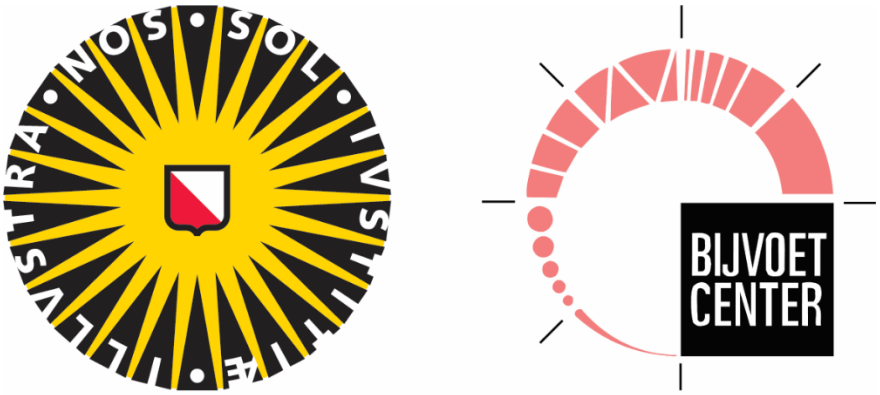


Utrecht University

Department of Chemistry
Bijvoet Centre for Biomolecular Research
NMR Spectroscopy Group

NMR studies of an enzyme-RNA complex that confers antimicrobial resistance

Major internship Report 2021-2022
Jochem de Waard
9669035

Molecular and Cellular Life Sciences (MSc)

Examiner:	Dr. Markus Weingarth
Daily supervisor:	Francesca Lavore, MSc
Second reviewer:	Dr. Hugo van Ingen

Utrecht, December 2022

Table of contents

1 Abstract

Resistance to antimicrobials is emerging at an ever-increasing pace. To ensure that our current spectrum of clinically used antimicrobials remain active, the development of inhibitors that block resistance mechanisms is a crucial strategy in the fight against antimicrobial resistance. Macrolides, lincosamides and streptogramins antibiotics are a class of antimicrobials that are effective against a plethora of bacteria. Bacteria gain resistance by modifying a specific nucleotide in the ribosomal-RNA (rRNA) of their ribosome, the binding site of these antibiotics. The enzymes responsible for this modification are erythromycin resistance methyltransferases (Erm) proteins. To prevent or overcome resistance development, a strategy is to develop inhibitors that prevent the interaction between Erm and rRNA. The effective development of such inhibitors requires a structural understanding of the Erm-rRNA interaction. Previous research created a minimal RNA substrate for Erm enzymes, however, when combining Erm and the 32 nucleotide RNA that resembles the natural fold of the rRNA the complex precipitates. The precipitate limits the structural investigation by conventional methods.

Here we study the non-crystalline precipitate that forms between ErmB and 32-mer RNA at atomic level using a combination of solution and solid-state nuclear magnetic resonance (NMR). Solution NMR enabled us to assign the backbone of the protein and transfer these to the ssNMR complex spectra. Using ssNMR, we show that the ErmB protein is highly ordered in the precipitate, while the RNA is heterogeneous and dynamic. However, the lack of clear signal changes between apo-ErmB and complex spectra; the similar dynamics of the protein in both states, and the peculiar behaviour of the RNA in the complex appear to question if the RNA binds specifically to ErmB in the precipitate. These observations raised the question whether the formed precipitate is actually representative of the actual binding interaction between Erm and RNA.

2 Layman summary

Pathogenic bacteria are becoming increasingly resistant to antibiotics, the drugs we use to fight them. Globally, the impact of resistance development among bacteria is on track to result in more deaths yearly than cancer. To combat resistance among bacteria the scientific community has mainly focussed on understanding the way bacteria obtain resistance for the used drugs and development of new, still effective, drugs. However, inhibiting further development of AMR for clinically important antibiotics, which is already realized for penicillin, is a strategy that needs to be expanded onto other antibiotic classes.

One class of antibiotics, for which AMR is emerging, was investigated within this research. These antibiotics are effective against a high variety of bacteria and are often used clinically. Bacteria gain resistance for these antibiotics by having a specific enzyme, catalysts for reactions in our body, modify the target of the antibiotics. Modification of the target makes them unable to bind and eliminate the bacteria. Inhibition of this enzyme could prevent further resistance development, which allows these antibiotics to remain a viable option clinically.

The modification site of the enzyme is RNA that is located in the bacterial ribosome, a complex molecular structure that creates proteins within the bacteria and is crucial for its survival. The enzyme can only modify the RNA, when the ribosome is not fully assembled. To design an inhibitor of the enzyme that modifies the target of the antibiotics, direct information is needed on which parts of the enzyme interact with which parts of the RNA. Therefore, researchers had to create a minimal RNA substrate for the enzyme. When this minimal RNA substrate and the enzyme are combined in a solution the two bind strongly together and precipitate into solids. Due to the solids, frequently used methods for gaining information on the interaction are unavailable.

Here the first structural data is reported on the enzyme-RNA interaction using a method that is able to obtain atomic level information on solids. The goal was to compare atomic level information of the enzyme alone with the atomic level information of the enzyme-RNA complex. It was expected that differences in the atoms that are involved in the interaction would be seen. Atoms that show differences can then be seen as the binding site of the enzyme-RNA interaction. However, the obtained atomic level information did not show differences for any atoms. Therefore, the data was more indicative of weak or unspecific binding between the enzyme and RNA, the opposite of what was expected. Moreover, the obtained information raised the question: are the solids that form upon combination of enzyme-RNA actually representative of the enzyme-RNA interaction that occurs in nature? For development of an inhibitor of the enzyme-RNA interaction atomic level information is a requirement and therefore different means to gain this information have to be investigated.

3 Introduction

3.1 Antimicrobial resistance

Antimicrobial resistance (AMR) is the phenomenon when microorganisms, such as bacteria, fungi and viruses, become insusceptible to previously effective drugs or other methods of extermination [1]. Since the discovery of the first antimicrobial drug, penicillin, and its first clinical use in 1942, AMR has been emerging across the globe at an ever-increasing pace. An investigation by the Antimicrobial Resistance Collaborators showed that in the year 2019 nearly 5 million deaths could be attributed to AMR [2], whilst, it is believed that this number will increase to approximately 50 million deaths by the year 2050 [3]. AMR is not only a threat to public health, since a reduction of 2-3.5% in gross domestic product (GDP) by the year 2050 can be expected, causing a huge financial burden worldwide [3]. Moreover, according to a World Health Organization (WHO) reported in 2022, development of new antimicrobial drugs is inadequate to address the rise of AMR [4], therefore, interventions have to be made. For example, global awareness has to be spread, sanitation has to be improved and unnecessary use of antimicrobials has to be avoided [5].

Nevertheless, the increasing concern remains: AMR has developed into a global threat and is leading to a pressing need for new solutions. Unravelling the mechanism of action of drugs already on the market, as well as, understanding the molecular basis of AMR are essential for future developments in the field. However, there is a significant lack of research on alternative strategies. Baym, Stone and Kishony (2016) wrote a review on multidrug evolutionary strategies to reverse antibiotic resistance. However, they also state that significant development and validation are needed for multidrug evolutionary strategies to become clinically relevant [6]. Meanwhile, others have re-evaluated and optimised the use of existing antibiotics to increase efficiency and reduce resistance development [7]. Beside the development of new antimicrobials effective against the resistant target, another alternative strategy is represented by the inhibition of the biochemical pathways leading to the development of resistance mechanisms, for which structural data of the involved substrates is a necessity.

3.2 Macrolides resistance mediated by Erythromycin Resistance Methyltransferases

Due to their fundamental role of protein synthesis in the cell, ribosomes are essential molecular machineries constituting one of the main antibiotic targets in bacteria. Macrolides, lincosamides and streptogramins (MLS) are a broad class of ribosome-targeting antibiotics, widely used in the clinic for the treatment of bacterial infections of the upper respiratory tract, skin and soft tissues caused by gram-positive bacteria, including *Streptococcus pneumoniae* and *Staphylococcus aureus [8, 9]*.

The first type of MLS antibiotic, erythromycin (figure 1A), was introduced for clinical use in 1952. Merely 1 year later, the first reports of resistance for MLS antibiotics were disclosed, highlighting the immediate adaption of bacteria [10]. In general, MLS antibiotics bind to the large ribosomal subunit, via the peptide exit tunnel, and prevent peptide elongation during protein synthesis and exploit a bactericidal mode of action which induces cell death [11].

Despite their different chemical structure (figure 1A/B/C) [12], MLS antibiotics target the ribosome in a similar manner and are inhibited by common mechanisms of resistance. Resistance to ribosome targeting antibiotics can arise from different mechanisms (figure 1D) [13], among which target modification (d) is the most clinically relevant for MLS antibiotics. In particular, resistance to MLS antibiotics is conferred by a family of structurally-conserved enzymes named erythromycin resistance methyltransferases (Erms), which are responsible for mono-or di-methylation of a specific adenine in the bacterial ribosome. The mechanism of MLS resistance has been elucidated by Svetlov et al. (2021)

using X-ray crystallography. They compared the crystal structure of erythromycin bound to the wild type ribosome (figure 1E) and bound to the Erm-dimethylated ribosome (figure 1F).

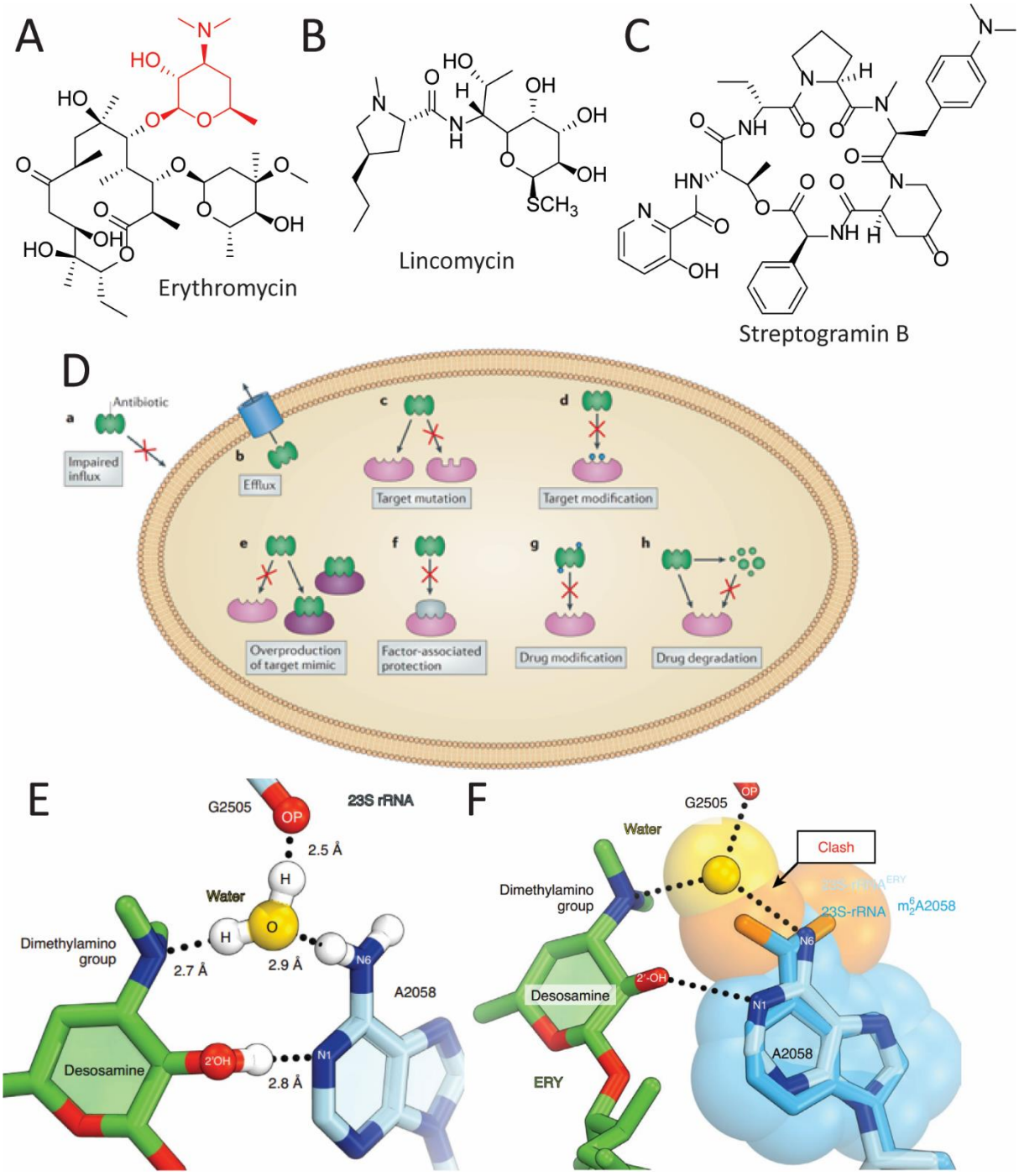


Figure 1: A) Structure of the macrolide antibiotic: erythromycin, with the desosamine-group highlighted in red. [12] B) Structure of the lincosamides antibiotic: lincomycin. [12] C) Structure of the streptogramin antibiotic: Streptogramin B. [12] D) Summary of different bacterial AMR mechanisms adapted from: [13]. Bacterial AMR mechanisms include (a) an impaired influx of the antibiotic or (b) forced efflux; (c) mutation or (d) modification of the antibiotics' target; (e) overproduction of an imitate of the target or (f) protection of the target by another ligand; and lastly, (g) in cell antibiotic modification or (h) degradation. E) A model of binding of the desosamine-group in erythromycin modelled by X-ray crystallography. The binding of erythromycin to residue A2058 of the 23S rRNA subunit of the bacterial ribosome is mediated through hydrogen bonding with a water molecule or directly between the desosamine-group and the residue. F) Binding between the desosamine-group of erythromycin and residue A2058 is inhibited due to placement of 2 methyl groups on residue A2058 by Erm-class methyltransferases. The methyl groups leave no space for a water molecule to mediate binding. Figure adapted from: [14].

The desosamine-group binds (figure 1A), coordinated by a water molecule, to a specific nucleotide (A2058 in E.coli numbering) in the 23S ribosomal-RNA (rRNA) subunit of the wild type bacterial ribosome (figure 1E) [14]. In the Erm-modified ribosome (figure 1F), the presence of two methyl groups does not prevent the macrolide binding due to steric hinderance, as speculated previously, but it provokes the loss of one important interaction between the drug and its target [14]. Without this important interaction the MLS antibiotics are unable to bind and won't have bactericidal effects. Thus, the bacteria have become resistant to MLS antibiotics.

The Erm gene is found in a variety of bacterial pathogens, including methicillin-resistant *Staphylococcus aureus* (MRSA) [15], *Streptococcus Pneumonia* [16] and Vancomycin-resistant *Enterococcus* (VRE) [17]. Erm is responsible for mono- or di-methylation of the specific nucleotide (A2058) in the 23S rRNA subunit of the bacterial ribosome. The methyl groups, are transferred from the ubiquitous S-adenosylmethionine (SAM) cofactor that is converted into S-adenosyl homocysteine (SAH). Methylation of nucleotide A2058 can only happen during the ribosome assembly phase because, Erm is unable to reach the nucleotide in the matured ribosome [14]. The reaction scheme shown in figure 2 is adapted from Svetlov et al. and shows the reaction that occurs when Erm proteins, together with the cofactor SAM, methylate nucleotide A2058 to obtain partial or complete MLS resistance.

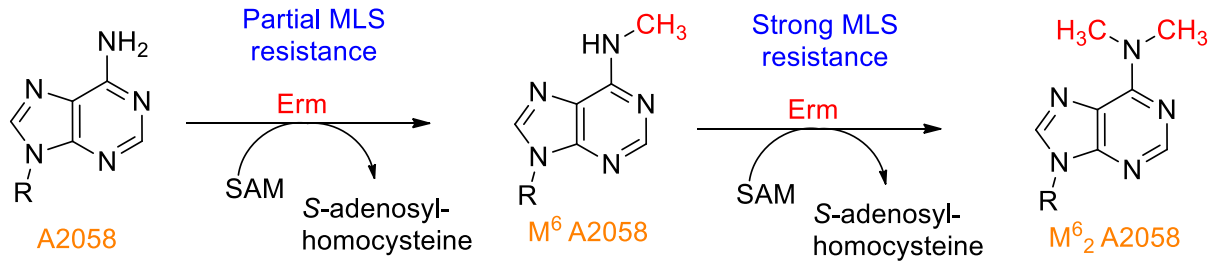


Figure 2: Reaction scheme for mono- and dimethylation of residue A2058 by Erm-class methyltransferases (adapted from: [14]). Upon monomethylation of residue A2058 partial resistance for MLS antibiotics is obtained, while dimethylation causes strong MLS resistance.

