° C h^{-1}. The fraction of native protein as obtained after a two state analysis of the data is shown as a function of the temperature.

For the *T. maritima* RbfA the signal intensity at 60° C is decreased by adding 0–2 M KCl. The neutralization of charge-charge interactions on the surface of the thermophilic RbfA protein led to destabilization of the protein fold. From these results it was concluded that for *T. maritima* RbfA the ionic interactions of charged amino acids on the protein surface are favorable for protein stability. Weakening of these interactions (by adding salt) resulted in decreased stability. The stability of the *T. maritima* RbfA was also determined as a function of guanidinium hydrochloride (Gd-HCl) concentration, which was used to compare stabilities of the protein at various salt concentrations.

The destabilization of the protein fold due to the denaturant Gd-HCl is close to the effect seen for KCl. It was also tested, if the variation of salt concentration in the presence of Gd-HCl had any effect on protein stability. The

results did not show further influence of protein stability due to salt concentrations in the presence of denaturant.

H. pylori					
KCl [mM]		0	20	50	100
Change of signal intensity [%]		0	18	31	42

T. maritima					
KCl [mM]	0	100	500	1,000	2,000
Change of signal intensity [%]	0	−8	−15	−26	−30

T. maritima			
Gd-HCl [mM]	0	100	500
Change of signal intensity [%]	0	−7	−19

Table 3.2
Change in signal intensity of the CD signal at 222 nm at various ion concentrations. Measurements were performed in phosphate buffer pH 6.5. The CD signal was monitored at the premelting temperature of 60° C.

Contact order

A recent study (Robinson-Rechavi and Godzik, 2005) suggested that the contact order (CO) is a major structural determinant of protein thermostability. The contact order is a structural feature describing the packing topology of proteins. It was initially introduced to explain differences in folding rates (Robinson-Rechavi and Godzik, 2005). For small proteins a linear relationship exists between the logarithm of the folding rate and the relative contact order as a metric of structural complexity (Sosnick and Pan, 2004). It was suggested that this structural feature allows distinguishing between thermophile and mesophile protein homologs (Robinson-Rechavi and Godzik, 2005). The relative contact order is the average sequence distance between tertiary contacts in the native protein structure normalized to the number of residues.

$$\text{Relative CO} = (1 / N_c\, L) \sum^{Nc} \Delta L_{i,j}$$

Where N_c is the total number of contacts within a contact threshold of 6 Å. $\Delta L_{i,j}$ is the number of residues separating the contacting residues i and j.

The relative contact order for the *E. coli* RbfA (mesophile), *M. pneumonia* RbfA (mesophile) and *T. maritima* RbfA (thermophile) were calculated based on the data of the published *M. pneumonia* (Rubin et al., 2004), *E. coli* (Huang et al., 2003) and the derived *T. maritima* 3D solution structure. The difference in contact order between the mesophilic *E. coli* and thermophilic *T. maritima* RbfA was calculated to be 0.046. The difference I relative contact order between the *M. pneumonia* and *T. maritima* RbfA was 0.047. In order to test if the sequence itself would have any influence on the contact order we also calculated the contact order for a hypothetical *T. maritima* RbfA-structure derived from a sequence homology model with the *E. coli* RbfA structure (Swiss model server). The relative contact order for the homology model turned out to be even lower, therefore ruling out any contribution of the different amino acid sequence to the increased contact order. The increase in contact order therefore seems to result from the different structural arrangement of the secondary structural elements in the *T. maritima* RbfA. The higher compactness of the protein fold therefore seems to be the main reason for the increased relative contact order.

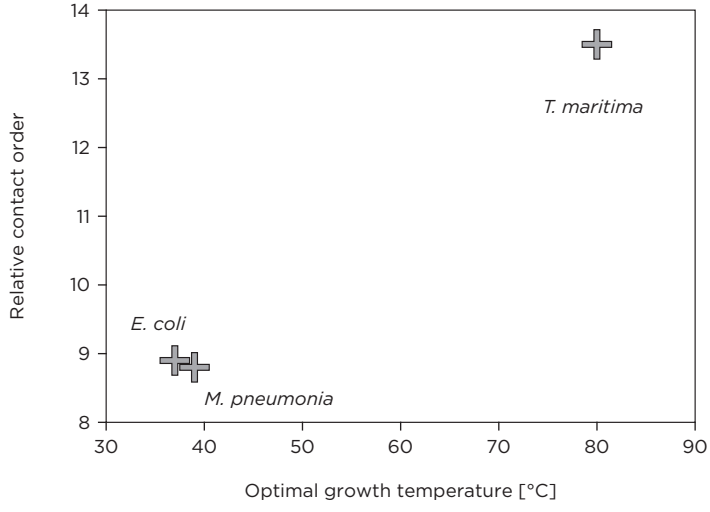


Figure 3.34
Relation between the optimal growth temperature and the relative contact order of RbfA from different bacteria. Each point corresponds to one experimental structure from the PDB. The PDB entries used are: 1KKG *(E. coli),* 1JOS *(M. pneumonia)* and the experimentally derived structure during this study *(T. maritima).*

Protein	chain length [AA]	growth temperature [°C]	relative contact order [%]
Csp B *B. subtilis*	66	34	7.5
Csp B *B. caldolyticus*	66	76	17.5
GAPDH *E. coli*	330	37	10
GAPDH *T. maritima*	332	80	13.9
RbfA *E. coli*	108	37	8.9
RbfA *T. maritima*	120	80	13.5
RbfA *M. pneumonia*	96	39	8.8
RbfA *E.coli* homology	108	–	8.2

Table 3.3 Relative contact orders for several thermophile/mesophile homolog proteins.

As a note of caution it should be kept in mind that the data for the hyperthermophile *T. maritima* protein was derived at temperatures of 40° C for technical reasons, not at the optimal growth temperature of *T. maritima* of 80° C.

Summary

The protein structures of *T. maritima* RbfA, *E. coli* RbfA and *M. pneumonia* RbfA show the same topology of secondary structural elements, whereas the three dimensional fold revealed several differences. The arrangement of the secondary structural elements in the three proteins varies in each case. The contact order was used as a structural determinant to address the question of protein fold compactness. It was possible to demonstrate that this structural feature distinguishes between the mesophile RbfA proteins from *E. coli*, *M. pneumonia* and *H. pylori* and the thermophile RbfA protein from *T. maritima*. The relative contact order is related to compactness of the structure and the topological complexity (Robinson-Rechavi and Godzik, 2005). This was in good agreement with our findings. The apparent reason for the increased compactness of the *T. maritima* RbfA is a slightly different arrangement of the secondary structural elements in the three dimensional fold, especially the parallel packing of the two α-helices α1 and α3. The increased compactness of the thermophile protein is also reflected in its diffusion constants.

Remarkably *T. maritima* RbfA, showed the same rotational correlation time as the mesophile *H. pylori* RbfA although the *T. maritima* RbfA is the larger protein (see chapter 3.3).

The contribution of charged residues to protein stability is believed to be rather small, but there is a broad agreement that surface exposed charged residues can be important for the stability of thermophilic proteins (Perl and Schmid, 2001). The comparison of charged residues in the RbfA homologs led to identification of several charged clusters within those proteins. However, the sequence variations made it too complex to identify charged residues that could contribute to protein stability. In simple approaches, distances between oppositely charged groups shorter than 6 Å are assumed to form stabilizing salt bridges. This approach turned out to be too simplistic and several analyses gave equivocal answers (Strop and Mayo, 2000; Sun, 1991; Serrano et al., 1990; Takano et al., 2000). The unfavorable loss in side chain entropy and the desolvation of the interacting charged groups both oppose the formation of pairwise salt bridges (Perl and Schmid, 2001). However, it was proposed that extended arrays of surface charges enhance stability (Karshikoff and Ladenstein, 1998). In such networks the entropic penalty would be smaller. It was suggested that thermophilic proteins are electrostatically better optimized and, in particular, that repulsive contacts are reduced (Karshikoff and Ladenstein, 1998; Spassov et al., 1994). This is in good agreement with our study. Here we favor the picture that extended arrays of charge-charge interactions, which are electrostatically optimized, lead to a more compact fold of the protein. This higher compactness ultimately results in thermostability through the cooperation of several components such as hydrophobic, van der Waals, coulombic and hydrogen bond interactions.

3.5 RNA binding of RbfA

Interactions with the helix I region of 16S rRNA

The fold of RbfA as deduced from NMR (Huang et al., 2003; Grimm and Wohnert, 2005; Rubin et al., 2003) and X-ray studies (Bonander et al., 2003) and the distribution of positive charges on the protein surface suggested that RbfA might bind to RNA. However, while it was known that RbfA binds to the 30S ribosomal subunit, it was never tested experimentally if RbfA is actually an RNA-binding protein. In addition, its binding site on the 30S ribosomal subunit was never identified. The identification of RbfA as a multicopy suppressor of the cold-sensitive phenotype of the C23U-mutation in helix I of the 16S rRNA pointed toward the helix I of 16S rRNA as a possible binding target of RbfA (Dammel and Noller, 1995). Later work (Inoue et al., 2003) pointed to multiple binding sites for RbfA on the 30S ribosomal subunit. In order to test if the RbfA protein is indeed an RNA-binding protein that interacts with the 16S rRNA, RNA binding assays using a combination of biophysical methods such as NMR, fluorescence quenching assays and gel shift experiments were carried out using the full-length *H. pylori* RbfA and RNA fragments derived from the 16S rRNA. The availability of *H. pylori* was essential for this kind of studies since in contrast to the *E. coli* and the *T. maritima* RbfA, it was soluble and stable in its full-length form and it contained a tryptophan in its sequence close to a patch of positively charged residues predicted to be the RNA-binding site of the protein.

Our initial binding studies were carried out with short RNA constructs, which mimic the helix I region of the 16S rRNA, to directly test the suggestion of Dammel and Noller that the helix I region is the RbfA binding site. Several RNA constructs were designed for this purpose. As a starting point a set of constructs was designed, which included nucleotides 1–26 (*E. coli* numbering) at the 5'-end of the 16S rRNA sequence form *H. pylori*. RNA I is a construct of 18 nucleotides (9–26) of the helix I stem loop structure (see Figure 3.35 B). RNA II also included the 5'-terminal ssRNA nucleotides of the 16S rRNA (nucleotides 1–26) (see Figure 3.35 C). Additionally, a construct was designed that included the C23U mutation that was initially responsible for the identification of RbfA (see Figure 3.35 D). This mutation in RNA III seems to strongly destabilize the helix I secondary structure and ultimately results in formation of a different secondary structure as predicted by the program Mfold (Zuker, 2003; Mathews et al., 1999).

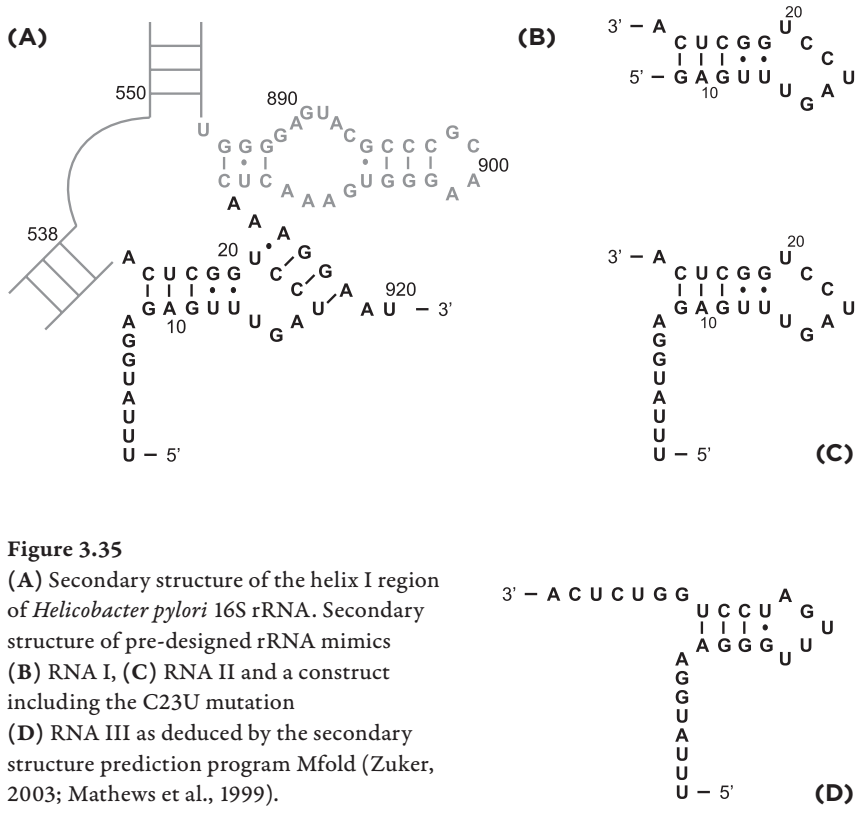


Figure 3.35
(A) Secondary structure of the helix I region of *Helicobacter pylori* 16S rRNA. Secondary structure of pre-designed rRNA mimics **(B)** RNA I, **(C)** RNA II and a construct including the C23U mutation **(D)** RNA III as deduced by the secondary structure prediction program Mfold (Zuker, 2003; Mathews et al., 1999).

As expected RNA I and RNA II fold into monomeric hairpin structures as indicated by native gel-electrophoresis (data not shown). They were found to contain a significant amount of helical secondary structure as shown by CD spectroscopy (see Figure 3.36 A). On the other hand RNA III appeared to be mostly single stranded as was indicated from the native gel electrophoresis (data not shown).

The circular dichrosim spectra were used to monitor structural changes in the RNA molecule. Folded RNA molecules containing A-form double helical elements exhibit characteristic bands between 320 nm and 180 nm (T. R. Sosnick et al., 2000). The presence of A-form helical elements in the RNA molecules was monitored by observing the shift of the peak maximum at 260 nm. To test the influence of divalent cations on the fold of the pre-designed RNA constructs Mg^{2+} titrations were performed.

The titrations suggested that the constructs RNA I and RNA II (data not shown) contain A-form helical elements which are further stabilized upon addition of 5 mM Mg^{2+} ions. A much smaller effect could be seen for the con-

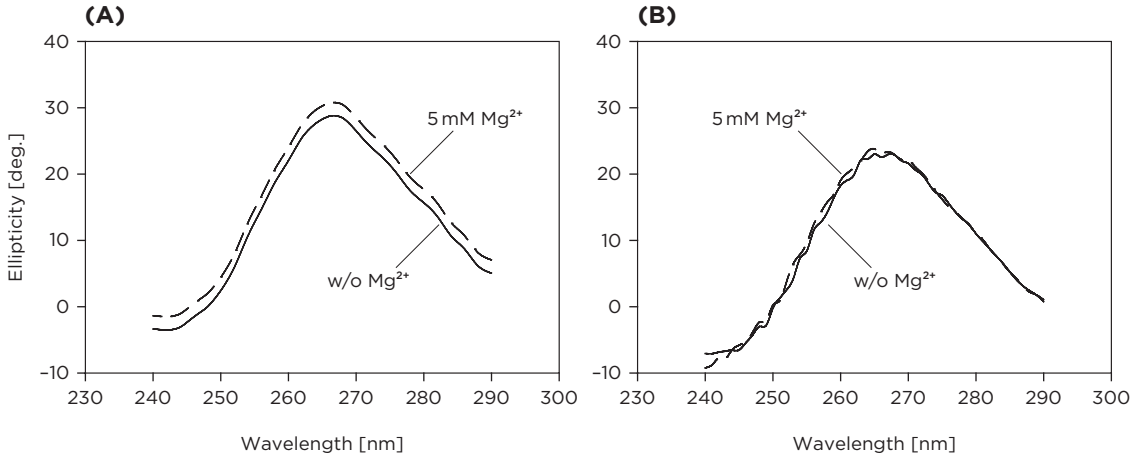


Figure 3.36
(A) CD spectra of RNA I titrations with Mg^{2+} cations. Solid line: Spectra of RNA I in the absence of divalent cations. Broken line: RNA I in the presence of 5 mM Mg^{2+} ions. **(B)** CD spectra of RNA III. Solid line: RNA III in the absence of divalent cations. Broken line: RNA III in the presence of 5 mM Mg^{2+} ions.

struct RNA III (see Figure 3.36 B), where the preformed secondary structure seemed not to be affected by the presence of Mg^{2+} cations.

Chemical shift perturbation experiments using NMR spectroscopy were performed to test RNA binding to *H. pylori* RbfA. Formation of complexes between ^{15}N-labeled RbfA and the unlabeled RNA constructs RNA I, RNA II and RNA III was monitored by using two dimensional ^{1}H-^{15}N-HSQC experiments as shown in Figure 3.37.

The addition of increasing amounts of RNA to a final molar excess of 4:1 lead to increasing changes of chemical shifts for a number of protein amide signals with all three RNAs (see Figure 3.38) whereas many other signals remained virtually unperturbed. The gradual changes of the chemical shifts for some of the residues were in agreement with the formation of RNA-protein complexes that are in fast exchange on the NMR time scale in all three cases, indicating binding constants in the high micromolar to millimolar range. The pattern of residues exhibiting chemical shift changes is also very similar for all three RNAs tested (see Figure 3.38).

The most pronounced chemical shift changes were found at the N-terminus of helix α1, β-strands β1 and β2 and their interconnecting loop and in α-helix α3.

The only difference was observed for RNA II including the single stranded region preceding helix I, where additional chemical shift changes were

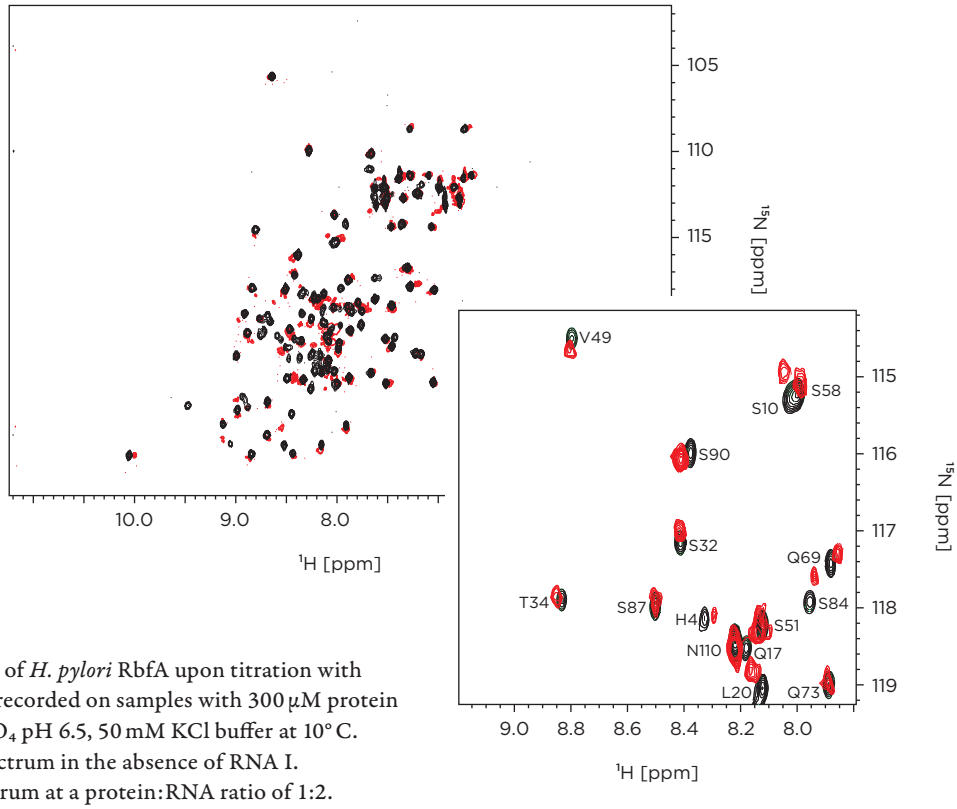


Figure 3.37
^{15}N edited HSQC spectra of *H. pylori* RbfA upon titration with RNA I. The spectra were recorded on samples with 300 µM protein samples in 25 mM KH_2PO_4 pH 6.5, 50 mM KCl buffer at 10° C.
Black: *H. pylori* RbfA spectrum in the absence of RNA I.
Red: *H. pylori* RbfA spectrum at a protein:RNA ratio of 1:2.

observed in β-strand β3. Remarkably, the affected residues form a contiguous surface of the protein (see Figure 3.39 A) and involve many of the highly conserved amino acids in the RbfA sequence (see Figure 3.39 B). It should be noted that the side chain amide group of the tryptophan residue W77 located in α-helix α3 exhibits significant chemical shift changes for all three RNAs.

The experiments described above established RbfA as a genuine RNA-binding protein. The rather low affinity for the substrates RNAs I–III and the similar binding behavior of substrates with very different secondary structure content indicate that the observed binding is rather unspecific.

Fluorescence quenching of trp 77 of H. pylori RbfA

Since the involvement of residue W77 in the RNA-binding was verified by NMR spectroscopy, fluorescence measurements were performed to probe the interaction of RNA with the tryptophan aromatic side chain and further determine the apparent equilibrium dissociation constant (K_D) values.

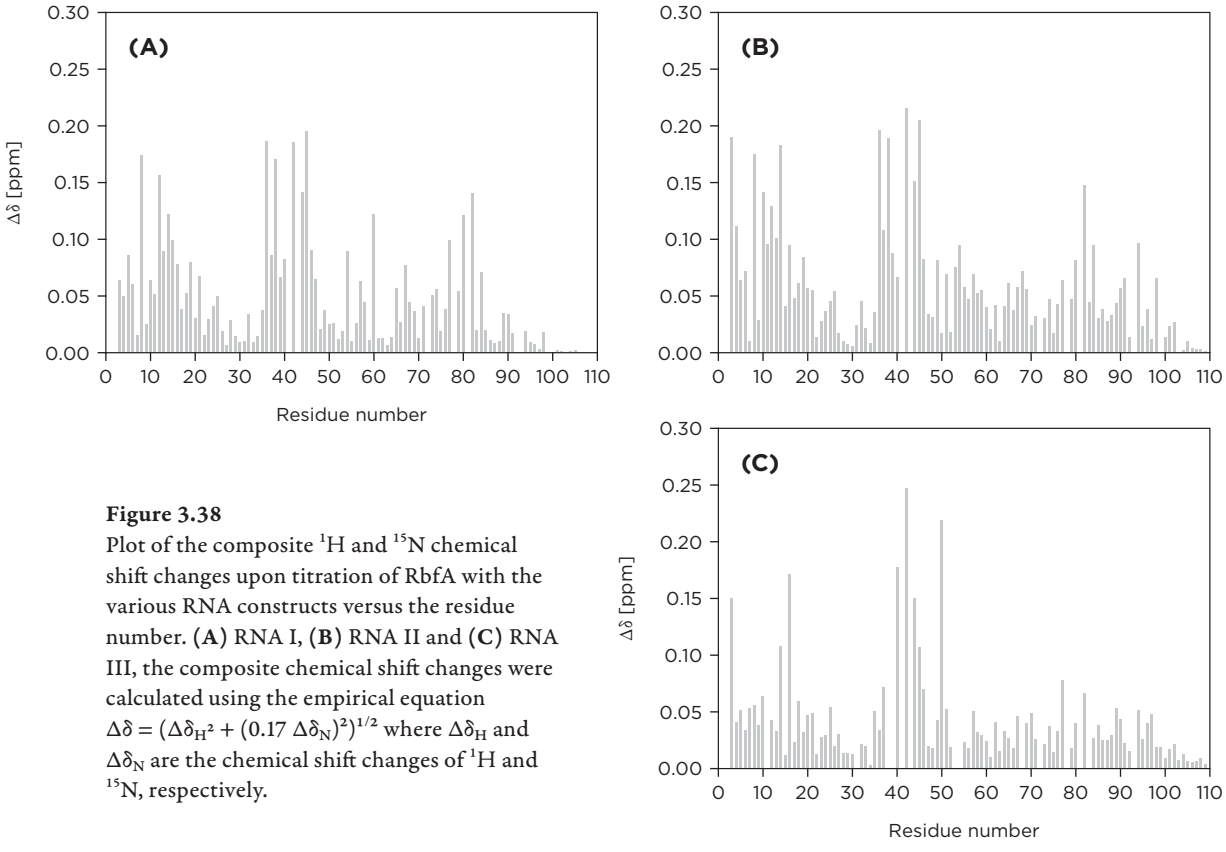
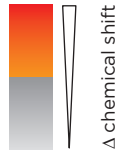
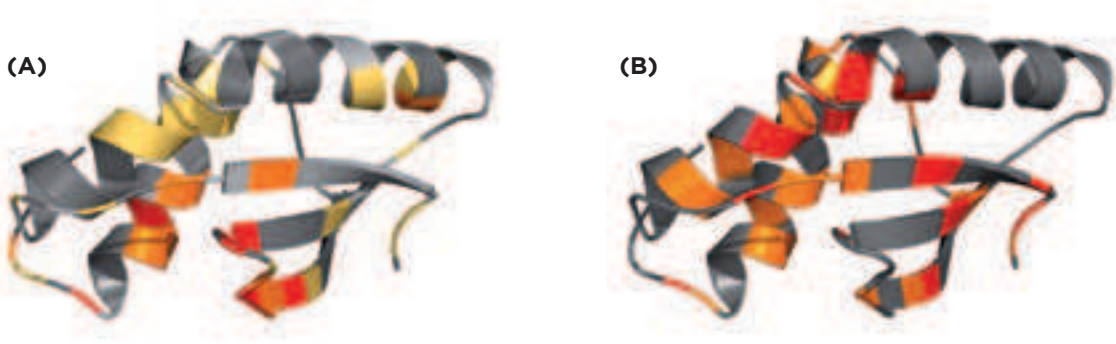


Figure 3.38
Plot of the composite ^{1}H and ^{15}N chemical shift changes upon titration of RbfA with the various RNA constructs versus the residue number. **(A)** RNA I, **(B)** RNA II and **(C)** RNA III, the composite chemical shift changes were calculated using the empirical equation $\Delta\delta = (\Delta\delta_H^2 + (0.17\ \Delta\delta_N)^2)^{1/2}$ where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift changes of ^{1}H and ^{15}N, respectively.

Figure 3.39
(A) Mapping of the composite ^{1}H and ^{15}N chemical shift changes on the solution structure of *T. maritima* RbfA. **(B)** Mapping of the conserved amino acid residues of the RbfA family onto the solution structure of *T. maritima* RbfA. Red: identical in multiple sequences. Orange: similar in multiple sequences.





