

# Optimisation and development of NMR spectroscopic methods for the investigation of biomolecular dynamics

Dissertation  
zur Erlangung des Doktorgrades  
der Naturwissenschaften

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

von  
**György Pintér**  
aus Szeged, Ungarn

Frankfurt, 2020

(D30)

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

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Prof. Dr. Harald Schwalbe  
Prof. Dr. Clemens Glaubitz

Datum der Disputation: 24.03.2021

Családomnak





|                            |    |
|----------------------------|----|
| ZUSAMMENFASSUNG .....      | 1  |
| SUMMARY .....              | 9  |
| LIST OF ABBREVIATIONS..... | 15 |

## **CHAPTER I. RADIO FREQUENCY HEATING AND NMR SPECTROSCOPY**

### **COMBINED TO STUDY PROTEIN FOLDING..... 17**

|  |           |
|--|-----------|
| <b>1. INTRODUCTION.....</b>  | <b>19</b> |
| 1.1. PROTEIN FOLDING.....  | 19        |
| 1.2. CHARACTERIZING PROTEIN FOLDING AND DYNAMICS BY NMR SPECTROSCOPY .....                                   | 20        |
| 1.2.1. Magnetization transfer methods.....   | 20        |
| 1.2.2. Real time NMR methods .....   | 21        |
| 1.3. TEMPERATURE JUMP SYSTEMS FOR NMR SPECTROMETER .....   | 21        |
| 1.4. MODELL PROTEIN: BARSTAR .....   | 23        |
| 1.4.1. Structure.....  | 23        |
| 1.4.2. Binding site.....   | 24        |
| 1.4.3. Overview of folding mechanism.....  | 25        |
| 1.4.4. Cold denaturation phenomena .....   | 29        |
| 1.4.5. Cold denaturation of barstar .....  | 30        |
| <b>2. MATERIALS AND METHODS .....</b>  | <b>32</b> |
| 2.1. PROTEIN PRODUCTION AND PURIFICATION.....  | 32        |
| 2.1.1. Synthesis of plasmid DNA .....  | 32        |
| 2.1.2. Transformation .....  | 32        |
| 2.1.3. Plasmid DNA purification .....  | 32        |
| 2.1.4. Site-directed mutagenesis .....   | 33        |
| 2.1.5. Protein production .....  | 34        |
| 2.2. NMR EXPERIMENTS FOR ASSIGNMENT.....   | 35        |
| 2.2.1. Chemical shift assignment of HN, C <sub>α</sub> , C <sub>β</sub> , CO in folded barstar .....         | 35        |
| 2.2.2. Chemical shift assignment of HN, C <sub>α</sub> , C <sub>β</sub> , CO in cold denatured barstar ..... | 35        |
| 2.3. T-JUMP NMR EXPERIMENT SETUP .....   | 36        |
| 2.3.1. Heating setup for T-jump experiments .....  | 36        |
| 2.3.2. Internal temperature sensor.....  | 37        |
| 2.3.3. Setting up cycling-2D NMR measurements .....  | 37        |
| 2.3.4. Heating cycles for equilibrium state .....  | 38        |
| 2.3.5. Heating cycles for fast folding pathway .....   | 38        |
| <b>3. RESULTS AND DISCUSSION .....</b>   | <b>39</b> |
| 3.1. MOTIVATION.....   | 39        |

|   |           |
|---|-----------|
| 3.2. OPTIMIZED T-JUMP SETUP AND SUITABLE MODEL PROTEIN .....              | 39        |
| 3.2.1. Protein sequence optimization for cycling T-jump experiments ..... | 40        |
| 3.2.2. Buffer optimization for heating efficiency.....                    | 40        |
| 3.2.3. Improvement of the duty-cycle of the radio frequency heating ..... | 41        |
| 3.3. T-JUMP EXPERIMENTS WITH ONE-DIMENSIONAL RT-NMR .....                 | 41        |
| 3.4. CYCLING T-JUMP EXPERIMENT TO MEASURE TWO-DIMENSIONAL SPECTRUM .....  | 42        |
| 3.5. COLD DENATURED STATE - CONFORMATIONAL HETEROGENEITY .....            | 43        |
| 3.6. SLOW PATHWAY – RATE LIMITING STEP .....                              | 44        |
| 3.6.1. Time resolved intensity changes .....                              | 44        |
| 3.7. INFLUENCE OF ISOMERIZATION ON THE FOLDING PATHWAY .....              | 47        |
| 3.8. FAST FOLDING PATHWAY .....   | 50        |
| 3.8.1. Heating cycles for double-jump experiment .....                    | 51        |
| 3.8.2. State-correlated spectroscopy pulse sequence .....                 | 52        |
| 3.8.3. State-correlated spectroscopy .....                                | 54        |
| 3.8.4. Intermediate on the fast folding pathway.....                      | 56        |
| <b>4. SUMMARY AND OUTLOOK .....</b>                                       | <b>59</b> |

**CHAPTER II. EFFECT OF PHOTOCAGED NUCLEOTIDE ON THE DNA DOUBLE  
HELIX.....61**

|   |           |
|---|-----------|
| <b>1. INTRODUCTION .....</b>  | <b>63</b> |
| 1.1. PHOTO LABILE PROTECTING GROUPS .....                                       | 63        |
| 1.2. CHARACTERIZING NUCLEIC ACID STABILITY .....                                | 64        |
| 1.2.1. Global stability characterized by melting temperature .....              | 64        |
| 1.2.2. Local base pair stability characterized by NMR spectroscopy .....        | 65        |
| 1.3. DETERMINATION OF ABSOLUTE OR RELATIVE CONFIGURATION OF A CHIRAL CENTRE.... | 65        |
| <b>2. MATERIALS AND METHODS.....</b>  | <b>67</b> |
| 2.1. SYNTHESIS OF PHOTOLABILE PROTECTED NUCLEOTIDES AND DNA .....               | 67        |
| 2.2. PREPARATION OF DOUBLE STRANDED DNA.....                                    | 67        |
| 2.3. CD SPECTRA AND MELTING POINT .....   | 67        |
| 2.4. MELTING POINT.....   | 68        |
| 2.5. NMR MEASUREMENTS.....  | 68        |
| 2.6. INVERSION RECOVERY EXPERIMENT.....   | 69        |
| 2.6.1. Sample preparation .....   | 69        |
| 2.6.2. Pulse sequence and experiment setup .....                                | 69        |
| 2.6.3. Analyses of inversion recovery experiment.....                           | 69        |
| <b>3. RESULTS .....</b>   | <b>71</b> |

|  |           |
|--|-----------|
| 3.1. DNA SEQUENCE DESIGN .....   | 71        |
| 3.2. EFFECT OF CAGES ON THE SECONDARY STRUCTURE AND MELTING POINT .....  | 71        |
| 3.3. SIGNAL ASSIGNMENT AND DETERMINATION OF ABSOLUTE CONFIGURATION ..... | 72        |
| 3.4. LOCALIZED EFFECT OF CAGE .....                                      | 75        |
| 3.5. STRUCTURE CALCULATION.....  | 77        |
| 3.6. MEDIUM THROUGH PUT SCREENING .....                                  | 78        |
| 3.7. MD SIMULATIONS.....   | 79        |
| 3.8. MD SIMULATIONS DATA COMPARED TO NMR DATA .....                      | 80        |
| <b>4. DISCUSSION – LIMITS OF SINGLE CAGE APPROACH .....</b>              | <b>82</b> |
| 4.1. THE EFFECT OF SIZE AND RELATIVE CONFIGURATION .....                 | 82        |
| 4.2. LIMITS OF SINGLE CAGE APPROACH, ALTERNATIVE WAYS TO REGULATE.....   | 86        |
| <b>5. SUMMARY AND OUTLOOK.....</b>                                       | <b>87</b> |

### **CHAPTER III. NMR SPECTROSCOPIC CHARACTERIZATION OF THE GCAA**

#### **RNA TETRALOOP..... 89**

|  |            |
|--|------------|
| <b>1. INTRODUCTION.....</b>  | <b>91</b>  |
| 1.1. MD SIMULATIONS OF RNA.....  | 91         |
| 1.2. SELECTING RNA TETRALOOPS AS REFERENCE SYSTEM .....                                | 92         |
| 1.3. DYNAMIC SYSTEMS AS REFERENCE .....  | 94         |
| <b>2. MATERIALS AND METHODS .....</b>  | <b>97</b>  |
| 2.1. TEMPLATE DNA PRODUCTION .....   | 97         |
| 2.1.1. Sequence design.....  | 97         |
| 2.1.2. Transformation .....  | 97         |
| 2.1.3. Plasmid amplification .....   | 97         |
| 2.1.4. Plasmid purification.....   | 98         |
| 2.1.5. Plasmid linearization .....   | 98         |
| 2.2. RNA PRODUCTION UTILIZING <i>IN VITRO</i> TRANSCRIPTION WITH T7 RNA POLYMERASE.... | 99         |
| 2.3. RNA PURIFICATION .....  | 100        |
| 2.4. NMR SPECTROSCOPY .....  | 102        |
| <b>3. RESULTS AND DISCUSSION .....</b>   | <b>103</b> |
| 3.1. CHEMICAL SHIFT ASSIGNMENT .....   | 103        |
| 3.1.1. Sequential connection through the phosphate backbone .....                      | 104        |
| 3.1.2. Assignment of ribose signals.....   | 104        |
| 3.1.3. Assignment of <sup>15</sup> N signals .....                                     | 106        |
| 3.1.4. Stereoselective assignment of proS and proR of H5' and H5'' signals.....        | 108        |
| 3.2. OBSERVATION OF TEMPERATURE DEPENDANT DYNAMICS.....                                | 109        |

|  |            |
|--|------------|
| 3.2.1. Canonical coordinates .....   | 110        |
| 3.3. TORSION ANGLES .....  | 112        |
| 3.3.1. Sugar puckering from homonuclear scalar coupling constants.....   | 113        |
| 3.3.2. Phosphate backbone relevant scalar couplings and cross-correlated relaxation rates..                        | 114        |
| 3.4. LONGITUDINAL RELAXATION AND HETERONUCLEAR NOE MEASUREMENTS .....  | 116        |
| <b>4. SUMMARY AND OUTLOOK .....</b>  | <b>118</b> |
| <b>REFERENCES .....</b>  | <b>121</b> |
| <b>ACKNOWLEDGEMENT .....</b>   | <b>135</b> |
| <b>APPENDIX .....</b>  | <b>137</b> |
| <b>1. CHEMICAL SHIFT ASSIGNMENT .....</b>  | <b>137</b> |
| 1.1. BARSTAR AT 303 K FOLDED AND AT 270 K UNFOLDED STATE .....   | 137        |
| 1.2. ( <i>R</i> )-NPPM AND ( <i>S</i> )-NPPM MODIFIED DNA DOUBLE STRAND.....                                       | 139        |
| 1.3. RESONANCE ASSIGNMENT OF GCAA 14NT TETRALOOP.....  | 144        |
| <b>2. PULSE SEQUENCES AND AU PROGRAMS.....</b>   | <b>146</b> |
| 2.1. AU PROGRAM TO SET UP CYCLING TWO DIMENSIONAL T-JUMP EXPERIMENT.....   | 146        |
| 2.2. PULSE SEQUENCE FOR STATE-CORRELATED T-JUMP EXPERIMENT .....   | 147        |
| 2.3. AU PROGRAM TO CONCATENATE TOGETHER TWO-DIMENSIONAL SPECTRUMS FROM<br>SERIES OF CYCLED T-JUMP EXPERIMENTS..... | 155        |
| <b>DECLARATIONS.....</b>   | <b>157</b> |
| <b>PUBLICATIONS .....</b>  | <b>161</b> |
| <b>CONFERENCE CONTRIBUTIONS.....</b>   | <b>161</b> |
| <b>CURRICULUM VITAE .....</b>  | <b>163</b> |

## ZUSAMMENFASSUNG

Die Kernspinresonanz-Spektroskopie (NMR-Spektroskopie) als biophysikalische Methode ermöglicht es, zeitabhängige Strukturänderungen von Biomolekülen, von vollständiger Rückfaltung bis hin zu lokalen Seitenkettenbewegungen, mit seitenspezifischer Auflösung zu charakterisieren. Sie bietet eine Vielzahl von Werkzeugen zur Untersuchung der inter- und intramolekularen Bewegungen von Biomolekülen auf verschiedenen Zeitskalen unter vergleichsweise nativen Bedingungen.

Methoden zur Charakterisierung der Dynamik von Biomolekülen durch NMR-Spektroskopie können in zwei Hauptkategorien unterteilt werden: Magnetisierungstransfer-basierte NMR-Spektroskopie und Echtzeit-NMR-Spektroskopie (RT-NMR). Erstere nutzt den Magnetisierungstransfer zwischen dem Grundzustand und sogenannten versteckten oder angeregten Zuständen von Biomolekülen, während letztere externe Trigger verwendet, um Biomoleküle aus dem thermodynamischen Gleichgewicht zu bringen, um anschließend die Relaxation zurück in ein thermodynamisches Gleichgewicht zu verfolgen.

Das erste Kapitel meiner Dissertation befasst sich mit der Entwicklung einer neuen RT-NMR-Methode zur Untersuchung der Proteinfaltung mittels rascher Temperaturänderung. Dabei wird in einem dafür entwickelten Probenkopf durch Anlegen einer Hochfrequenzstrahlung (200-220 MHz) an eine Probe mit hoher Salzkonzentration im NMR-Spektrometer ein Temperatursprung von 10-20 °C in 250-500 ms erreicht. Kalt-denaturierung, ein bekanntes Phänomen globulärer Proteine als Folge der hohen Temperaturabhängigkeit von Enthalpie und Entropie, ermöglicht in einigen Fällen die Entfaltung von Proteinen bei Temperaturen nahe oder oberhalb des Gefrierpunkts von Wasser. Barstar, das Inhibitorprotein der Barnase, ist ein gut untersuchtes Modellsystem zur Untersuchung der Proteinfaltung. Es kann leicht zwischen 0 und 5 °C mittels chaotroper Verbindungen wie 2-2,5 M Harnstoff kalt denaturiert werden, wobei bei 25 bis 30 °C unter denselben Bedingungen das Protein im gefalteten Zustand vorliegt. Für die Durchführung der Faltungsexperimente habe ich eine Cystein-freie Mutante erzeugt, welche ebenfalls in früheren Faltungsstudien verwendet wurde, um die Ausbildung nicht nativer Disulfidbrücken und einhergehende Fehlfaltungen zu vermeiden. Ich konnte zeigen, dass diese Mutante mehrere hundert Male ohne Abbau oder Aggregation denaturiert und rückgefaltet werden kann

Im NMR Spektrometer wird durch den Temperatursprung die Rückfaltung des kalt denaturierten Barstars initiiert, welche mittels RT-NMR verfolgt werden kann. Zusätzlich wurde in beiden Zuständen (denaturiert und gefaltet) eine Proteinrückgratresonanzzuweisung durchgeführt, um die Änderungen während der Rückfaltung auf einer seitenspezifischen Ebene verfolgen zu können. Erste kinetische Experimente unter Verwendung von Protonen-NMR zeigten langsame Änderungen der Signal-Intensität. Jedoch war die spektrale Auflösung – aufgrund starker Signalüberlappung und gleichzeitig vorhandener gefalteter und ungefalteter Zustände sowie möglicher Zwischenzustände – für detailliertere Analysen unzureichend.

Ich habe eine neue Methode entwickelt, die die reversible Rückfaltung von Barstar nutzt, um zweidimensionale NMR-Experimente durch Wiederholung von Temperatursprung-Experimenten (T-Sprung) aufzuzeichnen. Das Wiederholen der kinetischen Experimente mit unterschiedlichen Phasenzyklen und Evolutionszeitinkrementen für die indirekte Dimension in einer HSQC-Pulssequenz sowie addieren kinetischer Experimente – durch Verkettung derselben Zeitpunkte – ermöglicht die Aufzeichnung einer Zeitreihe von zweidimensionalen HSQC-Spektren. Diese Reihe von aufgenommenen Spektren hat die gleiche Zeitauflösung wie das kinetische Standard-Protonen-RT-NMR-Experiment, aber dank der zweiten Dimension können weitaus mehr seitenspezifische Informationen extrahiert werden.

Barstar hat zwei Prolin-Reste und von diesen beiden befindet sich die Tyr47-Pro48-Amidbindung im gefalteten Zustand in der seltenen *cis* Konformation. Durch NMR Experimente, welche selektiv Prolin-Reste detektieren, konnte ich jedoch zeigen, dass die Tyr47-Pro48-Amidbindung im ungefalteten Zustand hauptsächlich in *trans* vorliegt. In kinetischen Experimenten zeigten die zeitabhängigen Signalintensitätsänderungen, dass mehrere Reste entweder mono- oder doppelt exponentielles Verhalten im Zusammenhang mit der *trans-cis*-Isomerisierung aufweisen. Indem der Effekt der *trans-cis*-Isomerisierung in einem kryo-gekühlten Probenkopf mit höherem Probenvolumen verfolgt wurde, konnte auf dem Faltungsweg dank des höheren Signal-zu-Rausch-Verhältnisses ein stabiles Zwischenprodukt nachgewiesen werden. Diese Signale bildeten sich bereits in der Totzeit des Experiments und zeigten weiterhin einen monoexponentiellen Abfall. Die Signale des Zwischenprodukts konnten nicht zu einzelnen Resten zugeordnet werden, liegen jedoch in dem typischen chemischen Verschiebungsbereich von Amidien gefalteter Proteine.

Seitenspezifische Analysen der langsamen und schnellen kinetischen Veränderungen in Kombination mit in der Literatur bekannten Informationen ermöglichten die Skizzierung eines denkbaren Faltungsmechanismus. Die schnelle kinetische Veränderung kann dabei der Bildung der Sekundärstruktur zugeschrieben werden. Die langsame Kinetik resultiert aus der Bildung der finalen gefalteten Tertiärstruktur, welche durch die *trans-cis*-Isomerisierung der Tyr47-Pro48-Amidbindung verlangsamt wird. Der beobachtete Mangel an Intensitätsänderungen um die Bindungsstelle herum stimmt ebenfalls mit Literaturergebnissen überein, aus denen hervorgeht, dass das frühe Faltungsintermediat bereits aktiv ist und Barnase hemmen kann.

Weitere alternative, schnelle Faltungswege konnten durch Vermeidung des geschwindigkeitsbegrenzenden Schritts der *trans-cis*-Isomerisierung untersucht werden. Dazu wurde der Heizungszyklus geändert, sodass die Abkühlzeit so kurz wie möglich gehalten wurde und somit der Temperatursprung und die Initiierung der erneuten Faltung aus einem entfalteten Zustand – in dem die Tyr47-Pro48-Amidbindung größtenteils in der *cis*-Konformation vorliegt – stattfinden konnte. Dieser alternative Faltungsweg ist weitaus schneller, sodass bereits im ersten HSQC-Spektrum nach dem Temperatursprung ein vollständig gefalteter Zustand beobachtet wurde. Daher war es nicht möglich, die Faltung durch die zweidimensionale Temperatursprung-Methode zu verfolgen. Ich entwarf daher eine neue Pulssequenz basierend auf zustandskorrelierten und ZZ-Austausch Experimenten, um den entfalteten Zustand mit dem gefalteten Zustand zu korrelieren. In dieser neuen Pulssequenz findet die indirekte Dimensionsentwicklung noch bei der niedrigeren Temperatur, bei der Barstar im ungefalteten Zustand vorliegt, statt. Nach dem Temperatursprung wird unmittelbar vor der Detektion eine zusätzliche variable Zeit  $\tau$  eingeführt, in der sich das Protein weiter falten kann. Die Wiederholung des zustandskorrelierten Experiments mit unterschiedlichen Rückfaltungszeiten ermöglicht theoretisch die Extraktion kinetischer Informationen über den gesamten Faltungsprozess hinweg. Aufgrund der longitudinalen Relaxationszeit ist jedoch die Beobachtung des Faltungsprozesses mit dieser Methode auf einen engen Zeitraum begrenzt. Trotzdem konnte mit diesem zustandskorrelierten Experiment eine große Menge quantitativer Informationen über den schnellen Faltungsweg gewonnen werden. Mit diesem Experiment konnten unter anderem auch Zwischenprodukte auf dem schnellen Faltungsweg nachgewiesen werden. Darüber hinaus gibt es Austauschpeaks, die bei kurzer (5 ms) oder langer (200 ms) Rückfaltungszeit nicht nachweisbar sind,

während einige Reste auch bei einer Rückfaltungszeit von 200 ms noch deutliche entfaltete Peaks zeigen. Dies liefert den Beweis, dass verschiedene Teile von Barstar nicht parallel falten, sondern die Faltung schrittweise erfolgt.

Die Kombination von reversibler Kalt-denaturierung mit schnellem Temperatursprung innerhalb des NMR-Spektrometers ermöglichte die Untersuchung des Faltungsmechanismus von Barstar in atomarer Auflösung. Damit konnte ein wertvoller Einblick in die strukturellen Veränderungen von Barstar, die mit den verschiedenen Faltschritten auf zwei verschiedenen Faltungswegen einhergehen, gewonnen werden.

Zusammen mit den in den letzten Jahren beschriebenen Drucksprung-Echtzeit-NMR-Methoden zeigen die in dieser Arbeit mittels Temperatursprung gewonnenen Ergebnisse das hohe Potential dieser Methoden, um Proteinfaltung auf atomarer Ebene experimentell zu untersuchen und ein tieferes Verständnis von verschiedenen Faltungswegen zu erlangen.

In Kapitel II fasse ich meine Arbeit und Ergebnisse zusammen, die in einem Verbundprojekt im Rahmen des DFG Graduiertenprogramms CLiC erzielt wurden. Dieses Verbundprojekt hatte dabei zum Ziel, das mögliche Potenzial des Einzelnukleobasen-Cagings zu untersuchen, um die Helixbildung der Sekundärstruktur in Nukleinsäuren zu stören.

In meinem Projektteil untersuchte ich die 15-Basenpaare-lange doppelsträngige DNA, die am mittleren Cytidin (dC8) mit einer geänderten Schutzgruppe („cage“) vom NPE-Typ modifiziert war, wobei in  $\alpha$ -Position eine Phenylsubstitution angewendet wurde, was zu zwei Produkten – zu einem mit (*R*)-NPPM und zu einem anderen mit (*S*)-NPPM Schutzgruppe – führte. Die NMR-Untersuchung beschränkte sich dabei auf protonendetektierte Experimente, wie 2D-H,H-NOESY und 2D-H,H-TOCSY. Nachdem eine Zuordnung für mehr als 80% der Resonanzen erreicht war, wurde die relative Konfiguration des Käfigs bestimmt. Entsprechend der *R*- und *S*-Konfiguration der Schutzgruppe wurden basierend auf der zuvor bestimmten Struktur zwei Modellstrukturen erstellt, die die identische DNA-Sequenz hatten und in der gleichen Position mit NPE-Schutzgruppe modifiziert waren. Es wurden unterschiedliche NOE-Korrelationsmuster zwischen den Signalen der (*R*)-NPPM- und (*S*)-NPPM-Schutzgruppen und der DNA beobachtet. Diese Modelle lieferten mögliche beobachtbare NOE-Kontakte zwischen der Schutzgruppe und den DNA-Atomen, die sich zwischen



der *R*- und *S*-Konfiguration unterscheiden konnten. Im 2D-H,H-NOESY-Spektrum wurden einige dieser Kontakte identifiziert, die die Bestimmung der relativen Konfiguration der Schutzgruppen in den Proben ermöglichten. Ein Vergleich der Schmelzpunkte von (*R*)-NPPM-, (*S*)-NPPM-, (*R*)-NPE- und (*S*)-NPE-modifizierter DNA zeigte, dass die *S*-Konfiguration beider Schutzgruppen (NPPM und NPE) zu einer größeren Reduktion des Schmelzpunktes führt, obwohl für die NPPM-Schutzgruppe der Unterschied zwischen der *R*- und der *S*-Konfiguration geringer war. Dies ist höchstwahrscheinlich auf den geringeren Größenunterschied zwischen den Phenyl- und Nitrophenylgruppen zurückzuführen.

Um mögliche Änderungen in der DNA-Struktur sichtbar zu machen, wurde ein NMR-basiertes Modell mit der Schutzgruppe berechnet. Ohne ausreichende Einschränkungen (restraints) für die unabhängige Strukturberechnung wurden zusätzliche Einschränkungen der Diederwinkel und Wasserstoffbrücken für Basenpaare eingeführt. Diese zusätzlichen Einschränkungen galten für das modifizierte Basenpaar, für die zwei benachbarten Basenpaaren – sowohl in der 3'- als auch in der 5'-Richtung – sowie für das erste und letzte Basenpaar der DNA-Sequenz nicht. NOE-Einschränkungen für labile Protonen wurden in H<sub>2</sub>O-Puffer gemessen. Für nicht-labile Protonen wurden diese in D<sub>2</sub>O-Puffer gemessen. Um die Spin-Diffusion zu kompensieren, wurden zusätzlich verschiedene Mischzeiten für nicht-labile Protonen aufgezeichnet. Die Struktur zeigte, dass die NPPM-Schutzgruppe, obwohl sie im Vergleich zu der NPE-Schutzgruppe einen höheren sterischen Anspruch hat, immer noch in die große Furche passt und die Nukleobase nicht dazu zwingt, in das Lösungsmittel einzutauchen. Die Rückgrat-Winkel haben sich jedoch gegenüber den Standard-B-Helixwerten geändert, um die Modifikation zu berücksichtigen.

Weiterhin habe ich Inversion-Recovery-Experimente verwendet, um festzustellen, wie weit entfernt von der Einbaustelle der Schutzgruppe die Basenpaarstabilität beeinflusst wird. Dazu wurde mithilfe von diesen Inversion-Recovery-Experimenten die Austauschrate der labilen Imino-Protonen mit dem umgebenden Wasser bestimmt, welche ein direkter Indikator für die Basenpaarstabilität ist. Bei geringerer Stabilität wird die Wasserstoffbrückenbindung zwischen den Nukleobasen geschwächt und die Wahrscheinlichkeit eines Protonaustauschs mit dem Wasser erhöht. Im Vergleich zur nicht-modifizierten DNA-Helix wurde eine erhöhte Austauschrate bis zu zwei benachbarten Basenpaaren sowohl in der 3'- als auch in der 5'-Richtung beobachtet. Der

Unterschied zwischen den Austauschraten der geschützten und der nativen Helix ist direkt an der Modifikationsstelle am höchsten, während der Unterschied mit Entfernung der Reste schnell abnimmt. Dies erlaubt daher die Verwendung zusätzlicher Einschränkungen bei der Strukturberechnung für Basen, die von der Schutzgruppe nicht betroffen sind. Jedoch zeigt dies ebenfalls, dass die Schutzgruppe trotz des größeren sterischen Anspruchs immer noch nur Auswirkung auf Basenpaare in unmittelbarer Nähe hat. Dies legt eine Grenze für die maximale Wirkung nahe, die durch eine einzelne Schutzgruppe in einem DNA-Doppelstrang erreicht werden kann.

Durch Screening der Schmelztemperatur verschiedener Modifikationen – welche mittels Klick-Chemie in die DNA eingeführt wurden – und durch die Analyse mittels Moleküldynamik-Simulationen konnten stabilisierende Wechselwirkungen zwischen DNA und bestimmten Schutzgruppen charakterisiert werden. Dabei deckten sich die Simulationen mit den zuvor bei der NMR-Untersuchung der mit einer Schutzgruppe modifizierten DNA beobachteten Unterschieden.

Im Rahmen dieses interdisziplinären Forschungsprojektes konnten allgemeine Begrenzungen des Einzelschutzgruppenansatzes erörtert und generelle Designprinzipien zum Erreichen einer höheren Destabilisierung der DNA für photolabile Schutzgruppen – wie zum Beispiel die Präferenz der *S*-Konfiguration – aufgezeigt werden.

Das dritte Kapitel befasst sich mit der detaillierten strukturellen und dynamischen Charakterisierung von nicht-kanonisch gefalteten RNA Molekülen mittels NMR. Präzise und detaillierte experimentelle Daten über die Struktur und Dynamik solcher Moleküle sind – um sie für Moleküldynamik-Simulationen als Modellsysteme einsetzen zu können – durch die bislang geringe Anzahl vorhandener detaillierter experimenteller Daten von großem Interesse.

Ein ideales Modellsystem – wofür die Struktur und Dynamik mittels NMR experimentell bestimmt werden soll – sollte relativ klein sowie vorzugsweise biologisch relevant sein und dynamisches Verhalten aufweisen. Zu den häufigsten nicht-kanonischen Sekundärstrukturelemente gehören RNA-Tetra-loops. Für die Charakterisierung mittels NMR wurden die Sequenzen GCAA, CUUG und GAAG ausgewählt. Der Fokus meiner Arbeit lag dabei auf dem GCAA-Tetra-loop.

Nach der Optimierung der *In-vitro*-Transkription, um  $^{13}\text{C}$ ,  $^{15}\text{N}$ -Proben mit einheitlicher Isotopenmarkierung herzustellen, habe ich eine vollständige Resonanzzuweisung

durchgeführt. Dazu habe ich verschiedene dreidimensionale Experimente (3D-HCP und 3D-HCP-TOCSY) verwendet, um sequentielle Verbindungen entlang des Phosphat-Zucker-Rückgrats herzustellen. Die Ribosesignale wurden hauptsächlich basierend auf einem 3D-HCC-TOCSY-CCH-COSY-Experiment zugewiesen. Die C6H6 / C8H8- und N3 / N9-Signale wurden unter Verwendung einer zweidimensionalen Version von H(C)N-Spektren zugeordnet.

Nach vollständiger Zuordnung der Peaks im  $^{13}\text{C}$ -HSQC-Spektrum wurden Temperaturreihen von  $^{13}\text{C}$ -HSQC-Spektren gemessen, die temperaturabhängige chemische Verschiebungsänderungen der Ribosesignale sowohl in der Schleife als auch im schleifenschließenden Basenpaar zeigten. Die berechneten kanonischen Koordinaten deuteten bei mehreren Temperaturen auf temperaturabhängige Konformationsänderungen der Ribose hin. Zusätzliche Informationen über die Zuckerkonformation wurden durch Messen von  $^3J_{\text{HH}}$ -Skalarkopplungen unter Verwendung des 3D-HCC-TOCSY-CCH-E.COSY-Experiments gesammelt. Dies zeigte für die Ribose des Stamms solche Werte, die für die kanonische A-Form-Helix mit 3'-Endokonformation typisch sind. Für die Schleifennukleotiden fielen die Kopplungskonstanten jedoch zwischen typischen 3'-Endo- und 2'-Endo-Werten, was auf einen Konformationsaustausch zwischen den beiden Zuständen hindeutet.

Das RNA-Modellsystem wurde weiter charakterisiert, indem zusätzliche skalare Kopplungen des Phosphatrückgrats gemessen wurden. Zusätzlich wurden Relaxationsmessungen von  $R_1$ - und HetNOE-Messungen von C1'- und aromatischen Kernen durchgeführt. Es konnten mit diesen Experimenten weitere wertvolle Informationen über den GCAA-Tetraloop gewonnen werden.

Die erhaltenen Ergebnisse stimmen mit den verfügbaren Literaturinformationen zu diesem Tetraloop überein und legen die Existenz angeregter Zustände der Schleife nahe. Dieses dynamische Verhalten ist stark temperaturabhängig, was den GCAA-Tetraloop zu einem wertvollen Referenzsystem für MD-Simulationen macht. Es sind jedoch weitere experimentelle Untersuchungen erforderlich, um die für den GCAA-Tetraloop beobachtete temperaturabhängige Dynamik vollständig zu charakterisieren.



## SUMMARY

NMR spectroscopy has a great advantage, as a structural biology tool, to be able to characterize time dependant structural changes ranging from complete refolding to local side chain movements with site specific resolution. It offers a wide variety of tools to investigate the movements of biomolecules on different timescales in close to native environment.

Methods to characterise dynamic behaviour by NMR spectroscopy can be divided into two major categories - magnetization transfer based NMR spectroscopy and real-time NMR spectroscopy (RT-NMR). The former utilizes magnetization transfer between ground state and so called hidden or excited states of biomolecules, while the latter utilizes external triggers to rapidly move biomolecules out of equilibrium and follows their changes as they reach a new equilibrium.

In Chapter I. I discuss the development of a new RT-NMR method to investigate protein folding. Applying radio frequency irradiation (200-220 MHz) to a lossy sample inside the NMR spectrometer triggers rapid temperature changes (10-20 °C in 250-500 ms). Cold denaturation - a well-known phenomenon of globular proteins as a consequence of the high temperature dependence of enthalpy and entropy - allows in some cases unfolding of proteins at temperature close to or above the freezing point of water. Barstar, the inhibitor protein of barnase, is a well-studied model system to study protein folding and it can readily be cold denatured between 0-5 °C while maintaining folded state at 25-30 °C with the addition of chaotropic agents (2-2.5 M urea). For the kinetic experiments I generated a cysless mutant of barstar that has been often used in previous folding studies to avoid disulfide chemistry. I could show that this mutant can be reversibly refolded several hundred times without degradation or aggregation.

Inside the spectrometer the refolding of the cold denatured barstar is initiated by the rapid temperature jump. Additionally, a backbone resonance assignment was done in both states (denatured and folded) to be able to follow the changes during refolding at a site-specific level. Initial kinetic experiments using proton NMR spectra showed slow changes in intensity, but the spectral resolution was insufficient, due to the severe signal overlap of simultaneously existing folded, unfolded and possible intermediate states, for detailed analyses.

I developed a new method utilizing the reversible refolding of barstar to record two dimensional NMR experiments by repeating temperature jump (T-jump) experiments. Repeating the kinetic experiments with different phase cycle and evolution time incrementation for the indirect dimension points in an HSQC pulse sequence and post processing the series of kinetic experiments – by concatenating the same time points together – allows recording of time series of two-dimensional HSQC spectra. This series of spectra have the same time resolution as standard proton RT-NMR kinetic experiment but more site specific information can be extracted thanks to the second dimension.

Barstar has two proline residues and out of these two the Tyr47-Pro48 amide bond is in the rare *cis* conformation in the folded state. However, by measuring proline selective carbon detected experiments I could show that in the unfolded state the Tyr47-Pro48 amide bond adopts the major *trans* conformation. The time-dependant signal intensity changes in the kinetic experiments showed for several residues either mono or double exponential behaviour related to the *trans-cis* isomerization. Following the effect of *trans-cis* isomerization in a cryo-cooled probehead with higher sample volume a stable intermediate could be detected. These signals were already present in the dead-time of the experiment and showed mono exponential decay. The intermediate signals could not be assigned to specific residues, although they fall into the typical chemical shift range of folded proteins.

Site-specific analyses of the slower and faster phases of the kinetic changes combined with literature known information a plausible folding mechanism could be outlined. The fast phase is related to the secondary structure formation while the slow phase is related to the effect of the tertiary structure formation slowed down by the *trans-cis* isomerisation of the Tyr47-Pro48 amide bond. The observed lack of intensity changes around the binding site is also in agreement with literature results reporting that the early folding intermediate is already active and inhibits barnase.

Alternative, fast folding pathways could be studied by changing the heating cycle scheme to avoid the rate limiting step of *trans-cis* isomerization. Keeping the cooling time as short as possible, a non-equilibrium unfolded state is reached in which the Tyr47-Pro48 amide bond stays mostly in the *cis* conformation. The fast folding pathway is much faster and not possible to follow changes by the two-dimensional T-jump HSQC method, as folded state was already observed in the first HSQC spectrum after the rapid heating. I

designed a new pulse sequence based on state-correlated and ZZ-exchange type of experiments to correlate the unfolded state with the folded state. In this new pulse sequence indirect dimension evolution takes place still at the lower temperature followed by the rapid heating element. Additional  $\tau$  folding delay is introduced before detection takes place. Repeating the state-correlated experiment with different  $\tau$  times would theoretically allow extraction of kinetic information about the folding process. However, due to the overall time limit of longitudinal relaxation, the determination of precise kinetic information is practically limited. Nevertheless, this state-correlated spectrum contained vast amount of quantitative information about the fast folding pathway. Among others it has revealed that there are intermediates on the fast folding pathway as well. Furthermore, there are exchange peaks which are completely undetectable with short (5 ms) or long (200 ms) refolding time, while some residues still show intense unfolded peaks even with 200 ms refolding time. This provides evidence that different parts of barstar do not fold parallel, but proceed in a step-wise manner.

The combination of reversible cold denaturation with rapid temperature jump inside the NMR spectrometer allowed the study of the folding mechanism of barstar in site-specific detail. This allowed valuable insight into the structural changes associated with various folding steps on two different folding pathways.

Together with pressure jump real-time NMR methods described in recent years, the results obtained in this work by means of T-jump show the high potential of these methods to investigate protein folding experimentally at atomic level to gain a deeper understanding of protein folding mechanisms.

In Chapter II I summarise my work done and results achieved in a collaborative project under the CliC graduate program with the goal to investigate the possible potential of single nucleobase caging to disrupt the secondary structure helix formation in nucleic acids.