As mentioned above, the fight against AMR is in dire need of alternative strategies. One possibility for an alternative strategy could be to avoid the ribosomal modification by developing Erm inhibitors. As can be seen in the reaction scheme for MLS resistance (figure 2), there are three major contributors in this reaction: the Erm methylase, the ribosomal-RNA (rRNA) substrate, and the cofactor SAM. A potential Erm inhibitor could mimic either the cofactor or the rRNA substrate. An inhibitor based on an analogue of the cofactor SAM was already theorized in 2000 by Hanessian and Sgarbi. They used an ErmC' X-ray crystal structure to locate which important residues use hydrogen bonding for SAM binding and created multiple analogues of S-Adenosyl-L-homocysteine (SAH, the demethylated analogue of SAM). Unfortunately, of the ten analogues tested, the two which showed any binding only bound weakly [18]. Moreover, the cofactor SAM is used by many other methyltransferases [19] and will likely result in toxicity issues. This makes it difficult to create an inhibitor based on SAM binding site that is only specific to Erm.

The Erm-rRNA interaction is another possible target for the development of Erm inhibitors. Due to the loop alterations among the Erm proteins, compared to non-pathogenic methyltransferases [20], inhibition of the rRNA binding site is highly specific and will likely result in a low risk of toxicity. Erm inhibitors would bind to the rRNA binding site in Erm and prevent modification of the critical nucleotide A2058 of the 23S rRNA subunit, even in the presence of the cofactor. In this way retaining the activity of clinically relevant MLS antibiotics. The structural study of the Erm-rRNA interaction, using nuclear magnetic resonance (NMR) techniques, was the main goal of this research.

The system under study in this report is the complex between ErmB, a prevalent member of the Erm methyltransferase family, and a 32 nucleotide long RNA substrate, a minimal RNA substrate for Erm, designed to mimic the natural fold in the rRNA of the 23S rRNA subunit in the bacterial ribosome. The structural features of these two substrates are described in the next sections.

3.3 Erm methyltransferase family

During this study ErmB was chosen because, it is the most widespread Erm among bacteria. Sequence alignment of ErmB to related Erm proteins reveals highly conserved residues that could be crucial in cofactor or RNA recognition and binding (figure 3A). Prior research already highlighted important aspects of the biology and structure of the Erm family of methyltransferases. For instance, the structure of ErmAM was solved by solution-state nuclear magnetic resonance (NMR) [21], while, structures of ErmC' [22], ErmE [23] and Erm38 [24] have been solved using X-ray crystallography. The determined structures were used to find possible binding sites for Erm's cofactor (and analogues) and its substrate RNA [18, 24]. For Erm38 a possible RNA binding site has been simulated via molecular dynamics studies [24], which could aid structural studies in finding a binding site. Furthermore, sequence alignment and mutagenesis studies have identified certain residues that could be crucial to the proteins functioning [25-28].

Stretches of conserved residues have been identified amongst the methyltransferases (figure 3A). These conserved motifs have previously been named. Motifs IV, V, VI, VII, and VIII [26] have been identified as important in the ability of Erm proteins to recognize and bind the correct RNA sequence. Other conserved motifs are believed to be important in cofactor binding. Loops 1 and 12 are believed to be highly important in RNA substrate specificity. R. Bhujbalrao and R. Anand (2019) created different chimeras of *Bacillus subtilis* KsgA, a non-pathogenic methyltransferases in bacteria that is related to Erm, that more closely resembled Erm proteins to see if they could recreate the specificity of Erm. They found that by mutating or deleting parts of loop1 and loop12 they could mimic Erm's specificity for smaller RNA stretches, while KsgA is incapable of this. Based on their results they concluded that alterations in loop1 and loop12 are responsible for substrate selectivity among related methyltransferases [20].

As can be seen in the sequence alignment (figure 3A), ErmB and ErmAM differ only by 2 mutations. Therefore, it is very likely that the ErmAM structure is highly similar to ErmB's structure. Moreover, comparison of structures of solved Erm proteins has revealed high similarity between them. This makes apo-ErmB (unbound ErmB) a perfect candidate for homology-driven structure prediction using Alphafold [29, 30]. The predicted structure can be used to visualize findings and for structural comparison against solved structures of other Erm proteins. Important cofactor and RNA recognition sites, as well as the important recognition loops are visualized on the predicted structure (figure 3B). The predicted structure for ErmB shows the Rossmann-like α/β cofactor binding and catalytic N-terminal domain (residues 1-176), as well as, the helical C-terminal domain (residues 177-245) [23]. A predicted aligned error graph, that was provided when predicting the structure, reveals that Alphafolds' confidents only drops in the N- and C-terminal loops, however, the overall confidence in the predicted structure is high (figure S8).

When the predicted ErmB structure is aligned with the solved structures of the Erm proteins and KsgA shown in the sequence alignment the Root-Mean-Square deviation (RMSD) equals just 3.352Å. This indicates that the predicted Alphafold structure is highly representative of the actual structure, and useful for further data interpretation.

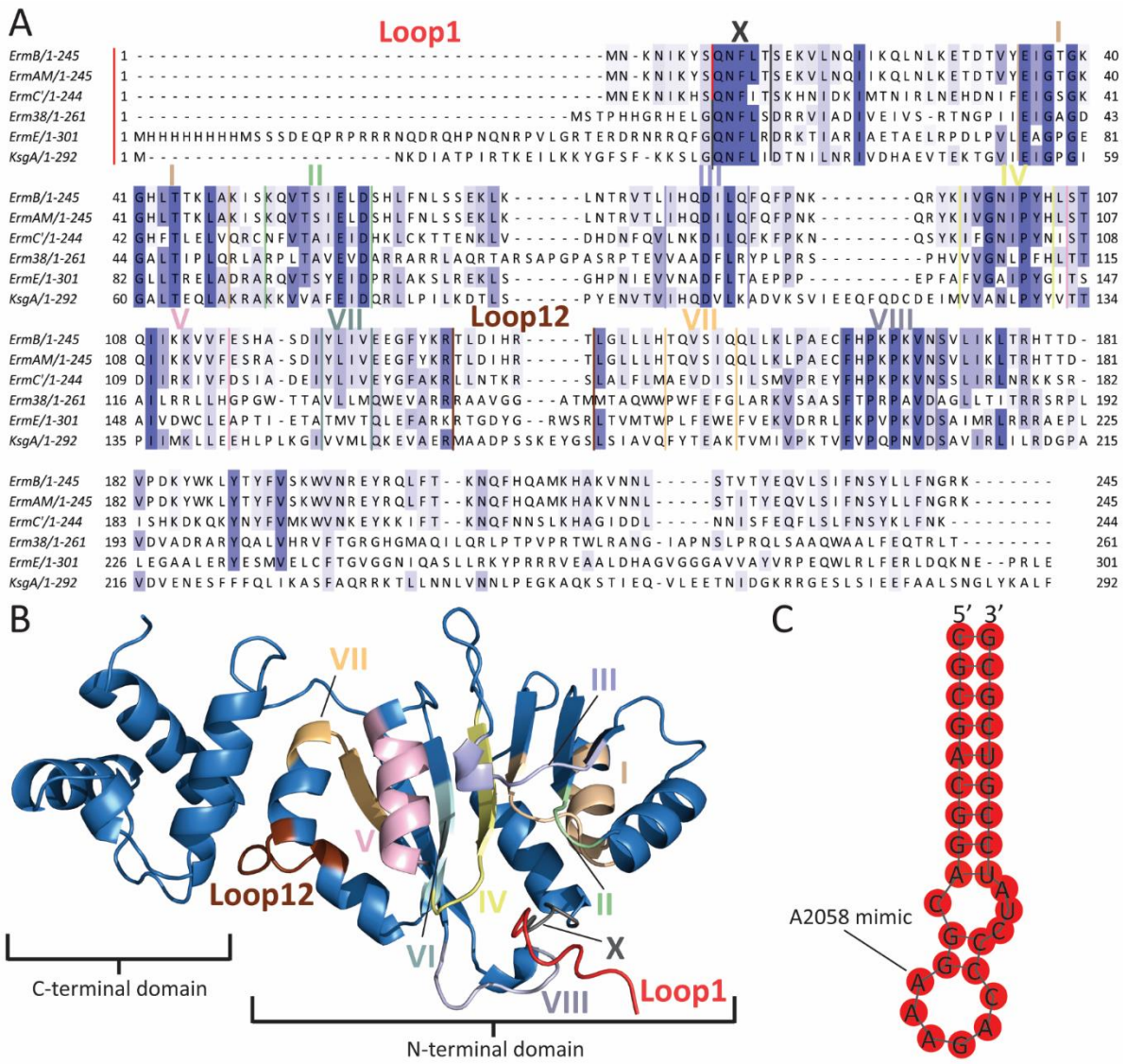


Figure 3: A) Sequence alignment between ErmB, other members of the Erm methyltransferase family and KsgA, a non-pathogenic bacterial methyltransferase. The previously identified motifs [26] and loops [20] that are important for cofactor/RNA recognition and binding are indicated according to their colour in the 3D model in B. B) Homology driven Alphafold predicted 3D model of ErmB based on its amino acid sequence. The important motifs and loops are coloured. C) Predicted secondary structure for 32-mer RNA, prediction performed with: [31].

3.4 Minimal RNA substrate for Erm

For the investigation on the Erm-rRNA interaction the RNA has to be defined as well. Since Erms can only methylate the specific adenosine before complete ribosome assembly, researchers had to create a minimal RNA substrate that can be methylated by Erms to investigate the Erm-RNA interaction. Previous research on ErmE revealed that the minimal RNA substrate, that resembles the natural fold of the 23S rRNA subunit of the bacterial ribosome, is a 27 nucleotide RNA sequence [32]. The minimal RNA substrate still showed methylation activity when combined with Erm and SAM. Moreover, this research also revealed important RNA characteristics required for recognition by Erm proteins [32]. For this structural investigation in the Erm-rRNA interaction a well-defined 32 nucleotide long RNA stretch was used [22], hereafter referred as 32-mer RNA (figure 3C) [31]. Our collaborators at the Schwalbe-lab in Frankfurt provided us with fully ^{15}N, ^{13}C labelled 32-mer RNA and specifically ^{15}N, ^{13}C adenosine labelled 32-mer RNA to help and elucidate a binding interaction. Adenosine labelling was chosen because an adenosine nucleotide gets methylated upon Erm modification, which makes it an important moiety to be investigated.

3.5 Study of ErmB-32-mer RNA complex

As previously stated, mutagenesis and molecular dynamic simulations have given some insights into the Erm-rRNA interaction. Although this is useful information, for drug design hard structural evidence is a necessity. However, to our knowledge, no direct structural or dynamical data has been reported on the binding site of the Erm-rRNA interaction. The reason why there is no direct structural data available on the interaction is the heavy precipitation that is observed when combining Erm with rRNA in relevant quantities, impeding structural investigations by most structural biology methods. Methylation activity was seen for the complex, but only when very low concentrations of RNA were added to the Erm protein in order to prevent precipitation [32]. Heavy precipitation rules out the use of X-ray crystallography and solution NMR, while the size of the complex (~40kDa) makes it too small for structure determination by cryo-electron microscopy (EM).

Due to the heavy precipitation of the complex, solid-state NMR (ssNMR) is the only structural elucidation method that can gain atomic insights into the ErmB-32-mer RNA interaction. ssNMR can obtain atomic resolution data, while also being able to probe the dynamics of the measured system. It does not require crystalline samples and can thus work on heterogenous samples, while, the linewidth is independent on the molecular weight of the system [33]. During this research we applied both solution NMR and ssNMR at high magnetic field up to 1200 MHz, to characterize the precipitate that forms between ErmB and 32-mer RNA. Solution NMR was first applied to assign the backbone of the protein, while collaborators assigned the resonances for the 32-mer RNA. Next, the complex was measured using ssNMR from both the ErmB perspective and the 32-mer RNA, which was enabled by isotopically labelling both compounds. We saw that the protein produces ssNMR spectra of good quality, but did not observe significant changes compared to the unbound solution NMR spectra. Spectra of the 32-mer RNA in the precipitate were of much lower quality. To investigate the odd behaviour of the 32-mer RNA and to elucidate a possible binding site the dynamics of both participants in the complex were determined. Efforts were also made to solubilize the complex.

4 Results

4.1 ErmB, 32-mer RNA and complex formation

As stated above, the fight against AMR is in high need of alternative strategies that inhibit AMR acquirement in clinically relevant bacteria. The aim of this research was to complete a structural characterization on the ErmB-rRNA binding interface, as well as, to gain insights on the dynamics of the complex. These potential results could then eventually aid the drug design towards ErmB-rRNA inhibitors, resulting in the restriction of acquiring more AMR against clinically relevant MLS antibiotics. Before any NMR experiments could be performed, ErmB had to be labelled with NMR active nuclei: ^{15}N and ^{13}C. However, the large amount of amino acids which make up the protein cause spectral crowding. Moreover, strong ^{1}H-^{1}H homonuclear dipole-dipole interactions affect relaxation in solution NMR and the line shapes of peaks in ssNMR [34, 35]. For more resolution in solution NMR and better line shapes in ssNMR a solution was needed. For ^{1}H-detected studies, we used perdeuterated (^{2}H, ^{13}C, ^{15}N) protein preparations to improve spectra resolution, which is based on dilution of the otherwise dense network of ^{1}H-^{1}H dipolar couplings [35]. The protein was expressed, purified and refolded using the described method in chapter 6.1, with a yield of approximately 20 mg of refolded ErmB per litre culture. During the expression it was quickly found that ErmB forms inclusion bodies and was therefore located in the insoluble fraction. This meant we needed to develop a refolding strategy [36]. The inclusion bodies were denatured using a 6M guanidinium solution. Dialysis and rapid dilution techniques were both applied to remove the denaturant [37], where rapid dilution using Amicon Ultra-15 Centrifugal Filter Unit tubes resulted in the highest protein recovery yield and optimal time requirement. During rapid dilution the protein starting concentration had considerable impact on the eventual yield. It was found that a starting concentration of approximately 1 mg/ml denatured ErmB resulted in the most time efficient method, while still retaining an acceptable protein recovery yield (~50%). SDS-gels showing cellular lysate before and after inducing expression, as well as after purification can be seen in figure S1A and S1B, respectively.

The used minimal RNA substrate for Erm; 32-mer RNA, was obtained either commercially or via our collaborators at the Schwalbe-lab in Frankfurt. After assembly of the complex between ErmB and 32-mer RNA it was tested by our collaborators at the Dedon-lab in Singapore to determine its binding constant, activity and specificity. Via isothermal titration calorimetry (ITC) the binding constant (Kd) of the complex was determined. The ITC curve shown was used to determine the Kd of the complex (figure 4A). Astonishingly, it was found that the Kd equals 49 nM, a remarkably high affinity. Moreover, the same binding constant was also observed by surface plasmon resonance (SPR) assays performed by our collaborators, confirming the ITC data (data not shown).

Moreover, the activity of the ErmB-32-Mer RNA complex was determined by looking at the amount of cofactor, S-adenosyl-methionine (SAM), that was turned into its demethylated counterpart, S-adenosyl-L-homocysteine (SAH). The graph revealed that an increased amount of SAM is converted into SAH with an increasing concentration of the 32-Mer RNA present (figure 4B). This is indicative for the formation of the actual complex, since otherwise methylation of the complex, shown by an increase in SAH concentration, would not occur and further proves that the 32-Mer RNA is the suitable substrate for ErmB.

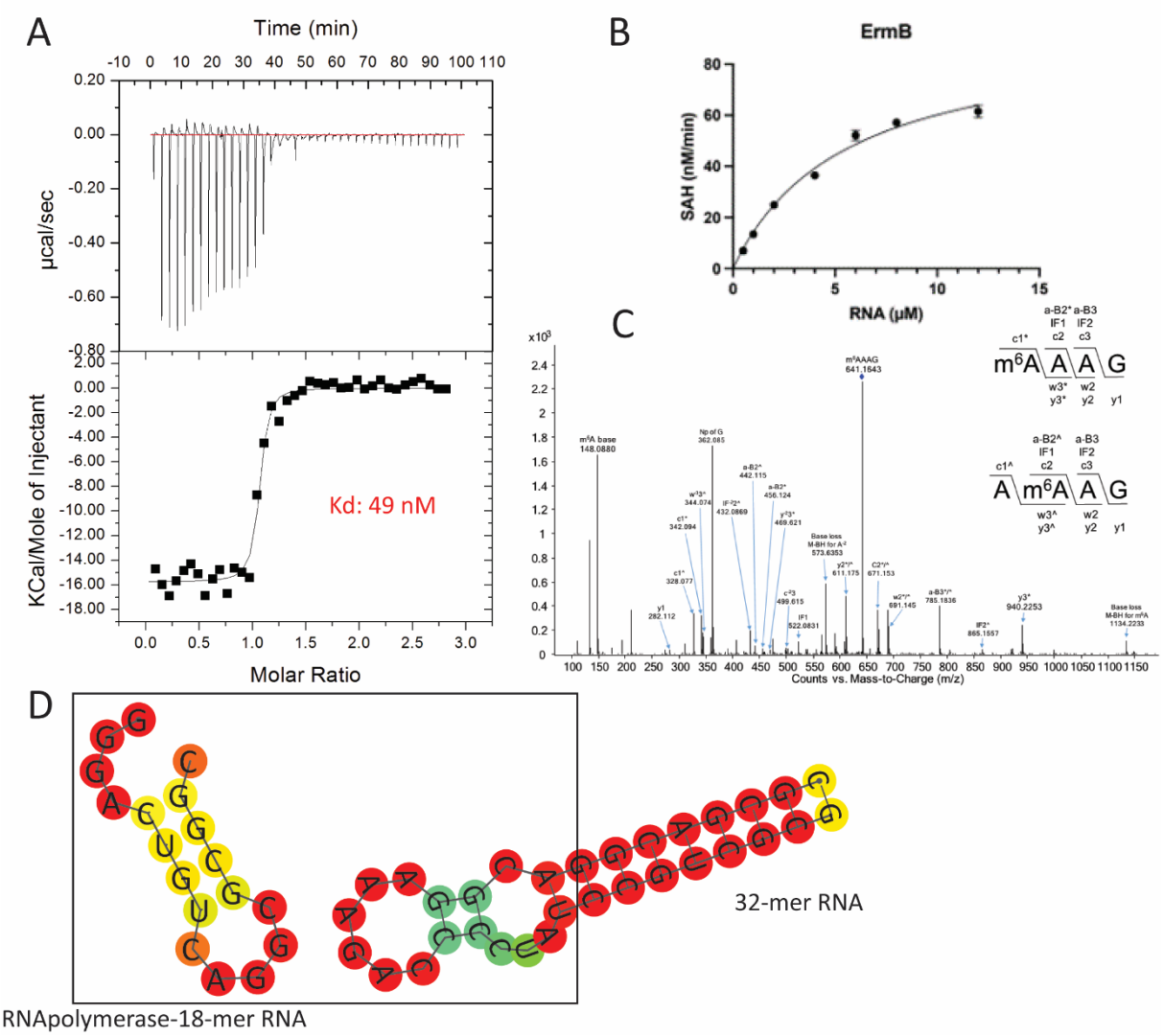


Figure 4: A) ITC data obtained by our collaborators at the Dedon-lab in Singapore on ErmB with the 32-mer RNA. They used 30 µM ErmB with 400 µM 32-mer RNA. The top panel of the ITC results shows corrected heat rate curves and the bottom panel shows under-the-curve fits. A binding constant of 49nM between ErmB and the 32-mer RNA was determined from the data. B) Conversion of cofactor SAM into SAH by ErmB with an increasing 32-mer RNA concentration, highlighting the activity of ErmB on the 32-mer RNA substrate. C) Mass spectrometry data on specific site-modification by ErmB. The fragment ions shown in the mass spectrum prove that only the adenosine that is representative of A2058 is modified by ErmB. D) Predicted secondary structures of 32-mer RNA and RNApolymerase-18-mer RNA. The hairpin fold adapted by 32-mer RNA is highly similar to the fold seen in the predicted secondary structure of the RNApolymerase-18-mer RNA, explaining why precipitation was also seen for this RNA substrate.

