Development of NMR spectroscopic methods for the characterization of RNA

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften

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

von

Robbin Schnieders

aus Baden/Schweiz

Frankfurt am Main

2020

(D30)

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

Dekan:	Prof. Dr. Clemens Glaubitz
Erster Gutachter:	Prof. Dr. Harald Schwalbe
Zweiter Gutachter:	Prof. Dr. Jens Wöhnert

Datum der Disputation:

Meiner Familie.

Table of Contents

Summary

This PhD thesis is dedicated to the extension of the portfolio of nuclear magnetic resonance (NMR) methods to characterize ribonucleic acids (RNAs). Only within the last few decades it has been realized that the cellular role of RNA goes well beyond the central dogma of molecular biology. In fact, RNA takes part in numerous cellular processes, executes numerous functions and acts either as a single player or in larger complexes, mostly RNA-protein complexes (RNPs) such as the ribosome or the spliceosome. This versatility in RNA function is coupled to a structural variety and the ability to adopt multiple long-lived and intricate conformations. Due to this high molecular complexity special demands are placed on the methods that are required for RNA structural characterization. With the ability to capture dynamics at atomic resolution and to measure under close to native conditions, NMR spectroscopy is undoubtedly a prime method for this purpose.

A general introduction to the current state of research, selected achievements as well as challenges in the field of NMR spectroscopy on RNA is given in **Chapter I**. This thesis is further composed of three independent chapters covering the three separate projects, which form the main body of work within the course of this thesis.

The imino group found in two of the four RNA nucleobases is generally considered to be the most powerful reporter group in the process of the NMR spectroscopic characterization of RNA. Its resonance assignment provides key information for a rapid determination of the RNA's secondary structure. This is possible, since the imino proton can only be detected, if it is protected from rapid solvent exchange through hydrogen bonding interactions or, in rare cases, steric shielding. Consequently, information on flexible regions of RNA that are not protected against solvent exchange cannot be derived using this NMR spy. It is a key finding of the thesis that nucleobase interactions can also be mapped through the amino groups, as they similarly take part in base pairing or RNA-ligand interactions. Notably, solvent exchange of the amino protons is always slower compared to the imino proton. Thus, ^{1}H,^{15}N resonances of the amino group can be detected even for dynamic regions of RNA. Moreover, focusing on characterizing amino groups of RNA nucleobases increases the number of available reporters as amino groups are present in three out of four RNA nucleobases.

However, there is a reason that up to work conducted in this thesis, amino groups have not been used for monitoring RNA nucleobases: the rate of the C-NH_2 bond rotation is most often close to the chemical shift differences of the two non-identical amino proton resonances, in particular for guanosines and adenosines, amino resonances regularly remain elusive in NMR spectra. Therefore, we developed experiments that excite double

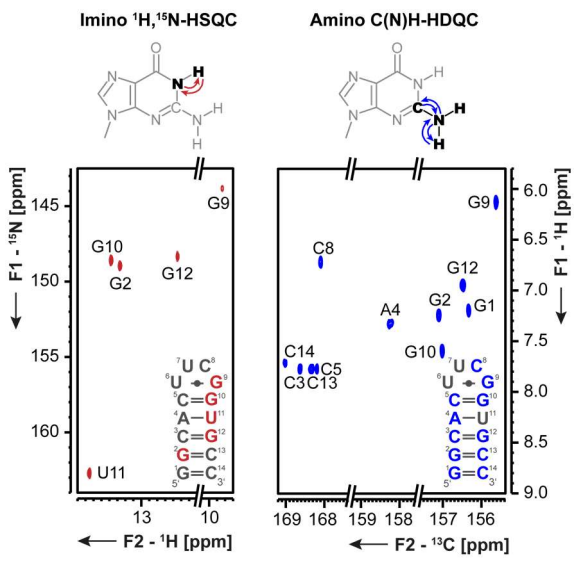


Figure I ¹H,¹⁵N-HSQC vs. C(N)H-HDQC spectra.

quantum (DQ) coherences of the two amino group protons and that further utilize ^{13}C-detection. Results on these experiments are discussed in **Chapter II** and show that the rotational exchange can be avoided by evolving double quantum instead of single quantum (SQ) coherence in the indirect dimension of a ^{13}C-detected C(N)H-HDQC experiment. The new experiment enables the detection of a full set of sharp amino resonances. The advantages of this experiment are immediately apparent when comparing the number of observable imino resonances in a classic ^{1}H,^{15}N-HSQC spectrum with the number of amino resonances obtained in the ^{13}C-detected C(N)H-HDQC spectrum of the same RNA (**Figure I**).

Furthermore, based on the newly available resonance assignment of amino groups, we developed a ^{13}C-detected "amino"-NOESY experiment to obtain precious additional structural restraints. The ^{13}C-detected "amino"-NOESY experiment enables the observation of NOE contacts that are not accessible using other ^{1}H-detected NOESY experiment. Among these new NOE contacts are valuable, inter-residual correlations, which are otherwise scarce in RNA due to the proton deficiency of its nucleobases. We showed that the newly obtained NOE contacts are especially important in the structure determination of RNAs with only few NOE restraints. Under such circumstances, the inclusion of the newly obtained amino NOE contacts lead to a significant improvement in the root-mean-square deviation (RMSD) of the three-dimensional structure of the 34 nts GTP class II aptamer (**Figure II**). Together the novel ^{13}C-detected NMR experiments developed within this PhD project provide a valuable alternative for the imino-based characterization of nucleobase interactions.

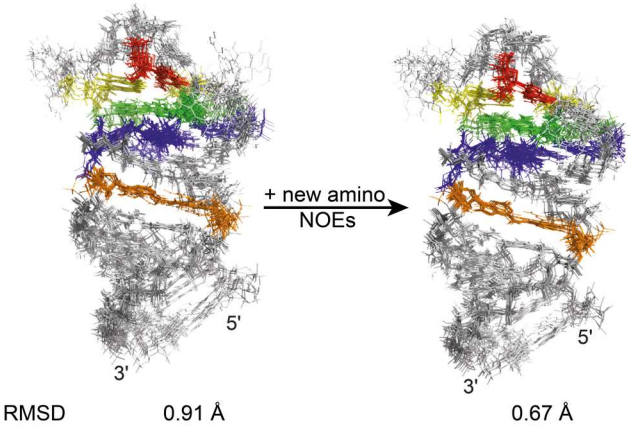


Figure II The impact of new amino NOE contacts.

While the ^{13}C-detected amino experiments address the problem of exchange processes, there are further challenges in NMR spectroscopy of RNA. With respect to other methods in structural biology such as X-ray crystallography or cryogenic electron microscopy

(cryo EM), the most severe drawback of NMR spectroscopy is most certainly the limitation in molecular size. This limitation results from increased resonance overlap as well as from unfavorable relaxation properties with increasing rotational correlation times, which causes signal broadening. As discussed in **Chapter III**, heteronuclear-detection schemes can potentially extend the size limitation towards the characterization of larger molecules as heteronuclei exhibit a slower transverse relaxation with respect to the proton nucleus. Among the heteronuclei, the ^{15}N-nucleus carries the most favorable relaxation properties. This is illustrated in **Figure III**, where predictions of the line width at half height in dependence of the rotational correlation time show that $^{1}H^{N}$-TROSY components increase more rapidly with molecular size than $^{15}N^{H}$-TROSY components. As a reduction in line width at half height is linked to a gain in signal intensity, these sharp lines might partly compensate for the sensitivity loss coming along with the low-gamma detection. Prompted by these predictions we performed different ^{15}N-detected H-N correlation experiments on RNA. To our knowledge these were the first reports on nitrogen direct-detected multi-dimensional NMR experiments for RNA. Among these experiments we identified the ^{15}N-detected BEST-TROSY experiment as the most sensitive ^{15}N-detected H-N correlation experiment. Therefore, this experiment was recorded along with its ^{1}H-detected counterpart for a set of different RNAs ranging from 5 kDa to 100 kDa in molecular size. By extracting ^{1}H- and ^{15}N-line width at half height, we were able to confirm the predicted trend that ^{1}H-line widths increase more rapidly with molecular size than ^{15}N-line widths (**Figure IV**). Moreover, we confirmed that the line width at half height is dependent on the type of base pairing: ^{15}N- and ^{1}H-lines in AU base pairs are generally sharper than the corresponding lines in GC base pairs. While we were able to observe an identical set of resonances for both, the ^{1}H- and the ^{15}N-detected BEST-TROSY experiments, the ^{1}H-detected experiment is more sensitive than the ^{15}N-detected experiment even for the largest RNA under study. However, this gap in sensitivity might be partially reduced when moving to even larger RNAs.

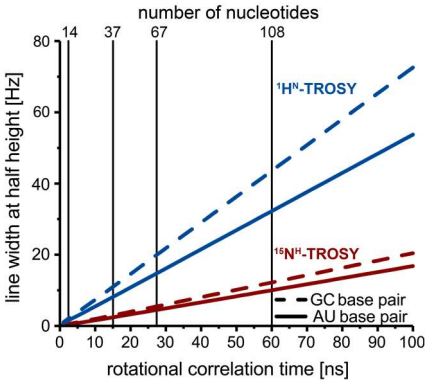


Figure III Predictions of $^{1}H^{N}$- and $^{15}N^{H}$-TROSY line widths.

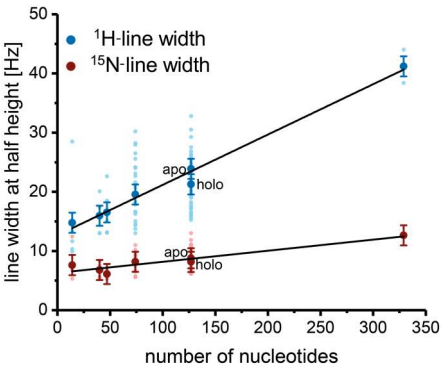


Figure IV Experimental $^{1}H^{N}$- and $^{15}N^{H}$-TROSY line widths.

Despite the continuous development of NMR methods problems remain, where NMR spectroscopy is reaching its limits. However, with the new advances in computational power, molecular dynamics (MD) simulations could potentially fill these gaps, which are beyond solely NMR experimental reach. To correctly describe the molecular properties of a system, well calibrated potential energy functions, so-called force fields, are needed.

This force field parametrization is still in development for the correct description of RNA and requires a variety of experimentally well-characterized RNAs to serve as reference systems. Experimental data derived by NMR spectroscopy is frequently used in the MD force field parametrization as the underlying and force field unbiased raw data are available. For this parametrization, tetraloops are the perfect model systems, as they combine a low molecular size with high structural diversity. The most frequently used structure for these purposes is an NMR structure of a 14 nts long hairpin RNA harboring a UUCG tetraloop. This three-dimensional structure was resolved in the Schwalbe lab and is based on a wealth of experimental data. However, there is the need for more force field parametrization reference systems, which are equally well characterized to describe the properties of the entire diverse class of RNA. As described in **Chapter IV**, we want to advance this force field development process with three new reference 14 nts RNAs exhibiting different tetraloop sequences. This PhD thesis focusses on one of these RNAs, a 14 nts RNA with a GAAG tetraloop. Building on initial work from my master's thesis, a complete resonance assignment was achieved for this RNA. Furthermore, the torsion angles β, γ, ε and χ have been experimentally determined through ^{3}J coupling constants and cross-correlated relaxation rates. A ^{1}H,^{13}C-HSQC temperature series revealed the presence of dynamics in the tetraloop, which are running into the intermediate exchange regime at lower temperatures. The temperature-dependent canonical coordinates show a transition from C3'- to C2'-endo ribose conformation at increased temperatures. Moreover, this is reflected in the ^{3}J(H,H) coupling constants of the tetraloop, which cannot be described by a pure C2'-endo conformation. Currently, we do not have a comprehensive description of the conformational dynamics of the GAAG tetraloop. We currently speculate that the GNRA type tetraloop is in a dynamic equilibrium with a second unknown state. Such conformational sampling is in line with the NOESY data as there are NOE contacts present that are typical for a GNRA structure, but also other NOE contacts that are not possible within this structure. Nevertheless, preliminary structure calculations were performed, yielding a structure bundle of an overall RMSD of 0.9 Å (**Figure V**). It is clear that further work beyond the scope of the PhD thesis, is required to determine and describe the conformational dynamics of this RNA.

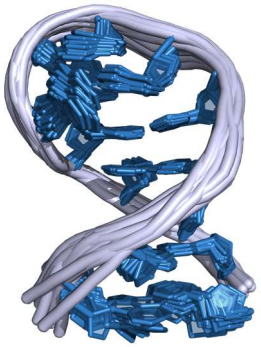


Figure V Structure bundle of the 14 nts RNA with GAAG tetraloop.

Zusammenfassung

Diese Doktorarbeit widmet sich der Erweiterung des Methodenportfolios der Kernspinresonanzspektroskopie (NMR Spektroskopie) zur Charakterisierung von Ribonukleinsäuren (RNAs). Erst in den letzten Jahrzehnten wurde erkannt, dass die zelluläre Rolle der RNA weit über das zentrale Dogma der Molekularbiologie hinausgeht. Tatsächlich sind RNAs an zahlreichen zellulären Prozessen beteiligt, erfüllen verschiedene Funktionen und agieren entweder als Einzelspieler oder in größeren Komplexen, insbesondere in RNA-Protein-Komplexen wie dem Ribosom oder dem Spleißosom. Diese Vielseitigkeit in der Funktion geht mit einer strukturellen Vielfalt und der Fähigkeit, mehrere langlebige und komplexe Konformationen annehmen zu können einher. Aufgrund dieser hohen molekularen Komplexität werden besondere Anforderungen an die Methoden gestellt, die zur strukturellen Charakterisierung von RNA verwendet werden. Mit der Fähigkeit, Moleküldynamiken auf atomarer Ebene zu erfassen und unter nahezu nativen Bedingungen zu messen, ist die NMR-Spektroskopie zweifellos eine sehr gut geeignete Methode für diesen Zweck.

Kapitel I beinhaltet eine allgemeine Einführung in den aktuellen Stand der Forschung, ausgewählte neue Ergebnisse sowie Herausforderungen auf dem Gebiet der NMR Spektroskopie an RNA. Abgesehen davon besteht die vorliegende Arbeit aus drei unabhängigen Kapiteln, die die drei separaten Projekte abdecken, die hauptsächlich im Rahmen dieser Doktorarbeit behandelt wurden.

Die wichtigste funktionelle Gruppe im Prozess der NMR-spektroskopischen Charakterisierung von RNA ist die Iminogruppe, die den experimentellen Goldstandard für eine schnelle Aufklärung der RNA Sekundärstruktur darstellt. Dies ist möglich, da das Iminoproton nur dann ein Signal hervorruft, wenn es durch Wasserstoffbrückenbindungen oder in seltenen Fällen durch sterische Abschirmung vor schnellem Lösungsmittelaustausch geschützt ist. Informationen über flexible Regionen der RNA sind mit diesem Goldstandard der NMR Spektroskopie jedoch nicht zugänglich. Stattdessen könnten Nukleobaseninteraktionen auch durch die Aminogruppen charakterisiert werden, da diese ebenfalls an Basenpaarungen oder RNA-Ligand-Interaktionen beteiligt sind. Insbesondere ist das Aminoproton unempfindlich gegenüber Lösungsmittelaustausch, was die Charakterisierung dynamischer Regionen der RNA ermöglichen würde. Außerdem erhöht sich dadurch die Anzahl der verfügbaren Reporter, da Aminogruppen in drei von vier Nukleobasen vorhanden sind, während Iminogruppen nur in Guanosinen und Uridinen vorkommen.

Es gibt jedoch einen Grund dafür, dass bis zu den Arbeiten in dieser Doktorarbeit, Aminogruppen nicht für die Charakterisierung von RNA-Nukleobasen verwendet

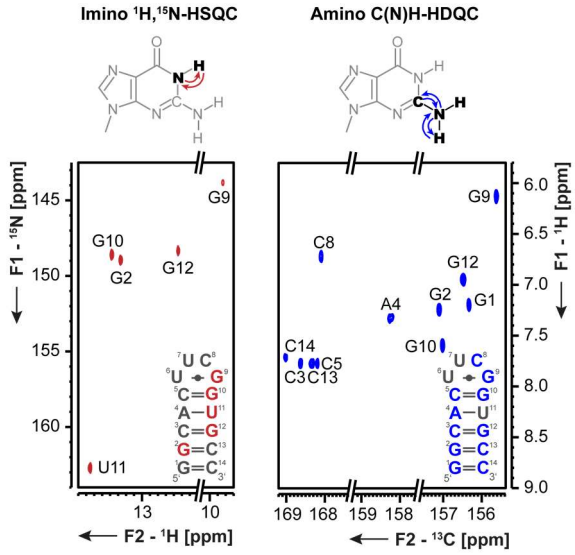


Abbildung I ^{1}H,^{15}N-HSQC vs. C(N)H-HDQC Spektren.

wurden: Da die Rotation der C-NH_2-Bindung häufig und besonders in Guanosinen und Adenosinen im intermediären Austauschregime auf der NMR-Zeitskala rotiert, sind Aminoresonanzen in NMR-Spektren meist schwer detektierbar. Um diesen Austausch zu umgehen, haben wir NMR-Experimente entwickelt, die Doppelquantenkohärenz (DQ) der beiden Aminoprotonen evolvieren sowie auf ^{13}C-Detektion basieren. Die Ergebnisse dieser Experimente werden in **Kapitel II** diskutiert und zeigen, dass die Effekte des Rotationsaustausches durch die Entwicklung von Doppelquantenkohärenz (DQ) anstelle von Einzelquantenkohärenz (SQ) in der indirekten Dimension eines ^{13}C-detektierten C(N)H-HDQC-Experiments, vermieden werden kann. Dies ermöglicht die Detektion eines vollständigen Satzes scharfer Aminoresonanzen, was mit keiner anderen ^{1}H-detektierten Methode möglich ist. Die Vorteile dieses Experiments sind sofort ersichtlich, wenn man die Anzahl der beobachtbaren Iminoresonanzen in einem klassischen ^{1}H,^{15}N-HSQC-Spektrum mit der Anzahl der Aminoresonanzen vergleicht, die mit dem ^{13}C-detektierten C(N)H-HDQC-Spektrum derselben RNA detektiert werden können (**Abbildung I**).

Darüber hinaus werden die Aminogruppen mit dem hier entwickelten ^{13}C-detektierten "Amino"-NOESY-Experiment in einen direkten strukturellen Zusammenhang gebracht. Dieses Experiment ermöglicht die Erfassung von NOE-Kontakten, die mit anderen ^{1}H-detektierten NOESY-Experiment nicht zugänglich sind. Unter diesen neuen NOE-Kontakten befinden sich wertvolle, interresiduale Korrelationen, die sonst in der RNA aufgrund des Protonenmangels ihrer Nukleobasen selten sind. Wir zeigen weiterhin, dass die neu gewonnenen NOE-Kontakte besonders wichtig für die Strukturbestimmung von RNAs sind, die einen Mangel an NOE-Kontakten aufweisen. Hier führte eine Einbeziehung der neu gewonnenen Amino-NOE-Kontakte zu einer signifikanten Verbesserung der root-mean-square deviation (RMSD) der dreidimensionalen Struktur des 34 Nukleotide (nts) langen GTP Klasse II Aptamers (**Abbildung II**). Zusammen stellen die neuartigen ^{13}C-detektierten NMR-Experimente, die im Rahmen

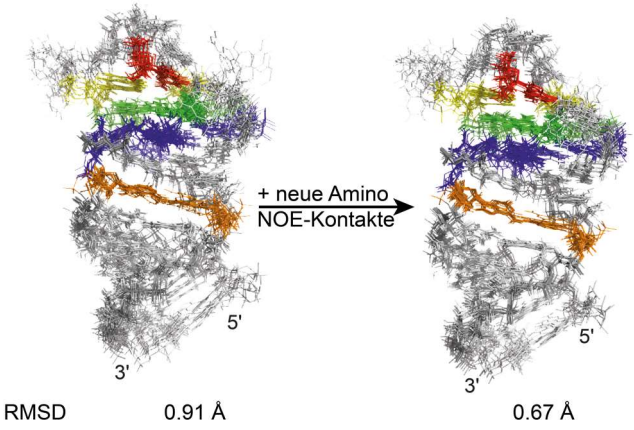


Abbildung II Einfluss der neuen Amino-NOE-Kontakte.

dieser Doktorarbeit entwickelt wurden, eine wertvolle Alternative für die iminobasierte Charakterisierung von Nukleobasen-Interaktionen dar.

Während die ^{13}C-detektierten Aminoexperimente das Problem der Austauschprozesse in RNA behandeln, gibt es weitere Herausforderungen in der NMR-Spektroskopie an RNA. Im Vergleich zu anderen Methoden der Strukturbiologie, wie der Röntgenkristallographie oder der Kryoelektronenmikroskopie (cryo-EM), ist die stärkste Einschränkung der NMR-Spektroskopie sicherlich die Begrenzung in der Größe der zu untersuchenden Moleküle. Diese Einschränkung ergibt sich aus dem zunehmenden Resonanzüberlapp sowie aus ungünstigen Relaxationseigenschaften mit steigender Rotationskorrelationszeit, die zu Signalverbreiterungen führt. Wie in **Kapitel III** beschrieben, können Heterokern-detektierte Experimente diese Größenbeschränkung potenziell auf die Charakterisierung größerer Moleküle ausdehnen, da Heterokerne eine langsamere transversale Relaxation im Vergleich mit Protonenkernen aufweisen. Unter den Heterokernen besitzt der ^{15}N-Kern die günstigsten Relaxationseigenschaften. Dies ist in der Vorhersage der Halbwertsbreite in Abhängigkeit von der Rotationskorrelationszeit in **Abbildung III** veranschaulicht, aus der hervorgeht, dass $^{1}H^{N}$-TROSY-Komponenten schneller mit der Molekülgröße zunehmen als $^{15}N^{H}$-TROSY-Komponenten. Da eine Verringerung der Halbwertsbreite mit einer Zunahme der Signalintensität verbunden ist, könnten diese scharfen Linien den Empfindlichkeitsverlust, der mit der Detektion von Kernen mit niedrigem gyromagnetischen Verhältnis einhergeht, teilweise kompensieren. Angeregt durch diese Vorhersagen führten wir verschiedene ^{15}N-detektierte H-N-Korrelationsexperimente an RNA durch. Nach unserem Wissen waren dies die ersten Berichte über ^{15}N direkt detektierte mehrdimensionale NMR-Experimente für RNA. Unter diesen Experimenten identifizierten wir das ^{15}N-detektierte BEST-TROSY-Experiment als das empfindlichste ^{15}N-detektierte H-N-Korrelationsexperiment. Daher wurde dieses Experiment zusammen mit seinem ^{1}H-detektierten Gegenstück für einen Satz verschiedener RNAs mit Molekülgrößen von 5 kDa bis 100 kDa durchgeführt. Durch die Messung der ^{1}H- und ^{15}N-Halbwertsbreiten konnten wir den vorhergesagten Trend bestätigen, dass ^{1}H-Linienbreiten schneller mit der Molekülgröße zunehmen, als ^{15}N-Linienbreiten (**Abbildung IV**). Darüber hinaus bestätigten wir, dass die Linienbreite auf halber Höhe von der Art der

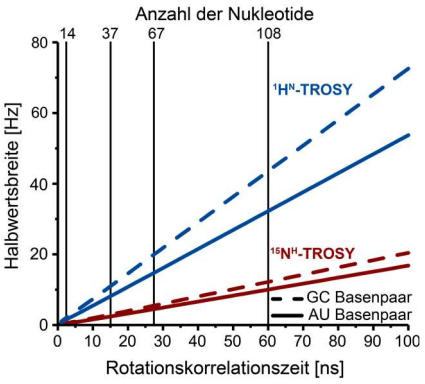


Abbildung III Berechnung der $^{1}H^{N}$- und $^{15}N^{H}$-TROSY Linienbreiten.

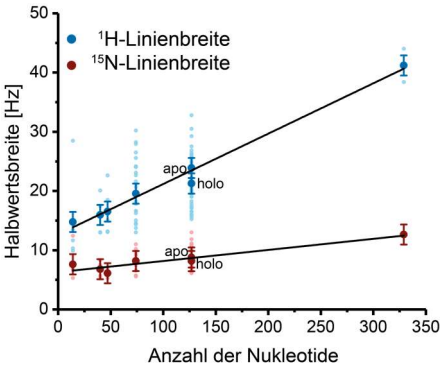


Abbildung IV Experimentell ermittelte $^{1}H^{N}$- und $^{15}N^{H}$-TROSY Linienbreiten.

Basenpaarung abhängt, was darin resultiert, dass ^{15}N- und ^{1}H-Linien in AU-Basenpaaren im Allgemeinen schärfer sind als die entsprechenden Linien in GC-Basenpaaren. Während wir für beide, das ^{1}H- und das ^{15}N-detektierte BEST-TROSY-Experiment, einen identischen Satz von Signalen beobachten konnten, war das ^{1}H-detektierte Experiment sensitiver als das ^{15}N-detektierte Experiment; selbst für die größte untersuchte RNA. Dieser Empfindlichkeitsunterschied könnte sich jedoch teilweise verringern, wenn man zur Charakterisierung noch größerer RNAs übergeht.

Trotz der kontinuierlichen Weiterentwicklung von NMR-Methoden gibt es immer noch Probleme, bei denen die NMR-Spektroskopie an ihre Grenzen stößt. Mit den stetigen Fortschritten in der Rechenleistung von Computern könnten Molekulardynamik-Simulationen (MD-Simulationen) diese Lücken, die über die alleinige experimentelle Reichweite der NMR hinausgehen, füllen. Um die molekularen Eigenschaften eines Systems korrekt zu beschreiben, werden gut kalibrierte potentielle Energiefunktionen, so genannte Kraftfelder, benötigt. Diese Kraftfeldparametrisierung befindet sich für die korrekte Beschreibung von RNA noch in der Entwicklung und erfordert eine Vielzahl von experimentell gut charakterisierten RNAs, die als Referenzsysteme dienen. Bei der MD-Kraftfeldparametrisierung werden häufig experimentelle Daten aus der NMR-Spektroskopie verwendet, da diese Rohdaten nicht durch ein Kraftfeld beeinflusst sind. Hier sind Tetraloops die perfekten Modellsysteme, da sie eine niedrige Molekülgröße mit einer hohen Strukturvielfalt kombinieren. Die für diese Zwecke am häufigsten verwendete Struktur ist eine NMR-Struktur einer 14 nts langen Haarnadel-RNA, die einen UUCG-Tetraloop beherbergt. Diese dreidimensionale Struktur wurde in der Schwalbe-Gruppe aufgelöst und basiert auf einer Fülle von experimentellen Daten. Es besteht jedoch der Bedarf an zusätzlichen Referenzsystemen für die MD-Kraftfeldparametrisierung, die ebenso gut charakterisiert sind, um die Eigenschaften der gesamten vielfältigen Klasse von RNA korrekt zu beschreiben. Wie in **Kapitel IV** beschrieben, wollen wir diesen Kraftfeldentwicklungsprozess mit drei neuen 14 nts-Referenz-RNAs, die unterschiedliche Tetraloop-Sequenzen aufweisen, unterstützen. Dieses Teilprojekt der vorliegenden Doktorarbeit konzentriert sich auf eine dieser RNAs, eine 14 nts RNA mit einem GAAG-Tetraloop. Aufbauend auf ersten Arbeiten aus meiner Masterarbeit wurde für diese RNA eine vollständige Resonanzzuordnung erreicht. Darüber hinaus wurden die Torsionswinkel β, γ, ε und χ experimentell durch ^{3}J-Kopplungskonstanten und kreuzkorrelierte Relaxationsraten bestimmt. Weiterhin zeigte eine ^{1}H,^{13}C-HSQC-Temperaturreihe das Vorhandensein von Dynamik im Bereich des Tetraloops, die bei niedrigeren Temperaturen in das intermediäre Austauschregime übergeht. Die temperaturabhängigen kanonischen Koordinaten zeigten bei höheren Temperaturen einen Übergang von C3'- zur C2'-endo-Ribose-Konformation. Außerdem spiegelt sich dies in den ^{3}J(H,H)-Kopplungskonstanten der Tetraloops wider, die nicht durch eine reine C2'-endo-Konformation beschrieben werden können. Gegenwärtig

haben wir keine vollständige Beschreibung der Konformationsdynamik des GAAG-Tetraloops. Wir spekulieren derzeit, dass sich dieser Tetraloop vom GNRA-Typ in einem dynamischen Gleichgewicht mit einem zweiten unbekannten Zustand befindet. Dieses konformationelle Gleichgewicht wird auch durch ^{1}H,^{1}H-NOESY-Daten untermauert, da sowohl NOE-Kontakte beobachtet werden, die typisch für eine GNRA-Struktur sind, als auch andere NOE-Kontakte, die innerhalb dieser Struktur nicht möglich sind. Dennoch wurden vorläufige Strukturberechnungen durchgeführt, die in einem Strukturbündel mit einem RMSD von 0.9 Å resultierten (**Abbildung V**). Diese vorläufige Struktur muss zukünftlich durch weitere Arbeiten, die nicht in den Rahmen dieser Arbeit fallen, finalisiert werden. Hierzu sollte die Strukturberechnung der RNA unter Berücksichtigung von mindestens zwei verschiedenen Zuständen durchgeführt werden.

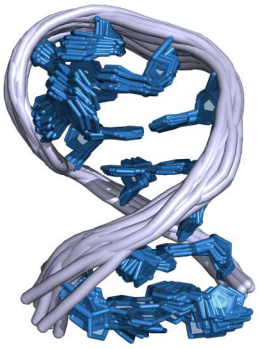


Abbildung V Strukturbündel der 14 nts RNA mit GAAG Tetraloop.

List of Abbreviations

2'-dG	2'-deoxyguanosine
ARIA	Ambiguous restraint for iterative assignment
BEST	Band-selective exciatation short-transient
BMRB	Biological Magnetic Resonance Bank
circRNA	Circular RNA
COSY	Correlation spectroscopy
cryo	Cryogenic
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbation
DD	Dipole-dipole
ddH_2O	Deionized and destilled water
DEAE	Diethylethanolamine
DEPC	Diethyl pyrocarbonate
DFHBI	(Z)-4-(3,5-Difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one
DNA	Deoxyribonucleic acid
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
FID	Free induction decay
fwdi	Forward-directed
Γ	Cross-correlated relaxation rate
GARP	Globally optimized alternating phase rectangular pulse
HDQC	Heteronuclear double quantum coherence
HDV	Hepatitis delta virus
HH	Hammerhead
HPLC	High pressure liquid chromatography
HSQC	Heteronuclear single quantum coherence
HZQC	Heteronuclear zero quantum coherence
IPAP	Inphase/antiphase
kb	Kilo base pairs
LB	Lysogeny Broth
lncRNA	Long non-coding RNA
MD	Molecular dynamics
miRNA	Micro RNA
MWCO	Molecular weight cut off
NMR	Nuclear magnetic resonance

NOESY	Nuclear overhauser enhancement spectroscopy
nts	Nucleotides
NUS	Non-uniform sampling
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEP	Preservation of equivalent pathways
RDC	Residual dipolar coupling
rel	relayed
RMSD	Root mean square deviation
RNA	Ribonucleic acid
rpm	Rounds per minute
S/N	Signal-to-noise ratio
SAH	S-adenosylhomocystein
SAM	S-adenosylmethionin
siRNA	Small interfering RNA
ST2-PT	Single transition-to-single transition polarization transfer
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
TALOS	Torsion angle likelihood obtained from chemical shift and sequence similarity
TBE	Tris borate EDTA
tmRNA	Transfer-messenger RNA
TOCSY	Total correlation spectroscopy
Tris	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimized spectroscopy
TSP	Trimethylsilylpropanoic acid-d_4

Chapter I: General Introduction

As opposed to earlier assumptions, it is nowadays widely known that the function of ribonucleic acid (RNA) goes well beyond encoding the amino acid sequence of proteins. In fact, in eukaryotes only around 2% of the transcribed RNA are translated into a certain protein. The remaining 98% are non-coding RNAs (ncRNAs). These ncRNAs conduct a large variety of functions in the cell - many of which are unknown until today. Over the last decade, a plethora of new functional RNAs have been discovered, both in prokaryotes, such as riboswitches[1], RNA thermometers[2] or transfer-messenger RNAs (tmRNAs)[3], and in eukaryotes, including small interfering RNAs (siRNAs)[4], microRNAs (miRNAs)[5], long non-coding RNAs (lncRNAs)[6], circular RNAs (circRNAs)[7] and others[8]. These RNAs can adopt intricate folded structures often required to exert their function. In contrast to proteins, where the biological active state is predominantly characterized by a single conformation, functional RNAs often adopt multiple long-lived conformations and interconversion between these states has been shown to be functionally relevant[9–13]. Moreover, the populations of the different states are often strongly dependent on external factors such as the temperature or the Mg^{2+} ion concentration[14]. This high complexity and variety in structure and function comes along with a rather dynamic character, which overall challenges the determination of a three-dimensional structure.

Generally, structures of RNAs can be determined using X-ray crystallography, NMR spectroscopy or cryogenic electron microscopy (cryo-EM). From 1976 on, the number of structures of biomolecules deposited in the PDB data bank[15] is ever increasing (Figure 1 A).

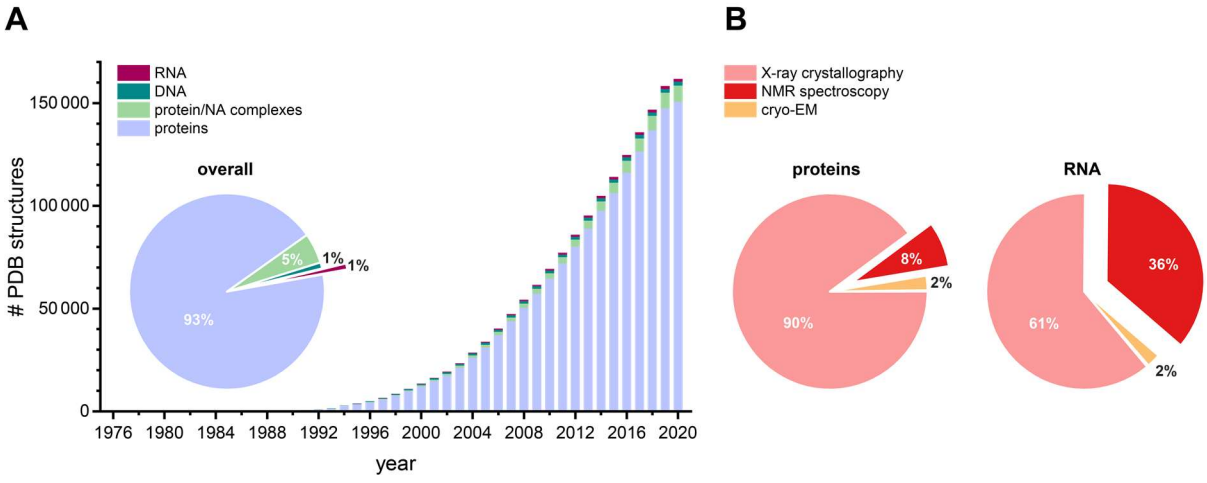


Figure 1 A Total number of biomolecular structures submitted to the PDB data bank (accessed 04/2020)[15] from 1976 to 2020. This includes three dimensional structures of proteins (light blue), protein/nucleic acid (NA) complexes (light green), DNA (dark green) and RNA (ruby colored). The circular chart shows the fractions of the respective biomolecules in the total number of structures. **B** Percentage of protein (left) and RNA (right) structures solved by X-ray crystallography (light red), NMR spectroscopy (red) and cryo-EM (orange).

The majority of the PDB entries represent protein structures (93%), while RNA structures account for only 1% (Figure 1 A, circular chart). Only ~8% of the protein structures have been solved by NMR spectroscopy. With 90% of all protein structures being solved by X-ray crystallography, this technique certainly dominates the field of protein structure determination (Figure 1 B, left). This holds not true to the same extent for structure determination of RNA, where 36% of all RNA structures deposited in the PDB data bank are based on NMR data (Figure 1 B, right), reflecting the particular importance of NMR spectroscopy in this field.

The observed trend towards NMR spectroscopy in RNA structure determination mainly results from the dynamic character of RNA. These dynamics cannot be captured in the solid state and furthermore result in a reluctance of RNAs to crystallize. As a consequence, NMR spectroscopy plays a significant role in the structure determination of RNA-protein-complexes[16–19] and of isolated RNAs[20–22] as will be addressed in the following paragraphs.

1.1 NMR spectroscopy on RNA – from major achievements to current challenges

If we were to look back over the past few years and were to report on the major contributions of NMR spectroscopy on the characterization of RNA, the RNA structure of a 155 nts long HIV-1 RNA packaging signal is probably the most prominent NMR solution structure, which has been determined using deuterium-edited NMR spectroscopy[23]. Along the line of RNA structure, a further important contribution has been made by Bailor *et al.*, who showed that the three-dimensional orientation of helices to each other in e.g. two-way junctions can be calculated from the RNA secondary structure through topological constraints[24]. Besides those findings on the overall structural level, important findings were also made on the molecular level with the observation of transient nucleobase mismatches involving tautomeric and anionic states[25]. These transient states are believed to mediate replication and translation errors. They are visualized using relaxation dispersion experiments in DNA and RNA by Kimsey *et al.*[25]. However, not only those short-lived states are known to have cellular functions. In fact, long-lived intermediates play a key role in the regulation of gene expression through riboswitches. This regulation occurs on the transcriptional as well as on the translational level and has been extensively studied using NMR spectroscopy. For transcriptional riboswitches, major contributions were the investigation of transcriptional intermediates revealing the important role of metastable states for RNA regulation[26] or the characterization of pause sites and ligand-dependent co-transcriptional folding, which is inevitable for a ligand response to the full-length guanine-sensing riboswitch[27]. Furthermore, the transcriptional metastable states have been studied providing insights on the kinetic demands for functionality through the determination of the life times of

these metastable states[28]. On the level of translation, the discovery of a switching mechanism involving three instead of two distinct conformations in the adenine-sensing riboswitch represents a key finding, which was discovered using NMR spectroscopy[14]. For this riboswitch, two ligand-free states are found, which are linked in a temperature-dependent pre-equilibrium out of which only one structure is capable of Mg^{2+}-dependent ligand binding[14].

When looking at all these studies of interesting systems of high biological relevance one should bear in mind that these findings are often not possible without the continuous development and improvement of NMR instrumentation, techniques and methods. Therefore, the following paragraph will highlight a couple of selected methodological achievements. In terms of instrumentation, the era of modern high resolution NMR spectroscopy started with the development of cryogenically cooled probes typically enhancing the detection sensitivity by a factor of 3 to 4[29,30]. This development allows recording of NMR experiments even at lower sample concentration within a reasonable amount of time, which is especially important in the field of biomolecular NMR. Non-uniform sampling (NUS) is another general technique that enables in particular high-dimensional (4D-7D) NMR experiments due to the massive reduction of measurement time. While NUS or predecessor techniques already emerged in 1987[31], the broad application of it started decades later with software and processing programs[32,33], which made NUS more broadly accessible (for overview see recent review article[34]). Using this technique many more-dimensional (3D, 4D, 5D) experiments become possible in the first place, such as 4D-NOESY experiments for the sequential assignment of long RNAs[35]. Another important development in the field of NMR spectroscopy on RNA has been the combination of NMR spectroscopy with laser pulses allowing light-induced real-time NMR studies of RNA folding[36,37]. In the field of pulse sequence development, the N-H coherence transfer module TROSY[38] has been a milestone in NMR spectroscopy of RNA. Here, it is used beyond single bond correlations in many important more dimensional NMR experiments for RNA, like the TROSY relayed HCCH-COSY experiment for adenines[39] and the TROSY-HCN experiment for the correlation of glycosidic and aromatic protons[40]. Another important NMR experiment for the characterization of hydrogen bonding in RNA represents the HNN-COSY experiment, which enables the detection of canonical[41] and non-canonical[42] base pairing.

Having demonstrated the power of NMR spectroscopy for the characterization of RNA, there still remain certain obstacles, which have to be overcome in future. In particular, high molecular weights and conformational interconversions of RNAs make their NMR spectroscopic investigations challenging. With increasing size of the RNA under study, the number of resonances as well as their peak width at half height increases leading to resonance overlap. Moreover, exchange of multiple conformational sub-states broadens peaks, particularly in the presence of divalent ions such as Mg^{2+} ions[43], which are, however, often inevitable for a correct RNA fold. All of the mentioned effects reduce the

maximum achievable resolution and in combination with an intrinsically low diversity in RNA building blocks (A, G, C and U, Figure 2 A) and the generally narrow ^{1}H-chemical shift dispersion, this already hampers the structural characterization of moderately sized RNAs. This is for example reflected in the H1'-to-H6/H8-protons region of a 2D ^{1}H,^{1}H-NOESY spectrum of a 41 nts long RNA, where signal overlap substantially impedes a complete data interpretation (Figure 2 B). This results in a molecular size limit for the structural characterization of an RNA at approximately 50 nts using ^{13}C-^{15}N labeling schemes, which is indispensable for a sophisticated structural study of a higher molecular weight RNAs[9].

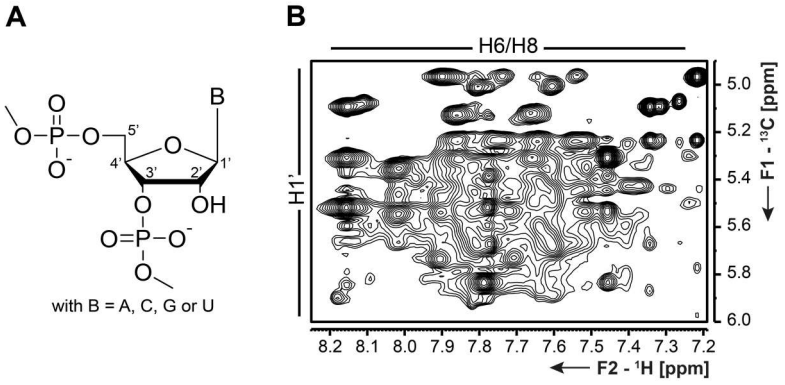


Figure 2 A Chemical structure of the ribose-phosphate backbone of an RNA, while "B" represents the respective nucleobase (A, C, G or U). **B** 2D ^{1}H,^{1}H-NOESY spectrum of a 41 nts terminator stem of the 2'-dG sensing riboswitch from *Mesoplasma florum*[44]. The figure has been adapted from literature[45].

Apart from the resolution, another key property for any kind of spectroscopy represents the signal-to-noise ratio (S/N ratio), which is decreased with increasing signal line width. Moreover, it is affected by the RNA concentration, which is often limited to the μM to low mM regime due to high costs in sample preparation, low yields, oligomerization preferences or solubility issues (particularly for high molecular weight RNAs). For the investigation of RNA-protein complexes this maximum concentration is usually even further reduced to the low μM range. Thus, it is evident that the experimental conditions should be optimized to yield both, a high resolution and a high S/N ratio in order to enable the characterization of RNAs beyond 50 nts.

One possible solution to achieve this are segmental or site-selective labeling strategies[46] as well as the incorporation of modified nucleotides as e.g. reporters. The most common ways to realize this are through T7 RNA polymerase mutants, which are able to incorporate a variety of modified rNTPs[47] or through solid phase synthesis, where a broad variety of modifications can be incorporated[48]. While the enzymatic method is independent in RNA length but not specific, the latter method allows a specific incorporation of nucleotides, however, at the expense of being limited in RNA length (~50 nts)[49]. For a combination of both desired features, ligation-based and chemo-enzymatic approaches are often used. 5'-ends can be for example modified post-transcriptionally using methyl transferases [50–52]. The post-transcriptional modification

of 3'-ends can be achieved through nucleotidyl transferases, adding nucleoside triphosphates[53–55], or T4 RNA ligase 1, adding 3',5'-bisphosphates [56,57]. More complex ligation based methods to segmentally label RNAs have been demonstrated by Tzakos *et al.*[22] and Duss *et al.*[58]. Both methods elegantly exploit ribozyme cleavage to obtain the required phosphorylation states of the RNA fragments, which are to be ligated. The approach from Tzakos *et al.* employs a template DNA including the desired 3'-fragment (donor), a hammerhead ribozyme type III (HH III), a hammerhead ribozyme type I (HH I) and the desired 5'-fragment (acceptor). By performing two separate *in vitro* transcriptions using differently labeled rNTPs and priming with GMP, donor and acceptor RNAs are obtained (enzymatic pathway, Figure 3 A). A subsequent ligation using *T4* DNA ligase yields two segmentally labeled RNAs with complementary labeling scheme.

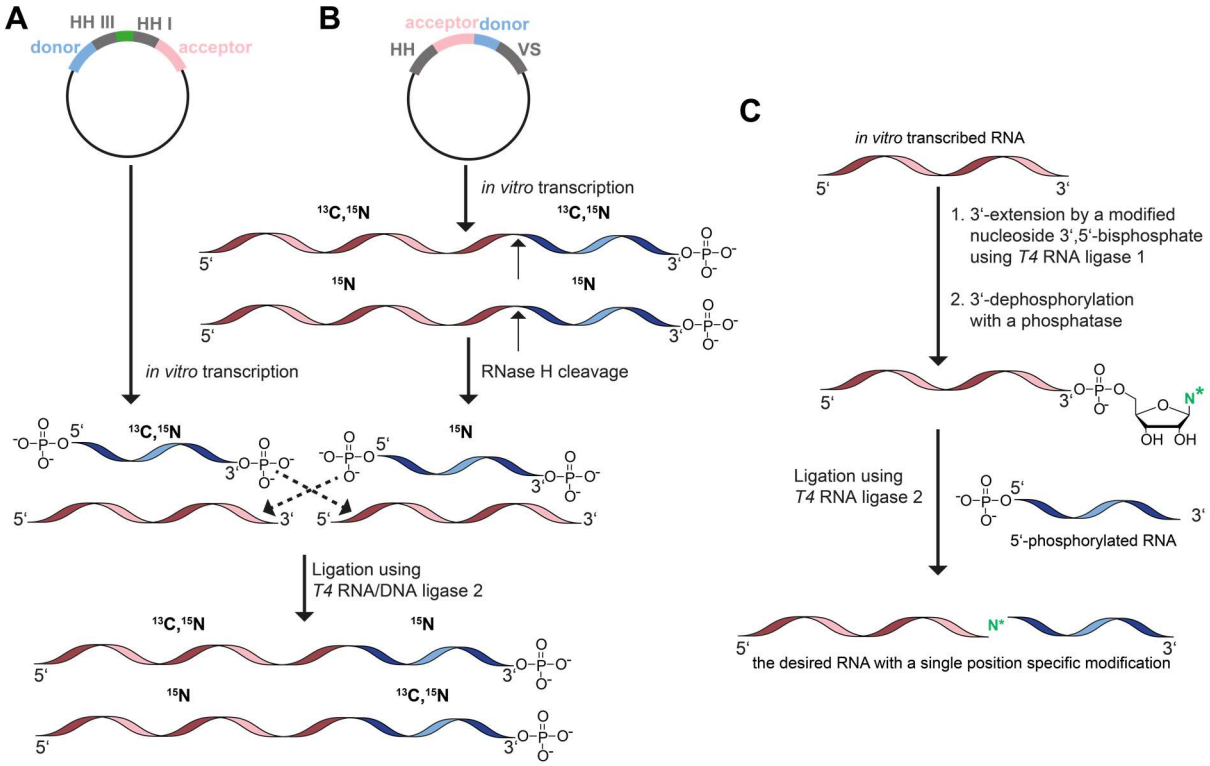


Figure 3 A Enzymatic scheme of the segmental labeling method by Tzakos *et al.*[22] and **B** by Duss *et al.*[58]. **C** Reaction overview of the chemo-enzymatic incorporation of a modified nucleotide N* (green) at a single position of an RNA in three steps. The figure has been adopted from Schnieders *et al.*[45]. The donor RNAs are labeled in blue and the acceptor RNAs are held in red.

The approach by Duss *et al.*[58] requires a DNA template encoding for a HH ribozyme, the desired full-length RNA sequence and a minimal recognition site for the VS ribozyme. After *in vitro* transcription and ribozyme cleavage the target RNA is further cleaved by RNase H at a desired position to yield the donor and acceptor RNAs (enzymatic pathway, Figure 3 B). If this process is run in parallel with different labeling strategies, a cross ligation with T4 RNA or DNA ligase yields two RNAs with complementary labeling schemes. While this method exhibits an additional enzymatic step as opposed to the approach by Tzakos *et al.*[22], it offers the possibility for the inclusion of multiple segments and is completely sequence independent. Both methods yield isotope-labeled segments,

which are limited to modifications accepted by the T7 RNA polymerase and which can in case of large RNAs still lead to resonance overlap.

However, if a site-selectively labeled RNA with a single incorporated labeled or modified nucleotide is desired, a different enzymatic pathway can be employed as was described by Keyhani *et al*.[59]. For this method two independently *in vitro* transcribed RNAs can be used, where the donor RNA has to be 5'-phosphorylated. The acceptor RNA is extended with a nucleoside 3'-5'-bisphosphate using T4 RNA ligase 1 (enzymatic pathway, Figure 3 C). This nucleoside 3'-5'-bisphosphate may be isotope-labeled or modified. After 3'-dephosphorylation acceptor and donor RNA are ligated with *T4* RNA ligase 2 yielding the final RNA with a single incorporated nucleotide. With respect to the two segmental labeling strategies overall more steps are required to finally yield the target RNA as e.g. the nucleoside 3'-5'-bisphosphate or the donor and acceptor RNAs have to be synthesized (not shown). However, using this approach, a broad variety of different modifications can be incorporated, which do not necessarily have to be accepted by the T7 RNA polymerase.

Especially the introduction of ^{19}F-nuclei as spy signals has recently gained in attention. The advantage here is that ^{19}F-nuclei are not naturally occurring in RNA and that the RNA structure stays unperturbed, with the exception of the incorporation of 2'-deoxy-2'-fluoro modified nucleotides, favoring the C3'-endo ribose conformation[60]. Thus, ^{19}F resonances are most commonly used as a reporter for structural changes, which can be efficiently tracked due to distinct chemical shift perturbations[61–63]. Despite these tempting properties, NMR studies using ^{19}F labeling in RNA were not extensively pursued for long. The reason for this is the large chemical shift anisotropy (CSA) of ^{19}F-nuclei, which leads to severe line-broadening for larger RNAs. However, with the continuous development of the synthesis of fluorinated and ^{13}C-labeled RNA building blocks and the introduction of a ^{13}C,^{19}F-TROSY-experiment this disadvantage was avoided[64].

All in all, with site-selective labeling methods, with few exceptions[65], the size limitation of NMR spectroscopy on RNA is pushed towards 150 - 200 nts[46]. Even though these methods provide a valuable route towards the characterization of larger RNAs, the time and costs associated with sample preparation are often multiplied, while at the same time RNA yields are modest. Furthermore, the desired target size of biologically relevant RNAs, e.g. long non-coding RNAs (lncRNAs), often exceeds 250 nts, requiring further development in biomolecular NMR spectroscopy.

Apart from the size limitation, which is primarily caused by resonance overlap due to the low ^{1}H chemical shift dispersion and unfavorable ^{1}H-relaxation properties, a further challenge of NMR spectroscopy represents the small number of ^{1}H in the nucleobases of RNA. This results in a usually low number of inter-residual long-range NOE contacts. On top of that, the few protons, which are present in the RNA nucleobases, are often involved in exchange processes including tautomerization, further reducing the number of reporter signals for the orientation of the nucleobase. This is particularly distinct in

regions of enhanced flexibility, where the exchange of the most prominent reporter, the imino proton, is broadened beyond detectability. However, since these flexible regions in particular are often involved in functional processes such as ligand or protein recognition[14,66,67], their characterization is essential.

1.2 Towards ^{13}C-detection

All mentioned challenges – resonance overlap, relaxation, low ^{1}H-density and chemical exchange – can be (partly) overcome by moving away from ^{1}H-direct detection, the experimental gold standard, towards heteronuclear detection; in particular ^{13}C-detection. One advantage in ^{13}C-detection is a typically higher chemical shift dispersion, spanning approximately 105 ppm in RNA. This higher dispersion in combination with a reduced peak width at half height due to more favorable relaxation properties, reduces the problem of signal overlap with ^{13}C- as opposed to ^{1}H-detected NMR experiments. This high possible resolution can be obtained by long sampling of the free induction decay (FID) in the direct dimension and thus without severely extending the required experimental time. These properties are especially valuable in the study of high molecular weight RNAs, bearing the potential to push the limit in molecular size further. Moreover, while the density of ^{1}H-nulei in the RNA nucleobases is low, they are composed of connected C-N fragments, resulting in a high carbon (and nitrogen) content. Furthermore, carbon nuclei are independent of solvent exchange rendering the characterization of flexible regions of RNA possible. This also features a higher tolerance towards changing experimental conditions as for example salt concentration, pH or temperature. Lastly, and even though being not as prominent in RNA when compared to proteins, nucleotides could also be fully deuterated without losing the reporter nucleus but at the same time exploiting the favorable relaxation properties[68].

However, the main disadvantage of ^{13}C-detection is the decreased sensitivity according to the lower gyromagnetic ratio. This is complemented with the necessity to decouple homonuclear ^{1}J(C,C) scalar couplings in case of uniform ^{13}C-labeling. While this decoupling involves a little more effort in terms of pulse programming, it is possible without sensitivity losses using inphase/ antiphase (IPAP)[69,70] and, even better, using S3E schemes[71] as well as selective decoupling during acquisition. As the potential value through additional information outcompetes the mentioned disadvantages, ^{13}C-direct detected experiments prove valuable alternatives and thus a few selected studies will be presented in the following paragraphs.

The problem of resonance overlap is particularly prominent within the ribose moiety, where all ^{1}H-resonances from the 2', 3', 4' and 5' positions are coming to resonance within 1-2 ppm resulting in an underrepresentation of these chemical shifts in the BMRB data bank[72]. The corresponding carbon resonances on the other side are spread over ~30 ppm, allowing a separation between 1', 4', 5' and 2'+3' chemical shifts. This is made

to use in the 3D (H)CC-TOCSY and (H)CPC-TOCSY experiments, which were introduced by Richter *et al.*[73] in 2011 (magnetization transfer Figure 4 A). With the (H)CC-TOCSY experiment a full resonance assignment of the carbon nuclei within the ribose can be achieved, where it is even possible to distinguish between C2' and C3' atoms through the length of the TOCSY transfer (Figure 4 B). The (H)CPC-experiment enables a sequential assignment through correlations between carbon atoms of two consecutive ribose moieties and the linking phosphorus, exploiting ^{n}J(C,P) coupling constants (Figure 4 C)

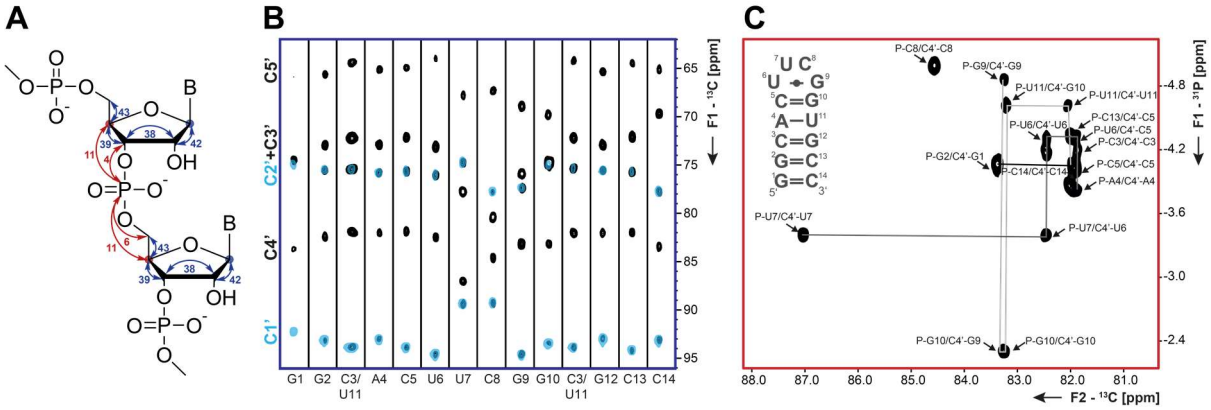


Figure 4 A Coherence transfer pathways for the ^{13}C-detected (H)CPC- (red) and (H)CC-TOCSY (blue) experiments. The *J*-coupling constants, which are used during the transfers are annotated. Carbon atoms, where the detection takes place on are marked with red and blue circles. **B** (H)CC-TOCSY spectrum of the 14 nts RNA with UUCG tetraloop with 15 ms (black) and 3 ms (light blue) TOCSY mixing times. **C** (H)CPC-spectrum of the 14 nts RNA with UUCG tetraloop. The figure has been adapted from Schnieders *et al.*[45].

In addition to the ribose moiety also the nucleobase can be assigned using ^{13}C-deteted NMR experiments. In the nucleobases of RNA, the low proton-density leads to long transfer pathways if ^{1}H-detected experiments are employed. Therefore, some of the most important NMR experiments for RNA, such as the 3D TROSY-relayed HCCH-COSY experiment for the correlation of adenosine H2C2 and H8C8[39] or the 3D HCCNH experiment for the correlation of the imino proton with H6C6/H8C8[74,75], are rather unsensitive, in particular for large RNAs. This problem can be circumvented by the ^{13}C-excited and -detected ^{13}C,^{15}N-HSQC experiment[76], with which a near to complete resonance assignment of the nucleobase can be achieved through a correlation of the different CN fragments via sizable ^{1}J(C,N) couplings. Furthermore, the resonances of the imino nitrogen atoms can be detected irrespective of the imino proton exchange leading to an increased number of signals in comparison to ^{1}H,^{15}N-HSQC spectra of the imino region (Figure 5 A and B). This insensitivity towards solvent exchange was exploited by Fürtig *et al.*[77] in 2016, where it is shown that the status of hydrogen bonding can be determined using a ^{13}C,^{15}N-spin filter HSQC experiment (coherence transfer pathway in Figure 5 C). This spin filter element leads to an exchange dependent modulation of the signal intensity. For imino protons involved in a hydrogen bond the sign of the resonance is inverted, while it stays unperturbed if the imino proton is not protected from exchange (Figure 5 D and E). Using a modified pseudo 3D version of the experiment one is further able to extract imino proton exchange rates[77].

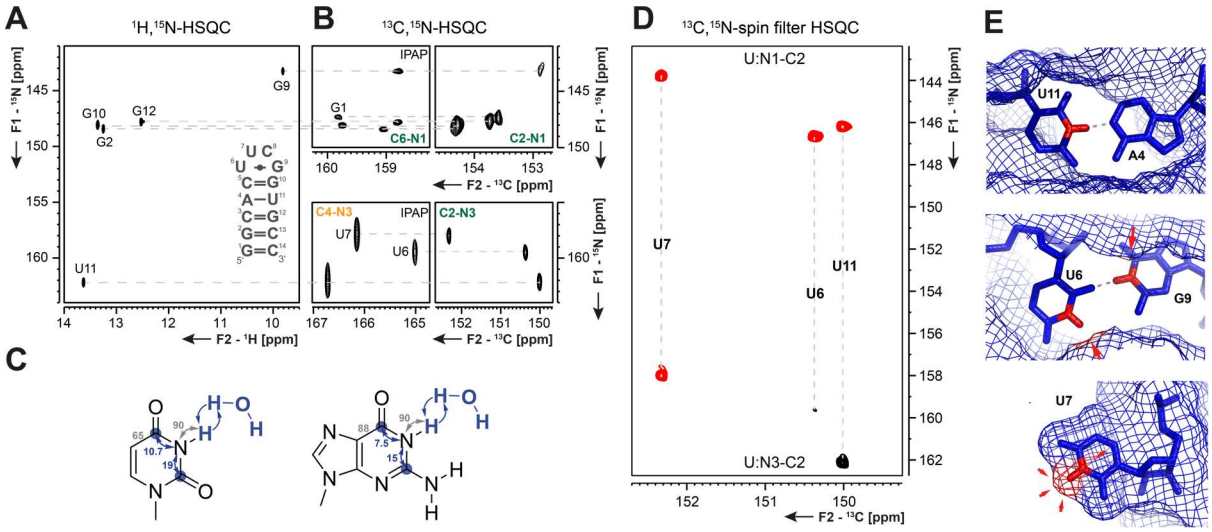


Figure 5 A ^{1}H,^{15}N-HSQC spectrum of the imino region of the 14 nts RNA with UUCG tetraloop. **B** ^{13}C,^{15}N-HSQC spectra of the imino region with and without IPAP for homonuclear CC-decoupling as indicated. The INEPT transfer times were calculated from a ^{1}J(C,N) coupling of 27 Hz (green) and 17 Hz (yellow). Horizontal dashed lines connect resonances with the same ^{15}N chemical shift. **C** Coherence transfer pathway of the ^{13}C,^{15}N-spinfilter HSQC experiment. The ^{1}J coupling constants, which are used in the INEPT transfer steps are given in blue, while inactive coupling constants for IPAP or the spin filter element are held in gray. The carbon nuclei, where the detection happens on are marked with blue circles. Solvent exchange is indicated. **D** ^{13}C,^{15}N-spin filter HSQC spectrum of the uridines U11, U6 and U7 of the 14 nts RNA with UUCG tetraloop. **E** Uridines U11 (upper panel, Watson-Crick), U6 (middle panel, sheared GU) and U7 (lower panel, solvent exposed) in their structural context. The solvent accessibility of the imino proton (red) is indicated. The figure has been adapted from Fürtig *et al.*[77] and Schnieders *et al.*[45].

Chapter II: Development of ^{13}C-detected NMR experiments for amino groups in RNA

The results presented in this chapter are summarized in the following publication[78]:

R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe, B. Fürtig, *Angew. Chemie - Int. Ed.* **2019**, *58*, 9140–9144

Author contributions are listed below:

R. Schnieders was involved in pulse sequence development and in project design, conducted all NMR experiments and the subsequent data analysis and was involved in manuscript writing

A. C. Wolter provided the 34 nts RNA aptamer, conducted structure calculations and was involved in manuscript writing.

C. Richter, J. Wöhnert, H. Schwalbe and B. Fürtig were involved in project design, in pulse sequence development and in manuscript writing.

This work was further outlined in the following minireview article[45]:

R. Schnieders, S. Keyhani, H. Schwalbe and B. Fürtig, *Chem. - A Eur. J.* **2020**, *26*, 102–113

Additionally, a detailed guide on how to set-up the ^{13}C-detected C(N)H-HDQC experiment and the "amino"-NOESY experiment proposed here has been submitted:

R. Schnieders, B. Knezic, H. Zetsche, A. Sudakov, T. Matzel, C. Richter, M. Hengesbach, H. Schwalbe and B. Fürtig, NMR spectroscopy of large functional RNAs - from sample preparation to low-gamma-detection, *Curr Protoc Nucleic Acid Chem* **2020**, *Manuscript submitted*.

2.1 Introduction

In NMR spectroscopy of RNA, imino groups play a pivotal role for a quick determination of the secondary structure since a signal for an imino proton is only observable if the imino group is protected through a hydrogen bond or in very rare cases through steric shielding[79,80]. Otherwise, the imino proton is in rapid exchange with solvent water and its resonance is broadened beyond detectability. While this enables a quick secondary structure determination on the one site, it concomitantly leads to an absence of imino proton-based information about unstructured regions of RNA on the other site. However, nucleobase interactions are not only mediated through imino protons. Apart from imino groups, also amino groups are involved in hydrogen bonding interactions as present in different canonical and non-canonical base pairings[81] as well as in RNA-protein or RNA-ligand interactions[66]. Moreover, amino groups are present at least in three out of four nucleotides; namely guanosines, adenosines and cytidines. Lastly, the chemical exchange of amino protons with solvent water is negligible slow even when not protected through hydrogen bonding[82], which in principle allows the characterization of dynamic RNAs. In NMR spectra, however, signals of amino groups appear very broad to the extent that these resonances are often beyond detectability. This line-broadening effect results from a restricted rotation around the C-NH_2 bond[83]. Especially in adenosines and guanosines this rotation is often in the intermediate exchange regime, where the exchange frequency is in the same range as the chemical shift difference between the two protons[83]. In line, only around 60% of all cytidine and 20% of all purine amino resonances are assigned in the BMRB database[72]. Also, structural information obtained by NOESY spectra is limited even if methods are used, which attempt to minimize the effects of the rotational exchange[83].

In this project, we developed two ^{13}C-detected experiments that facilitate the characterization of amino groups. The first experiment, the C(N)H-HDQC experiment, enables the detection of a full set of amino resonances for all amino group bearing nucleotides. It combines the evolution of exchange-independent ^{1}H double quantum (DQ) coherence in the indirect dimension with the direct-detection of a non-exchanging nucleus (^{13}C). As first results within this work show, this newly obtained chemical shift information can potentially be used in a chemical-shift-to-structure-relation. The second experiment proposed here is a ^{13}C-detected “amino”-NOESY experiment, which brings amino groups in direct structural context. In contrast to state-of-the-art methods the “amino”-NOESY experiment enables detection of additional inter-nucleotide correlations, as was demonstrated for a 34 nts RNA. Using these additional contacts in a structure calculation can increase the structures’ accuracy, especially when a lack of information for flexible regions of RNA is simulated by leaving out torsion angle restraints in non-helical parts. Taken together the new experiments promise a valuable addition to the standard imino proton-based NMR spectroscopical characterization of

RNA, facilitating the investigation of dynamic parts as well as non-canonical regions of the RNA under study.

2.1.1 From amino group exchange and attempts to reduce its effects

The exchange induced line-broadening for amino proton resonances is already evident in simple ^{1}H,^{15}N-HSQC spectra of the amino group region, where the set of resonances is, especially for hydrogen-bonded guanosines and adenosines, usually incomplete[9,83,84]. This effect is independent of the molecular size. Thus, even for a small 14 nts RNA only 1 out of 5 guanosine and no adenosine amino resonances are detectable, while cytidine amino signals arise as a complete set (Figure 6 A). Furthermore, the single observable guanosine resonance is close to the intermediate exchange regime and thus appears severely line broadened. Reports from literature indicate that the rotational frequencies are $k_{rot\text{-}CN}^{A} \sim 400 - 6500\ s^{-1}$[85], $k_{rot\text{-}CN}^{G} \sim 200 - 1100\ s^{-1}$[86] and $k_{rot\text{-}CN}^{C} \sim 40\ s^{-1}$[87] at ambient temperatures. For adenosines and guanosines involved in hydrogen bonds, the chemical shift differences $\Delta\omega$ match the rotational frequencies leading to line broadening. For hydrogen-bonded cytidines, the chemical shift difference is much larger with respect to the rotational frequency at ambient temperatures due to a more distinct partial double bond character in the C-NH_2 bond (Figure 6 B)[83]. Therefore, the rotation is in the slow exchange regime resulting in two sharp resonances per amino group (Figure 6 A).

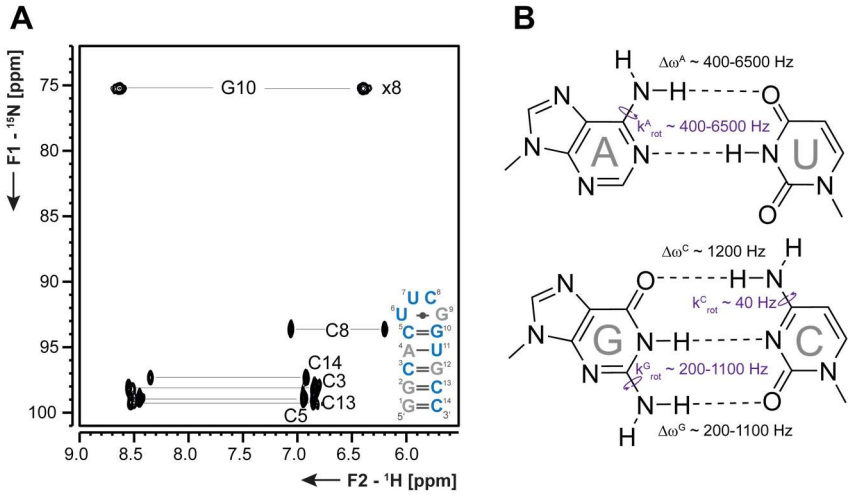


Figure 6 A ^{1}H,^{15}N-HSQC spectrum of the amino region of the 14 nts hairpin RNA with UUCG tetraloop. Nucleotides, where amino group resonances are not observed are held in gray. The spectrum was recorded with spectral windows of 16 ppm and 45 ppm in the direct ^{1}H- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 4.7 ppm, 150 ppm and 85 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Complex points were 1024 and 64 in the ^{1}H- and ^{15}N-dimensions, respectively. With 128 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 5 h. **B** AU- and GC-Watson-Crick base pairs with rotational frequencies k_{rot} around the C-NH_2 bond (purple) and the difference $\Delta\omega$ in amino proton chemical shifts in Hz.

Despite this lack of resonances for the greater part of amino groups, important structural information can still be obtained from amino groups. In ^{1}H,^{1}H-NOESY spectra for example, GC base pairs are often identified by the characteristic strong contacts between the guanosine imino proton and the cytidine amino protons[9,88]. Furthermore,

information for unusual nucleobase interactions can be obtained by amino HNN-COSY experiments, for example in AA base pairs[89] as well as in G-quadruplexes[90] or by correlating base pairing partners to non-exchangeable aromatic protons in e.g. sheared GA mismatches[91]. Additionally, amino groups can participate in hydrogen bonding interactions with the phosphate groups. These interactions can be determined by correlating the amino and the phosphate group through ^{2h}J(H,P) or ^{3h}J(N,P) coupling constants across the hydrogen bond.

Due to the high potential value of amino group correlations for structural characterization of RNAs, methods have been developed, which attempt to reduce the influence of the rotational exchange on signal line widths. The crux of these experiments lies in the H-N transfer method, where the INEPT scheme is replaced by CPMG-INEPT pulse trains, which effectively spin-lock the magnetization. For an effective quenching of the exchange contributions, it is crucial that the rate of 180° pulses is equal or higher than the rate of exchange. The prime examples here are the CPMG-HSQC experiment from Mulder *et al.*[92] and the CPMG-NOESY experiment from Mueller *et al.*[83]. As the latter experiment yields valuable structural information, it will be described in the following.

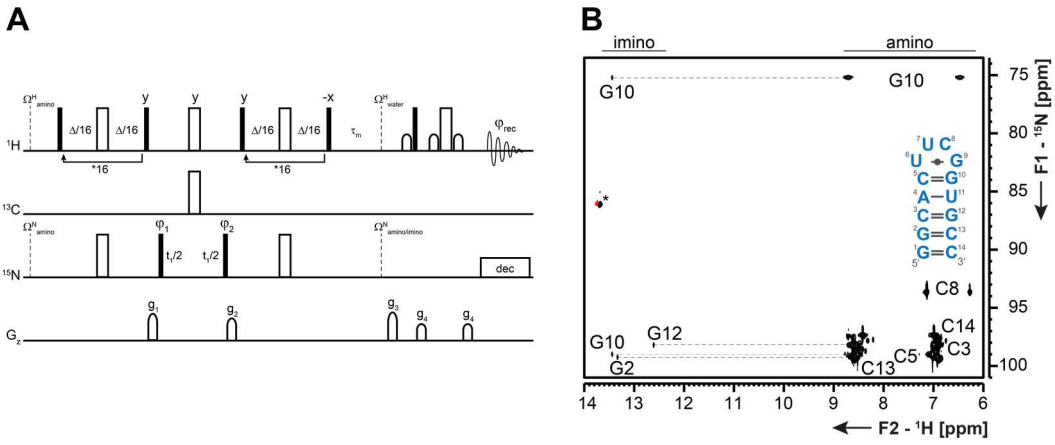


Figure 7 A Schematic pulse sequence of the CPMG-NOESY experiment. Narrow-filled bars represent rectangular 90° pulses while wide open bars mark rectangular 180° pulses. If not indicated otherwise the pulse phase is x. Other pulse phases are: φ_1 = x,-x, φ_2 = 4(x),4(-x) and φ_{rec} = x,-x,x,-x,-x,x,-x,x. Delays are: Δ = 2.78 ms = $1/4^1J$(N,H). Gradient pulses are applied for 1 ms with smooth square amplitude (SMSQ10.100) with strengths of 30% (g_1), 30% (g_2), 50% (g_3) and 40% (g_4). During acquisition ^{15}N nuclei are decoupled. **B** ^{1}H,^{15}N-CPMG-NOESY spectrum of the 14 nts RNA with UUCG tetraloop showing correlations between imino and amino groups. The resonance marked with an asterisk represents the folded U11 imino group diagonal peak. The spectrum was recorded with spectral windows of 25 ppm and 38 ppm in the direct ^{1}H- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 4.7 ppm, 150 ppm and 85 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.1 s and 0.03 s in the ^{1}H- and ^{15}N-dimensions, respectively. With 96 scans per increment and an inter-scan delay of 1.5 s, the experiment was recorded within 7.5 h. During acquisition ^{15}N nuclei were decoupled using GARP4 sequences[93].

In the indirect dimension of the CPMG-NOESY experiment, the chemical shift of amino nitrogen atoms evolves (pulse scheme in Figure 7 A). Those nitrogen atoms are not involved in exchange processes and thus give rise to resonances exhibiting sharp lines. They are then correlated via NOE to spatially close protons, where the signal is detected. This results in an H-N correlated NOESY spectrum, where the coherence of exchanging amino protons is in the transverse plane solely during the CPMG-transfer, which

represents the minimum amount of time. With this experiment, Mueller and coworkers were able to detect 19 NOE contacts, which were not observed in conventional ^{1}H,^{1}H-NOESY spectra of a 30 nts RNA[83].