In my part of the project I investigated the 15 base pair long double-stranded DNA, caged at the middle cytidine (dC8) with a modified NPE type cage where in the  $\alpha$  position a phenyl substitution is applied, resulting in *R* and *S* configuration of cages. The NMR investigation was limited to proton detected experiments such as 2D-H,H-NOESY and 2D-H,H-TOCSY. After establishing an assignment for more than 80% of the resonances, the relative configuration of the cage was determined. A model structure was built based

on the previously determined structure of the NPE caged DNA double helix with the exact same sequence. Different NOE correlation patterns have been observed between the signals of the (*R*)-NPPM and (*S*)-NPPM cages and the DNA. Two models have been built corresponding to the *R* and *S* configuration of the cage. These models provided possible observable NOE contacts between the cage and DNA atoms that could differ between *R* and *S* configuration. In the 2D-H,H-NOESY spectrum some of these contacts have been identified allowing the determination of the cages relative configuration in the samples. With relative configurations of the cages assigned the melting point of the samples were compared to the (*R*)-NPE and (*S*)-NPE cages. In both cases the *S* configuration showed better performance, although for NPPM cage, the difference between *R* and *S* configuration was smaller. This is most likely due to the smaller size difference between the phenyl and nitro-phenyl groups.

In order to visualize possible changes in the DNA structure, an NMR based model with the cage has been calculated. Due to the lack of sufficient restraints for independent structure calculation, additional dihedral angle and hydrogen bond restraints have been added for base pairs – except for the caged and two neighbouring base pairs in both the 3' and 5' directions and the first and last base pair, which were left to be freely adjusted based on NOE restraints. NOE restraints for labile protons were measured in H<sub>2</sub>O buffer while for non-labile protons in D<sub>2</sub>O buffer. Additionally, to compensate for the spin diffusion, three different mixing times have been recorded for non-labile protons. The structure revealed that the cage, even though it has more steric demand compared to the NPE cage, still fits into the major groove and does not force the nucleobase to flip out into the solvent. The backbone angles are changed from standard B-helix values to accommodate the modification.

Additionally, I used inversion recovery measurements to address how far from the incorporation site the effect of the cage has still influence on the base pair stability. The exchange rate of the labile imino protons with the surrounding water is related to the stability of the base pair. With lower stability the hydrogen bond is weakened and the likelihood for an exchange to occur is increased. Compared to the non-modified DNA helix, increased exchange rate was observed up to two neighbouring base pairs in both the 3' and the 5' directions. The difference between the exchange rates of the caged and native helix is the highest directly at the modification site, while the difference rapidly decays as the residues are further away. This supports the use of additional constraints in



structure calculation for residues unaffected by the caging group. The second observation is that even though we increased the size of the cage, the effect is still limited to base pairs in close proximity. This suggests a limit for the maximum effect achievable by a single caged nucleotide on the DNA double strand.

By melting temperature screening of various cage-modifications, which were introduced into the DNA using click chemistry, a limit of possible depression of melting temperature was found. Additionally, molecular dynamic simulations revealed stabilizing interactions between DNA and certain protective groups. The simulations coincided with the previously observed differences obtained by means of NMR spectroscopy on the structural effects of cages on the DNA.

As part of this interdisciplinary research project, general limitations of the single protecting group approach could be discussed and general design principles for achieving a higher destabilization of the DNA for photolabile protecting groups, such as the preference for the *S* configuration, were shown.

The 3rd chapter deals with the detailed structural and dynamic characterisation of a non-canonically folded RNA model system. Precise and detailed experimental data of model RNA systems are desired, as they are scarce, in order to be used as model systems in molecular dynamics simulations. On one hand the experimental data can be serve as a reference system to determine the quality of force fields, or sampling methods used in simulations on the other hand they can be indispensable source for fine-tuning force field parameters. I discuss the requirement of new well characterized RNA model systems for MD simulations – on one hand as a reference system to address the goodness of the existing force fields and sampling methods and on the other hand to provide force field independent parameters for fine tuning of existing force fields.

A suitable system should be relatively small, preferably biologically relevant with dynamic behaviour that should be also recaptured in the simulations. RNA tetraloops are one of the most common non-canonical secondary structural elements. It has been shown recently that tetraloops can be divided into different groups based on their geometry. With GNRA, UUCG and CUUG loops being the most common naturally occurring sequences, we selected GCAA, CUUG and GAAG to characterize by NMR spectroscopy to cover as much of the tetraloop geometry as possible while still focusing on biologically relevant systems. My research focused on the GCAA tetraloop.

After optimization of *in vitro* transcription to produce  $^{13}\text{C}$ ,  $^{15}\text{N}$  uniformly isotope labelled samples I did a full resonance assignment. I used different three-dimensional experiments (3D-HCP and 3D-HCP-TOCSY) to establish sequential connections along the phosphate-sugar backbone. The ribose signals were assigned mostly based on the 3D-HCC-TOCSY-CCH-COSY experiment. The C6H6/C8H8 and N3/N9 signals were assigned using a two-dimensional version of H(C)N spectra.

After establishing complete assignment of the peaks in the  $^{13}\text{C}$ -HSQC spectrum, temperature series of  $^{13}\text{C}$ -HSQC spectra were measured that showed high temperature dependant chemical shift changes of the ribose signals in the loop as well as in the closing base pair. Calculating and comparing the canonical coordinates at several temperatures suggested temperature dependant conformational changes of the ribose. Additional information about the sugar conformation was collected by measuring  $^3J_{\text{HH}}$  scalar couplings using the 3D-HCC-TOCSY-CCH-E.COSY experiment. This showed for the stem residues values that are typical for canonical A-form helix with 3'-endo conformation in agreement with the canonical coordinates result. However, for the loop residues the coupling constants fell between typical 3'-endo and 2'-endo values, suggesting conformational exchange between the two states.

Additional characterization of the RNA model system was carried out by measuring the additional set of scalar couplings for the phosphate backbone. Additionally, relaxation measurements of  $R_1$  and HetNOE of C1' and aromatic nuclei were done to gain additional information about the dynamics of the tetraloop.

Our results in accordance with the literature information available about this tetraloop, suggest the existence of excited states of the loop. This dynamic behaviour is highly temperature dependant, making the GCAA tetraloop a valuable reference system for MD simulations. Albeit further experimental results are required to fully assess the temperature dependant dynamics observed for the GCAA tetraloop.

## LIST OF ABBREVIATIONS

- AMBER – Assisted Model Building with Energy Refinement  
ARIA – Ambiguous Restraints for Iterative Assignment  
ATP – Adenosine-5'-triphosphate  
CD – Circular Dichroism  
CHARMM – Chemistry at Harvard Macromolecular Mechanics  
CTP – Cytidine-5'-triphosphate  
ddH<sub>2</sub>O - double distilled water  
DEAE – Diethylaminoethyl  
DEPC – Diethyl pyrocarbonate  
DTNB – 5,5'-dithiobis-(2-nitrobenzoic acid)  
DTT – Dithiothreitol  
eNOE – Exact Nuclear Overhauser Effect  
FRET – Förster Resonance Energy Transfer  
GdnDCI – deuterium exchanged guanidine hydrochloride  
GdnHCl – guanidine hydrochloride  
GTP – Guanosine-5'-triphosphate  
HDV ribozyme – Hepatitis delta virus ribozyme  
HSQC – Heteronuclear Single Quantum Coherence  
IAEDANS – 5-(2-[(iodoacetyl)amino]ethyl)amino)naphthalene-1-sulfonic acid  
LB – Lysogeny broth  
MD – Molecular Dynamics  
MW – Microwave  
MWCO – Molecular weight cut-off  
NMR - Nuclear Magnetic Resonance  
NOE – Nuclear Overhauser Effect  
NPAM – (Nitro-phenyl)(adamantyl)methyl  
NPPM – Nitro-Phenyl-Phenyl-Methyl  
OD<sub>600</sub> – Optical Density at 600 nm wavelength  
OPLS – Optimized Potentials for Liquid Simulations  
PAGE – Polyacrylamide gel electrophoresis  
PCR – Polymerase Chain Reaction  
RDC – Residual Dipolar Coupling

RF – Radio Frequency

rNTP – Ribonucleoside triphosphate

RP-IP-HPLC – Reverse Phase Ion-pair High Performance Liquid Chromatography

rpm – rotation per minute

RT-NMR – Real Time NMR spectroscopy

S/N – signal to noise

SOC media - Super Optimal broth with Catabolite repression

TRIS – tris(hydroxymethyl)aminomethane

TROSY – Transverse Relaxation Optimized Spectroscopy

UTP – Uridine-5'-triphosphate

YIPP – Yeast Inorganic Pyrophosphatase

# Chapter I.

## Radio frequency heating and NMR spectroscopy combined to study protein folding

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

**Pintér, G.;** Schwalbe, H. Refolding of Cold Denatured Barstar Induced by Radio Frequency Heating – New Method to Study Protein Folding by Real-Time NMR Spectroscopy. *Angewandte Chemie International Edition* – accepted article.  
DOI: 10.1002/anie.202006945



## 1. INTRODUCTION

### 1.1. PROTEIN FOLDING

Protein folding research is a central interest of biophysics since Cyrus Levinthal noted the paradox of folding happening on short timescales despite polypeptide chains – even average sized proteins – exhibit an astronomical number of possible conformations. Also severe diseases such as Alzheimer, haemodialysis-associated amyloidosis, prion diseases etc are related to misfolding of proteins.<sup>[1]</sup> Therefore, understanding the underlying mechanism of correct folding is important to be able to design effective treatments. These questions led to the development of various techniques to investigate protein folding and several theories to explain driving forces of protein folding.

The most common trigger to initiate protein folding is rapid mixing resulting in pH jump or reducing the concentration of denaturant.<sup>[2]</sup> Other triggers in special cases can also be applied such as heat either from electric discharge<sup>[3]</sup> or laser<sup>[4]</sup> either to study folding of cold denatured proteins or unfolding of proteins. Alternatively, photochemical triggers can be used, for example photolysis of CO-Fe(II) bond in cytochrome *c*.<sup>[5]</sup> The changes accompanying the protein as it folds or unfolds can be followed by different spectroscopic methods.<sup>[6]</sup> Tryptophan fluorescence, ANS fluorescence, FRET, UV/VIS absorption, near- and far UV CD spectroscopy, NMR, IR and Raman spectroscopy report on different levels of structural changes. Usually, to dissect the complete folding pathway, combinations of these methods are applied. This allows different time regimes and structural changes to be studied such as the secondary ( $\alpha$ -helix and  $\beta$ -sheet formation) structures formation and the final tertiary structure (relative position of the secondary structure elements) (re)arrangement at different time points during the folding process.

The vast amount of experimental results obtained in the last decades led to several theories about protein folding. Some of the more widely used ones are hydrophobic collapse<sup>[7]</sup>, diffusion-collision<sup>[8]</sup>, nucleation condensation<sup>[9]</sup>. Although none of these theories could universally explain protein folding, in several cases they were sufficient. Later energy landscape emerged as a uniform theory based on these previously mentioned theories. In an energy landscape all possible states are associated to their respective energy thereby creating a topological space. The advantage of this theory is

that it treats proteins as dynamic systems with several possible conformations in the unfolded, folded and intermediate states as well. It can explain different pathways with different energy minima. It can also explain misfolding and diseases related to protein states as possible kinetically trapped form or off-pathways with similar energy minima to the native state.

## 1.2. CHARACTERIZING PROTEIN FOLDING AND DYNAMICS BY NMR SPECTROSCOPY

NMR spectroscopy has been widely established to study protein folding. The applicable methods can be divided into two subgroups. One utilizes magnetization transfer during conformational changes in equilibrium state, these methods can include measurement of relaxation rates or amide exchanges. The second subgroup applies external triggers to move the sample out of equilibrium state and follow changes in real-time, albeit as is shown in Figure 1 this is usually limited to changes happening on the slower timescale of seconds to minutes.

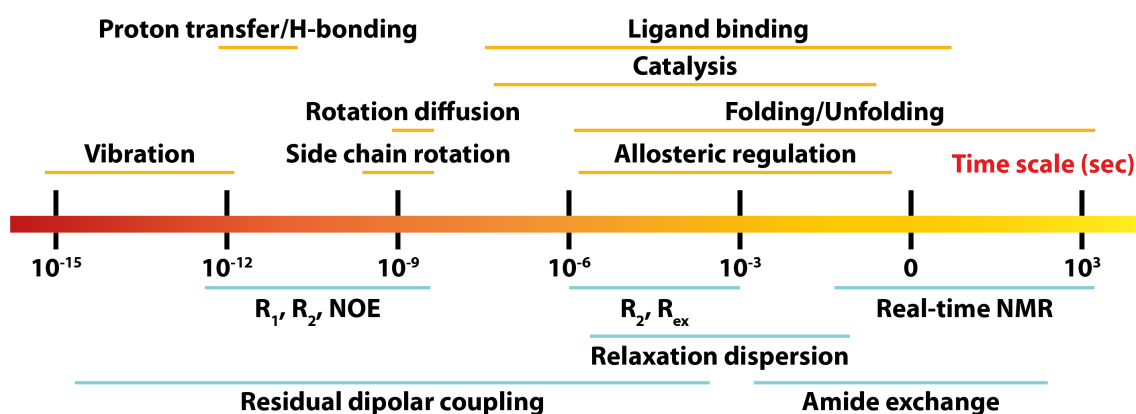


Figure 1 Overview of NMR timescales with corresponding NMR methods and/or parameters (light blue under timescale line) to study protein dynamics, structural changes and biomolecular reactions (light orange above timescale line). The Figure was adapted from Balbach et al.<sup>[10]</sup>

### 1.2.1. Magnetization transfer methods

Different aspects of protein folding and dynamics can be described by energy landscape, albeit simpler models are also suitable or more directly accessible. Nevertheless, NMR spectroscopy can be used to map the energy landscape in great detail. Providing atomistic resolution information about the structure of intermediates on the folding pathway or dynamic changes and excited states spanning from the near femtosecond regime up to minutes (see overview in Figure 1). Magnetization transfer methods rely on the existence of chemical exchange (in the fast exchange regime) between the major population (either



folded or unfolded) and the minor excited state(s). This excited state is directly not visible, but by magnetization transfer from the major population properties about it such as chemical shifts, RDC or the exchange rate between the two population can be studied. There are several methods to cover the different timescales of these changes. Usually a combination of some of the following methods are required: relaxation ( $R_1$ ,  $R_2$ , HetNOE), CPMG, rotating-frame relaxation ( $R_{1\rho}$ ,  $R_{ex}$ ) and saturation transfer to be able to completely address the different states of the protein ensemble<sup>[11–14]</sup>. Another important technique to characterize possible substructures on the folding pathway is hydrogen-deuterium exchange experiment, utilizing the exchange rate difference of amide protons involved in hydrogen bonds.<sup>[15]</sup>

When the interconversion between two conformations (e.g.: folded and unfolded state) is in the slow exchange regime ( $k_{ex} \ll \Delta\omega$ ) on the NMR timescale ZZ-exchange experiments can be used to directly measure the exchange rates.<sup>[16]</sup> By measuring the temperature dependence of the exchange rates and the population, the free energy differences of the ground states and the transition state can be obtained.

### 1.2.2. Real time NMR methods

The other major group of experiments are RT-NMR measurements. In this case the system in equilibrium is perturbed and the relaxation to the new equilibrium is followed by NMR spectroscopy. This is limited to changes that are happening on timescales greater than a few seconds, due to the inherent slowness of NMR spectroscopy. Triggers can be changing pH or denaturant concentration by rapid mixing<sup>[17–19]</sup>, releasing a compound by light irradiation<sup>[17,20–23]</sup>, changing the pressure or temperature. The latter two have the advantage that if the system of interest is reversible, the experiments can be readily repeated and data can be accumulated. In recent years technical advancement in pressure changing systems in combination with new pulse sequences led to highly detailed information about the folding of ubiquitin.<sup>[24–27]</sup> Temperature initiated kinetic measurements have been long conceptualized<sup>[28]</sup> but by technical limitations it has seen so far very limited application in the study of protein folding or dynamics.<sup>[29]</sup>

### 1.3. TEMPERATURE JUMP SYSTEMS FOR NMR SPECTROMETER

In the first few attempts to use temperature changes to trigger kinetic measurements inside the spectrometer cooled nitrogen gas flow was used to cool down the sample.<sup>[29–31]</sup> In these studies around -20K to -30K changes were achieved in around 5–6 seconds,

which limits the applicability to follow faster changes. Later systems were developed for extreme high temperature measurements utilizing either laser heating or inductive heating.<sup>[32–35]</sup> An impressive few hundred kelvin could be achieved by both methods, but the range of applicable pulse sequence was limited due to the lack of temperature stabilization and homogeneity. Ferguson and Haw compared these techniques and while laser heating is 2-3 times faster, the temperature homogeneity of radio frequency (RF) inductive heating was found to be much better.<sup>[36]</sup> The application of microwave heating to aqueous solutions have been first reported by Akasaka et al. in 1990 and an improved version was published in 1998.<sup>[37,38]</sup> In the latter publication the temperature jump was initiated by conventional 2.46 GHz microwave pulses generated by a 1.3 kW magnetron. Instead of a copper wire antenna the microwave is guided to the sample by a non-conducting dielectric resonator. To improve temperature homogeneity a mixing apparatus was inserted into the sample tube. This setup was capable of a jump as fast as 20 °C in 20 ms which is suitable to follow even faster protein folding steps. The downside was the lack of applicable Z-gradients due to the physical mixing of the sample. This limits the applicability to samples in D<sub>2</sub>O and hinders the use of modern pulse sequences. Different designs of radio frequency heating have been published and generally RF (< 1 GHz) or MW (> 1 GHz) volume heating is utilized.<sup>[39,40]</sup>

Generally, there is always a trade off in the design space of T-jump setups for NMR spectrometer. NMR is inherently insensitive and requires large volume of sample, leading to increased heating times. By reducing the volume faster heating can be achieved but at the expense of signal-to-noise. Two major factors contribute to the dead time of T-jump experiments: the actual heating time and the temperature homogenization time after the jump. While lower radio frequencies (< 1 GHz) have deeper penetration depth compared to higher radio frequencies (1-300 GHz), leading to more homogeneous heating, the energy transferred to heat is lower for water solvent.<sup>[41]</sup> The optimal frequency for a lossy water sample falls into the higher frequency regime (> 1 GHz) with the exact frequency depending on the salt concentration and temperature.<sup>[41]</sup> Also the applied frequency and homogeneity has direct influence on the sensitivity due to line broadening from temperature gradients in the sample.

The design of the T-jump equipment<sup>[42]</sup> used in this thesis, shown in Figure 2, applies heating pulses around 220 MHz, the exact value depends on the sample conductivity and the temperature influencing the tuning and matching. The inner radio frequency coil is a

cage coil designed to maximize the electric field component of the radio frequency. The small 2.5 mm diameter of the sample tube allows 80  $\mu$ l volume of sample to be used.

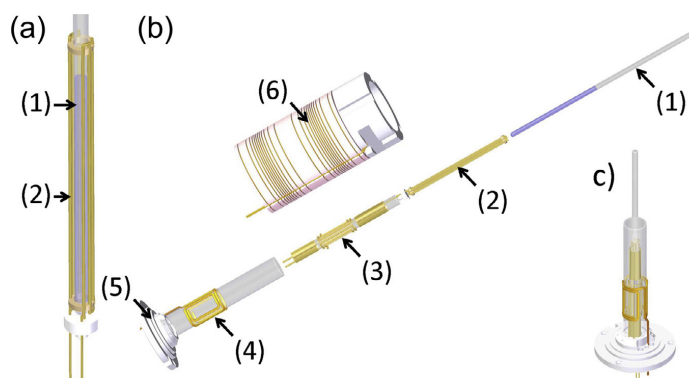


Figure 2 Temperature jump probehead published by Rinenthal et al., 2015<sup>[42]</sup>. (a) Cage coil/wire capacitor for rf-heating with sample tube (2.5 mm), (b) exploded view of the overall assembly with (1) sample tube, (2) 2.5 mm cage coil, (3) 5 mm double tuned ( $^1\text{H}$  &  $^2\text{H}$ ) rf NMR saddle coil, (4) 10 mm ( $^{15}\text{N}$ ) rf NMR saddle coil, (5) coil insert base and connection to main probe body, (6)  $z$  gradient coil. (c) Full assembly of coils in the T-jump probe ( $z$  gradient system not shown).

The major advantage of this setup is the ability to apply  $z$ -gradients, enabling water suppression for samples in  $\text{H}_2\text{O}$  buffer systems and limited phase cycle requirement by gradient selection in multi-dimensional NMR experiments. The latter is required for cycling 2D-NMR measurements designed during this project and described in chapter 3.4.

## 1.4. MODELL PROTEIN: BARSTAR

### 1.4.1. Structure

Barstar is the inhibitor protein of barnase first isolated from *Bacillus amyloliquefaciens* in 1967<sup>[43]</sup>. It inhibits the ribonuclease activity of barnase by sterically blocking the active site, forming a one to one complex<sup>[44]</sup>. This is required for the cell to survive until barnase is secreted. Barnase can only be co-expressed with barstar<sup>[45]</sup>. Barstar and barnase form one of the tightest known protein-protein complexes, with the dissociation constant of  $10^{-13}$  M<sup>[46]</sup> comparable to the affinity of streptavidin to biotin. Thanks to their relatively small size (barstar 89 aa, barnase 110 aa) and easy production both proteins have been widely used as a model system to study protein folding<sup>[47,48]</sup> and protein-protein interaction<sup>[49,50]</sup>

The three-dimensional structure of barstar is shown in Figure 3. It contains four  $\alpha$ -helices and a three stranded  $\beta$ -sheet. The small helix<sub>3</sub> and helix<sub>2</sub> are considered less stable and more flexible based on hydrogen-deuterium exchange experiment<sup>[51]</sup>. The proposed

theory is that helix<sub>2</sub> is involved in a rigid-body movement to facilitate binding of barnase. The  $\beta$ -strand<sub>2</sub> and  $\beta$ -strand<sub>3</sub> were shown to be very dynamic as well. Interestingly the relaxation values of the  $\beta$ -strands were exceptionally high compared to similar sized proteins. A later paper revising these unexpectedly high relaxation rates of found that the  $\beta$ -strand<sub>2</sub> and  $\beta$ -strand<sub>3</sub> are also involved in dimerization at high concentration of barstar<sup>[52]</sup>.

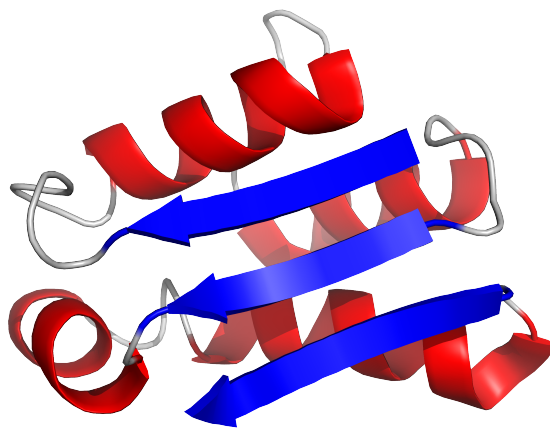


Figure 3 Three-dimensional structure determined by NMR spectroscopy of the C40A/C82A mutant barstar (PDB: 1AB7)<sup>[51]</sup>. Colour code of the secondary structure: blue –  $\beta$ -strands, red –  $\alpha$ -helices.

#### 1.4.2. Binding site

Residues involved in the binding between barnase and barstar have been investigated by X-ray crystallography<sup>[50,53]</sup> and different NMR spectroscopy techniques: chemical shift perturbation, hydrogen-deuterium exchange<sup>[54]</sup> and relaxation measurements<sup>[55]</sup> combined with single and double mutant cycles<sup>[56]</sup>. The crystal structures showed that mostly residues between 29-46 and 73, 76 of barstar are involved in the binding. It has been found that besides hydrogen bond interactions electrostatic interactions also play a crucial role. For example, replacing glutamic acid 76 by glutamine could form the same hydrogen bonds but resulted in reduced binding affinity. The chemical analyses of the residues in the binding surface area show interestingly high propensity towards charged and polar sidechains. More than 50% of the surface contains charged residues compared to other protein-protein interaction sites where usually non-polar residues are predominant with more than 50% and the rest of the residues are divided between hydrophobic, polar and charged residues. This is further supported by MD simulations, that the interaction between barstar and barnase is electrostatically optimized<sup>[57]</sup>. Replacing the side chains of the charged surface hotspots in any direction leads to reduced binding affinity, showing the evolutionary optimization of the tight binding

## 1.4.3. Overview of folding mechanism

The first detailed folding study on barstar was published by Schreiber et al. in 1993<sup>[58]</sup>. Stopped-flow measurements were conducted to follow the intrinsic fluorescence of the tryptophan residues of the C40A/C82A mutant of barstar. Various end concentration of urea and double jump experiments were conducted to dissect the possible folding pathways.

Their proposed scheme of folding (Figure 4.) can be divided into a slow folding (upper part) and fast folding pathway (lower part), which depends on Tyr47-Pro48 amide bond conformation. Using double jump experiments the effect of the trans-cis isomerization was separated, either using transiently or equilibrium denatured protein during stopped-flow measurement. On both folding pathway indirect evidence suggests an intermediate is formed before the first intermediate is detected. On the slow folding pathway two different intermediates ( $I_{T1}$  and  $I_{T2}$ ) were directly detected, while on the fast folding pathway only one ( $I_{C2}$ ).

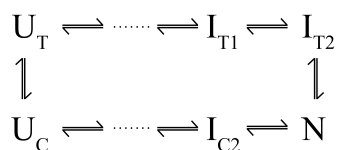


Figure 4 Proposed folding pathways of C40A/C82A mutant barstar by Fehrst.<sup>[58]</sup> U, I and N refers to unfolded, intermediate and native states respectively, while T and C refers to the *trans* and *cis* isomer state of the Tyr47-Pro48 bond. Upper part is the slow folding pathway, with  $I_{T2}$  suggested to be native like. The first step from the unfolded state on both pathways could not be directly detected, but indirect evidence was found.

The second major folding study on barstar was conducted by Shastry et al. and published in 1994<sup>[59]</sup>. In comparison to the first study wild type barstar was used with reducing agent DTT added to the buffer to prevent disulphide bridge formation. Furthermore, GdnHCl was used as chaotropic agent instead of urea. Their results were comparable to the folding of urea denatured barstar.

Investigation of the possible structure relation to the intermediate states was published in 1995 by Agashe et al.<sup>[7]</sup> Using stopped-flow system to trigger protein folding from denatured state combined with different detection methods. Secondary structure changes have been monitored by CD spectroscopy at 222 nm wavelength. They observed a double kinetic intensity changes and while intensity changes were observed, the amplitude of the intensity changes were higher in the fast phase, suggesting the formation of most of the secondary structure in this fast phase. Tertiary structure formation monitored by

ellipticity changes at 270 nm showed the opposite, almost completely forming in the slow phase. Monitoring the same kinetics by intrinsic fluorescence – stemming from W38, W44 and W53 – recovered its folded value by 70% in the slow and 30% in the fast phase.

Formation of solvent accessible hydrophobic surfaces, which is assumed to be due to non-specific collapse of the random coil during the burst phase within the dead time of stopped-flow measurements was proposed when studying binding of ANS to barstar. ANS is a fluorescent molecule which changes its fluorescence intensity at 460 nm upon binding to hydrophobic surfaces. Unfolded barstar does not bind ANS, while directly after the dead time of the mixing ANS binding was detected as the fluorescence changed compared to unfolded barstar, suggesting non-specific collapse of the random coil structure, as solvent accessible surface is formed. To gain information about the compactness of this burst-phase intermediate, barstar was labelled at the two cysteine residues with the dansyl-type fluorophore IAEDAN, and fluorescence energy transfer between the tryptophan residues and dansyl labels were monitored. The six pairs of Trp-dansyl correlation are distributed around the whole structure. In the burst phase 70% the energy transfer was restored compared to folded barstar, concluding that the burst phase intermediate is compact with no defined secondary structure.

Nölting et al. reported on the kinetics of the collapse of the unfolded structure of barstar, the first step on the folding pathway<sup>[3]</sup>. This step has been only indirectly observed as changes during the dead time of stopped-flow measurements. Barstar can be cold denatured (see in chapter 1.4.5), which allows rapid initiation of refolding by discharge heating in the microsecond and nanosecond time regime. C40A/C82A/P27A mutant barstar is less stable than wild-type or C40A/C82A mutant and a few percent of unfolded population can be reached without denaturant at 0 °C. The kinetics followed at different fluorescence wavelength showed that aromatic side chains are not completely buried but solvent accessibility is reduced. The hydrophobic collapse as the first step on the folding pathway was further supported by Nath et al.<sup>[60]</sup> They used a W38F/W44F mutant, to follow only the fluorescence of Trp53 which is completely buried in the fully folded protein. Their stopped-flow measurements showed that changes in fluorescence compared to unfolded barstar occur during the dead time of the experiment, due to the hydrophobic collapse of the unfolded chain.

Shastry et al. presented a detailed stopped-flow study to investigate the different folding pathways of barstar. From previous studies it was known that depending on the *cis-trans* isomerization a slow and fast folding pathway can be separated. Furthermore, the ratio of the slow and fast folding pathway can be influenced by the end concentration of the denaturing agent used in the folding experiment. By analysing the stopped-flow experiments following tryptophan and ANS fluorescence at different end concentrations of GdnHCl, the complete mechanism of folding was suggested to be divided into different pathways. They also observed the same native like intermediate on one specific pathway. This intermediate has been shown by Fehrst et al. to be active and readily inhibits barnase. While Shastry et al. found that it also binds ANS, which fully folded barstar does not. This suggest that the barnase binding site is already formed, while the core of barstar is still solvent accessible.

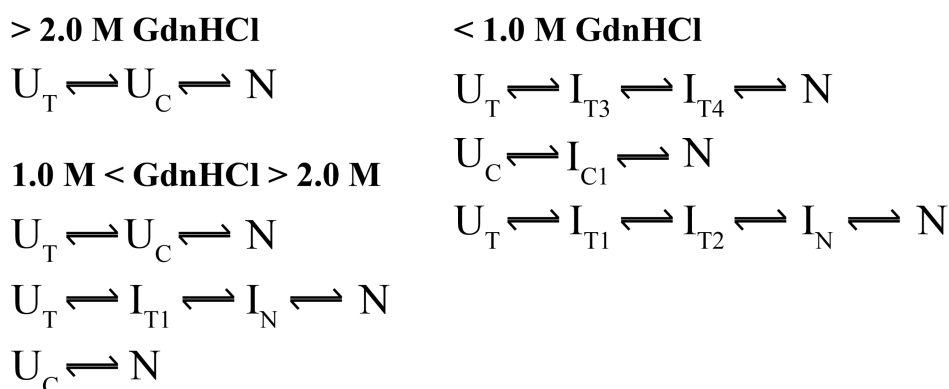


Figure 5 Proposed folding pathways of wt-barstar by Udgaonkar.<sup>[61]</sup> Depending on the concentration of GdnHCl different mechanisms are proposed. U, I, and N refers to unfolded, intermediate and native state, respectively, while subscript T and C refers to *trans* and *cis* conformation of the Tyr47-Pro48 amide bond, respectively. I<sub>N</sub> is an intermediate that was observed in previous studies as well, and was shown to be active by inhibiting barnase.<sup>[58]</sup>

Nölting et al. investigated the hierarchy of the folding by a detailed phi-value analysis. They investigated the mutant P27A/C40A/C82A barstar with additional mutations in different secondary structural elements. Combination of electric discharge T-jump experiments and stopped-flow measurements were used with detection of ellipticity changes at 222 nm. Combination of these kinetic measurements allowed the first estimation of which and to what extent secondary structures are formed at different time points of the folding process. In combination with mutations in different secondary structural elements the formation of these secondary structures during the folding process could be followed. The proposed mechanism of the folding in this study was nucleation-condensation where certain secondary structures are formed first and act as nucleus for other parts of the protein to fold. This mechanism is in agreement with previous studies,<sup>[7]</sup>



where intermediates have been shown to exhibit different levels of formed secondary structures and tertiary structures.

Bhuyan et al. did the first RT-NMR investigation of the slow folding part, which is assumed to be related to the *cis-trans* isomerization of Tyr47-Pro48 amide bond.<sup>[62]</sup> The slow folding part could be studied even with the long (~10 sec) dead time of their experiment. The sample was in 99% D<sub>2</sub>O buffer and only non-labile signals were used as reporters. Different reporter signals showed different time dependent signal intensity changes. This suggests that there must be parallel transition states of the different residues. The kinetics in this case were also dependent on the end concentration of GdnDCI. Parallel with this publication Killick et al. also investigated the slow folding phase of barstar.<sup>[63]</sup> In comparison they used the wild-type like P27A/C40A/C82A mutant barstar, and used manual mixing (dead time > 10 sec) and followed the changes by series of 2D-<sup>15</sup>N-HSQC. They did not report on multistate kinetics but could show the existence of stable intermediates on this pathway by detection of amide cross peaks related to neither unfolded nor folded signals. They tracked most of the changes to residues: A82, T85, V50, Y47 although smaller intensity changes of other residues could be observed.

Time resolved fluorescence measurements published by Sridevi et al.<sup>[64]</sup> investigated further the packing of the different intermediates on the folding pathway. Fluorescence anisotropy of Trp53 in the first detectable intermediate by stopped-flow measurement (after the burst-phase non-specific collapse) shows two types of rotational motion: a local one where the indole side chain rotates (~1 ns) and a slower (~7.2 ns) motion correlating with the global tumbling of the protein. The disappearance of the faster motion in the fully folded protein suggests that this intermediate does not have a fully formed tight core. While the slower overall tumbling time of the intermediate compared to the folded protein (~5.6 ns) suggests that the overall volume is 25% larger. Time dependent anisotropy shows that the fast rotation disappears before the protein completely folds, suggesting that the core packing occurs before the final secondary / tertiary structure is formed. However, the overall tumbling time change determined from relaxation rates matches with steady-state fluorescence and CD determined rates of the final folding.

An important aspect to consider when these studies are compared is the use of different constructs. In the Fehrst group mostly P27A/C40A/C82A wild-type like mutant was used while in Udgaonkar's lab the wt-barstar was preferred as standard. Later in a <sup>19</sup>F-NMR



study the C40A/C82A mutant was compared to the P27A/C40A/C82A mutant in equilibrium studies.<sup>[65]</sup> Different <sup>19</sup>F-Trp and <sup>19</sup>F-Phe labelling schemes were used. That study revealed the widespread effect of the P27A mutation on the stability and rigidity of the core. The use of different mutants makes subtle differences detected by different methods in different groups difficult to compare. However, a major picture about the folding pathways and the time scales can be concluded.

The results of the folding studies showed that barstar has multiple folding pathways, depending on the starting conformation of Tyr47-Pro48 amide bond and the end concentration of the denaturing agents, accompanied by multiple intermediates. The general steps on all folding pathways start with a sub millisecond collapse of the unfolded state, most likely without any secondary structure formation but likely with the formation of some specific interactions. In the next time regime of a few hundred milliseconds, one or more intermediates with different completeness of the final secondary structure are formed and parallel tertiary structures also start to form. In the final time regime, in the slow phase, the *trans-cis* isomerization of Tyr47-Pro48 amide bond takes place. These results can be explained in the framework of the nucleation condensation mechanism - as Fehrst suggested.

#### 1.4.4. Cold denaturation phenomena

Cold denaturation is nowadays accepted as a general property of globular proteins ranging from small proteins up to multi-domain proteins.<sup>[66]</sup> If we consider that heating increases enthalpy and entropy it is not surprising that proteins can be denatured by heat. By the same logic it would be expected that by decreasing temperature the protein stability would increase. These considerations from a thermodynamic point of view exclude the temperature dependence of enthalpy and entropy, although temperature dependence exists and it can be quite large even in biologically relevant temperature range (-5 °C to 35 °C). As Privalov showed the highest stability of a protein is reached where the entropy of folded and denatured state is equal and only the enthalpy difference stabilizes the protein. Furthermore, his calculations showed that a protein has higher cold denaturation temperature when there is large difference between the folded and unfolded state heat capacity and there is small difference between folded and unfolded state enthalpy. The mechanistic explanation for this phenomenon is due to the interaction between the hydrophobic side chains and the solvent.<sup>[67]</sup>

Challenges in constructing an experimental setup to study cold denaturation can be overcome by reducing the overall stability of the protein by changing parameters such as pressure, adding low concentration of chaotropic agents or cooling the water-protein system below freezing point to a metastable state. The latter can be done by increasing the relative surface to volume ratio e.g.: using capillaries or by encapsulating proteins in reverse micelles<sup>[68]</sup>.

Cold denaturation below water freezing point in capillaries has been used to study in atomistic detail the homodimeric repressor protein CyIR2 unfolding process, by determining the structure of the protein at different temperatures.<sup>[69]</sup> By dissecting the different interactions it could be shown that partially folded states have non-native interactions which likely lead to misfolding of the protein. Babu et al. also used cold denaturation in combination with NMR spectroscopy to follow the unfolding of the model protein ubiquitin.<sup>[68]</sup> Surprisingly cold denaturation of ubiquitin showed non-cooperative behaviour, unlike heat denaturation.

Determining the precise temperature dependence of enthalpy, entropy and heat capacity can be challenging for biomolecules. Cold denaturation has been used in combination of heat denaturation to map thermodynamic parameters of both folded and unfolded states at low and high temperature by Szyperski et al.<sup>[70]</sup> Using NMR spectroscopy has an advantage compared to standard melting curve measurements by other spectroscopic methods in that reporter signals of both folded and unfolded populations can be observed in parallel. This method allows the determination of the complete melting curve starting from cold denatured state up to the heat denatured state following both folded and unfolded population, allowing very precise determination of thermodynamic parameters.

