From the DEPARTMENT OF MEDICAL BIOCHEMISTRY AND BIOPHYSICS
Karolinska Institutet, Stockholm, Sweden

RIBOSOMAL RNA DYNAMICS STUDIED BY NMR-SPECTROSCOPY

Hampus Karlsson

Stockholm 2021

Ribosomal RNA dynamics studied by NMR-spectroscopy

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Hampus Karlsson

The thesis will be defended in public at Biomedicum Solnavägen 9, room: D1012, on Thursday 2021-04-15 at 13:00

Principal Supervisor:
Assoc. Prof. Dr. Katja Petzold
Karolinska Institutet
Department of Medical Biochemistry and
Biophysics
Division of Molecular Structural Biology

Co-supervisor(s):
Prof. Dr. Gunter Schneider
Karolinska Institutet
Department of Medical Biochemistry and
Biophysics
Division of Molecular Structural Biology

Opponent:
Prof. Dr. Harald Schwalbe
Johan Wolfgang Goethe-Universität
Institut für Organische Chemie und
Chemische Biologie
Zentrum für Biomolekulare
Magnetische Resonanz

Examination Board:
Prof. Dr. Lena Mäler
Stockholm University
Department of Biochemistry and Biophysics

Prof. Dr. Maria Selmer
Uppsala University
Department of Cell and Molecular Biology,
Structural Biology

Prof. Dr. Björn Högberg
Karolinska Institutet
Department of Medical Biochemistry and
Biophysics

To the curious undergraduate student

ABSTRACT

The ribosome is a large macromolecular machine that consists of both ribosomal proteins and ribosomal RNA (rRNA). It is a complex consisting of two subunits held together by non-covalent interactions, intersubunit bridges and some of these bridging interactions are mediated by the rRNA.

In this PhD-project the dynamics of such regions of rRNA, participating in intersubunit bridges and tertiary interaction within the rRNA have been investigated with solution state NMR-spectroscopy. These studies have been performed in the context of several miniaturized RNA systems, containing sequences of *E. coli* 16S rRNA. In particular, regions along helix 44 (h44), the penultimate stem of *E. coli* 16S rRNA have been studied. The stem-loop part of h44 has been studied in detail, this part of the rRNA contains a naturally occurring UUCG-loop and adenines participating in a tertiary interaction with helix (h8) in the 16S rRNA.

In order to characterize the dynamics of these RNA constructs with NMR-spectroscopy, purified RNA material in large amounts is a necessity. Because of this we have developed an RNA-sample production method (Paper I) as well an NMR-experiment method (Paper II) that we call SELOPE, a method that can reduce the need of using isotopically enriched RNA material for NMR-studies. The first chapter of this thesis introduces the underlying theory for RNA-sample preparation as well as alternative techniques compared to the ones used in Paper I. In a similar manner the underlying theory for the NMR-technique is introduced and with some emphasis on concepts crucial for understanding the SELOPE experiment, to contextualize Paper II. The usage of NMR-spectroscopy for the measurements of dynamics in RNA molecules is also introduced. The first chapter of the thesis also includes a description of the ribosome to help further understanding of Paper III. In addition, during chapter 2-5 of this thesis some work related to ^{1}H-R$_{1\rho}$ characterization of chemical exchange and cross-relaxation among RNA imino protons is described and discussed.

In Paper I, the development of an RNA-sample preparation method is described. The method is based on *in vitro* transcription of the wanted RNA sequence followed by a HPLC-purification procedure that uses two different HPLC techniques for the purification, both Reverse Phase Ion Pairing (RP-IP) and Ion Exchange (IE) HPLC. The complete method offers a robust and versatile alternative to other RNA sample preparation methods such as preparative gel electrophoresis techniques. In Paper II, we describe the development of an NMR pulse sequence that utilize a homonuclear magnetization transfer block in unlabeled RNA molecules. The pulse sequence can then for instance be used to transfer NMR signal of unwanted signals to other spectral regions and can for instance be used to remove the signal of pyrimidine H6s from the region of H6/H8/H2 in RNAs. In Paper III, the work of characterizing the stem-loop part of *E. coli* 16S rRNA h44 is described. This work both shows a UUCG-loop with dynamics on a millisecond time-scale as well as the dynamical behavior of a group of unpaired adenine bases, the study of dynamics of these adenines could aid the understanding of tertiary interactions within rRNA.

LIST OF SCIENTIFIC PAPERS

I. **Karlsson H.**, Baronti L., and Petzold K. (2020). A robust and versatile method for production and purification of large-scale RNA samples for structural biology. *RNA 26:1023-1037*.

II. Schlagnitweit J., Steiner E., **Karlsson H**., and Petzold, K. (2018) Efficient Detection of Structure and Dynamics in Unlabeled RNAs: The SELOPE Approach. *Chemistry a European Journal 24:6067-6070*.

III. **Karlsson H.**, Fontana C., Schlagnitweit J., Steiner E., and Petzold, K. UUCG-loop and A-minor motif dynamics in helix 44 of 16S ribosomal RNA of *E. coli* by NMR-spectroscopy. *Manuscript*

Additional papers not included in this thesis

Baronti L., **Karlsson H.**, Marušič M. and Petzold K. (2018). A guide to large-scale RNA sample preparation. *Analytical and bioanalytical chemistry, 410(14), 3239-3252*.

Karlsson, H., Feyrer H. and Petzold K. Production of structured RNA fragments by in vitro transcription and HPLC purification. *Manuscript*

CONTENTS

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
tRNA	Transfer RNA
rRNA	Ribosomal RNA
miRNA	MicroRNA
lncRNA	Long non-coding RNA
NMR	Nuclear Magnetic Resonance
mRNA	Messenger RNA
NTP	Nucleotide triphosphate
RNase	Ribonuclease
PAGE	Polyacrylamide gel-electrophoresis
TEMED	Tetramethylethylenediamine
LC	Liquid chromatography
MS	Mass spectrometry
HPLC	High performance liquid chromatography
FPLC	Fast performance liquid chromatography
N	Number of theoretical plates
H	Height of a theoretical plate
RP	Reverse phase
IP	Ion pairing
TEA	Triethylamine
IE	Ion exchange
LSU	Large subunit
SSU	Small subunit
UTR	Untranslated region
IF	Initiation factor
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
EF	Elongation factor

RF	Release factor
$\hat{I}_z$	Z spin angular momentum operator
$\hat{H}$	Hamiltonian operator
γ	Gyromagnetic ratio
ω_0	Larmor frequency
B_0	Magnetic field strength
NS	Number of scans
T_2	Transverse relaxation time
T_1	Longitudinal relaxation time
Ω	offset
ω_{eff}	Magnitude of effective filed
θ	Theta angle
FID	Free induction decay
ω_1	Field strength of radio pulse
ppm	Parts per million
NOE	Nuclear Overhauser effect
ROE	Rotating frame Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
ROESY	Rotating frame Overhauser effect spectroscopy
HSQC	Heteronuclear single quantum correlation
R_1	Longitudinal relaxation rate constant
R_2	Transverse relaxation rate constant
k_{ex}	Exchange rate constant
GS	Ground state
ES	Excited state
$R_{1\rho}$	Rotating frame relaxation rate constant
R_{ex}	Exchange contribution term
HCP	Heteronuclear cross-polarization
BMRB	Biological Magnetic Resonance Bank

1 INTRODUCTION

Nowadays it is well-known that nucleic acids can both carry information for protein synthesis and form complicated three dimensional structures. More than 60 years ago the *Central Dogma* was introduced by Francis Crick, that stated that *information* for protein synthesis was transferred from DNA to RNA that eventually became *translated* into proteins[1]. Later on, it was realized that RNA has many other functions of non-coding nature where the RNA forms functional molecules such as transfer RNA (tRNA), ribosomal RNA (rRNA), micro-RNA (miRNA) and long noncoding RNA (lncRNA)[2]. The versatile nature of RNA and its participation in both coding as well as complicated molecular processes such as catalysis, as in the ribosome has been the foundation for the idea of the *RNA world hypothesis,* which suggest that prior to a world where DNA is *transcribed* into RNA and translated into proteins, a world existed where everything was performed by RNA[3].

This increasing knowledge of the central role of non-coding RNA has led to extensive investigation of the structural biology and dynamical behavior of these molecules. The structure of ribosomes has been extensively investigated with techniques such as X-ray crystallography and cryo-EM technology over the decades that today have led to a highly detailed view of this complex molecular machine[4,5]. Also, more detailed studies using solution state NMR-spectroscopy have provided important information on the structure and dynamics on regions of rRNA, such as the ribosomal A-site[6,7]. As well as the other regions of the rRNA such as the Sarcin Ricin Loop, the target of the infamous ricin toxin[8]. Solution state NMR has also been applied to the studies of miRNAs[9,10] and also smaller *ribozymes* such as the Lead-dependent Ribozyme[11] and the Varkud satellite ribozyme[12].

This chapter elaborates on techniques used to perform such solution state NMR studies of various RNA molecules in the context of small RNA constructs and will serve as the underlying theory for Paper I-III in this thesis. Firstly, it will give an introduction to RNA sample preparation and purification using *in vitro* transcription and HPLC purification techniques. Secondly it will introduce the ribosome and important aspects as rRNA. It will also give an overview of the NMR-spectroscopy technique, its underlying theory and how it can be applied to study the dynamics of RNA molecules.

1.1 RNA sample production

Any biochemical study of RNA does require some quantity of RNA material to perform the study on. Structural investigations using techniques such as NMR spectroscopy or X-ray crystallography require significant amounts of RNA material, as pure and homogeneous in length and fold as possible. Biological experiments such as cell transfection assays or messenger RNA (mRNA) regulation experiments might put lower demands on the amount of RNA material but still high demands on for instance sequence length homogeneity to ensure correct RNA-RNA interactions and avoid off-target effects. There are many different ways to produce and purify RNA samples, various *in vitro* transcription systems are common alternatives as well as solid-phase synthesis. For the purification of impure RNA material,

different types of gel electrophoresis techniques are common and a wide variety of liquid chromatographic techniques on preparative scales as well. Research and development of more efficient and convenient RNA-sample production methods is still a highly important question and the research topic of constituent paper I of this PhD-thesis. The following sections will present an overview of the existing techniques and their applications.

1.1.1 *In vitro* transcription

In vitro transcription of RNA using RNA polymerase enzymes originating from different bacteriophages is still a very popular production method to obtain single stranded RNA molecules of various length in large amounts. T7 RNA polymerase is a common alternative and can be used for transcription from both linearized plasmids[13] and synthetic DNA-templates[14]. T7 RNA polymerase requires a double stranded promoter region and if a single stranded synthetic DNA-template is used the promoter region needs to be annealed with its reverse complementary DNA-strand[15]. The sequence of the T7 promoter region can vary and there exist several different classes of promoters, for the T7 based *in vitro* transcription used for RNA sample production throughout this PhD-project a version of the *class III promoters* has been used. The sequence of this T7 promoter and the appearance of a complete, synthetic single stranded DNA-template can be found in (Fig. 1.1a). The performance of the T7 RNA polymerase is sensitive to the concentration of reactants and cofactors of the transcription reaction and commonly $MgCl_2$, nucleotide triphosphates (NTPs) and DNA-template concentration are important factors to consider when optimizing the reaction conditions[14,15]. Examples of *in vitro* transcription optimization procedures can also be found in the material and methods section and supplementary information of constituent Paper I[16] of this PhD-thesis.

Another crucial factor that will affect the ultimate yield of transcription reactions, is the sequence itself of the RNA to be transcribed and T7 RNA polymerase prefers RNA-sequences with purine rich starting sequences, one example of highly transcribing starting sequence is 5´-GGGAG-3´[17] and commonly inclusion of at least two starting guanines is beneficial. T7 transcription from linear templates means that the enzyme will continue to add nucleotides all the way to the 5´-end of the DNA-template (Fig. 1.1a) and then leave the template when it reaches the end, this is called *run off transcription*. It is common that the polymerase adds a few extra nucleotides at the 3´-end of the *RNA transcript,* nucleotides that were not coded by the DNA-template, this is a well-known drawback of the T7 RNA polymerase[14]. The problems of the T7 based RNA production such as non-templated nucleotide addition at the 3´-end of the transcript and reaction yields highly dependent on the sequence of the RNA have prompted the development of several improvements and modified versions of this RNA production technique.
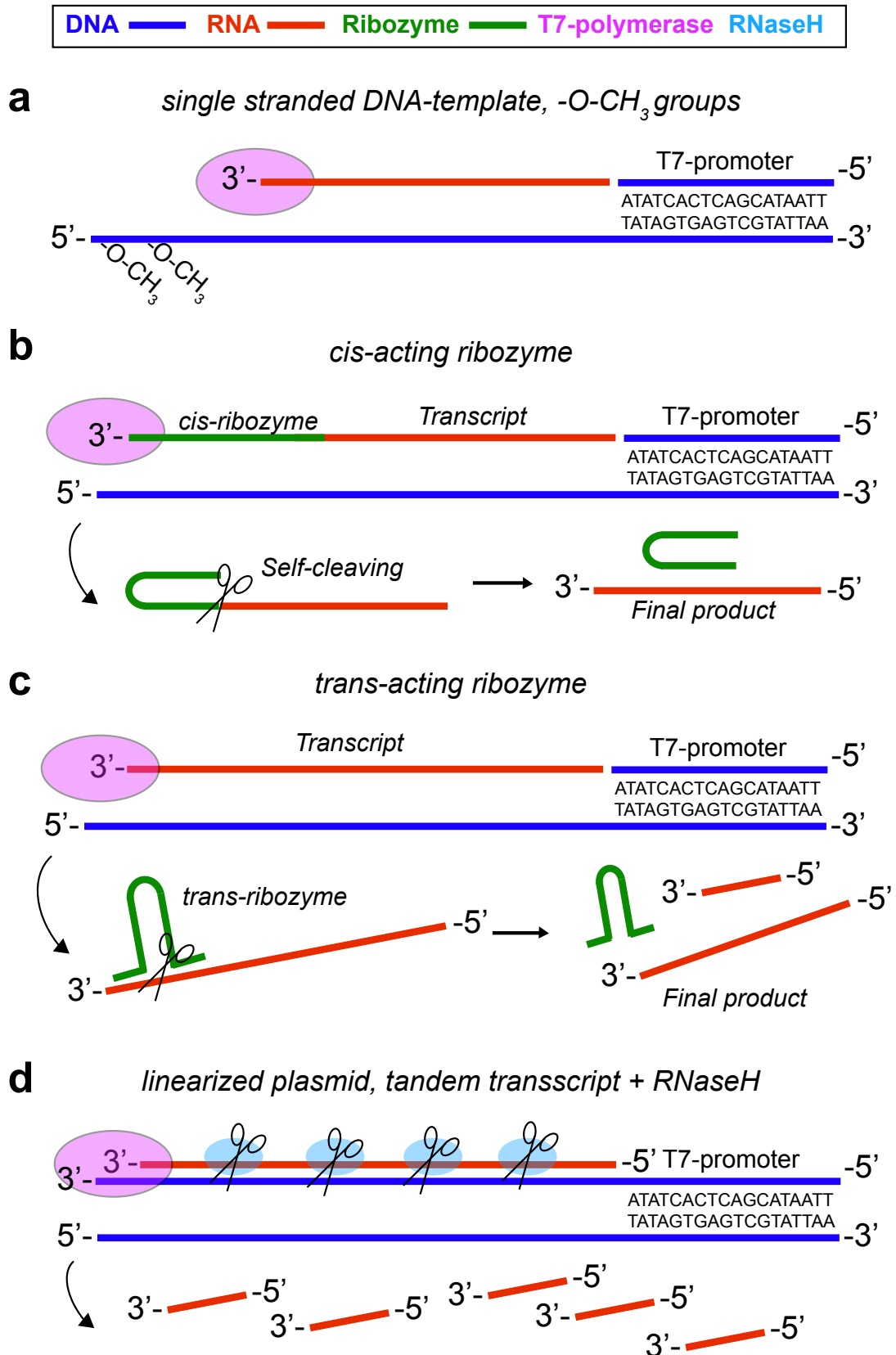
Figure. 1.1. Overview of *in vitro* transcription techniques. **a,** Transcription from single-stranded DNA-template with methyl groups to prevent non-templated nucleotide addition. **b,** Usage of a cis-acting ribozyme to increase 3´-end homogeneity. **c,** Usage of *trans*-acting ribozyme to increase 3´-end homogeneity. **d,** Transcription of polycistronic RNA cleaved by RNaseH.

To reduce the problem of addition of non-templated nucleotides at the 3′-end of the RNA transcript, methoxy groups can be added to the C2′-position of the ribose part of the last two nucleotides at the 5′-end of the DNA-template[18,19]. This is an option when transcribing from a synthetically made DNA-template when the methoxy groups can be added during synthesis of the DNA-template. If it is not possible to change the starting nucleotides of the wanted RNA transcript, such that it starts with two guanines or another sequence beneficial for transcription there are other alternatives to boost the yield of the transcription and addition of solvents such as DMSO can increase the yield of the reaction[20]. Addition of DMSO to the reaction mixture and usage of methoxy groups on the transcription template are examples of modern improvements to the *in vitro* transcription technique that has been implemented in the RNA sample production method described in constituent Paper I of this PhD-thesis.

Since limited yields due to an unsuitable starting sequence have been a common problem for labs using T7 transcription, a multitude of methods utilizing ribozymes (RNA molecules with enzymatic activities) or *ribonuclease* enzymes (RNases, RNA-cleaving enzymes) have been described in literature, see also (Fig. 1.1b, c & d). Common to these methods are that the RNA to be transcribed is designed with starting sequences that are beneficial for transcription yield, such as the aforementioned 5′-GGGAG-3′ or similar, the transcript is then cleaved to yield the final, wanted RNA molecule. These methods can also aid in reducing 3′-end inhomogeneity of the RNA transcript due to non-templated nucleotide addition, since the part containing the unwanted nucleotides can be cleaved of. *Cis-acting* ribozymes are ribozymes that cleave the same RNA sequence that they are a part of and *trans-acting* ribozymes are ribozymes that cleave other RNA molecules. An example of a ribozyme that could be added in *cis* (i.e. added as an additional part, 3′ relative to a wanted RNA sequence and then cleaved off) is the *hepatitis delta ribozyme,* this ribozyme has no other upstream sequence demands than to cut after any base that is not a G to make the cleavage reaction work, it is however large (87 nucleotides long) and it leaves a cyclic phosphate group at the 3′-end of the wanted cleavage product[21,22]. If a transcribed RNA is intended for NMR-studies and it is transcribed using nucleotides with expensive, NMR-active isotopes such as ^{15}N or ^{13}C a *cis*-acting ribozyme might not be a good option since a large proportion of the expensive nucleotides will form the ribozyme. An alternative is then to use a *trans*-acting ribozyme. The *hammerhead* ribozyme can be used *cis*-acting but also *trans*-acting, which would open up for the possibility to use a non-isotopically labeled ribozyme that then for instance could cleave the 3′-end of a wanted ^{15}N/^{13}C labelled transcript, the hammerhead ribozyme does however have some sequence demands on the target transcript[23]. Another example of a ribozyme that can be used *trans*-acting is the antigenomic delta ribozyme[24]. Sophisticated methods utilizing both ribozymes and ribonucleases such as RNaseH for RNA cleavage combined with RNA ligation steps also exist, such methods allow for combining isotopically labeled RNA transcripts with non-labelled ones i.e. creating partially labeled RNA transcripts[25]. RNaseH based methods utilizing tandem repeats of the wanted RNA sequence such as the method described by Feyrer et al. [26]. is also an alternative (Fig. 1.1d). In this method, a long RNA transcript containing some 20 or more repeats of the

wanted sequence is transcribed from a linearized plasmid, the crude transcript is then cleaved using RNaseH guided by small guiding oligos targeting the intersections of the repeats[26].

1.1.2 Solid phase synthesis

As an alternative to the enzymatic production of RNA, solid phase synthesis can be used. The chemical synthesis of an RNA molecule offers the possibility of incorporating modified and non-natural nucleotides into the growing RNA polymer. Solid phase synthesis of both DNA and RNA is done 3′ to 5′ (as opposed to polymerase-based methods where the RNA polymer is created 5′ to 3′) and is based on phosphoramidite chemistry[27]. The synthesis cycle is based on the nucleophilic reaction of a free hydroxy group on the 5′ carbon of the ribose part of a nucleoside with the 3′ phosphoramidite moiety of the next nucleotide, since the process is based nucleophilic attacks other groups such as 2′ hydroxy groups and nitrogens in nucleobases needs *protecting groups* during the synthesis[28]. The completion of the synthesis cycle needs removal and addition of protecting groups at different time points and each step has different solvent demands, this makes the complete synthesis quite complicated and a little beyond the scope of this paragraph as well as my knowledge. Caruthers[27] and Somoza[28] as well as the solid phase synthesis paragraph of the RNA sample production review written by our lab by Baronti et al.[29] are however good literature resources to dive deeper into this topic. On the contrary to *in vitro* transcription techniques, which can be used for the production of really long transcripts, >10000 nucleotides[13] solid phase synthesis is more suited for the creation of shorter RNAs, from below 10 nucleotides up to ~80 nucleotides[30].

1.1.3 Analytical gel electrophoresis

Both enzymatic and chemical synthesis of oligonucleotides will generate unwanted side products to some extent, in order to evaluate how abundant different RNA/DNA molecules are, *analytical gel electrophoresis* can be used and this classic technique is still a major tool in any lab working with nucleic acids or proteins. Electrophoresis is based on the separation of charged molecules by forcing them to migrate through a gel matrix in a buffer under the impact of an electric potential[31]. Size separated molecules are then detected on the gel by *staining* with a dye that binds the RNA/DNA. The velocity of migration through the gel is proportional to the applied voltage and will depend on strength of the applied electric field, the total charge of the molecule as well as size properties of both the molecule and the gel matrix[32].

The gel matrix for electrophoresis typically consists of agarose or polymerized acrylamide. Polyacrylamide Gel Electrophoresis (PAGE) with higher percentages ~12-20% of polyacrylamide is suitable for separation of shorter nucleic acids. The polyacrylamide mixture used to create the gels are typically a mixture of a high proportion (such as 19:1) of *acrylamide* that creates linear chains and *bis-acrylamide,* which connect several linear chains, these chemicals when polymerized together will form a net like gel matrix with a particular *pore size* dependent on concentration[33]. The polymerization can be achieved by using persulfate ions and N,N,N,N-tetramethylethylenediamine (TEMED) but it is also possible to use riboflavin in combination with light for the polymerization reaction[34]. Nucleic acids can adopt different

folds during electrophoretic separation, different folds might have different migration properties on the gel and to reduce this problem *denaturing PAGE* can be used[35]. During such separations denaturing agents are added to the gel mix to reduce the impact of molecular folds and secondary structures one the separation and a classic option is the addition of urea to aid denaturation[36]. It is interesting to see other alternative denaturing agents that where up for discussion in the literature during the early days of RNA gel electrophoresis, among the alternatives suggested were organomercury compounds, formaldehyde etcetera[37]. These are highly toxic compounds in addition to the usage of acrylamide, which itself is very toxic.

Another important aspect of analytical gel electrophoresis is the detection of the separated bands, this is commonly done by submerging the gel in a solution of some UV-fluorescent or strongly colored compound that will bind to the nucleic acid, this is the staining procedure. An old and very common option for doing this has been *Ethidium Bromide*, which bind nucleic acids and is fluorescent in the presence of UV-light. Already since the introduction of this compound it has been proposed to bind to nucleic acids by intercalating between the base pairs[38] and to be a mutagenic by inducing *reading frame shifts*[39]. The longstanding discussion regarding the toxicity of ethidium bromide has prompted the development of less toxic alternatives and nowadays, modern alternatives such as the SYBR® series of stains are available, which have been shown to be less toxic[40]. Another example of usage of a staining dye that doesn't require UV-light is toluidine blue, and staining protocol examples can be found in literature[41].

1.1.4 Preparative gel electrophoresis

Denaturing, analytical PAGE performed with a high percentage (such as 20 % w/v) of acrylamide and a high concentration of urea (8 M) and ultimately stained with ethidium bromide or for instance SYBR Gold is a good way to analyze the performance of *in vitro* transcription reactions and purification methods. Electrophoretic separation of RNA can also be done on a larger scale and this is still a common way to *purify* RNA material produced in higher amounts. Preparative PAGE purifications of RNA are typically done on large gels, requiring hundreds of milliliters of polyacrylamide mixture followed be an *elution step* where the RNA is allowed to spontaneously diffuse out of the gel matrix or it is driven out of the gel with an electroelution apparatus[13,42]. As with all purification methods there are both pros and cons and for preparative PAGE purification, commonly the time aspect, low efficiency of elution and presence of acrylamide in the finished RNA preparation are highlighted as drawbacks of preparative PAGE[43]. Because of the inconvenience of detection, excision and elution of the RNA from slab gels, methods using gel electrophoresis where the separated RNA is continuously eluted and detected from the gel has been developed[44,45].

1.1.5 Liquid chromatography techniques

Liquid chromatography (LC) techniques have a wide range of applications and perhaps it is a technique that is most commonly associated with pure analytical chemistry and the analysis of low molecular weight compounds in various matrices, nowadays also commonly combined

with powerful detection techniques such as mass spectrometry (MS). Liquid chromatography techniques can however also be a powerful tool for more preparative purification setups for nucleic acids, which is the topic for the coming sections. The fundamental principle of the technique is still the same though, regardless if it is applied to nucleic acids or small drug compounds. The usage of chromatographic techniques for the special purpose of purifying RNA is also discussed in more detail in constituent Paper I[16] of this thesis as well as in the review by Baronti et al.[29].

The separation in liquid chromatography techniques such as High Performance Liquid Chromatography (HPLC) or Fast Performance Liquid Chromatography (FPLC) is based on dissolving an *analyte* in a *mobile phase* and then pump the mobile phase through a column containing a solid material, a *stationary phase.* The analytes, or the compounds to be separated will then interact with the stationary phase and their elution from the column will be delayed to different degrees, depending on how strongly they interact with the stationary phase. It is common that efficiency of a chromatographic column is described by the *number of theoretical plates* (N) and the height of a plate or *plate height* (H), the plate height is basically the column length (L) divided by the number of theoretical plates[46].

$$H = \frac{L}{N} \qquad (1)$$

The values of H and N are thus theoretical descriptors of the performance of a chromatographic column and are not related to existence of actual physical plates in the column, instead they are theoretical values and the word "plates" is an old word remaining from times when distillation columns with plates were used for separations[46]. There are multiple expressions utilizing the elution time or *retention time* of the analyte and different widths of the chromatographic peak to calculate values of N and H[46]. Another theoretical framework for the description of column efficiency is the *Van Deemter equation,* which describes H as a function of the *linear velocity* or the flow of the mobile phase[47]. The Van Deemter equation has the following appearance.

$$H = A + \frac{B}{u} + u(C_s + C_m) \qquad (2)$$

In which A is a term taking into account that analyte molecules will take pathways of different lengths through the stationary phase, a *multiple paths* term. B is a term describing *longitudinal diffusion* of the molecule and the C terms are *mass transfer terms* for the mobile and stationary phase respectively and *u* represents the linear velocity of the mobile phase. The details of some of these terms, especially the mass transfers are complicated and a little beyond the scope of this section but Skoog[46] and Andrés[47] elaborates more on this. Andrés[47] does also provide a unit less version (which yields H divided by the particle diameter of the stationary phase) of the Van Deemter equation with approximate values for the different terms, a plot using these approximate values show their respective contribution to the plate height is shown in (Fig. 1.2).

Another useful expression for HPLC usage is the one for calculation of a scaling factor for mobile phase flowrates when switching between columns of different diameters and it has the following appearance.

$$scaling\ factor = \left(\frac{d_l}{d_s}\right)^2 \qquad (3)$$

In which d_l is the diameter of the larger column and d_s the diameter of the small column. This expression I learned from HPLC workshop organized by Sigma-Aldrich in 2015, unfortunately I do not have any proper reference for it and I think it should be considered a rule of thumb and it should be recognized that switching column diameters actually is a more complicated problem.
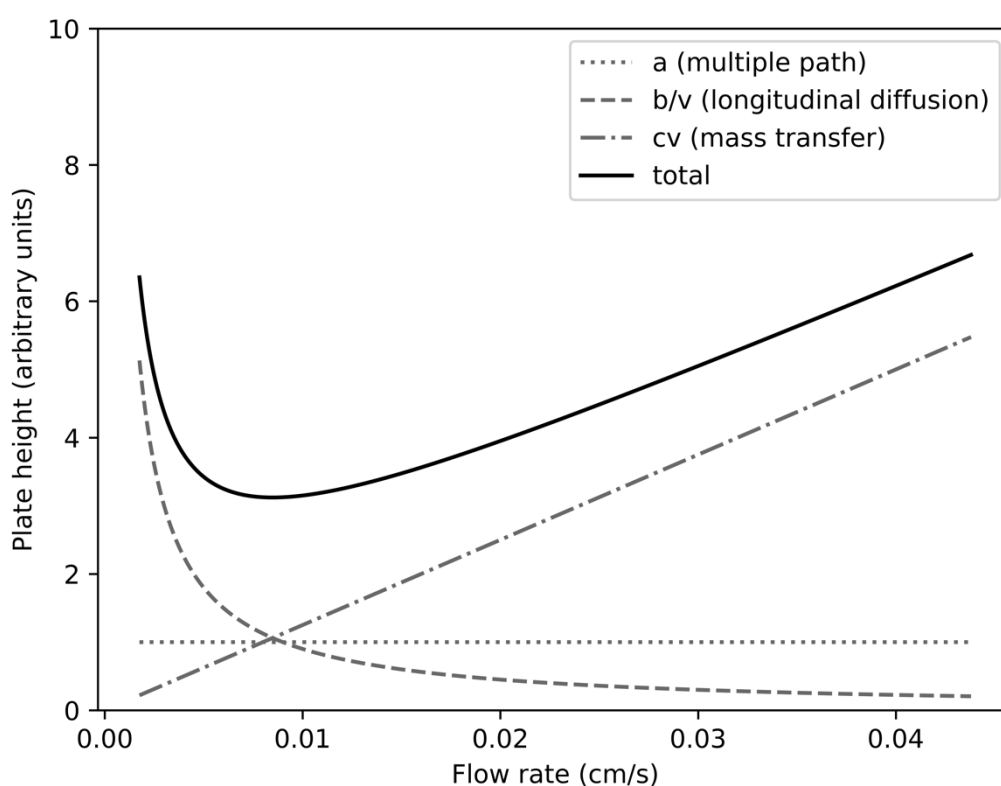

Figure. 1.2. A Van Deemter curve created based on the theory and approximate values presented by Andrés[47]. The black curve that is the sum of the three contributing terms indicates that there is an optimum flow rate at which theoretical plate height is at minimum. At higher flow rates the linear dependency of the mass transfer term (i.e. the analyte will not interact sufficiently with the stationary phase) will reduce plate height.