However, despite the frequency of 180° pulses in the CPMG-INEPT step exceeding the rotational frequency, the spectrum still shows an incomplete set of signals, as is demonstrated for a 14 nts hairpin RNA (Figure 7 B). The only correlations visible in this case are contacts between amino groups and the imino proton, which is involved in the respective nucleobase interaction. Other correlations to e.g. H1' protons cannot be observed. This observation can be attributed to exchange contributions during free evolution periods.

Thus, contributions by exchange can be eliminated by omitting free evolution periods of spins involved in the exchange. This can be accomplished by moving from ^{1}H-detected towards heteronuclear-detected NMR experiments since the chemical shifts of the nitrogen (and carbon) atoms attached or adjacent to the protons undergoing chemical exchange are hardly affected by exchange. Heteronuclear detection is nowadays possible as there has been a constant development in cryogenic probes as well as probes optimized for heteronuclear detection[30]. Here, especially carbon direct-detection is highly important and has been applied in CNN-COSY and CN spin-filter HSQC experiments to eliminate exchange contributions caused by solvent exchange in the characterization of imino groups in RNA[94].

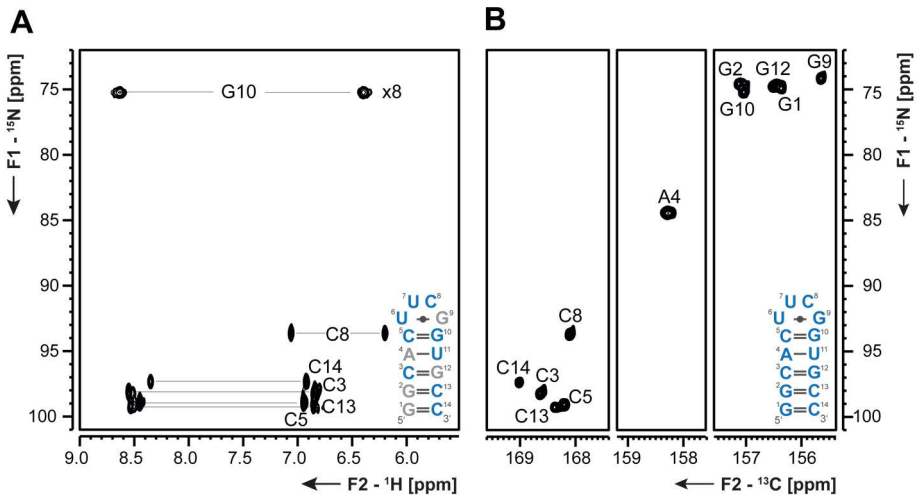


Figure 8 A ^{1}H,^{15}N-HSQC spectrum of the amino region of the 14 nts RNA with UUCG tetraloop. The spectrum is already shown in Figure 6 A, where experimental details can be found in the figure caption. **B** ^{13}C,^{15}N-HSQC spectrum with IPAP decoupling scheme of the amino region of the 14 nts RNA with UUCG tetraloop. The spectrum has been processed separately for cytidines, adenosines and guanosines. The spectrum was recorded with spectral windows of 40 ppm and 33 ppm in the direct ^{13}C- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 160 ppm, 4.7 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 512 and 28 in the ^{13}C- and ^{15}N-dimensions, respectively. With 48 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 4 h.

The simplest ^{13}C-detected 2D experiment, which is the ^{13}C,^{15}N-HSQC experiment[76], is unaffected by the rotational exchange. Thus, a full set of resonances that are frequency labeled with the amino nitrogen chemical shift during t_1, can be observed (Figure 8 B).

While the ability to detect those C-N correlations is certainly beneficial as opposed to ^{1}H-detected NMR experiments, direct observation of the amino protons still remains the ultimate goal.

2.1.2 Learning from protein experiments

Amino groups are not only found in RNA nucleobases, but under physiological conditions in their protonated form ($-NH_3^+$) also in lysine and arginine residues in proteins. They are known to be involved in many important interactions like hydrogen bonding, salt bridges or cation-π-interactions[95,96]. Both amino acids feature solvent as well as rotational exchange. For lysine residues, solvent exchange severely broadens resonances in conventional ^{1}H,^{15}N-HSQC experiments. This is impressively avoided in ^{1}H,^{15}N-HISQC experiments by the evolution of ^{15}N in-phase coherence and concomitant ^{1}H decoupling (WALTZ) in the indirect dimension of the H-N correlation experiment[97]. Furthermore, the C-N bond rotation in lysine residues was characterized in conjunction with molecular dynamic (MD) simulations, which reveals a dynamic nature of lysine hydrogen bonding interactions[98].

In the detection of arginine residues it becomes evident that resonances are not only broadened from solvent exchange but also from the rotation around the N^{ε}-C^{ζ} bond, which is often close to the intermediate exchange regime[99]. As a result, correlations to $N^{\eta'}$ and $N^{\eta''}$ are often broadened to the extent that signal overlap prevents unambiguous resonance assignment or resonances are even broadened beyond detectability[99]. Human ubiquitin, for example, harbors four arginine residues[100,101], whose H^{ε}-N^{ε} correlations can be observed in a ^{1}H,^{15}N-HSQC spectrum, while the H^{η}-N^{η} resonances appear exchange broadened (Figure 9 A). The same holds true for the correlation of the exchanging site, N^{η} with the covalently attached carbon C^{ζ} in a ^{13}C-detected ^{13}C,^{15}N-HSQC spectrum (Figure 9 B)[99,102].

As was shown by Mackenzie *et al.*[103], the rotation-induced signal broadening can be avoided by evolving ^{15}N DQ instead of SQ coherence in the indirect dimension of the ^{13}C-^{15}N correlation experiment. With this, the sum of the ^{15}N chemical shifts $\Omega_{\eta'} + \Omega_{\eta''}$ is detected. This sum remains constant as it is independent of the rotation, which interchanges the two exchanging sites into each other. For ubiquitin, this results in a ^{13}C,^{15}N-HDQC spectrum, where C^{ζ}-N^{η} correlations can be observed for all four arginine residues (Figure 9 C). The DQ signals come to resonance at the sum of the chemical shift differences from the carrier frequency[103,104] and are processed in a way that they appear at the mean $^{15}N^{\eta}$ chemical shifts. They exhibit a sharp line, again showing their independence of the rotational motion[103]. The authors further state that it is indispensable for a correct application of this experiment, that the magnetization of the exchanging site does at no point in time exist as SQ in the transverse plane[103]. As soon as this is the case, the effect of the rotational exchange comes into play.

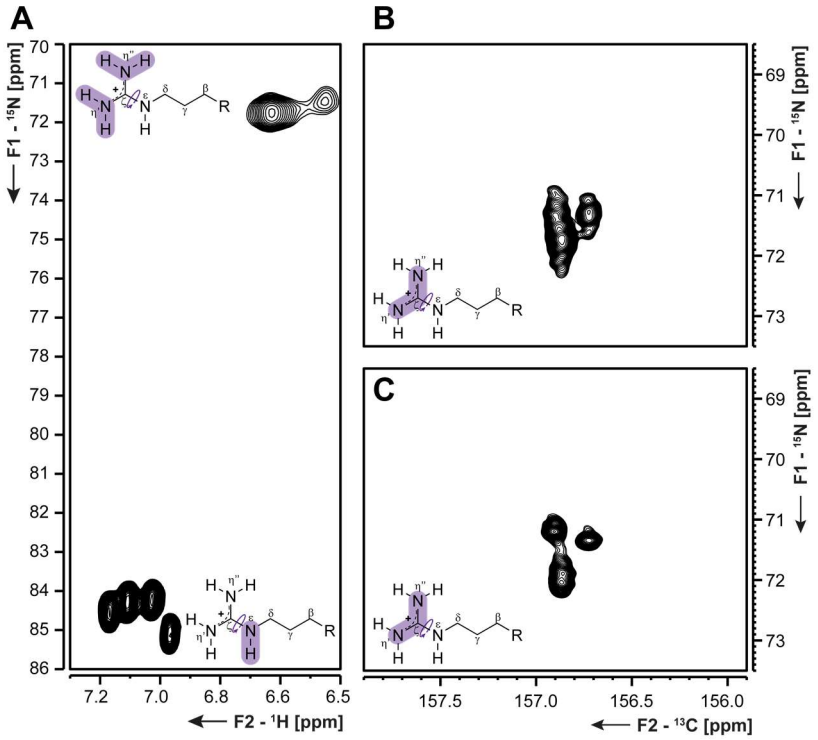


Figure 9 A ^{1}H,^{15}N-HSQC spectrum of the H-N$^{\varepsilon}$ and H-N$^{\eta}$ regions of ^{13}C,^{15}N-labeled ubiquitin[100,101]. The spectrum was recorded with spectral windows of 16 ppm and 80 ppm in the direct ^{1}H- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 4.7 ppm, 49.5 ppm and 100 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.07 s and 0.05 s in the ^{1}H- and ^{15}N-dimensions, respectively. With 8 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 1.3 h. **B** ^{13}C,^{15}N-HSQC spectrum and **C** ^{13}C,^{15}N-HDQC spectrum of the C$^{\zeta}$-N$^{\eta}$ correlation for the arginine residues in ubiquitin. The spectra were recorded with spectral windows of 15 ppm and 8 ppm in the direct ^{13}C- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 156 ppm, 4.7 ppm and 71 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.11 s and 0.07 s in the ^{13}C- and ^{15}N-dimensions, respectively. With 80 scans per increment and an inter-scan delay of 3.7 s, the experiment in **B** was recorded within 6.5 h. With 160 scans per increment and an inter-scan delay of 3.7 s, the experiment in **C** was recorded within 12.7 h.

The very same principal of evolving DQ coherence to avoid exchange contributions was applied within this PhD project for amino groups in RNA. Here, not ^{15}N but ^{1}H DQ coherence is evolved in the indirect dimension of the ^{13}C-detected C(N)H-HDQC experiment. Furthermore, amino groups are brought in direct structural context with the within this PhD project developed ^{13}C-detected "amino"-NOESY experiment, which yields NOE contacts that are not accessible with the previously introduced CPMG-NOESY experiment.

2.2 Materials and Methods

2.2.1 The RNAs under study

The ^{13}C-detected C(N)H-HDQC and the "amino"-NOESY experiments were mainly optimized using the 14 nts RNA with UUCG tetraloop[105,106] and the 34 nts GTP class II aptamer RNA (Figure 10 box, Table 1)[107,108].

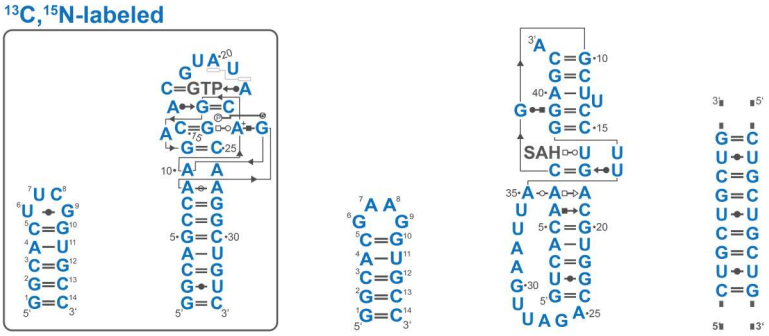


Figure 10 Secondary structures of the 14 (UUCG) and 34 nts RNAs (left box) as well as the 14 (GAAG), 43 and 329 nts long RNAs (right). Nucleotides in blue mark ^{13}C-^{15}N-labeling while unlabeled nucleotides are held in gray (GTP and SAH).

The C(N)H-HDQC experiment was further applied to the 14 nts RNA with GAAG tetraloop[109], the 43 nts SAM/SAH-binding riboswitch[110,111] and the 329 nts CUG repeat RNA[112,113] (Figure 10 right, Table 1). These results were used for a chemical-shift-to-structure analysis. Additionally, preliminary experiments were conducted for the 34 nts GTP class II aptamer RNA in the absence of ligand GTP and with and without Mg^{2+} ions for a potential follow-up project.

Table 1 List of RNAs under study including their sequences from 5' to 3' and their isotope labeling schemes.

length	sequence	isotope-labeling
14 nts	5'-pppGGCACUUCGGUGCC-3'	uniformly ^{13}C,^{15}N
34 nts	5'-pppGGCAGCCAGAAGAGCACGUAUACGCAAGGCUGUCp$^{2'3'cycl}$-3'	uniformly ^{13}C,^{15}N
14 nts	5'-pppGGCACGAAGGUGCCp$^{2'3'cycl}$-3'	uniformly ^{13}C,^{15}N
43 nts	5'-pppGUCACAACGGCUUCCUGACGUGGCAGAUUGAAUUAUUGGAGCAp$^{2'3'cycl}$-3'	uniformly ^{13}C,^{15}N
329 nts	5'-GGGAGACCGGCAGAUCUGAUAUCAUCGAUGAAUU(CUG)$_{97}$GGGG-3'	uniformly ^{13}C,^{15}N

2.2.2 Sample preparation

14 nts UUCG RNA

The uniformly ^{13}C,^{15}N labeled 14 nts hairpin RNA with cUUCGg tetraloop was both prepared via *in vitro* transcription and purchased from Silantes (Munich, Germany). The NMR sample contained 300 µM of RNA in 25 mM potassium phosphate buffer (pH 6.4) and 10% D_2O. Spectra were referenced externally to DSS and the resonance assignment was taken from previously published data (BMRB 5705)[105].

34 nts RNA

The uniformly ^{13}C,^{15}N-labeled GTP class II aptamer was prepared by Antje C. Wolter via run-off *in vitro* transcription from linearized plasmid DNA using T7 RNA polymerase as described in literature[107]. To generate 3' homogeneity, a self-splicing Hammerhead ribozyme was used[114]. The ^{13}C,^{15}N-labeled rNTPs were obtained commercially (Silantes, Munich Germany). The NMR sample contained 25 mM potassium phosphate buffer (pH 6.3), 25 mM KCl, 2 mM $Mg(OAc)_2$, 0.98 mM RNA, 1.96 mM unlabeled GTP and 5% D_2O. The resonance assignment was taken from the literature (BMRB 25661)[107]. The resonance assignment for adenosine C6 atoms was obtained using a 3D TROSY relayed HCCH-COSY experiment[39].

14 nts GAAG RNA

The uniformly ^{13}C,^{15}N-labeled 14 nts RNA with GAAG tetraloop was prepared as described in chapter 4.2.2.

43 nts RNA

The uniformly ^{13}C,^{15}N-labeled SAM/SAH-binding riboswitch was prepared by Katharina Weickhmann via run-off *in vitro* transcription from linearized plasmid DNA using T7 RNA polymerase as described in literature[111]. The ^{13}C,^{15}N-labeled rNTPs were obtained commercially (Silantes, Munich Germany). To generate 3' homogeneity, a self-splicing HDV ribozyme was used[115]. Purification and folding was performed as described earlier[111]. The NMR sample contained 25 mM potassium phosphate buffer (pH 6.2), 50 mM KCl, 2 mM $Mg(OAc)_2$, 0.8 mM RNA, 2 mM unlabeled SAH and 8% D_2O. The resonance assignment was taken from the literature (BMRB 27452)[110].

329 nts RNA

The uniformly ^{13}C,^{15}N labeled 329 nts CUG repeat RNA was provided by Dr. Matthias Görlach and was prepared via run-off *in vitro* transcription from linearized plasmid DNA using T7 RNA polymerase as described[112]. Purification and RNA folding was conducted using previously published protocols[116]. The sample used for NMR spectroscopy measurements contained 700 μM RNA, 25 mM potassium phosphate (pH 6.4), 50 mM KCl and 8% D_2O. The resonance assignment was taken from the literature[112].

2.2.3 NMR spectroscopy

All NMR spectra were recorded at 298 K on a 800 MHz (18.8 T) Bruker NMR spectrometer, equipped with a ^{13}C-optimized 5 mm, z-axis gradient ^{13}C, ^{15}N [^{1}H]-TXO cryogenic probe, where ^{13}C and ^{15}N are recorded on the inner and ^{1}H is detected on the outer coil. To also show applicability on the more broadly accessible ^{1}H optimized probes, the ^{13}C-detected C(N)H-HDQC experiment for the 34 nts class II aptamer was additionally applied on a 700 MHz NMR spectrometer. This spectrometer is equipped with a 5 mm, z-axis gradient ^{1}H [^{13}C, ^{15}N, ^{31}P]-QCI cryogenic probe where ^{1}H is detected on the inner and ^{13}C and ^{15}N are recorded on the outer coil. Spectra were processed with Topspin (version 3.5 pl6) and analyzed using Sparky[117].

In order to conduct the experiments developed here with the maximum achievable sensitivity, selective pulses, transfer methods and transfer delays have been optimized as discussed in chapter 2.3.2. Optimal conditions are given in the figure caption of Figure 12.

2.2.4 Structure calculation

Structure calculations for the 34 nts GTP class II aptamer RNA were conducted by Antje C. Wolter with cyana 3.98 using different sets of restraints. Among these were a full set of restraints, which was described earlier[108] that include 863 upper limit NOE contacts, 15 lower limit NOE restraints, 168 hydrogen bond distance constraints, 60 ribose conformation restraints, 107 backbone torsion angle restraints as well as 27 restraints for the glycosidic torsion angle. Additionally, a reduced set of restraints was used, which included all of the NOE restraints and hydrogen bond distance constraints as well as all torsion angle restraints for helical parts of the RNA (residues 1-6 and 28-34). This reduced set consisted of 23 ribose conformation restraints, 66 backbone torsion angle restraints as well as 12 restraints for the glycosidic torsion angle. The influence of 22 newly derived NOE contacts, which resulted from the ^{13}C-detected “amino”-NOESY experiment proposed here (list in Appendix 1 Table 17), was tested by including them in structure calculations with the previously described full set and with the reduced set of experimental data. The new NOE restraints were included with upper distance limits of 6.5 Å. For residues 2-12,14-20,22-33, excluding the closing base pair G1:C34 and the flexible residues A13 and U21, RMSD values were calculated with MOLMOL[118]. Structures were visualized with MOLMOL and Pymol (DeLano Scientific LLC; Schrödinger, Inc).

2.3 Results and Discussion

2.3.1 Pulse sequence analysis

The schematic coherence transfer pathways of both ^{13}C-detected experiments developed in the course of this PhD project, the C(N)H-HDQC and the "amino"-NOESY experiment are shown in Figure 11 A and B, respectively. In the following, these experiments will be described using the product operator formalism while omitting the IPAP scheme for simplicity.

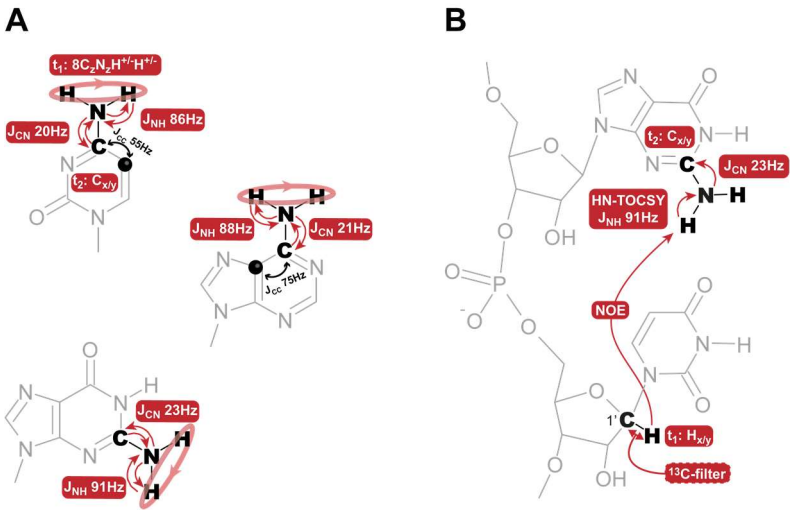


Figure 11 Schematic coherence transfer pathways for the ^{13}C-detected C(N)H-HDQC experiment (**A**) and the "amino"-NOESY experiment with optional ^{13}C-filter (**B**). The size of the ^1J scalar coupling constants is given. If not indicated otherwise, the transfer method is INEPT[119]. Coherences that evolve during the direct (t_2) and indirect (t_1) dimensions are given. Nuclei that are involved in IPAP decoupling schemes are marked with a black circle. The figure has been adapted from Schnieders *et al.*[78].

The C(N)H-HDQC experiment involves ^{13}C-excitation with a subsequent amino-selective CN-INEPT transfer step. The generated coherence at time point **a** is thus $2C_zN_y$ (pulse scheme Figure 12 A, pulse program in Appendix 1 5.3.3). Coherence is further transferred by an NH-INEPT step, from the amino nitrogen to both amino protons H^1 and H^2:

$$2C_zN_y \xrightarrow{\pi J(N,H^1)\Delta} 2C_zN_y\cos(\pi J(N,H^1)\Delta) - 4C_zN_xH_z^1\sin(\pi J(N,H^1)\Delta)$$

As $^1J(N,H^1) = {}^1J(N,H^2)$ in NH_2 groups, the following 4 terms emerge:

$$\xrightarrow{\pi J(N,H^2)\Delta} 2C_zN_y\cos^2(\pi J(N,H)\Delta) - 4C_zN_xH_z^2\cos(\pi J(N,H)\Delta)\sin(\pi J(N,H)\Delta)$$
$$- 4C_zN_xH_z^1\sin(\pi J(N,H)\Delta)\cos(\pi J(N,H)\Delta) - 8C_zN_yH_z^1H_z^2\sin^2(\pi J(N,H)\Delta)$$

With $\Delta = 1/2{}^1J(N,H)$ and a ^{15}N $90°_x$ pulse (pulse scheme Figure 12 A and B) the longitudinal 4 spin order $8C_zN_zH^1{}_zH^2{}_z$ is obtained at time point **b**:

$$\longrightarrow -8C_zN_yH_z^1H_z^2 \xrightarrow{90°_x(N)} -8C_zN_zH_z^1H_z^2$$

The following 90° pulse on protons transforms the coherence to $-8C_zN_zH_y^1H_y^2$, which is equal to $-8C_zN_z(ZQ_x - DQ_x)$, when considering the correlations depicted in Table 2.

Table 2 Relation between ZQ/DQ coherence and the cartesian product operators.

annotation	relation
ZQ_x	$0.5(H_x^1H_x^2 + H_y^1H_y^2)$
ZQ_y	$0.5(H_y^1H_x^2 - H_x^1H_y^2)$
DQ_x	$0.5(H_x^1H_x^2 - H_y^1H_y^2)$
DQ_y	$0.5(H_x^1H_y^2 + H_y^1H_x^2)$
Ω_{ZQ}	$\Omega_{H^1} - \Omega_{H^2}$
Ω_{DQ}	$\Omega_{H^1} + \Omega_{H^2}$

During t_1, proton chemical shift evolves:

$$-8C_zN_z[ZQ_x \cos\Omega_{ZQ}t_1 + ZQ_y \sin\Omega_{ZQ}t_1 - DQ_x \cos\Omega_{DQ}t_1 - DQ_y \sin\Omega_{DQ}t_1]$$

With the relations from Table 2 and after the ^{1}H 90°$_x$ pulse the obtained coherences are:

$$-8C_zN_z\begin{bmatrix} 0.5(H_x^1H_x^2 + H_z^1H_z^2)\cos\Omega_{ZQ}t_1 + 0.5(H_z^1H_x^2 - H_x^1H_z^2)\sin\Omega_{ZQ}t_1 \\ -0.5(H_x^1H_x^2 - H_z^1H_z^2)\cos\Omega_{DQ}t_1 - 0.5(H_z^1H_x^2 - H_x^1H_z^2)\sin\Omega_{DQ}t_1 \end{bmatrix}$$

The application of a gradient pulse dephases magnetization that exhibits a transverse component. Thus, the resulting coherence (neglecting the factor 0.5) at time point **c** after the ^{15}N 90°$_x$ pulse can be described as:

$$8C_zN_yH_z^1H_z^2\cos\Omega_{ZQ}t_1 + 8C_zN_yH_z^1H_z^2\cos\Omega_{DQ}t_1$$

By appropriate inversion of the receiver phase and concomitant cycling of the ^{15}N 90° pulse before t_1, the second term is selected, which is modulated with the ^{1}H DQ chemical shift. With two subsequent reverse INEPT steps (H-N and N-C), the coherence is transferred back to the amino group attached carbon nucleus, where the FID is recorded. The frequency discrimination in the indirect dimension is achieved by cycling of the ^{1}H 90° pulse before t_1 by 45°, as was described previously[120].

It is to note here, that usually the exact chemical shift of a resonance has to be calculated tediously in DQ/ZQ-spectroscopy[103,104], as signals come to resonance depending on their offset from the carrier frequency. Thus, when changing the carrier frequency, the peak positions will change accordingly, which complicates the choice of carrier frequencies and spectral windows while setting up the experiment. In order to avoid this and to facilitate the setup, we applied the following trick: The increment for chemical shift evolution in the indirect dimension is divided by 2. With this slight modification, signals come to resonance at the mean chemical shift of the exchanging sites, irrespective of the position of the carrier frequency.

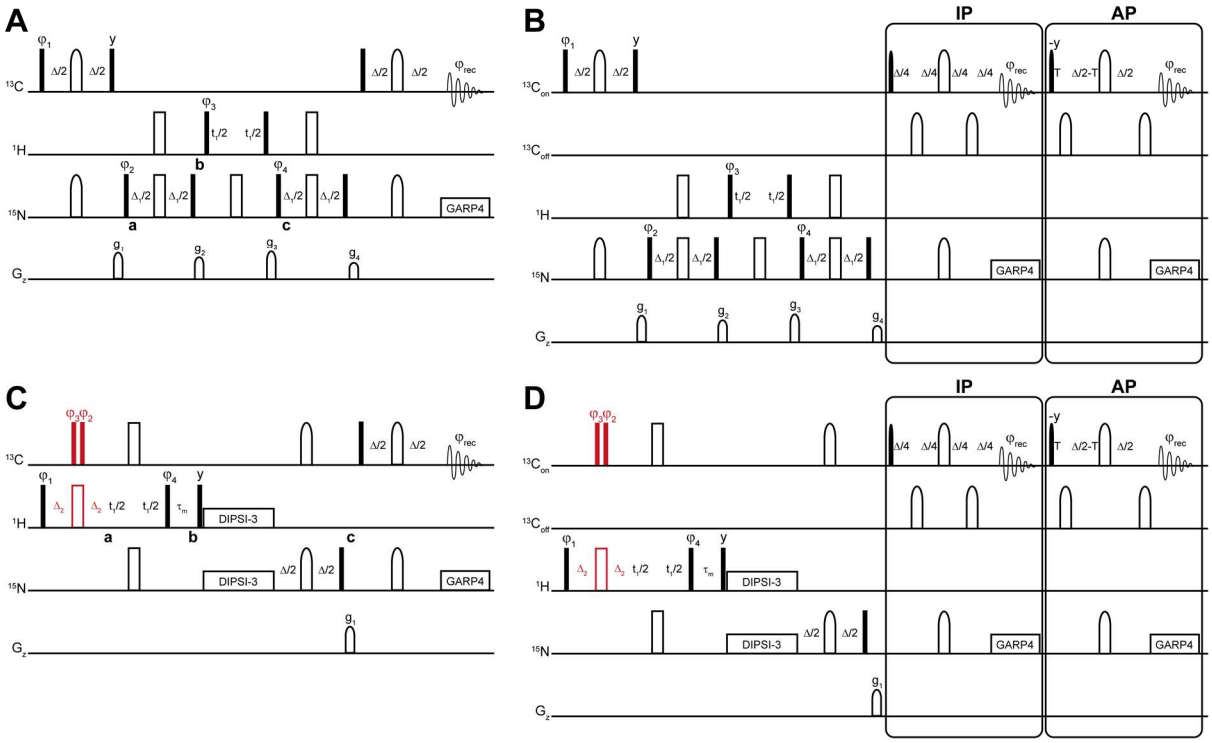


Figure 12 Schematic representation of the pulse sequence for **A** the ^{13}C-detected C(N)H-HDQC experiment without IPAP and **B** the ^{13}C-detected C(N)H-HDQC experiment with IPAP for homonuclear decoupling. Pulse scheme for the "amino"-NOESY experiment with optional ^{13}C-filter (red) without **C** and with **D** IPAP for homonuclear CC-decoupling. All following parameters like pulse lengths refer to 800 MHz NMR spectrometers. The default phase is x. Narrow-filled bars represent rectangular 90° pulses, while wide open bars mark rectangular 180° pulses. Gradient pulses and selective 180° pulses (^{13}C: 0.4 ms Q3, on-/ off-resonant and ^{15}N: 1.2 ms Reburp2[121]) are represented using semi-elliptic unfilled shapes. Semi-elliptic narrow-filled shapes describe selective 90° pulses (^{13}C: 0.6 ms Q5, on resonant). ^{15}N-nuclei are decoupled during acquisition using asynchronous GARP4 sequences[93]. Gradient pulses were applied with a smoothed square amplitude (SMSQ10.100) for 1 ms and 100% gradient strength corresponds to 53 G/cm. Incrementation of the delay t_1 occurred with half of the dwell time in the C(N)H-HDQC experiment (**A** and **B**). For this experiment frequency discrimination in the indirect dimension was achieved by cycling φ_3 by 45°. Delays are: $\Delta = 1/2{}^1J(C,N) = 19.23$ ms, $\Delta_1 = 1/2{}^1J(N,H) = 4.17$ ms, $T = 1/4{}^1J(C,C) = 4.5$ ms (cytidine) or 3.3 ms (adenosine), $\Delta_2 = 1/2{}^1J(C,H) = 6.25$ ms and $\tau_M = 150$ ms. Pulse phases for **A** and **B** are: φ_1 = x,-x, φ_2 = (4)x,(4)-x, φ_3 = (2)x,(2)y,(2)–x,(2)-y, φ_4 = (8)x,(8)-x and φ_{rec} = x,(2)-x,x,-x,(2)x,(2)-x,(2)x,-x,x,(2)-x,x. Pulse phases for **C** and **D** are: φ_1 = x,-x, φ_2 = (4)x,(4)-x, φ_3 = (2)x,(2)-x, φ_4 = (8)x,(8)-x and φ_{rec} = x,(2)-x,x,-x,(2)x,(2)-x,(2)x,-x,x,(2)-x,x,-x,(2)x,-x,x,(2)-x,(2)x,(2)-x,x,-x,(2)x,-x. Gradient strengths for **A** and **B** are 50% (g_1), 60% (g_2), 48% (g_3) and 35% (g_4). Gradient strength for g_1 in **C** and **D** is 35%. For the H-N transfer in **C** and **D** a 7.8 ms DIPSI-3 sequence[122] was applied at 3.5 kHz.

The "amino"-NOESY experiment (schematic coherence transfer pathway in Figure 11 A) is ^{1}H-excited and ^{13}C-detected, which allows a reduction of the inter-scan delay according to the shorter ^{1}H T_1 relaxation time. ^{1}H-excitation and a subsequent ^{13}C-filter transfers coherence to -H_y for all ^{13}C-bound protons at time point **a** (pulse scheme Figure 12 C and D, pulse program in Appendix 1 5.3.4). If the ^{13}C-filter is not applied, all protons are along -H_y as well. Thus, this case will not be discussed separately. After ^{1}H chemical shift evolution the magnetization is aligned along the z-axis.

$$-H_y^1 \cos(\Omega_{H^1} t_1) + H_x^1 \sin(\Omega_{H^1} t_1) \xrightarrow{90°_x(H)} -H_z^1 \cos(\Omega_{H^1} t_1) + H_x^1 \sin(\Omega_{H^1} t_1)$$

Via NOE, magnetization is transferred for longitudinal proton nuclei to protons within a distance of 5 Å during τ_m. This proton magnetization of a second proton is modulated

with the chemical shift of the first proton (time point **b**) and is brought in the transverse plane through application of a ^{1}H 90°$_{y}$ pulse.

$$-\mathrm{H}_x^2 \cos(\Omega_{\mathrm{H}^1} \mathrm{t}_1) \xrightarrow{\mathrm{HN-TOCSY}} -\mathrm{N}_\mathrm{x} \cos(\Omega_{\mathrm{H}^1} \mathrm{t}_1)$$

Since relaxation through the C-N bond rotation comes into effect as soon as amino protons are in the transverse plane, a hetero HN-TOCSY sequence was used for the H-N transfer, which effectively spin-locks the magnetization. In our hands, the TOCSY sequence was superior to INEPT and CPMG transfer methods (data not shown). With a following NC-INEPT step, coherence is further transferred to the neighboring carbon, so that at time point **c** the longitudinal two spin order $-2\mathrm{N_z C_z} \cos(\Omega_{\mathrm{H}^1} \mathrm{t}_1)$ results, which is then refocused to in-phase carbon coherence with another CN-INEPT step. Detection happens on this quaternary carbon, exhibiting a very sharp resonance. At a closer inspection of the pulse schemes of the ^{1}H-detected CPMG-NOESY and the ^{13}C-detected "amino"-NOESY experiment proposed here (Figure 7 and Figure 12), it is evident that the time during which the coherence of the exchanging species is in the transverse plane is equal. Potential advantages are thus expected to arise from the detection on the sharp ^{13}C resonance of the quaternary carbon and the unnecessity of water suppression schemes, which decrease signals around the water resonance such as H1' protons.

Finally, both experiments proposed here are selective for amino groups by utilizing selective 180° nitrogen pulses in the CN-transfer steps. Moreover, selective 180° carbon pulses are used in these transfer delays in order to avoid further relocation of the coherence to neighboring carbon atoms in cytidines and adenosines.

2.3.2 Optimization of the CN-transfer delay

To guarantee the highest achievable sensitivity for both ^{13}C-detected experiments developed here, the CN-transfer delay was optimized in an amino-selective ^{13}C,^{15}N-HSQC experiment (schematic pulse sequence and pulse program in Appendix 5.3.1) using the 14 nts RNA with UUCG tetraloop. Here, the first increment was recorded for transfer delays deduced from ^{1}J(C,N) coupling constants between 12 and 40 Hz (Figure 13 left). Integration of the characteristic regions for cytidines, adenosines and guanosines yielded optimal coupling constants of 24 Hz (cytidines and guanosines) and 26 Hz (adenosines) for the calculation of the transfer delays. It is to note here, that especially for guanosines, the differences between 22 and 34 Hz are minute. The coupling constants in the nucleobase are not expected to change dramatically between different RNAs and were thus used for all RNAs under study. Especially for larger RNAs, however, optimal transfer times might be shorter due to the enhanced relaxation.

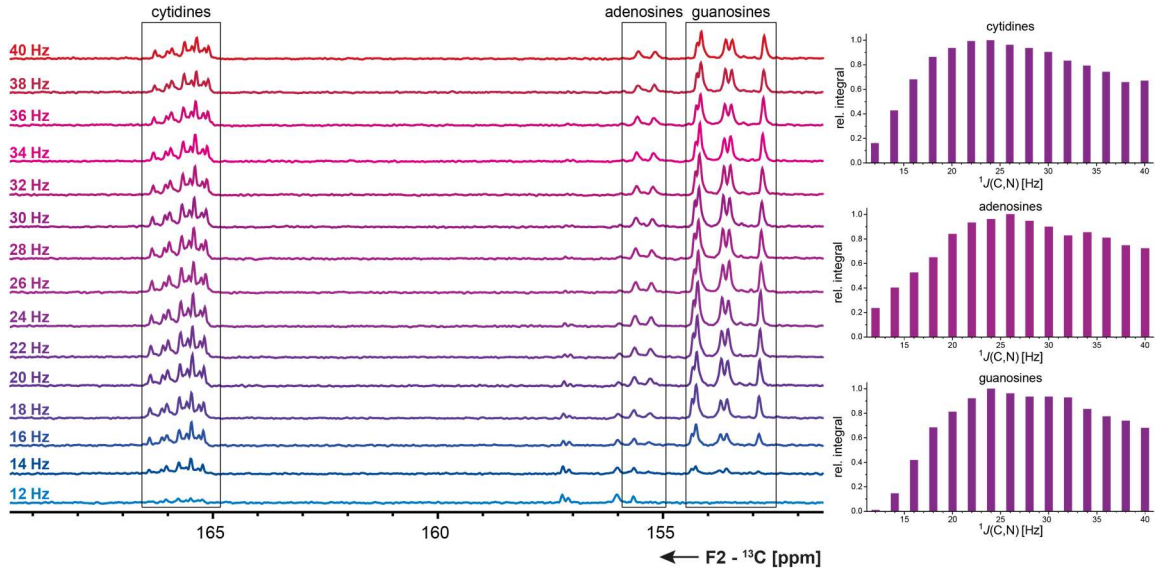


Figure 13 First increments of the amino-selective CN-HSQC experiment with transfer delays calculated for coupling constants between 12 and 40 Hz, in order to optimize transfer efficiency. Areas, where cytidines, adenosines and guanosines come to resonance are annotated accordingly and were integrated (right) and plotted in dependence of the ^1J(C,N) coupling constants that were used for the calculation of the transfer time = $1/2^1J$(C,N). To minimize errors, experiments were conducted with 1024 scans per increment. Carrier frequencies were 156 ppm, 7.1 ppm and 96 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. The direct dimension was recorded with a spectral window of 50 ppm and an acquisition time of 0.1 s. With an inter-scan delay of 3 s, each experiment was recorded within 1 h.

2.3.3 Application of the ^{13}C-detected C(N)H-HDQC experiment

The ^{13}C-detected C(N)H-HDQC experiment was first applied using the 14 nts RNA with UUCG tetraloop. To evaluate the influence of the DQ evolution, the C(N)H experiment was similarly conducted, but with the evolution of SQ coherence in the indirect dimension. This experiment will be termed C(N)H-HSQC in the following (schematic pulse sequence in Appendix 1 Figure 58 and pulse program in chapter 5.3.3). Furthermore, ^{1}H,^{15}N-HSQC and amino-selective ^{13}C,^{15}N-HSQC spectra were recorded for comparison.

As described in chapter 2.1.1 the ^{1}H,^{15}N-HSQC spectrum (Figure 14 A) yields an incomplete set of signals, while in the ^{13}C,^{15}N-HSQC spectrum (Figure 14 B), all correlations for the amino group bearing nucleotides can be observed. In the ^{13}C-detected C(N)H-HSQC spectrum (Figure 14 C), again not all correlations are present. Here, similar to the ^{1}H,^{15}N-HSQC experiment, all cytidine resonances and the resonances for G10 can be detected. Additionally, signals for G1 and G9 are observable with the C(N)H-HSQC experiment, which are not present in the ^{1}H,^{15}N-HSQC spectrum. This is due to the fact that the proton magnetization in the C(N)H-HSQC experiment exists only in the transverse plane during the chemical shift evolution and stays along the z-axis otherwise. By contrast, in the ^{1}H,^{15}N-HSQC experiment, the proton magnetization is in the x/y plane throughout the whole INEPT transfer step, before chemical shift is evolved. During this delay of approximately 5 ms, relaxation has already erased the signals for G1 and G9 despite the choice of a suitable CPMG cycle. Through the evolution of DQ instead of SQ chemical shift in the indirect dimension of the ^{13}C-detected C(N)H-HDQC experiment, the

effect of rotational exchange can be avoided, thus resulting in a spectrum with a full set of sharp amino resonances for all of the amino group bearing nucleotides (Figure 14 D).

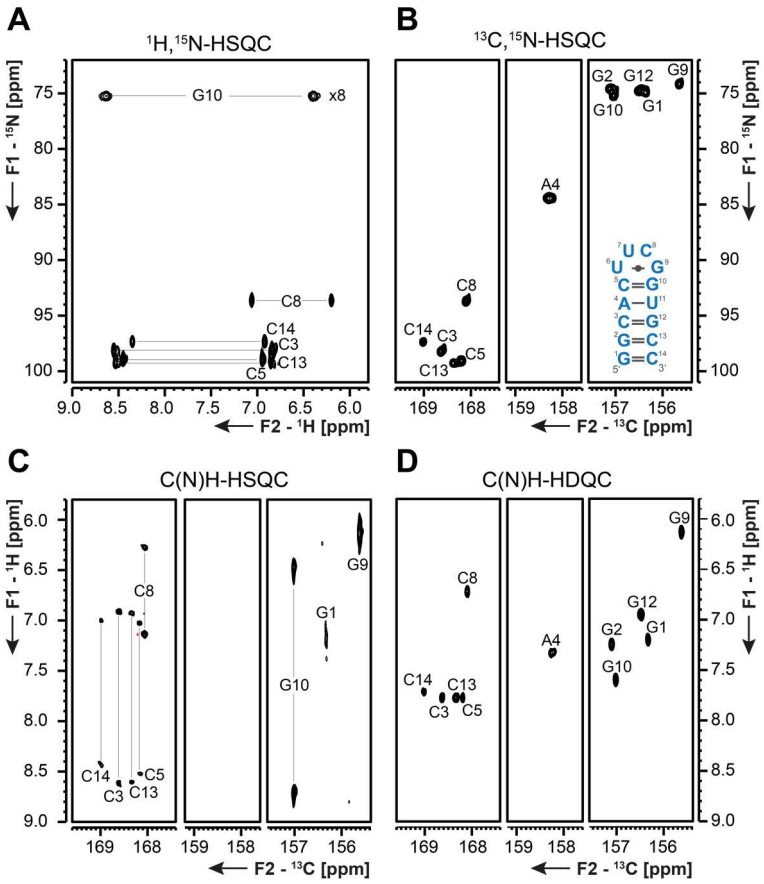


Figure 14 A ^{1}H,^{15}N-HSQC and **B** ^{13}C,^{15}N-HSQC spectra of the 14 nts hairpin RNA. Experimental details are depicted in Figure 6 and Figure 8, respectively. **C** ^{13}C-detected C(N)H-HSQC spectrum of the 14 nts RNA with IPAP decoupling. The spectrum has been processed separately for cytidines, adenosines and guanosines and was recorded with spectral windows of 100 ppm and 3.2 ppm in the ^{13}C- and ^{1}H-dimensions, respectively. Carrier frequencies were 160 ppm, 7.3 ppm and 86 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 2048 and 80 in the ^{13}C- and ^{1}H-dimensions, respectively. With 64 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 8 h. **D** ^{13}C-detected C(N)H-HDQC spectra of the 14 nts RNA with UUCG tetraloop. Three different spectra have been recorded with optimized parameters for cytidines, adenosines and guanosines. However, it is also possible to record one spectrum with averaged parameters for all three nucleotides (Appendix Figure 59). The spectra were recorded with a spectral window of 100 ppm and 2048 complex points in the ^{13}C-dimension. Carrier frequencies were 160 ppm and 86 ppm for ^{13}C and ^{15}N, respectively and the inter-scan delay was 2.5 s. The cytidine-optimized spectrum was recorded with IPAP decoupling scheme and a spectral window of 1.3 ppm and 16 complex points in the ^{1}H-dimension. The ^{1}H carrier frequency was 7.1 ppm. With 64 scans per increment, the experiment was recorded within 1.7 h. The adenosine-optimized spectrum was recorded with IPAP decoupling scheme and a spectral window of 0.3 ppm and 4 complex points in the ^{1}H-dimension. The ^{1}H carrier frequency was 7.2 ppm. With 64 scans per increment, the experiment was recorded within 0.5 h. The guanosine-optimized spectrum was recorded with a spectral window of 1.8 ppm and 22 complex points in the ^{1}H-dimension. The ^{1}H carrier frequency was 7.2 ppm. With 128 scans per increment, the experiment was recorded within 4.5 h. The figure was adapted from Schnieders *et al.*[78].

In order to quantify the line-broadening effect of the rotation, line widths were determined in the indirect dimension of DQ and SQ spectra. This was conducted in spectra that exhibit a high resolution of 9 Hz in the indirect ^{1}H-dimension to enable the detection of small variations in line widths. The comparison in Figure 15 clearly shows that line widths of amino protons in DQ spectra are generally significantly smaller than the corresponding ^{1}H-line widths in SQ spectra, verifying that DQ line widths are independent of the C-NH_2 bond rotation. Furthermore, DQ line widths are of the same

size, irrespective of the corresponding residue being embedded in the stem or in the tetraloop of the 14 nts hairpin RNA. This confirms that the exchange with solvent water is negligible slow for amino protons in RNA, as was also described in literature[82].

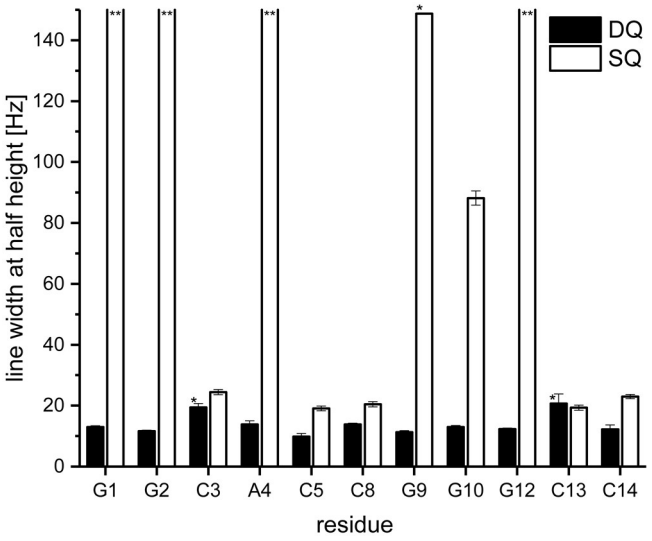


Figure 15 Comparison of SQ (open bars) and DQ (filled bars) ^{1}H line widths at half height. Line widths and the corresponding errors were determined using the deconvolution tool in Topspin. The SQ and DQ spectra used here were acquired with an ultra-high resolution of ~9 Hz in the indirect dimension and were, if not marked with an asterisk, processed without window function. Signals, which are broadened beyond detectability must exhibit a broader line width, than the broadest observable signal (G9) and are shown for representation purposes (marked with a double asterisk). SQ line widths represent the mean values from ^{1}H,^{15}N-HSQC and C(N)H-HSQC spectra. The figure was adapted from Schnieders *et al*.[78].

After having implemented the ^{13}C-detected C(N)H-HDQC experiment on a small reference RNA, the experiment was applied to a 34 nts GTP-binding aptamer[107,108]. Here, again an incomplete set of amino group signals was observed in the ^{1}H,^{15}N-HSQC spectrum (Figure 16 A). While all cytidine resonances were detectable, only 50% of all adenosine and 40% of all guanosine signals were observed. Notably here is that guanosines involved in unusual interactions (e.g. base quartets or triplets, Figure 16 E) do all give rise to a correlation. This is especially interesting for residue G12, where the amino group is involved in two hydrogen bonds. The C-NH_2 bond rotation is thus slowed down so that two sharp peaks arise. The amino protons of residue A11 appear strongly downfield shifted, which is due to the protonation of this nucleotide.

As opposed to the ^{1}H,^{15}N-HSQC spectrum, the ^{13}C,^{15}N-HSQC spectrum again harbors a full set of resonances (Figure 17 A), as it is independent of the rotational exchange. When extending the ^{13}C,^{15}N-HSQC experiment by an additional INEPT N-H transfer, obtaining the C(N)H-HSQC experiment (Figure 17 B), again the same incomplete set of signals is detected. Apart from an additional signal for G1, as was also observed for the 14 nts hairpin RNA, the resonances are identical with the ones obtained in the ^{1}H,^{15}N-HSQC spectrum.

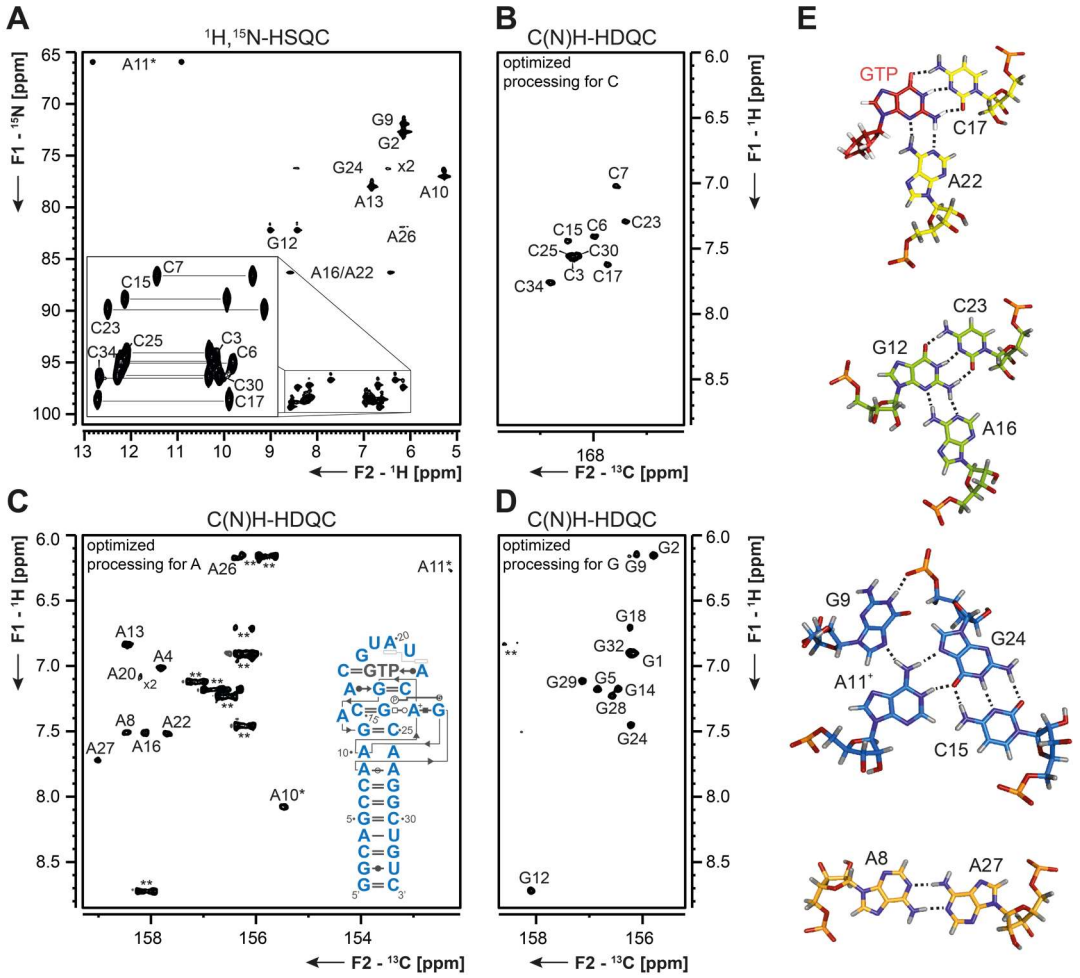


Figure 16 A ^{1}H,^{15}N-HSQC and **B, C, D** C(N)H-HDQC spectra of the 34 nts GTP class II aptamer RNA[107,108]. The ^{1}H,^{15}N-HSQC spectrum was recorded with spectral windows of 24 ppm and 40 ppm in the direct ^{1}H- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 4.7 ppm, 150 ppm and 85 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Complex points were 2048 and 128 in the ^{1}H- and ^{15}N-dimensions, respectively. With 8 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 0.7 h. Spectra in **B**, **C** and **D** were processed separately for cytidines **B**, adenosines **C** and guanosines **D** but are originating from the same dataset. The spectrum was recorded with spectral windows of 50 ppm and 5 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 154 ppm, 7.1 ppm and 90 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 160 in the ^{13}C- and ^{1}H-dimensions, respectively. With 80 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 19 h. Signals marked with an asterisk are folded into the spectrum. Signals marked with "**" in **C** and **D** arise from processing and can be attributed to guanosines and adenosines, respectively. The secondary structure of the GTP class II aptamer is shown in **C** where arrows mark the backbone direction from 5' to 3' and nucleobase interactions are annotated using the Leontis and Westhof nomenclature[123]. **E** Different structural elements of the 34 nts GTP class II aptamer[107,108]. The figure has been adapted from Schnieders *et al.*[78].

By contrast, a full set of resonances for all amino group bearing nucleotides is observed in the ^{13}C-detected C(N)H-HDQC spectrum (Figure 16 B, C and D). Furthermore, the resonances show a significant deviation in ^{1}H-DQ chemical shifts, which might be correlated to the respective mode of interaction, as residues involved in Watson-Crick like interactions seem to be clustered (G5, G28 and G29), while other residues, which are involved in unusual interactions like G12, are strongly shifted.

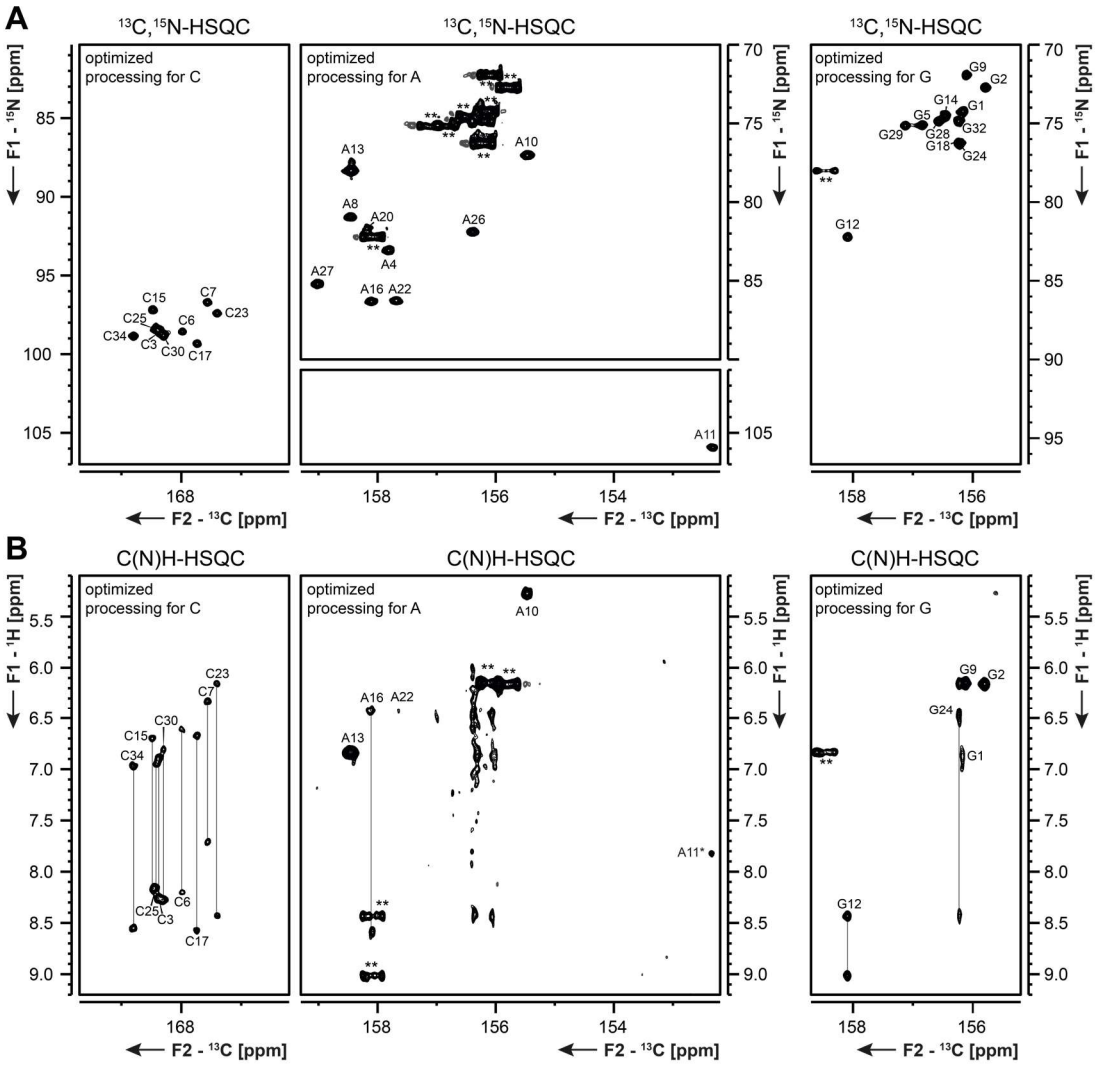


Figure 17 A ^{13}C,^{15}N-HSQC and **B** ^{13}C-detected C(N)H-HSQC spectra of the 34 nts GTP class II aptamer[107,108]. Spectra were processed differently for cytidines, adenosines and guanosines. The spectrum in **A** was recorded with spectral windows of 50 ppm and 43 ppm in the direct ^{13}C- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 160 ppm, 4.7 ppm and 87 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 128 in the ^{13}C- and ^{15}N-dimensions, respectively. With 48 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 9.3 h. The spectrum in **B** was recorded with spectral windows of 50 ppm and 5 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 166 ppm, 7.1 ppm and 87 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 160 in the ^{13}C- and ^{1}H-dimensions, respectively. With 80 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 19.3 h. The figure has been adapted from Schnieders *et al.*[78].

In order to further assess the impact of the RNA's structure on the ^{1}H-DQ chemical shifts, the C(N)H-HDQC experiment was applied to three additional RNAs; a 14 nts hairpin RNA with GAAG tetraloop[109], the 43 nts SAM/SAH riboswitch[110,111] and a 329 nts CUG repeat RNA[112,113] (spectra in Appendix 1 Figure 61, Figure 62 and Figure 63). After calculation of a mean value of the ^{1}H-DQ chemical shift for all GC- and AU-Watson-Crick base pairs over all RNAs, the absolute difference in the ^{1}H-DQ chemical shift from this mean value was plotted for each nucleotide (Figure 18). Even though this study was only based on a limited number of RNAs, it is evident that canonical and non-canonical interactions can be distinguished to a certain extend only based on the chemical shift information (Figure 18). This discrimination becomes very clear in both 14 nts RNAs, where the tetraloop

^{1}H DQ chemical shifts strongly deviate from the chemical shifts of nucleotides embedded in the stem (Figure 18 A and C). For other RNAs like the 34 nts GTP class II aptamer, the discrimination between canonical and non-canonical interactions is not always possible, as is especially observed for non-canonical interactions in adenosines (e.g. A16, A20, Figure 18 B). This effect might arise from an underrepresentation of adenosines involved in canonical interactions, from which the mean value is calculated. In the RNAs under study only three adenosines are involved in AU Watson-Crick base pairs. Similarly, non-canonical interactions for cytidines are not particularly covered in this study, as only one of the cytidine residues is not involved in Watson-Crick like interactions. The overview in Figure 18 F shows that the discrimination is most reliable for guanosines, which can be attributed to the fact that experimental data are available to a greater extend for both, canonical and non-canonical interactions.

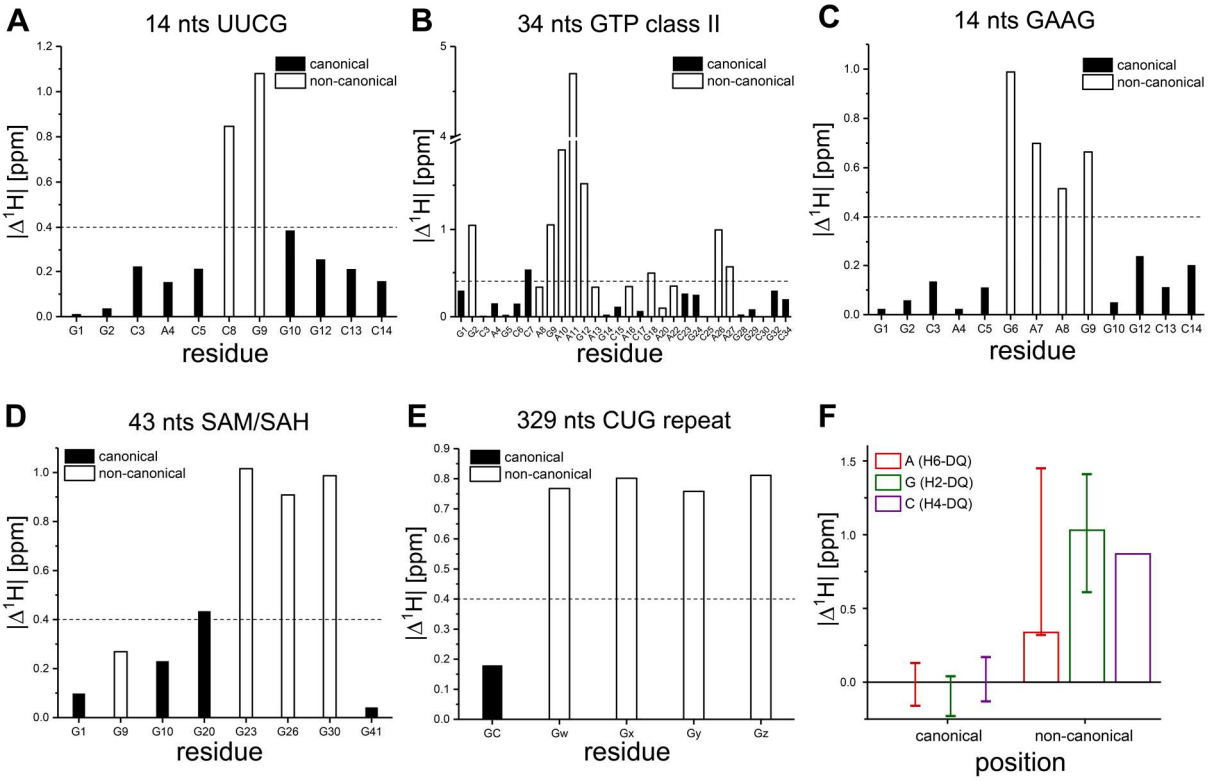


Figure 18 Absolute difference in ^{1}H-DQ chemical shifts for nucleotides involved in canonical and non-canonical nucleobase interactions for **A** the 14 nts RNA with UUCG tetraloop[105,106], **B** the GTP class II aptamer[107,108], **C** the 14 nts RNA with GAAG tetraloop[109], **D** the 43 nts SAM/SAH-binding riboswitch[110,111] and **E** the 329 nts CUG repeat RNA with guanosine rich 5' and 3' overhangs[112,113]. The mean values for chemical shifts of canonical nucleobase interactions have been calculated over all RNAs for cytidines, adenosines and guanosines separately. The dashed line represents two times the standard deviation. **F** Absolute ^{1}H-DQ chemical shift difference for canonical and non-canonical nucleobase interactions over all RNAs for adenosines, guanosines and cytidines. The figure was adapted from Schnieders *et al.*[45,78].

As this chemical-shift-to-structure relation is only conducted with experimental data of five different RNAs, the results are preliminary. However, the data indicate that amino proton chemical shifts have a high potential for a chemical shift based structural prediction, if the number of RNAs under study is increased through application of the ^{13}C-detected C(N)H-HDQC experiment.

2.3.4 Application of the ^{13}C-detected "amino"-NOESY experiment

Apart from the previously described method to obtain indirect structural information through chemical shift information, amino groups were also brought in direct structural context with the ^{13}C-detected "amino"-NOESY experiment developed here. Again, the experiment was first implemented using the 14 nts RNA with UUCG tetraloop (Figure 19 A, gray). For base-paired nucleotides, the unmodified "amino"-NOESY experiment typically yields NOE contacts between the amino groups and the imino protons, which are involved in the corresponding hydrogen bond interaction. Furthermore, correlations that exhibit very broad lines are observable. As these correlations are the SQ-diagonal peaks, their information content is not particularly high. Moreover, these diagonal peaks come to resonance in a region, where interesting NOE contacts, like e.g. NH_2 to H1' protons are expected. For this reason, the undesired SQ diagonal peaks are removed in an alternative version of the "amino"-NOESY experiment (Figure 19 A, red). Here, a ^{13}C-filter prior to the NOE transfer selects for protons, which are bound to ^{13}C-nuclei, while all other correlations do not give rise to a signal. This enables detection of additional NOE contacts, as the resonance overlap is reduced through the elimination of the broad SQ diagonal peaks.

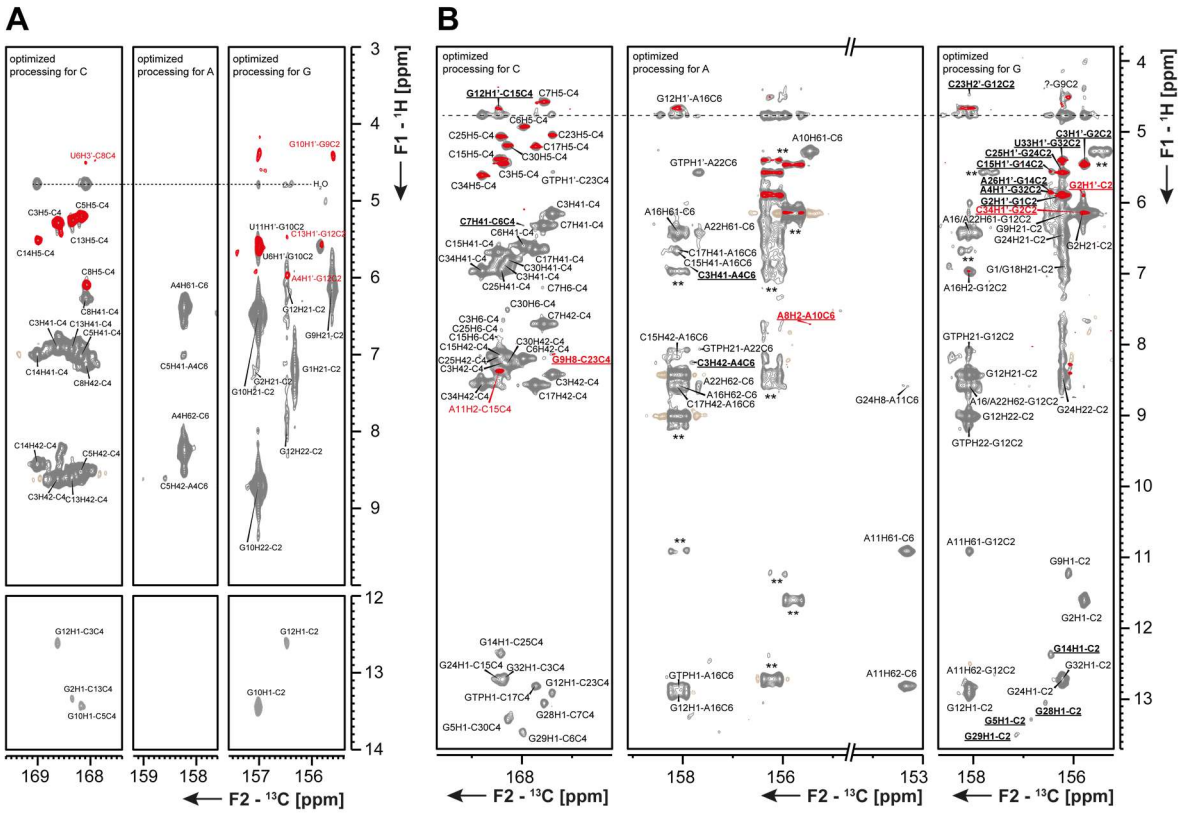


Figure 19 Overlay of ^{13}C-detected "amino"-NOESY spectra without (gray) and with ^{13}C-filter (red) of **A** the 14 nts RNA featuring UUCG tetraloop[105,106] and **B** the GTP class II aptamer RNA[107,108]. Spectra are processed differently for cytidines, adenosines and guanosines. The water resonance is marked with a dashed line and pseudo doublets annotated with "**" appear from processing. NOE contacts, which are not observed in conventional ^{1}H-detected NOESY experiments are written in bold and underlined font. For experimental details see Appendix 1 chapter 5.1.1.

After having successfully implemented both NOESY experiments using a small hairpin RNA, they were further applied to the 34 nts GTP class II aptamer[107,108] (Figure 19 B). Overall, 22 additional NOE contacts could be detected, which were not observed using conventional ^{1}H-detected NOESY experiments, including the CPMG-NOESY experiment, which represents the state-of-the-art NOESY experiment for amino groups. Here, it is interesting that those new NOE contacts are often inter-residual amino-group-to-H1' correlations, which are especially valuable, as inter-residual contacts are scarce in RNA due to the low proton density in the nucleobases of RNA. In comparison to the CPMG-NOESY experiment, the observed benefits of the "amino"-NOESY experiment lie mainly in the detection of H1' contacts. This might be a confirmation for the hypothesis that advantages can result from the possibility to omit water suppression schemes in ^{13}C-detected NMR experiments.

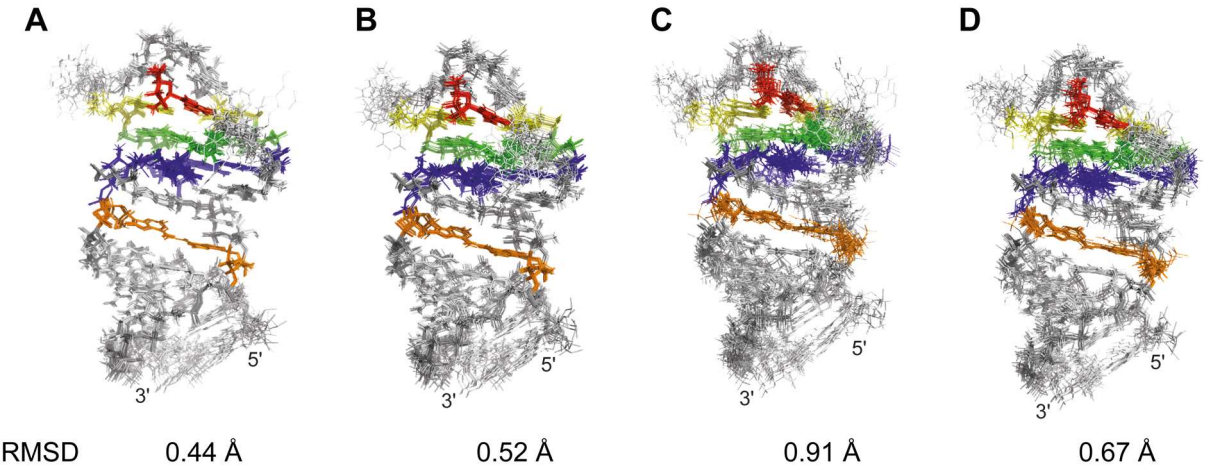


Figure 20 Bundles of the 10 structures with the lowest target function of the 34 nts GTP class II aptamer with the corresponding RMSD value. Structure calculations contain: **A** the original set of data, **B** the original set of data plus the 22 newly determined NOE contacts, **C** the reduced set of data, where torsion angle restraints are omitted for non-helical regions, and **D** the reduced set of data plus the 22 newly determined NOE contacts. The colors of the different base pairing interactions are matching the structures in Figure 16 E.

As a high-resolution NMR structure was available for the 34 nts RNA, the effect of the newly determined NOE contacts on the structure's precision was studied. Hence, Antje C. Wolter from the lab of Prof. Jens Wöhnert conducted structure calculations with and without the newly determined amino NOE contacts. Here, it was observed that the mere inclusion of the NOE contacts did not lead to an improvement in the structure's precision (Figure 20 A and B). This can be attributed to the fact that the original structure determination was already conducted on the basis of a plethora of experimental data that includes on average 25 distance restraints per residue from NOE data. On top of that, several torsion angle restraints are available for this system (60 ribose puckers, 27 glycosidic bond angles, 107 backbone angles). As this vast amount of data is not necessarily accessible for less structured RNAs, a lack of data was simulated by repeating structure calculations without torsion angle restraints in non-helical parts of the RNA. Here, the inclusion of the NOE data from previously exchange broadened amino groups lead to a significant improvement in the structure's RMSD value from 0.91 Å to 0.67 Å

(Figure 20 C and D). Most likely, this effect can be ascribed to the inter-nucleotide character of the new NOE data. The "amino"-NOESY experiment is thus expected to be especially valuable for systems, where the amount of accessible experimental data is limited, and which e.g. exhibit flexible regions, where NOE data is scarce.

2.4 Conclusion and Outlook

Within this PhD project, two ^{13}C-detected NMR experiments have been developed in order to exploit the full potential of an amino-based structural characterization of RNA. In comparison to the experimental gold-standard in NMR spectroscopy of RNA, the imino proton-based characterization, the characterization of amino groups exhibits several advantages: Due to the considerably slow chemical exchange of amino protons with solvent water[82], amino resonances can be observed not only for flexible regions of RNA but also for a broad temperature range. These are already two advantages of major impact towards e.g. structural studies in living cells. Additionally, amino groups are embedded in three out of four nucleotides while imino groups are only found in guanosines and uridines. Together with fact that, as opposed to imino groups, amino groups can also participate in sugar-edge or Hoogsteen interactions, an amino-based structural analysis has the potential to evolve into a powerful characterization routine. With the ^{13}C-detected C(N)H-HDQC experiment developed here, a full set of sharp amino resonances can be detected; a feature, which is unachieved with any other NMR experiment due to a rotational exchange in the intermediate exchange regime. When comparing the sheer amount of imino and amino resonances, which are observed for a 14 nts RNA in an imino ^{1}H,^{15}N-HSQC and in an amino C(N)H-HDQC experiment, it is evident that this number is increased for the latter (Figure 21). Here, resonances for 11 out of 14 nucleotides are observed, while only five signals appear in the imino ^{1}H,^{15}N-HSQC spectrum. This simple comparison demonstrates the future impact, which might be achieved through the application of the C(N)H-HDQC experiment.

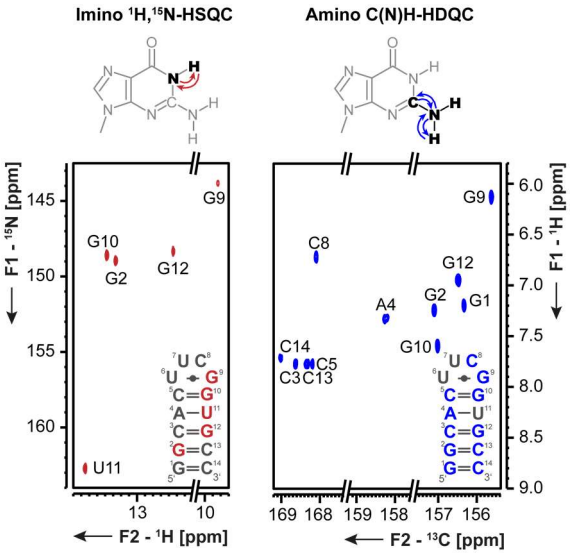


Figure 21 Comparison of a ^{1}H,^{15}N-HSQC spectrum (left) and a C(N)H-HDQC spectrum (right) of the 14 nts RNA with UUCG tetraloop. Nucleotides for which a resonance is observed in the corresponding spectrum are written in red (left) or blue (right) in the secondary structure. The coherence transfer pathways are indicated at the top. The figure has been adapted from Schnieders *et al.*[78].