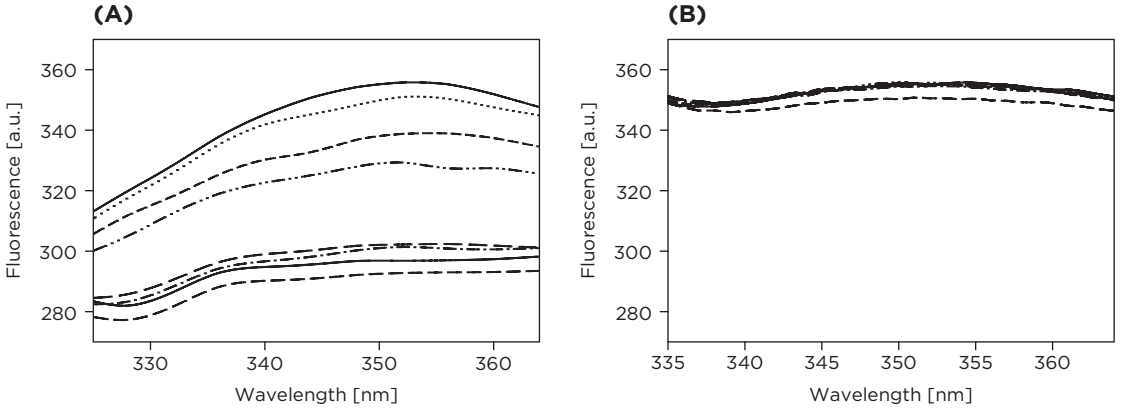


Figure 3.40
Fluorescence changes upon titration of *H. pylori* RbfA and **(A)** RNA I and a polyU ssRNA **(B)**. During the titration *H. pylori* RbfA had a concentration of 1 μM. The RNA constructs were titrated with concentrations of (0, 1, 5, 10, 20, 35, 50 and 100 μM). The experiments were performed at 20° C.

The fluorescence of *H. pylori* RbfA was measured at 20° C with a total protein concentration of 1 μM in phosphate buffer at pH 6.5 in the presence of 50 mM KCl. The binding of the RNA constructs mimicking the helix I region rRNA was directly monitored by following the decrease in fluorescence emission at 350 nm. Figure 3.40 (A) shows the decrease in emission intensity, in a saturable fashion, upon the addition of increasing amounts of the respective RNA construct.

The direction of this fluorescence change indicates that the tryptophan side chain at position 77 in the *H. pylori* RbfA is buried in a hydrophobic environment upon binding of the RNA. In order to test for specific binding to the designed RNA constructs, a nine nucleotide polyU ssRNA was also titrated. In this case, no quenching of the tryptophan side chain signal could be detected (see Figure 3.40 B).

The fluorescence data were consistent with the initial NMR data, where the tryptophan side chain signal displayed chemical shift changes in the conducted RNA titrations. By fitting the data to a hyperbolic function, the dissociation constants were extracted. For the fitting procedure the simplest possible mechanism was adopted, ignoring the possibility of multi step or multiple binding site interactions:

$$\text{Protein} + \text{RNA} \leftrightarrow \text{Protein} \cdot \text{RNA}$$

The fluorescence data were fitted with the formula described in experimental procedures.

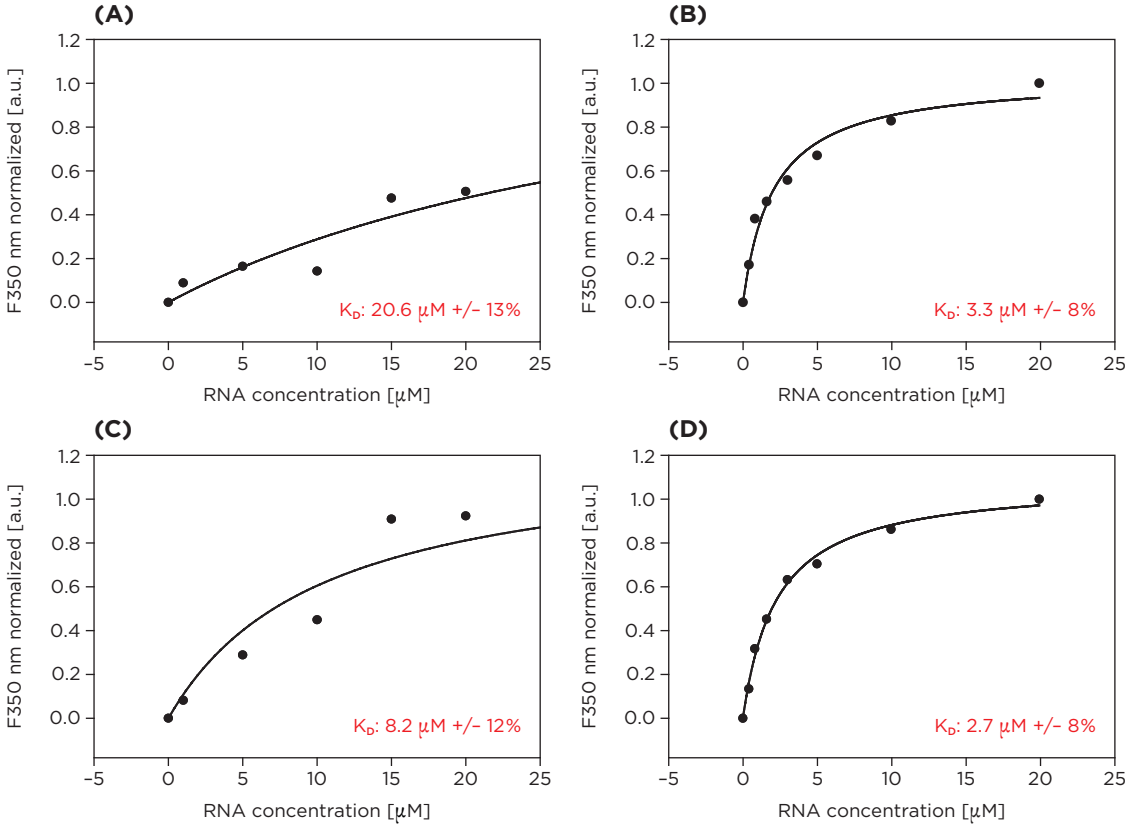
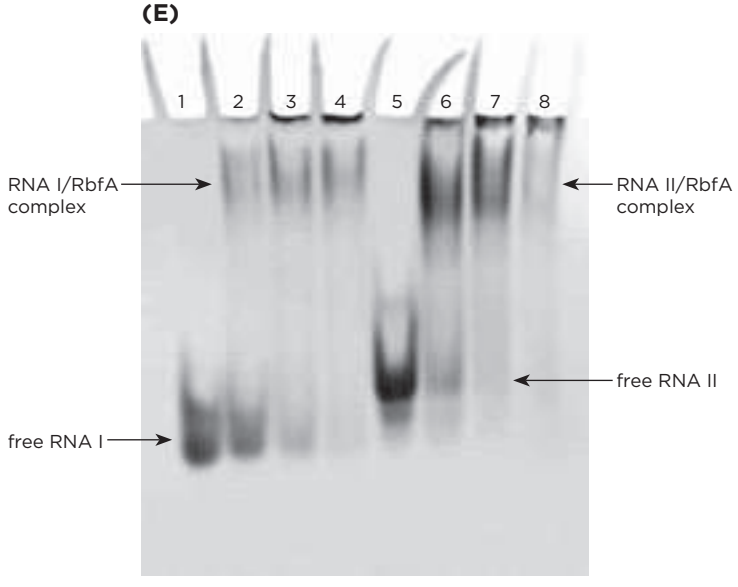


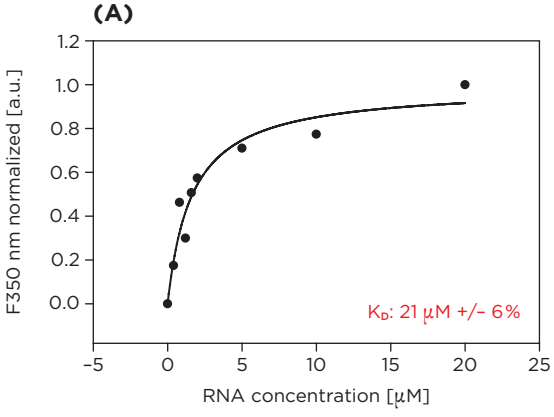
Figure 3.41
(A)–(D) Curve fittings of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). **(A)** and **(B)** titration of RNA I in the absence and presence of Mg^{2+} cations, respectively. **(C)** and **(D)** titration of RNA II in the absence and presence of Mg^{2+} cations, respectively. **(E)** Native gel shift analysis of the interaction between *H. pylori* RbfA and RNA. Lane 1: shows RNA I without protein present. For lanes 2–4 the concentration ratios were 1:2, 1:6 and 1:8. Lane 5: shows RNA II in the absence of protein. For lanes 6–8 the concentration ratios were 1:2, 1:6 and 1:8 respectively.

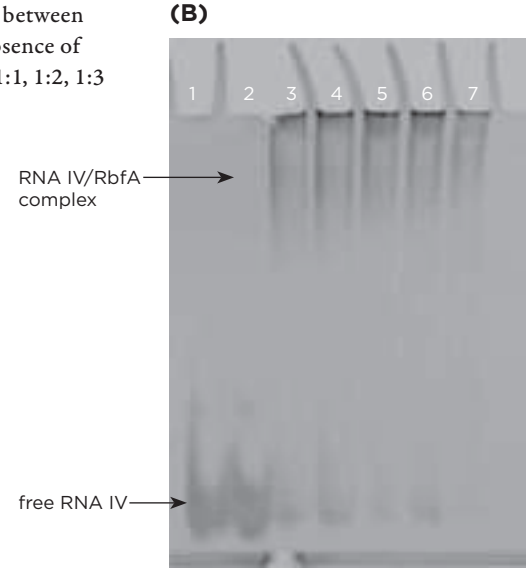


For RNA I and RNA II the K_D values were ~20 µM and ~8 µM, respectively (see Figure 3.41 A and C and Table 3.4). Since the NMR data suggested an influence of the ssRNA of the 5'-end, a nine nucleotide long ssRNA (UAUAUGGAG, nucleotides 1–9 of the 16S rRNA) which resembled the 5'-end nucleotides of the 16S rRNA was titrated. A quenching effect of less than 10 % of the absolute fluorescence signal was detected for the ssRNA (data not shown). Therefore no interaction between the RNA and the RbfA protein was suggested. It was also tested if divalent cations such as Mg^{2+} would have any influence on the binding affinity of the RNA constructs. Thus, the titrations were repeated in the presence of 5 mM Mg^{2+}. The binding affinity increased in the presence of divalent cations to 3.3 µM and 2.7 µM for RNA I and RNA II, respectively. The fluorescence data for RNA III in the presence of cations gave a binding constant of ~2.6 µM. The binding of RNA with the RbfA protein was also analyzed by native gel mobility shift assays. The RNA concentration was fixed, whereas the protein concentration varied from RNA:protein ratios of 1:0 to 1:8.

Figure 3.41 (E) shows the native PAGE analysis of the RNA I/RbfA and RNA II/RbfA interaction. The formation of the complex gradually increased with the increase of RbfA concentrations. Saturation of the complex formation was reached at a protein: RNA ratio of 1:6 for both of the RNA constructs.

Figure 3.42
(A) Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA IV concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). The experiments were conducted at 20° C in the presence of 5 mM Mg^{2+} cations. **(B)** Native gel shift analysis of the interaction between *H. pylori* RbfA and RNA IV. Lane 1 and 2: RNA IV in the absence of protein. For lanes 6–8: the concentration ratios were 1:0.5, 1:1, 1:2, 1:3 and 1:4, respectively.





Stabilizing helix I

In order to test if exchange between conformational states of the model RNAs has inhibited the binding interaction, a new construct was designed which stabilized the stem-loop structure of helix I (RNA IV).

The helix I was stabilized by additional GC base pairs at the 5'-end of the helix and by replacing the GU base pair adjacent to the hairpin loop with a GC base pair. Fluorescence quenching experiments were performed on the stabilized RNA IV (see Figure 3.42 A) and revealed a K_D value of ~2 μM. The effect of divalent cations is much less pronounced for the stabilized helix I. The presence of Mg^{2+} ions (5 mM) increased the binding constant from 2.7 μM to 2.1 μM. In addition, native gel mobility shift analysis was performed and the experiments showed that the RNA:protein complex formation gradually increased with the increase of the protein concentration (Figure 3.42 B). Saturation of the RbfA:RNA complex was already reached at a protein:RNA ratio of 1:3 rather than a ratio of 1:6 as seen for the wild type RNA sequences. NMR titrations were conducted which verified the binding of the constructs and showed intermediate exchange on the NMR time scale at a temperature of 10° C. This made it impossible to extract chemical shift perturbation data from the NMR spectra due to the extensive line broadening (data not shown). The results for the stabilized helix I mimic indicated an increase in binding affinity to the RbfA protein, compared to the initial RNA mimics.

NMR on RNA

To test the binding interactions of the RNA constructs, NMR spectroscopy was employed to identify the regions of the RNA nucleotides interacting with the protein. Chemical shift perturbations in the RNA were mapped upon addition of *H. pylori* RbfA. In case of the RNA, 1D ^{1}H-iminoproton resonance spectra were used to follow molecular interactions between nucleotides of the respective constructs and protein residues. The resonances of imino protons were assigned using 2D NOESY experiments. The imino proton region of guanines (H1) and uracils (H3) between 10–15 ppm contain valuable information about base pair interactions in RNA molecules (Furtig et al., 2003). Since only those protons which are involved in base pair interactions can be observed, the imino proton spectra provide a fingerprint of each base pair within the helical part of the RNA construct. By identifying the number of imino proton signals, it was essentially possible to count the number of base pairs. In addition, the chemical shift values of imino protons of Watson-Crick base pairs tend to be found in the region of 12–15 ppm, whereas imino

protons of non-canonical base pairs experience an upfield shift. The starting point of determining the hydrogen bonding pattern is to sequentially assign the well resolved imino resonances in the 2D NOESY experiment. The assignment strategy is based on the observation of NOE contacts. Watson Crick A:U base pairs are easily identified through their strong NOE cross peak between H2 protons of adenine and the H3 imino proton of uridine. In Watson Crick G:C base pairs, the H1 imino proton of guanine shows a strong NOE contact to the amino protons of the base pairing cytosine. For non-canonical base pairs specific NOE contacts can be identified as well. The imino ^{1}H NMR spectrum of construct RNA IV yielded well dispersed resonance peaks of all 6 base pairs predicted by the secondary structure folding program Mfold (Zucker, 2003; Mathews et al., 1999). One additional resonance could be observed at 12.1 ppm. This resonance seemed to result from a base pair interaction involving nucleotides within the hairpin loop. Assignment of the imino protons were extracted from NOESY spectra recorded at 700 µM RNA concentration using different mixing times optimized for the respective RNA construct. For the RNA IV construct a mixing time of 80 ms was used. The assignments initially concentrated on

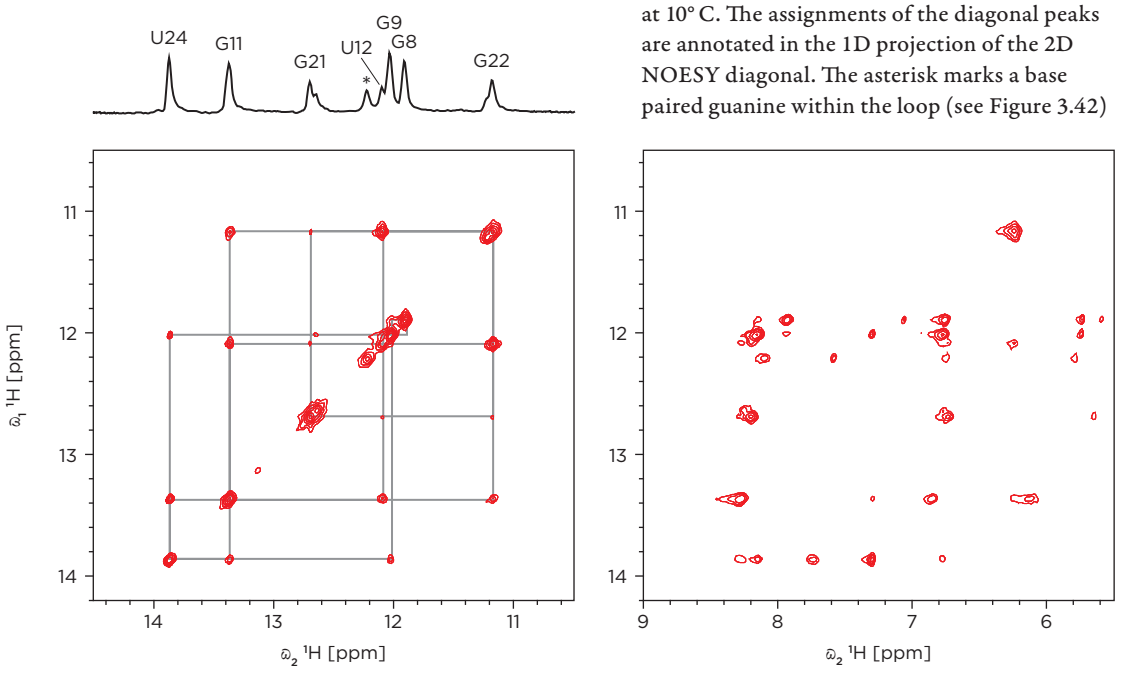


Figure 3.43
2D NOESY spectra of RNA IV. The spectra were recorded with samples containing 500 µM RNA in 25 mM KH_2PO_4 pH 6.5, 50 mM KCl buffer at 10° C. The assignments of the diagonal peaks are annotated in the 1D projection of the 2D NOESY diagonal. The asterisk marks a base paired guanine within the loop (see Figure 3.42)

the imino proton diagonal peak region. The cross peaks could be used for connecting the neighboring imino peaks. Additionally the NOESY spectrum region at 5.3–8.5 ppm was used to identify the nucleotides by their base pair specific NOE contacts of the respective base pair (see Figure 3.43).

The G22:U12 base pair signals were easily identified through the non-Watson-Crick base pair chemical shift and a strong cross peak between G22 H1 and U12 H3 imino protons. The resonance of U24 H3 was identified due to its strong NOE to the A10 amino H2 proton at 7.35 ppm and by sequential NOEs to G11 and G9 imino protons.

The Watson Crick GC base pair guanine H1 protons were assigned by comparing the internucleotide NOE signals of the imino proton region and the NOEs cross peaks to the cytosine H41/H42 protons. Identification of neighboring NOE contacts completed the assignment process.

H. pylori RbfA was then added to RNA IV and a series of 1D ^{1}H NMR spectra were recorded at RNA:protein ratios of 1:0, 1:0.25, 1:0.7 and 1:1 (see Figure 3.44). As the molar ratio of RbfA protein increases, peaks in the imino

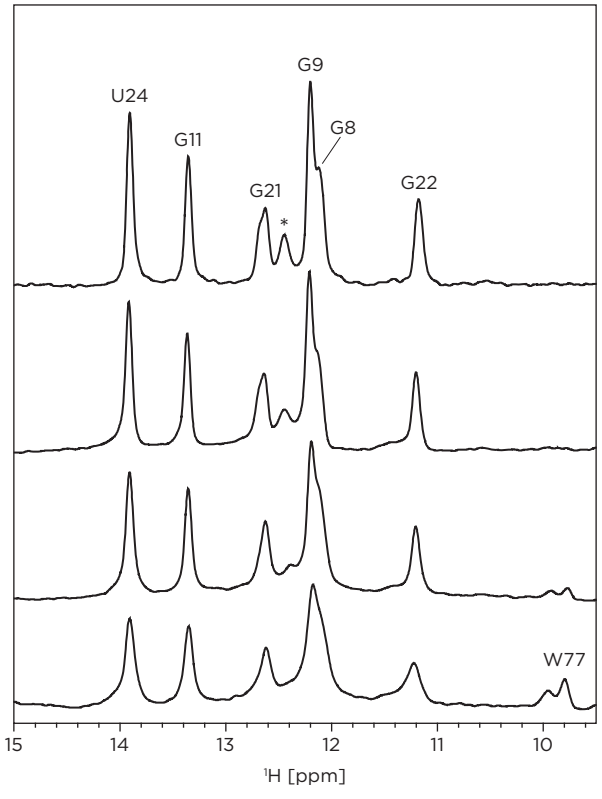


Figure 3.44
1D ^{1}H imino proton spectra of the RNA IV construct upon titration with *H. pylori* RbfA. The spectra were recorded on RNA samples with 200 μM RNA concentration at 10° C. RNA:protein ratios used in the titration were 1:0, 1:0.25, 1:0.7 and 1:1.

region broaden significantly, indicating that the free RNA and the RNA-protein complex were in intermediate exchange on the NMR time scale. Peaks G21 and G22 show the most pronounced line-broadening even at low protein concentrations.

Interactions with the central pseudoknot of 16S rRNA

The pseudoknot structure

Previous studies showed that the disruption of the central pseudoknot structure involving helix I and helix II of the 16S rRNA abolishes ribosomal activity (Poot et al., 1998). In order to test if the central pseudoknot is a better binding target for the RbfA protein then the RNA constructs containing only the helix I, new constructs were designed that should be able to form a pseudoknot structure (see Figure 3.45) with structural similarities to the central pseudoknot region of the 16S rRNA. First, a construct was designed which included the helix I nucleotides 9–26 (*E. coli* numbering) followed by a seven nucleotide poly uracil linker and the nucleotides 912–920 (*E. coli* numbering) which are supposed to base pair with nucleotides 17–20 to form the helix II of the pseudoknot structure (RNA V). The next construct which included the ssRNA at the 5'-end of the 16S rRNA was designed (RNA VI) to account for influences of the ssRNA to the binding interaction.

To test the influence of the polyU linker region, a construct was designed with a nine uracil nucleotide linker (RNA VII). As shown for the helix I region a set of constructs was designed which stabilize the helix I (RNA VIII) and helix II (RNA IX) within the constructs. All constructs were tested for their secondary structure with the secondary structure prediction program Mfold (Zucker, 2003; Mathews et al., 1999). All constructs were predicted to form the helix I secondary structure.

The binding of RbfA to the RNA constructs was first analyzed by fluorescence quenching experiments to evaluate the apparent equilibrium binding constants (see Figure 3.46). For the construct RNA V the measurements revealed that the presence of the helix II increased the binding affinity by a factor of ~3 compared to the RNA constructs mimicking the helix I region. The influence of divalent cations was tested by repeating the titrations in the presence of 5 mM Mg^{2+} ions. The apparent binding constant increases in the presence of divalent cations by a factor of 4 from 4.1 μM to 1.1 μM for RNA V. For the RNA VI, which includes the ssRNA at the 5'-end a smaller effect of the presence of divalent cations was found.

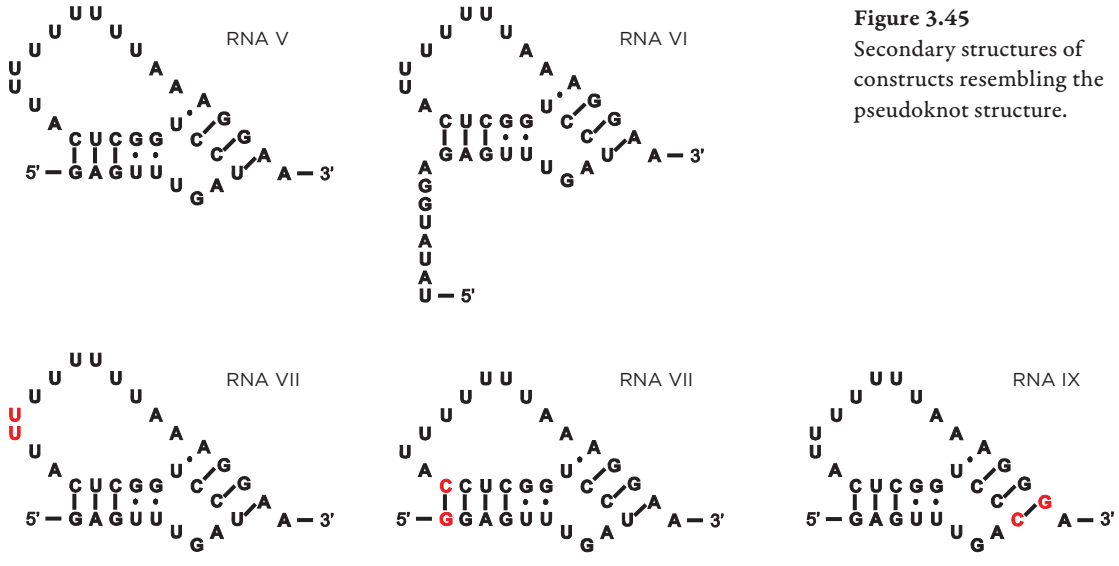


Figure 3.45
Secondary structures of constructs resembling the pseudoknot structure.

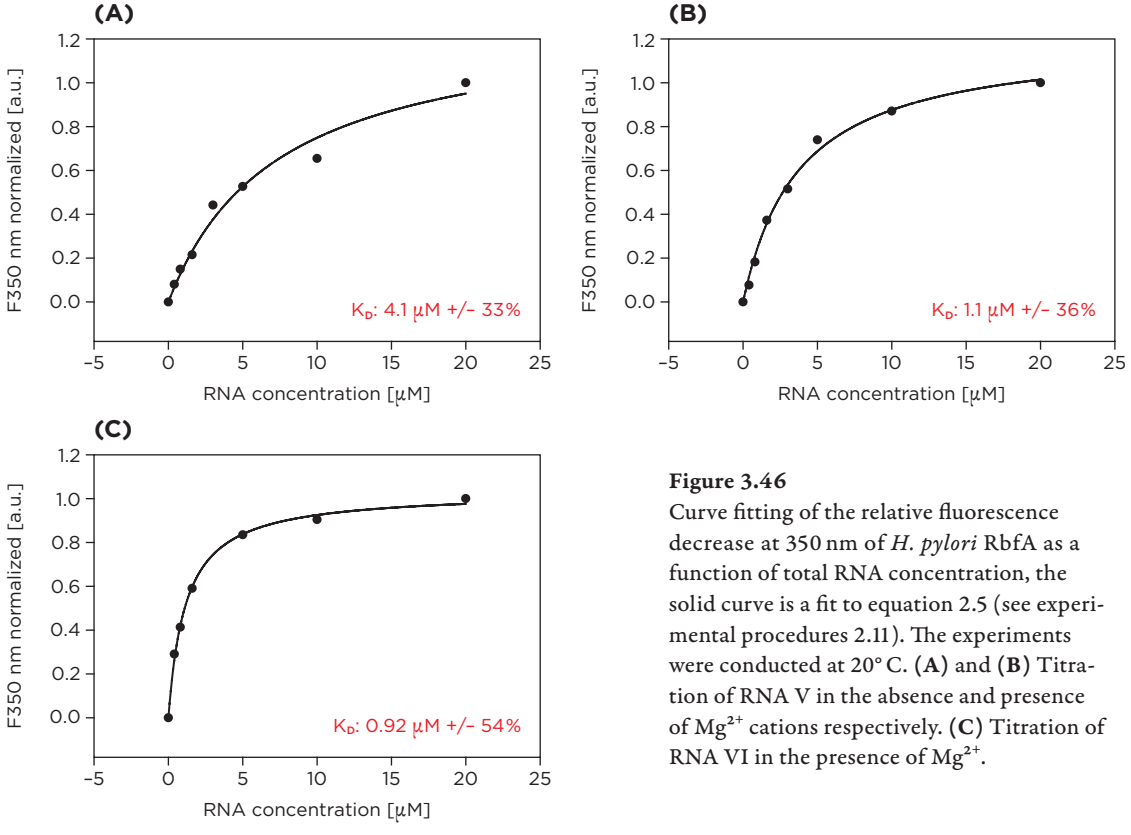


Figure 3.46
Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). The experiments were conducted at 20° C. **(A)** and **(B)** Titration of RNA V in the absence and presence of Mg^{2+} cations respectively. **(C)** Titration of RNA VI in the presence of Mg^{2+}.