1.1.6 FPLC methods

FPLC, where columns packed with gel like materials such as sepharose is used as stationary phase can be used for both protein purifications and nucleic acid purifications and there are a number of FPLC based methods available in literature. Lukavsky and Puglisi[48] have described a FPLC based RNA production method in which RNA is transcribed from a linearized plasmid then purified with FPLC with a stationary phase that separates the RNA based on the size

exclusion principle (SEC), (Fig. 1.3a). McKenna et al. also describes an *in vitro* transcription based protocol that utilize size exclusion FPLC for the subsequent purification[49,50]. As an alternative to the size exclusion mechanism for the separation, ion exchange (IE) in which the anionic analyte and another anion competitively interacts with a cationic stationary phase (Fig. 1.3c) can also be combined with the FPLC technique for RNA purifications and usage of such columns have been described by Easton et al.[51] and Koubek et al.[52].

1.1.7 HPLC methods – Reverse Phase Ion Pairing systems

Reverse Phase HPLC (RP-HPLC) in which columns with a hydrophobic stationary phase are used can be applied for nucleic acid separation and the negatively charged nucleic acid is commonly allowed to form an ion pair with a hydrophobic cation that can interact with the stationary phase (Fig. 1.3b), creating Reverse Phase Ion Paring HPLC (RP-IP-HPLC)[53,54]. The RP-IP-HPLC method when applied to the separation of smaller amounts RNA can achieve high resolving power even of longer RNA molecules, when injections are done with smaller amounts (the ~10 µg range) which is very little material[55]. Other RP-IP-HPLC examples where purification of larger amounts is also possible to find in literature, McCarthy et al.[56] shows how purification of larger amounts of synthetic RNA can be done using RP-IP technology and tetraethyl ammonium ions as the ion pairing partner for creation of the lipophilic complex. Murray et al.[57] display example of a tetrabutylammonium hydrogen sulphate-based RP-IP buffer system that has been the basis for the RP-IP purification step for the RNA purification method described in constituent Paper I of this thesis.
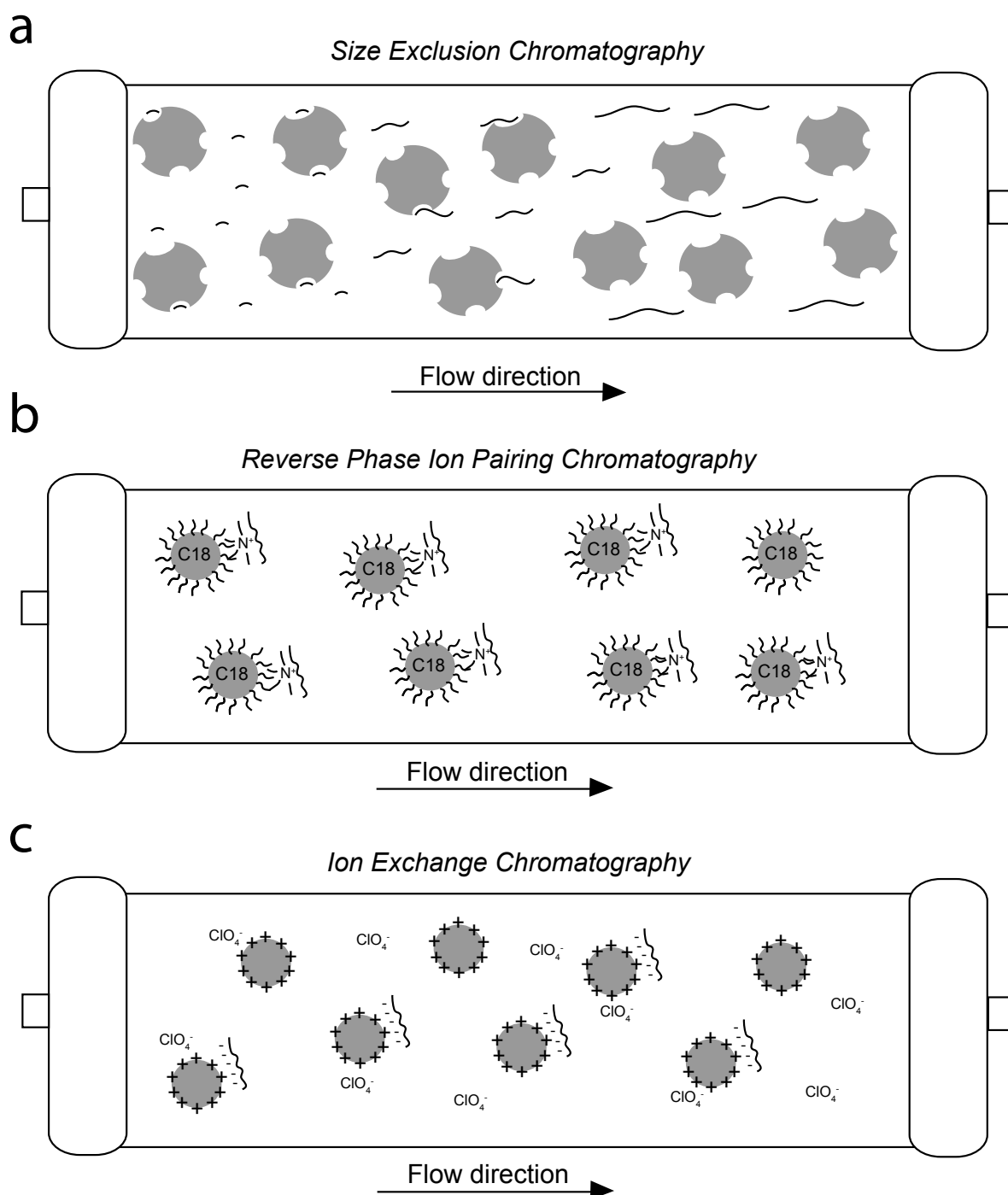
Figure 1.3. Separation mechanisms in liquid chromatography. **a,** In Size Exclusion Chromatography smaller molecules will enter pores in the stationary phase to a higher extent than larger molecules, thus smaller molecules will elute with longer retention time than larger ones. **b,** In Reverse Phase Ion Pairing chromatography negatively charged analytes (e.g. RNA) will form ion pairs with a lipophilic cation (typically quaternary ammonium compounds) the aliphatic chains of cations will make the complex lipophilic and enable interaction with the hydrophobic stationary phase of a reverse phase column, the complex can be eluted with organic solvent, typically acetonitrile. **c,** In Ion Exchange chromatography the anionic analyte (RNA) will directly interact with a cationic stationary phase, the analyte is eluted from the stationary phase using a gradient of a competitively binding hydrophilic anion such as Cl^- or ClO_4^-.

1.1.8 HPLC methods – Ion Exchange HPLC systems

Ion exchange HPLC (IE-HPLC) is the other major HPLC method for the purification of nucleic acids and perhaps more extensively used for larger, more preparative setups of the HPLC method. Andersson et al.[58] provide a detailed example of how the IE method can be used for the purification of RNA, especially for the structural biology techniques. They also importantly discuss the impact of secondary structures on IE nucleic acid separations and the effect of different cations on the elution behavior. It is also possible to combine the usage of IE-HPLC for purification of *in vitro* transcribed RNA and utilize *trans*-cleavage from a hammerhead ribozyme[23]. Other good examples of usage of IE-HPLC for RNA production in combination with ribozymes and ribonucleases comes from Duss et al.[25] who use IE-HPLC for the purification of segmentally labeled RNAs.

1.2 The ribosome

The ribosome is a highly complex component of the cell and it is responsible for the crucial task of translating the genetic information into proteins. It operates together with a large set of additional factors and energy supplying molecules to perform its important job. The following sections introduce the anatomy of the ribosome, how it performs translation of proteins and some other central aspects of it.

1.2.1 Structure of the ribosome

The ribosome is a large macromolecular machine that translates the messenger RNA (mRNA) into proteins. Ribosomes consist of two subunits, a large and a small subunit and it is the large subunit that catalyze the peptide bond formation and the small subunit that decodes the mRNA[4]. The individual subunits are in turn complexes of ribosomal RNA (rRNA) and proteins. Eukaryotic ribosomes contain three rRNAs, the 5S, 5.8S and the 25S rRNA in their large subunit (LSU), one chain of rRNA, the *18S* rRNA in their small subunit (SSU) and in total ~80 ribosomal proteins and all these components together constitute the *80S* eukaryotic ribosome[59]. The LSU of the bacterial ribosome consists of two rRNAs, the 23S and 5S and more than 30 proteins, that together creates the *50S* LSU. The SSU of bacterial ribosome consists of one rRNA, the *16S* rRNA and more than 20 proteins that together form the *30S* SSU, the 50S LSU and 30S SSU together then creates the *70S* bacterial ribosome[60]. An overview picture of a 70S bacterial ribosome is shown in (Fig. 1.4). Another interesting category of ribosomes are the mitochondrial ribosomes in eukaryotes, which have a higher proportion of ribosomal proteins compared to rRNA and that are responsible for the translation of the proteins in mitochondria[61]. The 16S rRNA of the SSU in *E. coli* bacteria is the rRNA that has been under study in this PhD-project and more in particular, certain regions of the 16S rRNA (see section 2.4). As highlighted by Ramakrishnan[5] in the early days of ribosomal research it was not known exactly what role the RNA in the ribosome played, if it just kept the proteins together or if it participated in the actual protein production. Nowadays, thanks to many high-resolution X-ray crystallographic and cryo-EM structures and biochemistry studies,

the details of how the ribosomal proteins are embedded in the rRNA and what part that play what role is known. High-resolution X-ray crystallographic structures from the early 2000's, especially from the Ramakrishnan lab but also others have shed light on the details of how transfer RNA (tRNA) molecules interacted with the mRNA in the small 30S subunit and that the 16S rRNA played a crucial role in the selection between correct and incorrect tRNA molecules[62–64]. X-ray crystallographic structures of complete 70S ribosomes from that period provided information of how metal ions participated in the non-covalent interactions between the subunits, the *intersubunit bridges*[65]. Since then a multitude of ribosomal structures have been presented, Schmeing and Ramakrishnan[4] and Wilson and Cate[59] gives very good overviews of this. In addition to that I've also found several more recent high-resolution cryo-EM structures very interesting and the study by Fischer et al.[66] provides a ribosome structure in which all rRNA modifications[67] are visualized. Recent high-resolution cryo-EM studies have also been able to show the same for human ribosomes[68].

The Bacterial Ribosome

Small subunit | *Large subunit*
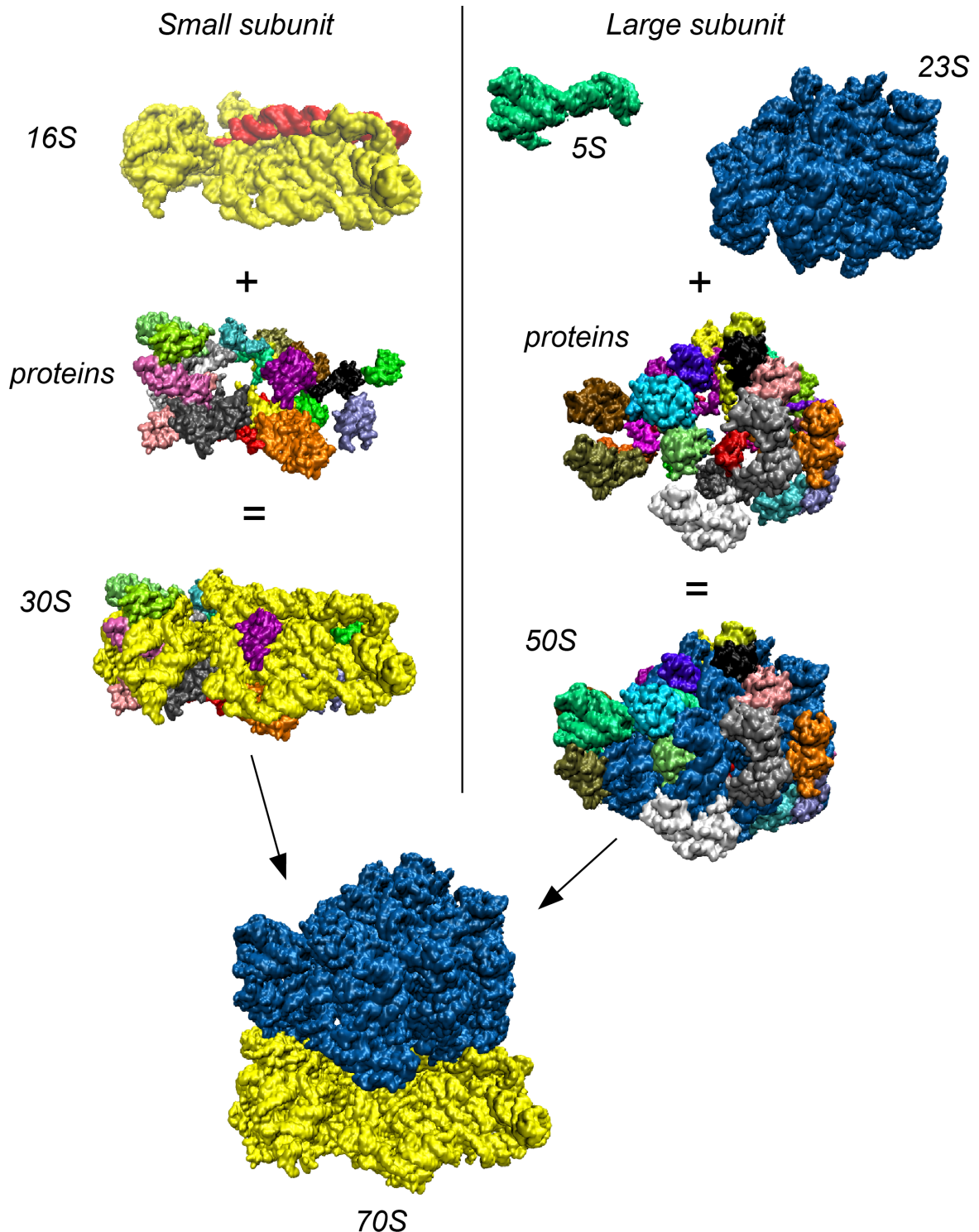
Figure 1.4. Structure of the Bacterial ribosome. Both the LSU and the SSU are complexes of ribosomal RNAs and ribosomal proteins, in total there are three ribosomal RNAs and ~50 ribosomal proteins. In the complete 70S structure the whole 30S subunit (rRNA and proteins) has been colored yellow and in the same manner the whole 50S subunit colored blue. Notice in the upper left picture of the 16S RNA that helix 44 is colored red. Figure based on PDB id 4wf1 and has been created using VMD[69].

1.2.2 Translation

Translation in bacteria is *initiated* by the binding of the mRNA to the 30S subunit, a particular sequence called the *Shine-Dalgermo* sequence in the *5′ untranslated region* (5′-UTR) of the mRNA will base pair with a sequence complementary region (the *anti-Shine Dalgermo sequence*) close to the 3′-end of the 16S rRNA[70]. A special type of tRNA molecule called *fMet-tRNA^fmet* adds to this complex and base pairs with a particular starting codon triplet, *AUG* in the mRNA, three different additional proteins or *initiation factors* help to stabilize the complex between 30S, mRNA and fMet-tRNA[fmet71]. In the early steps of the translation process, the two subunits also associate to form 70S complexes, the initiation factors are involved in this step too[72]. Once 70S ribosomes are formed the *elongation* process is started. Another protein factor, *Elongation Factor Tu* (EF-Tu) help transporting new tRNA molecules loaded with amino acids, *aminoacyl-tRNA* (aa-tRNA) to the translating ribosome, when EF-Tu delivers a new tRNA to the ribosome it uses energy supplied by the hydrolysis of guanosine triphosphate (GTP). Once the correct tRNA molecule has been delivered the peptide bond is formed by catalysis in the *peptidyl transferase center* (PTC) in the LSU[73]. After peptide bond formation, *translocation* occurs, that is the subsequent movement of the mRNA and tRNA molecules between the *A, P* and *E* site on the ribosome, during translocation the two subunits move relative to each other, the translocation process is catalyzed by *Elongation Factor G* (EF-G)[74,75]. At this point the elongation process of the protein can continue with the help of EF-Tu and EF-G until the ribosome encounters a *stop codon* in the mRNA sequence. An example of one stop codon in bacteria is *UAG,* once the ribosome encounters such a codon in the mRNA, several different *release factors* will bind to the ribosome complex instead and help release the newly formed protein from the ribosome[76]. After finishing the translation and release of the protein, *ribosomal recycling factors* in combination with EF-G splits the LSU and SSU of the 70S complex a part again and the individual subunits are now free again[77]. A simplistic overview of the translation process in bacteria is show in (Fig. 1.5).
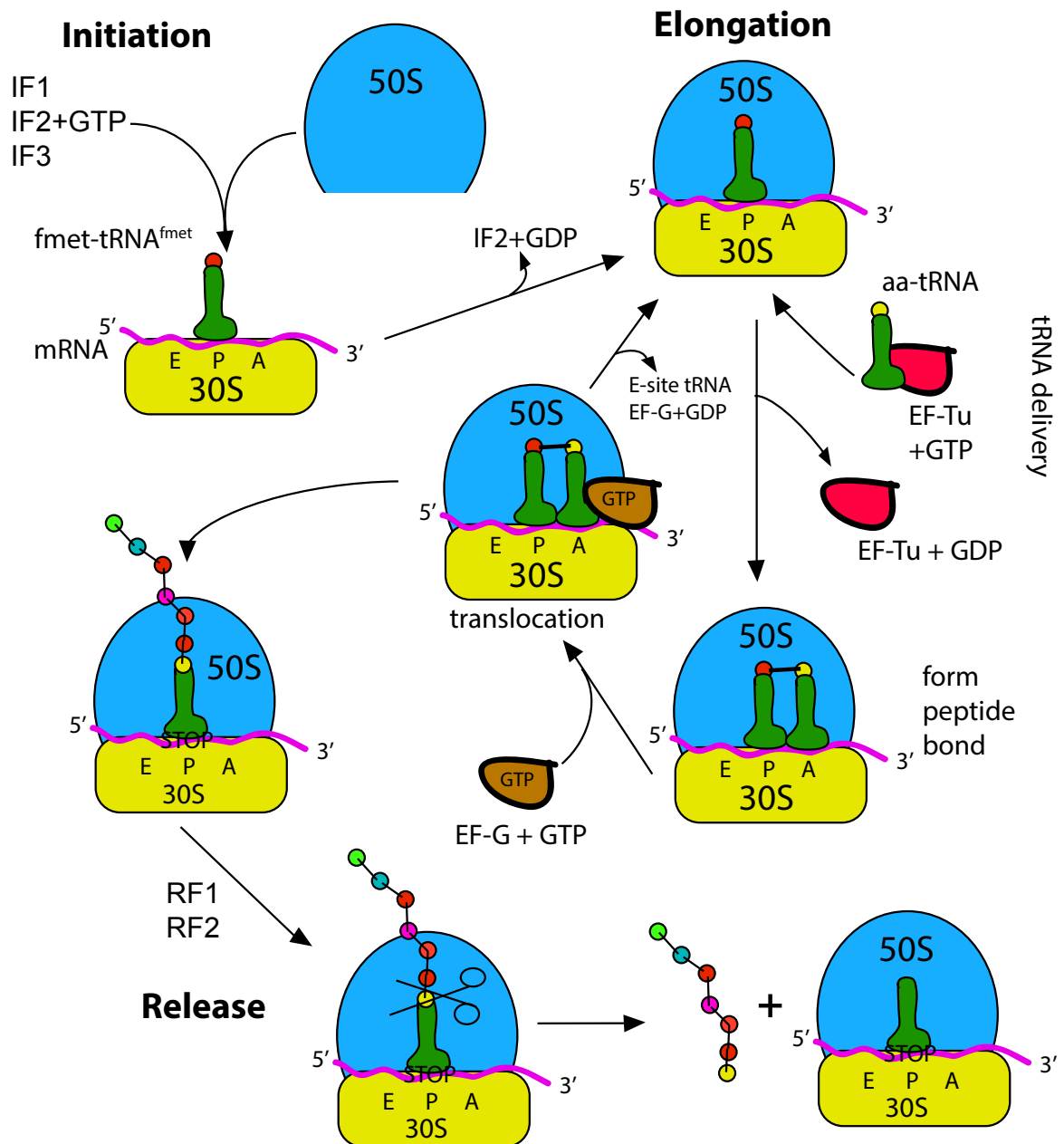
Figure 1.5. A simplistic overview of the bacterial translation process. During the initiation phase the 70S ribosome is assembled together with fmet-tRNA^fmet and mRNA with the help from initiation factors and GTP. During the elongation phase the actual protein is synthesized with the help from elongation factor Tu and elongation factor G and energy supplied from GTP. Once a stop codon is encountered in the mRNA, release factors are recruited to the complex and help release the protein from the complex. After release, ribosomal recycling can occur in which the 70S is once again disassembled into individual subunits, that part is not shown in this figure.

1.2.3   16S rRNA

As introduced in the two previous sections, the 16S rRNA of the SSU play an important role in both positioning the mRNA correctly in the subunit and in helping to select the correct tRNA molecule. In addition to that the 16S rRNA participates in several crucial intersubunit bridging interactions, which is the topic of the next section. In order to be able to perform this wide variety of tasks some parts of the 16S rRNA are highly conserved, it is also important that the 16S rRNA will maintain a particular secondary structure (base pairing pattern) and a particular

tertiary structure (three-dimensional structure) so that crucial parts are exposed to their interaction partners correctly. The maturation process of the 16S rRNA and assembly into 30S subunits is complicated and involves extensive processing of the crude rRNA transcripts by RNA-cleaving enzymes, ribonucleases (RNases) and not all details of this process is fully described yet[78]. Although detailed structural data or knowledge of subunit maturation weren't available in the early days of ribosomal research, a good understanding of the secondary structure of the 16S rRNA existed already in the 80's thanks to experiments with chemical probing techniques[79]. A secondary structure map of the *E. coli* 16S rRNA is shown in (Fig. 1.6). As aforementioned the 16S rRNA participates in several important ribosome functions (e.g. mRNA decoding and positioning) the parts of the rRNA that performs these functions, for instance stabilizing the decoding of the mRNA, which is done by two particular adenine bases (A1492 and A1493) in the ribosomal A site, are highly conserved[6,62]. There are also other regions of the 16S rRNA that are not highly conserved and more sequence varied and the 16S rRNA contains nine such regions that are called *hypervariable regions*[80,81]. The stabilization of the mRNA decoding and selection of the right tRNA molecule is a big, long standing question within the ribosome research field. Early high-resolution structures, such as those by Ogle et al.[62] showed in detail how the 16S rRNA participates is protecting the interaction between the mRNA and tRNA. Later on, multiple high-resolution structures have contributed with more information on this question, sometimes causing debate[82] and still today, new structures are published continuing the investigation of this question[83]. In the context of this discussion of tRNA selection and rejection and the correct base pairing between the mRNA and tRNA the phenomenon of *keto-enol tautomerism*[84] becomes highly relevant. Tautomerism processes are suitable for study by NMR-spectroscopy and have thus been extensively investigated with that technique[85,86].

1.2.4   Helix 44 & intersubunit bridges

In (Fig. 1.6) a secondary structure map of the *E. coli* 16S rRNA is shown, the secondary structure map shows which regions of the rRNA that are base paired and thus forms helical regions of the rRNA. Close to the 3´-end of the rRNA there is a particular helix, *helix 44* (h44) or sometimes referred to as the *penultimate stem*. Regions of h44 have been extensively studied in this PhD-project. In the mature, folded 30S subunit h44 is located on the interface area between the two subunits as can be seen in (Fig. 1.4). The fact that h44 is positioned on the interface area of the SSU allows it to interact with the LSU and along h44 several important intersubunit bridges are located, the B2a, B3, B5 and B6 bridge are positioned there[87]. The regions of h44 that participates in these bridge interactions are indicated in (Fig. 1.6). Early high-resolution structures of the 70S ribosome, such as the work by Selmer et al. provided details of the properties of intersubunit interactions and emphasized the involvement of magnesium ions in several of the bridging interactions[65]. More recent work utilizing time resolved cryo-EM techniques have then provided insight in the order of which the bridges are formed. Studies by Shaikh et al. have showed that for instance some of the central bridges such as the B2a and B3, located along h44 are some of the earliest forming bridges during subunit association[88]. Biological experiments have also contributed to the knowledge about the

intersubunit bridges and mutation studies and phenotypical characterization have for instance showed that mutations of the B6 bridge had little impact on subunit association and growth rate of bacteria whereas mutating the B3 bridge had large effects[89]. Some bridges along h44 needs to maintain contact with the LSU during the translocation process, in order to do so h44 goes through a change in shape during the translocation process and exhibit slightly different conformations in ribosome structures with different subunit rotations[90].
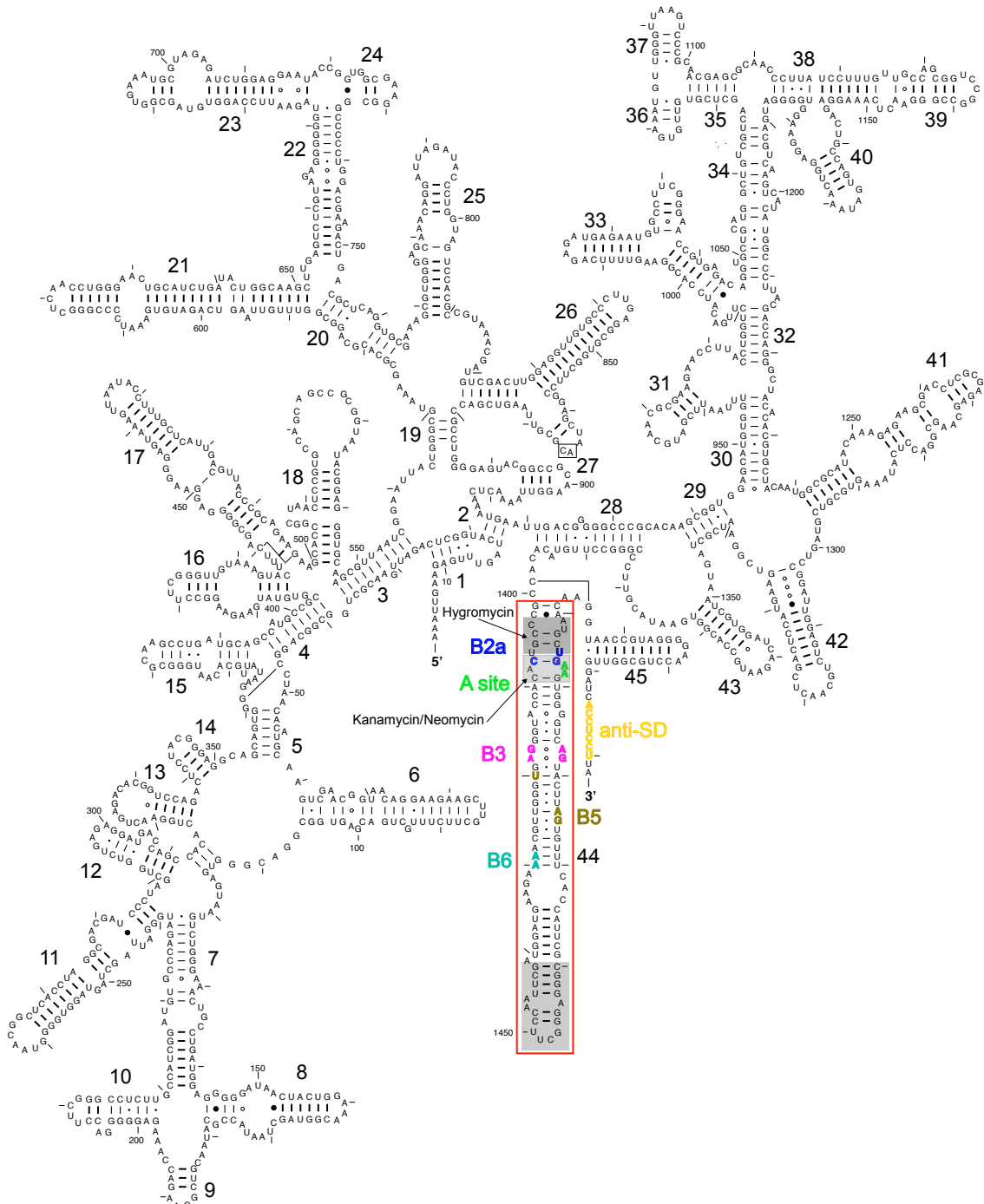

Figure 1.6. Secondary structure map of *E.coli* 16S rRNA. The black numbers represent helices. Helix 44 is boxed in red and nucleotides representing different bridging interactions located along h44 are indicated in different colors. The gray, loop part of h44 is a region that has been extensively studied in this PhD-project. Close to the 3´-end, the anti Shine-Dalgermo sequence is indicated. Gray boxes are also used to indicate the binding sites of certain aminoglycoside antibiotics in the region of the B2a bridge and the A-site. Figure based on RNA secondary structure map, used with permission from the Comparative RNA Website: https://crw-site.chemistry.gatech.edu/SIM/4C/16S/.