#### 1.4.5. Cold denaturation of barstar

Barstar has been shown to undergo cold denaturation in presence of denaturing agents, which however cannot destabilize the protein at the temperature where it is most stable.<sup>[71-73]</sup> The main reason found in case of barstar for cold denaturation as Privalov's theory suggests is the high  $\Delta C_p$  of  $1.46 \pm 0.07 \text{ kcal mol}^{-1} \text{ K}^{-1}$  and the low  $\Delta H_{298K}$  of  $4.0 \text{ kcal mol}^{-1}$ . Another protein exhibiting similar thermodynamic parameters is myoglobin, for which cold denaturation has been experimentally shown as well.<sup>[74]</sup> Nölting et al. also suggested a mechanistic explanation for the relatively high cold denaturation temperature of barstar.<sup>[73]</sup> The major contribution is suspected to be the

interaction between side chains and solvent molecules, that decrease the hydrophobic effect at lower temperatures. In case of barstar the especially low ratio of electrostatic contribution compared to the hydrophobic contribution be driving the cold denaturation. Barstar exhibits residual secondary structures in the unfolded state, according to the results of CD spectroscopy measurements, independent of whether unfolding has been achieved by cold denaturation, chaotropic agents or by heat. Wong et al. published a detailed characterization of the cold denatured state of barstar by NMR spectroscopy.<sup>[72]</sup> Besides a complete resonance assignment of the mutant C40A/C82A barstar in the cold denatured state, detailed analyses of the residual secondary structure has been carried out. Average chemical shift deviation from sequence corrected shifts<sup>[75]</sup> showed three clusters with residual secondary structure: Leu26 to Gly31, Trp38 to Glu46 and Glu52 to Gln55. As a second measure of secondary structure preference  $\sigma_{N\alpha}/\sigma_{\alpha N}$  was determined as the intensity ratio of  $H\alpha_{(i+1)}-HN_{(i+1)}$  to  $H\alpha_{(i)}-HN_{(i+1)}$  revealing similarly three regions: Ser12 to Lys21, Tyr29 to Glu46 and Leu51 to Phe56 that align with the chemical shift analysis. The analyses of  $^3J_{HN-H\alpha}$  couplings of these clusters showed preference to helical structures, which in case of the first two regions matches with the folded secondary structure, but Leu51 to Phe56 is in the  $\beta$ -sheet region. The first two regions might require less folding space to sample during folding, hence speeding up the folding of these regions and might therefore serve as a nucleation site.

In this part of the thesis technical development of a T-jump probehead is applied in combination with cold denaturation to gain further insight into the folding mechanism of barstar. Pulse sequence development for the measurement of  $^{15}N$ -HSQC spectrum in real time and shorter duty cycle during the T-jump part allowed residue specific characterization of the effect of *cis-trans* isomerization has on the slow folding pathway. A pulse sequence to generate state-correlated spectrum was developed to gain insight into the earlier steps on the fast folding pathway, thereby revealing intermediates along the fast folding pathway of barstar.

## 2. MATERIALS AND METHODS

### 2.1. PROTEIN PRODUCTION AND PURIFICATION

#### 2.1.1. Synthesis of plasmid DNA

Wild-type barstar sequence (codon optimized for expression in *E.Coli*) has been chemically synthesized and cloned into the pET-3a vector with cloning site NdeI/BamHI by GenScript (USA). The lyophilized plasmid (4 µg) was dissolved in ddH<sub>2</sub>O and stored at -20°C. Sequence of wt-barstar:

MKKAVINGEQIRISISDLHQTLKKELALPEYYGENLDALWDCLTGWVEYPLVLE  
WRQFEQSKQLTENGAESVLQVFREAKAEGCDITIILS

#### 2.1.2. Transformation

The plasmid has been amplified in DH5α (New England Biolabs, USA) *E.coli* cells and transformed for protein production into the BL21(DE3)pLysS *E.coli* cell line. For transformation 50 µl of cells have been thawed on ice for 10 minutes. The cells have been supplemented with 3 µl plasmid (70 ng/µl) and incubated on ice for 30 minutes. Thereafter, cells have been heat-shocked for 45 second to facilitate the uptake of the plasmid. The mixture was incubated on ice for 5 minutes and then it was supplemented with 950 µl SOC media. The mixture was incubated at 37 °C and shaken at 250 rpm for one hour. Dilution series of 1:1, 1:10, 1:100 and 1:1000 have been plated onto selection plates containing ampicillin. The plates were incubated at 37 °C overnight to facilitate colony growth.

#### 2.1.3. Plasmid DNA purification

5 ml LB medium was supplemented with 100 µg/ml ampicillin and was inoculated with a freshly picked cell colony. The mixture was incubated at 37 °C and 160 rpm for 6-16 hours. The plasmids were isolated and purified with the QIAquick<sup>®</sup> Mini Prep Kit (Qiagen, Germany) according to the manufacturer manual. Plasmid sequence was controlled by standard sequencing using T7-pET-mod primer (Eurofins Genomics, Luxembourg).

## 2.1.4. Site-directed mutagenesis

Site directed mutagenesis was used to generate double alanine to cysteine mutant (C40A/C82A) of barstar with the following amino acid sequence:

*MKKAVINGEQIRISIDLHQLKELALPEYYGENLDALWDAL<sup>A</sup>TGWVEYPLVLE  
WRQFEQSKQLTENGAESVLQVLFREAKAEG<sup>A</sup>DITIILS*

The primers were designed manually. The desired mutation was placed in the middle of the primers and in both 5'- and 3'-direction 12-15 nucleotide overhang was designed in consideration of an optimal melting point of 64 °C and closing base pairs of GC desired, the primers are listed in Table 1. The melting temperature was calculated with Oligo Calc.<sup>[76]</sup> The PCR mixture shown in Table 2 was thermocycled according to Table 3.

Table 1 List of primers to introduce C40 to A40 and C82 to A82 point mutations by site directed mutagenesis.

| Primer name                      | Primer sequence                            |
|----------------------------------|--|
| Reverse primer for C40A mutation | 5'-catagtactccgC <sup>c</sup> cagcgcag-3'  |
| Forward primer for C40A mutation | 5'-ctggcgctgG <sup>c</sup> gggagtactatg-3' |
| Reverse primer for C82A mutation | 5'-gaaccagcG <sup>c</sup> gtattccacc-3'    |
| Forward primer for C82A mutation | 5'-gggtggaatacG <sup>c</sup> gctgttc-3'    |

Table 2 Composition of PCR mixture used in site directed mutagenesis.

| Components                 | Final amount / concentration |
|----------------------------|------------------------------|
| Plasmid template           | 70 ng                        |
| 5x Phusion HF or GC buffer | 4 µl                         |
| Reverse primer             | 0.5 µM                       |
| Forward primer             | 0.5 µM                       |
| dNTP mix                   | 200 µM                       |
| ddH <sub>2</sub> O         | 10.5                         |
| Phusion polymerase         | 0.5 U                        |

Table 3 Thermocycle program for site directed mutagenesis.

| Number of cycles | Temperature | Time     |
|------------------|-------------|----------|
| 1                | 95 °C       | 1 min    |
| 18               | 95 °C       | 1 min    |
|                  | 65 °C       | 1 min    |
|                  | 72 °C       | 7 min    |
| 1                | 72 °C       | 7 min    |
| Pause            | 4 °C        | infinite |

To remove non modified template plasmid Dpn I digestion was used. 10 U of the restriction enzyme was added to the reaction mixture, it was gently mixed and incubated at 37 °C for 2 hours, followed by heat inactivation at 80 °C. The modified plasmid was transformed according to the protocol described in chapter 2.1.2.

### 2.1.5. Protein production

Protein production was performed in BL21(DE3)pLysS *E.coli* cells transformed with the desired plasmids. Preculture of 5 ml LB media was supplemented with 100 µg/ml ampicillin. It was inoculated with freshly plated cells or cryo-stocks. The cells were grown at 37 °C and 160 rpm for 6 hours. The preculture was transferred into 100 ml M9 minimal media. Isotope labelling was achieved by supplementation of 1 g/l <sup>15</sup>N-NH<sub>4</sub>Cl and / or 2 g/l <sup>13</sup>C-glucose as the only nitrogen and / or carbon source. The cells were grown for 4-6 hours until cell density reached value of one OD600 (optical density at 600 nm wavelength). The media was transferred to preparative scale of one litre M9 minimal media with the corresponding isotope supplements. The starting cell density was adjusted to 0.1 OD600. The cells were grown at 37 °C and 140 rpm. The protein production was induced by addition of 1.0 mM IPTG at 0.8 OD600 cell density and was further incubated at 20 °C and 140 rpm for 16 hours. The cells were harvested at 4000 rpm and 4 °C for 20 min. Cells were resuspended in minimal volume of 50 mM Tris-HCl, pH 8.0 buffer supplemented with cOmplete protease inhibitor (Roche, Switzerland). Cell membrane was disrupted by microfluidizer operating at 1000 bar. Cell pellet was centrifuged at 16000 rpm and at 4 °C for 40 min. Barstar is found in the supernatant, which was saturated to 40% (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> concentration (required amount was calculated by EnCor Biotechnology Inc. online tool for 20 °C). The precipitate was separated by centrifugation at 16000 rpm at 20 °C for 30 min. The supernatant was further saturated to 80% (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> concentration. The precipitate containing barstar was centrifuged at 16000 rpm at 20 °C. It was dissolved in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl buffer and gel filtered on a Superdex 75 26/600 column in an ÄKTA system with automatic fractionation, collecting 7 ml fractions. The fractions were analysed by PAA gel-electrophoresis and was visualized by Coomassie dye staining. The barstar containing fractions were pooled and concentrated up to 25 mg/ml concentration in centrifugal concentrator (Vivaspin® Sartorius AG, Germany) with 3000 dalton MWCO. Next purification step was done on a Resource™ Q 6 ml anion exchange columns with gradient separation.

The barstar containing fractions were concentrated in centrifugal concentrator. Final polishing step was gel filtration on a Superdex Increase 75 10/300 GL column: this step was combined with buffer exchange to 10 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub>, pH 6.4, 130 mM KCl buffer.

Final sample concentration (2.5 - 3 mM) was achieved by concentrating the sample in centrifugal concentrator (Vivaspin®, MWCO 3000).

## 2.2. NMR EXPERIMENTS FOR ASSIGNMENT

All NMR experiments were measured on spectrometers with fields strengths from 14 Tesla up to 21.1 Tesla (corresponding to 600-900 MHz  $^1\text{H}$  Larmor frequency respectively) equipped with cryoprobes with radio frequency channels optimized for  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$  Larmor frequencies. Topspin version of 3.2 to 4.0 was used for the measurements and data processing. Further analyses was done with CcpNMR Analysis v2.<sup>[77]</sup>

### 2.2.1. Chemical shift assignment of HN, $C_\alpha$ , $C_\beta$ , CO in folded barstar

Chemical shift assignment was based on literature published assignment<sup>[51,78]</sup> of wt- and C40A/C82A mutant barstar. The backbone assignment was done on the C82A mutant barstar with standard BEST-TROSY type of HNCACB, HNcoCACB experiments taken from Bruker library. The full chemical shift assignment is found in the Appendix. C40A/C82A mutant barstar was assigned by comparison of HSQC spectra to C82A barstar assignment and literature data. Small shifts have been observed compared to literature published values due to the different buffer conditions.

### 2.2.2. Chemical shift assignment of HN, $C_\alpha$ , $C_\beta$ , CO in cold denatured barstar

Cold denatured barstar showed typical low chemical shift dispersion in  $^1\text{H}$  dimension of standard  $^{15}\text{N}$ -HSQC correlation spectrum of unfolded protein. All measurements required to assign barstar in the cold denatured state has been recorded at 270 K. The unfolded state leads to lower dispersion and the low temperature reduces the S/N arising from increased  $T_2$  relaxation time. To overcome these limitations BEST-TROSY type of HNCACB, HNcoCACB, HNCO and HNcaCO spectrum<sup>[79]</sup> were recorded with 25% NUS, for  $^{15}\text{N}$  indirect dimension 0.008 and for  $^{13}\text{C}$  dimension 0.05  $T_2$  exponential weighted sampling distribution was used to improve the resolution within the measurement time limit. Additionally, hNcacoNH spectra<sup>[79]</sup> was also recorded, utilizing the higher dispersion of the  $^{15}\text{N}$  chemical shifts of the amide backbone. The full chemical shift assignment is found in the Appendix. Not all residues could be assigned, due to line broadening beyond detection.

Table 4 Parameters used to measure the 3D experiments for cold denatured state assignment.

| Pulse program      | Spectral width of F1 (ppm) | Spectral width of F2 (ppm) | Spectral width of F3 (ppm) | Number of indirect points in F1 | Number of indirect points in F2 | Number of indirect points in F3 | Carrier frequency of F1 (ppm) | Carrier frequency of F2 (ppm) | Carrier frequency of F3 (ppm) | Number of scans |
|--------------------|----------------------------|----------------------------|----------------------------|---------------------------------|---------------------------------|---------------------------------|-------------------------------|-------------------------------|-------------------------------|-----------------|
| b_trhncacbgp3d.2   | 65                         | 22                         | 12                         | 200                             | 180                             | 1024                            | 43                            | 119.2                         | 4.7                           | 32              |
| b_trhncocacbgp3d.2 | 65                         | 22                         | 12                         | 200                             | 180                             | 1024                            | 43                            | 119.2                         | 4.7                           | 32              |
| b_trhncogp3d.2     | 7                          | 22                         | 12                         | 90                              | 180                             | 1344                            | 172.5                         | 119.2                         | 4.7                           | 8               |
| b_trhncocanhgp3d   | 7                          | 22                         | 12                         | 90                              | 180                             | 1344                            | 176                           | 130                           | 4.7                           | 32              |
| b_trhncocanhgp3d.2 | 22                         | 22                         | 12                         | 160                             | 180                             | 1344                            | 119.2                         | 119.2                         | 4.7                           | 32              |

### 2.3. T-JUMP NMR EXPERIMENT SETUP

#### 2.3.1. Heating setup for T-jump experiments

Setting up T-jump experiments require the optimization of the heating pulses. The tuning and matching were optimized at the midpoint of the desired temperature jump. The carrier frequency is determined at the lowest and highest temperature of the jump – as is shown in Figure 7b. The power level of the heating pulses, the number of pulses was optimized to achieve desired T-jump for each sample individually. Stable temperature of the sample after T-jump element was also optimized, if several FIDs are collected over longer time period for example in case of long kinetic measurements by cycling 2D-HSQC kinetics. Stable temperature requires the optimization of the heating unit – see in Figure 6a – the manual mode state length and power output, which depends mostly on the final temperature and the delay before heating pulses are applied also was adjusted. Additionally, after the rapid heating block, stabilizing RF heating pulses (loop 4) were applied. Power level, length and number of these pulses were optimized.



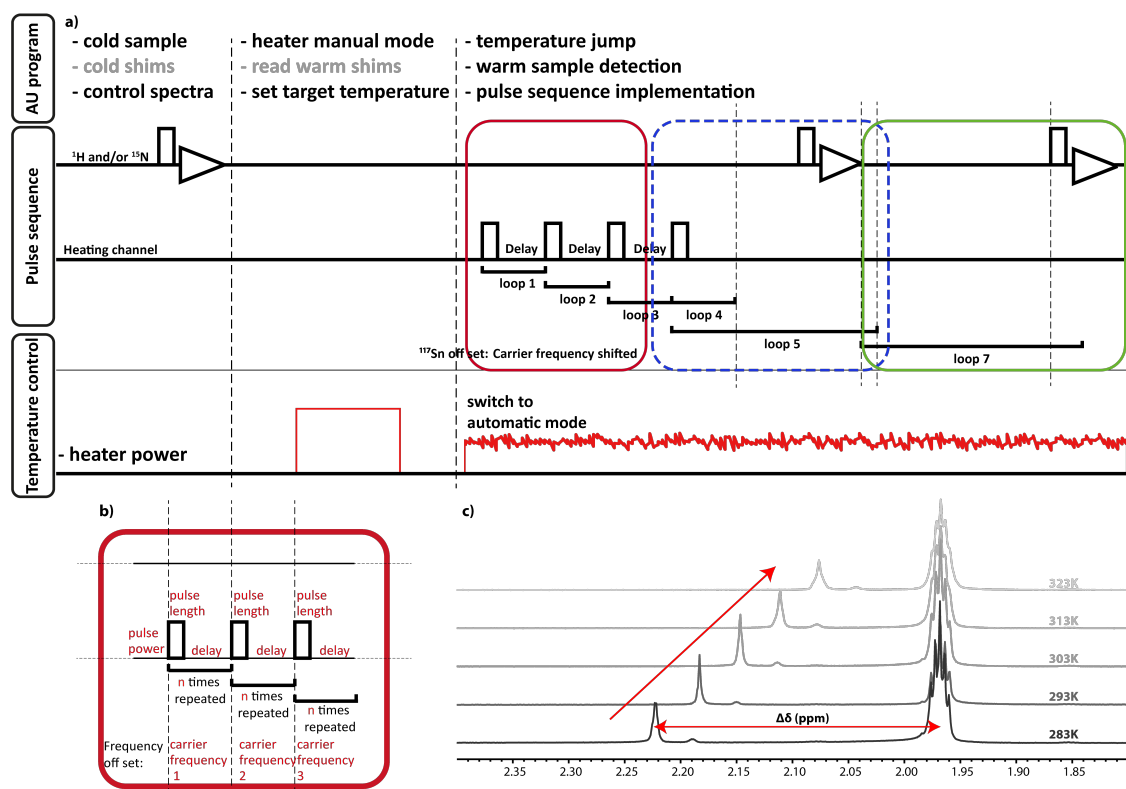


Figure 6 general schematic of a single T-jump experiment with relevant heating parameters. a) T-jump measurement overview showing. Upper lanes represent channels with standard coils to irradiate Larmor frequencies of  $^1\text{H}$  and/or  $^{15}\text{N}$ . Heating channel shows possible timing of heating pulses. The temperature control lane shows the timing of the automatic and manual heating, controlled by an AU program. b) Detailed parameter overview of the T-jump element highlighted with red in a). c) Example of internal temperature determination as difference between temperature dependant and independent signal.

### 2.3.2. Internal temperature sensor

To follow the temperature changes during the kinetic measurement, an internal temperature “sensor” has to be calibrated. Well resolved temperature independent and dependant signals are selected. At different temperatures, the chemical shift difference (Figure 6c) is measured and by plotting  $\Delta\delta^1\text{H}$  against the temperature, a linear calibration curve is generated. Using the calibration curve, the real temperature can be determined from the spectrum.

### 2.3.3. Setting up cycling-2D NMR measurements

With optimized standard NMR parameters (pulse calibration, tuning and matching, shimming) and optimized heating setup the cycling-2D NMR experiments have to be setup. Each time-domain point of a two-dimensional NMR spectra is recorded as a separate kinetic experiment. An AU program has been written, see in Appendix, to setup the series of kinetic experiments with parameters changed for the correct phase cycle and  $t_1$  evolution time. The experiments are started with simple python script accommodating

temperature changes for cooling and required delay times are optimized for equilibrium or non-equilibrium denatured starting state of barstar, detailed in the following subchapters. The AU program to control the variable temperature unit of the spectrometer and synchronise with the pulse sequence is achieved by an AU program, written by Thorsten Marquardsen, Bruker Biospin GMBH.

#### 2.3.4. Heating cycles for equilibrium state

The sample was kept one hour at low temperature – denaturing conditions – to ensure the equilibrium conformational state of the proline residues. In each kinetic experiment measurement time at high temperature – folding conditions – were 20 minutes. After cooling, the sample was kept at low temperature – denaturing conditions – for 30 minutes before next T-jump kinetic experiment was done. The first four cycles of kinetic experiments were used as dummy scans to ensure all kinetic experiments were starting from same proline conformational state.

#### 2.3.5. Heating cycles for fast folding pathway

Sample was kept one hour at high temperature – at folding temperature – to ensure the equilibrium conformational state of the proline residues. The cooling time was kept at 2.5 minutes before T-jump is initiated directly after 2.5 minutes. In each kinetic experiment measurement time at high temperature – folding conditions – were 2 minutes. The cooling time was short 2.5 minutes ensuring that equilibrium conformation of the unfolded state is NOT reached before next T-jump kinetic experiment (Tyr47-Pro48 is mostly in *cis* conformation at the beginning of the T-jump experiment). The first four cycles of kinetic experiments were used as dummy scans to ensure all kinetic experiments were starting from same proline conformational state.

### **3. RESULTS AND DISCUSSION**

#### **3.1. MOTIVATION**

Protein folding is a major focus of biophysics even 50 years after Levinthal described the protein folding paradox for the first time. There have been numerous methods developed to study the underlying mechanism. Kinetic experiments are among the most intriguing to study not just the equilibrium unfolded and folded state, but also the intermediate states during the folding process and their activation energies. Usually chemical mixing, pressure or light is used to shift the equilibrium from an unfolded state to the direction of a folded state. These methods have already been successfully combined with NMR spectroscopy which offers detailed structural information. Albeit temperature jump can also be used, so far it has seen limited application in combination with NMR spectroscopy. In this part of the thesis I will present the development of combining rapid temperature changes, as a trigger for kinetic measurements, with high resolution NMR spectroscopy to investigate protein folding mechanism.

#### **3.2. OPTIMIZED T-JUMP SETUP AND SUITABLE MODEL PROTEIN**

Required optimization to study protein folding via NMR utilizing a T-jump can be divided into two major parts: instrument and sample. On the one side the spectrometer had to be optimized to significantly decrease the time needed for a temperature jump, without rebuilding the existing T-jump probehead.

On the other side, a suitable target protein needed to be selected and the sample composition had to be optimized for T-jump experiments. The target protein should cold denature above water freezing point and fold under conditions that can be achieved with the available experimental setup. Preferentially can be recycled between folded and unfolded state without aggregating. Furthermore, additional knowledge about the folding mechanism was desired to be able to compare obtained results to proposed folding mechanisms and results of previous studies obtained with different methods. Barstar was fitting these criteria to a great extent. Albeit wt-barstar could not be cycled between folded and unfolded state without certain amount of aggregation. Therefore, two cysteine were mutated to alanine to solve this problem. Finally, also the buffer conditions were optimized to match the applied radio frequency and maximize T-jump efficiency.

## 3.2.1. Protein sequence optimization for cycling T-jump experiments

Temperature cycling of wild-type barstar between unfolded and folded states has been performed at varying protein concentrations from 100  $\mu\text{M}$  to 3 mM in a thermocycler (see Figure 7 for thermo cycling scheme). Afterwards the samples have been analysed by analytical size exclusion chromatography. As it is clearly visible by the chromatograms in Figure 7 wt-barstar starts to aggregate even at the lowest tested, 300  $\mu\text{M}$  concentration.

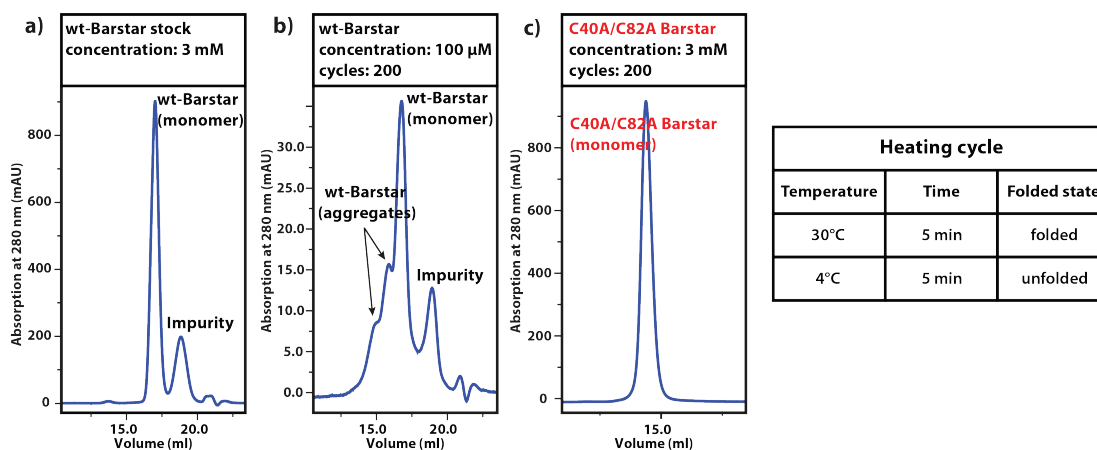


Figure 7 Size-exclusion chromatogram of a) 3 mM wt-barstar b) 100  $\mu\text{M}$  wt-barstar after 100 heating cycles c) 3.0 mM C40A/C82A barstar mutant after 198 cycles. All samples were in 10 mM  $\text{K}_2\text{H}_2\text{P}_2\text{O}_7$ , pH 6.4, 115 mM KCl and 2.2 M urea buffer - with additional 5 mM DTT for wild type.

It was clear that the aggregates of wt-barstar are mainly dimers, tetramers. A plausible explanation is that during the unfolded state exposed Cys40 and Cys82 residues can form intramolecular disulphide bridges. According to literature a cysteine-less mutant version of barstar is widely used to avoid disulphide chemistry during folding studies.<sup>[3,58]</sup> By point mutation (see in chapter 2.1.4) generated C40A/C82A mutant showed no aggregation even without reducing agent after several hundred heating cycles at very high 3 mM protein concentration, in agreement with literature. This opened the possibility to design NMR experiments utilizing the reversibility of the system. Such cycling T-jump experiments are discussed in detail in later chapters.

## 3.2.2. Buffer optimization for heating efficiency

Optimal T-jump experiments require “matching” the concentration of salt<sup>[41,42,80]</sup> to the applied radio frequency. The used 220 MHz radio frequency for heating has its optimal value between 100 mM and 130 mM salt concentration.<sup>[42]</sup> Applicability if the literature results have been verified by a series of KCl solution without protein, urea and buffer.

Therefore with the final protein ( from 2.5 mM to 3 mM for  $^{15}\text{N}$  and  $^2\text{H}^{15}\text{N}$  isotope labelled protein samples respectively) in buffer (10 mM  $\text{K}_x\text{H}_y\text{PO}_4$ , pH 6.4) with urea (2.2 M and 1.7 M for  $^{15}\text{N}$  and  $^2\text{H}^{15}\text{N}$  isotope labelled protein samples respectively) concentration of KCl was tested between 110-130 mM in 5 mM steps. The optimal value was found to be at KCl concentration of 115 mM.

### 3.2.3. Improvement of the duty-cycle of the radio frequency heating

The T-jump setup used in this project could previously achieve a 20 °C jump in one second.<sup>[42]</sup> This was limited by the maximum duty cycle of the power amplifier (300 W in frequency range 200-220 MHz). The probehead's cage coil used for heating has no limit regarding applied duty cycle, and could even support a continuous mode, but the power amplifier presented a bottleneck. By exchanging the power amplifier from 300 W to 1000 W the duty cycle limit could be increased. Increase of the duty cycle up to 90% was tested and no limitation in power output was observed. It is shown in Figure 8 as the duty cycle is increased the overall time needed for T-jump is decreasing. By reducing the time for a T-jump state-correlated spectroscopy experiments as detailed in chapter 3.8.3. have been made possible.

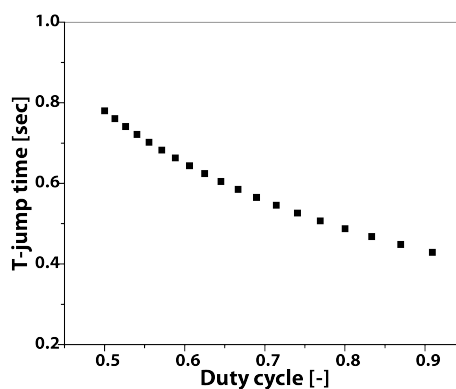


Figure 8 Correlation between phase cycle and T-jump time (dead-time), achieved by 1000 W power amplifier. Each experiment was recorded with 100  $\mu\text{s}$  pulses with 80 W power and decreasing delay between heating pulses.

### 3.3. T-JUMP EXPERIMENTS WITH ONE-DIMENSIONAL RT-NMR

The easiest experiment to record in RT-NMR is a simple one-dimensional proton spectrum. This has several downsides in case of following the refolding process of barstar. Even considering a single state 89 amide signals are expected in the 6-10 ppm range. Even in case of the simplest folding pathway with no stable intermediate the two states – folded and unfolded – present would lead to possible 188 amide signals. The resolution of such spectrum is limited as can be seen in Figure 9. Even in the spectra

before T-jump not all signals can be resolved, due to the limited dispersion of amide signals in the denatured state.

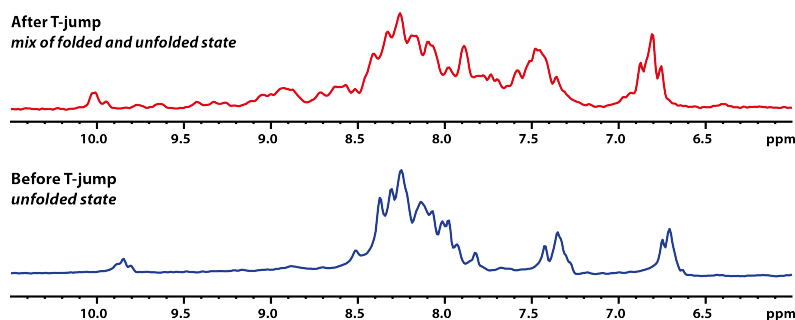


Figure 9  $^{15}\text{N}$ -filtered  $^1\text{H}$ -NMR spectra of barstar before T-jump (blue) and during refolding (red) after T-jump initiated refolding.

In the spectra after T-jump both states exist in parallel and unfolded signals severely overlap with folded signals. Therefore, this type of experiment without well resolved reporter signals was not suitable for detailed study of barstar folding.

### 3.4. CYCLING T-JUMP EXPERIMENT TO MEASURE TWO-DIMENSIONAL SPECTRUM

Introducing an additional dimension to separate signals would diminish the resolution problem of one-dimensional experiments. The combination of reversible cold denaturation and the stable T-jump system allows recording of two-dimensional NMR experiments by repeating the same folding experiment and recording several real-time experiments. In each repetition the pulse sequence, phase cycle or the evolution time can be changed. Schematic of this method where  $t_1$  evolution time and phase cycle is modulated in each real-time NMR experiment is shown in Figure 10. Concatenating the corresponding time points from each experiment results in a standard two-dimensional spectrum. By applying this method, a series of HSQC spectra could be recorded that gave reporter signals of barstar at different timepoints of the folding process. Acquired spectra and extracted kinetic information can be found in chapter 3.6.

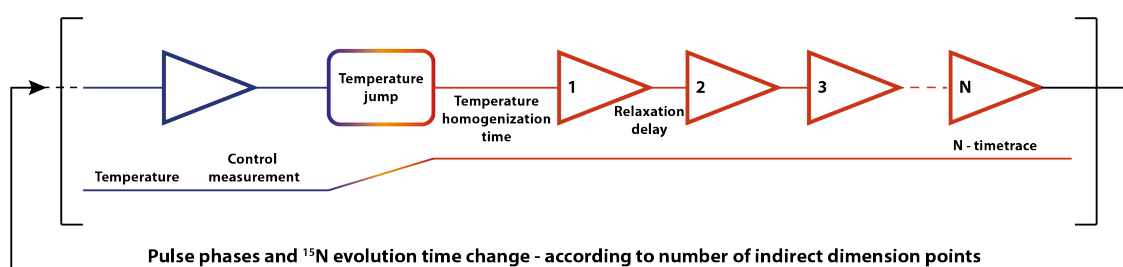


Figure 10 Schematic of two-dimensional NMR by cycling. Between the cycles the  $t_1$  increment delay and the pulses phases have to be changed according to the desired spectral width in the indirect dimension and to the States-TPPI quadrature detection.<sup>[81]</sup>

The pulse sequence to record such HSQC spectrum and the script to set up series of kinetic experiments for cycling RT-NMR is found in the appendix. The script to concatenate together the same time points of the RT-NMR spectra into a time resolved series of standard HSQC spectra can also be found in the appendix.

### 3.5. COLD DENATURED STATE - CONFORMATIONAL HETEROGENEITY

The assignment of the cold denatured state of barstar was published for amide residues of the backbone.<sup>[72]</sup> To be able to investigate the conformational state of the Tyr47-Pro48 and Leu26-Pro27 amide bonds additional nuclei assignment was required. Furthermore, slight shifts of the backbone due to different buffer conditions and low dispersion in cold denatured state did not allow direct transfer of the published assignment.

Proline residues have been shown to be important on the folding pathway of barstar,<sup>[7,58,62]</sup> and the Tyr47-Pro48 amide bond is in the unusual *cis* conformation in the folded state.<sup>[78]</sup> <sup>13</sup>C-detected proline selective CaCON type experiments were measured at different temperatures.<sup>[82]</sup> In the unfolded state four different cross peaks were observed Figure 11a shows that both Tyr47-Pro48 and Leu26-Pro27 amide bonds have mixed conformational states, while in the folded state (Figure 11b) only two cross peaks were observed.

Considering that the thermodynamically favoured conformation is the *trans* conformation and in the denatured form free rotation is not hindered the major conformation of Tyr47-Pro48 amide bond is likely *trans*. The mixed conformational state suggests that different folding pathways exist, depending on the starting conformation.

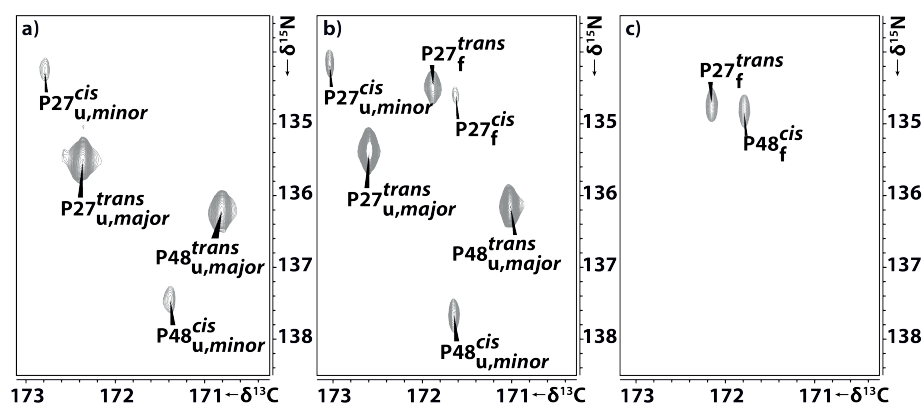


Figure 11 <sup>13</sup>C-detected proline selective CaCON type NMR spectrum<sup>[82]</sup> at a) 270K b) 288K c) 303K. a) cold denatured C40A/C82A barstar shows mixed conformational state of both proline residue. b) mixed folded and unfolded state shows 6 correlation while c) folded barstar shows only two conformational state.

According to literature a rate limiting step on the folding pathway exist with a timescale of a few minutes.<sup>[7,58]</sup> This is typical for proline isomerization. Experiments described in literature were all conducted from barstar denatured by chaotropic agents such as urea or GdnHCl. Secondary chemical shift analysis of cold denatured and urea denatured barstar show that propensities for certain secondary structures are different. Therefore, differences in folding pathways can be expected to be observed by different methods studying the effect of proline isomerization on the folding process. This is no surprise as the two ways of denaturation have different effects on the stable thermodynamic state. This is due to differences in solvent environment and the interaction between the protein and the solvent.<sup>[66]</sup>

### 3.6. SLOW PATHWAY – RATE LIMITING STEP

Equilibrium denatured barstar was followed by two dimensional RT-NMR, using HSQC pulse sequence adapted to T-jump measurements as shown in Figure 10. This has allowed recording of a time series of 510 HSQC spectra. The heating and cooling cycles were set to be long enough to reach an equilibrium state between T-jumps.

#### 3.6.1. Time resolved intensity changes

The analysis of the time dependant signal intensity of assigned and resolved peaks could be fitted either with mono- or double exponential function. Representative curves are shown in Figure 12. Although the majority of signals show either mono- or double exponential behaviour, the amplitude of the curves differs between residues. Furthermore, the faster phase of double exponential time traces shows different rates, while the slow phases are generally the same order (see Figure 13). Mono exponential curves, except for Q72, show similar rates compared to the slow phase of double exponential curves, therefore they are likely to be involved in the same step of the folding pathway. The difference between the folding rates of folded and unfolded signals are indirect evidence for intermediate states on the folding pathway.



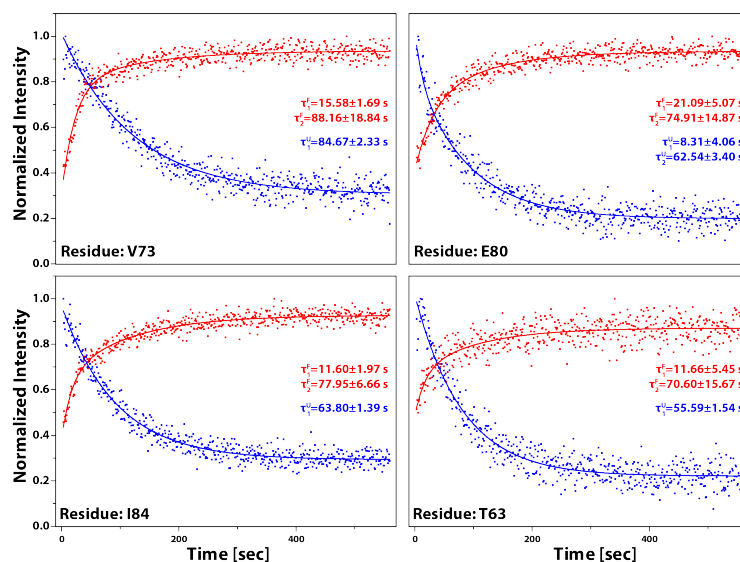


Figure 12 Time dependant signal intensities of representative residues. Colour coding: red is the folded signal and blue is the unfolded signal of the same residue shown below the time traces. Lifetime  $\tau$  from mono- and double exponential fits are shown.