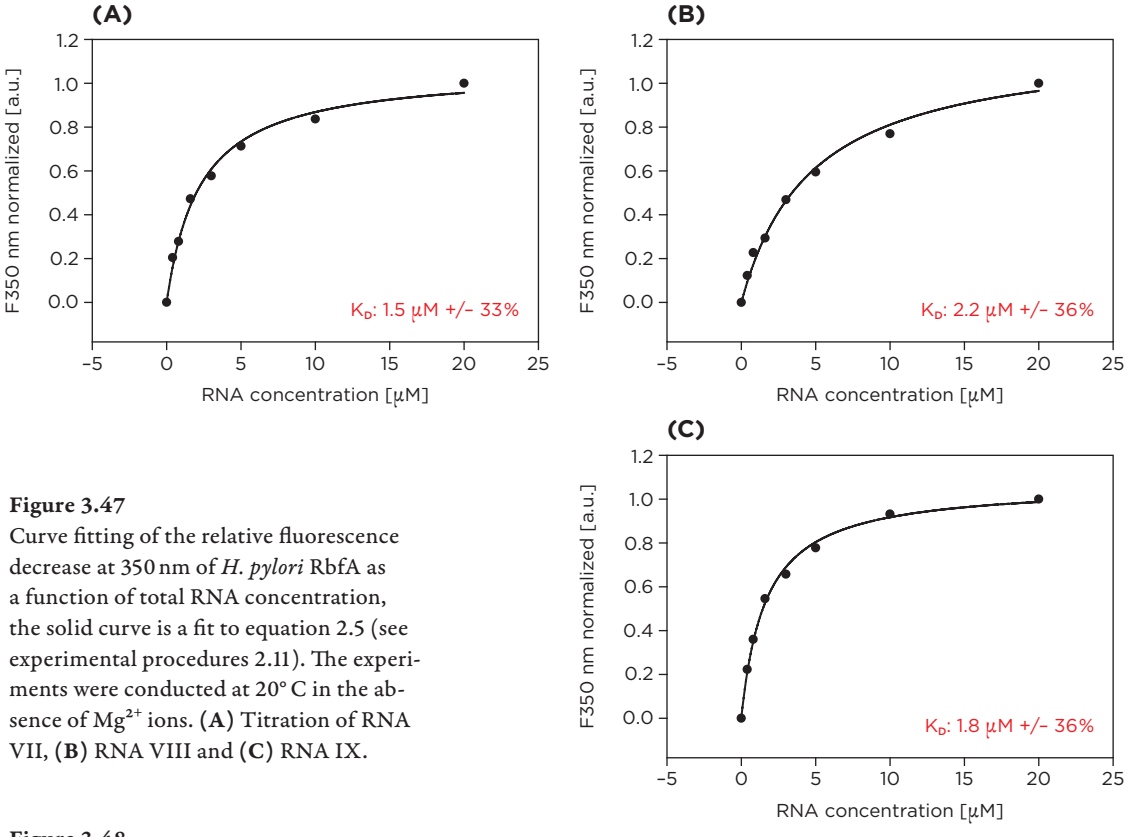
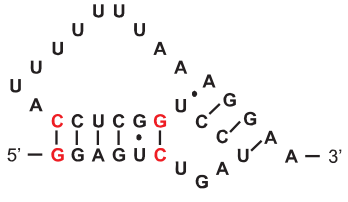
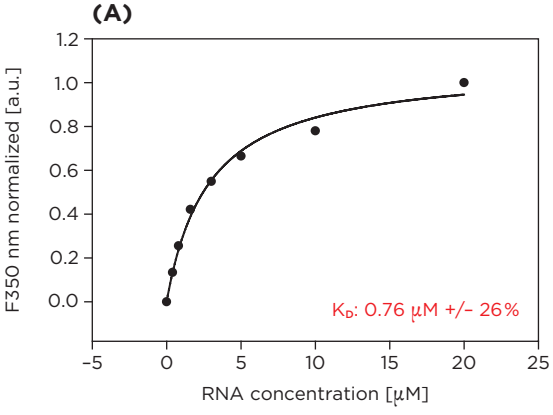
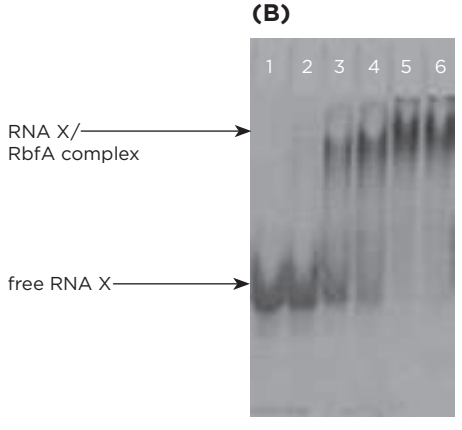


Figure 3.47
Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). The experiments were conducted at 20° C in the absence of Mg^{2+} ions. **(A)** Titration of RNA VII, **(B)** RNA VIII and **(C)** RNA IX.

Figure 3.48
Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). The experiments were conducted at 20° C. **(A)** Titration of RNA X. **(B)** Native gel shift analysis of the interaction between *H. pylori* RbfA and RNA X. Lane 1 and 2: RNA X in the absence of protein. For lanes 3–6: the concentration ratios were 1:1, 1:2, 1:3 and 1:4, respectively.







The binding affinity increases only slightly from 1.3 μM to 0.9 μM. The fluorescence experiments with the mutated constructs showed that RNA VII (1.10 μM) has roughly the same binding constant in the presence of Mg^{2+} ion as RNA V (1.14 μM), indicating that there is no influence of the length of the polyU linker on the binding affinity. Stabilization of the 3'-end of helix I or stabilization of helix II did not lead to a significant increase in the binding affinity (see Figure 3.47).

However, a new construct was designed which includes additional GC base pairs at the end of the helix I and replacing the GU base pair adjacent to the loop with a GC base pair (RNA X) (see Figure 3.48). This RNA construct showed in the fluorescence quenching experiments an increased binding affinity with KD values of ~800 nM. This result was supported by native gel analysis, in which the complex formation was saturated at a protein:RNA ration of 1:3. The fluorescence data showed no significant influence of the divalent cations in the binding affinity, as was already seen for the stabilized helix I construct RNA IV. The results clearly indicated that the RNA constructs including the pseudoknot structure interact with the RbfA protein. Furthermore the results showed that the RNA constructs resembling the pseudoknot possessed a slightly higher binding affinity for the RbfA protein.

The RNA construct RNA X was then selected for NMR titrations. The 1D ^{1}H imino proton spectra showed a subset of peaks that are similar to the ones already identified for the RNA IV construct containing only helix I. Three additional peaks can be seen at positions 11.74 ppm, 12.22 ppm and 12.32 ppm. In order to assign the RNA nucleotides to their respective resonances a 2D NOESY experiment was used (see Figure 3.49). The base paired resonances of helix I could be identified by their neighboring NOE signals to neighboring base pairs and showed a pattern similar to the one observed for the RNA IV (see Figure 3.43). The additional resonances at positions 11.74 ppm and 12.22 ppm could be identified as signals from GC base pairs by comparison of the NOESY spectrum region between 5.3 ppm and 8.5 ppm. The third additional signal at 12.32 ppm did not show specific NOEs to identify it without a doubt. For all of the additional peaks no internucleotide cross signals could be identified.

One reason was the signal overlap in the respective region. In addition, the helix II formation might not be conformationally stable. Conformational exchange leading to enhanced line broadening made it difficult to observe any sequential NOE cross peaks. With the assignments of the 2D NOESY experiment the 1D ^{1}H imino proton spectra from the titrations were analyzed (see Figure 3.50). RbfA was added to RNA X and a series of 1D ^{1}H NMR

spectra were collected at RNA:protein ratios of 1:0, 1:0.25, 1:0.7 and 1:1. As the molar ratio of RbfA protein increases, peaks in the imino region broaden significantly, indicating that the free RNA and the RNA-protein complex were in intermediate exchange on the NMR time scale. Imino proton resonance U24 shifted upon addition of RbfA, whereas all guanine H1 imino signals showed significant line broadening.

In order to test the influence of divalent cations the RNA X construct was titrated with increasing concentrations of Mg^{2+} ions and 1D ^1H imino proton spectra were recorded at each step. This NMR data are in good agreement with the fluorescence data where the stabilized constructs show no significant influence on the RNA constructs. The imino resonance of G21 shows line broadening upon titration with Mg^{2+}, the other signals stay unperturbed with increasing Mg^{2+} concentration. These results suggested that the RNA constructs resembling the pseudoknot structure interact with the RbfA protein in a similar manner as the constructs resembling the helix I.

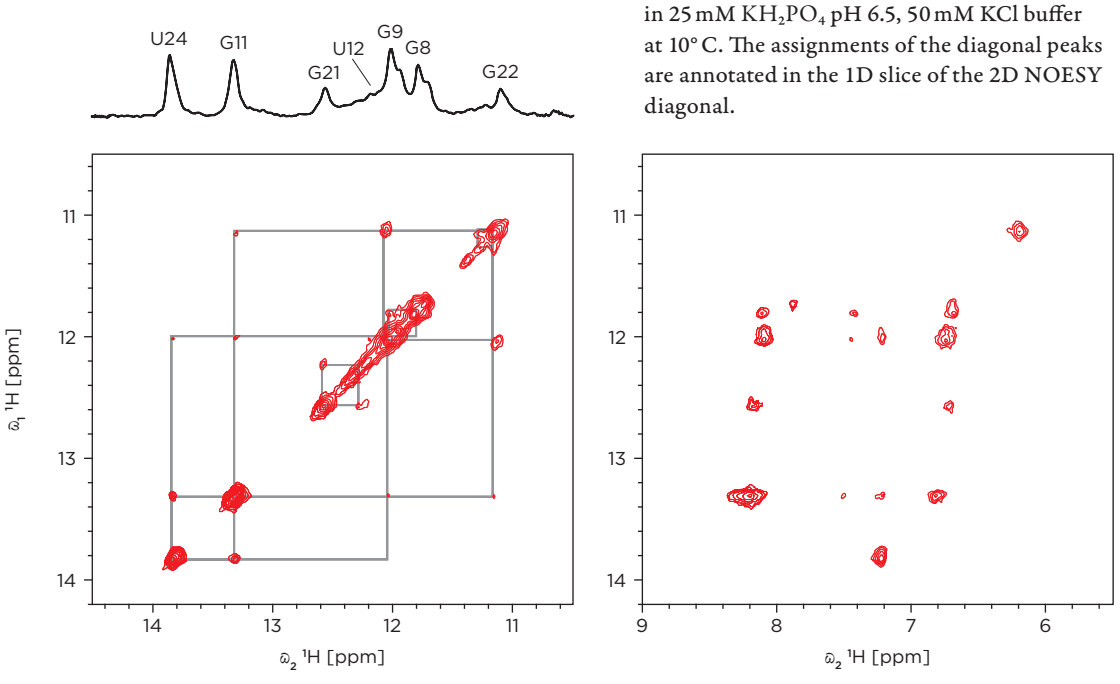


Figure 3.49
2D NOESY spectra of RNA X. The spectra were recorded on samples with 500 µM RNA samples in 25 mM KH_2PO_4 pH 6.5, 50 mM KCl buffer at 10° C. The assignments of the diagonal peaks are annotated in the 1D slice of the 2D NOESY diagonal.

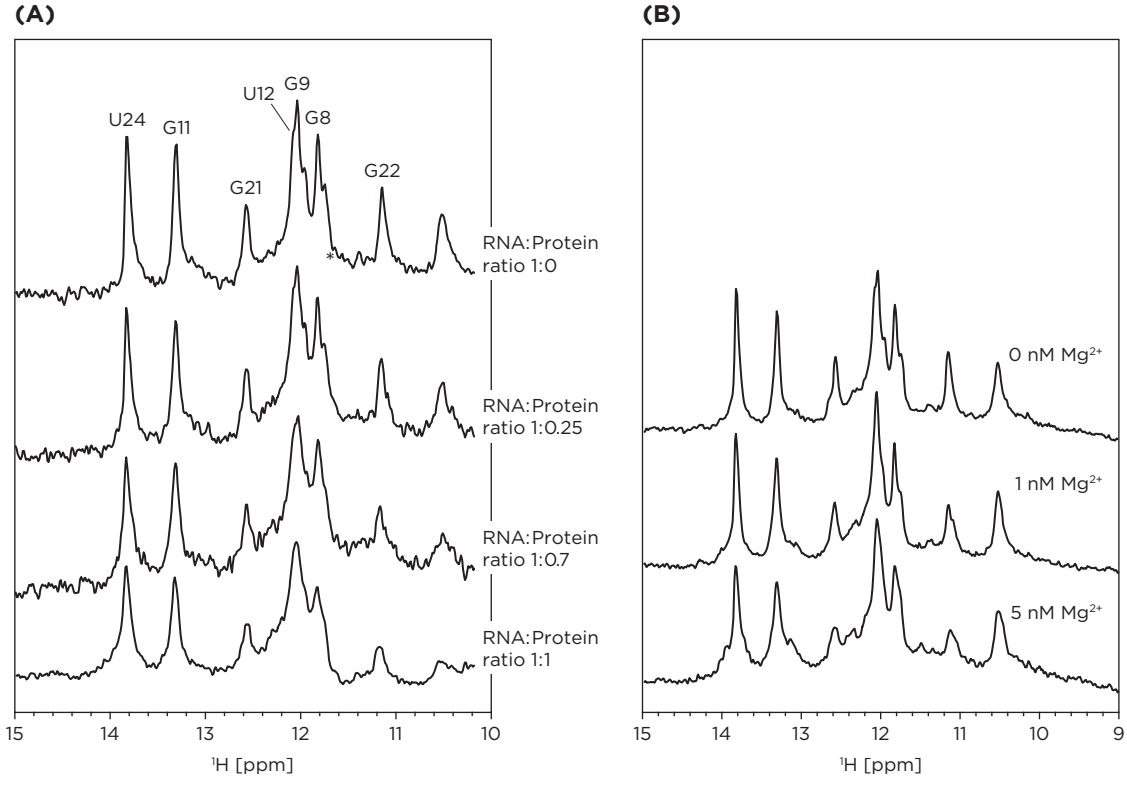


Figure 3.50
(A) 1D ^{1}H imino proton spectra of the RNA X construct upon titration with *H. pylori* RbfA. The spectra were recorded on samples with 200 µM RNA concentration at 10° C. RNA:protein ratios used in the titration were 1:0, 1:0.25, 1:0.7 and 1:1.
(B) 1D ^{1}H imino proton spectra of the RNA X construct upon titration with Mg^{2+}. The spectra were recorded on RNA samples with 200 µM RNA concentration at 10° C. Mg^{2+} concentrations used in the titration were 1 mM and 5 mM.

Interactions with the pre-16S rRNA

The 17S rRNA

Since the RbfA protein was suggested to be involved in the last maturation steps of the 16S rRNA (Bylund et al., 1998), a series of constructs was designed which were derived from the pre-16S rRNA (17S rRNA) sequences.

Previous studies suggested an alternative RNA fold of the immature *E. coli* 16S rRNA (Klein et al., 1985; Li et al., 1999). This alternative fold involves nucleotides from the pre 16S rRNA sequence, which form an alternative helical fold involving nucleotides from the helix I region. Similar to the situation in *E. coli*, nucleotides in the pre-sequences of *Helicobacter pylori*

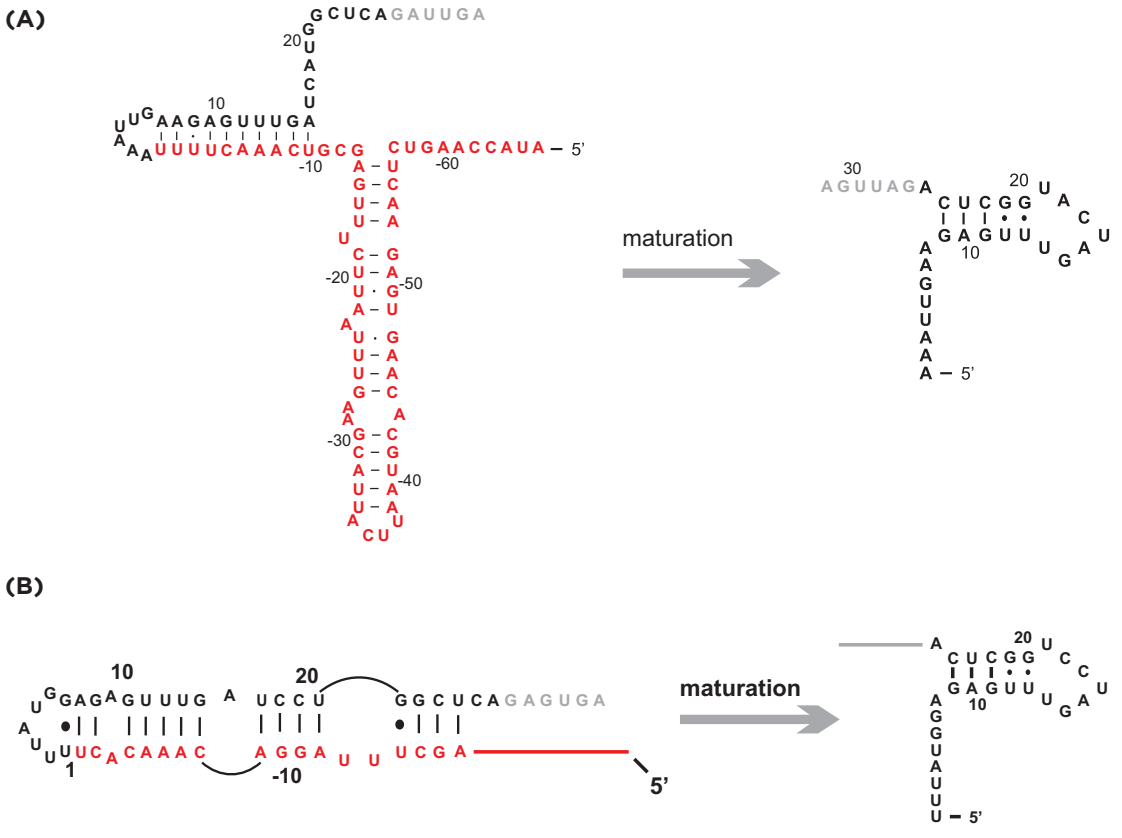


Figure 3.51
Fold of the pre 16S rRNA sequences from *E. coli* **(A)** and *Helicobacter pylori* **(B)**. The nucleotides from the pre 16S rRNA, the helix I region and the 16S rRNA are colored red, black and grey respectively.

16S rRNA can form an alternative helix with nucleotides of helix I despite a complete lack of sequence conservation in the pre-16S rRNA sequences of the two bacteria (see Figure 3.51). The possible alternative helix in *H. pylori* is longer then the one observed for *E. coli* and includes nucleotides 1–25 of the 16S rRNA and nucleotides –18 to –1 of the pre-sequence.

The first RNA constructs (RNA XI) included the nucleotides 1–26 from the helix I region and nucleotides –1 to –18 from the pre-16S rRNA sequence to form a hairpin loop structure (see Figure 3.52).

Furthermore a RNA construct was designed which included the helix I nucleotides 1–15 and the pre-sequence from nucleotide –1 to –18 of the

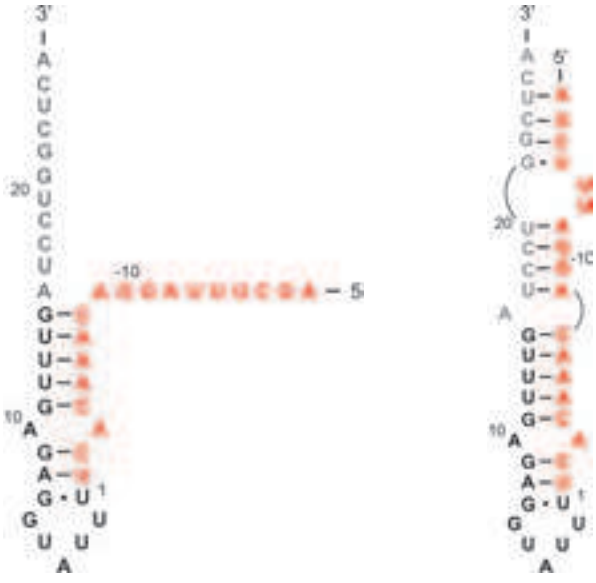


Figure 3.52
Secondary structure of the pre 16S rRNA constructs. Left: RNA XII; Right: RNA XI

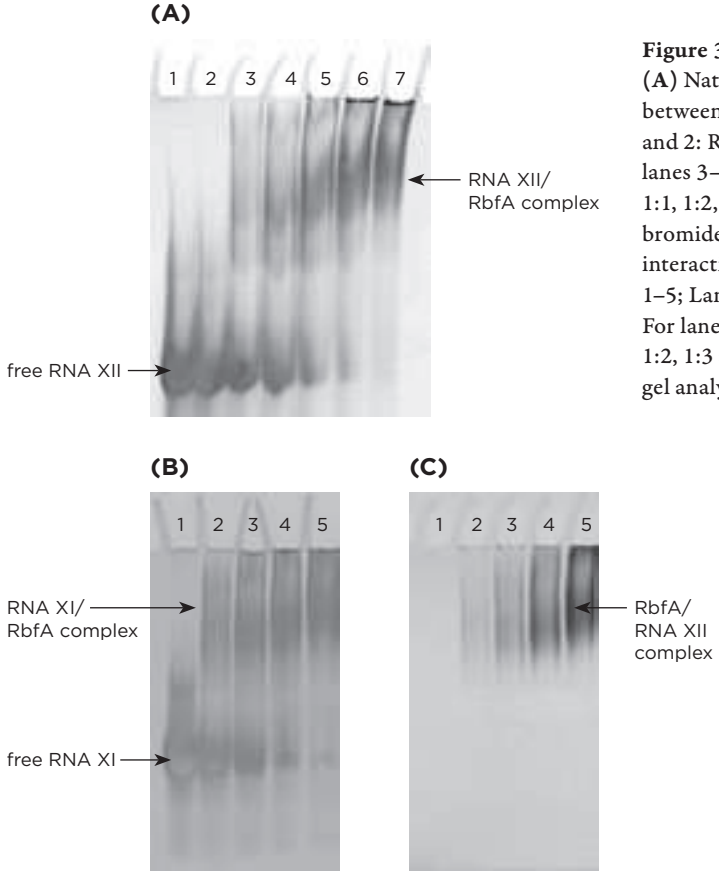


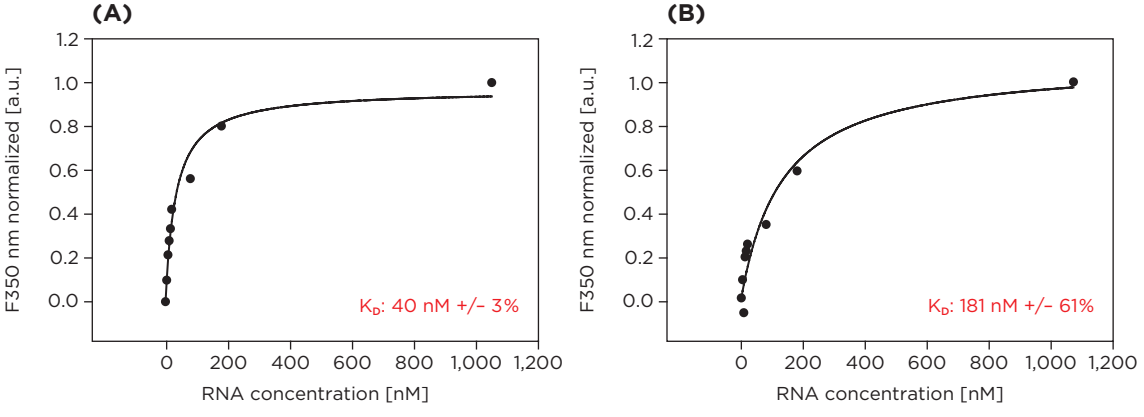
Figure 3.53
(A) Native gel shift analysis of the interaction between *H. pylori* RbfA and RNA XII. Lane 1 and 2: RNA XII in the absence of protein. For lanes 3–6: the concentration ratios were 1:0.5, 1:1, 1:2, 1:3 and 1:4, respectively. **(B)** Ethidium bromide stained native gel shift analysis of the interaction between *H. pylori* and RNA XI. Lane 1–5; Lane 1: RNA XI in the absence of protein. For lanes 2–5: concentration ratios were 1:1, 1:2, 1:3 and 1:4. **(C)** Coomassie stained native gel analysis of the same gel as in (B).

H. pylori 16S rRNA sequence (RNA XII). The binding of the RNA constructs including the pre-sequence was also analyzed by gel mobility shift assay and fluorescence quenching. Figure 3.53 shows the native PAGE analysis of RbfA/RNA XI and RbfA/RNA XII interactions. With increasing amounts of protein the formation of the complex was clearly visible, even at RNA:protein ratios of 1:2 suggesting a higher binding affinity compared to the helix I region constructs.

To clearly identify the protein as the binding partner, the native gels were subsequently stained with Coomassie brilliant blue to detect the protein. Figure 3.53 (C) shows the presence of the protein in the detected complex bands. This suggested a higher binding affinity compared to the helix I region constructs. Fluorescence titrations were performed to evaluate the apparent equilibrium constants of the pre 16S rRNA sequences (see Figure 3.54). The results of the initial titration indicated a much higher binding affinity compared to the mature 16S rRNA sequences used before (see Figure 3.54 A). Therefore the RNA concentration range was changed to concentrations in the nanomolar range. The new titration range provided data which could be fitted with equation 2.5 (experimental procedures). The extracted KD values showed a significant increase in affinity by a factor of ~20 compared to the constructs resembling helix I. Both RNA constructs showed binding affinities in the lower nanomolar range.

The titrations were repeated in the presence of 5 mM Mg^{2+}. The influence of the divalent cations was quite large in case for RNA XI, were the binding constant was increased from 181 nM

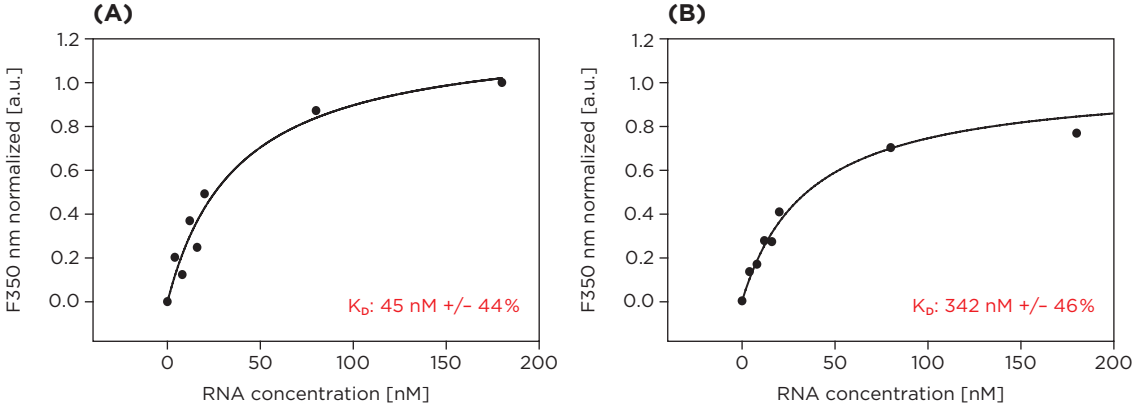
Figure 3.54
Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). The experiments were conducted at 20° C in the absence of Mg^{2+}. **(A)** Titration of RNA XII with RNA concentrations in the nanomolar range. **(B)** Titration of RNA XI.



to 45 nM. For the RNA XII construct the effect is almost negligible. The binding affinity changes from 40 nM to 41 nM. In order to characterize the involvement of the single stranded 5'-end of construct RNA XII for binding, new constructs were designed which include the nucleotides 1–15 of the helix I region and nucleotides –1 to –8 from the immature 16S rRNA. The 10 nucleotides at the 5' end were either exchanged by their complementary bases (e.g. A ≥ G) (RNA XIII) or exchanged for a poly uracile sequence (RNA XIV) (see Figure 3.55).