Lastly, the collaborators performed two experiments on the complex to determine its binding and site-modification specificity. First, the modified 32-mer RNA was injected into a mass spectrometer to confirm that the mass for the adenosine ion that is modified has increased by the mass of two methyl groups. Only the adenosine fragment ion showed an increase of exactly two methyl groups, thus confirming site-specificity for ErmB modification on the 32-mer RNA (figure 4C). The second experiments entailed precipitation tests on ErmB to determine its specificity. In these tests ErmB was added to different RNA and DNA substrates in a 1:1.5 ratio, to ensure complete protein saturation by the RNA or DNA substrates. Table 1 lists the tests substrates and whether they showed precipitation.

Table 1: Description and results of the RNA and DNA substrates used by collaborators during precipitation tests

Substrate	Description	Precipitation
32-mer RNA	Positive control	Yes
32-mer DNA	Same as RNA, but DNA version	No
23S rRNA	23S subunit of *e. faecalis* OG1RF strain rRNA (weakly methylated by ErmB in the presence of SAM)	Yes
13-mer RNA	Random RNA sequence	No
Poly-CA RNA	5'-CCCAAAACCCCAAAACCC-3'	No
RNApolymerase-18-mer RNA	5'-GGGACUGUCAGGCGCGGC-3'	Yes
ErmB K111A DNA primer	DNA primer for the K111A point mutation in ErmB	No

The 32-mer RNA and the 23S rRNA from *E. faecalis* show precipitation when combined with ErmB, as expected (table 1). Interestingly, the RNApolymerase-18-mer RNA displays precipitation as well, even though this is not a natural substrate for ErmB. To further investigate this a secondary structure prediction on both the 32-mer and the RNApolymerase-18-mer were performed (figure 4D) [31]. When comparing the secondary structure predictions one can see the resemblance between the 32-mer and RNApolymerase-18-mer. The natural fold of the 23S rRNA subunit resembles that of an hairpin, thus named a hairpin loop. The literature reveals that this hairpin loop is a recognition element of Erm proteins and is required for the methylation activity [32]. The hairpin loop that is formed by base-pairing in the 18-mer closely resembles the hairpin loop in the 32-mer that is expected to bind ErmB, explaining why the 18-mer also showed precipitation.

4.2 Solution-NMR backbone assignment of ErmB

These results obtained from our collaborators enticed us to further investigate this interaction by applying ssNMR. Preliminary experiments showed us that the ssNMR spectrum of ErmB in complex with 32-mer RNA is highly similar to the solution NMR spectrum of apo-ErmB (unbound ErmB). Therefore, we decided to use solution NMR assignments to analyse the ssNMR probe of ErmB in the precipitate. ErmAM was already assigned in solution, however, the assignments were back then (1997) not deposited and had to be repeated for ErmB [21]. *De novo* backbone chemical shift assignments were obtained using five J-coupling based 3D solution NMR experiments (HNCα, HNCOCα, HNCO, HNCαCO, HNCβ) (figure 5A). The HNCα/HNCOCα and HNCO/HNCαCO sequential walks enabled backbone connectivity and HNCβ enabled easier amino acid identification (figure S2A/B). The five experiments enabled the assignment of 164 of the total 245 amino acids in the sequence of ErmB, corresponding to around 70%. Some amides (NH) within the backbone likely had fast exchange of protons with the used solvent. When this exchange is faster than the NMR time scale peaks get broadened, resulting in them getting lost in the noise. Therefore, complete assignment was not accomplished. Interestingly, the mostly unassigned segment of the sequence is thought to be involved in cofactor binding (figure 5B) [26]. A possible explanation for this could be that this region displays certain chemical exchange events that are close to the NMR experimental timescale. When this happens, different states are probed between measurements. The different states can cancel each other out and therefore the peaks get lost in the noise [38].

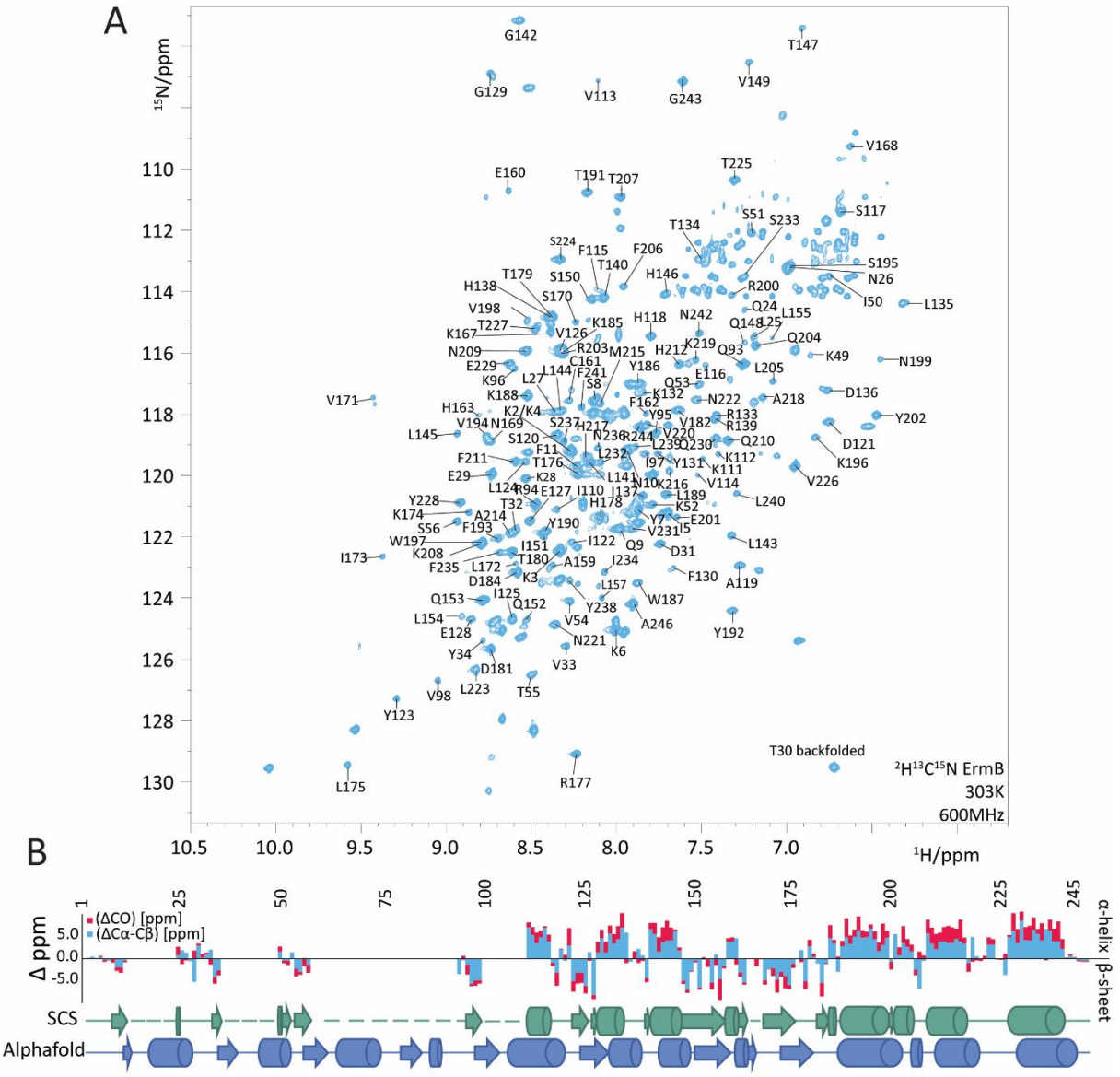


Figure 5: A) Solution-state backbone chemical shift assignments for apo-ErmB. B) Probability-based protein secondary structure identification based on Cα/Cβ and CO chemical shifts. Positive values correspond to α-helices, while, negative values correspond to β-sheets. When the secondary structure identification is compared against the secondary structures of ErmB predicted by Alphafold, a high similarity between the two is seen, confirming the validity of the Alphafold predicted structure and the backbone chemical shifts assignments.

The secondary structure elements of a folded protein can be inferred from the Cα/Cβ chemical shift values of different individual amino acids, due to their strong correlation [39]. The Cα/Cβ and CO chemical shifts of the individual amino acids were compared against the secondary structure elements predicted by the Alphafold model to check both the validity of the predicted structure, as well as the assignments (figure 5B). The probability-based protein secondary structure identification agrees well with the Alphafold homology driven predicted structure, further confirming both the validity of the assignment and the accuracy of the predicted structure (figure 5B). The backbone chemical shift assignments, and the Alphafold predicted structure, can now be used for further data interpretation.

4.3 Solution-NMR studies on the ErmB-SAM binding site

The cofactor, SAM, which donates the methyl group that is placed on residue A2058 has been suggested as a possible inhibitory site [18], but to our knowledge no Erm-inhibiting drug has been designed based on the cofactor binding site. The SAM binding site has already been identified by X-ray crystallography [22]. However, in this research investigating chemical shift perturbations (CSPs) upon the addition of the cofactor could be useful to identify some of the unassigned residues as coming from the cofactor binding region. Moreover, the extent of CSPs that are observable for ErmB was probed by addition of the cofactor. Therefore, a small titration series of SAH (the demethylated analogue of SAM) into an ErmB containing solution was performed. SAH had to be used instead of SAM, due to the instability of SAM in aqueous solutions. An overlay of the apo-ErmB spectrum and the ErmB+SAH (1:15) spectrum reveals the CSPs (figure 6A). Even though SAH was used instead of SAM for this titration, SAM showed similar CSPs when it was added in large excess (to prevent too much degradation) to ErmB (figure S3).

The residues that show CSPs above 0.1 ppm (figure 6B) are, as expected, located in or around motifs that have been identified to be important for cofactor binding [26]. Interestingly, two residues, Q24 and L25, show CSPs, although they are not located in, or near, any of the identified important motifs. However, these residues are located in the cofactor binding domain, and therefore, their chemical environment likely changes upon SAH binding. The CSPs of the assigned residues in the important cofactor binding motifs are as expected and confirm that, also for ErmB, the previously identified motifs X (S8, Q9, N10), I (Y34) and II (S56) are indeed involved in cofactor binding. However, CSPs are also seen for residues in motifs V (L112, E116), VI (D121, I122, Y123, L124, E127-F130) and VIII (A159, F162, N169), which are believed to be involved in the catalytical process. Since the reaction involves both the cofactor and the substrate RNA it makes sense that conformational changes occur in the catalytic region when the cofactor is bound, since they would need to be in close proximity. To compare the behaviour of the N- and C-terminal domain upon SAH binding the average CSP for residues in both domains were calculated. The N-terminal domain residues (left of red line, figure 6B) showed an average CSP of 0.06 ppm, while, the C-terminal domain residues (right of red line, figure 6B) showed an average CSP 0.02 ppm, highlighting the importance of the entire N-terminal domain on SAH binding. Lastly, particular unassigned peaks also showed significant shifts or complete disappearance upon SAH binding (figure 6A, peaks labelled 1-8). Their shifts indicate that these residues are involved in cofactor binding and that further validates the solution-state assignments.

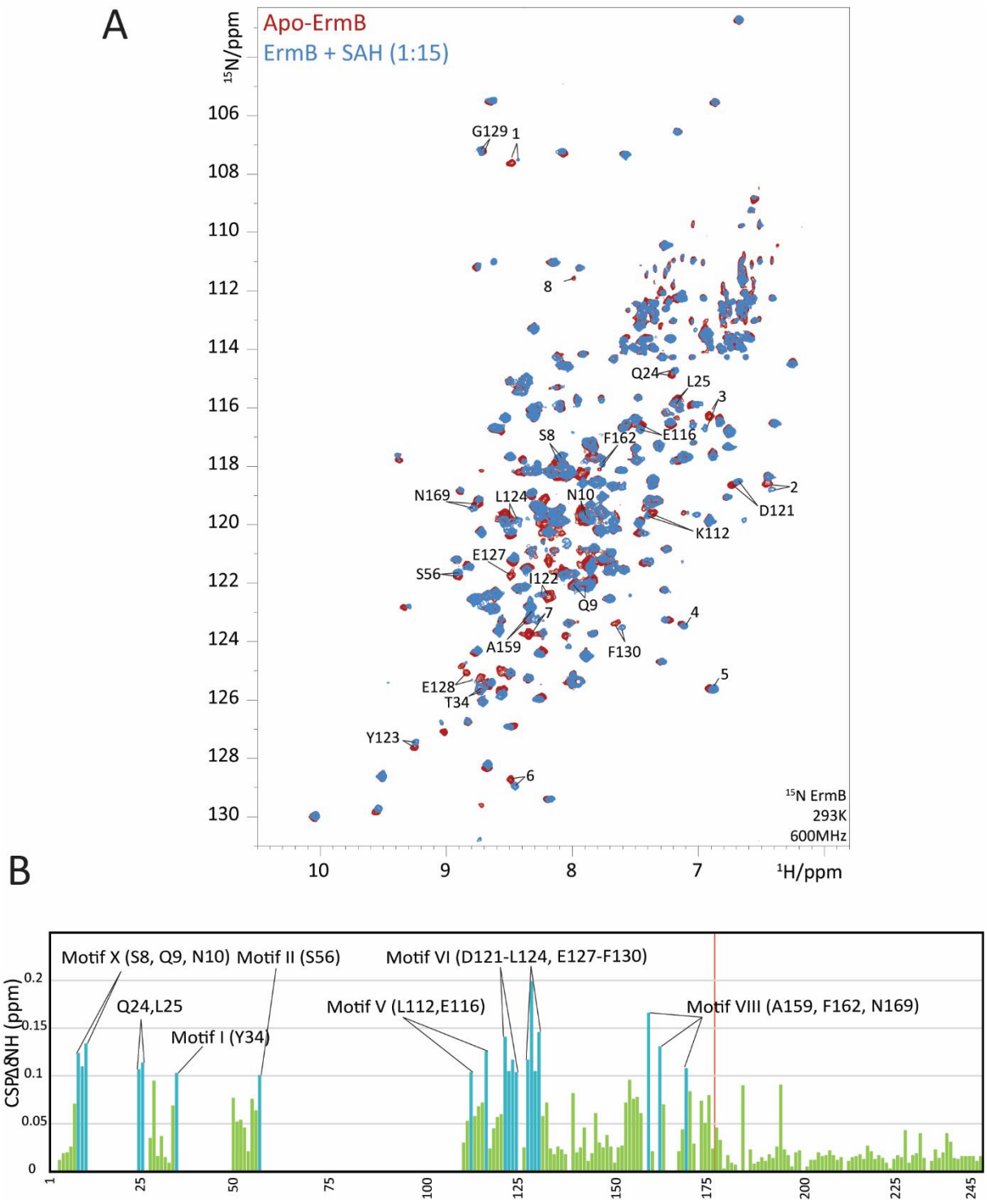


Figure 6: A) NH solution NMR spectrum of apo-ErmB (red) overlayed with the solution NMR spectrum of ErmB in complex with SAH (blue), the demethylated cofactor. The annotated peaks show significant CSPs upon addition of SAH to an ErmB containing solution. B) CSPs graph for all the assigned residues. A cut-off for significance of a CSP was put at 0.1 ppm difference. The residues that show CSPs above this cut-off are coloured light blue. Left of the red vertical line are residues located in the N-terminal domain, while residues on the right of the red vertical line are located in the C-terminal domain, highlighting the importance of the N-terminal domain for cofactor binding.