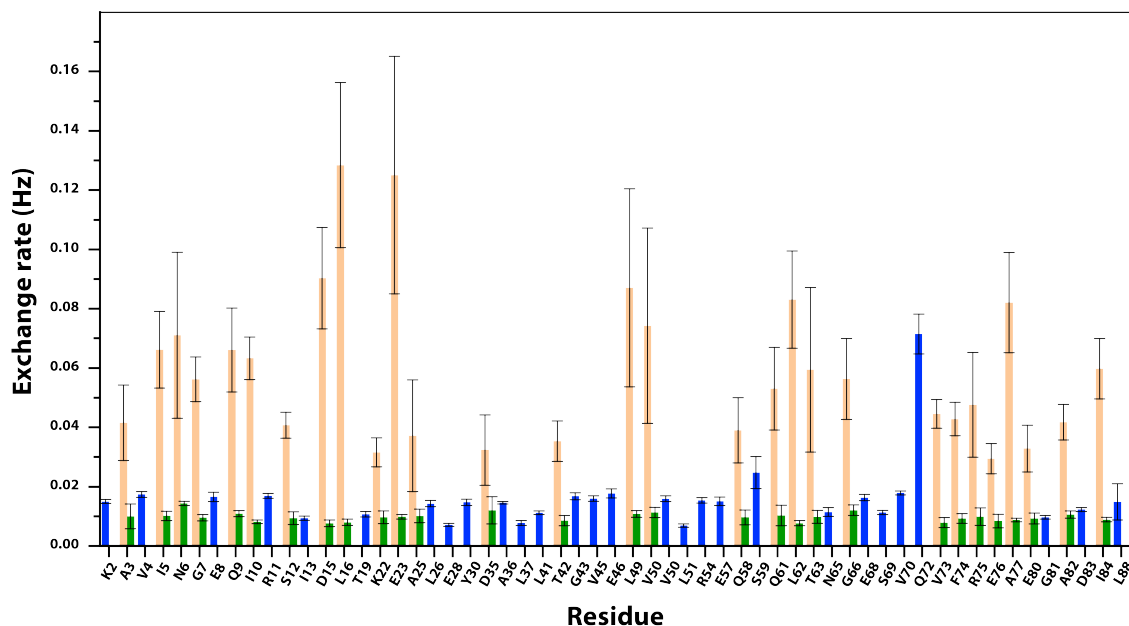


Figure 13 Bar diagram of exchange rates determined by fitting well resolved folded peaks with mono (blue)- or double-exponential (green slow phase and orange fast phase) decay functions. Error bars represent standard error of fit.

### 3.6.2. VTU initiated refolding – detection of intermediate signals

The rate limiting step of the slow folding pathway takes up to a few minutes, this allows following structural changes during this step with long dead times. Such RT-NMR measurements of the refolding process from urea denatured P26A/C40A/C82A mutant barstar using series of HSQC spectrum has been published<sup>[63,83]</sup>. In this case refolding was initiated by manual dilution of urea concentration resulting in 100 seconds of dead time and it took an additional 100 seconds to record the first HSQC spectrum. Even with

the relatively low time resolution and very long dead time allowed detection of intermediate signals. Similar idea was implemented for cold denatured barstar. Modern VTU units could achieve heating dead time of 80-100 seconds which is suitable to study the changes happening during the rate limiting step of the refolding of cold denatured barstar. The advantage of using this setup is the higher sample volume and cryoprobes used for detection resulting in higher S/N. The aim of this experimental setup was to observe possible intermediate signals.

Series of HSQC spectra were recorded and we observed similar folding behaviour as in the T-jump experiments. The first spectrum recorded after refolding was initiated still showed high population of unfolded signals. Analysing the time dependant signals showed mono exponential decay. The time constants of the mono exponential decay fit matched very well with the slow phase of double exponential decay fit of the T-jump experiments. This means the faster phase of double exponential kinetics could not be observed due to the much longer dead time of the experimental setup but the higher sensitivity allowed the detection of intermediate signals. The T-jump experiment is better suited for kinetic analysis even for folding kinetics slowed down by the *trans-cis* isomerization.

On Figure 14 the upper two spectrum shows the first HSQC spectrum recorded after refolding was initiated. There are clear intermediate signals present and their time dependant signal intensities showed mono exponential decay. Most likely these intermediates formed during an earlier step on the folding pathway, as no build up curve was observed. The chemical shifts could not be assigned to specific residues, but their chemical shift range falls into the range of amide proton chemical shift representative of structured proteins. This suggests that these intermediates have folded structure but different from the final folded state.

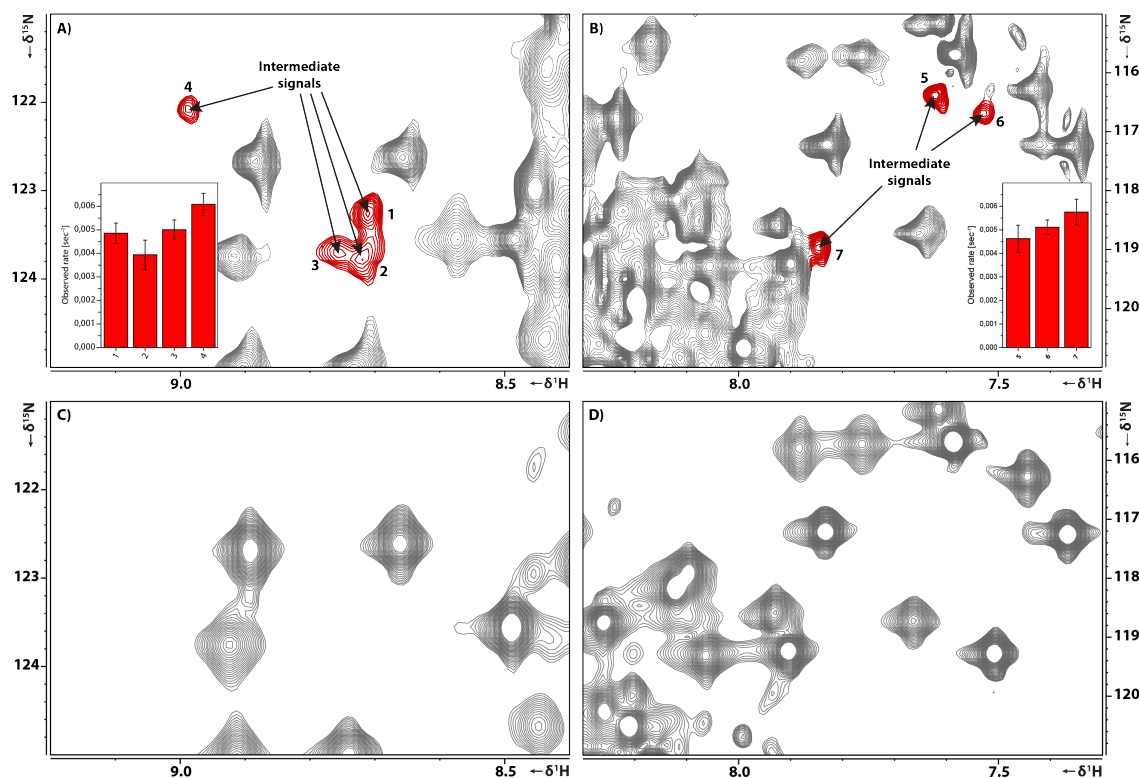


Figure 14 NMR spectrum of representative regions of barstar. The spectrum was recorded on a 700 MHz spectrometer equipped with triple resonance cryoprobe. The upper two spectrum a) and b) shows clear intermediate signals, and the bottom two spectrum c) and d) shows the same region after folding was finished with no signals at the position of the intermediate signals. Bar diagrams as inlet show the corresponding rates from mono exponential fit of the time traces of intermediate signals.

In comparison to urea initiated refolding experiment to follow the *trans-cis* isomerization of Tyr47-Pro48 in the mutant P27A/C40A/C82A barstar showed lower population of unfolded signals but intermediate signals were also observed. The limited number of unfolded signals detected by Fersht et al. could be due to the overall lower stability (around 70% compared to C40A/C82A mutant)<sup>[65]</sup> of the mutant used in their study. The reduced stability could lower the energy level of the transient state therefore speeding up the refolding process. Interestingly the intermediate signals detected in their experiment shows proton chemical shifts in the unfolded region, while the intermediates detected in our experiments clearly come from structured region. These differences could be attributed to the P27A mutation and the different denatured state. It would be interesting to repeat T-jump experiments with the P27A/C40A/C82A mutant to have a clear comparison of the kinetics in the future.

### 3.7. INFLUENCE OF ISOMERIZATION ON THE FOLDING PATHWAY

The slow and fast phases of the exponential decay fits were mapped onto the three dimensional structure (see in Figure 17). Our theory is that the residues involved in the

slow phase are involved in the stabilization of the tertiary structure contacts. We concluded that barstar forms its final tertiary structure in this last step. For example, the helix<sub>1</sub> and the strand<sub>1</sub> of the  $\beta$ -sheet likely gets closer together as only residues with contacts between these two building blocks show changes. Other secondary structure elements show similar effect, which could be accounted to the relative orientation changes or formation of stabilizing of these building blocks. Work from Fehrst<sup>[58]</sup> and Udgaonkar<sup>[7]</sup> showed stopped-flow measurements using different detection techniques that the tertiary structure formation is the last step on the folding pathway, in agreement with our observation. Such as the tryptophan fluorescence experiments which report mostly on the changes around the Trp53 residue, buried in the hydrophobic core in the folded state. The time constants of the fluorescence exponential kinetic falls into the same range as the slow phase time constant in our experiment. This suggest that the Trp53 residue gets buried in the rigid core of barstar as the last step of the final tertiary structure formation. Tertiary structure changes were also monitored by ANS fluorescence which changes upon binding to hydrophobic surfaces of barstar and by ellipticity changes at 270 nm<sup>[7]</sup> wavelength, in previous studies. Udgaonkar et al. reported that the ANS cannot bind to the hydrophobic surface neither in denatured nor in fully folded barstar, but it binds to intermediate states. Upon formation of the final tertiary structure ANS is pushed out of the hydrophobic core and cannot bind to barstar. The observed widespread effect on the structure, mostly on residues in contact with other secondary structural elements was to a lesser degree observed also in RT-NMR studies on the P27A/C40A/C82A mutant barstar as well.<sup>[63,83]</sup> The differences between these experiments could be attributed to the destabilization by P27A mutation which could influence the stability of the intermediate on the folding pathway compared to the C40A/C82A mutant. Udgaonkar also performed RT-NMR study on the folding of barstar from GdnDCI denatured state which was limited to <sup>1</sup>H-NMR spectroscopy but sufficient reporter signals could be analysed. Their results showed multistate kinetics, in agreement with our fast and slow phase kinetics.

If we look at residues involved in the fast phase, they are mostly limited to parts of the sequence with defined secondary structure. This would require further analysis but there is a plausible explanation that certain secondary structures are formed later and can be still observed in our experiment. The formation of secondary structure is supported by the findings of Udgaonkar et al., where they characterized one late intermediate structure

by ANS binding experiment and using CD spectroscopy 222 and 270 nm wavelengths. The molar ellipticity intensity at 222 nm, reporter on secondary structures, is only 20% of the final intensity, showing the lack of complete secondary structure. They concluded that not all secondary structure is formed in this state, but secondary structures are only partially completed and more is formed in a later step. Additionally, steady state fluorescence and the correlation time measurements also suggest the still ongoing process of secondary structure formation.<sup>[64]</sup> Interestingly in early studies, they have shown that the early intermediate with the wrong *trans* Tyr47-Pro48 amide bond is already active. Therefore, it is usually referred to as native-like intermediate. Interestingly the fast phase of kinetics is mostly observed on residues with defined secondary structure and avoid the binding site (residue 29-46 around helix<sub>2</sub>) of barnase. Additionally, fragments of barnase consisting the binding site are able to partially inhibit barnase suggesting that the final folded structure is not required for functionality. The lack of any changes around the binding site in our observation is in agreement with intermediates being active. The binding site is formed, capable of inhibiting barnase, while the other parts of the protein still forms its final structure. Thanks to the atomic resolution of NMR spectroscopy in combination of T-jump setup we could gain more insight into the final stages of the folding pathway.

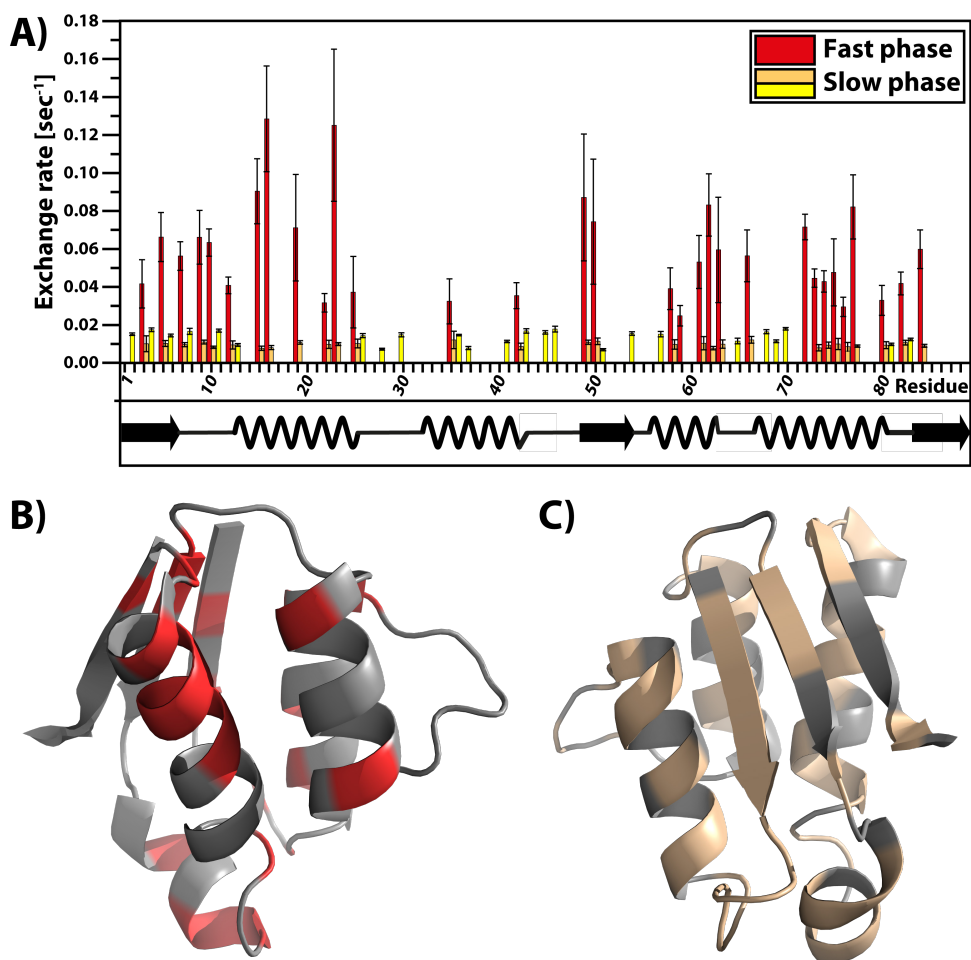


Figure 15 A) Bar diagram of double and mono exponential fit results of the build-up kinetics of final folded signals in sequential order. Under the bar diagram secondary structure representation of barstar. B) and C) are the three dimensional structure of barstar (PDB: 1AB7)<sup>[51]</sup>, with residues in red for fast and teal for slow phases of the time resolved NMR signal intensity.

### 3.8. FAST FOLDING PATHWAY

Barstar has several steps and stable intermediates on the folding pathway and depending on the starting conformation of the Tyr47-Pro48 amide bond has at least two different folding pathways, a slow and a fast for *trans* and *cis* conformation, respectively. In stopped-flow measurements double jump experiment is a quite common method. In the first jump the protein is transiently unfolded – keeping proline residues in the native state. Then in a second jump protein folding is initiated from a non-equilibrium denatured state. This method circumvents rate limiting step of *cis-trans* isomerization pushing the folding mechanism towards the faster pathway.

I applied the same idea for temperature-initiated folding of barstar. Transiently unfolded protein can be generated if the cooling time is kept relatively short compared to the refolding of proline residues from native state (Tyr47-Pro48*cis* and Leu26-Pro27*trans*)

to cold denatured equilibrium state (mixed proline conformation, see Figure 11). Transiently denatured barstar can then be refolded by rapid radio frequency heating. This allowed the design of new state-correlated experiments with T-jump NMR spectroscopy to study the faster folding pathway explained in detail in the following chapters.

### 3.8.1. Heating cycles for double-jump experiment

First heating cycles had to be optimised whether it is possible to transiently denature barstar. The most critical parts are how fast the *cis-trans* isomerization of the amide bond is and how fast the sample can be cooled down.

The cooling time was monitored by series of  $^1\text{H}$ -spectrum following the DSS signal changes. The signal of DSS is temperature independent but due to the spectrometer being locked on the deuterium frequency of HDO – which is temperature dependant – the DSS signal shifts based on the temperature changes. If the final temperature is reached in the sample, the chemical shift of the DSS signal must be constant and show no linewidth broadening. The changes observed on DSS signal is shown in Figure 16. With the adjusted P.I.D. control of the VTU unit (P=7.4 I=6.0 D=2.0), the required time to cool down the sample (80  $\mu\text{l}$  volume) by 20  $^\circ\text{C}$  was 2.5 minutes.

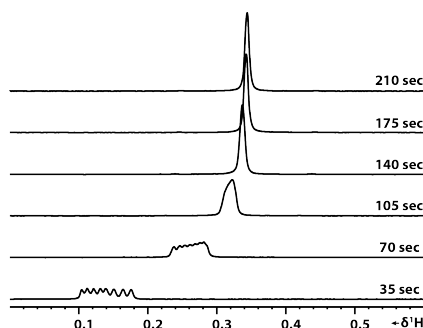


Figure 16 Time series of  $^1\text{H}$  NMR spectrum, focusing on the DSS signal as the temperature is changing. Each spectrum is an average of 8 scans. After 175 sec the signal is not changing, the temperature is stable. The temperature was changed from 293 K to 273 K.

To address how fast *cis-trans* isomerization proceeds in the cold denatured state an additional test was made by utilizing the two-dimensional cycling HSQC T-jump measurements, described in chapter 3.4. In Figure 17 the spectrums before and after the T-jumps are shown, with either enough equilibration time (30 min) at low temperature or with minimal cooling time (2.5 min) between the T-jump experiment repetitions. In case of the short cooling time, a minimal, below 5%, folded population is still observable before the T-jump but it can be neglected for later experiments. If we compare the spectra after the T-jump in case of the shorter cooling time, the folding is almost finished within



the experiment dead time. This means the Tyr47-Pro48 amide bond remained mostly in *cis* conformation before RF heating was applied, avoiding the rate limiting step of *trans* to *cis* isomerization. Double jump experiments are suitable to shift the folding mechanism to the faster pathway by eliminating the rate-limiting step of *trans-cis* isomerization.

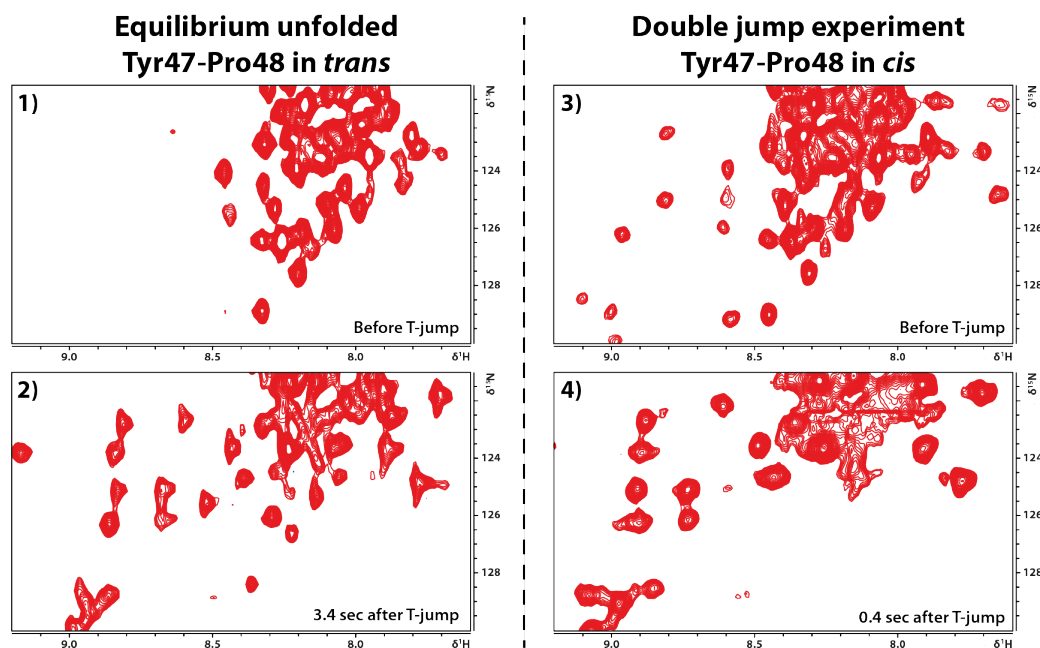


Figure 17 Comparing the effect of short and long cooling time on the folding mechanism of barstar. All spectrum were recorded according to pulse scheme shown in Figure 10 written for HSQC T-jump experiment. In 1) and 3) shows the spectrum before T-jump was executed and in 3) lower contour level was used to show that less than 5% of folded barstar is still visible. 2) and 4) shows the same spectral region as in 1) and 3) after a short delay. Comparing 2) and 4) shows that depending on the starting conformation of the amide bond higher population of unfolded signals are observed in 2).

### 3.8.2. State-correlated spectroscopy pulse sequence

State-correlated (SC) spectroscopy was first introduced by the group of Akasaka in 1990. It was similar to a NOESY or EXSY type of spectrum, but during the mixing time a physical parameter change is introduced<sup>[38]</sup>. Conceptually any physical parameter which would change the sample equilibrium can be used. Although to record a two-dimensional spectrum these changes must be reversible and faster than the  $T_1$  relaxation time of the sample. The most suitable parameters are pressure or temperature. Akasaka showed that T-jump (using an adjusted microwave heating coil around the sample) can be utilized and recorded the first SC-spectrum of water, correlating the chemical shift changes of water at two different temperatures.

The limiting factor in their experimental setup was the requirement of  $D_2O$  samples. Also  $^1H$  nuclei are an order of magnitude faster relaxing compared to other heteronuclei ( $^{13}C$



or  $^{15}\text{N}$ ). If we look into other pulse sequences for correlations between exchanging conformational states,  $^{15}\text{N}$ -ZZ-exchange experiments<sup>[84,85]</sup> correlates two slowly exchanging conformational state. It has the advantage of higher resolution thanks to the nitrogen dimension additionally, the longer relaxation time of the  $^{15}\text{N}_z$  spin order is favourable, allowing longer mixing times to be used.

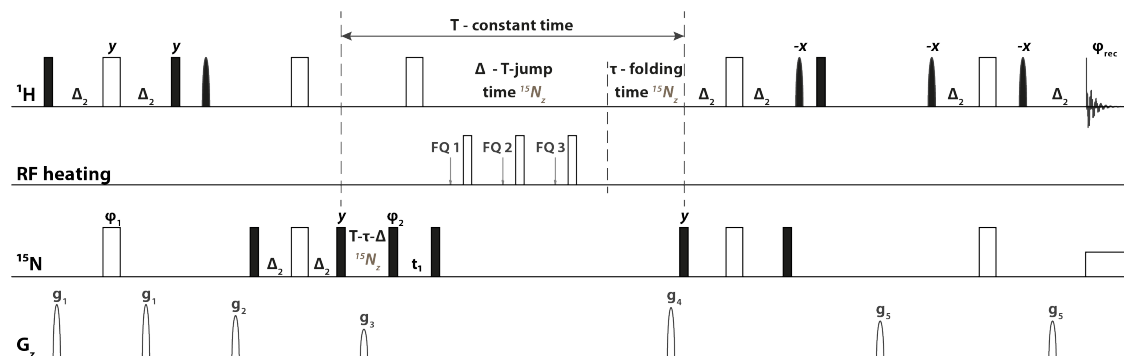


Figure 18 Pulse scheme for state-correlated (SC) T-jump experiment. Narrow filled bars represent  $90^\circ$  pulses, wide unfilled bars represent rectangular  $180^\circ$  pulses. Gradients are shown as semi-elliptic unfilled shapes. Semi-elliptic filled shapes represent selective  $90^\circ$  pulses ( $^{13}\text{C}$  Q5 600  $\mu\text{s}$ , on resonant). If not depicted otherwise pulses are applied along the x-axis. During acquisition  $^{15}\text{N}$ -nuclei is decoupled using asynchronous GARP4 sequences<sup>[86]</sup>. Gradient pulses is applied for 1 ms with a smoothed square amplitude (SMSQ10.100) and 100% gradient strengths corresponds to 53 G/cm, the gradients are: g1 (50%), g2 (80%), g3 (20%), g4 (40%), g5 (30%). Pulse phases are the following:  $\phi_1 = 2(x)$ ,  $2(y)$ ,  $2(-x)$ ,  $2(-y)$ ;  $\phi_2 = x$ ,  $-x$ ,  $y$ ,  $-y$ ,  $-x$ ,  $x$ ,  $-y$ ,  $y$ ;  $\phi_{\text{rec}} = x$ ,  $-x$ . The delay  $\Delta_2 = 1/(4 \cdot J_{\text{NH}})$  2.71 ms,  $\Delta_1$  was adjusted according to desired temperature jump typically 200-500 ms corresponding to 10-25°C jump;  $\tau$  is adjusted according to required folding time. During T-jump element FQ 1, FQ 2, FQ 3 are frequency jumps to compensate for the temperature changes inside the sample.

I designed a new pulse sequence based on the combination of the SC- and  $^{15}\text{N}$ -ZZ-exchange experiments, also needs to be mentioned that Charlier et al. used a similar approach for P-jump experiments recently but with slightly different timing of the P-jump element.<sup>[27]</sup> The new pulse sequence is shown in Figure 18. First the  $^1\text{H}$  magnetization is transferred to the nitrogen and changed to the slowest relaxing spin order  $\text{N}_z$  in order to minimize the loss of magnetization during the compensation time (T-t- $\Delta$ ). The compensation time is introduced to keep the pulse sequence length constant between experiments with different  $\tau$  folding time. Afterwards  $^{15}\text{N}_z$  spin order is transferred to  $\text{N}_y\text{H}_z$  transverse spin order to evolve the nitrogen chemical shifts (the indirect dimension). After the  $t_1$  evolution magnetization is transferred back again to  $\text{N}_z$  spin order to minimize the signal loss due to relaxation during the T-jump part and t folding time. Finally, the magnetization is transferred back to  $\text{H}_x\text{N}_z$  spin order to let  $^1\text{H}$  chemical shift evolve.

### 3.8.3. State-correlated spectroscopy

With the optimized heating cycle and the new pulse sequence SC-spectrum was recorded. Although rapid jumps for 10-20°C can be achieved within a few hundred milliseconds, the length of the jump can still be close to the  $T_1$  relaxation time of proteins. Therefore, the compensation time in the pulse sequence was set to zero ( $T-\tau-\Delta$ ) in order to maximize the S/N. An overview of a state-correlated experiment can be seen in Figure 19. Control measurements before and after the exchange spectrum were also recorded, with the exact same pulse sequence except for the RF heating pulse power was set to zero. This results in a similar spectrum to a standard HSQC but with the same relaxation effect as in the SC-spectrum.

The T-jump time was 390 ms to achieve a jump from 273 K to 298 K. At the end temperature the population of unfolded signals were below 5%. The refolding time was set to 200 ms, and the complete pulse sequence with detection was 700 ms long. The experiment was repeated twice to achieve a total number of scans of four for sufficient S/N. The assignment of a state correlated spectrum is achieved by comparison of reference spectrum recorded at low and high temperature. An example peak assignment is shown in Figure 20.

If the folding process is finished before detection takes place, only cross peaks with  $^{15}\text{N}_U\text{-}^1\text{H}_F$  chemical shifts should be visible, if the folding process is not yet complete  $^{15}\text{N}_U\text{-}^1\text{H}_U$  or possible  $^{15}\text{N}_U\text{-}^1\text{H}_{\text{intermediate}}$  signals could be observed. The maximum folding time that is feasible is limited by the  $T_1$  relaxation time of the protein. Unfortunately, the folding of barstar even on the faster pathway is longer than 600 ms therefore complete refolding during an exchange spectrum recording could not be achieved. This is supported by the presence of unfolded signals in the SC spectrum even after 200 ms folding time (dark blue colour code in Figure 19). Interestingly some signals were not observed at the expected  $^{15}\text{N}_U\text{-}^1\text{H}_F$  position (yellow peaks in Figure 19), even though these residues are observed even at low and also at high temperature. This means these signals are not undetectable due to the fast relaxation, but likely present in an intermediate state. Additionally, there were also low intensity folded peaks observed from residual folded peaks at low temperature (orange in Figure 19), due to the shortened cooling time.

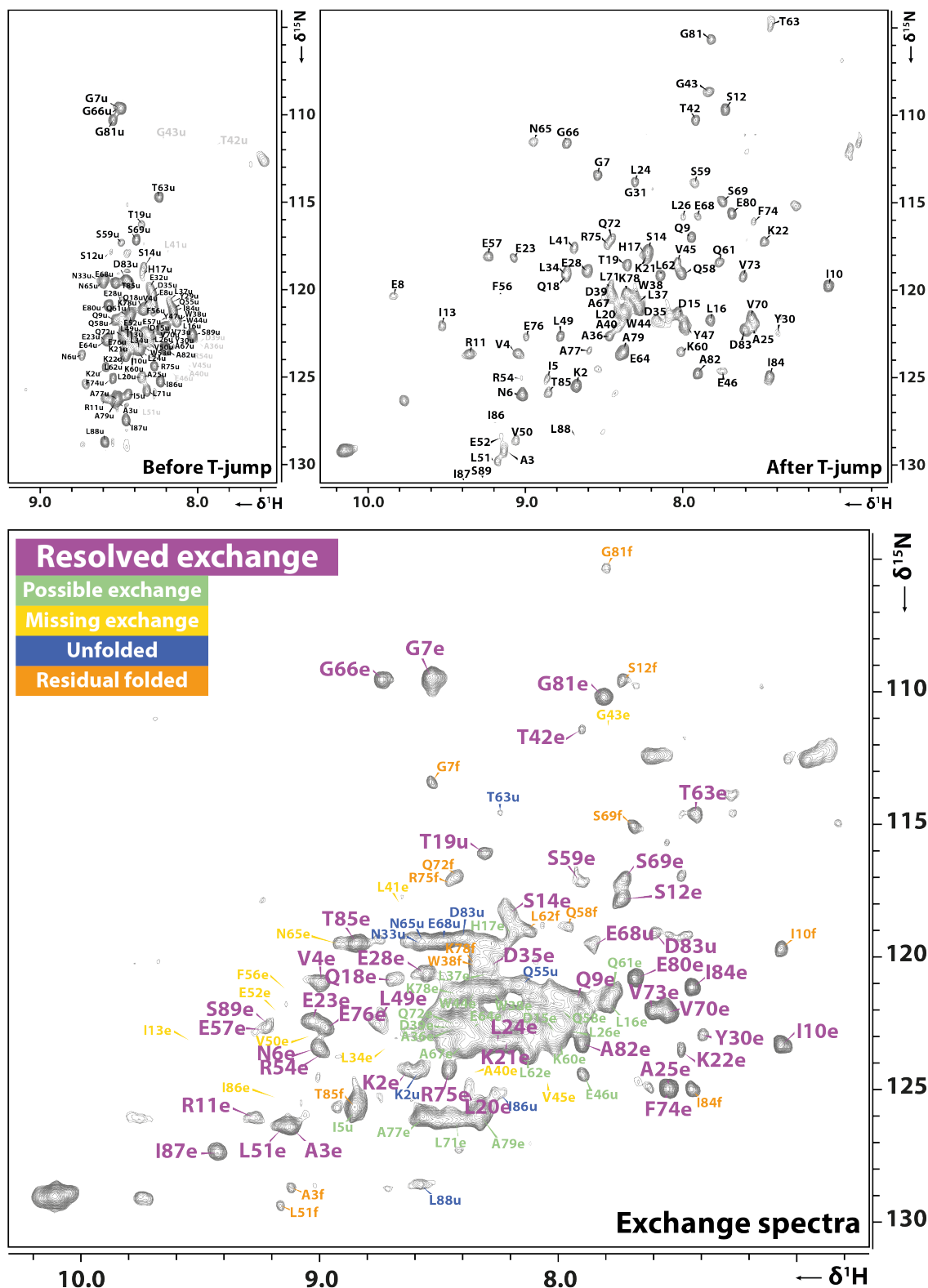


Figure 19 Overview picture of state-correlated T-jump spectrum with reference spectra before and after SC-spectrum. All the spectra were recorded with pulse sequence shown in Figure 18. The  $\tau$  folding time was set to 200 ms,  $\Delta$ -T-jump time was 390 ms corresponding to 25°C temperature jump, the compensation delay: T- $\Delta$ - $\tau$  was set to zero to achieve higher S/N. The upper left spectrum was measured with RF heating pulses set to zero at low temperature before the SC-spectrum, the upper right spectrum was measured with RF heating pulses set to zero at high temperature after the SC-spectrum was recorded. The lower spectrum shows the SC spectrum. The colour coding of the labels is shown in the upper left corner of the spectrum.

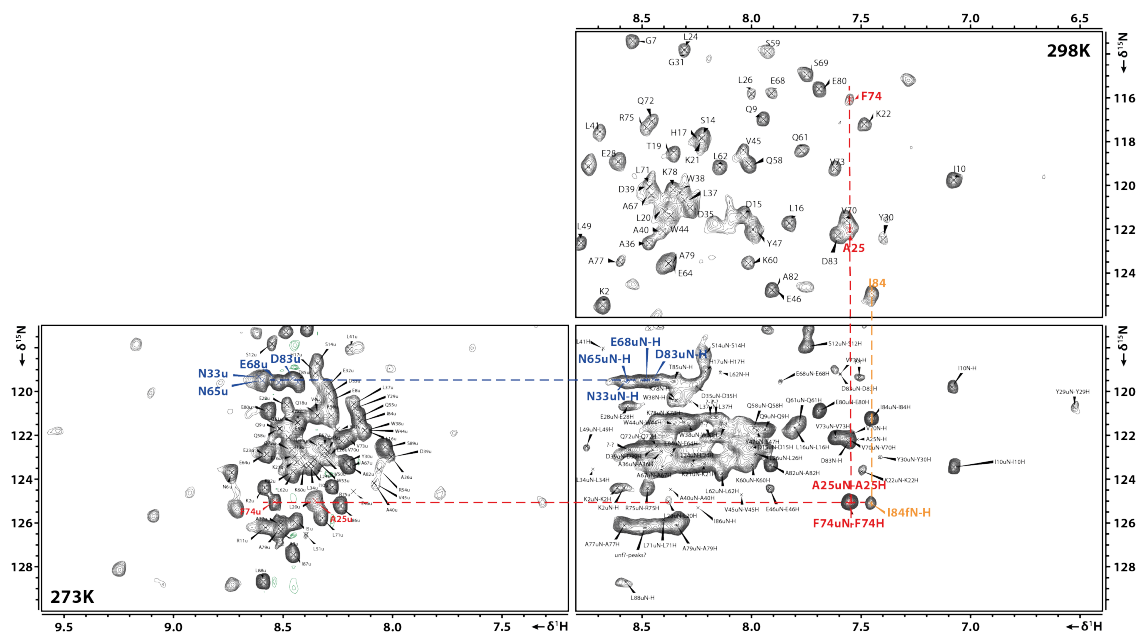


Figure 20 Assignment strategy for SC-spectrum. The SC-spectrum has  $^{15}\text{N}$  chemical shifts from the cold denatured state at low temperature and  $^1\text{H}$  chemical shifts from the final temperature. Reference spectrum at low and high temperature was recorded to aid the assignment.

### 3.8.4. Intermediate on the fast folding pathway

State-correlated spectrums recorded with 5 ms, 100 ms and 200 ms  $\tau$  folding times are like snapshots of the folding process at different time points. We can qualitatively analyse these snapshots that certain residues are rapidly reaching folded state. Interestingly, there are peaks for which only the exchange peak is observed (and no unfolded signals of the same residue) e.g.: I84, T85 etc, but their intensity is still increasing as the folding is increasing. While some residues e.g.: E68, E80, D83, Q55 etc even at the longest  $\tau = 200$  ms have high population of unfolded signals. In case of residues such as S69 and T63 both folded and unfolded signals are observable at 5 ms folding time, but the unfolded signals are unobservable in spectrum recorded with  $\tau = 200$  ms.

These differences between the residue's population of folded and unfolded signals mean that certain parts of barstar fold earlier while some other parts must fold slower. These findings already suggest the existence of intermediate states on the fast folding pathway. Additionally, there are residues from which no signal was detected, neither as unfolded nor as exchange peak. These residues are likely still in intermediate state(s).

This is in agreement with stopped-flow double jump experiment results indicating several intermediates also on the fast folding pathway.<sup>[7,58]</sup> Literature results showed that the earliest step on the folding pathway is the hydrophobic collapse of the unfolded polymer

chain with time constant below one millisecond and with no secondary structure. This has been shown by burst phase changes in the tryptophan fluorescence, while ellipticity measurements at 222 nm reporting on secondary structure formation showed no changes in the burst phase of the stopped-flow experiment. Direct evidence has been provided later by electrical discharge T-jump setup where the sub millisecond event could be detected reporting a rapid change of barstar, but no formation of secondary structure was detected.