1.2.5 Antibiotics targeting the ribosome

The ribosome is a very central component in the cell and it is a prominent target for antibiotics, as erroneous or interrupted protein synthesis is devastating to the organism. A wide variety of low molecular weight compounds bind to different sites of the ribosome. Initially discovered as one of the first class of compounds with activity against *M. tuberculosis* the *aminoglycoside* group of antibiotics has become a famous class of ribosome targeting compounds that binds to h44 in the decoding site of the 16S rRNA in the SSU and induces misreading of the mRNA or inhibiting translocation[91,92]. Within the class of aminoglycoside antibiotics there is then different variants of these compounds that target the rRNA slightly differently and for instance *neomycin* binds close to the two special adenines involved in decoding of the mRNA whereas aminoglycosides such as *hygromycin B* binds slightly further away, they both interact with h44 though[93]. The binding regions of these antibiotics are indicated in the secondary structure map in (Fig. 1.6).

Although the aminoglycoside category of antibiotics is still in clinical use, their use is associated with quite severe side effects and both hearing loss and toxic effects on the kidneys are infamous side effects of these compounds[94]. The mechanism behind the problem of hearing loss is interesting and it is believed that the hearing loss occurs due to the unwanted, high binding affinity of aminoglycosides to the human mitochondrial ribosome and that the fact that the aminoglycosides target the mitochondrial ribosome is due to higher structural similarity between the bacterial and mitochondrial ribosome due closer ancestry[95]. As with most antibiotics, resistance mechanisms have evolved as a response to the use of aminoglycoside compounds and resistance can occur by several means, drug efflux, changed modification pattern of the rRNA and in *M.tuberculosis,* mutation of the rRNA in response to streptomycin has been observed, which is interesting and a special case since *M.tuberculosis* only has one copy of its rRNA gene[91]. The interesting aspect of *M. tuberculosis* only having one copy of its rRNA gene whereas for instance *E. coli* has seven is that these two evolutionary strategies offers both pros and cons, in *M. tuberculosis* a mutation in the rRNA can offer resistance to an antibiotic and in *E. coli* multiple gene copies offers increased flexibility and adaptability to different environmental factors[96].

Another class of antibiotics that target the SSU is the *tetracycline* group of compounds, that primarily interacts with the decoding site but at a slightly different position compared to the aminoglycoside group, the tetracyclines inhibit bacterial protein synthesis by interacting with the A site and preventing tRNA molecules to bind to that site[97]. Other groups of antibiotics also target important parts of the LSU as well. *Chloramphenicol* belonging to the chemical group of *phenicols* is known to interact in the region of PTC in the LSU and prevent the aminoacyl part of aa-tRNA to bind there and form the peptide bond[98]. The *macrolide* group of antibiotics bind to the so called *nascent peptide exit tunnel* and block the newly synthesized peptide from exiting the ribosome[99]. As with antibiotic compounds targeting the SSU resistance to the macrolide and phenicol group of antibiotics is also a problem and also these groups of antibiotics can be pumped out of the bacterial cell by *efflux pumps* and chemical modifications to rRNA nucleobases in their binding regions can also impair their desired effects[100]. The

majority of antibiotic classes, still in use today in some form, where discovered before the 1960's and until the early 2000's not many new antibiotic classes were discovered[101]. Throughout the 20th century when these antibiotics became more widely used, antibiotic resistance came along with it. But as pointed out by Lin et al., resistance to antibiotics is nothing new, as bacterially produced antibiotics is something that has been around for billions of years and has been a way for bacteria to fight each other on the evolutionary battlefield[102]. Yet it is becoming a big problem for us humans today. There are however still new, interesting, natural compounds being discovered such as the peptide compound *Klebsazolicin* that target the ribosome, such novel compounds could offer alternative chemical scaffolds to build new antibiotics on[102]. Biotechnological efforts to screen the rRNA for new, sensitive, potential antibiotic target sites could also aid in progressing the development of new antibiotics[103].

1.3   Nuclear Magnetic Resonance

The Nuclear Magnetic Resonance (NMR) phenomenon and the associated NMR-spectroscopy technique is a complicated and extensive topic to try to understand. It is also possible to choose different theoretical frameworks to describe the technique. To fully explain the results of all NMR-experiment in great detail, a quantum mechanical approach is an absolute necessity. In this chapter I will avoid quantum mechanics as much as possible and instead try to stick with a classical (i.e. non-quantum mechanical) model to paint my picture of NMR mostly using a classical description and the so-called vector model for bulk magnetization as well as phenomenological Bloch equations for various relaxation phenomena. Nevertheless, at certain points I feel obliged to try to include quantum mechanical concepts to some extent, I will do that from a more reflective and qualitative view point.

1.3.1   Spin & Bulk Magnetization

Unless referenced otherwise this first section is based on the book by Levitt[104], which I think offers a very comprehensive but yet accessible introduction to this topic. Atomic nuclei possess a property called *spin*, it is considered a quantum mechanical version of angular momentum. The spin of a nucleus is characterized by quantum numbers, I and M. The number I will dictate the range of numbers that M can take, which in turn is the number of discrete states the spin can adopt in a magnetic field. For nuclei with I = ½ there are two possible values of M = -½ and ½ and thus two possible discrete states. Some examples of nuclei with I = ½ are:

$$^{1}H, \, ^{13}C, \, ^{15}N, \, ^{31}P, \, ^{19}F$$
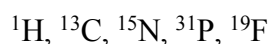
Of which ^{1}H, ^{31}P and ^{19}F are the dominating and stable isotopes of these elements and hence easy to study by NMR. ^{13}C, ^{15}N are not the naturally dominating isotopes of these elements[105] so when one wants to study these elements it becomes more complicated and isotope enrichment or long data acquisition times are necessary.

Regarding the two discrete states of spin-1/2 nuclei they are commonly referred to as the α and the β state and these states are the *eigenfunctions* or *eigenstates* of the spin angular momentum operator, $\hat{I}_z$. As emphasized by Levitt, a spin can exist in a linear combination of these two states and it is not mandatory that the spin is in any of the two eigenstates. Regardless of the detailed nature of the behavior of NMR-active spins in an external magnetic field, it is common to depict them as small magnets with an arrow indicating their magnetic moment and associate the α state as spins being parallel to the external magnetic field and the β state as spins being antiparallel to the external magnetic field, as illustrated in (Fig. 1.7).
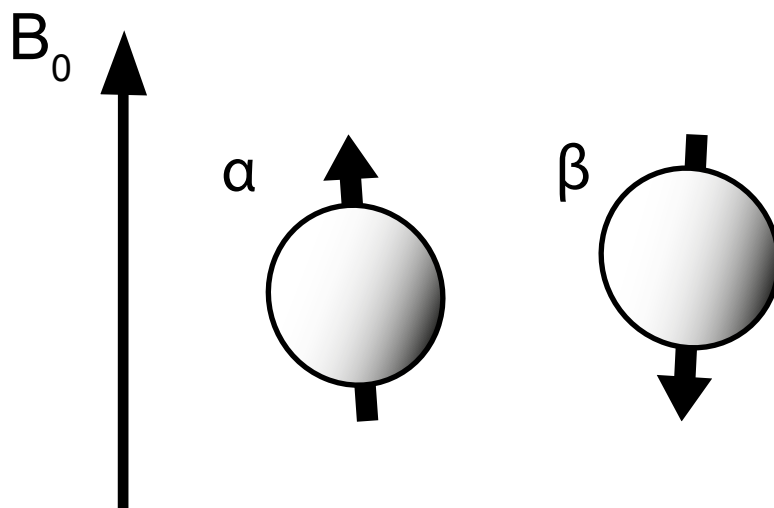

Figure 1.7. Nuclear spins with arrows indicating their nuclear magnetic moment, oriented in an external magnetic field (B_0). The α state is parallel to the external magnetic field whereas the β state is antiparallel.

The analysis of the spin angular momentum operator provides the eigenstates for the spin in the magnetic field. Similar analysis of the *Hamiltonian* operator, $\hat{H}$ will provide the energy levels associated with the different states in the shape of its eigenvalues, this matter will be further addressed on the next page.

Once an NMR-sample with spin-1/2 nuclei is put inside an NMR-magnet, several interesting things will happen. The spins will *precess* at a frequency proportional to the magnetic field, the precession is commonly depicted as the magnetic moments of the spins (i.e. their arrows in Fig. 1.7) rotating around the length axis of the external magnetic field at some angle, the frequency is called the Larmor frequency[106]. In textbooks about NMR it is also common to make analogies between this *precession* movement and the movement of macroscopic objects such as gyroscopes[107]. The constant of proportionality for this precession frequency is called the gyromagnetic ratio, commonly γ (gamma) is used as symbol. Some examples of gyromagnetic ratios, recreated from[105] are presented in Table 1 below.

Nucleus	spin	$\gamma/(2\pi)$ (MHz/Tesla)	Abundance (%)
^{1}H	1/2	42.577	99.99
^{13}C	1/2	10.708	1.07
^{15}N	1/2	-4.318	0.37
^{19}F	1/2	17.25	100
^{31}P	1/2	40.08	100

Table 1. Gyromagnetic ratios for some spin-1/2 nuclei, here given in MHz/Tesla. A proton will according to these values precess at a frequency of 600 MHz in a magnetic field of a strength of 14.1 Tesla, it is common that the field strengths of NMR-spectrometers are given in MHz instead of Tesla.

Since the gyromagnetic ratio now has been introduced, I would like to return to the matter of the two different energy levels for the α and the β state. As aforementioned they will be the eigenvalues of the *Hamiltonian* and they can be calculated as ±1/2 times the Larmor frequency, ω_0. The Larmor frequency is given by:

$$\omega_0 = -\gamma \times B_0 \qquad (4)$$

The difference in energy between the states will thus be equal to the Larmor frequency and in order to get these energy calculations correct one also need to include Planck's constant, h (J·s) in the calculations. An example calculation for a proton in a 11.7 T (B_0 in the formula above is the magnetic field strength) magnet will thus be:

$$\Delta E = 42.577 \times 10^6 \times 6.62607015 \times 10^{-34} \times 11.7 = 3.3 \times 10^{-25} J$$

Notice here that γ in Hz/T is used and h in units of J·s and not rad/T and ℏ. What is interesting to notice is that the difference in energy will depend on two factors, the gyromagnetic ratio and the magnetic fields strength. This means that for instance the energy difference for a proton spin will be larger than for a carbon spin in the same magnetic field. It also means that one can increase the energy difference between the states by using a stronger magnetic field. Both of these facts are important later on in this chapter and for the moment the fact that a stronger magnetic field increases the ΔE will be very important as the equilibrium populations of the states will be discussed.

Until now it has not emphasized that the α state when the spins are parallel with the external magnetic field is the state of low energy and the β state when they are antiparallel are the high energy state. When an NMR-sample is put inside a magnet the spins will align themselves with the field of the magnet and a small excess of the spins will prefer the low energy α state, it is

this excess of spins in the α state that ultimately will generate the NMR-signal once perturbed by a radio pulse. The fraction of spins in the low energy state will follow the Boltzmann distribution and after a Taylor series expansion it can be written like this, as explained by Levitt[108] & Cavanagh[105].

$$N_\alpha = \frac{1}{2} + \frac{1}{4}\frac{h\gamma B_0}{k_B T}$$ (5)

The units for h, γ and B_0 are the same as before, k_B is the Boltzmann constant (J/K) and T is the absolute temperature in Kelvin. The important thing with this formula is that it shows that the excess of the low energy state and hence the NMR-signal will increase linearly with the magnetic field strength, B_0. This is not the whole story though. If one look at the actual NMR-signal in relation to the signal noise, the *signal to noise ratio* (S/N) things become more complicated. As highlighted by for instance Lee and coworkers[109] if one takes into consideration the behavior of signal and noise in the radio receiver coil in relation to magnetic field strength, the actual signal to noise ratio will depend on magnetic field as:

$$S/N \sim B_0{}^{3/2}$$ (6)

And with some further search through the NMR-literature it is possible to find more complicated and useful expressions for the S/N ratio Griesinger et al.[110] presents the following formula for this:

$$S/N \sim N\,\gamma_{exc}\,\gamma_{det}{}^{3/2}\,B_0{}^{3/2}\,(NS)^{1/2}\,\frac{T_2}{T}$$ (7)

In which N is the number of spins, $\gamma_{exc}/\gamma_{det}$ are the gyromagnetic ratios for nuclei used for excitation and detection B_0 is a before, the magnetic field strength, NS is the *number of scans* and T_2 is the transverse relaxation time (see section 1.3.8) and T is the absolute temperature.

1.3.2 The vector model & effective fields

Unless referenced otherwise this section is based on the book chapter by Keeler[106], yet I've also found other literature resources such as[111] very good for the description of this topic as well.

As introduced in the previous section, once an NMR-sample is inserted into an NMR-magnet the spins will align themselves with the external magnetic field with a small excess of the spins in the low energy α state, parallel to the magnetic field B_0.

The combined magnetic moments of this excess of spins pointing in one particular direction, will create the *net magnetization vector*. The classical vector model is all about how this net magnetization vector can be modulated, twisted and turned during precession and under the

influence of radio pulses and various relaxation processes. Irradiating the NMR-sample with radio pulses will make spins flip between their α and β state and make the distribution of spins go away from their equilibrium values introduced in section 1.3.1. The vector model ignores the details of the α and β state and instead focusses on the behavior of the bulk magnetization.

In the vector model, the net equilibrium magnetization is easily visualized as an arrow (vector) along the z-axis (i.e. parallel to B_0) in a cartesian coordinate system (Fig. 1.8a). It is very common to work with the vector model in a *rotating frame of reference*. This means that the coordinate system in which the magnetization vector exists, rotate around its z-axis at some frequency (Fig. 1.8b). In the vector model, radio pulses applied to the sample will either be applied along the x or the y-axis, or some combination thereof (this corresponds to the phase of the pulse and is a setting that can be controlled in the NMR-spectrometer). It is common to allow the rotating frame to rotate at the same frequency as the frequency of the radio pulse. The difference between the radio frequency and the Larmor frequency (eq. 4 in section 1.3.1) is called the *offset* and is commonly referred to by the letter Ω. In the rotating reference frame, in the presence of a radio frequency field applied along either the x or the y-axis the magnetic fields felt by spins will come both from the external field of the magnet (B_0) as well as from the radio frequency field irradiation. The combination of these two magnetic fields will create the *effective field* this will be the field that the *magnetization vector* feels and precess around. In the presence of a radio pulse the effective field will not be aligned along the z-axis and if the difference between the frequency of the radio pulse and the Larmor precession is zero, i.e. the offset ($\Omega = 0$) the spins/magnetization vector will not feel the external field at all and only that of the radio pulse, this is illustrated in (Fig. 1.8c). One crucial thing to emphasize is that the magnetization vector of a spin in the rotating frame will rotate around the z-axis if its offset is non-zero i.e. if $\Omega \neq 0$ for a spin relative to the frequency of the applied pulse its vector will precess in the x,y-plane with a frequency = Ω. This means that if $\Omega = 0$ the magnetization vector will remain statically along the x or y-axis in the rotating frame of reference.
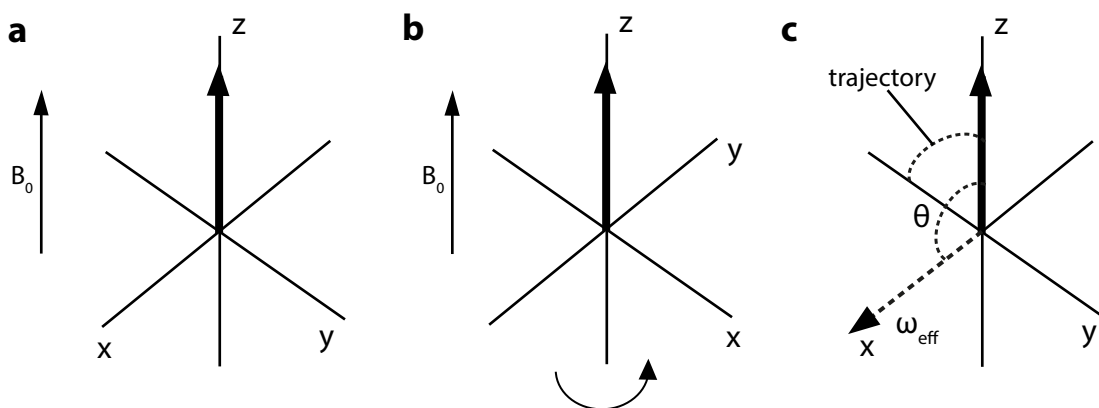

Figure 1.8. Net magnetization vector. a) The net magnetization vector for the equilibrium magnetization, aligned along the +z-axis. b) In the commonly used rotating frame of reference the coordinate system rotates around the z-axis at some frequency, commonly that of the applied radio pulse. c) A radio frequency pulse is applied along the x-axis, with a frequency equal to the Larmor frequency ($\Omega = 0$) the theta angle then becomes $\pi/2$ rad = 90° and the magnetization vector will follow the indicated trajectory and rotate towards the negative y-axis, notice that B_0 has been taken away in this figure to indicate that the spins will not feel the external magnetic field because of $\Omega = 0$.

Some properties of the effective field are given by the following formulas, adapted from Keeler[106]. These formulae are very central to this chapter, to my understanding of NMR and a prerequisite for the later sections of this chapter discussing $R_{1\rho}$ relaxations. The magnitude of the effective field (ω_{eff}) is given by:

$$\omega_{eff} = \sqrt{\omega_1^2 + \Omega^2} \qquad (8)$$

In which ω_1 is the magnitude the magnetic field from by the radio pulse expressed in Hz and Ω is as before the offset, i.e. the difference between the frequency of the radio pulse and the Larmor precession frequency. Regarding the expressing of the magnetic field from the radio pulse in Hz. I will come back to that in the next section. The other important formula for the effective field is:

$$\tan \theta = \frac{\omega_1}{\Omega} \qquad (9)$$

Of which I much rather prefer the expression:

$$\theta = \frac{\pi}{2} - \arctan\left(\frac{\Omega}{\omega_1}\right) \qquad (10)$$

Equation 10 is the *theta angle* and that is the angle between the effective field and the z-axis in the rotating reference frame, in the example in (Fig. 1.8c) when $\Omega = 0$ theta will be $\pi/2$ radians, that means the effective field will be aligned along the x- axis in the x,y-plane and have no z-component. In such example the net magnetization vector will leave the z-axis and start rotate around the effective field along x-axis (Fig. 1.8c).

1.3.3 Fourier Transform NMR & Pulse basics

In the early days of NMR-spectroscopy, continuous wave NMR-spectroscopy was used. This meant that an NMR-sample was put inside a magnet and that the sample was irradiated with a continuous radio frequency field. The field strength of the magnet or the radio frequency was then varied[112]. Here one can connect to the offset term (Ω) in section 1.3.2, varying the radio frequency is equal to varying Ω and once $\Omega = 0$ the spins will only feel the magnetic field of the radio wave, they will *resonate* and absorb maximum of the radio frequency energy.

Later on, Fourier Transform NMR (FT-NMR) was invented[113]. In FT-NMR all nuclear spins are excited at once, meaning that the magnetization vector of all spins will be rotated down in the x,y-plane, the *transverse plane*. Once the magnetization vector is rotated down in transverse plane the radio pulse is turned off. Then the magnetization vector will rotate in the transverse plane perpendicular to the z-axis/main magnetic field of the NMR-spectrometer, the magnetization vector will then induce a small, oscillating electric current in a receiver coil. This

induced current will then decay away due to relaxation mechanisms (section 1.3.8), this oscillating decaying current is called a *free induction decay* (FID)[107].

In FT-NMR the FID is then *Fourier Transformed* to give the final spectrum. Several parameters of how the recording of the FID is done will affect important properties of the final spectrum. The magnitude of the FID signal is recorded or *sampled* at discrete data points. The distance between the sampled points will decide the range of frequencies that is recorded, the *spectral width*[114]. The *Nyquist Sampling theorem* also dictates how often a periodic signal should be sampled in order to accurately describe all the containing signals, it states that the highest frequency that will be correctly sampled will be equal to the inverse of two times the distance between the sampling points[114]. The duration, or how long the FID is, will affect the resolution of Fourier Transformed spectrum and the resolution will be the inverse of the acquisition time[115]. A schematic example of a one-dimensional (1D) pulse acquire FT-NMR experiment is outlined in (Fig. 1.9).
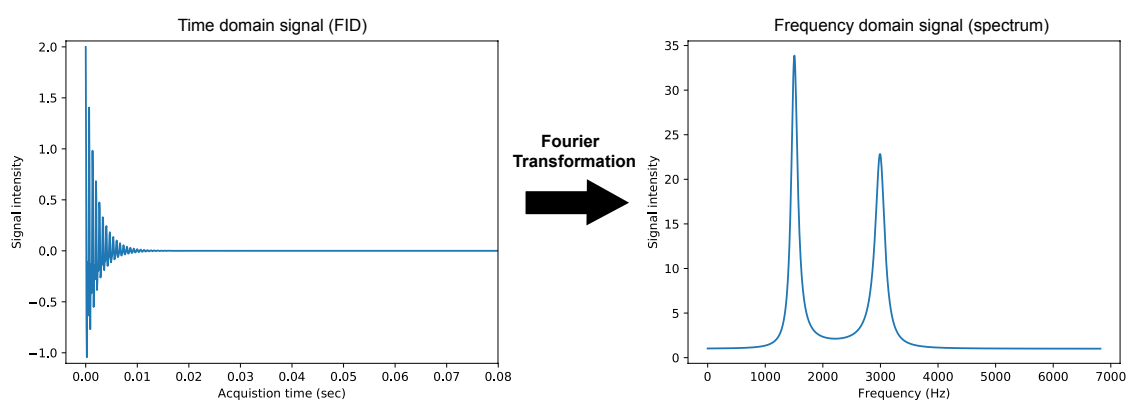

Figure 1.9. Schematic FT-NMR. A free induction decay and the corresponding Fourier Transformed spectrum.

An example of a simple NMR-experiment could be a 1D pulse acquisition experiment, such an experiment could yield a FID and spectrum very similar to (Fig. 1.9). In such experiment the sample is irradiated by one strong, short radio frequency pulse, immediately after a FID is recorded. The radio pulse is centered around some frequency in the frequency range of interest and will excite all spins within the spectral width, depending on properties of the pulse.

The applied radio pulse will rotate the magnetization vector as previously illustrated in (section 1.3.2, Fig. 1.8c). If the radio pulse is strong (high power in Watt) the magnetization vector will rotate fast and if the power is low it will rotate slow. A pulse that rotate the magnetization vector $90° = \pi/2$, is called a *ninety-degree* pulse, if one pulse for longer the magnetization will rotate, 180°, 270° and so on. The inverse of a 360° pulse is equal to the field strength of the pulse in Hz (ω_1), this can also be stated as:

$$Pulse\ strength\ in\ Hz\ = \frac{1}{4 \times P_{90}} \qquad (11)$$

This then connect back to equation 10 in section 1.3.2. If one considers a strong pulse where ω_1 is large, θ will be close to π/2 even if the offset (Ω) becomes big, such pulse is called a hard pulse[116]. A hard pulse will rotate the magnetization vector for spins with different Ω almost perfectly perpendicular their effective field located along either the x or y-axis. A pulse with low power (small ω_1) will not rotate spins with large Ω perfectly perpendicular to the transverse axes, in fact if Ω is large enough relative ω_1 the magnetization vector will not be rotated or excited at all, such a pulse is called a *soft pulse*[116].

Commonly an NMR-spectrometer is setup with a default power level for each of the radio channels. If one calibrates the duration of a ninety-degree pulse at this power level one can calculate the power level that generates a pulse with a particular field strength using the following formula[117]:

$$\Delta dB \ = \ 20 \ \times \ log_{10} \left(\frac{P_{new}}{P_{ref}} \right) \qquad (12)$$

This equation requires some additional explanation. P_{new} is the duration of the 90°-pulse that corresponds to the wanted field strength (eq. 11), P_{ref} is the duration of the initially calibrated 90°-pulse at for instance the default power level. ΔdB is the difference in decibels between these power levels, it is common to state NMR-pulse power levels in decibels, this is called *attenuation* and represents how much the power is reduced or "attenuated" relative to the default power level of the channel[118]. It is also common praxis in NMR-literature and documentation to state power levels as the corresponding to the duration of a particular 90°-pulse, this can be extremely confusing when introduced to it, an example can be:

1. A probe specification sheet states that the maximum allowed pulse duration is 400 ms @ 35 μs

2. The default power level for that channel is 8.6W corresponding to -9.34 dB

3. If the duration of a 90°-pulse at default power level (i.e. 8.6W = -9.34 dB) is 7.8 μs

4. Then the power level for a 35 μs pulse is calculated using (eq. 12) is 0.43W = 3.69 dB

5. So, 400 ms @ 35 μs, actually means 400 ms @ 0.43W/3.69 dB

1.3.4 Chemical Shift

Nuclei inside a molecule will have different electronic environments and different electron density surrounding them. High electron density surrounding the nucleus will shield it from the external magnetic field of an NMR-magnet and it will precess at a lower Larmor frequency, this means that the B_0 term in equation 1 will be felt less by the nucleus. The opposite case with lower electron density will of course increase the Larmor frequency. Because the Larmor frequency will change in NMR-spectrometers with different field strengths and inaccuracies in

B_0, it is routine to express the frequency as a fraction relative to a reference compound, the obtained unitless number is called the *chemical shift* (eq. 13) and is given as parts per million, *ppm*[119].

$$\delta = \left(\frac{\upsilon_{sample} - \upsilon_{reference}}{\upsilon_{reference}}\right) \times 10^6 \qquad (13)$$

In equation 13 above υ_{sample} represents the Larmor frequency in Hz of the sample and $\upsilon_{reference}$ the frequency of the reference compound. The electronegativity of surrounding elements, temperature, buffer conditions and possible aromaticity and molecular structure will all influence the chemical shift of nuclei. Chemical shift of nuclei in various functional chemical groups is possible to find in a huge number of textbooks. For biological molecules it can be more complicated to find coherent chemical shift information. For RNA molecules one can find chemical shifts in the *Biological Magnetic Resonance Bank* or short just *BMRB*[120]. In this database researchers can deposit chemical shifts of different biological molecules. One can then look through the database and correlate chemical shifts with structural features that can be extracted from the *Protein Data Bank*[121] or the *FRAbase*[122] for instance. An example of average chemical shifts extracted from the BMRB-database and correlating with base pair properties of nuclei in nucleobases of RNA are exemplified in (Table 2) below. This will be further discussed in section 1.3.10.

Nucleobase	Nucleus	Atom	Base pair	Chemical shift (ppm)	Standard deviation	Number of data points
Adenine	^{1}H	H8	Watson-Crick	7.93	0.23	73
Adenine	^{13}C	C8	Watson-Crick	139.72	0.69	73
Guanine	^{1}H	H8	Watson-Crick	7.59	0.37	141
Guanine	^{13}C	C8	Watson-Crick	137.13	1.36	141
Cytosine	^{1}H	H6	Watson-Crick	7.65	0.21	116

Cytosine	^{13}C	C6	Watson-Crick	141.08	0.99	116
Uridine	^{1}H	H6	Watson-Crick	7.81	0.16	74
Uridine	^{13}C	C6	Watson-Crick	142.02	0.70	74

Table 2. Average chemical shifts of some aromatic nuclei of RNA extracted from the BMRB database and correlated to base pairing properties from the FRAbase. As one can see there is very little difference in chemical shift between the ^{1}H for different bases whereas for instance the ^{13}C chemical shift between adenine and guanine C8s differ slightly more.

1.3.5 J-coupling

The chemical shift provides important information on molecular structure and presence of certain functional groups in a molecule. Another important parameter that provides information on molecular structure and also allows for *transfer* of magnetization between NMR-active nuclei is *scalar coupling* or *J-coupling,* J-coupling will split NMR-signals into several peaks separated by a constant value, the *coupling constant*[123]. J-coupling is most commonly first encountered in undergraduate organic chemistry where one can use its resulting peak splitting patterns to elucidate the connectivity of protons in small organic molecules. Here I instead want to focus on the possibility of using J-coupling for *polarization transfer* or transfer of magnetization between spins. This is a matter that is complicated and hard to give justice in this context but it is very important to NMR-spectroscopy. Unless referenced otherwise this section is based on the book by Keeler[124]. If a J-coupled spin for instance a proton, an H6 in a cytosine nucleobase (that is J-coupled to the neighboring H5 with a coupling constant of ~7 Hz, see Fig. 1.10) is excited with a 90°-pulse applied along the x-axis this will generate magnetization of this spin along the minus y-axis. This scenario is similar to rotation of the magnetization vector in (Fig. 1.8c).
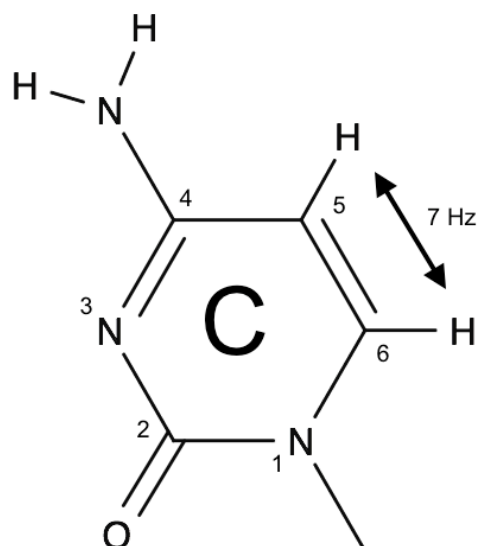
Figure 1.10. Cytosine nucleobase where atom numbers has been indicated as well as the J-coupling constant between H5 and H6.