Figure 3.55
Secondary structure of RNA constructs with alternating sequences at the 5' end of the 16S rRNA ssRNA part.

Figure 3.56
Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see experimental procedures 2.11). The experiments were conducted at 20° C in the absence of Mg^{2+}. **(A)** Titration of RNA XIII. **(B)** Titration of RNA XIV.



The constructs RNA XIII and RNA XIV were as well tested for binding to the *H. pylori* RbfA protein. For the construct RNA XIII the measurements revealed that the presence of the complementary bases did not influence the fluorescence quenching.

The presence of the poly uracile sequence at the 5'-end of the RNA XIV construct decreased the binding affinity significantly. Still both constructs exhibit an increased binding affinity by a factor of 20 (RNA XVIII) and 3 (RNA XIV) compared to the RNA constructs mimicking the helix I region.

In order to test the influence of the dsRNA part of construct RNA XI, a RNA construct was designed which lacked the ssRNA at the 5'-end (RNA XV). This short hairpin loop construct was as well tested for interaction with the RbfA protein from *H. pylori*. The fluorescence and native gel mobility shift analysis showed interactions with the protein in a similar manner as was shown for the constructs RNA XI and RNA XII. It also showed binding affinities in the same nanomolar range (Figure 3.57).

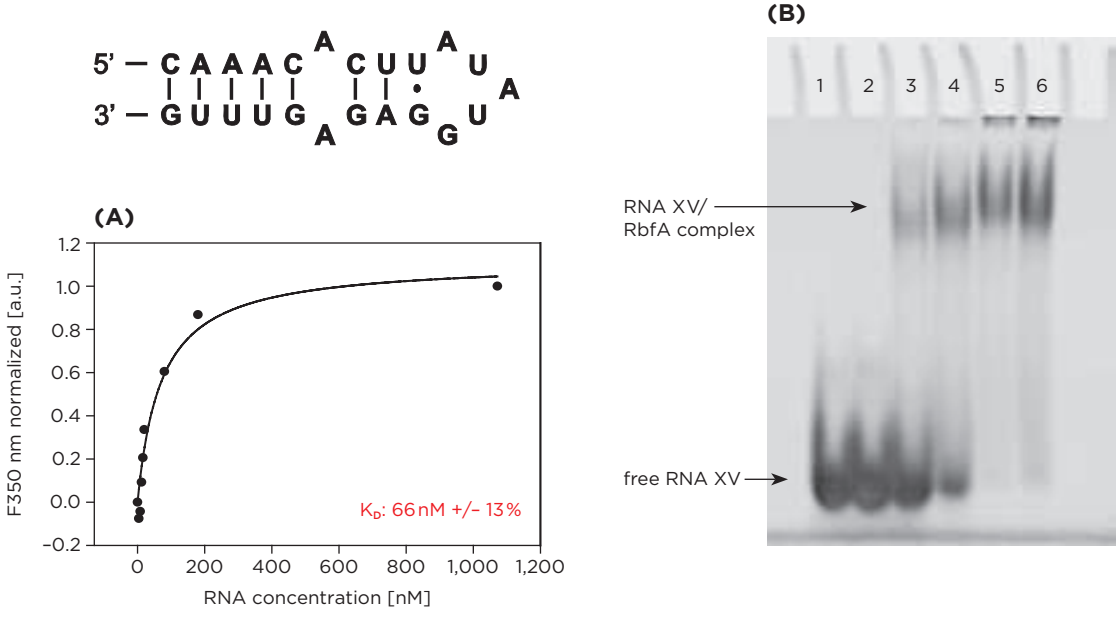


Figure 3.57
Curve fitting of the relative fluorescence decrease at 350 nm of *H. pylori* RbfA as a function of total RNA concentration, the solid curve is a fit to equation 2.5 (see *Experimental Procedures* 2.11). The experiments were conducted at 20° C in the absence of Mg^{2+} ions. **(A)** Titration of RNA XV **(B)** Native gel shift analysis of the interaction between *H. pylori* RbfA and RNA XV. Lane 1 and 2: RNA XV in the absence of protein. For lanes 3–6: the concentration ratios were 1:1, 1:2, 1:3 and 1:4, respectively.

This shorter construct was then chosen for further investigation using NMR techniques. RNA titration experiments were performed and monitored by a series of 1D ^{1}H spectra. To test for any influence of divalent cations on the folding of the helix loop structure the RNA XV was titrated with Mg^{2+} ions, although the fluorescence data showed that the binding constant did not change significantly. The experiments showed that the presence of low concentrations of the divalent cation Mg^{2+} lead to the appearance of novel signals, greatly improved the resolution and minimized signal crowding in the imino region (see Figure 3.58).

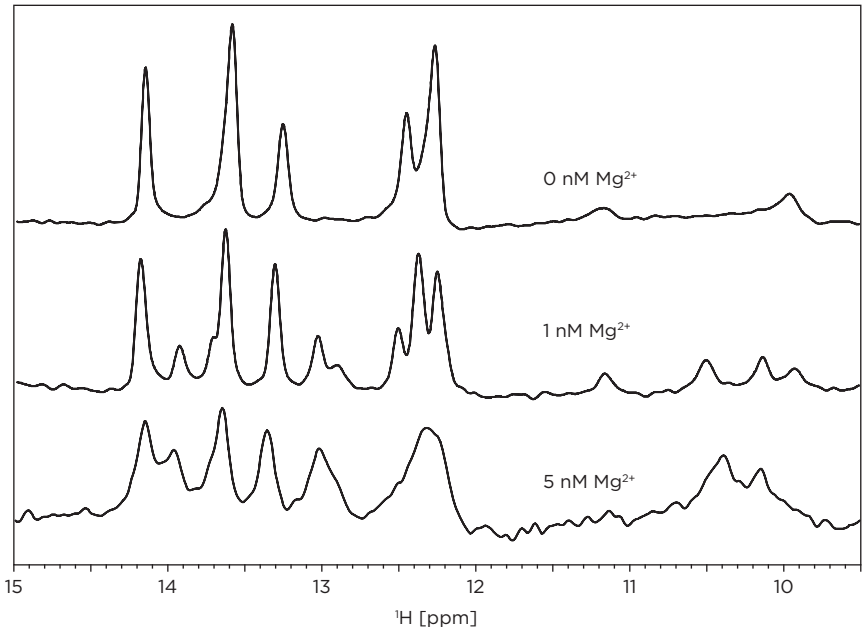


Figure 3.58
1D ^{1}H imino proton spectra of the RNA XV construct upon titration with Mg^{2+}. The spectra were recorded on samples with 200 μM RNA concentration at 10° C. Mg^{2+} concentrations used in the titration were 1 mM and 5 mM.

RNA XV showed 6 out of the 9 expected imino proton signals in the absence of divalent cations. In the presence of Mg^{2+} ions the number of imino proton signals which are well resolved increased to 11. This suggested that upon addition of Mg^{2+} ions the hairpin loop structure folds into different conformations, which are in chemical exchange to each other. Further addition of Mg^{2+} ions lead to extensive line broadening. A 2D NOESY experiment was conducted to assign the imino proton region. Lack of cross peaks between the diagonal imino proton resonances made it difficult to assign the resonances without a doubt. Several resonances were also subjected to signal overlap. A 2D NOESY experiment was repeated in the presence of Mg^{2+} cations. Again

the lack of cross peaks between the diagonal imino proton resonances made it impossible to correctly assign the resonances. The construct was then used in a titration experiment with the *H. pylori* RbfA protein (see Figure 3.59). The 1D spectrum showed extensive line broadening for all resonances, suggesting intermediate exchange on the NMR time scale.

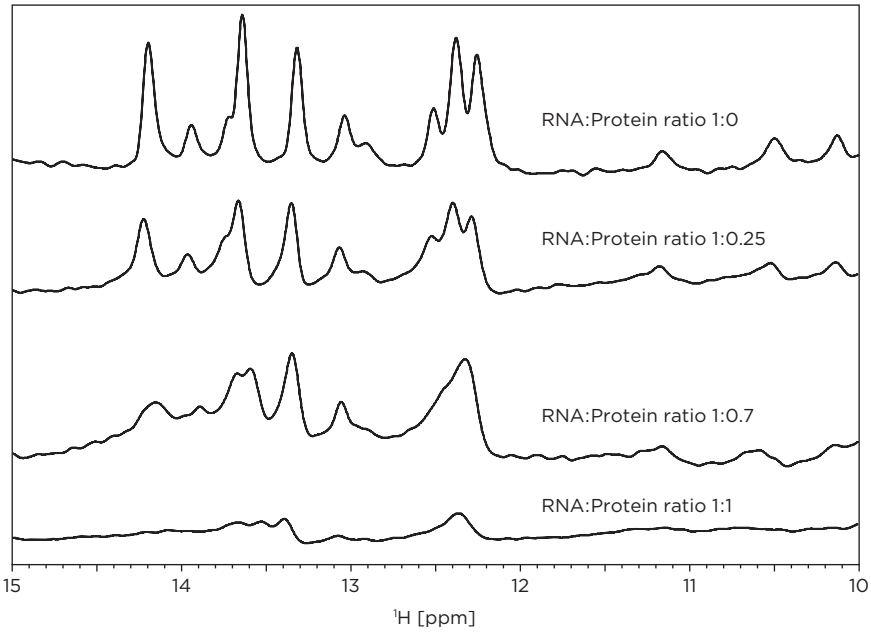


Figure 3.59
1D ^{1}H imino proton spectra of the RNA XV construct upon titration with *H. pylori* RbfA. The spectra were recorded on samples with an RNA concentration of 200 µM at 10° C. RNA:protein ratios used in the titration were 1:0, 1:0.25, 1:0.7 and 1:1.

In order to test the interactions of the most favorable RNA constructs with the *H. pylori* protein, titrations were performed on ^{15}N labeled *H. pylori* RbfA and monitored by a series of 2D ^{15}N-edited HSQC spectra. The protein concentration was fixed and the RNA concentrations were varied resulting in RNA:protein ratios from 1:0 to 1:3. The first HSQC spectra at a ratio of 1:0.25 revealed intermediate exchange on the NMR time scale, where due to extensive line broadening no resonances could be detected except resonances at the C-terminal end of the RbfA protein. Therefore, the temperature was raised from 10° C to 25° C to push the exchange rate into fast exchange on the NMR time scale. The ^{15}N-edited HSQC spectra at 25° C could then be used to identify the RbfA residues involved in the interaction with the RNA XV construct.

Figure 3.60 A shows an overlay of the ^{15}N edited spectra at RNA:protein ratios of 1:0 and 1:2. By plotting the chemical shift changes against the

sequence it could be shown that the premature RNA constructs bind at the similar regions of the RbfA protein as was show for the constructs of the mature 16S rRNA constructs mimicking the helix I region.

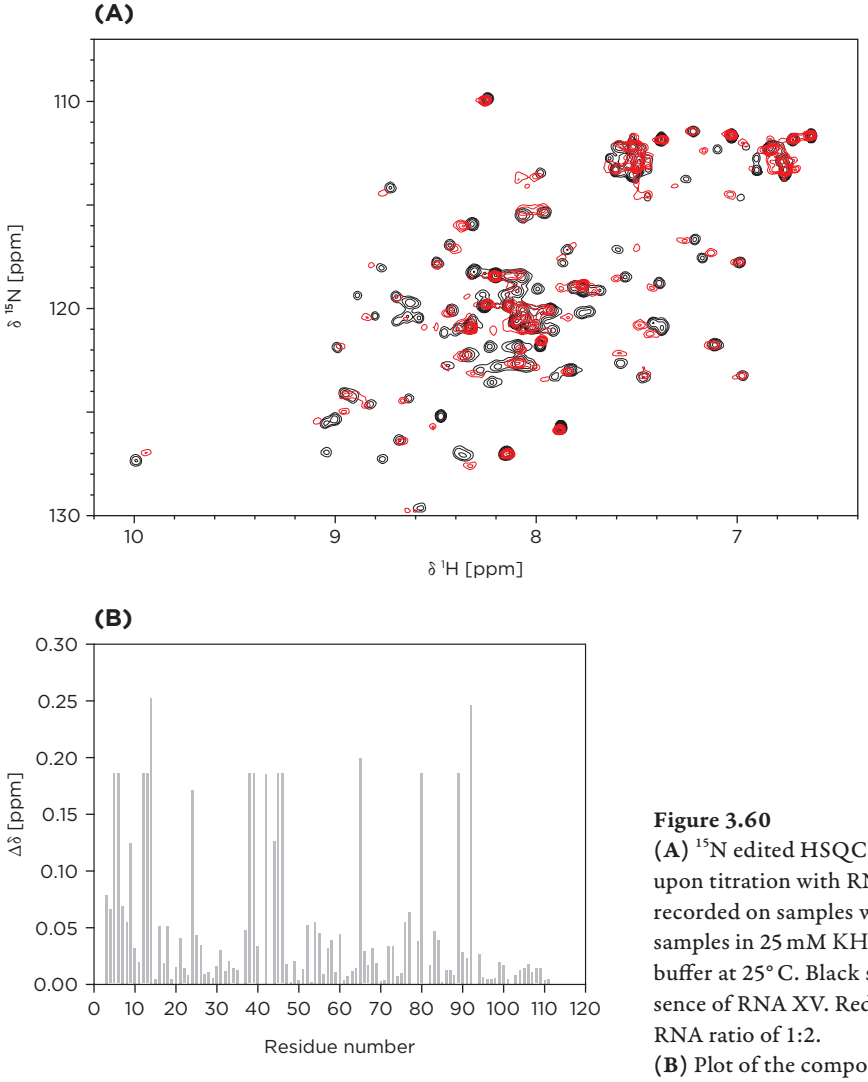


Figure 3.60
(A) ^{15}N edited HSQC spectra of *H. pylori* RbfA upon titration with RNA XV. The spectra were recorded on samples with 300 μM protein samples in 25 mM KH_2PO_4 pH 6.5, 50 mM KCl buffer at 25° C. Black spectrum is in the absence of RNA XV. Red spectrum is at a protein: RNA ratio of 1:2.
(B) Plot of the composite ^{1}H and ^{15}N chemical shift changes upon titration of RbfA with construct RNA XV versus the residue number. The composite chemical shift changes were calculated using the empirical equation $\Delta\delta = (\Delta\delta_{H^2} + (0.17\ \Delta\delta_N)^2)^{1/2}$ where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift changes of ^{1}H and ^{15}N, respectively.

Summary

Here, we used a combination of several biophysical methods to demonstrate that RbfA is an RNA-binding protein. The RNA-binding surface of *H. pylori* RbfA is made up mostly by amino acids of helix α1, β-strands β1 and β 2, helix α3 and the loop region connecting the β-sheets β1 and β2. These amino acids form a continuous surface of the protein and include many conserved basic residues as expected for an RNA-binding protein.

Comparison with the available three dimensional structures of KH domain nucleic acid complexes, namely those of hnRNP K, NOVA-2, SF1 and FBP (Braddock et al., 2002a, 2002b; Lewis et al., 2000; Liu et al., 2001) showed one common feature, these KH domain type I folds bind to ssDNA or ssRNA. A core motif of four nucleotides, mainly pyrimidines, could be identified in each case. These nucleotides are recognized by the KH domain type I fold (βααββα) topology (Backe et al., 2005). Helices α1 and α2, the GXXG motif and strand β2 of the KH domain interact with the DNA or RNA molecules (Liu et al., 2001, Backe et al., 2005). The ssRNA or ssDNA molecules lie in a narrow groove between the GXXG motif and a variable loop connecting strands β2 and β3 (in type I KH domain proteins). The GXXG motif sticks out of the protein surface and is solvent exposed. The large RNA binding surfaces are hydrophobic, due to numerous conserved aliphatic side chains that are located at the surface of the protein. In addition, conserved positively charged side chains flanking the hydrophobic RNA binding groove, form compensating electrostatic interactions with the solvent-exposed phosphate backbone of the RNA.

In contrast, the AXG motif between the helices α2 and α3, which is highly conserved within the RbfA family, is significantly less exposed to the solvent. This structural attribute favors our results that the AXG motif plays only a minor role in RNA binding. On the other hand the variable loop connecting strands β1 and β2 (in type II KH domains) seems as well to play a major role in binding to the dsRNA in the RbfA protein. This specifically includes the positively charged, conserved amino acids within this loop. The RbfA protein from *H. pylori* with a KH domain type II folding topology was shown to preferentially bind to structured RNAs and the binding surface is placed just opposite to the binding surface of ssDNA or ssRNA molecules in the type I KH domain folds. Previous studies showed that the sequence specific binding of ssRNA or ssDNA involves mostly specific contacts between the nucleic acids and a hydrophobic surface of the proteins. In the RbfA protein, however, several charged amino acids seem to contribute to the binding as well.

We therefore favor a picture that RbfA recognizes the structural fold of the central pseudoknot region, instead of binding to a specific RNA sequence. The fluorescence quenching experiments showed that the affinity of RbfA for RNA-oligonucleotides resembling helix I of 16S rRNA is rather low. It binds with higher affinity to RNAs that are able to fold into a pseudoknot-structure that resembles the central pseudoknot of 16S rRNA. Since these RNA mimics still show a binding affinity around 1 μM it seems that the presence of the helix II, however, plays a less pronounced role in the binding affinity. Intriguingly, mutations that stabilize the intrinsically unstable secondary structure of the helix I and helix II lead to an increased binding affinity (see Table 3.4). These results suggest that the RbfA protein binds to a preformed native RNA secondary structure instead of acting as a chaperone for the helix I region of the 16S rRNA.

However, RNA constructs mimicking the precursor 16S rRNA sequences have affinities in the lower nM range.

Table 3.4
Binding affinities of RNA constructs.

RNA construct	$K_D - Mg^{2+}$ [μM]	$K_D + Mg^{2+}$ [μM]	Chemical exchange in protein titrations (NMR)	Chemical exchange in RNA titrations (NMR)
RNA I	20.6	3.31	fast	
RNA II	8.23	2.71	fast	
RNA III		2.62	fast	
RNA IV	2.65	2.17	intermediate / fast	fast
RNA V	4.12	1.14	intermediate	
RNA VI	1.28	0.92	intermediate	
RNA VII	1.57	1.07	intermediate	
RNA VIII	2.22	1.23	intermediate	
RNA IX	1.78	0.99	intermediate	
RNA X	0.762	0.813	intermediate	fast
RNA XI	0.181	0.045	intermediate	
RNA XII	0.040	0.041	intermediate	
RNA XIII	0.045		intermediate	
RNA XIV	0.342		intermediate	
RNA XV	0.66	0.071	intermediate	intermediate

For the *E. coli* precursor rRNA it was shown that the helix I region forms a helix-loop structure involving nucleotides 1–15 of the helix I region and nucleotides –1 to –8 from the immature 16S rRNA (Li et al., 1999).

For the leader region of the *H. pylori* 16S rRNA, we found that an alternative helix could extend even further and include the full helix I region. Although the RbfA binds to the full length constructs, removal of 8 base pairs at the end of the helix results in higher binding affinities. This suggested that the base pairs adjacent to the loop play the major role in binding the alternative helix fold (see Table 3.4). These results suggest that the alternative helix fold of the precursor sequences is probably the native target for the RbfA protein.

Conclusions

Conclusions

In this thesis the three dimensional solution structure of the RbfA protein from *T. maritima* is presented. The protein possesses the folding topology of type II KH domain. The KH domain proteins are known as RNA or DNA binders (Liu et al., 2001; Backe et al., 2005). In the RbfA family, the GXXG motif that is commonly found in KH-domain proteins is replaced by the less solvent exposed AXG motif. The relaxation rates analyzed for the *T. maritima* RbfA identified flexible N- and C-termini as well as several flexible regions in the core of the protein. To gain insights into the function of the RbfA protein the RbfA protein from *H. pylori* was used, which in contrast to the genetically well characterized *E. coli* and the *T. maritima* protein could be isolated in its functional relevant full-length form. For comparison of the two RbfA proteins several biophysical methods were used, which verified the close resemblance of the two proteins. The results for the two proteins were then compared to other RbfA protein family members. These revealed significant differences in the structural features between mesophile and thermophile proteins. Calculation of the relative contact order revealed that the thermophile RbfA protein from *T. maritima* has a more compact fold. In addition CD-spectroscopic experiments revealed that protein surface charge-charge interactions play a major role in the stabilization of the thermophile proteins.

In a second a part of this thesis I describe results which demonstrate that the KH domain type II RbfA protein indeed binds to the 16S rRNA in a non-sequence specific way. The protein employs ionic interactions as well as contacts via a hydrophobic surface including amino acid residues in helix $\alpha 1$, strands $\beta 1$ and $\beta 2$, the variable loop connecting the strands and helix $\alpha 3$. These findings assign RbfA a novel role in RNA binding. Thus far, KH domain proteins were known to bind sequence specific ssRNA and ssDNA molecules (Liu et al., 2001; Backe et al., 2005). Our results show the RbfA KH domain as a binder of structured RNAs but not of single stranded RNA sequences. In addition the role of RbfA seems to be to recognize a folded structure rather then a specific sequence.

Genetic evidence pointed towards an interaction of RbfA with the 5' terminal helix of the 16S rRNA (Dammel and Noller, 1995). However, our data show that RbfA binds only weakly to RNAs mimicking either helix I or the fold of the central pseudoknot region of the 16S rRNA. Instead RbfA binds much tighter to a hairpin structure that would correspond to an alternative

helix formed between nucleotides of the pre-rRNA sequence and nucleotides of the helix I, which cannot coexist with the structure of the mature 16S rRNA and a folded central pseudoknot (shown in Figure 4.1). By binding to this precursor structure the RbfA protein possibly influences the 16S rRNA processing and thereby might be involved in the formation of the 5' helix. In agreement with this model is the finding that the lack of RbfA leads to a reduced processing of the 17S to 16S rRNA, thus a role of RbfA in the late maturation step of the small ribosomal subunit was suggested (Bylund et al., 1998). RbfA binding to the precursor structure during maturation might be important for protecting and presenting the correct cleavage site for the endoribonuclease RNAseG. It possibly destabilizes the basepair adjacent to the loop in the precursor helix including the position of nucleotide U1—the cleavage site for RNAseG (Li et al., 1999)—and renders its accessibility for cleavage. In addition, our findings that RbfA preferentially binds to structured RNA mimics of the folded central pseudoknot of 16S rRNA suggested a second role for RbfA in capturing a native structure and protect it from rearrangements into non-native interactions.

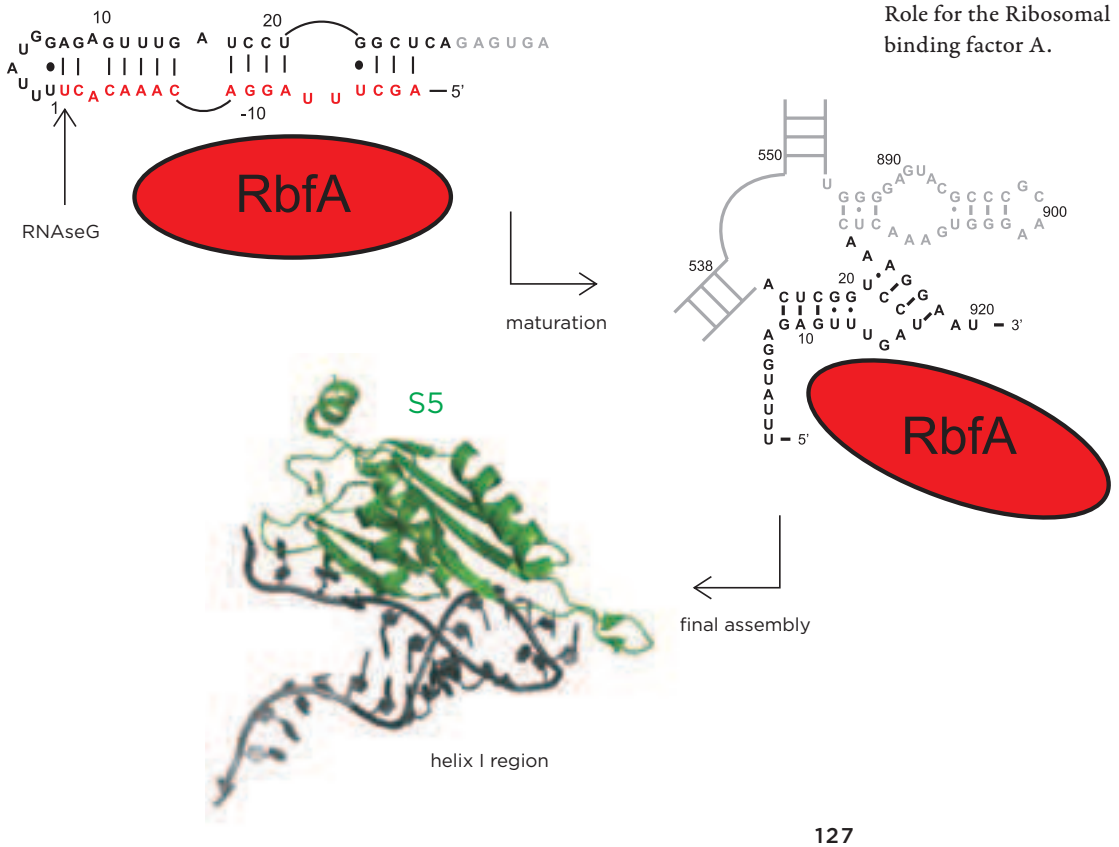


Figure 4.1
Role for the Ribosomal binding factor A.

The RbfA protein then pre-organizes the binding site for ribosomal protein S5 by capturing and protecting the native helix I RNA fold essential for the ribosome assembly, acting as a surrogate for ribosomal proteins. This would be a necessary step during maturation and or upon a cold shock to prevent misfolding of the 16S rRNA.