4.4 ssNMR studies of ErmB in the complex

After assigning the apo-state of ErmB, the precipitate was measured using Magic-Angle-Spinning (MAS) ssNMR. Upon addition of the 32-mer RNA to an ErmB containing solution, precipitate immediately formed. The precipitate was then transferred into a 1.3mm ssNMR rotor via centrifugation. A proton detected dipolar cross polarization-based 2D NH fingerprint spectrum was recorded at a magnetic field of 1200 MHz and a MAS frequency of 60 kHz (figure 7A). The spectrum is of high quality for ssNMR on a protein consisting of 245 amino acid and it seems that the protein establishes in a well-defined state (figure 7A). Moreover, a phase-alternated recoupling irradiation scheme (PARIS) experiment, which is a carbon-carbon correlation experiment based on proton-driven spin diffusion [40] obtained at 700 MHz with a MAS frequency of 42 kHz on ErmB reveals high spectral quality with good sensitivity and resolution (figure S7). The high spectral quality is indicative of the protein being in a well-defined state. Moreover, a well-defined state suggests that the complex is likely of specific nature.

The significant similarity and overlap between solution- and solid-state NH fingerprint spectra allowed us to transfer assignments by overlapping the solution-state spectrum with a solid-state spectrum. This enabled not having to redo the assignments in solid-state, which would not have been easily accessible. Conventional assignments could likely only be achieved by obtaining higher dimension spectra, which would require long measurement times and data processing. Moreover, the instability of the protein, that was observed after a week of MAS, makes measuring for longer periods difficult. With the peaks lined up a dipolar based 3D ssNMR experiment (CαNH) was used to validate the transferred assignments based on their Cα chemical shifts. Overlaying the spectra allowed us to transfer 28 assignments to the more isolated peaks in the ssNMR NH spectrum (figure 7A). Higher resolution assignment spectra are required to distinguish peaks in the crowded regions of the spectrum. Moreover, due to the broad line shape seen in ssNMR spectra even some of the more isolated peaks in the ssNMR spectrum can still consist of multiple residues when it is overlapped with the solution NH spectrum (figure 7C). Transferring assignments was more difficult due to this observation and makes interpretation of possible CSPs more challenging.

CSPs are highly sensitive to structural changes and report on changes in the chemical environment of that part of the protein [41]. If a certain residue displays CSPs, it is likely that this residue is located near, or is in contact with the RNA. However, due to the much broader peaks observed in ssNMR, we could only analyse signals in the isolated spectral regions using 2D ssNMR spectroscopy (figure 7A). Possibly interesting peaks from in or surrounding an identified motif, that could be interpreted as showing CSPs that are above the average CSPs of the transferred assignments, are: V113, S117, T147 and V168 (figure 7A/D). Furthermore, residue D136, located in the important loop12 [20], seems to completely disappear in the ssNMR spectra (figure 7A). This disappearance could be a result of binding with the RNA or due to the loop being highly mobile and therefore not being properly sampled by cross-polarization based experiments. Even though the mentioned peaks could be slightly shifted, most solution peaks still overlap with the solid peak, indicating weak or unspecific binding. Moreover, discrepancies between spectra obtained on different samples, with otherwise identical preparation conditions, argue against the formation of a specifically bound complex and make it impossible to reliably interpret CSPs (figure 7B). Therefore, other means to uncover a possible ErmB-rRNA binding site had to be explored.

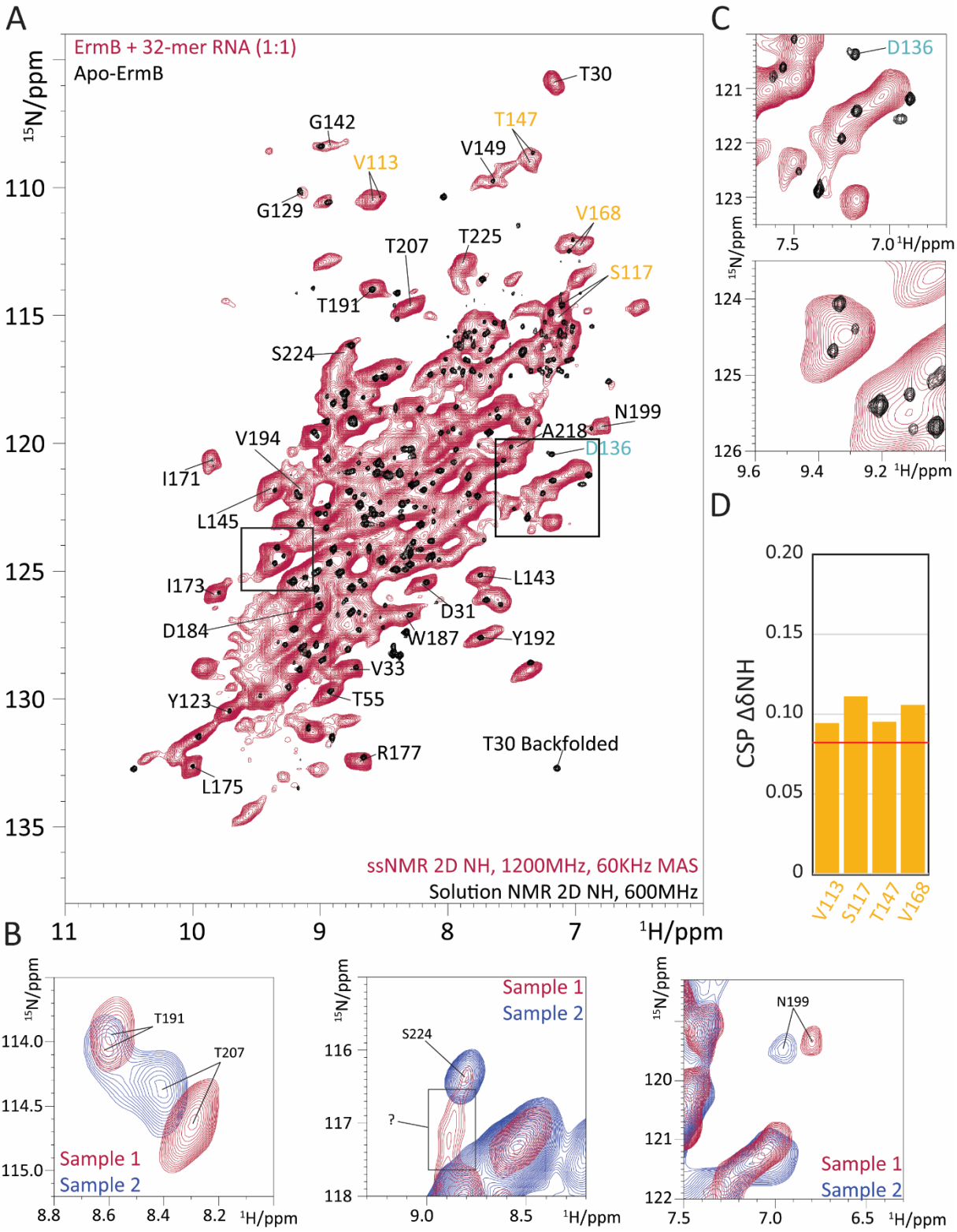


Figure 7: A) ssNMR NH fingerprint spectrum of ErmB + 32-mer RNA (dark red) overlayed with solution NMR NH fingerprint spectrum of apo-ErmB (black). CSPs were difficult to interpret due to the broad lines. However, peaks in important motifs and loops in the C-terminal domain that show possible CSPs are annotated in yellow. B) Discrepancies in the obtained spectra observed between similarly prepared samples measured under the same conditions, making interpretation of CSPs more difficult and indicating unspecific binding of 32-mer RNA to ErmB. C) Visualization of the more isolated peaks that are still covered by multiple solution NMR peaks. D) Possible CSPs quantified for residues in important motifs (yellow). The red line indicates the average CSP for all transferred residues.

4.5 ssNMR studies on 32-mer RNA in the complex

Another approach for obtaining structural data on the complex is to investigate the 32-mer RNA in the complex. ^{15}N and ^{13}C Labelled RNA is a requirement for NMR investigation and was provided by our collaborators at the Schwalbe-lab. Both fully ^{15}N, ^{13}C labelled and specifically ^{15}N, ^{13}C adenosine labelled RNA were measured. Adenosine labelling was chosen, since A2058 gets methylated by Erm. However, when measuring the adenosine labelled RNA it quickly became apparent that the RNA behaves differently compared to the protein. The same type of experiment was recorded for the 32-mer RNA; proton detected dipolar cross polarization based NH and CH spectra were recorded at a magnetic field of 1200 MHz and a MAS frequency of 60 kHz (figure 8A/B/C).

The sample showed poor sensitivity, which resulted in long measurement times. The NH spectrum of adenosine labelled 32-mer RNA was recorded in 27 hours (figure 8A/B), while the protein NH spectrum was recorded in approximately 10 hours, indicating the different behaviour of the protein and 32-mer RNA in the complex. Adenosine specifically labelled 32-mer RNA was measured in this spectrum, four different peaks are expected according to the overlayed NH solution Heteronuclear Single Quantum Coherence (HSQC) spectrum (figure 8A). However, a theoretical maximum of seven different peaks could be observed, since one NH moiety was sampled within this spectral range and seven adenosines are labelled (figure 8C). In the ssNMR NH spectrum seven peaks are sampled, therefore, the 32-mer RNA might undergo significant structural changes when forming a complex with ErmB, implied by the possible CSPs.

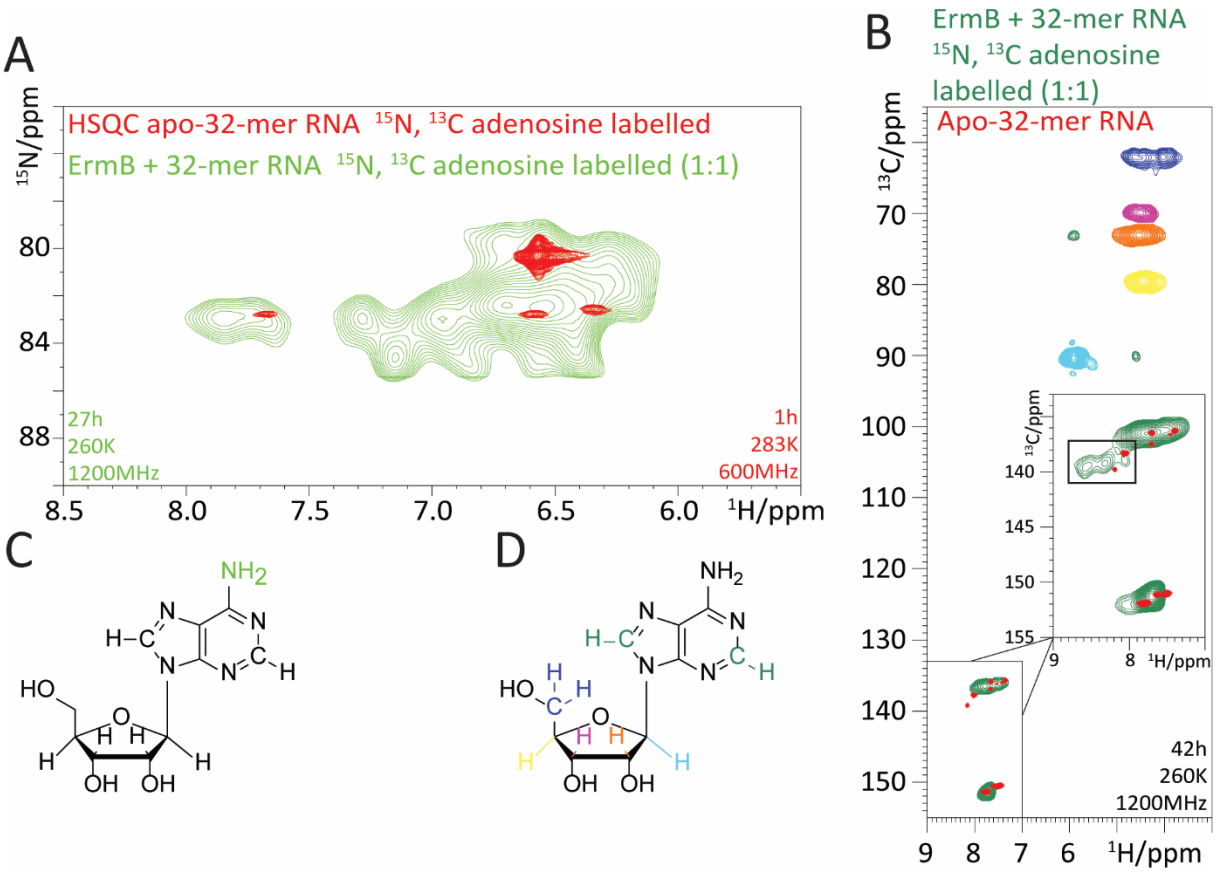


Figure 8: A) NH ssNMR spectrum of ^{15}N, ^{13}C adenosine labelled 32-mer RNA (light green) overlayed with solution NMR NH HSQC (red) B) CH ssNMR spectrum of ^{15}N, ^{13}C adenosine labelled 32-mer RNA (multi coloured according to C-H moieties shown in E) overlayed with solution NMR spectrum of the apo-32-mer RNA (red). The insert highlights possible CSPs, as well as the fact that solution-NMR can sample more states per adenosine. C) Adenosine ribonucleotide with the N-H moiety coloured (light green) that is sampled by the experiment shown in A. D) Adenosine ribonucleotide with the C-H moieties coloured (multi colour) according to the peak colour shown in B.

The CH spectrum highlights the same behaviour as was seen for the NH spectrum (figure 8B). However, some peaks can be attributed to different C-H moieties in the adenosine ribonucleotide based on the chemical shifts reported for a singular adenosine ribonucleotide [42] (figure 8B/D). A maximum of seven different states, per C-H moiety, could in theory be sampled, because there are seven different adenosines labelled in the 32-mer RNA. However, in the ssNMR spectrum it looks like each C-H moiety was only sampled once, even when the ssNMR spectrum was measured for a staggering 42 hours. Moreover, the broadness of the peaks likely results from slightly different chemical environments per adenosine nucleotide. The part of the ssNMR spectrum that corresponds to the adenosine aromatic C-H moieties is overlayed with the solution spectrum of the apo-32-mer RNA. In the solution-NMR spectrum the different states were sampled, since more peaks can be seen. CSPs were anticipated, since it is expected that the secondary structure of the 32-mer RNA would change upon binding with ErmB, especially surrounding the adenosine that mimics A2058. However, the overlap between solution- and solid-state reveals similar spectra. Only the outlined peaks could be interpreted as showing CSP's (figure 8B). Nevertheless, proper interpretation of the RNA ssNMR spectra and possible CSPs requires solution-state assignments, which still have to be provided by the collaborators, thus no further remarks can be made on the individual adenosines of the RNA.

Spectra obtained on the fully ^{15}N, ^{13}C labelled RNA, in complex with ErmB, demonstrated low sensitivity and resolution as well, making a comparison between solution- and solid-state spectra without assignments impossible. Consequently, dynamical data on both the protein side, as well as the 32-mer RNA side, had to be obtained to find indications of a possible binding site and to explain the bad spectral quality of the RNA in complex.

4.6 Dynamic studies on ErmB-32mer-RNA complex

Due to the lack of observed CSPs, discrepancies in the protein spectra, the poor spectral quality of the RNA spectra and to possibly indicate a ErmB-32-mer RNA binding site, dynamical data was obtained of the protein in complex, the 32-mer RNA in complex and the apo-protein. Dynamical data on the complex was obtained by performing ^{15}N slow rotating-frame relaxation ($R_{1\rho}$) experiments on samples with either labelled protein or RNA. In this type of experiment an relaxation element is added to the pulse scheme, and the duration of this relaxation element is incrementally extended in a series of 2D experiments. The signal intensity then scales according to the motion of the measured part of the complex. The decay in intensity can then be converted to relaxation rates, higher relaxation rates indicating increased dynamics. ^{15}N $R_{1\rho}$ relaxation is predominantly sensitive to motions in the microsecond range. Due to the relatively high sensitivity of the protein spectra, site-resolved dynamics could be probed with 2D $^{1}H^{15}N$ experiments that include the relaxation element. For proper integration it was crucial that only well resolved peaks are included in this analysis. The ^{15}N $R_{1\rho}$ for the protein in complex overall aligns with expectations; residues in secondary structure features (helixes or sheets) are overall more rigid (lower relaxation rates) compared to residues probed in loops (higher relaxation rates) (figure 9A/B). However, residues T207, S224 and T225, amino acids that reside in loops in the C-terminal domain of the protein, relaxed relatively slowly, indicating a rigid state (figure 9A/B). Slow relaxation could be indicative of these loops being bound to RNA, however, this needs to be compared with apo-ErmB dynamics data.