This scenario is unfortunately not possible to describe further using only rotations of a magnetization vector, instead it is necessary to use *product operators*. Product operators are quantum mechanical operators describing the magnetizations evolving in J-coupled spin systems, in a similar manner as the spin angular momentum operator ($î_z$) introduced in section 1.3.1. The properties of and how product operators will respond to different pulses, delays etcetera is best to look up in the books such as Levitt[125] or Keeler[124]. These literature sources provide some pragmatic tools to work with this and some simple rules adapted from[124] of how magnetizations will evolve both among non-J-coupled and J-coupled spins, are shown in Table 3 below.

X	$\rightarrow$	Z	$\rightarrow$	-Y	$\rightarrow$	-Z	$\rightarrow$	Y
Y	$\rightarrow$	Z	$\rightarrow$	X	$\rightarrow$	-Z	$\rightarrow$	-X
Z	$\rightarrow$	X	$\rightarrow$	Y	$\rightarrow$	-X	$\rightarrow$	-Y
ZZ	$\rightarrow$	X	$\rightarrow$	YZ	$\rightarrow$	-X	$\rightarrow$	-YZ
ZZ	$\rightarrow$	Y	$\rightarrow$	-XZ	$\rightarrow$	-Y	$\rightarrow$	XZ

Table 3. Rules for how product operators will transform during rotations around different axis. The leftmost, boxed column represent an axis that the operator rotates around and the letters in the other columns represent subscripts for operators. The three first rows show how operators will transform for a single non J-coupled spin, rotations around the X and Y-axis will depend on the presence of a radio pulse. For instance, in the first row if an operator $î_z$ (subscript Z) is subjected to a 90°-rotation (a radio pulse) around the X-axis (first row, leftmost column) it will transform into operator -$î_y$. In a similar manner using the second row, if an operator $î_x$ is subjected to a 90°-

rotation (radio pulse) around the Y-axis will transform into -$\hat{i}_z$. The two bottom rows shows how product operators will evolve if a J-coupled spin is excited and rotated down in the X,Y-plane and left to precess freely for some time. If the terms in the rightmost column are left to evolve they will evolve back into the terms in the third column. It is important to emphasize that when product operators such as YZ ($\hat{i}_{yz}$) are created, the individual components Y and Z will obey the rotation rules in the first three rows, this means that if a 90°-pulse along X is applied to YZ it will transform into -ZY.

Regarding the buildup rates of the different operators the starting operator will decay as a cosine function and the forming operator will build up as a sine function. For rotations around the X and Y-axis (induced by radio pulses, first two rows in table 3) the rotation frequency will be the pulse strength in Hz (ω_1), see equation 11 in section 1.3.3. For rotations around the z-axis in (third row, table 3), the frequency will be the offset (Ω), in both cases multiplied by 2π. For J-couplings (bottom two rows in table 3) the frequency will be the coupling constant (Hz) multiplied by π. I have found the simple rules introduced in table 3 mighty powerful during my PhD-years and able to explain the structure of many NMR-experiments. If one wants to go in greater detail regarding this matter and explore the evolution of J-coupled system in more in depth and possibly in the presence of radio pulses one could work with the Liouville-Von-Neuman equation and Hamiltonians for both J-couplings and radio pulses, I've found the papers by Allard et al. regarding this very interesting[126,127]. At this point it is time to return the J-coupled H6 spin introduced in the beginning of this section. At the time I left it, it had just been excited by a 90°-pulse and $\hat{i}_{z(H6)}$ had turned into -$\hat{i}_{y(H6)}$. Notice that I now added (H6) in the subscript to indicate that this operator/part of operator belongs to this H6 proton. If -$\hat{i}_{y(H6)}$ is now allowed to *evolve* for some time t, the following will happen:

$$-\hat{i}_{y(H6)} \rightarrow -cos(\pi J\, \hat{i}_{y(H6)}) + sin(\pi J\, \hat{i}_{x(H6)}\, \hat{i}_{z(H5)}) \qquad (14)$$

If one waits ½ J, like in (Fig. 1.11) below -$\hat{i}_{y(H6)}$ will completely decay away and a maximum of $\hat{i}_{x(H6)}\hat{i}_{z(H5)}$ will build up. In realistic situations one doesn't only have one H6 spin to consider, in reality one would have a number of spins that was excited by the initial 90°-pulse and they will all have different Ω relative to the applied pulse.
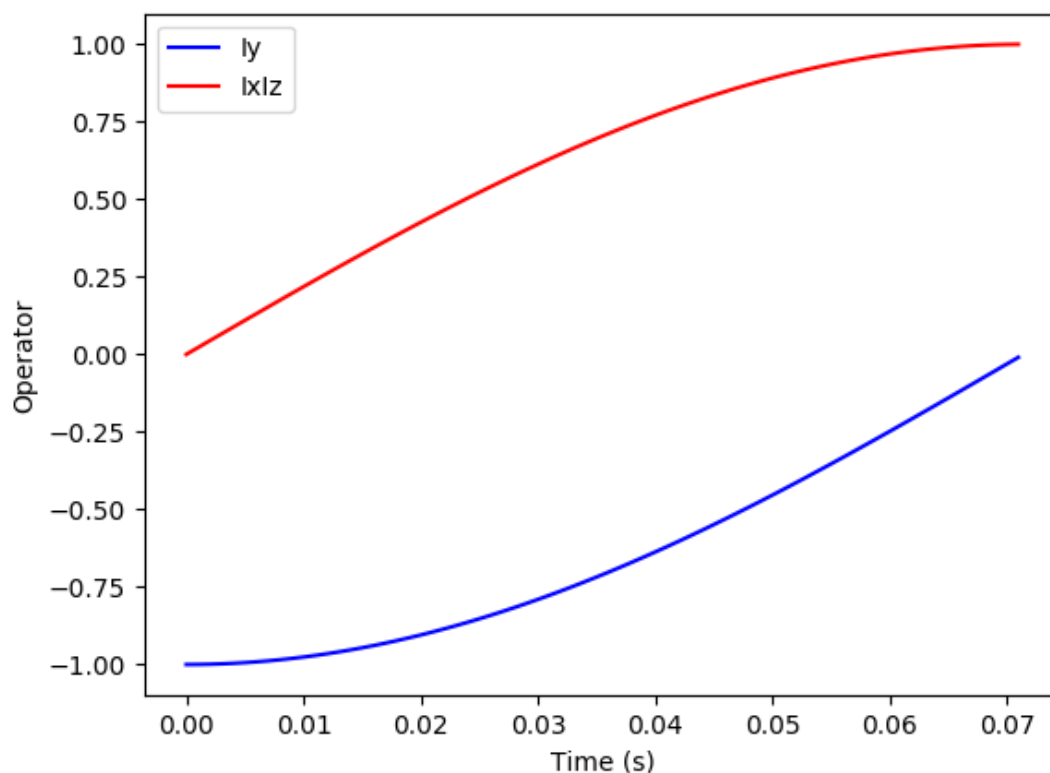
Figure 1.11. Example of allowing the operator $-\hat{i}_{y(H6)}$ of the coupled spin H6 evolve freely for 70 ms during which it disappears and instead $\hat{i}_{x(H6)}\hat{i}_{z(H5)}$ will be created.

If one were to wait for the J-coupling of the H6 spin to evolve, these other spins with non-zero Ω will start to rotate in the x,y-plane and create all kinds $\hat{i}_x$ and $\hat{i}_y$ operators, that might not be wanted. To avoid this one can use a 180°-degree pulse or a *spin-echo*, this is illustrated in (Fig. 1.12).
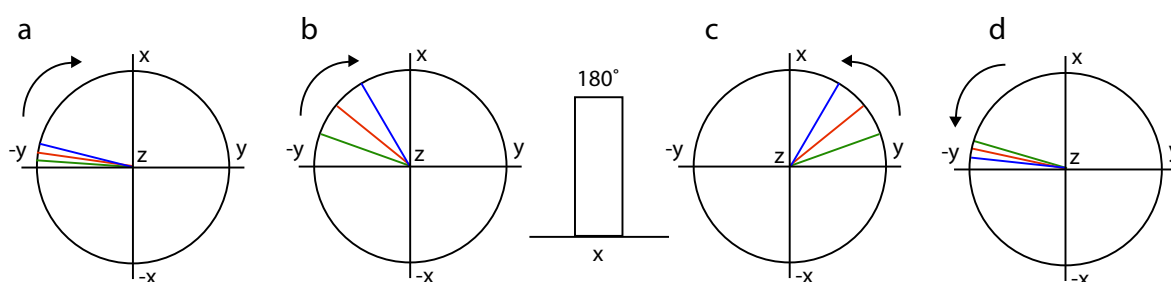
Figure 1.12. A schematic spin-echo that will refocus operators. a) Three spins with different Ω will start to rotate in the x,y-plane from $-\hat{i}_y$, according to the third row in table 3. b) A moment later they have rotated to create both some $\hat{i}_x$ and $-\hat{i}_y$ at that point (between b and c) a 180°-pulse is applied along the x-axis. c) The 180°-pulse will rotate all $-\hat{i}_y$ around the x-axis, according to table 3, first row and create $\hat{i}_y$. However, this $\hat{i}_y$ will not continue towards the +y-axis, it will instead go back to the -y-axis and become $-\hat{i}_y$ again because the refocusing effect of the 180°-pulse. d) After waiting another while all operators will have rotated back to where they started becoming $-\hat{i}_y$ again.

With some of the theory of J-coupling transfers introduced, I would like to turn to a complete example of a multi pulse coherence transfer pulse block, the "homonuclear *INEPT*" transfer step. This is used in the *SELOPE approach* described by constituent Paper II of this thesis and allows for spectral decrowding in non-isotopically labeled RNA samples[128]. The complete appearance of this pulse program block is shown in (Fig. 1.13).
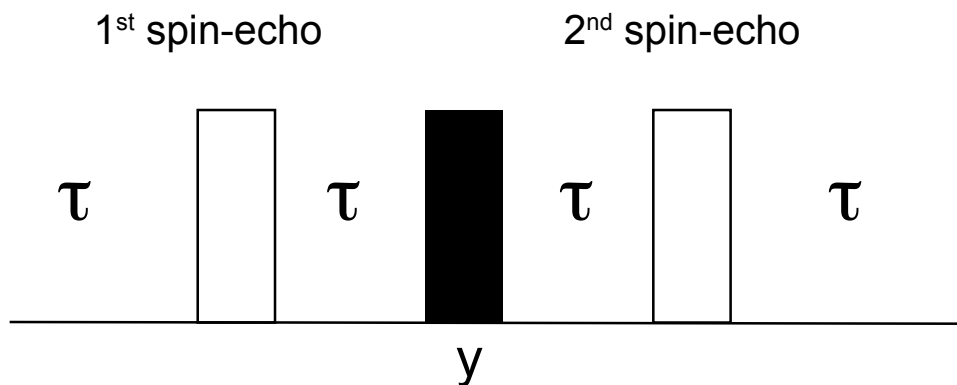
1st spin-echo   2nd spin-echo

$$\tau \quad | \quad \tau \quad \blacksquare \quad \tau \quad | \quad \tau$$

y

Figure 1.13. The "homonuclear INEPT" transfer block used in the SELOPE approach. It consists of two spin-echoes surrounding a 90°-pulse. If no phase is indicated it implicitly means the phase of the pulse is x.

If one now again turns to the example of the excited, H6 spin (J-coupled to H5) in a cytosine in RNA. As one can see in (table 2) its chemical shift is similar to that of H8 and H2 of adenines and guanines and if one excites this spectral region with a 90°-degree pulse of x-phase, H8:s and H2:s will be excited along with it. So, after such 90°-degree pulse one will have $-\hat{\imath}_{y(H6)}$, $-\hat{\imath}_{y(H8)}$, $-\hat{\imath}_{y(H2)}$ with various offsets, what will now happen if these operators enter the transfer block shown in (Fig. 1.13) and the delay τ is set to $1/4J_{H6/H5} \approx 35$ ms. During the first spin-echo the $-\hat{\imath}_{y(H8)}$, $-\hat{\imath}_{y(H2)}$ operators will be refocused by the 180°-degree pulse and remain unchanged while the $-\hat{\imath}_{y(H6)}$ operator will be unaffected by the 180°-degree pulse and the coupling will evolve and transform $-\hat{\imath}_{y(H6)}$ into the product operator $\hat{\imath}_{x(H6)}\hat{\imath}_{z(H5)}$ as outlined in equation 14. At that point a hard 90°-degree pulse of y phase is applied, this will now affect both parts of the product operator $\hat{\imath}_{x(H6)}\hat{\imath}_{z(H5)}$ because both parts of this product operator is a proton spin (product operators of course exist for coupled spins of different elements, see Keeler[124] for this). How the parts of the product operator will be affected by the 90°-degree pulse one can figure out using the rules in Table 3 again. Each individual part of the product operator $\hat{\imath}_{x(H6)}\hat{\imath}_{z(H5)}$ will obey the rules in Table 3 and once again I would like to emphasize that product operators for protons J-coupled to heteronuclei ^{13}C, ^{15}N, ^{31}P etcetera will also follow these rotation rules as long as the pulses are applied on the channel for the heteronucleus. But in this case the 90°-degree pulse along y will transform $\hat{\imath}_{x(H6)}\hat{\imath}_{z(H5)}$ into $-\hat{\imath}_{z(H6)}\hat{\imath}_{x(H5)}$. During the second spin echo the $-\hat{\imath}_{y(H8)}$, $-\hat{\imath}_{y(H2)}$ will once again be refocused by the 180°-degree pulse and remain unchanged while $-\hat{\imath}_{z(H6)}\hat{\imath}_{x(H5)}$ will now transform into $-\hat{\imath}_{y(H5)}$ due to the J-coupling. So, by using the trick of the "homonuclear INEPT" transfer in (Fig. 1.13), all magnetization from the H6 is now transferred to the J-coupled H5. As can be seen in Table 2, the chemical shift of H5:s are very different compared the H6/H8/H2 chemical shift, this means that by transferring the magnetization to H5 the region of H6/H8/H2, the aromatic region now becomes less crowed

and it is now possible to observe and study H8/H2 signals that previously were overlapping with H6:s. The transfer step described here and used in the SELOPE approach of course also transfers the signal from H6 of uridines in a similar manner and, depending on the sugar pucker, the signal from H1′ to H2′.

1.3.6 Nuclear Overhauser Effect

In order to make my picture of NMR less incomplete I feel obliged to address another NMR-related phenomenon, in addition to chemical shifts and J-couplings and that is the *Nuclear Overhauser Effect* (NOE). The NOE is an interaction through space and it is a *cross-relaxation* phenomenon, *relaxation* will be further discussed in section 1.3.8. For now, relaxation can be thought of as spins going from the high energy (β state) to the low energy (α state) and as mentioned in the earlier sections it is the excess of spins in the α state that will yield the NMR-signal and transitions from the α to the β state can be induced by a radio pulse. For the NOE, spins that are close to each other in space can go through relaxation processes such that the α and the β state change for both spins simultaneously, such transitions are called *zero quantum transitions* and *double quantum transitions* such transitions cannot be induced by radio pulses, transitions where one the state of one spin change is called *single quantum transition* and can be induced via a radio pulse[129]. An example illustrating how the α and β populations in a cytosine nucleobase can change via a double quantum transition is shown in (Fig. 1.14).
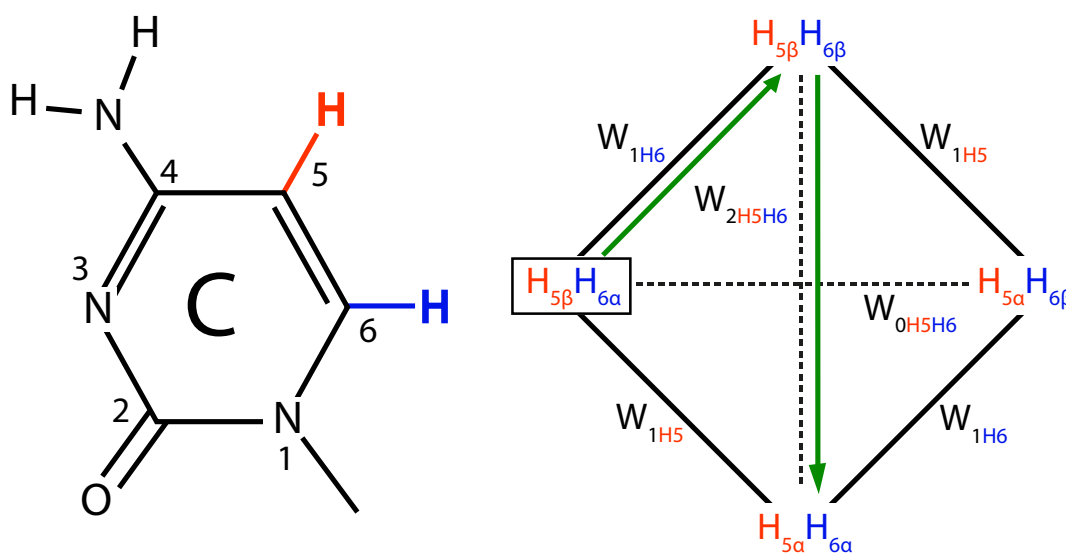

Figure 1.14. NOE in a cytosine nucleobase. (Left) The H5 and H6 protons are close in space since they are covalently bound to a flat aromatic ring. (Right) a diagram showing possible transition between α and β state of for the two protons. The solid diagonal lines indicate single quantum transitions possible to achieve using a radio pulse, the horizontal line indicates a zero quantum transition and the vertical, dashed line a double quantum transition, W_{xx} represent different transition probabilities, the diagram is based on the theory presented in[129]. If one considers a cytosine base in which the H5 is in the β state and the H6 in the α state (the box in the right diagram) if a radio pulse is applied selectively to the H6 it will go through a single quantum transition (green arrow/W_{1H6}) to the $H_{5\beta}H_{6\beta}$ state, this state can then go through a double quantum transition (W_{2H5H6}) to the $H_{5\alpha}H_{6\alpha}$ state in

which both spins have relaxed to their low energy state. That is one example of how the NOE can work and by doing the double quantum transition to the $H_{5\alpha}H_{6\alpha}$ state the population/excess of spins in the $H_{5\alpha}$ state has increased and as stated in section 1.3.1 that excess is the origin of the NMR-signal so the subsequent H5 signal will increase if one where to excite the H5 spins again.

The transition probabilities and hence the magnitude and rate that these cross-relaxation phenomena occur with, depends in an intricate manner on the distance between the two spins as $1/r^6$, the presence of random motions at different frequencies the *spectral densities,* the so-called *correlation time* (which is the time it takes for the molecule to rotate one radian in the solution) and the gyromagnetic ratios of the spins[129,130]. The NOE phenomenon can occur both along the axis of the main magnetic field (z-axis), as *longitudinal cross relaxation* and also within the transverse plane (x,y-plane) as transverse cross-relaxation. The cross-relaxation phenomenon that occurs along the z-axis is utilized in the classic NOESY experiment[131]. Some more details on practical applications of the NOESY experiment in the RNA context is given later in section 1.3.10. Cross-relaxation in the transverse plane can only occur in the presence of a static radio frequency irradiation, a so-called *spin-lock* (this will be further discussed in section 1.3.9) the NOE in the transverse plane in the presence of a spin-lock is referred to as a *Rotating frame Overhauser Effect* or *ROE* and is utilized in the ROESY experiment[131,132]. The ROE is important in later sections (1.3.9 and 4.4) since it can give rise to cross-relaxation artefacts in ^{1}H-$R_{1\rho}$ experiments[133]. This is further discussed in these later sections.

1.3.7 2D NMR-spectroscopy

In (Fig. 1.9) previously, a schematic FID and spectrum that could come from a 1D pulse acquisition experiment was shown, that was a good example of 1D NMR-spectroscopy. In two dimensional, *2D NMR-spectroscopy* or any N-dimensional NMR-spectroscopy the *pulse sequences* and data acquisition is more complicated. The basis of 2D NMR-spectroscopy is to record several 1D spectra that are then Fourier Transformed in a particular manner to get the final 2D-spectrum. The typical 2D NMR-experiment consists of a *preparation* period such as a 90°-pulse, followed by and incrementable *evolution* period *t1*, followed by a *mixing* period where some kind of information transfer between spins happen, all this ultimately followed by direct detection/acquisition during *t2*[134]. To illustrate what happens during a 2D NMR-experiment, a semi quantitative illustration is presented in (Fig. 1.15).
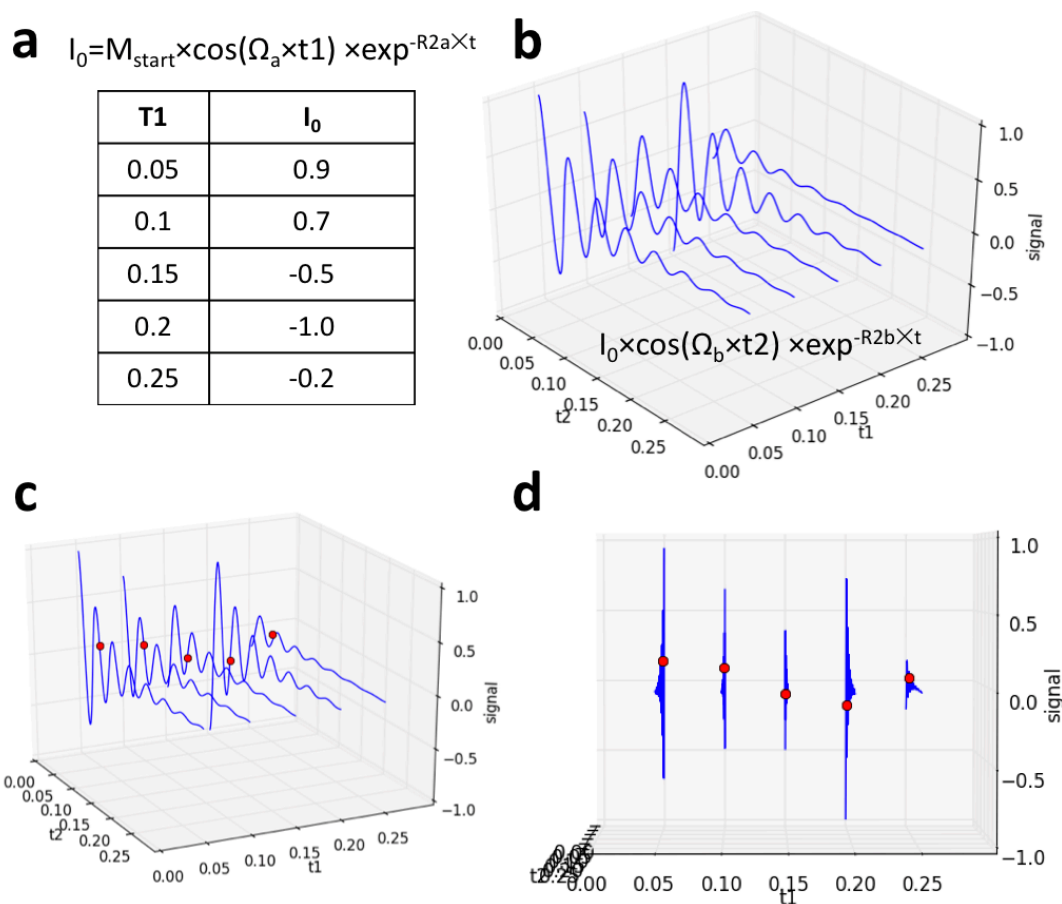
Figure 1.15. Schematic representation of 2D time domain data, generated with Python. a) During preparation and evolution some starting value of magnetization (here M_{start}) will be allowed to precess in the x,y-plane relative to the detector of the NMR-spectrometer. If one were to measure the signal magnitude along one of the axis, let's say the y-axis the amount of signal along that y-axis will follow a cosine function and depend on the frequency of precession (Ω_a) and the duration of evolution time, t1. b) Depending on the duration of t1 the magnitude of magnetization available at the start of the directly detected FID i.e. I_0 will vary. This gives the series of FIDs in b). c & d) If one now looks along the red dots positioned at t2=0.05 as a function of t1 one can clearly see an oscillating periodic function, $\cos(\Omega_a t1)$. This is the principle of 2D NMR that the evolution period t1 will give this oscillation shared between all directly detected FIDs.

The possibility to transfer magnetization between spins via J-coupling and through space via the NOE, previously introduced in section 1.3.5 and 1.3.6 allows for the creation of 2D NMR-spectra in which a multitude of information can be contained. By using a coherence transfer step, via J-coupling to a heteronucleus such as ^{13}C or ^{15}N, the chemical shift of these nuclei can evolve during the t1 period, then one will observe the chemical shift of different nuclei in the two different Fourier transformed dimensions, this is the basis for commonly used heteronuclear correlation experiments such as HSQCs and HMBCs etcetera[135]. By allowing cross-relaxation to happen during a mixing time, cross peaks which magnitude depends on proximity of the spins will be observed as in the NOESY experiment[131].

Once again, the process outlined in (Fig. 1.15) is a simplification. Just as directly detected FIDs are detected relative to two receivers (quadrature detection) to create a complex input for the Fourier Transform the same is necessary in the indirect dimension, this can for example be done with States[136].

1.3.8 Relaxation & Exchange

When a NMR-sample is put inside a magnet and excited with a strong radio pulse the magnetization vector of all spins will be rotated into the x,y-plane the vectors will then rotate in the plane relative to a detector and create a FID as illustrated in section 1.3.3. Ultimately the signal will die out, it will *decay* away and eventually the equilibrium magnetization along the +z-axis will be reestablished. This happens due to *relaxation*. As emphasized in many previous sections in this chapter I am once again approaching a topic that is complicated and hard to give justice in this context but it is an absolute necessity to address it since it is very central to this thesis and PhD-project.

The major reason why the FID signal decays is because the signal or magnetization present in the x,y-plane disappears, this is called *transverse relaxation* and is described by the time T_2 (s)[137]. The other, central type of relaxation is called *longitudinal relaxation* described the value T_1 (s) and is the time for the magnetization to return to its equilibrium value along the +z-axis[137]. From now on I will work with these values in the shape of their inverse values R_1 and R_2 (both in units of s^{-1} then), referred to as the longitudinal and transverse *relaxation rate constants*. The detailed mechanisms that underlies the longitudinal and transverse relaxation phenomena, molecular tumbling, interactions between the magnetic moments of spins and presence of motion at different frequencies are complicated processes and an extensive topic that will not be addressed much further here, for more background on this matter, see[130,137]. Instead this section will introduce some pragmatic tools for working with relaxation and that I've found very meaningful during my PhD-years and that is phenomenological *Bloch equations*. Originally introduced by Bloch in 1946[138] these classical (classical in the sense, non-quantum mechanical) differential equations allows for a very complete description of longitudinal and transverse relaxation process and impact of radio pulses on bulk magnetization. The rotating frame Bloch equations can be written as in equation 15[139].

$$\frac{d}{dt}\begin{bmatrix} M_x \\ M_y \\ M_z \\ 1 \end{bmatrix} = \begin{bmatrix} -R_2 & \Omega & 0 & 0 \\ -\Omega & -R_2 & \omega_1 & 0 \\ 0 & -\omega_1 & -R_1 & R_1 M_z^0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} M_x \\ M_y \\ M_z \\ 1 \end{bmatrix} \qquad (15)$$

This is a set of phenomenological, linear, coupled, ordinary differential equations. Phenomenological means that they describe the overall behavior of the bulk magnetization and not describe the underlying mechanisms in detail. Bloch equations can greatly help in explaining and understanding experimental results for everything from inversion recovery experiments to pulse rotations etcetera. Solving these systems of differential equations require some eigenvalue analysis, this becomes rapidly very complicated when the systems become larger and numerical methods is the only option. This is a topic of great scientific interest also within the NMR community. Murase[139–141] provide a good introduction to this topic, I've also found these internet resources highly meaningful as well[142143]. Personally, I also think that Scipy implementations of the matrix exponential is a convenient way to work with these equations[144145].