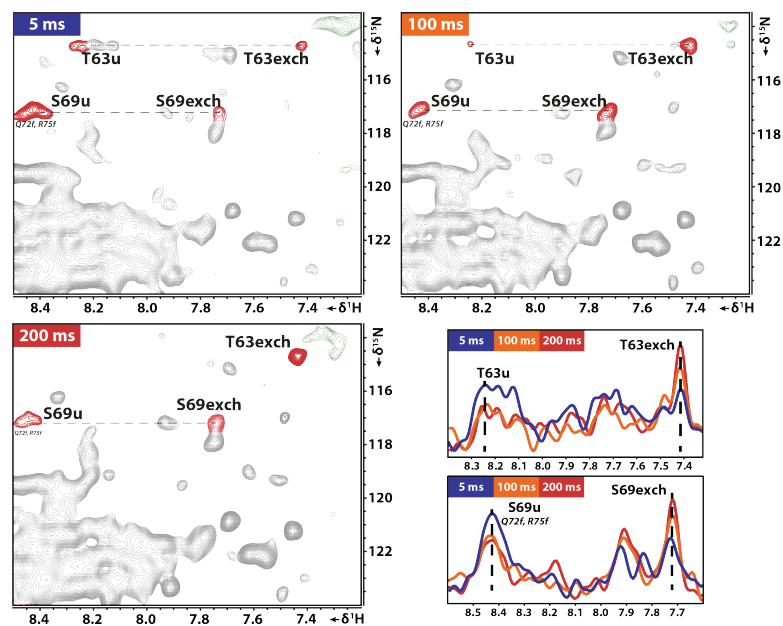


Figure 21 State-correlated spectrum with different  $\tau$  folding times (show in upper left corner). T63 and S69 unfolded and exchange peaks are highlighted. In the lower right corner overlay of projections from the cross peaks is shown with colour codes showing the corresponding  $\tau$  folding time. The unfolded peak intensity is decreasing, while the exchange peak intensity is increasing with increasing folding time.

From this collapsed state starts the formation of some secondary structures, which was shown to be in hundreds of milliseconds to second.

The earliest reporter signals from state-correlated spectrum are limited to the shortest 5 ms folding time. Partially folded state can be observed as some residues show already high intensity exchange peaks even with 5 ms  $\tau$  folding time. Although the intensity of exchange peaks has increased with longer folding times and unfolded signal intensities have decreased, no complete folded state could be reached. Although the reference spectrum recorded after the exchange experiment showed no sign of unfolded peaks, shown in Figure 19, meaning that the upper time limit of the folding is  $\sim 2$  seconds.

Fehrst et al. proposed condensation-nucleation as mechanism for the folding of barstar<sup>[3]</sup>, while hydrophobic collapse mechanism was suggested by Udgaonkar<sup>[7]</sup>. These two

theories are not mutually exclusive. Actually, hydrophobic collapse could explain the very first step (collapse of the unfolded polymerase chain with no secondary structure) of the folding pathway while nucleation condensation the later steps (different population of secondary structures of stable intermediates). In the initial collapse and thereof partially formed secondary structures serve as a core for the rest of the protein to fold onto. If we would consider a more complex mechanism where several parallel pathways exist, and the intermediate states are considered as conformational ensemble a more complex energy landscape could better describe all the possible folding mechanism.

#### 4. SUMMARY AND OUTLOOK

Protein folding studies usually rely on stopped-flow systems, which offer high time resolution but limited site specific information. In comparison NMR spectroscopy has the advantage of atomic resolution but usually only amenable to study slow folding systems. Recent developments in pressure jump systems showed to possibility to overcome the limited time resolution of NMR spectroscopy by using model protein ubiquitin that can reversibly denatured and refolded combined with new pulse sequences leading to very detailed structural description of the folding mechanism.

In this chapter I have shown that the C40A/C82A mutant of barstar can be readily “switched” between denatured and folded state by changing the temperature between 0-25 °C. Applying radio frequency irradiation inside the NMR spectrometer built into the probehead and as coherent part of pulse sequence, high spectral and time resolution series of two-dimensional NMR spectrum can be recorded. This allowed the observation of the slow folding kinetics at very high spectral resolution. This process is due to the *trans-cis* isomerization of Tyr47-Pro48 amide bond. By utilizing higher sensitivity cryoprobes stable intermediate signals were also shown on the folding pathway. Furthermore, analyses of the time dependant signal intensities of amide signals reporting on almost all residues could be achieved thanks to the now feasible cycling two dimensional T-jump measurement. The intensity changes showed both mono- and biexponential behaviour reporting on the different structural changes during the slow folding of barstar.

Utilizing a different heating cycle, I could circumvent the effect of *trans-cis* isomerization, allowing the investigate on of the alternative fast folding pathway. To overcome the time resolution limit of cycling 2D T-jump HSQC measurements, state-correlated pulse sequence was developed which correlated unfolded <sup>15</sup>N signals to folded <sup>1</sup>H signals. This experiment provided evidence for intermediates on the fast folding pathway by qualitative analyses of spectrums.

Some aspects of the T-jump setup could be further improved in future. As it stands now the optimized heating time is between 250-500 ms corresponding to 10-25 °C temperature change. This is still relatively long compared to usual protein folding times, especially to observe the early steps in the time regime of a few milliseconds. Although SC-spectroscopy is promising, for quantitative description of the changes even shorter

heating time is desired. Theoretically higher frequency heating pulses would allow more efficient heating of lossy water samples but could increase temperature heterogeneity after T-jump. Additionally, the maximum power level of the heating pulses could be increased if smaller radius of the sample tube is used as a compromise with reduced sensitivity. Alternatively, the  $^{15}\text{N}$  channel of the probehead could be changed to  $^{13}\text{C}$  utilizing the higher sensitivity and longer longitudinal relaxation time of methyl groups nuclei. This could allow for example in case of the state-correlated spectrum longer  $\tau$  folding times to be used, allowing barstar to reach complete folding during the measurement and the possibility to record more time points for quantitative analyses.

Reversible cold denaturation in combination with temperature jump NMR probehead was for the first time utilized to gain site specific information about protein folding. In combination with pressure jump developments in the recent years, this could open the possibilities to study the structure of stable intermediates, even short lived ones, on an atomistic level along the complete folding process of proteins.



## Chapter II.

# Effect of photocaged nucleotide on the DNA double helix

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

P. Seyfried<sup>‡</sup>, M. Heinz<sup>‡</sup>, **G. Pintér<sup>‡</sup>**, D.-P. Klötzner, Y. Becker, M. Bolte, H. R. A. Jonker, L. S. Stelzl, G. Hummer, H. Schwalbe, A. Heckel, *Chemistry – A European Journal* **2018**, *24*, 17568–17576.

<sup>‡</sup>These authors contributed equally to this work, working on different experimental part of the work and analysing corresponding data and equally contributing to the writing of the manuscript.



## 1. INTRODUCTION

### 1.1. PHOTO LABILE PROTECTING GROUPS

Photo labile groups are light sensitive groups that undergo chemical reaction upon irradiation with certain electromagnetic wavelength. This reaction can be either irreversible (photocages)<sup>[87,88]</sup> or reversible (photoswitches)<sup>[89–94]</sup>. Photocages of biomolecules (RNA, DNA, proteins)<sup>[95–98]</sup> or biologically active compounds (neurotransmitters<sup>[99–101]</sup>, calcium<sup>[102,103]</sup>, antibiotics<sup>[104]</sup>, metabolites<sup>[105,106]</sup> etc) are widely used in biological<sup>[107,108]</sup>, biochemical, biophysical research allowing the release of the active compound with high spatial and temporal control. Both *in vivo* and *in vitro* application of caged oligonucleotides have been shown. *In vivo* gene expression control was achieved by caged siRNA<sup>[109]</sup>, antisense<sup>[110]</sup> or aptamer strategies, also photolabile caging strategies have been successfully incorporated into *in vitro* studies of RNA<sup>[111,112]</sup> and DNA<sup>[113]</sup> folding .

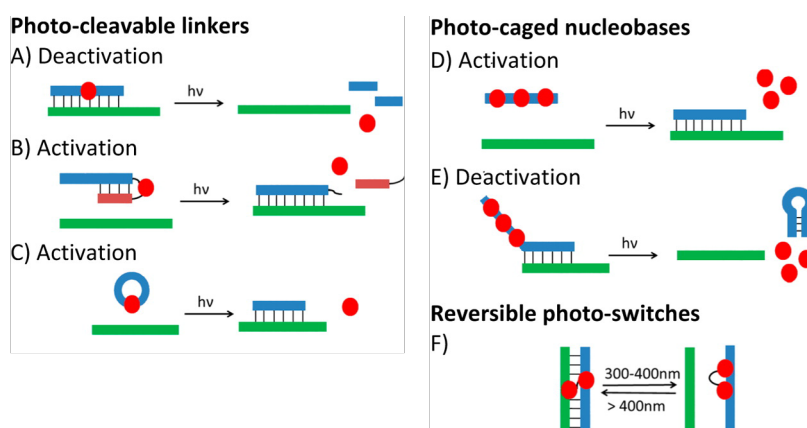


Figure 22 Different caging approaches of oligonucleotides using light sensitive protecting groups. First major approach involves photo-cleavable linkers used in a) oligomer antisense sequence b) self-hybridizing loop structure and c) cyclic oligonucleotide. Second major approach uses photo-caged nucleobases either to d) activate or e) deactivate hybridization of double stranded oligonucleotides. Third approach involves photo-switches to reversibly disrupt hybridization. Adapted with permission from Ref.<sup>[114]</sup> Copyright 2013 American Chemical Society.

In all cases a clear on and off state is desired, examples of different caging strategies are shown in Figure 22. The two most common caging sites are either the phosphate backbone, with minimal disruption of the secondary structure, or directly caging the nucleobase and thus interrupting base pairing. The latter achieves better destabilization of stable secondary structures, but still requires multiple cages to fully disrupt longer sequences. Multiple cages require longer irradiation time reducing the release efficiency of the native oligonucleotides while risking damaging the sample. Therefore, the aim of

this study was to investigate in depth the effect of single nucleobase cages and apply our new understanding to develop more efficient cages compared to existing photolabile protecting groups for oligonucleotides.

## 1.2. CHARACTERIZING NUCLEIC ACID STABILITY

Internal and external interactions that drive folding or interactions between biomolecules with its surrounding such as solvent, ions, inhibitors or substrates are all driven by changes on the energy landscape. Quantification of these interactions is crucial to understand the driving forces in detail. Conformational changes accompany DNA/RNA as they regulate gene expression (RNA thermometers<sup>[115–117]</sup>, folding of riboswitches<sup>[112,118,119]</sup>) or the enzymatic functions of ribozymes<sup>[120,121]</sup>. Similarly, to understand the effect of photocages introduced into an RNA/DNA system investigating the local changes on the energy landscape contain valuable information. The photolabile group can lead to different destabilizing interactions, such as disrupting the  $\pi$ - $\pi$  stacking of the nucleobases, force the phosphate-sugar backbone into a different conformation or interrupt the hydrogen bond formation between the base pairs. These local changes have a global effect on the stability of the double stranded RNA/DNA which can be qualitatively studied. Thermodynamic parameters of equilibrium state can be measured by isothermal calorimetry (ITC), differential scanning calorimetry (DSC) or by measuring the thermal melting curve detected by ultraviolet-visible light (UV-VIS) or circular dichroism (CD) changes. While NMR spectroscopy can be used to investigate the local effects by measuring the exchange rate of the hydrogen bonded imino signals.

### 1.2.1. Global stability characterized by melting temperature

The global stability of oligonucleotides can be measured by several techniques. Differential scanning calorimetry measures the required heat to keep sample and reference cells at the same temperature; or optical spectroscopy measures changes in absorbance of UV/VIS light while the temperature of the sample is slowly increased.<sup>[122,123]</sup> The latter results in a melting curve with a typical baseline at low ( $T \ll T_m$ ) and high temperature ( $T \gg T_m$ ) corresponding to hybridized or dissociated state of two complementary DNA sequence, respectively. The melting point ( $T_m$ ) is defined as the temperature where half of the duplex is still annealed. Usually a simple fitting of the low and the high temperature baseline is used to determine the  $T_m$ . Additional information can be extracted from melting curves by transforming the raw data into

temperature dependant fraction folded graph to extract the enthalpy and entropy upon unfolding of double stranded DNA.

### 1.2.2. Local base pair stability characterized by NMR spectroscopy

NMR spectroscopy has become invaluable tool in dissecting the thermodynamic parameters of local structures at atomic level. Oligonucleotides form canonical or non-canonical base pairs through hydrogen bonds. NMR spectroscopy offers several methods to address the base pair stability. Labile imino protons are in exchange with bulk water. This exchange characterizes the local stability. Usually measurements are done in the EX2 (bimolecular) regime, where the observed exchange rate shows linear catalyst dependence.<sup>[124–126]</sup>

An easy qualitative assessment of the stability of the base pairs comes from the linewidth of the imino proton signals. Quantitative analysis is hindered by other factors such as viscosity, ionic strength, sample homogeneity etc influencing linewidth, but to compare imino signals in the same probe it is suitable and straightforward. Direct real time measurement by H-D exchange can also be used but requires tedious sample preparation and limited by the dead time of the mixing.

Selective magnetization transfer from the bulk water can also be used. There have been several implementation of this technique.<sup>[117,127,128]</sup> Usually the pulse sequence involves selective flipping of water magnetization which is followed by a  $\tau$  mixing time, during which the imino proton can exchange with the water, followed by a simple 90° pulse as a readout of the imino protons. Measuring several  $\tau$  delays the exchange rate can be extracted. Changing the external catalyst concentration<sup>[128]</sup> or the temperature<sup>[117]</sup> (or the latest technique combines the advantage of both method, by measuring at two catalyst concentration with variable temperatures)<sup>[127]</sup> and repeating the exchange rate measurement can yield detailed thermodynamic information associated with the opening of the base pairs.

## 1.3. DETERMINATION OF ABSOLUTE OR RELATIVE CONFIGURATION OF A CHIRAL CENTRE

Stereochemistry is a crucial aspect of biomolecules. It revolves around the spatial arrangement of atoms in molecules with the same molecular formula. Biomolecules such as amino acids, nucleotides or phospholipids are all present in nature with only a single

configuration. Interaction between molecules of different configuration or with whole biological systems can lead to drastically different interactions, therefore determination of the absolute configuration of *S* or *R* according to the CIP rules is an important aspect of chemistry. If there are only one chiral centre in the molecules the two molecules are called enantiomers. Enantiomers physicochemical properties are the same, only show differences interacting with enantiomerically pure compounds or light (e.g.: circularly polarized light).

There are several techniques to determine the configuration of a chiral centre.<sup>[129]</sup> First, enantiomerically pure compounds have to be obtained. Either by separation of the enantiomers (e.g.: chiral chromatography) or by stereoselective synthesis. If crystallization is possible X-ray diffraction can be directly used to determine absolute configuration, albeit it can be useful to derivatize the compound in question and crystalize the now diastereomeric compounds.<sup>[130]</sup> If no crystal can be obtained, chiroptical spectroscopy mostly CD spectroscopy can be used<sup>[131]</sup>, albeit it requires the knowledge of all possible conformers present in the sample (e.g.: from *ab initio* simulations).

NMR spectroscopy shows different set of signals for diastereomers, or for enantiomers in optically active environment. If one of the chiral centre's configuration is known, the relative configuration of the second chiral centre can be determined, given that any possible conformation of the molecule is available as a model. Derivatizing agents are usually have one bulky group, to hinder possible rotations and to limit possible conformers, conformational space; an aromatic group to produce anisotropic effect leading to large chemical shift changes and a third group to allow reaction with the compound in question.<sup>[132-134]</sup> Additionally scalar couplings and/or NOE contacts can also be used, especially if possible rotamers are limited and models of possible conformations are available.<sup>[135]</sup> Usually a combination of the previously outlined methods are required if absolute or relative configuration of a chiral centre has to be determined with high certainty.

## 2. MATERIALS AND METHODS

### 2.1. SYNTHESIS OF PHOTOLABILE PROTECTED NUCLEOTIDES AND DNA

The synthesis of the NPPM caged DNA strand was done by **Dr. Dean-Paulos Klötzner** and **Yvonne Becker**. Detailed synthesis route of the caged nucleotide and the solid phase synthesis used to incorporate the caged nucleotide into the required DNA sequence is described in the publication supplementary information.<sup>[136]</sup> In case of NPPM caged DNA, the stereoselective separation was achieved with the racemic mixture of the precursor compound before further reaction were done. In case of NPAM caged DNA strand, the separation of the diastereomeric mixture was achieved after incorporation into the DNA strand via solid phase synthesis.

### 2.2. PREPARATION OF DOUBLE STRANDED DNA

Counter strand with no caging group was purchased from Eurofins Genomics. It was diluted with ddH<sub>2</sub>O to 1 mM concentration and was afterwards HPLC purified by Elke Stirmal. After HPLC purification the sample was freeze dried. The pellet was dissolved in minimal volume of water. The sample concentration was measured by UV absorption spectroscopy.

Annealing was achieved using 1:1 ratio of the two single stranded DNA with addition of 100 mM KCl by heating to 90 °C for 2 minutes and cooling it down on ice for 1 hour. Desalting was achieved using centrifugal concentration (Vivaspin® 2, MWCO 3kDa) with ddH<sub>2</sub>O as a solvent. Buffer exchange was achieved by using centrifugal concentrator with the required buffer depending on the measurement.

### 2.3. CD SPECTRA AND MELTING POINT

CD spectra and melting point of the annealed (*S*)-NPPM<sub>d</sub>C8, (*R*)-NPPM<sub>d</sub>C8, NPAM(A)<sub>d</sub>T9 and NPAM(B)<sub>d</sub>T9 caged double stranded DNA were measured with a JASCO spectropolarimeter. Sample concentration was 10 μM in phosphate buffer (25 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub>, 50 mM KCl, pH 6.8) spectrometer setting was 320 nm to 220 nm, scanning speed of 50 nm/min, band width of 1 nm, temperature of 293K and 10 scans were accumulated.

## 2.4. MELTING POINT

The melting point of the double stranded DNA helices were determined by temperature dependant ellipticity measurement at 272 nm. All spectra were measured with a JASCO spectropolarimeter. All samples were 20  $\mu$ M concentration in phosphate buffer (25 mM  $K_xH_yPO_4$ , 50 mM KCl, pH 6.8). The temperature range was from 5  $^{\circ}$ C to 95  $^{\circ}$ C, data points were measured every 0.5  $^{\circ}$ C with a ramp up rate of 0.5  $^{\circ}$ C/min. The data was analysed with OriginLab according to reference<sup>[122]</sup> using lower and upper baseline fit to determine the melting point.

## 2.5. NMR MEASUREMENTS

All spectra were recorded on 600/700/800/950 MHz Bruker NMR spectrometer equipped with cryogenically cooled HCN  $z$ -gradient or HCPN  $z$ -gradient probehead. All pulse sequences were taken from the Bruker library. For resonance assignment NOESY spectra of samples in 10% or 99.9% D<sub>2</sub>O containing buffer was used. The mixing time was set to 150 ms and 512 points were recorded in the indirect dimension. For TOCSY spectra the isotropic mixing time (achieved with DIPSI2) was set to 80 ms or 100 ms. Fourier transformation was done with linear prediction and zero filling up to 1024 points in indirect dimension and in both dimension sine bell square function weighting was applied with sine bell shift value of two and three for TOCSY and NOESY spectra, respectively.

The NOESY spectra for structure calculation were recorded in 99.9% D<sub>2</sub>O or 10% D<sub>2</sub>O phosphate buffer: 25 mM  $K_xH_yPO_4$ , 50 mM KCl, pH 6.8. To address spin diffusion during structure calculation three different mixing times have been measured. (*S*)-NPPM caged DNA: in D<sub>2</sub>O with 50 ms (800 MHz), 100 ms (800 MHz), 150 ms (950 MHz) and in H<sub>2</sub>O 150 ms (700 MHz); (*R*)-NPPM caged DNA: in D<sub>2</sub>O with 50 ms (950 MHz), 100 ms (950 MHz), 150 ms (950 MHz) and in H<sub>2</sub>O 150 ms (600 MHz). All the spectra were processed with Bruker Topspin. In the indirect dimension 512 or 608 point have been recorded, and were digitally filled up with zero points (no linear prediction was applied) to 1024 points and both indirect and direct dimension was multiplied by qsine function with sine bell shift of 2.5 prior Fourier transformation. The spectra were manually phase corrected by zero and first order phase correction in both dimensions. No baseline correction or digital water suppression has been used.



## 2.6. INVERSION RECOVERY EXPERIMENT

### 2.6.1. Sample preparation

After annealing and desalting with centrifugal concentrator the sample was buffer exchanged to “high catalyst” (122 mM  $K_2H_2P_4O_{10}$ , 100 mM KCl, pH 6.8) buffer. The sample was diluted with the buffer to 2000  $\mu$ l and concentrated to 200  $\mu$ l, it was repeated 10 times. The final sample volume was adjusted to 300  $\mu$ l by addition of the buffer and 10% v/v  $D_2O$  and 100  $\mu$ M DSS was added to the sample as internal reference.

### 2.6.2. Pulse sequence and experiment setup

All the inversion recovery measurements have been done on 600 MHz Bruker NMR spectrometer equipped with 5 mm HCN  $z$ -gradient cryogenic probehead (liquid  $N_2$  or He cooled). Pseudo-2D version of the inversion recovery experiment was used. The pulse sequence started with a selective  $180^\circ$  RE-BURP pulse set to the water frequency followed by  $\tau$  mixing delay. During the  $\tau$  delay gradients with power level of 5% of the maximum of 55 Gauss/cm field strength was applied to reduce radiation damping. Hard watergate element was used for water suppression. 16 to 24 different  $\tau$  delays have been measured and they were varied from 4  $\mu$ s to 6000 ms. 6 seconds relaxation delay was chosen to be sufficiently long for the water relaxation. The measurement were done at 298 K or 293 K as high temperature and at 288 K or 283 K as low temperature according to Rinnenthal et al.<sup>[117]</sup>

### 2.6.3. Analyses of inversion recovery experiment

The data analyses can be conducted in different ways, as they are described in the references.<sup>[116,117,127]</sup> The most detailed and precise analyses requires a temperature series of two samples at different catalyst concentration. As the first sample of NPPM caged DNA (not shown) was a racemic mixture and due to severe signals overlap the data analyses failed. The NPAM caged DNA could be measured and analysed. The non-racemic mixture of NPPM caged DNA was admitted only to qualitative assessment by measuring the exchange rates only at two different temperatures. The imino proton signal intensity from pseudo two dimensional data with varying mixing time ( $\tau_m$ ) was fit according to Eq. 1.

$$\frac{I_H(t_m)}{I_H(0)} - 1 = -2k_{ex} \frac{\exp(-R_{1,H}t_m) - \exp(-R_{1,w}t_m)}{R_{1,w} - R_{1,H}} \quad \text{Eq. 1}$$

where  $I_H$  is the imino proton intensity at a given mixing time,  $R_{1,H}$  and  $R_{1,w}$  are longitudinal relaxation rate of imino proton and water proton, respectively. The latter two parameters are allowed to adjust freely during fitting.

### 3. RESULTS

#### 3.1. DNA SEQUENCE DESIGN

The test DNA sequence to compare the effect of different cages was based on the previously published research from the group of Schwalbe.<sup>[137]</sup> The sequence was chosen to have mixed AT and GC base pairs, to be stable and to achieve good signal dispersion. As it has been discussed in the introduction, oligonucleotides due to the lack of diverse building blocks and/or secondary structures, are usually hindered with low spectral resolution. Therefore, the sequence shown in Figure 23 has been chosen and a cytidine was modified in case of the NPPM cage and for the NPPA cage, in which case the synthesis was not possible for cytidine, a thymidine nucleobase was modified at T9 position in the sequence.

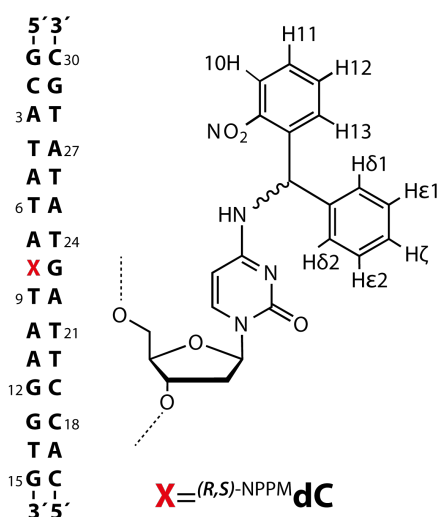


Figure 23 NPPM caged nucleotide structure and the caged DNA helix with the numbering scheme used in this thesis.

#### 3.2. EFFECT OF CAGES ON THE SECONDARY STRUCTURE AND MELTING POINT

All samples showed the typical CD curve of B-form helix secondary structure. All samples showed a minimum at 249-251 nm and a maximum at 273-275 nm. These results show that the cages cannot completely disrupt the secondary structure B-helix form of the double stranded DNA.

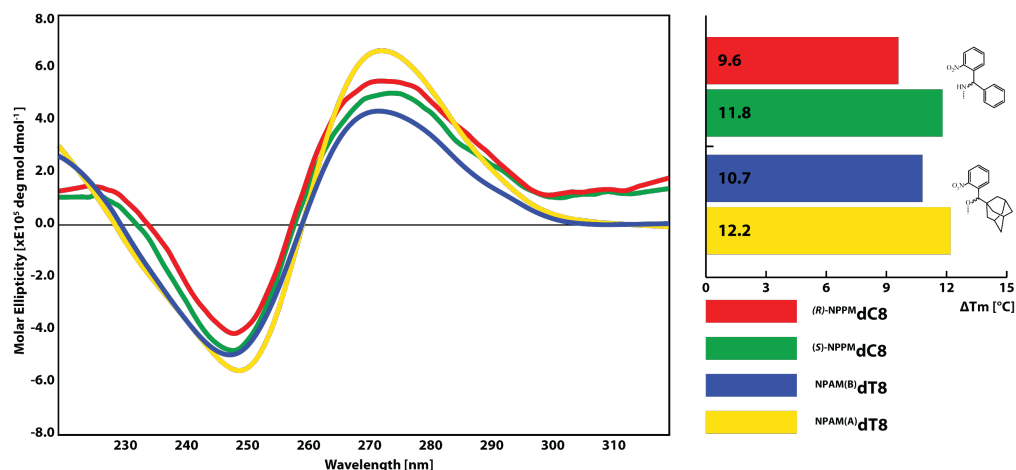


Figure 24 CD spectra and melting point changes compared to the unmodified double stranded DNA. Colour code is shown at bottom right and follows as  $(S)$ -NPPM dC8 (green),  $(R)$ -NPPM dC8 (red), NPAM(A) dT9 (yellow) and NPAM(B) dT9 (blue).

The melting point of the DNA double helices decreased between 9.6 to 12.2 °C, see exact details in Figure 24. In both cases – NPPM and NPPA – there was difference between the melting point of the  $S$  and  $R$  configuration of the cages. The slightly bigger NPPA cage showed higher reduction in melting point.

### 3.3. SIGNAL ASSIGNMENT AND DETERMINATION OF ABSOLUTE CONFIGURATION

The assignment of the non-isotope labelled nucleic acids were based on proton NMR spectroscopy. This involved  $^1\text{H}$ ,  $^1\text{H}$ -NOESY measured in buffer containing 10%  $\text{D}_2\text{O}$  as well as 99.99%  $\text{D}_2\text{O}$ , in the latter case different (50 ms, 100 ms, 150 ms) mixing times were also measured. To confirm assignments  $^1\text{H}$ ,  $^1\text{H}$ -TOCSY with mixing time of 70 ms were also used. The sequential walk assignment of the imino proton region is shown in Figure 25.  $^1\text{H}$ -NMR of the imino region from the caged  $(R)$ -NPPM dC9 and the  $(S)$ -NPPM dC9 and the native sequence with assignment is shown in Figure 26. G23 shows upfield shift of the imino signal as it is base paired to the caged nucleotide. Additionally, signals from base pairs from close proximity to the cage such as T9, T24, T6 show pronounced chemical shift changes compared to the native sequence, while other signals do not shift. Followed by the assignment of the adenine H2 signals and the amine groups, correlations to the H1' signal of the sugar backbone was established. Starting from the H1' signals, sugar-aromatic sequential walk was followed. The assignment was based on previously described strategies.<sup>[138]</sup>

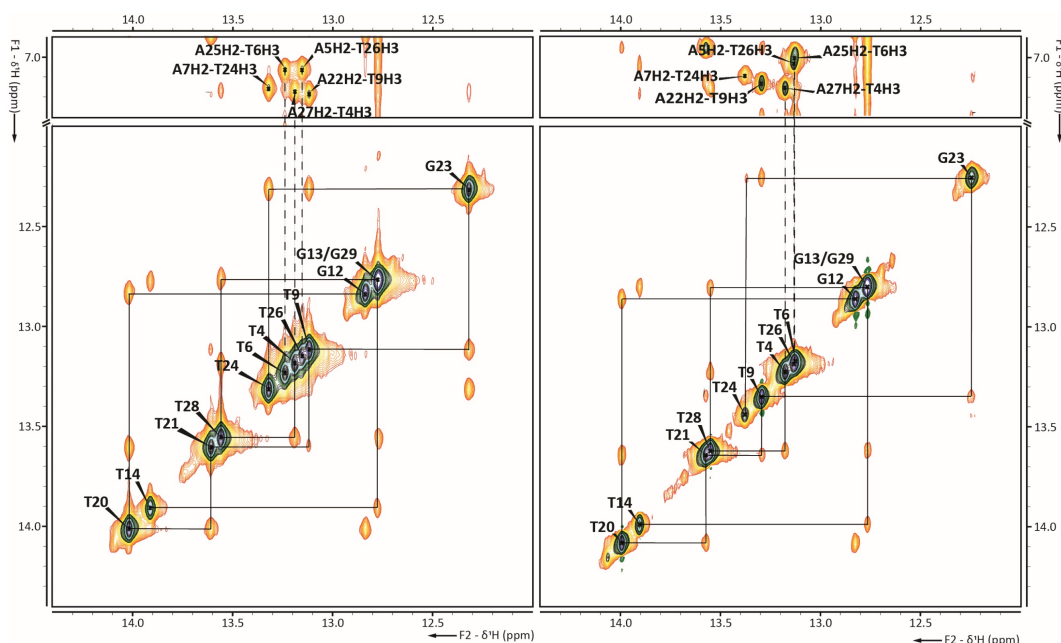


Figure 25 Imino-imino and imino-aromatic region of  $^1\text{H},^1\text{H}$ -NOESY spectra of (*R*)-NPPM-dC8 (left) and (*S*)-NPPM-dC8 (right) caged double helix DNA. The numbering scheme is shown in Figure 23. The neighbouring imino resonance walk is highlighted with black lines and dashed lines show assignment of the adenine H2 signals of the base pairs of the thymidine residues.

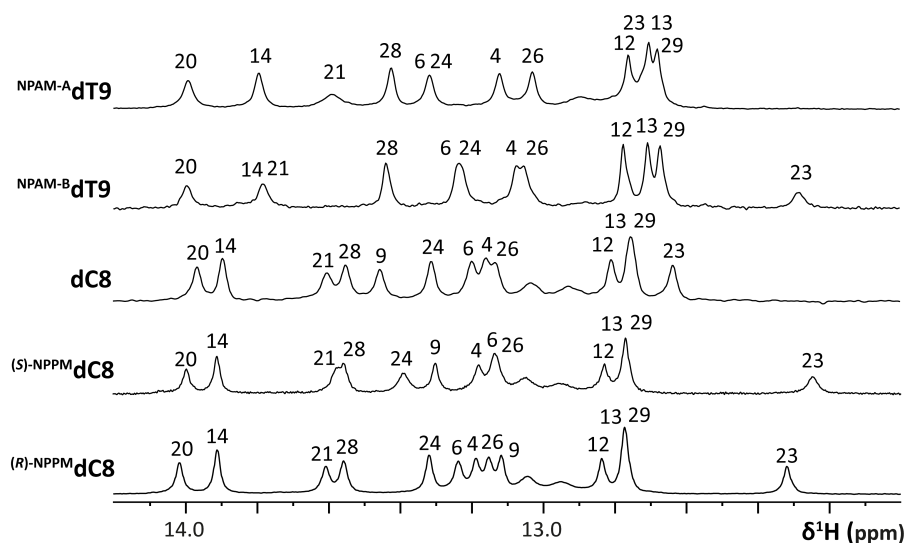


Figure 26 Imino region of  $^1\text{H}$ -NMR spectra of unmodified (middle), NPAM-A-dT9, NPAM-B-dT9, (*S*)-NPPM-dC8 and (*R*)-NPPM-dC8 caged double helix DNA. The numbering scheme is shown in Figure 23. The G23 imino signal is upfield shifted due to the cage on its base pair C8.

The NPPM cage has a chiral centre leading to two different possible configurations. While NMR spectroscopy cannot differentiate between enantiomers, when the cage is attached to the DNA, which is also optically active, the final products are diastereomers. The NMR spectra of these compounds have different set of signals, as it can be observed already on the imino signals shown in Figure 26. It has been shown before<sup>[139]</sup> if three

dimensional models are combined with specific NOE contacts, the relative configuration of the cages in the two diastereomers can be determined.

With the almost full assignment of proton chemical shifts (see list of chemical shifts in Appendix chapter 1.2) of both compounds the relative configuration of the cage was established. A model was built (shown in Figure 27 lower part) based on the  $(S)$ -NPEdC8 and  $(R)$ -NPEdC8 NMR structure published by Steinert et al. Analysing the structural model showed several possible atoms with distances, that could be theoretically observed in the NOE spectrum. Additionally, this required the assignment not just the DNA signals but of the different aromatic signals of the phenyl and nitrophenyl groups. Useful contacts had been identified between the nitrophenyl group, with four well resolved proton signals, and the DNA. The phenyl group signals have low spectral resolution, due to the free rotation of the group leading to very similar chemical shifts of all five aromatic proton signals, making it difficult to identify specific contacts. In the well resolved methyl-aromatic correlation region (shown in Figure 27 upper part) easily identified patterns were observed, which matched very well with the two model structures.

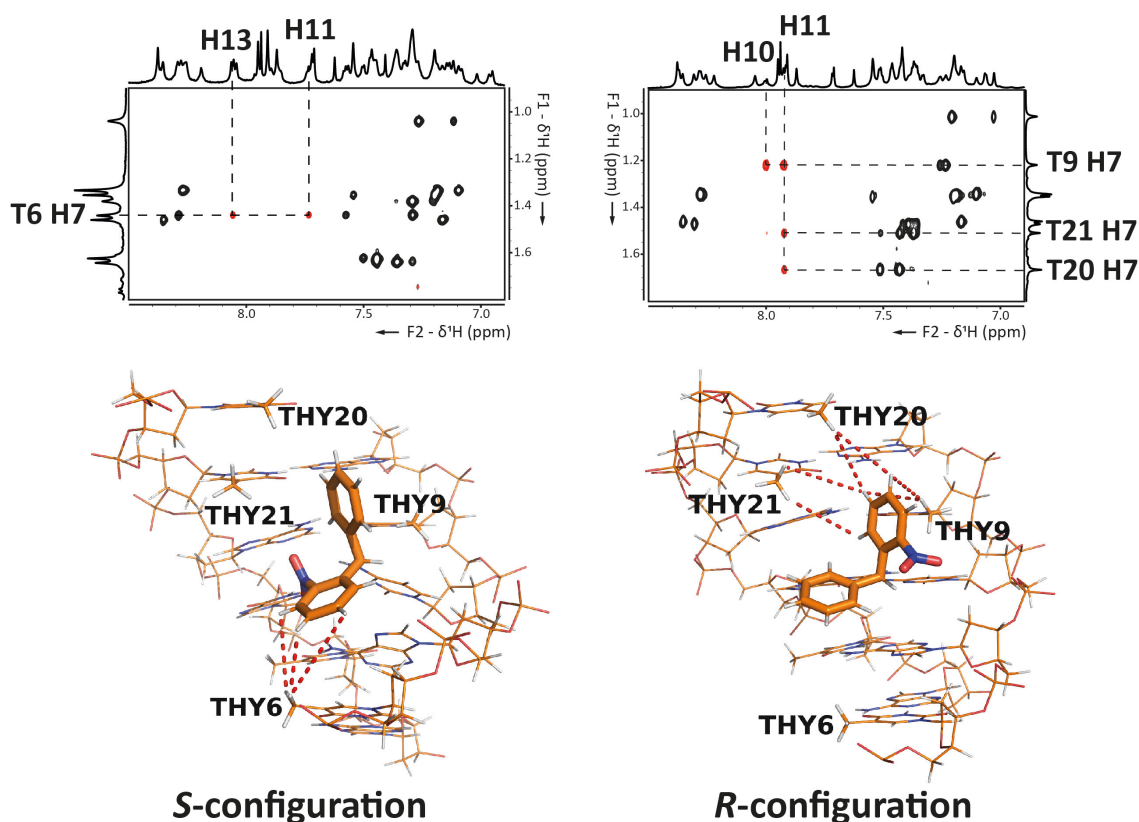


Figure 27 Upper part: aromatic (6.9-8.5 ppm) methyl (0.9-1.8 ppm) correlation region of NOE spectra. The aromatic signals highlighted are from the nitro-phenyl group of  $(S)$ -NPPMdC8 and  $(R)$ -NPPMdC8 on the left and right side respectively. On the lower part are the models built on previous results published by Steinert et al.<sup>[136]</sup> The dashed line on the three dimensional model highlights the observed NOE contacts.