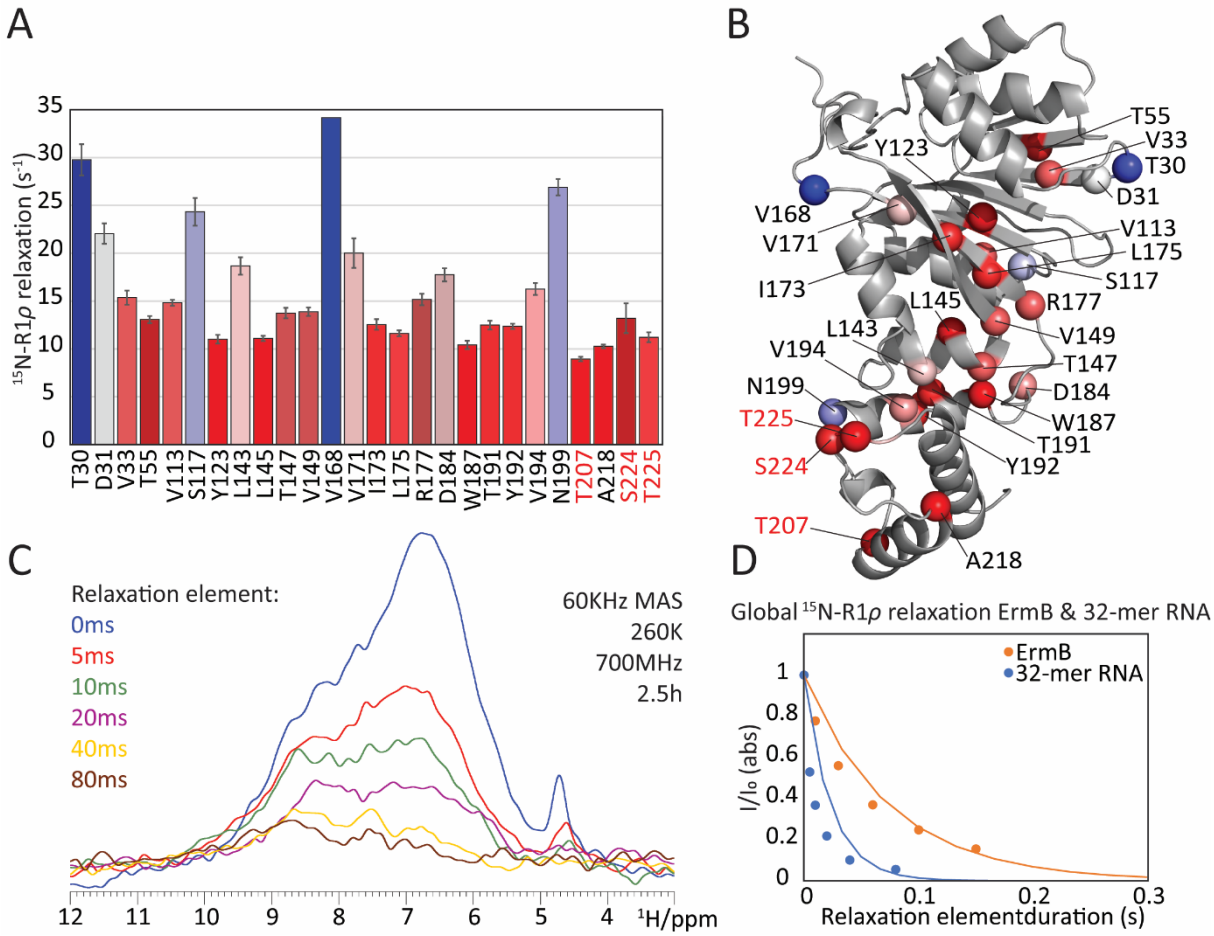


Figure 9: A) ^{15}N $R_{1\rho}$ relaxation data on ErmB in complex with 32-mer RNA. A lower relaxation rate ($R_{1\rho}$) means that the residue is more rigid, whilst, a higher $R_{1\rho}$ means that the residue is more mobile. In the Alphafold predicted structure in B) the residues are coloured according to their $R_{1\rho}$ values (red = rigid, blue = mobile). Most residues show expected $R_{1\rho}$ values, except for T207, S224 and T225, which are highly rigid, while being located in loops, which usually are more mobile. C) 1D ^{15}N-^{1}H $R_{1\rho}$ spectra with different relaxation element lengths reporting on the overall dynamics of the 32-mer RNA. The chosen relaxation lengths are already quite short compared to the ones chosen for ErmB, yet, the 32-mer RNA still shows fast decay of signal. D) A comparison between the overall $R_{1\rho}$ dynamics of ErmB (orange) and the 32-mer RNA (blue). From the steepness of the curves it quickly becomes apparent that the 32-mer RNA relaxes much faster compared to ErmB.

For the 32-mer RNA in the complex site-resolved dynamics were unachievable due to the low sensitivity. Therefore, the dynamics for the RNA could only be probed globally by looking at the overall decay in 1D ^{15}N-^{1}H experiments with an included relaxation element. Recording one 1D spectrum took 2.5 hours to record, highlighting the low sensitivity of the sample. The chosen delay introduced by the relaxation element ranged from 0ms to 80ms, which is quite short. Especially compared to the protein, in which the delay ranged from 0ms to 150ms. This already highlights that the RNA in complex relaxes much faster, which was confirmed by the quick disappearing of signal in the overlayed 1D spectra. Moreover, the signal was already almost completely disappeared when the relaxation delay was only 40ms (figure 9C). Global dynamics were also obtained for the protein by extracting 1D spectra from the 2D spectra obtained for site-specific ^{15}N $R_{1\rho}$ relaxation and looking at the bulk intensity loss. In figure 9D it can be seen that ErmB demonstrates much slower relaxation globally, compared to the 32-mer RNA in the complex. The fast relaxation of the RNA in the complex indicates that the 32-mer RNA is highly mobile, even while engaged in the complex. The fast relaxation conflicts with the ITC data that displayed a low binding constant and thus enthalpy driven binding. With this type of binding it would be expected that both parties are relatively rigid when bound to each other, which was not the case for the 32-mer RNA (figure 9D).

Lastly, dynamical data was also obtained on apo-ErmB using solution-state NMR. Pseudo 3D HSQC experiments were recorded, in which the loss in peak intensities in reports on the relaxation rate. The ^{15}N longitudinal relaxation rate (R_1), ^{15}N transverse relaxation rate (R_2) and the R_1 over R_2 values per residue are shown in figure 10A,B and C, respectively.

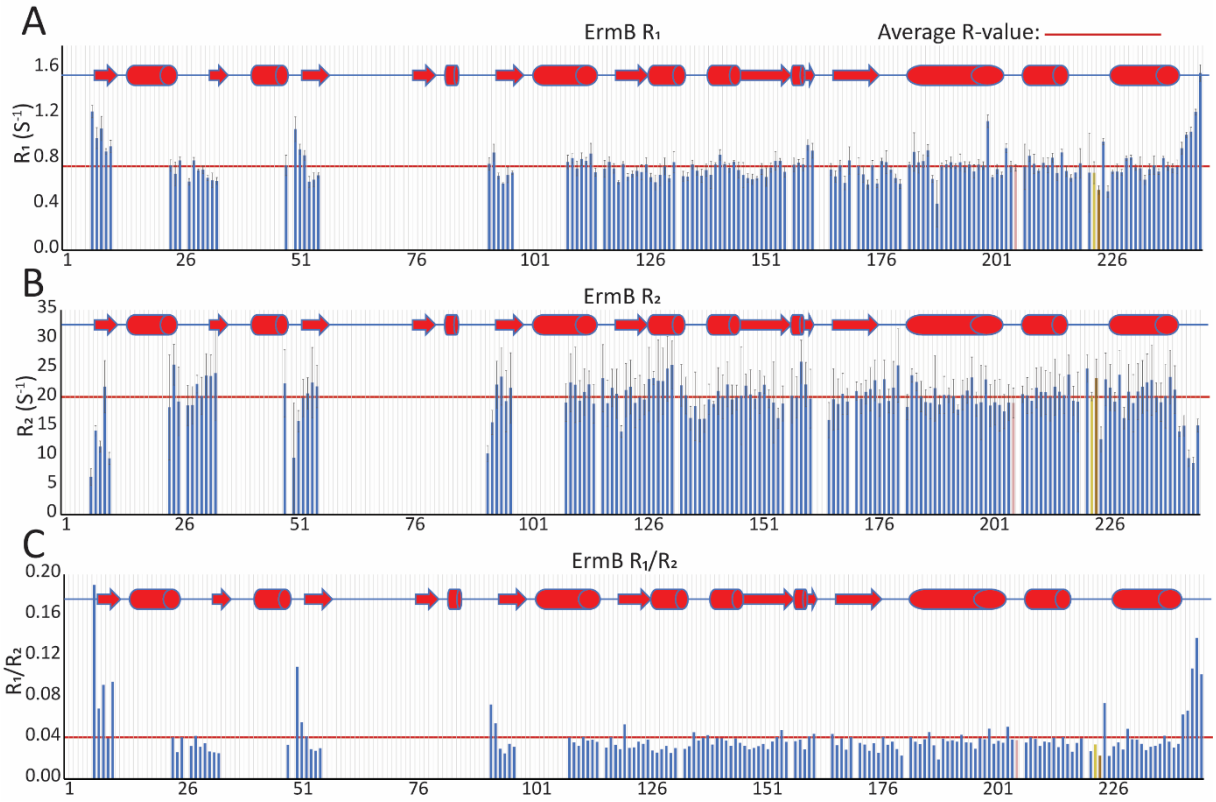


Figure 10: A) ^{15}N longitudinal relaxation rate (R_1), B) ^{15}N transverse relaxation rate (R_2) and C) R_1 over R_2 values determined per residue for apo-ErmB. Interesting residues from the ssNMR $R_{1\rho}$ relaxation studies are highlighted in pink (T207), yellow (S224) and brown (T225). The red line drawn through the graph is the average R value for that type of relaxation. The observed R-values for each amino acid overall agree with the secondary element within which they are located. Moreover, the presented data cannot argue for the interesting residues from ssNMR $R_{1\rho}$ relaxation studies becoming more rigid upon binding with the 32-mer RNA.

Overall all residues of apo-ErmB behave as expected; C- and N-terminal regions are more dynamic (higher R1, lower R2 and higher R1/R2) as are loops, while, structural elements and the core of the protein have R values around the average (figure 10A/B/C). A qualitative comparison between R_2 relaxation and $R_{1\rho}$ is only possible with the available data. If a R_2 relaxation rate for an individual residue is above the average for all residues it can be seen as having slower dynamics. When a R_2 relaxation rate for an individual residue is below the average for all residues it can be seen as having faster dynamics. As stated above, residues T207, S224 and T225 showed reduced mobility when in complex with 32-mer RNA (figure 9A/B). In the apo-state residues T207, S224 and T225 reveal average mobility, since all three are close to the average (figure 10B). Moreover, the error bars are large, making reliable interpretation more difficult. Therefore, it cannot be confidently concluded that residues T207, S224 and T225 become more rigid upon complex formation. The presented data raises the following question: is the precipitate that forms actually representative of the Erm-rRNA interaction?

5 Discussion and outlook

The fight against AMR, the phenomenon in which microorganisms become insusceptible to previously effective drugs [1], is in dire need of alternative strategies that prevent further development. Since, in 2019 nearly 5 million deaths were attributable to AMR [2], and it is believed that this number will increase to 50 million deaths annually by 2050 [3]. The alternative strategy investigated within this research thesis focusses on inhibition of development of AMR, for clinically relevant antibiotics: macrolides, lincosamides and streptogramins, or MLS antibiotics. The MLS antibiotics are often used to treat bacterial infections from *Streptococcus pneumoniae* and *Staphylococcus aureus* [8], highlighting their clinical relevance. Bacteria acquire resistance via target modification (figure 1D), by placing two methyl groups on residue A2058 in the 23S rRNA subunit of the bacterial ribosome; the binding site for MLS antibiotics [14] (figure 1E/F). Target modification is performed by Erm proteins, a family of structurally-conserved enzymes. In order to inhibit the development of AMR for MLS antibiotics a structural understanding of the Erm-rRNA interaction is a requirement. For a more accessible investigation into the Erm proteins and their interaction with rRNA Vester et al. (1998) progressively truncated the part of the domain of the 23S rRNA subunit in which nucleotide A2058 was located and measured whether Erm proteins would still modify the rRNA [32]. They found that they could cut domain V (625 nucleotides) down to just 32 nucleotides and Erm could still recognize and methylate it. Therefore, it was believed that the precipitate that forms between Erm proteins and small stretches derived from the original rRNA domain are representative of the complex and were used here for a structural investigation using ssNMR. However, the results presented within this research thesis are not necessarily indicative of specifically bound complex between ErmB and 32-mer RNA.

5.1 Solution NMR allows assignment of ErmB and CSPs confirm cofactor binding site

Before discussing the implications of the results obtained during the structural investigation on the ErmB-32-mer RNA complex using ssNMR, the solution NMR data will be discussed. Using a set of five 3D solution NMR experiments nearly 70% of all 245 amino acids could be assigned. The main missing stretch of amino acids (figure 5B) corresponds to the cofactor binding domain. A possible explanation for why specifically these amino acids are missing in the assignments of apo-ErmB is that this region of the protein has a certain mobility due to it needing to adapt to fit the cofactor and bring it close to the rRNA. Due to the mobility and chemical exchange some resonances might get sampled in two different conformations whose chemical shift cancel each other out, resulting in the peak getting lost in the noise. This agrees with findings during a structural investigation on apo-ErmAM, another member of the Erm methyltransferase family that is highly similar to ErmB, using solution NMR, in which they also saw exchange broadening in the absence of the cofactor [21]. Furthermore, upon addition of the demethylated cofactor (SAH) significant CSPs were seen for unassigned peaks, indicating that they are representative of amino acids in the cofactor binding domain. The observed CSPs upon addition of SAH confirm not only the solution NMR assignments, but also that the expressed ErmB protein has the correct fold and is representative of Erm methyltransferases. While the observed CSPs are significant, and thus indicative of SAH binding, the maximum CSPs are around 0.2 ppm, which would be too low to distinguish in ssNMR spectra due to the intrinsic problem of broad line shapes. This makes possible CSPs interpretation in ssNMR spectra a lot harder for ErmB.

5.2 ssNMR reveals first atomic insights on ErmB-32-mer RNA complex

Next, the structural investigation on the ErmB-32-mer RNA complex will be discussed. The initial tests on the complex to determine its binding constant, activity and specificity performed by our collaborators highlight the formation of a specifically bound complex. Therefore, it was expected to see significant CSPs when measuring ErmB in complex with the 32-mer RNA. However, the solution NH fingerprint spectrum of ErmB overlaps well with the solid NH fingerprint spectrum of ErmB in complex with the 32-mer RNA, which indicates the exact opposite of the low binding constant; a non-specifically bound 32-mer RNA to the protein (figure 7A). The unspecific binding of the 32-mer RNA to ErmB is further substantiated by the inconsistency shown between different samples that should result in the same NH fingerprint spectra (figure 7B). The only small difference between the two measured samples shown in figure 7B is the fact that for sample 1 ^{15}N, ^{13}C adenosine labelled RNA was used and for sample 2 unlabelled RNA. However, there should be no difference in binding and therefore this does not explain the observed discrepancies between the NH spectra, thus indicating unspecific binding of 32-mer RNA to ErmB.

Another compelling argument for the formation of an unspecifically bound complex between 32-mer RNA and ErmB is the behaviour of the 32-mer RNA when it is in complex with ErmB. The ssNMR spectra that were obtained on ^{15}N, ^{13}C (adenosine) labelled RNA show surprisingly low sensitivity and resolution (figure 8A/B) , even with long measurement times compared to ErmB. This makes possible CSP interpretation challenging, however, more peaks were sampled in the solid-state for the NH spectrum (figure 8A), while less peaks were sampled in the CH spectrum, compared to solution spectra. For proper interpretation of these spectra the solution state assignments of the 32-mer RNA are a requirement and these are yet to be provided by our collaborators at the Schwalbe-lab in Frankfurt.

Even though the NMR results indicate formation of an unspecifically bound complex between ErmB and the 32-mer RNA, the preliminary tests performed by our collaborators indicate otherwise. The methylation activity of ErmB on the 32-mer RNA was confirmed by screening SAH concentrations when ErmB and 32-mer RNA were combined in the presence of the cofactor SAM (figure 4B). Furthermore, the specificity of methylation was confirmed by mass spectrometry (figure 4C) and the performed precipitation tests highlight that ErmB only binds RNA with a specific fold that mimics the natural fold of the rRNA in the 23S subunit of the bacterial ribosome. A possible explanation for these findings by our collaborators could be that the high net positive charge of ErmB (+20), unspecifically interacts with the negatively charged phosphate groups in the backbone of the 32-mer RNA (figure 11A). The high net positive charge of the protein was determined by calculating the titration curve and interpreting it at a pH of 6.5, the used pH of the buffer in which the complex was formed for ssNMR studies (figure 11B). The methylation activity that is seen can come from the 32-mer RNA being in the correctly bound mode, making the methyl group transfer possible, while in other conformations it is not. An unspecifically bound complex could also explain why precipitation was also seen for the RNApolymerase-18-mer RNA (table 1, figure 4D). The surface of the predicted structure coloured according to the charge calculated for that region reveals a lot of positive regions, especially in the C-terminal domain and the deemed important loop12 [20]. It is possible that these residues are forming unspecific interactions with the phosphate groups in the backbone of the 32-mer RNA, but the hairpin fold of the 32-mer RNA, or alike, is clearly still a necessity for it to bind, even if it is unspecific.