Using the newly obtained chemical shift information, preliminary structural predictions can be undertaken based on the a chemical-shift-to-structure relation. Here, the amino

proton chemical shifts are especially sensitive towards hydrogen bonding interactions, as they directly participate in the hydrogen bonds of multiple canonical and non-canonical nucleobase interactions[81]. This approach has the potential to be developed into a sophisticated tool in structural biology. However, a sufficient amount of experimental data has to be gathered beforehand by the application of the C(N)H-HDQC experiment. Over the last decades, chemical shift based analyses like this are already commonly applied in protein research[124–127] and resulted in powerful programs like TALOS[128,129]. As opposed to proteins, chemical shift based structural predictions for RNA are still in their infancy and need to be further developed in order to perform reliably. In this field, the amino proton DQ chemical shifts can most certainly contribute if the C(N)H-HDQC experiment is followed up on and chemical shift data is deposited in the BMRB database[72]. The challenges in this proposition will, among other things like the amount of data needed, most certainly lie in the dependence of the proton chemical shifts on buffer conditions (salt content, pH value) or the temperature. Similar complications were observed and corrected for other nuclei like non-exchanging protons[130] or carbon atoms[131]. Comparable correction methods would have to be applied to the amino DQ chemical shifts in order to be able to use the chemical shift information.

Besides this newly obtained chemical shift information, amino groups were also brought in direct structural context with the ^{13}C-detected "amino"-NOESY experiment developed here. This experiment is not based on the indirect detection of DQ coherence, but on the effective spin-lock of proton SQ coherence in a TOCSY transfer that yields the correlation of amino groups to protons within a distance of 5 Å. The comparison of this experiment with conventional NOESY and state-of-the-art CPMG NOESY experiments yielded 22 previously not observable NOE contacts, which were shown to significantly improve structure calculations of a 34 nts RNA. Due to the inter-residual character of most of these new NOE correlations, the "amino"-NOESY experiment is expected to have major impacts where inter-nucleotide contacts between the nucleobase and the ribose backbone are rare, as is for example the case for nucleotides with an imino proton resonance that is broadened beyond detectability.

Together, both experiments proposed here promote amino protons as an additional standard reporter for nucleobase interactions in nucleic acids. Due to this high potential of amino-based structural analysis of RNA, first experiments were conducted in order to lay foundations for potential follow-up projects. Therefore, the implementation of a ^{13}C-detected zero quantum C(N)H-HZQC experiment and the characterization of RNA in the ligand-free apo state in order to characterize a potential preformation of the ligand binding pocket are being outlined in the following paragraphs.

2.4.1 From DQ to ZQ

To begin with, instead of DQ coherence, ZQ coherence could be detected in a ^{13}C-detected C(N)H-HZQC experiment. With the combination of DQ and ZQ experiments the SQ chemical shifts could potentially be back calculated. This approach of combining information from HDQC and HZQC experiments was already demonstrated in N_z exchange experiments and resulted in valuable complementary datasets[104].

The selection of the ZQ component is easily achieved by adapting the receiver phase in a way that the magnetization modulated with the ZQ chemical shift is accumulated:

$$8C_zN_yH_z^1H_z^2 \cos \Omega_{ZQ}t_1 + 8C_zN_yH_z^1H_z^2 \cos \Omega_{DQ}t_1$$

The resulting spectrum (Figure 22 A) shows axial peaks wherever the SQ proton chemical shifts are identical ($\Delta\Omega = 0$) like for residue G9 of the 14 nts RNA with UUCG tetraloop. G10 exhibits two resonances in HSQC spectra (Figure 6 A) and yields signals at $+\Delta\Omega$ and $-\Delta\Omega$ in the HZQC spectrum. From the peak intensities it is, however, evident that the HZQC experiment is not independent of the rotational exchange around the C-NH_2 bond, as the intensity profile fits the one obtained in SQ experiments, where several resonances are broadened beyond detectability (Figure 22 B). This can be attributed to the fact that the ZQ chemical shift is not constant, while the protons continuously interchange into each other ($\Omega_1 - \Omega_2 \neq \Omega_2 - \Omega_1$) and thus the HZQC approach was not further followed up on.

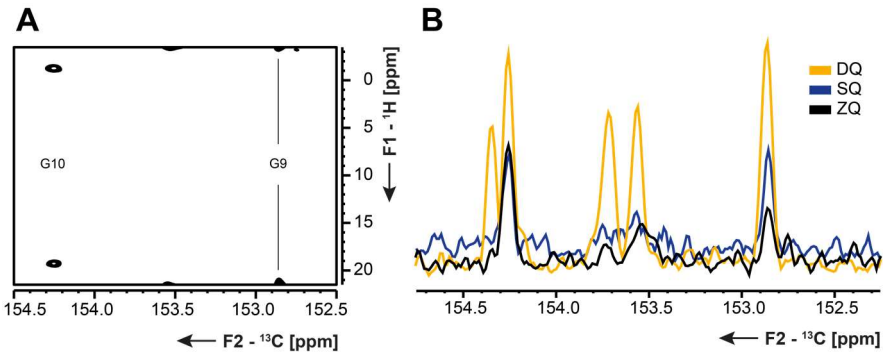


Figure 22 A ^{13}C-detected C(N)H-HZQC spectrum of the 14 nts RNA with UUCG tetraloop. The spectrum was recorded with spectral windows of 50 ppm and 25 ppm in the direct ^{13}C- and indirect ^{1}H-dimension, respectively. Carrier frequencies were 154 ppm, 9.0 ppm and 90 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.1 s in the direct and 0.002 s in the indirect dimensions. With an inter-scan delay of 1.8 s and 16 scans per increment, the experiment was recorded within 40 min. **B** Overlay of positive projections of DQ, ZQ and SQ spectra.

2.4.2 The C(N)H-HDQC experiment for the characterization of apo states

A further field of application of the ^{13}C-detected C(N)H-HDQC experiment could be the study of aptamers in their ligand-free (apo) states. Those states are often not characterized in detail due to their dynamic nature, which makes imino-based assignments not accessible. Their study would allow a closer inspection of potential preformation of the ligand binding pocket or a mapping of structural changes that the RNA has to undergo upon ligand-binding. As such a preformation was observed for the

34 nts GTP class II aptamer RNA[108], but has not been further characterized, this system would be an interesting target. Therefore, preliminary experiments have been conducted with the RNA in buffer and in the presence of 2 mM Mg^{2+} ions, as preformation has previously observed upon addition of Mg^{2+} ions[108].

From the comparison of ^{1}H,^{15}N-HSQC spectra it is evident that addition of Mg^{2+} ions indeed induces a preformation of the ligand binding pocket as several reporter resonances for the complex interactions, e.g. G12 or A11, appear (Figure 23). Other resonances, which are marked with a purple box, are neither belonging to the ligand-bound state nor to the Mg^{2+}- and ligand-free state and are thus particularly interesting.

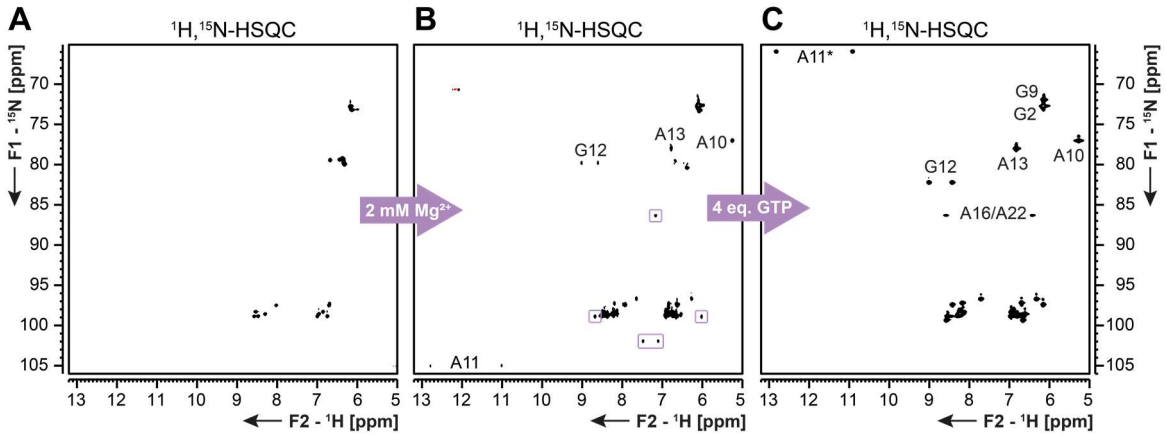


Figure 23 ^{1}H,^{15}N-HSQC spectra of the GTP class II aptamer RNA in the free form (**A**), in the presence of 2 mM Mg^{2+} ions (**B**) and in the presence of Mg^{2+} ions and ligand GTP (**C**). The spectrum in **A** was recorded with spectral windows of 24 ppm and 40 ppm in the direct ^{1}H- and in the indirect ^{15}N-dimensions. Carrier frequencies were 4.7 ppm (^{1}H), 150 ppm (^{13}C) and 85 ppm (^{15}N). The acquisition times were 0.1 s in the direct and 0.04 s in the indirect dimension. With 16 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 1.5 h. The spectrum in **B** was recorded with spectral windows of 24 ppm and 44 ppm in the direct ^{1}H- and in the indirect ^{15}N-dimensions. Carrier frequencies were 4.7 ppm (^{1}H), 150 ppm (^{13}C) and 85 ppm (^{15}N). The acquisition times were 0.1 s in the direct and 0.04 s in the indirect dimension. With 4 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 20 min. Experimental details for the spectrum in **C** are given in Figure 16.

In comparison to the ^{1}H,^{15}N-HSQC spectra, the ^{13}C-detected ^{13}C,^{15}N-HSQC spectra again provide an increased number of resonances (Figure 24). This effect is especially observed for resonances of guanosines and adenosines. However, in order to be able to track the changes and bring them in structural context, the resonances of the RNA in ligand-free form in the absence and in the presence of Mg^{2+} ions have to be assigned. This can be achieved by HNCO type experiments for guanosines, the TROSY relayed HCCH-COSY experiment[39] for adenosines and ^{1}H,^{1}H-NOESY experiments for cytidines. Chemical shift assignment can be further complemented by ^{13}C,^{15}N-HSQC spectra that correlate the carbon nucleus attached to the amino group to neighboring nitrogen nuclei. This resonance assignment was not conducted within this PhD thesis and is subject to further studies.

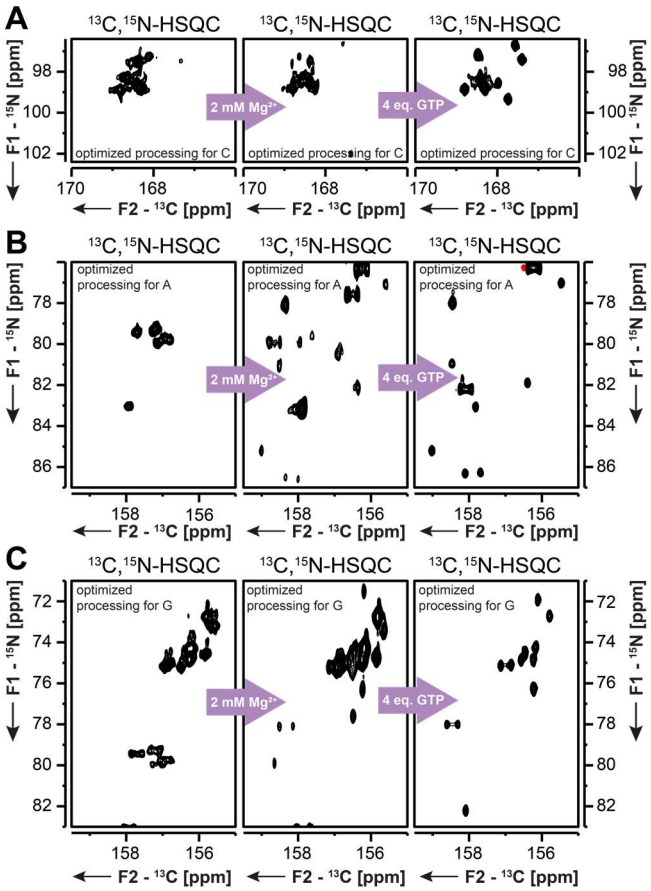


Figure 24 ^{13}C,^{15}N-HSQC spectra of the 34 nts GTP class II aptamer RNA of the amino region for **A** cytidines, **B** adenosines and **C** guanosines in the free form (left), in the presence of 2 mM Mg^{2+} ions (middle) and in the presence of Mg^{2+} and ligand GTP (right). The spectra have been processed differently for **A**, **B** and **C**. Both experiments in the left and middle panel were recorded with spectral windows of 50 ppm (^{13}C) and 32 ppm (^{15}N). Carrier frequencies were set to 160 ppm, 4.7 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.05 s for the indirect dimension. With an inter-scan delay of 2.5 s and 80 scans per increment, the experiment in the left panel was recorded within 15.5 h. With an inter-scan delay of 2.5 s and 128 scans per increment, the experiment in the middle panel was recorded within 24.5 h. Experimental details for spectra in the right panel are depicted in the figure caption of Figure 17.

Despite the lack of chemical shift assignment, the ^{13}C-detected C(N)H-HDQC experiment was conducted for the 34 nts GTP class II aptamer RNA in the absence and presence of Mg^{2+} and GTP (Figure 25) to estimate the amount of additional information, which can potentially be obtained with this experiment. Due to a poor S/N ratio, the number of resonances is reduced in comparison to the ^{13}C,^{15}N-HSQC spectrum for the ligand-free RNA. This is especially prominent for cytidines and adenosines. On the other side, for guanosines almost all of the expected resonances are obtained in the DQ spectrum, which can be attributed to the fact that the experiment yields generally the highest S/N ratio for guanosine residues. In comparison to the ^{1}H,^{15}N-HSQC experiments, additional ^{1}H chemical shift information can still be acquired. Therefore, the detailed characterization of the preformation in the GTP class II aptamer RNA can certainly benefit from the application of the ^{13}C-detected C(N)H-HDQC experiment in future. To tackle this challenging task of being able to track structural changes while a mixture of states is present, the S/N ratio of the C(N)H-HDQC experiment needs to be maximized. This can

be achieved by recording the experiment separately for each nucleotide type (cytidines, adenosines and guanosines) with the respective optimal parameters for scalar coupling constants instead of recording one experiment with averaged delays and pulses.

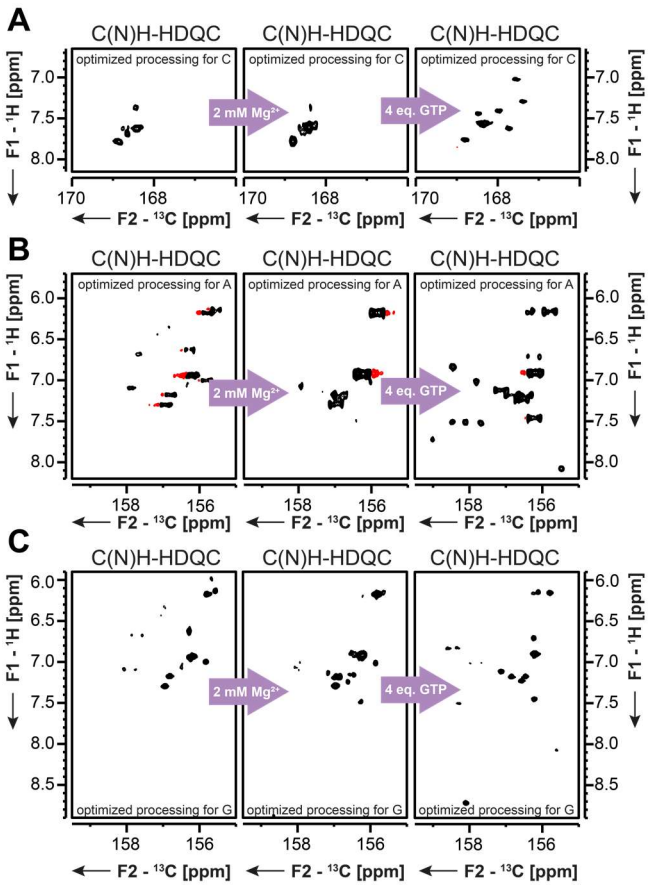


Figure 25 ^{13}C-detected C(N)H-HDQC spectra of the 34 nts GTP class II aptamer RNA of the amino region for **A** cytidines, **B** adenosines and **C** guanosines in the free form (left), in the presence of 2 mM Mg^{2+} ions (middle) and in the presence of Mg^{2+} and ligand GTP (right). The spectra have been processed with the optimized processing parameters for **A**, **B** and **C**. Experimental details for spectra in the left and middle panel can be taken from Appendix 1 chapter 5.1.2. Experimental details for spectra in the right panel are depicted in the figure caption of Figure 17.

Chapter III: ^{15}N-detected H-N correlation experiments for RNA

The results presented in this chapter are summarized in the following publication[132]:

R. Schnieders, C. Richter, S. Warhaut, V. de Jesus, S. Keyhani, E. Duchardt-Ferner, H. Keller, J. Wöhnert, L. T. Kuhn, A. L. Breeze, W. Bermel, H. Schwalbe and B. Fürtig, *J. Biomol. NMR* **2017**, *69*, 31–44.

The fundamentals of this project have already been laid in the course of my master thesis. However, the main part that comprises of the optimization of the ^{15}N-detected BEST-TROSY pulse sequence discussed here, the application of this enhanced pulse sequence to the RNAs under study, the final line width predictions, the evaluation of the magnetic field dependence and the manuscript writing were conducted during my PhD thesis.

Author contributions are listed below:

R. Schnieders was involved in project design, conducted all experiments, was involved in pulse sequence optimization and in manuscript writing.

S. Warhaut, V. de Jesus, S. Keyhani, E. Duchardt-Ferner, H. Keller and J. Wöhnert provided NMR samples and were involved in manuscript writing.

L. T. Kuhn and A. L. Breeze provided access to their 950 MHz NMR spectrometer, where experiments for the magnetic field dependence were conducted and were involved in manuscript writing.

C. Richter and W. Bermel helped with pulse sequence optimization and were involved in manuscript writing.

H. Schwalbe and B. Fürtig designed the project were involved in manuscript writing.

This work was further outlined in the following minireview article[45]:

R. Schnieders, S. Keyhani, H. Schwalbe and B. Fürtig, *Chem. - A Eur. J.* **2020**, *26*, 102–113

Additionally, a detailed guide on how to set-up the ^{15}N-detected BEST-TROSY experiment proposed here has been submitted:

R. Schnieders, B. Knezic, H. Zetsche, A. Sudakov, T. Matzel, C. Richter, M. Hengesbach, H. Schwalbe and B. Fürtig, NMR spectroscopy of large functional RNAs - from sample preparation to low-gamma-detection, *Curr Protoc Nucleic Acid Chem* **2020**, *Manuscript submitted*.

3.1 Introduction

While NMR spectroscopy is undoubtedly a powerful method to study RNA, there are certain challenges in the field, which have to be met in the coming decades. One of the most critical aspects is the limitation in molecular size of the RNA under study due to resonance overlap. The reasons for this signal overlap are: (i) increased number of resonances, (ii) poor signal distribution due to chemically similar building blocks and (iii) broad lines for larger RNAs due to increased relaxation rates. With the continuous development of higher magnetic field NMR spectrometers and advances in elaborate sample preparation by segmental labeling[133] or position-specific modification of RNA[59], the size limitation is pushed further towards the characterization of higher molecular weight RNAs[20,46,65] or RNA-protein complexes[16]. Furthermore, this progress can be supported by moving towards the detection on slower relaxing nuclei, like ^{13}C or ^{15}N, and thus attenuating the line broadening for larger RNAs. These heteronuclear-detection schemes are nowadays feasible, as there has been a constant development in cryogenic probes as well as probes optimized for ^{13}C- or ^{15}N-detection[30], which compensate for the reduced sensitivity due to the lower gyromagnetic ratios.

As opposed to carbon-direct detection, which is widely applied to biomolecules since 2000[134] and ever since increasing in importance, the field of nitrogen-direct detection was only introduced quite recently in multidimensional biomolecular NMR spectroscopy. Pioneers on this field are Koh Takeuchi and Gerhard Wagner, who have developed several ^{15}N-detected NMR experiments for proteins, starting with ^{15}N-detected CAN and CON experiments for the resonance assignment of the protein backbone in 2010[135]. Those experiments exploit the favorable relaxation properties of ^{15}N-nuclei and avoid complicated decoupling schemes for homonuclear CC-decoupling (IPAP/DIPAP) as necessary for the respective ^{13}C-detected experiments. With this they found that the median signal-to-noise ratio (S/N ratio) for the ^{15}N-detected CAN and CON experiments is higher in comparison to their ^{13}C-detected counterparts, even though it should be reduced by a factor of 4, when only the gyromagnetic ratios are considered. ^{15}N-direct detection schemes for proteins were then followed up with the introduction of hCAN[136] and TROSY-HSQC experiments[137,138]. Especially the latter yielded promising results due to the very sharp resonances in the ^{15}N-dimension and the fact that protein deuteration is not necessarily required. This enables the study of large proteins which are not deuterated easily. Furthermore, it was shown that the ^{15}N-detected TROSY experiment is less affected by high (physiological) salt concentrations when compared to the ^{1}H-detected equivalent. This leads to an almost comparable sensitivity of the ^{1}H- and the ^{15}N-detected TROSY experiments for a large, non-deuterated protein at high salt concentrations[138].

Prompted by these results from Takeuchi and coworkers for proteins, we wanted to test applicability and suitability of ^{15}N direct-detection on RNA. In the beginning of this thesis, multidimensional ^{15}N-detected NMR experiments were not yet applied to RNA. With the

application of different ^{15}N-detected H-N correlation experiments to RNA and the investigation of the effect of molecular size on sensitivity, resolution and relaxation behavior, this has been achieved in the course of the presented study[132]. The present chapter describes the application of a ^{15}N-detected HSQC, a TROSY and a BEST-TROSY experiment to RNAs ranging from 14 to 329 nts. A thorough line widths analysis reveals that resonances in the ^{15}N-dimension stay sharper than the corresponding resonances in the ^{1}H-dimension with rising molecular size. In addition to the effect of molecular size, also the influence of base pairing interactions, solvent exchange, magnetic field strength and conformational equilibria on the resolution are analyzed. With the TROSY experiment as the central element in this study, its underlying principles are described in the following paragraphs.

3.1.1 TROSY in an N-H spin system

In NMR spectroscopy of large biomolecules, in particular of proteins, transverse relaxation-optimized spectroscopy (TROSY)[38] is frequently used to conduct H-N correlation experiments and is particularly favorable for large molecules at high magnetic fields. As it is described in the following, these experiments exploit cross-correlation between dipolar relaxation and relaxation through chemical shift anisotropy (CSA). This cross-correlation can physically be explained through the fact that both relaxation mechanisms are dependent on molecular tumbling. This common ground inevitably entails a certain degree of correlation. This correlation can either lead to a compensation or a reinforcement of the relaxation of a certain nucleus. Whether the relaxation is compensated or reinforced is dependent on the spin state so that a nucleus, such as ^{15}N, relaxes differently depending on the spin state of the coupled nucleus, e.g. H^{α} and H^{β}. This difference in relaxation behavior consequently leads to different line widths for the peaks in a coupled doublet, which would be observed in a coupled ^{1}H,^{15}N-HSQC spectrum, as shown schematically in Figure 26 A. As this interference is present for both, the ^{15}N and ^{1}H nuclei, the following properties are observed for all four resonances: One peak is broad in both dimensions (1), two resonances are narrow in one and broad in the other dimension (2+3) and another one is narrow in both dimensions (4). This peak (4) is called the TROSY-peak, while the peak that relaxes quickly in both dimensions (1) is usually referred to as the anti-TROSY peak. If decoupling is applied in both dimensions, all line widths are averaged which results in the familiar decoupled ^{1}H,^{15}N-HSQC spectrum (Figure 26 B). However, if the TROSY-peak (1) would be selected (Figure 26 C) and would then be shifted about $1/2^{1}J$(N,H) in both dimensions to the correct position, a similar spectrum would result exhibiting signals with narrower lines than in the ^{1}H,^{15}N-HSQC spectrum. This is realized in TROSY experiments. One might expect that this selection of only one out of four resonances would lead to a decrease in sensitivity by a factor of four. However, when it is considered that in a Lorentzian line the intensity of a

resonance is correlated with its peak width, as shown in equations [1]-[3], the loss of sensitivity is compensated depending on the degree of cross-correlation.

$$S(\omega) = \frac{L}{((\omega - \Omega)^2 + L^2)} \quad [1]$$

$$S(\Omega) = \frac{1}{L} \; ; \; \omega_{1/2} = 2L \quad [2],[3]$$

$S(\omega)$ = signal at frequency ω, Ω = chemical shift, $S(\Omega)$ = signal height at chemical shift Ω, ω = frequency, $\omega_{1/2}$ = peak width at half height in rad s^{-1}

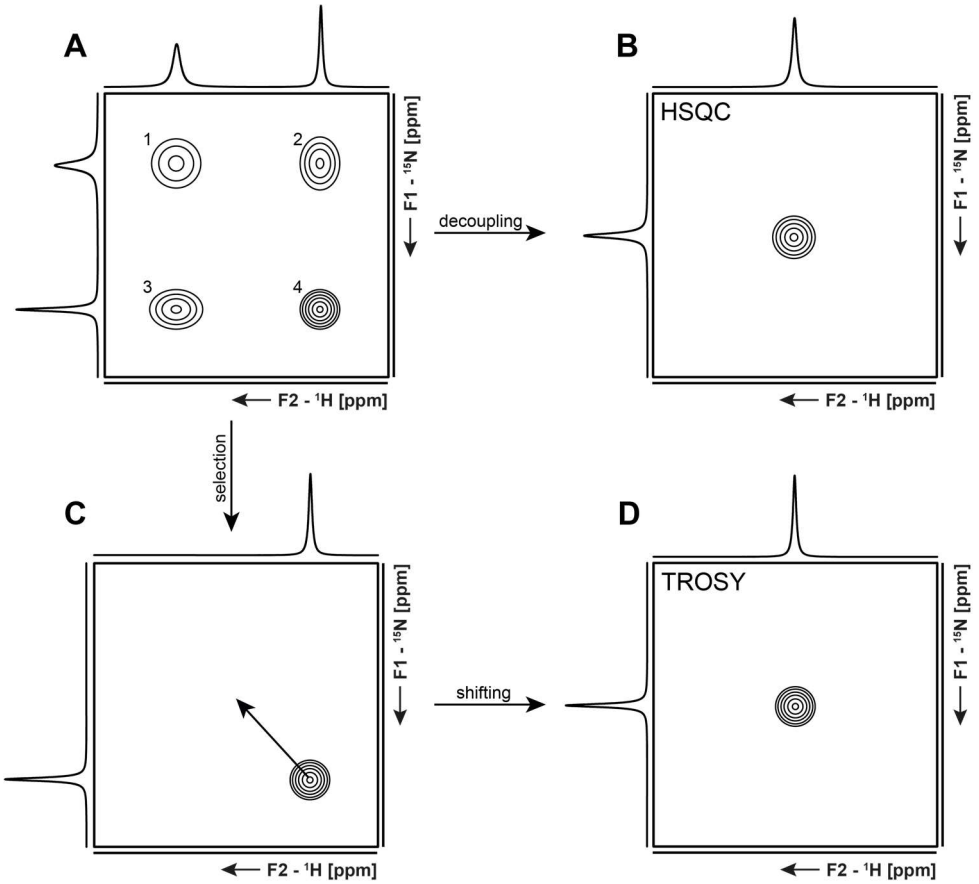


Figure 26 The TROSY effect for an N-H spin system in schematic ^{1}H,^{15}N-correlation spectra. **A** Coupled schematic ^{1}H,^{15}N-HSQC spectrum with constructive or destructive interference of relaxation through chemical-shift-anisotropy (CSA) and dipolar interactions leading to the anti-TROSY peak (1), both semi-TROSY peaks (2+3) and the TROSY-peak (4). **B** In both dimensions decoupled schematic ^{1}H,^{15}N-HSQC spectrum that leads to an averaged line width of the signal. **C** Schematic ^{1}H,^{15}N-HSQC spectrum where the TROSY-peak (4) is selected through phase cycling. **D** Schematic ^{1}H,^{15}N-TROSY-HSQC spectrum, where the TROSY peak is shifted by $1/2^{1}J_{NH}$ in both dimensions to the correct peak position.

The selection of the TROSY-peak can be achieved by e.g. echo-antiecho gradient selection or single transition-to-single transition polarization transfer (ST2-PT)[139]. The latter is very similar to the preservation of equivalent pathways (PEP)[140] and relies on the combination of two separately recorded FIDs, where the phases of certain pulses are shifted so that a given combination only yields the desired transition. Furthermore, this experiment can be conducted with band-selective proton pulses, which leave the water resonance unperturbed along the z-axis. This allows a reduction of the inter-scan delay,

as the T_1 relaxation is enhanced. Therefore, this approach is named band-selective excitation short transient (BEST).

As it is crucial for this work, the following paragraphs will describe the relaxation behavior of the TROSY-peak in the ^{15}N- and ^{1}H-dimension. The ^{1}H- and ^{15}N-TROSY transverse relaxation rates R_2^{TROSY} are generally described as the sum of relaxation contributions through dipolar relaxation R_2^{dip}, relaxation through CSA R_2^{csa} and relaxation through remote protons R_2^{HH}. This is then reduced by the destructive interference of relaxation through CSA and dipolar relaxation R_2^{int} as described above:

$$R_{2H}^{TROSY} = R_{2H}^{dip} + R_{2H}^{csa} + R_{2H}^{HH} - R_{2H}^{int} \quad [4]$$

$$R_{2N}^{TROSY} = R_{2N}^{dip} + R_{2N}^{csa} + R_{1N}^{HH} - R_{2N}^{int} \quad [5]$$

From these relations it can be easily extracted that the larger the R_2^{int} term gets, the smaller will be the overall relaxation rate, which results in an increased resolution. In order to identify the crucial parameters, the individual contributions to the transverse relaxation of TROSY components are themselves described in equations [6]-[15].

The transverse relaxation of a spin I through dipolar interactions is affected by the spectral density at different frequencies. This also involves the spectral density of spin S as the relaxation of spin I is dependent on the spin state of spin S. Furthermore, this relaxation rate decreases rapidly with increasing distances between spins I and S ($\sim 1/r_{IS}^6$).

$$R_{2I}^{dip} = \frac{p_{IS}^2}{8}[4J(0) + 3J(\omega_I) + J(|\omega_I - \omega_S|) + 6J(\omega_S) + 6J(\omega_I + \omega_S)\,] \quad [6]$$

R_{2Idip} = transverse relaxation of spin I due to dipolar interactions, J(ω) = spectral density, $\omega_{I/S}$ = Larmor frequency of spins I/S

The relaxation through CSA increases with increasing static magnetic field strengths B_0:

$$R_{2I}^{csa} = \frac{\delta_I^2}{6}[4J(0) + 3J(\omega_I)] + \frac{\delta_S^2}{2}\,J(\omega_S) \quad [7]$$

R_{2Icsa} = transverse relaxation of spin I due to chemical shift anisotropy

Furthermore, remote protons contribute to the transverse relaxation of ^{1}H and ^{15}N nuclei. The extent of the contribution is again dependent on the internuclear distances ($\sim 1/r_{HH}^6$):

$$R_{2H}^{HH} = \frac{p_{HH}^2}{8}[5J(0) + 9J(\omega_H) + 6J(2\omega_H)\,] \quad [8]$$

R_{2IHH} = transverse relaxation of H due to remote protons

$$R_{1N}^{HH} = \frac{p_{HH}^2}{8}[J(0) + 3J(\omega_H) + 6J(2\omega_H)\,] \quad [9]$$

R_{1HH} = relaxation of N due to remote protons

As apparent from the following equations, parameters, which affect this interference relaxation term are for example the magnetic field strength B_0, the rotational correlation time τ_C of the RNA or the gyromagnetic ratio γ_I of the nucleus under study:

$$R_{2I}^{int} = \frac{Cp_{IS}\delta_I[4J(0) + 3J(\omega_I)]}{\sqrt{8} \cdot \sqrt{6}} \quad [10]$$

R_{2Iint} = relaxation due to cross correlation of CSA and dipole-dipole interactions

$$p_{IS} = \left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_I\gamma_S\hbar}{r_{IS}^3} \quad [11]$$

$$p_{HH} = \left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_H\gamma_H\hbar}{r_{HH}^3} \quad [12]$$

p_{IS} = dipole-dipole relaxation, μ_0 = vacuum permeability, γ_I = gyromagnetic ratio of spin I, γ_S = gyromagnetic ratio of spin S, $\hbar$ = Planck constant divided by 2π, r_{IS} = distance between spins I and S

The spectral density can be seen as the amount of motion present at a given frequency ω, which also reflects the number of transitions at that frequency:

$$J(\omega) = \frac{2\tau_C}{5(1 + (\tau_C\omega)^2)} \quad [13]$$

τ_C = rotational correlation time, ω = spectrometer frequency

$$\delta_I = \left(\frac{1}{\sqrt{3}}\right)\gamma_I B_0 \Delta\sigma_I \quad [14]$$

δ_I = relaxation caused by CSA, B_0 = magnetic field strength, $\Delta\sigma_I$ = chemical shift anisotropy

$$C = (3\cos^2\theta - 1) \quad [15]$$

C = angular dependence of cross correlation, θ = angle between the tensor axis of dipole-dipole and CSA interactions

Therefore, the parameters, which are significant for the relaxation behavior and which can be controlled to a certain extent, are identified as B_0, τ_C and γ. Their effect on the transverse relaxation is visualized in simulations of the relaxation behavior in response to increasingly large RNA molecules and magnetic field strengths in the ^{15}N- and in the ^{1}H-dimensions of a TROSY experiment. This is done separately for GC and AU base pairs, as the parameters for the CSA tensors and N-H bond lengths differ (Table 3).

Table 3 Overview of the different parameters for the CSA of ^{1}H and ^{15}N, the N-H bond length and the angle between tensor axis of dipole-dipole and CSA interactions for GC and AU base pairs, respectively.

	GC base pair	**AU base pair**
$\Delta\sigma_N$[141]	-116.85 ppm	-102.75 ppm
$\Delta\sigma_H$[142]	19.65 ppm	20.45 ppm
r_{NH}[141]	1.043 Å	1.057 Å
θ_N[141]	13.1°	11.4°
θ_H[143]	1.9°	2.3°

For better comparison, the transverse relaxation R_2 can be translated into the more familiar peak width at half height $\omega_{1/2}$ using the following relation:

$$\omega_{1/2} = \frac{R_2}{\pi} \qquad [16]$$

In accordance with the predictions for proteins[137], the calculations show a more pronounced increase in the $^{1}H^{N}$-TROSY line width than in the $^{15}N^{H}$-TROSY line width with rising rotational correlation time τ_C, which can be seen as a measure for molecular size (Figure 27 A). This parameter does not necessarily correlate with the RNA length, but examples for a 14, a 37, a 67 and a 108 nts long RNA indicate in which range rotational correlation times are to be expected (Figure 27 A).

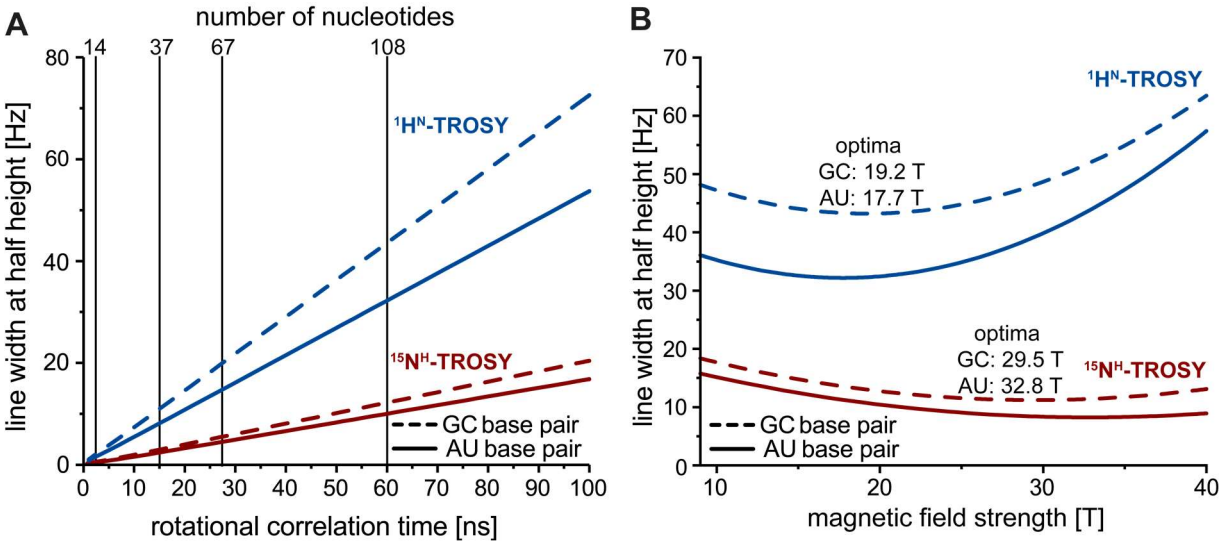


Figure 27 Simulated line width at half height in the dependence of **A** the rotational correlation time and **B** the magnetic field strength for $^{1}H^{N}$- (blue) and $^{15}N^{H}$-TROSY (red) components in GC- (dashed line) and AU-Watson-Crick base pairs (solid line).Line widths at half height have been simulated using equations [4]-[15] and parameters from Table 3. $^{1}H^{N}$-^{1}H distances were measured from U11-H3 and G12-H3 in the NMR structure of the 14 nts RNA with UUCG tetraloop (2KOC)[106] and are: U11: 2.2 Å, 3.8 Å, 2.9 Å, 4.2 Å, 4.8 Å, 3.9 Å, 4.5 Å, 4.6 Å, 4.5 Å, 4.9 Å and G12: 2.2 Å, 2.4 Å, 3.9 Å, 3.5 Å, 4.9 Å, 4.7 Å, 4.8 Å, 4.9 Å, 4.5 Å, 3.8 Å, 4.1 Å. **A** Simulations of the line width at half height have been conducted for a magnetic field strength of 18.8 T corresponding to a ^{1}H-frequency of 800 MHz. Exemplary rotational correlation times for 14, 37, 67 and 108 nts long RNAs have been calculated from the corresponding PDB structure (2KOC[106], 2LHP[144], 2NC1[145] and 2NBX[145]) using HYDRONMR[146]. **B** Line width at half height have been calculated with a rotational correlation time of 60 ns. Optima of the magnetic field strength are given.Furthermore, the imino group is expected to exhibit sharper lines in both dimensions for AU compared to GC base pairs, due to the differences in parameters (Table 3). However, as effects of imino proton exchange with water are neglected in the simulations, line width differences in AU and

GC base pairs might be compensated, especially in the ^{1}H-dimension. The predicted resolution as a function of the magnetic field strength shows that an optimum in resolution is achieved at lower magnetic fields for $^{1}H^{N}$-TROSY than for $^{15}N^{H}$-TROSY signals (Figure 27 B). For the first, the maximum resolution is predicted at around 18 T corresponding to an 800 MHz NMR spectrometer, while the latter reaches a maximum at approximately 31 T corresponding to a 1300 MHz NMR spectrometer. However, for the $^{15}N^{H}$-TROSY line widths the curve profile is flat, suggesting only minute resolution improvements of around 1.5 Hz, when moving from 800 MHz to 1300 MHz spectrometers. Again, GC and AU base pairs exhibit slightly different optima in magnetic field strength due to a different relaxation behavior.

3.2 Materials and Methods

3.2.1 The RNAs under study

^{15}N-detected H-N correlation experiments were conducted on a set of different RNAs ranging in size from 14 to 329 nts with different isotope labeling schemes and phosphorylation states (Figure 28, Table 4).

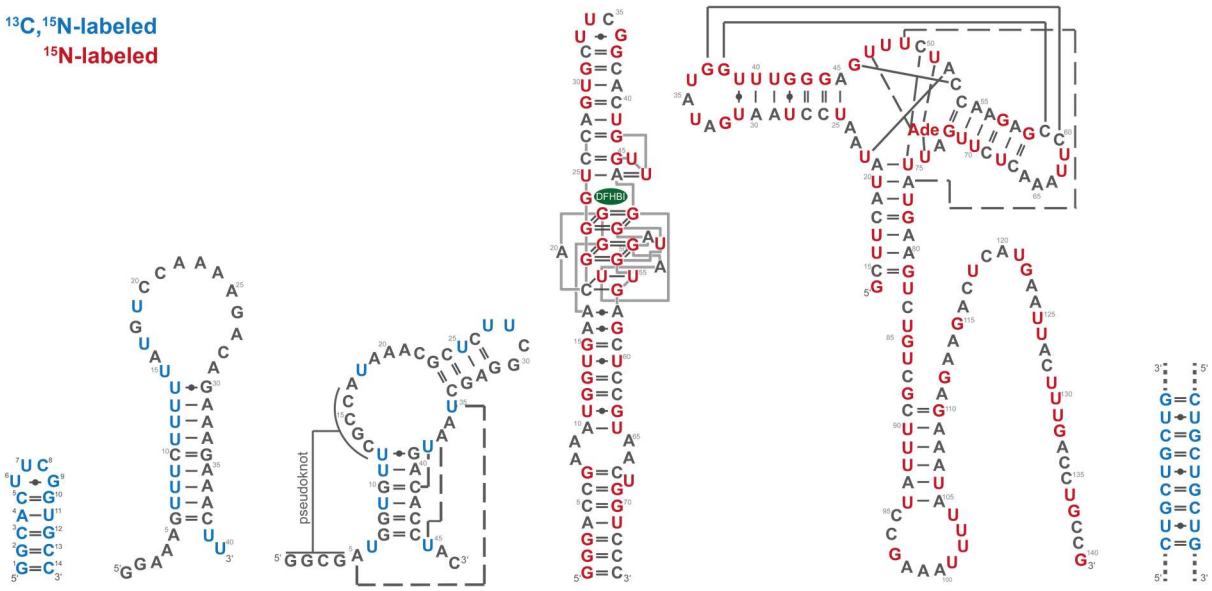


Figure 28 Secondary structures of the 14, 40, 47, 74, 127 and 329 nts long RNAs (from left to right). Nucleotides in blue mark ^{13}C-^{15}N-labeling, ^{15}N-labeled nucleotides are written in red and unlabeled nucleotides are held in gray.

The RNAs do not only vary in size but also in diversity regarding their structural features and functions, so that overall a broad spectrum of different properties is covered. Among the structural features are for example a G-quadruplex in the 74 nts Spinach RNA, the complex tertiary interactions in the 127 nts adenine-sensing riboswitch from *Vibrio vulnificus*[14] involving a kissing loop interaction, or the pseudoknot formation in the 47 nts fluoride riboswitch from *Bacillus cereus*[147]. Furthermore, ligand-binding interactions are being studied as the 127 nts riboswitch is characterized in the presence and in the absence of its ligand adenine.

Table 4 List of RNAs under study including their sequences from 5' to 3' and their isotope labeling schemes.

length	sequence	labeling scheme
14 nts	5'-pppGGCACUUCGGUGCC-3'	uniformly ^{13}C,^{15}N
40 nts	5'-pppGGAAAGUUUCUUUUUAUGUCCAAAAGACAGAAAGAAACUU-3'	^{13}C,^{15}N U
47 nts	5'-pppGGCGAUGGUGUUCGCCAUAAACGCUCUUCGGAGCUAAUGACACC UACp$^{2'3'cycl}$-3'	^{13}C,^{15}N U
74 nts	5'-pppGGGACCGAAAUGGUGAAGGACGGGUCCAGUGCUUCGGCAGUGU UGAGUAGAGUGUGAGCUCCGUAACUGGUCCC-3'	^{15}N U & G
127 nts	5'-GCUUCAUAUAAUCCUAAUGAUAUGGUUUGGGAGUUUCUACCAAGA GCCUUAAACUCUUGAUUAUGAAGUCUGUCGCUUUAUCCGAAAUUUUA UAAAGAGAAGACUCAUGAAUUACUUUGACCUGCCG-3'	^{15}N U &G
329 nts	5'-GGGAGACCGGCAGAUCUGAUAUCAUCGAUGAAUU(CUG)$_{97}$GGGG-3'	uniformly ^{13}C-^{15}N

3.2.2 Sample preparation

14 nts RNA

The uniformly ^{13}C,^{15}N labeled 14 nts hairpin RNA with cUUCGg tetraloop was prepared as described in chapter 2.2.2.

40 nts RNA

The 40 nts ^{13}C,^{15}N uridine labeled expression platform of the 2'-dG sensing riboswitch from *Mesoplasma fluorum*[26] was prepared by Dr. Sara Keyhani via *in vitro* transcription from PCR-amplified DNA fragments using T7 RNA polymerase. The transcription was performed as described in the literature[94] with the following modifications: The PCR was conducted using 0.5 μM of each primer and the conditions for the *in vitro* transcription were 100 mM Tris-glutamic acid (pH 8.1), 2 mM spermidine, 20 mM dithiothreitol (DTT), 2% (v/v) PCR mixture, 15 mM $Mg(OAc)_2$, 7 mM of each rNTP (with ^{13}C,^{15}N rUTP; Silantes, Munich), 20% (v/v) DMSO, 0.2 U/ml yeast inorganic pyrophosphatase (YIPP, NEB) and 144 nM T7 RNA polymerase (P266L)[148]. The RNA was further purified using a buffer exchange protocol described in the literature[149]. The NMR sample contained 2 mM RNA, 25 mM potassium phosphate (pH 6.2) and 8% D_2O. The uridine imino resonance assignment was provided by Dr. Sara Keyhani (unpublished data).

47 nts RNA

The 47 nts ^{13}C,^{15}N uridine labeled fluoride riboswitch from *Bacillus cereus*[147] was prepared by Dr. Heiko Keller via run-off *in vitro* transcription from linearized plasmid DNA using T7 RNA polymerase[148]. ^{13}C,^{15}N-labeled rUTP was purchased from Silantes (Munich, Germany). A self-splicing HDV ribozyme[150] was used to generate 3'-end homogeneity. Preparation and purification of the RNA were performed as described in literature[151]. RNA folding was conducted in NMR buffer (50 mM potassium acetate, pH 6.1) by heating the mixture to 95°C for 10 min followed by an incubation on ice for 30 min. To complex residual Mg^{2+} ions, the RNA was incubated with 2.5 mM EDTA for 10 min at room temperature and then transferred to NMR buffer, reducing the concentration of EDTA below 100 nM. The NMR sample contained 1 mM RNA, 50 mM potassium phosphate (pH 6.1), 10 mM $Mg(OAc)_2$ and 7.5% D_2O. The uridine imino resonance assignment was provided by Dr. Heiko Keller (unpublished data).

74 nts RNA

The 74 nts ^{15}N uridine and guanosine labeled Spinach aptamer was synthesized by Dr. Elke Duchardt-Ferner via run-off *in vitro* transcription from linearized plasmid DNA using T7 RNA polymerase[148]. The RNA forms an internal G-quadruplex upon binding of the fluorogenic ligand 3,5-dihydro-4-fluoro-benzylidene-imidazolinone (DFHBI)[152]. Conditions for transcription and purification are described in the literature[153]. ^{15}N-labeled rUTPs/rGTPs were obtained commercially from Silantes (Munich, Germany) and DFHBI was purchased from Lucerna Inc (New York, USA). RNA folding was conducted by heating to 95°C for 5 min and subsequently diluting 5-fold with ice-cold water (snap cooling). After transferring the RNA into NMR buffer, the sample contained 200 µM RNA, 40 mM HEPES (pH 7.5), 125 mM KCl, 5 mM DFHBI and 8% D_2O. The uridine and guanosine imino resonance assignment was provided by Dr. Elke Duchardt-Ferner (unpublished data).

127 nts RNA

The 127 nts ^{15}N uridine and guanosine labeled full length adenine-sensing riboswitch from *Vibrio vulnificus*[14] was prepared by Dr. Sven Warhaut via *in vitro* transcription from PCR-amplified DNA fragment using T7 RNA polymerase as described the in literature[149]. ^{15}N-labeled rUTP/rGTP was obtained commercially from Silantes (Munich, Germany). A self-splicing 5'-hammerhead ribozyme was used to generate 5'-end homogeneity. Purification occurred via DEAE column as described in the literature[154] followed by a preparative polyacrylamide gel electrophoresis using standard protocols. RNA folding was conducted in water at an RNA concentration of 0.3 mM by heating to 95°C for 5 min and subsequently diluting 10-fold with ice-cold water (snap cooling). After transferring into NMR buffer, the sample contained 500 µM RNA, 25 mM potassium phosphate (pH 6.2), 50 mM KCl, 5 mM $MgCl_2$, 10% D_2O and 100 µM DSS (for referencing). Characterization followed in the absence and presence of 1.4 eq ^{13}C,^{15}N labeled ligand adenine which was synthesized as described[155]. The resonance assignment was taken from the literature[14].

329 nts RNA

The uniformly ^{13}C,^{15}N labeled 329 nts CUG repeat RNA was provided by Dr. Matthias Görlach and was prepared as described in chapter 2.2.2.

3.2.3 NMR spectroscopy

Both, ^{15}N- and ^{1}H-detected H-N correlation experiments were conducted on 800 MHz Bruker NMR spectrometers (18.8 T). ^{15}N-detection was conducted using a 5 mm, z-axis

gradient ^{13}C, ^{15}N [^{1}H]-TXO cryogenic probe, where ^{13}C and ^{15}N are on the inner coil and ^{1}H is on the outer coil. The probe is optimized for ^{13}C. When compared to the respective TCI probe, sensitivity for heteronuclear detection is increased by a factor of at least 2 for the TXO-probe due to a higher filling factor and cold preamplifiers for both nuclei (^{13}C and ^{15}N). All ^{1}H-detected NMR experiments were conducted with a 5 mm, z-axis gradient ^{1}H [^{13}C, ^{15}N]-TCI cryogenic probe, where ^{1}H is on the inner coil and ^{13}C and ^{15}N are on the outer coil (without cold ^{15}N preamplifier). This change of spectrometer was necessary to guarantee comparability between ^{1}H- and ^{15}N-detected experiments, since the probes are similarly optimized for the respective nucleus. To investigate the effect of the magnetic field strength on the ^{1}H and ^{15}N line widths, BEST-TROSY experiments were additionally conducted on a 950 MHz Bruker NMR spectrometer (22.3 T) equipped with a 5 mm, z-axis gradient ^{13}C, ^{15}N [^{1}H]-TXO cryogenic probe (^{15}N optimized) for the 14 nts, 127 nts and 329 nts RNAs. If not indicated otherwise, experiments were conducted at 298 K. Topspin version 3.5 pl6 was used and data was further evaluated using Sparky[117].

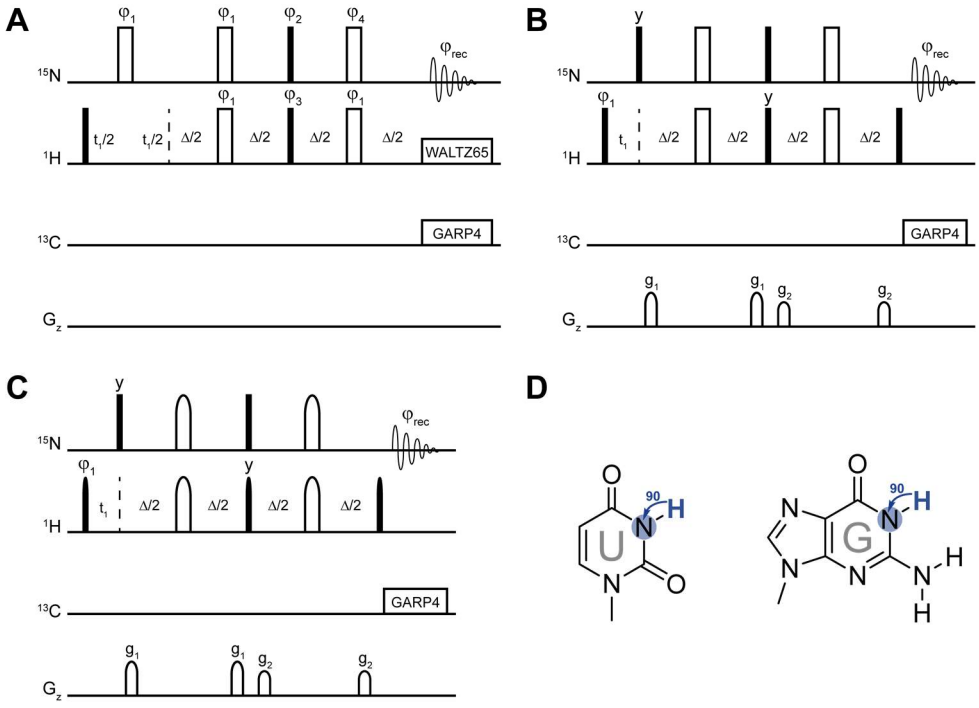


Figure 29 Pulse schemes for **A** the ^{15}N-detected HSQC experiment, **B** the ^{15}N-detected TROSY experiment and **C** the ^{15}N-detected BEST-TROSY experiment. Narrow-filled bars represent rectangular 90° pulses, while wide open bars mark rectangular 180° pulses. Selective 90° pulses are represented by narrow-filled semi elliptic shapes and selective 180° pulses by wide open semi-elliptic shapes. Pulse lengths and shapes for selective pulses at 800 MHz are: ^{1}H: 1275 µs EBurp2.tr, EBurp2 and EBurp2.tr[121] (90° pulses) and 1050 µs Reburp2[121] (180° pulses); ^{15}N: 1500 µs Reburp2[121] (180° pulses). If not indicated otherwise the pulse phase is x. Other pulse phases are; **A** φ_1 = 2(x),2(-x), φ_2 = 4(x),4(y),4(-x),4(-y), φ_3 = y,-y, φ_4 = 2(x),2(-x),2(y),2(-y), φ_{rec} = x,-x,x,-x,y,-y,y,-y,-x,x,-x,x,-y,y,-y,y and **B/ C** φ_1 = y,-y,-x,x / -y,y,x,-x, φ_{rec} = x,-x,-y,y. Delays are: Δ = 5.56 ms = $1/2^{1}J$(N,H). Gradient pulses in **B** and **C** are applied for 1 ms with smooth square amplitude (SMSQ10.100) with strengths of 37% (g_1) and 19% (g_2), where 100% corresponds to 53 G/cm. During acquisition ^{1}H nuclei are decoupled using WALTZ65 sequences[160] (**A**) and asynchronous GARP4 sequences[93] are applied to decouple ^{13}C nuclei (**A**, **B**, **C**). **D** Magnetization transfer schemes for ^{15}N-detected H-N correlation experiments for the imino groups of uridines and guanosines. The nitrogen, where the magnetization is detected, is marked with a blue circle. The ^{1}J(N,H) coupling constant of 90 Hz is given.

The ^{15}N-detected HSQC experiment (Figure 29A) was recorded with a standard pulse sequence[156] with additional ^{13}C-decoupling during the acquisition (for the pulse program see Appendix 1). The ^{15}N-detected TROSY-experiment (Figure 29B) was conducted using the pulse scheme from Takeuchi *et al.*[137] that was published for the application to proteins (for the pulse program see Appendix 1). This pulse sequence was modified to yield the ^{15}N-detected BEST-TROSY experiment (Figure 29C), which was conducted for all RNAs under study (for the pulse program see Appendix 1). For the ^{1}H-detected reference BEST-TROSY experiment, the standard Bruker pulse sequence was used[157,158]. ^{15}N T_2 times were measured using a CPMG-based approach as a pseudo 3D and evaluated with Dynamics Center (Bruker). ^{1}H T_2 times were recorded as a pseudo 2D using a CPMG-sequence as well. The imino proton exchange rates were extracted from a pseudo 2D inversion recovery experiment[159].

To allow a better comparability to the ^{1}H-detected experiments, the direct and indirect dimensions of the ^{15}N-detected experiments were transposed. S/N ratios were determined from 1D rows of the 2D spectra using the “sino” function in Topspin. The highest occurring value of the signal region was divided by the RMSD of the noise region. For determination of line widths, spectra were processed without a filter function in the respective direct dimension, so that the maximal achievable resolution is not biased. Lines were fitted with a Lorentzian line (see equation [1]) in order to extract the line width.

3.3 Results and Discussion

3.3.1 ^{15}N-detected H-N correlation experiments – Which one to use?

In order to determine the most sensitive ^{15}N-detected H-N correlation experiment for RNA, three different experiments, namely an HSQC, a TROSY and a BEST-TROSY experiment, were applied to a small reference RNA. As the favorable relaxation properties of ^{15}N-nuclei should be exploited to its full extend while at the same time guaranteeing comparability to ^{1}H-detected experiments, several considerations had to be made beforehand. To define the optimal acquisition times in the direct and indirect dimensions of H-N correlation experiments, the T_2 times for the 14 nts RNA with UUCG tetraloop were determined using Carr-Purcell-Meiboom-Gill (CPMG) experiments (Table 5). It is to note here that using this method especially proton T_2 times might not be accurate, as the CPMG approach likely quenches some of the imino proton exchange. Therefore, proton T_2 times determined here represent an upper limit and the effective T_2 time is most likely shorter. However, as the purpose of this T_2 time determination is to estimate optimal FID sampling times, the accuracy is well enough in this case.

Table 5 $^{1}H^{N}$ and $^{15}N^{H}$ T_2 times for the 14 nts hairpin RNA, which were determined using CPMG sequences. Optimal acquisition times for the direct (F2 aq_{opt}) and indirect (F1 aq_{opt}) dimensions, which correspond to 3*T_2 and 1.26*T_2, respectively, are given.

residue	$^{1}H^{N}$ T_2 [s]	$^{15}N^{H}$ T_2 [s]
G2	0.07	0.18
G9	0.02	0.17
G10	0.07	0.21
U11	0.09	0.25
G12	0.10	0.29
F2 aq_{opt}	0.30	0.87
F1 aq_{opt}	0.13	0.37

The T_2 times for $^{1}H^{N}$ vary between 0.02 s^{-1} and 0.10 s^{-1} for G9 and G12, while the $^{15}N^{H}$ T_2 times lie between 0.17 s^{-1} and 0.29 s^{-1}. From this, FID sampling times can be estimated, which have to strike the balance between S/N ratio and resolution. In the direct dimension, the FID is usually sampled for a time as long as 3*T_2, while in the indirect dimension a maximum S/N ratio is achieved at 1.26*T_2[161]. With this information at hand, optimal acquisition times for ^{1}H-detected experiments are 0.3 s in the direct and 0.37 s in the indirect dimension. For ^{15}N-detected experiments, best acquisition times are 0.87 s in the direct and 0.13 s in the indirect dimension. However, as ^{13}C-decoupling has to be applied during the acquisition in ^{15}N-detected experiments, the FID sampling time was reduced to 0.3 s (~1*T_2) to prevent sample heating. These conditions combine the highest S/N ratio with the maximum achievable resolution.

Having identified the optimal sampling times, a ^{1}H-detected BEST-TROSY experiment and three different ^{15}N-detected H-N correlation experiments were applied to the 14 nts RNA with UUCG tetraloop. Among the ^{15}N-detected experiments, the TROSY scheme has

already been established for proteins[137] and the HSQC experiment is a standard pulse sequence[156]. However, the ^{15}N-detected BEST-TROSY experiment represents a modification of the TROSY experiment, where band-selective proton pulses are applied to the imino protons between 9 and 15 ppm, leaving the water resonance unperturbed and thus allowing a rapid sampling with inter-scan delays as short as 0.2 s. The resulting spectra (Figure 30) were compared in terms of the resolution in the ^{15}N-dimension as well as the sensitivity.

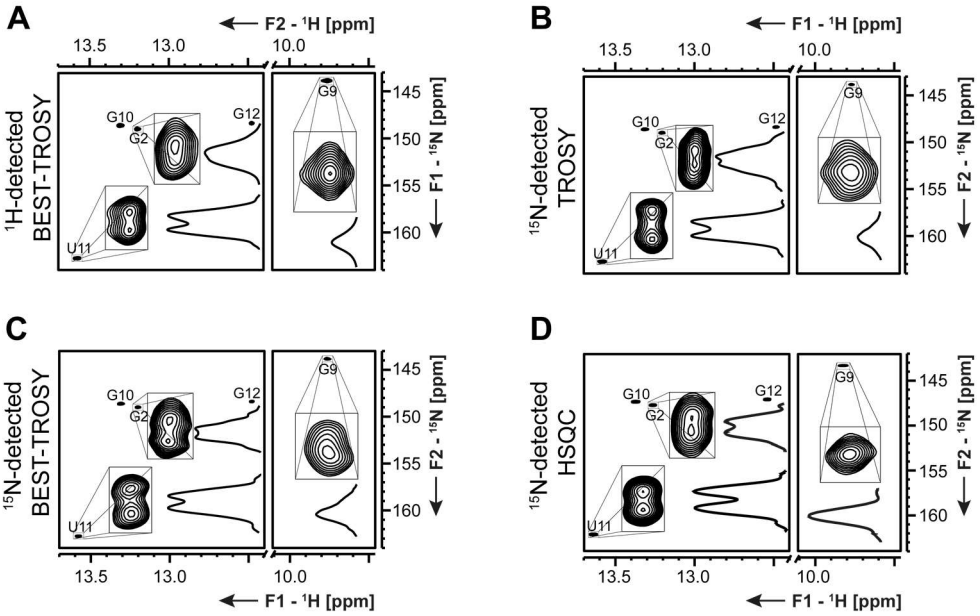


Figure 30 **A** ^{1}H-detected BEST TROSY, **B** ^{15}N-detected TROSY, **C** ^{15}N-detected BEST-TROSY and **D** ^{15}N-detected HSQC spectra of the 14 nts hairpin RNA. 2D close-ups as well as 1D traces are shown for the signals of U11, G2 and G9. All experiments were conducted at 298 K and at 800 MHz spectrometers. The spectrum in panel **A** was recorded with spectral windows of 24 ppm in both dimensions. Carrier frequencies were set to 4.7 ppm, 150 ppm and 153 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.3 s in both, the ^{1}H- and ^{15}N-dimensions. With 32 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded within 8 h. The spectra in panel **B** and **C** were recorded with spectral windows of 35 ppm in the direct ^{15}N-dimension and 4.3 ppm in the indirect ^{1}H-dimension. Carrier frequencies were set to 155 ppm, 160 ppm and 11.68 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct ^{15}N-dimension, during which ^{13}C nuclei were decoupled using GARP4 sequences[93], and 0.126 s in the indirect ^{1}H-dimension. With 112 scans per increment and an inter-scan delay of 1 s, the ^{15}N-detected TROSY spectrum (**B**) was recorded within 1 d 14 h. The ^{15}N-detected BEST-TROSY spectrum (**C**) on the other side was acquired with 64 scans per increment and an inter-scan delay of 0.3 s, which resulted in an overall experimental time of 9.5 h. The spectrum in panel **D** was recorded with spectral windows of 35 ppm in the direct ^{15}N-dimension and 4.4 ppm in the indirect ^{1}H-dimension. Carrier frequencies were set to 155 ppm, 150 ppm and 11.8 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct ^{15}N-dimension, during which ^{13}C nuclei were decoupled using GARP4 sequences[93], and 0.04 s in the indirect ^{1}H-dimension. The acquisition time in the ^{1}H-dimension was reduced to save measurement time and thus the resolution can only be compared in the ^{15}N-dimension. With 128 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded within 15 h. The figure has been modified from Schnieders *et al*[132].

Importantly, all of the applied H-N correlation experiments result in the same set of imino signals. Additionally, it is possible to observe all expected cytidine amino resonances with the ^{15}N-detected HSQC experiment (see Appendix 2 Figure 64). Furthermore, the high ^{15}N-resolution becomes evident, as the doublet structure of the resonances belonging to Watson-Crick base pairs can be resolved, which is caused by the small ^{2h}J(N,N) coupling across the hydrogen bond. In line with the previously determined T_2 times, the resolution

cannot be improved by increasing the acquisition time, as it already reaches a plateau and further sampling would only result in a decrease in sensitivity (Figure 31).

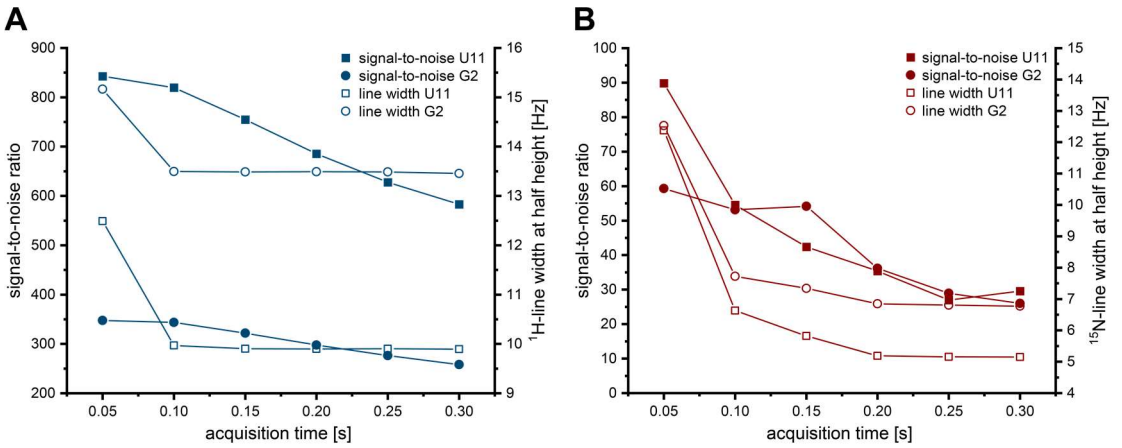


Figure 31 Signal-to-noise ratio and line width of U11 and G2 in dependence of the acquisition time for **A** the ^{1}H-detected and **B** the ^{15}N-detected BEST-TROSY experiments. The S/N ratio has been determined from 1D rows using the "sino" function in Topspin.

To detect resolution differences, ^{15}N line widths were extracted for all of the resonances. Among the ^{15}N-detected experiments, the HSQC experiment clearly exhibited the highest resolution with line widths of ^{15}N-LW(HSQC) = 4.9 ± 0.6 Hz for Watson-Crick base pairs. The TROSY experiments on the other side only yielded line widths of ^{15}N-LW(TROSY) = 6.1 ± 0.8 Hz and ^{15}N-LW(BEST-TROSY) = 6.4 ± 0.7 Hz. This difference was even more prominent for the resonance of the non-canonical GU base pair (G9). Here, the line width at half height was 4.9 Hz in the HSQC spectrum, which is exceeded by far in the TROSY spectra with ^{15}N-LW(TROSY) = 12.5 Hz and ^{15}N-LW(BEST-TROSY) = 11.5 Hz. This indicates that lines in the HSQC experiment are less influenced by imino proton exchange than in the TROSY experiments. This seems reasonable when it is considered that in the HSQC experiment a pure in-phase operator $N_{x/y}$ evolves, while an anti-phase operator $N_{x/y}H_z$ is acquired in the TROSY experiments. Moreover, for this short RNA the TROSY effect is expected to be rather small and thus yields a minor benefit in resolution by selecting the slowly relaxing line of the coupled multiplet. This modest advantage might not compensate for the increased impact of solvent exchange in the TROSY experiments. However, the resolution advantage in the HSQC experiment is expected to vanish for RNAs with larger rotational correlation times τ_c. Indeed, when the ^{15}N resonances of the 40 nts RNA are compared (Figure 32 B), the resolution in the BEST-TROSY and in the HSQC experiment is equal.

In terms of sensitivity, the relative S/N ratio per unit time is three times lower in the ^{15}N-detected HSQC experiment in comparison to the ^{15}N-detected BEST-TROSY experiment. The relative sensitivity increases from 1 in the HSQC experiment to 1.4 in the TROSY experiment to 2.9 in the BEST-TROSY experiment. This is also reflected in the overlay of 1D rows of the TROSY and BEST-TROSY experiments (Figure 32 A), where a similar S/N ratio is achieved, even though the BEST-TROSY experiment is only recorded with around

half (64 instead of 112) transients per increment. Having identified the BEST-TROSY experiment as the most sensitive ^{15}N-detected H-N correlation experiment, it was compared to the ^{1}H-detected BEST-TROSY experiment.

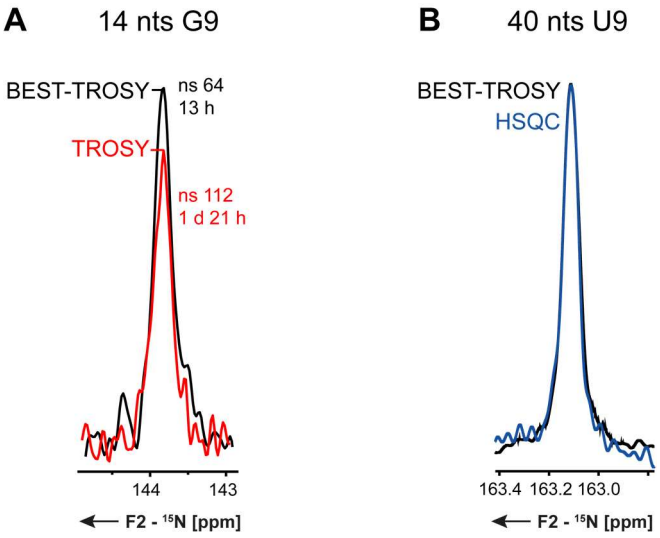


Figure 32 A Overlay of 1D rows for signal G9 in the ^{15}N-detected TROSY (red) and BEST-TROSY (black) experiments. Spectra were scaled to the same amplitude of noise. The figure has been adapted from Schnieders *et al.*[132]. **B** Overlay of 1D rows for the signal U9 of the 40 nts RNA in the ^{15}N-detected BEST-TROSY (black) and HSQC (blue) spectra. For better comparison of the line width at half height resonances were plotted to the same intensity.

In comparison to its ^{1}H-detected counterpart, the ^{15}N-detected BEST TROSY experiment is by a factor of 30 less sensitive. This is well in agreement with expectations since the S/N ratio increases with $\gamma_{\mathrm{det}}^{3/2}$, which would result in a theoretical factor of 32 between S/N ratios. Furthermore, there is no significant difference in the ^{15}N-resolution, as the FID is sampled for the same time throughout all experiments. As ^{15}N direct-detection promises high resolution, especially for RNAs of increasing molecular size, ^{1}H and ^{15}N line width were determined in ^{15}N- and ^{1}H-detected BEST TROSY spectra for a set of different RNAs exhibiting a broad variety of structural features and ranging from 5 kDa to 100 kDa in their molecular size.

3.3.2 Effect of molecular size on line widths

Within the process of applying the ^{1}H- and ^{15}N-detected BEST-TROSY experiments to various RNAs, an interesting feature was observed for the 127 nts long adenine-sensing riboswitch. Here, ligand binding could be mapped, as the ^{2h}J(N,N) coupling across the hydrogen bond to the ^{13}C,^{15}N-labeled ligand adenine could be resolved in the ^{15}N-dimensions of both BEST-TROSY experiments (Figure 33). In the ^{1}H-detected experiment this could only achieved by sampling the indirect ^{15}N-dimension with an unusually high resolution. This coupling of around 6-7 Hz is typically exploited in HNN-COSY experiments, where Watson-Crick base pairs can be identified[41]. With this size the ^{2h}J(N,N) coupling exceeds the intra-base ^{2}J(N3,N1) and ^{2}J(N1,N2) coupling constants,

which are only at 2.2 Hz[162] and can thus not be detected even for the smallest RNA (14 nts) under study.