During chemical or conformational exchange an atomic nucleus will switch back and forth between different environments, due to the different chemical shifts/frequencies in these two situations relaxation rates will increase and the appearance of NMR-spectra will change[146]. The rate of an exchange process is given by the exchange rate constant:

$$k_{ex} = k_{AB} + k_{BA} \qquad (16)$$

Which is the sum of the forward, k_{AB} (Hz) and backward, k_{BA} (Hz) rate constants. The forward and backward rate constants are related to populations of the nuclei in the two different states as:

$$k_{AB} = p_B \times k_{ex} \qquad (17)$$

$$k_{BA} = p_A \times k_{ex} \qquad (18)$$

In which p_A is the proportion of nuclei in the major state and p_B the proportion in the minor state. The inverses of the rate constants will give the average lifetime of the respective states. The exchange process is a chemical equilibrium process and the relationship between the free energy and equilibrium constant is then given by[147]:

$$\Delta G^\circ = -RT \ln K_{eq} \qquad (19)$$

In which R is the gas constant (8.314 J/K/mol) and T the absolute temperature in Kelvin and K_{eq} the equilibrium constant. Using this formula and keeping in mind that $K_{eq} = p_b/p_A$ and that $p_A = 1-p_b$ one can have a look at p_B as a function of difference in free energy, as in (Fig. 1.16) below.
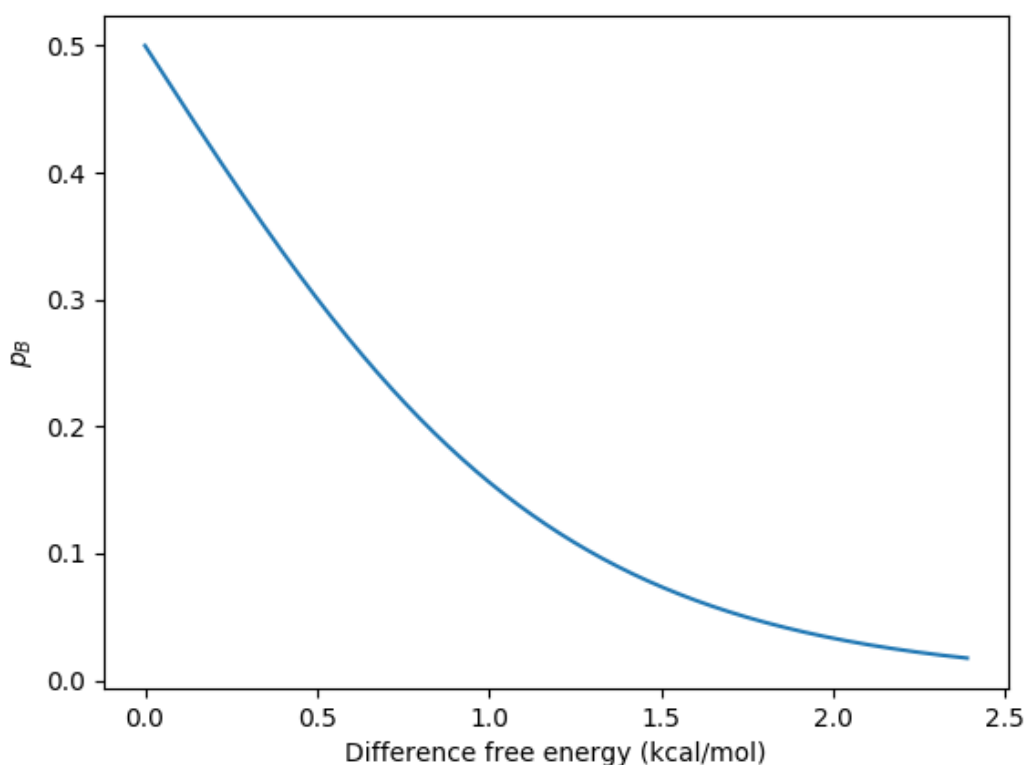
Figure 1.16. Population B, the alternative state as a function of difference in free energy.

From (Fig. 1.16) above one can see that for equilibria where the populations are getting smaller the difference in free energy becomes much higher, such low populated states, higher in free energy are commonly referred to as *excited states* in the NMR context[148,149].

The next issue to tackle is the impact of exchange between these low populated excited states and how it affects the NMR-spectrum. To address this, I borrow the following equation from[146].

$$\frac{d}{dt}\begin{bmatrix} M_A \\ M_B \end{bmatrix} = \begin{bmatrix} -R_{2A} + i\omega_A - k_{AB} & k_{BA} \\ k_{AB} & -R_{2B} + i\omega_B - k_{BA} \end{bmatrix}\begin{bmatrix} M_A \\ M_B \end{bmatrix} \tag{20}$$

And solve it with the SciPy matrix exponential[144,145] and subject the results to a Fourier transform using the NumPy[150] library to create an overview of the impact of exchange on the NMR-spectrum, the final results of doing this is shown in (Fig. 1.17). In equation 20, k_{AB} and k_{BA} are as before the forward and backward exchange rate constants and R_2 (Hz) is the transverse relaxation rate constants and $\omega_{A/B}$ is the offsets of the A and B spin in rad/s
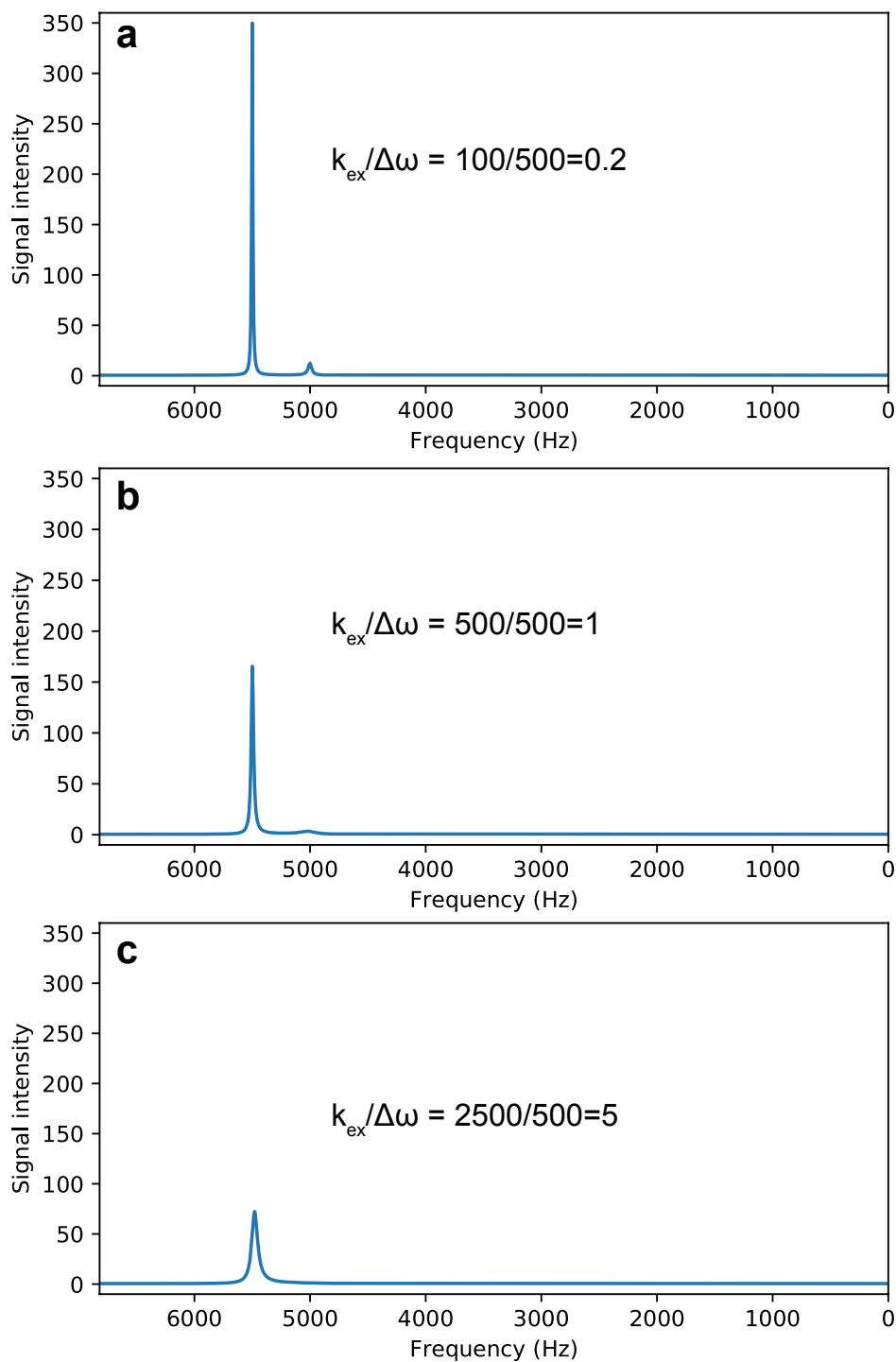
Figure 1.17. Effect of exchange on a 1D NMR-spectrum where the alternative conformer is present at abundance of 10%. As the exchange rate increase relative to the difference in frequency between the conformers, the peaks will become broadened. In panel c where exchange is fast relative to the frequency difference the small peak is not visible anymore, in such situations one can anyway study the invisible peak with sensitive NMR-experiments such as the $R_{1\rho}$-experiment described in section 1.3.9.

Equation 20 used to simulate the NMR-spectra in (Fig. 1.17) was sufficient for that purpose. A more extensive description of the impact of exchange on NMR-relaxation, both longitudinal and transverse and in the presence of radio frequency irradiation was given by McConnell in

1958 who extended the original Bloch equations with terms for exchange[151]. The rotating frame *Bloch-McConnell* equation can be written as:

$$
\frac{d}{dt}
\begin{bmatrix}
M_{AX} \\
M_{AY} \\
M_{AZ} \\
M_{BX} \\
M_{BY} \\
M_{BZ} \\
1
\end{bmatrix}
= -
\begin{bmatrix}
R_{2A}+k_{AB} & \Omega_A & 0 & k_{BA} & 0 & 0 & 0 \\
-\Omega_A & R_{2A}+k_{AB} & \omega_1 & 0 & k_{BA} & 0 & 0 \\
0 & -\omega_1 & R_{1A}+k_{AB} & 0 & 0 & k_{BA} & R_{1A}M_{0A} \\
k_{AB} & 0 & 0 & R_{2B}+k_{BA} & \Omega_B & 0 & 0 \\
0 & k_{AB} & 0 & -\Omega_B & R_{2B}+k_{BA} & \omega_1 & 0 \\
0 & 0 & k_{AB} & 0 & -\omega_1 & R_{1B}+k_{BA} & R_{1B}M_{0B} \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
M_{AX} \\
M_{AY} \\
M_{AZ} \\
M_{BX} \\
M_{BY} \\
M_{BZ} \\
1
\end{bmatrix}
$$

(21)

As can be seen, solving this system of equations using numerical integrators or implementations of the exponential matrix found in software packages such as Python or MATLAB, will provide the evolution of magnetization along X, Y and Z-axis for the A and B state as a function of relaxation and exchange rate constants. This will be important in the next section 1.3.9 in the discussion of the $R_{1\rho}$-experiment.

Since methods for solving these equations relying on numerical integrators or various implementations of the exponential matrix isn't always a nice thing, a lot of literature providing equations that partly describes or approximate the solution have been published. As aforementioned, Murase[139–141] provides a nice mathematical overview of the subject and for instance the Palmer lab have then published a number of studies providing solutions and expressions suitable for analyzing experimental data (especially $R_{1\rho}$-data) without having to rely on the computer heavy numerical solutions[152,153]. The expressions provided[153] have been used extensively for the analysis of $R_{1\rho}$-data in this PhD-project and within some contexts (e.g. here) it is referred to as the *Laguerre equation* it has the following appearance for a 2-state exchange process:

$$
R_{1\rho} = R_1 cos^2\theta + R_2 sin^2\theta + \frac{sin^2\theta\, p_A p_B \delta^2 k_{ex}}{\frac{\omega_A^2 \omega_B^2}{\omega_E^2} + k_{ex}^2 - sin^2\theta\, p_A p_B \delta^2 \left(1 + \frac{2k_{ex}^2(p_A\omega_A^2 + p_B\omega_B^2)}{\omega_A^2\omega_B^2 + \omega_E^2 k_{ex}^2}\right)}
$$

(22)

In this expression R_1, R_2 are the relaxation rate constants introduced in the beginning of this section. θ is the theta angle introduced in (eq. 10, section 1.3.2), p_a, p_b are the respective populations of the A and B state, k_{ex} is the exchange rate constant $\omega_{A/B}$ is the effective field (eq. 8) for the respective state δ is the difference in chemical shift between the A and B state and ω_E is the effective field calculated as for the observed signal. Using this equation, one can extract a multitude of information from $R_{1\rho}$-data (the expression and usage of it is explained further in the next section, 1.3.9). Another mathematical expression I would like to introduce is:

$$R_{1\rho} = R_1 cos^2\theta + sin^2\theta \left(R_2 + \frac{\phi_{ex}k_{ex}}{k_{ex}^2 + \omega_{eff}^2} \right) \qquad (23)$$

Which is an expression originating from[154] here (23) written as in[155]. The meaning of R_1, R_2, θ and k_{ex} are the same as before, Φ_{ex} contains terms for populations and chemical shifts. This expression can be used describe exchange process that are fast but the primary reason for introducing it here is that during this PhD-project it has also been used to analyze $R_{1\rho}$ -data, only consisting of so-called on resonance data (section 1.3.9 & 4.4 will discuss this expression further).

1.3.9 The R$_{1\rho}$-experiment

The R$_{1\rho}$-experiment is a type of *relaxation dispersion* experiment, a category of NMR-experiment that measures dispersion (i.e. spread) of relaxation. The dispersion can be measured with these experiments and it can be controlled by experimental parameters such as radio pulse strength and frequencies, it will also depend on and be influenced by dynamical processes in molecules. There are a number of different types of relaxation dispersion experiments suitable for characterizing dynamics within biomolecules, such as nucleic acids and proteins[156,157]. In its simplest form, an $R_{1\rho}$-experiment consists of a few pulses and a so-called *spin-lock*, which is basically also a long continuous pulse. A schematic representation of a $R_{1\rho}$-pulse sequence is shown in (Fig. 1.18) below.
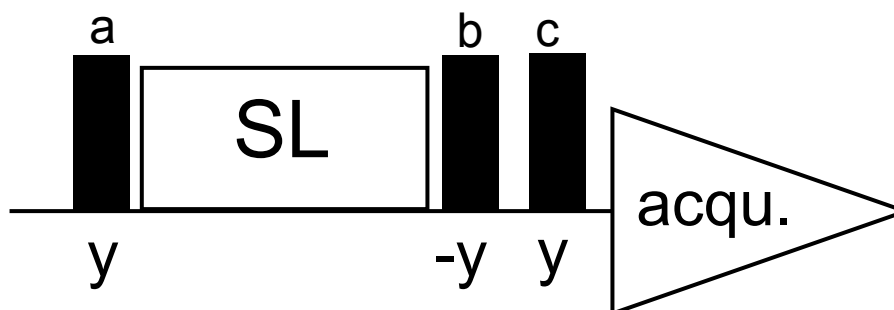

Figure 1.18. A schematic representation of a very simple R$_{1\rho}$ pulse sequence. A pulse (a) of phase y will rotate the magnetization vector towards the x-axis (see section 1.3.2) immediately after, a spin-lock is applied that is a long pulse along the x-axis. Since the spin-lock is applied along the same axis as the magnetization vector the vector will not rotate in the x,y-plane in the rotating frame of reference it will appear static i.e. spin-locked (assuming the frequency difference of the spin-lock and spin is zero, Ω = 0). During the spin-lock the magnetization will relax under the influence of exchange and the radio field from the spin-lock. After the spin-lock, a pulse (b) of -y phase rotates the magnetization back to the z-axis. Ultimately another pulse (c) rotates the magnetization vector down again to the x,y-plane in this schematic example. The magnetization is then detected and a FID is recorded during acquisition. Several experiments with different duration and strength of the spin-lock is then recorded the decay of the signal in the final spectrum is observed and measured. The intensity of the signal in the spectrum will decay monoexponentially the decay constant is the R$_{1\rho}$ value.

The basics of the $R_{1\rho}$-experiment is outlined in Figure 1.18 but the figure does not really explain what is going on during the spin-lock. The best way to approach that is go back to section 1.3.1 and 1.3.2 and once again look at equation 4 and 8. Equation 8 describes the effective field that a spin will experience in the presence of a radio pulse, equation 4 states that precession frequency is proportional to an experienced magnetic field. Equation 8 applies perfectly well for the spin-locked magnetization in the example in (Fig. 1.18), equation 4 applies as well and the magnetization vector will precess around the effective field from the radio pulse, in the rotating frame. If one now considers the situation where the magnetization is spin-locked along the x-axis in (Fig. 1.18) and the frequency of the applied spin-lock/pulse is the same as the spin i.e. $\Omega = 0$, then the magnitude of the effective field from the spin-lock that the spin will experience will only be ω_1, the other term in equation 8. If this spin then exchanges into another chemical environment with another chemical shift the Ω term will no longer be zero, the spin in that environment will not be *on-resonant* with the spin-lock and it will experience an effective field different from the initial state. Equation 4 and 8 is thus the key to a semi-quantitative understanding of the $R_{1\rho}$-experiment. When a spin exchange back and forth between two different environments the Ω term will change, it will experience different effective fields and precess at different frequencies in these different environments. This different precession frequency in the different environments will *dephase* the magnetization and increase relaxation. Due to the square dependence of the terms in equation 8 one can also see that by increasing the strength of applied spin-lock one can compensate for a difference in Ω between the environments by increasing the power enough the difference in Ω becomes negligible and the additional relaxation vanishes. This is the essence of the $R_{1\rho}$-experiment.

Although the previous paragraph gives a more elaborate explanation of what happens during the spin-lock it is still far from a complete description of this complicated topic. To proceed the $R_{1\rho}$ value need to be explained in more strict detail. The $R_{1\rho}$ value is yet another *relaxation rate constant* it is the relaxation rate constant of magnetization projected onto the effective field of the spin-lock, within the rotating frame of reference and will have units of (s^{-1})[155,158]. The word *projected* here is important and it means that the relaxation of all three of the individual X,Y,Z magnetization components will contribute to the $R_{1\rho}$ value, an attempt to illustrate this is done in (Fig 1.19). The decay of the individual X,Y,Z components will in turn depend upon R_1, R_2 and exchange under influence of the spin-lock, thus $R_{1\rho}$ is a composite value and can be written as:

$$R_{1\rho} = R_1 cos^2\theta + R_2 sin^2\theta + R_{ex} sin^2\theta \quad (24)$$

In which R_1, R_2 are the longitudinal and transverse relaxation rate constants as introduced in section 1.3.8 and R_{ex} is the exchange contribution (Hz) to the $R_{1\rho}$ -value. The R_{ex} term can be one out of several mathematical expressions one example is the last part of the *Laguerre equation* (equation 22) introduced in the previous section. At this point one can also clearly see the connection to the Bloch-McConnell equations in the previous section, as the equations will

provide the evolution of the individual X,Y,Z components. So, by solving the system of equations and calculate projected magnetization as a function of R_1, R_2, exchange rate constants, populations, chemical shift differences, radio frequency field strength and duration one can fit experimental data and obtain values of these parameters, which then can be used to make structural/chemical/biological conclusions of the system under study.
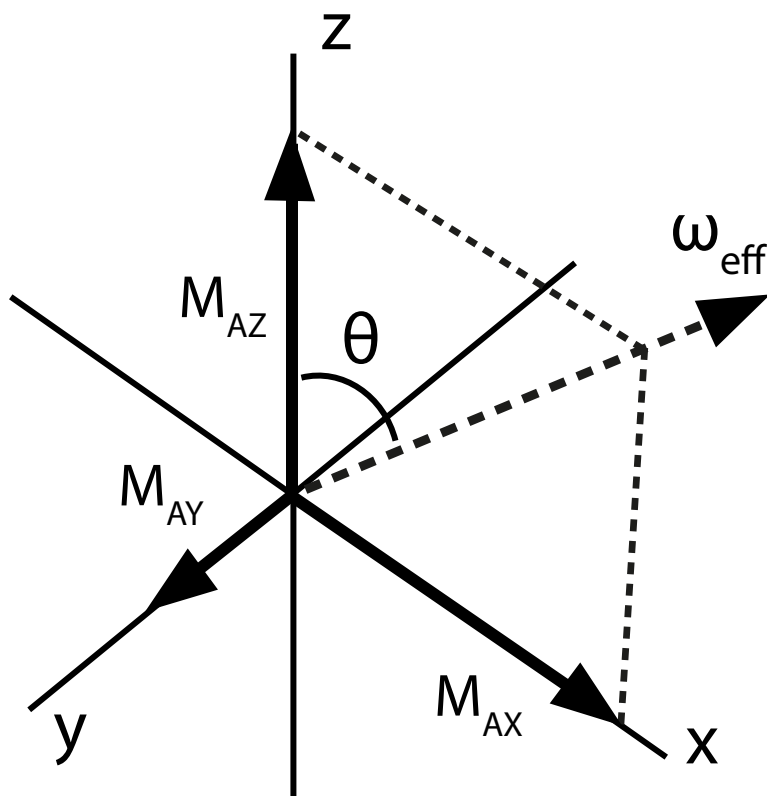

Figure 1.19. Projection of magnetization components onto effective field within the rotating frame of reference. The effective field of a spin-lock that is NOT on-resonant is shown as a dashed arrow (ω_{eff}) stretching away from the origin. The effective field is characterized by a theta angle depending on how off-resonant it is and the strength of the radio field according to equation 10 in section 1.3.2. X,Y,Z magnetization components are projected onto the effective field (dashed lines). The different magnetization components will relax depending on R_1, R_2 and exchange and the decay rate of the projection or "shadow" on ω_{eff} will ultimately be $R_{1\rho}$.

To get a complete description of an exchange process it is common that one does a large number of short $R_{1\rho}$-experiments where one varies both the offset (Ω) and the strength (ω_1) of the spin-lock. The resulting data from all experiments are then fitted to some model such as the Bloch-McConnell equation or Laguerre equation introduced previously. The fitting of the data will provide the interesting parameters, an example of a data set simulated with the Bloch-McConnell equation, then fitted with the Laguerre equation is shown in (Fig. 1.20).
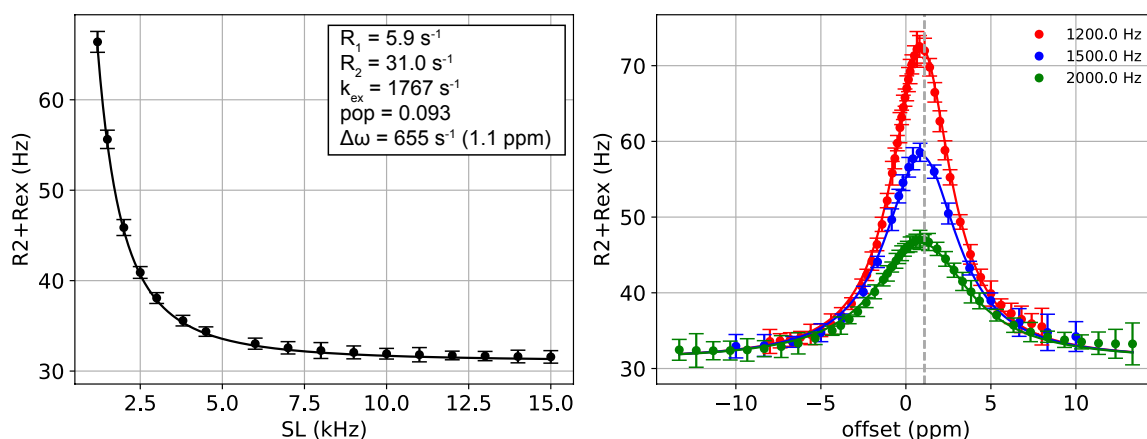
Figure 1.20. A data set simulated with the Bloch-McConnell equations using the following parameters as input: $R_1 = 2.5$ s^{-1}, $R_2 = 25.0$ s^{-1}, $k_{ex} = 2500$ s^{-1}, pop 0.1, $\Delta\omega = 500$ s^{-1} (notice that $k_{ex} = 2500$ s^{-1}, pop 0.1, $\Delta\omega = 500$ s^{-1} are the same parameters as used in Figure 1.17c in which the peak of the B state became invisible). The data set has been fitted with the Laguerre equation the fitted parameters are shown in the small box in the left panel. The left panel shows on-resonance data in which the frequency of the spin-lock is the same as the observable peak (which is close to the major/A/ground state) in that case the Ω term is zero and then the strength/magnitude of the spin-lock is varied, hence the values in kHz on the x-axis represent magnetic field strength of the radio pulse and could equally well be written in some other unit describing magnetic field strength such as gauss or tesla but convention says frequency units. The units on the y-axis says $R_2 + R_{ex}$, which is equal to $R_{1\rho}$ in the on-resonance case, this can be understood from eq. 10 & 21 since θ approaches $\pi/2$ as the spin-lock in on-resonance. The left panel thus show a decrease in relaxation due to suppression of exchange contribution in response to higher spin-lock strength. The right panel shows off-resonance data in which the spin-lock strength is kept constant within a series of experiments but instead the offset (Ω) of the spin-lock is varied instead. During the frequency variation, when the spin-lock instead is on-resonant with minor/B/excited state the measured $R_{1\rho}$-value at that point will increase this effect is even more visible in the "off-resonance plot" in the right panel as the R_1-contribution is subtracted from the $R_{1\rho}$-value. The fitted value where this seem to occur in the right plot is indicated by a gray, dashed line (+655 s^{-1} = 1.1 ppm in this example, which is quite close to the simulated value, +500 s^{-1}). Notice in the right plot that the unit here on the x-axis does represent a frequency difference, y-axis magnitude of relaxation and color the spin-lock strength.

The simulated $R_{1\rho}$-data set shown in (Fig. 1.20) was intended to exemplify a realistic proton $R_{1\rho}$-dataset. By performing both the *on-resonance* and *off-resonance* experiments as in (Fig. 1.20) and fit the data one can get information about presence of the invisible peaks in (Fig. 1.17c).

The preceding paragraphs of this section gives an introduction to some basics of the $R_{1\rho}$-experiment, how it works, can be used and what information that can be obtained from it. (Fig. 1.18) gave a schematic introduction to the $R_{1\rho}$-pulse sequence, I would like to conclude this section though by elaborating a little on more realistic and complicated $R_{1\rho}$-pulse sequences. To start with a pulse sequence originating from[159,160] and extensively used in this PhD-project and by our lab to measure $R_{1\rho}$ on heteronuclei such as ^{13}C and ^{15}N, it is shown in (Fig. 1.21). The pulse sequence utilizes so-called *Heteronuclear Cross-Polarization (HCP)* transfer steps[161] to transfer magnetization from the ^1H to the ^{13}C that is then spin-locked like in (Fig. 1.18).
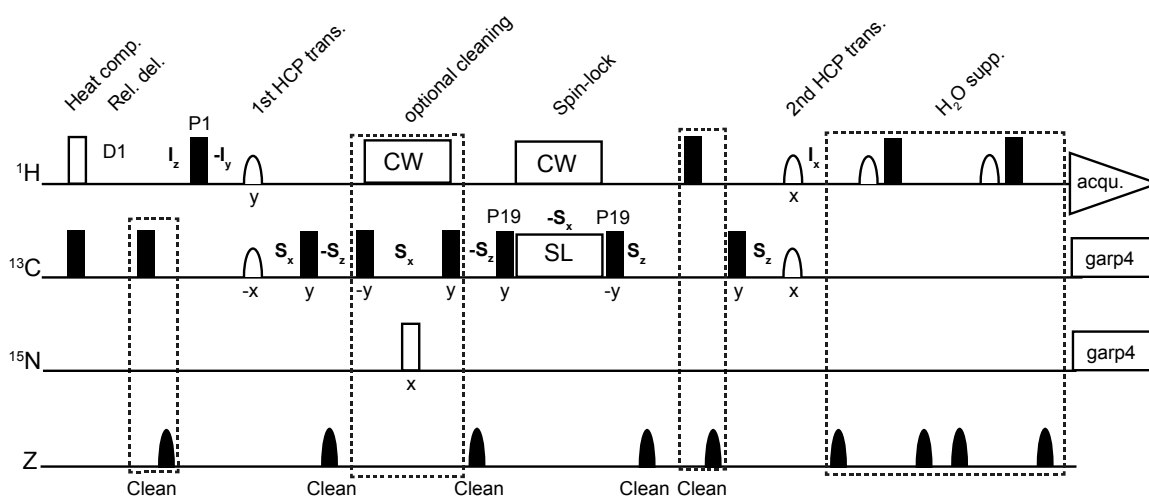
Figure 1.21. The $^{13}C/^{15}N$ $R_{1\rho}$ pulse sequence typically used in this PhD-project, as can be seen it is much more complicated than the schematic pulse scheme in Figure 1.18, yet in principle it is very similar. Indicated in the pulse sequence are the different operators present at different steps in the sequence, this can be understood by following the phases of pulses and schemes for product operator that can be found in articles such as the one from Nolis[161] used in this case. Once magnetization has been transferred to the carbon/nitrogen nucleus it is spin-locked and one can see in this pulse sequence the pulses called P19, which correspond to the pulse a & b in Figure 1.18. In this pulse-sequence the J-coupled proton used for transfer is irradiated during the heteronuclear spin-lock to avoid artifacts. Notice in the beginning of the pulse sequence there is a "heat compensation" step, here the sample is irradiated with a pulse of a frequency far away from the studied spectral region but that will still mean a power/ energy input into the sample that will heat it, this compensates for when spin-locks of different durations are used. The bottom line shows gradient pulses applied along the Z-axis of the sample, that destroys any magnetization present in the X,Y-plane. Such pulses are used for "cleaning" i.e. to get rid of potential unwanted magnetizations present at different steps in the pulse sequence.

Fig. 1.21 exemplifies a pulse sequence for heteronuclear $R_{1\rho}$-measurements another type of $R_{1\rho}$-pulse sequence that have been central to this PhD-project is the proton $R_{1\rho}$-sequences containing a "homonuclear INEPT" or as we call it, a *SELOPE* transfer step[128]. The structure of such a pulse sequence is shown in (Fig. 1.22). The example in (Fig. 1.22) shows the SELOPE ^1H-$R_{1\rho}$ pulse sequence in a version that could be used to measure $R_{1\rho}$ in a 1D manner on

aromatic purine protons that are spectrally overlapping with pyrimidine H6 protons by transferring the signal from H6 to the upfield shifted H5s.
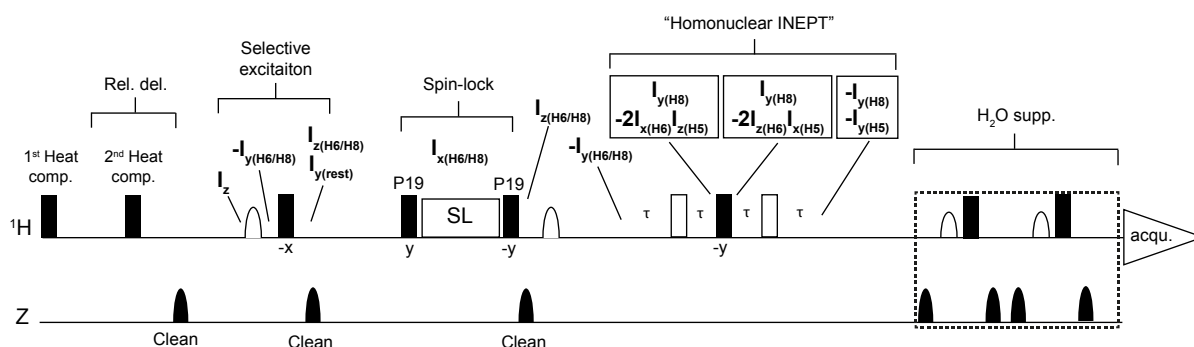

Figure 1.22. 1D ^{1}H-R$_{1\rho}$ SELOPE pulse-sequence for the measurement of aromatic purine protons. The picture shows which operators that are present at different time points. The aromatic proton region of RNA molecules contacting H8:s of purines, H2:s of adenines and H6:s of pyrimidine (see section 1.3.10) will be selectively excited. I$_x$ magnetization of the H6, H8 and possibly H2:s and amino protons will all enter the spin-lock part, if acquisition would be done immediately after the spin-lock as in the simplified schematic example in Fig. 1.18 the resulting spectrum would be over-crowded and impossible to use. Thanks to the "homonuclear INEPT"/SELOPE transfer step (see section 1.3.5) at the end, signal from H6:s will be transferred to H5:s much more upfield in the spectrum, hence it will be possible to observe pure signals from H8:s and H2:s and calculate R$_{1\rho}$ values.