A possible misfolding of the central pseudoknot that connects the 16S rRNA 5'-domain, the central-domain and the 3'-domains has a profound effect on the overall stability of the ribosomal subunit (Poot et al., 1998). Such a misfolding of the central pseudoknot structure leads to inhibition of translation *in vivo,* due to the inability to form 70S initiation complexes after the loss of several ribosomal proteins (Poot et al., 1998). These include the proteins S2, S18 and S21, where S2 and S21 are dependent on the ribosomal protein S5 to bind (Traub and Nomura, 1968; Culver, 2003; Mizushima and Nomura, 1970; Held et al., 1974). The ribosomal protein S5 is known to bind in the late assembly steps to the helix I region of the 16S rRNA (Traub and Nomura, 1968; Culver, 2003; Mizushima and Nomura, 1970; Held et al., 1974). The S5 protein interacts with the RNA via a surface that comprises the extended loop between strands β1 and β2 and the N-terminal part of helix α1 (Bycroft et al., 1995). These regions contain a number of conserved basic residues, which make non-specific ionic interactions with the phosphate backbone of the RNA. Due to this similar binding mode and binding site we favor a synergistic model of S5 and RbfA where the RbfA pre-organizes the binding site of S5 prior to the final assembly step of the 30S initiation complex. Then S5 replaces RbfA in late maturation leading to the assembly of a functional 30S initiation complex. Whether, a delta RbfA mutant then lacks the S5 protein in the 30S ribosomal subunits needs to be proven.

Appendix

Appendix

A.1 NMR Parameter

Table A.1
Acquisition and processing parameters for the Experiments of *T. maritima* RbfA

Experiment	Frequency[a]	SW1[b]	td1[c]	SW2[d]	td2[e]	SW3[f]	td3[g]
pseudo 2D watersled	600		33	12	4,096		
15N HSQC	700	25	256	12	1,024		
13C HSQC	700	70	900	12	2,048		
13C CT HSQC ali	700	70	400	12	2,048		
13C CT HSQC aro	700	35	400	12	2,048		
HNCO	700	16	96	25	76	14	1,024
HNCA	700	36	160	25	76	14	1,024
HNCACB	700	74	180	25	76	14	1,024
CBCACONH	700	74	180	25	76	14	1,024
HBHACBCACONH	700	6	256	25	76	14	1,024
HNCACO	600	11	128	25	80	14	2,048
HCCCONH	700	8	160	40	128	14	1,024
HCCH TOCSY	700	8	160	48	160	14	1,024
HCCH COSY	700	7.5	160	48	160	14	1,024
15N NOESY HSQC	900	12	200	24	100	14	1,024
13C NOESY HSQC ali	900	11	200	26	100	14	2,048
13C NOESY HSQC aro	700	12	200	26	100	14	2,048
HNHA	600	26	80	12	120	14	2,048
T1 HSQC	600	26	256	14	2,048		
T2 HSQC	600	26	256	14	2,048		
HET NOE	600	26	256	14	2,048		
IPAP HSQC aniso	800	5	400	12	1,024		
IPAP HSQC iso	800	5	400	12	1,024		
HNCA rdc aniso	700	16	200	25	64	14	1,024
HNCA rdc iso	700	16	200	25	64	14	1,024

Experiment	n[h]	d1[i]	mix[j]	15N offset[k]	13C offset[l]	ω1[m]	ω2[n]	ω3[o]
pseudo 2D watersled	512	3						
15N HSQC	16	1.5		119		512	2,048	
13C HSQC	8	1.2			38	1,024	1,024	
13C CT HSQC ali	32	1.2			38	1,024	1,024	
13C CT HSQC aro	16	1.2			120	1,024	1,024	
HNCO	8	1.5		119	174	512	256	2,048
HNCA	8	2		119	54	512	256	2,048
HNCACB	16	1.5		119	39	512	256	2,048
CBCACONH	16	1.5		119	39	256	512	2,048
HBHACBCACONH	8	1.3		119	39	256	512	2,048
HNCACO	8	1.5		119	174	512	256	2,048
HCCCONH	8	1		119	25	512	256	2,048
HCCH TOCSY	8	1.2		119	39	512	256	2,048
HCCH COSY	8	1.2		119	39	512	256	2,048
15N NOESY HSQC	8	2.2	120	119		512	256	2,048
13C NOESY HSQC ali	8	1.5	120	119	38	512	512	2,048
13C NOESY HSQC	8	1.8	100	119	125	512	512	2,048
HNHA	16	1.5		119	54	512	256	2,048
T1 HSQC	8	2		119		256	2,048	
T2 HSQC	8	2		119		256	2,048	
HET NOE	8	2		119		256	2,048	
IPAP HSQC aniso	48	1.2		119		1,024	2,048	
IPAP HSQC iso	24	1.2		119		1,024	2,048	
HNCA rdc aniso	8	1.3		119	54	512	256	2,048
HNCA rdc iso	8	1.3		119	54	512	256	2,048

(a) Proton frequency in MHz. (b) Spectral width in the first dimension in ppm. (c) Number of points in the first dimension in points. (d) Spectral width in the second dimension in ppm. (e) Number of points in the second dimension in points. (f) Spectral width in the third dimension in ppm. (g) Number of points in the third dimension in points. (h) Number of transients. (i) Relaxation delay in seconds. (j) Mixing time in milliseconds. (k) Carrier frequency in ppm. (l) Carrier frequency in ppm. (m) Number of points after Fourier transformation in the first dimension in points. (n) Number of points after Fourier transformation in the second dimension in points. (o) Number of points after Fourier transformation in the third dimension in points.

Table A.2
Acquisition and processing parameters for the Experiments with the *H. pylori* RbfA

Experiment	Frequency[a]	SW1[b]	td1[c]	SW2[d]	td2[e]	SW3[f]	td3[g]
pseudo 2D watersled	600		33	12	4,096		
15N HSQC	700	26	256	12	2,048		
13C HSQC	700	75	256	14	2,048		
13C CT HSQC ali	700	75	256	16	2,048		
HNCO	700	16	96	25	76	14	1,024
HNCA	700	36	160	25	76	14	1,024
HNCACB	700	74	180	25	76	14	1,024
CBCACONH	700	74	180	25	76	14	1,024
HBHACBCACONH	700	6	256	25	76	14	1,024
15N NOESY HSQC	900	12	200	24	100	14	1,024
HNHA	600	26	72	12	144	14	2,048
T1 HSQC	600	26	256	14	2,048		
T2 HSQC	600	26	256	14	2,048		
HET NOE	600	26	256	14	2,048		
IPAP HSQC aniso	700	5	400	12	1,024		
IPAP HSQC iso	700	5	400	12	1,024		
Titration HSQC	600–800	24	256	12	1,024		

Experiment	n[h]	d1[i]	mix[j]	15N offset[k]	13C offset[l]	ω1[m]	ω2[n]	ω3[o]
pseudo 2D watersled	512	3						
15N HSQC	8	1.3		116		512	2,048	
13C HSQC	8	1.4			40	1,024	1,024	
13C CT HSQC ali	8	1.4			40	1,024	1,024	
HNCO	8	1.5		116	174	512	256	2,048
HNCA	8	1.5		116	54	512	256	2,048
HNCACB	8	1.5		116	39	512	256	2,048
CBCACONH	8	1.5		116	39	256	512	2,048
HBHACBCACONH	8	1.3		116	39	256	512	2,048

Experiment	n[h]	d1[i]	mix[j]	15N offset[k]	13C offset[l]	ω1[m]	ω2[n]	ω3[o]
15N NOESY HSQC	8	2.2	120	116		512	256	2,048
HNHA	16	1.2		116	40	512	256	2,048
T1 HSQC	8	2		116		256	2,048	
T2 HSQC	8	2		116		256	2,048	
HET NOE	8	2		116		256	2,048	
IPAP HSQC aniso	24	1.2		116		1,024	2,048	
IPAP HSQC iso	24	1.2		116		1,024	2,048	
Titration HSQC	16	1.5		116		512	2,048	

(a) Proton frequency in MHz. (b) Spectral width in the first dimension in ppm. (c) Number of points in the first dimension in points. (d) Spectral width in the second dimension in ppm. (e) Number of points in the second dimension in points. (f) Spectral width in the third dimension in ppm. (g) Number of points in the third dimension in points. (h) Number of transients. (i) Relaxation delay in seconds. (j) Mixing time in milliseconds. (k) Carrier frequency in ppm. (l) Carrier frequency in ppm. (m) Number of points after Fourier transformation in the first dimension in points. (n) Number of points after Fourier transformation in the second dimension in points. (o) Number of points after Fourier transformation in the third dimension in points.

Table A.3
Acquisition and processing parameters for the Experiments with the RNA molecules.

Experiment	Frequency[a]	SW1[b]	td1[c]	SW2[d]	td2[e]
1D watergate RNA	600	22	8,192		
2D 1H1H NOESY RNA	600	14	800	22	2,048

Experiment	n[h]	d1[i]	mix[j]	ω1[m]	ω2[n]
1D watergate RNA	32–256	2		2,048	
2D 1H1H NOESY RNA	32	1.8	80	1,024	2,048

(a) Proton frequency in MHz. (b) Spectral width in the first dimension in ppm. (c) Number of points in the first dimension in points. (d) Spectral width in the second dimension in ppm. (e) Number of points in the second dimension in points. (h) Number of transients. (i) Relaxation delay in seconds. (j) Mixing time in milliseconds. (m) Number of points after Fourier transformation in the first dimension in points. (n) Number of points after Fourier transformation in the second dimension in points.

A.2 Chemical shift

Table A.4
Chemical shift for *H. pylori* RbfA at 298K in ppm.

Residue	HN	N	HA	CA	HB2	HB3	CB
M1							
N2			4.639		2.747	2.669	
A3	8.443	125.215	4.166	49.952	1.232		16.292
H4	8.326	117.968	4.51	53.306	3.125	3.022	26.674
K5	8.208	122.703	4.141	54.124	1.715	1.646	30.262
E6	8.491	121.271	4.201	54.296	1.981	1.878	27.056
R7	8.224	121.707	4.287	53.479	1.792	1.697	27.438
L8	8.235	123.7	4.209	53.823	1.542		39.424
E9	8.727	119.592	4.346	56.016			26.751
S10	8.019	115.29	4.347	57.22	3.839		60.723
N11	8.358	120.932	4.433	51.242	2.918		36.523
L12	8.539	120.834	4.02	55.586	1.74	1.637	39.042
L13	8.129	119.303	3.977	57.22			38.813
E14	7.587	117.221	3.977	56.704	2.007		26.522
L15	8.011	118.927	4.054	55.027			40.416
L16	8.64	119.776		54.124			38.737
Q17	8.16	118.467	3.943	56.102	2.153		25.911
E18	7.83	119.215	3.882	56.532	2.136	2.041	26.369
A19	8.302	123.14		51.414			17.208
L20	8.106	118.971	4.201	55.887	1.371		37.286
A21	7.233	117.657		51.242			15.146
S22	7.17	112.301	4.184	56.704	3.736	3.452	61.258
L23	7.139	121.706		52.102			38.889
N24	8.679	119.462	4.587	50.64	2.781	2.695	35.148
D25	7.005	117.906	4.527	50.769	2.446		39.195
S26	8.97	125.07	4.218	58.038	3.865		60.113
E27	8.443	120.212	4.011	55.371	1.775	1.448	26.903
L28	7.319	114.043	4.286	52.059			40.187

Residue	HN	N	HA	CA	HB2	HB3	CB
N29	8.004	113.545	4.355	51.543	3.056	2.713	35.53
S30	7.241	111.304	4.303	54.253	3.934	3.822	61.182
L31	7.013	123.326	4.295	51.93	1.706	1.577	41.256
S32	8.412	117.064	4.803	55.156	3.684		61.563
V33	8.679	126.441	4.519	58.425	1.86		29.804
T34	8.813	117.849	4.192	60.016			65.533
K35	7.422	118.841	4.424	53.608	1.732	1.645	32.018
V36	8.577	120.585	4.829	57.694			32.476
E37	9.096	125.724	4.828	51.844			30.262
C38	9.466	124.696		56.016			26.14
S39	8.695	120.398		55.586			61.487
K40	8.884	124.386	4.484	55.93	3.908		29.194
G41	8.823	111.272	3.873/3.771	42.64			
K42	7.88	113.982		54.124			
H43							
H44	7.744	119.839	5.026	52.231	2.704	2.523	30.033
A45	8.491	126.877	4.854	47.801	0.837		18.658
Y46	8.733	120.634	4.493	54.898	2.807		37.057
V47	8.853	124.821	5.155	58.941			29.346
F48	8.82	127.5		53.866			36.904
V49	8.774	114.357	5.302	55.371	1.817		32.858
L50	8.404	127.436	4.622	50.726	2.024	1.232	37.515
S51	8.098	118.217	4.303	56.489			62.25
S52	8.978	121.831	4.57	54.468	3.942	3.796	61.258
D53	7.972	123.388		50.726			39.195
H54	9.065	126.816	4.115	56.79	3.168		26.522
K55	8.404	120.77	3.916	56.747	1.697		28.507
I56	7.186	121.644	3.435	60.403	2.102		33.316
L57	7.72	119.153		55.414			37.439
S58	7.972	115.169		59.199			59.731
K59	7.634	122.764	3.822	57.737	1.938	1.577	30.721
L60	7.949	121.023		55.027			37.439

Residue	HN	N	HA	CA	HB2	HB3	CB
K61	7.264	116.66	4.131	56.661	2.024		29.346
K62	7.592	118.467	4.106	55.973	1.843		29.423
A63	7.485	120.771		48.919			15.299
E64	7.422	121.208		58.898			26.98
G65	8.609	105.51	3.865/3.727	44.532			
L66	7.485	123.388	4.063	54.855	1.826	1.431	39.5
I67	8.145	121.769	3.461	62.941	1.904		35.454
R68	8.892	119.339	3.891	57.478	1.852		27.59
Q69	7.87	117.346	3.925	56.317	2.162	2.067	25.376
F70	7.833	120.336	3.985	58.984			35.836
V71	8.31	119.319		64.403			28.736
L72							
Q73	7.87	118.904	3.813	55.543	1.852		25.224
A74	8.012	120.148	3.856	50.984	0.974		16.444
S75	7.619	109.995		56.446			61.105
G76	7.642	110.929		43.113			
W77							
F78							
K79	8.452	121.333	4.141	54.296	3.28	1.732	27.056
C80	8.247	121.769		53.479			27.209
P81							
K82	7.394	120.678	3.968	54.425	1.732	1.516	29.652
L83	8.475	123.139	4.717	50.769	1.491	0.346	41.18
S84	7.932	117.8	4.484	54.726	3.658		61.869
F85	8.924	124.137	5.207	54.769	2.816	2.592	37.744
V86	8.852	120.461	4.372	56.618	1.465		32.858
S87	8.491	117.907	4.708	57.364	3.822	3.745	61.411
D88	8.664	124.446	4.648	52.102	2.617	2.386	39.729
N89	8.349	121.956	4.665	50.898	2.755		35.836
S90	8.357	115.913	4.149	57.694	3.83		60.647
L91	8.011	122.726	4.218	53.091	1.594	1.508	38.966
E92	7.941	120.024		54.855			26.98

Residue	HN	N	HA	CA	HB2	HB3	CB
K93			4.088		1.732		
Q94	8.09	119.9	4.132	53.694	2.007	1.913	26.14
L95	8.066	122.017	4.192	53.005	1.585	1.431	39.424
R96	8.043	120.959	4.218	53.608	1.775	1.688	27.667
L97	8.121	122.64	4.166	53.651	1.603	1.456	39.195
D98	8.239	119.65	4.424	52.145	2.592		38.049
A99	7.846	123.016	4.157	50.425	1.276		16.139
I100	7.767	118.716	3.917	59.242	1.697		35.606
F101	8.074	122.454	4.458	55.672	3.056	2.927	36.446
N102	8.121	119.775	4.527	50.855	2.695	2.609	35.759
E103	8.058	120.895	4.123	54.296	1.946	1.869	27.285
I104	7.965	121.456	3.968	58.941	1.757		35.606
A105	8.133	126.938	4.174	49.866	1.25		16.139
K106	8.051	120.397	4.183	53.78	1.757	1.688	30.11
G107	8.247	109.808	3.917/3.839	42.382			
K108	8.091	120.648	4.261	53.435	1.757	1.654	30.186
D109	8.326	120.896	4.501	51.844	2.635	2.497	38.279
N110	8.207	118.404	4.648	50.425	2.755	2.6	36.599
D111	7.886	125.817		53.22			39.271

Table A.5
Chemical shift for RbfA protein from *T. maritima* at 313K in ppm.

Residue	HN	N	CO	HA	CA	HB2	HB3	CB	HG2	HG3
M1										
N2										
P3			175.39	4.38	62.341	2.354	1.952			
A4	8.217	121.992	176.66	4.28	51.631	1.396		16.652		
Y5	7.869	118.996	174.58	4.551	56.599	3.156	3.083	36.46		
R6	7.963	121.153			55.473			28.378		
K7	8.104	121.998	174.87	4	57.468	1.887		31.226	1.422	
A8	8.037	120.847	178.13	4.493	52.076	1.496		16.453		
M9	7.981	118.483	175.22	4.185	55.804	1.938	1.852	30.431	1.705	
L10	8.112	121.973	177.06	4.173	56.003	1.688	1.596	40.104		
E11	8.815	118.778	176.49	3.628	59.117	2.341	2.093	26.788	2.491	2.161
S12	7.877	113.292	175.21	4.41	59.562	4.113		61.237		
E13	8.188	122.234	177.15	4.235	57.129	2.292	2.135	26.986	2.521	2.44
I14	8.668	118.674	174.78	3.598	64.483	1.97		35.334	0.825	0.693
Q15	8.28	119.822	175.3	3.676	59.24	2.6	2.167	26.531	2.471	2.374
K16	7.698	117.345	177.98	4.161	57.739	2.067		30.63	1.74	1.534
L17	8.361	119.895	177.53	4.36	55.861	2.164	1.516	41.107		
L18	8.901	119.118	177.23	4.065	56.467	1.857	0.873	39.716	1.936	
M19	7.669	116.956	176.96	4.594	56.136	2.535	2.331	29.247	2.983	2.917
E20	7.731	119.834	177.33	4.18	57.256	2.443	2.304	26.986	2.643	2.085
A21	8.44	123.127	177.48	3.015	53.083	0.886		16.055		
L22	8.404	117.34	176.28	4.098	56.015	2.357	2.033	40.309	1.752	
Q23	7.343	112.793	175.53	3.953	55.963	2.344	2.226	25.869	2.728	2.63
Q24	7.534	114.908	174	4.227	54.047	2.418	2.05	27.185	2.574	2.479
L25	7.246	121.421	174.53	3.662	53.665	1.626	1.294	39.914		
R26	8.28	124.476	173.38	4.366	53.287	1.82	1.675	27.384		
D27	7.619	122.425			50.173					
P28			175.33	4.476	62.209	2.366	2.025	30.157		
R29	8.551	116.992	174.82	4.177	54.739	1.839	1.689	28.775	1.654	
L30	7.469	118.723	174.26	4.383	51.998	1.675	1.487	41.702	1.434	

Residue	HN	N	CO	HA	CA	HB2	HB3	CB	HG2	HG3
K31	8.994	128.537	175.04	4.562	52.818	2.027	1.797	28.444	1.373	
K32	7.837	120.518	175.42	3.668	58.702	1.832	1.778	30.763	1.272	
D33	8.567	115.001	174.18	4.352	53.752	2.481	2.095	37.189		
F34	7.572	115.97	172.99	4.63	53.344	2.713	2.656	37.122		
V35	7.269	121.787	172.49	4.324	60.712	2.04		31.172	1.085	0.412
T36	8.332	119.85	170.9	4.23	59.387	3.922		69.796	1.073	
F37	8.505	125.429	173	5.614	53.046	3.09	2.85	34.804		
S38	8.753	121.549	173.26	4.211	58.22	3.884	3.779	61.767		
R39	7.498	116.622	171.5	4.632	54.175	1.851		31.491	1.503	
V40	8.581	123.681	172.1	4.906	59.295	1.977		32.499	0.86	0.753
E41	9.43	126.271	174.28	4.893	51.998	2.128	1.985	30.1	2.23	
L42	9.922	129.499	175.2	4.994	52.34	1.768	1.599	42.025		
S43	8.878	119.474	175.12	4.368	56.782	4.259	3.685	62.043		
K44	8.927	123.104	175.45	4.053	57.894	1.918		30.034	1.585	1.54
D45	7.879	114.488	173.81	4.504	51.031	3.083	2.63	38.315		
K46	8.021	113.387	172.74	3.579	56.362	2.101	1.929	27.516	1.417	1.302
R47	7.584	116.162	174.54	3.918	56.07	1.61	1.386		1.255	
Y48	7.971	115.248	171.36	5.339	54.431	2.931	2.623	39.971		
A89	9.035	123.119	171.75	4.911	48.479	0.937		18.109		
D50	8.698	124.531	172.53	5.224	51.264	2.892	2.335	39.839		
V51	9.076	124.285	172.16	4.574	59.24	1.534		30.642	0.674	0.361
Y52	8.821	127.064	173.57	4.974	56.295	2.906	2.658	37.454		
V53	9.099	122.497	173.66	4.81	60.285	2.084		31.768	1.218	0.866
S54	9.36	124.587	170.79	5.063	55.036	3.86		62.545		
F55	8.615	124.348	172.34	5.455	54.791	2.833	2.734	42.426		
L56	7.471	129.817	171.44	4.351	51.834	1.36	1.123	38.255	1.412	
G57	7.115	109.294	171.3	4.225 4.151	42.687					
T58	8.86	113.13		4.73	58.852	4.876		66.65	1.447	
P59			177.61	4.278	63.895	2.494	2.031	29.971		
E60	8.65	117.903	177.05	4.056	58.223	2.124	1.95	26.456	2.507	2.327
E61	7.765	120.4	178.06	4.005	56.751	2.283	1.901	28.113	2.414	

Residue	HN	N	CO	HA	CA	HB2	HB3	CB	HG2	HG3
R62	8.558	124.095	175.8	3.915	58.322	3.162		27.98		
K63	7.855	119.508	177.23	3.76	58.2	1.984	1.928	30.365	1.533	1.389
E64	7.814	118.841	176.06	4.02	56.99	2.033	1.975	26.854	2.339	2.226
T65										
V66										
E67	7.748	120.18	177.45	4.163	57.982	2.215	2.12	26.655	2.444	2.306
I68	8.101	120.725	177.02	3.497	63.988	1.901		35.941	0.93	0.865
L69	8.023	119.797	176.46	3.752	56.403	1.066	0.47	36.93	1.712	
N70	8.248	114.554	178.21	4.93	54.85	2.902	2.804			
R71	8.229	121.809	175.47	4.3	57.45	2.073		27.683	1.996	1.788
A72	7.663	121.664	176.68	4.95	49.423	1.649		17.049		
K73	7.704	119.46	176.81	4.571	59.57	2.526	2.458	29.173	1.618	1.875
G74	9.291	111.848	174.23	4.086 4.046	45.171					
F75	7.822	124.72	175.94	4.391	58.974	3.331	3.183	36.394		
F76	7.756	116.613	175.1	3.799	60.935	3.136	2.967	37.255		
R77	8.381	119.141	176.09	3.878	58.818	2.237	2.054	28.378	1.631	
T78	8.127	114.725	173.61	3.931	64.493	4.213		66.888	1.269	
F79	7.959	122.604	177.09	3.855	60.317	2.892	2.578	37.586		
I80	8.271	118.34	175.08	3.77	62.857	2.157		34.946	1.759 1.594	0.954
A81	8.405	121.867	178.05	4.02	53.585	1.506		16.254		
K82	7.925	114.127	176.29	4.071	56.38	1.753	1.68	30.564	1.441	1.371
N83	7.403	115.199	173.13	4.538	53.859	2.555	2.384	39.706		
L84	8.128	119.529	173.72	4.51	52.686	1.627		40.633	1.294	
R85	7.81	118.487	173.92	4.055	54.886	1.879		25.86	1.575	
L86	7.884	118.141	174.71	4.348	52.317	1.423		41.296		
Y87	8.086	120.493	173.94	4.409	57.465	3.044	2.979	36.857		
V88	7.576	117.538	172.47	4.054	59.331	1.839		31.238	0.852	0.795
A89	8.502	128.818			47.391			16.585		
P90			173.2	4.242	60.337	1.878	1.49	28.858		
E91	7.891	119.211	173.14	4.103	54.809	2.056	1.987	28.245	2.342	

Residue	HN	N	CO	HA	CA	HB2	HB3	CB	HG2	HG3
I92	8.31	126.579	172.71	4.554	57.018	1.741		35.807	1.416 1.196	0.012
R93	8.117	124.609	171.36	4.44	53.411	2.035	1.452	30.895		
F94	8.495	118.807	173.47	5.339	54.987	2.65	2.58	38.712		
Y95	9.116	119.978	173.37	4.723	54.48	3.015	1.835	41.097		
E96	9.162	120.356	173.9	4.258	55.798	2.067		29.769	2.431	2.056
D97	8.746	123.008	173.92	4.891	51.774	2.917	2.372	38.911		
K98	8.376	124.201	175.12	4.396	54.492	1.991	1.739	30.696	1.374	
G99	8.428	108.791	172.42	4.011 3.93	43.593					
I100	7.984	119.549	174.53	4.167	59.422	1.872		36.734	1.441 1.181	0.904
E101	8.389	124.166	174.2	4.281	54.678	2.043	1.966	27.649		
A102	8.159	124.381	176.05	4.299	50.844	1.394		17.182		
S103	8.085	113.947	173.17	4.417	56.931	3.891		61.731		
V104	7.915	121.261	174.59	4.087	61.369	2.114		30.63	0.923	
K105	8.08	122.805	175.15	4.277	55.014	1.778		30.564		
V106	7.839	119.272	174.34	3.967	61.314	2.059		30.498	0.935	0.887
H107	8.343	120.261	172.84	4.55	54.693	3.298	3.239	26.655		
Q108	8.232	119.932	174.24	4.227	55.009	2.068		27.251		
L109	8.085	122.316	175.37	4.274	54.016	1.632		40.302		
L110	8.064	121.567	175.02	4.314	54.148	1.735		40.17		
V111	7.859	119.659	174.73	4.051	61.237	2.048		30.564	0.883	
Q112	8.184	122.473	176.32	4.311	54.55	2.085		27.251		
L113	8.137	121.843	175.72	4.321	53.605	1.667	1.556	40.567		
G114	8.178	108.183	171.66	3.965 3.881	43.408					
Y115	7.811	119.476	172.33	4.552	55.539	2.94		37.122		
D116	8.164	123.294			49.047			39.441		
P117			174.97	4.339	61.536	2.277	1.961			
L118	8.134	119.46	173.23	4.306	53.366	1.743	1.602	39.706		
K119	7.753	120.542	173.29	4.34	54.17	1.868	1.709	31.359		
D120										

Residue	CG	CG2	HD2	HD3	CD1	CD2	HE	CE	HZ	CZ
M1										
N2										
P3										
A4										
Y5			7.146		130.507					
R6										
K7										
A8										
M9							3.265			
L10			0.843	0.809						
E11										
S12										
E13	33.868									
I14	28.301	21.992	0.664							
Q15	31.641									
K16	22.734		1.784				3.049			
L17			1.021	0.969	21.435					
L18			0.639	0.38						
M19										
E20	33.682									
A21										
L22	24.961		1.136	0.983	22.92	24.03				
Q23	31.27									
Q24	31.827									
L25			0.598	0.439	22.92	21.62				
R26										
D27										
P28										
R29			3.238	3.159						
L30			0.867		24.219					
K31			1.549							
K32			1.528							

Residue	CG	CG2	HD2	HD3	CD1	CD2	HE	CE	HZ	CZ
D33										
F34			7.007		128.077		7.351	128.638		
V35	21.064	19.951								
T36	19.394									
F37			7.114		128.747		7.204	128.77		
S38										
R39			2.938							
V40	19.023	18.837								
E41	33.868									
L42			0.861	0.835	23.105	22.92				
S43										
K44			1.756				3.058			
D45										
K46			1.675	1.59						
R47			3.087							
Y48			6.935		131.058					
A89										
D50										
V51	18.837									
Y52			6.908		130.6		6.515	114.979		
V53	20.693	19.208								
S54										
F55			6.854		127.927		7.484	129.962		
L56			0.775	0.686	22.363	21.80				
G57										
T58	19.951									
P59										
E60	34.61									
E61										
R62										
K63	22.549		1.752		27.188					
E64										

Residue	CG	CG2	HD2	HD3	CD1	CD2	HE	CE	HZ	CZ
T65										
V66										
E67										
I68	27.002	19.023	0.666		12.175					
L69			0.609	0.456	20.879	25.51				
N70										
R71			3.332	3.238						
A72										
K73			1.995				2.117			
G74										
F75			6.646		129.381		7.001	128.975	7.34	127.59
F76										
R77			3.291	3.229						
T78	19.765									
F79			7.411		129.577		6.935	128.949		
I80	27.002	19.032	0.911		11.619					
A81										
K82			1.565				2.882			
N83										
L84			0.812	0.792	23.291					
R85			3.204							
L86			0.868	0.84	23.105	21.06				
Y87			7.129				6.837	115.552		
V88	18.837									
A89										
P90										
E91										
I92	16.629		0.558							
R93										
F94			6.476		127.947		6.791	129.476		
Y95			6.632		130.462		6.663	115.58		
E96										

Residue	CG	CG2	HD2	HD3	CD1	CD2	HE	CE	HZ	CZ
D97										
K98			1.487				3.019			
G99										
I100	19.23									
E101										
A102										
S103										
V104										
K105										
V106	19.023									
H107			7.257		117.559					
Q108										
L109										
L110										
V111										
Q112										
L113										
G114										
Y115										
D116										
P117										
L118										
K119										
D120										

Table A.6
Chemical shift of RNA constructs RNA IV and RNA VI at 283K in ppm.