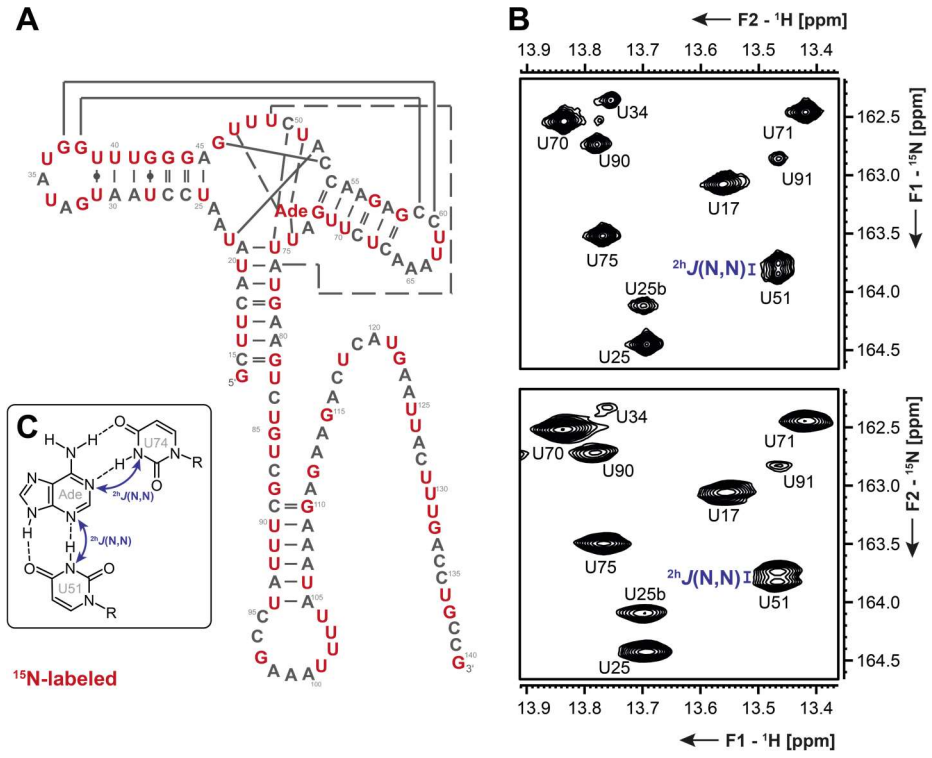


Figure 33 A Secondary structure of the ligand bound state of the 127 nts long adenine-sensing riboswitch. ^{15}N-labeled nucleotides are written in red and tertiary interactions are indicated with gray lines. **B** ^{1}H- (upper spectrum) and ^{15}N-detected BEST-TROSY spectra of the ligand bound state of the 127 nts RNA at 298 K (for full spectrum see Figure 65 C). The assignment was taken from the literature[14]. The ^{1}H-detected experiment was recorded with spectral windows of 25 ppm and 28 ppm in the direct ^{1}H- and indirect ^{15}N-dimensions, respectively. Carrier frequencies were set to 4.7 ppm, 101 ppm and 153 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.3 s in both, the ^{1}H- and ^{15}N-dimensions. With 64 scans per increment and an inter-scan delay of 0.5 s, the experiment was recorded within 20.5 h. The ^{15}N-detected BEST-TROSY experiment was recorded with spectral windows of 35 ppm in the direct ^{15}N-dimension and 5.2 ppm in the indirect ^{1}H-dimension. Carrier frequencies were set to 155 ppm, 120 ppm and 12.2 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct ^{15}N-dimension, during which ^{13}C nuclei were decoupled using GARP4 sequences[93], and 0.03 s in the indirect ^{1}H-dimension. With 576 scans per increment and an inter-scan delay of 0.2 s, the experiment was recorded within 20 h. The ^{2h}J(N,N) coupling is marked in blue for residue U51, which is part of the binding pocket as schematically represented in panel **C**[163]. The figure has been modified from Schnieders *et al*[132].

The 127 nts RNA was not only studied in the presence (Figure 33), but also in the absence of ligand (Appendix 2 Figure 65), which revealed an identical set of signals for both, the ^{1}H- and the ^{15}N-detected BEST-TROSY experiments. This also holds true for all the other RNAs under study, which exhibited the same set of resonances in both experiments, irrespective of their structural motifs (Figure 34). Furthermore, the ^{2h}J(N,N) coupling across the hydrogen bond was not observed for these RNAs, as either the labeling scheme did not allow it (40, 47 and 74 nts) or the signal line widths exceeded the inter-base scalar coupling in the ^{15}N-dimension, which was the case for the 329 nts RNA (Figure 34 D).

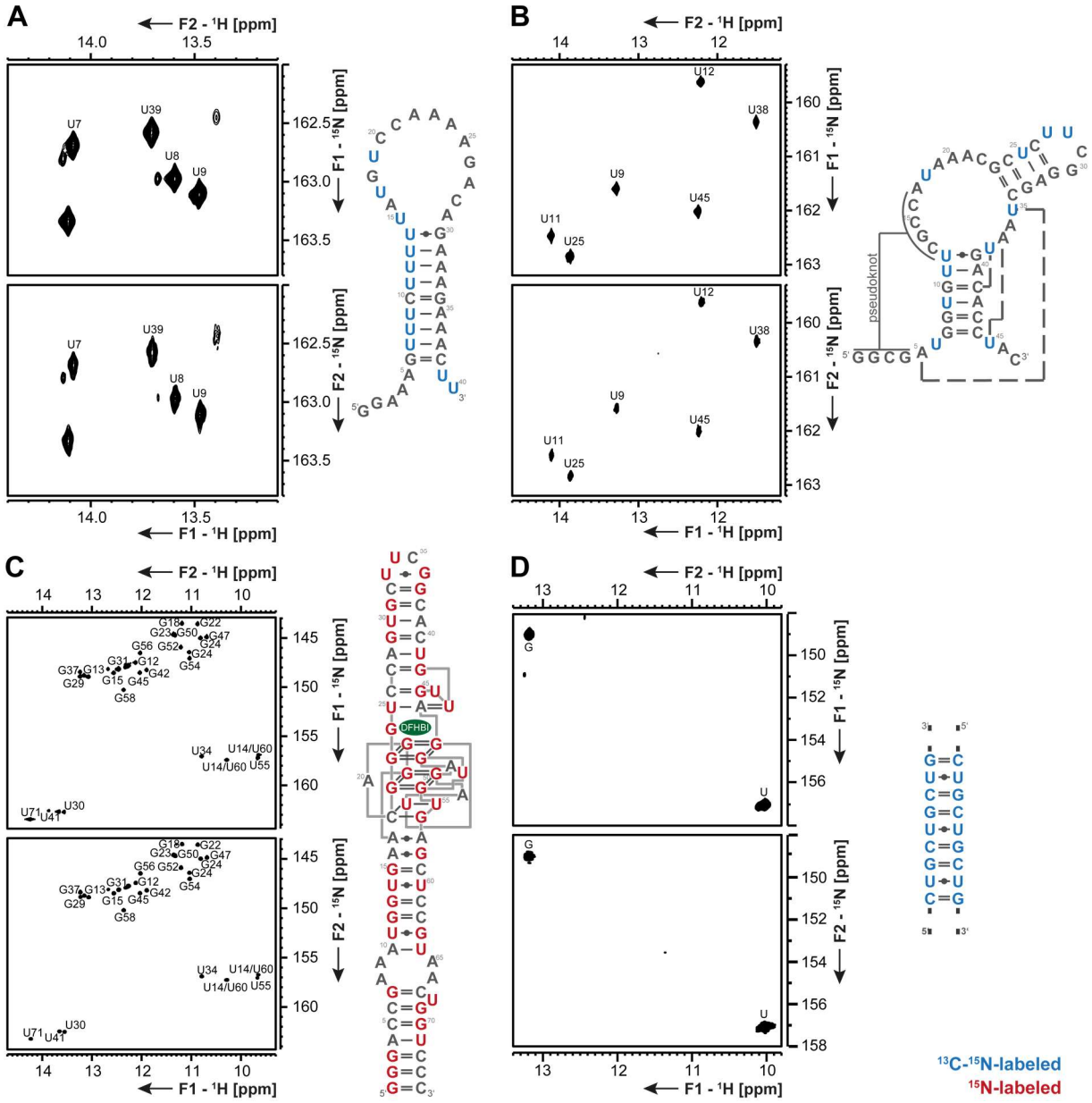


Figure 34 ^{1}H-detected (upper panel) and ^{15}N-detected (lower panel) BEST-TROSY experiments of the **A** 40 nts terminator stem of the 2'dG riboswitch[94], the **B** 47 nts fluorid riboswitch[147], the **C** 74 nts Spinach aptamer[152] and the **D** 329 nts CUG repeat RNA[112] with their secondary structures. ^{13}C,^{15}N-labeled nucleotides are held in blue while ^{15}N-labeled nucleotides are shown in red. The common experimental parameters of the ^{1}H-detected BEST-TROSY experiment were: The spectral width in the ^{1}H-dimension was 24 ppm and carrier frequencies for ^{1}H and ^{13}C were 4.7 ppm and 101 ppm, respectively. The acquisition time in the direct ^{1}H-dimension was 0.3 s. All of the ^{15}N-detected BEST-TROSY experiments were recorded with a spectral window of 35 ppm in the ^{15}N-dimension, while the carrier frequencies were 155 ppm (^{15}N) and 150 ppm (^{13}C). During the acquisition (0.3 s) carbon decoupling was performed using GARP4 sequences[93] with a field strength of 1.8 kHz. The remaining parameters for ^{1}H- and ^{15}N-detected experiments can be extracted from Appendix 2 chapter 6.1.1. The figure has been modified from Schnieders *et al*[132].

The signal line widths were extracted from all spectra in the direct ^{1}H- and ^{15}N dimensions from the ^{1}H- and ^{15}N-detected BEST-TROSY spectra, respectively (Figure 35 B). A comparison of the line widths and the theoretical predictions (Figure 35 A) reveals that the general trend that ^{15}N line widths increase slower than ^{1}H line widths with the molecular size can be confirmed. However, the line widths in both dimensions are a little larger than theoretically predicted, which is due to the fact that neither

conformational nor chemical exchange were considered in the simulations, as these contributions strongly depend on the respective system under study, which renders an inclusion in a general prediction impossible.

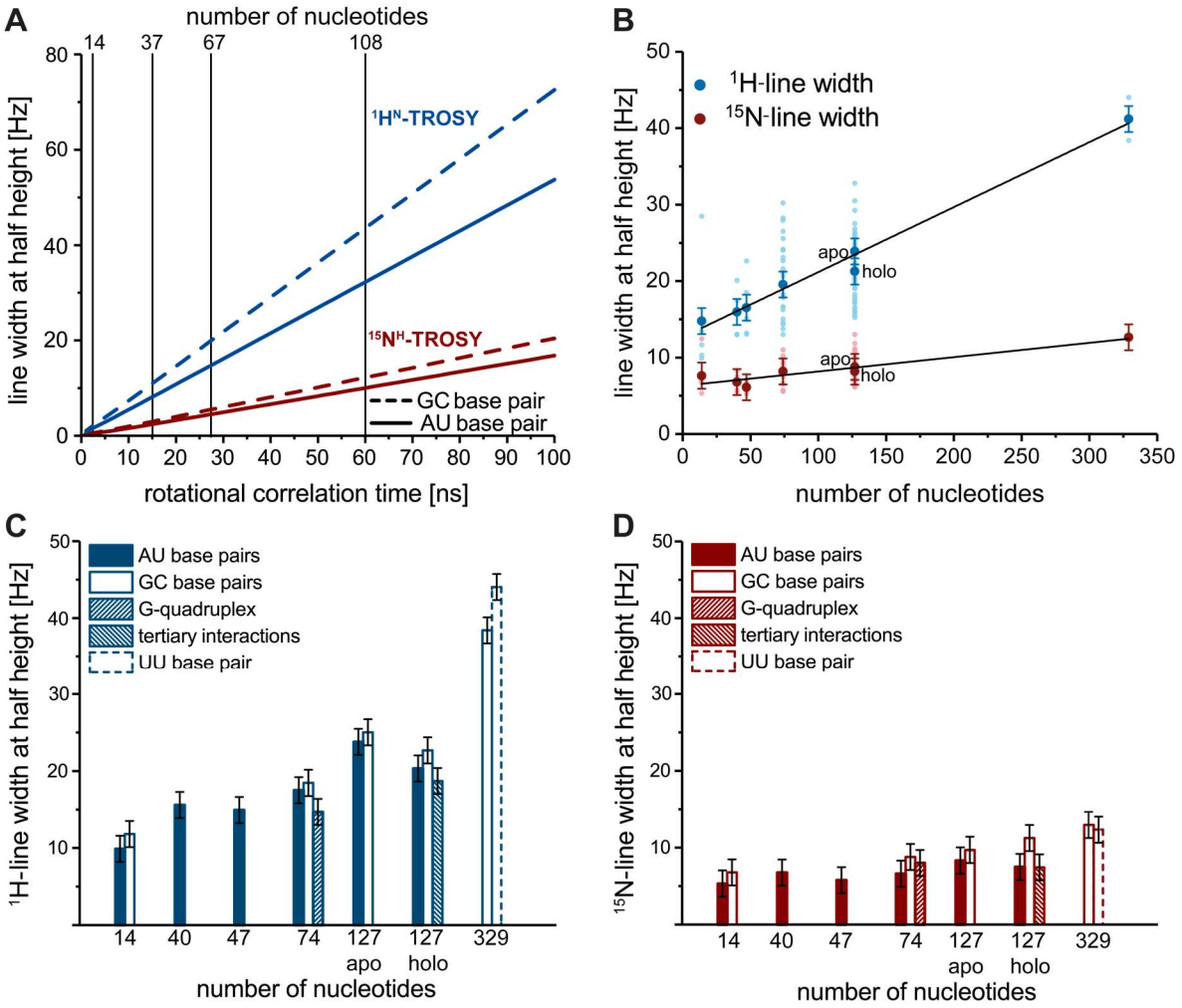


Figure 35 A Simulated line width at half height in the dependence of the rotational correlation time for $^{1}H^{N}$- (blue) and $^{15}N^{H}$-TROSY (red) components in GC- (dashed line) and AU-Watson-Crick base pairs (solid line). Exemplary rotational correlation times for a 14, a 37, a 67 and a 108 nts RNA are marked and were extracted from their PDB files (2KOC[106], 2LHP[144], 2NC1[145] and 2NBX[145], respectively) using HYDRONMR[146]. Further details are described in Figure 27. **B** Experimentally determined ^{1}H (blue) and ^{15}N line widths (red) at half height for a 14[106], a 40[26], a 47[147], a 127[14] and a 329 nts long RNA[112]. The 127 nts adenine-sensing riboswitch was studied in the ligand-free (apo) and in the ligand-bound form (holo). All line widths from which the mean values (bright colored circles) were determined are shown as light colored circles. The experimental error represents the maximal possible resolution, which is 1.7 Hz in both dimensions. **C** Classification of ^{1}H (blue) and **D** ^{15}N line widths (red) into different interaction type categories, namely AU (filled bars) and GC Watson-Crick base pairs (open bars), G-quadruplex resonances (up-hashed bars), residues involved in tertiary interactions (down-hashed bars) and UU base pairs (dashed bars). Errors again correspond to the maximum achievable resolution of 1.7 Hz. The figure was adapted from Schnieders *et al.*[132].

The observation that imino proton exchange has a significant effect especially on ^{1}H line widths is reflected in the distribution of ^{1}H line widths within a single RNA. This distribution ranges for example from ~13 Hz to ~30 Hz for the 74 nts RNA, while ^{15}N line widths of the same RNA only range between ~6 Hz and ~11 Hz. When determining imino proton exchange rates for the 14 nts RNA, it becomes very clear that this solvent exchange is already contributing up to 29% (G10) to the line widths of Watson-Crick base pairs (Table 6). Furthermore, the better the base pair is protected, the less the exchange

contributes to the line width. This can be observed for residue G12, which is well embedded in the middle of the stem in the 14 nts hairpin RNA and consequently exhibits the lowest exchange contribution (9%) to the line width.

Table 6 Overview of imino proton exchange rates and the line widths of the imino proton signals of the 14 nts hairpin RNA[106]. Exchange rates were determined using inversion recovery experiments[159]. Contributions of exchange rates to the line widths are given in percent.

residue	^{1}H exchange rate	^{1}H line width	contribution
G2	2.1 s^{-1} ± 0.7 s^{-1}	13.5 s^{-1}	16%
G10	3.4 s^{-1} ± 0.8 s^{-1}	11.7 s^{-1}	29%
U11	2.3 s^{-1} ± 0.7 s^{-1}	9.9 s^{-1}	23%
G12	0.9 s^{-1} ± 0.6 s^{-1}	10.3 s^{-1}	9%

To further investigate the dependence of resolution on secondary structure motifs, the line widths were classified into different interaction type categories (Figure 35 C and D). Here, the general trend that AU base pairs tend to give rise to sharper lines than GC base pairs in both, ^{15}N- and ^{1}H-dimensions, could be confirmed. Additionally, the solvent exchange contributions on ^{1}H-line widths again could be observed for G-quadruplex resonances of the 74 nts Spinach aptamer. Imino protons that are involved in G-quadruplex interactions are highly shielded from solvent exchange, which is especially visible when the solvent is exchanged from H_2O to D_2O. Here, the signals of the imino protons, which are involved in standard base pairs vanish within seconds, whereas resonances of imino protons in G-quadruplexes can still be observed after hours[164]. This exchange behavior leads to very narrow ^{1}H signals for the G-quadruplex imino protons (Figure 35 C), while the ^{15}N line widths of the same imino group are not particularly affected (Figure 35 D). A similar relation can be observed for the UU base pair in the 329 nts CUG repeat RNA. The imino proton exchange is generally enhanced in UU base pairs in comparison to GC base pairs[165]. For this reason, the ^{1}H line width of the UU base pair is larger than the one of the GC base pair in the 329 nts RNA. This effect is again diminished for ^{15}N line widths, where no significant difference between the two interaction types can be detected.

3.3.3 Effect of magnetic field strength

As a higher magnetic field strength is expected to have a positive effect on ^{15}N line widths up to 30 T, the ^{15}N-detected BEST-TROSY experiment was repeated for the 14 nts hairpin RNA, the 127 nts ligand-bound state of the adenine-sensing riboswitch and for the 329 nts CUG repeat RNA at a 950 MHz NMR spectrometer (Figure 36 A, B and D). To guarantee comparability, the experiments were repeated using the identical parameters, while adapting the selective pulses as well as the number of increments in order to reach the same resolution according to the higher magnetic field.

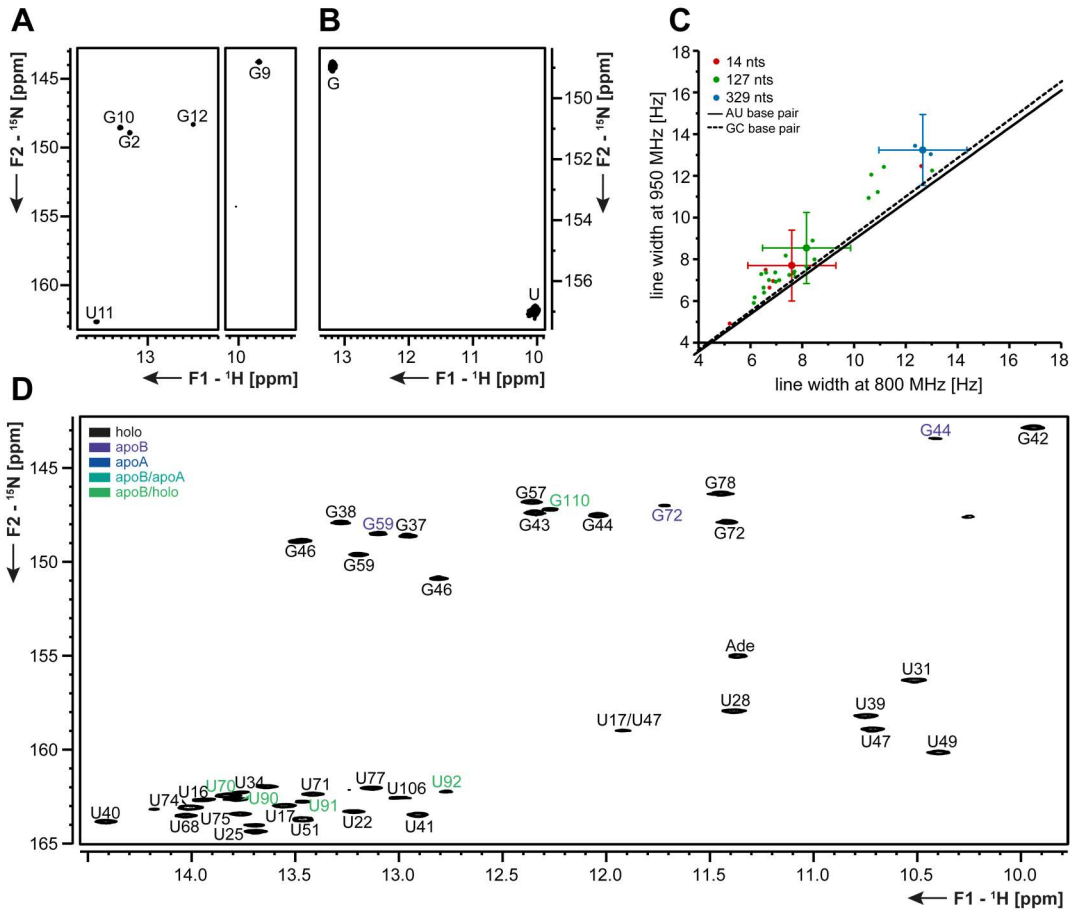


Figure 36 A ^{15}N-detected BEST-TROSY spectrum of the 14 nts hairpin RNA[106] at 950 MHz. The experiment was recorded with spectral windows of 35 ppm in the direct ^{15}N-dimension and 4.3 ppm in the indirect ^{1}H-dimension. Carrier frequencies were set to 155 ppm, 150 ppm and 11.7 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct ^{15}N-dimension, during which ^{13}C nuclei were decoupled using GARP4 sequences[93], and 0.06 s in the indirect ^{1}H-dimension. With 64 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded within 6 h. **B** ^{15}N-detected BEST-TROSY spectrum of the 329 nts CUG repeat[112] at 950 MHz. The experiment was recorded with spectral windows of 35 ppm in the direct ^{15}N-dimension and 3.5 ppm in the indirect ^{1}H-dimension. Carrier frequencies were set to 155 ppm, 150 ppm and 11.6 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct ^{15}N-dimension, during which ^{13}C nuclei were decoupled using GARP4 sequences[93], and 0.03 s in the indirect ^{1}H-dimension. With 640 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded within 22 h. **C** Line widths of the 14 (red), the 127 (green) and the 329 nts (blue) RNAs determined at 950 MHz and 800 MHz plotted against each other. Big circles represent the respective mean values determined from the individual line widths that are shown as small circles. The error represents the maximal achievable resolution of 1.7 Hz. Theoretical predictions of line widths for GC (dashed line) and AU base pairs (solid line) are shown in black. **D** ^{15}N-detected BEST-TROSY spectrum of the 127 nts riboswitch[14] at 950 MHz. The experiment was recorded with spectral windows of 35 ppm in the direct ^{15}N-dimension and 5.0 ppm in the indirect ^{1}H-dimension. Carrier frequencies were set to 155 ppm, 150 ppm and 12.2 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct ^{15}N-dimension, during which ^{13}C nuclei were decoupled using GARP4 sequences[93], and 0.02 s in the indirect ^{1}H-dimension. With 576 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded within 21 h. The figure was adapted from Schnieders *et al.*[132].

The sensitivity of these experiments at 950 MHz is increased by a factor of ~1.5, which results from the higher magnetic field strength itself[137,166] in combination with a 30% increase, as the TXO probe at the 950 MHz NMR spectrometer is optimized for ^{15}N as opposed to ^{13}C nuclei. In contrast, the expected resolution increase that is caused by the higher magnetic field strength was not observed for any of the RNAs under study (Figure 36 C). This behavior can be explained as the expected resolution increase is only at around 10% for an RNA with a rotational correlation time of 60 ns. This is even below

the maximum achievable resolution of 1.7 Hz and thus the effects are predicted to be rather small. However, as also no improvement can be detected for the largest RNA under study (329 nts), it has to be assumed that exploiting even larger rotational correlation times or magnetic field strengths might allow the observation of a positive effect in ^{15}N resolution. However, at a closer inspection the resolution at 950 MHz is getting worse in comparison to the spectra recorded at 800 MHz, which can currently only be explained by exchange processes that are present in all RNAs and that are closer to the intermediate exchange regime at 950 MHz. Another possible explanation for this deviation of experiment and prediction might be that the parameters for the predictions do not precisely match the RNAs under study. This might be the case e.g. for the CSA, which has a strong influence on the magnetic field strength optimum and can only be determined in the solid state and might deviate in solution.

3.4 Conclusions and Outlook

Within this PhD project different ^{15}N-detected H-N correlation experiments were applied to RNAs to investigate the effect of molecular size and magnetic field strength on sensitivity, resolution and relaxation behavior. The experiments presented here remain the first multidimensional nitrogen direct-detected NMR experiments, which have been applied to RNA up to today. We were able to identify the ^{15}N-detected BEST-TROSY experiment as the most sensitive H-N correlation experiment, which increases the sensitivity in comparison to the previously published ^{15}N-detected TROSY experiment[137] by a factor of 2.

To investigate the spectral resolution in dependence of the molecular size, the ^{15}N-detected BEST-TROSY experiment was recorded along with its ^{1}H-detected analogue for a set of different RNAs ranging from 5 kDa to 100 kDa. Here, we were able to confirm the theoretical predictions that ^{15}N line widths increase slower with the molecular size than ^{1}H line widths and analyzed the different effects of various structural motifs on the resolution. The very high ^{15}N-resolution is especially reflected in the ability to resolve small scalar ^{2h}J(N,N) couplings across the hydrogen bonds of Watson-Crick base pairs or ligand-binding sites. In the light of the goal to push the limitations of NMR spectroscopy on biomolecules towards higher molecular weight RNAs, this observation is highly important. Therefore, when only considering resolution, ^{15}N-detection might be favorable for larger RNAs due to the very high ^{15}N resolution, which comes for free in terms of measurement time in the direct dimension. However, in a more realistic picture, ^{15}N-detection suffers from the loss of sensitivity in comparison to the ^{1}H-detected counterpart experiments. This is especially problematic when considering that with the latter, the identical ^{15}N-resolution can be achieved when sampling the same number of increments in the indirect ^{15}N-dimension. Furthermore, the predicted magnetic field dependence cannot be confirmed as ^{15}N resonances tend to exhibit slightly larger line widths at a higher magnetic field strength (950 MHz instead of 800 MHz); an effect, which should be the contrary according to predictions and might result from exchange processes that are present in all RNAs under study.

Altogether, ^{15}N direct-detected BEST-TROSY experiments do not provide any advantage in comparison to the ^{1}H-detected counterpart under the experimental conditions applied here. However, in principal this could be achieved by changing the experimental set-up as follows: The nitrogen direct-detected experiments should be conducted on a TXO probe optimized for ^{15}N-detection. Furthermore, moving towards *in vivo* conditions will favor ^{15}N-detection due to the higher salt tolerance when compared to ^{1}H-detected experiments[138]. Lastly, an RNA exhibiting an even higher rotational correlation time will further increase the gap between ^{1}H and ^{15}N-resolution, which will at some point favor nitrogen direct-detection. Following this approach, a nearly identical sensitivity of ^{1}H- and ^{15}N-detected TROSY experiments was reached in proteins[138].

Moving away from the BEST-TROSY experiment, the potential advantage of ^{15}N-detection does not only lie in an increased resolution but also in the independence of solvent exchange processes, which are in RNA especially prominent for the imino protons. This benefit cannot be exploited to its full extend with the ^{15}N-detected BEST-TROSY experiment, as the operator, which evolves the nitrogen chemical shift still exhibits an H_z component. Therefore, regarding the resolution, the ^{15}N-detected HSQC experiment outcompetes all other H-N correlation experiments for low molecular weight RNAs. With a further optimization of this experiment in terms of maximizing the sensitivity by e.g. implementing selective pulses to enable fast pulsing, additional imino proton resonances might be observed that are not accessible in ^{1}H-detected experiments within a reasonable experimental time.

Chapter IV: Structural characterization of RNA tetraloops for force field calibration

The results presented in this chapter are also compiled in a manuscript in preparation:

R. Schnieders‡, A. Oxenfarth‡, G. Pintér‡, C. Richter, H. R. Jonker, B. Fürtig and H. Schwalbe, *Manuscript in preparation*.

Author contributions are listed below:

R. Schnieders, A. Oxenfarth and G. Pintér contributed equally to this project with having studied a different tetraloop each.

H. R. Jonker helped with structure calculations.

C. Richter, B. Fürtig and H. Schwalbe conducted the project design and helped with pulse sequence optimization.

Preface:

In the course of this chapter the results of all three tetraloops, which are part of this project, will be compared to each other. It has to be mentioned here that the results of the 14 nts RNA with CUUG tetraloop have been generated by Andreas Oxenfarth and the results of the GCAA tetraloop was studied by György Pintér. The results for the GAAG tetraloop were generated by myself. These contributions are also marked at the appropriate positions in the text.

14 nts RNA with GAAG tetraloop:

The fundamentals of this project have already been laid in the course of my master thesis. This includes the optimization of the GAAG sample preparation and a full resonance assignment except for the quaternary carbon and tertiary nitrogen atoms of the nucleobases. Furthermore, the ^{n}J(C,P) and preliminary ^{3}J(H,H) coupling constants have been already determined. All further coupling constants, the temperature-dependent chemical shift analysis, all cross-correlated relaxation rates and structure calculations are subject of this PhD thesis.

4.1 Introduction

As has been outlined in Chapter I, NMR spectroscopy is undoubtedly a powerful tool in the structure determination of RNA. As opposed to cryogenic electron microscopy or X-ray crystallography, solution NMR spectroscopy is applied under close to native conditions, rendering the characterization of molecular dynamics possible. This feature is especially important for RNA due to the complex nature of RNA structure. Moreover, multiple torsion angles of the RNA backbone can also be determined by NMR spectroscopy, which makes NMR data a valuable resource in the development of computational methods for structural predictions. With the increase in computational power that is for example needed for extended molecular dynamics simulations (MD simulations), this field gains more and more in importance and can potentially be applied as an orthogonal method to fill the gaps, which are beyond experimental reach. In protein research for example MD simulations already complement experimental observations quite successfully as can be seen exemplary in the predictions of the effects of mutations on the stability and function of a protein[167,168]. This nowadays achieved interdisciplinarity has been preceded by a still ongoing history of optimizations of the underlying potential energy functions that describe the properties of a given protein in the correct way.

In contrast to protein computational biology, the development of these so-called force fields for RNA is still extensively worked on. One of the barriers here is the lack of sufficiently well characterized RNAs, which can be used for the parametrization of the force fields. The most commonly used reference system is the 14 nts RNA with UUCG tetraloop, as a high resolution NMR structure[106] is available for this RNA which is based on a plethora of experimental data.

The goal of this project is to extend the number of potential reference systems, which are equally well characterized as the 14 nts RNA with UUCG tetraloop. Therefore, RNAs with different tetraloops, namely the GCAA, the GAAG and the CUUG tetraloops with their individual native closing base pairs, were designed. The tetraloops were embedded in the same stem sequence as present in the 14 nts RNA with UUCG tetraloop. While the GCAA and CUUG tetraloops are being studied by György Pintér and Andreas Oxenfarth, the scope of this PhD thesis involves the extensive structural characterization of the GAAG tetraloop. The fundament of this process is the sample and a full resonance assignment, which were optimized and initiated within my master's thesis. The continuation of this includes the determination of various torsion angles based on *J* coupling constants and cross-correlated relaxation rates as well as the evaluation of NOESY spectra, which was subsequently rounded off by a structure calculation yielding a preliminary three-dimensional structure with an RMSD value of 0.9 Å.

4.1.1 MD simulations on RNA

With the constant increase in computational power, MD simulations of biomolecules are gaining importance. MD simulations shed light on biomolecular processes and ways of function as the movement of single atoms can be followed. As of today, MD simulations provide useful mechanistic insights of highly complex systems such as the ribosome[169] or the spliceosome[170]. For this kind of research, however, experimental data are indispensable to validate the simulation results on the one hand and to calibrate the underlying potential energy functions on the other hand. For this, NMR data can be valuable as is reflected in the fact that a comparison with NMR experiments was crucial to correctly describe the tendency to form α-helices in proteins[171–173].

This also holds true for the force field development for MD simulations of RNA. In particular, the structure of a 14 nts RNA with UUCG tetraloop with the PDB entry code 2KOC and its underlying wealth of experimental data[106] are commonly used to validate and calibrate potential energy functions. The 14 nts RNA is suitable for this purpose not only because of the plethora of available experimental data but also due to the tetraloop it enharbours. Tetraloops generally combine a high structural diversity with a low molecular size, which is important for the force field development as MD simulations are limited in computational power and struggle with the precise folding of non-canonical structural elements. Here, milestones were set by e.g. Chen *et al.* in 2013[174] and Giambaşu *et al.* in 2015[175], as they further improved existing potential energy functions and analyzed the remaining problems as for example the incorrect description in the backbone conformation of residues G9-G10[175]. Also, an overstabilization of helical regions still remains as a problem of current potential energy functions[176], which can only be approached with an increasing amount of experimental data especially for non-canonical structural elements of RNA. Moreover, a detailed characterization of potentially existing different states will be of importance, as latest results from MD simulations indicate that even for the supposedly rather rigid UUCG tetraloop a second minor populated state might exist[177]. The described sampling of motion in the UUCG tetraloop is further supported by independent studies, as will be raised in the next section[178–180].

4.1.2 The UUCG 14 nts RNA structures

The small 14 nts long RNA with UUCG tetraloop is not only used as a reference system for force field reparametrization but is also widely applied as a test system in the development of NMR methods and pulse sequences[77,181–183]. The complete resonance assignment was published in 2004 by Fürtig *et al.*[105] followed by the NMR structure in 2009 by Nozinovic *et al.* (PDB 2KOC, Figure 37 A)[106]. This solution NMR structure was not only determined based on mere distance restraints derived from NOE data but included torsion angle restraints derived from cross-correlated relaxation rates and *J*-

coupling constants. As a result, nearly a quarter of the 400 restraints for the structure calculation were torsion angle restraints. Finally, the measurement of residual dipolar couplings (RDCs) as well as the determination of NOE contacts involving the 2'-OH groups allowed the final high resolution structure determination with an overall RMSD of 0.37 Å[106].

However, with more than ten years having passed since the determination of the 14 nts RNA structure, new methods and improved force fields have been developed, unravelling discrepancies between the 2KOC structure and observations from promising new perspectives for future structure determinations. Criticism was raised for example in 2016 by Bermejo, Clore and Schwieters stating that the distance between base pairs in the stem of the 14 nts RNA is, with 4.33 Å, too large and that there is with more than 40% an unusual high number of backbone outliers[184]. This was significantly improved through structure calculations using Xplor-NIH[185,186] with the new force field RNA-ff1[184]. Another very promising development was the introduction of exact NOEs (eNOEs) that enable the definition of exact proton-proton distances rather than upper distance restraints in structure calculations[187]. An eNOE-based structure calculation of the 14 nts RNA with UUCG tetraloop was conducted in 2018 by Nichols *et al.*[188] and resulted in a high resolution structure featuring an overall RMSD of 0.44 Å (Figure 37 B). This accuracy is comparable to that obtained in a structure calculation based on conventional NOE contacts and dihedral angles[106], which underlines the impact of eNOE data for small RNAs. In this structural study, they even described the bundle using a two-state model (Figure 37 C red and blue), which decreased the target function compared to a one-state model.

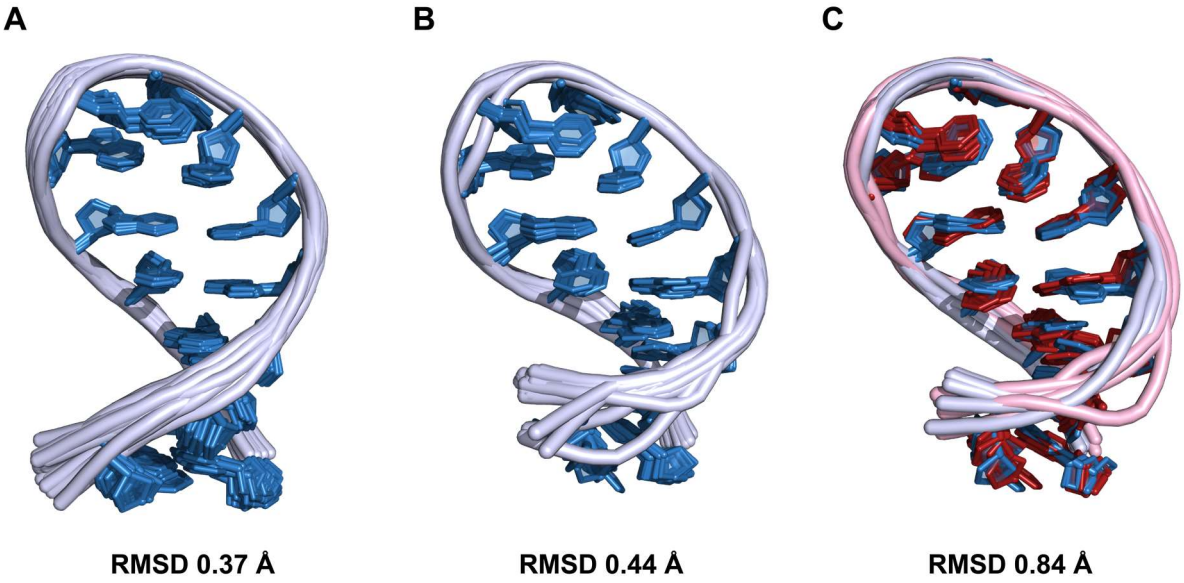


Figure 37 A The 2KOC[106], **B** the eNOE-only based single-state 6BY4[188] and **C** the eNOE-only based two-state 6BY5 (red and blue)[188] structure bundles. The overall RMSD is given for each structure.

As the primary goal of this study is to generate more reference systems, which are characterized in depth using state-of-the-art NMR methods, both approaches discussed

here, namely the Xplor-NIH based refinement and the application of eNOEs should be considered in this process.

4.1.3 New reference systems – an overview

In order to generate new reference systems for the reparametrization of potential energy functions for MD simulations, several considerations had to be made according to the desired properties of the system. For instance, the tetraloop should strike a balance between a dynamic character and a sufficient stability to allow high quality structural characterization by NMR spectroscopy. At the same time, it should be naturally occurring.

In order to fulfill the requirements, tetraloop stabilities were predicted using the mfold webtool[189] for all possible combinations of tetraloops (Figure 38). The predictions were conducted using a target sequence, where the tetraloop is embedded in the stem sequence of the 14 nts RNA with UUCG tetraloop.

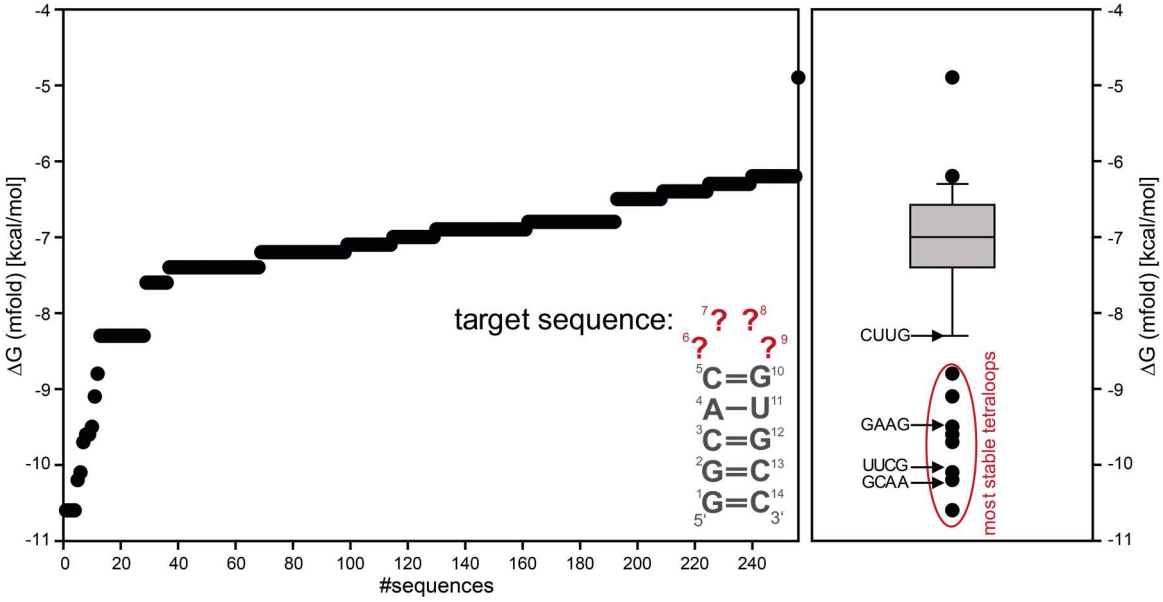


Figure 38 Stabilities for all 256 possible combinations of tetraloops in a 14 nts hairpin RNA with the given target sequence. The stabilities of the tetraloops under study here (UUCG, GAAG, GCAA and CUUG) are marked in the right plot. The gray box plot covers all tetraloop sequences that are not shown in individual data points. The stabilities were predicted with mfold[189].

In conjunction with the natural occurrence of different tetraloop sequences (Figure 39), three tetraloops were selected next to the commonly used UUCG tetraloop, namely the GCAA, GAAG and CUUG tetraloops. To generate a high diversity, the tetraloops differ in natural occurrence and stability. While the GCAA tetraloop for example combines a high natural abundance with a high predicted stability, the CUUG tetraloop is not among the most stable tetraloops and exhibits a moderate natural abundance.

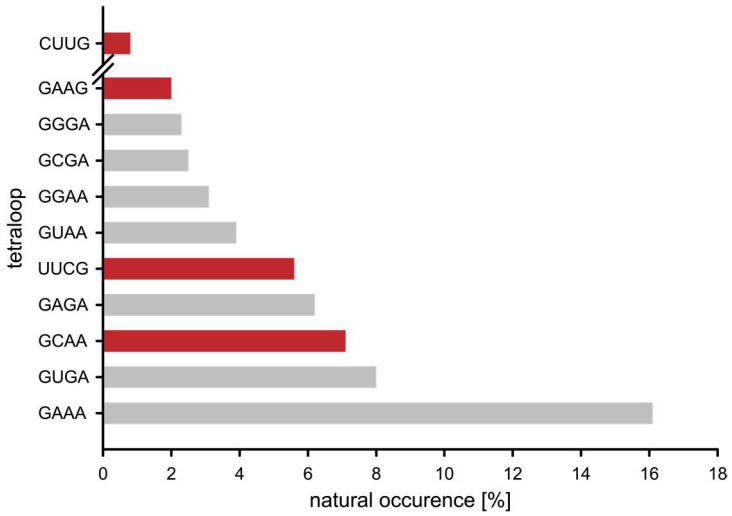


Figure 39 Natural occurrence of different tetraloops within a database of secondary structures that includes biologically relevant RNAs, such as rRNAs, tRNAs and introns[190,191]. Tetraloops, which are subject of this study as well as the previously studied UUCG tetraloop are shown in red.

Apart from the stability and natural abundance of the tetraloops, also the amount of the available experimental NMR data was considered in the selection process. The tetraloops under study have already been part of a structural characterization by NMR spectroscopy[109,192,193]. However, all structure calculations were conducted with a sparse set of experimentally derived restraints. As can be seen in Table 7, a maximum number of 13 distance and dihedral restraints per nucleotide have been used in the structure determinations. However, this study as part of the present PhD thesis, aims at around 25 restraints per residue, which is in the range of the number of experimental constraints available for the 14 nts RNA with UUCG tetraloop.

Table 7 Overview of the experimental parameters used in the structure determination of the UUCG[106], GAAG[109], GCAA[192] and CUUG[193] tetraloops. Distance restraints and dihedral restraints were combined in the calculation of the number of restraints per nucleotide. [a] The 32 RDCs were also included in this calculation. The overall RMSD is given for each structure.

tetraloop	length	distance restraints	dihedral restraints	restraints per residue	RMSD
UUCG	14 nts	267	96	~29[a]	0.4 Å
GAAG	12 nts	127	34	~13	1.0 Å
GCAA	12 nts	86	33	~10	0.6 Å
CUUG	12 nts	82	23	~9	1.3 Å

The structure bundles for the four tetraloops in Figure 40 as well as the RMSD (Table 7) show that the deviation between the different structures is smallest for the UUCG tetraloop. Furthermore, it is evident that the tetraloop structures of the GAAG and GCAA tetraloops are similar (Figure 40 B and C). Both loops adopt a conformation, which is typical for GNRA tetraloops and features π–π-stacking of three nucleobases and a GA base pair[194], even though the GAAG tetraloop does technically not belong to that class. Remarkably, in relation to the number of restraints that were used to determine the RNA structures featuring GAAG, GCAA and CUUG tetraloops, the precisions (RMSD values) of

the tetraloop structures are rather high. This might be indicative of a bias caused by the used force fields on the final RNA structure bundles.

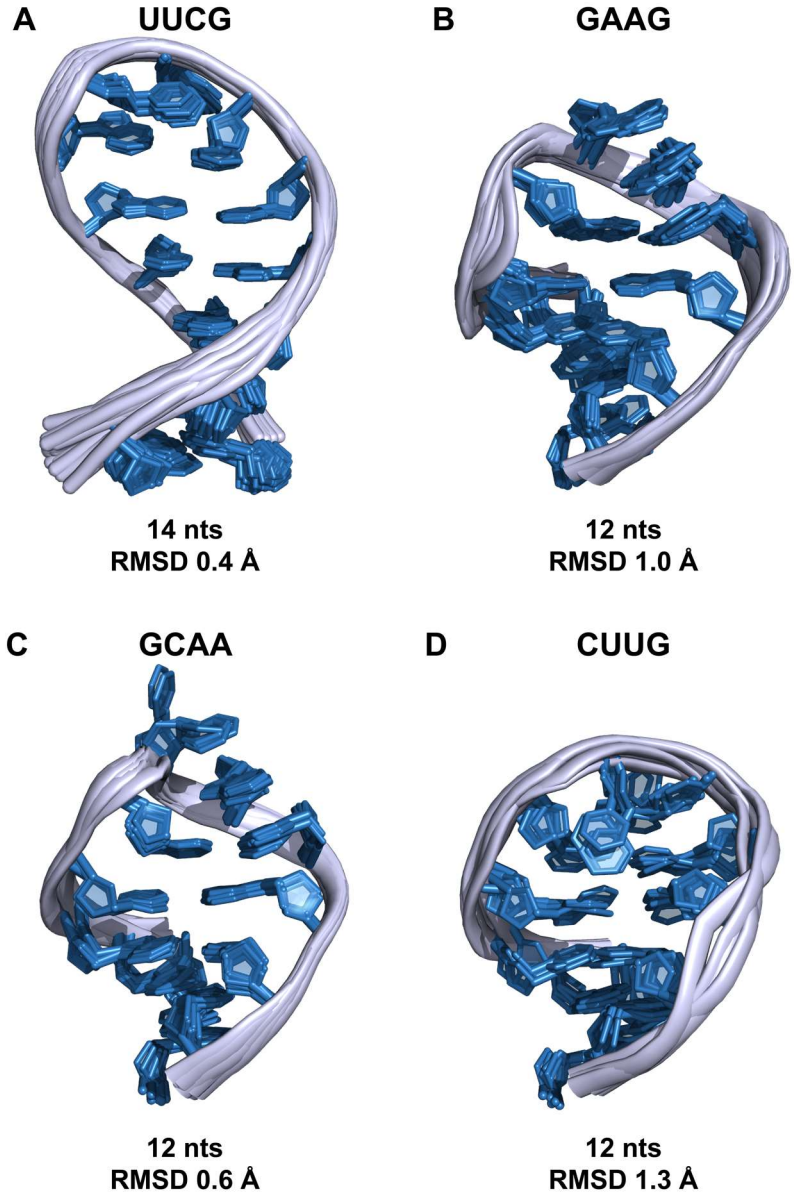


Figure 40 Structure bundles for **A** the UUCG (2KOC)[106], **B** the GAAG (2F87)[109], C the GCAA (1ZIH)[192] and **D** the CUUG (1RNG)[193] tetraloops. The number of nucleotides and the RMSD of the NMR structures are given for each RNA.

4.2 Materials and Methods

4.2.1 The RNAs under study

Three 14 nts RNAs with different tetraloop sequences, namely cGAAGg, cGCAAg and gCUUGc, were studied within this project (Figure 41). Subject of this thesis was the structural characterization of the GAAG tetraloop, while the GCAA and CUUG tetraloops were studied by György Pintér and Andreas Oxenfarth, respectively.

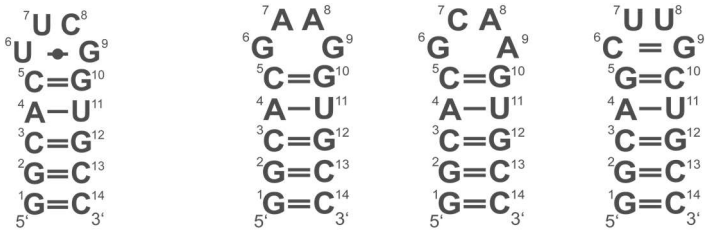


Figure 41 Secondary structures of the 14 nts RNAs with different tetraloops. The reference system with UUCG tetraloop (left) is thoroughly characterized, 14 nts RNAs with GAAG, GCAA (György Pintér) and CUUG (Andreas Oxenfarth) tetraloops are structurally characterized within this project.

4.2.2 Sample preparation

In the following, the process of sample preparation is briefly described for the 14 nts RNA with GAAG tetraloop. This process was optimized within my master thesis and thus this chapter only includes a summary of the optimized conditions.

The 14 nts RNA was transcribed from plasmid DNA via *in vitro* transcription using T7 RNA polymerase (P266L mutant)[148]. The vector contained an ampicillin resistance and a self-cleaving HDV ribozyme to generate 3'-end homogeneity[115] as depicted in Figure 42. This plasmid will be referred to as GAAG plasmid DNA in the following.

Figure 42 Sequence of the insert in vector pUC57 featuring the restriction sites EcoRI and SmaI (blue and yellow), the T7 promotor (green), the transcribed sequences for the 14 nts GAAG RNA and the self-cleaving HDV ribozyme[115]. The plasmid DNA was obtained commercially (Gene Script, USA).

Transformation

50 µl of competent *E.coli* DH5α cells (New England Biolabs, USA) were supplemented with 100 ng of GAAG plasmid DNA. The mixture was incubated on ice for 30 min. After a heat shock for 45 s at 42°C during which the plasmid DNA is transferred into the cells, the mixture was incubated on ice for 10 min. The cells were supplemented with 950 µl LB medium (Carl Roth, Germany; 5 g/l yeast extract, 10 g/l tryptone, 10 g/l NaCl) and incubated for 1 h at 37°C. 50 µl and 100 µl of the mixture were plated on LB agar plates

(LB medium, 15 g/l agar and 100 μg/ml ampicillin). The plates were incubated at 37°C for 12 h.

Amplification of the GAAG plasmid DNA

4 ml of LB medium including 100 μg/ml ampicillin was inoculated with a single clone of cells from the agar plates and incubated for 6 h at 37°C and 120 rpm. The cells were harvested at 4,000 g and 4°C for 5 min, washed with 1 ml LB medium and resuspended in 1 ml LB medium.

2 l autoclaved LB medium was supplemented with 100 μg/ml ampicillin and 1 ml of the cell suspension. The mixture was incubated at 120 rpm and 37°C for 12 h. Cells were harvested at 4,000 g and 4°C for 15 min.

Purification of the GAAG plasmid DNA

To extract the plasmid DNA a NucleoBond plasmid purification kit (Macherey-Nagel, Germany) was used. The cell pellet was resuspended in 90 ml of buffer S1 (50 mM TRIS, 10 mM EDTA, 100 μg/ml RNase A, pH 8.0 with HCl). Cells were lysed with 90 ml of buffer S2 (200 mM NaOH, 1% SDS (w/v)) for 5 min after which proteins and cell fragments were precipitated through addition of 90 ml of ice-cold buffer S3 (2.8 M KOAc, pH 5.1 with AcOH). The mixture was incubated for 10 min and centrifuged at 12,000 g and 4°C for 50 min. The pellet was discarded and the supernatant filtered. Megaprep Nucleobond AX2000 columns (1 per l cell culture) were equilibrated with 20 ml buffer N2 (100 mM TRIS, 15% EtOH (v/v), 700 mM KCl, 1.5% TritonX100 (v/v), pH 6.3 with H_3PO_4). The filtrate was loaded onto the columns and the columns were washed with 2 x 35 ml buffer N3 (100 mM TRIS, 15% EtOH (v/v), 1.15 M KCl, pH 6.3 with H_3PO_4). Through an increase in pH, the GAAG plasmid DNA was eluted with 25 ml of buffer N5 (100 mM TRIS, 15% EtOH (v/v), 1 M KCl, pH 8.5 with H_3PO_4). The eluate was supplemented with 18 ml isopropyl alcohol and centrifuged at 10,000 g and 4°C for 30 min to precipitate the GAAG plasmid DNA. The supernatant was discarded, the pellet was washed with 1 ml ice-cold 70% EtOH and air-dried. The plasmid DNA was analyzed via analytical agarose gel electrophoresis, via UV/vis spectroscopy (NanoDrop photometer) and sequencing (Eurofins, Germany).

Linearization

In preparation of the subsequent *in vitro* transcriptions, the plasmid DNA was linearized using restriction enzyme SmaI. Here, it is vital that the digestion occurs quantitatively, as circular plasmid DNA leads to a synthesis of long RNA fragments through the T7 RNA

polymerase, as it does not fall off the DNA template. The digestion was conducted with 120 units of SmaI per mg DNA and in the presence of CutSmart Buffer (New England Biolabs, USA, 50 mM KOAc, 20 mM TRIS-OAc, 10 mM MgOAc, 100 μg/ml BSA, pH 7.9) for 12 h at room temperature. The completeness of the digestion was controlled via analytical agarose gel electrophoresis. If necessary, the reaction was incubated for another 4 h upon addition of more enzyme (units depending on amount of circular plasmid DNA). The enzyme was removed in a phenol-chloroform-isoamylalcohol extraction (PCI extraction), where the reaction mixture was washed twice with 1 sample volume of PCI solution followed by washing with 1 sample volume of chloroform to remove traces of phenol. The DNA was precipitated with 0.3 M NaOAc (pH 5.5) and 2.5 sample volumes of EtOH (abs.). After incubation at -20°C for 12 h the mixture was centrifuged for 20 min at 4,000 g and 4°C. The supernatant was discarded, the pellet air-dried and resolved in ddH_2O. The concentration was determined via UV/vis spectroscopy (NanoDrop photometer) and the purity was analyzed with analytical agarose gel electrophoresis. If necessary, the stock solution was diluted to approximately 2000 ng/μl.

In vitro transcriptions

Typical reaction conditions for the *in vitro* transcription are summarized in Table 8. Whenever new stock solutions were used, the reaction conditions were reevaluated in a 25 μl volume test transcription. Preparative transcriptions were conducted in a 25 ml scale using either not isotope-labeled rNTPs (Carl Roth, Germany) or ^{13}C,^{15}N-labeled rNTPs (Silantes, Germany).

Table 8 Reaction conditions for the *in vitro* transcription of the 14 nts GAAG RNA. Conditions have been optimized within my master thesis. In total 6 mM rNTPs have been used and concentrations of the single rNTPs have been adapted according to the percentage in the transcribed sequence (written in brackets).

component	concentration
TRIS glutamate buffer	2.5x
DTT	20 mM
spermidine	2 mM
DNA	750 μg
$Mg(OAc)_2$	25 mM
rATP (17%)	1.02 mM
rUTP (17%)	1.02 mM
rGTP (35%)	2.10 mM
rCTP (31%)	1.86 mM
T7 RNA polymerase	9.6 μg/ml

RNA purification

The RNA was purified with a DEAE chromatography and a subsequent HPLC purification, which was conducted by Elke Stirnal.

DEAE Sepharose (GE Healthcare, United Kingdom) was treated with 0.1% DEPC solution for 12 h to inactivate RNases. The column was washed with steaming water and equilibrated with 0.1 M NaOAc solution (pH 5.5). The transcription was loaded onto the column and washed with 5 x 10 ml of 0.6 M, 1 M and 2 M NaOAc solutions (pH 5.5). The RNA is eluted with increasing salt concentration from 0.6 M to 2 M. In order to remove potential remaining RNA traces from the column after the RNA purifications, it was washed with 3 x 10 ml of 3 M NaOAc solution (pH 5.5). The RNA content was analyzed with UV/vis spectroscopy and analytical polyacrylamide gel electrophoresis (PAGE). RNA molecules in fractions that contain 14 nts RNA were precipitated by adding 4 sample volumes of EtOH (abs.). The NaOAc concentration was diluted to 0.6 M. Precipitation occurred at -20°C for 12 h. After centrifugation at 8000 g and 4°C for 45 min, the pellets were air-dried and dissolved in ddH_2O.

The RNA was further purified by HPLC chromatography, which was conducted by Elke Stirnal. The purification was conducted at 60°C using a Perfectsil RP18 column (10 x 250 mm) and a 50 mM potassium phosphate buffer (buffer A). The RNA was eluted using a gradient (Table 9) with increasing amounts of acetonitrile (buffer B: 60% acetonitrile, 50 mM potassium phosphate).

Table 9 Gradient of the HPLC chromatography. The purification was conducted at a constant flow rate of 5 ml/min.

time [min]	buffer A [%]	buffer B [%]
0	100	0
5	65	35
30	62	38
35	0	100
40	0	100
43	100	0
55	100	0

The purified RNA was lyophilized and washed with ddH_2O in a centrifugal concentrator (Vivaspin, GE Healthcare, 1k molecular weight cut-off (MWCO)) and subsequently precipitated upon addition of 5 sample volumes of 2% $LiClO_4$ solution in acetone to remove remaining HPLC buffer. The precipitation with $LiClO_4$ was repeated three times. The RNA was transferred into NMR buffer (50 mM potassium phosphate, pH 6.4) and concentrated to a volume of 270 µl. Depending on the type of required sample, either 5% D_2O was added or the solution was lyophilized and refilled with 270 µl D_2O (99.95%, process repeated three times). RNA folding occurred through heating the NMR sample at 95°C for 5 min and subsequently cooling down on ice. The final RNA concentration was determined by UV/vis spectroscopy (NanoDrop photometer). The purity of the sample was analyzed via analytical denaturing PAGE and the homogeneity of the RNA fold was tested in an analytical native PAGE. The NMR sample was transferred into a Shigemi tube which was previously treated with 0.1% DEPC solution. An overview of the different

GAAG NMR samples used within this project and their RNA concentration is given in Table 10.

Table 10 Summary of the different 14 nts GAAG NMR samples, which were used in the course of this project. Labeling schemes, RNA concentrations and solvents are given.

name	labeling scheme	concentration	solvent
sample#1	not labeled	1.26 mM	95% H_2O/ 5% D_2O
sample#2	not labeled	0.56 mM	100% D_2O
sample#3	uniformly ^{13}C,^{15}N	1.24 mM	95% H_2O/ 5% D_2O
sample#4	uniformly ^{13}C,^{15}N	0.64 mM	100% D_2O

4.2.3 NMR spectroscopy

All NMR spectra were recorded on Bruker NMR spectrometers equipped with cryogenic probes as listed in Table 11. If not indicated otherwise, the spectra were recorded at 298 K. ^1H chemical shifts were referenced externally to DSS and ^{13}C and ^{15}N chemical shifts were indirectly referenced from the ^1H chemical shift as described earlier[195]. Spectra were processed using Topspin (version 3.5 pl6) and further evaluated using Sparky[117].

Table 11 Overview of the different NMR spectrometers with the respective field strength and probe. The NMR experiments which have been conducted within this project are given for the respective NMR spectrometer.

field strength [MHz]	probe	experiments
599.88	cryo TCI ^1H[^{13}C,^{15}N]	TROSY rel. HCCH-COSY, Γ-HCNCH
600.13	cryo TCI ^1H[^{13}C,^{15}N]	^{13}C,^{15}N-HSQC
600.17	Prodigy TCI ^1H/^{19}F[^{13}C,^{15}N]	coupled ^1H,^{13}C-HSQC, Γ-HCCH, ^1H,^1H-NOESY (50, 100, 150 ms, 100% D_2O), ^1H,^1H-NOESY (50, 100, 150 ms, 95% H_2O/ 5% D_2O)
600.03	cryo TCI ^1H[^{13}C,^{31}P]	P-FIDS
700.17	cryo QCI ^1H[^{13}C,^{15}N,^{31}P]	^{13}C,^{15}N-HSQC, lr-^1H,^{15}N-HSQC, lr-^1H,^{13}C-HSQC, Γ-HCNCH, *fwd* HCC-TOCSY-CCH-E.COSY, C5'-selective ^1H,^{13}C-HSQC

Chemical shift perturbations (CSPs) were calculated from ^1H,^{13}C-HSQC spectra between cold (274 K) and hot (308 K) temperatures using the following equation with a scaling factor of 0.25 for ^{13}C-chemical shifts according to the gyromagnetic ratio[196]:

$$\Delta\delta = \sqrt{(\Delta\delta_H)^2 + 0.25 \cdot (\Delta\delta_C)^2} \qquad [16]$$

4.2.4 Structure calculation

Structure calculations were performed using ARIA (version 1.2)[197] including CNS 1.1. The force field was "dna_rna_allatom". NOE distances were calibrated and corrected for spin diffusion using standard protocols. Structure calculations were performed starting from a linear template and generating 100 starting structures. In an iterative process, the

NOE distances were successively calibrated until the tolerance for distance violations was 0.1 Å in the final iteration cycle number eight. In this final run, 200 structures were calculated, out of which the 20 structures with the lowest energy were finalized in a water refinement. Furthermore, a solvent exchange correction was applied for cross peak intensities from NOESY spectra, which were recorded on RNA predominantly solved in H_2O (95% H_2O, 5% D_2O). Additionally, the stereospecific assignment of the H5' protons was allowed to be swapped during the structure calculation. Structures were visualized using Pymol (DeLano Scientific LLC; Schrödinger, Inc).

4.3 Results and Discussion

In the following results and discussion section, the results for the 14 nts RNA with GAAG tetraloop as product of this PhD thesis, will be described in conjunction with the data from the UUCG tetraloop as reference system as well as with the GCAA and CUUG tetraloops. All data shown for the latter two tetraloops were generated by György Pintér and Andreas Oxenfarth, respectively, and are shown here only for comparison purposes. Furthermore, the contributions are marked at the appropriate positions.

4.3.1 Resonance assignment of the 14 nts RNA with GAAG tetraloop

An extensive ^{1}H, ^{13}C and ^{15}N resonance assignment of the 14 nts RNA with GAAG tetraloop was already achieved within my master thesis. During this PhD project, this was completed to a full resonance assignment including the quaternary carbon and tertiary nitrogen atoms of the nucleobase, which will be described in the following separately for the four nucleobases.

Generally, the assignment was started by correlating an atom with an assigned chemical shift, like the aromatic protons (H6 or H8 for pyrimidines and purines) or the imino protons of uridines and guanosines, with a quaternary carbon or the tertiary nitrogen atoms. This was then supplemented with a variety of ^{13}C,^{15}N-HSQC experiments using different coupling constants for the calculation of the INEPT transfer time as well as the C(N)H-HDQC experiment (introduced in Chapter II).

For the purine nucleobases, N7 and N9 nitrogen atoms were assigned based on the H8 chemical shifts using a long-range ^{1}H,^{15}N-HSQC experiment (Figure 43 and Figure 44) with a transfer delay according to a coupling constant of 15 Hz. Using ^{13}C,^{15}N-HSQC and long-range ^{1}H,^{13}C-HSQC spectra, C4 and C5 atoms can be assigned. For adenosines this can be complemented with the correlations between H2 and C4 as well as C6 in a long-range ^{1}H,^{13}C-HSQC spectrum (Figure 43).

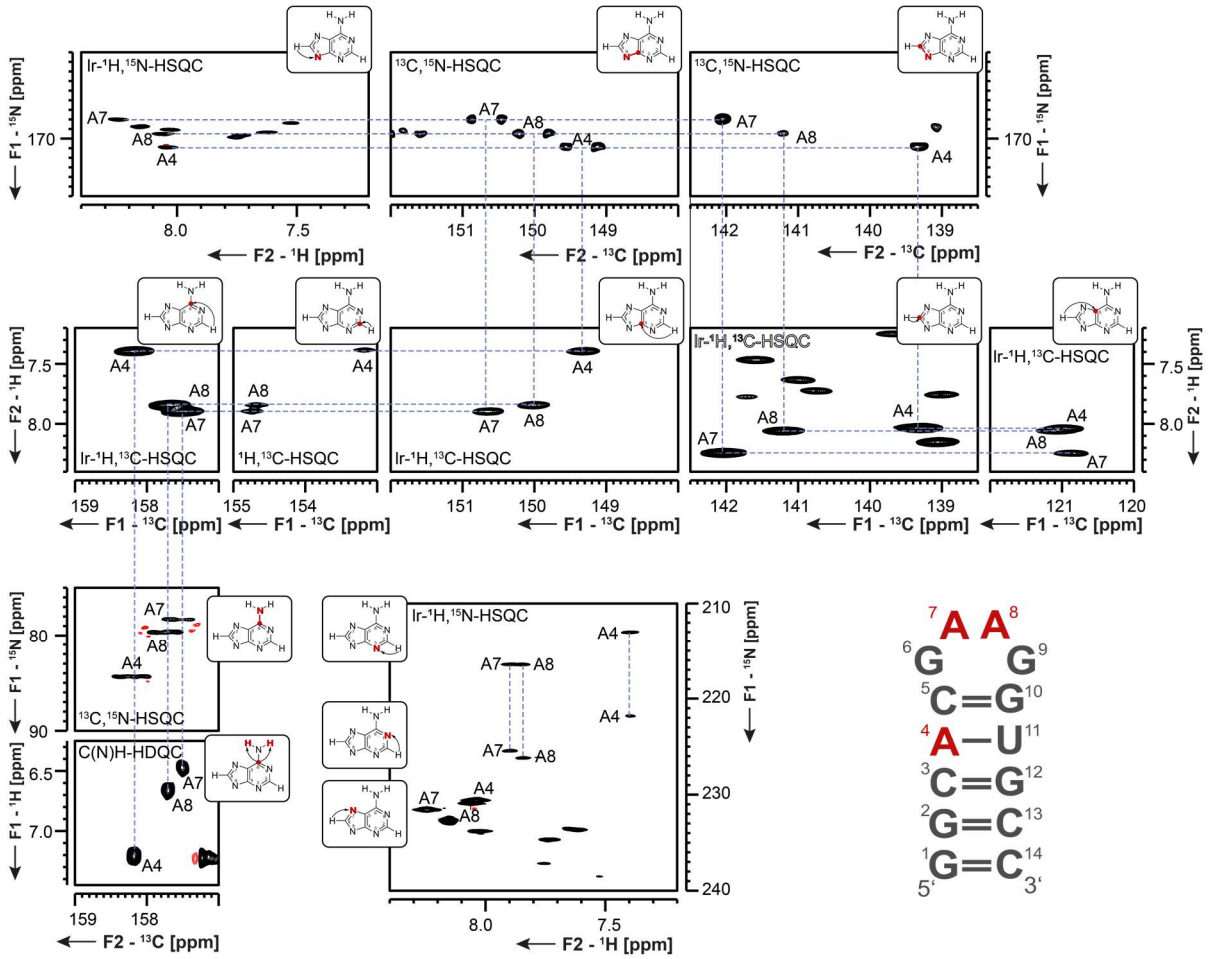


Figure 43 Nucleobase resonance assignment of the tertiary nitrogen and quaternary carbon atoms in adenosines of the 14 nts RNA with GAAG tetraloop using long-range $^{1}H,^{15}N$-HSQC and $^{1}H,^{13}C$-HSQC, $^{13}C,^{15}N$-HSQC[76] and C(N)H-HDQC[78] spectra. Positive signals are held in black while negative resonances are colored in red. Resonances that are not relevant for the adenosine assignment are shown in gray. The walk through the C-N fragments is indicated with blue dashed lines. The correlations, which are visible in the respective spectrum, are highlighted in the adenine structure next to the spectrum. The secondary structure of the 14 nts RNA highlights all adenosine residues in red. Assignments are given. Experimental details can be taken from Appendix 3 chapter 7.1.1 and the caption of Figure 61.

$^{13}C,^{15}N$-HSQC experiments with different transfer delays were then used to assign the remaining C-N fragments C6N1, N1C2, C2N3 and C2N2 in guanosines (Figure 44). The nucleobase assignment was completed with the assignment of the ^{1}H-DQ chemical shifts of the amino protons using the ^{13}C-detected C(N)H-HDQC experiment (see also Chapter II). The assignment was verified using the H8C8N9C4 experiment[105], the TROSY relayed HCCH-COSY experiment[39] as well as imino and amino HNCO type experiments (data not shown).

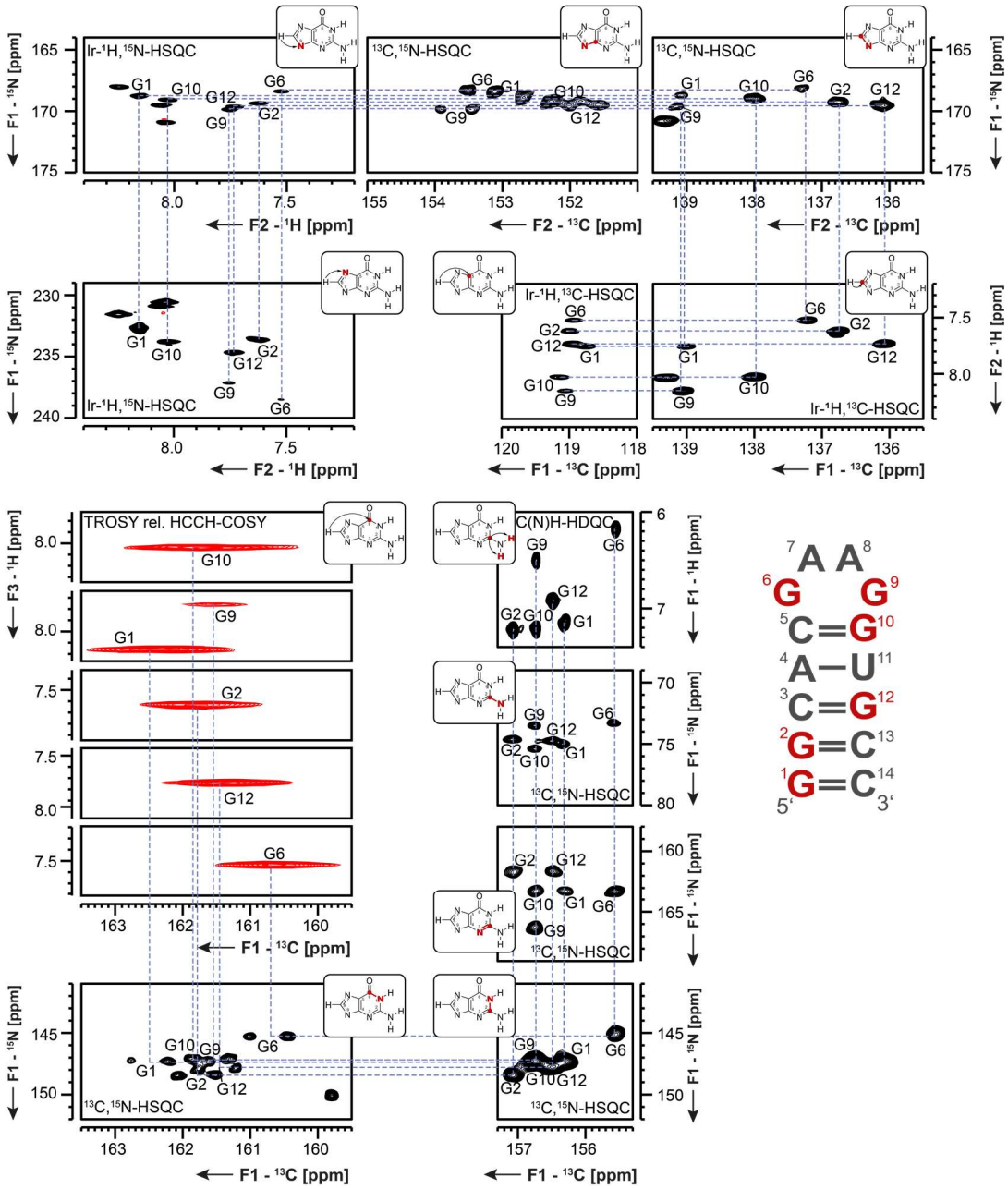


Figure 44 Nucleobase resonance assignment of the tertiary nitrogen and quaternary carbon atoms in guanosines of the 14 nts RNA with GAAG tetraloop using long-range $^{1}H,^{15}N$-HSQC and $^{1}H,^{13}C$-HSQC, $^{13}C,^{15}N$-HSQC[76], C(N)H-HDQC[78] and 3D TROSY relayed HCCH-COSY experiments[39]. Positive signals are held in black while negative resonances are colored in red. Resonances that are not relevant for the guanosine assignment are shown in gray. The walk through the C-N fragments is indicated with blue dashed lines. The correlations, which are visible in the respective spectrum, are highlighted in the guanine structure next to the spectrum. The secondary structure of the 14 nts RNA highlights all guanosine residues in red. Assignments are given. Experimental details can be taken from Appendix 3 chapter 7.1.1 and the caption of Figure 61.

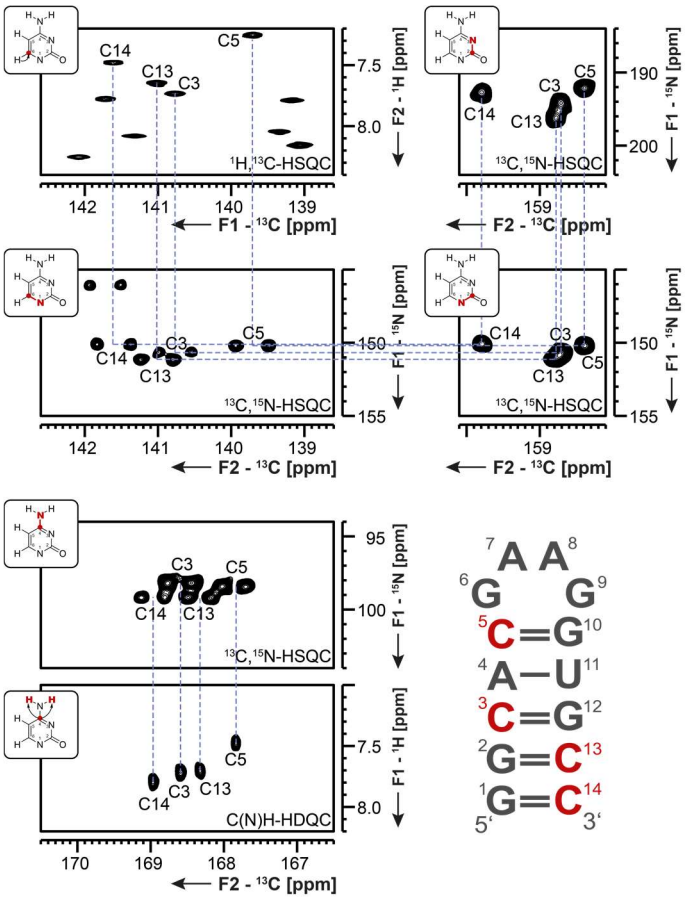


Figure 45 Nucleobase resonance assignment of the tertiary nitrogen and quaternary carbon atoms in cytidines of the 14 nts RNA with GAAG tetraloop using ^{1}H,^{13}C-HSQC, ^{13}C,^{15}N-HSQC[76] and C(N)H-HDQC[78] experiments. Positive signals are held in black while negative resonances are colored in red. Resonances that are not relevant for the cytidine assignment are shown in gray. The walk through the C-N fragments is indicated with blue dashed lines. The correlations, which are visible in the respective spectrum, are highlighted in the cytosine structure next to the spectrum. The secondary structure of the 14 nts RNA highlights all cytidine residues in red. Assignments are given. Experimental details can be taken from Appendix 3 chapter 7.1.1 and the caption of Figure 61.