1.3.10 NMR-studies of RNA

In order to characterize the dynamics of RNA molecules in a meaningful manner it is necessary to assign the individual NMR-signals in the spectra, i.e. to figure which signal that belongs to which atom in the molecule. This can be a quite challenging task and sometimes it is not possible to assign all signals in a spectrum for a larger molecule. To succeed with this, it is necessary to use nuclei from different parts of the RNA-molecule that have different chemical shift characteristics, an overview of some useful nuclei in RNA nucleobases and ribose parts that have been used for assignments in this PhD-project is shown in (Fig. 1.23). The chemical shifts shown in (Fig. 1.23) have been taken from the BMRB database[120] base pairing information has been taken form the FRAbase[122]. There are many pitfalls to this approach, first one needs to consider the fact that not all chemical shifts submitted in the BMRB database are correctly calibrated[162], leading to large spreads in especially chemical shifts of heteronuclei[163], secondly using FRAbase for the elucidation of base pairing partners only make use of secondary structure of the RNA hence structural details and their impact on the chemical shift are not taken into account. Nevertheless, the chemical shift presented in (Fig. 1.23) gives a good overview of chemical shifts for some popular nuclei in RNA. Detailed chemical shift information can of course be found in literature as well[164,165]. There are many strategies and NMR-experiments that can be used to assign the signals in NMR-spectra of RNA. In this PhD-project a central and very important experiment for doing this is the NOESY experiment introduced in section 1.3.6. As explained in section 1.3.6 the NOE will depend on the proximity of nuclei in space, so nuclei that are closer to each other will give stronger signals in the NOESY spectrum, some nuclei that are close to each other within an RNA helix are shown in (Fig. 1.24).
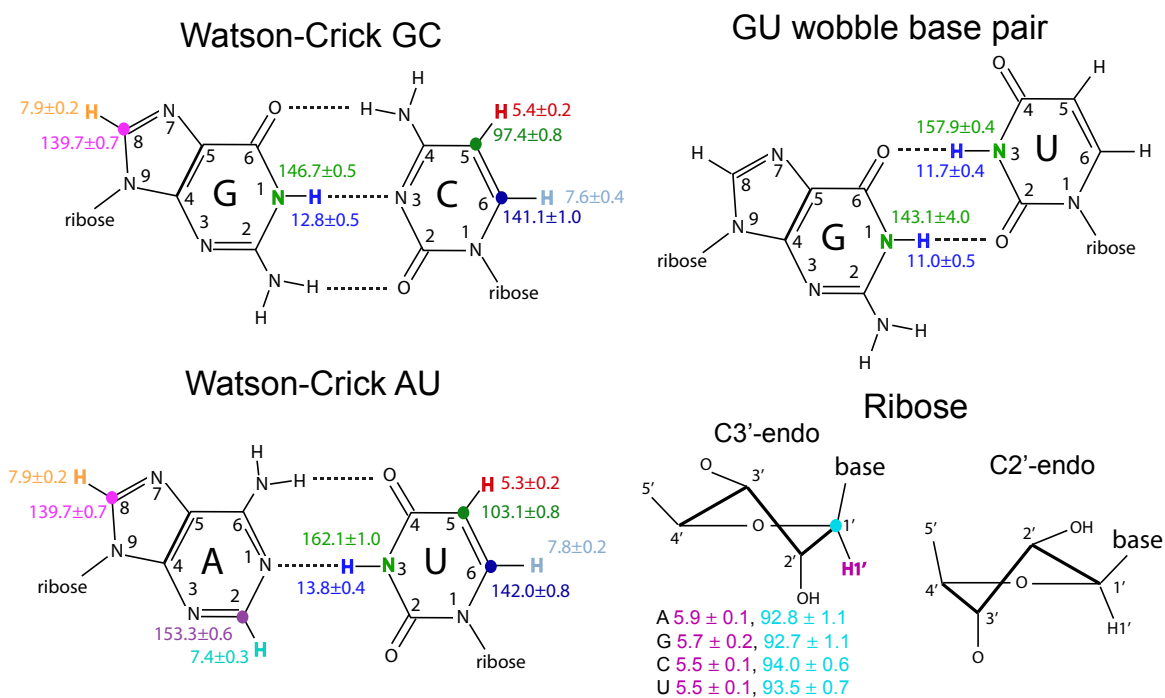
Figure 1.23. Example of chemical shifts for some common nuclei in RNA, chemical shifts extracted from the BMRB database and FRAbase used for base pairing information. For the ribose part chemical shifts are presented for the H1' and C1' nuclei, these have been taken from Watson-Crick base pairs in helical elements of the RNA structures and in helical parts of RNA the ribose is assumed to have a C3´-endo conformation, hence the chemical shifts presented are put under C3´-endo figure. In non-canonical base pairs, bulges and loops it is common that the ribose ring adopts a C2´-endo conformation in such conformations typically the C1´ chemical shift is upfield shifted (lower ppm value).
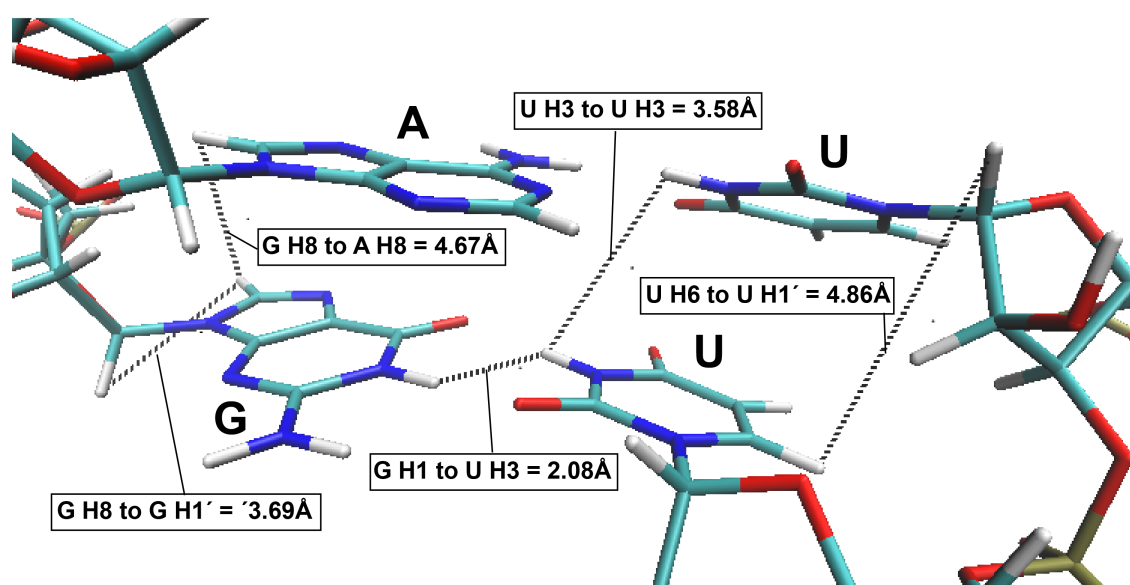
Figure 1.24. Examples of some distances within an RNA helix, that typically give rise to signals in the NOESY type of experiments, units are in Ångström (Å). For numbering of atoms see figure 1.23. To the left can be seen the distance of a guanine H8 to the H1´ of its own ribose (3.69 Å) and to the H8 of a neighboring adenine base. In the middle is shown the distance (2.08 Å) between the guanine H1 and uridine H3 (imino protons) in a GU wobble

base pair as well as the distance (3.58 Å) between the uridine H3 and the H3 of the neighboring uridine nucleotide. To the right is shown the distance (4.86 Å) from a uridine H6 to the H1' of the neighboring uridine. All the indicated distances usually gives signals in the NOESY experiment of RNA molecules and they are very useful for performing "sequential assignments" when one elucidates the order of nucleotides in the molecule thanks to the possibility of observing NOE signals between atoms in different nucleotides. Figure based on PDB id 1s23 and created using VMD[69].

NOEs from the distances shown in (Fig. 1.24) will give rise to *cross-peaks* in the RNA NOESY spectrum, a schematic NOESY spectrum is shown in (Fig. 1.25). To succeed with the assignments using the NOESY spectrum it is common that one uses a set of heteronuclear correlation experiments such as HSQC spectra's between ^{1}H,^{15}N and ^{1}H,^{13}C. As could be seen in (Fig. 1.23) the heteronuclei at different positions in different base pairs have different chemical shifts, this can be used to determine the base identities of two nucleobases that are close to each other and give rise to NOESY cross peaks. Using the NOESY experiment and a set of correlation experiments it is usually possible to figure out secondary structure of the RNA molecules and base identities for different signals.

Once an RNA molecule has been assigned one can perform a more thorough investigation of structure and dynamics and there is a lot of scientific literature describing this for a wide variety of RNA molecules. The famous and unusually stable UUCG tetra-loop found in many RNA systems has been thoroughly structurally investigated[165,166]. The dynamics of this system has also been investigated on faster time scales by NMR[167] and by combining NMR and computational methods[168,169].

The stability of the UUCG-loop is also affected by closing base pair under it[170]. These already existing studies of the UUCG-loop are highly relevant for the work presented in Paper III of this thesis.
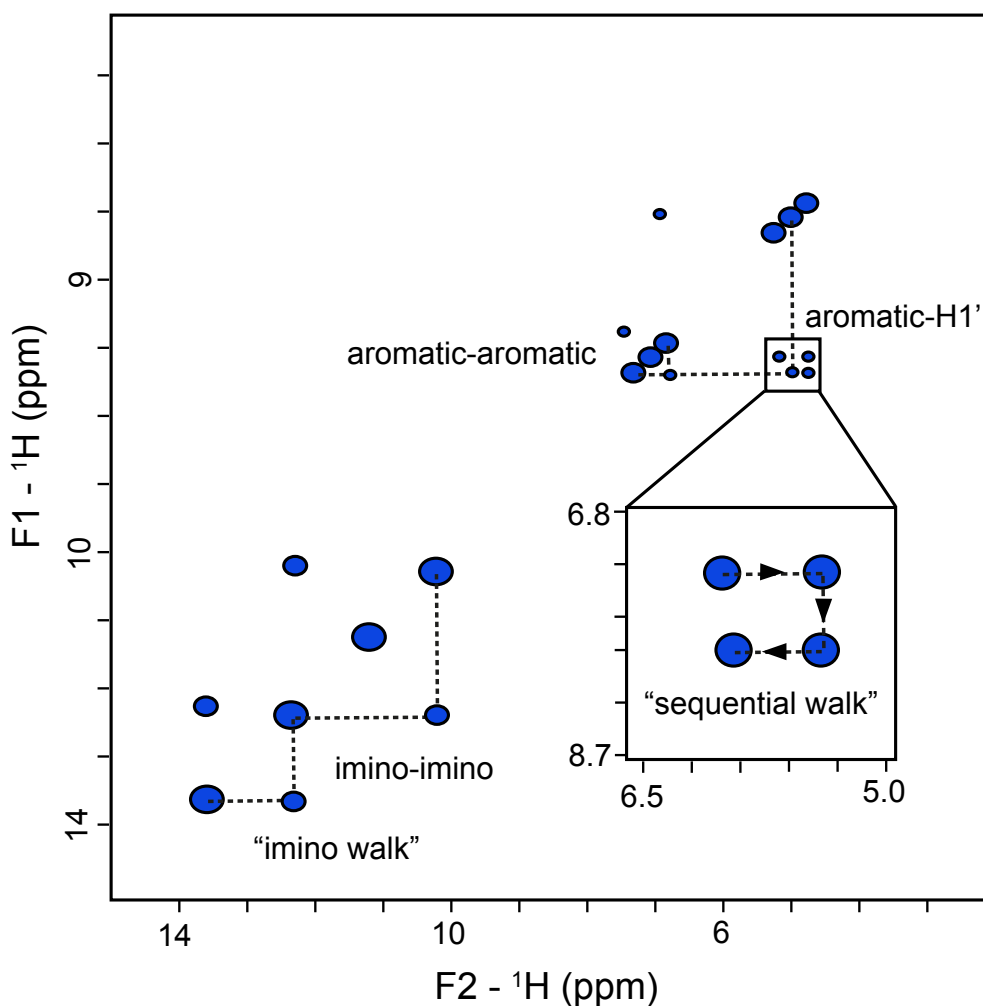

Figure 1.26. A schematic NOESY spectrum showing some crucial regions that are useful for the assignment of RNA. To the left in the region around ~9-15 ppm the signals from the imino protons will be. In helical segments of RNA with strong base pairing where the imino protons are protected from solvent exchange, off diagonal or "cross-peaks" from imino protons in base pairs neighboring each other will be visible. In such cases it is possible to do a "imino walk" and see a stretch of base pairs stacked on top of each other in the helix, in that way the secondary structure of the helix can be determined. Another region that is useful is aromatic region around 7.0-8.5 ppm where it can be possible to see cross peaks between aromatic nuclei H8/H6 immediately between the bases, that can also be used for sequential assignments. In the region ~5.0-6.5 in F2, the direct dimension and the region ~7.0-8.5 ppm in F1, the indirect dimension it is possible to find cross-peaks between H1' and H6/H8 this can be used for a sequential walk when aromatic protons generate signal both with its own H1' and neighboring H1's.

Concerning interesting parts of the ribosome and ribosomal RNA (rRNA), the structure of a part of the 23S rRNA of the large subunit that is the target site for two dangerous toxins, the α-sarcin and ricin toxin has been structurally investigated by NMR, showing its structure in detail[8]. Another important part the ribosome that has been heavily investigated is the ribosomal A-site that is the binding site for the aminoglycoside category of antibiotics[100]. This part of the 16S rRNA of the small ribosomal subunit was originally structurally investigated by NMR in the context of a small 27-mer RNA construct in complex with aminoglycoside antibiotics[6]. This was at a time point where no other high resolution structural data of the ribosome was

available. Later on, after publication of high resolution crystal structures the accuracy of the original NMR solution structures was instigated again[171]. The dynamics of the ribosomal A-site has then been more thoroughly investigated by NMR after that, showing that the unpaired adenines (A1492 & A1493) (can be seen in secondary structure map in Fig. 1.6) becomes more helical while U1495 gets expelled from the helical structure during an exchange process[1727]. The ribosomal A-site is crucial for the selection process of the correct tRNA molecule during the translation process[173]. Regarding the question of accuracy during translation another very important question is the existence of tautomeric versions of nucleobases[84]. Possible tautomeric processes in RNA and various nucleobases have been extensively investigated by NMR[174175–177]. Tautomeric processes in nucleobases can change the geometry of base pairs and in GU wobble base pairs tautomerism can cause the geometry to change to that of a Watson Crick GC base pair leading to errors during decoding of the nucleic acid sequnece[85,86,178]. Some reflections on tautomerism processes in nucleic acids will be discussed in later chapters of this thesis. Other highly interesting studies of chemical exchange processes in nucleic acids are for instance the studies of solvent exchange to investigate melting phenomena in *RNA thermometers*[179]. As well as older studies of solvent exchange in tRNA molecules[180].

A lot of the NMR investigations done on RNA systems are done using solution state NMR and small, miniaturized RNA systems suitable for that experimental technique. That is also the way RNA systems have been studied in this PhD-project. There are however some interesting studies where attempts have been made to study larger systems with solution state NMR and that is especially interesting in the context of the ribosome. Although the studies performed focused on ribosomal proteins there has been studies made where complete 70S ribosome where put in solution state NMR-spectrometer, in such studies signal from only some of the most flexible ribosomal proteins are possible to observe[181]. In order to study less flexible parts of the ribosome using NMR-spectroscopy it is necessary to use solid state NMR techniques and spin the sample at high speeds but nevertheless there have been some studies done on the ribosome using that strategy[182,183].

2 AIMS

The underlying work for this PhD-thesis has aimed to both gain an insight into the structure and dynamics of certain regions of rRNA connecting the large and small ribosomal subunits together as well as to develop novel biochemical methods and NMR-experiments for studies of RNA.

2.1 Paper I

The aim of Paper I was thus to develop a more modern sample production method relying on improvements done to the *in vitro* transcription method and usage of High Performance Liquid Chromatography (HPLC) for purification of the RNA. An important aspect of the aim for this work was that the final method should be robust and flexible and suitable for people with limited experience working with liquid chromatography.

2.2 Paper II

The aim of Paper II was to develop an NMR-experiment that utilize the possibility to transfer signal between J-coupled proton spin systems in RNA molecules and to "decrowd" certain spectral regions. Fulfillment of this aim and usage of such method means that one can perform NMR-measurements on proton signals in crowded regions in spectra of non-isotopically labeled RNA molecules, thus alleviating the need for expensive NMR-active isotopes.

2.3 Paper III

In accordance with the aim to characterize certain regions of rRNA on the interface area between the two ribosomal subunits we have investigated several regions along helix 44 of the *E. coli* 16S rRNA, in the context of small (~10 kDa) RNA molecules using solution state NMR-spectroscopy. In Paper III emphasis and effort have been put on the investigation of the stem-loop part of helix 44. The aim of studying this region have been to characterize the dynamics of the whole stem-loop part that contains a naturally occurring UUCG-loop and several adenine nucleotides that in their native 70S ribosomal context, bulge out and form a so-called A-minor motif interaction with helix 8 of the 16S rRNA. This could aid in getting a better understanding of the behavior of A-minor motifs as a tertiary interaction in rRNA.

2.4 Studies of exchange and observation of cross-relaxation during imino ^{1}H-$R_{1\rho}$-experiments

The headline of this paragraph is not directly connected to only one of the original, individual aims of this PhD-project, instead one can say it is loosely connected to the aim of developing new NMR-experiments as well as to the study of dynamical rRNA regions. Yet, in the following sections, in addition to the topics related to point 2.1-2.3, I would also like to include and discuss some additional data collected on imino protons in RNA samples modeling the B6 intersubunit bridge in *E. coli* and the B3 bridge in the human cytosolic ribosome. Although the data has been collected on samples with the original intention to characterize conformational

exchange processes the results and discussions related to this point will concern chemical exchange processes and cross-relaxation artefacts during ^{1}H- $R_{1\rho}$ experiments.

3 MATERIAL & METHODS

The following sections in this chapter only concerns the limited work of studies of exchange processes among imino protons in certain RNA molecules modelling regions of rRNA.

3.1 Studies of exchange and observation of cross-relaxation during imino ^{1}H-$R_{1\rho}$-experiments

3.1.1 RNA hairpin constructs

For these experiments a non-isotopically labeled 29-mer RNA hairpin construct modelling the B6 region of *E. coli* 16S rRNA was *in vitro* transcribed with the methods described in paper I of this thesis. It was HPLC-purified using a test IE buffer system utilizing NaCl (1.2 M) for elution and 25 % formamide to aid denaturation. In a similar manner a 39-mer RNA construct modeling the B3 region of human cytosolic ribosome was produced and used.

3.1.2 Imino ^{1}H-$R_{1\rho}$-experiments

Imino proton $R_{1\rho}$-measurements were performed on non-isotopically labeled samples using a pulse sequence in which the imino proton of interest was selectively excited with an Eburp2.1000 shaped pulse, followed a by a hard pulse and a purge gradient. This was then followed by an on- or off-resonant spin-lock (hard pulses were used to rotate the magnetization to the theta-angle of the spin-lock). Ultimately a Watergate, water suppression block was used for water suppression prior to detection. The pulse sequence contained heat-compensation blocks at two different power levels, one at the power level of the current spin-lock used and one at a different compensatory power level to compensate for heating between experiments when spin-locks of widely different strengths were used. Experiments were performed in a 1D manner and spin-locks up to 150 ms duration and 15 kHz strength were used. Measurements were done both on- and off-resonance, spin-lock powers and offsets were varied in a systematic manner and commonly each signal decay was sampled with 6-8 data points.

3.1.3 Water inversion exchange experiment

Based on the water selective inversion recovery experiment previously used by Rinnenthal et al.[179], a simplified version of such an experiment was tested. This was done by modifying a zgesgp sequence, implemented with decoupling of nitrogen and carbon during acquisition. It was modified in such a way that the standard squa100.1000 selective 180°-pulses used in the excitation sculpting water suppression scheme in zgesgp was also used for the water selective inversion prior to the hard 90°-pulse. An exchange delay with bipolar gradients (5%) was also used[179]. The final appearance of this pulse (Fig. 3.1) sequence was thus, water selective 180°-pulse (squa100.1000), exchange delay with bipolar gradients (5%), water selective 180°-pulse again (squa100.1000) then followed by the normal zgesgp sequence (hard pulse and excitation sculpting and in this case decoupling). Only exchange delays below 300 ms was tested.
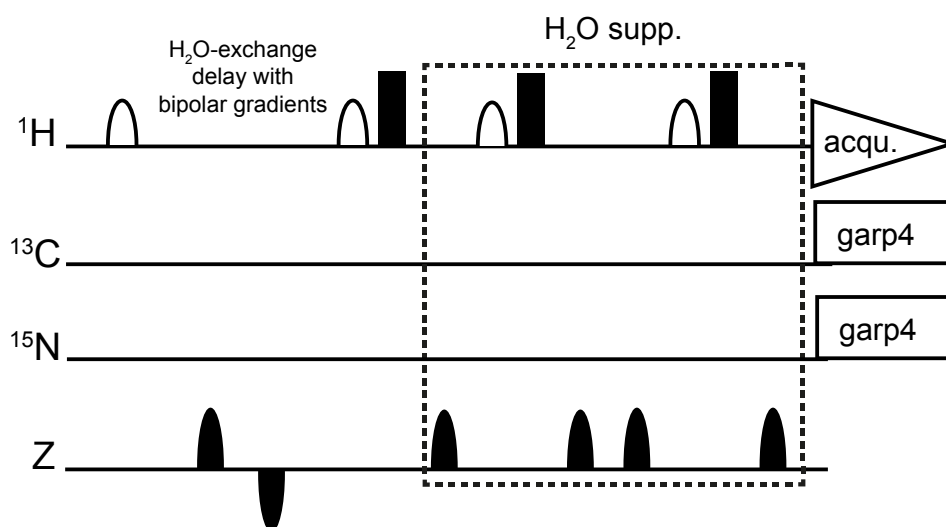
Figure 3.1. Schematic overview of the pulse sequence used for investigation of H₂O-exchange. In the beginning of the pulse sequence, water selective 180°-pulses (white, half ellipse shapes) are used for water inversion. In the variable delay between the selective pulses water exchange occurs, during this period bipolar gradients (opposite signs) was applied. After water selection and exchange delay, the pulse sequence is a normal ^{1}H 1D experiment with excitation using a hard 90°-pulse and an excitation sculpting water suppression block. By selectively inverting the water in the beginning, signals that are exchanging with water will show a 180° phase shift depending on the duration of the exchange delay and magnitude of the exchange.

3.1.4   Data processing

$R_{1\rho}$ data was processed in Topspin 3.5pl7 in combination with Python 2.7 scripts written in-house for macro generation and fitting of monoexponential decays. Standard deviations of the obtained $R_{1\rho}$-values were calculated using a Monte-Carlo error estimation approach using the signal to noise ratio from the individual 1D spectra as input. Imino proton $R_{1\rho}$ data sets were then further analyzed with equation 22 & 23 from section 1.3.8 in this thesis. In addition to this imino proton $R_{1\rho}$-data was also analyzed using a Bloch-style equation obtained by combining terms for exchange with terms for cross-relaxation[126,133]. When attempting to fit the data using this model the offset to the cross-relaxing spin was fixed whereas cross-relaxation rate constants were allowed to vary freely, during least squares optimizations. SciPy[144] implementations of the matrix exponential was used to solve the equation and SciPy library was also used for least squares fitting of the data.

4 RESULTS AND DISCUSSIONS

4.1 Paper I

In Paper I, we have developed an RNA-sample preparation method based on a sequence design step, *in vitro* transcription using T7 RNA-polymerase and addition of DMSO in the reaction mixture and C2´-methoxy groups on the DNA-template[18–20]. The sequence design step is based on utilizing MC-Fold[184] for the investigating and comparison of the impact of sequence on folding and stability properties of RNA molecules of different lengths. Once a suitable sequence has been decided it is transcribed and transcription is followed by a workup procedure in which possible inorganic pyrophosphate (iPP) is dissolved with an EDTA solution[185]. We tested several different concentrations of EDTA and compared dissolving performance relative inorganic pyrophosphatase enzyme, it was concluded that EDTA solutions with concentrations as low as 25 mM, efficiently dissolved the iPP-pellet after 5 minutes in room temperature. Usage of EDTA is a simple and affordable alternative to the usage of recombinant inorganic pyrophosphatase[186].

After dissolving any potential, residual iPP-pellet the material is filtered and then subjected to a two-step HPLC purification procedure, consisting of both a reverse phase ion pairing step (RP-IP) and an ion exchange step (IE). We tested and optimized a variety of gradients, typically consisting of a short period where material is allowed to bind to the column, followed by a main separation gradient using elution buffer percentages in the ~20-40% range (typical for both RP-IP and IE) followed by a period where the column is flushed with 100% of elution buffer to elute long products. The optimization and development of these gradients also provided estimates of what amounts that were suitable to inject on the two different columns with maintained, acceptable separation. Typically, RNA material up to 0.6 mg/injection was used in the RP-IP step, where resolution demands were lower and slightly less up to ~0.4 mg/injection for the more high-resolution separations on the IE column. Since *in vitro* transcription reactions on a 10 ml or larger scale, such as those used in paper I may produce several mg of the wanted transcription product, distribution of the RNA material between several injections on the two HPLC columns was a necessity.

The RP-IP HPLC step in Paper I utilizes a buffer system with tetrabutylammonium hydrogen sulphate as ion-pairing agent that will form a lipophilic ion-pairing complex with the RNA and bind to a C18 reverse phase HPLC-column, elution is then achieved using acetonitrile[57]. The RP-IP purification step is done at room temperature and serve as crude, initial purification of the material. During the development and testing of the method it was concluded that by pre-purifying the material with RP-IP one reduces to load and contamination of the IE column by avoiding the injection of a lot of short deletion products, possible microscopic, residual iPP particles and other contaminants on the IE column. The RP-IP gradients presented in Paper I has proved to work very consistently for a wide size range of RNA molecules.

After RP-IP purification of the material, the material is subjected to an IE purification step. The impact of secondary structures is important when it comes to IE purification of RNA[58], and in

our IE purification step the column temperatures (75°C) is increased to reduce the impact of secondary structures. However, for GC-rich RNA molecules, especially in HPLC buffer systems with high salt concentrations melting temperatures close to these values can occur, thus a large proportion of the molecules will not be in an unfolded state. To compensate for a possible reduction in separation due secondary structures etcetera, also in this step the material is distributed between several injections, most often more injections than in the RP-IP step and in that way, it is possible to obtain HPLC fractions were the RNA material is highly pure. These fractions are then pooled together and concentrated down and buffer exchanged into a final sample.

4.2   Paper II

Based on the idea and will to measure ^{1}H-R$_{1\rho}$ on non-isotopically labeled RNA samples, we developed a proton NMR-experiment that take advantage of homonuclear, proton J-couplings and use a coherence transfer step to rid certain spectral regions of unwanted signals.

This strategy, that we call the *SELOPE* strategy increases the accessibility of measurable signals in unlabeled RNAs, since signals can be moved to other spectral regions. This opens up for the possibility that after a selective excitation, for instance transfer signal from the aromatic pyrimidine H6 protons to the considerably upfield shifted H5 protons and then observe H8 protons (lacking ^{1}H J-coupling partner) that otherwise would overlap with H6 signals. Vice versa signal of aromatic H5 protons can be transferred downfield to the region of H6/H8/NH2s, allowing the study of H1´ protons in the sugar phosphate backbone of the RNA.

Excitation of the aromatic region ~7-8 ppm will inevitably also excite amino protons, luckily in unlabeled RNA the vicinity to quadrupolar nucleus ^{14}N for these protons will increase their transverse relaxation rates so that their signal will disappear during the duration of the coherence transfer step.

The SELOPE transfer principle can readily be combined with other NMR pulse sequence elements where a spectral decrowding would be useful. In this PhD-project the transfer step has commonly been preceded by a spin-lock for the measurement ^{1}H-R$_{1\rho}$ in a 1D manner, but also just a 2D experiment with a t1 delay preceding the transfer step, showing a purine diagonal and H6/H5 correlation cross peaks is a time efficient experiment that aid in the RNA assignment process.

^{1}H-R$_{1\rho}$ versions of the SELOPE experiment applied to unlabeled samples also gives the advantage that measurement is done on a ^{1}H that lack ^{13}C J-coupling partner (at least 98.93% of them) this opens up for the possibility to measure with lower spin-locking fields without having problems of artefacts arising from heteronuclear J-coupling.