Base atom	RNA IV	RNA VI
G8H1	11.91	11.82
G8H21	7.93	8.11
G8H22	6.75	6.684
G9H1	12.02	11.96
G9H21	8.16	8.1
G9H22	6.78	6.74
A10	7.3	7.23
G11H1	13.37	13.32
G11H21	8.3	8.22
G11H22	6.85	6.82
U12H3	12.09	12.04
G22H1	12.69	12.57
G22QH2	8.2	8.16
G21H1	6.73	6.183
G21H21	11.17	11.13
G21H22	6.241	6.72
U17H3	13.87	13.84

A.3 Experimental data

Table A.7
Experimental restraints for *T. maritima* RbfA at 313K in Hz.

Residue	^{3}J (HNHA)	RDC H_N-N	RDC Hα-Cα
M1			
N2			
P3			
A4		2.47	13.68
Y5			4.95
R6	6.06	4.14	4.84
K7		8.26	17.17
A8		–2.63	–7.44
M9	5.59	11.27	6.44
L10			1.64
E11	3.07	10.32	23.67
S12	4.57	13.99	
E13		14.07	20.19
I14	4.44	12.64	–18.59
Q15		–0.3	–12.13
K16	4.59	15.82	4.77
L17		–1.77	–2.38
L18	3.62	8.28	
M19	4.98	7.74	–22.69
E20		14.25	–9.75
A21	3.51	10.65	–20.27
L22	2.37	4.12	26.86
Q23	4.61	11.71	–12.9
Q24		12.19	–3.24
L25	5.00	–6.55	17.82
R26		–9.96	22.63
D27	5.81	–8.78	16.38
P28			

Residue	^{3}J (HNHA)	RDC H_N-N	RDC Hα-Cα
R29	8.22	–4.6	–19.65
L30		3	1.64
K31	7.90	2.11	0.8
K32		3.8	3.99
D33	4.37	11.54	–27.52
F34		–13.88	48.75
V35	8.31	4.68	–13.01
T36		2.68	–1.7
F37	9.17	–6.59	3.28
S38	9.31	–6.7	30.91
R39	9.80	–6.4	10.61
V40	9.77	–9.3	7.32
E41	9.81	–0.01	25.97
L42	6.77	–16.03	4.93
S43	4.47	3.29	–6.43
K44	2.39	6.27	16.28
D45	8.39	11.06	–17.11
K46	6.97	–0.01	9
R47		8.71	0.06
Y48	8.61	4.64	–16.98
A89	9.14	–0.58	–16.98
D50	8.56	5.43	2.38
V51	8.78	–14.96	4.8
Y52	9.14	–2.15	28.37
V53	8.93	–14.15	
S54	10.74	5.6	–16.32
F55	9.17	4.8	–17.93
L56		–0.58	9.81
G57		8.58	5.75
T58	5.71	5.59	4.88
P59			
E60	4.26	4.75	17.89

Residue	^{3}J (HNHA)	RDC H_N-N	RDC Hα-Cα
E61		1.13	25.98
R62	3.95	3.73	
K63		–0.76	
E64	3.99	–3.81	
T65	3.62	–1.07	
V66		–1.55	
E67		11.21	
I68	3.85	–3.69	17.75
L69		–7.59	–33.33
N70	4.38	10.6	–25.15
R71		2.04	21.88
A72	9.13	–4.85	13.8
K73		6.25	31.75
G74	6.34	5.86	10.52
F75	4.62	12.45	–20.25
F76	4.61	5.44	8.87
R77	3.50	–0.3	25.23
T78	3.88	8.99	–36.45
F79	3.16	14.66	–9.82
I80	5.02	0.78	30.8
A81	3.10	4.05	6.53
K82	5.02	15.11	–8.22
N83	8.32	5.78	–12.92
L84			4.85
R85	6.90	–3.02	1.59
L86	7.84	7.58	–20.15
Y87		5.31	–34.13
V88	8.94	–1.46	2.49
A89	6.33	–3.65	20.23
P90			
E91		1.002	–18.6
I92	9.21	–1.49	–8.89

Residue	^{3}J (HNHA)	RDC H_N-N	RDC Hα-Cα
R93	9.56	–4.67	–5.63
F94	9.04	–12.94	0.03
Y95	9.98	–10.98	24.27
E96	8.70	–15.84	0.05
D97	7.18	4.63	–0.13
K98	7.28	–0.02	11.42
G99	6.42	6.32	
I100			4.13
E101		–0.19	0.04
A102		0.34	–0.87
S103		–2.27	2.4
V104		11.659	
K105		–0.97	
V106		–0.01	
H107			
Q108		3.82	
L109		–6.073	
L110			
V111		1.042	
Q112		–0.338	
L113			
G114		–2.66	0.83
Y115			
D116		–0.91	13.07
P117			
L118		–1.89	
K119			2.39
D120		–1.4	

Table A.8
Relaxation parameter R1, R2 and Het NOE from *T. maritima* RbfA.

Residue	R1[a]	R2[a]	Het NOE[a]	R1[b]	R2[b]	Het NOE[b]
M1						
N2						
P3						
A4				1.6990	7.1040	
Y5				1.8190	8.1770	
R6				1.8600	5.8450	0.4785
K7				1.9310	6.5610	0.6179
A8				1.8410	7.4880	0.6570
M9	1.5523	6.7935	0.5389	1.8980	7.4000	0.7079
L10	1.7343	7.5815	0.6848	1.8630	7.7310	0.7821
E11	1.6189	8.3752	0.7103	1.7920	8.4120	0.7755
S12	1.5765	7.0225	0.7617	1.7730	7.9900	0.7562
E13						
I14	1.5340	7.0872	0.7398	1.7450	7.8080	0.7998
Q15	1.6008	7.7519	0.7219	1.7380	6.1920	
K16	1.6584	7.3260	0.7148	1.7450	7.9970	0.8150
L17	1.5684	7.0900	0.7120	1.8440	8.3000	0.7793
L18	1.5446	7.6394	0.7070	1.8440	8.2420	0.7665
M19	1.6300	7.4571	0.7028	1.7280	7.8710	0.7806
E20				1.7670	8.3120	0.7988
A21	1.5596	7.0126	0.7838	1.8331	7.1899	0.7288
L22	1.5312	6.8306	0.7947	1.8091	7.9890	0.7978
Q23	1.5263	6.7069	0.7295	1.7330	7.9410	0.7614
Q24	1.6532	6.8446	0.6526	1.6850	7.3380	0.7333
L25	1.5645	6.5617	0.6909	1.7220	7.2340	0.6978
R26	1.2299	4.8828	0.0417	1.0508	1.7270	0.2060
D27	1.2575	3.9448	0.1090	1.2057	4.2197	0.2212
P28						
R29	1.4144	6.4516	0.6573	1.6870	7.0650	0.6778

Residue	R1[a]	R2[a]	Het NOE[a]	R1[b]	R2[b]	Het NOE[b]
L30	1.5097	6.4475	0.5964	1.5460	6.9090	0.6343
K31	1.5700	6.9930	0.6163	1.7030	9.5680	0.7348
K32				1.6370	6.6830	0.6223
D33	1.5550	6.5445	0.7322	1.7740	7.3150	0.7732
F34	1.6592	6.7843	0.6598	1.7040	7.2620	0.7478
V35	1.6418	6.7889	0.7852	1.8842	7.7900	0.7971
T36	1.5501	7.2046	0.7219	1.8420	8.3000	0.7793
F37	1.5610	6.1690	0.8634	1.7337	6.7233	0.6947
S38	1.6069	6.7476	0.6232	1.6480	7.1100	0.7810
R39	1.6532	6.8213	0.6931	1.7390	7.3410	0.7568
V40	1.5430	6.2035	0.7252	1.7420	7.0690	0.7445
E41	1.5065	6.8966	0.7592	1.7160	7.6850	0.7837
L42	1.6348	7.4627	0.7399	1.6920	8.6190	0.7832
S43	1.6280	7.2150	0.7514	1.9110	7.8450	0.7552
K44	1.5750	6.9784	0.6559	2.0330	8.2340	0.7690
D45	1.5586	7.5700	0.6337	1.8320	8.2930	0.7491
K46	1.5370	7.4294	0.7250	1.8720	8.1380	0.7679
R47	1.5667	7.2780	0.6708	1.7901	7.5490	0.7921
Y48	1.6889	7.4239	0.7336	1.9110	7.9090	0.8049
A89	1.6556	6.6138	0.8022	1.8070	7.5390	0.7740
D50	1.5990	6.6756	0.6989	1.6850	7.3200	0.7721
V51	1.6044	6.3654	0.7856	1.8161	7.2840	0.7954
Y52	1.6064	7.7042	0.6236	1.7310	8.4450	0.7625
V53	1.6311	6.6357	0.7416	1.8590	7.3207	0.7958
S54	1.5957	7.5700	0.6914	1.7450	7.6760	0.7930
F55	1.5758	6.5920	0.7791	1.7520	7.2700	0.7341
L56	1.5181	7.1736	0.5977	1.6340	9.3090	0.7253
G57	1.2475	5.2274	0.1665	1.2450	6.6990	0.1579
T58	1.6625	6.3291	0.6162	1.7510	7.4180	0.7690
P59						
E60	1.6667	7.0822	0.6977	1.7280	7.8380	0.8193

Residue	R1[a]	R2[a]	Het NOE[a]	R1[b]	R2[b]	Het NOE[b]
E61				1.8410	7.5510	0.7695
R62	1.5860	7.3368	0.7940	1.8203	8.7740	0.7675
K63	1.5413	5.4915	0.0761			
E64	1.6176	5.6022	0.1604			
T65	1.0038	3.5625	0.2970			
V66	1.6734	7.2046	0.7934	1.8440	8.2880	0.8170
E67				1.8160	8.5990	0.8726
I68	1.7042	7.0522	0.6859	1.7680	8.3190	0.7586
L69	1.5969	8.9526	0.7181	1.8310	8.8680	0.7951
N70	1.5274	7.4794	0.7193	1.6970	8.0730	0.8112
R71	1.1521	4.3309	0.2141	1.1200	2.5545	0.2720
A72	1.2478	4.3290	0.3740	1.7320	4.9050	0.6604
K73				1.8510	5.3320	0.6144
G74	1.5440	7.2569	0.7229	2.0635	8.9290	0.8700
F75	1.5202	6.9735	0.7648	1.7750	7.8460	0.8046
F76	1.6404	6.8446	0.8701	1.8519	7.9260	0.8006
R77	1.6404	7.6746	0.7779	1.9060	8.4520	0.8000
T78	1.6319	7.3260	0.7046	1.7580	7.7440	0.7630
F79	1.6725	8.0000	0.7082	1.8416	8.6350	0.8068
I80	1.6152	7.9365	0.6200	1.8130	8.3560	0.7954
A81	1.6215	7.1531	0.6641	1.9250	8.8370	0.7734
K82	1.6202	7.5245	0.7733	1.6660	8.0430	0.7814
N83	1.6604	6.8213	0.7381	1.7130	7.0950	0.7646
L84						
R85				1.7610	9.5090	0.6301
L86	1.3806	6.9930	0.5620	1.6590	7.3790	0.6457
Y87				1.6853	7.1710	0.5914
V88	1.5237	8.0645	0.5103	1.5482	7.0510	0.5757
A89	1.3831	6.6313	0.5744	1.5600	7.1050	0.5243
P90						
E91				1.7050	7.0790	0.7035

Residue	R1[a]	R2[a]	Het NOE[a]	R1[b]	R2[b]	Het NOE[b]
I92	1.5382	6.8729	0.6914	1.7851	7.4310	0.7794
R93	1.5170	7.6161	0.7658	1.8050	7.3510	0.7899
F94	1.5632	7.1994	0.7834	1.8080	7.7260	0.7568
Y95	1.5602	6.9396	0.7783	1.8360	7.7960	0.8154
E96	1.5726	6.5402	0.7016	1.7680	7.1400	0.7477
D97	1.6995	7.1174	0.5865	1.7251	7.2874	0.7841
K98	1.6351	4.7962	0.2254	1.6438	4.9050	
G99	1.5753	5.4765	0.4570	1.7000	6.6370	0.4364
I100				1.6300	4.8120	0.3779
E101	1.4648	2.7778	0.1517	1.5280	3.2350	0.1919
A102				1.3304	2.2140	–0.1490
S103	1.3776	2.2931	0.0100	1.4080		–0.2313
V104				1.3570	2.3660	–0.2320
K105				1.3660	2.4250	–0.4653
V106	1.4477	4.8733	2.3084e-3			
H107					3.4180	
Q108						
L109						
L110						
V111				1.1770	2.0250	–0.6855
Q112				1.1470		–0.7326
L113						
G114	1.4241	2.4060	0.0623			–0.7290
Y115				1.1230	1.1780	–0.7630
D116	1.0167	2.2563	0.1471	0.8926	1.6760	–0.8630
P117						
L118	1.0666	1.4438	–0.8200	0.8260	1.8719	–0.6960
K119						
D120	1.0738	1.4172	–0.5350	0.6704	1.0070	–0.8480

(a) Relaxation parameter at 303K. (b) Relaxation parameter at 313K.

Table A.9
Experimental restraints and relaxation parameter for *H.pylori* RbfA

Residue	^{3}J (HNHA)[a]	RDC H_N-N	R1[b]	R2[b]	Het NOE[b]
M1					
N2					
A3	5.9500		1.3730	3.5255	
H4			1.7540	4.3860	–0.5840
K5	3.6400	–1.501	1.7260	4.2635	0.2572
E6	5.3900		1.3310	5.9050	
R7	6.7900		1.4930	5.4650	0.1015
L8	6.1100	2.602	1.5780	5.6050	0.4561
E9			1.5830	7.9600	0.4971
S10	3.0400		1.5240	7.9100	0.5504
N11					0.6572
L12			1.8930		0.5683
L13		–0.043	1.6070	8.3350	
E14			1.4060	9.4150	0.8638
L15	4.4300	2.691	1.5450	7.3700	0.6412
L16		2.112	1.4810	9.0550	0.7513
Q17		–2.083	1.3310		0.7910
E18	4.8600		1.6690	8.1150	0.9024
A19	3.5400	–0.037	1.5870	7.8550	0.8765
L20		1.867	1.6960	8.9550	0.7930
A21	3.3200	–3.915	1.5640	8.5050	0.8929
S22	4.6300		1.8130	8.6350	0.8674
L23	4.4500	–1.197			0.7990
N24	6.0900	–2.958	1.9670	8.4500	0.9352
D25	6.0200	0.496	1.0810	5.3700	0.9009
S26	3.3200		1.7710	7.7500	0.9502
E27	7.0000	4.392	1.6020	8.1200	0.2759
L28	8.3300		1.5260	7.4150	0.5494
N29	3.6100	–9.795	1.5700	8.3850	0.8497
S30	8.9300		1.4690		0.8819

Residue	^{3}J (HNHA)[a]	RDC H_N-N	R1[b]	R2[b]	Het NOE[b]
L31	4.7900	1.325	1.5230	7.8900	0.9275
S32	8.2400	6.176	1.3410	8.3100	0.9192
V33	8.2300	1.326	1.5940	8.7900	0.8451
T34	9.2900	7.752	1.8060	8.1900	0.9512
K35		2.083	1.5770	7.4050	0.7823
V36	5.0500	–0.441	1.5650	7.9200	0.8268
E37	8.6000	–1.754	1.5260	7.3550	0.7726
C38			1.7520	9.0400	0.7442
S39	7.1240		1.6870	8.2150	
K40					
G41		6.207			0.9054
K42		–7.191			0.7491
H43					0.7888
H44			1.8410	8.2950	0.7922
A45			1.7150	10.0950	
Y46		–0.691			0.8489
V47			1.2059	7.5700	
F48	6.5800	–4.256	1.4050	7.6400	0.9478
V49	8.3200	–6.257	1.1040	8.8000	0.7506
L50	8.4300	–4.94	1.5440	8.6450	
S51	9.7200	3.408	1.6820	7.2750	0.8321
S52	6.3600			9.0950	0.7196
D53	4.7700	–0.784	1.5490	8.0600	0.8844
H54	4.6600		1.6950	6.4400	
K55	4.7600		1.6800	1.9160	
I56	4.4600	–6.125	1.1140	8.1000	0.7965
L57	4.8500	–8.963	1.3000	8.8150	0.8117
S58	4.2400		1.5420	9.7000	0.8186
K59	4.0900	–3.816	1.5820	8.2300	0.8134
L60	5.1000	–4.761	1.5740	7.1450	
K61		–12.88	1.5690	9.1550	0.9264
K62	2.8300	–7.55	1.7040	8.0950	0.9910

Residue	^{3}J (HNHA)[a]	RDC H_N-N	R1[b]	R2[b]	Het NOE[b]
A63	7.6200	–6.502	1.4960	8.3200	0.9072
E64		–5.861	1.6870	8.6000	
G65	5.2700		1.7160	5.0725	0.8744
L66	4.8100	–10.222	1.6430	8.2750	0.8333
I67		–3.532	1.6240	6.9500	0.9065
R68	3.7900	1.359	1.5940	8.3150	0.8315
Q69		–8.749	1.4910	9.7200	0.8600
F70	4.1000	–11.211	1.5360	8.0550	0.7831
V71			1.7310	8.9300	
L72					0.6637
Q73		–8.233	1.4810		0.6719
A74		–7.063		6.8490	0.7906
S75	6.0600		1.7060	9.6550	0.8690
G76		2.045			0.8904
W77		10.25	1.4490	10.1400	0.9330
F78					0.8732
K79					0.7293
C80		1.795		7.5800	0.6325
P81					
K82		0.956	1.4580	10.0050	0.9138
L83	7.6700	1.589	1.6960	2.3520	0.9755
S84	10.2000	–1.292	1.6220	8.7400	0.9511
F85	8.5100	4.28	1.6030	6.9200	1.0773
V86	9.9100	0.211	1.5490	8.7500	0.8270
S87	4.0700		1.7880	7.4950	
D88	5.2300	2.925	1.7200	7.8450	0.4916
N89	6.5500			7.3950	0.4524
S90	4.4500	–11.802		7.4850	
L91	6.1770	–3.715	1.6920	6.4450	0.6256
E92	5.0900	–5.617	1.7420	5.8850	–0.9337
K93					0.6722
Q94	4.7300	–2.656		4.4285	0.8858

Residue	^{3}J (HNHA)[a]	RDC H_N-N	R1[b]	R2[b]	Het NOE[b]
L95	4.7900	–1.451	1.8010	5.1750	0.8777
R96					0.9207
L97	5.2900	–1.363	1.8480	4.6745	0.8857
D98	5.6400	–2.182	1.6980	4.5550	0.6773
A99	5.4300	–1.23	1.1700	4.4215	0.7490
I100	6.3300		1.0980	4.2010	0.6763
F101	4.8800	–1.167	1.2180	4.7250	0.5170
N102	6.1400	–0.591	1.6030	4.4285	
E103	5.7900	–0.545	1.2500	3.8905	0.4400
I104	6.2200	0.034	1.2020	3.2155	0.4354
A105	5.1800	0.403	1.1810	4.4685	
K106	5.8800	1.472	1.1510	3.0915	0.3695
G107	6.6300	–1.965	1.2470	2.3600	0.2601
K108	5.9700	0.75	1.3350	1.9465	0.3479
D109	6.5100		1.1560	1.6665	0.2921
N110	7.5600	0.297	1.0490	1.5050	0.2886
D111	7.5500	0.046	0.4990	1.3520	0.3051

(a) Experimental restraints in Hz at 293K. (b) Relaxation parameters at 293K.

References

1. Agalarov, S. C., Sridhar Prasad, G., Funke, P. M., Stout, C. D. and Williamson, J. R. (2000). Structure of the S15,S6,S18-rRNA complex: assembly of the 30S ribosome central domain. *Science* **288,** 107–13.

2. Agrawal, R. K., Penczek, P., Grassucci, R. A., Li, Y., Leith, A., Nierhaus, K. H. and Frank, J. (1996). Direct visualization of A-, P-, and E-site transfer RNAs in the Escherichia coli ribosome. *Science* **271,** 1000–2.

3. Alix, J. H. and Guerin, M. F. (1993). Mutant DnaK chaperones cause ribosome assembly defects in Escherichia coli. *Proc Natl Acad Sci USA* **90,** 9725–9.

4. Altieri, A. S. and Byrd, R. A. (1995). Randomization approach to water suppression in multidimensional NMR using pulsed field gradients. *J Magn Reson B* **107,** 260–6.

5. Anantharaman, V. and Aravind, L. (2002). The PRC-barrel: a widespread, conserved domain shared by photosynthetic reaction center subunits and proteins of RNA metabolism. *Genome Biol* **3,** RESEARCH0061.

6. Ashley, C. T., Jr., Wilkinson, K. D., Reines, D. and Warren, S. T. (1993). FMR1 protein: conserved RNP family domains and selective RNA binding. *Science* **262,** 563–6.

7. Backe, P. H., Messias, A. C., Ravelli, R. B., Sattler, M. and Cusack, S. (2005). X-ray crystallographic and NMR studies of the third KH domain of hnRNP K in complex with single-stranded nucleic acids. *Structure* **13,** 1055–67.

8. Balzer, M. and Wagner, R. (1998). Mutations in the leader region of ribosomal RNA operons cause structurally defective 30 S ribosomes as revealed by in vivo structural probing. *J Mol Biol* **276,** 547–57.

9. Bartels, C., Xia, T. H., Billeter, M., Guntert, P. and Wuthrich, K. (1995). The Program Xeasy for Computer-Supported NMR Spectral-Analysis of Biological Macromolecules. *Journal of Biomolecular NMR* **6,** 1–10.

10. Belanger, F., Theberge-Julien, G., Cunningham, P. R. and Brakier-Gingras, L. (2005). A functional relationship between helix 1 and the 900 tetraloop of 16S ribosomal RNA within the bacterial ribosome. *RNA* **11,** 906–13.

11. Besancon, W. and Wagner, R. (1999). Characterization of transient RNA-RNA interactions important for the facilitated structure formation of bacterial ribosomal 16S RNA. *Nucleic Acids Res* **27,** 4353–62.

12. Bodenhausen, G. and Ruben, D. J. (1980). Natural Abundance N-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chemical Physics Letters* **69,** 185–189.

13. Bollen, A., Helser, T., Yamada, T. and Davies, J. (1969). Altered ribosomes in antibiotic-resistant mutants of E. coli. *Cold Spring Harb Symp Quant Biol* **34,** 95–100.

14. Bollen, A. and Herzog, A. (1970). The ribosomal protein altered in spectinomycin resistant Escherichia coli. *FEBS Lett* **6,** 69–72.

15. Boni, I. V., Isaeva, D. M., Musychenko, M. L. and Tzareva, N. V. (1991). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res* **19,** 155–62.

16. Braddock, D. T., Baber, J. L., Levens, D. and Clore, G. M. (2002). Molecular basis of sequence-specific single-stranded DNA recognition by KH domains: solution structure of a complex between hnRNP K KH3 and single-stranded DNA. *Embo J* **21,** 3476–85.

17. Braddock, D. T., Louis, J. M., Baber, J. L., Levens, D. and Clore, G. M. (2002). Structure and dynamics of KH domains from FBP bound to single-stranded DNA. *Nature* **415,** 1051–6.

18. Brandi, A., Spurio, R., Gualerzi, C. O. and Pon, C. L. (1999). Massive presence of the Escherichia coli ‚major cold-shock protein' CspA under non-stress conditions. *Embo J* **18,** 1653–9.

19. Brimacombe, R. (1988). The emerging three-dimensional structure and function of 16S ribosomal RNA. *Biochemistry* **27,** 4207–14.

20. Brimacombe, R. (1995). The structure of ribosomal RNA: a three-dimensional jigsaw puzzle. *Eur J Biochem* **230,** 365–83.

21. Brimacombe, R., Atmadja, J., Stiege, W. and Schuler, D. (1988). A detailed model of the three-dimensional structure of Escherichia coli 16 S ribosomal RNA in situ in the 30 S subunit. *J Mol Biol* **199,** 115–36.

22. Brink, M. F., Verbeet, M. P. and de Boer, H. A. (1993). Formation of the central pseudoknot in 16S rRNA is essential for initiation of translation. *Embo J* **12,** 3987–96.

23. Britton, R. A., Powell, B. S., Court, D. L. and Lupski, J. R. (1997). Characterization of mutations affecting the Escherichia coli essential GTPase era that suppress two temperature-sensitive dnaG alleles. *J Bacteriol* **179,** 4575–82.

24. Britton, R. A., Powell, B. S., Dasgupta, S., Sun, Q., Margolin, W., Lupski, J. R. and Court, D. L. (1998). Cell cycle arrest in Era GTPase mutants: a potential growth rate-regulated checkpoint in Escherichia coli. *Mol Microbiol* **27,** 739–50.

25. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54,** 905–21.