The relative configuration was established by the combination of the two models with different configuration of the cage and comparison of the different NOE contact pattern observed in the spectrum. In case of the *R*-configuration of the cage, H11 showed correlation to the T9, T20 and T21 methyl groups and H10 to the T9 methyl group. While for the *S*-configuration no correlations to the previously mentioned residues were observed, but to the T6 methyl group which was not observable for the (*R*)-NPPM cage. These results were conclusive to determine the cages relative configuration.

### 3.4. LOCALIZED EFFECT OF CAGE

Inversion recovery experiments were limited as the stereoselective synthesis of the compounds posed challenges, and the separation of stereoselectively synthesized compounds were difficult to produce in high quantities required for several NMR samples. Additionally, even the slight modification of the cage compared to the NPE cage led to some overlaps in the imino regions compared to the spectrum of NPE caged DNA.

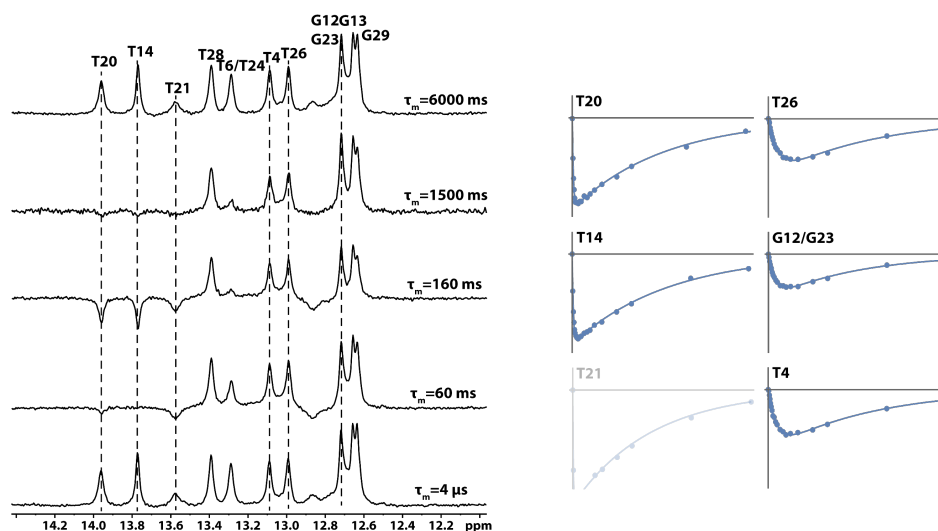


Figure 28 Example  $^1\text{H}$ -NMR spectra slices from pseudo two-dimensional inversion recovery experiment. The sample was NPAM-A caged DNA, at 293 K. Next to the spectra the  $\tau_m$  delay is shown. Dashed lines are highlighting example peaks to show fittings result of selected peaks on the right. The T21 imino proton is too labile to be measured by inversion recovery.

In Figure 27 example measurements are shown. On the left side the imino proton region of the NPAM-A caged DNA is shown with varying  $\tau_m$  mixing time. Notice that with certain mixing time some residues show negative signal while others are still positive. This shows up on the normalized intensity versus  $\tau_m$  mixing time curves as a deeper dip. If the S/N is not sufficient, due to the line broadening almost beyond detection, e.g.: T21

the fitting procedure will fail. In these cases, either lower temperature or lower external catalyst should be used.

Linear scaling of the measured exchange rate by external catalyst concentration have been demonstrated on the native DNA samples shown in Figure 29. This measurement was used as a control experiment to establish that the buffer concentration used in this study remains in the EX2 (bimolecular) regime, which is a required assumption for the data analyses.

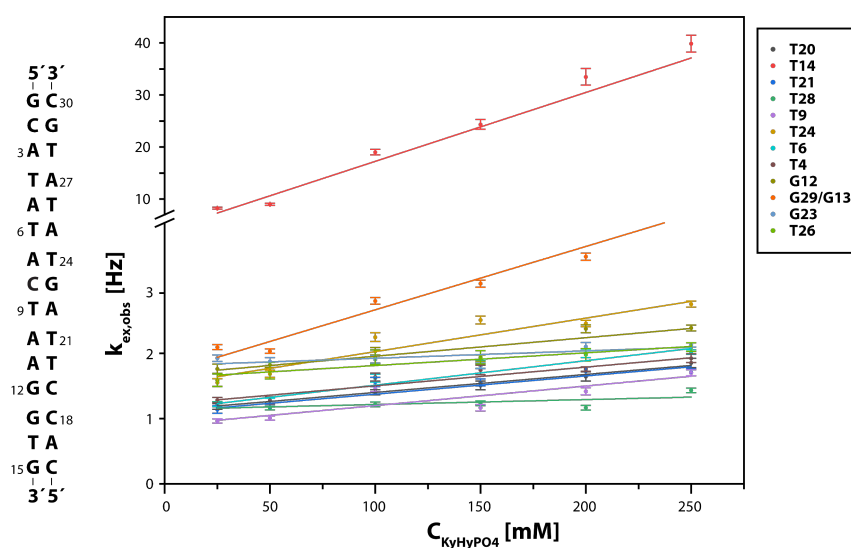


Figure 29 Catalyst titration effect on the observed exchange rate measured by inversion recovery experiment at 293K. On the left shown the DNA sequence used in the experiment. The error bars represent fitting error of the inversion recovery experiment. All nucleotides exchange rate show in the range of 25 mM to 250 mM linear dependence.

Full thermodynamic analysis was not possible either due to limited amount of sample resulting in unfeasible time demands for the measurement of complete temperature series. Additional attempt was made by using higher concentration of racemic mixture of the samples, but it was hindered by the severe signal overlap, prohibiting the extraction of signal intensities. For the NPAM cage due to the different modification site as compared to NPPM lead to different imino signal distribution with several overlaps. For the NPAM cage analyses different sequence design could elevate this hindrance, but the effect of the cage on shielding or deshielding is unpredictable, making it rather difficult to predict a suitable test sequence. Nevertheless, in Figure 30 the exchange rates measured at 298 K in “low catalyst” buffer is shown in order of the sequence for the NPAM-B caged DNA, grey filled bars. Showing pronounced destabilization effect on G23, T21 and T20, and a smaller effect on T24, while no difference is observed compared to the native sequence shown with white empty bars.



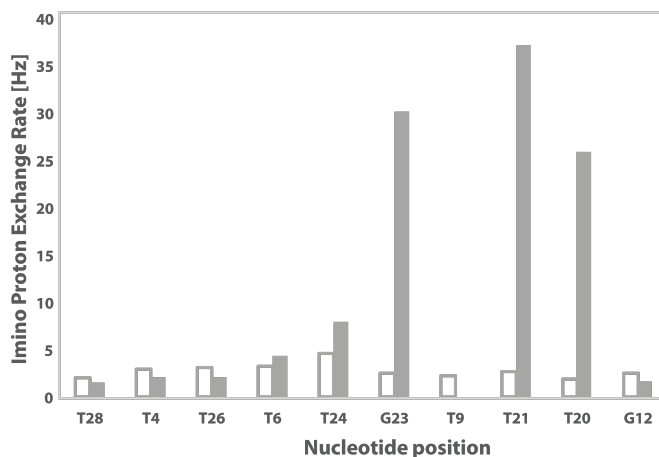


Figure 30 Exchange rates of imino protons measured at 298 K in “low catalyst” concentration buffer (5 mM  $K_xH_yPO_4$ , pH 6.8, 244 mM KCl) by inversion recovery experiment. The exchange rate was calculated according to Eq. 1. The empty bars and grey filled bars represent the value measured on native DNA duplex and NPAM-B caged DNA duplex, respectively. T9 could not be analysed for caged DNA, as there is no imino proton on T9 due to the cage

### 3.5. STRUCTURE CALCULATION

The structure calculation has been done by **Dr. Hendrik R. A. Jonker** with CNS<sup>[140]</sup> 1.1 to generate conformers and then using ARIA 1.2 with nucleic acid forcefield protocols. It also included OPLS parameters<sup>[141]</sup> for charges and non-bonded parameters. Standard annealing protocols were, including NOE distance classifications separately for exchangeable protons, and for non-exchangeable protons, for which spin-diffusion correction (using three different mixing times) was also used. The NPPM cage topology and parameters were manually generated with the help of standard nucleic acid and protein forcefields.

Additional manual restraints have been used to compensate for the limited amount of restraints, as dihedral angles are not readily determined for unlabelled samples and NOE constraints density per residue in nucleic acids is generally low compared to proteins. These additional constraints were dihedral angles, base pair hydrogen bonding and planarity, but were limited to the upper and lower 5 base pairs: 2-7 and 11-14. The middle 3 and the closing base pairs were left free to be changed during the calculation. The peaks have been manually picked and assigned to a large extent, exception of signals where no clear assignment was possible. The spin diffusion correction was achieved by using NOE cross peak intensities of different mixing times (50 ms, 100 ms, 150 ms) in 99.9%  $D_2O$  buffer. Single mixing time (150 ms) was used for the labile protons measured in 90%  $H_2O$  buffer. The chemical shift tolerance was to 0.015 ppm for proton dimensions.

Fifty starting structures were generated from linear templates. Seven iterations of structure calculations have been done. In each iteration fifty structures were generated. Based on the new structures the NOE distance restraints were recalibrated before the next iteration based on the ten lowest energy structures. The violation tolerance was stepwise reduced to the final 0.1 Å. In the final 8<sup>th</sup> iteration 200 structures were calculated and out of this the twenty lowest energy structures were refined in explicit water.<sup>[142]</sup> During structure calculation different temperature stages were used from 10000 K down to 50 K and in each step the force constant for NOE restraints were set stepwise from 0 to 50 kcal mol<sup>-1</sup> Å<sup>-2</sup>.

Table 5 NMR Structure calculation statistics for (S)-NPPM<sub>dC8</sub> and (R)-NPPM<sub>dC8</sub> caged double helix DNA.

|                                   | (S)-NPPM <sub>dC8</sub> | (R)-NPPM <sub>dC8</sub> |
|-----------------------------------|-------------------------|-------------------------|
| <b>Hydrogen bonds</b>             | 23                      | 23                      |
| <b>Base planarity</b>             | 10                      | 10                      |
| <b>Torsion angles</b>             | 232                     | 232                     |
| <b>NOE distance restraints</b>    | 764                     | 823                     |
| • <i>Intra</i>                    | 411                     | 468                     |
| • <i>Sequential</i>               | 226                     | 261                     |
| • <i>Medium and long range</i>    | 61                      | 68                      |
| <b>Ambiguous</b>                  | 66                      | 26                      |
| <b>Violations</b>                 |                         |                         |
| • <i>Distances (&gt;0.4 Å)</i>    | 0                       | 0                       |
| • <i>Dihedral angles (&gt;5°)</i> | 1                       | 1                       |
| <b>Average RMSD to mean</b>       | 1.56 Å                  | 1.27 Å                  |

### 3.6. MEDIUM THROUGH PUT SCREENING

**Patrick Seyfried** conducted the synthesis and the melting curve measurements shown in Figure 31 and discussed in this chapter.

To address the effect of different size and relative configuration of modifications listed in in Figure 31 two different types of photolabile groups were modified. The following compounds were synthesized: 1-(2-nitrophenyl)but-3-yn (NPBY) or 7-(diethylamino)-4-(but-3-yn-1-yl)coumarin (DEACBY) caged thymidine inserted into the sequence (shown in Figure 31 left side) at the centre position. After the solid phase synthesis, the different functional groups (Figure 31 **b-j**) were incorporated onto the photolabile group via click chemistry and finally annealed to the DNA counter strand. The melting point of the DNA double helix was measured by UV absorption changes.

All the synthesis and measurement details can be found in the reference supplementary information.<sup>[136]</sup>

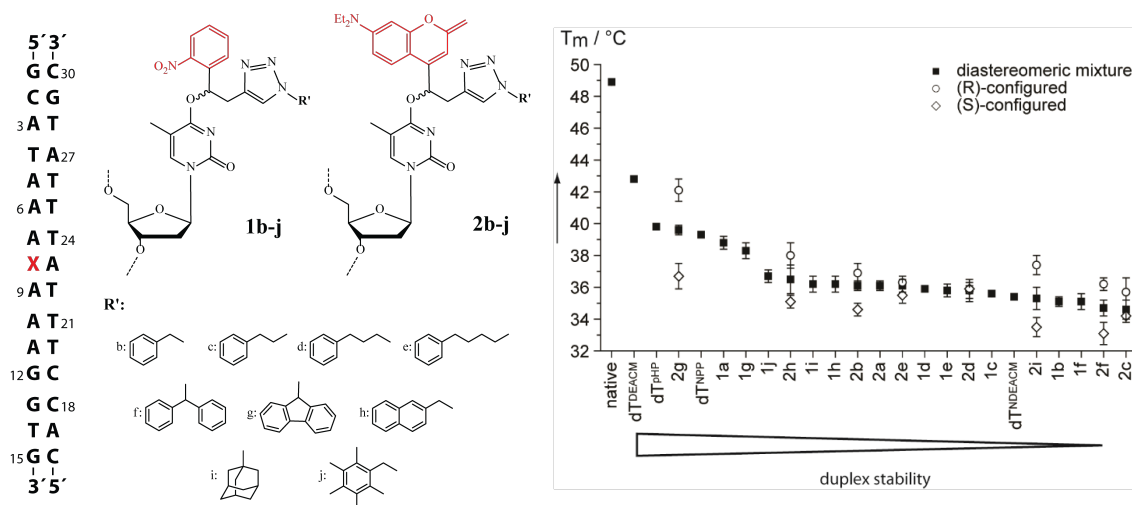


Figure 31 On the left shown the DNA sequence with X highlighting the position of the modified caged nucleotide. In the middle part shown schematics of the modified caged dT. Each modified nucleotide was incorporated by solid phase synthesis. **1b-j** are modified nitro-phenyl cages, while **2b-j** cages are modified coumarin cages. On the left side the melting points of the modified cages incorporated into sequence are shown. In case of coumarin cages stereoselective synthesis was used and the different melting point of the *S* and *R* configuration is shown with circles (*R*) and rhombi (*S*). The graph with the melting points was published by Seyfried et al. in the reference.<sup>[136]</sup>

### 3.7. MD SIMULATIONS

The MD simulations have been conducted by **Marcel Heinz**. Short summary of the simulations are described here, more in depth details can be found in reference supplementary.<sup>[136]</sup> The first step was to parametrize the new cages for further MD simulations. This has been achieved by quantum mechanics (QM) calculations, assuming the sugar and phosphate backbone does not change significantly only the nucleobase with the cage were used in this step. Afterwards the new cage was covalently attached to the sugar. The dsDNA was put in a water-box with additional NaCl up to the concentration of 160 mM. All the molecular mechanic simulations were performed with Amber 14 and 16 packages. First energy minimization was done in three steps to avoid steric clashes. Five to ten independent simulations were run, each up to 200 ns, resulting in a total simulations time of one to two  $\mu$ s. The trajectories were clustered with GROMACS algorithm based on the RMSD (cut off was set to 1.25 Å) of the caged and the neighbouring residues, plus the complementary residues. In case of (*S*)-NPPM this resulted in two major clusters with 51.1% and 28.5% populations, while for (*R*)-NPPM one major cluster could be identified with 59.9% population.

## 3.8. MD SIMULATIONS DATA COMPARED TO NMR DATA

To assess the quality of the MD simulations detailed analyses could be done, in form of comparing the available NOE data to the distances in the MD simulated structures. NOE distances were calculated according to Eq. 2 for the MD trajectories, where  $r$  is the average distances of each trajectory  $r_i$  distance.

$$r = \langle r_i^{-6} \rangle^{-\frac{1}{6}} \quad \text{Eq. 2}$$

The range of distances are taken from the structure calculation done with ARIA<sup>[143]</sup> according to Eq. 3, where  $d$  is the approximate distance value based on the NOE cross peak intensity.

$$\text{range} = d \pm 0.125d^2 \quad \text{Eq. 3}$$

The values were compared between contacts from the cage to the DNA and contacts within the caged residue, up to 6.5 Å, as NOE upper limit. The correlation plots are shown in Figure 32 (for (*S*)-NPPM) and in Figure 33 (for (*R*)-NPPM).

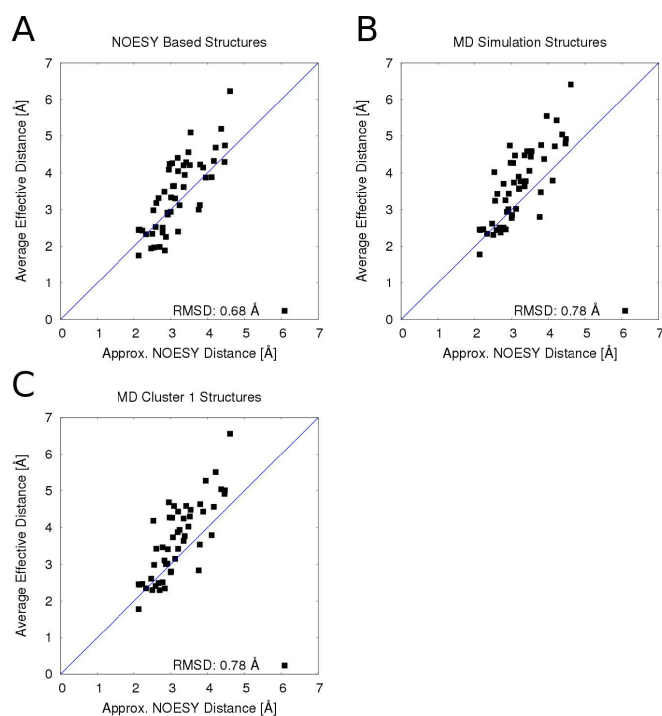


Figure 32 Correlations plots of (*S*)-NPPM caged DNA. The correlations are shown for approximate NOE distances and the distances calculated from MD trajectories. In all cases the cage-cage and cage DNA distances were used in the comparison. In A the structure bundle of NMR based structure was used as comparison, in B and C the direct NOE intensities were used. In B all structures from MD simulations were used while in C only trajectories from cluster 1 were used. The figure was taken from the reference supplementary information published by Seyfried et al.<sup>[136]</sup>

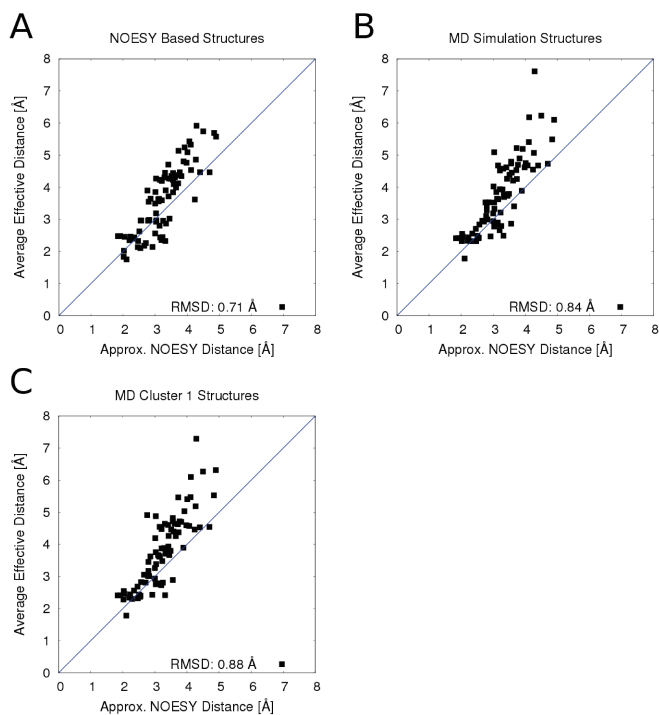


Figure 33 Correlations plots of (*R*)-NPPM caged DNA. The correlations are shown for approximate NOE distances and the distances calculated from MD trajectories. In all cases the cage-cage and cage DNA distances were used in the comparison. In A the structure bundle of NMR based structure was used as comparison, in B and C the direct NOE intensities were used. In B all structures from MD simulations were used while in C only trajectories from cluster 1 were used. The figure was taken from the reference supplementary information published by Seyfried et al.<sup>[136]</sup>

Visual comparison between cluster 1 and cluster 2 to the NMR structure ensemble leads to a plausible explanation as the cage free rotation around the C-N bond has two limits in an about 45° angle, which are exactly observed in cluster 1 and 2, while NMR structure shows an average of the two. In case of (*R*)-NPPM the major cluster has better agreement with NOE distances, and visual comparison also shows good agreement.

## 4. DISCUSSION – LIMITS OF SINGLE CAGE APPROACH

In the following chapter the effect a single cage approach to maximize destabilization of DNA double strand is discussed using the results of NMR spectroscopic investigation of (*R*)-NPPM, (*S*)-NPPM, NPAM-A and NPAM-B (synthesized by Dean Klötzner and Yvonne Becker) measured and analysed by György Pintér. Additionally, a medium throughput screening was conducted by Patrick Seyfried and the molecular dynamic simulations were done by Marcel Heinz to investigate other aspects of DNA destabilization by a photolabile protecting group.

### 4.1. THE EFFECT OF SIZE AND RELATIVE CONFIGURATION

Steinert et al. has shown in case of the most commonly used nucleotide cage NPE, that it is beneficial to not use the racemic mixture, but instead the *S* configuration.<sup>[137]</sup> For the first time they observed difference between destabilization of the (*R*)- and (*S*)-NPE caged DNA duplex. Therefore, we also studied the effect of the relative configuration and how different it can interact with the duplex structure. The second effect we were interested in is the size of the cage, whether a more sterically demanding group could further destabilize the DNA duplex structure.

The first part of the study focused on a comprehensive NMR structural characterization of the (*R*)-NPPM and (*S*)-NPPM cages, which are modified NPE cages. The NPE cage was modified in the alpha position, where the methyl group was changed to a larger phenyl moiety. Additionally, a less thorough NMR investigation was done on the NPAM-A and NPAM-B cages where the NPE cage was modified in the alpha position by replacing the methyl group with a larger adamantyl group.

The melting point of (*R*)-NPPM and (*S*)-NPPM caged DNA showed further depression of melting point in comparison to the NPE cage. Although the overall CD spectrum still showed a typical curve of a B-helix secondary structure (Figure 24) conformation, it concludes that the cage cannot completely disrupt the helix and the overall conformation stays intact. The <sup>1</sup>H-NMR spectrum (Figure 26) shows the imino proton region, which indicates stable fold of the secondary structure, as imino proton signals are only observable when stable hydrogen bond of the base pairs is formed. Albeit the directly involved base pair G23 imino signal and the neighbouring imino signals exhibit larger linewidth in comparison to the native sequence. This is due to the lower stability of the base pair and the increased exchange rate with the bulk water. After the relative

configuration determination of the NPPM cages (see detailed discussion in chapter 3.3), the different effect of the *S* and *R* configuration was investigated.

In comparison to the (*S*)-NPE or (*R*)-NPE both (*S*)-NPPM and (*R*)-NPPM shows better destabilization. The difference between reduced temperature for (*S*)-NPE and (*R*)-NPE is 4.4 °C while for (*S*)-NPPM and (*R*)-NPPM this difference drops to 2.2 °C. The most likely explanation is the smaller size difference and chemical similarity between the 2-nitro-phenyl and phenyl moiety in the NPPM cage, while for NPE the size difference between the nitro-phenyl group and methyl group is larger. This means for the application of cage compounds, that if the substituents are similar sized, racemic mixture is likely to work almost as good as a stereochemically pure compound, considering destabilization. If the substituents have larger steric differences, one of the compounds is likely to have higher impact on the DNA double strand destabilization, as is the case also for example for compound **2i** shown in Figure 31.

Interestingly in both the NMR investigation and in the melting point screening study the *S* configuration of the cages show more efficient destabilization. Regarding destabilization difference between the *R* and *S* configuration, it can vary drastically from very minimal below 0.5 °C to as large as above 4 °C, as is the case for **2i** and **2d**. Furthermore, the uncaging efficiency measurements, defined as a product of molar extinction coefficient ( $\epsilon$ ) and uncaging quantum yield ( $\phi$ ), yields for (*S*)-**2f** 258 M<sup>-1</sup>·cm<sup>-1</sup>, and for (*R*)-**2f** 170 M<sup>-1</sup>·cm<sup>-1</sup>. The *S* configuration is not just better in destabilization, but also the uncaging efficiency is superior. This difference, and preference regarding the *S* configuration could be explained by looking at the NMR model structure built previously for (*R*)- and (*S*)-NPE<sup>[137]</sup> and in this study for (*R*)- and (*S*)-NPPM and also taking into account the results of MD simulations.

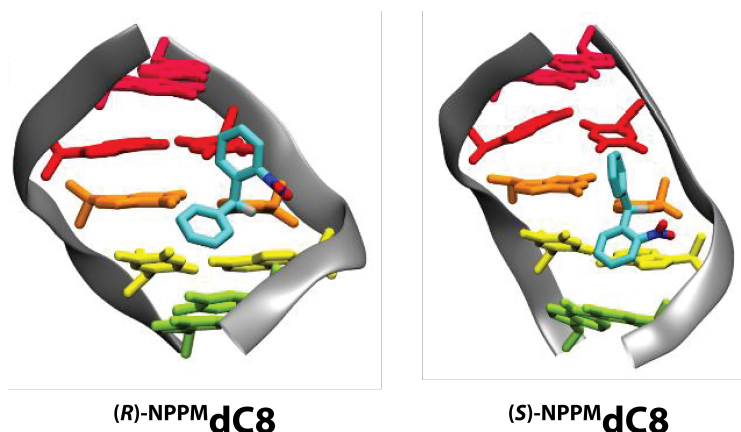


Figure 34 NMR model structure of *(R)*-NPPM dC8 and *(S)*-NPPM dC8 caged DNA double helix with sequence shown in Figure 23. The middle five pair of nucleotides are shown. The nitrophenyl group points to the T9 and to A7 in case of *(R)*-NPPM dC8 and *(S)*-NPPM dC8, respectively. The figure was adapted from ref.<sup>[136]</sup>

In all structures, either NMR determined or by MD simulation calculated, the cages fitted in all cases in the large groove of the helix and did not flip out into the solvent. Therefore, the theoretically negative thermodynamic effect of a hydrophobic cage penetrating into the solvent, further destabilizing the helix was not observed, even though more and more sterically demanding groups were tested. Although, the secondary structure was never completely disturbed, the caging group induced tension in the backbone angles as they deviated from optimal values in accord to accommodate the modification, leading to the reduced stability of the structure. These deviations from standard backbone angles led to slightly disturbed nucleobase positions, disturbing the stabilizing  $\pi$ - $\pi$  stacking effect as well as hydrogen bond formation. The NPPM-cage NMR structure bundle as well as the MD simulations trajectories showed restricted rotation of the modification around the C-N bond. In case of the *(S)*-NPPM the MD simulations shows two major clusters counting for almost 90% of the population. Interestingly the NMR based model showed an average position the cage positions observed in the MD simulation clusters. A likely explanation is that both clusters are real but as NOE data stems from an ensemble average, the two conformational equilibrium cannot be distinguished leading to an average position in the structure calculation.

In contrast what the size of the DEACM cage compared to the NPE cage would indicate, it was less destabilizing, the opposite of the observed trend of decreasing melting point as the size of the cage was increasing. MD simulations could also give a hint why the DEACM cage showed less destabilization. In the MD simulated structure shown in Figure 35, the aromatic coumarin points to the nucleobase and stacks on top of the neighbouring nucleobase. This leads to a stabilizing  $\pi$ - $\pi$  interaction compensating for the



destabilizing effect stemming from disrupted hydrogen bond and the strain on the phosphate backbone.

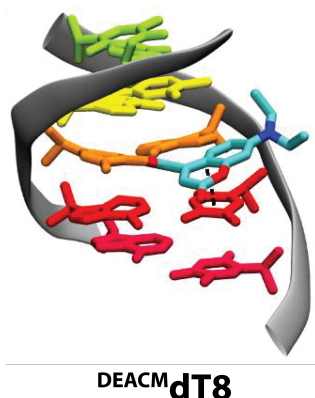


Figure 35 MD structure of <sup>DEACM</sup>dT8 caged DNA double helix with sequence shown in Figure 31. The middle three pair of nucleotides are shown in. The overall secondary structure is intact, but the coumarin penetrates into the structure and likely forms a stabilizing  $\pi$ - $\pi$  interaction with the neighbouring nucleobase. The figure was adapted from ref.<sup>[136]</sup>

All experiments pointed into the direction that even though increasing size of the cage can further depress the melting point, the overall structure of the double helix could never be completely disrupted. NMR measurements of the opening rates of the nucleobases further supported the localized effect of the cages. The selective water inversion recovery experiments showed increased exchange rate of the imino protons with water up to two base pairs away from the caging site. Also, the increase in exchange rate compared to the native strand decays rapidly as it gets further away from the modification. This has also been the case for NPE cages investigated by Steinert et al., and in my studies for NPPM and NPAM cages. Even though in our case detailed thermodynamic analyses was not possible, the even bulkier NPAM cage also showed a restricted effect up to two neighbouring base pairs. This behaviour is similar to the effect of bulges or mismatches which are common secondary structural elements in nucleic acids. These secondary structure elements followed by a stable base pair can still form an overall double helix secondary structure and does not influence the overall conformation or the stability of the double helix apart from a few neighbouring nucleotides.

Generally, the size of the modifications improved the destabilization, compared to the unmodified DEACM- or NPE-cage, although there was a limit on how far steric effects affected the base pair stability. The medium throughput screening of different size and types of substituents (shown in Figure 31) utilizing click chemistry was used to test the maximum destabilization possibility by more and more sterically demanding

modifications. The melting points showed a decreasing trend but with a plateau at 34 °C. to the effect attached to the photolabile group. The inversion recovery experiments already suggested limited and localized destabilization of the double helix, which could explain the melting point plateau that was seen in the medium throughput screening, as the melting point reports on the overall stability, if the cage can only affect a maximum of five base pairs, the final melting point is mostly determined by the rest of the unaffected base pairs in the helix.

#### 4.2. LIMITS OF SINGLE CAGE APPROACH, ALTERNATIVE WAYS TO REGULATE

In this detailed study we could show that there are still possible improvements regarding the destabilizing effect of single cages and a better candidate has been found in (*S*)-**2f**. Also, the importance of the relative configuration of the cage both for the better uncaging efficiency and for the destabilization has been shown. The limit of increasing size of the cage approach has been observed. Therefore, inhibition of longer double helices to achieve gene expression control without multiple cages still impose a challenge but the number of required cages can be reduced based on our results. This is already a great achievement if the aim is to reduce the required laser irradiation time and further improving spatial and temporal control. In further studies other physical-chemical parameters of cages could be improved such as uncaging efficiency to further reduce irradiation time. Alternatively, well positioned single cages have been shown to be sufficient for gene expression control if the cages inhibit crucial interaction site, as discussed in the introduction.

## 5. SUMMARY AND OUTLOOK

Photolabile protecting groups are invaluable tools in several fields of biophysical and biological research. They allow high spatial and temporal control by releasing native compounds upon laser irradiation. By inserting the caging group into an RNA or DNA its function or structure can be altered. One way is to disturb the double helix, therefore making it inactive until irradiation. Usually this approach requires several caging groups to be inserted for complete on-off effect, which leads to high irradiation times and to the possible damage of the sample. Therefore, more efficient destabilization of the double helix by photocages is desired for optimal applications.

The effect of the increasing size and the difference between *S* and *R* configuration of multiple cages have been investigated in our study. The use of NMR spectroscopy in combination with MD simulations and click chemistry based screening allowed detailed investigation of the DNA double helix destabilization by the introduction of different single nucleobase cages. Generally, the higher spatial demand of the cage led to higher depression of the melting point, but a plateau was observed in this trend. Measuring exchange rates of imino protons with bulk water provided evidence that the effect of a single cage is limited up to two neighbouring base pairs, suggesting a maximum obtainable destabilization. Furthermore, the NMR structure and MD simulated structures all showed that the cage fits into the major groove but is restricted in rotations. Also revealing that the backbone angles and nucleobase positions deviate from optimal value to accommodate the modification. Additionally, the effect *S* and *R* configuration of a cages showed a trend of *S* configuration being more efficient in destabilization of the DNA. Although, the difference between *R* and *S* configuration can vary from almost negligible to a very pronounced effect. Finally, this detailed study led to the development of the most efficient single nucleobase cage with improved photochemical properties up to now, but also revealed the limits of single-cage destabilization approach.



# Chapter III.

## NMR spectroscopic characterization of the GCAA RNA tetraloop

The results presented in this chapter are also partially compiled in different format, as a manuscript in preparation:

R. Schnieders<sup>‡</sup>, A. Oxenfarth<sup>‡</sup>, **G. Pintér<sup>‡</sup>**, C. Richter, B. Fürtig and H. Schwalbe, *manuscript in preparation.*

<sup>‡</sup>These authors contributed equally to this work by investigating different RNA tetraloop sequences.

C. Richter, B. Fürtig and H. Schwalbe designed the project and assisted with pulse sequence optimization.

Some part of this chapter is also published in the PhD thesis of Robbin Schnieders with the following title: Development of NMR spectroscopic methods for the characterization of RNA



## 1. INTRODUCTION

RNA fulfils variety of functions in the cell, from conveying genetic information (mRNA) to transport function (tRNA), it can act as a catalyst (ribozymes) and form complexes (ribosomal RNA)<sup>[144]</sup> to regulate gene expression (lncRNA,<sup>[145]</sup> microRNA<sup>[146]</sup>, riboswitches<sup>[147]</sup>). The goal of structural biology is to understand the function-structure relation, which requires detailed description of the primary-, secondary- and tertiary structure. It has been long recognised that biomolecules, to fulfil their function, cannot be looked at as a rigid structure, but an ensemble of multiple conformers. NMR spectroscopy and MD simulations are extremely valuable tools to investigate the dynamic changes, interconversion of conformers and interaction between the molecules at an atomistic level. While NMR spectroscopy can provide experimental evidence about dynamics at certain timescales, MD simulation can serve as an orthogonal method to understand underlying conformational changes at possible different timescales.

### 1.1. MD SIMULATIONS OF RNA

Quantum mechanical simulations on biological systems even on the smallest scale are not feasible up to now. An alternative solution to address systems containing hundreds of atoms emerged as molecular mechanics and molecular dynamics simulations. In these simulations a physical model so called force field is used to describe the different inter- and intramolecular interactions of atoms or building blocks, that in the end define how a biomolecule behaves in the simulation. Force fields simultaneously describe bonds, angles, van der Waals, electrostatic, hydrogen bonding and solvent interactions by potentials. These terms require detailed characterization and optimization before any larger systems should be simulated, considering the computational costs. For proteins there are existing force fields, the two major force fields developed are CHARMM<sup>[148]</sup> and AMBER<sup>[149]</sup> and both are robust and can be generally used for simulations. An example of their goodness is represented by even an earlier version of the CHARMM force field being able to *de novo* reversibly fold 12 structurally diverse peptides in 1 ms simulation time.<sup>[150]</sup> In comparison these force fields for nucleic acids especially for RNA are still lagging behind. Comprehensive reviews about the state of MD simulations of RNA have been published in recent years.<sup>[151,152]</sup> It seems the community agrees that there are still no generally accepted force field, and further adjustment to the physical description of the interactions, the force fields have to be optimized. One approach is the

use of experimental parameters either to optimize or to test the force fields. The advantage of NMR spectroscopy lies in its unique feature to be able to characterize dynamics of the system. Reference systems need to be reasonably small such as tetranucleotides.<sup>[153–155]</sup> These provide useful information that can identify problems in force fields especially regarding single stranded, unstructured RNA sequences. They are cost efficient and NMR data can be directly compared to simulation results, but they lack H-bonding and base pairing which are key elements in RNA interactions. The smallest systems to test simultaneously canonical A-helix formation and non-canonical elements are tetraloops, containing a short stem sequence and a four nucleotide long loop as a cap.<sup>[156]</sup> Using these test systems, several comprehensive comparison between available sampling techniques and force fields have been published.<sup>[157–159]</sup> These results revealed the issue of too short simulation times usually used in other studies. While using NMR or X-ray structures as a starting point, in a short time they seem stable, but over longer simulation time the structures deviate and do not return to the accepted structure, meaning other non-experimental structures show lower energy minima in the simulation. Additionally, these reviews pointed out the importance of directly comparing simulations to NMR parameters rather than to a single model structure based on NMR parameters. Up to now GNRA<sup>[160,161]</sup> and UUCG<sup>[156]</sup> tetraloops were often used as reference system, although the former lacks direct information about scalar coupling, cross correlated relaxation rates, which are completely independent of the force field. The UUCG tetraloop is extensively characterized and numerous NMR parameters are available. This short list highlights the lack of suitable experimental data to optimize and test new force fields.

## 1.2. SELECTING RNA TETRALOOPS AS REFERENCE SYSTEM

RNA can adopt complex secondary and tertiary structural elements. Among these the most common ones are hairpin loops connecting the 3'- and 5'-ends of the double helix, with the most common loop size of four nucleotides. The most common sequences, which are highly conserved in nature are GNRA, UUCG and to a lesser extent CUUG and GAAG.<sup>[162,163]</sup> Also recent clustering of known tetraloop structures revealed that there are 8 major folding groups, shown in Figure 36.<sup>[164]</sup> Based on this structural map and already available literature data we were looking at possible model systems. Considering this map UUCG covers the lower left corner, while CUUG falls into the middle and GCAA and GAAG belongs to the largest group of GNRA folds. Bottaro et



al. also revealed the importance of categorizing tetraloops based on their structure instead of their sequence as for example GAAG falls also into the GNRA folding group. Considering the selection of the three new model tetraloop system (GCAA, GAAG, CUUG), which have been characterized already to some extent, would cover most of the structural map of tetraloops.