The pyrimidine chemical shift assignment was initiated with correlations to the known chemical shift of the C6 resonance in ^{13}C-detected ^{13}C,^{15}N-HSQC spectra (Figure 45 and Figure 46). The C-N fragments of the pyrimidines were further linked with ^{13}C,^{15}N-HSQC spectra using different times for the INEPT-transfer. With this the following linkages are achieved: C6N1, N1C2, C2N3, N3C4 (only in uridines) and C4N4 (only in cytidines). The ^{1}H-DQ chemical shifts of the amino protons in cytidines were obtained using the ^{13}C-detected C(N)H-HDQC experiment. The pyrimidine nucleobase assignment was further complemented using amino and imino HNCO type experiments (data not shown).

The overall chemical shift assignment for the 14 nts RNA with GAAG tetraloop is given in Table 18 - Table 20 in the Appendix 3. Based on the complete resonance assignment, the structural characterization of the 14 nts RNA with GAAG tetraloop was conducted.

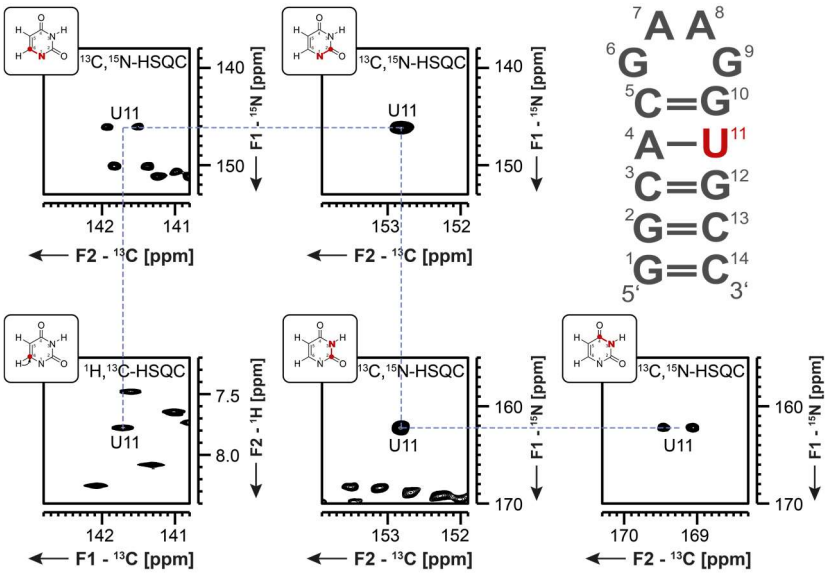


Figure 46 Nucleobase resonance assignment of the tertiary nitrogen and quaternary carbon atoms in the uridine of the 14 nts RNA with GAAG tetraloop using ^{1}H,^{13}C-HSQC and ^{13}C,^{15}N-HSQC[76] experiments. Resonances that are not relevant for the uridine assignment are shown in gray. The walk through the C-N fragments is indicated with blue dashed lines. The correlations, which are visible in the respective spectrum, are highlighted in the uracil structure next to the spectrum. The secondary structure of the 14 nts RNA highlights the single uridine residue in red. Assignments are given. Experimental details can be taken from Appendix 3 chapter 7.1.1.

4.3.2 RNA conformation – external influences

The RNA fold is often influenced by external effects like temperature, pH or ionic strength. Here, in particular Mg^{2+} ions are crucial for the RNA structure and function. To obtain information about the potential influences of these external effects ^{1}H,^{13}C-HSQC spectra were recorded under different experimental conditions; namely at different temperatures as well as Mg^{2+} ion concentrations.

A ^{1}H,^{13}C-HSQC temperature series was conducted for all tetraloops under study including the 14 nts RNA with UUCG tetraloop (Figure 47). The spectra of the UUCG as well as the CUUG tetraloop show no significant temperature dependence of the chemical shifts. However, in the ^{1}H,^{13}C-HSQC spectra of the GAAG and the GCAA tetraloop RNAs, two effects are observed. First, signals, which correspond to residues of the tetraloops, are shifting with varying temperatures. Furthermore, several signal intensities decrease beyond detectability at lower temperatures (e.g. A7 for GAAG or C7 for GCAA). This is evidence for the presence of dynamics in the GAAG and GCAA tetraloops that drift into the intermediate exchange regime on the NMR timescale at lower temperatures. For both mentioned tetraloops, it was not possible to reach the slow exchange regime, neither by measuring at temperatures as low as 268 K nor by conducting experiments at a 1.1 GHz NMR spectrometer at 268 K (GAAG tetraloop, data not shown).

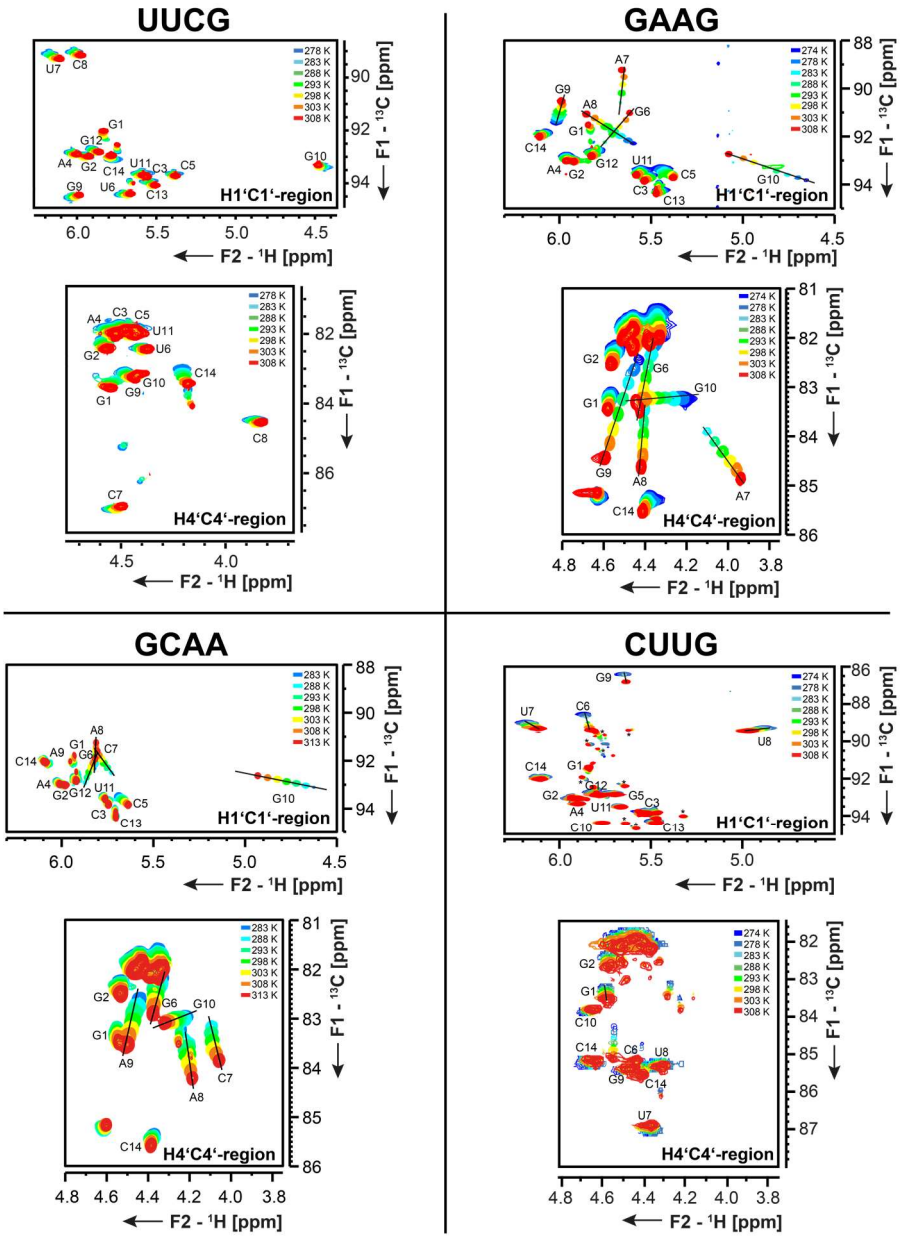


Figure 47 ^{1}H,^{13}C-HSQC spectra for the temperature series for the 14 nts RNAs with UUCG (upper left), GAAG (upper right), GCAA (lower left) and CUUG (lower right) tetraloops. For each RNA H1'C1'- and H4'C4'-regions are shown. Peaks are color coded in a rainbow gradient from cold (blue) to hot (red) temperatures. All spectra were recorded using a standard constant-time ^{1}H,^{13}C-HSQC experiment with the maximum achievable resolution in the indirect dimension at 600 MHz.

This suggests that both, the CUUG and the UUCG tetraloops are more rigid than the GAAG and GCAA tetraloops. These dynamics are quantified for the GAAG tetraloop using the CSPs calculated from the ribose ^{1}H and ^{13}C chemical shifts according to equation [16] in chapter 4.2.3 (Figure 48). Here, a clear increase in $\Delta\delta$ above the threshold (dashed line) is observed for almost all ribose positions of the tetraloop nucleotides being indicative of motion throughout the whole ribose moiety of all tetraloop nucleotides.

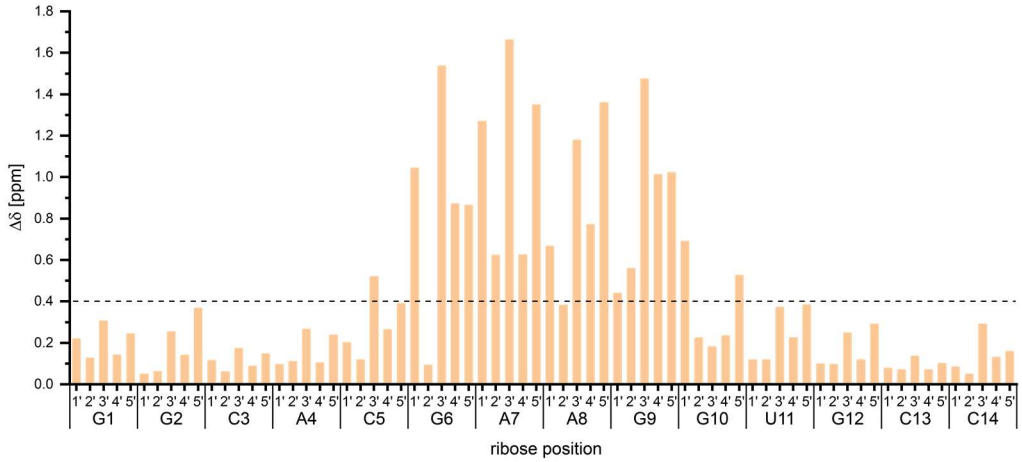


Figure 48 Chemical shift perturbations (CSPs) between cold (274 K) and hot (308 K) temperatures for all ribose resonances plotted against the sequence of the of the 14 nts with GAAG tetraloop. CSPs were calculated using equation [16] in chapter 4.2.3[125]. The threshold value of 0.4 ppm (dashed horizontal line) equals the standard deviation (1*σ) of the mean CSP.

The present chemical shift changes are notably distinct in the ^{13}C-dimension (e.g. GAAG G9 ~2 ppm for H4'C4), which might be indicative of dynamics in the ribose pucker. In order to test this hypothesis, the temperature dependent canonical coordinates[198,199] were calculated for all tetraloops under study (Figure 49). The canonical coordinates represent an empirical way to calculate the exocyclic torsion angle γ and the pseudorotation phase P from the carbon chemical shifts of the ribose. The latter is a combination of the ribose torsion angles and adopts 0° in the C3'- and 180° in the C2'-endo conformation[200]. However, as the canonical coordinates are usually used for a qualitative determination of the ribose pucker, the border between the two ribose conformations is set at ~70°[199]. For the more rigid UUCG and CUUG tetraloops, the plots of the canonical coordinates show indeed no change in the ribose conformation at different temperatures. Here, the coordinates of the UUCG tetraloop only shift a minimum, while the CUUG tetraloop shows more distinct shifts in the canonical coordinates that are nevertheless still comparable to the shifts of the canonical nucleotides. By contrast, for the GCAA and GAAG tetraloops the canonical coordinates change significantly with rising temperatures. Here, remarkable changes between 50° - 100° are observed in the pseudorotation phase (P_{FIT}), suggesting that most of the nucleotides that are part of the tetraloop cross the border between C3'-endo and C2'-endo conformation. This is indicative of a temperature dependent equilibrium between C2'- and C3'-endo conformation that is strikingly shifted towards the non-canonical C2'-endo conformation at higher temperatures. Another interesting observation was made for residue G10 of the closing base pair, which shows a different temperature-dependence of the canonical coordinates than the nucleotides of the tetraloop. This difference is particularly evident in the angle γ, which is decreasing with higher temperatures for residue G10, while it is increasing for all other nucleotides. Predominantly, this is observed for residue G10 of the GAAG tetraloop and might be indicative of a hinge function which is carried out by this residue. As the canonical

coordinates are an empirical chemical shift based measure of the ribose pucker, this hypothesis has to be further validated through directly measurable observables such as ^{3}J(H,H) coupling constants or $\Gamma_{CH,CH}$ cross-correlated relaxation rates (chapter 4.3.3).

Furthermore, it should be mentioned that this temperature-dependence is not only limited to the ribose moiety but is also observed in ^{1}H,^{13}C-HSQC spectra of the aromatic region, where for the GAAG tetraloop predominantly resonances of residues G6, A8 and G9 are shifting with different temperatures (Appendix 3 Figure 66). This indicates that the temperature-dependent structural changes also include a rearrangement of the nucleobases resulting in e.g. a different stacking behavior.

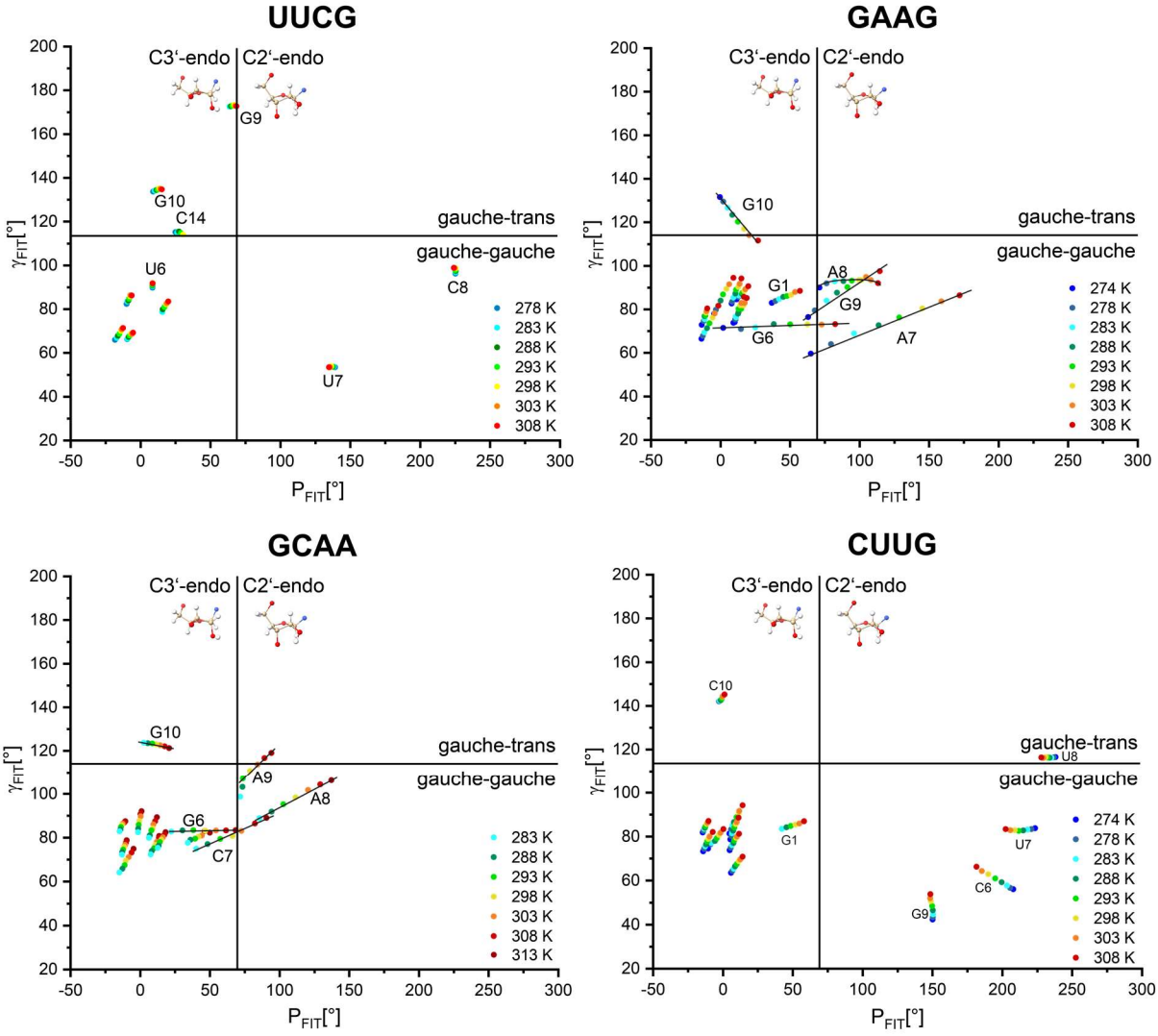


Figure 49 Temperature-dependent canonical coordinates[198,199] for the 14 nts RNAs with UUCG (upper left), GAAG (upper right), GCAA (lower left) and CUUG (lower right) tetraloops. The borders for the different conformations of the exocyclic torsion angle γ (gauche-gauche or gauche-trans) as well as the ribose pucker (C3'-endo and C2'-endo) are given. Temperatures are rainbow color coded from blue (cold) to red (hot).

Due to the polyanionic character of RNA, its fold is especially dependent on the presence of cations[201,202]. Here, in particular Mg^{2+} ions are crucial as they not only stabilize the RNA structure in general[203], but are also indispensable for many regulatory[204] or

catalytically active RNAs[205]. Therefore, a titration with $MgCl_2$ was performed for the 14 nts RNA with GAAG tetraloop (Figure 50 B). From ^{1}H-1D spectra of the imino region (data not shown) it is apparent that there is no significant structural rearrangement observable upon addition of Mg^{2+} ions. At a closer inspection of the ^{1}H,^{13}C-HSQC spectra, it is evident that the presence of an increasing concentration of Mg^{2+} ions has a similar effect on the RNA signals as decreasing temperatures (Figure 50 A and B). This observation is in line with expectations as both, low temperatures and Mg^{2+} ions, are expected to rigidify the system. The only difference here is that upon addition of Mg^{2+} ions the equilibrium between different conformations is not passing the intermediate exchange regime, since all resonances are observable even at high Mg^{2+} concentrations (corresponding to the low temperature conformation). This is also reflected in a temperature series at very high Mg^{2+} concentrations (18 mM), where resonances are observable even at 274 K (Figure 50 C). At this extreme Mg^{2+} concentration, the chemical shift changes with decreasing temperature are not as distinct as for the RNA in the absence of Mg^{2+} ions (Figure 50 A vs. C). This is indicative of the formation of a final more rigid structure at 274 K and at 18 mM Mg^{2+} ions.

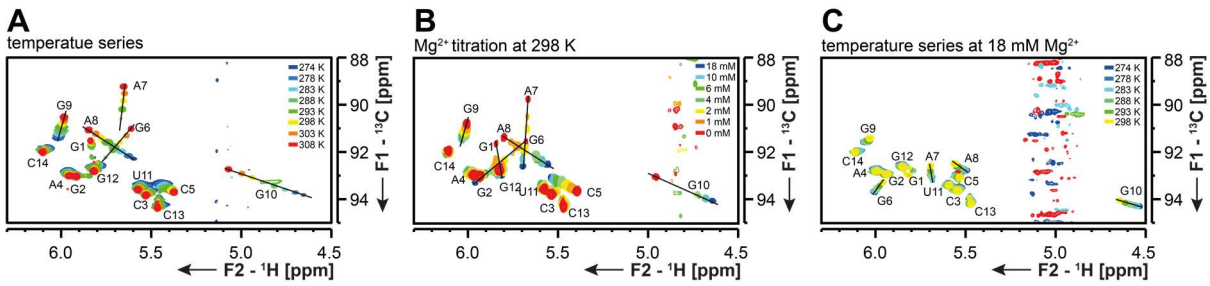


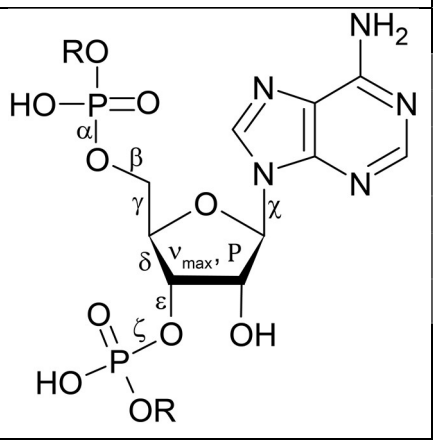
Figure 50 ^{1}H,^{13}C-HSQC spectra of the 14 nts GAAG RNA under different external conditions. **A** Temperature series from 308 K (red) to 274 K (blue) as also shown in Figure 47. **B** Mg^{2+} titration from 0 mM (red) to 18 mM (blue) $MgCl_2$. **C** Temperature series at 18 mM $MgCl_2$ from 298 K (yellow) to 274 K (blue). All spectra were recorded using standard ^{1}H,^{13}C-HSQC sequences featuring a constant-time evolution period.

As the aim of this study was the characterization of a more dynamic system compared to the UUCG tetraloop RNA and as the chemical shift dispersion decreases massively upon addition of Mg^{2+} ions, the structural characterization of all tetraloops was carried out in the absence of Mg^{2+} ions at elevated (298 K – 308 K) temperatures.

4.3.3 Determination of torsion angles in RNA

Besides chemical shift information or internuclear distances, also angular information can be obtained by NMR spectroscopy. In addition to six backbone torsion angles and one glycosidic torsion angle, the ribose conformation is crucial for a structure determination. The following chapters will describe the determination of several torsion angles by NMR spectroscopy. The overview in Table 12 shows the different experimental parameters that are relevant for the determination of the respective torsion angle.

Table 12 Overview of the different torsion angles and the NMR parameters which are used to determine them. The angles are shown exemplary in the chemical structure of adenosine (right).

angle	experimental parameter	
α	$\Gamma^{DD,CSA}_{C5'H5'/H5'',P}$ and $\Gamma^{DD,CSA}_{C4'H4',P}$ [183]	
β	^3J(H5'/H5'',P) and ^3J(C4',P) coupling constants[206,207]	
γ	^3J(H4',H5'/H5'')[208,209] and ^3J(C3',H5'/H5'')[210]	
δ	^3J(H3',H4')[208,209], ^3J(C5',H3') and ^3J(C2',H4') coupling constants[211]	
ε	^3J(H3',P), ^3J(C2',P) and ^3J(C4',P) coupling constants[206,207]	
ζ	$\Gamma^{DD,CSA}_{C3'H3',P_{i+1}}$ and $\Gamma^{DD,CSA}_{C4'H4',P_{i+1}}$ [183]	
χ	$\Gamma^{DD,DD}_{C8H8,C1'H1'}$ (purines), $\Gamma^{DD,DD}_{C6H6,C1'H1'}$ (pyrimidines)[181,212]	
P	^3J(H1',H2') and ^3J(H3',H4') [208,209]/ $\Gamma^{DD,DD}_{C1'H1',C2'H2'}$ and $\Gamma^{DD,DD}_{C3'H3',C4'H4'}$[213,214]	

Determination of the ribose conformation

In canonical regions of RNA, the ribose usually adopts the C3'-endo conformation, while the C2'-endo conformation can be present in non-canonical regions of RNA. Generally, three different parameters can be used in order to determine the ribose pucker, namely carbon chemical shifts used for the calculation of canonical coordinates[198,199], ^3J(H,H) coupling constants (Karplus relations, see Appendix 3 Figure 67 A)[208,209] and $\Gamma^{DD,DD}_{CH,CH}$ cross-correlated relaxation rates[213,214].

^3J(H,H) coupling constants were determined using 3D *forward directed* HCC-TOCSY-CCH-E.COSY spectra (pulse program in Appendix 3 chapter 7.3.3, exemplary plane see Appendix 3 Figure 67 B)[208,209]. For the determination of the ribose conformation, the ^3J(H1',H2') and ^3J(H3',H4') coupling constants are relevant and were thus determined for all 14 nts hairpin RNAs. The results in Figure 51 indicate that all stem residues adopt the C3'-endo conformation (^3J(H1',H2') = ~1 Hz, ^3J(H3',H4') = ~9 Hz). All tetraloop nucleotides of the 14 nts RNA with CUUG tetraloop as well as residues U7 and C8 of the 14 nts RNA with UUCG tetraloop can be clearly assigned to the C2'-endo conformation (^3J(H1',H2') = ~8 Hz, ^3J(H3',H4') = ~1 Hz, Figure 51 A and D). These results are in line with the empirically determined canonical coordinates (Figure 49 A and D). For the 14 nts RNAs with GAAG and GCAA tetraloops, the ^3J(H,H) coupling constants do not allow an unambiguous interpretation. Here, both coupling constants adopt values that are neither matching the C2'-endo nor the C3'-endo conformation (Figure 51 B and C). Similar as with the canonical coordinates, this observation is hinting towards an equilibrium between C2'-endo and C3'-endo conformations, which could be verified through the repetition of the experiment at lower temperatures. However, as the resonances of the tetraloop residues are shifting into the intermediate exchange regime at lower temperatures, which broadens signals beyond detectability, this approach was not further followed up on.

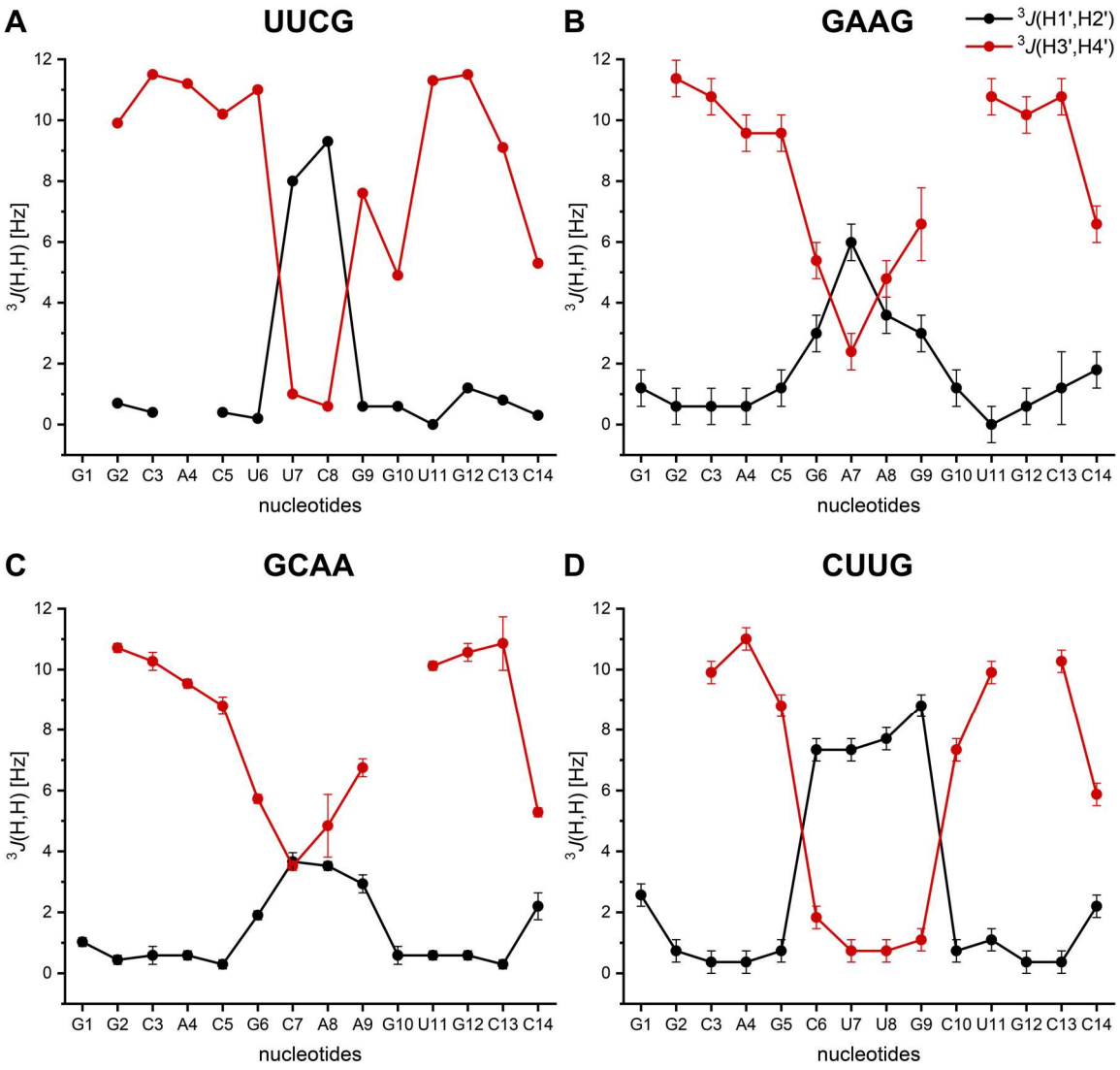


Figure 51 ^3J(H1',H2') and ^3J(H3',H4')coupling constants for the 14 nts RNAs with **A** UUCG, **B** GAAG, **C** GCAA and **D** CUUG tetraloops. The coupling constants are listed in Appendix 3 Table 21. Errors arise from the fitting procedure, while the minimum error corresponds to the respective digital resolution. ^3J(H,H) coupling constants of the 14 nts RNA with UUCG tetraloop were taken from literature[106]. ^3J(H,H) coupling constants of the 14 nts RNAs with GCAA and CUUG tetraloops were provided by György Pintér and Andreas Oxenfarth, respectively.

The ribose pucker can further be determined through $\Gamma^{DD,DD}_{CH,CH}$ cross-correlated relaxation rates[213,214] (Figure 52 D, 2D pulse sequences in Appendix 3 chapter 7.4.2). The overview in Figure 52 shows that the results from ^3J(H,H) coupling constants and canonical coordinates are unambiguously reflected in the $\Gamma^{DD,DD}_{CH,CH}$ cross-correlated relaxation rates for the 14 nts RNA with UUCG tetraloop (Figure 52 A). For the 14 nts RNA with CUUG tetraloop, the same trend is observed in the loop nucleotides (Figure 52 C). However, the set of $\Gamma^{DD,DD}_{CH,CH}$ cross-correlated relaxation rates is incomplete and needs to be further complemented through measurement of the 3D *forward directed quantitative* Γ-HCCH-TOCSY experiment. The same holds true for the 14 nts RNA with GAAG tetraloop, where not even the $\Gamma^{DD,DD}_{C1'H1',C2'H2'}$ cross-correlated relaxation rates can be accurately determined for the tetraloop nucleotides due to a poor S/N ratio (Figure 52 B). As the $\Gamma^{DD,DD}_{CH,CH}$ cross-

correlated relaxation rates are essential for a closing statement about the underlying motion in the tetraloops, the final phase angle of pseudorotation was not yet determined. Depending on the $\Gamma_{CH,CH}^{DD,DD}$ cross-correlated relaxation rates, it will further become apparent, whether the description of the ribose pucker by a single pseudorotation phase is sufficient.

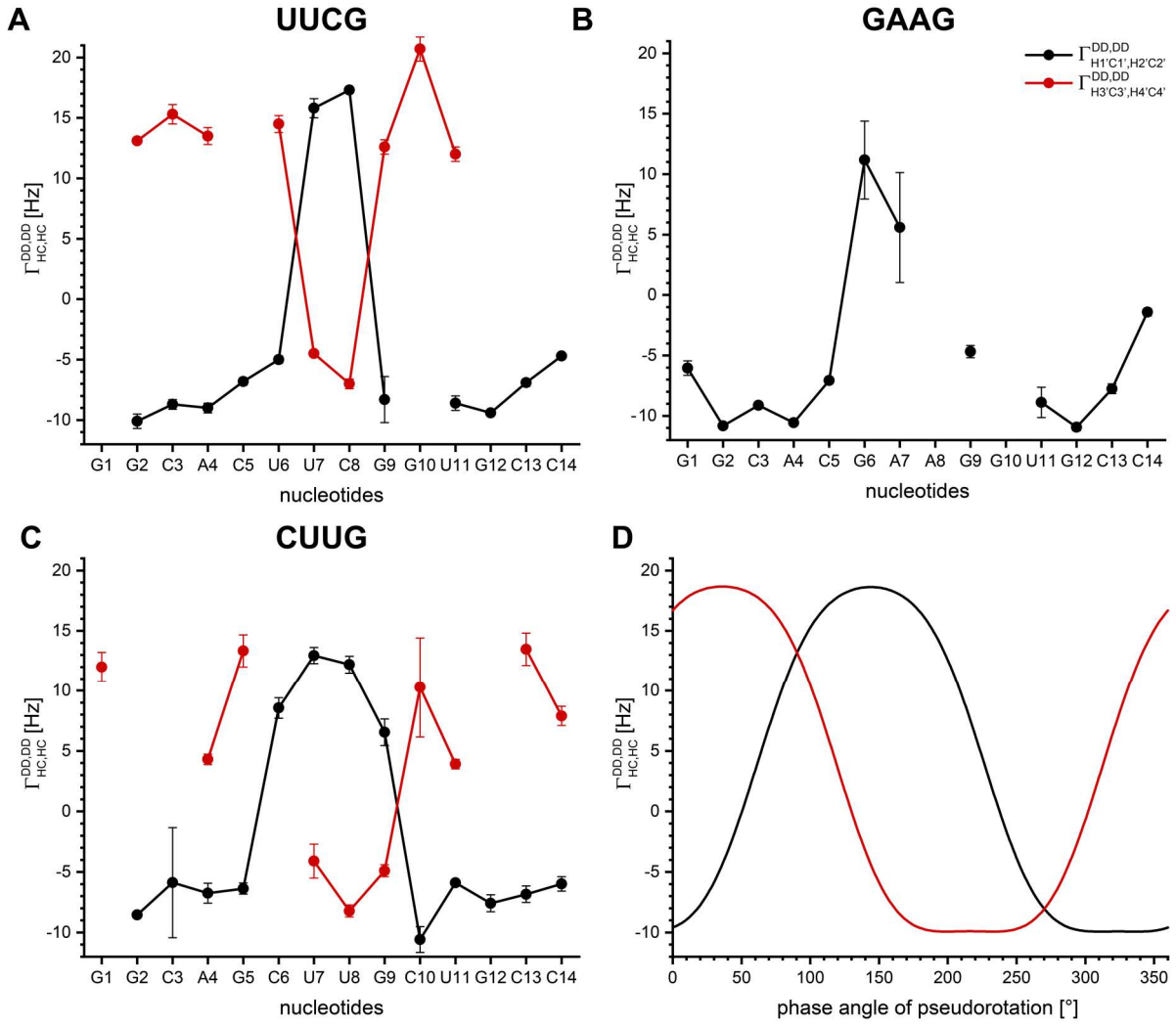


Figure 52 $\Gamma^{DD,DD}_{C1'H1',C2'H2'}$ and $\Gamma^{DD,DD}_{C3'H3',C4'H4'}$ cross correlated relaxation rates for the 14 nts RNA with **A** UUCG tetraloop, **B** GAAG tetraloop and **C** CUUG tetraloop. The rates for the UUCG tetraloop were taken from the literature[106]. The rates for the CUUG tetraloop were provided by Andreas Oxenfarth. The cross-correlated relaxation rates are listed in Appendix 3 Table 22. **D** Plot of the $\Gamma^{DD,DD}_{C1'H1',C2'H2'}$ and $\Gamma^{DD,DD}_{C3'H3',C4'H4'}$ cross correlated relaxation rates in dependence of the phase angle of pseudorotation.

Determination of the glycosidic angle χ

The glycosidic torsion angle χ is determined via cross-correlated relaxation rates, as described earlier (Figure 53 A)[212]. For the calculation of the final angle the order parameter S^2 as well as the rotational correlation time τ_c have to be known. As this is not yet the case for the tetraloops under study, the exact angle cannot be determined.

However, as the glycosidic torsion angle usually adopts either the syn (40-80°) or the anti (180-240°) conformation and the cross-correlated relaxation rates are clearly distinguishable between these different conformations, irrespective of the order parameter S^2 (Figure 53 B), a first estimate on the angle can be made using the rotational correlation time of the UUCG 14 nts RNA (2.3 ns[215]) and the S^2 values of the 14 nts RNA with UUCG tetraloop for the stem residues[182].

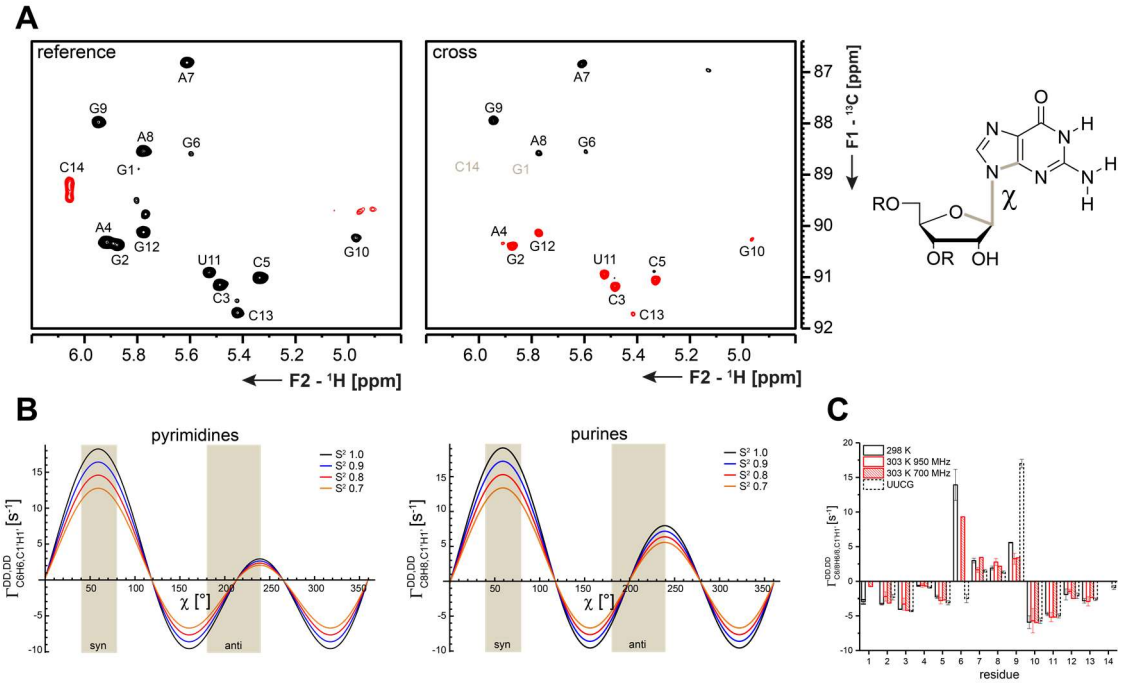


Figure 53 A Reference and cross experiment for the determination of the $\mathbf{\Gamma^{DD,DD}_{C6H6,C1'H1'}}$, and $\mathbf{\Gamma^{DD,DD}_{C8H8,C1'H1'}}$, cross-correlated relaxation rates to determine the glycosidic torsion angle χ (right)[212]. The spectra were recorded at 700 MHz with spectral windows of 20 ppm and 8 ppm in the direct ^{1}H and indirect ^{13}C dimension, respectively. Carrier frequencies were 4.7 ppm, 89 ppm and 161 ppm for ^{1}H, ^{13}C and ^{15}N. Acquisition times were 0.07 s in the direct and 0.02 s in the indirect dimension. The mixing time τ_m was 20 ms. With 128 scans per increment and an inter-scan delay of 1 s, the reference experiment was recorded within 3 h. With 1024 scans per increment and an inter-scan delay of 1 s, the cross experiment was recorded within 1 d 4 h. The pulse programs in Appendix 3 chapter 7.3.1 and 7.3.2 were used. **B** Cross-correlated relaxation rates $\mathbf{\Gamma^{DD,DD}_{C6H6,C1'H1'}}$, and $\mathbf{\Gamma^{DD,DD}_{C8H8,C1'H1'}}$, for pyrimidines and purines in dependence of torsion angle χ for different order parameters S^2. The areas for syn and anti conformation are highlighted. **C** Experimentally determined cross-correlated relaxation rates for the 14 nts RNA with GAAG tetraloop in comparison to the UUCG tetraloop. Rates were determined at 298 K (black) and 303 K (red). Errors result from multiple measurements at different mixing times τ_m (20 ms and 25 ms).

In order to rule out external effects, Γ-HCNCH cross-correlated relaxation rates of the 14 nts RNA with GAAG tetraloop were determined at different temperatures (298 and 303 K) and field strength (700 and 950 MHz). The rates are well in agreement and show that the glycosidic angle χ exhibits a tendency towards the unusual syn conformation for residue G6 (Figure 53 C). The remaining nucleotides adopt the anti conformation. For the GCAA and CUUG 14 nts RNAs the χ angle is in the anti conformation for all nucleotides. The Γ-HCNCH cross-correlated relaxation rates as well as the proposed conformation of the glycosidic angle χ are summarized for the different tetraloops in Table 13.

Table 13 Cross-correlated relaxation rates $\Gamma^{DD,DD}_{C6/8H6/8,C1'H1'}$ for the 14 nts with UUCG, GAAG, GCAA and CUUG tetraloops. The rates for the UUCG tetraloop at 298 K were taken from the literature[212]. The rates for the GAAG and GCAA tetraloops were measured at 298 K at a 600 MHz NMR spectrometer. The rates for the CUUG tetraloop were measured at 308 K at a 600 MHz NMR spectrometer. The $\Gamma^{DD,DD}_{C6/8H6/8,C1'H1'}$ rates of the GCAA and CUUG tetraloop RNAs were determined by György Pintér and Andreas Oxenfarth, respectively. ^1J(C,H) coupling constants for the 14 nts RNA with GAAG tetraloop are given in Appendix 3 Table 23. For the 14 nts RNA with GCAA and CUUG tetraloops coupling constants of ^1J(C,H) = 216 Hz (purines) and ^1J(C,H) = 184 Hz (pyrimidines) were used. Due to poor S/N ratio the cross-correlated relaxation rates could not be determined for residues marked with an asterisk "*". The error was determined by performing the experiment several times at different mixing times. The pulse program given in Appendix 3 chapters 7.3.1 and 7.3.2 were used for all tetraloops.

residue	cUUCGg		cGAAGg		cGCAAg		gCUUGc	
	$\Gamma^{DD,DD}_{C6/8H6/8,C1'H1'}$	χ	$\Gamma^{DD,DD}_{C6/8H6/8,C1'H1'}$	χ	$\Gamma^{DD,DD}_{C6/8H6/8,C1'H1'}$	χ	$\Gamma^{DD,DD}_{C6/8H6/8,C1'H1'}$	χ
G1			-3.0 ± 0.4 Hz	anti	-*		-1.0 Hz	anti
G2	-2.3 ± 0.4 Hz	anti	-3.3 ± 0.2 Hz	anti	-2.5 ± 0.4 Hz	anti	-3.3 Hz	anti
C3	-4.3 ± 0.1 Hz	anti	-4.0 ± 0.1 Hz	anti	-3.5 ± 0.3 Hz	anti	-2.7 Hz	anti
A4	-0.9 ± 0.1 Hz	anti	-0.7 ± 0.2 Hz	anti	-0.7 ± 0.2 Hz	anti	-1.9 Hz	anti
5	-3.1 ± 0.3 Hz	anti	-2.3 ± 0.3 Hz	anti	-*		-1.2 Hz	anti
6	-2.5 ± 0.6 Hz	anti	**13.9 ± 2.3 Hz**	**syn**	-*		1.5 Hz	anti
7	1.5 ± 0.2 Hz	anti	3.0 ± 0.4 Hz	anti	-*		1.0 Hz	anti
8	1.3 ± 0.2 Hz	anti	1.9 ± 0.3 Hz	anti	2.6 ± 0.2 Hz	anti	0.7 Hz	anti
9	**17.0 ± 0.6 Hz**	**syn**	5.6 ± 0.1 Hz	anti	-*		-6.7 Hz	anti
10	-5.7 ± 0.4 Hz	anti	-5.9 ± 1.0 Hz	anti	-*		-4.6 Hz	anti
U11	-5.0 ± 0.3 Hz	anti	-4.7 ± 0.3 Hz	anti	-4.0 ± 0.1 Hz	anti	-3.0 Hz	anti
G12	-2.1 ± 0.4 Hz	anti	-1.9 ± 0.8 Hz	anti	-2.0 ± 0.5 Hz	anti	2.0 Hz	anti
C13	-2.6 ± 0.2 Hz	anti	-2.8 ± 0.4 Hz	anti	-1.7 ± 0.4 Hz	anti	-1.8 Hz	anti
C14	-0.9 ± 0.3 Hz	anti	-*		-*		-*	

Determination of the angles β and ε

The backbone angles β and ε can be determined through ^3J(C,P) and ^3J(H,P) coupling constants (Table 12). ^3J(C,P) coupling constants were determined through quantitative *J*-HCP-experiments[207] and ^3J(H,P) coupling constants by P-FIDS-HSQC experiments (fitting of coupling constants from doublets and singlets). The ^3J(C,P) coupling constants have been determined for all tetraloops under study, while the ^3J(H,P) coupling constants and the final angles β and ε were only determined for the 14 nts RNA with GAAG tetraloop (Table 14).

In canonical RNA, torsion angles β and ε are usually 180° and 209° large, respectively[207]. For the 14 nts RNA with GAAG tetraloop, angle β is in this range throughout the stem (except for G1 and G10). However, as the angles are deviating, the β angle was included with an error of 40° in the structure calculations. In the tetraloop, the β angle of residue A8 is in a non-canonical conformation. The angle ε adopts the canonical conformation for all nucleotides in the stem and in the tetraloop, apart from A7, which is with 166° close to the torsion angle of B-form RNA (150°)[207].

Table 14 Angles β and ε for the 14 nts RNAs with UUCG and GAAG tetraloops. The angles for the UUCG tetraloop were taken from the literature[106]. The angles for the GAAG tetraloop were determined from ^3J(C,P) ad ^3J(H,P) coupling constants (see Appendix 3 Table 24 and Table 25). Missing values could not be determined due to resonance overlap.

residue	cUUCGg		cGAAGg	
	β	ε	β	ε
G1	180°	228°	230°	212°
G2	180°		178°	217°
C3	177°	194°	185°	209°
A4		189°		218°
C5	176°	220°	193°	212°
6	176°	207°	193°	210°
7	180°	300°	166°	166°
8	184°	215°	148°	214°
9		212°	198°	214°
G10		194°	129°	209°
U11	184°	212°	202°	212°
G12		217°	178°	210°
C13	180°	217°	193°	211°
C14	175°		190°	226°

Determination of the angle γ

The angle γ can be qualitatively determined from ^3J(H4',H5'), ^3J(H4',H5''), ^2J(C4',H5') and ^2J(C4',H5'') coupling constants[210]. Using these coupling constants (Table 15) and chemical shift information (Figure 54 A), the H5' protons were stereospecifically assigned (Figure 54 B) for the 14 nts RNA with GAAG tetraloop.

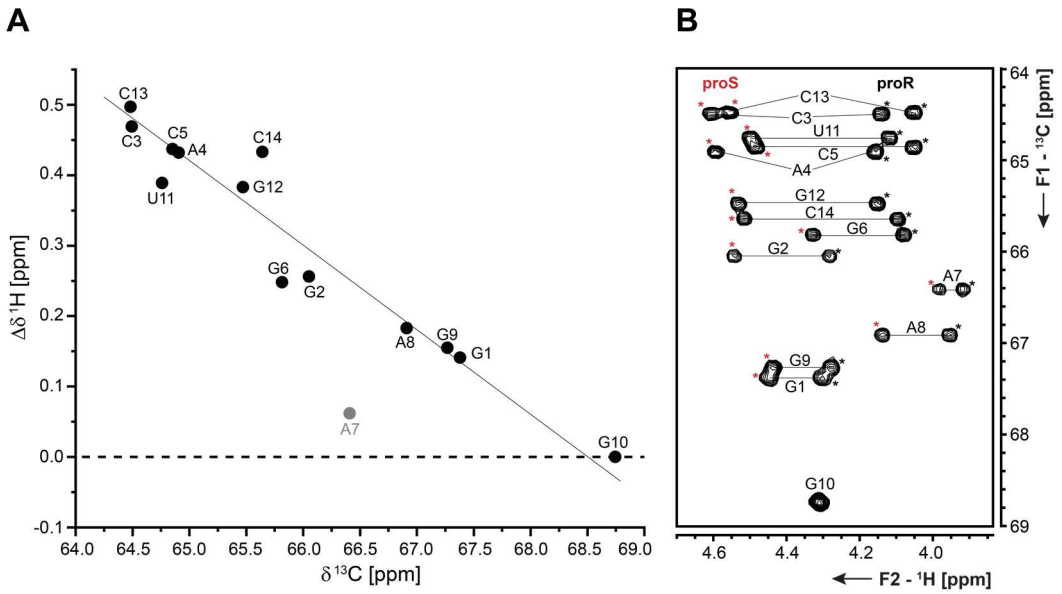


Figure 54 A Plot of the difference in H5' (proS-proR) chemical shifts in dependence of the C5' chemical shifts for the 14 nts RNA with GAAG tetraloop. **B** ^{1}H,^{13}C-HSQC spectrum of the C5'-region of the 14 nts RNA with GAAG tetraloop including the stereospecific assignment. H5'(proS) resonances are marked with a red asterisk, while a black asterisk marks the H5'(proR) resonances. The assignment was conducted based on the plot in A and based on the *J*-coupling constants in Table 15.

Due to a poor S/N ratio and asymmetry in peaks, the ^3J(H4',H5'/H5'') coupling constants were only qualitatively categorized into "small" and "large". This crude discrimination is sufficient for the stereospecific assignment and the determination of angle γ [210]. In the 14 nts RNA with GAAG tetraloop, angle γ adopts the gauche(+) rotameric state (60°) in all nucleotides. The trans rotamer would feature a ~10 Hz ^3J(H4',H5'(proR)) coupling constant, which would be still detectable, despite the general inaccuracies in measurement of the ^3J(H4',H5') coupling constants observed here. Only for residue G10 the angle γ cannot be determined, as the H5'(proR) and H5'(proS) chemical shifts are degenerate. Here, it would be possible to repeat both experiments at a different temperature (e.g. 308 K) as the chemical shift perturbations in response to temperature differ for the two protons, thus removing the degeneracy.

Table 15 ^2J(C4',H5') and ^3J(H4',H5') coupling constants and the qualitatively derived angle γ for the 14 nts RNAs with UUCG and GAAG tetraloops at 298 K. The coupling constants for the UUCG tetraloop were taken from the literature[106]. Here, absolute values are given, as the Karplus relation does not allow negative values for ^3J(H,H) coupling constants. ^2J(C4',H5') coupling constants of the 14 nts RNA with GAAG tetraloop were determined from a selective ^{1}H,^{13}C-HSQC spectrum[210] and the ^3J(H4',H5') coupling constants were determined via a 3D *forward directed* HCC-TOCSY-CCH-E.COSY experiment[208,209]. The ^3J(H4',H5') coupling constants are marked with "s" for small (<5 Hz) and "l" for large (>5 Hz) values. For residue G10 coupling constants could not be determined due to chemical shift degeneracy of H5'(proS) and H5'(proR).

residue	cUUCGg					cGAAGg				
	^2J(C4',H5') [Hz]		^3J(H4',H5') [Hz]		γ	^2J(C4',H5') [Hz]		^3J(H4',H5') [Hz]		γ
	proS	proR	proS	proR		proS	proR	proS	proR	
G1	0.3	2.0	5.2	0.6	60°	-1.9	-0.6	s	s	60°
G2	-0.4	1.6	2.7	1.7	60°	-2.6	-0.6	s	s	60°
C3						-4.0	0.8	s		60°
A4						-3.2	1.1	s	s	60°
C5	0.7	2.8	1.5	1.0	60°	-2.6	0.9	s	s	60°
6	-1.4	2.3	0.6	0.6	60°	-3.3	1.4	s	s	60°
7	-1.0	3.8	1.2	0.3	60°	-2.6	2.5	s	s	60°
8	-0.9	3.4	1.1	0.3	60°	-2.0	-0.1	s	s	60°
9	2.7	-0.8	4.6	2.8	180°	-1.5	-0.7	s	s	60°
G10										
U11	-2.1	3.3	4.5	0.9	60°	-3.4	1.0	s	s	60°
G12							0.5	s	s	60°
C13	-1.3	3.2	2.8	1.7	60°	-4.0	0.5	s	s	60°
C14	-0.6	3.0	1.3	0.8	60°			s	s	60°

4.3.4 On the way to a three-dimensional structure – NOESY spectra

The most commonly used restraints in structure calculations are ^{1}H-^{1}H distances that are derived from NOESY spectra. Here, a set of 2D ^{1}H,^{1}H-NOESY spectra was recorded with different mixing times (50 ms, 100 ms and 150 ms) for both, the 14 nts RNA with GAAG tetraloop solved in 100% D_2O and solved in 95% H_2O and 5% D_2O. From the NMR samples dissolved in mostly H_2O, NOE contacts to protons that participate in solvent exchange (NH, NH_2 and 2'OH groups) are obtained, while the NOESY spectra of the RNA in D_2O yielded NOE contacts between non-exchanging protons. When evaluating the

^{1}H,^{1}H-NOESY spectra, it becomes evident that the amino groups of the tetraloop nucleotides are particularly important, as they give rise to a fair amount of restraints in a region where cross-peaks are otherwise scarce. Especially the amino protons of G6 give rise to overall 12 valuable inter-residual NOE-correlations (Figure 55 C). These correlations were further confirmed at a different temperature in order to eliminate cross peaks which might arise if the chemical shift of this amino group and e.g. a 2'-OH group was degenerate (data not shown).

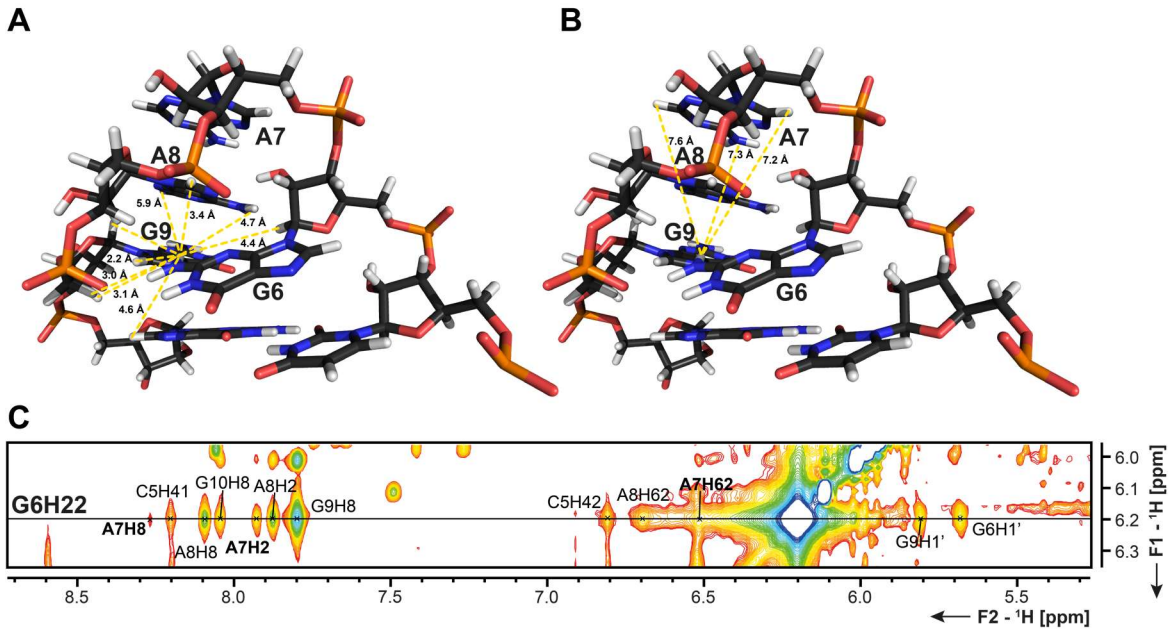


Figure 55 A GAAG tetraloop from the NMR structure in the literature[109] with G6-NH_2 NOE contacts (yellow dashed line), which are observed in the 2D ^{1}H,^{1}H NOESY spectrum in **C** and which are in agreement with the proposed tetraloop structure. **B** The structure from **A** with G6-NH_2 NOE contacts (yellow dashed line), which are observed in the 2D ^{1}H,^{1}H-NOESY spectrum in **C** and which are not in agreement with the proposed tetraloop structure. **C** Section of a ^{1}H,^{1}H-NOESY spectrum of the 14 nts RNA with GAAG tetraloop showing the cross-peaks for the NH_2-group of G6. Bold labels mark contacts, which are shown in **B**. The spectrum has been recorded using sample#1 at 298 K with a mixing time of 150 ms. The spectrum was recorded with spectral windows of 18.6 ppm and 12.5 ppm in the direct and indirect ^{1}H-dimensions, respectively. The ^{1}H carrier frequency was 4.7 ppm. Acquisition times were 0.09 s in the direct and 0.05 s in the indirect dimension. With an inter-scan delay of 1.1 s and 80 scans per increment, the experiment was recorded within 22 h.

When mapping the 12 G6-NH_2 NOE contacts on the GAAG structure, which was previously solved based on NMR data[109], most of the cross-peaks observed here are in well agreement with this structure (Figure 55 A). Surprisingly, a closer inspection shows that three NOE distance restraints between residues G6 and A7 are not in line with the reference structure (Figure 55 B). These peaks are presumably not a result of spin-diffusion, as most of them are also present in a ^{1}H,^{1}H-NOESY spectrum with 50 ms mixing time (data not shown). As the NMR-structure of the GAAG tetraloop, which is described in the literature, adopts the consensus GNRA tetraloop structure, the observations made here indicate that either a conformational equilibrium between the GNRA consensus structure and a second structure is formed or that only one, but a completely different tetraloop structure is adopted that gives rise to all observed NOE contacts. Furthermore, it was proposed earlier that G6 and G9 form a single-bond non-canonical G-G mismatch

with a hydrogen bond that involves G6NH$_2$ and G9N7[109]. This assumption was based on a strong cross-peak between G6NH$_2$ and G9H8[109]. While this strong correlation can be confirmed within this study (Figure 55 C), no further evidence for the existence of the hydrogen bond could be detected in HNN-COSY experiments. On the contrary, the amino protons do only give rise to a single resonance, indicating that the C-N bond rotation is in the fast exchange regime. This bond rotation is usually in the intermediate exchange regime at room temperature if one of the amino protons is involved in a stable hydrogen bond (see Chapter II). However, the single G6NH$_2$ resonance is also still observable at 274 K. Additionally, the familiar downfield shift of a hydrogen bonded amino proton is not observed; in fact the G6NH$_2$ resonance represents the most upfield shifted amino proton resonance of this RNA (Figure 61 A). Based on these observations and the discussions in Chapter II it can be concluded that a stable hydrogen bond between the amino group of G6 and G9 nitrogen N7 is most likely not formed.

The peak lists, which were used in the structure calculations, are listed in Appendix 3 chapter 7.5.3.

4.3.5 The structure of the 14 nts RNA with GAAG tetraloop

Even though the main objective of this study is the generation of primary experimental data for the reparametrization of potential energy functions and despite the very likely presence of a conformational equilibrium, the RNA structure of the 14 nts RNA with GAAG tetraloop was calculated in order to visually compare the experimental and the MD structures. Therefore, structure calculations were conducted for the 14 nts RNA with GAAG tetraloop with the distance constraints that are derived from the NOESY data and the torsion angle restraints determined here (lists see Appendix 3 chapter 7.5). In comparison with the experimental parameters, which were used in the structure determination of the 14 nts RNA with UUCG tetraloop (29 restraints per residue), the sheer amount of restraints that is used in the structure determination of the GAAG tetraloop in this work, is with 31 restraints per residue in a similar range (Table 16). Hydrogen bond and base pair planarity restraints (lists see Appendix 3 chapter 7.5) were derived from HNN-COSY spectra[41]. As the process of gathering experimental data is not yet finished, the structure presented here is still preliminary.

Table 16 Overview of the experimental parameters used in the structure determination of the UUCG[106] tetraloop in comparison to the experimental parameters used for the preliminary structure determination of the GAAG tetraloop. Distance restraints and dihedral restraints were combined in the calculation of the number of restraints per nucleotide. [a] The 32 RDCs were also included in this calculation. The overall RMSD is given.

tetraloop	length	distance restraints	dihedral restraints	restraints per residue	RMSD
UUCG	14 nts	267	96	~29[a]	0.4 Å
GAAG	14 nts	387	49	~31	0.9 Å

As expected from mfold predictions[189] and the comparison with the 14 nts RNA with UUCG tetraloop, the 14 nts RNA structure consists of a stem, which is comprised of five Watson-Crick base pairs and which is topped by a GAAG tetraloop (Figure 56 A). Here, the GAAG tetraloop is not adopting the proposed GNRA type structure, where A7, A8 and G9 would stack on top of each other and a G6-G9 mismatch would be formed. Instead, a stacking interaction between G6 and A7 is observed, while A8 and G9 are tilted outwards (Figure 56 B). The overall RMSD of the structure bundle of the 20 structures with the lowest energy is 0.9 Å. However, it should be underlined that temperature series, ^{3}J(H,H) coupling constants and NOE cross peaks all hint towards the existence of an alternative conformation of the GAAG tetraloop. As a result, the three-dimensional structure presented here might not reflect the reality. Therefore, the structure determination should rather be repeated using a two-state model.

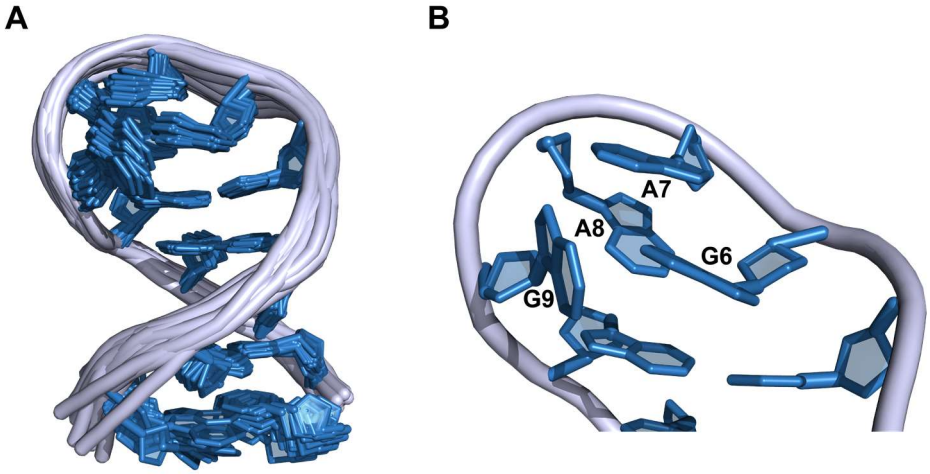


Figure 56 A Superposition of the 20 structures with the lowest energy of the 14 nts RNA with GAAG tetraloop. Lists of NOE distances, dihedral restraints, hydrogen bonding restraints, base planarity restraints and the distance violations are given in Appendix 3 chapter 7.5. **B** Close-up of the GAAG tetraloop structure.

4.4 Conclusions and Outlook

Within this project, three 14 nts long RNAs featuring different tetraloop sequences, namely GAAG, GCAA and CUUG are structurally characterized. For the construct with GAAG tetraloop, which was subject of this PhD thesis, a complete resonance assignment was achieved. Various coupling constants and cross-correlated relaxation rates have been measured, which allowed the characterization of the dihedral angles β, γ, ε and χ. Structure calculations with these torsion angles as restraints as well as with ^{1}H-^{1}H distance restraints derived from 2D ^{1}H,^{1}H-NOESY spectra (in 100% D_2O and 95% H_2O/5%D_2O) were used for a structure calculation and resulted in a preliminary structure bundle with an overall RMSD of 0.9 Å. Furthermore, the ribose pucker was analyzed based on carbon chemical shifts, ^{3}J(H,H) coupling constants and $\Gamma_{CH,CH}^{DD,DD}$ cross-correlated relaxation rates. At this point, the results are indicative of the existence of a temperature dependent equilibrium between C2'-endo and C3'-endo conformation for the nucleotides of the tetraloop. However, this has to be verified by measuring a complete set of $\Gamma_{CH,CH}^{DD,DD}$ cross-correlated relaxation rates, as this structural parameter promises the most accurate result. The reason for this is that, unlike the ^{3}J(H,H) coupling constants, cross-correlated relaxation rates do not depend on the Karplus parametrization. In addition to the phase angle of pseudorotation, the dihedral angles α and ζ still need to be characterized through the measurement of $\Gamma_{CH,P}^{DD,CSA}$ cross-correlated relaxation rates[183]. These torsion angles should then lead to a refined three-dimensional structure, which can be further complemented by residual dipolar couplings (RDCs). Additionally, structural information on the NOE-contacts involving the 2'-OH groups is not yet reported on. This is due to the fact that the signals of the 2'-OH groups are not observed at any temperature, even at reduced buffer concentrations (25 mM potassiumphosphate). A further reduction to 10 mM potassiumphosphate buffer might allow the detection of the 2'-OH groups, as those are observed for the 14 nts RNA with CUUG tetraloop, which is generally studied under these conditions.

For the structure determination itself it should be questioned whether the systems under study within this project (especially the GAAG and GCAA tetraloops) are adequately described by a single conformation. This might not be the case due to dynamics, which are observed in the tetraloops. These dynamics are indicative of at least one or even more additional mutually exchanging states and should be further characterized by e.g. relaxation dispersion experiments ($R_{1\rho}$) and the measurement of eNOE data. Also, the order parameter S^2 is to be determined in the course of the dynamic characterization, which can be then used to refine the angle χ. Furthermore, the final structure should be refined using XPLOR-NIH[186].

The primary goal of this project is, however, not the final structural characterization, as this structure might be biased by the force field itself, which is used to calculate the three-dimensional structure. Instead, the accumulation of raw data for different RNAs with

different structural features is the core of this work. While the tetraloops that are studied here already exhibit structural diversity, the future perspective of this project would move away from tetraloops towards different structural motifs like pentaloops or hexaloops. This is a more challenging task due to the additional loop nucleotide(s) and could be performed using the identical stem nucleotide sequence as apparent in the 14 nts RNAs within the present study. The loop sequences could then be altered, while considering the respective native closing base pairs.

References

[1] W. Winkler, A. Nahvi, R. R. Breaker, *Nature* **2002**, *419*, 952–956.

[2] S. Altuvia, D. Kornitzer, D. Teff, A. B. Oppenheim, *J. Mol. Biol.* **1989**, *210*, 265–280.

[3] B. K. Ray, D. Apirion, *Mol. Genet. Genomics* **1979**, *174*, 25–32.

[4] A. J. Hamilton, D. C. Baulcombe, *Science (80-.).* **1999**, *286*, 950–952.

[5] R. C. Lee, R. L. Feinbaum, V. Ambros, *Cell* **1993**, *75*, 843–854.

[6] P. Kapranov, J. Cheng, S. Dike, D. A. Nix, R. Duttagupta, A. T. Willingham, P. F. Stadler, J. Hertel, J. Hackermüller, I. L. Hofacker, et al., *Science (80-.).* **2007**, *316*, 1484–1488.

[7] J. M. Nigro, K. R. Cho, E. R. Fearon, S. E. Kern, J. M. Ruppert, J. D. Oliner, K. W. Kinzler, B. Vogelstein, *Cell* **1991**, *64*, 607–613.

[8] M. Ghildiyal, P. D. Zamore, *Nat. Rev.* **2009**, *10*, 94–108.

[9] B. Fürtig, C. Richter, J. Wöhnert, H. Schwalbe, *Chembiochem* **2003**, *4*, 936–62.

[10] M. S. Marek, A. Johnson-Buck, N. G. Walter, *Phys. Chem. Chem. Phys* **2011**, *13*, 11524–11537.

[11] C. Höbartner, M.-O. Ebert, B. Jaun, R. Micura, *Angew. Chemie - Int. Ed.* **2002**, *41*, 605–609.

[12] M. A. Juen, C. H. Wunderlich, F. Nußbaumer, M. Tollinger, G. Kontaxis, R. Konrat, D. F. Hansen, C. Kreutz, *Angew. Chemie - Int. Ed.* **2016**, *55*, 12008–12012.

[13] B. Fürtig, C. Richter, P. Schell, P. Wenter, S. Pitsch, H. Schwalbe, *RNA Biol.* **2008**, *5*, 41–48.

[14] A. Reining, S. Nozinovic, K. Schlepckow, F. Buhr, B. Fürtig, H. Schwalbe, *Nature* **2013**, *499*, 355–9.

[15] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.

[16] T. Carlomagno, *J. Magn. Reson.* **2014**, *241*, 126–136.

[17] J. Hennig, C. Militti, G. M. Popowicz, I. Wang, M. Sonntag, A. Geerlof, F. Gabel, F. Gebauer, M. Sattler, *Nature* **2014**, *515*, 287–290.

[18] T. Afroz, Z. Cienikova, A. Cléry, F. H. T. Allain, *Methods Enzymol.* **2015**, *558*, 235–278.

[19] A. Lapinaite, B. Simon, L. Skjaerven, M. Rakwalska-bange, F. Gabel, T. Carlomagno, *Nature* **2013**, *502*, 519–523.

[20] S. C. Keane, V. Van, H. M. Frank, C. A. Sciandra, S. McCowin, J. Santos, X. Heng, M. F. Summers, *Proc. Natl. Acad. Sci.* **2016**, *113*, 13033–13038.

[21] K. D. Mouzakis, E. A. Dethoff, M. Tonelli, H. Al-Hashimi, S. E. Butcher, *Biophys. J.* **2015**, *108*, 644–654.

[22] A. G. Tzakos, L. E. Easton, P. J. Lukavsky, *Nat. Protoc.* **2007**, *2*, 2139–2147.

[23] S. C. Keane, X. Heng, K. Lu, V. Ramakrishnan, G. Carter, S. Barton, A. Hosic, A. Florwick, J. Santos, N. C. Bolden, et al., *Science (80-.).* **2015**, *348*, 917–922.

[24] M. H. Bailor, X. Sun, H. M. Al-Hashimi, *Science (80-.).* **2010**, *327*, 202–206.

[25] I. J. Kimsey, K. Petzold, B. Sathyamoorthy, Z. W. Stein, H. M. Al-hashimi, *Nature* **2015**, *519*, 315–320.

[26] C. Helmling, A. Wacker, M. T. Wolfinger, I. L. Hofacker, M. Hengesbach, B. Fürtig, H. Schwalbe, *J. Am. Chem. Soc.* **2017**, *139*, 2647–2656.

[27] H. Steinert, F. Sochor, A. Wacker, J. Buck, C. Helmling, F. Hiller, S. Keyhani, J. Noeske, S. Grimm, M. M. Rudolph, et al., *Elife* **2017**, 1–18.

[28] C. Helmling, D. Klötzner, F. Sochor, R. A. Mooney, A. Wacker, R. Landick, B. Fürtig, A. Heckel, H. Schwalbe, *Nat. Commun.* **2018**, *9*, 944.

[29] P. Styles, N. F. Soffe, C. A. Scott, *J. Magn. Reson.* **1989**, *84*, 376–378.

[30] H. Kovacs, D. Moskau, M. Spraul, *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 131–155.

[31] J. C. J. Barna, E. D. Laue, M. R. Mayger, J. Skilling, S. J. P. Worrall, **1987**, *77*, 69–77.

[32] K. Kazimierczuk, V. Y. Orekhov, *Angew. Chem., Int. Ed.* **2011**, *50*, 5556–5559.

[33] V. Y. Orekhov, V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 271–292.

[34] S. Robson, H. Arthanari, S. G. Hyberts, G. Wagner, *Methods Enzymol.* **2019**, *614*, 263–291.

[35] J. Stanek, P. Podbevsek, W. Kozminski, J. Plavec, M. Cevec, *J Biomol NMR* **2013**, *57*, 1–9.

[36] T. Kühn, H. Schwalbe, *J. Am. Chem. Soc.* **2000**, *122*, 6169–6174.