Efforts have been made to solubilize the complex by increasing the salt concentration to lessen the effect of electrostatic interactions. Increasing the salt concentration to 500 mM NaCl resulted in no precipitation when the 32-mer RNA was added to a buffer containing ErmB, even when going to stoichiometries of 1:1.5. However, no significant CSPs could be seen in the spectra (figure S5). The high salt concentration likely perturbs the 32-mer RNA's fold, which causes ErmB to not bind (unspecifically) and, therefore, no precipitation was observed. An interesting investigation could be to alter the pH of the used buffer in order to significantly change the net charge of the protein and therefore change its interaction capabilities with the 32-mer RNA. Beforehand, it is crucial to investigate whether ErmB and the 32-mer RNA retain the same fold when in the different pH; this can be checked by obtaining solution NMR spectra and observing whether there are major differences in peak distribution.

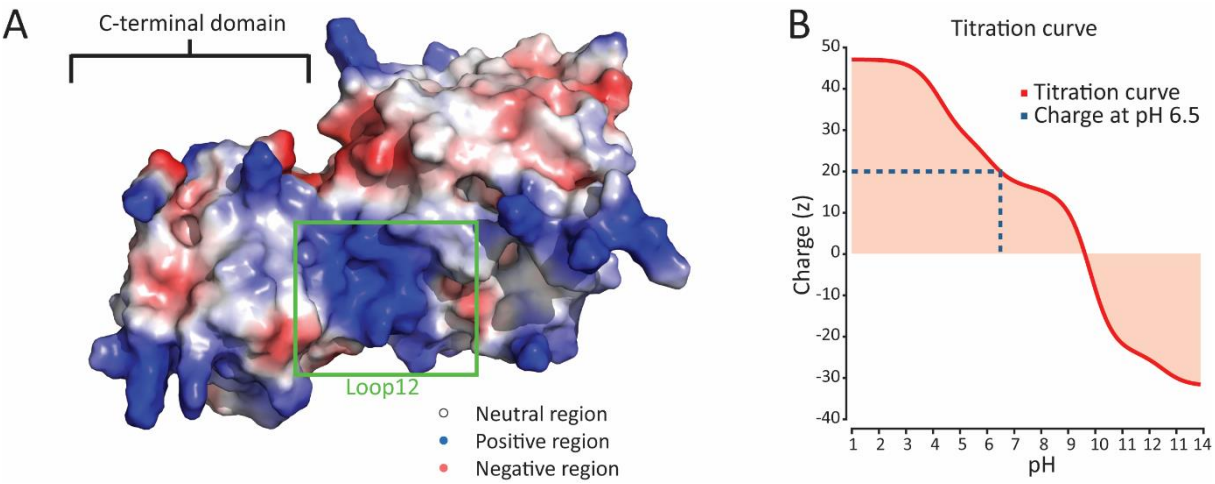


Figure 11: A) The surface of the Alphafold predicted structure of ErmB coloured according to the charge calculated for that region. A high count of positively charged regions (blue) can be seen on the surface of ErmB, especially in the C-terminal domain and the important loop12 [20]. It is theorized that RNA, with the correct hairpin fold, can unspecifically interact with the positively charged residues via negatively charged phosphate groups in the backbone. B) Calculated titration curve via protpi.ch used to interpret the net charge of the protein at the pH (6.5) of the used buffer to form the precipitation.

5.3 NMR Dynamics data explains bad sensitivity of 32-mer RNA in complex with ErmB

The ^{15}N $R_{1\rho}$ dynamics data further backs up the claim that an unspecifically bound complex forms between ErmB and the 32-mer RNA. While the data shown in figure 9A are mostly showing values as expected: residues in α-helices and β-sheets are more rigid, and residues in loops are more flexible. However, as mentioned above, residues T207, S224 and T225 are more rigid than expected, which could mean that these residues are engaged in hydrogen bonds with nucleotides from the RNA. Nevertheless, when comparing the dynamics of these three residues to the apo-state of the protein (figure 10A/B/C) it quickly becomes apparent that that conclusion cannot be confidently made, since in the apo-state the dynamics of residues T207, S224 and T225 are matched with the average R_2 dynamics of the protein and the error bars are large (figure 10B). This makes it difficult to conclude that either of these three residues show reduced mobility when engaged in binding with the 32-mer RNA.

In contrast, the ^{15}N $R_{1\rho}$ dynamics data, from RNA, are showing very high mobility (figure 9C), explaining the bad sensitivity of the sample. Nonetheless, for a specifically bound ligand it would be expected that both molecules are overall more rigid when engaged in complex, thus, further backing up the claim that the formed precipitate comes from unspecific binding between ErmB and the 32-mer RNA.

5.4 Conclusions and future perspective

The results presented in this research thesis are the first hard structural data on the ErmB-32-mer RNA complex, which was enabled by applying ssNMR techniques on the precipitate that forms when ErmB and 32-mer RNA are combined. For a long time it has been assumed that the formed precipitate is representative of the actual Erm-rRNA interaction when Erm modifies residue A2058 in the 23S rRNA subunit of the bacterial ribosome. However, the data presented here could allude to the contrary. The absent CSPs in the ssNMR NH spectrum when it is compared to the solution NH spectrum, discrepancies in the ssNMR spectra on the protein and the unexpected behaviour of the RNA lead to the conclusion that the interaction between ErmB and the 32-mer RNA could be of unspecific nature. The precipitate that forms could come from unspecific interactions between positively charged moieties of ErmB with the negatively charged phosphate backbone of the 32-mer RNA and a RNA secondary structure that closely resembles the natural fold of the 23S rRNA subunit.

However, to completely confirm unspecific binding the RNA side of the complex has to be further investigated. In this thesis the RNA side was only looked at shortly. Where possible, the assignments of the apo-32-mer RNA are a necessity for proper interpretation of the obtained spectra, but, these are yet to be provided by our collaborators. Moreover, the intrinsic difficulties that come with proton-detected ssNMR on such a large protein (30kDa, 245 amino acids) will always make this investigation by ssNMR difficult. A possible solution could be to specifically label certain amino acids of ErmB, this reduces spectral crowding and makes interpretation of the spectra more accessible. It is crucial to pick the correct amino acids to label, since, only these residues will experience different chemical environments when engaged in the complex. Another approach for acquiring structural data on the ErmB-rRNA interaction could be cryo-EM. The ErmB-32-mer RNA is believed to be too small for a cryo-EM structural investigation. However, when increasing the size of the RNA to for instance; domain V of the 23S rRNA subunit of the bacterial ribosome, which is the largest RNA substrate that Erm can still methylate [32], cryo-EM might be able to elucidate structural information on it. Moreover, KsgA, a non-pathogenic bacterial methyltransferase, in complex with the 30S ribosomal subunit has been characterized structurally by cryo-EM [43]. Therefore, applying cryo-EM for elucidating the ErmB-rRNA binding interface could be a possibility. Additionally, more structural data is needed to elucidate whether the formed precipitate is indeed representative of the ErmB-rRNA binding interface. Likewise, more conclusive structural information is a necessity to develop specific inhibitors of the Erm-rRNA interaction.

6 Materials and Methods

6.1 ErmB expression and purification

The *ErmB* gene was cloned into a pNIC28-CH vector with a C-terminal His6 tag and transformed into *E. Coli* BL21(DE3)Star (Invitrogen) competent cells. The amino acid sequence of ErmB can be found in figure S6A. Perdeuterated (U-^{2}H, ^{13}C, ^{15}N) ErmB for NMR experiments was expressed in cytoplasmic inclusion bodies by growing the cells in 100% D_2O M9 minimal medium, containing 0.5 g/L of $^{15}NH_4Cl$, 2 g/L of D-glucose-$^{13}C_6D_7$ and 50 µg/ml of kanamycin. (U-^{15}N) ErmB for NMR experiments was expressed in cytoplasmic inclusion bodies by growing the cells in 100% H_2O M9 minimal medium, containing 0.5 g/L of $^{15}NH_4Cl$, 2 g/L of glucose-C_6H_7 and 50 µg/ml of kanamycin. Unlabelled ErmB for RNA ssNMR experiments was expressed in cytoplasmic inclusion bodies by growing the cells in LB medium, containing 0.5 g/L of NH_4Cl, 2 g/L of glucose-C_6H_7 and 50 µg/ml of kanamycin. Cells were grown to an OD_{600} of 0.9 (37°C, 250 rpm shake) at which protein expression was induced with 0.5 mM IPTG, after 15 min cold shock at 4°C. After 17 h of expression, cells were harvested (4000 rpm, 20 min, 4°C) and resuspended in PBS buffer containing lysozyme and protease inhibitor cocktail, then lysed by freeze-thaw, French press and sonication cycles in an ice bath. Inclusion bodies were isolated by centrifugation (15.000 rpm, 30 min, 4°C) and solubilized with the unfolding buffer (6 M GdmCl, 50 mM Tris, 100 mM NaCl, 1 mM βME, pH 7.5). The soluble fraction containing unfolded ErmB was loaded on a 5 mL-HisTrap FF column pre-equilibrated with unfolding buffer and eluted with a gradient reaching 500 mM imidazole in unfolding buffer (expression and purification yield ~40 mg/L of culture). The purified ErmB was refolded by fast dilution with refolding buffer (50 mM sodium phosphate, 100 mM NaCl, 1 mM βME, pH 6.5) using 10-kDa molecular weight cut-off (MWCO) Amicon Ultra-15 Centrifugal Filter Unit (Merck Millipore) carrying out 8 x 10-fold dilutions at 4°C, starting from a concentration of ErmB of 1.5 mg/mL (refolding yield ~50%). ErmB stocks in refolding buffer were protected from bacterial growth with 0.01 $\%_{V/V}$ NaN_3, flash frozen and stored at -20°C.

6.2 Solution NMR

Solution NMR samples were prepared by thawing sufficient ErmB stocks for the wanted concentration, the stocks were concentrated to the correct volume. Solution NMR spectra on uniformly ^{2}H, ^{15}N, ^{13}C (U-^{2}H, ^{13}C, ^{15}N) labelled ErmB or uniformly ^{15}N (U-^{15}N) labelled ErmB were obtained on either a Bruker Avance III 600 MHz or Avance Neo 900 MHz spectrometer. Both spectrometers are equipped with a cryoprobe to increase signal-to-noise ratios. 1-Dimensional, 2-dimensional and 3-dimensional experiments were all processed using Bruker Topspin 4.1.3. While, Collaborative Computational Project for NMR V3.1.0 (CCPnmr) was used to analyse the spectra [44].

6.2.1 Backbone assignments ErmB

A 5mm glass NMR tube (Wilmar) was filled with (U-^{2}H, ^{13}C, ^{15}N) labelled ErmB (150 µM), 0.01% (v/v) NaN_3 to prevent bacterial growth and 10% (v/v) D_2O for the lock signal. Backbone assignment of the NH fingerprint spectra of (U-^{2}H, ^{13}C, ^{15}N) labelled ErmB was obtained by recording five 3D J-coupling based experiments at T_{set} 303K: HNCα [45], HNCOCα [45], HNCO [45], HNCαCO [46] and HNCβ [47]. The HNCα/HNCOCα and HNCO/HNCαCO pairs were used for the sequential backbone walk, while, the HNCβ was used to help identify individual amino acids. Deuterium decoupling was applied to all the experiments and the Watergate pulse sequence was used for water suppression.

6.2.2 Probability-based protein secondary structure identification

Chemical shifts of amino acids are strongly correlated to the secondary structure features of the protein. Both Cα/Cβ and CO chemical shifts can be used to calculate the secondary structure propensity (P) of individual amino acids. The following equations can be used to calculate the P values:

$$P_{C\alpha/C\beta} = \left(C\alpha_{experimental} - Ca_{random\ coil\ value}\right) - \left(C\beta_{experimental} - C\beta_{random\ coil\ value}\right)$$

$$P_{CO} = \left(CO_{experimental} - CO_{random\ coil\ value}\right)$$

$P_{C\alpha/C\beta}$ or P_{CO} is the likelihood that this amino acid is located in a α-helix, β-strand or random coil. The $C\alpha_{experimental}$, $C\beta_{experimental}$ or $CO_{experimental}$ values are determined from the 3D experiments and the random coil values are average values that were reported for all amino acids [39].

6.2.3 Chemical Shift Perturbations (CSPs)

CSPs upon binding of SAH for (U-^{15}N) labelled ErmB were determined by first recording a NH fingerprint spectrum of apo-ErmB, next adding 1:3 ratio SAH and lastly 1:15 ratio SAH, to ensure that peaks could be followed correctly. CCPnmr was used to quickly interpret CSPs for every amino acid [48], which applied the following formula to correct for the larger chemical shift dispersion of ^{15}N:

$$CSP = \sqrt{\left(\frac{\Delta N}{6.5}\right)^2 + (\Delta H)^2}$$

The division factor 6.5 was introduced for all amino acids to account for the larger chemical shift dispersion of ^{15}N [41].

6.2.3 ^{15}N T_1 and T_2 relaxation

For solution-NMR experiments T_1 (longitudinal relaxation) and T_2 (transverse relaxation) experiments were measured to probe fast and intermediate motion. Site-specific ^{15}N T_1 and T_2 relaxation data on (U-^{2}H, ^{13}C, ^{15}N) labelled apo-ErmB was probed by obtaining pseudo 3D Heteronuclear Single Quantum Coherence (HSQC) spectra on a 600 MHz spectrometer. In these spectra the T_1 or T_2 relaxation is encoded in the third dimension by varying peak intensities (dependent on the dynamics of that residue) in each HSQC. 2D ^{1}H^{15}N spectra were then processed and extracted from the pseudo 3D spectrum using Bruker Topspin 4.1.3. In the T_1 pseudo 3D experiment the following relaxation delays were applied: 10, 200, 400, 600, 800, 1000 and 1400 ms, with 400, 600 and 1000 ms measured as duplicates. In the T_2 pseudo 3D experiment the following relaxation delays were applied: 16, 33, 48, 64, 80 and 110 ms, with 48 and 110 ms measured as duplicates. The extracted 2D ^{1}H^{15}N spectra were then integrated using PINT version 2.1.0 [49]. The line shape of peaks in the 2D ^{1}H^{15}N spectra from the T_1 experiment were fitted to a Gaussian shape. While, the line shape of peaks in the 2D ^{1}H^{15}N spectra from the T2 experiment were fitted to a Galore shape. In both experiments the integrated peak volumes were optimized and fit to a decaying exponential function of relaxation delay by PINT.

6.3 Solid-State NMR

ssNMR spectra on either unlabelled or (U-^{2}H, ^{13}C, ^{15}N) labelled ErmB in complex with either unlabelled 32-mer RNA, uniformly ^{15}N, ^{13}C labelled 32-mer RNA or ^{15}N, ^{13}C-adenosine only labelled 32-mer RNA were performed on either a Bruker 700 MHz AVANCE III or Bruker 1200 MHz AVANCE NEO spectrometer. All ^{1}H-detected ssNMR experiments were performed at 60 kHz MAS spinning frequency. Solution NH assignments were transferred to solid state by overlaying the spectra and using a 3D-CαNH experiment based on N-Cα specific-CP transfer for confirmation [50]. The CαNH was recorded at 700 MHz with a real sample temperature of 293 K. Low power PISSARRO decoupling scheme was applied in all dimensions setting the decoupling amplitude at 15 kHz i.e. at 1/4 of MAS [51].

6.3.1 ssNMR sample preparation

Formation of the complex was achieved by adding the 32-mer RNA, obtained either commercially from Sigma Aldrich (unlabelled) or via our collaborators (^{15}N,^{13}C (adenosine) labelled) to an ErmB containing solution (1.1:1). The nucleotide sequence of the 32-mer RNA can be found in figure S6B. The complex was measured using a 1.3mm solid-state NMR rotor, to sufficiently fill the rotor about 60 nmol of each were added together in solution. The complex would almost immediately precipitate and was incubated for 2 hours at 4°C. The solids were harvested by ultracentrifugation (40.000 xg, 15 min, 4°C) and placed in a 1.3mm zirconia ssNMR rotor using the appropriate tools.

6.3.2 ^{15}N $T_{1\rho}$ relaxation on ErmB in the complex

For ssNMR relaxation experiments $T_{1\rho}$, or decay of magnetization along the applied spin lock pulse in the rotating frame, was measured which reports on the µs dynamics of the system. Site-specific relaxation data was available for isolated signals in the NH fingerprint ssNMR spectrum of (U-^{2}H, ^{13}C, ^{15}N) labelled ErmB. Relaxation was probed by 2D ^{1}H-^{15}N experiments acquired at 1200 MHz, using a ^{15}N spin-lock field of 19 kHz and relaxation increments of 0, 10, 30, 60, 100 and 150 ms. Peak intensities were normalized and used to estimate $T_{1\rho}$ using the following equation:

$$i = e^{\left(\frac{-\tau}{T_{1\rho}}\right)}$$

Where i = peak intensity and τ = duration relaxation element. $T_{1\rho}$ was then converted into an relaxation rate by applying the following equation:

$$R_{1\rho} = \frac{1}{T_{1\rho}}$$

Only isolated peaks could be investigated, while per peak the processing parameters had to be exactly the same to be able to compare peak intensities. Peak intensities were extracted using Bruker Topspin 4.1.3. $T_{1\rho}$ and $R_{1\rho}$ were calculated using Graphpad Prism version 9.4.1 for Windows, with points fitted using a non-linear fit.