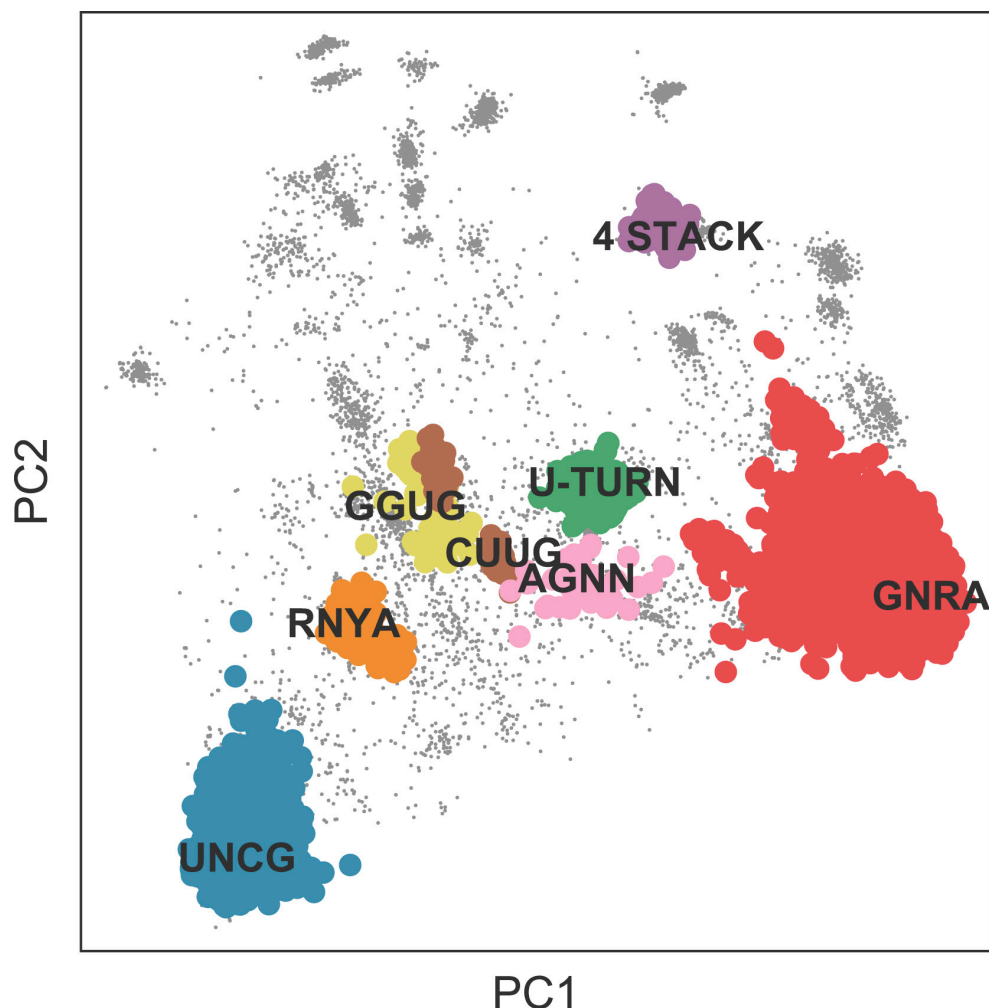


Figure 36 Principle component analysis of RNA tetraloop structures. The figure is taken from reference.<sup>[164]</sup> Our selected three model RNA tetraloop with sequence CUUG, GAAG and GCAA fall into the brown CUUG motif and red GNRA group, while the previously well-characterized UUCG fits to the UNCG cluster.

The CUUG, GAAG and GCAA tetraloops are not just interesting systems as model systems for MD simulations but are also functional elements in biologically relevant RNA structures. These tetraloops have been previously characterized and their structure was determined by NMR spectroscopy, based on NOE restraints, and limited amount of dihedral angle constraints derived from scalar couplings. Thanks to the advancement in NMR spectroscopic characterization of nucleic acids several new constraints and parameters can be determined.<sup>[165–169]</sup> In comparison, the UUCG tetraloop structure is

based on NOE contacts, dihedral restraints for all six backbone dihedral angles, the *syn* / *anti* conformation of the glycosidic bond is determined, RDC information was also used and the hydroxyl groups were characterized (Figure 37a). Additionally, in recent years a second structure in 1.106 Å agreement with the 2KOC structure was calculated based on exact NOEs (eNOEs) (Figure 37b).

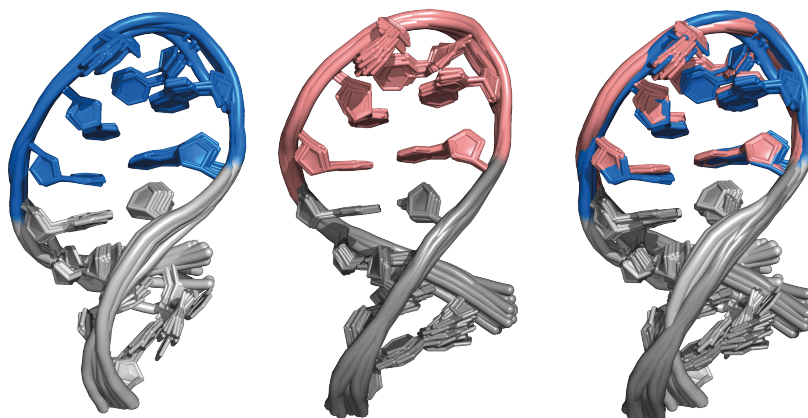


Figure 37 UUCG tetraloop structures determined by NMR spectroscopy. The blue structure is from 6BY4<sup>[170]</sup> and the red is 2KOC<sup>[156]</sup> from the PDB database. The two structural ensembles have 1.106 Å agreement. The red structure was calculated using NOE, RDC and dihedral angle restraints determined by scalar couplings and cross-correlated relaxation rates. The blue structure was determined using eNOE constraints.

Interestingly, the general consensus in literature about the rigidity of the UUCG tetraloop structure has also been questioned.<sup>[171]</sup> Bottaro et al. incorporated extensive eNOE measurements into MD simulations. The precise distance restraints from eNOE data could navigate MD simulations and an alternative loop conformation was revealed which could also fit to the NOE data. These recent results highlight the need for independent NMR parameters for force field optimization. It also demonstrates the power of combining NMR experimental data with MD simulations to address underlying dynamic behaviour of RNA molecules.

### 1.3. DYNAMIC SYSTEMS AS REFERENCE

The GCAA RNA tetraloop belongs to one of the most common hairpin family of GNRA. It is also a common tetraloop to test MD simulation developments,<sup>[172,173]</sup> although it is far less characterized than the UUCG tetraloop and only limited amount of independent NMR parameters are available. It is an intriguing system as the NMR based structure using NOE constraints and limited torsion angle constraints based on coupling constant measurements suggest a possible alternative conformation of the loop.

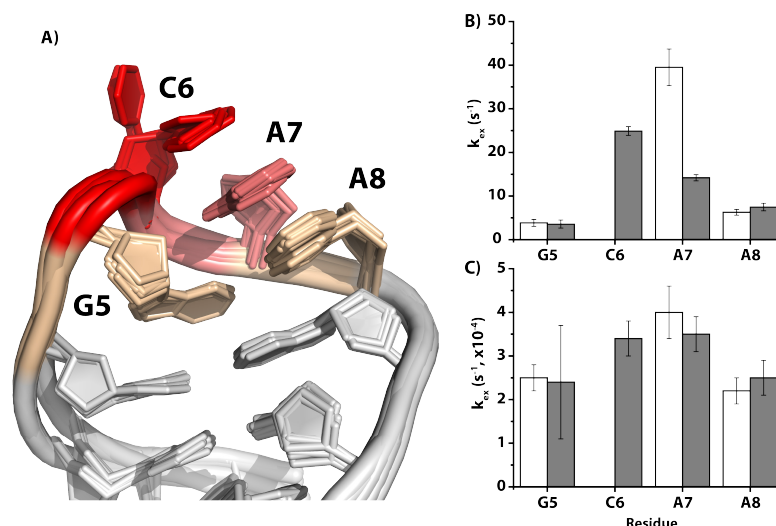


Figure 38 A) Structure of the GCAA tetraloop, taken from PDB database:1ZIH.<sup>[160]</sup> B) and C) are conformational exchange rates derived from B) exchange contribution to transversal relaxation rates and C) from relaxation dispersion analysis. The data was taken from reference.<sup>[174]</sup> Both the structure and the relaxation data suggest extensive conformational exchange for the loop residues.

Relaxation studies further supported conformational exchange of the loop on the  $\mu$ s to ms timescale. Usually NMR relaxation studies of RNA is limited due to the strong homonuclear coupling between neighbouring carbons and the extensive cross-correlated relaxation occurring between the neighbouring C-H vectors in the ribose. Nevertheless, first studies on uniformly labelled GCAA tetraloop revealed based on <sup>13</sup>C  $R_1$ ,  $R_{2\rho}$  and hetNOE relaxation extensive chemical exchange, albeit the authors give warning that some inconsistency exists in the data, likely due to neighbouring atom interference. The second relaxation study utilized a newly developed biochemical route to produce, with high selectivity, C2' and C4' isotope labelled samples. This allowed more extensive relaxation studies on the sugars by removing the complication introduced by neighbouring <sup>13</sup>C atoms.<sup>[174]</sup> In this study model free analysis of  $R_1$ ,  $R_{1\rho}$  and hetNOE revealed also extensive exchange in loop residues (Figure 38b) and additional faster exchange has been derived from relaxation dispersion measurements (Figure 38c). Other spectroscopic methods also showed possible dynamics for the GNRA tetraloop motif with a possible explanation of alternative stacking of the loop residues on top of the closing base pairs.<sup>[175,176]</sup> Utilizing chemically synthesized nucleotide derivatives as fluorescence reporters about base stacking suggests the parallel existence of 3' and 5' stacking direction of loop residues.

The GCAA tetraloop is an often-used reference system for MD simulations. Additionally, it shows extensive dynamics in the loop for possible fine-tuning of force

fields. Albeit direct, unbiased NMR spectroscopic parameters are lacking. Therefore, additional unbiased measurable parameters for MD simulations, that characterize possible alternative conformations and exchanges in the loop is desired. These parameters would include further scalar couplings, cross-correlated relaxation rates and possible residual dipolar couplings.

Therefore, we aimed to provide new NMR experimental parameters and to further analyse the structure of the GCAA, CUUG and GAAG model tetraloops in more details. As a collaborative approach my focus is on the GCAA tetraloop, while the CUUG tetraloop is characterized by Andreas Oxenfarth and the GAAG tetraloop is investigated by Robbin Schnieders. In the following chapter I discuss only the results about the GCAA tetraloop while comparison between the structures is foreseen in a common publication.

## 2. MATERIALS AND METHODS

### 2.1. TEMPLATE DNA PRODUCTION

#### 2.1.1. Sequence design

The plasmid with designed DNA sequence shown in Figure 39 was inserted into high copy yield pUC59 vector and was synthesized by GenScript (Germany). T7 promoter region is required for further application for *in vitro* transcription. HDV ribozyme sequence is used to produce homogeneous 3'-end<sup>[177]</sup>, while T7 polymerase has high tendency to produce slightly longer RNA by attaching additional nucleotides at the 3'-end. SmaI restriction site is used to linearize the amplified circularized plasmid DNA.

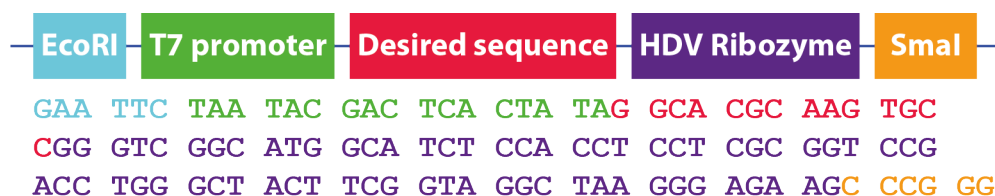


Figure 39 DNA sequence design, inserted into pUC57 vector. Green is T7 promoter region, red is the desired RNA construct sequence, purple is a self-cleaving HDV ribozyme and blue and orange are EcoRI and SmaI restriction sites, respectively.

#### 2.1.2. Transformation

The plasmid was transformed into DH5 $\alpha$  *E. coli* cells (New England Biolabs, USA). 50  $\mu$ l of cells were thawed on ice for 10 minutes. The cells were supplemented with 3  $\mu$ l plasmid (70 ng/ $\mu$ l) and incubated on ice for 30 minutes. The cells were heat-shocked for 45 second to facilitate the uptake of the plasmid. The mixture was incubated on ice for 5 minutes until it was supplemented with 950  $\mu$ l LB media. The mixture was incubated at 37 °C and shaken at 160 rpm for one hour. Dilution series of 1, 1:10, 1:100 and 1:1000 was plated onto selection plates containing ampicillin. The plates were incubated at 37 °C overnight.

#### 2.1.3. Plasmid amplification

The plasmid DNA was amplified by using DH5 $\alpha$  *E. coli* cells transformed with desired DNA sequence. 4 ml of LB medium was supplemented with 100 mg/ml ampicillin and was inoculated with a clone of the freshly plated cells. The cells were grown at 37 °C and 140 rpm for 6 h at 37 °C and 160 rpm. The main culture of 2000 ml of LB media was

supplemented with 100 µg/ml ampicillin. It was inoculated with 4 ml of the preculture. The cells were incubated at 37 °C and 140 rpm overnight.

#### 2.1.4. Plasmid purification

The cells were pelleted by centrifugation at 4000 rpm at 4 °C for 20 min. The plasmid DNA purification consists of a chemical cell lysis step followed by a manual anion exchange chromatography, according to the following protocol.

The cell pellet was resuspended in buffer S1 (50 mM TRIS, 10 mM EDTA, 100 µg/ml RNase A, pH 8.0). The cells were chemically lysed with 90 ml buffer S2 (200 mM NaOH, 1% SDS (w/v)) until the solution became clear, up to a maximum of 5 minutes. Cell debris and proteins were precipitated by the addition of 90 ml S3 buffer (2.8 M KOAc, pH 5.1 with AcOH) incubated for 10 minutes.

The soluble phase was separated by centrifugation at 12000 g at 4 °C for 50 min. Megaprep Nucleobond AX2000 column (Macherey-Nagel, Germany) was equilibrated with 20 ml buffer N2 (100 mM TRIS, 15% EtOH (v/v), 700 mM KCl, 1.5% TritonX100 (v/v), pH 6.3 adjusted with H<sub>3</sub>PO<sub>4</sub>). The supernatant was first filtrated and then added to the previously equilibrated column – one column was used per one litre main culture. The column was washed twice with 35 ml of buffer N3 (100 mM TRIS, 15% EtOH (v/v), 1.15 M KCl, pH 6.3 adjusted with H<sub>3</sub>PO<sub>4</sub>). The bound plasmid DNA was eluted from the column with buffer N5 (100 mM TRIS, 15% EtOH (v/v), 1 M KCl, pH 8.5 adjusted with H<sub>3</sub>PO<sub>4</sub>) by increasing the pH.

The plasmid DNA was concentrated by addition of 18 ml isopropyl alcohol. The precipitated plasmid was centrifuged at 10000 g at 4 °C for 40 min. The pellet was washed with 1 ml ice-cold 70% EtOH and was air dried. The DNA plasmid was tested by agarose gel electrophoresis for contaminations, and the sequence integrity was analysed by sequencing (Eurofins, Germany). The yield of the plasmid DNA was determined by UV/VIS absorption measurement on a NanoDrop One spectrophotometer.

#### 2.1.5. Plasmid linearization

Linearization was achieved by using SmaI restriction enzyme to cut the circular plasmid after the HDV ribozyme, position shown in Figure 39. 120 units of SmaI (New England Biolabs, USA) per mg DNA was used with CutSmart Buffer provided by manufacturer (New England Biolabs, USA, 50 mM KOAc, 20 mM TRIS-OAc, 10 mM MgOAc,

100 µg/ml BSA, pH 7.9) and DNA was incubated for 12 hours at room temperature. The digestion was controlled by analytical agarose gel electrophoresis, in case of incomplete digestion additional enzyme and further incubation time was used.

To separate the enzyme from the now linearized plasmid phenol-chloroform-isoamyl alcohol (PCI) extraction was used. The reaction mixture was washed twice with one sample volume of PCI mixture (24:24:1 v/v/v saturated with Tris-HCl, pH 8.0, 1 mM EDTA; Sigma-Aldrich). To remove any traces of phenol (inhibiting further *in vitro* transcription by T7 polymerase) one sample volume of chloroform was used.

The DNA was precipitated with 0.3 M NaOAc, pH 5.5 and 2.5 sample volume of absolute EtOH. The mixture was incubated at -20 °C overnight. The mixture was centrifuged at 6000 g for 20 min. The pellet was air dried and dissolved in ddH<sub>2</sub>O to reach concentration about 2000 ng/µl. The concentration was measured by UV/VIS spectroscopy with NanoDrop One spectrophotometer, also allowing the additional control of phenol contamination.

## 2.2. RNA PRODUCTION UTILIZING *IN VITRO* TRANSCRIPTION WITH T7 RNA POLYMERASE

To produce milligram quantities of isotope labelled and non-isotope labelled RNA with desired sequence *in vitro* transcription was utilized. T7 RNA polymerase<sup>[178]</sup> in combination with a DNA template (see Figure 39) and rNTPs can yield the desired sequence.

The optimal conditions of the *in vitro* transcription were optimized in 25 µl or 50 µl scale and analysed by polyacrylamide gel electrophoresis (PAGE). The optimized parameters included Mg(OAc)<sub>2</sub> concentration in the range of 5-70 mM first in 10 mM steps parallel with template optimization in smaller 5 mM steps, plasmid DNA template in range of 200-400 µg/ml, total rNTP concentration between 4-8 mM in 2 mM steps. The optimized values can be found in Table 6.



Table 6 Optimized reaction condition for *in vitro* transcription of GCAA 14 nucleotide RNA tetraloop. Optimization was done in 25  $\mu$ l scale, while the final preparative transcription was between 20-35 ml volume. The % value behind the rNTPs refers to the rNTP distribution in the sequence shown in Figure 39, considering the transcribed region only.

| <b>Reagent</b>        | <b>Concentration</b> |
|-----------------------|----------------------|
| TRIS glutamate buffer | 200 mM               |
| DTT                   | 20 mM                |
| Spermidine            | 2 mM                 |
| Plasmid DNA template  | 300 mg/ml            |
| Mg(OAc) <sub>2</sub>  | 25 mM                |
| rATP (14%)            | 0.84 mM              |
| rUTP (19%)            | 1.14 mM              |
| rGTP (33%)            | 1.98 mM              |
| rCTP (33%)            | 1.98 mM              |
| T7 RNA polymerase     | 9.6 mg/ml            |

Reagents for the transcription were added together on ice and T7 RNA polymerase was always added last to the reaction mixture. The mixture was incubated at 37 °C for 16 h and preparative scale mixture was additionally mixed at 140 rpm. In case of preparative scale transcription one unit/ml Yeast Inorganic Pyrophosphatase (YIPP, New England Biolabs, USA) was added to the mixture after one hour of incubation time to hydrolyse the insoluble magnesium pyrophosphate by-product.

### 2.3. RNA PURIFICATION

The RNA purification involved two major steps. First the RNA was separated from leftover unused rNTPs and enzyme by manual ion exchange chromatography using diethylaminoethyl (DEAE) resin. The mostly RNA containing sample was further purified by RP-IP-HPLC, conducted by Elke Stirnal. In this step self-cleaved HDV ribozyme, uncleaved RNA constructs were separated from the 14 nucleotide RNA.

The DEAE Sepharose (Cytiva, formerly GE Healthcare, USA) was treated with 0.1% diethyl pyrocarbonate (DEPC) solution overnight to denature any possible RNase contamination. The column was washed with boiling hot water to remove any residual DEPC. Two column volume of 0.1 M NaOAc solution, pH 5.5 was used to equilibrate the column. The crude transcription mixture was centrifuged at 4000 g and the supernatant was loaded onto the column. The column was washed with 0.1 M NaOAc, pH 5.5 until the flow through showed no signal in UV/VIS absorption measurement at 260 nm, all unused rNTP is washed away. The RNA of interest was eluted with increasing salt concentration of 0.6 M, 1 M and 2 M NaOAc, pH 5.5 solutions using 5 column volume of each concentration, collecting roughly 5 ml fractions. Any leftover RNA was washed from the column with a final 3 M NaOAc, pH 5.5 solution using 3



column volume. The fractions have been analysed by UV/VIS absorption spectroscopy and by analytical PAGE. Fractions containing the 14-nucleotide long RNA of interest were diluted to 1 M NaOAc, pH 5.5 concentration and precipitated with addition of 4 volumes of EtOH (abs), incubated for 16 hours at -20 °C. The RNA pellet was centrifuged down at 8000 g, for 30 min at 4 °C. The pellet was air-dried and dissolved in ddH<sub>2</sub>O to roughly 100 OD concentration measured by UV/VIS absorption.

The RP-IP-HPLC step was done by Elke Stiral. The chromatography was conducted at 60 °C on a Perfectsil RP18 column size of 10 mm diameter and 250 mm length. Buffer A: was 50 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub>, pH 5.9 buffer, 2 mM tetrabutylammonium hydrogen sulfate and buffer B: was 60% acetonitrile with 50 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub>, pH 5.9, 2 mM tetrabutylammonium hydrogen sulfate. Buffer B leads to an increasing acetonitrile concentration during the gradient separation. The purification was done with constant flow rate of 5 ml/min and a gradient profile described in Table 7 was applied to separate the 14 nucleotide long RNA.

Table 7 Gradient profile for RP-IP-HPLC purification of 14 nucleotide RNA on Perfectsil RP18 preparative scale column. Buffer A: 50 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub> and buffer B: 60% acetonitrile with 50 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub> buffer.

| <i>Time [min]</i> | <i>buffer A [%]</i> | <i>buffer B [%]</i> |
|-------------------|---------------------|---------------------|
| 0                 | 100                 | 0                   |
| 5                 | 65                  | 35                  |
| 30                | 62                  | 38                  |
| 35                | 0                   | 100                 |
| 40                | 0                   | 100                 |
| 43                | 100                 | 0                   |
| 55                | 100                 | 0                   |

The RNA after the HPLC purification was lyophilized. The RNA was dissolved in ddH<sub>2</sub>O and residual solvents from HPLC purification was removed by several cycles of dilution and concentration in centrifugal concentrator Vivaspin® with 2000 MWCO cut-off (Sartorius AG, Germany). The sample was precipitated with an addition of 5 volumes of 2% LiClO<sub>4</sub> acetone solution to ensure no residual buffer or salt remained. The pellet was dissolved in minimal volume of ddH<sub>2</sub>O and the precipitation was repeated two more times. The sample was again washed with ddH<sub>2</sub>O in centrifugal concentrator in 5 cycles of dilution to 2 ml and concentration to 200 µl. The sample buffer was changed by centrifugal concentrator in 5 rounds to 50 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub>, pH 6.4 buffer and concentrated to 270 µl. The RNA was refolded by heating to 95 °C for 5 minutes and cooling down on ice. The purity of the RNA was analysed by PAGE, and the concentration was

determined by UV/VIS absorption spectroscopy on NanoDrop One. The homogeneity of the folding was tested by native PAGE analyses. 5-10v/v% of D<sub>2</sub>O was added to the sample and transferred into a Shigemi NMR tube. Samples in ~99.95% D<sub>2</sub>O was achieved by three cycles of lyophilization and after each lyophilization dissolving in 99% D<sub>2</sub>O and after the last lyophilization in 99.98% D<sub>2</sub>O.

#### 2.4. NMR SPECTROSCOPY

All NMR spectra were recorded on Bruker NMR spectrometers of 600 MHz, 700 MHz, equipped with cryogenically cooled triple resonance <sup>1</sup>H/<sup>15</sup>N/<sup>13</sup>C or <sup>1</sup>H/<sup>13</sup>C/<sup>31</sup>P or quadruple resonance <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P probes. Chemical shift have been externally referenced to the methyl signal of DSS, while <sup>13</sup>C, <sup>15</sup>N and <sup>31</sup>P chemical shifts were indirectly referenced according to reference.<sup>[179]</sup> All spectra were processed with Topspin version 3.2 to 4.0.1. Further analyses have been done with Sparky.<sup>[180]</sup>

### 3. RESULTS AND DISCUSSION

#### 3.1. CHEMICAL SHIFT ASSIGNMENT

As a first step of the analyses  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^{31}\text{P}$  resonances have to be assigned before further analyses could be done, I used IUPAC nomenclature for the numbering schemes in this chapter and it is shown in.

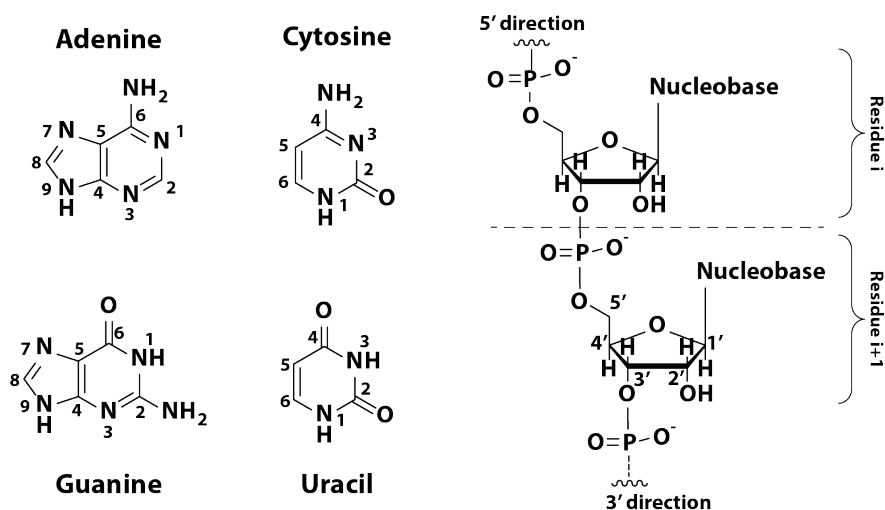


Figure 40 Numbering schemes used in this chapter, according to the IUPAC nomenclature.<sup>[181]</sup>

Nucleotides can be divided into a sugar moiety and a nucleobase. The former has generally low proton chemical shift dispersion and high proton density, while the latter has higher proton chemical shift dispersion but limited number of signals. Additionally, nucleobases have several labile protons, which can be assigned only by measuring samples in  $\text{H}_2\text{O}$  buffer. Utilizing 2D-H,H-NOESY imino protons can be assigned by sequential walk, correlation between the neighbouring imino protons in helical regions. Establishing a connection from imino protons to aromatic signals and amine signals in 2D-H,H-NOESY serves as a starting point in other spectra. For example and 2D-H(C)N<sup>[182]</sup> spectra can directly correlate assigned aromatic H8/H6 signals to H1' in the sugar. For complete assignment, several spectra were used to assign all possible signals and to confirm the assignment. The list of the spectra are: 2D-H(C)N<sup>[182]</sup>, 3D-HCC-TOCSY-CCH-COSY<sup>[183]</sup>, 3D-HCP-TOCSY<sup>[184]</sup>, 3D-HCP<sup>[185]</sup>. Heteronuclear correlations were measured on uniformly isotope labelled samples, while proton homonuclear spectra were measured on non-isotope labelled sample.

## 3.1.1. Sequential connection through the phosphate backbone

The sequential assignment of the backbone was achieved with 3D-HCP and 3D-HCP-TOCSY spectrum, the former correlates the C4'-H4' while the latter the well resolved H1'-C1' resonances of the sugar moieties the chemical shift its own  $^{31}\text{P}_{(i)}$  and its preceding  $^{31}\text{P}_{(i-1)}$  resonances. The major difference between the two pulse sequences is the use of isotropic mixing to transfer the carbon magnetization along the sugar ring, reaching the C1'. In both pulse sequences first, the magnetization is transferred from the proton resonances to the carbon resonances and from there to the phosphorus (all steps achieved by INEPT transfer), where chemical shift evolution takes place. After magnetization is transferred back to the carbon resonances in case of the 3D-HCP-TOCSY an isotropic mixing takes place to transfer the magnetization to the C1' or to C4' by a reverse INEPT transfer in the 3D-HCP pulse sequence. The complete walk of the 3D-HCP-TOCSY spectrum is shown in Figure 41.

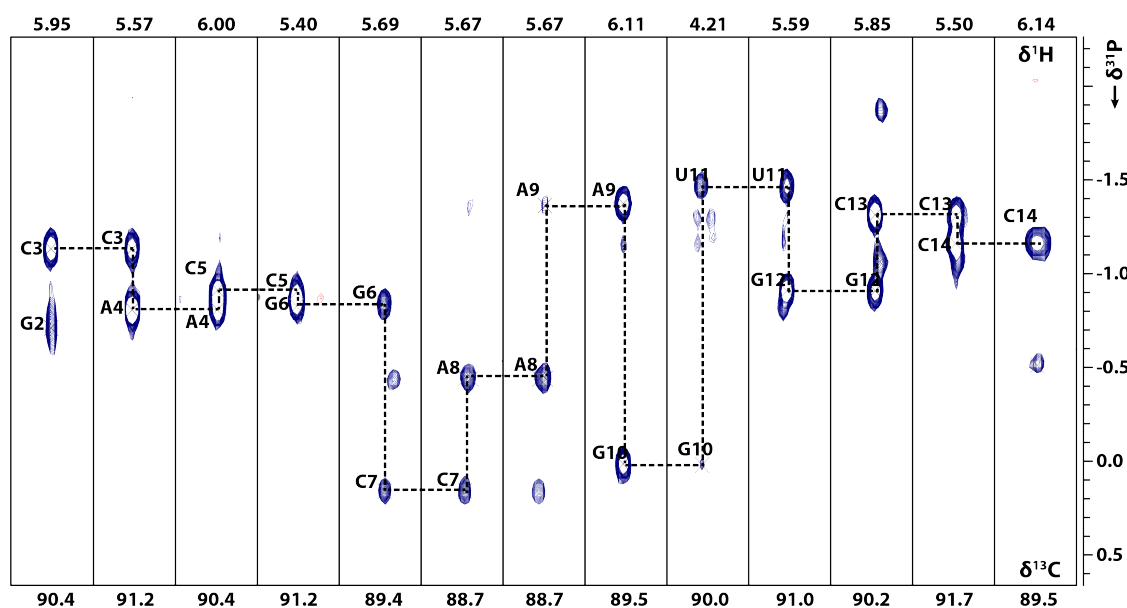


Figure 41 Sequential walk along the phosphate backbone of 14 nucleotide long RNA with GCAA tetraloop. The HCP-TOCSY spectrum correlates the well resolved H1'-C1' resonances to the  $^{31}\text{P}_{(i)}$  and  $^{31}\text{P}_{(i-1)}$  signals. The spectrum was measured at 308 K, with 2048, 64, 96 points in the  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{31}\text{P}$  dimensions, respectively. The A9 H1'-C1'- $^{31}\text{P}_{(i-1)}$  signal is much less intense in comparison to the other signals, likely the  $^3J_{\text{CP}}$  coupling constant is different, as it is in the tetraloop part.

## 3.1.2. Assignment of ribose signals

After the sequential connection was established and parallel the C1'-H1' and C4'-H4' of the sugar resonances were assigned. The other sugar resonances, namely C2'-H2', C3'-H3', C5'-H5' and C5'-H5'' were assigned using the 3D-HCC-TOCSY-CCH-COSY. In this spectrum the  $^{13}\text{C}$  and  $^1\text{H}$  resonances of each sugar

moieties are correlated to the corresponding C1'-H1' resonances, resolving in high resolution the otherwise crowded C2'-H2' and C3'-H3' resonances, an example plane is shown in Figure 42.

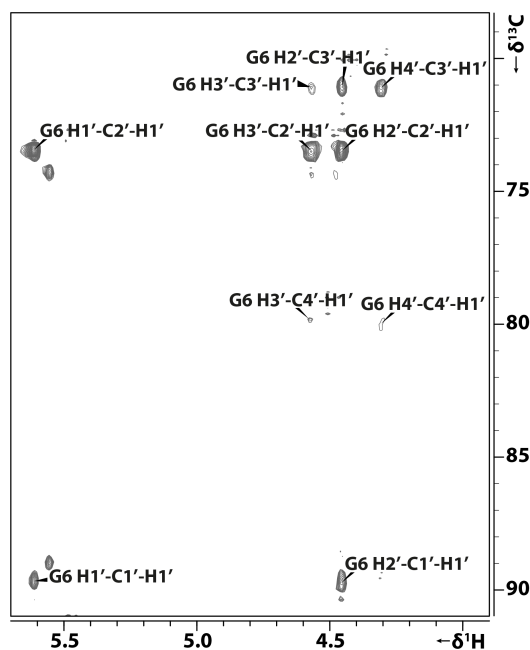


Figure 42 Example plane from 3D-HCC-TOCSY-CCH-COSY spectrum. The plane is at 5.60 ppm, backward directed and autocorrelation peaks are also observed. The pulse sequence was taken from the Bruker Topspin library (na\_hchfwdigp3d). The spectrum was measured with 1024, 68, 96 points in F3, F2, F1 dimensions, respectively at 298 K on a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled sample in  $\text{D}_2\text{O}$  buffer (50 mM  $\text{K}_x\text{H}_y\text{PO}_4$ , 25mM KCl, pH 6.4) with 0.5 mM sample concentration and accumulating 24 scans, in 54 hours.

To achieve correlation through the complete sugar backbone the pulse sequence starts with the evolution of the proton magnetization which is then transferred to the attached carbon via INEPT transfer. This is followed by an isotropic mixing and matching constant time evolution period of the carbon resonances resulting in forward direction, and finally the magnetization is transferred back to the proton where detection takes place. Forward direction can further reduce overlap of signals, by suppressing backward directed and autocorrelation peaks in the spectrum. Forward direction was attempted but so far failed in suppressing the previously mentioned redundant signals. This is likely due to the requirement of the very precise optimized combination of the constant time evolution of the carbon indirect dimension and the isotropic mixing element used to transfer magnetization.<sup>[183]</sup>

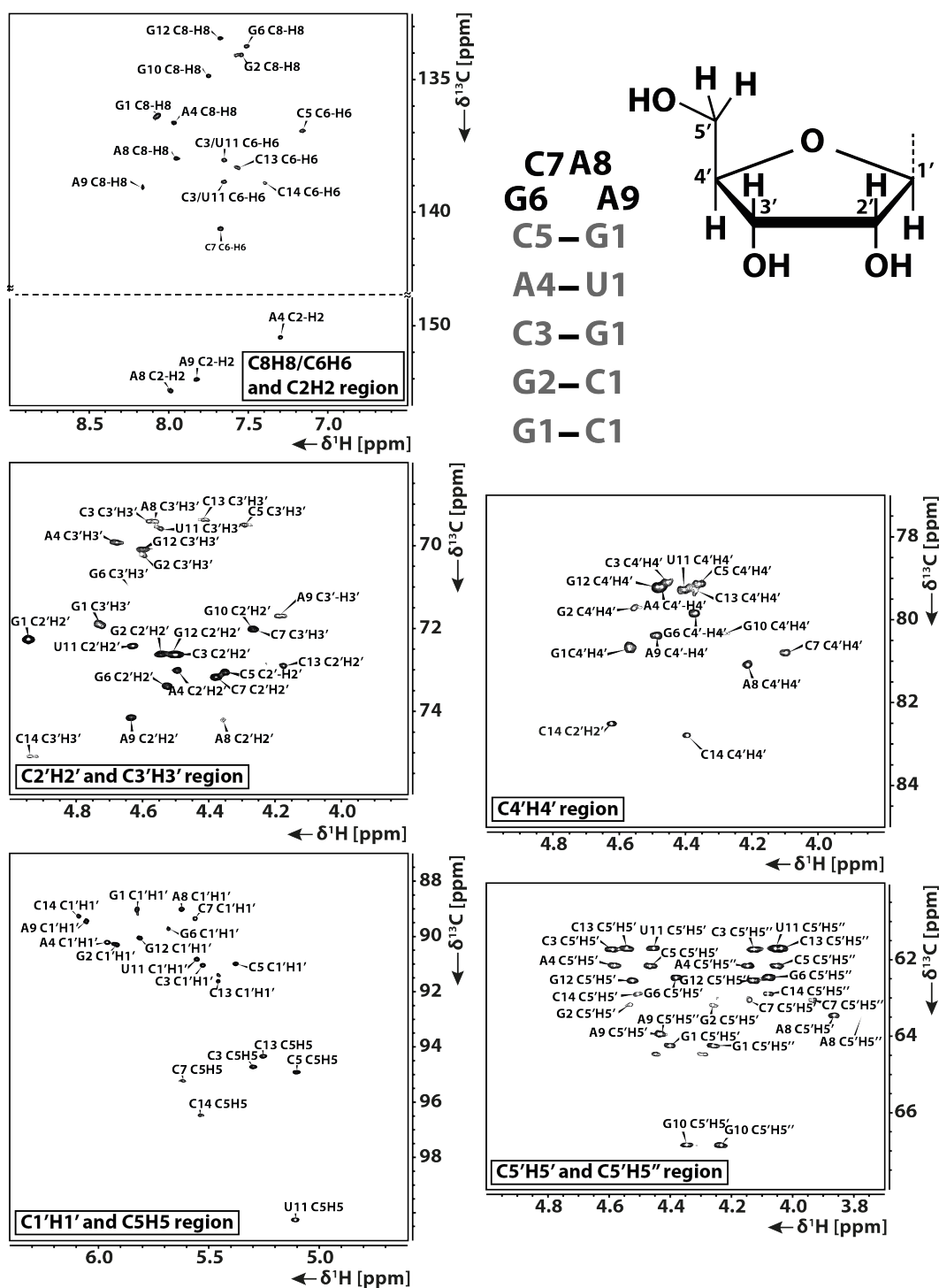


Figure 43 Complete resonance assignment of all protons and its covalently bound carbons in the sugar moieties and nucleobases. The complete chemical shift list can be found in Appendix 1.3. The spectrum was measured at 298 K on a non-isotope labelled sample  $\text{D}_2\text{O}$  buffer (50 mM  $\text{K}_x\text{H}_y\text{PO}_4$ , 25mM KCl, pH 6.4) with 0.7 mM sample concentration.