4.3   Paper III

In Paper III we characterize the dynamics of the stem-loop part of the penultimate stem (h44) in *E. coli* 16S rRNA. This is done by extensive solution state, R$_{1\rho}$-measurements on a 25-mer

RNA-molecule containing the sequence of the loop. It is shown in our system that contains a UUCG-loop and a group of three poorly base-paired adenines further down in the helix that both nucleotides in the UUCG-loop and among the adenines are highly dynamic. It is well recognized that the UUCG-loop is known to be a unusually stable RNA-loop[187–190], yet the UUCG-loop in our system clearly displays dynamics with excited state lifetimes in the 1 ms range. The NMR-assignments of ^{1}H and ^{13}C nuclei in our UUCG-loop, in position 6 of pyrimidine bases, and position 8 of the guanine, as well as H1´/C1´nuclei agrees very well with assignments possible to find in litterature[165]. The structural change associated with the dynamical process in the UUCG-loop of our system, rearrange the base-pairing pattern in the loop part. The GU-base pair, typical to the UUCG-loop with the guanine base in a *syn*-conformation relative to its ribose part, is believed to rotate around its χ-bond during the exchange process instead form a canonical Watson-Crick base pair with a cytosine base preceding its original **U**UCG base-paring partner, forming a tri-loop and a necessary switch in the base pairing further down the helix. This rearrangement is supported in paper III by the usage additional RNA molecules, containing additional GC-base pairs that stabilize the hypothesized secondary structures of the conformational exchange process and trap the different states.

The group of adenines further down in the helix, that in the 70S ribosome bulge out and form a tertiary interaction with h8 in the 16S rRNA (see for instance PDB id 4wf1 & 5afi) are shown in paper III to display faster dynamics than many nucleotides in the UUCG-loop. In fact, several of the nuclei display so fast dynamics that attempts to clearly extract chemical differences from $R_{1\rho}$-off-resonance data have been unsuccessful. Yet for a few data sets this have been possible and for instance A9C1´ (A1447, in *E. coli* 16S rRNA) and U7C6 (U1445) indicate that these nucleotides become more helical during the dynamical process. The additional GC-base pairs in the modified RNA molecules that trap the different states in the conformational exchange process, does affect the dynamics in both the UUCG-loop and among the nucleotides in the region of the adenines, this is interpreted as a further support that the conformational exchange process among the adenines and the loop are coupled to some extent.

4.4 Studies of exchange and observation of cross-relaxation during imino ^{1}H-$R_{1\rho}$-experiments

The following sections in this chapter briefly presents and discuss the results from some limited characterizations of exchange and cross-relaxation processes among imino protons in our RNA-systems that have modeled certain rRNA regions.

4.4.1 Investigating cross-relaxation in a 39-mer RNA

During the characterization of conformational dynamics in in a 39-mer RNA construct (Fig. 4.1) modelling the B3 intersubunit bridge in human cytosolic ribosomes, ^{15}N-$R_{1\rho}$ experiments where preformed on imino nitrogens in the construct. On one particular imino nitrogen U29N3, in a GU-wobble base pair, a highly down field shifted (+31.3 ppm) excited state was observed (data not shown). This was interpreted as an indication of possible keto-enol tautomerism processes, as previously described by Kimsey et al.[85,86,178]. Since tautomerism in nucleobases is a highly interesting and widely discussed phenomenon[84,174,191] and ^{1}H-$R_{1\rho}$ methodology was easy at hand, it was tested to measure an on- and off-resonance data set on the corresponding imino proton in a non-isotopically labeled version of the sample (Fig. 4.1).
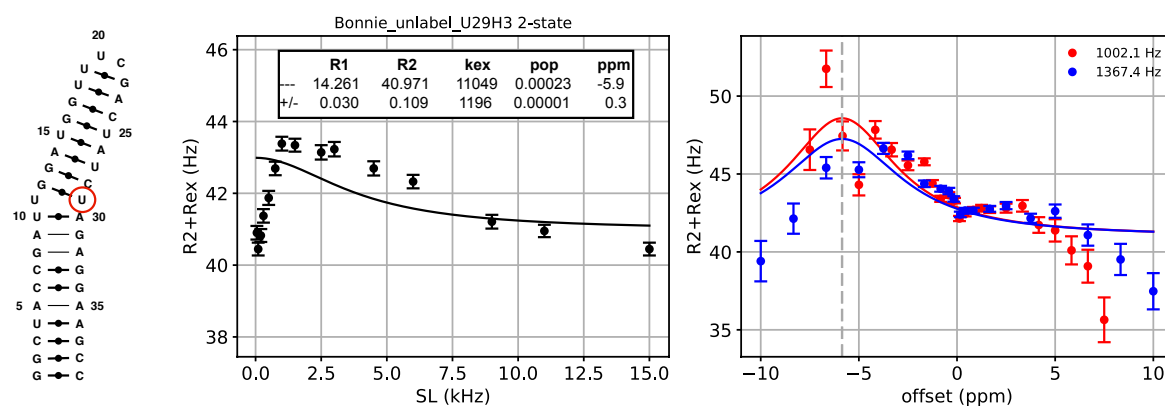
Figure 4.1. The secondary structure of the 39-mer construct is shown to the left, U29 is indicated with a red circle. To the right is shown a ^{1}H-$R_{1\rho}$ data set measured on the imino proton (H3) of U29, the data has been fitted to a 2-state "Laguerre equation". Except from the debatable appearance of the off-resonance curve, special attention should be payed to the increase in relaxation observed as a response to increased spin-lock power in the beginning of the on-resonance curve (in the region of spin-lock strengths ranging from 50.0 Hz – 1 kHz).

Initially after acquiring the data set shown in (Fig. 4.1) the steep increase in $R_{1\rho}$ in the beginning of the curve was very puzzling, it was acquired perfectly on-resonance so potential off-resonance effects was ruled out. After some time, it was hypothesized that increase in $R_{1\rho}$ could be due to cross-relaxation with the G12H1, i.e. a ROE artefact. As previously pointed out in (Fig. 1.24 in section 1.3.10) imino protons in GU-wobble base pairs are close in space. Once this was realized that cross-relaxation with the other imino proton was a potential culprit attempts were made to find a better mathematical model to fit the data with, one that take cross-relaxation into account. This was done by combining the terms for exchange from the Bloch-McConnell[151] equation with the terms for cross-relaxation as described by Allard et al.[126] and Lundström & Akke[133]. This yields a pretty large relaxation matrix containing exchange and cross-relaxation terms in the off-diagonal elements. The results of attempting to fit the data with such a model is shown in (Fig. 4.2).
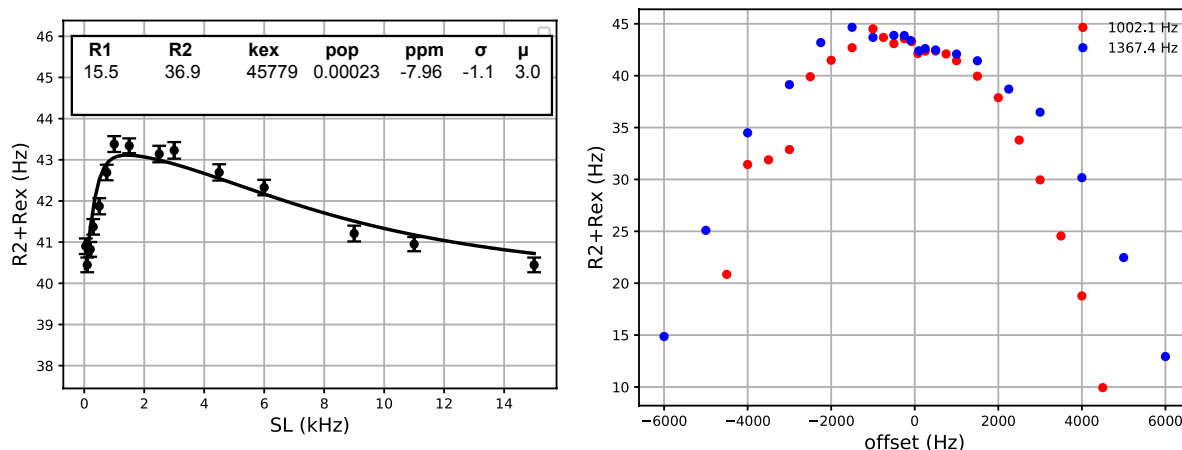
Figure 4.2. Fit of the imino proton $R_{1\rho}$ data of U29H3 in the 39-mer construct to a model taking into account cross-relaxation to the neighboring imino proton in the GU-wobble base pair G12H1. The offset to the cross-relaxing spin has been fixed to -752 Hz (taken from imino proton assignments) all other parameters has been allowed to float freely during the fit. R_1, R_2, k_{ex} are in units of s^{-1}, population $0.00023 = 0.023\%$, σ represents the longitudinal cross-relaxation rate constant and the μ transverse cross-relaxation rate, both in units of s^{-1}. Although this data set is not to be considered really conclusive, the major progress in testing this model was to be able to describe the rising part of the on-resonance curve as well as that the fitted values for the cross-relaxation constants were of reasonable magnitudes.

During the tests with this imino proton $R_{1\rho}$ data, possible water exchange was also considered. To investigate that, a simplified, 1D version of the water inversion experiment used by Rinnenthal et al.[179] was tested and the result of that experiment is shown in (Fig. 4.3). As can be seen in (Fig. 4.3) the imino proton signals of U29H3 is clearly exchanging with water as do G12H1, as both these signals becomes more rapidly inverted compared to for instance the signals of guanines from more stable regions of the construct. No real, quantitative fitting of H_2O-exchange rate constants have been attempted for the data in (Fig. 4.3) as possibly a suboptimal water inversion pulse was used and only a few shorter exchange delays was tested.

So, at this point it became clearer how incredibly complicated the exchange and cross-relaxation process for these imino protons are, also considering that both exchange and possible cross-relaxation with H_2O should be considered[179].
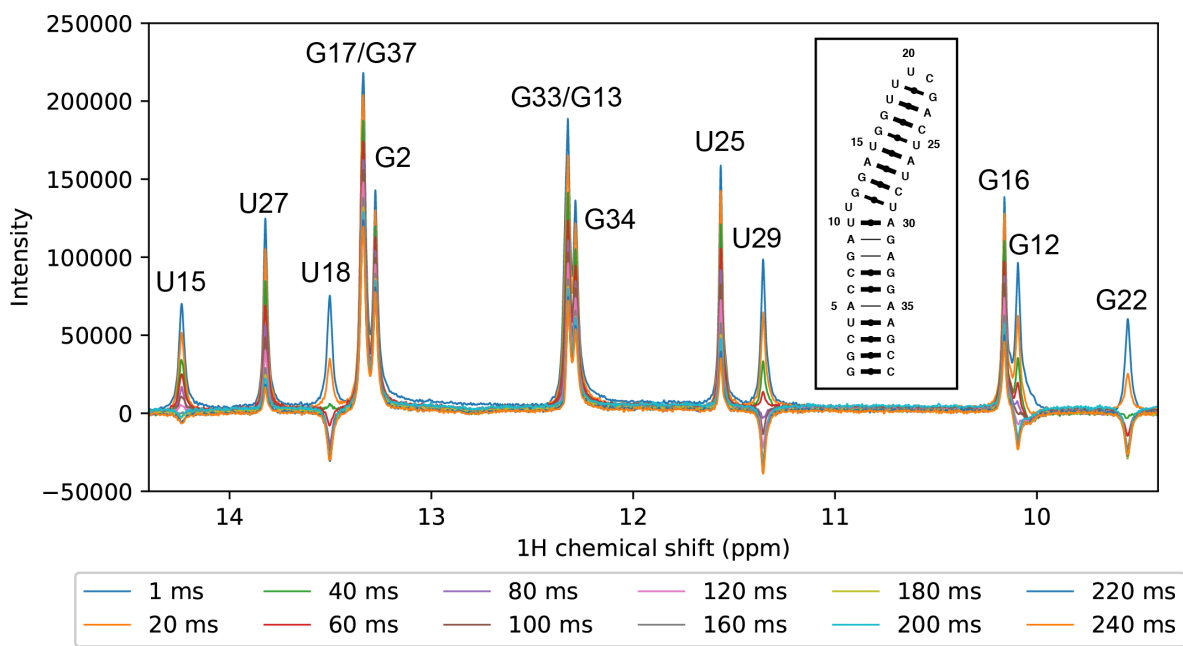
Figure 4.3. A series of 1D spectra from a water inversion experiment performed on the 39-mer RNA construct. The water signal is selectively inverted followed by a delay, during the delay "inverted" water molecules will exchange into the molecule, once subjected to a hard pulse and detection exchanged protons will give rise to peaks 180 degree out of phase, the longer the exchange delay and the faster water exchange is, peaks become more inverted. From these series of spectra, it is clear that U29H3, G12H1, G22H1, and U18H3 exchange with water and become inverted.

4.4.2   Fast exchange in imino protons in GC-base pair of a 29-mer RNA

In addition to the interesting imino proton in the GU-wobble base pair discussed in the previous section, 4.4.1 I would briefly, also like to address some other aspects of imino proton $R_{1\rho}$ data observed on a imino proton in a GC base pair in a 29-mer RNA construct modeling the B6 intersubunit bridge in the *E. coli* ribosome. The experimental data is shown in (Fig. 4.4). The data in (Fig. 4.4) is only on-resonance ^{1}H-$R_{1\rho}$ data (off-resonance data has not been acquired) and it has been fitted to equation (23 in section 1.3.8), it is indicative of an exchange process in the >30 kHz range. What is interesting with this imino proton data set is that otherwise in this construct there is very little relaxation dispersion and it is not indicative of extensive conformational exchange process, so a more "open" base pair or base pair rearrangements is not very likely for this construct at least the region of this GC base pair, yet the imino proton seem to participate in some kind of fast exchange process. The interesting question is, what could these exchange processes be? Could this be hydrogen transfers/tautomeric processes[192] in GC base pairs, giving rise to this on-resonance curve? To what extent can tunneling phenomena[193] play a role in such processes? Regarding the reliability of the exchange rate constant when just fitting on-resonance to equation 23 from section 1.3.8, an example from a brief simulation explaining this further is included in (Fig. 4.5).
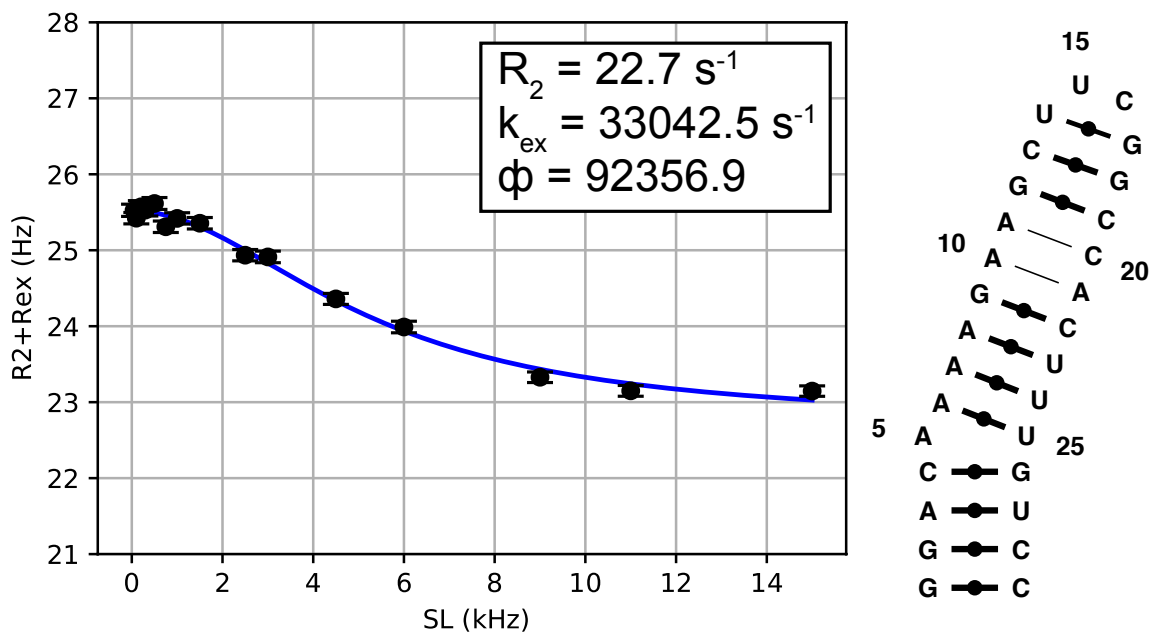
Figure 4.4. On-resonance imino proton $R_{1\rho}$ relaxation dispersion curve acquired on the G18H1 imino proton, a GC base pair that otherwise seem stable but clearly the imino proton seem to participate in a fast exchange process. The data has been fitted to equation 23 in section 1.3.8.
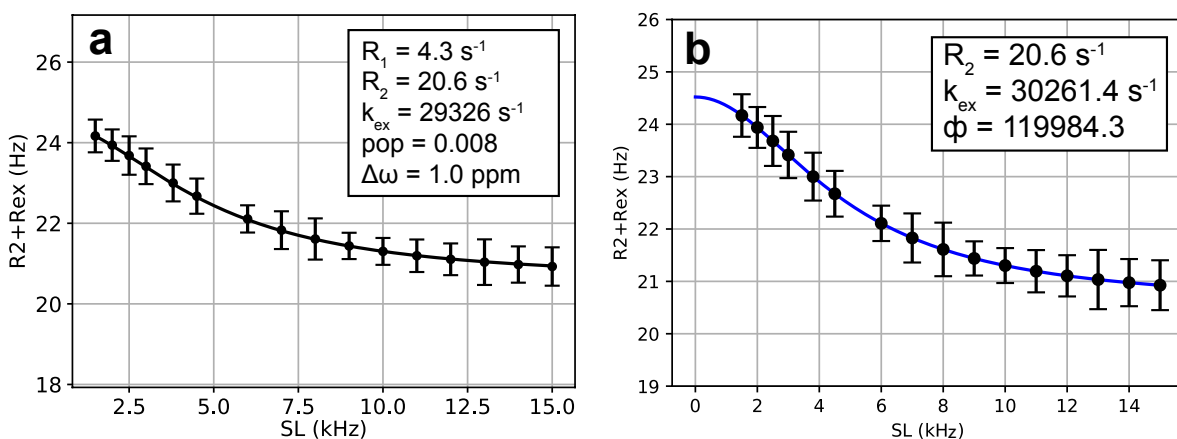

Figure 4.5. Simulations of fast exchange. A synthetic ^{1}H-$R_{1\rho}$ data set has been generated with the following set of parameters: R_1 = 4.0 s^{-1}, R_2 = 20.0 s^{-1}, k_{ex} = 30000.0 s^{-1}, population = 0.015, offset = -450.0 s^{-1}, spin-lock duration 80 ms, off-resonance spin-lock powers 1500 Hz and 2000 Hz. **a)** The synthetic data set has been fitted with the 2-state Laguerre equation (only fit and on-resonance part shown). **b)** Only the on-resonance part of the synthetic data set has been fitted to equation 23 in section 1.3.8. The point of showing fits of this simulated data set is to emphasize the validity of the fitted exchange rate constant (k_{ex}) when only fitting on-resonance data and using equation 23 in section 1.3.8.

5   CONCLUSIONS AND PERSPECTIVES

In this project we have used NMR-spectroscopy to characterize dynamical processes in small RNA molecules containing sequences of h44 in *E. coli* 16S rRNA. Especially the stem-loop part of *E. coli* h44 has been characterized. In addition to this we have developed an RNA-sample preparation method and new NMR-experiment allowing increased investigation of non-isotopically labeled RNA molecules.

5.1   Paper I

The HPLC-based RNA-sample production method developed and described in paper I, is a good alternative to other methods based on preparative PAGE or FPLC for purification and it has been used for the production of several RNA-samples in the lab and has proven useful and easy to use. The possibility to skip or only do the RP-IP step offers flexibility and by distributing the RNA material between multiple injections pure HPLC-fractions can be obtained. The method is also suitable for inexperienced HPLC users.

Even if this RNA-sample production method is a viable alternative to many other existing RNA-production -purification methods (i.e. PAGE, FPLC, ribozymes etc.) there is still room for continued development of RNA-sample production techniques. The resolution of nucleic acid separation in the purification techniques is probably the key point where improvements would be needed.

5.2   Paper II

The SELOPE experiment has clearly pin pointed the possibility to use coherence transfer methods to be able to study otherwise overlapped signals in RNA. This is useful since unlabeled nucleic acid material can be used more effectively for structure and dynamics measurements, saving money when usage of $^{13}C/^{15}N$ enriched nucleotides can be avoided.

The success of the SELOPE approach and what signals it might make available, of course depends on the chemical shifts of the signals in the RNA and is thus highly sample depended. As sample size increase, signal overlap is likely to increase as well, even with a SELOPE block. So, even if the SELOPE approach solves many problems there is of course a continued need for more techniques that allows investigation of larger RNAs with solution state NMR.

5.3   Paper III

Over all in Paper III, we observe an extensive conformational rearrangement of the stem-loop of *E. coli* h44, that includes a large structural change in the UUCG-loop, a change in base-paring pattern and a more helical conformation of several nucleotides in the region of the adenines that in the ribosome form a tertiary interaction with h8. We conclude that this rearrangement might interrupt the tertiary interaction with h8, allowing h44 to become more flexible and more easily interact with the LSU during translation.

To further investigate the behavior of the h44 stem-loop, exchanging conformers and the tertiary interaction between h44 and h8 within 70S ribosomal context, more experiments are needed. At the timepoint of writing this PhD-thesis, this is being done. 70S ribosomes where the 16S rRNA has been modified to contain the additional GC-base pairs shown in Paper III has been produced. These modified ribosomes have also been isolated using a rRNA streptavidin aptamer technology[194]. The isolated ribosomes have also been characterized *in vitro* together with the Sanyal lab at Uppsala University, in addition some initial cryo-EM data have been collected together with the Amunts lab at Stockholm University. This is still very much work in progress but at a later stage it will help us to give an even more detailed picture of the structure and behavior of the h44 stem-loop in *E. coli*, at that point it is likely that the additional data will be included in and extend Paper III.

5.4 Studies of exchange and observation of cross-relaxation during imino ^{1}H-$R_{1\rho}$-experiments

The investigations of exchange and cross-relaxation phenomena among imino protons is a work that somehow have connected to both the ribosomal RNA characterization as well as to the development and discussions of ^{1}H-$R_{1\rho}$ methodology. It is a very intriguing topic that poses many questions and I still believe that cross-relaxing imino protons in GU-wobble base pairs can be a suitable system for continued development and investigation of mathematical formalism for exchange intertwined with cross-relaxation.

It is also interesting to see how literature concerning this will develop in the future. Will someone be able to experimentally observe and characterize in great detail an imino proton participating in tautomerism processes, H_2O-exchange, cross-relaxation with both neighboring spins and water all at the same time in the same experiment?

6 ACKNOWLEDGEMENTS

7 REFERENCES

1.   Thieffry, D. & Sarkar, S. Forty years under the central dogma. *Trends in biochemical sciences*, **23**,312-316 (1998).

2.   Ostankovitch, M. & Pyle, A. M. Noncoding RNAs: A Story of Networks and Long-Distance Relationships. *Journal of Molecular Biology* **425**, 3577–3581 (2013).

3.   Higgs, P. G., & Lehman, N. The RNA World: molecular cooperation at the origins of life. *Nature Reviews Genetics*, **16**, 7-17 (2015).

4.   Schmeing, T. M. What recent ribosome structures have revealed about the mechanism of translation. *Nature*, **461**, 1234-1242 (2009).

5.   Ramakrishnan, V. The Ribosome Emerges from a Black Box. *Cell* **159**, 979–984 (2014).

6.   Fourmy, D., Recht, I.M., Blanchard, S.C. and Puglisi, J.D. Structure of the A Site of Escherichia coli 16S Ribosomal RNA Complexed with an Aminoglycoside Antibiotic. *Science* **274**, 1367-1371 (1996).

7.   Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature* **491**, 724–728 (2012).

8.   Szewczak, A.A. and Moore, P.B. The Sarcin/Ricin Loop, a Modular RNA. J*ournal of Molecular Biology*, **247**, 81-98 (1995).

9.   Cevec, M., Thibaudeau, C. & Plavec, J. NMR structure of the let-7 miRNA interacting with the site LCS1 of lin-41 mRNA from Caenorhabditis elegans. *Nucleic Acids Research* **38**, 7814–7821 (2010).

10. Baronti, L. *et al.* Base-pair conformational switch modulates miR-34a targeting of Sirt1 mRNA. *Nature* **583**, 139–144 (2020).

11. Hoogstraten, C. G., Legault, P., & Pardi, A. NMR Solution Structure of the Lead-dependent Ribozyme: Evidence for Dynamics in RNA Catalysis. *Journal of Molecular Biology*, **284**, 337-350 (1998).

12. Hoffmann, B. *et al.* NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proceedings of the National Academy of Sciences* **100**, 7003–7008 (2003).

13. Beckert, B. & Masquida, B. Synthesis of RNA by In Vitro Transcription. in *RNA* (ed. Nielsen, H.) vol. 703 pp. 29–41 (Humana Press, 2011).

14. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Research* **15**, 8783–8798 (1987).

15. Milligan, J. F. & Uhlenbeck, O. C. [5] Synthesis of small RNAs using T7 RNA polymerase. in *Methods in Enzymology* vol. 180 pp. 51–62 (Elsevier, 1989).

16. Karlsson, H., Baronti, L. & Petzold, K. A robust and versatile method for production and purification of large-scale RNA samples for structural biology. *RNA* **26**, 1023–1037 (2020).

17. Price, S. R., Ito, N., Oubridge, C., Avis, J. M. & Nagai, K. Crystallization of RNA-protein complexes I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *Journal of Molecular Biology* **249**, 398–408 (1995).

18. Kao, C., Zheng, M. & RüDisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3′ terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268–1272 (1999).

19. Helmling, C. *et al.* Rapid NMR screening of RNA secondary structure and binding. *Journal of Biomolecular NMR* **63**, 67–76 (2015).

20. Chen, Z. & Zhang, Y. Dimethyl sulfoxide targets phage RNA polymerases to promote transcription. *Biochemical and Biophysical Research Communications* **333**, 664–670 (2005).

21. Ferre-D'Amare, A. R. & Doudna, J. A. Use of Cis- and Trans-Ribozymes to Remove 5' and 3' Heterogeneities From Milligrams of In Vitro Transcribed RNA. *Nucleic Acids Research* **24**, 977–978 (1996).

22. Schurer, H. A universal method to produce in vitro transcripts with homogeneous 3' ends. *Nucleic Acids Research* **30**, 56e–556 (2002).

23. Shields, T. P., Mollova, E., Marie, L. Ste., Hansen, M. R. & Pardi, A. High-performance liquid chromatography purification of homogenous-length RNA produced by trans cleavage with a hammerhead ribozyme. *RNA* **5**, 1259–1267 (1999).

24. Wichlacz, A. Generating in vitro transcripts with homogenous 3' ends using trans-acting antigenomic delta ribozyme. *Nucleic Acids Research* **32**, 39e–339 (2004).

25. Duss, O., Maris, C., von Schroetter, C. & Allain, F. H.-T. A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. *Nucleic Acids Research* **38**, e188–e188 (2010).

26. Feyrer, H., Munteanu, R., Baronti, L. & Petzold, K. One-Pot Production of RNA in High Yield and Purity Through Cleaving Tandem Transcripts. *Molecules* **25**, 1142 (2020).

27. Caruthers, M. H. A brief review of DNA and RNA chemical synthesis. *Biochemical Society Transactions* **39**, 575–580 (2011).

28. Somoza, Á. Protecting groups for RNA synthesis: an increasing need for selective preparative methods. *Chem. Soc. Rev.* **37**, 2668 (2008).

29. Baronti, L., Karlsson, H., Marušič, M. & Petzold, K. A guide to large-scale RNA sample preparation. *Analytical and Bioanalytical Chemistry* **410**, 3239–3252 (2018).

30. Dominguez, C., Schubert, M., Duss, O., Ravindranathan, S. & Allain, F. H.-T. Structure determination and dynamics of protein–RNA complexes by NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* **58**, 1–61 (2011).

31. Skoog, D.A., Holler, F.J. and Crouch, S.R. Ch. 30 Capillary Electrophoresis, Capillary Electro-chromatography, and Field-Flow Fractionation. in *Principles of Instrumental Analysis* pp. 867-892 (Brooks/Cole, Cengage Learning, 2007).

32. Barril, P. and Nates, S. Gel-Electrophoresis and Its Applications. in (ed. Magdeldin, S.) *Gel Electrophoresis: Principles and Basics.* pp. 15-33 (InTech Croatia, 2015).

33. Shi, Q. and Jackowski, G. One-dimensional polyacrylamide gel electrophoresis. in (ed. Hames, B.D.) *Gel Electrophoresis of Proteins: A Practical Approach*. pp. 1-8 (Oxford University Press,1998).

34. Chrambach, A. & Rodbard, D. Polyacrylamide Gel Electrophoresis. *Science* **172**, 440-451 (1971).

35. Rio, D. C., Ares, M., Hannon, G. J. & Nilsen, T. W. Polyacrylamide Gel Electrophoresis of RNA. *Cold Spring Harbor Protocols* **2010**, 1-6 (2010).

36. Reijnders, L., Sloof, P., Sival, J. & Borst, P. Gel electrophoresis of RNA under denaturing conditions. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **324**, 320–333 (1973).

37. Lehrach, H., Diamond, D., Wozney, J.M and Boedtker, H. RNA Molecular Weight Determinations by Gel Eelctrophoresis under Denatruing Conditions, a Critical Reexamination. *Biochemistry*, **16**, 4743-4751 (1977).

38. Waring, M. J. Complex formation between ethidium bromide and nucleic acids. *Journal of Molecular Biology* **13**, 269–282 (1965).