6.3.3 Bulk $T_{1\rho}$ relaxation on RNA in the complex

Site-specific relaxation data was unachievable for the (^{13}C,^{15}N Adenine)-32mer-RNA in complex due to the low sample sensitivity. Relaxation was thus probed by 1D ^{1}H-^{15}N experiments at 700MHz, using a ^{15}N spin-lock field of 20kHz and spin-lock durations of 0, 5, 10, 20, 40, and 80 ms. All spectra were processed equally and peak intensities were extracted at 6.7 ppm (highest point of the peak, figure 9C). Intensities were normalized and the above mentioned equations were used, as well as Graphpad Prism version 9.4.1.

7 Acknowledgements

First of all, I would like to thank Markus Weingarth for providing me with the opportunity to do an internship in the Weingarth-lab. His insights and expertise in ssNMR helped me tremendously throughout my internship. Secondly, I would like to thank Francesca Lavore for her daily supervision, her general guidance and our fruitful discussions in the laboratory and throughout this internship. Thirdly, I want to thank Hugo van Ingen for helping us set up solution NMR experiments. Lastly, the Dedon-lab is thanked for performing the initial experiments on the ErmB-32-mer RNA interaction and the Schwalbe-lab is thanked for providing us with labelled 32-mer RNA.

8 References

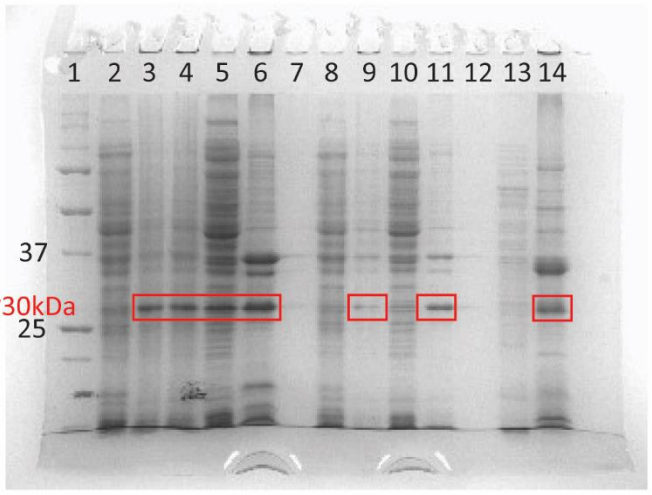
1. Morrison, L. and T.R. Zembower, *Antimicrobial Resistance.* Gastrointestinal Endoscopy Clinics of North America, 2020. **30**(4): p. 619-635.
2. Murray, C.J., et al., *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* The Lancet, 2022. **399**(10325): p. 629-655.
3. O'Neill, J., *Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations.* Review on Antimicrobial Resistance, 2014.
4. (WHO), W.H.O., *2021 Antibacterial Agents in Clinical and Preclinical Development: an overview and analysis*. 2022.
5. O'Neill, J., *Tackling drug-resistant infections globally: final report and recommendations*. 2016: Government of the United Kingdom.
6. Baym, M., L.K. Stone, and R. Kishony, *Multidrug evolutionary strategies to reverse antibiotic resistance.* Science, 2016. **351**(6268).
7. Abdul-Aziz, M., et al., *Applying Pharmacokinetic/Pharmacodynamic Principles in Critically Ill Patients: Optimizing Efficacy and Reducing Resistance Development.* Seminars in Respiratory and Critical Care Medicine, 2015. **36**(01): p. 136-153.
8. Golkar, T., M. Zielinski, and A.M. Berghuis, *Look and Outlook on Enzyme-Mediated Macrolide Resistance.* Frontiers in Microbiology, 2018. **9**.
9. Tsui, W.H., et al., *Dual effects of MLS antibiotics: transcriptional modulation and interactions on the ribosome.* Chem Biol, 2004. **11**(9): p. 1307-16.
10. Duval, R.E., M. Grare, and B. Demore, *Fight Against Antimicrobial Resistance: We Always Need New Antibacterials but for Right Bacteria.* Molecules, 2019. **24**(17).
11. Tenson, T., M. Lovmar, and M. Ehrenberg, *The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome.* J Mol Biol, 2003. **330**(5): p. 1005-14.
12. Claudio O. Gualerzi, L.B., Attilio Fabbretti, and Cynthia L. Pon, *Antibiotics Targets, Mechanisms and Resistance*. 2014, Weinheim: Wiley-VCH.
13. Wilson, D.N., *Ribosome-targeting antibiotics and mechanisms of bacterial resistance.* Nature Reviews Microbiology, 2014. **12**(1): p. 35-48.
14. Svetlov, M.S., et al., *Structure of Erm-modified 70S ribosome reveals the mechanism of macrolide resistance.* Nature Chemical Biology, 2021. **17**(4): p. 412-+.
15. Saribas, Z., F. Tunckanat, and A. Pinar, *Prevalence of erm genes encoding macrolide-lincosamide-streptogramin (MLS) resistance among clinical isolates of Staphylococcus aureus in a Turkish university hospital.* Clin Microbiol Infect, 2006. **12**(8): p. 797-9.
16. Okitsu, N., et al., *Characterization of ermB gene transposition by Tn1545 and Tn917 in macrolide-resistant Streptococcus pneumoniae isolates.* J Clin Microbiol, 2005. **43**(1): p. 168-73.
17. Manson, J.M., et al., *A clonal lineage of VanA-type Enterococcus faecalis predominates in vancomycin-resistant Enterococci isolated in New Zealand.* Antimicrob Agents Chemother, 2003. **47**(1): p. 204-10.
18. Hanessian, S. and P.W.M. Sgarbi, *Design and synthesis of mimics of S-adenosyl-L-homocysteine as potential inhibitors of erythromycin methyltransferases.* Bioorganic & Medicinal Chemistry Letters, 2000. **10**(5): p. 433-437.
19. Roje, S., *S-Adenosyl-L-methionine: beyond the universal methyl group donor.* Phytochemistry, 2006. **67**(15): p. 1686-98.
20. Bhujbalrao, R. and R. Anand, *Deciphering Determinants in Ribosomal Methyltransferases That Confer Antimicrobial Resistance.* J Am Chem Soc, 2019. **141**(4): p. 1425-1429.
21. Yu, L., et al., *Solution structure of an rRNA methyltransferase (ErmAM) that confers macrolide-lincosamide-streptogramin antibiotic resistance.* Nat Struct Biol, 1997. **4**(6): p. 483-9.
22. Schluckebier, G., et al., *The 2.2 angstrom structure of the rRNA methyltransferase ErmC ' and its complexes with cofactor and cofactor analogs: Implications for the reaction mechanism.* Journal of Molecular Biology, 1999. **289**(2): p. 277-291.

23. Stsiapanava, A. and M. Selmer, *Crystal structure of ErmE - 23S rRNA methyltransferase in macrolide resistance.* Sci Rep, 2019. **9**(1): p. 14607.
24. Goh, B.C., et al., *Crystal structure and functional analysis of mycobacterial erythromycin resistance methyltransferase Erm38 reveals its RNA-binding site.* J Biol Chem, 2022. **298**(2): p. 101571.
25. Lee, H.J., S.T. Jhang, and H.J. Jin, *Potential Target Site for Inhibitors in MLSB Antibiotic Resistance.* Antibiotics (Basel), 2021. **10**(3).
26. Farrow, K.A., et al., *Identification of essential residues in the Erm(B) rRNA methyltransferase of Clostridium perfringens.* Antimicrob Agents Chemother, 2002. **46**(5): p. 1253-61.
27. Maravic, G., et al., *Alanine-scanning mutagenesis of the predicted rRNA-binding domain of ErmC ' redefines the substrate-binding site and suggests a model for protein-RNA interactions.* Nucleic Acids Research, 2003. **31**(16): p. 4941-4949.
28. Rowe, S.J., et al., *Shared requirements for key residues in the antibiotic resistance enzymes ErmC and ErmE suggest a common mode of RNA recognition.* Journal of Biological Chemistry, 2020. **295**(51): p. 17476-17485.
29. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-+.
30. Varadi, M., et al., *AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models.* Nucleic Acids Research, 2022. **50**(D1): p. D439-D444.
31. Gruber, A.R., et al., *The Vienna RNA websuite.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W70-4.
32. Vester, B., et al., *ErmE methyltransferase recognition elements in RNA substrates.* J Mol Biol, 1998. **282**(2): p. 255-64.
33. Renault, M., A. Cukkemane, and M. Baldus, *Solid-State NMR Spectroscopy on Complex Biomolecules.* Angewandte Chemie-International Edition, 2010. **49**(45): p. 8346-8357.
34. Long, Z., S.H. Park, and S.J. Opella, *Effects of deuteration on solid-state NMR spectra of single peptide crystals and oriented protein samples.* J Magn Reson, 2019. **309**: p. 106613.
35. Sattler, M. and S.W. Fesik, *Use of deuterium labeling in NMR: Overcoming a sizeable problem.* Structure, 1996. **4**(11): p. 1245-1249.
36. Clark, E.D.B., *Refolding of recombinant proteins.* Curr Opin Biotechnol, 1998. **9**(2): p. 157-63.
37. Yamaguchi, H. and M. Miyazaki, *Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies.* Biomolecules, 2014. **4**(1): p. 235-51.
38. Keeler, J., *Understanding NMR Spectroscopy*. 2010, United Kingdom: John Whiley & Sons Ltd.
39. Wang, Y. and O. Jardetzky, *Probability-based protein secondary structure identification using combined NMR chemical-shift data.* Protein Sci, 2002. **11**(4): p. 852-61.
40. Weingarth, M., et al., *Improved magnetization transfer in solid-state NMR with fast magic angle spinning.* Chemical Physics Letters, 2009. **469**(4-6): p. 342-348.
41. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding.* Prog Nucl Magn Reson Spectrosc, 2013. **73**: p. 1-16.
42. Ciuffreda, P., S. Casati, and A. Manzocchi, *Complete (1)H and (13)C NMR spectral assignment of alpha- and beta-adenosine, 2'-deoxyadenosine and their acetate derivatives.* Magn Reson Chem, 2007. **45**(9): p. 781-4.
43. Boehringer, D., et al., *Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis.* J Biol Chem, 2012. **287**(13): p. 10453-10459.
44. Skinner, S.P., et al., *CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis.* J Biomol NMR, 2016. **66**(2): p. 111-124.
45. Grzesiek, S. and A. Bax, *IMPROVED 3D TRIPLE-RESONANCE NMR TECHNIQUES APPLIED TO A 31-KDA PROTEIN.* Journal of Magnetic Resonance, 1992. **96**(2): p. 432-440.
46. Clubb, R.T., V. Thanabal, and G. Wagner, *A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C' chemical shifts in 15N・13C-labelled proteins.* Journal of Magnetic Resonance (1969), 1992. **97**(1): p. 213-217.

47. Grzesiek, S. and A. Bax, *AN EFFICIENT EXPERIMENT FOR SEQUENTIAL BACKBONE ASSIGNMENT OF MEDIUM-SIZED ISOTOPICALLY ENRICHED PROTEINS.* Journal of Magnetic Resonance, 1992. **99**(1): p. 201-207.
48. Mureddu, L. and G.W. Vuister, *Simple high-resolution NMR spectroscopy as a tool in molecular biology.* Febs Journal, 2019. **286**(11): p. 2035-2042.
49. Niklasson, M., et al., *Comprehensive analysis of NMR data using advanced line shape fitting.* J Biomol NMR, 2017. **69**(2): p. 93-99.
50. Baldus, M., et al., *Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems.* Molecular Physics, 1998. **95**(6): p. 1197-1207.
51. Weingarth, M., P. Tekely, and G. Bodenhausen, *Efficient heteronuclear decoupling by quenching rotary resonance in solid-state NMR.* Chemical Physics Letters, 2008. **466**(4-6): p. 247-251.

9 Supplementary information

A



1. Marker
2. $^{15}N/^{13}C$ ErmB before expression
3. $^{15}N/^{13}C$ ErmB after expression
4. $^{15}N/^{13}C$ ErmB after expression
5. $^{15}N/^{13}C$ ErmB supernatant
6. $^{15}N/^{13}C$ ErmB pellet
7. -
8. ^{15}N ErmB before expression
9. ^{15}N ErmB after expression
10. ^{15}N ErmB supernatant
11. ^{15}N ErmB pellet
12. -
13. Unlabelled ErmB supernatant
14. Unlabelled ErmB pellet

B

1. Marker
2. -
3. $^{15}N/^{13}C$ ErmB initial
4. Wash-fractions
5. Flowthrough-fractions
6. Fractions 17-18
7. Fractions 19-20
8. Fractions 21-22
9. Fractions 23-24
10. Fractions 25-26
11. Fractions 27-28
12. Fractions 29-30
13. Fractions 31-32
14. Fractions 33-34
15. Fractions 35-36

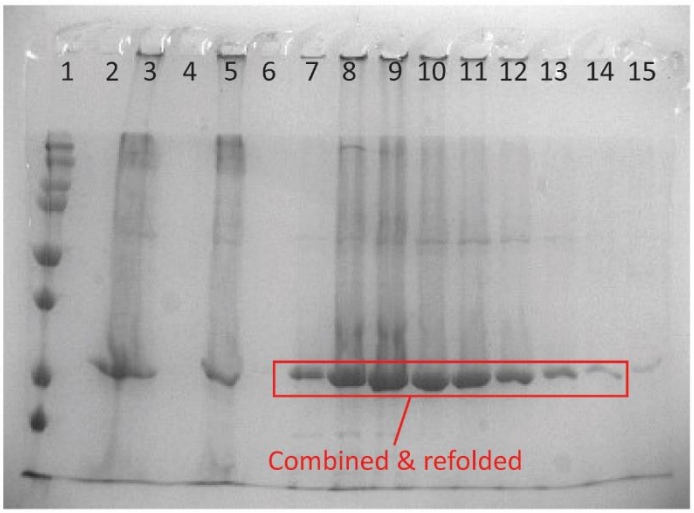


Figure S1: A) SDS-gel showing the clear appearance of a protein at the wanted molecular weight (~30kDa) before and after inducing expression for a $^{15}N/^{13}C$, ^{15}N and unlabelled ErmB culture (red). Each lane is marked and right of the gel contents per lane are shown. B) SDS-gel after purification by HisTrap FF column. The combined and refolded fractions are highlighted in red. Each lane is marked and left of the gel contents per lane are shown.

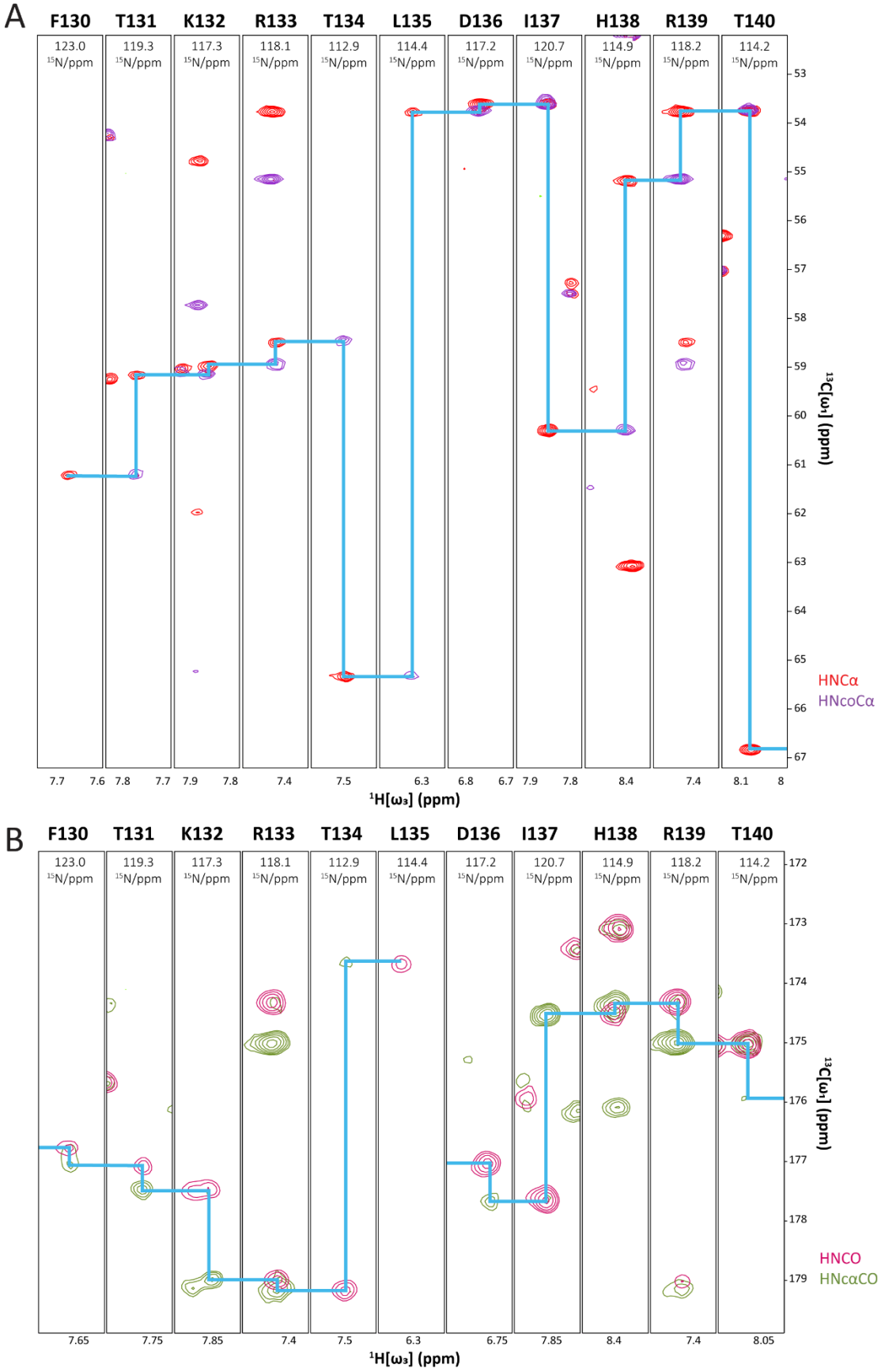


Figure S2: HNCα/HNcoCα (A) and HNCO/HNcαCO (B) sequential walks shown for amino acids F130-T140. These four experiments enables the connection of carbon (Cα and CO) resonances to the NH resonance of residue i, as well as residue i-1 to NH resonance of residue i. Overlaying the different spectra enables connectivity between residues i and i-1 sequentially.