[37] P. Wenter, B. Fürtig, A. Hainard, H. Schwalbe, S. Pitsch, *Angew. Chem., Int. Ed.* **2005**, *44*, 2600–2603.

[38] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci.* **1997**, *94*, 12366–12371.

[39] B. Simon, K. Zanier, M. Sattler, *J. Biomol. NMR* **2001**, *20*, 173–176.

[40] R. Riek, K. Pervushin, C. Fernández, M. Kainosho, K. Wüthrich, **2001**, 4507–4513.

[41] A. J. Dingley, S. Grzesiek, *J. Am. Chem. Soc.* **1998**, *7863*, 714–718.

[42] J. Wöhnert, A. J. Dingley, M. Stoldt, M. Görlach, S. Grzesiek, L. R. Brown, *Nucleic Acids Res.* **1999**, *27*, 3104–3110.

[43] M. C. Erat, H. Kovacs, R. K. O. Sigel, *J. Inorg. Biochem.* **2010**, *104*, 611–613.

[44] J. N. Kim, A. Roth, R. R. Breaker, *Proc. Natl. Acad. Sci.* **2007**, *104*, 16092–16097.

[45] R. Schnieders, S. Keyhani, H. Schwalbe, B. Fürtig, *Chem. - A Eur. J.* **2020**, *26*, 102–113.

[46] L. J. Alvarado, R. M. Leblanc, A. P. Longhini, S. C. Keane, N. Jain, Z. F. Yildiz, B. S. Tolbert, V. M. D. Souza, M. F. Summers, C. Kreutz, et al., *Chembiochem* **2014**, *15*, 1573–1577.

[47] L. J. Alvarado, A. P. Longhini, R. M. Leblanc, B. Chen, C. Kreutz, T. K. Dayie, *Methods Enzymol.* **2014**, *549*, 133–162.

[48] S. Quant, R. W. Wechselberger, M. A. Wolter, K.-H. Wörner, P. Schell, J. W. Engels, C. Griesinger, H. Schwalbe, **1994**, *35*, 6649–6652.

[49] L. Jud, R. Micura, *Chem. Eur. J.* **2017**, *23*, 3406–3413.

[50] J. M. Holstein, L. Anhäuser, A. Rentmeister, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10899–10903.

[51] F. Muttach, F. Mäsing, A. Studer, A. Rentmeister, *Chem. Eur. J.* **2017**, *23*, 5988–5993.

[52] D. Schulz, J. M. Holstein, A. Rentmeister, *Angew. Chemie - Int. Ed.* **2013**, *52*, 7874–7878.

[53] V. Rosemeyer, A. Laubrock, R. Seibl, *Anal. Biochem.* **1995**, *224*, 446–449.

[54] G. Martin, W. Keller, *RNA* **1998**, *4*, 226–230.

[55] M.-L. Winz, A. Samanta, D. Benzinger, A. Jäschke, *Nucleic Acids Res.* **2012**, *40*, e78.

[56] T. E. England, O. C. Uhlenbeck, *Biochemistry* **1978**, *17*, 2069–2076.

[57] J. R. Barrio, M. del C. G. Barrio, N. J. Leonard, T. E. England, O. C. Uhlenbeck, *Biochemistry* **1978**, *17*, 2077–2081.

[58] O. Duss, C. Maris, C. Von Schroetter, F. H. T. Allain, *Nucleic Acids Res.* **2010**, *38*, DOI 10.1093/nar/gkq756.

[59] S. Keyhani, T. Goldau, A. Blümler, A. Heckel, H. Schwalbe, *Angew. Chemie - Int. Ed.* **2018**, *57*, 12017–12021.

[60] B. Reif, V. Wittmann, H. Schwalbe, C. Griesinger, K. Wörner, K. Jahn-Hofmann, J. W. Engels, W. Bermel, *Helv. Chim. Acta* **1997**, *80*, 1952–1971.

[61] D. Graber, H. Moroder, R. Micura, *J. Am. Chem. Soc.* **2008**, *130*, 17230–17231.

[62] B. Puffer, C. Kreutz, U. Rieder, M.-O. Ebert, R. Konrat, R. Micura, *Nucleic Acids Res.* **2009**, *37*, 7728–7740.

[63] F. Sochor, R. Silvers, D. Müller, C. Richter, B. Fürtig, H. Schwalbe, *J. Biomol. NMR* **2016**, *64*, 63–74.

[64] A. Boeszoermenyi, S. Chhabra, A. Dubey, D. L. Radeva, N. T. Burdzhiev, C. D. Chanev, O. I. Petrov, V. M. Gelev, M. Zhang, C. Anklin, et al., *Nat. Methods* **2019**, *16*, 333–340.

[65] K. Lu, X. Heng, L. Garyu, S. Monti, E. L. Garcia, S. Kharytonchyk, B. Dorjsuren, G. Kulandaivel, S. Jones, A. Hiremath, et al., *Science (80-.).* **2011**, *334*, 242–245.

[66] H. Schwalbe, J. Buck, B. Fürtig, J. Noeske, J. Wöhnert, *Angew. Chemie - Int. Ed.* **2007**, *46*, 1212–1219.

[67] A. Lingel, B. Simon, E. Izaurralde, M. Sattler, *Nat. Struct. Mol. Biol.* **2004**, *11*, 576–

577.

[68] K. Lu, Y. Miyazaki, M. F. Summers, *J. Biomol. NMR* **2010**, *46*, 113–125.

[69] A. Hammarstroem, G. Otting, *J. Am. Chem. Soc.* **1994**, *116*, 8847–8848.

[70] M. Ottiger, F. Delaglio, A. Bax, *J. Magn. Reson.* **1998**, 373–378.

[71] A. Meissner, J. Ø. Duus, O. W. Sørensen, *J. Magn. Reson.* **1997**, *128*, 92–97.

[72] E. L. Ulrich, H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, et al., *Nucleic Acids Res.* **2008**, *36*, 402–408.

[73] C. Richter, H. Kovacs, J. Buck, A. Wacker, B. Fürtig, W. Bermel, H. Schwalbe, *J. Biomol. NMR* **2010**, *47*, 259–269.

[74] V. Sklenár, T. Dieckmann, S. E. Butcher, J. Feigon, *J. Biomol. NMR* **1996**, *7*, 83–7.

[75] J. Wöhnert, R. Ramachandran, M. Görlach, L. R. Brown, *J. Magn. Reson.* **1999**, *139*, 430–433.

[76] R. Fiala, V. Sklenár, *J. Biomol. NMR* **2007**, *39*, 153–163.

[77] B. Fürtig, R. Schnieders, C. Richter, H. Zetzsche, S. Keyhani, C. Helmling, H. Kovacs, H. Schwalbe, *J. Biomol. NMR* **2016**, *64*, 207–221.

[78] R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe, B. Fürtig, *Angew. Chemie - Int. Ed.* **2019**, *58*, 9140–9144.

[79] D. R. Kearns, D. J. Patel, R. G. Shulman, *Nature* **1971**, *229*, 338–339.

[80] A. Wacker, J. Buck, D. Mathieu, C. Richter, J. Wo, H. Schwalbe, *Nucleic Acids Res.* **2011**, *39*, 6802–6812.

[81] N. B. Leontis, J. Stombaugh, E. Westhof, *Nucleic Acids Res.* **2002**, *30*, 3497–3531.

[82] A. Kettani, M. Gue, J. Leroy, *J. Am. Chem. Soc.* **1997**, *119*, 1108–1115.

[83] L. Mueller, P. Legault, A. Pardi, *J. Am. Chem. Soc.* **1995**, *117*, 11043–11048.

[84] A. C. Wolter, A. Pianu, J. Kremser, E. Strebitzer, R. Schnieders, B. Fürtig, C. Kreutz, E. Duchardt-Ferner, J. Wöhnert, *Biomol. NMR Assign.* **2019**, DOI 10.1007/s12104-019-09892-z.

[85] R. Michalczyk, I. M. Russu, *Biophys. J.* **1999**, *76*, 2679–2686.

[86] M. Adrian, F. R. Winnerdy, B. Heddi, A. T. Phan, *Biophys. J.* **2017**, *113*, 775–784.

[87] L. D. Williams, N. G. Williams, B. R. Shaw, *J. Am. Chem. Soc.* **1990**, *112*, 829–833.

[88] H. A. Heus, A. Pardi, *J. Am. Chem. Soc.* **1991**, *113*, 4360–4361.

[89] A. Majumdar, A. Kettani, E. Skripkin, *J. Biomol. NMR* **1999**, *14*, 67–70.

[90] A. Majumdar, Y. Gosser, D. J. Patel, *J. Biomol. NMR* **2001**, *21*, 289–306.

[91] D. J. Patel, A. Kettani, A. Majumdar, E. Skripkin, *J Biomol NMR* **2001**, *19*, 103–113.

[92] F. A. A. Mulder, C. A. E. M. Spronk, M. Slijper, R. Kaptein, R. Boelens, *J. Biomol. NMR* **1996**, *8*, 223–228.

[93] A. J. Shaka, P. B. Barker, R. Freeman, *J. Magn. Reson.* **1985**, *64*, 547–552.

[94] B. Fürtig, R. Schnieders, C. Richter, H. Zetzsche, S. Keyhani, C. Helmling, H. Kovacs, H. Schwalbe, *J. Biomol. NMR* **2016**, DOI 10.1007/s10858-016-0021-5.

[95] C. L. Borders, J. A. Broadwater, P. A. Bekeny, J. E. Salmon, A. S. Lee, A. M. Eldridge, V. B. Pett, *Protein Sci.* **1994**, *3*, 541–548.

[96] J. H. Tomlinson, S. Ullah, P. E. Hansen, M. P. Williamson, *J. Am. Chem. Soc.* **2009**, *131*, 4674–4684.

[97] J. Iwahara, Y. S. Jung, G. M. Clore, *J. Am. Chem. Soc.* **2007**, *129*, 2971–2980.

[98] A. Esadze, D.-W. Li, T. Wang, R. Brüschweiler, J. Iwahara, *J. Am. Chem. Soc.* **2011**, *133*, 909–919.

[99] K. Gerecht, A. M. Figueiredo, D. F. Hansen, *Chem. Commun.* **2017**, *53*, 10062–10065.

[100] G. Goldstein, M. Scheid, U. Hammerling, E. A. Boyse, D. H. Schlesinger, H. D. Niall, *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 11–15.

[101] S. Vijay-Kumar, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **1987**, *194*, 531–544.

[102] N. D. Werbeck, J. Kirkpatrick, D. F. Hansen, *Angew. Chemie - Int. Ed.* **2013**, *52*, 3145–3147.

[103] H. W. Mackenzie, D. F. Hansen, *J. Biomol. NMR* **2017**, *69*, 123–132.

[104] S. A. Robson, R. Peterson, L. S. Bouchard, V. A. Villareal, R. T. Clubb, *J. Am. Chem. Soc.* **2010**, *132*, 9522–9523.

[105] B. Fürtig, C. Richter, W. Bermel, H. Schwalbe, *J. Biomol. NMR* **2004**, *28*, 69–79.

[106] S. Nozinovic, B. Fürtig, H. R. A. Jonker, C. Richter, H. Schwalbe, *Nucleic Acids Res.* **2010**, *38*, 683–94.

[107] A. C. Wolter, E. Duchardt-Ferner, A. H. Nasiri, K. Hantke, C. H. Wunderlich, C. Kreutz, J. Wöhnert, *Biomol. NMR Assign.* **2016**, *10*, 101–105.

[108] A. C. Wolter, A. K. Weickhmann, A. H. Nasiri, K. Hantke, O. Ohlenschläger, C. H. Wunderlich, C. Kreutz, E. Duchardt-Ferner, J. Wöhnert, *Angew. Chemie - Int. Ed.* **2017**, *56*, 401–404.

[109] K. Okada, M. Takahashi, T. Sakamoto, G. Kawai, K. Nakamura, A. Kanai, *Nucleosides, Nucleotides, and Nucleic Acids* **2006**, *25*, 383–395.

[110] A. K. Weickhmann, H. Keller, E. Duchardt-Ferner, E. Strebitzer, M. A. Juen, J. Kremser, J. P. Wurm, C. Kreutz, J. Wöhnert, *Biomol. NMR Assign.* **2018**, *12*, 329–334.

[111] A. K. Weickhmann, H. Keller, J. P. Wurm, E. Strebitzer, A. M. Juen, J. Kremser, Z. Weinberg, C. Kreutz, E. Duchardt-Ferner, J. Wöhnert, *Nucleic Acids Res.* **2018**, 1–12.

[112] J. Leppert, C. R. Urbinati, S. Häfner, O. Ohlenschläger, M. S. Swanson, M. Görlach, R. Ramachandran, *Nucleic Acids Res.* **2004**, *32*, 1177–1183.

[113] K. Riedel, C. Herbst, S. Häfner, J. Leppert, O. Ohlenschläger, M. S. Swanson, M. Görlach, R. Ramachandran, *Angew. Chemie - Int. Ed.* **2006**, *45*, 5620–5623.

[114] C. A. Grosshans, T. R. Cech, *Nucleic Acids Res.* **1991**, *19*, 3875–3880.

[115] H. Schürer, K. Lang, J. Schuster, M. Mörl, *Nucleic Acids Res.* **2002**, *30*, e56.

[116] M. Stoldt, J. Wöhnert, M. Görlach, L. R. Brown, *EMBO J.* **1999**, *18*, 6508–6521.

[117] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325–1327.

[118] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graph.* **1996**, *14*, 51–55.

[119] G. A. Morris, R. Freeman, *J. Am. Chem. Soc.* **1979**, *101*, 760–762.

[120] A. Bax, R. Freeman, T. A. Frenkiel, M. H. Levitt, *J. Magn. Reson.* **1981**, *43*, 478–483.

[121] H. Geen, R. Freeman, *J. Magn. Reson.* **1991**, *93*, 93–141.

[122] A. J. Shaka, C. J. Lee, A. Pines, *J. Magn. Reson.* **1988**, *77*, 274–293.

[123] N. B. Leontis, E. Westhof, *RNA* **2001**, *7*, 499–512.

[124] K. Ösapay, D. A. Case, *J. Am. Chem. Soc.* **1991**, *113*, 9436–9444.

[125] T. Asakura, K. Taoka, M. Demura, M. P. Williamson, *J. Biomol. NMR* **1995**, *6*, 227–236.

[126] D. Sharma, K. Rajarathnam, *J. Biomol. NMR* **2000**, *18*, 165–171.

[127] S. P. Mielke, V. V. Krishnan, *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *54*, 141–165.

[128] G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **1999**, *13*, 289–302.

[129] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213–223.

[130] S. Barton, X. Heng, B. A. Johnson, M. F. Summers, *J. Biomol. NMR* **2013**, *55*, 33–46.

[131] T. Aeschbacher, M. Schubert, F. H. T. Allain, *J. Biomol. NMR* **2012**, *52*, 179–190.

[132] R. Schnieders, C. Richter, S. Warhaut, V. De Jesus, S. Keyhani, E. Duchardt-Ferner, H. Keller, J. Wöhnert, L. T. Kuhn, A. L. Breeze, et al., *J. Biomol. NMR* **2017**, *69*, 31–44.

[133] O. Duss, P. J. Lukavsky, F. H. T. Allain, *Isot. labeling Biomol. NMR, Adv. Exp. Med. Biol.* **2012**, *992*, 121–142.

[134] Z. Serber, C. Richter, D. Moskau, J. Bo, T. Gerfin, D. Marek, L. Baselgia, F. Laukien, A. S. Stern, J. C. Hoch, et al., *J. Am. Chem. Soc.* **2000**, *122*, 3554–3555.

[135] K. Takeuchi, G. Heffron, D. P. Frueh, G. Wagner, *J. Biomol. NMR* **2010**, *47*, 271–282.

[136] M. Gal, K. A. Edmonds, A. G. Milbradt, K. Takeuchi, G. Wagner, *J. Biomol. NMR* **2011**, *51*, 497–504.

[137] K. Takeuchi, H. Arthanari, I. Shimada, G. Wagner, *J. Biomol. NMR* **2015**, *63*, 1–9.

[138] K. Takeuchi, H. Arthanari, M. Imai, G. Wagner, *J. Biomol. NMR* **2016**, *64*, 143–151.

[139] K. V. Pervushin, G. Wider, K. Wüthrich, *J. Biomol. NMR* **1998**, *12*, 345–348.

[140] M. Rance, J. P. Loria, A. G. Palmer, *J. Magn. Reson.* **1999**, *136*, 92–101.

[141] A. Grishaev, L. Yao, J. Ying, A. Pardi, A. Bax, *J. Am. Chem. Soc.* **2009**, *131*, 9490–9491.

[142] J. Dittmer, C.-H. Kim, G. Bodenhausen, *J. Biomol. NMR* **2003**, *26*, 259–275.

[143] J. Czernek, *J. Phys. Chem.* **2001**, *105*, 1357–1365.

[144] M. Ziegeler, M. Cevec, C. Richter, H. Schwalbe, *Chembiochem* **2012**, *13*, 2100–2112.

[145] S. Imai, P. Kumar, C. U. T. Hellen, V. M. D. Souza, G. Wagner, *Nat. Struct. Mol. Biol.* **2016**, *23*, 859–864.

[146] J. G. de la Torre, M. L. Huertas, B. Carrasco, *J. Magn. Reson.* **2000**, *147*, 138–146.

[147] B. Zhao, A. L. Hansen, Q. Zhang, *J. Am. Chem. Soc.* **2014**, *136*, 20–23.

[148] J. Guillerez, P. J. Lopez, F. Proux, H. Launay, M. Dreyfus, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 5958–63.

[149] C. Helmling, S. Keyhani, F. Sochor, B. Fürtig, M. Hengesbach, H. Schwalbe, **2015**, 67–76.

[150] S. C. Walker, J. M. Avis, G. L. Conn, *Nucleic Acids Res.* **2003**, *31*, 1–6.

[151] R. Silvers, H. Keller, H. Schwalbe, M. Hengesbach, *Chembiochem* **2015**, *16*, 1109–1114.

[152] J. S. Paige, K. Y. Wu, S. R. Jaffrey, *Science (80-.).* **2011**, *333*, 642–646.

[153] E. Duchardt-Ferner, S. R. Gottstein-Schmidtke, J. E. Weigand, O. Ohlenschläger, J.-P. Wurm, C. Hammann, B. Suess, J. Wöhnert, *Angew. Chemie - Int. Ed.* **2016**, *55*, 1527–1530.

[154] M. Stoldt, J. Wöhnert, M. Görlach, L. R. Brown, *EMBO J.* **1998**, *17*, 6377–6384.

[155] J. Noeske, C. Richter, M. A. Grundl, H. R. Nasiri, H. Schwalbe, J. Wöhnert, *Proc. Natl. Acad. Sci.* **2005**, *102*, 1372–1377.

[156] A. Bax, G. A. Morris, *J. Magn. Reson.* **1981**, *42*, 501–505.

[157] A. Favier, B. Brutscher, *J. Biomol. NMR* **2011**, *49*, 9–15.

[158] Z. Solyom, M. Schwarten, L. Geist, R. Konrat, D. Willbold, B. Brutscher, *J. Biomol. NMR* **2013**, *55*, 311–321.

[159] J. Rinnenthal, B. Klinkert, F. Narberhaus, H. Schwalbe, *Nucleic Acids Res.* **2010**, *38*, 3834–47.

[160] Z. Zhou, R. Kümmerle, X. Qiu, D. Redwine, R. Cong, A. Taha, D. Baugh, B. Winniford, *J. Magn. Reson.* **2007**, *187*, 225–33.

[161] D. Rovnyak, J. C. Hoch, A. S. Stern, G. Wagner, *J. Biomol. NMR* **2004**, *30*, 1–10.

[162] S. S. Wijmenga, B. N. M. van Buuren, *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287–387.

[163] A. Serganov, Y. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A. T. Phan, C. Hobartner, R. Micura, R. R. Breaker, D. J. Patel, *Chem. Biol.* **2004**, *11*, 1729–1741.

[164] M. Adrian, B. Heddi, A. T. Phan, *Methods* **2012**, *57*, 11–24.

[165] P. K. Bhattacharya, J. Cha, J. K. Barton, *Nucleic Acids Res.* **2002**, *30*, 4740–4750.

[166] K. Takeuchi, H. Arthanari, G. Wagner, *J. Biomol. NMR* **2016**, *66*, 221–225.

[167] T. A. Hopf, J. B. Ingraham, F. J. Poelwijk, C. P. I. Schärfe, M. Springer, C. Sander, D. S. Marks, *Nat. Biol.* **2017**, *35*, 128–135.

[168] K. A. Matreyek, L. M. Starita, J. J. Stephany, B. Martin, M. A. Chiasson, V. E. Gray, M. Kircher, A. Khechaduri, J. N. Dines, R. J. Hause, et al., *Nat. Genet.* **2018**, *50*, 874–882.

[169] M. Krepl, P. Stadlbauer, M. Otyepka, *J. Phys. Chem. Lett.* **2014**, *5*, 1771–1782.

[170] L. Casalino, G. Palermo, A. Spinello, U. Rothlisberger, A. Magistrato, *Proc. Natl. Acad. Sci.* **2018**, *115*, 6584–6589.

[171] J. Graf, P. H. Nguyen, G. Stock, H. Schwalbe, *J. Am. Chem. Soc.* **2007**, *129*, 1635–1641.

[172] R. B. Best, N. Buchete, G. Hummer, *Biophys. J.* **2008**, *95*, 7–9.

[173] R. B. Best, G. Hummer, *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

[174] A. A. Chen, A. E. García, *Proc. Natl. Acad. Sci.* **2013**, *110*, 16820–16825.

[175] G. M. Giambasu, D. M. York, D. A. Case, *RNA* **2015**, *21*, 963–974.

[176] D. Tan, S. Piana, R. M. Dirks, D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E1346–E1355.

[177] S. Bottaro, P. J. Nichols, B. Vögeli, M. Parrinello, K. Lindorff-Larsen, *https://doi.org/10.1101/690412* **2020**, 1–18.

[178] A. N. Borkar, P. Vallurupalli, C. Camilloni, L. E. Kay, M. Vendruscolo, *Phys Chem Chem Phys* **2017**, *19*, 2797–2804.

[179] C. Hartlmüller, J. C. Günther, A. C. Wolter, J. Wöhnert, M. Sattler, T. Madl, *Sci. Rep.* **2017**, *7*, 1–10.

[180] P. J. Nichols, A. Born, M. A. Henen, D. Strotz, C. N. Celestine, P. Güntert, B. Vögeli, *Chembiochem* **2018**, *19*, 1695–1701.

[181] E. Duchardt, C. Richter, O. Ohlenschläger, M. Görlach, J. Wöhnert, H. Schwalbe, *J. Am. Chem. Soc.* **2004**, *126*, 1962–1970.

[182] J. Ferner, A. Villa, E. Duchardt, E. Widjajakusuma, J. Wöhnert, G. Stock, H. Schwalbe, *Nucleic Acids Res.* **2008**, *36*, 1928–1940.

[183] S. Nozinovic, C. Richter, J. Rinnenthal, B. Fürtig, E. Duchardt-Ferner, J. E. Weigand, H. Schwalbe, *J. Am. Chem. Soc.* **2010**, *132*, 10318–10329.

[184] G. A. Bermejo, G. M. Clore, C. D. Schwieters, *Structure* **2016**, *24*, 806–815.

[185] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65–73.

[186] C. D. Schwieters, J. J. Kuszewski, G. M. Clore, *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47–62.

[187] B. Vögeli, T. F. Segawa, D. Leitz, A. Sobol, A. Choutko, D. Trzesniak, W. van Gunsteren, R. Riek, *J. Am. Chem. Soc.* **2009**, *131*, 17215–17225.

[188] P. J. Nichols, M. A. Henen, A. Born, D. Strotz, P. Güntert, B. Vögeli, *Commun. Biol.*

2018, *1*, 1–11.

[189] M. Zuker, *Nucleic Acids Res.* **2003**, *31*, 3406–3415.

[190] J. P. Sheehy, A. R. Davis, B. M. Znosko, *RNA* **2010**, *16*, 417–429.

[191] T. Dale, R. Smith, M. J. Serra, *RNA* **2000**, *6*, 608–615.

[192] F. M. Jucker, H. A. Heus, P. F. Yip, E. H. M. Moors, A. Pardi, *J. Mol. Biol.* **1996**, *264*, 968–980.

[193] F. M. Jucker, A. Pardi, *Biochemistry* **1995**, *34*, 14416–14427.

[194] H. A. Heus, A. Pardi, *Science (80-.).* **1991**, *256*, 191–194.

[195] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. D. Jane, E. Oldfield, J. L. Markley, B. D. Sykes, *J. Biomol. NMR* **1995**, *6*, 135–140.

[196] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

[197] J. P. Linge, S. I. O'Donoghue, M. Nilges, *Methods Enzymol.* **2001**, *339*, 71–90.

[198] M. Ebrahimi, P. Rossi, C. Rogers, G. S. Harbison, *J. Magn. Reson.* **2001**, *150*, 1–9.

[199] A. V. Cherepanov, C. Glaubitz, H. Schwalbe, *Angew. Chemie Int. Ed.* **2010**, *49*, 4747–4750.

[200] C. Altona, M. Sundaralingam, *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.

[201] D. E. Draper, *RNA* **2004**, *10*, 335–343.

[202] D. E. Draper, *Biophys. J.* **2008**, *95*, 5489–5495.

[203] J. C. Bowman, T. K. Lenz, N. V Hud, L. D. Williams, *Curr. Opin. Struct. Biol.* **2012**, *22*, 262–272.

[204] A. R. Ferré-D'Amaré, W. C. Winkler, *Met Ions Life Sci.* **2011**, *9*, 141–173.

[205] W. G. Scott, J. T. Finch, A. Klug, *Cell* **1995**, *81*, 991–1002.

[206] H. Schwalbe, W. Samstag, J. W. Engels, W. Bermel, C. Griesinger, *J. Biomol. NMR* **1993**, *3*, 479–486.

[207] C. Richter, B. Reif, K. Wörner, S. Quant, J. P. Marino, W. Engels, C. Griesinger, H. Schwalbe, *J. Biomol. NMR* **1998**, *2*, 223–230.

[208] H. Schwalbe, J. P. Marino, S. J. Glaser, C. Griesinger, *J. Am. Chem. Soc.* **1995**, 7251–7252.

[209] S. J. Glaser, H. Schwalbe, J. P. Marino, C. Griesinger, *J. Magn. Reson. Ser. B* **1996**, *180*, 160–180.

[210] J. P. Marino, H. Schwalbe, S. J. Glaser, C. Griesinger, *J. Am. Chem. Soc.* **1996**, *7863*, 4388–4395.

[211] E. Duchardt, C. Richter, B. Reif, S. J. Glaser, J. W. Engels, C. Griesinger, H. Schwalbe, *J. Biomol. NMR* **2001**, *21*, 117–126.

[212] J. Rinnenthal, C. Richter, J. Ferner, E. Duchardt, H. Schwalbe, *J. Biomol. NMR* **2007**,

39, 17–29.

[213] I. C. Felli, C. Richter, C. Griesinger, H. Schwalbe, *J. Am. Chem. Soc.* **1999**, *121*, 1956–1957.

[214] C. Richter, C. Griesinger, I. Felli, P. T. Cole, G. Varani, H. Schwalbe, **1999**, 241–250.

[215] E. Duchardt, H. Schwalbe, *J. Biomol. NMR* **2005**, *32*, 295–308.

[216] A. G. Haasnoot, M. de Leeuw, H. de Leeuw, C. Altona, *Org. Magn. Reson.* **1981**, *15*, 43–52.

Chapter V: Appendix - Supporting Information for ^{13}C-detection

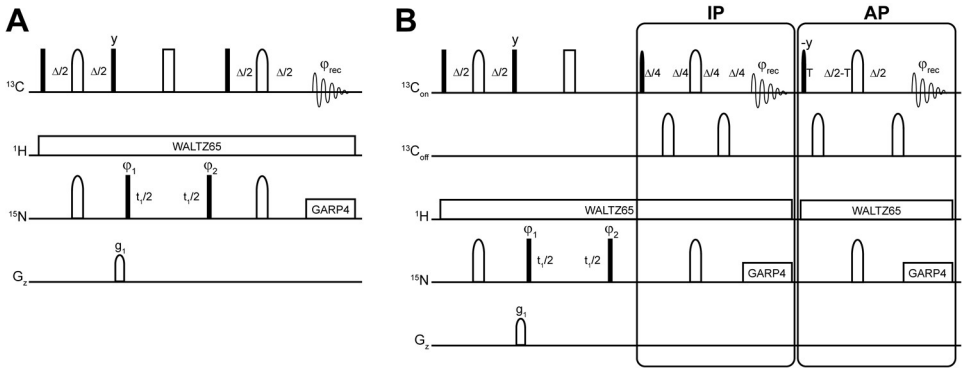


Figure 57 Schematic pulse sequence of the ^{13}C-detected ^{13}C,^{15}N-HSQC experiment without **A** and with **B** IPAP scheme for homonuclear CC-decoupling. Narrow-filled bars represent rectangular 90° pulses while wide open bars mark rectangular 180° pulses. Gradient pulses and selective 180° pulses (^{13}C: 0.4 ms Q3, on-/ off-resonant and ^{15}N: 1.2 ms Reburp2[121]) are represented using semi-elliptic unfilled shapes. Semi-elliptic narrow-filled shapes describe selective 90° pulses (^{13}C: 0.6 ms Q5, on resonant). The default phase is x. ^{15}N- and ^{1}H-nuclei are decoupled using asynchronous GARP4[93] and WALTZ65 sequences[160], respectively. Gradient pulses were applied with a smoothed square amplitude (SMSQ10.100) for 1 ms and 100% gradient strength corresponds to 53 G/cm. Delays are: $\Delta = 1/2{}^1J(C,N) = 19.23$ ms, $T = 1/4{}^1J(C,C) = 4.5$ ms (cytidine) or 3.3 ms (adenosine). Pulse phases are: φ_1 = x,-x, φ_2 = (4)x,(4)-x, and φ_{rec} = x,-x,x,-x,-x,x,-x,x. Gradient strength is 50% (g_1).

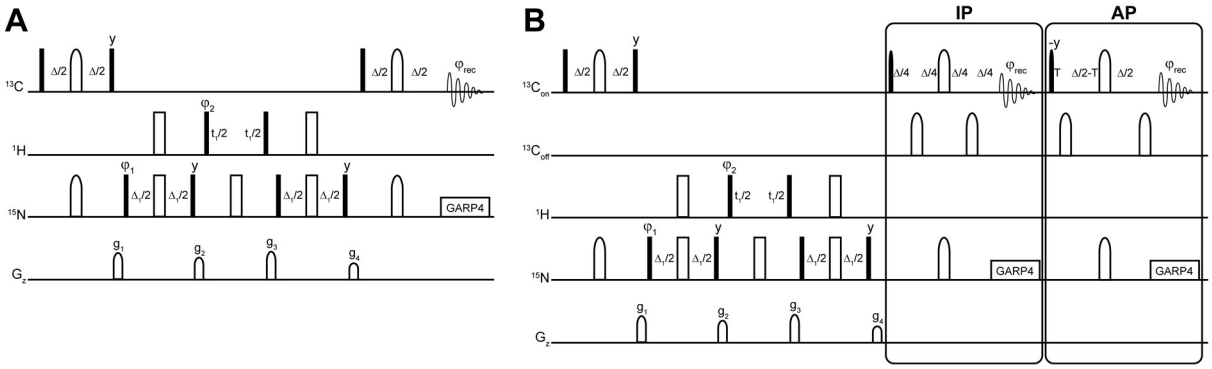


Figure 58 Schematic pulse sequence of the ^{13}C-detected C(N)H-HSQC experiment without **A** and with **B** IPAP scheme for homonuclear CC-decoupling. Narrow-filled bars represent rectangular 90° pulses while wide open bars mark rectangular 180° pulses. Gradient pulses and selective 180° pulses (^{13}C: 0.4 ms Q3, on-/ off-resonant and ^{15}N: 1.2 ms Reburp2[121]) are represented using semi-elliptic unfilled shapes. Semi-elliptic narrow-filled shapes describe selective 90° pulses (^{13}C: 0.6 ms Q5, on resonant). The default phase is x. ^{15}N-nuclei are decoupled during acquisition using asynchronous GARP4 sequences[93]. Gradient pulses were applied with a smoothed square amplitude (SMSQ10.100) for 1 ms and 100% gradient strength corresponds to 53 G/cm. Delays are: $\Delta = 1/2{}^1J(C,N) = 19.23$ ms, $\Delta_1 = 1/4{}^1J(N,H) = 2.08$ ms, $T = 1/4{}^1J(C,C) = 4.5$ ms (cytidine) or 3.3 ms (adenosine). Pulse phases are: φ_1 = (4)x,(4)-x, φ_2 = x,-x, and φ_{rec} = x,-x,x,-x,-x,x,-x,x. Gradient strengths are 50% (g_1), 60% (g_2), 48% (g_3) and 35% (g_4).

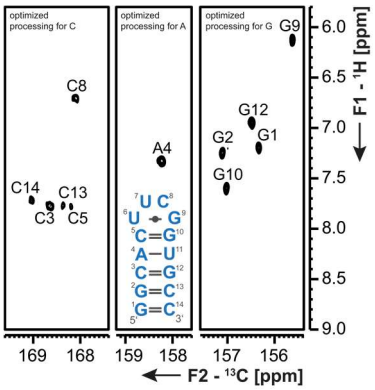


Figure 59 The ^{13}C-detected C(N)H-HDQC spectrum with IPAP decoupling of the 14 nts RNA with UUCG tetraloop. The spectrum has been processed differently for cytidines, adenosines and guanosines, respectively. The spectrum was recorded with spectral windows of 100 ppm and 2 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 160 ppm, 6.9 ppm and 86 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 2048 and 50 in the ^{13}C- and ^{1}H-dimensions, respectively. With 64 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 5 h.

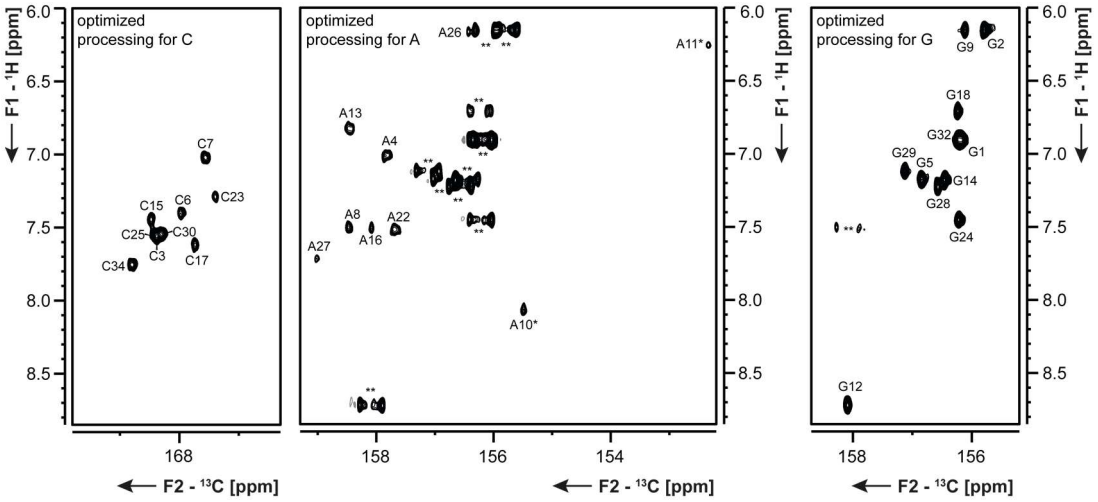


Figure 60 ^{13}C-detected C(N)H-HDQC experiment of the 34 nts GTP class II aptamer[107,108] measured on a 700 MHz NMR spectrometer equipped with a ^{1}H [^{13}C, ^{15}N, ^{31}P]-QCI cryogenic probe, where ^{13}C and ^{15}N channels are on the outer and the ^{1}H channel is on the inner coil. The spectrum has been processed differently for cytidines, adenosines and guanosines, respectively. It was recorded with spectral windows of 50 ppm and 2.8 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 160 ppm, 7.4 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 76 in the ^{13}C- and ^{1}H-dimensions, respectively. With 224 scans per increment and an inter-scan delay of 2.3 s, the experiment was recorded within 23.3 h.

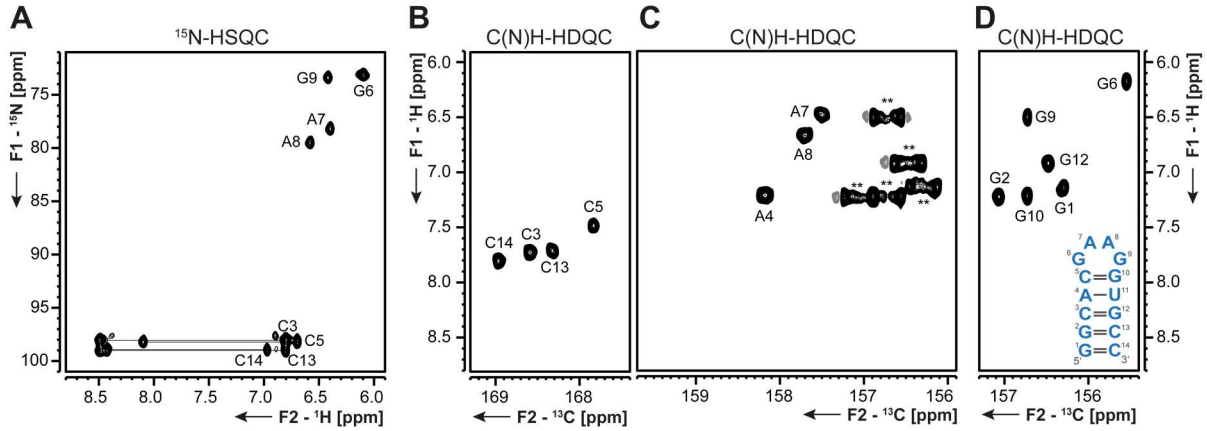


Figure 61 ^{1}H,^{15}N-HSQC spectrum (**A**) and C(N)H-HDQC spectra (**B**, **C**, **D**) of the 14 nts GAAG RNA. **A** was recorded with spectral windows of 12 ppm (^{1}H) and 41 ppm (^{15}N). Carrier frequencies were set to 4.7 ppm, 150 ppm and 85 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.07 s for the direct and 0.01 s for the indirect dimension. With an inter-scan delay of 1 s and 8 scans per increment, the experiment was recorded within 10 min. Spectra in **B**, **C** and **D** are stemming from the same experiment but are processed differently. The experiment was recorded with spectral windows of 100 ppm (^{13}C) and 4 ppm (^{1}H). Carrier frequencies were set to 160 ppm, 7.0 ppm and 86 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.01 s for the indirect dimension. With an inter-scan delay of 2.5 s and 64 scans per increment, the experiment was recorded within 4 h.

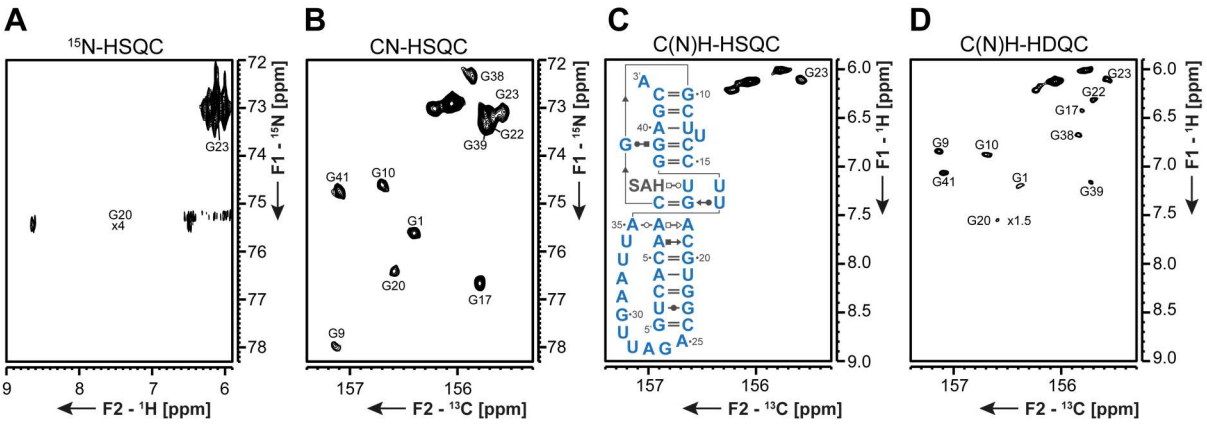


Figure 62 ^{1}H,^{15}N-HSQC (**A**), CN-HSQC (**B**), C(N)H-HSQC (**C**) and C(N)H-HDQC spectra (**D**) of the 43 nts SAM/SAH binding riboswitch[110,111]. **A** was recorded with spectral windows of 25 ppm (^{1}H) and 34 ppm (^{15}N). Carrier frequencies were set to 4.7 ppm, 150 ppm and 85 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.03 s for the indirect dimension. With an inter-scan delay of 1 s and 128 scans per increment, the experiment was recorded within 6 h. **B** was recorded with spectral windows of 50 ppm (^{13}C) and 8 ppm (^{15}N). Carrier frequencies were set to 156 ppm, 7.1 ppm and 75 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.03 s for the indirect dimension. With an inter-scan delay of 2.5 s and 256 scans per increment, the experiment was recorded within 6.5 h. **C** and **D** were recorded with spectral windows of 100 ppm (^{13}C) and 4.2 ppm (^{1}H). Carrier frequencies were set to 160 ppm, 7.0 ppm and 75 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.01 s for the indirect dimension. With an inter-scan delay of 2.5 s and 416 scans per increment, each experiment was recorded within 27 h.

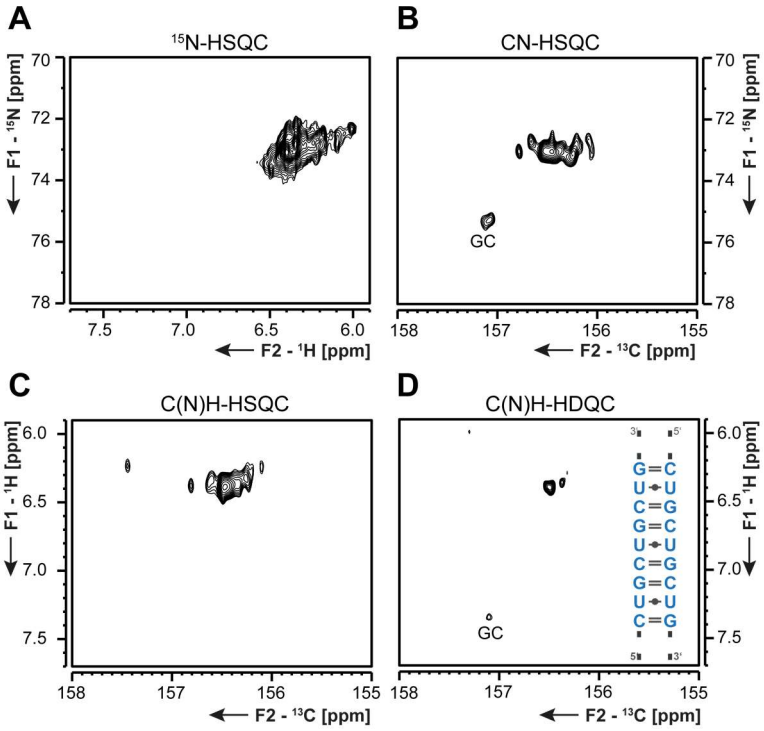


Figure 63 ^{1}H,^{15}N-HSQC (**A**), CN-HSQC (**B**), C(N)H-HSQC (**C**) and C(N)H-HDQC spectra (**D**) of the 329 nts CUG repeat RNA[112,113]. **A** was recorded with spectral windows of 16 ppm (^{1}H) and 30 ppm (^{15}N). Carrier frequencies were set to 4.7 ppm, 156 ppm and 85 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.08 s for the direct and 0.02 s for the indirect dimension. With an inter-scan delay of 1 s and 112 scans per increment, the experiment was recorded within 4 h. **B** was recorded with spectral windows of 50 ppm (^{13}C) and 28 ppm (^{15}N). Carrier frequencies were set to 156 ppm, 7.1 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.02 s for the indirect dimension. With an inter-scan delay of 2.9 s and 128 scans per increment, the experiment was recorded within 12.5 h. **C** and **D** were recorded with spectral windows of 100 ppm (^{13}C) and 5 ppm (^{1}H). Carrier frequencies were set to 156 ppm, 7.0 ppm and 75 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.01 s for the indirect dimension. With an inter-scan delay of 3 s and 512 scans per increment, each experiment was recorded within 32.5 h.

5.1 Detailed figure captions

5.1.1 Experimental parameters for spectra in Figure 19

A The "amino"-NOESY experiment without ^{13}C-filter was recorded with spectral windows of 100 ppm and 10.5 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 160 ppm, 8.8 ppm and 86 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 2048 and 208 in the ^{13}C- and ^{1}H-dimensions, respectively. With 208 scans per increment and an inter-scan delay of 2.5 s, the experiment was recorded within 2 d 20 h. The "amino"-NOESY experiment with ^{13}C-filter was recorded with spectral windows of 70 ppm and 4.4 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 160 ppm, 6.2 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 88 in the ^{13}C- and ^{1}H-dimensions, respectively. With 1024 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 2 d 18 h.

B The "amino"-NOESY experiment without ^{13}C-filter was recorded with spectral windows of 70 ppm and 12 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 160 ppm, 8.7 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 448 in the ^{13}C- and ^{1}H-dimensions, respectively. With 128 scans per increment and an inter-scan delay of 1.2 s, the experiment was recorded within 2 d. The "amino"-NOESY experiment with ^{13}C-filter was recorded with spectral windows of 70 ppm and 5.4 ppm in the direct ^{13}C- and indirect ^{1}H-dimensions, respectively. Carrier frequencies were set to 160 ppm, 6.0 ppm and 85 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Complex points were 1024 and 208 in the ^{13}C- and ^{1}H-dimensions, respectively. With 256 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 1 d 15 h.

5.1.2 Experimental parameters for spectra in Figure 25

Left panels: C(N)H-HDQC spectra of the 34 nts RNA in absence of GTP and Mg^{2+} ions. The experiment was recorded with spectral windows of 50 ppm (^{13}C) and 5 ppm (^{1}H). Carrier frequencies were set to 154 ppm, 7.1 ppm and 90 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.04 s for the indirect dimension. With an inter-scan delay of 2.5 s and 96 scans per increment, the experiment was recorded within 23 h.

Middle panels: C(N)H-HDQC spectra of the 34 nts RNA in absence of GTP but in the presence of 2 mM Mg^{2+} ions. The experiment was recorded with spectral windows of 50 ppm (^{13}C) and 4.2 ppm (^{1}H). Carrier frequencies were set to 154 ppm, 6.9 ppm and 90 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.10 s for the direct and 0.03 s for the indirect dimension. With an inter-scan delay of 2.4 s and 80 scans per increment, the experiment was recorded within 13 h.

5.2 34 nts RNA – New NOE contacts

Table 17 Through the ^{13}C-detected "amino"-NOESY experiment newly determined NOE contacts of the 34 nts GTP class II aptamer RNA[107,108]. NOE contacts were used with an upper distance limit of 6.5 Å. For the amino group a pseudo atom "Q" was used.

residue	nucleotide	atom	residue	nucleotide	atom	distance
1	G	Q2	2	G	H1'	6.50 Å
2	G	Q2	2	G	H1'	6.50 Å
2	G	Q2	3	C	H1'	6.50 Å
2	G	Q2	34	C	H1'	6.50 Å
3	C	H41	4	A	Q6	6.50 Å
3	C	H42	4	A	Q6	6.50 Å
4	A	H1'	32	G	Q2	6.50 Å
5	G	H1	5	G	Q2	6.50 Å
6	C	Q4	7	C	H41	6.50 Å
6	C	Q4	7	C	H42	6.50 Å
8	A	H2	10	A	Q6	6.50 Å
9	G	H1'	23	C	Q4	6.50 Å
9	G	H8	23	C	Q4	6.50 Å
12	G	H1'	15	C	Q4	6.50 Å
12	G	Q2	23	C	H2'	6.50 Å
14	G	H1	14	G	Q2	6.50 Å
14	G	Q2	15	C	H1'	6.50 Å
14	G	Q2	26	A	H1'	6.50 Å
24	G	Q2	25	C	H1'	6.50 Å
28	G	H1	28	G	Q2	6.50 Å
29	G	H1	29	G	Q2	6.50 Å
32	G	Q2	33	U	H1'	6.50 Å

5.3 Pulse programs

5.3.1 ^{13}C-detected amino-selective CN-HSQC

```
;c_con_iasqsp.rs2
;avance-version (12/01/11)
;CON
;2D sequence with 13C detected correlation for triple
resonance using inept transfer steps
;
;     F1(C) -> F3(N,t1) -> F1(C,t2)
;
;on/off resonance 13C pulses using shaped pulses
;phase sensitive (t1)

prosol relations=<triple_c>
#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p22=p21*2"
"p2=p1*2"
"d11=30m"
"d12=20u"
"d23=1/4*cnst4"

"d0=3u"
"in0=inf1/2"

"DELTA=d0*2+p2"
"DELTA1=d23/2-p12/2-p31/4"
"DELTA2=d23-d22-p12-p31/2"
"DELTA3=d23-p31/2"

1 ze
  d11 pl12:f2
2 d11 do:f2 do:f3
3 d1 pl3:f3 pl1:f1
  50u UNBLKGRAD
  d12 cpd2:f2
  (p1 ph1):f1
  DELTA3 pl30:f3
```

```
 (center (p12:sp24 ph1):f1 (p31:sp31 ph1):f3 )
 DELTA3 pl1:f1
 (p1 ph2):f1
 p16:gp1
 d16 pl3:f3
 (p21 ph3):f3
 d0
 (p2 ph5):f1
 d0
 (p22 ph1):f3
 DELTA pl0:f1
 (p21 ph4):f3

if "l0 %2 == 1"
   {
   (p11:sp23 ph1)
   DELTA1
   (p12:sp26 ph1)
   DELTA1
   (center (p12:sp24 ph1) (p31:sp31 ph1):f3)
   DELTA1
   (p12:sp26 ph1)
   DELTA1 pl16:f3
   }
 else
   {
   (p11:sp23 ph6)
   d22
   (p12:sp26 ph1)
   DELTA2
   (center (p12:sp24 ph1) (p31:sp31 ph1):f3 )
   DELTA1
   DELTA1 pl16:f3
   (p12:sp26 ph1)
   }
 4u BLKGRAD
 go=2 ph31 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2
   F1I(iu0, 2)
   F1PH(calph(ph3, +90), caldel(d0, +in0))
exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=3
ph31=0 2 0 2 2 0 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD
;pl16: f3 channel - power level for CPD
;sp23: f1 channel - shaped pulse 90 degree (on resonance)
;sp24: f1 channel - shaped pulse 180 degree (on resonance)
;sp26: f1 channel - shaped pulse 180 degree (C off resonance)
;p11: f1 channel - 90 degree shaped pulse
;p12: f1 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse [1 ms]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (F1 in 2D) [3 μs]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 ms]
;d12: delay for power switching [20 μs]
;d16: delay for homospoil/gradient recovery
;d22: 1/(4J(CC))
;d23: 1/(4J(CN))
;o1p: C on chemical shift
;inf1: 1/SW(N) = 2 * DW(N)
;in0: 1/(2 * SW(N)) = DW(N)
;nd0: 2
;ns: 8 * n
;ds: >= 32
;td1: number of experiments in F1
;FnMODE: States-TPPI (or TPPI) in F1
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;for z-only gradients:
;gpz1: 50%
;use gradient files:
;gpnam1: SMSQ10.100
```

5.3.2 ^{13}C-detected C(N)H-HSQC experiment

```
;c_coniasqsp.rs22.8
;avance-version (12/01/11)
;2D sequence with 13C detected correlation for triple
resonance using inept transfer steps
;on/off resonance 13C pulses using shaped pulses
;phase sensitive (t1)

prosol relations=<triple_c>
```

```
#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p22=p21*2"
"p2=p1*2"
"d11=30m"
"d12=20u"
"d23=1/4*cnst4"
"d24=1/8*cnst5"

"d0=3u"
"in0=inf1/2"

"DELTA=d0*2+p22"
"DELTA1=d24-d12-4u-p22/2"
"DELTA2=d23-p31/2"
"DELTA3=d24-p22/2"
"DELTA4=d23/2-p12/2-p31/4"
"DELTA5=d23-d22-p12-p31/2"

1 ze
 d11 pl12:f2
2 d11 do:f2 do:f3
3 d1 pl3:f3 pl1:f1
 50u UNBLKGRAD
 (p1 ph1):f1
 DELTA2
 (center (p12:sp24 ph1):f1 (p31:sp31 ph1):f3)
 DELTA2 pl1:f1
 (p1 ph2):f1
 p16:gp1
 d16 pl2:f2 pl3:f3
 (p21 ph4):f3
 DELTA3
 (center (p22 ph1):f3 (p4 ph1):f2)
 DELTA3 pl3:f3
 (p21 ph2):f3
 p16:gp3
 d16
 (p3 ph3):f2
 d0
 (p22 ph5):f3
 d0
 (p4 ph1):f2
 DELTA
 (p3 ph1):f2
 p16:gp4
 d16
 (p21 ph1):f3
 DELTA3
 (center (p22 ph1):f3 (p4 ph1):f2)
 DELTA3
 (p21 ph2):f3
 p16:gp2
 d16

 if "l0 %2 == 1"
  {
  (p11:sp23 ph1)
  DELTA4
  (p12:sp26 ph1)
  DELTA4
  (center (p12:sp24 ph1) (p31:sp31 ph1):f3)
  DELTA4
  (p12:sp26 ph1)
  DELTA4 pl16:f3
  }
 else
  {
  (p11:sp23 ph6)
  d22
  (p12:sp26 ph1)
  DELTA5
  (center (p12:sp24 ph1) (p31:sp31 ph1):f3 )
  DELTA4
  DELTA4 pl16:f3
  (p12:sp26 ph1)
  }
 4u BLKGRAD
 go=2 ph31 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2
  F1I(iu0, 2)
  F1PH(ip3, id0)
exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=3
ph31=0 2 0 2 2 0 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl16: f3 channel - power level for CPD
;sp23: f1 channel - shaped pulse 90 degree (on
resonance)
;sp24: f1 channel - shaped pulse 180 degree (on
resonance)
;p11: f1 channel - 90 degree shaped pulse
;p12: f1 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse [1 ms]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (F1 in 2D) [3 µs]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 ms]
;d12: delay for power switching [2 µs]
```

```
;d16: delay for homospoil/gradient recovery
;d23: 1/(4J(CN))
;d24: 1/(8J(NH))
;o1p: C chemical shift
;inf1: 1/SW(N) = 2 * DW(N)
;in0: 1/(2 * SW(N)) =  DW(N)
;nd0: 2
;ns: 8 * n
;ds: >= 32
;td1: number of experiments in F1
;FnMODE: States-TPPI (or TPPI) in F1
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;for z-only gradients:
;gpz1: 50%
;gpz2: 45%
;gpz3: 55%
;gpz4: 40%
;use gradient files:
;gpnam1: SMSQ10.100
```

5.3.3 ^{13}C-detected C(N)H-HDQC experiment

```
;c_con_iadqsp.rs.11a
;avance-version (12/01/11)
;2D sequence with 13C detected correlation for triple resonance using inept transfer steps
;on/off resonance 13C pulses using shaped pulses
;phase sensitive (t1)

prosol relations=<triple_c>
#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p22=p21*2"
"p2=p1*2"
"d11=30m"
"d12=20u"
"d23=1/4*cnst4"
"d24=1/4*cnst5"

"d0=3u"
"in0=inf1/4"

"DELTA=d0*2+p22"
"DELTA1=d24-d12-4u-p22/2"
"DELTA2=d23-p31/2"
"DELTA3=d24-p22/2"
"DELTA4=d23/2-p12/2-p31/4"
"DELTA5=d23-d22-p12-p31/2"

1 ze
 d11 pl12:f2
2 d11 do:f2 do:f3
3 d1 pl3:f3 pl1:f1
 50u UNBLKGRAD
 (p1 ph10):f1
 DELTA2 pl30:f3 pl0:f1
 (center (p12:sp24 ph1):f1 (p31:sp31 ph1):f3)
 DELTA2 pl1:f1
 (p1 ph2):f1
 p16:gp1
 d16 pl2:f2 pl3:f3
 (p21 ph4):f3
 DELTA3
 (center (p22 ph1):f3 (p4 ph1):f2)
 DELTA3 pl3:f3
 (p21 ph1):f3
 p16:gp3
 d16
 (p3 ph3):f2
 d0
 (p22 ph5):f3
 d0
 (p4 ph1):f2
 DELTA
 (p3 ph1):f2
 p16:gp4
 d16
 (p21 ph7):f3
 DELTA3
(center (p22 ph1):f3 (p4 ph1):f2)
 DELTA3
 (p21 ph1):f3
  p16:gp2
  d16 pl30:f3 pl0:f1

if "l0 %2 == 1"
   {
   (p11:sp23 ph1)
   DELTA4
   (p12:sp26 ph1)
   DELTA4
   (center (p12:sp24 ph1) (p31:sp31 ph1):f3 )
   DELTA4
   (p12:sp26 ph1)
   DELTA4 pl16:f3
   }
 else
   {
```

```
   (p11:sp23 ph6)
   d22
   (p12:sp26 ph1)
   DELTA5
   (center (p12:sp24 ph1) (p31:sp31 ph1):f3 )
   DELTA4
   DELTA4 pl16:f3
   (p12:sp26 ph1)
   }
 4u BLKGRAD
 go=2 ph31 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2
   F1I(iu0, 2)
   F1PH(ip3, id0)
exit

ph1=0
ph2=1
ph3=(8) 0 0 2 2 4 4 6 6
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=3
ph7=0 0 0 0 0 0 0 0
  2 2 2 2 2 2 2 2
ph10=0 2
ph31=0 2 2 0 2 0 0 2
   2 0 0 2 0 2 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl16: f3 channel - power level for CPD
;sp23: f1 channel - shaped pulse 90 degree (on
resonance)
;sp24: f1 channel - shaped pulse 180 degree (on
resonance)
;p11: f1 channel - 90 degree shaped pulse
;p12: f1 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse [1 ms]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (F1 in 2D) [3 µs]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 ms]
;d12: delay for power switching [20 µs]
;d16: delay for homospoil/gradient recovery
;d23: 1/(4J(CN))
;d24: 1/(4J(NH))
;o1p: C chemical shift
;inf1: 1/SW(N) = 2 * DW(N)
;in0: 1/(2 * SW(N)) = DW(N)
;nd0: 2
;ns: 16 * n
;ds: >= 32
;td1: number of experiments in F1
;FnMODE: States-TPPI (or TPPI) in F1
;cpd3: decoupling according to sequence defined by
cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling
sequence

;for z-only gradients:
;gpz1: 50%
;gpz2: 45%
;gpz3: 55%
;gpz4: 40%
;use gradient files:
;gpnam1: SMSQ10.100
```

5.3.4 ^{13}C-detected “amino”-NOESY experiment

```
;c_con_noesy_xfiasqspcp.rs2
;avance-version (12/01/11)
;2D sequence with 13C detected correlation for triple
resonance using inept transfer steps
;on/off resonance 13C pulses using shaped pulses
;phase sensitive (t1)

prosol relations=<triple_c>
#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p22=p21*2"
"p2=p1*2"
"d11=30m"
"d12=20u"
"d2=1/2*cnst2"
"d23=1/4*cnst4"

"d0=3u"
"in0=inf1/2"

"DELTA3=d23-p31/2-8u"
"DELTA4=d23/2-p12/2-p31/4"
"DELTA5=d23-d22-p12-p31/2"
"DELTA6=d2-d0*2-p22"

1 ze
  d11 pl12:f2
2 d11 do:f2 do:f3
3 d1 pl3:f3 pl2:f2 pl1:f1
  50u UNBLKGRAD
  (p3 ph3):f2
  d2 fq=cnst19(bf ppm):f1
```

```
 (center (p4 ph1):f2 (p1 ph5 p1 ph4):f1)
 DELTA6
 d0
 (center (p22 ph1):f3 (p2 ph1):f1)
 d0
 (p3 ph7):f2
 d8 fq=cnst20(bf ppm):f1
 (p3 ph2):f2
 4u
 4u pl26:f2 pl23:f3                    ;HN INEPT
          ;begin DIPSI3:f2, DIPSI3:f3
4 (p25*2.722 ph11):f2 (p25*2.722 ph13):f3
 (p25*4.389 ph12):f2 (p25*4.389 ph14):f3
 (p25*2.778 ph11):f2 (p25*2.778 ph13):f3
 (p25*3.056 ph12):f2 (p25*3.056 ph14):f3
 (p25*0.333 ph11):f2 (p25*0.333 ph13):f3
 (p25*2.556 ph12):f2 (p25*2.556 ph14):f3
 (p25*4.000 ph11):f2 (p25*4.000 ph13):f3
 (p25*2.722 ph12):f2 (p25*2.722 ph14):f3
 (p25*4.111 ph11):f2 (p25*4.111 ph13):f3
 (p25*3.778 ph12):f2 (p25*3.778 ph14):f3
 (p25*3.889 ph11):f2 (p25*3.889 ph13):f3
 (p25*2.889 ph12):f2 (p25*2.889 ph14):f3
 (p25*3.000 ph11):f2 (p25*3.000 ph13):f3
 (p25*0.333 ph12):f2 (p25*0.333 ph14):f3
 (p25*2.500 ph11):f2 (p25*2.500 ph13):f3
 (p25*4.050 ph12):f2 (p25*4.050 ph14):f3
 (p25*2.830 ph11):f2 (p25*2.830 ph13):f3
 (p25*4.389 ph12):f2 (p25*4.389 ph14):f3

 (p25*2.722 ph12):f2 (p25*2.722 ph14):f3
 (p25*4.389 ph11):f2 (p25*4.389 ph13):f3
 (p25*2.778 ph12):f2 (p25*2.778 ph14):f3
 (p25*3.056 ph11):f2 (p25*3.056 ph13):f3
 (p25*0.333 ph12):f2 (p25*0.333 ph14):f3
 (p25*2.556 ph11):f2 (p25*2.556 ph13):f3
 (p25*4.000 ph12):f2 (p25*4.000 ph14):f3
 (p25*2.722 ph11):f2 (p25*2.722 ph13):f3
 (p25*4.111 ph12):f2 (p25*4.111 ph14):f3
 (p25*3.778 ph11):f2 (p25*3.778 ph13):f3
 (p25*3.889 ph12):f2 (p25*3.889 ph14):f3
 (p25*2.889 ph11):f2 (p25*2.889 ph13):f3
 (p25*3.000 ph12):f2 (p25*3.000 ph14):f3
 (p25*0.333 ph11):f2 (p25*0.333 ph13):f3
 (p25*2.500 ph12):f2 (p25*2.500 ph14):f3
 (p25*4.050 ph11):f2 (p25*4.050 ph13):f3
 (p25*2.830 ph12):f2 (p25*2.830 ph14):f3
 (p25*4.389 ph11):f2 (p25*4.389 ph13):f3
          ;end DIPSI3:f2, DIPSI3:f3
 4u
 DELTA3 pl0:f3 pl30:f1
 4u
 (center (p12:sp24 ph1):f1 (p31:sp31 ph1):f3 )
 4u
 DELTA3 pl3:f3    pl1:f1
 4u                                    ;NyCz
 (p21 ph1):f3
 p16:gp2
 d16 pl0:f3 pl30:f1
  if "l0 %2 == 1"
    {
    (p11:sp23 ph1)
    DELTA4
    (p12:sp26 ph1)
    DELTA4
    (center (p12:sp24 ph1) (p31:sp31 ph1):f3)
    DELTA4
    (p12:sp26 ph1)
    DELTA4 pl16:f3
    }
  else
    {
    (p11:sp23 ph6)
    d22
    (p12:sp26 ph1)
    DELTA5
    (center (p12:sp24 ph1) (p31:sp31 ph1):f3 )
    DELTA4
    DELTA4 pl16:f3
    (p12:sp26 ph1)
    }
 4u BLKGRAD
 go=2 ph31 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2
    F1I(iu0, 2)
    F1PH(ip3, id0)
exit

ph1=0
ph2=1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=3
ph7=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph10=0 2
ph11=0
ph12=2
ph13=0
ph14=2
ph31=0 2 2 0 2 0 0 2 2 0 0 2 0 2 2 0
   2 0 0 2 0 2 2 0 0 2 2 0 2 0 0 2

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl16: f3 channel - power level for CPD
;sp23: f1 channel - shaped pulse 90 degree (on
resonance)
```

;sp24: f1 channel - shaped pulse 180 degree (on resonance)
;p11: f1 channel - 90 degree shaped pulse
;p12: f1 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse [1 ms]
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (F1 in 2D) [3 μs]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 ms]
;d12: delay for power switching [20 μs]
;d16: delay for homospoil/gradient recovery
;d2: 1/(2J(CH))
;d23: 1/(4J(CN))
;o1p: C chemical shift
;inf1: 1/SW(N) = 2 * DW(N)
;in0: 1/(2 * SW(N)) = DW(N)
;nd0: 2
;ns: 32 * n
;ds: >= 32
;td1: number of experiments in F1
;FnMODE: States-TPPI (or TPPI) in F1
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;for z-only gradients:
;gpz1: 35%
;use gradient files:
;gpnam1: SMSQ10.100

Chapter VI: Appendix 2 - Supporting Information for ^{15}N-detection

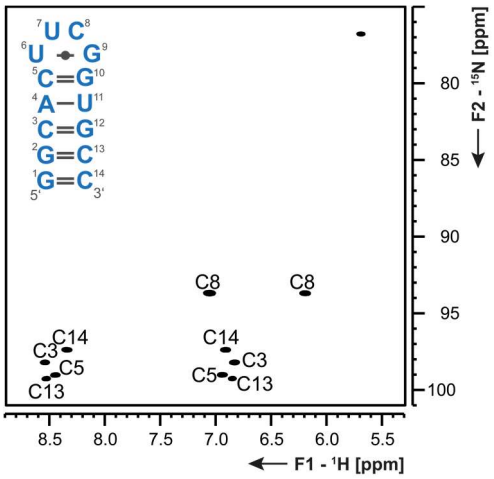


Figure 64 The ^{15}N-detected HSQC spectrum of the amino region was recorded with spectral widths of 33 ppm in the direct ^{15}N- and 4.6 ppm and indirect ^{1}H-dimension. The direct dimension was recorded with 264 complex points while in the indirect dimension 64 complex points were conducted. Carrier frequencies were 86 ppm (^{15}N), 7.1 ppm (^{1}H) and 150 ppm (^{13}C). With 128 scans per increment and an inter-scan delay of 1 s, an overall experimental time of 5 h resulted. The figure has been adapted from Schnieders *et al.*[132].

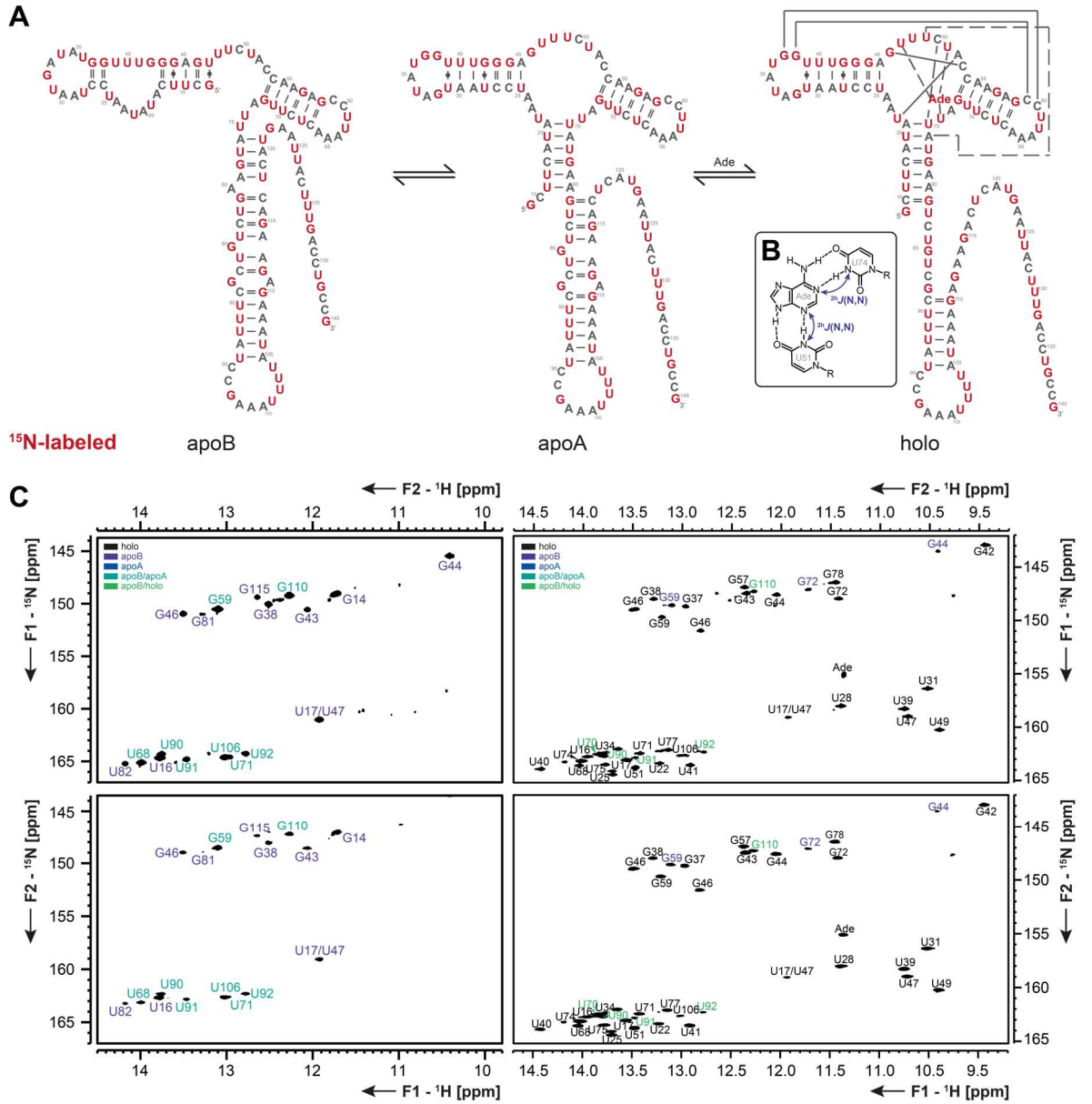


Figure 65 A Secondary structures of the 127 nts adenine-sensing riboswitch with its three conformations apoB (left, binding incompetent), apoA (middle, binding competent) and holo (right, ligand-bound)[14]. All nucleotides written in red are ^{15}N-labeled. **B** Schematic representation of the ligand binding pocket with interactions to residues U51 and U74[163]. The ^{2h}J(N,N) coupling across the hydrogen bond is indicated. **B** ^{1}H- (top) and ^{15}N-detected (bottom) BEST-TROSY spectra of the 127 nts riboswitch in the absence (left) and in the presence (right) of ligand adenine. The ^{1}H-detected spectrum without ligand was recorded with spectral widths of 25 ppm in the direct ^{1}H- and 28 ppm in the indirect ^{15}N-dimension. Carrier frequencies were 4.7 ppm, 101 ppm and 153 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.3 s in in the direct and 0.06 s in the indirect ^{15}N-dimension. With 64 scans per increment and an inter-scan delay of 0.5 s, the experiment was recorded for 4 h. The ^{15}N-detected spectrum in the absence of ligand was recorded with spectral widths of 35 ppm in the direct ^{15}N- and 4.3 ppm in the indirect ^{1}H-dimension. Carrier frequencies were 153 ppm, 150 ppm and 12.2 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. Acquisition times were 0.3 s in the direct dimension and 0.02 s in the indirect dimension. GARP4 decoupling sequences[93] were applied during acquisition to decouple carbon nuclei with 1.8 kHz. With 368 scans per increment and a relaxation delay of 0.3 s the experiment was conducted for 23 h. The ^{1}H-detected spectrum in the presence of ligand adenine was recorded with spectral widths of 25 ppm (^{1}H) and 28 ppm (^{15}N). Carrier frequencies were 4.7 ppm, 101 ppm and 153 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. The experiment was conducted with 64 scans per increment and acquisition times of 0.3 s in the direct and 0.3 s in the indirect dimension. With an inter-scan delay of 0.5 s, the experiment was recorded for 20.5 h. The ^{15}N-detected BEST-TROSY experiment in presence of ligand was recorded with spectral widths of 35 ppm in the direct and 5.2 ppm in the indirect dimension. Carrier frequencies were 155 ppm, 120 ppm and 12.2 ppm for ^{15}N, ^{13}C and ^{1}H, respectively. With acquisition times of 0.3 s in the ^{15}N- and 0.03 s in the ^{1}H-dimension, 576 number of scans per increment and an inter-scan delay of 0.2 s, the experiment was conducted for 20 h. The figure was adapted from Schnieders *et al.*[132].

6.1 Detailed figure captions

6.1.1 Experimental parameters for spectra in Figure 34:

A The ^{1}H-detected BEST-TROSY spectrum (upper panel) of the 40 nts expression platform of the 2'-dG sensing riboswitch[26] was recorded with a spectral width of 4 ppm in the indirect ^{15}N dimension, while the carrier frequency was 162 ppm for ^{15}N. The acquisition time was 0.3 s in the indirect dimension. With 16 scans per increment and an inter-scan delay of 0.5 s, the experiment was recorded for 1 h. The ^{15}N-detected BEST-TROSY spectrum (lower panel) was recorded with a spectral width of 3.3 ppm in the indirect ^{1}H-dimension. The corresponding carrier frequency was 13.8 ppm and the ^{1}H acquisition time was 0.15 s. With 16 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded for 2.5 h.