26. Bycroft, M., Grunert, S., Murzin, A. G., Proctor, M. and St Johnston, D. (1995). NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *Embo J* **14,** 3563–71.

27. Bylund, G. O., Lovgren, J. M. and Wikstrom, P. M. (2001). Characterization of mutations in the metY-nusA-infB operon that suppress the slow growth of a DeltarimM mutant. *J Bacteriol* **183,** 6095–106.

28. Bylund, G. O., Persson, B. C., Lundberg, L. A. and Wikstrom, P. M. (1997). A novel ribosome-associated protein is important for efficient translation in Escherichia coli. *J Bacteriol* **179,** 4567–74.

29. Bylund, G. O., Wipemo, L. C., Lundberg, L. A. and Wikstrom, P. M. (1998). RimM and RbfA are essential for efficient processing of 16S rRNA in Escherichia coli. *J Bacteriol* **180,** 73–82.

30. Bystrom, A. S., Hjalmarsson, K. J., Wikstrom, P. M. and Bjork, G. R. (1983). The nucleotide sequence of an Escherichia coli operon containing genes for the tRNA(m1G)methyltransferase, the ribosomal proteins S16 and L19 and a 21-K polypeptide. *Embo J* **2,** 899–905.

31. Capel, M. S., Engelman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Sillers, I. Y. and et al. (1987). A complete mapping of the proteins in the small ribosomal subunit of Escherichia coli. *Science* **238,** 1403–6.

32. Capel, M. S., Kjeldgaard, M., Engelman, D. M. and Moore, P. B. (1988). Positions of S2, S13, S16, S17, S19 and S21 in the 30 S ribosomal subunit of Escherichia coli. *J Mol Biol* **200,** 65–87.

33. Carter, A. P., Clemons, W. M., Jr., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T. and Ramakrishnan, V. (2001). Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* **291,** 498–501.

34. Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T. and Ramakrishnan, V. (2000). Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407,** 340–8.

35. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. and Noller, H. F. (1999). X-ray crystal structures of 70S ribosome functional complexes. *Science* **285,** 2095–104.

36. Cech, T. R. (1986). RNA as an enzyme. *Sci Am* **255,** 64–75.

37. Chakravarty, S. and Varadarajan, R. (2002). Elucidation of factors responsible for enhanced thermal stability of proteins: A structural genomics based study. *Biochemistry* **41,** 8152–8161.

38. Charollais, J., Pflieger, D., Vinh, J., Dreyfus, M. and Iost, I. (2003). The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli. *Mol Microbiol* **48,** 1253–65.

39. Chen, X., Court, D. L. and Ji, X. (1999). Crystal structure of ERA: a GTPase-dependent cell cycle regulator containing an RNA binding motif. *Proc Natl Acad Sci USA* **96,** 8396–401.

40. Clubb, R. T., Thanabal, V. and Wagner, G. (1992). A New 3d Hn(Ca)Ha Experiment for Obtaining Fingerprint H(N)-H-Alpha Cross Peaks in N-15-Labeled and C-13-Labeled Proteins. *Journal of Biomolecular NMR* **2,** 203–210.

41. Culver, G. M. (2003). Assembly of the 30S ribosomal subunit. *Biopolymers* **68,** 234–49.

42. Culver, G. M. and Noller, H. F. (1999). Efficient reconstitution of functional Escherichia coli 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins. *RNA* **5,** 832–43.

43. Dahlberg, A. E., Dahlberg, J. E., Lund, E., Tokimatsu, H., Rabson, A. B., Calvert, P. C., Reynolds, F. and Zahalak, M. (1978). Processing of the 5‘ end of Escherichia coli 16S ribosomal RNA. *Proc Natl Acad Sci USA* **75,** 3598–602.

44. Daigle, D. M. and Brown, E. D. (2004). Studies of the interaction of Escherichia coli YjeQ with the ribosome in vitro. *J Bacteriol* **186,** 1381–7.

45. Daigle, D. M., Rossi, L., Berghuis, A. M., Aravind, L., Koonin, E. V. and Brown, E. D. (2002). YjeQ, an essential, conserved, uncharacterized protein from Escherichia coli, is an unusual GTPase with circularly permuted G-motifs and marked burst kinetics. *Biochemistry* **41,** 11109–17.

46. Dammel, C. S. and Noller, H. F. (1993). A cold-sensitive mutation in 16S rRNA provides evidence for helical switching in ribosome assembly. *Genes Dev* **7,** 660–70.

47. Dammel, C. S. and Noller, H. F. (1995). Suppression of a cold-sensitive mutation in 16S rRNA by overexpression of a novel ribosome-binding factor, RbfA. *Genes Dev* **9,** 626–37.

48. Davis, A. L., Keeler, J., Laue, E. D. and Moskau, D. (1992). Experiments for Recording Pure-Absorption Heteronuclear Correlation Spectra Using Pulsed Field Gradients. *Journal of Magnetic Resonance* **98,** 207–216.

49. Dekio, S. and Takata, R. (1969). Genetic studies of the ribosomal proteins in Escherichia. coli. II. Altered 30s ribosomal protein component specific to spectinomycin resistant mutants. *Mol Gen Genet* **105,** 219–24.

50. Del Campo, M., Kaya, Y. and Ofengand, J. (2001). Identification and site of action of the remaining four putative pseudouridine synthases in Escherichia coli. *RNA* **7,** 1603–15.

51. DeLano, W. L. (2002). The PyMol Molecular Graphics System. DeLano Scientific, San Carlos CA.

52. Dosset, P., Hus, J. C., Marion, D. and Blackledge, M. (2001). A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J Biomol NMR* **20,** 223–31.

53. El Hage, A. and Alix, J. H. (2004). Authentic precursors to ribosomal subunits accumulate in Escherichia coli in the absence of functional DnaK chaperone. *Mol Microbiol* **51,** 189–201.

54. El Hage, A., Sbai, M. and Alix, J. H. (2001). The chaperonin GroEL and other heat-shock proteins, besides DnaK, participate in ribosome biogenesis in Escherichia coli. *Mol Gen Genet* **264,** 796–808.

55. Fox, G. E. and Woese, C. R. (1975). The architecture of 5S rRNA and its relation to function. *J Mol Evol* **6,** 61–76.

56. Frank, J. and Agrawal, R. K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* **406,** 318–22.

57. Fukushima, H., Martin, C. E., Iida, H., Kitajima, Y. and Thompson, G. A., Jr. (1976). Changes in membrane lipid composition during temperature adaptation by a thermotolerant strain of Tetrahymena pyriformis. *Biochim Biophys Acta* **431**, 165–79.

58. Fulco, A. J. (1974). Metabolic alterations of fatty acids. *Annu Rev Biochem* **43,** 215–41.

59. Garcia de la Torre, J., Huertas, M. L. and Carrasco, B. (2000). HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic claculations. *J Magn Reson B* **147,** 138–146.

60. Gayathri, C., Bothnerby, A. A., Vanzijl, P. C. M. and Maclean, C. (1982). Dipolar Magnetic-Field Effects in NMR-Spectra of Liquids. *Chemical Physics Letters* **87,** 192–196.

61. Gibson, T. J., Thompson, J. D. and Heringa, J. (1993). The KH domain occurs in a diverse set of RNA-binding proteins that include the antiterminator NusA and is probably involved in binding to nucleic acid. *FEBS Lett* **324,** 361–6.

62. Ginsburg, D. and Steitz, J. A. (1975). The 30 S ribosomal precursor RNA from Escherichia coli. A primary transcript containing 23 S, 16 S, and 5 S sequences. *J Biol Chem* **250,** 5647–54.

63. Giuliodori, A. M., Brandi, A., Gualerzi, C. O. and Pon, C. L. (2004). Preferential translation of cold-shock mRNAs during cold adaptation. *RNA* **10,** 265–76.

64. Goddard, T. D. and G., K. D. Sparky 3. University of California, San Francisco.

65. Gollop, N. and March, P. E. (1991). A GTP-binding protein (Era) has an essential role in growth rate and cell cycle control in Escherichia coli. *J Bacteriol* **173,** 2265–70.

66. Gopal, B., Haire, L. F., Gamblin, S. J., Dodson, E. J., Lane, A. N., Papavinasasundaram, K. G., Colston, M. J. and Dodson, G. (2001). Crystal structure of the transcription elongation/anti-termination factor NusA from Mycobacterium tuberculosis at 1.7 A resolution. *J Mol Biol* **314,** 1087–95.

67. Grimm, S. K. and Wohnert, J. (2005). NMR assignments of the cold-shock protein ribosome-binding factor A (RbfA) from Thermotoga maritima. *J Biomol NMR* **31,** 73–4.

68. Grishin, N. V. (2001). KH domain: one motif, two folds. *Nucleic Acids Res* **29,** 638–43.

69. Grondek, J. F. and Culver, G. M. (2004). Assembly of the 30S ribosomal subunit: positioning ribosomal protein S13 in the S7 assembly branch. *RNA* **10,** 1861–6.

70. Grzesiek, S., Anglister, J. and Bax, A. (1993). Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in C-13/N-15-Enriched Proteins by Isotropic Mixing of C-13 Magnetization. *Journal of Magnetic Resonance Series B* **101,** 114–119.

71. Grzesiek, S. and Bax, A. (1992). Correlating Backbone Amide and Side-Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. *Journal of the American Chemical Society* **114,** 6291–6293.

72. Grzesiek, S. and Bax, A. (1992). Improved 3d Triple-Resonance NMR Techniques Applied to a 31-Kda Protein. *Journal of Magnetic Resonance* **96,** 432–440.

73. Gualerzi, C. O., Giuliodori, A. M. and Pon, C. L. (2003). Transcriptional and post-transcriptional control of cold-shock genes. *J Mol Biol* **331,** 527–39.

74. Gualerzi, C. O. and Pon, C. L. (1990). Initiation of mRNA translation in prokaryotes. *Biochemistry* **29,** 5881–9.

75. Guntert, P., Mumenthaler, C. and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *Journal of Molecular Biology* **273,** 283–298.

76. Guthrie, C., Nashimoto, H. and Nomura, M. (1969). Studies on the assembly of ribosomes in vivo. *Cold Spring Harb Symp Quant Biol* **34,** 69–75.

77. Hang, J. Q. and Zhao, G. (2003). Characterization of the 16S rRNA- and membrane-binding domains of Streptococcus pneumoniae Era GTPase: structural and functional implications. *Eur J Biochem* **270,** 4164–72.

78. Hartz, D., Binkley, J., Hollingsworth, T. and Gold, L. (1990). Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3. *Genes Dev* **4,** 1790–800.

79. Hayes, F., Hayes, D., Fellner, P. and Ehresmann, C. (1971). Additional nucleotide sequences in precursor 16S ribosomal RNA from Escherichia coli. *Nat New Biol* **232,** 54–5.

80. Heilek, G. M. and Noller, H. F. (1996). Site-directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S5. *Science* **272,** 1659–62.

81. Held, W. A., Ballou, B., Mizushima, S. and Nomura, M. (1974). Assembly mapping of 30 S ribosomal proteins from Escherichia coli. Further studies. *J Biol Chem* **249,** 3103–11.

82. Held, W. A., Gette, W. R. and Nomura, M. (1974). Role of 16S ribosomal ribonucleic acid and the 30S ribosomal protein S12 in the initiation of natural messenger ribonucleic acid translation. *Biochemistry* **13,** 2115–22.

83. Held, W. A., Mizushima, S. and Nomura, M. (1973). Reconstitution of Escherichia coli 30 S ribosomal subunits from purified molecular components. *J Biol Chem* **248,** 5720–30.

84. Helser, T. L., Davies, J. E. and Dahlberg, J. E. (1971). Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in Escherichia coli. *Nat New Biol* **233,** 12–4.

85. Henikoff, S. and Henikoff, J. G. (1992). Amino-Acid Substitution Matrices from Protein Blocks. *Proceedings of the National Academy of Sciences of the United States of America* **89,** 10915–10919.

86. Herrmann, T., Guntert, P. and Wuthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of Molecular Biology* **319,** 209–227.

87. Hill, W. E., Tassanakajohn, A. and Tapprich, W. E. (1990). Interaction of tRNA with domain II of 23S rRNA. *Biochim Biophys Acta* 1050, 45–50.

88. Himeno, H., Hanawa-Suetsugu, K., Kimura, T., Takagi, K., Sugiyama, W., Shirata, S., Mikami, T., Odagiri, F., Osanai, Y., Watanabe, D., Goto, S., Kalachnyuk, L., Ushida, C. and Muto, A. (2004). A novel GTPase activated by the small subunit of ribosome. *Nucleic Acids Res* **32,** 5303–9.

89. Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F. and Hartl, F. U. (1999). Identification of in vivo substrates of the chaperonin GroEL. *Nature* **402,** 147–54.

90. Huang, Y. J., Swapna, G. V., Rajan, P. K., Ke, H., Xia, B., Shukla, K., Inouye, M. and Montelione, G. T. (2003). Solution NMR structure of ribosome-binding factor A (RbfA), a cold-shock adaptation protein from Escherichia coli. *J Mol Biol* **327,** 521–36.

91. Inoue, K., Alsina, J., Chen, J. and Inouye, M. (2003). Suppression of defective ribosome assembly in a rbfA deletion mutant by overexpression of Era, an essential GTPase in Escherichia coli. *Mol Microbiol* **48,** 1005–16.

92. Iost, I. and Dreyfus, M. (1994). mRNAs can be stabilized by DEAD-box proteins. *Nature* **372,** 193–6.

93. Ishii, Y., Markus, M. A. and Tycko, R. (2001). Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *Journal of Biomolecular NMR* **21,** 141–151.

94. Jiang, W., Fang, L. and Inouye, M. (1996). The role of the 5'-end untranslated region of the mRNA for CspA, the major cold-shock protein of Escherichia coli, in cold-shock adaptation. *J Bacteriol* **178,** 4919–25.

95. Jiang, W., Hou, Y. and Inouye, M. (1997). CspA, the major cold-shock protein of Escherichia coli, is an RNA chaperone. *J Biol Chem* **272,** 196–202.

96. Jones, P. G. and Inouye, M. (1996). RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol Microbiol* **21,** 1207–18.

97. Jones, P. G., Mitta, M., Kim, Y., Jiang, W. and Inouye, M. (1996). Cold shock induces a major ribosomal-associated protein that unwinds double-stranded RNA in Escherichia coli. *Proc Natl Acad Sci USA* **93,** 76–80.

98. Kaczanowska, M. and Ryden-Aulin, M. (2005). The YrdC protein—a putative ribosome maturation factor. *Biochim Biophys Acta* **1727,** 87–96.

99. Karshikoff, A. and Ladenstein, R. (1998). Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Protein Eng* **11,** 867–72.

100. Kasai, R., Kitajima, Y., Martin, C. E., Nozawa, Y., Skriver, L. and Thompson, G. A., Jr. (1976). Molecular control of membrane properties during temperature acclimation. Membrane fluidity regulation of fatty acid desaturase action? *Biochemistry* **15,** 5228–33.

101. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. and Formankay, J. D. (1993). A Gradient-Enhanced Hcch Tocsy Experiment for Recording Side-Chain H-1 and C-13 Correlations in H_2O Samples of Proteins. *Journal of Magnetic Resonance Series B* **101,** 333–337.

102. Kay, L. E., Xu, G. Y. and Yamazaki, T. (1994). Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H_2O Saturation. *Journal of Magnetic Resonance Series A* **109,** 129–133.

103. Kim, D. F. and Green, R. (1999). Base-pairing between 23S rRNA and tRNA in the ribosomal A site. *Mol Cell* **4,** 859–64.

104. King, T. C. and Schlessinger, D. (1983). S1 nuclease mapping analysis of ribosomal RNA processing in wild type and processing deficient Escherichia coli. *J Biol Chem* **258,** 12034–42.

105. Kisselev, L. L. and Buckingham, R. H. (2000). Translational termination comes of age. *Trends Biochem Sci* **25,** 561–6.

106. Klein, B. K., Staden, A. and Schlessinger, D. (1985). Alternative conformations in Escherichia coli 16S ribosomal RNA. *Proc Natl Acad Sci USA* **82,** 3539–42.

107. Koradi, R., Billeter, M. and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graphics* **14,** 51–55.

108. Kreider, G. and Brownstein, B. L. (1972). Ribosomal proteins involved in the suppression of streptomycin dependence in Escherichia coli. *J Bacteriol* **109,** 780–5.

109. Krzyzosiak, W., Denman, R., Nurse, K., Hellmann, W., Boublik, M., Gehrke, C. W., Agris, P. F. and Ofengand, J. (1987). In vitro synthesis of 16S ribosomal RNA containing single base changes and assembly into a functional 30S ribosome. *Biochemistry* **26,** 2353–64.

110. Kuboniwa, H., Grzesiek, S., Delaglio, F. and Bax, A. (1994). Measurement of H-N-H-Alpha J-Couplings in Calcium-Free Calmodulin Using New 2d and 3d Water-Flip-Back Methods. *Journal of Biomolecular NMR* **4,** 871–878.

111. Kumar, S., Ma, B., Tsai, C. J. and Nussinov, R. (2000). Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins* **38,** 368–83.

112. Langer, J. A., Engelman, D. M. and Moore, P. B. (1978). Neutron-scattering studies of the ribosome of Escherichia coli: a provisional map of the locations of proteins S3, S4, S5, S7, S8 and S9 in the 30 S subunit. *J Mol Biol* **119,** 463–85.

113. Laskowski, R. A., Moss, D. S. and Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231,** 1049–67.

114. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8,** 477–86.

115. Leijonmarck, M., Eriksson, S. and Liljas, A. (1980). Crystal structure of a ribosomal component at 2.6 A resolution. *Nature* **286,** 824–6.

116. Leonard, K. R. and Lake, J. A. (1979). Ribosome structure: hand determination by electron microscopy of 30 S subunits. *J Mol Biol* **129,** 155–63.

117. Lerner, C. G., Gulati, P. S. and Inouye, M. (1995). Cold-sensitive conditional mutations in Era, an essential Escherichia coli GTPase, isolated by localized random polymerase chain reaction mutagenesis. *FEMS Microbiol Lett* **126,** 291–8.

118. Lewis, H. A., Musunuru, K., Jensen, K. B., Edo, C., Chen, H., Darnell, R. B. and Burley, S. K. (2000). Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* **100,** 323–32.

119. Li, Z., Pandit, S. and Deutscher, M. P. (1999). RNase G (CafA protein) and RNase E are both required for the 5‘ maturation of 16S ribosomal RNA. *Embo J* **18,** 2878–85.

120. Li, Z., Pandit, S. and Deutscher, M. P. (1999). RNase G (CafA protein) and RNase E are both required for the 5’ maturation of 16S ribosomal RNA. *Embo J* **18,** 2878–85.

121. Lin, Y. P., Sharer, J. D. and March, P. E. (1994). GTPase-dependent signaling in bacteria: characterization of a membrane-binding site for era in Escherichia coli. *J Bacteriol* **176,** 44–9.

122. Lindahl, L. (1975). Intermediates and time kinetics of the in vivo assembly of Escherichia coli ribosomes. *J Mol Biol* **92,** 15–37.

123. Linge, J. P. and Nilges, M. (1999). Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. *J Biomol NMR* **13,** 51–9.

124. Linge, J. P., O’Donoghue, S. I. and Nilges, M. (2001). Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol* **339,** 71–90.

125. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M. and Nilges, M. (2003). Refinement of protein structures in explicit solvent. *Proteins* **50,** 496–506.

126. Liu, Z., Luyten, I., Bottomley, M. J., Messias, A. C., Houngninou-Molango, S., Sprangers, R., Zanier, K., Kramer, A. and Sattler, M. (2001). Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science* **294,** 1098–102.

127. Lopez, M. M., Yutani, K. and Makhatadze, G. I. (2001). Interactions of the cold shock protein CspB from Bacillus subtilis with single-stranded DNA. Importance of the T base content and position within the template. *J Biol Chem* **276,** 15511–8.

128. Lu, Q. and Inouye, M. (1998). The gene for 16S rRNA methyltransferase (ksgA) functions as a multicopy suppressor for a cold-sensitive mutant of era, an essential RAS-like GTP-binding protein in Escherichia coli. *J Bacteriol* **180,** 5243–6.

129. Maki, T., Hasegawa, H., Watarai, H. and Ueda, K. (2004). Classification for dimethylarsenate-decomposing bacteria using a restrict fragment length polymorphism analysis of 16S rRNA genes. *Anal Sci* **20,** 61–8.

130. Mangiarotti, G., Apirion, D., Schlessinger, D. and Silengo, L. (1968). Biosynthetic precursors of 30S and 50S ribosomal particles in Escherichia coli. *Biochemistry* **7,** 456–72.

131. Mangiarotti, G., Turco, E., Ponzetto, A. and Altruda, F. (1974). Precursor 16S RNA in active 30S ribosomes. *Nature* **247,** 147–8.

132. March, P. E., Lerner, C. G., Ahnn, J., Cui, X. and Inouye, M. (1988). The Escherichia coli Ras-like protein (Era) has GTPase activity and is essential for cell growth. *Oncogene* **2,** 539–44.

133. Marsh, R. C. and Parmeggiani, A. (1973). Requirement of proteins S5 and S9 from 30S subunits for the ribosome-dependent GTPase activity of elongation factor G. *Proc Natl Acad Sci USA* **70,** 151–5.

134. Martin, C. E., Hiramitsu, K., Kitajima, Y., Nozawa, Y., Skriver, L. and Thompson, G. A. (1976). Molecular control of membrane properties during temperature acclimation. Fatty acid desaturase regulation of membrane fluidity in acclimating Tetrahymena cells. *Biochemistry* **15,** 5218–27.

135. Mathews, D. H., Sabina, J., Zuker, M. and Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology* **288,** 911–940.

136. Meier, T. I., Peery, R. B., Jaskunas, S. R. and Zhao, G. (1999). 16S rRNA is bound to era of Streptococcus pneumoniae. *J Bacteriol* **181,** 5242–9.

137. Meier, T. I., Peery, R. B., McAllister, K. A. and Zhao, G. (2000). Era GTPase of Escherichia coli: binding to 16S rRNA and modulation of GTPase activity by RNA and carbohydrates. *Microbiology* **146** (Pt 5), 1071–83.

138. Mizushima, S. and Nomura, M. (1970). Assembly mapping of 30S ribosomal proteins from E. coli. *Nature* **226,** 1214.

139. Moazed, D. and Noller, H. F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. *Nature* **342,** 142–8.

140. Moore, P. B., Engelman, D. M. and Schoenborn, B. P. (1975). A neutron scattering study of the distribution of protein and RNA in the 30 S ribosomal subunit of Escherichia coli. *J Mol Biol* **91,** 101–20.

141. Mori, H., Dammel, C., Becker, E., Triman, K. and Noller, H. F. (1990). Single base alterations upstream of the E. coli 16S rRNA coding region result in temperature-sensitive 16S rRNA expression. *Biochim Biophys Acta* **1050,** 323–7.

142. Mori, S., Abeygunawardana, C., Johnson, M. O. and Vanzijl, P. C. M. (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,** 94–98.

143. Muhandiram, D. R. and Kay, L. E. (1994). Gradient-Enhanced Triple-Resonance 3-Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance Series B* **103,** 203–216.

144. Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *Embo J* **12,** 861–7.

145. Muth, G. W., Ortoleva-Donnelly, L. and Strobel, S. A. (2000). A single adenosine with a neutral pKa in the ribosomal peptidyl transferase center. *Science* **289,** 947–50.

146. Nashimoto, H., Held, W., Kaltschmidt, E. and Nomura, M. (1971). Structure and function of bacterial ribosomes. XII. Accumulation of 21 s particles by some cold-sensitive mutants of Escherichia coli. *J Mol Biol* **62,** 121–38.

147. Nashimoto, H., Miura, A., Saito, H. and Uchida, H. (1985). Suppressors of temperature-sensitive mutations in a ribosomal protein gene, rpsL (S12), of Escherichia coli K12. *Mol Gen Genet* **199,** 381–7.

148. Nashimoto, M. (1993). 3' truncated tRNAArg is essential for in vitro specific cleavage of partially synthesized mouse 18S rRNA. *Nucleic Acids Res* **21,** 4696–702.

149. Nierhaus, K. H. (1991). The assembly of prokaryotic ribosomes. *Biochimie* **73,** 739–55.

150. Nierhaus, K. H., Bordasch, K. and Homann, H. E. (1973). Ribosomal proteins. 43. In vivo assembly of Escherichia coli ribosomal proteins. *J Mol Biol* **74,** 587–97.

151. Nikolaev, N., Silengo, L. and Schlessinger, D. (1973). A role for ribonuclease 3 in processing of ribosomal ribonucleic acid and messenger ribonucleic acid precursors in Escherichia coli. *J Biol Chem* **248,** 7967–9.

152. Nikulin, A., Serganov, A., Ennifar, E., Tishchenko, S., Nevskaya, N., Shepard, W., Portier, C., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S. and Dumas, P. (2000). Crystal structure of the S15-rRNA complex. *Nat Struct Biol* **7,** 273–7.

153. Nilges, M. (1997). Ambiguous distance data in the calculation of NMR structures. *Fold Des* **2,** 53–7.

154. Nissen, P., Hansen, J., Ban, N., Moore, P. B. and Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* **289,** 920–30.

155. Noller, H. F. (1991). Ribosomal RNA and translation. *Annu Rev Biochem* **60,** 191–227.

156. Noller, H. F., Hoffarth, V. and Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256,** 1416–9.

157. Noller, H. F., Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Lancaster, L., Dallas, A., Fredrick, K., Earnest, T. N. and Cate, J. H. (2001). Structure of the ribosome at 5.5 A resolution and its interactions with functional ligands. *Cold Spring Harb Symp Quant Biol* **66,** 57–66.

158. Nomura, M. (1973). Assembly of bacterial ribosomes. *Science* **179,** 864–73.

159. Oakes, M. I., Clark, M. W., Henderson, E. and Lake, J. A. (1986). DNA hybridization electron microscopy: ribosomal RNA nucleotides 1392-1407 are exposed in the cleft of the small subunit. *Proc Natl Acad Sci USA* **83,** 275–9.

160. Ottiger, M., Delaglio, F., Marquardt, J. L., Tjandra, N. and Bax, A. (1998). Measurement of dipolar couplings for methylene and methyl sites in weakly oriented macromolecules and their use in structure determination. *Journal of Magnetic Resonance* **134,** 365–369.

161. Palmer, A. G., Cavanagh, J., Wright, P. E. and Rance, M. (1991). Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation NMR-Spectroscopy. *Journal of Magnetic Resonance* **93,** 151–170.

162. Perl, D. and Schmid, F. X. (2001). Electrostatic stabilization of a thermophilic cold shock protein. *J Mol Biol* **313,** 343–57.

163. Persson, B. C., Bylund, G. O., Berg, D. E. and Wikstrom, P. M. (1995). Functional analysis of the ffh-trmD region of the Escherichia coli chromosome by using reverse genetics. *J Bacteriol* **177,** 5554–60.

164. Piepersberg, W., Bock, A. and Wittmann, H. G. (1975). Effect of different mutations in ribosomal protein S5 of Escherichia coli on translational fidelity. *Mol Gen Genet* **140,** 91–100.