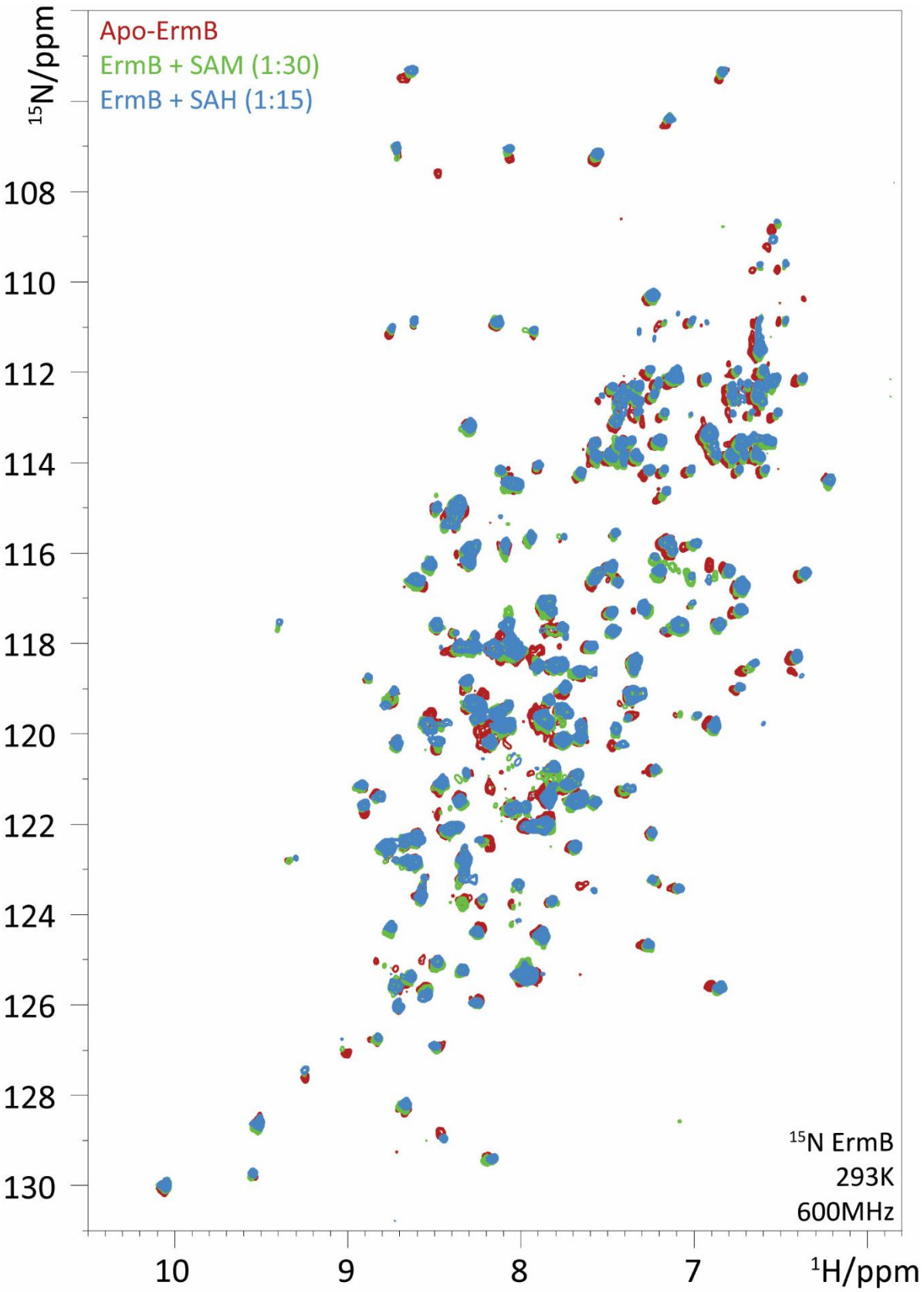


Figure S3: NH solution NMR spectrum of apo-ErmB (red) overlayed with the solution NMR spectrum of ErmB in complex with SAM (green) and SAH (blue), the demethylated cofactor. The solution NMR spectrum of ErmB in complex with SAM and SAH show high similarity, highlighting the same binding mode and proving that SAH can also be used instead of SAM to prevent too fast degradation of the cofactor.

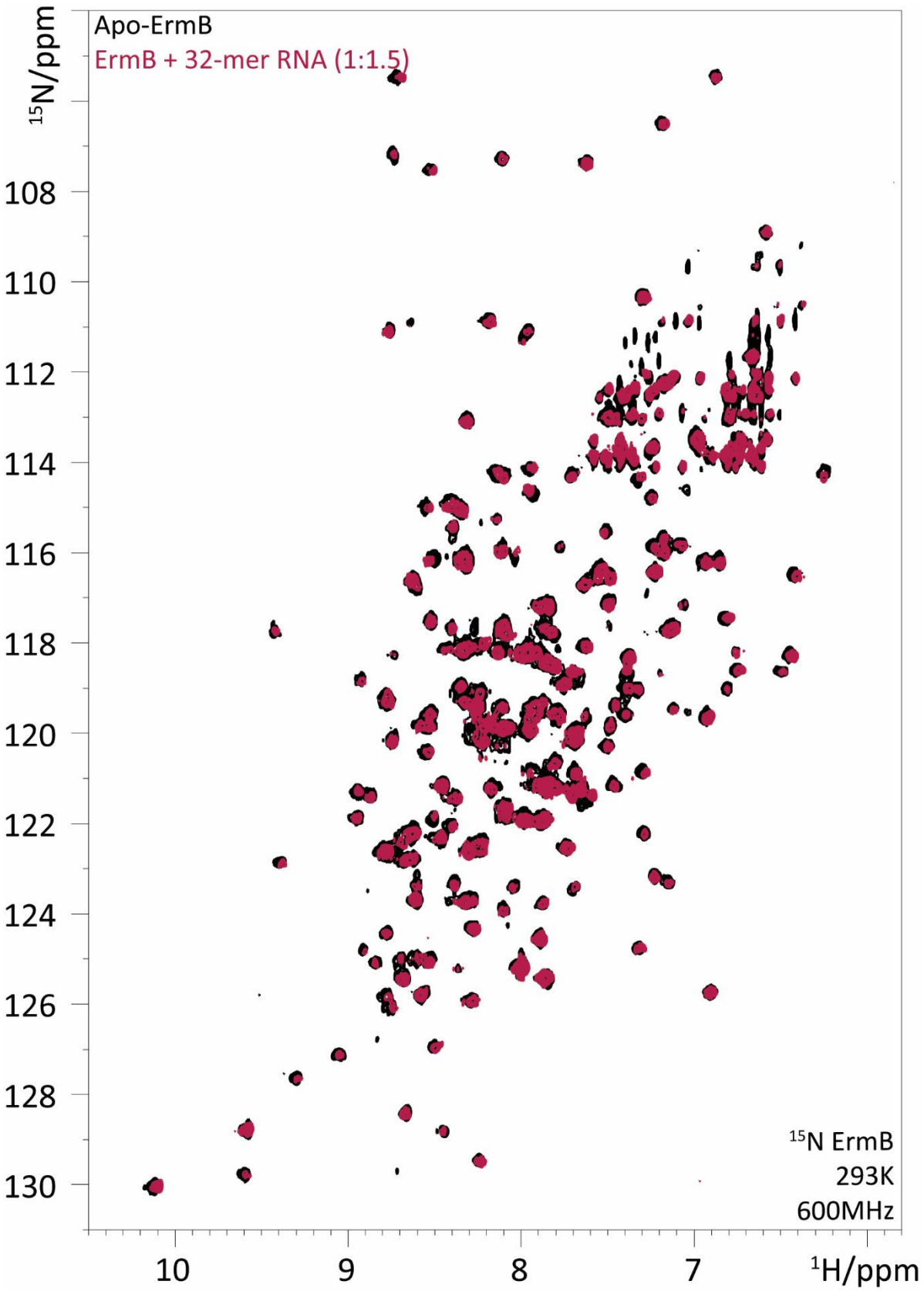


Figure S5: NH solution NMR spectrum of apo-ErmB (black) overlayed with the solution NMR spectrum of ErmB with 32-mer RNA (dark red) present under high salt concentrations (500 mM NaCl). The spectra overlap completely and no significant CSPs can be seen. It is likely that the 32-mer RNA does not adapt its distinct hairpin fold and therefore cannot (unspecifically) bind to ErmB.

A **Amino acid sequence ErmB**

```
          10         20         30         40         50
MNKNIKYSQN FLTSEKVLNQ IIKQLNLKET DTVYEIGTGK GHLTTKLAKI
          60         70         80         90        100
SKQVTSIELD SHLFNLSSEK LKLNTRVTLI HQDILQFQFP NKQRYKIVGN
         110        120        130        140        150
IPYHLSTQII KKVVFESHAS DIYLIVEEGF YKRTLDIHRT LGLLLHTQVS
         160        170        180        190        200
IQQLLKLPAE CFHPKPKVNS VLIKLTRHTT DVPDKYWKLY TYFVSKWVNR
         210        220        230        240
EYRQLFTKNQ FHQAMKHAKV NNLSTVTYEQ VLSIFNSYLL FNGRK
```

B **Nucleotide sequence 32-mer RNA**

```
          10         20         30
CGCGACGGAC GGAAAGACCC CUAUCCGUCG CG
```

Figure S6: A) Amino acid sequence ErmB. B) Nucleotide sequence 32-mer RNA

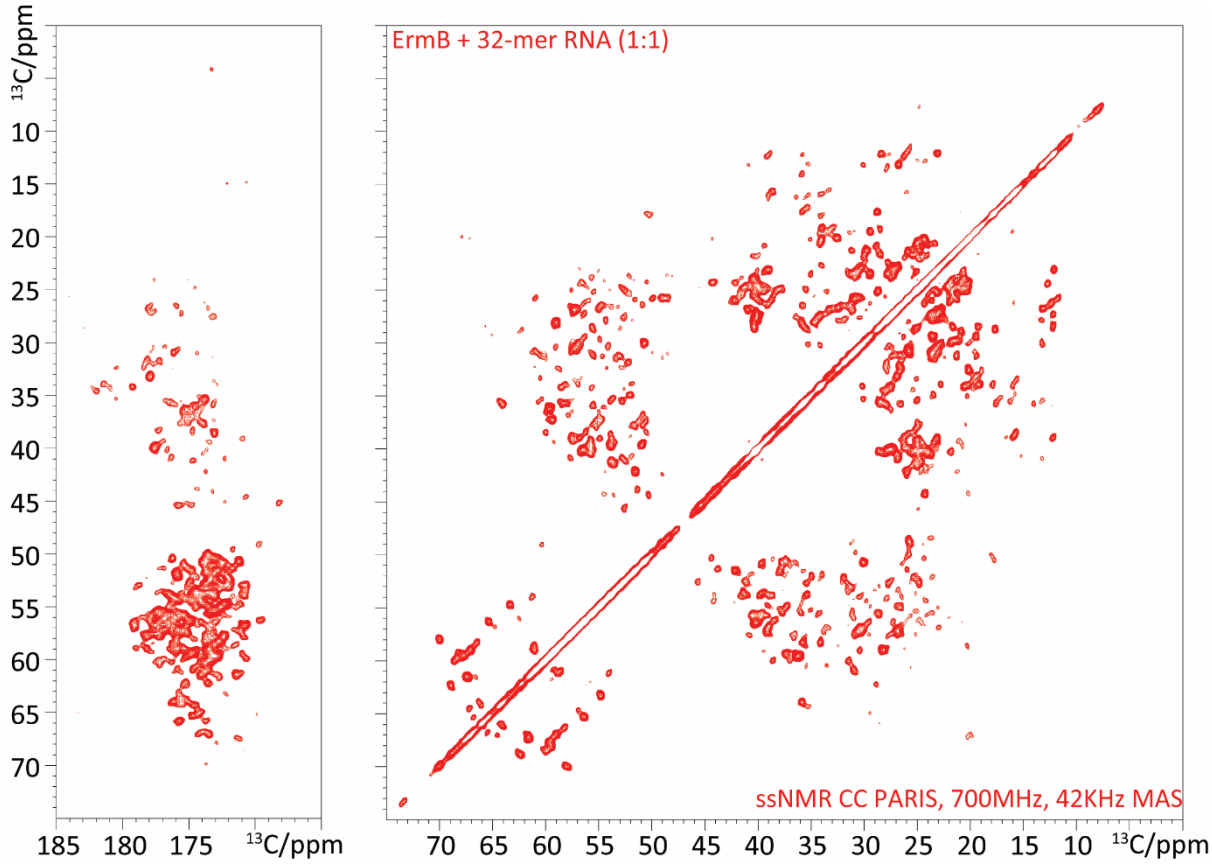


Figure S7: 2D CC correlated PARIS spectrum of ErmB in complex with 32-mer RNA. The left panel shows the CC correlations for CO with other C-atoms per amino acids. The right panel shows C-atom correlations within each amino acid, excluding CO. High spectral quality is observed in this PARIS spectrum for a relatively large protein such as ErmB.

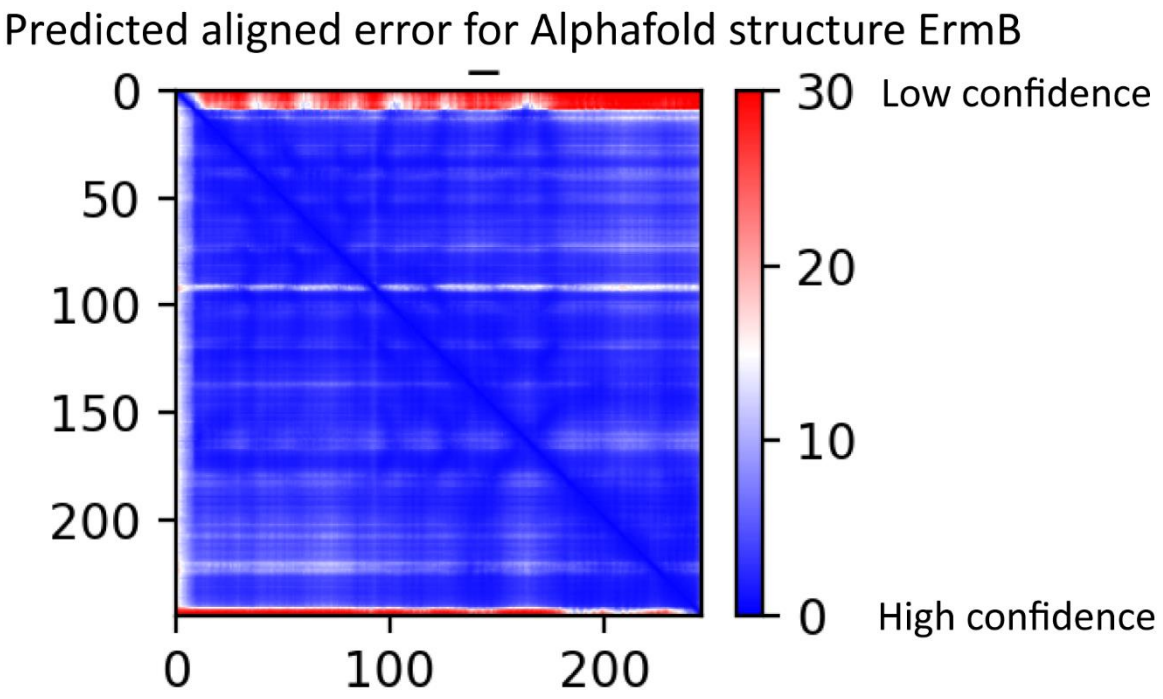


Figure S8: Predicted aligned error graph for the structure of ErmB predicted by Alphafold. Each line represents amino acid, red colour means that Alphafold is low in confidence for that residue, while, blue colour means that Alphafold is high in confidence for that residue. Only the N- and C-terminal domain show low confidence.