B The ^{1}H-detected BEST-TROSY spectrum (upper panel) of a 47 nts fluoride riboswitch[147] was recorded with a spectral width of 4.5 ppm in the indirect ^{15}N-dimension and the corresponding carrier frequency was 161 ppm. The acquisition time was set to 0.3 s in the ^{15}N-dimension. With 16 scans per increment and an inter-scan delay of 0.4 s, the experiment was recorded for 1 h. The ^{15}N-detected BEST-TROSY spectrum (lower panel) was recorded with a spectral width of 2.8 ppm in the indirect ^{1}H-dimension. The corresponding carrier frequency was 12.5 ppm and the ^{1}H acquisition time was 0.15 s. With 32 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded for 4.5 h.

C The ^{1}H-detected BEST-TROSY spectrum of a 74 nts Spinach RNA[152] was recorded with a spectral width of 22 ppm in the indirect ^{15}N-dimension, while the corresponding carrier frequency was 154 ppm and the ^{15}N acquisition time was 0.3 s. With 32 scans per increment and an inter-scan delay of 0.4 s, the experiment was recorded for 10 h. The ^{15}N-detected BEST-TROSY spectrum (lower panel) was recorded with a spectral width of 5.3 ppm in the indirect ^{1}H-dimension. The corresponding carrier frequency was 12.0 ppm and the acquisition time was 0.15 s. With 128 scans per increment and an inter-scan delay of 0.3 s, the experiment was recorded for 33 h.

D The ^{1}H-detected BEST-TROSY spectrum of a 329 nts CUG repeat RNA[112] was recorded with a spectral width of 11.4 ppm in the indirect ^{15}N-dimension. The corresponding carrier frequency was 153 ppm and the acquisition time was 0.07 s. With 48 scans per increment and an inter-scan delay of 0.5 s, the experiment was conducted for 1.5 h. The ^{15}N-detected BEST-TROSY spectrum (lower panel) was recorded with a spectral width of 4.2 ppm in the indirect ^{1}H-dimension. The corresponding carrier frequency was 11.6 ppm and the acquisition time was 0.03 s. With 64 scans per increment and an inter-scan delay of 0.2 s, the experiment was recorded for 20 h.

6.2 Pulse programs

6.2.1 ^{15}N-detected HSQC experiment

```
;hxinepph
;avance-version (12/01/11)
;2D heteronuclear shift correlation
;with refocussing of chem. shifts
;phase sensitive
;
;A. Bax & G.A. Morris, J. Magn. Reson. 42, 501 (1981)

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p2=p1*2"
"p4=p3*2"
"d3=1s/(cnst2*cnst11)"
"d4=(1s/(cnst2*4))"
"d11=30m"
"d12=20u"
"d0=3u"

"in0=inf1/2"

"DELTA=d0*2+p2"

1 ze
2 d11
3 d1 do:f2  do:f3
 d12 pl2:f2 pl1:f1
 (p3 ph1):f2
 d0
 p2 ph4
 d0
 d4
 (center (p4 ph2):f2 (p2 ph4):f1 )
 d4
 DELTA
 (p3 ph3):f2 (p1 ph5)
 d3
 (center (p4 ph2):f2 (p2 ph6):f1 )
 d3 pl12:f2 pl16:f3
 go=2 ph31 cpd2:f2 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2 F1PH(calph(ph3, +90),
caldel(d0, +in0))
exit

ph1=0
ph2=0 0 2 2
ph3=1 3
ph4=0 0 2 2
ph5=0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph6=0 0 2 2 1 1 3 3
ph31=0 2 0 2 1 3 1 3 2 0 2 0 3 1 3 1

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;d0 : incremented delay (2D)         [3 µs]
;d1 : relaxation delay; 1-5 * T1
;d3 : 1/(6J(XH))  XH, XH2, XH3 positive
;    1/(4J(XH))  XH only
;  1/(3J(XH))  XH, XH3 positive, XH2 negative
;d4 : 1/(4J(XH))
;d11: delay for disk I/O                [30 ms]
;d12: delay for power switching     [20 µs]
;cnst2: = J(XH)
;cnst11: = 6  XH, XH2, XH3 positive
;        4  XH only
;        3  XH, XH3 positive, XH2 negative
;inf1: 1/SW(H) = 2 * DW(H)
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;ns: 2 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ
;cpd2: decoupling according to sequence defined by
cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling
sequence
```

6.2.2 ^{15}N-detected TROSY experiment

```
;2D 15N shift correlation
;with refocussing of chemical shifts
;phase sensitive

;(A. Bax & G.A. Morris, J. Magn. Reson. 42, 501 (1981))

prosol relations=<triple>

#include <Avance.incl>
#include <Grad.incl>
```

```
#include <Delay.incl>

"p2=p1*2"
"p22=p21*2"
"d11=30m"
"d12=20u"
"d26=1s/(cnst4*4)"

"in0=inf1"
"d0=in0/2-p21*4/PI"

"DELTA1=d26-p16-d16"
"DELTA2=d26-p16-d16-p1*2/PI"
"DELTA3=d26-p16-d16-de"

"cnst62=-((bf3*cnst19)-(o3*1000000))"
"cnst63=-(cnst62*inf1/1000000)*360.0"


"l0=0"
"acqt0=0"
baseopt_echo

1 ze
2 d11 do:f2
3 d1
 d12 pl1:f1 pl12:f2 pl3:f3
 50u UNBLKGRAD

 if "l0 %2 == 1"
   {
   (p21 ph3):f3
   }
 else
   {
   (p21 ph4):f3
   }
 d0
 (p1 ph5):f1
 DELTA1
 p16:gp1
 d16
 4u
 (center (p2 ph1):f1 (p22 ph1):f3 )
 4u
 p16:gp1
 d16
 DELTA1
 (p1 ph1):f1
 (p21 ph2):f3
 DELTA2
 p16:gp2
 d16
 4u
 (center (p2 ph1):f1 (p22 ph1):f3 )
 4u
 p16:gp2
 d16
 DELTA3
 4u BLKGRAD
 (p21 ph6):f3

 go=2 ph31 cpd2:f2
 d11 do:f2 mc #0 to 2
   F1EA(calph(ph5, +180) & calph(ph6, +180) & calclc(l0, 1), caldel(d0, +in0) & calph(ph3, cnst63) & calph(ph4, cnst63))
exit

ph1=0
ph2=1
ph3=1 3 0 2
ph4=3 1 0 2
ph5=3
ph6=2
ph31=0 2 3 1

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD
;pl16: f3 channel - power level for CPD
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p21: f3 channel -  90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                [30 ms]
;d12: delay for power switching     [20 μs]
;d16: delay for homospoil/gradient recovery
;d20: incremented delay (2D)          [3 μs]
;d25: 1/(6J(NH))  NH, NH2 positive
;    1/(4J(NH))  NH only
;    1/(3J(NH))  NH positive, NH2 negative
;d26: 1/(4J(NH))
;cnst4: = J(NH)
;cnst11: = 6  NH, NH2 positive
;        4  NH only
;        3  NH positive, NH2 negative
;inf1: 1/SW(H) = 2 * DW(H)
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;ns: 2 * n
;ds: >=32
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
```

```
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence
;for z-only gradients:
;gpz1: 31%
;gpz2: 19%
;use gradient files:
;gpnam1: SMSQ10.100
```

6.2.3 ^{15}N-detected BEST-TROSY experiment

```
;2D 15N shift correlation
;with refocussing of chemical shifts
;phase sensitive
;(A. Bax & G.A. Morris, J. Magn. Reson. 42, 501 (1981))

prosol relations=<triple_n>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"d11=30m"
"d12=20u"
"d26=1s/(cnst4*4)"

"in0=inf1"
"d0=in0/2"

"DELTA1=d26-larger(p2,p42)/2-p16-d16"
"DELTA2=d26-larger(p2,p42)/2-p16-d16-p1*2/PI"
"DELTA3=d26-larger(p2,p42)/2-p16-d16-de"

"spoff26=bf3*(cnst19/1000000)-o3"
"spoff28=bf3*(cnst19/1000000)-o3"
"spoff29=bf3*(cnst19/1000000)-o3"

"spoal26=0.5"
"spoal28=0"
"spoal29=1"

"cnst62=-((bf3*cnst19)-(o3*1000000))"
"cnst63=-(cnst62*inf1/1000000)*360.0"

"l0=0"
"acqt0=0"
baseopt_echo

1 ze
2 d11 do:f2
3 d1
 d12 pl1:f1 pl12:f2 pl3:f3
 50u UNBLKGRAD
 if "l0 %2 == 1"
   {
   (p43:sp29 ph3):f3
   }
 else
   {
   (p43:sp29 ph4):f3
   }
 d0
 (p1 ph5):f1
 DELTA1
 p16:gp1
 d16
 4u
 (center (p2 ph1):f1 (p42:sp26 ph1):f3 )
 4u
 p16:gp1
 d16
 DELTA1
 (p1 ph1):f1
 (p43:sp28 ph2):f3
 DELTA2
 p16:gp2
 d16
 4u
 (center (p2 ph1):f1 (p42:sp26 ph1):f3 )
 4u
 p16:gp2
 d16
 DELTA3
 4u BLKGRAD
 (p43:sp29 ph6):f3
 go=2 ph31 cpd2:f2
 d11 do:f2 mc #0 to 2
   F1EA(calph(ph5, +180) & calph(ph6, +180) & calclc(l0, 1), caldel(d0, +in0) & calph(ph3, cnst63) & calph(ph4, cnst63))
exit

ph1=0
ph2=1
ph3=1 3 0 2
ph4=3 1 0 2
ph5=3
ph6=2
ph31=0 2 3 1
```

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;pl16: f3 channel - power level for CPD/BB decoupling
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p21: f3 channel - 90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30 ms]
;d12: delay for power switching [20 μs]
;d16: delay for homospoil/gradient recovery
;d20: incremented delay (2D) [3 μs]
;d25: 1/(6J(NH)) NH, NH2 positive
; 1/(4J(NH)) NH only
; 1/(3J(NH)) NH positive, NH2 negative
;d26: 1/(4J(NH))
;cnst4: = J(NH)
;cnst11: = 6 NH, NH2 positive
; 4 NH only
; 3 NH positive, NH2 negative
;inf1: 1/SW(H) = 2 * DW(H)
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;ns: 2 * n
;ds: >=32
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence
;for z-only gradients:
;gpz1: 31%
;gpz2: 19%
;use gradient files:
;gpnam1: SMSQ10.100

Chapter VII: Appendix 3 - Supporting Information for the structural characterization of RNA tetraloops

7.1 Chemical shift assignment of the 14 nts RNA with GAAG tetraloop

7.1.1 Experimental details for experiments in Figure 44 - Figure 46

Long-range ^{1}H,^{13}C-HSQC experiment

The spectrum was recorded with a spectral width of 11 ppm in the direct ^{1}H-dimension and 85 ppm in the indirect ^{13}C dimension. Carrier frequencies were 4.7 ppm, 132.5 ppm and 130 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.07 s in the direct and 0.02 s in the indirect dimension. The INEPT-transfer time was 25 ms (J(H,C) = 20 Hz).With 8 scans per increment and an inter-scan delay of 1 s, the experiment was recorded for 1.5 h.

Long-range ^{1}H,^{15}N-HSQC experiment

The spectrum was recorded with a spectral width of 16 ppm in the direct ^{1}H-dimension and 80 ppm in the indirect ^{15}N-dimension. Carrier frequencies were 4.7 ppm, 150 ppm and 200 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.05 s in the direct and 0.05 s in the indirect dimension. The INEPT-transfer time was 33 ms (J(H,N) = 15 Hz).With 8 scans per increment and an inter-scan delay of 1 s, the experiment was recorded for 1.5 h.

^{13}C,^{15}N-HSQC experiments

The spectrum of the amino region was recorded with a spectral width of 70 ppm in the direct ^{13}C-dimension and 80 ppm in the indirect ^{15}N-dimension. Carrier frequencies were 150 ppm, 4.7 ppm and 90 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.04 s in the direct and 0.01 s in the indirect dimension. The INEPT-transfer time was 18.5 ms (^{1}J(C,N) = 27 Hz).With 64 scans per increment and an inter-scan delay of 2 s, the experiment was recorded within 5 h. The spectra for the pyrimidine C6N1 correlations and the guanosine C6N1 correlations were recorded with a spectral width of 90 ppm in the direct ^{13}C-dimension and 100 ppm in the indirect ^{15}N-dimension. Carrier frequencies were 140 ppm, 4.7 ppm and 190 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.08 s in the direct and 0.02 s in the indirect dimension. The INEPT-transfer time were 30 ms (^{1}J(C,N) = 17 Hz) and 22 ms (^{1}J(C,N) = 23 Hz), respectively. With 80 scans per

increment and an inter-scan delay of 1.5 s, the experiments were recorded within 7 h. The spectrum for all other correlations was recorded with a spectral width of 70 ppm in the direct ^{13}C-dimension and 100 ppm in the indirect ^{15}N-dimension. Carrier frequencies were 150 ppm, 4.7 ppm and 185 ppm for ^{13}C, ^{1}H and ^{15}N, respectively. Acquisition times were 0.04 s in the direct and 0.01 s in the indirect dimension. The INEPT-transfer time was 18.5 ms (^{1}J(C,N) = 27 Hz).With 64 scans per increment and an inter-scan delay of 2 s, the experiment was recorded within 5 h.

TROSY relayed HCCH-COSY experiment

The 3D spectrum was recorded with a spectral width of 24 ppm in the direct ^{1}H-dimension and 55 ppm in the indirect ^{13}C-dimensions, respectively. Carrier frequencies were 4.7 ppm, 140 ppm and 150 ppm for ^{1}H, ^{13}C and ^{15}N, respectively. Acquisition times were 0.07 s in the direct and 0.006 s (F2) and 0.003 s (F1) in the indirect dimensions. With 16 scans per increment and an inter-scan delay of 1.4 s, the experiment was recorded within 1 d 16 h.

7.1.2 Resonance overview

Table 18 ^{1}H resonance assignment of the 14 nts RNA with GAAG tetraloop. Chemical shifts are given in ppm at 298 K and are referenced externally to DSS. Chemical shifts, which are written in italic letters are at 278 K. Nuclei which are not present in the respective nucleobase are held in light orange. Chemical shifts of NH_2 groups correspond to ^{1}H-DQ chemical shifts.

	H1	H2	H3	H5	H6	H8	NH_2	H1'	H2'	H3'	H4'	H5'1	H5'2
G1	*12.92*	-	-	-	-	8.16	7.14	5.84	4.96	4.73	4.58	4.44	4.30
G2	*13.34*	-	-	-	-	7.63	7.23	5.93	4.56	4.59	4.58	4.54	4.28
C3	-	-	-	5.31	7.73	-	7.72	5.53	4.50	4.58	4.47	4.61	4.14
A4	-	7.41	-	-	-	8.05	7.21	5.96	4.50	4.67	4.48	4.59	4.16
C5	-	-	-	5.11	7.25	-	7.48	5.39	4.26	4.28	4.33	4.49	4.05
G6	*10.11*	-	-	-	-	7.55	6.18	5.67	4.53	4.61	4.42	4.33	4.08
A7	-	7.92	-	-	-	8.25	6.49	5.66	4.56	4.55	4.00	3.98	3.92
A8	-	7.86	-	-	-	8.08	6.67	5.80	4.66	4.76	4.43	4.14	3.95
G9		-	-	-	-	7.79	6.50	6.00	4.68	4.91	4.57	4.43	4.28
G10	*12.90*	-	-	-	-	8.03	7.22	4.96	4.46	4.35	4.39	4.31	
U11	-	-	*13.89*	5.14	7.78	-	-	5.57	4.67	4.56	4.46	4.50	4.11
G12	*12.71*	-	-	-	-	7.74	6.93	5.82	4.52	4.59	4.51	4.53	4.15
C13	-	-	-	5.27	7.64	-	7.70	5.47	4.19	4.42	4.39	4.55	4.05
C14	-	-	-	5.56	7.48	-	7.79	6.10	4.64	4.95	4.41	4.52	4.09

Table 19 ^{13}C resonance assignment of the 14 nts RNA with GAAG tetraloop. Chemical shifts are given in ppm at 298 K and are referenced as described by Wishart *et al.*[195]. Nuclei which are not present in the respective nucleobase are held in light orange.

	C2	C4	C5	C6	C8	C1'	C2'	C3'	C4'	C5'
G1	156.3	152.3	118.6	162.5	139.1	91.6	75.1	74.8	83.4	67.4
G2	157.1	151.5	118.6	161.8	136.8	93.1	75.4	73.1	82.5	66.1
C3	158.7	168.6	97.6	140.8	-	93.8	75.4	72.2	81.8	64.5
A4	153.2	149.3	120.3	158.2	139.4	93.0	75.8	72.7	82.0	64.9
C5	158.4	167.9	97.6	139.7	-	93.7	75.9	72.3	81.9	64.9
G6	155.6	153.3	118.6	160.8	137.3	91.6	75.6	74.7	83.1	65.8
A7	154.8	150.7	120.2	157.6	142.1	89.8	76.8	76.5	84.6	66.4
A8	154.7	150.0	120.5	157.8	141.3	91.4	76.2	75.9	84.3	66.9
G9	156.8	153.7	118.7	161.6	139.2	90.8	76.1	75.7	83.9	67.3
G10	156.7	152.1	118.7	161.9	138.0	93.1	74.9	74.6	83.3	68.7
U11	152.8	169.3	103.0	141.7	-	93.6	75.2	72.5	82.1	64.8
G12	156.8	151.8	118.6	161.5	136.1	92.8	75.5	72.9	82.0	65.5
C13	158.8	168.3	97.2	141.0	-	94.4	75.7	72.2	82.1	64.5
C14	159.8	169.0	99.3	141.6	-	92.0	85.2	77.8	85.5	65.6

Table 20 ^{15}N and ^{31}P resonance assignment of the 14 nts RNA with GAAG tetraloop. Chemical shifts are given in ppm at 298 K and are referenced as described by Wishart *et al.*[195]. Nuclei which are not present in the respective nucleobase are held in light orange.

	N1	N2	N3	N4	N6	N7	N9	P
G1	147.3	75.1	163.3	-	-	232.7	168.8	-1.133
G2	148.5	74.7	161.8	-	-	233.7	169.4	-0.727
C3	150.7	-	197.1	98.2	-	-	-	-1.255
A4	221.8	-	213.0	-	84.4	230.6	170.9	-0.934
C5	150.3	-	196.1	98.5	-	-	-	-1.014
G6	142.2	73.3	163.4	-	-	238.5	168.4	-0.763
A7	225.5	-	216.5	-	78.3	231.6	168.1	0.204
A8	226.2	-	216.5	-	79.7	230.9	169.6	-0.598
G9	147.1	73.6	166.3	-	-	237.2	169.9	-1.189
G10	147.5	75.4	163.3	-	-	233.8	169.1	0.247
U11	146.2	-	162.3	-	-	-	-	-1.470
G12	146.2	74.8	162.3	-	-	234.7	169.7	-0.939
C13	151.2	-	198.2	99.2	-	-	-	-1.404
C14	150.1	-	196.4	99.2	-	-	-	-1.234 -1.545

7.2 Temperature-dependence of the GAAG tetraloop

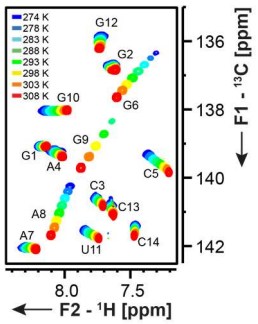


Figure 66 $^{1}H,^{13}C$-HSQC spectra of the aromatic region for the temperature series of the 14 nts RNA with GAAG tetraloop. Peaks are color coded in a rainbow gradient from cold (blue) to hot (red) temperatures. All spectra were recorded using a standard constant-time $^{1}H,^{13}C$-HSQC pulse sequence with the maximum achievable resolution in the indirect dimension at 600 MHz.

7.3 Pulse programs

7.3.1 Γ-HCNCH reference experiment

```
;ddG_CN_final_ref_104

prosol relations=<triple_na>
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"          ;1H
"p4=p3*2"          ;13C
"p22=p21*2"        ;15N

"d0=3u"
"d4=1.25m"   ;H8C8 Tranfer 1J=216Hz
"d5=1.5m"     ;H1'C1' Transfer, 1J=170Hz
"d10=3u"
"d11=30m"
"d13=4u"
"d23=18m"          ;CN Transfer 1J=11 Hz
"d22=16m"
"d27=2.94m"        ;reference experiment
"d28=d27/4-p22/2"

"in0=inf1/2"
"d20=d23-d13-p16-d16-p24"
"in20=in0"

"DELTA1=d22-d13-p16-d16-larger(p14,p30)/2"
"DELTA2=d23-p14/2-d16-p16"
"DELTA3=d25/4-p22/2-p25/4"
"DELTA4=d25/4-p16-d16-d28-p22/2-p25/4"
"DELTA5=d25/4-p16-d16-p22/2-p25/4"
"DELTA6=d25/4-d28-p22/2-p25/4"
"DELTA7=d23-d16-p16-d0-p24"
"TAU=d13+p16+d16+4u"

"spoff3=0"
"spoff9=0"

1 ze
 d11 pl12:f2 pl16:f3
2 d11 do:f2 do:f3
3 d12 pl9:f1
 d1 cw:f1 ph29
 d13 do:f1
 d12 pl1:f1
 d12 fq=cnst21(bf ppm):f2          ;C8, 138ppm
 d13 fq=cnst18(bf ppm):f3          ;N9, 155ppm
 (p1 ph1)                                  ;  Start
Transfer H8 to C8
 d4 pl2:f2 pl1:f1
 (center (p2 ph1) (p4 ph1):f2 )    ;  p14,  1.5ms  on
C8
 d4 UNBLKGRAD
 (p1 ph2)
 4u
 (p11:sp1 ph13:r):f1
 d13
 p16:gp1
 d16 pl2:f2
```

```
 (p3 ph3):f2                                    ;    Start
Transfer C8 to N9
 DELTA1
 d13
 p16:gp2
 d16
 (center (p14:sp3 ph1):f2 (p30:sp9 ph1):f3 )
 d13
 p16:gp2
 d16 pl2:f2 pl3:f3
 DELTA1

 (p3 ph2):f2                                    ;    Start
transfer N9 to C1'
 d13 fq=cnst22(bf ppm):f2              ;on C1', 90ppm
 4u
 d13
 p16:gp11
 d16
 (p21 ph5):f3
 DELTA2
 d13
 p16:gp6
 d16
 (center (p14:sp3 ph1):f2 (p22 ph1):f3 )
 d13
 p16:gp6
 d16 pl2:f2 pl3:f3
 DELTA2
 (p21 ph2):f3
  d13 fq=cnst23(bf ppm):f2             ;middle  C1'  and
C8, 114ppm
 d13 fq=cnst19(bf ppm):f3             ;middle  N7  and
N9, 185ppm
 4u
 (p11:sp1 ph13:r):f1
 d13
 p16:gp12
 d16
 (p3 ph10):f2                                   ;start
DQZQ evolution time
 4u
 DELTA3
 (p22 ph1):f3
 d28 pl1:f1
 (p2 ph1)
 DELTA4
 p16:gp7
 d16
 (p25:sp7 ph12):f2
 DELTA5
 p16:gp7 pl2:f2
 d16
 (p22 ph1):f3
 d28
 (p2 ph1)
 DELTA6

 (p3 ph11):f2                                                ;start
C8N9 ruecktransfer
 d13 fq=cnst21(bf ppm):f2             ;on C8, 138ppm
 d13 fq=cnst18(bf ppm):f3             ;N9, 155ppm
 4u
 d13
 p16:gp10
 d16
 (p21 ph2):f3
 DELTA1
 d13
 p16:gp8
 d16
 (center (p14:sp3 ph6):f2 (p30:sp9 ph1):f3 )
 d13
 p16:gp8
 d16 pl2:f2 pl3:f3
 DELTA1
 (p21 ph1):f3
 d13 fq=cnst22(bf ppm):f2             ;on C1', 90ppm
 4u
 d13
 p16:gp5
 d16

 (p3 ph7):f2                                                 ;start
N9C1' ruecktransfer
 d0
 (center (p2 ph1) (p24:sp5 ph1):f2 )
 DELTA7
 p16:gp3*EA*-1
 d16
 (center (p14:sp3 ph1):f2 (p22 ph1):f3 )
 d13
 p16:gp3*EA
 d16
 d20
 (p24:sp5 ph1):f2
 4u pl2:f2
 (center (p1 ph1) (p3 ph8):f2 )       ;start       C1'H1'
ruecktransfer
 d5
 (center (p2 ph1) (p4 ph1):f2 )
 d5
 (center (p1 ph2) (p3 ph9):f2 )
 d5
 (center (p2 ph1) (p4 ph1):f2 )
 d5
 (p1 ph1)
  TAU
 (p2 ph1)
 d13
```

```
 p16:gp4
 d16 pl12:f2 pl16:f3
 4u  BLKGRAD
 go=2 ph31 cpd2:f2 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2
   F1EA(calgrad(EA) & calph(ph9, +180), caldel(d0, +in0) & caldel(d20, -in20) & calph(ph6, +180) & calph(ph7, +180) & calph(ph31, +180))
exit

ph1=0
ph2=1
ph3=0
ph4=0
ph5=0 2
ph6=0
ph7=1 1 1 1 3 3 3 3
ph8=0 0 2 2
ph9=3 3 1 1
ph10=0 0 0 0 0 0 0 0
   2 2 2 2 2 2 2 2
ph11=1 1 1 1 1 1 1 1
   1 1 1 1 1 1 1 1
   3 3 3 3 3 3 3 3
   3 3 3 3 3 3 3 3
ph12=0
ph13=0
ph29=0
ph31=0 2 2 0 2 0 0 2

;pl0 : 120db
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD
;pl16: f3 channel - power level for CPD
;sp3: f2 channel - shaped pulse 180 degree (on resonance)
;sp5: f2 channel - shaped pulse 180 degree (off resonance)
;sp7: f2 channel - shaped pulse 180 degree (Q3_rna_c68c1.1)
;sp9: f3 channel - shaped pulse 180 degree (on resonance)
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse  [1 ms]
;p21: f3 channel -  90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p30: f3 channel - 180 degree shaped pulse for inversion
;d0 : incremented delay                  [3 us]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J(HC))   for H8':          [1.25 ms]
;d11: delay for disk I/O                 [30 ms]
;d13: short delay                         [4 us]
;d16: delay for homospoil/gradient recovery
;d20: decremented delay
;d22: 1/(4J(CN))                        [16 ms]
;and constant time delay T(C) = 1/(J(CC)) with J(CC)=40 Hz
;d25: mixing time cross correlated relaxation rate (20-30ms)
;in0: 1/(2 * SW(N)) =  DW(N)
;nd0: 2
;nd10: 2
;NS: 8 * n
;DS: >= 32
;td1: number of experiments in F1
;FnMODE: echo-antiecho in F1
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;cnst18: N9 [161ppm]
;cnst19: middle N7 N9 [185ppm]
;cnst21: on C8 [138ppm]
;cnst22: on C1' [90ppm]
;cnst23: middle C8 and C1' [112ppm]


;for z-only gradients:
;gpz1: -40%
;gpz2: 25%
;gpz3: 40%
;gpz4: 20.1%
;gpz5: 50%
;gpz6: 20%
;gpz7: 20%
;gpz8: 20%
;gpz10: 65%
;gpz11: 50%
;gpz12: -40%

;use gradient files:
;gpname: SMSQ10.100
```

7.3.2 Γ-HCNCH cross experiment

```
;ddG_CN_final_cross_104

prosol relations=<triple_na>
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"          ;1H
"p4=p3*2"          ;13C
"p22=p21*2"        ;15N

"d0=3u"
"d4=1.25m"         ;H8C8 Tranfer 1J=216Hz
"d5=1.5m"          ;H1'C1' Transfer, 1J=170Hz
"d10=3u"
"d11=30m"
"d13=4u"
"d23=18m"          ;CN Transfer 1J=11 Hz
"d22=16m"
"d27=0 us"         ;cross experiment:

"in0=inf1/2"
"d20=d23-d13-p16-d16-p24-p14/2"
"in20=in0"

"DELTA1=d22-d13-p16-d16-larger(p14,p30)/2"
"DELTA2=d23-p14/2-d16-p16"
"DELTA3=d25/4-p22/2-p25/4"
"DELTA4=d25/4-p16-d16-p22/2-p25/4"
"DELTA5=d25/4-p16-d16-p22/2-p25/4"
"DELTA6=d25/4-p22/2-p25/4"
"DELTA7=d23-d16-p16-d0-p24-p14/2"
"TAU=d13+p16+d16+4u"

"spoff3=0"
"spoff9=0"

1 ze
 d11 pl12:f2 pl16:f3
2 d11 do:f2 do:f3
3 d12 pl9:f1
 d1 cw:f1 ph29
 d13 do:f1
 d12 pl1:f1
 d12 fq=cnst21(bf ppm):f2          ;on C8, 138ppm
 d13 fq=cnst18(bf ppm):f3          ;N9, 155ppm
 (p1 ph1)                          ;  Start
Transfer H8 to C8
 d4 pl2:f2 pl1:f1
 (center (p2 ph1) (p4 ph1):f2 )    ;  p14,  1.5ms  on
C8
 d4 UNBLKGRAD
 (p1 ph2)
 4u
 (p11:sp1 ph13:r):f1
 d13 pl1:f1
 p16:gp1
 d16 pl2:f2

 (p3 ph3):f2                       ;  Start
Transfer C8 to N9
 DELTA1
 d13
 p16:gp2
 d16
 (center (p14:sp3 ph1):f2 (p30:sp9 ph1):f3 )
 d13
 p16:gp2
 d16 pl2:f2 pl3:f3
 DELTA1

 (p3 ph2):f2                       ;  Start
transfer N9 to C1'
 d13 fq=cnst22(bf ppm):f2          ;on C1', 90ppm
 4u
 d13
 p16:gp11
 d16
 (p21 ph5):f3
 DELTA2
 d13
 p16:gp6
 d16
 (center (p14:sp3 ph1):f2 (p22 ph1):f3 )
 d13
 p16:gp6
 d16 pl2:f2 pl3:f3
 DELTA2
 (p21 ph2):f3
 d13 fq=cnst23(bf ppm):f2          ;middle  C1'  and
C8, 114ppm
 d13 fq=cnst19(bf ppm):f3          ;middle  N7  and
N9, 185ppm
 4u
 (p11:sp1 ph13:r):f1
 d13 pl1:f1
 p16:gp12
 d16
 (p3 ph10):f2                      ;start
DQZQ evolution time
 4u
 DELTA3
 (center (p22 ph1):f3 (p2 ph1))
 DELTA4
 p16:gp7
 d16
```

```
(p25:sp7 ph12):f2
DELTA5 pl2:f2
p16:gp7
d16
(center (p22 ph1):f3 (p2 ph1))
DELTA6

(p3 ph11):f2                                    ;start
C8N9 ruecktransfer
d13 fq=cnst21(bf ppm):f2            ;on C8, 138ppm
d13 fq=cnst18(bf ppm):f3            ;N9, 155ppm
4u
d13
p16:gp10
d16
(p21 ph2):f3
DELTA1
d13
p16:gp8
d16
(center (p14:sp3 ph6):f2 (p30:sp9 ph1):f3 )
d13
p16:gp8
d16 pl2:f2 pl3:f3
DELTA1
(p21 ph1):f3
d13 fq=cnst22(bf ppm):f2            ;on C1', 90ppm
4u
d13
p16:gp5
d16

(p3 ph7):f2                                     ;start
N9C1' ruecktransfer
d0
(center (p2 ph1) (p24:sp5 ph1):f2 )
DELTA7
p16:gp3*EA*-1
d16
(center (p14:sp3 ph1):f2 (p22 ph1):f3 )
d13
p16:gp3*EA
d16
d20
(p24:sp5 ph1):f2
4u pl2:f2
(center (p1 ph1) (p3 ph8):f2 )      ;start      C1'H1'
ruecktransfer
d5
(center (p2 ph1) (p4 ph1):f2 )
d5
(center (p1 ph2) (p3 ph9):f2 )
d5
(center (p2 ph1) (p4 ph1):f2 )
d5
(p1 ph1)
 TAU
(p2 ph1)
d13
p16:gp4
d16 pl12:f2 pl16:f3
4u  BLKGRAD
go=2 ph31 cpd2:f2 cpd3:f3
d11 do:f2 do:f3 mc #0 to 2
   F1EA(calgrad(EA) & calph(ph9, +180), caldel(d0,
+in0) & caldel(d20, -in20) & calph(ph6, +180) &
calph(ph7, +180) & calph(ph31, +180))
exit


ph1=0
ph2=1
ph3=0
ph4=0
ph5=0 2
ph6=0
ph7=1 1 1 1 3 3 3 3
ph8=0 0 2 2
ph9=3 3 1 1
ph10=0 0 0 0 0 0 0 0
   2 2 2 2 2 2 2 2
ph11=1 1 1 1 1 1 1 1
   1 1 1 1 1 1 1 1
   3 3 3 3 3 3 3 3
   3 3 3 3 3 3 3 3
ph12=0
ph13=0
ph29=0
ph31=0 2 2 0 2 0 0 2

;pl0 : 120db
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD
;pl16: f3 channel - power level for CPD
;sp3: f2 channel - shaped pulse 180 degree (on
resonance)
;sp5: f2 channel - shaped pulse 180 degree (off
resonance)
;sp7: f2 channel - shaped pulse 180 degree
(Q3_rna_c68c1.1)
;sp9: f3 channel - shaped pulse 180 degree (on
resonance)
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse
;p16: homospoil/gradient pulse    [1 ms]
```

```
;p21: f3 channel -  90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p28: f1 channel - trim pulse            [1 ms]
;p30: f3 channel - 180 degree shaped pulse for inversion
;d0 : incremented delay                  [3 us]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J(HC))    for H8':             [1.25 ms]
;d11: delay for disk I/O                 [30 ms]
;d13: short delay                        [4 us]
;d16: delay for homospoil/gradient recovery
;d22: 1/(4J(CN))                         [16 ms]
;and constant time delay T(C) = 1/(J(CC)) with J(CC)=40 Hz
;d25: mixing time cross correlated relaxation rate (20-30ms)
;in0: 1/(2 * SW(N)) =  DW(N)
;nd0: 2
;in10: 1/(2 * SW(C)) = DW(C)
;nd10: 2
;in30: = in10
;NS: 8 * n
;DS: >= 32
;td1: number of experiments in F1
;FnMODE: echo-antiecho in F1
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;cnst18: N9 [161ppm]
;cnst19: middle N7 N9 [185ppm]
;cnst21: on C8 [138ppm]
;cnst22: on C1' [90ppm]
;cnst23: middle C8 and C1' [112ppm]

;for z-only gradients:
;gpz1: -40%
;gpz2: 25%
;gpz3: 40%
;gpz4: 20.1%
;gpz5: 50%
;gpz6: 20%
;gpz7: 20%
;gpz8: 20%
;gpz10: 65%
;gpz11: 50%
;gpz12: -40%

;use gradient files:
;gpname: SMSQ10.100
```

7.3.3 forward directed HCC-TOCSY-CCH-E.COSY experiment

```
;fwhcchtocecos3d (geht, modifiziert für TopSpin2.1, ric 281108)

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p4=p3*2"
"p22=2*p21"

"d25=((p2-p3)/2)"
"d10=d22/2"
"d30=d22/2"
"d18=d23/2"
"d7=d23/2-d2-p1"
"d6=d24-p16-d19"          ;fuer DNA = 0.75m und fuer RNA = 1.5m
"d5=d24-p2-p16-d19"  ;Unterschied CH und CH2
"d0=3u"
"d4=1.5m"

"DELTA=d4-d0*2-p22-p4"
"in0=inf1/2"
"in10=inf2/2"
"in30=in10"
"l4=d8/((p27+4u)*115.132)"

aqseq 312

1 ze
  d11 pl12:f2 pl16:f3
2 d11 do:f2 do:f3
3 d12 do:f2  pl9:f1
  d1 cw:f1
  4u do:f1
  20u pl1:f1 pl3:f3
  (p1 ph21)
   d0 pl2:f2
  (p22 ph1):f3
   DELTA
  (p4 ph0):f2
  d0
  (p2 ph0)
  d4
```

```
 (p1 ph1)
 50u UNBLKGRAD
 p16:gp1
 d19
 (p3 ph2):f2
 d6
 p16:gp2
 d19
 (p4 ph0):f2
 (p2 ph0)
 p16:gp2
 d19
 d5 pl8:f2
7 (p27*3.556 ph28):f2 ;(p27+4u)*115.132*l4 mixing
time
 (p27*4.556 ph29):f2
 (p27*3.222 ph28):f2
 (p27*3.167 ph29):f2
 (p27*0.333 ph28):f2
 (p27*2.722 ph29):f2
 (p27*4.167 ph28):f2
 (p27*2.944 ph29):f2
 (p27*4.111 ph28):f2
 (p27*3.556 ph29):f2
 (p27*4.556 ph28):f2
 (p27*3.222 ph29):f2
 (p27*3.167 ph28):f2
 (p27*0.333 ph29):f2
 (p27*2.722 ph28):f2
 (p27*4.167 ph29):f2
 (p27*2.944 ph28):f2
 (p27*4.111 ph29):f2
 (p27*3.556 ph29):f2
 (p27*4.556 ph28):f2
 (p27*3.222 ph29):f2
 (p27*3.167 ph28):f2
 (p27*0.333 ph29):f2
 (p27*2.722 ph28):f2
 (p27*4.167 ph29):f2
 (p27*2.944 ph28):f2
 (p27*4.111 ph29):f2
 (p27*3.556 ph28):f2
 (p27*4.556 ph29):f2
 (p27*3.222 ph28):f2
 (p27*3.167 ph29):f2
 (p27*0.333 ph28):f2
 (p27*2.722 ph29):f2
 (p27*4.167 ph28):f2
 (p27*2.944 ph29):f2
 (p27*4.111 ph28):f2
 lo to 7 times l4
 4u
 d10 pl2:f2
 (p4 ph4):f2
 d30
 4u
 (p3 ph5):f2
 d18
 (p4 ph6):f2
 d7
 (p1 ph9):f1
 4u
 (p1 ph12):f1
 d2
 (p2 ph0):f1 (d25 p3 ph7):f2
 4u BLKGRAD
 d2 pl12:f2 pl16:f3
 go=2 ph31 cpd2:f2 cpd3:f3
 d11 do:f2 do:f3 mc #0 to 2
   F1PH(ip21, id0)
   F2PH(rd0 & ip4 & ip5 & ip6 & ip7, id10 & dd30)
 exit

ph0 = 0
ph1 = 1
ph2 = 0 2
ph4 = 0
ph5 = 0 0 0 0 2 2 2 2 2
ph6 = 0
ph7 = 0
ph9 = (360) 135
ph12= (360) 0
ph21=0 0 2 2
ph28= 0
ph29= 2
ph31= 0 2 2 0

;d19 : gradient recovery [200u]
;d22 : CC-Constant time [8.3ms]
;d23 : 1/(4JCC) [6.25m]
;d24 : fuer DNA = 0.75m und fuer RNA = 1.5m (CH or
CH2 Group)
;d2 : 1/2JCH (3m)
;d8 : tocsy mixing time (C1'/C4' = 9.21m) (C1'/C5' =
13.5m)
;for z-only gradients:
;gpz1: 80%
;gpz2: 25%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
```

7.3.4 2D Γ-HCCH reference/ cross experiment

```
;gamhcch2d

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p4=p3*2"

"d0=4u"
"d4=1.6m"
"d11=30m"
"d13=4u"
"d23=25m"
"d31=d9/4"
"d10=d22/2"

"DELTA=d4-p19-d16-10u"
"DELTA1=d22/2-d0-p2"
"DELTA2=d25/2-p19-d16-10u"
"d27=d23/4+d31"
"d28=d23/4-p19-d16-10u-d31"
"d29=d23/4-p19-d16-10u+d31"
"d30=d23/4-d31"

"in0=inf1/2" ;eingefuegt rs
"in10=in0"

1 ze
 d11 pl12:f2
2 d11 do:f2
3 d12 pl9:f1
 d1 cw:f1
 d13 do:f1
 d12 pl1:f1
 (p1 ph11)
 DELTA UNBLKGRAD
 p19:gp1
 d16 pl2:f2
 10u
 (center (p2 ph10):f1 (p4 ph13):f2 )
 10u
 p19:gp1
 d16
 DELTA
 (p1 ph12)
 4u
 p16:gp2
 d16
 (p3 ph3):f2
 d0
 (p2 ph1)
 DELTA1
 (p4 ph1):f2
 d10
 (p3 ph5):f2        ; relaxation
 d27
 (p2 ph1)
 d28
 p19:gp3
 d16
 10u
 (p4 ph1):f2
 10u
 p19:gp3
 d16
 d29
 (p2 ph11)
 d30
 (p3 ph6):f2        ; relaxation end
 DELTA2
 p19:gp4
 d16
 10u
 (p4 ph1):f2
 10u
 p19:gp4
 d16
 DELTA2
 (p3 ph4):f2
 4u
 (p1 ph2):f1
 DELTA
 p19:gp5
 d16
 10u
 (center (p2 ph1):f1 (p4 ph1):f2 )
 10u
 p19:gp5
 d16
 DELTA pl12:f2
 4u BLKGRAD
 go=2 ph31 cpd2:f2
 d11 do:f2 mc #0 to 2
        F1PH(ip3 & ip13, id0 & dd10)
exit

ph1=0
ph10=0
ph11=0
ph2=1
ph12=1
```

ph3=0 2 ;d9 ref 3.36m, cross 0s
ph4=0 0 2 2
ph5=0 0 0 0 2 2 2 2
ph6=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph8=0
ph13=0
ph31=0 2 2 0 2 0 0 2 2 0 0 2 0 2 2 0

;d22 [6.25m]

7.4 Coupling constants and cross-correlated relaxation rates

7.4.1 ^3J(H,H) coupling constants

Table 21 ^3J(H,H) coupling constants for the 14 nts RNA with UUCG, GAAG, GCAA and CUUG tetraloops. The coupling constants were determined from 3D HCC-TOCSY-CCH-E.COSY spectra. Spectra were measured at 298 K for the UUCG and GAAG tetraloops and at 308 K for the GCAA and CUUG tetraloops. The error was determined from the fitting procedure. Asterisks mark coupling constants, where an error could not be determined. Here, the error was set to the maximum achievable digital resolution.

residue	cUUCGg			cGAAGg			cGCAAg		gCUUGc		
	$^3J(H_{1'},H_{2'})$ [Hz]	$^3J(H_{2'},H_{3'})$ [Hz]	$^3J(H_{3'},H_{4'})$ [Hz]	$^3J(H_{1'},H_{2'})$ [Hz]	$^3J(H_{2'},H_{3'})$ [Hz]	$^3J(H_{3'},H_{4'})$ [Hz]	$^3J(H_{1'},H_{2'})$ [Hz]	$^3J(H_{3'},H_{4'})$ [Hz]	$^3J(H_{1'},H_{2'})$ [Hz]	$^3J(H_{2'},H_{3'})$ [Hz]	$^3J(H_{3'},H_{4'})$ [Hz]
G1				1.2 ± 0.6			1.0 ± 0.2		2.6 ± 0.4*	7.0 ± 0.7	
G2	0.7	5.6	9.9	0.6 ± 0.6*	5.4 ± 4.8	11.4 ± 0.6*	0.4 ± 0.2*	10.7 ± 0.2	0.7 ± 0.4*	3.7 ± 2.2	
C3	0.4		11.5	0.6 ± 0.6*	4.2 ± 0.6	10.8 ± 0.6*	0.6 ± 0.3	10.3 ± 0.3	0.4 ± 0.4*	1.8 ± 1.8	9.9 ± 0.4*
A4			11.2	0.6 ± 0.6*	3.6 ± 0.6*	9.6 ± 0.6	0.6 ± 0.2*	9.5 ± 0.2	0.4 ± 0.4*	4.0 ± 0.4*	11.0 ± 0.4
5	0.4	3.3	10.2	1.2 ± 0.6*	7.2 ± 0.6	9.6 ± 0.6	0.3 ± 0.2*	8.8 ± 0.3	0.7 ± 0.4*	5.1 ± 0.4	8.8 ± 0.4
6	0.2	5.5	11.0	3.0 ± 0.6*	6.0 ± 3.0	5.4 ± 0.6	1.9 ± 0.2*	5.7 ± 0.2	7.3± 0.4*	4.4 ± 0.4*	1.8 ± 0.4*
7	8.0	5.2	1.0	6.0 ± 0.6*		2.4 ± 0.6*	3.7 ± 0.3	3.5 ± 0.2	7.3 ± 0.4	4.0 ± 0.4	0.7 ± 0.4*
8	9.3	4.9	0.6	3.6 ± 0.6	5.4 ± 0.6	4.8 ± 0.6	3.5 ± 0.2	4.8 ± 1.0	7.7 ± 0.4*	4.0 ± 0.4*	0.7 ± 0.4*
9	0.6	3.4	7.6	3.0 ± 0.6*	4.2 ± 0.6*	6.6 ± 1.2	3.0 ± 0.3	6.8 ± 0.3	8.8 ± 0.4	3.3 ± 0.4*	1.1 ± 0.4
G10	0.6		4.9	1.2 ± 0.6			0.6 ± 0.3		0.7 ± 0.4*	5.1 ± 2.7	7.3 ± 0.4*
U11	0.0		11.3	0.0 ± 0.6	3.6 ± 0.6	10.8 ± 0.6*	0.6 ± 0.2	10.1 ± 0.2	1.1 ± 0.4*	2.9 ± 0.4	9.9 ± 0.4
G12	1.2		11.5	0.6 ± 0.6*	1.8 ± 0.6	10.2 ± 0.6*	0.6 ± 0.2	10.6 ± 0.3	0.4 ± 0.4*	10.3 ± 0.4*	
C13	0.8	5.8	9.1	1.2 ± 1.2	4.8 ± 1.8	10.8 ± 0.6*	0.3 ± 0.2*	10.9 ± 0.9	0.4 ± 0.4*	4.0 ± 0.4*	10.3 ± 0.4
C14	0.3	5.5	5.3	1.8 ± 0.6*	6.0 ± 0.6*	6.6 ± 0.6*	2.2 ± 0.4	5.3 ± 0.2	2.2 ± 0.4	5.5 ± 0.4*	5.9 ± 0.4*

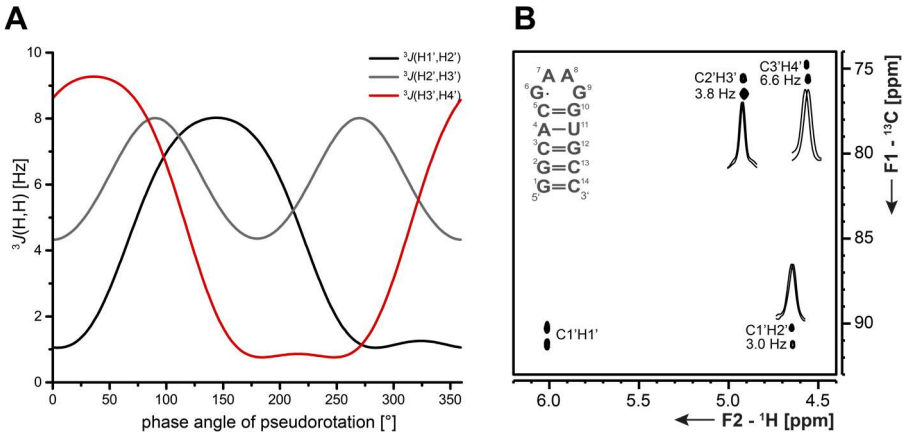


Figure 67 A Karplus curves for the ^3J(H1',H2'), ^3J(H2',H3') and ^3J(H3',H4') coupling constants, which were calculated as described in literature[216]. **B** The G9 2D plane of a 3D forward directed HCC-TOCSY-CCH-E.COSY spectrum. The experiment was recorded with a spectral width of 14 ppm in the direct ^{1}H-dimension, 35 ppm in the indirect ^{13}C-dimension (F2) and 3.3 ppm in the indirect ^{1}H-dimension. Carrier frequencies were 4.7 ppm, 76.5 ppm and 0.2 ppm for ^{1}H, ^{13}C and ^{31}P, respectively. Acquisition times were 0.10 s in the direct and 0.007 s (F2) and 0.022 s (F1) in the indirect dimensions. With 16 scans per increment and an inter-scan delay of 1 s, the experiment was recorded within 1 d 19 h.

E.COSY python script

```
from de.bruker.nmr.an.dcon1d import DconUtil
from de.bruker.nmr.jutil.misc import UtilDiscFunc
from de.bruker.nmr.jview.data import NMRData
from de.bruker.nmr.mfw.root import *
from math import sqrt
import os
low1, high1 = str(sys.argv[1]).split("-")
low2, high2 = str(sys.argv[2]).split("-")
XCMD("dpl") # Get limits from currently visible region
f1p = float(GETPAR("2 F1P"))
f2p = float(GETPAR("2 F2P"))
hz = float(GETPAR("2s HZpPT"))
#print hz
oldDataset = CURDATA()
range1 = CURDATA()
range1[2] = "1"+low1
#RSR(str(sys.argv[1]),str(sys.argv[1]),"y")
XCMD("f2sum %s %s %s" %(low1, high1, "1"+low1))
RE(range1)
procdata1 = list(GETPROCDATA(f1p,f2p))
#print procdata1
#print len(procdata1)
CLOSEWIN(CURDATA())
RE(oldDataset)
oldDataset = CURDATA()
range2 = CURDATA()
range2[2] = "1"+low2
#RSR(str(sys.argv[2]),str(sys.argv[2]),"y")
XCMD("f2sum %s %s %s" %(low2, high2, "1"+low2))
RE(range2)
procdata2 = list(GETPROCDATA(f1p,f2p))
#print procdata2
#print len(procdata2)
CLOSEWIN(CURDATA())
RE(oldDataset)
diffs = []
shifts = []
for x in range(len(procdata1)):
        summe = sum(abs(procdata1[i] -
procdata2[i]) for i in range(len(procdata1)))
        diffs.append(summe)
        tmp = procdata1.pop(0)
        procdata1.append(tmp)
for x in range(len(procdata1)):
        shift = x * hz
        shifts.append(shift)
#print diffs
#print shifts
disprop = GET_DISPLAY_PROPS(xStart = str(0), xEnd =
str(max(shifts)), xUnit = 'Hz', xLegend='shift',
yLegend='difference')
DISPLAY_DATALIST_XY([diffs,], [shifts,], propList =
[disprop,], winTitle = "Dreck")
#print disprop
minimum_diffs = min(diffs)
for y in range(len(diffs)):
        if diffs[y] == minimum_diffs:
                minimum = shifts[y]
devis = []
for z in range(len(diffs)):
        if diffs[z] <= 1.02*minimum_diffs:
                devis.append(shifts[z])
```

```
if minimum-min(devis) >= max(devis)-minimum:
        devi=minimum-min(devis)
else:
        devi=max(devis)-minimum
result="ECOSY analysis of rows "+str(sys.argv[1])+" and "+str(sys.argv[2])+"\n\nCoupling constant: "+str(round(minimum, 3))+" +- "+str(round(devi, 3))+" Hz."
VIEWTEXT(title="Result", header="", text=result, modal=1)
```

7.4.2 $\Gamma^{DD,DD}_{CH,CH}$ cross-correlated relaxation rates

Table 22 $\mathbf{\Gamma^{DD,DD}_{CH,CH}}$ cross-correlated relaxation rates for the 14 nts RNAs with UUCG, GAAG and CUUG tetraloops. The rates were determined from 2D Γ-HCCH experiments[213] at 298 K (UUCG, GAAG) and 308 K (CUUG). The error represents the standard deviation. Asterisks mark rates, where a mean value and the standard deviation could not be determined and an error of 10% was assumed.

residue	cUUCGg		cGAAGg		gCUUGc	
	$\mathbf{\Gamma^{DD,DD}_{C1'H1',C2'H2'}}$ **[Hz]**	$\mathbf{\Gamma^{DD,DD}_{C3'H3',C4'H4'}}$ **[Hz]**	$\mathbf{\Gamma^{DD,DD}_{C1'H1',C2'H2'}}$ **[Hz]**	$\mathbf{\Gamma^{DD,DD}_{C3'H3',C4'H4'}}$ **[Hz]**	$\mathbf{\Gamma^{DD,DD}_{C1'H1',C2'H2'}}$ **[Hz]**	$\mathbf{\Gamma^{DD,DD}_{C3'H3',C4'H4'}}$ **[Hz]**
G1			-6.0 ± 0.6*			12.0 ± 1.2*
G2	-10.1 ± 0.6	13.1 ± 0.1	-10.8 ± 0.2		-8.6 ± 0.3	
C3	-8.7 ± 0.4	15.3 ± 0.8	-9.1 ± 0.2		-5.9 ± 4.6	
A4	-9.0 ± 0.4	13.5 ± 0.7	-10.6 ± 0.1		-6.8 ± 0.9	4.3 ± 0.5*
5	-6.8 ± 0.3		-7.1 ± 0.3		-6.4 ± 0.5	13.3 ± 1.4*
6	-5.0 ± 0.3	14.5 ± 0.7	11.2 ± 3.3		8.6 ± 0.9*	
7	15.8 ± 0.8	-4.5 ± 0.2	5.6 ± 4.6		13.0 ± 0.7	-4.1 ± 1.4
8	17.3 ± 0.2	-7.0 ± 0.4			12.2 ± 0.7	-8.2 ± 0.9*
9	-8.3 ± 1.9	12.6 ± 0.6	-4.7 ± 0.5*		6.5 ± 1.1	-4.9 ± 0.5*
10		20.7 ± 1.0			-10.6 ± 1.1*	10.3 ± 4.2
U11	-8.6 ± 0.6	12.0 ± 0.6	-8.9 ± 1.3		-5.9 ± 0.2	3.9 ± 0.4*
G12	-9.4 ± 0.2		-10.9 ± 0.3		-7.6 ± 0.7	
C13	-6.9 ± 0.3		-7.8 ± 0.4		-6.8 ± 0.7*	13.5 ± 1.4*
C14	-4.7 ± 0.1		-1.4 ± 0.3		-6.0 ± 0.6*	7.9 ± 0.8*

7.4.3 ^{1}J(C,H) coupling constants

Table 23 ^{1}J(C,H) coupling constants for the 14 nts RNA with GAAG tetraloop. Fields marked with an asterisk correspond to residues where the respective coupling constant could not be determined due to signal overlap. Coupling constants were measured from coupled ^{1}H,^{13}C-HSQC spectra at 298 K and 600 MHz.

	^{1}J(C6/8,H6/8) [Hz]	**^{1}J(C1',H1') [Hz]**	**^{1}J(C2',H2') [Hz]**	**^{1}J(C3',H3') [Hz]**	**^{1}J(C4',H4') [Hz]**	**^{1}J(C5',H5'/H5'') [Hz]**
G1	226.3	177.1	159.4	148.6	149.5	156.4/148.7
G2	223.3	177.1	160.0	148.3	145.2	160.9/-*
C3	182.5	180.7	-*	145.5	147.7	158.1/146.9
A4	218.5	175.9	141.0	144.7	144.0	160.0/147.1
C5	183.7	179.5	141.8	137.3	144.0	157.9/-*
G6	221.5	172.3	-*	-*	147.0	-*/148.2
A7	223.3	169.9	-*	-*	150.6	-*/146.4
A8	225.1	171.7	160.5	-*	152.4	151.5/155.9
G9	227.5	170.5	156.6	155.2	150.0	154.9/148.7
G10	225.7	175.9	157.3	148.8	153.0	147.0
U11	184.3	181.3	161.5	147.3	147.0	157.2/147.6
G12	221.5	176.5	156.0	145.5	145.8	156.4/-*
C13	178.9	180.1	160.3	145.2	147.0	156.8/-*
C14	181.3	175.3	167.5	169.1	148.2	157.4/150.4

7.4.4 ^{3}J(C,P) and ^{3}J(H,P) coupling constants

Table 24 ^{3}J(C,P) coupling constants for the 14 nts RNAs with UUCG, GAAG and CUUG tetraloops, which were determined using the quantitative *J*-HCP- experiments. The coupling constants for the UUCG tetraloops were taken from literature[106] and were determined at 298 K. The coupling constants for the GAAG RNA were determined at 298 K and 700 MHz. The coupling constants for the CUUG tetraloop were determined by Andreas Oxenfarth at 308 K and 600 MHz.

residue	**cUUCGg**			**cGAAGg**			**gCUUGc**		
	$^{3}J(C2'_i,P_{i+1})$ [Hz]	**$^{3}J(C4'_i,P_{i+1})$ [Hz]**	**$^{3}J(C4'_i,P_i)$ [Hz]**	**$^{3}J(C2'_i,P_{i+1})$ [Hz]**	**$^{3}J(C4'_i,P_{i+1})$ [Hz]**	**$^{3}J(C4'_i,P_i)$ [Hz]**	**$^{3}J(C2'_i,P_{i+1})$ [Hz]**	**$^{3}J(C4'_i,P_{i+1})$ [Hz]**	**$^{3}J(C4'_i,P_i)$ [Hz]**
G1		6.0		2.2	8.7				7.0
G2			7.8	1.0	10.1	9.7		9.0	8.5
C3	0.8			0.8		11.1		9.8	9.4
A4				0.8	8.6			6.7	6.6
5	0.9			0.9	8.9	10.3		9.9	5.8
6	0.9	9.5	9.6	1.5	6.4	9.6		7.5	7.9
7	3.5		6.3	3.6	3.6	7.3	2.3	6.8	3.8
8	6.1	7.1	9.3	2.3	5.4	7.8	4.9	8.3	1.6
9	2.2	7.7	8.0	1.9	6.2	9.3	1.7	5.8	5.2
G10	0.9		6.1	1.0	8.7	.3	5.1	3.4	8.2
U11				0.8	10.4	6.7		8.4	9.3
G12	0.7	6.5		0.7	8.3	11.9		10.6	11.5
C13	0.6			0.4	10.9	11.2		7.5	9.9
C14			7.7	0.9	1.1	9.8		9.5	

Table 25 ^{3}J(H,P) coupling constants for the 14 nts RNAs with UUCG and GAAG tetraloops, which were determined at 298 K. The coupling constants for the 14 nts RNA with UUCG tetraloop were taken from literature[106]. The coupling constants for the 14 nts RNA with GAAG tetraloop were determined using P-FIDS-HSQC experiments[206].

residue	cUUCGg			cGAAGg		
	$^{3}J(H_{3'i},P_{i+1})$ [Hz]	$^{3}J(H_{5'i},P_i)$ [Hz]	$^{3}J(H_{5''i},P_i)$ [Hz]	$^{3}J(H_{3'i},P_{i+1})$ [Hz]	$^{3}J(H_{5'i},P_i)$ [Hz]	$^{3}J(H_{5''i},P_i)$ [Hz]
G1		6.9	5.7	7.8	3.9	3.3
G2		2.1	2.0	9.0	2.8	3.1
C3		0.5	1.7	7.6	3.2	1.9
A4	3.8			9.4	3.0	3.1
5	9.4	1.8	3.0	8.2	4.6	3.0
6	7.3	1.7	2.9	7.6	3.9	1.5
7	3.7	3.7	5.7		0.4	2.9
8	7.5	2.3	1.3	8.1	4.6	5.1
9	7.5	9.0	8.6	8.1	4.7	1.8
G10		2.3	1.3	7.5		
U11	8.0	1.7		8.2	4.0	0.5
G12	7.8			7.9	2.6	3.0
C13	9.3	1.7	1.7	8.0	4.4	0.2
C14		1.8	3.1	11.3	3.4	1.2

7.5 The GAAG structure

7.5.1 Base pair planarity and hydrogen bond restraints

```
! Planarity restraints to maintain RNA base pair planarity
!
!
! G1-C14 WC
group
selection = ((resid 1 and name N1) or (resid  1 and name N3) or (resid 1 and name C5)
      or (resid 14 and name N1) or (resid 14 and name N3) or (resid 14 and name C5))
weight = 10 end
!
! G2-C13 WC
group
selection = ((resid  2 and name N1) or (resid  2 and name N3) or (resid  2 and name C5)
      or (resid 13 and name N1) or (resid 13 and name N3) or (resid 13 and name C5))
weight = 25 end
!
! C3-G12 WC
group
selection = ((resid  3 and name N1) or (resid  3 and name N3) or (resid  3 and name C5)
      or (resid 12 and name N1) or (resid 12 and name N3) or (resid 12 and name C5))
weight = 25 end
!
! A4-U11 WC
group
selection = ((resid  4 and name N1) or (resid  4 and name N3) or (resid  4 and name C5)
      or (resid 11 and name N1) or (resid 11 and name N3) or (resid 11 and name C5))
weight = 25 end
!
```

```
! C5-G10 WC
group
selection = ((resid  5 and name N1) or (resid  5 and name N3) or (resid  5 and name C5)
      or (resid 10 and name N1) or (resid 10 and name N3) or (resid 10 and name C5))
weight = 10 end

! Hydrogen bonds
!
assign ( resid 14 and name N3 )( resid 1 and name N1 ) 2.90 0.6 0.5
assign ( resid 14 and name N3 )( resid 1 and name H1 ) 1.90 0.5 0.5
assign ( resid 14 and name N4 )( resid 1 and name O6 ) 2.90 0.6 0.5
assign ( resid 14 and name H42)( resid 1 and name O6 ) 1.90 0.5 0.5
assign ( resid 14 and name O2 )( resid 1 and name N2 ) 2.90 0.6 0.5
assign ( resid 14 and name O2 )( resid 1 and name H22) 1.90 0.5 0.5
!
assign ( resid 13 and name N3 )( resid 2 and name N1 ) 2.90 0.5 0.4
assign ( resid 13 and name N3 )( resid 2 and name H1 ) 1.90 0.4 0.4
assign ( resid 13 and name N4 )( resid 2 and name O6 ) 2.90 0.5 0.4
assign ( resid 13 and name H42)( resid 2 and name O6 ) 1.90 0.4 0.4
assign ( resid 13 and name O2 )( resid 2 and name N2 ) 2.90 0.5 0.4
assign ( resid 13 and name O2 )( resid 2 and name H22) 1.90 0.4 0.4
!
assign ( resid 3 and name N3 )( resid 12 and name N1 ) 2.90 0.5 0.4
assign ( resid 3 and name N3 )( resid 12 and name H1 ) 1.90 0.4 0.4
assign ( resid 3 and name N4 )( resid 12 and name O6 ) 2.90 0.5 0.4
assign ( resid 3 and name H42)( resid 12 and name O6 ) 1.90 0.4 0.4
assign ( resid 3 and name O2 )( resid 12 and name N2 ) 2.90 0.5 0.4
assign ( resid 3 and name O2 )( resid 12 and name H22) 1.90 0.4 0.4
!
assign ( resid 4 and name N1 )( resid 11 and name N3 ) 2.90 0.5 0.4
assign ( resid 4 and name N1 )( resid 11 and name H3 ) 1.90 0.4 0.4
assign ( resid 4 and name N6 )( resid 11 and name O4 ) 2.90 0.5 0.4
assign ( resid 4 and name H62)( resid 11 and name O4 ) 1.90 0.4 0.4
!
assign ( resid 5 and name N3 )( resid 10 and name N1 ) 2.90 0.6 0.5
assign ( resid 5 and name N3 )( resid 10 and name H1 ) 1.90 0.5 0.5
assign ( resid 5 and name N4 )( resid 10 and name O6 ) 2.90 0.6 0.5
assign ( resid 5 and name H42)( resid 10 and name O6 ) 1.90 0.5 0.5
assign ( resid 5 and name O2 )( resid 10 and name N2 ) 2.90 0.6 0.5
assign ( resid 5 and name O2 )( resid 10 and name H22) 1.90 0.5 0.5
```

7.5.2 Dihedral restraints

```
! Torsion angle restraints
!
! nu0  (broad definition based on CCR and 3J coupling data 3 ± 22 or -21 ± 22)
                    MIDDLE OUT FOR NU0-NU4
assign (resid  1 and name C4') (resid  1 and name O4') (resid  1 and name C1') (resid  1 and name C2') 1   3 22 2
assign (resid  2 and name C4') (resid  2 and name O4') (resid  2 and name C1') (resid  2 and name C2') 1   3 22 2
assign (resid  3 and name C4') (resid  3 and name O4') (resid  3 and name C1') (resid  3 and name C2') 1   3 22 2
assign (resid  4 and name C4') (resid  4 and name O4') (resid  4 and name C1') (resid  4 and name C2') 1   3 22 2
assign (resid  5 and name C4') (resid  5 and name O4') (resid  5 and name C1') (resid  5 and name C2') 1   3 22 2
assign (resid 10 and name C4') (resid 10 and name O4') (resid 10 and name C1') (resid 10 and name C2') 1   3 22 2
assign (resid 11 and name C4') (resid 11 and name O4') (resid 11 and name C1') (resid 11 and name C2') 1   3 22 2
```

```
assign (resid 12 and name C4') (resid 12 and name O4') (resid 12 and name C1') (resid 12 and name C2') 1   3 22 2
assign (resid 13 and name C4') (resid 13 and name O4') (resid 13 and name C1') (resid 13 and name C2') 1   3 22 2
assign (resid 14 and name C4') (resid 14 and name O4') (resid 14 and name C1') (resid 14 and name C2') 1   3 22 2
!
! nu1  (broad definition based on CCR and 3J coupling data -25 ± 20 or 35 ± 16)
assign (resid  1 and name O4') (resid  1 and name C1') (resid  1 and name C2') (resid  1 and name C3') 1 -25 20 2
assign (resid  2 and name O4') (resid  2 and name C1') (resid  2 and name C2') (resid  2 and name C3') 1 -25 20 2
assign (resid  3 and name O4') (resid  3 and name C1') (resid  3 and name C2') (resid  3 and name C3') 1 -25 20 2
assign (resid  4 and name O4') (resid  4 and name C1') (resid  4 and name C2') (resid  4 and name C3') 1 -25 20 2
assign (resid  5 and name O4') (resid  5 and name C1') (resid  5 and name C2') (resid  5 and name C3') 1 -25 20 2
assign (resid 10 and name O4') (resid 10 and name C1') (resid 10 and name C2') (resid 10 and name C3') 1 -25 20 2
assign (resid 11 and name O4') (resid 11 and name C1') (resid 11 and name C2') (resid 11 and name C3') 1 -25 20 2
assign (resid 12 and name O4') (resid 12 and name C1') (resid 12 and name C2') (resid 12 and name C3') 1 -25 20 2
assign (resid 13 and name O4') (resid 13 and name C1') (resid 13 and name C2') (resid 13 and name C3') 1 -25 20 2
assign (resid 14 and name O4') (resid 14 and name C1') (resid 14 and name C2') (resid 14 and name C3') 1 -25 20 2
!
! nu2  (broad definition based on CCR and 3J coupling data 36 ± 14 or -35 ± 15)
assign (resid  1 and name C1') (resid  1 and name C2') (resid  1 and name C3') (resid  1 and name C4') 1  36 14 2
assign (resid  2 and name C1') (resid  2 and name C2') (resid  2 and name C3') (resid  2 and name C4') 1  36 14 2
assign (resid  3 and name C1') (resid  3 and name C2') (resid  3 and name C3') (resid  3 and name C4') 1  36 14 2
assign (resid  4 and name C1') (resid  4 and name C2') (resid  4 and name C3') (resid  4 and name C4') 1  36 14 2
assign (resid  5 and name C1') (resid  5 and name C2') (resid  5 and name C3') (resid  5 and name C4') 1  36 14 2
assign (resid 10 and name C1') (resid 10 and name C2') (resid 10 and name C3') (resid 10 and name C4') 1  36 14 2
assign (resid 11 and name C1') (resid 11 and name C2') (resid 11 and name C3') (resid 11 and name C4') 1  36 14 2
assign (resid 12 and name C1') (resid 12 and name C2') (resid 12 and name C3') (resid 12 and name C4') 1  36 14 2
assign (resid 13 and name C1') (resid 13 and name C2') (resid 13 and name C3') (resid 13 and name C4') 1  36 14 2
assign (resid 14 and name C1') (resid 14 and name C2') (resid 14 and name C3') (resid 14 and name C4') 1  36 14 2
!
! nu3  (broad definition based on CCR and 3J coupling data -35 ± 15 or 24 ± 23)
assign (resid  1 and name C2') (resid  1 and name C3') (resid  1 and name C4') (resid  1 and name O4') 1 -35 15 2
assign (resid  2 and name C2') (resid  2 and name C3') (resid  2 and name C4') (resid  2 and name O4') 1 -35 15 2
assign (resid  3 and name C2') (resid  3 and name C3') (resid  3 and name C4') (resid  3 and name O4') 1 -35 15 2
assign (resid  4 and name C2') (resid  4 and name C3') (resid  4 and name C4') (resid  4 and name O4') 1 -35 15 2
assign (resid  5 and name C2') (resid  5 and name C3') (resid  5 and name C4') (resid  5 and name O4') 1 -35 15 2
assign (resid 10 and name C2') (resid 10 and name C3') (resid 10 and name C4') (resid 10 and name O4') 1 -35 15 2
assign (resid 11 and name C2') (resid 11 and name C3') (resid 11 and name C4') (resid 11 and name O4') 1 -35 15 2
assign (resid 12 and name C2') (resid 12 and name C3') (resid 12 and name C4') (resid 12 and name O4') 1 -35 15 2
assign (resid 13 and name C2') (resid 13 and name C3') (resid 13 and name C4') (resid 13 and name O4') 1 -35 15 2
assign (resid 14 and name C2') (resid 14 and name C3') (resid 14 and name C4') (resid 14 and name O4') 1 -35 15 2
!
! nu4  (broad definition based on CCR and 3J coupling data 20 ± 23 or -2 ± 25)
assign (resid  1 and name C3') (resid  1 and name C4') (resid  1 and name O4') (resid  1 and name C1') 1  20 23 2
assign (resid  2 and name C3') (resid  2 and name C4') (resid  2 and name O4') (resid  2 and name C1') 1  20 23 2
assign (resid  3 and name C3') (resid  3 and name C4') (resid  3 and name O4') (resid  3 and name C1') 1  20 23 2
assign (resid  4 and name C3') (resid  4 and name C4') (resid  4 and name O4') (resid  4 and name C1') 1  20 23 2
assign (resid  5 and name C3') (resid  5 and name C4') (resid  5 and name O4') (resid  5 and name C1') 1  20 23 2
assign (resid 10 and name C3') (resid 10 and name C4') (resid 10 and name O4') (resid 10 and name C1') 1  20 23 2
assign (resid 11 and name C3') (resid 11 and name C4') (resid 11 and name O4') (resid 11 and name C1') 1  20 23 2
assign (resid 12 and name C3') (resid 12 and name C4') (resid 12 and name O4') (resid 12 and name C1') 1  20 23 2
assign (resid 13 and name C3') (resid 13 and name C4') (resid 13 and name O4') (resid 13 and name C1') 1  20 23 2
assign (resid 14 and name C3') (resid 14 and name C4') (resid 14 and name O4') (resid 14 and name C1') 1  20 23 2
!
! chi (from CCR)
                    ONLY STEM 200 +/- 40 (or 50)
assign (resid  2 and name O4') (resid  2 and name C1') (resid  2 and name N9) (resid  2 and name C4) 1 200 50 2
```

```
assign (resid  3 and name O4') (resid  3 and name C1') (resid  3 and name N1) (resid  3 and name C2) 1 200 50 2
assign (resid  4 and name O4') (resid  4 and name C1') (resid  4 and name N9) (resid  4 and name C4) 1 200 50 2
assign (resid  5 and name O4') (resid  5 and name C1') (resid  5 and name N1) (resid  5 and name C2) 1 200 50 2
assign (resid  7 and name O4') (resid  7 and name C1') (resid  7 and name N9) (resid  7 and name C4) 1 200 40 2
assign (resid  8 and name O4') (resid  8 and name C1') (resid  8 and name N9) (resid  8 and name C4) 1 200 40 2
assign (resid  9 and name O4') (resid  9 and name C1') (resid  9 and name N9) (resid  9 and name C4) 1 200 40 2
assign (resid 11 and name O4') (resid 11 and name C1') (resid 11 and name N1) (resid 11 and name C2) 1 200 50 2
assign (resid 12 and name O4') (resid 12 and name C1') (resid 12 and name N9) (resid 12 and name C4) 1 200 50 2
assign (resid 13 and name O4') (resid 13 and name C1') (resid 13 and name N1) (resid 13 and name C2) 1 200 50 2
assign (resid 14 and name O4') (resid 14 and name C1') (resid 14 and name N1) (resid 14 and name C2) 1 200 50 2
!
! beta (from 3J)
assign (resid  2 and name P) (resid  2 and name O5') (resid  2 and name C5') (resid  2 and name C4') 1  178 40 2
assign (resid  3 and name P) (resid  3 and name O5') (resid  3 and name C5') (resid  3 and name C4') 1  185 40 2
assign (resid  5 and name P) (resid  5 and name O5') (resid  5 and name C5') (resid  5 and name C4') 1  193 40 2
assign (resid  6 and name P) (resid  6 and name O5') (resid  6 and name C5') (resid  6 and name C4') 1  193 40 2
assign (resid  7 and name P) (resid  7 and name O5') (resid  7 and name C5') (resid  7 and name C4') 1  166 60 2
assign (resid  8 and name P) (resid  8 and name O5') (resid  8 and name C5') (resid  8 and name C4') 1  148 60 2
assign (resid  9 and name P) (resid  9 and name O5') (resid  9 and name C5') (resid  9 and name C4') 1  198 60 2
assign (resid 10 and name P) (resid 10 and name O5') (resid 10 and name C5') (resid 10 and name C4') 1  129 60 2
assign (resid 11 and name P) (resid 11 and name O5') (resid 11 and name C5') (resid 11 and name C4') 1  202 40 2
assign (resid 12 and name P) (resid 12 and name O5') (resid 12 and name C5') (resid 12 and name C4') 1  178 40 2
assign (resid 13 and name P) (resid 13 and name O5') (resid 13 and name C5') (resid 13 and name C4') 1  193 40 2
assign (resid 14 and name P) (resid 14 and name O5') (resid 14 and name C5') (resid 14 and name C4') 1  190 40 2
!
! epsilon (from 3J)
assign (resid  1 and name C4') (resid  1 and name C3') (resid  1 and name O3') (resid  2 and name P) 1 -148 40 2
assign (resid  2 and name C4') (resid  2 and name C3') (resid  2 and name O3') (resid  3 and name P) 1 -143 40 2
assign (resid  3 and name C4') (resid  3 and name C3') (resid  3 and name O3') (resid  4 and name P) 1 -151 40 2
assign (resid  4 and name C4') (resid  4 and name C3') (resid  4 and name O3') (resid  5 and name P) 1 -142 45 2
assign (resid  5 and name C4') (resid  5 and name C3') (resid  5 and name O3') (resid  6 and name P) 1 -148 40 2
assign (resid  6 and name C4') (resid  6 and name C3') (resid  6 and name O3') (resid  7 and name P) 1 -150 60 2
assign (resid  7 and name C4') (resid  7 and name C3') (resid  7 and name O3') (resid  8 and name P) 1  166 60 2
assign (resid  8 and name C4') (resid  8 and name C3') (resid  8 and name O3') (resid  9 and name P) 1 -146 60 2
assign (resid  9 and name C4') (resid  9 and name C3') (resid  9 and name O3') (resid 10 and name P) 1 -146 60 2
assign (resid 10 and name C4') (resid 10 and name C3') (resid 10 and name O3') (resid 11 and name P) 1 -151 40 2
assign (resid 11 and name C4') (resid 11 and name C3') (resid 11 and name O3') (resid 12 and name P) 1 -148 40 2
assign (resid 12 and name C4') (resid 12 and name C3') (resid 12 and name O3') (resid 13 and name P) 1 -150 40 2
assign (resid 13 and name C4') (resid 13 and name C3') (resid 13 and name O3') (resid 14 and name P) 1 -149 40 2
!
! alpha (from CCR)
assign (resid  2 and name O3') (resid  3 and name P) (resid  3 and name O5') (resid  3 and name C5') 1 295 35 2
assign (resid  4 and name O3') (resid  5 and name P) (resid  5 and name O5') (resid  5 and name C5') 1 295 35 2
assign (resid 10 and name O3') (resid 11 and name P) (resid 11 and name O5') (resid 11 and name C5') 1 295 35 2
assign (resid 11 and name O3') (resid 12 and name P) (resid 12 and name O5') (resid 12 and name C5') 1 295 35 2
assign (resid 12 and name O3') (resid 13 and name P) (resid 13 and name O5') (resid 13 and name C5') 1 295 35 2
assign (resid 13 and name O3') (resid 14 and name P) (resid 14 and name O5') (resid 14 and name C5') 1 295 35 2
!
! zeta (from CCR)
assign (resid  3 and name C3') (resid  3 and name O3') (resid  4 and name P) (resid  4 and name O5') 1  290 50 2
assign (resid  4 and name C3') (resid  4 and name O3') (resid  5 and name P) (resid  5 and name O5') 1  290 50 2
assign (resid 11 and name C3') (resid 11 and name O3') (resid 12 and name P) (resid 12 and name O5') 1  290 50 2
assign (resid 12 and name C3') (resid 12 and name O3') (resid 13 and name P) (resid 13 and name O5') 1  290 50 2
assign (resid 13 and name C3') (resid 13 and name O3') (resid 14 and name P) (resid 14 and name O5') 1  290 50 2
!
```