165. Piepersberg, W., Bock, A., Yaguchi, M. and Wittmann, H. G. (1975). Genetic position and amino acid replacements of several mutations in ribosomal protein S5 from Escherichia coli. *Mol Gen Genet* **143,** 43–52.

166. Piotto, M., Saudek, V. and Sklenar, V. (1992). Gradient-Tailored Excitation for Single-Quantum Nmr-Spectroscopy of Aqueous-Solutions. *Journal of Biomolecular NMR* **2,** 661–665.

167. Plaxco, K. W., Simons, K. T. and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* **277,** 985–94.

168. Pleij, C. W., Rietveld, K. and Bosch, L. (1985). A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res* **13,** 1717–31.

169. Poot, R. A., van den Worm, S. H., Pleij, C. W. and van Duin, J. (1998). Base complementarity in helix 2 of the central pseudoknot in 16S rRNA is essential for ribosome functioning. *Nucleic Acids Res* **26,** 549–53.

170. Powell, B. S., Court, D. L., Inada, T., Nakamura, Y., Michotey, V., Cui, X., Reizer, A., Saier, M. H., Jr. and Reizer, J. (1995). Novel proteins of the phosphotransferase system encoded within the rpoN operon of Escherichia coli. Enzyme IIANtr affects growth on organic nitrogen and the conditional lethality of an erats mutant. *J Biol Chem* **270,** 4822–39.

171. Powers, T., Daubresse, G. and Noller, H. F. (1993). Dynamics of in vitro assembly of 16 S rRNA into 30 S ribosomal subunits. *J Mol Biol* **232,** 362–74.

172. Ramakrishnan, V. and White, S. W. (1998). Ribosomal protein structures: insights into the architecture, machinery and evolution of the ribosome. *Trends Biochem Sci* **23,** 208–12.

173. Robinson-Rechavi, M., Alibes, A. and Godzik, A. (2006). Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: lessons from structural genomics of Thermotoga maritima. *J Mol Biol* **356,** 547–57.

174. Robinson-Rechavi, M. and Godzik, A. (2005). Structural genomics of thermotoga maritima proteins shows that contact order is a major determinant of protein thermostability. *Structure* **13,** 857–60.

175. Rodnina, M. V., Savelsbergh, A., Katunin, V. I. and Wintermeyer, W. (1997). Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* **385,** 37–41.

176. Rohl, R. and Nierhaus, K. H. (1982). Assembly map of the large subunit (50S) of Escherichia coli ribosomes. *Proc Natl Acad Sci USA* **79,** 729–33.

177. Rosset, R. and Gorini, L. (1969). A ribosomal ambiguity mutation. *J Mol Biol* **39,** 95–112.

178. Rubin, S. M., Pelton, J. G., Yokota, H., Kim, R. and Wemmer, D. E. (2003). Solution structure of a putative ribosome binding protein from Mycoplasma pneumoniae and comparison to a distant homolog. *J Struct Funct Genomics* **4,** 235–43.

179. Sambrook, J. and Pollack, R. (1974). Basic methodology for cell culture—cell transformation. *Methods Enzymol* **32,** 583–92.

180. Saupe, A. and Englert, G. (1963). High-Resolution Nuclear Magnetic Resonance Spectra of Orientated Molecules. *Physical Review Letters* **11,** 462.

181. Sayed, A., Matsuyama, S. and Inouye, M. (1999). Era, an essential Escherichia coli small G-protein, binds to the 30S ribosomal subunit. *Biochem Biophys Res Commun* **264,** 51–4.

182. Schaup, H. W., Green, M. and Kurland, C. G. (1971). Molecular interactions of ribosomal components. II. Site-specific complex formation between 30S proteins and ribosomal RNA. *Mol Gen Genet* **112,** 1–8.

183. Schindelin, H., Marahiel, M. A. and Heinemann, U. (1993). Universal nucleic acid-binding domain revealed by crystal structure of the B. subtilis major cold-shock protein. *Nature* **364,** 164–8.

184. Schleucher, J., Sattler, M. and Griesinger, C. (1993). Coherence Selection by Gradients without Signal Attenuation—Application to the 3-Dimensional Hnco Experiment. *Angewandte Chemie—International Edition in English* **32,** 1489–1491.

185. Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J., Sorensen, O. W. and Griesinger, C. (1994). A General Enhancement Scheme in Heteronuclear Multidimensional NMR Employing Pulsed-Field Gradients. *Journal of Biomolecular NMR* **4,** 301–306.

186. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **102,** 615–23.

187. Serrano, L., Horovitz, A., Avron, B., Bycroft, M. and Fersht, A. R. (1990). Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles. *Biochemistry* **29,** 9343–52.

188. Shaka, A. J., Barker, P. B. and Freeman, R. (1987). 3-Spin Effects in Broad-Band Decoupling. *Journal of Magnetic Resonance* **71,** 520–531.

189. Shaka, A. J., Lee, C. J. and Pines, A. (1988). Iterative Schemes for Bilinear Operators—Application to Spin Decoupling. *Journal of Magnetic Resonance* **77,** 274–293.

190. Shaka, A. J. and Pines, A. (1987). Symmetrical Phase-Alternating Composite Pulses. *Journal of Magnetic Resonance* **71,** 495–503.

191. Shine, J. and Dalgarno, L. (1974). The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA* **71,** 1342–6.

192. Siomi, H., Matunis, M. J., Michael, W. M. and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res* **21,** 1193–8.

193. Sklenar, V. and Bax, A. (1987). Spin-Echo Water Suppression for the Generation of Pure-Phase Two-Dimensional NMR-Spectra. *Journal of Magnetic Resonance* **74,** 469–479.

194. Sklenar, V., Peterson, R. D., Rejante, M. R. and Feigon, J. (1994). Correlation of Nucleotide Base and Sugar Protons in a N-15-Labeled Hiv-1 RNA Oligonucleotide by H-1-N-15 Hsqc Experiments. *Journal of Biomolecular NMR* **4,** 117–122.

195. Smith, D. W., Crowder, S. W. and Reich, N. O. (1992). In vivo specificity of EcoRI DNA methyltransferase. *Nucleic Acids Res* **20,** 6091–6.

196. Sohlberg, B., Lundberg, U., Hartl, F. U. and von Gabain, A. (1993). Functional interaction of heat shock protein GroEL with an RNase E-like activity in Escherichia coli. *Proc Natl Acad Sci USA* **90,** 277–81.

197. Sosnick, T. R., Fang, X. and Shelton, V. M. (2000). Application of circular dichroism to study RNA folding transitions. *Methods Enzymol* **317,** 393–409.

198. Sosnick, T. R. and Pan, T. (2004). Reduced contact order and RNA folding rates. *J Mol Biol* **342,** 1359–65.

199. Spassov, V. Z., Karshikoff, A. D. and Ladenstein, R. (1994). Optimization of the electrostatic interactions in proteins of different functional and folding type. *Protein Sci* **3,** 1556–69.

200. Spera, S., Ikura, M. and Bax, A. (1991). Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment. *J Biomol NMR* **1,** 155–65.

201. Stark, H., Mueller, F., Orlova, E. V., Schatz, M., Dube, P., Erdemir, T., Zemlin, F., Brimacombe, R. and van Heel, M. (1995). The 70S Escherichia coli ribosome at 23 A resolution: fitting the ribosomal RNA. *Structure* **3,** 815–21.

202. Steitz, J. A. and Tycowski, K. T. (1995). Small RNA chaperones for ribosome biogenesis. *Science* **270,** 1626–7.

203. Stejskal, E. O. and Tanner, J. E. (1965). Spin Diffusion Measurements—Spin Echoes in Presence of a Time-Dependent Field Gradient. *Journal of Chemical Physics* **42,** 288.

204. Stoffler, G., Deusser, E., Wittmann, H. G. and Apirion, D. (1971). Ribosomal proteins. XIX. Altered S5 ribosomal protein in an Escherichia coli revertant from streptomycin dependence to independence. *Mol Gen Genet* **111,** 334–41.

205. Stoffler-Meilicke, M. and Stoffler, G. (1987). The topography of ribosomal proteins on the surface of the 30S subunit of Escherichia coli. *Biochimie* **69,** 1049–64.

206. Strop, P. and Mayo, S. L. (2000). Contribution of surface salt bridges to protein stability. *Biochemistry* **39,** 1251–5.

207. Sun, D. P., Sauer, U., Nicholson, H. and Matthews, B. W. (1991). Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis. *Biochemistry* **30,** 7142–53.

208. Swapna, G. V., Shukla, K., Huang, Y. J., Ke, H., Xia, B., Inouye, M. and Montelione, G. T. (2001). Resonance assignments for cold-shock protein ribosome-binding factor A (RbfA) from Escherichia coli. *J Biomol NMR* **21,** 389–90.

209. Szilagyi, A. and Zavodszky, P. (2000). Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure* **8,** 493–504.

210. Takano, K., Tsuchimori, K., Yamagata, Y. and Yutani, K. (2000). Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry* **39,** 12375–81.

211. Tanner, N. K. and Linder, P. (2001). DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell* **8,** 251–62.

212. Teplova, M., Tereshko, V., Sanishvili, R., Joachimiak, A., Bushueva, T., Anderson, W. F. and Egli, M. (2000). The structure of the yrdC gene product from Escherichia coli reveals a new fold and suggests a role in RNA binding. *Protein Sci* **9,** 2557–66.

213. Theissen, G., Thelen, L. and Wagner, R. (1993). Some base substitutions in the leader of an Escherichia coli ribosomal RNA operon affect the structure and function of ribosomes. Evidence for a transient scaffold function of the rRNA leader. *J Mol Biol* **233,** 203–18.

214. Thieringer, H. A., Jones, P. G. and Inouye, M. (1998). Cold shock and adaptation. *Bioessays* **20,** 49–57.

215. Toone, W. M., Rudd, K. E. and Friesen, J. D. (1991). deaD, a new Escherichia coli gene encoding a presumed ATP-dependent RNA helicase, can suppress a mutation in rpsB, the gene encoding ribosomal protein S2. *J Bacteriol* **173,** 3291–302.

216. Traub, P. and Nomura, M. (1968). Structure and function of Escherichia coli ribosomes. I. Partial fractionation of the functionally active ribosomal proteins and reconstitution of artificial subribosomal particles. *J Mol Biol* **34,** 575–93.

217. Tscherne, J. S., Nurse, K., Popienick, P., Michel, H., Sochacki, M. and Ofengand, J. (1999). Purification, cloning, and characterization of the 16S RNA m5C967 methyltransferase from Escherichia coli. *Biochemistry* **38,** 1884–92.

218. Tscherne, J. S., Nurse, K., Popienick, P. and Ofengand, J. (1999). Purification, cloning, and characterization of the 16 S RNA m2G1207 methyltransferase from Escherichia coli. *J Biol Chem* **274,** 924–9.

219. van Buul, C. P. and van Knippenberg, P. H. (1985). Nucleotide sequence of the ksgA gene of Escherichia coli: comparison of methyltransferases effecting dimethylation of adenosine in ribosomal RNA. *Gene* **38,** 65–72.

220. Venema, J. and Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. *Annu Rev Genet* **33,** 261–311.

221. Vuister, G. W. and Bax, A. (1993). Quantitative J Correlation—a New Approach for Measuring Homonuclear 3-Bond J(H(N)H(Alpha) Coupling-Constants in N-15-Enriched Proteins. *Journal of the American Chemical Society* **115,** 7772–7777.

222. Vuister, G. W. and Bax, A. (1994). Measurement of 4-Bond H(N)-H-Alpha J-Couplings in Staphylococcal Nuclease. *Journal of Biomolecular NMR* **4,** 193–200.

223. Wang, H., Eberstadt, M., Olejniczak, E. T., Meadows, R. P. and Fesik, S. W. (1998). A liquid crystalline medium for measuring residual dipolar couplings over a wide range of temperatures. *Journal of Biomolecular NMR* **12,** 443–446.

224. Wilson, K. S., Appelt, K., Badger, J., Tanaka, I. and White, S. W. (1986). Crystal structure of a prokaryotic ribosomal protein. *Proc Natl Acad Sci USA* **83,** 7251–5.

225. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* **407,** 327–39.

226. Wireman, J. W. and Sypherd, P. S. (1974). In vitro assembly of 30S ribosomal particles from precursor 16S RNA of Escherichia coli. *Nature* **247,** 552–4.

227. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. and Sykes, B. D. (1995). H-1, C-13 and N-15 Chemical-Shift Referencing in Biomolecular NMR. *Journal of Biomolecular NMR* **6,** 135–140.

228. Wishart, D. S. and Sykes, B. D. (1994). Chemical shifts as a tool for structure determination. *Methods Enzymol* **239,** 363–92.

229. Wittekind, M. and Mueller, L. (1993). Hncacb, a High-Sensitivity 3d NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha-Carbon and Beta-Carbon Resonances in Proteins. *Journal of Magnetic Resonance Series B* **101,** 201–205.

230. Woese, C. R., Gutell, R., Gupta, R. and Noller, H. F. (1983). Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. *Microbiol Rev* **47,** 621–69.

231. Wrzesinski, J., Nurse, K., Bakin, A., Lane, B. G. and Ofengand, J. (1995). A dual-specificity pseudouridine synthase: an Escherichia coli synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). *RNA* **1,** 437–48.

232. Wunderlich, F., Speth, V., Batz, W. and Kleinig, H. (1973). Membranes of Tetrahymena. 3. The effect of temperature on membrane core structures and fatty acid composition of Tetrahymena cells. *Biochim Biophys Acta* **298,** 39–49.

233. Xia, B., Ke, H., Jiang, W. and Inouye, M. (2002). The Cold Box stem-loop proximal to the 5'-end of the Escherichia coli cspA gene stabilizes its mRNA at low temperature. *J Biol Chem* **277,** 6005–11.

234. Xia, B., Ke, H., Shinde, U. and Inouye, M. (2003). The role of RbfA in 16S rRNA processing and cell growth at low temperature in Escherichia coli. *J Mol Biol* **332,** 575–84.

235. Yamanaka, K. and Inouye, M. (2001). Induction of CspA, an E. coli major cold-shock protein, upon nutritional upshift at 37 degrees C. *Genes Cells* **6,** 279–90.

236. Yeh, Y. C., Traut, R. R. and Lee, J. C. (1986). Protein topography of the 40 S ribosomal subunit from Saccharomyces cerevisiae as shown by chemical cross-linking. *J Biol Chem* **261,** 14148–53.

237. Young, R. A. and Steitz, J. A. (1978). Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA. *Proc Natl Acad Sci USA* **75,** 3593–7.

238. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. and Noller, H. F. (2001). Crystal structure of the ribosome at 5.5 A resolution. *Science* **292,** 883–96.

239. Zengel, J. M. and Lindahl, L. (1994). Diverse mechanisms for regulating ribosomal protein synthesis in Escherichia coli. *Prog Nucleic Acid Res Mol Biol* **47,** 331–70.

240. Zhou, H. X. and Dong, F. (2003). Electrostatic contributions to the stability of a thermophilic cold shock protein. *Biophys J* **84,** 2216–22.

241. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31,** 3406–3415.

Zusammenfassung

In der vorliegenden Dissertationsschrift wurde die Raumstruktur eines 120 Aminosäure langen Proteins mittels mehrdimensionaler, heteronuklearer NMR-Spektroskopie an ^{15}N- bzw. $^{15}N/^{13}C$-isotopenmarkierten Proteinproben bestimmt. Es handelt sich dabei um das Kälteschockprotein „ribosome binding factor A" (RbfA) aus *Thermotoga maritima*. Die Struktur in Lösung von *Thermotoga maritima* RbfA in freier Form wurde mit einer Präzision von 0.99 Å für die durchschnittliche Abweichung der Schweratomkoordinaten zur mittleren Struktur einer Konformerenschar bestimmt. Dazu wurden aus NMR-Experimenten gewonnene „Nuclear Overhauser Effekt" Abstandseinschränkungen, dihedrale Winkeleinschränkungen aus Kopplungskonstanen und dipolare Restkopplungen verwendet. Das Protein wies eine Faltungstopolgie auf, die ähnlich der einer K-Homologie-Domäne ist. Solche KH-Domänen wurden als Bindungsmotiv für RNA-Liganden identifiziert. In der RbfA-Familie ist die für KH-Domänen signifikante GXXG-Sequenz, die ein Helix-Schlaufe-Helix-Motiv bildet, durch die AXG-Sequenz ersetzt. Die Analyse der ^{15}N-Relaxationsraten zeigte eine hohe Flexibilität der N- und C-Termini auf, sowie mehrere Regionen im Inneren des Proteins, deren Relaxationsparameter auf eine hohe Flexibilität schliessen lassen. Dies bezieht sich auf das C-terminale Ende der Helix α1 und das Helix-Schlaufe-Helix-Motiv der AXG-Sequenz.

Das RbfA-Protein bindet an die Helix-I-Region der 16S-rRNA. Um den Bindungsmodus zwischen RbfA und der 16S-rRNA näher zu untersuchen, wurde ein zweites Konstrukt von RbfA aus *Helicobacter pylori* verwendet. Um die beiden Proteine aus *T. maritima* und *H. pylori* miteinander zu vergleichen wurden NMR-spektroskopische Experimente durchgeführt. Beide Proteine weisen dieselbe Anordnung und Faltung der Sekundärstruktur auf und besitzen dieselben dynamischen Eigenschaften. Daraus wurde der Schluss gezogen, dass beide Proteine dieselbe drei-dimensionale Faltung einer KH-Domäne aufweisen. Dies wurde bestätigt durch den Vergleich mit der schon publizierten RbfA-Struktur des *Escherichia coli* Proteins, das dieselben Charakteristika aufweist. Dabei wurden jedoch Unterschiede zwischen dem thermophilen *T. maritima* RbfA und dem mesophilen *E. coli* RbfA aufgedeckt. Das *T. maritima* RbfA weist eine weitaus kompaktere Faltung auf, was mittels der Berechnung der „Contact Order" bewiesen werden konnte. Unter Zuhilfenahme der CD-Spektroskopie wurde zudem aufgezeigt, dass das mesophile *H. pylori* RbfA Protein eine für die Faltung destabilisierende Verteilung von Oberflächenladungen besitzt, während das thermophile *T. maritima* RbfA eine stabilitätsfördernde Verteilung von Oberflächenladungen hat. Diese Oberflächenladungen ermöglichen es dem thermophilen

Protein eine kompaktere Struktur einzunehmen und durch das Zusammenspiel verschiedener Faktoren somit eine höhere Stabilität zu erreichen.

Im zweiten Teil der Arbeit wurde der RNA-Ligand für das RbfA-Protein identifiziert. Die Sekundär-Struktur der RNA im ungebundenen Zustand wurde mittels NMR spektroskopischer Daten ermittelt. Es konnte gezeigt werden, dass das RbfA-Protein mit der Helix-I-Region der 16S-rRNA interagiert. Die apparenten Bindungskostanten wurden mittels Fluoreszenzspektroskopie bestimmt. Daraus ergab sich eine Bindungskonstante des RbfA-Proteins an die Helix I (Nukleotide 1–26) von ~20 μM. RNA-Konstrukte, die mittels einer Poly-Uracil-Schlaufe den Pseudoknoten der Helix-I-Region imitieren, wurden ebenso auf ihre Bindungseigenschaften zum RbfA-Protein untersucht. Deren Analyse ergab eine Bindungskonstante von ~1 μM. Konstrukte, deren Helix I durch Basenpaaraustausch stabilisiert wurde, wiesen eine Bindungskonstante im Bereich von ~1 μM auf. Dies führte zur Vermutung, dass das RbfA-Protein an eine schon in ihrer nativen Struktur gefaltete RNA-Helix bindet. Desweitern wurde untersucht ob die Prä-RNA-Sequenz der unreifen 16S-rRNA mit dem RbfA-Protein interagiert. Diese Untersuchungen führten zur Identifizierung der nativen Bindungsstelle des RbfA-Proteins an seinen RNA-Liganden. Die Faltung der Prä-Sequenz, welche die Nukleotide –18–1 umfasst, führt zur Bildung einer alternative Helixfaltung in der unreifen 16S-rRNA, die mit der Helix I nicht koexistieren kann. Die durch Fluoreszenzspektroskopie erhaltenen apparenten Bindungskonstanten liegen im unteren nanomolaren Bereich. Dies ließ am Ende dieser Studie den Schluss zu, dass das RbfA-Protein durch die Bindung an die alternative Helix der Prä-Sequenz die korrekte Schnittstelle der unreifen rRNA der Endoribonuklease RNAse G präsentiert. Daduch wird die Faltung in eine native Helix-I-Konformation ermöglicht, an die im Laufe des Aufbaus der kleinen ribosomalen Untereinheit das S5-Protein bindet.

Danksagung

In den letzten vier Jahren, die ich am Institut für Organische Chemie und Chemische Biologie zubrachte, konnte ich unglaublich viele Erfahrungen sammeln. Da dies nicht ohne die Hilfe von vielen Kollegen vonstattengeht, versteht sich von selbst. Aus diesem Grund möchte ich an dieser Stelle meinen Dank aussprechen an alle, die an der Fertigstellung dieser Arbeit beteiligt waren.

Allen voran möchte ich meinem Doktorvater Prof. Dr. Harald Schwalbe danken, der unablässlich dafür arbeitet, hervorragende Arbeitsbedingungen zu schaffen und jederzeit für Anregungen und Ratschläge zur Verfügung stand. Mein besonderer Dank geht an meinen Betreuer Prof. Dr. Jens Wöhnert, dessen Themenstellung ich bearbeitet habe. Dabei zeichnete sich schon früh eine beispiellose Zusammenarbeit ab, die ihren Höhepunkt im letzten Jahr meiner Doktorarbeit in San Antonio erreichte. In der Zeit, die ich mit Jens in San Antonio verbrachte, konnte ich ungemein viel dazulernen. Für das Vertrauen, mich an dieser Zeit teilnehmen zu lassen, möchte ich mich nochmal extra bedanken.

Nun folgend möchte ich mich bei all den Leuten bedanken, die während meiner Zeit in Frankfurt meinen Weg begleiteten. Dr. Christian Richter danke ich für die hervoragende Unterstützung bei allen Fragen, vor allem der, die sich mit den NMR-Spektrometern befassten. Dr. Henry Jonker danke ich, dass er mit mir zusammen alle Strukturrechnungen ansetzte. Prof. Dr. Hashim Al-Hashimi danke ich, dass er die manchmal nur schwach flackernde Lust an der Wissenschaft immer zu einer hoch auflodernden Flamme formen konnte. Dr. Holger Berk danke ich für die vielen Tipps im Biolabor, die einem das Leben einfacher machen. Dr. Oliver Ohlenschläger danke ich, mir die ersten Schritte der Strukturrechnung näher gebracht zu haben, und von diesem Punkt an immer für Fragen erreichbar war. Dem „Boys Room“, welcher aus Boris Fürtig, Christian Schlörb, Kai Schlepckow, Sergey Ilin, Jan-Peter Ferner, Jürgen Graf und später noch Martin Hähnke bestand, möchte ich für die gemeinsame Zeit danken. Darunter gab es Tage, die sich für den Rest meines Lebens in meine Erinnerungen eingebrannt haben. Dem „Girls Room“, welcher aus Dr. Julia Wirmer, Dr. Elke Duchardt, Karla Werner, Dr. Aphrodite Anastasiadis-Pool, Dr. Emily Collins und später noch Janina Buck, Vijayalaxmi Manoharan, Gerd Nielsen und Hamid Nasiri bestand, möchte ich für die gemeinsame facettenreiche Zeit danken. Auch wenn Ihr dem „Boys Room“ nie das Wasser reichen konntet. Dem „Middle Room“, welcher aus Stefan Rehm, Jonas Noeske, Dr. Karin Abarca Heidemann, Jitendra Kumar und Sridhar Sreeramulu bestand, danke ich für die gemeinsame Zeit. Wobei auch gemeinsame sportliche Intressen nicht unerwähnt bleiben sollten. Karin danke ich im Besonderen für die gemeinsame Zeit in San An-

tonio. Elke Stirnal und Sarah Mensch danke ich für die Unterstützung in allen Laborfragen. Dem Sekretariatsteam Anna Paulus und Elena Hartmann danke ich für die Unterstützung in allen Belangen, die sich im Laufe der Doktorarbeit ansammeln. Ich danke auch allen hier nicht namentlich erwähnten früheren und derzeitigen Mitgliedern des Arbeitskreises für die gute Zusammenarbeit und stete Hilfsbereitschaft.

Im Besonderen danke ich dem „Hinck Lab", das mich in San Antonio sehr herzlich in ihrem Labor aufnahm. Bei dieser Gelegenheit möchte ich auch Belinda Leal danken, die unser Labor in San Antonio organisiert.

Besonderer Dank geht auch an meine Feunde und an meine Familie, die eher unfreiwillig an der Fertigstellung dieser Arbeit teilhaben durften. Insbesondere möchte ich meiner Freundin Sibylle Bosshammer danken, die für die manchmal bitter nötige Ablenkung sorgte.

List of Publications

Grimm, S. K., Jonker, H. R., Ohlenschlager, O., Wohnert, J. (2007). Comparison of thermophilic and mesophilic proteins of the RbfA family. Manuscript to be submitted.

Grimm, S. K. and Wohnert, J. (2007). Ribosome binding factor A (RbfA) preferentially binds to an alternative helix of the pre-16S rRNA in vitro. Manuscript to be submitted.

Jonker, H. R., Ilin, S., Grimm, S. K., Wohnert, J., Schwalbe, H. (2007). L11 domain rearrangement upon binding to RNA and thiostrepton studied by NMR spectroscopy. *Nucleic Acid Res.* **35** (2), 441–54.

Grimm, S. K., and Wohnert, J. (2005). NMR assignment of the cold shock protein ribosome-binding factor A (RbfA) from Thermotoga maritima. *J Biomol NMR* **31,** 73–74.

Curriculum vitæ

Steffen Kaspar Grimm

PERSONAL DATA

Born: December 31st, 1974, Offenbach/Germany, (German citizen)

EDUCATION

Since 2006	*University of Texas Health Science Center San Antonio* Student associate with Prof. Jens Wöhnert San Antonio, Texas, United States of America
Since 2003	*University of Frankfurt/Main, Germany* PhD student in Biochemistry with Prof. Dr. H. Schwalbe "Structure, RNA-binding specificity and function of RbfA—a bacterial cold shock protein"
2003	*University of Frankfurt/Main, Germany* Diploma in Biochemistry with Prof. Dr. H. Schwalbe "Biochemical and NMR-spectroscopic experiments on the L11-RNA-thiostrepton complex"
2000	*University of Frankfurt/Main, Germany* Intermediate diploma in Biochemistry
1998–2003	*University of Frankfurt/Main, Germany* Biochemistry student
1998	*Technical University of Darmstadt, Germany* Chemistry student
1994–1997	*Ludwig Geissler Schule Hanau/Main, Germany* A-level College (Higher education entrance qualification)
1991–1994	*Hoechst AG Offenbach/Main, Germany* Apprenticeship chemical laboratory worker
1987–1991	*Einhardschule Seligenstadt/Main, Germany* Secondary school (GCSE)
1987–1987	*Gerhart-Hauptmann-Schule Seligenstadt/Main, Germany* Secondary school orientation stage
1985–1985	*Walinusschule Seligenstadt/Main, Germany* Primary school