39. Macgregor, J. T. & Johnson, I. J. In vitro metabolic activation of ethidium bromide and other phenanthridinium compounds: Mutagenic activity in Salmonella typhimurium. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **48**, 103–107 (1977).

40. Kirsanov, K. I., Lesovaya, E. A., Yakubovskaya, M. G. & Belitsky, G. A. SYBR Gold and SYBR Green II are not mutagenic in the Ames test. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **699**, 1–4 (2010).

41. Petrov, A., Tsa, A. & Puglisi, J. D. Analysis of RNA by Analytical Polyacrylamide Gel Electrophoresis. in *Methods in Enzymology* vol. 530 301–313 (Elsevier, 2013).

42. Petrov, A., Wu, T., Puglisi, E. V. & Puglisi, J. D. RNA Purification by Preparative Polyacrylamide Gel Electrophoresis. in *Methods in Enzymology* vol. 530 315–330 (Elsevier, 2013).

43. Kanwal, F., Chen, T., Zhang, Y., Simair, A. & Lu, C. A Modified In Vitro Transcription Approach to Improve RNA Synthesis and Ribozyme Cleavage Efficiency. *Mol Biotechnol* **61**, 469–476 (2019).

44. Hagen, F. S. & Young, E. T. Preparative polyacrylamide gel electrophoresis of ribonucleic acid. Identification of multiple molecular species of bacteriophage T7 lysozyme messenger ribonucleic acid. *Biochemistry* **13**, 3394–3400 (1974).

45. Cunningham, L. Preparative-scale purification of RNA using an efficient method which combines gel electrophoresis and column chromatography. *Nucleic Acids Research* **24**, 3647–3648 (1996).

46. Skoog, D.A., Holler, F.J. and Crouch, S.R. Ch. 26 An introduction to Chromatographic Separations in *Principles of Instrumental Analysis* pp. 763-787 (Brooks/Cole, Cengage Learning, 2007).

47. Andrés, A., Broeckhoven, K. & Desmet, G. Methods for the experimental characterization and analysis of the efficiency and speed of chromatographic columns: A step-by-step tutorial. *Analytica Chimica Acta* **894**, 20–34 (2015).

48. Lukavsky, P. J. Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides. *RNA* **10**, 889–893 (2004).

49. Kim, I., McKenna, S. A., Viani Puglisi, E. & Puglisi, J. D. Rapid purification of RNAs using fast performance liquid chromatography (FPLC). *RNA* **13**, 289–294 (2006).

50. McKenna, S. A. *et al.* Purification and characterization of transcribed RNAs using gel filtration chromatography. *Nature Protocols* **2**, 3270–3277 (2007).

51. Easton, L. E., Shibata, Y. & Lukavsky, P. J. Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. *RNA* **16**, 647–653 (2010).

52. Koubek, J., Lin, K. F., Chen, Y. R., Cheng, R. P. & Huang, J. J. T. Strong anion-exchange fast performance liquid chromatography as a versatile tool for preparation and purification of RNA produced by in vitro transcription. *RNA* **19**, 1449–1459 (2013).

53. McGinnis, A. C., Chen, B. & Bartlett, M. G. Chromatographic methods for the determination of therapeutic oligonucleotides. *Journal of Chromatography B* **883–884**, 76–94 (2012).

54. Martins, R., Queiroz, J. A. & Sousa, F. Ribonucleic acid purification. *Journal of Chromatography A* **1355**, 1–14 (2014).

55. Huang, Z., Jayaseelan, S., Hebert, J., Seo, H. & Niu, L. Single-nucleotide resolution of RNAs up to 59 nucleotides by high-performance liquid chromatography. *Analytical Biochemistry* **435**, 35–43 (2013).

56. McCarthy, S. M., Gilar, M. & Gebler, J. Reversed-phase ion-pair liquid chromatography analysis and purification of small interfering RNA. *Analytical Biochemistry* **390**, 181–188 (2009).

57. Murray, J.B., Collier, K.C. and Arnold, J.R.P. A General Purification Procedure for Chemically Synthesized Oligoribonucleotides. *Analytical Biochemistry*, **218**, 177-184 (1994).

58. Andersson, A.C., Scaringe, S.A., Earp, B.E. and Frederick, C.A. HPLC purification of RNA for crystallography and NMR. *RNA,* **2**, 110-117 (1996).

59. Wilson, D. N. & Doudna Cate, J. H. The Structure and Function of the Eukaryotic Ribosome. *Cold Spring Harbor Perspectives in Biology* **4**, a011536–a011536 (2012).

60. Schuwirth, B. S. *et al.* Structures of the Bacterial Ribosome at 3.5 Å Resolution. *Science* **310**, 827–834 (2005).

61. Amunts, A., Brown, A., Toots, J., Scheres, S. H. W. & Ramakrishnan, V. The structure of the human mitochondrial ribosome. *Science* **348**, 95–98 (2015).

62. Ogle, J. M. *et al.* Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit. **292**, 7 (2001).

63. Yusupov, M. M. *et al.* Crystal Structure of the Ribosome at 5.5 Å Resolution. **292**, 15 (2001).

64. Schluenzen, F. *et al.* Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution. *Cell* **102**, 615-623 (2000).

65. Selmer, M. Structure of the 70S Ribosome Complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).

66. Fischer, N. *et al.* Structure of the E. coli ribosome–EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520**, 567–570 (2015).

67. Decatur, W. A. & Fournier, M. J. rRNA modifications and ribosome function. *Trends in Biochemical Sciences* **27**, 344–351 (2002).

68. Natchiar, S. K., Myasnikov, A. G., Kratzat, H., Hazemann, I. & Klaholz, B. P. Visualization of chemical modifications in the human 80S ribosome structure. *Nature,* **551**, 472-477 (2017).

69. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).

70. Nakagawa, S., Niimura, Y., Miura, K. -i. & Gojobori, T. Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proceedings of the National Academy of Sciences* **107**, 6382–6387 (2010).

71. Antoun, A., Pavlov, M. Y., Lovmar, M. & Ehrenberg, M. How Initiation Factors Maximize the Accuracy of tRNA Selection in Initiation of Bacterial Protein Synthesis. *Molecular Cell* **23**, 183–193 (2006).

72. Huang, C., Mandava, C. S. & Sanyal, S. The Ribosomal Stalk Plays a Key Role in IF2-Mediated Association of the Ribosomal Subunits. *Journal of Molecular Biology* **399**, 145–153 (2010).

73. Voorhees, R. M., Weixlbaumer, A., Loakes, D., Kelley, A. C. & Ramakrishnan, V. Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol* **16**, 528–533 (2009).

74. Pulk, A. & Cate, J. H. D. Control of Ribosomal Subunit Rotation by Elongation Factor G. *Science* **340**, 1235970 (2013).

75. Frank, J., Gao, H., Sengupta, J., Gao, N. & Taylor, D. J. The process of mRNA–tRNA translocation. *PNAS*, **104**, 19671-19678 (2007).

76. Dunkle, J. A. & Cate, J. H. D. Ribosome Structure and Dynamics During Translocation and Termination. *Annual Review of Biophysics* **39**, 227–244 (2010).

77. Fu, Z. *et al.* Key Intermediates in Ribosome Recycling Visualized by Time-Resolved Cryoelectron Microscopy. *Structure* **24**, 2092–2101 (2016).

78. Smith, B. A., Gupta, N., Denny, K. & Culver, G. M. Characterization of 16S rRNA Processing with Pre-30S Subunit Assembly Intermediates from E. coli. *Journal of Molecular Biology* **430**, 1745–1759 (2018).

79. Moazed, D., Stern, S. & Noller, H. F. Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension. *Journal of Molecular Biology* **187**, 399–416 (1986).

80. Chakravorty, S., Helb, D., Burday, M., Connell, N. & Alland, D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods* **69**, 330–339 (2007).

81. Yang, B., Wang, Y. & Qian, P.-Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* **17**, 135 (2016).

82. Pavlov, M. Y., Liljas, A. & Ehrenberg, M. A recent intermezzo at the Ribosome Club. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**, 20160185 (2017).

83. Loveland, A. B., Demo, G. & Korostelev, A. A. Cryo-EM of elongating ribosome with EF-Tu•GTP elucidates tRNA proofreading. *Nature* **584**, 640–645 (2020).

84. Topal, M. D. & Fresco, J. R. Complementary base pairing and the origin of substitution mutations. *Nature* **263**, 285–289 (1976).

85. Kimsey, I. J., Petzold, K., Sathyamoorthy, B., Stein, Z. W. & Al-Hashimi, H. M. Visualizing transient Watson–Crick-like mispairs in DNA and RNA duplexes. *Nature* **519**, 315–320 (2015).

86. Kimsey, I. J. *et al.* Dynamic basis for dG•dT misincorporation via tautomerization and ionization. *Nature* **554**, 195–201 (2018).

87. Liu, Q. & Fredrick, K. Intersubunit Bridges of the Bacterial Ribosome. *Journal of Molecular Biology* **428**, 2146–2164 (2016).

88. Shaikh, T. R. *et al.* Initial bridges between two ribosomal subunits are formed within 9.4 milliseconds, as studied by time-resolved cryo-EM. *Proc Natl Acad Sci USA* **111**, 9822–9827 (2014).

89. Sun, Q. Mutations in the intersubunit bridge regions of 16S rRNA affect decoding and subunit–subunit interactions on the 70S ribosome. *Nucleic Acids Research* **39**, 3321-3330.

90. Zhang, W., Dunkle, J. A. & Cate, J. H. D. Structures of the Ribosome in Intermediate States of Ratcheting. *Science* **325**, 1014–1017 (2009).

91. Magnet, S. & Blanchard, J. S. Molecular Insights into Aminoglycoside Action and Resistance. *Chem. Rev.* **105**, 477–498 (2005).

92. Holm, M., Mandava, C. S., Ehrenberg, M. & Sanyal, S. The mechanism of error induction by the antibiotic viomycin provides insight into the fidelity mechanism of translation. *eLife* **8**, e46124 (2019).

93. Dibrov, S. M., Parsons, J. & Hermann, T. A model for the study of ligand binding to the ribosomal RNA helix h44. *Nucleic Acids Research* **38**, 4458–4465 (2010).

94. Forge, A. & Schacht, J. Aminoglycoside Antibiotics. *Audiology and Neurotology* **5**, 3-22 (2000).

95. Hobbie, S. N. *et al.* Genetic analysis of interactions with eukaryotic rRNA identify the mitoribosome as target in aminoglycoside ototoxicity. *Proceedings of the National Academy of Sciences* **105**, 20888–20893 (2008).

96. Condon, C. N., Liveris, D., Squires, C., Schwartz, I. & Squires, C. L. rRNA Operon Multiplicity in Escherichia coli and the Physiological Implications of rrn Inactivation. *J. BACTERIOL.* **177**, 5 (1995).

97. Brodersen, D. E. *et al.* The Structural Basis for the Action of the Antibiotics Tetracycline, Pactamycin, and Hygromycin B on the 30S Ribosomal Subunit. *Cell* **103**, 1143-1154 (2000).

98. Svetlov, M. S. *et al.* High-resolution crystal structures of ribosome-bound chloramphenicol and erythromycin provide the ultimate basis for their competition. *RNA* **25**, 600–606 (2019).

99. Llano-Sotelo, B. *et al.* Binding and Action of CEM-101, a New Fluoroketolide Antibiotic That Inhibits Protein Synthesis. *AAC* **54**, 4961–4970 (2010).

100. Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Reviews Microbiology* **12**, 35–48 (2014).

101. Fischbach, M. A. & Walsh, C. T. Antibiotics for Emerging Pathogens. **325**, 6 (2009).

102. Lin, J., Zhou, D., Steitz, T. A., Polikanov, Y. S. & Gagnon, M. G. Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annual Review of Biochemistry* **87**, 451–478 (2018).

103. Yassin, A., Fredrick, K. & Mankin, A. S. Deleterious mutations in small subunit ribosomal RNA identify functional sites and potential targets for antibiotics. *Proceedings of the National Academy of Sciences* **102**, 16620–16625 (2005).

104. Levitt, M.H. Ch. 6 Review of Quantum Mechanics. in *Spin Dynamics: Basics of Nuclear Magnetic Resonance* pp. 127-168 (John Wiley & Sons, 2006).

105. Cavanagh, J., Fairbrother, W.J., Palmer A.G. and Skelton, N.J. Ch. 1 Classical NMR Spectroscopy. in *Protein NMR Spectroscopy: Principles and Practice* pp. 1-28 (Elsivier Academic Press, 2007).

106. Keeler, J. Ch. 4 The Vector Model. in *Understanding NMR Spectroscopy,* 2nd ed. pp. 47-75 (John Wiley & Sons Ltd., 2010).

107. Skoog, D.A., Holler, F.J. and Crouch, S.R. Ch. 19 Nuclear Magnetic Resonance Spectroscopy in *Principles of Instrumental Analysis* pp. 498-549 (Brooks/Cole, Cengage Learning, 2007).

108. Levitt, M.H Ch. 10 Ensemble of Spins-1/2. in *Spin Dynamics: Basics of Nuclear Magnetic Resonance* pp. 273-313 (John Wiley & Sons, 2006).

109. Lee, J. H., Okuno, Y. & Cavagnero, S. Sensitivity enhancement in solution NMR: Emerging ideas and new frontiers. *Journal of Magnetic Resonance* **241**, 18–31 (2014).

110. Griesinger, C., Schwalbe, H., Schleucher, J. and Sattler, M. Ch. 3 Proton-Detected Heteronuclear and Multidimensional NMR. in (ed. Croasmun, W.R. & Carlsson, M.K.) *Two-Dimensional NMR Spectroscopy* pp. 458-580 (VCH Publishers, Inc., 1994)

111. Callaghan, P.T. Ch. 2 Introductory Nuclear Magnetic Resonance. in *Principles of Nuclear Magnetic Resonance Microscopy* pp. 25-91 (Oxford University Press, 1991).

112. Freeman, R. Continuous wave spectroscopy. in *A Handbook of Nuclear Magnetic Resonance* pp. 52-58 (Longman Group UK Limited, 1988).

113. Ernst, R.R. and Andersson, W.A. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Review of Scientific Instruments* **37**, 93-102 (1966).

114. Cavanagh, J., Fairbrother, W.J., Palmer A.G. and Skelton, N.J. Ch. 3 Experimental Aspects of NMR Spectroscopy. in *Protein NMR Spectroscopy: Principles and Practice* pp. 124-136 (Elsivier Academic Press, 2007).

115. Derome, A.E. Ch. 2 Why Bother With Pulse NMR. in *Modern NMR Techniques for Chemistry Research* pp. 9-29 (Pergamon Books Ltd., 1987).

116. Freeman, R. Shaped radiofrequency pulses in high resolution NMR. *Journal of Progress in Nuclear Magnetic Resonance Spectroscopy* **32**, 59-106 (1998).

117. Topspin3.5pl7 auxiliary program "calcpowlev" source code.

118. Keeler, J. How the spectrometer works. in *Understanding NMR Spectroscopy,* 2nd ed. pp. 483-493 (John Wiley & Sons Ltd., 2010).

119. Harris, R. K., Becker, E. D., Goodfellow, R. & Granger, P. (IUPAC Recommendations 2001). *Pure and Applied Chemistry* 24 (2001).

120. Ulrich, E. L. *et al.* BioMagResBank. *Nucleic Acids Research* **36**, D402–D408 (2007).

121. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. The Protein Data Bank. *Nucleic Acids Research* **28**, 235-242 (2000).

122. Popenda, M. *et al.* RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics* **11**, 231 (2010).

123. Bagby, S. J Coupling. in (ed. Roberts, G.C.K) *Encyclopedia of Biophysics* pp. 1163-1173 (Springer Berlin Heidelberg, 2013).

124. Keeler, J. Ch. 7 Product operators. in *Understanding NMR Spectroscopy*, 2nd ed. pp. 139-181 (John Wiley & Sons Ltd., 2010).

125. Levitt, M.H. Ch. 12 Homonuclear AX System. in *Spin Dynamics: Basics of Nuclear Magnetic Resonance* pp. 340-385 (John Wiley & Sons, 2006).

126. Allard, P., Helgstrand, M. & Härd, T. A Method for Simulation of NOESY, ROESY, and Off-Resonance ROESY Spectra. *Journal of Magnetic Resonance* **129**, 19–29 (1997).

127. Allard, P., Helgstrand, M. & Härd, T. The Complete Homogeneous Master Equation for a Heteronuclear Two-Spin System in the Basis of Cartesian Product Operators. *Journal of Magnetic Resonance* **134**, 7–16 (1998).

128. Schlagnitweit, J., Steiner, E., Karlsson, H. & Petzold, K. Efficient Detection of Structure and Dynamics in Unlabeled RNAs: The SELOPE Approach. *Chemistry - A European Journal* **24**, 6067–6070 (2018).

129. Neuhaus, D. and Williamson, M. Ch. 2 The Steady-State NOE for Two Spins. in *The Nuclear Overhauser Effect in Structural and Conformational Analysis* pp. 23-62 (VCH Publishers, Inc. 1989).

130. Canet, D. Chapter 1. Introduction to Nuclear Spin Cross-relaxation and Cross-correlation Phenomena in Liquids. in *New Developments in NMR* (ed. Canet, D.) pp. 1–60 (Royal Society of Chemistry, 2017).

131. Keeler, J. Ch. 9 Relaxation and the NOE. in *Understanding NMR Spectroscopy*, 2nd ed. pp. 242-316 (John Wiley & Sons Ltd., 2010).

132. Bax, A. & Grzesiek, S. ROESY. in *Encyclopedia of Magnetic Resonance* (ed. Harris, R. K.) (John Wiley & Sons, Ltd, 2007).

133. Lundström, P. & Akke, M. Off-resonance rotating-frame amide proton spin relaxation experiments measuring microsecond chemical exchange in proteins. *Journal of Biomolecular NMR* **32**, 163–173 (2005).

134. Gray, G.A. Introduction to Multidimensional NMR methods. in (ed. Croasmun, W.R. & Carlsson, M.K.) *Two-Dimensional NMR Spectroscopy* pp. 1-12 (VCH Publishers, Inc., 1994)

135. Keeler, J. Ch. 8 Two-dimensional NMR. in *Understanding NMR Spectroscopy*, 2nd ed. pp. 184-239 (John Wiley & Sons Ltd., 2010).

136. States, D. J., Haberkorn, R. A. & Ruben, D. J. A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *Journal of Magnetic Resonance (1969)* **48**, 286–292 (1982).

137. Kowalewski, J. and Mäler, L. Nuclear Spin-Relaxation in Liquids. (CRC Press, 2018).

138. Bloch, F. Nuclear induction. *Physical review.* **70**, 460-474 (1946).

139. Murase, K. A theoretical and numerical consideration of the longitudinal and transverse relaxations in the rotating frame. *Magnetic Resonance Imaging* **31**, 1544–1558 (2013).

140. Murase, K. & Tanki, N. Numerical solutions to the time-dependent Bloch equations revisited. *Magnetic Resonance Imaging* **29**, 126–131 (2011).

141. Murase, K. Numerical Analysis of the Magnetization Behavior in Magnetic Resonance Imaging in the Presence of Multiple Chemical Exchange Pools. *Open Journal of Applied Sciences* **07**, 1–14 (2017).

142. Tidell, C. How to solve systems of differential equations. https://www.youtube.com/watch?v=aJlSjNAB-2Y&t=530s (accessed 2021, February 18).

143. Strang, G. (2005) MIT OpenCourseWare, 18.06 Linear Algebra, lecture 23. Differential Equations and exp(At). https://www.youtube.com/watch?v=IZqwi0wJovM&list=PLE7DDD91010BC51F8&index=24&t=690s (accessed 2021, February 18).

144. https://www.scipy.org

145. Holst, A. and Ufnarovski, V. Ch. 6 Spectral Theory. in *Matrix Theory* pp. 141-164 (Studentlitteratur, 2014).

146. Palmer, A. G., Kroenke, C. D. & Patrick Loria, J. Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules. in *Methods in Enzymology* vol. 339 204–238 (Elsevier, 2001).

147. Ladbury J.E. Thermodynamics of Biomolecular Intractions. in (e.d Roberts, G.C.K.) *Encyclopedia of Biophysics* pp. 2589-2606 (Springer-Verlag Berlin Heidelberg, 2012).

148. Korzhnev, D. M., Orekhov, V. Y., Dahlquist, F. W. & Kay, L. E. Off-resonance R1ρ relaxation outside of the fast exchange limit: An experimental study of a cavity mutant of T4 lysozyme. *Journal of Biomolecular NMR* **26**, 39-48 (2003).

149. Al-Hashimi, H. M. NMR studies of nucleic acid dynamics. *Journal of Magnetic Resonance* **237**, 191–204 (2013).

150. https://numpy.org

151. McConnell, H. M. Reaction Rates by Nuclear Magnetic Resonance. *The Journal of Chemical Physics* **28**, 430–431 (1958).

152. Trott, O. & Palmer, A. G. R1ρ Relaxation outside of the Fast-Exchange Limit. *Journal of Magnetic Resonance* **154**, 157–160 (2002).

153. Miloushev, V. Z. & Palmer, A. G. R1ρ relaxation for two-site chemical exchange: General approximations and some exact solutions. *Journal of Magnetic Resonance* **177**, 221–227 (2005).

154. Davis, D.G., Perlman, M.E. and London, R.E. Direct Measurements of Dissociation-Rate Constant for Inhibitor-Enzyme Complexes via T1rho and T2 (CPMG) Methods. *Journal of Magnetic Resonance, series B* **104**, 266-275 (1994).

155. Bothe, J. R., Stein, Z. W. & Al-Hashimi, H. M. Evaluating the uncertainty in exchange parameters determined from off-resonance R1ρ relaxation dispersion for systems in fast exchange. *Journal of Magnetic Resonance* **244**, 18–29 (2014).

156. Bothe, J. R. *et al.* Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nature Methods* **8**, 919–931 (2011).

157. Zhuravleva, A. & Korzhnev, D. M. Protein folding by NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **100**, 52–77 (2017).

158. Palmer, A. G. & Massi, F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chemical Reviews* **106**, 1700–1719 (2006).

159. Korzhnev, D. M., Orekhov, V. Yu. & Kay, L. E. Off-Resonance R $_{1\rho}$ NMR Studies of Exchange Dynamics in Proteins with Low Spin-Lock Fields: An Application to a Fyn SH3 Domain. *Journal of the American Chemical Society* **127**, 713–721 (2005).

160. Hansen, D. F., Vallurupalli, P. & Kay, L. E. An Improved ^{15}N Relaxation Dispersion Experiment for the Measurement of Millisecond Time-Scale Dynamics in Proteins [†]. *J. Phys. Chem. B* **112**, 5898–5904 (2008).

161. Nolis, P. & Parella, T. Solution-State NMR Experiments Based on Heteronuclear Cross-Polarization. *CAC* **3**, 47–68 (2007).

162. Wishart, D. S. *et al.* 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135–140 (1995).

163. Aeschbacher, T., Schubert, M. & Allain, F. H.-T. A procedure to validate and correct the 13C chemical shift calibration of RNA datasets. *J Biomol NMR* **52**, 179–190 (2012).

164. Fürtig, B., Richter, C., Wöhnert, J. & Schwalbe, H. NMR Spectroscopy of RNA. *ChemBioChem* **4**, 936–962 (2003).

165. Fürtig, B., Richter, C., Bermel, W. & Schwalbe, H. New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUCG tetraloop. *J Biomol NMR* **28**, 69–79 (2004).

166. Nozinovic, S., Fürtig, B., Jonker, H. R. A., Richter, C. & Schwalbe, H. High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Research* **38**, 683–694 (2010).

167. Akke, M., Fiala, R., Jiang, F., Patel, D. and Palmer III, A.G. Base dynamics in a UUCG tetraloop RNA hairpin characterized by 15N spin relaxation: Correlations with structure and stability. *RNA* **3**, 702-709 (1997).

168. Borkar, A. N., Vallurupalli, P., Camilloni, C., Kay, L. E. & Vendruscolo, M. Simultaneous NMR characterisation of multiple minima in the free energy landscape of an RNA UUCG tetraloop. *Phys. Chem. Chem. Phys.* **19**, 2797–2804 (2017).

169. Bottaro, S., Nichols, P. J., Vögeli, B., Parrinello, M. & Lindorff-Larsen, K. Integrating NMR and simulations reveals motions in the UUCG tetraloop. *Nucleic Acids Research* **48**, 5839–5848 (2020).

170. Williams, D. J. & Hall, K. B. Experimental and computational studies of the G[UUCG]C RNA tetraloop11Edited by I. Tinoco. *Journal of Molecular Biology* **297**, 1045–1061 (2000).

171. Lynch, S. R., Gonzalez, R. L. & Puglisi, J. D. Comparison of X-Ray Crystal Structure of the 30S Subunit-Antibiotic Complex with NMR Structure of Decoding Site Oligonucleotide-Paromomycin Complex. *Structure* **11**, 43–53 (2003).

172. Zeng, X., Chugh, J., Casiano-Negroni, A., Al-Hashimi, H. M. & Brooks, C. L. Flipping of the Ribosomal A-Site Adenines Provides a Basis for tRNA Selection. *Journal of Molecular Biology* **426**, 3201–3213 (2014).

173. Carter, A. P. *et al.* Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340–348 (2000).

174. Ruterjans, H., Kaun, E., Hull, W.E. and Limbach H.H. Evidence for tautomerism in nucleic acid base pairs 1H NMR study of 15N labeled tRNA. *Nucleic Acids Research.* **10**, 7027-7039 (1982).

175. Sowers, L. C., Goodman, M. F., Eritja, R., Kaplan, B. & Fazakerley, G. V. Ionized and wobble base-pairing for bromouracil-guanine in equilibrium under physiological conditions. *Journal of Molecular Biology* **205**, 437–447 (1989).

176. Cho, B. P., Kadlubar, F. F., Culp, S. J. & Evans, F. E. Nitrogen-15 nuclear magnetic resonance studies on the tautomerism of 8-hydroxy-2'-deoxyguanosine, 8-hydroxyguanosine, and other C8-substituted guanine nucleosides. *Chemical Research in Toxicology* **3**, 445–452 (1990).

177. Sowers S, L. C., Eritjag, R., Kaplans, B., Goodmany, M. & Fazakerlyll, G. V. Equilibrium between a Wobble and Ionized Base PairFormed between Fluorouracil and Guanine in DNA as Studied by Proton and Fluorine NMR. *The Journal of Biological Chemistry* **263**, 14794-14801 (1988).

178. Szymanski, E. S., Kimsey, I. J. & Al-Hashimi, H. M. Direct NMR Evidence that Transient Tautomeric and Anionic States in dG·dT Form Watson–Crick-like Base Pairs. *Journal of the American Chemical Society* **139**, 4326–4329 (2017).

179. Rinnenthal, J., Klinkert, B., Narberhaus, F. & Schwalbe, H. Direct observation of the temperature-induced melting process of the Salmonella fourU RNA thermometer at base-pair resolution. *Nucleic Acids Research* **38**, 3834–3847 (2010).

180. Johnston, P. D. & Redfield, A. G. An NMR study of the exchange rates for protons involved in the secondary and tertiary structure of yeast tRNA^phe. *Nucleic Acids Research* **4**, 3599–3615 (1977).

181. Christodoulou, J. *et al.* Heteronuclear NMR investigations of dynamic regions of intact Escherichia coli ribosomes. *Proceedings of the National Academy of Sciences* **101**, 10949–10954 (2004).

182. Gelis, I. *et al.* Solid-state NMR enhanced by dynamic nuclear polarization as a novel tool for ribosome structural biology. *J Biomol NMR* **56**, 85–93 (2013).

183. Barbet-Massin, E., Huang, C.-T., Daebel, V., Hsu, S.-T. D. & Reif, B. Site-Specific Solid-State NMR Studies of "Trigger Factor" in Complex with the Large Ribosomal Subunit 50S. *Angew. Chem. Int. Ed.* **54**, 4367–4369 (2015).

184. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**, 51–55 (2008).

185. Shopsowitz, K. E., Roh, Y. H., Deng, Z. J., Morton, S. W. & Hammond, P. T. RNAi-Microsponges Form through Self-Assembly of the Organic and Inorganic Products of Transcription. *Small* **10**, 1623–1633 (2014).

186. Rupert, P.B. and Ferre-DAmare A.R. Crytallisation of tbhe hairpin ribozyme: illustrative protocols. *Methods in Molecular Biology* **252**, 303-311 (2004).

187. Cheong, C., Varani, G. and Tinoco, I. Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. *Nature* **346**, 680-682 (1990).

188. Tuerk, C. *et al.* CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes. *Proceedings of the National Academy of Sciences* **85**, 1364–1368 (1988).

189. Allain, F. H.-T. & Varani, G. Structure of the P1 Helix from Group I Self-splicing Introns. *Journal of Molecular Biology* **250**, 333–353 (1995).

190. Ennifar, E. *et al.* The crystal structure of UUCG tetraloop 1 1Edited by J Doudna. *Journal of Molecular Biology* **304**, 35–42 (2000).

191. Singh, V., Fedeles, B. I. & Essigmann, J. M. Role of tautomerism in RNA biochemistry. *RNA* **21**, 1–13 (2015).

192. Tolosa, S., Sansón, J. A. & Hidalgo, A. Mechanisms for guanine–cytosine tautomeric equilibrium in solution via steered molecular dynamic simulations. *Journal of Molecular Liquids* **251**, 308–316 (2018).

193. Hans-Heinriclhimbach, B. Y. J. Kinetic Study of Hydrogen Tunnelling in meso-Tetraphenylporphine by Nuclear Magnetic Resonance Lineshape Analysis and Selective Tl$_\rho$-Relaxation Time Measurements. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics.* **75**, 752-766 (1979).

194. Golovina, A. Y., Bogdanov, A. A., Dontsova, O. A. & Sergiev, P. V. Purification of 30S ribosomal subunit by streptavidin affinity chromatography. *Biochimie* **92**, 914–917 (2010).