! gamma (qualitative from J)
assign (resid 1 and name O5') (resid 1 and name C5') (resid 1 and name C4') (resid 1 and name C3') 1 60 30 2
assign (resid 2 and name O5') (resid 2 and name C5') (resid 2 and name C4') (resid 2 and name C3') 1 60 30 2
assign (resid 3 and name O5') (resid 3 and name C5') (resid 3 and name C4') (resid 3 and name C3') 1 60 30 2
assign (resid 4 and name O5') (resid 4 and name C5') (resid 4 and name C4') (resid 4 and name C3') 1 60 30 2
assign (resid 5 and name O5') (resid 5 and name C5') (resid 5 and name C4') (resid 5 and name C3') 1 60 30 2
assign (resid 6 and name O5') (resid 6 and name C5') (resid 6 and name C4') (resid 6 and name C3') 1 60 30 2
assign (resid 7 and name O5') (resid 7 and name C5') (resid 7 and name C4') (resid 7 and name C3') 1 60 30 2
assign (resid 8 and name O5') (resid 8 and name C5') (resid 8 and name C4') (resid 8 and name C3') 1 60 30 2
assign (resid 9 and name O5') (resid 9 and name C5') (resid 9 and name C4') (resid 9 and name C3') 1 60 30 2
assign (resid 11 and name O5') (resid 11 and name C5') (resid 11 and name C4') (resid 11 and name C3') 1 60 30 2
assign (resid 12 and name O5') (resid 12 and name C5') (resid 12 and name C4') (resid 12 and name C3') 1 60 30 2
assign (resid 13 and name O5') (resid 13 and name C5') (resid 13 and name C4') (resid 13 and name C3') 1 60 30 2
assign (resid 14 and name O5') (resid 14 and name C5') (resid 14 and name C4') (resid 14 and name C3') 1 60 30 2

7.5.3 NOESY peak lists

NOESY 50 ms 100% D_2O

Assignment	w1	w2	Data Height
?-?	4.587	7.776	3168376
?-?	4.173	4.634	13320453
?-?	4.034	4.586	1548118
?-?	4.302	4.769	1274019
?-?	7.78	8.077	127939
?-?	4.176	7.778	648497
?-?	4.531	7.775	548528
?-?	4.625	7.778	1838357
?-?	4.662	4.769	4289302
?-?	4.468	4.943	1004695
?-?	4.489	7.776	322278
?-G1H1'	4.581	5.866	176523
?-G1H2'	4.584	4.987	417026
?-G2H1'	4.645	5.957	191814
?-G2H1'	4.584	5.957	1526277
?-G2H8	4.586	7.664	371824
?-C3H1'	4.175	5.565	62361
?-C3H1'	4.652	5.565	64038
?-C3H1'	4.618	5.565	220389
?-C3H1'	4.491	5.565	420553
?-C3H5	4.586	5.338	241357
?-C3H5	4.648	5.338	177854
?-A4H1'	4.521	5.991	1709448
?-A4H8	4.526	8.08	2649933
?-A4H8	4.614	8.081	558012
?-C5H1'	4.52	5.427	225736
?-C5H1'	4.365	5.427	422741
?-C5H5	4.521	5.144	105148
?-C5H6	4.52	7.3	1567452
?-G6H1'	4.033	5.719	68148
?-G6H8	4.367	7.559	183059
?-A7H1'	4.034	5.693	391529
?-A7H1'	4.587	5.692	1305818
?-A7H2	4.587	7.926	65456
?-A7H8	4.029	8.287	188199
?-A7H8	4.589	8.287	1470504
?-A8H1'	4.397	5.807	21513
?-A8H1'	4.589	5.808	200229
?-A8H1'	4.944	5.808	76346
?-A8H2	4.587	7.872	102074
?-A8H2	4.468	7.873	52505
?-A8H3'	4.415	4.77	865478
?-A8H8	4.585	8.096	1169201
?-A8H8	4.034	8.097	92848
?-A8H8	3.944	8.096	88377
?-G9H1'	4.938	6.033	162214
?-G9H3'	4.679	4.939	891988
?-G10H8	4.941	8.073	531537
?-G10H8	4.587	8.072	431554
?-G10H8	4.382	8.072	543039
?-U11H1'	4.536	5.601	112226
?-U11H1'	4.475	5.602	505528
?-U11H6	4.335	7.821	105304
?-U11H6	4.467	7.823	1948198
?-G12H1'	4.554	5.852	1013048
?-C13H1'	4.545	5.498	186205
?-C14H5	4.449	5.577	90580
G1H1'-G1H8	5.866	8.196	223759
G1H1'-G2H1'	5.863	5.956	391782
G1H1'-G2H8	5.868	7.663	358965
G1H2'-G1H1'	4.989	5.864	1449723
G1H2'-G1H8	4.99	8.199	673507
G1H2'-G2H1'	4.988	5.958	186140
G1H2'-G2H8	4.989	7.663	1936046
G1H3'-G1H1'	4.769	5.866	201907
G1H3'-G1H2'	4.77	4.987	2132162
G1H3'-G1H8	4.772	8.2	278449
G1H3'-G2H8	4.769	7.663	353544
G1H4'-G1H1'	4.609	5.865	592986
G1H4'-G1H2'	4.606	4.988	405127
G1H4'-G1H8	4.61	8.196	97586
G1H5'1-G1H1'	4.482	5.866	72074
G1H5'1-G1H2'	4.484	4.987	99049
G1H5'1-G1H3'	4.491	4.769	1752012
G1H5'1-G1H8	4.484	8.203	136304
G1H5'2-G1H1'	4.336	5.868	78890
G1H5'2-G1H2'	4.338	4.989	112479
G1H5'2-G1H3'	4.336	4.771	1233610
G1H5'2-G1H5'1	4.333	4.489	4552521
G1H5'2-G1H8	4.33	8.201	444382
G2H1'-G2H8	5.957	7.664	124378
G2H1'-C3H6	5.958	7.776	58442
G2H2'-C3H1'	4.585	5.566	229600
G2H3'-G1H2'	4.642	4.987	198332
G2H3'-G2H8	4.644	7.664	430879
G2H5'2-G1H2'	4.302	4.987	200623

G2H5'2-G2H1'	4.297	5.951	59423
G2H5'2-G2H5'1	4.303	4.589	9800865
G2H5'2-G2H8	4.302	7.663	411790
G2H5'2-C3H5	4.302	5.331	42367
G2H8-G1H8	7.66	8.205	50792
C3H1'-C3H6	5.564	7.774	166832
C3H1'-A4H8	5.567	8.079	78670
C3H2'-C3H1'	4.526	5.565	1255932
C3H2'-C3H5	4.53	5.336	37048
C3H3'-C3H5	4.619	5.339	95037
C3H5-C3H1'	5.337	5.562	83941
C3H5-C3H6	5.339	7.776	3015545
A4H1'-A4H2	5.981	7.428	43043
A4H1'-A4H8	5.991	8.08	145276
A4H2'-A4H2	4.527	7.431	76703
A4H2'-A4H3'	4.521	4.705	1636732
A4H3'-A4H1'	4.708	5.991	102492
A4H3'-A4H8	4.709	8.08	734509
A4H3'-C5H5	4.709	5.143	66978
A4H3'-C5H6	4.707	7.299	269531
A4H5'1-A4H1'	4.619	5.991	83692
A4H5'2-A4H1'	4.182	5.996	52237
A4H5'2-A4H3'	4.185	4.695	853711
A4H5'2-A4H8	4.185	8.081	374857
C5H1'-A4H2	5.427	7.432	263658
C5H1'-C5H6	5.428	7.3	140599
C5H2'-A4H2	4.285	7.432	8075
C5H2'-C5H1'	4.286	5.427	1193519
C5H2'-C5H6	4.291	7.299	331210
C5H2'-G6H1'	4.286	5.72	134621
C5H2'-G6H8	4.287	7.56	819558
C5H3'-C5H1'	4.333	5.427	216403
C5H3'-C5H5	4.332	5.142	75308
C5H3'-C5H6	4.329	7.301	832494
C5H3'-G6H8	4.329	7.559	351112
C5H4'-A4H1'	4.37	5.993	25982
C5H4'-C5H5	4.364	5.142	57610
C5H4'-C5H6	4.358	7.3	458347
C5H5'2-C5H5'1	4.082	4.514	13821472
C5H5'2-C5H6	4.082	7.299	257375
C5H6-A4H2	7.305	7.433	74685
C5H6-A4H8	7.303	8.079	50118
C5H6-G6H8	7.302	7.559	69442
G6H1'-G6H8	5.721	7.56	128017
G6H1'-A7H8	5.722	8.287	211431
G6H2'-G6H1'	4.547	5.72	869103
G6H2'-G6H8	4.547	7.56	884140
G6H2'-A7H8	4.551	8.287	359131
G6H3'-G6H1'	4.646	5.72	127287
G6H3'-G6H8	4.646	7.56	615286
G6H3'-A7H8	4.647	8.287	237673
G6H4'-G6H1'	4.442	5.72	400986
G6H4'-G6H8	4.444	7.56	102546
G6H4'-A7H8	4.445	8.287	71750
G6H5'1-G6H1'	4.365	5.719	61401
G6H5'2-C5H1'	4.112	5.429	41850
G6H5'2-G6H1'	4.108	5.719	27949
G6H5'2-G6H3'	4.112	4.648	1457796
G6H5'2-G6H4'	4.108	4.441	3303269
G6H5'2-G6H5'1	4.108	4.363	3853778
G6H5'2-G6H8	4.108	7.56	239552
A7H1'-A7H2	5.693	7.929	56874
A7H1'-A7H8	5.693	8.287	215204
A7H1'-A8H1'	5.691	5.809	178136
A7H1'-A8H2	5.69	7.872	28558
A7H1'-A8H8	5.693	8.095	60918
A7H5'2-?	3.948	4.585	1332617
A7H5'2-G6H1'	3.951	5.719	34491
A7H5'2-A7H1'	3.95	5.694	108395
A7H5'2-A7H5'1	3.956	4.028	8444842
A7H5'2-A7H8	3.947	8.287	132900
A8H1'-A7H2	5.809	7.926	178218
A8H1'-A8H2	5.809	7.867	58782
A8H1'-A8H8	5.808	8.096	391920
A8H1'-G9H1'	5.808	6.031	46416
A8H1'-G9H8	5.809	7.784	43235
A8H1'-G10H8	5.81	8.072	38787
A8H2'-A8H1'	4.657	5.808	225351
A8H2'-A8H2	4.646	7.873	54810
A8H2'-A8H8	4.657	8.096	572280
A8H3'-A8H1'	4.764	5.809	87049
A8H3'-A8H8	4.766	8.097	326236
A8H4'-A8H1'	4.451	5.809	459777
A8H4'-A8H3'	4.454	4.768	1579277
A8H4'-A8H8	4.452	8.096	113907
A8H5'1-A7H1'	4.172	5.694	62492
A8H5'1-A8H1'	4.171	5.807	39299
A8H5'1-A8H3'	4.174	4.77	909073
A8H5'1-A8H4'	4.171	4.453	2467448
A8H5'1-A8H8	4.173	8.098	228522
A8H5'2-A7H1'	3.978	5.693	146404
A8H5'2-A8H1'	3.974	5.808	56124
A8H5'2-A8H2'	3.975	4.654	483927
A8H5'2-A8H3'	3.975	4.77	920071
A8H5'2-A8H4'	3.975	4.451	2150844
A8H5'2-A8H5'1	3.976	4.174	12286096
A8H5'2-A8H8	3.973	8.096	201283
G9H1'-A8H2	6.032	7.873	78547
G9H1'-G9H8	6.033	7.784	355045
G9H1'-G10H8	6.032	8.073	61741
G9H2'-G9H1'	4.679	6.033	491718
?-G9H8	4.94	7.784	670592
G9H4'-G9H1'	4.586	6.033	480528
G9H4'-G9H3'	4.59	4.937	158387
G9H5'1-G9H1'	4.471	6.033	77911
G9H5'1-G9H8	4.468	7.785	306013
G9H5'2-A8H1'	4.307	5.812	87923
G9H5'2-A8H2	4.299	7.872	62967
G9H5'2-G9H1'	4.303	6.033	50384
G9H5'2-G9H3'	4.302	4.939	613452
G9H5'2-G9H5'1	4.303	4.472	11292125
G9H5'2-G9H8	4.301	7.782	205713
G9H8-A8H8	7.787	8.095	46288
G10H1'-U11H5	4.944	5.164	98749
G10H1'-U11H6	4.945	7.821	82318
G10H2'-G10H8	4.467	8.074	392262
G10H2'-U11H5	4.467	5.166	193635
G10H3'-?	4.364	4.944	136506
G10H3'-U11H5	4.371	5.165	210873
G10H3'-U11H6	4.373	7.822	335231
G10H4'-?	4.391	4.943	229339
G10H5'2-?	4.339	4.944	237007
G10H5'2-G10H8	4.336	8.072	1904868
U11H1'-U11H6	5.602	7.823	137008
U11H1'-G12H1'	5.605	5.847	61571
U11H1'-G12H8	5.602	7.778	72077
U11H2'-U11H1'	4.69	5.603	1088920
U11H2'-U11H6	4.689	7.82	197139
U11H2'-G12H1'	4.69	5.852	158957
U11H2'-G12H8	4.689	7.778	2637950
U11H3'-U11H1'	4.595	5.602	212560
U11H3'-U11H5	4.595	5.166	101160

U11H3'-U11H6	4.597	7.822	933948
U11H5'1-U11H6	4.524	7.823	215561
U11H5'2-U11H1'	4.142	5.602	46787
U11H5'2-U11H4'	4.145	4.479	2969370
U11H5'2-U11H5'1	4.143	4.528	3908668
U11H5'2-U11H6	4.142	7.823	318477
U11H6-G10H8	7.821	8.072	76888
G12H1'-A4H2	5.852	7.432	236993
G12H1'-G12H8	5.852	7.778	164826
G12H2'-C13H5	4.545	5.299	151787
G12H2'-C13H6	4.548	7.687	1546344
G12H3'-G12H1'	4.621	5.852	270509
G12H3'-C13H5	4.625	5.299	142970
G12H3'-C13H6	4.625	7.688	303667
G12H4'-G12H1'	4.522	5.853	551817
G12H5'2-G12H5'1	4.179	4.566	2098642
C13H1'-C13H6	5.497	7.69	119857
C13H2'-C13H1'	4.214	5.497	907803
C13H2'-C13H6	4.214	7.685	132229
C13H2'-C14H1'	4.214	6.13	48812
C13H2'-C14H5	4.214	5.578	85356
C13H2'-C14H6	4.213	7.513	877591
C13H3'-C13H1'	4.452	5.498	187185
C13H3'-C13H5	4.461	5.299	43223
C13H3'-C13H6	4.458	7.691	741528
C13H3'-C13H6	4.458	7.691	741528
C13H3'-C14H6	4.453	7.513	177832
C13H4'-G12H1'	4.402	5.852	90113
C13H4'-C13H1'	4.412	5.498	326596
C13H4'-C13H5	4.414	5.299	20807
C13H4'-C13H6	4.415	7.688	273972
C13H5-C13H1'	5.297	5.503	56272
C13H5-C13H6	5.299	7.684	1268676
C13H5-C14H5	5.297	5.577	101566
C13H5'1-C13H6	4.586	7.689	267990
C13H5'2-C13H1'	4.077	5.503	38842
C13H5'2-C13H4'	4.084	4.412	1546655
C13H5'2-C13H5'1	4.084	4.582	8047310
C13H5'2-C13H6	4.083	7.689	203446
C14H1'-C14H6	6.127	7.515	90571
C14H2'-C14H1'	4.665	6.13	504140
C14H2'-C14H3'	4.67	4.978	261011
C14H2'-C14H6	4.664	7.513	254697
C14H3'-C14H1'	4.984	6.127	51851
C14H3'-C14H6	4.981	7.513	176042
C14H4'-C14H1'	4.437	6.13	385249
C14H4'-C14H6	4.435	7.513	148840
C14H5-C14H6	5.58	7.513	3984598
C14H5'1-C14H1'	4.542	6.13	55429
C14H5'1-C14H3'	4.544	4.98	63120
C14H5'1-C14H6	4.546	7.515	109330
C14H5'2-C14H1'	4.126	6.126	57371
C14H5'2-C14H3'	4.123	4.977	193156
C14H5'2-C14H5'1	4.124	4.544	4165163
C14H5'2-C14H6	4.123	7.513	89012
C14H6-C13H6	7.517	7.69	60267

NOESY 100 ms 100% D_2O

Assignment	w1	w2	Data Height
?-?	4.531	7.775	1110120
?-?	4.173	4.634	15853014
?-?	4.034	4.586	2920887
?-?	4.302	4.769	1678237
?-?	4.459	5.453	17868
?-?	7.779	8.077	228123
?-?	4.625	7.778	3322120
?-?	4.587	7.776	5697724
?-?	4.468	4.943	2132275
?-?	4.662	4.769	4390494
?-?	4.489	7.776	653103
?-?	4.176	7.778	1228627
?-G1H1'	4.581	5.866	328497
?-G1H2'	4.584	4.987	869476
?-G2H1'	4.645	5.957	354676
?-G2H1'	4.584	5.957	2793266
?-G2H8	4.586	7.664	739825
?-C3H1'	4.491	5.565	762810
?-C3H1'	4.618	5.565	458682
?-C3H1'	4.652	5.565	130026
?-C3H1'	4.175	5.565	144161
?-C3H5	4.586	5.338	504064
?-C3H5	4.648	5.338	311259
?-A4H1'	4.521	5.991	3158921
?-A4H1'	4.322	5.991	12884
?-A4H8	4.526	8.08	4887303
?-A4H8	4.614	8.081	1176542
?-C5H1'	4.52	5.427	471340
?-C5H1'	4.365	5.427	807221
?-C5H5	4.521	5.144	311604
?-C5H6	4.52	7.299	2832788
?-G6H1'	4.033	5.719	136067
?-G6H1'	4.592	5.722	199570
?-G6H8	4.367	7.559	383711
?-A7H1'	4.034	5.693	724009
?-A7H1'	4.587	5.692	1121731
?-A7H1'	4.45	5.694	128620
?-A7H2	4.457	7.927	80762
?-A7H2	4.587	7.926	107041
?-A7H8	4.031	8.286	410480
?-A7H8	4.589	8.287	2881774
?-A8H1'	4.337	5.809	79398
?-A8H1'	4.027	5.812	59756
?-A8H1'	4.397	5.807	59715
?-A8H1'	4.589	5.808	354189
?-A8H1'	4.944	5.809	177815
A7H4'-A8H2	4.033	7.872	93789
?-A8H2	4.943	7.874	64905
?-A8H2	4.587	7.872	225508
?-A8H2	4.468	7.873	152868
?-A8H2	4.391	7.872	50985
?-A8H3'	4.415	4.77	872124
?-A8H3'	4.037	4.769	392271
?-A8H8	3.944	8.096	192072
?-A8H8	4.031	8.096	221633
?-A8H8	4.585	8.096	2217231
?-G9H1'	4.94	6.033	346266
?-G9H3'	4.679	4.939	1663320
?-G10H8	4.382	8.072	1063302
?-G10H8	4.587	8.072	824715
?-G10H8	4.941	8.072	996127
?-U11H1'	4.475	5.602	934079
?-U11H1'	4.536	5.601	196208
?-U11H6	4.467	7.823	3454351
?-U11H6	4.335	7.821	222626
?-G12H1'	4.554	5.852	1860847
?-C13H1'	4.545	5.498	398580
?-C14H5	4.449	5.577	183963
G1H1'-G1H8	5.866	8.196	372989
G1H1'-G2H1'	5.863	5.956	376558
G1H1'-G2H8	5.868	7.663	688346

G1H2'-G1H1'	4.989	5.866	2975845
G1H2'-G1H8	4.989	8.198	1120758
G1H2'-G2H1'	4.988	5.957	369184
G1H2'-G2H8	4.989	7.663	3434428
G1H3'-G1H1'	4.769	5.866	381294
G1H3'-G1H2'	4.77	4.987	3745167
G1H3'-G1H8	4.771	8.198	480540
G1H3'-G2H8	4.769	7.663	652015
G1H4'-G1H1'	4.609	5.865	1108101
G1H4'-G1H2'	4.606	4.988	787045
G1H4'-G1H3'	4.616	4.769	4299606
G1H4'-G1H8	4.61	8.196	179716
G1H5'1-G1H1'	4.482	5.866	130956
G1H5'1-G1H2'	4.484	4.987	192096
G1H5'1-G1H3'	4.491	4.769	1523412
G1H5'1-G1H8	4.488	8.202	267347
G1H5'2-G1H1'	4.336	5.868	99393
G1H5'2-G1H2'	4.338	4.989	186889
G1H5'2-G1H3'	4.336	4.771	1546574
G1H5'2-G1H5'1	4.333	4.489	8880227
G1H5'2-G1H8	4.333	8.203	689449
G2H1'-G2H8	5.957	7.664	265919
G2H1'-C3H6	5.958	7.776	148270
G2H2'-C3H1'	4.585	5.566	489820
G2H3'-G1H2'	4.642	4.987	134502
G2H3'-G2H8	4.644	7.664	787976
G2H5'2-G1H1'	4.302	5.864	139177
G2H5'2-G1H2'	4.302	4.987	418964
G2H5'2-G2H1'	4.298	5.957	85638
G2H5'2-G2H5'1	4.302	4.589	14144141
G2H5'2-G2H8	4.302	7.663	695805
G2H8-G1H8	7.661	8.206	67635
C3H1'-C3H6	5.565	7.775	309109
C3H1'-A4H8	5.564	8.08	148508
C3H2'-C3H1'	4.526	5.565	2309896
C3H2'-C3H5	4.53	5.336	69922
C3H3'-C3H5	4.619	5.339	234861
C3H5-G2H8	5.337	7.664	136038
C3H5-C3H1'	5.339	5.561	110106
C3H5-C3H6	5.338	7.773	1283361
C3H5-A4H8	5.337	8.079	36478
C3H5'2-G2H1'	4.165	5.951	42162
A4H1'-A4H2	5.994	7.433	68747
A4H1'-A4H8	5.992	8.08	232221
A4H1'-C5H6	5.992	7.299	72591
A4H2'-A4H2	4.52	7.433	145245
A4H2'-A4H3'	4.522	4.704	4525181
A4H3'-A4H1'	4.706	5.99	275110
A4H3'-A4H8	4.709	8.08	1331277
A4H3'-C5H1'	4.71	5.429	37054
A4H3'-C5H5	4.709	5.143	164965
A4H3'-C5H6	4.708	7.3	503207
A4H5'1-A4H1'	4.619	5.991	138950
A4H5'2-A4H1'	4.186	5.99	111496
A4H5'2-A4H3'	4.182	4.695	1386686
A4H5'2-A4H8	4.185	8.081	705992
A4H5'2-C5H6	4.195	7.297	62808
C5H1'-A4H1'	5.425	5.991	48557
C5H1'-A4H2	5.428	7.432	477952
C5H1'-C5H6	5.426	7.299	278980
C5H1'-G6H1'	5.428	5.721	66881
C5H1'-G6H8	5.427	7.561	37101
C5H2'-A4H2	4.285	7.432	39437
C5H2'-C5H1'	4.286	5.427	2223452
C5H2'-C5H5	4.284	5.145	61294
C5H2'-C5H6	4.291	7.299	636640
C5H2'-G6H1'	4.286	5.72	240635
C5H2'-G6H8	4.287	7.56	1547571
C5H3'-C5H1'	4.333	5.427	427867
C5H3'-C5H5	4.332	5.142	135937
C5H3'-C5H6	4.329	7.301	1462068
C5H3'-G6H1'	4.332	5.719	74650
C5H3'-G6H8	4.329	7.559	640063
C5H4'-A4H1'	4.37	5.993	40756
C5H4'-C5H5	4.364	5.142	77024
C5H4'-C5H6	4.358	7.3	832578
C5H5-A4H8	5.143	8.081	140805
C5H5-C5H1'	5.142	5.424	51067
C5H5-C5H6	5.144	7.3	74427
C5H5-G6H8	5.143	7.562	41206
C5H5'2-A4H1'	4.08	5.992	39356
C5H5'2-C5H1'	4.084	5.428	98377
C5H5'2-C5H5	4.086	5.143	30757
C5H5'2-C5H5'1	4.082	4.514	15304585
C5H5'2-C5H6	4.083	7.3	473583
C5H6-A4H2	7.305	7.433	93967
C5H6-A4H8	7.303	8.079	85407
C5H6-G6H8	7.302	7.559	140598
G6H1'-G6H8	5.719	7.559	288043
G6H1'-A7H8	5.722	8.287	438524
G6H2'-G6H1'	4.547	5.72	1537641
G6H2'-G6H8	4.547	7.56	1595813
G6H2'-A7H8	4.551	8.287	731505
G6H3'-G6H1'	4.646	5.72	263671
G6H3'-G6H8	4.646	7.56	1167089
G6H3'-A7H1'	4.648	5.695	76486
G6H3'-A7H8	4.647	8.287	492738
G6H4'-C5H1'	4.438	5.429	39684
G6H4'-G6H1'	4.442	5.72	746873
G6H4'-G6H8	4.444	7.56	192224
G6H4'-A7H8	4.445	8.287	141816
G6H5'1-G6H1'	4.359	5.72	133103
G6H5'1-A7H8	4.363	8.286	64422
G6H5'2-C5H1'	4.108	5.427	61572
G6H5'2-G6H1'	4.113	5.721	80408
G6H5'2-G6H3'	4.111	4.642	2051540
G6H5'2-G6H4'	4.108	4.441	5520561
G6H5'2-G6H5'1	4.108	4.36	10899509
G6H5'2-G6H8	4.108	7.56	425244
G6H5'2-A7H8	4.115	8.289	62827
G6H8-A7H8	7.567	8.285	50374
A7H1'-A7H2	5.689	7.926	90025
A7H1'-A7H8	5.693	8.287	438395
A7H1'-A8H1'	5.691	5.809	180879
A7H1'-A8H2	5.692	7.872	55665
A7H1'-A8H8	5.693	8.097	131330
A7H5'2-?	3.948	4.585	2048276
A7H5'2-G6H1'	3.951	5.719	60775
A7H5'2-A7H1'	3.95	5.694	183258
A7H5'2-A7H5'1	3.949	4.031	13744019
A7H5'2-A7H8	3.946	8.287	295443
A7H5'2-A8H2	3.961	7.876	41355
A8H1'-A7H2	5.808	7.926	366086
A8H1'-A8H2	5.809	7.874	60039
A8H1'-A8H8	5.809	8.097	758354
A8H1'-G9H1'	5.808	6.031	69428
A8H1'-G9H8	5.805	7.788	59180
A8H1'-G10H8	5.81	8.072	95854
A8H2'-A8H1'	4.657	5.808	1065818
A8H2'-A8H2	4.656	7.871	60334
A8H2'-A8H8	4.657	8.096	1092102
A8H3'-A8H1'	4.764	5.809	163657

A8H3'-A8H8	4.765	8.097	619796
A8H3'-G9H8	4.766	7.788	121327
A8H4'-A8H1'	4.451	5.809	876717
A8H4'-A8H3'	4.454	4.768	2167526
A8H4'-A8H8	4.452	8.096	255240
A8H5'1-A7H1'	4.172	5.694	108223
A8H5'1-A7H8	4.171	8.288	42617
A8H5'1-A8H1'	4.171	5.807	105614
A8H5'1-A8H3'	4.174	4.77	1173771
A8H5'1-A8H4'	4.174	4.451	4155089
A8H5'1-A8H8	4.173	8.098	447637
A8H5'2-A7H1'	3.978	5.693	269344
A8H5'2-A8H1'	3.974	5.808	88491
A8H5'2-A8H2'	3.975	4.654	544902
A8H5'2-A8H3'	3.974	4.77	1344546
A8H5'2-A8H4'	3.975	4.451	3417407
A8H5'2-A8H5'1	3.976	4.173	12641415
A8H5'2-A8H8	3.974	8.097	431536
A8H5'2-G9H8	3.975	7.785	65955
G9H1'-A8H2	6.033	7.872	106956
G9H1'-G9H8	6.033	7.784	715736
G9H1'-G10H8	6.032	8.072	91186
G9H2'-G9H1'	4.678	6.033	1403044
?-G9H8	4.94	7.784	1269260
G9H4'-G9H1'	4.586	6.033	927770
G9H4'-G9H3'	4.59	4.937	918778
G9H5'1-G9H1'	4.474	6.033	162530
G9H5'1-G9H8	4.466	7.785	634964
G9H5'2-A8H1'	4.303	5.808	134484
G9H5'2-A8H2	4.302	7.872	146628
G9H5'2-G9H1'	4.301	6.032	132754
G9H5'2-G9H3'	4.302	4.939	1206070
G9H5'2-G9H5'1	4.303	4.472	13639921
G9H5'2-G9H8	4.301	7.782	418895
G9H8-A8H8	7.787	8.095	91750
G10H1'-U11H5	4.946	5.161	88728
G10H1'-U11H6	4.945	7.821	180854
G10H2'-G10H8	4.468	8.073	765868
G10H2'-U11H5	4.467	5.166	410901
G10H3'-?	4.364	4.944	437825
G10H3'-U11H1'	4.371	5.6	70368
G10H3'-U11H5	4.371	5.165	408126
G10H3'-U11H6	4.373	7.822	655553
G10H4'-?	4.391	4.943	572614
G10H4'-G9H1'	4.393	6.032	53985
G10H5'2-?	4.339	4.944	596407
G10H5'2-G9H1'	4.342	6.035	72386
G10H5'2-G10H8	4.336	8.073	3505202
G10H5'2-U11H5	4.325	5.165	129671
U11H1'-A4H2	5.604	7.429	64957
U11H1'-U11H6	5.602	7.823	287988
U11H1'-G12H1'	5.603	5.852	47826
U11H1'-G12H8	5.602	7.779	180371
U11H2'-A4H2	4.692	7.432	80773
U11H2'-U11H1'	4.69	5.603	2074348
U11H2'-U11H5	4.69	5.165	40356
U11H2'-U11H6	4.689	7.822	374221
U11H2'-G12H1'	4.69	5.852	340669
U11H2'-G12H8	4.689	7.778	4703485
U11H3'-U11H1'	4.595	5.602	435318
U11H3'-U11H5	4.595	5.163	196777
U11H3'-U11H6	4.597	7.822	1653765
U11H5-G10H8	5.167	8.072	147210
U11H5-U11H6	5.167	7.824	329162
U11H5'1-U11H5	4.523	5.167	43258
U11H5'1-U11H6	4.524	7.823	456400
U11H5'2-U11H1'	4.142	5.602	113202
U11H5'2-U11H4'	4.145	4.479	4048196
U11H5'2-U11H5	4.145	5.169	53280
U11H5'2-U11H5'1	4.143	4.528	14549177
U11H5'2-U11H6	4.142	7.823	525155
U11H6-G10H8	7.821	8.072	72304
G12H1'-A4H2	5.852	7.432	428519
G12H1'-G12H8	5.852	7.778	260754
G12H2'-C13H5	4.545	5.299	359605
G12H2'-C13H6	4.548	7.687	2731820
G12H3'-G12H1'	4.621	5.852	558972
G12H3'-C13H5	4.625	5.299	311483
G12H3'-C13H6	4.625	7.688	617824
G12H4'-G12H1'	4.522	5.853	1048033
G12H5'2-U11H1'	4.184	5.603	32208
G12H5'2-G12H1'	4.178	5.851	108916
G12H5'2-G12H5'1	4.179	4.562	12205586
G12H5'2-C13H5	4.184	5.3	53058
C13H1'-C13H6	5.497	7.688	238846
C13H2'-C13H1'	4.214	5.497	1722416
C13H2'-C13H6	4.214	7.685	255562
C13H2'-C14H1'	4.213	6.131	112988
C13H2'-C14H5	4.214	5.578	173652
C13H2'-C14H6	4.213	7.513	1619074
C13H3'-C13H1'	4.452	5.498	425791
C13H3'-C13H5	4.461	5.299	116532
C13H3'-C13H6	4.458	7.691	1310030
C13H3'-C13H6	4.458	7.691	1310030
C13H3'-C14H6	4.453	7.513	363333
C13H4'-G12H1'	4.402	5.852	44292
C13H4'-C13H1'	4.412	5.498	648985
C13H4'-C13H5	4.414	5.299	46885
C13H4'-C13H6	4.415	7.688	500404
C13H5-G12H1'	5.305	5.851	48486
C13H5-C13H1'	5.298	5.501	108689
C13H5-C13H6	5.299	7.687	2267305
C13H5-C14H5	5.298	5.578	152529
C13H5'1-C13H1'	4.584	5.498	104616
C13H5'1-C13H6	4.586	7.689	619519
C13H5'2-G12H1'	4.083	5.847	56532
C13H5'2-C13H1'	4.085	5.498	158857
C13H5'2-C13H4'	4.084	4.412	2793776
C13H5'2-C13H5	4.083	5.302	48225
C13H5'2-C13H5'1	4.084	4.582	9182961
C13H5'2-C13H6	4.083	7.689	345544
C13H5'2-C14H6	4.081	7.514	49366
C14H1'-C14H6	6.13	7.512	162358
C14H2'-C14H1'	4.665	6.13	799380
C14H2'-C14H3'	4.67	4.978	766612
C14H2'-C14H6	4.664	7.513	444949
C14H3'-C14H1'	4.985	6.129	105400
C14H3'-C14H6	4.98	7.513	328578
C14H4'-C14H1'	4.437	6.13	744501
C14H4'-C14H3'	4.442	4.978	406379
C14H4'-C14H6	4.435	7.513	292784
C14H5-C13H6	5.582	7.688	112845
C14H5-C14H6	5.58	7.513	3748064
C14H5'1-C14H1'	4.542	6.13	95831
C14H5'1-C14H3'	4.544	4.98	206040
C14H5'1-C14H6	4.546	7.515	243473
C14H5'2-C13H1'	4.123	5.496	78951
C14H5'2-C14H1'	4.123	6.13	69358
C14H5'2-C14H3'	4.123	4.977	392557
C14H5'2-C14H5'1	4.124	4.544	9580568
C14H5'2-C14H6	4.123	7.513	190133
C14H6-C13H6	7.517	7.69	113565

NOESY 150 ms 100% D_2O

Assignment	w1	w2	Data Height
?-?	4.176	7.778	1704753
?-?	4.173	4.634	17647860
?-?	4.531	7.775	1657439
?-?	4.459	5.453	45647
?-?	4.587	7.776	7685771
?-?	7.78	8.077	339341
?-?	4.468	4.943	3089303
?-?	4.034	4.586	3785084
?-?	4.489	7.776	947787
?-?	4.625	7.778	4553761
?-?	4.302	4.769	1960623
?-?	4.662	4.769	4554081
?-G1H1'	4.581	5.866	520894
?-G1H2'	4.584	4.987	1315131
?-G2H1'	4.645	5.957	599270
?-G2H1'	4.584	5.957	3925409
?-G2H8	4.586	7.664	1130845
?-C3H1'	4.618	5.565	712106
?-C3H1'	4.175	5.565	267831
?-C3H1'	4.652	5.565	223298
?-C3H1'	4.491	5.565	1124088
?-C3H5	4.648	5.338	514289
?-C3H5	4.586	5.338	871246
?-A4H1'	4.521	5.991	4418470
?-A4H1'	4.322	5.991	67688
?-A4H2	4.551	7.432	102233
?-A4H8	4.614	8.081	1729731
?-A4H8	4.526	8.08	6664769
?-C5H1'	4.52	5.427	727789
?-C5H1'	4.365	5.427	1183440
?-C5H5	4.521	5.144	537592
?-C5H6	4.52	7.299	3856120
?-G6H1'	4.033	5.719	175046
?-G6H1'	4.592	5.722	281062
?-G6H8	4.367	7.559	747718
?-A7H1'	4.45	5.694	215461
?-A7H1'	4.587	5.692	1736321
?-A7H1'	4.034	5.693	1040182
?-A7H2	4.457	7.927	103715
?-A7H2	4.587	7.926	151119
?-A7H8	4.589	8.287	4037153
?-A7H8	4.032	8.287	594166
?-A8H1'	4.033	5.808	73952
?-A8H1'	4.589	5.808	524087
?-A8H1'	4.944	5.809	268181
?-A8H1'	4.397	5.807	133976
?-A8H1'	4.337	5.809	116899
?-A8H2	4.381	7.872	70112
?-A8H2	4.945	7.873	86094
?-A8H2	4.587	7.872	322117
A7H4'-A8H2	4.033	7.872	134178
?-A8H2	4.468	7.873	253383
?-A8H3'	4.415	4.77	919045
?-A8H3'	4.037	4.769	358357
?-A8H8	4.585	8.096	3145415
?-A8H8	4.034	8.097	298309
?-A8H8	3.944	8.096	296670
?-G9H1'	4.939	6.033	519850
?-G9H3'	4.679	4.939	2531527
?-G10H8	4.587	8.072	1191194
?-G10H8	4.382	8.072	1530883
?-G10H8	4.943	8.072	1378558
?-U11H1'	4.536	5.601	354354
?-U11H1'	4.475	5.602	1357156
?-U11H6	4.467	7.823	4583926
?-U11H6	4.335	7.821	356631
?-G12H1'	4.554	5.852	2612272
?-C13H1'	4.545	5.498	572740
?-C14H5	4.449	5.577	316721
G1H1'-G1H8	5.865	8.195	540190
G1H1'-G2H1'	5.863	5.956	511104
G1H1'-G2H8	5.868	7.663	971601
G1H2'-G1H1'	4.989	5.866	4017892
G1H2'-G1H8	4.988	8.194	1489930
G1H2'-G2H1'	4.989	5.957	508960
G1H2'-G2H8	4.989	7.663	4506360
G1H3'-G1H1'	4.769	5.866	571428
G1H3'-G1H2'	4.77	4.987	4785034
G1H3'-G1H8	4.768	8.196	634709
G1H3'-G2H8	4.769	7.663	878990
G1H4'-G1H1'	4.609	5.865	1617444
G1H4'-G1H2'	4.606	4.988	1272831
G1H4'-G1H3'	4.614	4.769	4373050
G1H4'-G1H8	4.61	8.196	262298
G1H5'1-G1H1'	4.482	5.866	218525
G1H5'1-G1H2'	4.484	4.987	324460
G1H5'1-G1H3'	4.491	4.769	1672744
G1H5'1-G1H8	4.485	8.197	339345
G1H5'2-G1H1'	4.336	5.868	200531
G1H5'2-G1H2'	4.338	4.989	338352
G1H5'2-G1H3'	4.336	4.771	1700110
G1H5'2-G1H5'1	4.333	4.489	10643453
G1H5'2-G1H8	4.332	8.196	778827
G2H1'-G2H8	5.957	7.664	345644
G2H1'-C3H6	5.958	7.776	259654
G2H2'-C3H1'	4.585	5.566	727769
G2H3'-G1H2'	4.642	4.987	475358
G2H3'-G2H8	4.644	7.664	1055278
G2H5'2-G1H1'	4.302	5.864	247978
G2H5'2-G1H2'	4.302	4.987	703018
G2H5'2-G2H1'	4.301	5.957	174909
G2H5'2-G2H5'1	4.302	4.588	18424036
G2H5'2-G2H8	4.302	7.663	967997
G2H5'2-C3H5	4.302	5.336	71496
G2H8-G1H8	7.662	8.199	79244
C3H1'-G2H1'	5.571	5.96	43574
C3H1'-C3H6	5.564	7.774	383805
C3H1'-A4H1'	5.566	5.992	42997
C3H1'-A4H8	5.565	8.081	244234
C3H2'-C3H1'	4.526	5.565	3224251
C3H2'-C3H5	4.53	5.336	145712
C3H3'-C3H5	4.619	5.339	374832
C3H5-G2H1'	5.338	5.956	73445
C3H5-G2H8	5.337	7.664	143910
C3H5-C3H1'	5.335	5.56	189968
C3H5-C3H6	5.338	7.773	2563646
C3H5-A4H8	5.336	8.08	74968
C3H5'2-G2H1'	4.174	5.958	45305
C3H5'2-C3H5	4.171	5.334	75796
A4H1'-A4H2	5.993	7.432	136153
A4H1'-A4H8	5.991	8.08	413316
A4H1'-C5H6	5.991	7.301	132795
A4H2-?	7.431	7.78	55664
A4H2'-A4H2	4.52	7.433	205019
A4H2'-A4H3'	4.521	4.705	5229518
A4H3'-A4H1'	4.708	5.991	437912
A4H3'-A4H8	4.708	8.08	1793955

A4H3'-C5H1'	4.71	5.429	41759
A4H3'-C5H5	4.709	5.143	307716
A4H3'-C5H6	4.708	7.299	701638
A4H5'1-A4H1'	4.619	5.991	257318
A4H5'1-C5H6	4.624	7.299	41890
A4H5'2-A4H1'	4.185	5.991	151589
A4H5'2-A4H3'	4.185	4.695	1857387
A4H5'2-A4H8	4.185	8.081	987373
A4H5'2-C5H5	4.18	5.143	69023
A4H5'2-C5H6	4.186	7.3	89425
C5H1'-A4H1'	5.427	5.992	80895
C5H1'-A4H2	5.427	7.432	684597
C5H1'-C5H6	5.427	7.3	373227
C5H1'-G6H1'	5.428	5.721	102518
C5H1'-G6H8	5.427	7.56	83567
C5H2'-A4H2	4.285	7.432	52695
C5H2'-C5H1'	4.286	5.427	3091338
C5H2'-C5H5	4.284	5.145	84899
C5H2'-C5H6	4.291	7.299	908107
C5H2'-G6H1'	4.286	5.72	355387
C5H2'-G6H8	4.287	7.56	2119137
C5H2'-A7H8	4.289	8.287	60603
C5H3'-C5H1'	4.333	5.427	673179
C5H3'-C5H5	4.332	5.142	203425
C5H3'-C5H6	4.329	7.301	1911326
C5H3'-G6H1'	4.332	5.719	122098
C5H3'-G6H8	4.329	7.559	906887
C5H4'-A4H1'	4.37	5.993	85394
C5H4'-A4H2	4.364	7.429	40982
C5H4'-C5H5	4.364	5.142	129648
C5H4'-C5H6	4.358	7.3	1155587
C5H5-A4H1'	5.145	5.991	52775
C5H5-A4H8	5.143	8.081	223393
C5H5-C5H1'	5.148	5.429	100341
C5H5-C5H6	5.144	7.3	3693499
C5H5-G6H8	5.144	7.561	89289
C5H5'2-A4H1'	4.079	5.993	90095
C5H5'2-C5H1'	4.084	5.428	181718
C5H5'2-C5H5	4.086	5.143	72848
C5H5'2-C5H5'1	4.082	4.514	15856886
C5H5'2-C5H6	4.083	7.3	669132
C5H6-A4H2	7.305	7.433	145576
C5H6-A4H8	7.303	8.079	142248
C5H6-G6H8	7.302	7.559	205932
G6H1'-G6H8	5.721	7.56	415135
G6H1'-A7H8	5.722	8.287	579418
G6H2'-G6H1'	4.547	5.72	2164036
G6H2'-G6H8	4.547	7.56	2230589
G6H2'-A7H8	4.551	8.287	1049665
G6H3'-G6H1'	4.646	5.72	432854
G6H3'-G6H8	4.646	7.56	1593957
G6H3'-A7H1'	4.648	5.695	112974
G6H3'-A7H8	4.647	8.287	658282
G6H4'-C5H1'	4.438	5.429	77939
G6H4'-G6H1'	4.442	5.72	1108472
G6H4'-G6H8	4.444	7.56	299659
G6H4'-A7H8	4.445	8.287	242165
G6H5'1-G6H1'	4.365	5.719	234455
G6H5'1-A7H8	4.367	8.288	90631
G6H5'2-C5H1'	4.108	5.427	105131
G6H5'2-G6H1'	4.108	5.719	150565
G6H5'2-G6H3'	4.111	4.642	2534031
G6H5'2-G6H4'	4.108	4.441	7391901
G6H5'2-G6H5'1	4.108	4.363	13535790
G6H5'2-G6H8	4.108	7.56	607356
G6H5'2-A7H8	4.111	8.286	91590
G6H8-A7H8	7.559	8.286	67646
A7H1'-A7H2	5.693	7.926	126471
A7H1'-A7H8	5.693	8.287	639770
A7H1'-A8H1'	5.691	5.809	274900
A7H1'-A8H2	5.69	7.872	91727
A7H1'-A8H8	5.695	8.097	186422
A7H5'2-?	3.948	4.586	2529299
A7H5'2-G6H1'	3.951	5.719	106375
A7H5'2-A7H1'	3.95	5.694	297325
A7H5'2-A7H5'1	3.948	4.032	15425060
A7H5'2-A7H8	3.947	8.287	421011
A7H5'2-A8H2	3.953	7.87	70084
A8H1'-A7H2	5.809	7.926	498417
A8H1'-A8H2	5.807	7.872	103255
A8H1'-A8H8	5.809	8.097	1116646
A8H1'-G9H1'	5.808	6.031	149625
A8H1'-G9H8	5.809	7.784	82784
A8H1'-G10H8	5.81	8.072	125201
A8H2'-A8H1'	4.657	5.808	1524417
A8H2'-A8H2	4.656	7.871	113788
A8H2'-A8H8	4.657	8.096	1522170
A8H3'-A8H1'	4.764	5.809	250364
A8H3'-A8H8	4.765	8.097	790239
A8H3'-G9H8	4.766	7.788	200797
A8H4'-A8H1'	4.451	5.809	1240869
A8H4'-A8H3'	4.454	4.768	2634961
A8H4'-A8H8	4.452	8.096	433897
A8H5'1-A7H1'	4.172	5.694	214424
A8H5'1-A7H8	4.172	8.289	46155
A8H5'1-A8H1'	4.171	5.807	159264
A8H5'1-A8H3'	4.174	4.77	1473995
A8H5'1-A8H4'	4.171	4.453	5646099
A8H5'1-A8H8	4.173	8.098	608031
A8H5'2-A7H1'	3.978	5.693	379873
A8H5'2-A8H1'	3.974	5.808	160154
A8H5'2-A8H2'	3.975	4.654	719548
A8H5'2-A8H3'	3.975	4.77	1596178
A8H5'2-A8H4'	3.975	4.451	4460844
A8H5'2-A8H5'1	3.976	4.173	12833786
A8H5'2-A8H8	3.973	8.096	623316
A8H5'2-G9H8	3.975	7.785	91850
G9H1'-A7H2	6.031	7.926	40777
G9H1'-A8H2	6.033	7.872	214804
G9H1'-G9H8	6.033	7.784	1055984
G9H1'-G10H8	6.034	8.072	154538
G9H2'-G9H1'	4.679	6.033	1911816
?-G9H8	4.941	7.784	1755538
G9H4'-G9H1'	4.586	6.033	1302228
G9H4'-G9H3'	4.59	4.937	1337802
G9H5'1-G9H1'	4.471	6.033	274865
G9H5'1-G9H8	4.467	7.785	926837
G9H5'2-A8H1'	4.303	5.808	189585
G9H5'2-A8H2	4.299	7.872	189379
G9H5'2-G9H1'	4.303	6.033	186620
G9H5'2-G9H3'	4.302	4.939	1713563
G9H5'2-G9H5'1	4.303	4.472	18055820
G9H5'2-G9H8	4.301	7.782	617056
G9H8-A8H8	7.787	8.095	164414
G10H1'-U11H5	4.944	5.164	131957
G10H1'-U11H6	4.945	7.821	234389
G10H2'-G10H8	4.468	8.073	1126525
G10H2'-U11H5	4.467	5.166	680250
G10H3'-?	4.364	4.944	728800
G10H3'-U11H1'	4.371	5.6	92443
G10H3'-U11H5	4.371	5.165	605847
G10H3'-U11H6	4.373	7.822	919338

G10H4'-?	4.391	4.943	915546
G10H4'-G9H1'	4.39	6.033	92987
G10H5'2-?	4.339	4.944	973693
G10H5'2-G9H1'	4.342	6.035	113603
G10H5'2-G10H8	4.336	8.072	4913200
G10H5'2-U11H5	4.325	5.165	192849
U11H1'-A4H2	5.604	7.431	68515
U11H1'-U11H6	5.602	7.823	379692
U11H1'-G12H1'	5.603	5.852	122312
U11H1'-G12H8	5.602	7.778	297858
U11H2'-A4H2	4.692	7.432	74455
U11H2'-U11H1'	4.69	5.603	2862641
U11H2'-U11H5	4.69	5.165	88477
U11H2'-U11H6	4.689	7.822	546849
U11H2'-G12H1'	4.69	5.852	468180
U11H2'-G12H8	4.689	7.778	6265446
U11H3'-U11H1'	4.595	5.602	653707
U11H3'-U11H5	4.595	5.166	316394
U11H3'-U11H6	4.597	7.822	2229437
U11H5-G10H8	5.167	8.072	189257
U11H5-U11H1'	5.165	5.606	81577
U11H5-U11H6	5.167	7.822	3610838
U11H5'1-U11H5	4.523	5.167	118873
U11H5'1-U11H6	4.524	7.823	657148
U11H5'2-U11H1'	4.142	5.602	163467
U11H5'2-U11H4'	4.145	4.479	5729914
U11H5'2-U11H5	4.145	5.169	72307
U11H5'2-U11H5'1	4.143	4.528	13271156
U11H5'2-U11H6	4.142	7.823	760536
U11H6-G10H8	7.821	8.072	162458
G12H1'-A4H2	5.852	7.432	622075
G12H1'-G12H8	5.852	7.778	398583
G12H2'-C13H5	4.545	5.299	603310
G12H2'-C13H6	4.548	7.687	3679863
G12H3'-G12H1'	4.621	5.852	829607
G12H3'-C13H5	4.625	5.299	453341
G12H3'-C13H6	4.625	7.688	833898
G12H4'-G12H1'	4.522	5.853	1488752
G12H5'2-U11H1'	4.184	5.603	73147
G12H5'2-G12H1'	4.178	5.851	184014
G12H5'2-G12H5'1	4.177	4.561	10994290
G12H5'2-C13H5	4.18	5.3	74372
C13H1'-G12H1'	5.498	5.855	61839
C13H1'-C13H6	5.498	7.687	347497
C13H1'-C14H6	5.496	7.514	60033
C13H2'-C13H1'	4.214	5.497	2390649
C13H2'-C13H5	4.21	5.298	54230
C13H2'-C13H6	4.214	7.685	358855
C13H2'-C14H1'	4.214	6.13	184532
C13H2'-C14H5	4.214	5.578	288813
C13H2'-C14H6	4.213	7.513	2193021
C13H3'-C13H1'	4.452	5.498	611863
C13H3'-C13H5	4.461	5.299	212606
C13H3'-C13H6	4.458	7.691	1743125
C13H3'-C13H6	4.458	7.691	1743125
C13H3'-C14H6	4.453	7.513	523099
C13H4'-G12H1'	4.402	5.852	86381
C13H4'-C13H1'	4.412	5.498	954732
C13H4'-C13H5	4.414	5.299	71367
C13H4'-C13H6	4.415	7.688	717398
C13H5-G12H1'	5.3	5.852	51316
C13H5-C13H1'	5.297	5.503	128057
C13H5-C13H6	5.299	7.687	2649882
C13H5-C14H5	5.298	5.578	244282
C13H5'1-C13H1'	4.584	5.498	195668
C13H5'1-C13H6	4.586	7.689	928925
C13H5'2-G12H1'	4.094	5.85	74122
C13H5'2-C13H1'	4.085	5.498	226097
C13H5'2-C13H4'	4.084	4.412	3876045
C13H5'2-C13H5	4.082	5.297	64558
C13H5'2-C13H5'1	4.084	4.582	10965963
C13H5'2-C13H6	4.083	7.689	522080
C13H5'2-C14H6	4.081	7.514	76580
C14H1'-C14H6	6.129	7.512	258078
C14H2'-C14H1'	4.665	6.13	1306087
C14H2'-C14H3'	4.67	4.978	1219231
C14H2'-C14H6	4.664	7.513	667627
C14H3'-C14H1'	4.985	6.129	145936
C14H3'-C14H5	4.982	5.581	58258
C14H3'-C14H6	4.981	7.513	506388
C14H4'-C14H1'	4.437	6.13	1047326
C14H4'-C14H3'	4.442	4.978	688952
C14H4'-C14H6	4.435	7.513	415835
C14H5-C13H6	5.582	7.688	147894
C14H5-C14H6	5.58	7.513	3614897
C14H5'1-C14H1'	4.542	6.13	147229
C14H5'1-C14H3'	4.544	4.98	424380
C14H5'1-C14H6	4.546	7.515	360768
C14H5'2-C13H1'	4.123	5.496	108510
C14H5'2-C14H1'	4.119	6.13	107664
C14H5'2-C14H3'	4.123	4.977	586089
C14H5'2-C14H5'1	4.124	4.544	13047518
C14H5'2-C14H6	4.123	7.513	252873
C14H6-C13H6	7.517	7.69	134909

NOESY 100 ms 95%H_2O/5%D_2O

Assignment	w1	w2	Data Height
?-G2H1	8.592	13.334	534314
?-G2H1	6.912	13.334	1522570
?-G2H1	7.244	13.335	310678
?-G2H1'	7.15	5.94	568263
?-G6H22	4.293	6.2	34034
?-G6H22	4.961	6.202	252827
?-G6H22	4.444	6.2	71764
?-G6H22	4.571	6.198	75980
?-U11H3	6.351	13.698	56188
?-C14H1'	7.154	6.115	126640
G1H8-C14H41	8.167	8.596	180931
G2H1'-G2H1	5.939	13.338	34749
G2H8-C3H41	7.638	8.595	57104
C3H1'-G2H1	5.556	13.335	184174
C3H1'-G12H1	5.548	12.672	39515
C3H2'-G12H1	4.518	12.681	66009
C3H41-U11H3	8.595	13.69	23677
C3H41-G12H1	8.594	12.681	4379595
C3H42-G2H8	6.911	7.63	111169
C3H42-C3H41	6.909	8.595	39651412
C3H42-C3H6	6.907	7.75	351065
C3H42-A4H2	6.914	7.416	64363
C3H42-A4H8	6.91	8.057	29879
C3H42-U11H3	6.906	13.696	24852
C3H42-G12H1	6.909	12.681	1813033
C3H5-G2H1	5.321	13.327	49795
C3H5-C3H41	5.319	8.596	1703624
C3H5-C3H42	5.319	6.911	4195033
C3H5-G12H1	5.316	12.68	68180
C3H6-G2H1	7.731	13.335	169292
C3H6-C3H41	7.744	8.594	109180
C3H6-G12H1	7.746	12.679	134135
A4H1'-G12H1	5.975	12.677	216964

A4H2-C5H41	7.405	8.196	47879
A4H2-U11H3	7.415	13.697	1013930
A4H2-G12H1	7.416	12.681	169260
A4H8-G12H1	8.052	12.684	142485
C5H41-U11H3	8.197	13.696	97174
C5H42-?	6.807	8.062	300301
C5H42-A4H2	6.803	7.414	52106
C5H42-C5H41	6.806	8.201	28796224
C5H42-C5H6	6.805	7.265	276748
C5H42-G6H8	6.804	7.558	40983
C5H42-U11H3	6.82	13.694	35815
C5H5-C5H41	5.123	8.201	1391860
C5H5-C5H42	5.123	6.808	4209950
C5H6-C5H41	7.263	8.201	70401
G6H1'-C5H42	5.678	6.804	80903
G6H1'-G6H22	5.68	6.2	134199
G6H1'-A8H62	5.684	6.697	441172
G6H22-C5H41	6.201	8.199	113088
G6H22-C5H42	6.198	6.805	130296
G6H22-A7H2	6.2	7.929	195787
G6H22-A7H62	6.198	6.515	277305
G6H22-A7H8	6.206	8.266	65625
G6H22-A8H2	6.197	7.876	498011
G6H22-A8H62	6.199	6.694	185252
G6H22-A8H8	6.198	8.094	513935
G6H22-G9H8	6.198	7.798	1833334
G6H22-G10H8	6.2	8.043	272577
G6H8-C5H41	7.559	8.202	49350
A7H62-A7H2	6.513	7.929	64879
A7H62-A7H8	6.506	8.265	35485
A7H62-A8H2	6.516	7.877	96854
A8H1'-?	5.804	6.518	184034
A8H1'-G6H22	5.803	6.199	174824
A8H2'-G6H22	4.678	6.2	224860
A8H5'2-G6H22	3.965	6.203	50853
A8H62-A8H2	6.698	7.878	70141
G9H1'-A7H62	6.015	6.522	182900
U11H1'-U11H3	5.584	13.703	61284
G12H1-G2H1	12.679	13.334	521625
G12H1-U11H3	12.67	13.701	25461
G12H1'-U11H3	5.829	13.697	53736
G12H8-C13H41	7.755	8.54	61010
C13H1'-G2H1	5.479	13.334	36240
C13H1'-G12H1	5.478	12.682	97772
C13H2'-G2H1	4.195	13.329	38120
C13H41-G2H1	8.529	13.334	3133022
C13H41-G12H1	8.532	12.681	384262
C13H42-C13H41	6.912	8.532	28411958
C13H42-C13H6	6.913	7.656	282067
C13H42-C14H42	6.912	7.083	938596
C13H42-C14H6	6.912	7.492	52192
C13H5-G2H1	5.283	13.334	80021
C13H5-C13H41	5.281	8.531	1212016
C13H5-C13H42	5.284	6.912	4373447
C13H5-C14H42	5.28	7.083	205005
C13H6-G2H1	7.643	13.332	207325
C13H6-G12H1	7.655	12.677	102684
C13H6-C13H41	7.652	8.532	71924
C13H6-C14H41	7.655	8.588	58017
C14H1'-G2H1	6.111	13.332	71963
C14H42-G2H1	7.084	13.334	442577
C14H42-C14H41	7.081	8.583	14503791
C14H42-C14H6	7.081	7.492	200733
C14H5-C13H41	5.563	8.533	213601
C14H5-C13H42	5.563	6.915	179011
C14H5-C14H41	5.564	8.584	782527
C14H5-C14H42	5.565	7.083	2467294
C14H6-G2H1	7.487	13.335	275466
C14H6-C13H41	7.484	8.532	39380
C14H6-C14H41	7.485	8.582	34529

NOESY 150 ms 95%H_2O/5%D_2O

Assignment	w1	w2	Data Height
?-G2H1	7.244	13.335	514378
?-G2H1	6.912	13.334	2716646
?-G2H1	8.592	13.334	867343
?-G2H1'	7.15	5.94	1100062
?-U11H3	8.429	13.699	55827
?-U11H3	6.674	13.7	62943
?-U11H3	7.758	13.699	73353
?-U11H3	6.414	13.696	103320
?-U11H3	7.941	13.698	86989
?-U11H3	6.351	13.698	92037
?-C13H41	7.197	8.532	120617
?-C14H1'	7.154	6.115	208880
G1H8-C14H41	8.16	8.592	218944
G2H1'-G2H1	5.936	13.334	86659
G2H2'-G2H1	4.568	13.336	80204
G2H8-C3H41	7.638	8.595	101831
C3H1'-G2H1	5.553	13.334	323829
C3H1'-G12H1	5.549	12.685	84220
C3H2'-G12H1	4.515	12.685	139831
C3H41-U11H3	8.593	13.698	54083
C3H41-G12H1	8.594	12.681	5580934
C3H42-G2H8	6.911	7.639	384802
C3H42-C3H41	6.909	8.595	42478644
C3H42-C3H6	6.908	7.751	621527
C3H42-A4H2	6.909	7.415	120563
C3H42-A4H8	6.91	8.057	119947
C3H42-U11H3	6.902	13.7	58913
C3H42-G12H1	6.909	12.681	2967406
C3H5-G2H1	5.312	13.338	120153
C3H5-C3H41	5.319	8.596	2632053
C3H5-C3H42	5.319	6.911	5622916
C3H5-G12H1	5.318	12.677	121927
C3H6-G2H1	7.74	13.334	275764
C3H6-C3H41	7.745	8.593	259392
C3H6-G12H1	7.744	12.68	171007
A4H1'-G12H1	5.976	12.678	390308
A4H2-C5H41	7.412	8.199	89286
A4H2-U11H3	7.417	13.699	1414045
A4H2-G12H1	7.417	12.68	272346
A4H8-G12H1	8.052	12.684	206748
C5H41-U11H3	8.193	13.7	133260
C5H42-?	6.807	8.062	418589
C5H42-A4H2	6.803	7.414	162817
C5H42-C5H41	6.806	8.201	38652256
C5H42-C5H6	6.805	7.265	606921
C5H42-G6H8	6.804	7.558	124739
C5H42-A7H2	6.803	7.932	100239
C5H42-U11H3	6.808	13.698	63127
C5H5-C5H41	5.123	8.201	2316793
C5H5-C5H42	5.123	6.809	5594017
C5H6-C5H41	7.263	8.201	153368
G6H1'-C5H42	5.68	6.81	234144
G6H1'-G6H22	5.68	6.196	650862
G6H1'-A8H62	5.682	6.696	715873
G6H1'-G9H22	5.68	6.526	266305
G6H2'-A8H62	4.533	6.697	122515
G6H22-C5H41	6.198	8.202	212185

G6H22-C5H42	6.195	6.808	290184
G6H22-A7H2	6.197	7.928	403174
G6H22-A7H62	6.2	6.514	323764
G6H22-A7H8	6.201	8.268	103316
G6H22-A8H2	6.197	7.876	989190
G6H22-A8H62	6.197	6.695	283575
G6H22-A8H8	6.199	8.093	1018007
G6H22-G9H8	6.198	7.799	3714239
G6H22-G10H8	6.197	8.044	535393
G6H8-C5H41	7.559	8.202	90730
A7H62-A7H2	6.513	7.929	285794
A7H62-A7H8	6.515	8.266	82948
A7H62-A8H2	6.516	7.877	348130
A8H1'-G6H22	5.811	6.198	720030
A8H2'-G6H22	4.664	6.197	142764
A8H3'-G6H22	4.757	6.201	563953
A8H5'2-G6H22	3.965	6.203	72243
A8H62-G6H8	6.695	7.558	78026
A8H62-A7H8	6.698	8.268	152588
A8H62-A8H2	6.692	7.874	187390
G9H1'-A7H62	6.01	6.518	127192
G9H22-G9H8	6.538	7.795	95526
G9H22-G10H8	6.525	8.04	104117
U11H1'-U11H3	5.581	13.703	67997
U11H5-U11H3	5.148	13.698	68162
G12H1-G2H1	12.679	13.334	783865
G12H1-U11H3	12.67	13.701	47887
G12H1'-U11H3	5.83	13.697	55558
G12H8-C13H41	7.757	8.527	97411
C13H1'-G2H1	5.479	13.334	123245
C13H1'-G12H1	5.479	12.681	236072
C13H2'-G2H1	4.197	13.336	84230
C13H41-G2H1	8.529	13.334	4465796
C13H41-G12H1	8.534	12.683	553399
C13H42-C13H41	6.912	8.532	34061888
C13H42-C13H6	6.913	7.656	507482
C13H42-C14H42	6.912	7.083	1420454
C13H42-C14H6	6.912	7.492	99321
C13H5-G2H1	5.281	13.333	184066
C13H5-C13H41	5.281	8.531	1969024
C13H5-C13H42	5.284	6.912	5485721
C13H5-C14H42	5.28	7.084	304419
C13H6-G2H1	7.646	13.334	361487
C13H6-G12H1	7.655	12.677	192287
C13H6-C13H41	7.654	8.529	133628
C13H6-C14H41	7.655	8.588	132406
C14H1'-G2H1	6.112	13.332	142499
C14H42-G2H1	7.082	13.333	779092
C14H42-C13H6	7.084	7.65	168222
C14H42-C14H41	7.081	8.583	20202284
C14H42-C14H6	7.081	7.491	526425
C14H5-C13H41	5.564	8.531	352241
C14H5-C13H42	5.561	6.912	288910
C14H5-C14H41	5.564	8.584	1427902
C14H5-C14H42	5.565	7.083	3731812
C14H6-G2H1	7.487	13.335	423713
C14H6-C13H41	7.484	8.532	64588
C14H6-C14H41	7.485	8.582	96837

7.5.4 Violations

Dihedral violations

Dexp= 60.000 Dave= 100.869 Viol= -10.869 #viol= 20 6 GUA O4' 6 GUA C1' 6 GUA N9 6 GUA C4
Dexp= 200.000 Dave= -111.475 Viol= -8.525 #viol= 20 8 ADE O4' 8 ADE C1' 8 ADE N9 8 ADE C4
Dexp= 151.000 Dave= -143.788 Viol= -5.212 #viol= 3 8 ADE P 8 ADE O5' 8 ADE C5' 8 ADE C4'

Distance violations

Rexp= 6.500 Rave= 6.826 Viol= -0.326 #viol= 2 (5 CYT H2') (7 ADE H8)
Rexp= 6.500 Rave= 6.815 Viol= -0.315 #viol= 1 (7 ADE H5'2) (6 GUA H1')
Rexp= 6.500 Rave= 6.812 Viol= -0.312 #viol= 1 (7 ADE H1') (8 ADE H2)

Acknowledgements

Publications

B. Fürtig, **R. Schnieders**, C. Richter, H. Zetzsche, S. Keyhani, C. Helmling, H. Kovacs and H. Schwalbe (2016). Direct ^{13}C-detected NMR experiments for mapping and characterization of hydrogen bonds in RNA, *J. Biomol. NMR*, 64(3):207-21.

R. Schnieders, C. Richter, S. Warhaut, V. de Jesus, S. Keyhani, E. Duchardt-Ferner, H. Keller, J. Wöhnert, L. T. Kuhn, A. L. Breeze, W. Bermel, H. Schwalbe and B. Fürtig (2017). Evaluation of ^{15}N-detected H – N correlation experiments on increasingly large RNAs, *J. Biomol. NMR*, 69(1):31-44.

A. C. Wolter, A. Pianu, J. Kremser, E. Strebitzer, **R. Schnieders**, B. Fürtig, C.Kreutz, E. Duchardt-Ferner and J. Wöhnert (2019). NMR resonance assignments for the GTP-binding RNA aptamer 9-12 in complex with GTP, *Biomol. NMR Assign.*, 13:281-86.

R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe and B. Fürtig (2019). Novel ^{13}C-detected NMR experiments for the precise detection of RNA structure, *Angew. Chemie – Int. Ed.*, 58:9140-44.

R. Schnieders, S. Keyhani, H. Schwalbe and B. Fürtig (2020). More than proton detection—New avenues for NMR spectroscopy of RNA, *Chem. Eur. J.*, 26:102-13.

R. Schnieders, B. Knezic, H. Zetsche, A. Sudakov, T. Matzel, C. Richter, M. Hengesbach, H. Schwalbe and B. Fürtig (2020). NMR spectroscopy of large functional RNAs - from sample preparation to low-gamma-detection, *Manuscript submitted*.

O. Binas, V. de Jesus, T. Landgraf, A. E. Völklein, J. Bains, H. Berg, T. Biedenbänder, D. Hymon, J. Martins, A. Niesteruk, A. Oxenfarth, N. S. Qureshi, T. Schamber, **R. Schnieders**, A. Tröster, A. Wacker, J. Wirmer-Bartoschek, M. Wirtz, K. Azzaoui, C. Richter, S. Sreeramulu, M. J. J. Blommers and H. Schwalbe (2020). ^{19}F-NMR-based fragment screening of biologically active RNA, *Manuscript submitted*.

A. Wacker, J. E. Weigand, S. R. Akabayov, N. Altincekik, J. K. Bains, ……, **R. Schnieders**, H-Schwalbe*, A. Simba-Lahuasi, S. Sreeramulu, E. Stirnal, A. Sudakov, J.-N. Tants, B. S. Tolbert, J. Vögele, L. Weiß, J. Wirmer-Bartoschek, M. Wirtz Martin, J. Wöhnert, H. Zetzsche (2020). Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy and DMS footprinting analysis, *Manuscript submitted*.

N. S. Qureshi, T. Matzel, E. Cetiner, **R. Schnieders**, H. R. A. Jonker, B. Fürtig and H. Schwalbe (2020). NMR structure of D3-D4 domains of *Vibrio vulnificus* ribosomal protein S1 provides insights into molecular recognition of single-stranded RNAs, *Manuscript in preparation*.

Conference contributions

R. Schnieders, G. Pintér, H. Jonker, L. Stelzl, G. Hummer and H. Schwalbe, NMR structure determination of RNA tetraloops as reference for MD force field reparametrization, **Poster presentation**, EUROMAR, Warsaw, Poland, July 2017.

R. Schnieders, G. Pintér, H. Jonker, L. Stelzl, G. Hummer and H. Schwalbe, NMR structure determination of RNA tetraloops as reference for MD force field reparametrization, **Poster presentation**, iNEXT Workshop: Structural biology approaches for drug development, Budapest, Hungary, September 2018.

R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe and B. Fürtig, Development of heteronuclear-detected NMR-experiments for characterization of NH- and NH_2-groups in RNA, **Oral presentation**, Experimental Nuclear Magnetic Resonance Conference (ENC), Pacific Grove, CA, USA, April 2019.

R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe and B. Fürtig, Development of heteronuclear-detected NMR-experiments for characterization of NH- and NH_2-groups in RNA, **Oral presentation**, Gordon Research Conference on Computational Aspects of Biomolecular NMR: From Advances of Pulse Sequence Design to the Molecular Movie by Biomolecular NMR, Les Diablerets, Switzerland, June 2019.

R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe and B. Fürtig, Development of heteronuclear-detected NMR-experiments for characterization of NH- and NH_2-groups in RNA, **Oral presentation**, 11th BMRZ Symposium, Bad Homburg, Germany, August 2019.

R. Schnieders, A. C. Wolter, C. Richter, J. Wöhnert, H. Schwalbe and B. Fürtig, Development of heteronuclear-detected NMR-experiments for characterization of NH- and NH_2-groups in RNA, **Poster presentation**, EUROISMAR, Berlin, Germany, August 2019.

Curriculum vitae