### 3.1.3. Assignment of $^{15}\text{N}$ signals

The nitrogen atoms can be divided into two groups. One with attached labile protons such as imino and amine groups and the other attached and the other include nitrogen forming the glycosidic bond. Labile proton assignment was conducted on a sample in 90% v/v

H<sub>2</sub>O buffer using HNN-COSY spectrum. In this spectrum the imino protons, previously assigned from NOESY spectrum, is correlated through  $J_{\text{NH}}$  scalar couplings to both the donor and acceptor nitrogen.

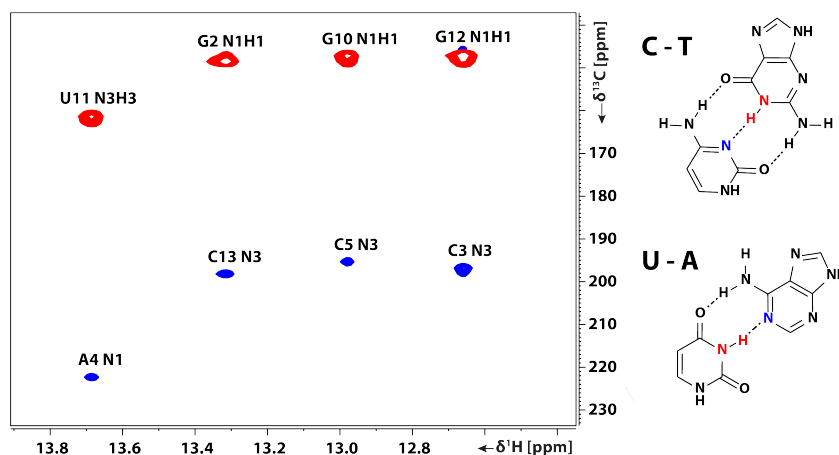


Figure 44 HNN-COSY spectrum used to assign <sup>15</sup>N chemical shifts for imino proton donors (red) and acceptor (blue). The pulse sequence was taken from the Bruker Topspin library (na\_hnncosygpphwg). The spectrum was measured with 2048, 128 points in F2, F1 dimensions, respectively at 298 K on a uniformly <sup>13</sup>C, <sup>15</sup>N labelled sample in 90% v/v H<sub>2</sub>O buffer (50 mM K<sub>x</sub>H<sub>y</sub>PO<sub>4</sub>, 25mM KCl, pH 6.4) with 0.5 mM sample concentration and accumulating 32 scans, in 2 hours.

Nitrogen atoms (N1 for pyrimidine and N3 for purine nucleobases) involved in the glycosidic bond are assigned through a two-dimensional version of the H(C)N spectrum (shown in Figure 45). This spectrum correlates the aromatic H8/H6 signals of the nucleobase to the H1' signals of the ribose moiety through transfer of the magnetization sequentially as shown in Figure 45 schematic via INEPT and reverse INEPT steps utilizing the large  $^1J_{\text{CH}}$  and  $^1J_{\text{CN}}$  couplings. The transfer delays were optimized to have both aromatic and sugar signals observable, although the transfer efficiency was not optimal for all residues (e.g.: G10 or C7) resulting in lower signal intensity compared to other signals.

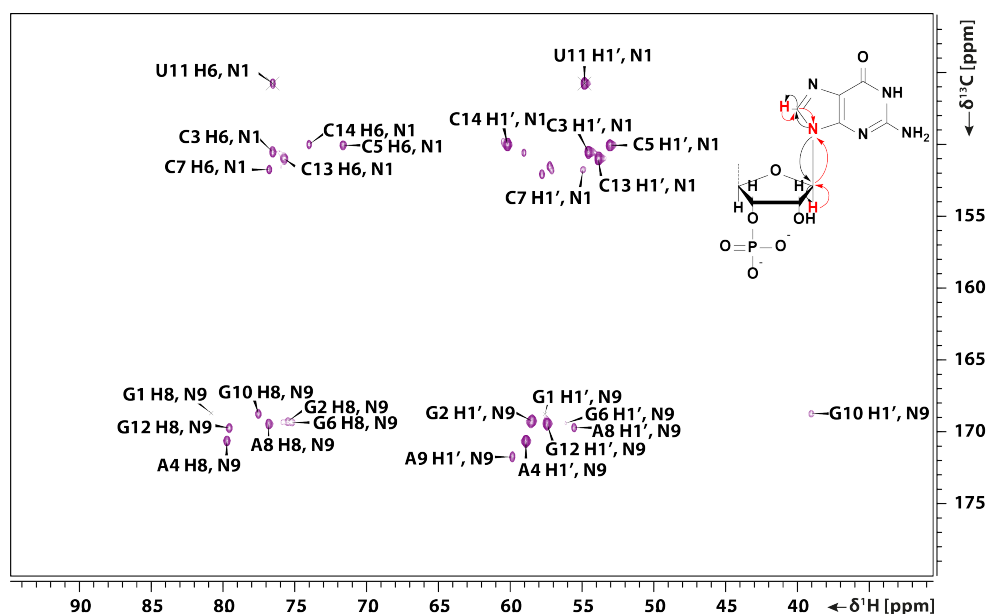


Figure 45 H(C)N spectrum used to assign  $^{15}\text{N}$  chemical shifts of N1 and N3. The pulse sequence was taken from the Bruker Topspin library (na\_hcnetgpsi3d). The spectrum was measured with 1024, 256 points in F2, F1 dimensions, respectively at 298 K on a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled sample in 90% v/v  $\text{H}_2\text{O}$  buffer (50 mM  $\text{K}_x\text{H}_y\text{PO}_4$ , 25mM KCl, pH 6.4) with 0.5 mM sample concentration and accumulating 16 scans, in 1.5 hours.

### 3.1.4. Stereoselective assignment of proS and proR of H5' and H5'' signals

Stereoselective assignment of the H5' and H5'' signals is important as they also define backbone angle  $\gamma$  ( $\text{O}5'-\text{C}5'-\text{C}4'-\text{C}3'$ ) and  $\beta$  ( $\text{P}5'-\text{O}5'-\text{C}5'-\text{C}4'$ ), which are determined by measuring  $^3J_{\text{H}5'-\text{P}}$  coupling. Furthermore, the NOE contacts used in structure calculations can be more accurately used if H5' and H5'' are stereospecifically assigned. This can be achieved by measuring coupling constants of  $^2J_{\text{C}5'\text{H}5'}/^2J_{\text{C}5'\text{H}5''}$  and  $^3J_{\text{H}4'\text{H}5'}/^3J_{\text{H}4'\text{H}5''}$  or utilizing the proton and carbon chemical shifts of H5'/H5'' and C5'. Chemical shift analyses based on previous observation can give preliminary assignment. The analyses of the difference of the H5'-H5'' plotted against the C5' chemical shifts can be used.<sup>[169]</sup>



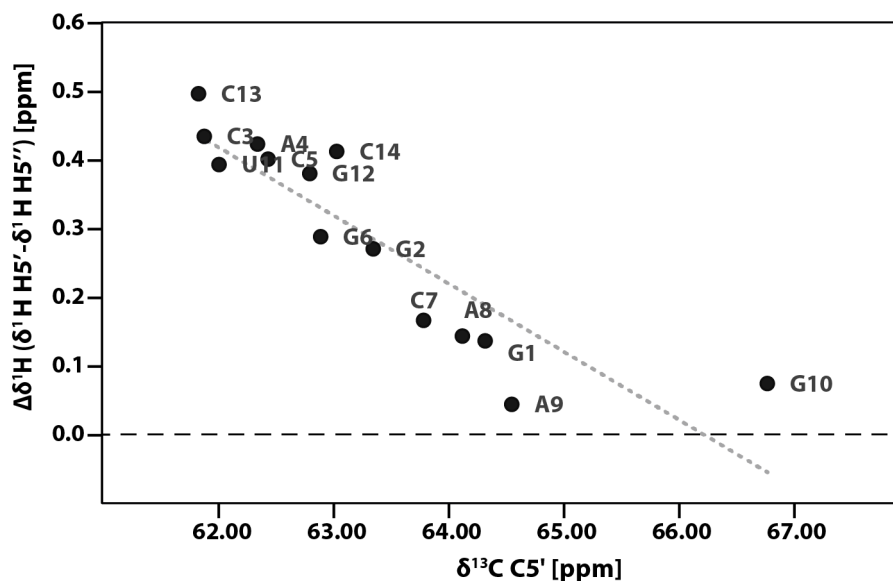


Figure 46 H5' and H5'' chemical shifts difference plotted against the C5' chemical shifts. G10 likely adopts the “non-canonical” state where the proR is upfield shifted, while all other residues follow the canonical trend of proS being upfield shifted.

Based on these analysis shown in Figure 46 all nucleotides retain the “canonical”, in double helix conformation observed, as proS is the upfield shifted signal except for G10. These results are based only on empirical observations and to verify these correlations further  ${}^2J_{C5'H5'}/{}^2J_{C5'H5''}$  and  ${}^3J_{H4'H5'}/{}^3J_{H4'H5''}$  coupling constants have to be measured.

### 3.2. OBSERVATION OF TEMPERATURE DEPENDANT DYNAMICS

Sugar pucker is an important measure for the structure of the nucleotides, as in a standard A-helix conformation this is strictly limited to the C3'-endo conformation. In any non-canonical conformation, the sugar-phosphate backbone changes its conformation. NMR spectroscopy offers different methods to determine the sugar pucker. A qualitative analysis can be fulfilled by analysing the carbon chemical shifts, based on empirical investigation of the chemical shift changes.<sup>[186,187]</sup> A modified version of the previously described 3D-HCC-TOCSY-CCH-COSY pulse sequence where coupling constants are not refocused and utilizing the high coupling constant value of  ${}^1J_{CH}$  scalar coupling, even small  ${}^3J_{HH}$  homonuclear coupling constants can be determined. Finally, the most precise method, albeit the least sensitive measurement is based on determination of the cross-correlated relaxation rates.

## 3.2.1. Canonical coordinates

Canonical coordinates can be used directly after the full assignment of the resonances of the sugar moieties are established. It is based on an empirical calculation using the chemical shifts of the sugar backbone according to Eq. 4 and Eq. 5.<sup>[186]</sup>

$$can1^* = -14.7 \cdot \delta_{C1'} + 22.1 \cdot \delta_{C2'} + 13.2 \cdot \delta_{C3'} + 6.5 \cdot \delta_{C4'} - 2.9 \cdot \delta_{C5'} - 1595 = P_{FIT} \quad \text{Eq. 4}$$

$$can2^* = 9.8 \cdot \delta_{C1'} + 16.5 \cdot \delta_{C2'} - 0.5 \cdot \delta_{C3'} - 1.7 \cdot \delta_{C4'} + 13.5 \cdot \delta_{C5'} - 2781 = \gamma_{FIT} \quad \text{Eq. 5}$$

This results in two parameters describing the sugar conformation: the pseudo rotation phase and the exocyclic angle  $\gamma$ . The results of the calculation at different temperatures can be seen in Figure 48. If the pseudorotation phase P (see Figure 47) angle is below  $72^\circ$  the sugar conformation falls into the C3'-endo conformation, while above  $72^\circ$  it is C2'-endo, as these conformations are at local energy minima, compared to exo form.

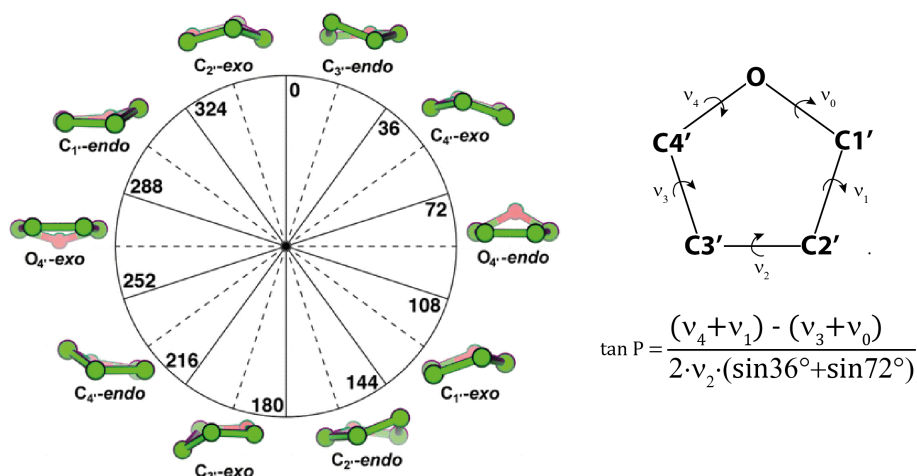


Figure 47 Pseudorotation phase cycle of ribose based on Altona-Sundaralingam definition<sup>[188]</sup> the figure was adopted from reference<sup>[189]</sup>

For standard A helix conformation in RNA, the ribose adopts C3'-endo conformation. In non-canonical regions like the loop A7, A8 and G9 adopt the C2'-endo conformation. Interestingly with decreasing temperature all riboses in the loop move towards the C3'-endo conformation, and A7 crosses this barrier. The torsion angle  $\gamma$  can be best described by using Newman projections along the C4'-C5' bond.

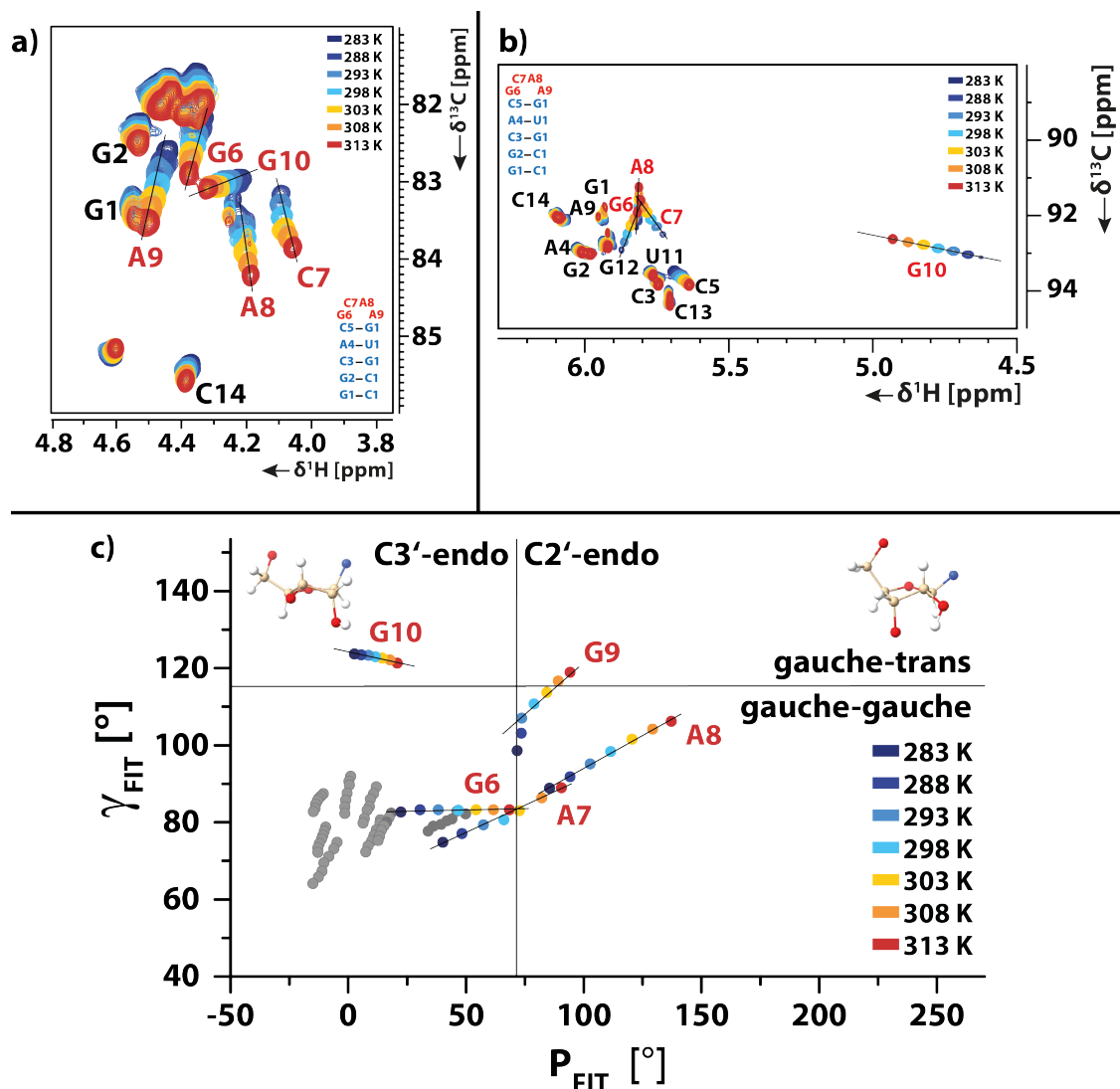


Figure 48 Temperature series of  $^{13}\text{C}, ^1\text{H}$ -HSQC spectrum showing a) the region of the spectrum with C4' signals and b) with C1' region of GCAA 14 nt long RNA tetraloop, respectively. The colour coding of the respective temperatures is shown in the figure. c) is the calculated temperature dependant canonical coordinates according to Eq. 4 and Eq. 5.

Gauche-trans and gauche-gauche conformations are energetically favourable and in canonical A-helix secondary structure  $\gamma$  is found in the gauche-gauche conformation. The loop signals and G10 from the closing base pair show significant temperature dependant chemical shift changes see in Figure 48 a) and b). This is likely due to dynamic changes in the sugar backbone, as the sugar pucker changes with temperature based on the canonical coordinates, the closing base pair G10 and at higher temperatures residue G9 adopt the gauche-trans conformation, and G9 and A7 crosses the border between C3'-endo and C2'-endo depending on the temperature.

The temperature dependence of the sugar pucker is an intriguing feature of the GCAA tetraloop. In comparison, the well-studied UUCG tetraloop<sup>[156]</sup> shows very minimal

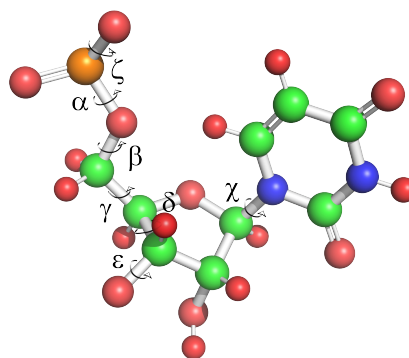
temperature dependence (unpublished results from Robbin Schnieders), similar to the stem signals. This temperature dependence suggests at least two conformations exist in fast exchange (only one set of signals were observed), that differ in their sugar conformations and change their population depending on the temperature. Additionally, the loop signals linewidth become broader at lower temperatures, below 283 K the signals are broadened beyond detection (not shown results). This is due to the exchange rate slows down and reaches the intermediate exchange regime. The slow exchange regime could not be reached by cooling it down to -5 °C. Recently, a new theory that even the very rigid UUCG tetraloop has an alternative conformation<sup>[170,171]</sup> have been suggested. In this alternative form the non-Watson Crick CG base pair is not formed and the C8 and G9 nucleobases flip into the solution. This has been confirmed using eNOE data in combination with MD simulations. Previous NMR studies of the GCAA, CUUG and GAAG tetraloops by Pardi et al.,<sup>[190]</sup> all resulted in a single conformation based on NOE structure calculations. Considering recent developments in RNA energy landscapes and our observations of temperature dependant chemical shift changes and canonical coordinates also point into the direction that these structures should be re-evaluated.

### 3.3. TORSION ANGLES

NMR is suitable to study the dihedral angles of oligonucleotides. This can give information about changes compared to canonical backbone conformation around the loop residues and the closing base pair. Oligonucleotides have a total of six torsion angles that define the backbone and the glycosidic torsion angle  $\chi$ , that defines the rotation of the nucleobase compared to the ribose ring. The angles can be determined either based on measured scalar couplings and using a suitable Karplus equation to calculate the angle, or cross-correlated relaxation rates can be used to calculate the angle. The latter has the advantage of being free from empirical parametrization, but the sensitivity of these experiments is generally low. A list of the angles and their possible corresponding NMR parameters for determination are listed in Table 8.

Table 8 List of torsion angles found in oligonucleotides and their corresponding NMR parameters to determine them. The angles are illustrated on a single uridine nucleotide.

| Angle      | Experimental parameters  |
|------------|--|
| $\alpha$   | $\Gamma_{C4'H4',P}^{DD,CSA} \mid \Gamma_{C5'H5'/H5'',P}^{DD,CSA}$        |
| $\beta$    | ${}^3J_{H5'/H5'',P} \mid {}^3J_{C4',Pi}$                                 |
| $\gamma$   | ${}^3J_{H4',H5'/H5''} \mid {}^3J_{C3',H5'/H5''}$                         |
| $\delta$   | ${}^3J_{H3',H4'} \mid {}^3J_{C5',H3'} \mid {}^3J_{C2',H4'}$              |
| $\epsilon$ | ${}^3J_{H3',P} \mid {}^3J_{C2',Pi} \mid {}^3J_{C4',P(i+1)}$              |
| $\zeta$    | $\Gamma_{C4'H4',P_{i+1}}^{DD,CSA} \mid \Gamma_{C3'H3',P_{i+1}}^{DD,CSA}$ |
| $\chi$     | $\Gamma_{C8H8,C1'H1'}^{DD,DD} \mid \Gamma_{C6H6,C1'H1'}^{DD,DD}$         |



### 3.3.1. Sugar puckering from homonuclear scalar coupling constants

${}^3J_{HH}$  scalar couplings in the ribose ring provide information about the sugar conformation. Generally due to the linewidth observed in biomolecules small  ${}^3J_{HH}$  couplings cannot be directly measured. For RNA, the three dimensional HCC-TOCSY-CCH-E.COSY spectrum developed by Schwalbe et al.,<sup>[183]</sup> circumvents this problem by not removing the large  ${}^1J_{CH}$  coupling and utilizing it to develop signals in E.COSY pattern. An example peak is shown in Figure 49 where the different coupling constants are visible around the corresponding resonance frequency.

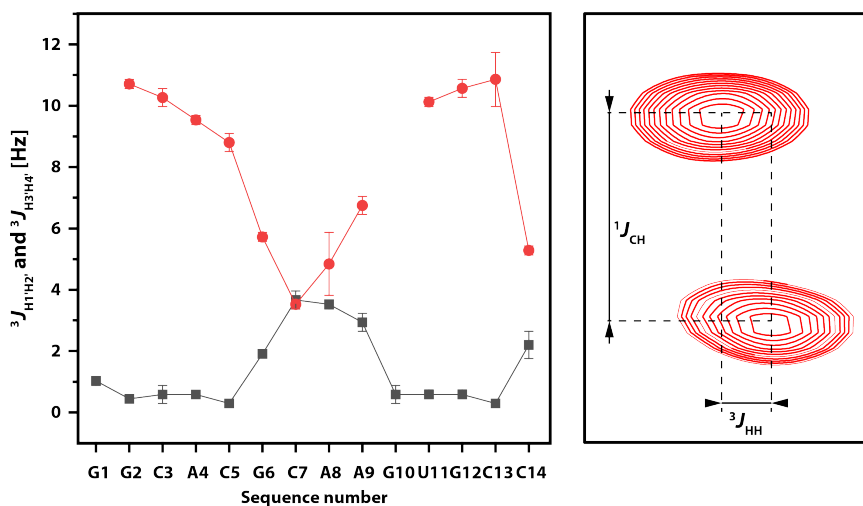


Figure 49 Measurement of  ${}^3J_{H1'H2'}$  (black square) and  ${}^3J_{H3'H4'}$  (red dot) scalar couplings measured with HCC-TOCSY-CCH-E.COSY pulse sequence. Values obtained are shown on the left, and an example peak with  ${}^1J_{CH}$  coupling in the indirect dimension and the  ${}^3J_{HH}$  coupling as a difference between peak centre position is shown. The measurement was done at 308 K.

The  ${}^3J_{H1'H2'}$  coupling constant in canonical double helix region falls between 0-2 Hz, while the  ${}^3J_{H3'H4'}$  coupling constant is usually around  $\sim 10$  Hz. These values are typical for C3'-endo conformation of the ribose. This agrees with measured values for the stem (G2-C5, U11-C13). In case of the C2'-endo conformation the coupling constants show the reverse trend as  ${}^3J_{H1'H2'}$  would be small 0-2 Hz and  ${}^3J_{H3'H4'}$  is bigger about  $\sim 10$  Hz. In the UUCG tetraloop U7 and C8 signals fall into the C2'-endo conformation as with the coupling constants are 8.0 and 9.3 Hz for  ${}^3J_{H1'H2'}$  and 1.0 and 0.6 for  ${}^3J_{H3'H4'}$ , respectively. For the GCAA tetraloop, all four loop signals show an average value for both coupling constants between the expected small (0-2 Hz) and large (10 Hz) value. This can only be explained as an average value stemming from different conformations, resulting in an ensemble observed parameter. This is in good agreement with the temperature dependant canonical coordinate trend observed previously.

### 3.3.2. Phosphate backbone relevant scalar couplings and cross-correlated relaxation rates

The phosphate backbone angles  $\varepsilon$  and  $\beta$  requires the measurement of  ${}^3J_{H,P}$  and/or  ${}^3J_{C,P}$  scalar couplings. The latter for C2' and C4' atoms can be measured with quantitative HCP<sup>[192]</sup> experiment. This result in an HSQC like spectrum where the signal intensities are modulated according to  $\sin^2(\pi J(C,P)\tau_{CT})$  or  $\cos^2(\pi J(C,P)\tau_{CT})$  in case of cross and reference experiment, respectively. Using Eq. 6 we can determine the  $J_{CP}$  scalar coupling using the intensities of the cross and reference experiments.

$$J_{CP} = \frac{1}{\pi\tau_{CT}} \arctg \left( \sqrt{\frac{I_{cross}}{I_{ref}} \cdot \frac{nS_{ref}}{nS_{cross}}} \right) \quad \text{Eq. 6}$$

The values are shown in Figure 50 for the C2' signals. Note that signals from the stem were not observed in the cross experiment, this is due to the generally  $< 2$  Hz value of  ${}^3J_{C2',P}$  scalar coupling constant, which is normal for canonical A-form helix conformation. The very small coupling constants lead to inefficient magnetization transfer and finally to the loss of signals. These values for the loop signals already suggest that it adopts a non-canonical conformation.

Additionally, the two-dimensional version of this measurement does not allow determination of the coupling constants of the C4' atom as it is coupled to both  $P_{(i)}$  and  $P_{(i+1)}$  and couplings could not be differentiated. This could be overcome by the introduction of the phosphorus indirect dimension and measuring the quantitative HCP

spectrum as a three-dimensional spectrum, or if the resolution allows as a two dimensional proton-phosphorous spectrum, H(C)P.

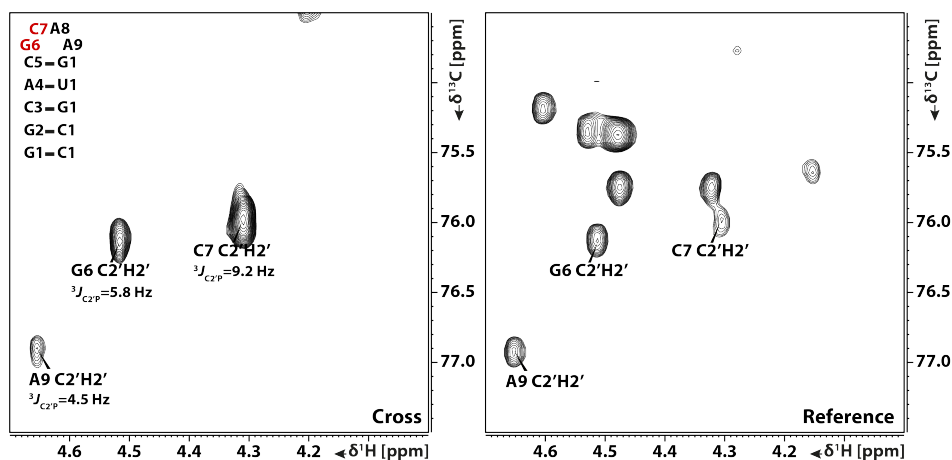


Figure 50 C2' region of quantitative HC(P) spectrum of cross (left) and reference (right) experiments. Both experiments were measured on a 500  $\mu$ M sample in 50 mM  $K_xH_yPO_4$ , pH 6.4 buffer at 308 K with 128 scans and 32 scans accumulated for cross and reference experiment, respectively with 256 points recorded in the indirect dimension, the constant time for coupling evolution was 12.5 ms. The  $^3J_{C2',P}$  scalar couplings were calculated according to Eq. 6 and are shown on the spectrum under the peak label.

The only method to determine the  $\alpha$  and  $\zeta$  angles is the measurement of the following cross correlated relaxation rates:  $\Gamma_{C4'H4',P}^{DD,CSA}$ ,  $\Gamma_{C5'H5'/H5'',P}^{DD,CSA}$ ,  $\Gamma_{C4'H4',P_{i+1}}^{DD,CSA}$ ,  $\Gamma_{C3'H3',P_{i+1}}^{DD,CSA}$ , because the phosphorus atom is only bound to NMR inactive oxygen atoms. This method has been developed by Nozinovic et al.<sup>[167]</sup> It requires the measurement of a reference experiment and a cross experiment. During the pulse sequence the DQC/ZQC spin order of  $H_zC_yP_y$  evolves into two terms modulated by  $\sinh(\Gamma_{CH,P}^{DD,CSA}T_m)$  or  $\cosh(\Gamma_{CH,P}^{DD,CSA}T_m)$  due to the cross correlated relaxation effect taking place during the  $T_m$  mixing time. By coherence selection in the cross and reference experiment the different operators are selected. Due to the severe insensitivity of the cross experiment and likely due to the conformational dynamics it was not yet successful (see Figure 51) to calculate the relaxation rates and the corresponding angles.

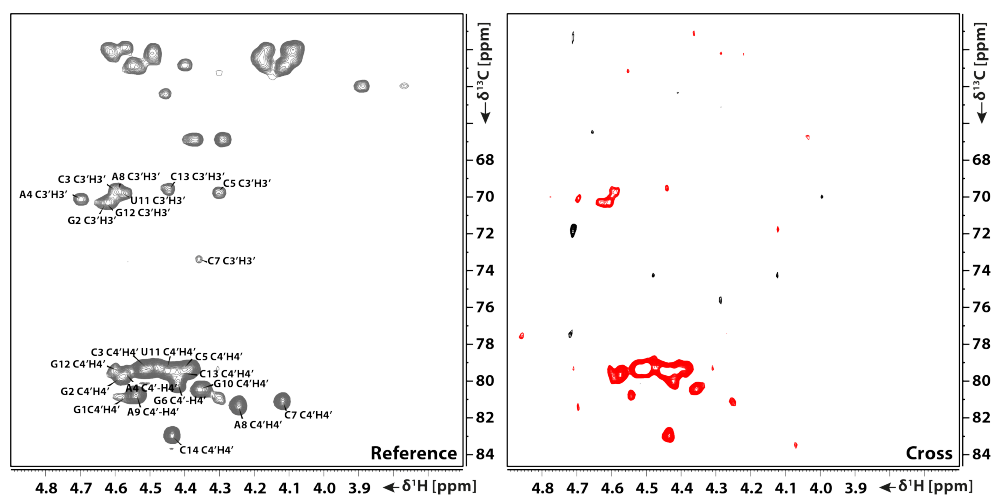


Figure 51 Preliminary results of  $\Gamma$ -HCP measurements. Due to the insensitivity of the measurement in the two-dimensional version only weak signals of C3' sugar signals were observed, while strong C4' sugar signals are observed but due to the coupling to both 3' and 5' direction to the phosphate three-dimensional measurements are required. Both experiments were measured on a 500  $\mu$ M sample in 50 mM  $K_xH_yPO_4$ , pH 6.4 buffer with 380 scans accumulated at 308 K and 80 points were measured in the indirect dimension, the mixing time was 5 ms. The total measurement time was 16 hours.

The most intense peaks observed in the two-dimensional spectrums are C4' sugar signals, but due to the coupling to both  $P_{(i)}$  and  $P_{(i-1)}$  nuclei no meaningful cross-correlated relaxation rate can be determined. Therefore, the three-dimensional version of the experiment has to be used, probably setting the number of increments to one in the carbon indirect dimension and recording the lower resolution proton-phosphorus two-dimensional spectrum H(C)P.

### 3.4. LONGITUDINAL RELAXATION AND HETERONUCLEAR NOE MEASUREMENTS

Relaxation measurements can also provide invaluable information about the underlying dynamics of RNAs. The different methods such as ZZ-exchange<sup>[193]</sup>, H(D) exchange and relaxation dispersion<sup>[194]</sup> provide information about dynamics happening on the millisecond to second timescale. Alternatively, for very fast ps-ns timescale dynamics, the use of standard longitudinal and transverse relaxation with heteronuclear NOE using for example the Lipari-Szabo model free formalism<sup>[195]</sup> can provide information.



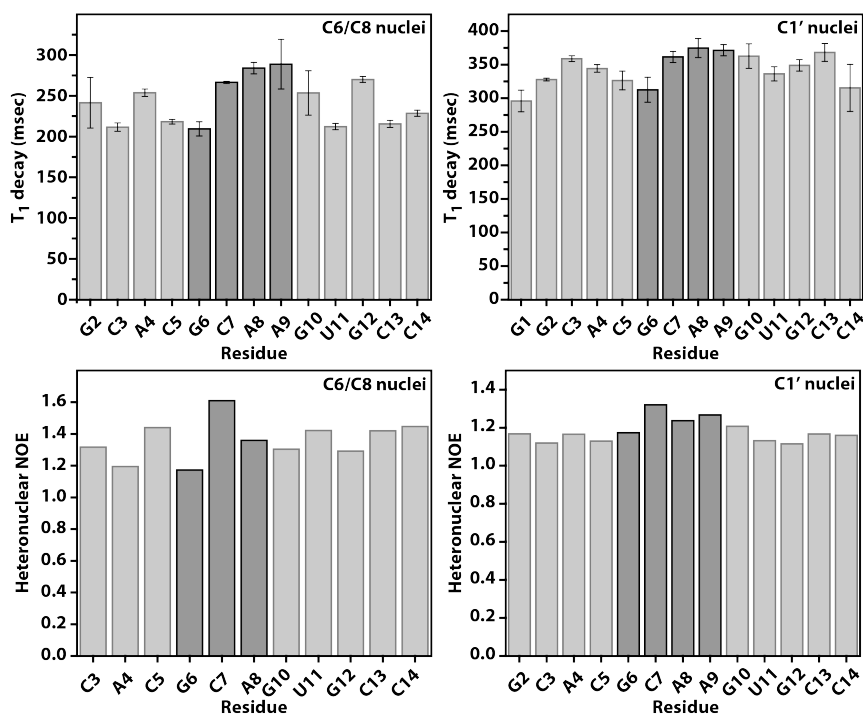


Figure 52 T<sub>1</sub> relaxation of aromatic residues (top left) and sugar C1' nuclei (top right) and the heteronuclear NOE on aromatic residues (bottom left) and sugar C1' nuclei (bottom right). Dark grey bars are the loop nucleotides.

Relaxation measurements and especially transverse relaxation of carbons are limited in uniformly <sup>13</sup>C, <sup>15</sup>N labelled nucleic acids. Intensive homonuclear <sup>13</sup>C scalar coupling networks in both the aromatic and ribose moieties limits transverse relaxation measurements. Usually instead of CPMG type of measurements R1ρ is measured.

These preliminary relaxation measurements show no outstanding difference between the loop and stem nuclei relaxation behaviour. In literature relaxation dispersion on selectively <sup>13</sup>C labelled C2' and C4' nuclei in the sugar moieties revealed on the millisecond timescale exchanges, in agreement with observation of scalar couplings.<sup>[174]</sup> This would indicate that for more details about so called excited states (ES) relaxation dispersion measurements should be conducted.<sup>[194]</sup>

## 4. SUMMARY AND OUTLOOK

MD simulations of RNA is an important complementary tool to experimental research, but force fields used in simulations require further fine tuning. Force fields are developed and tested on experimental results of crystal structures, NMR structures and parameters. The latter has the advantage to be able to describe dynamic systems, which should also be potentially present in the results of MD simulations as differently populated trajectories. In this project our aim is to make three new very well characterized model systems (GCAA, GAAG and CUUG tetraloops), that would further stimulate computer-based research of RNA. These tetraloops have been selected based on their natural occurrence in longer functional RNA molecules and their suggested conformational dynamics.

Recent findings showed a low-populated alternative conformation of the UUCG tetraloop, which was considered a very rigid, conformationally homogenous system. Our findings show a much more dynamic system with the GCAA tetraloop. The chemical shifts of the loop signals show remarkably high temperature dependence. At 278 K all the loop signals and the closing base pair signals are undetectable, due to the intermediate exchange between different conformation(s). With the complete assignment temperature dependant canonical coordinates were calculated suggesting conformational changes in the sugar puckering.

The scalar  $^3J_{\text{HH}}$  couplings of the sugar moieties also supported different population of sugar puckering conformations, as the loop residues at 308 K show average values of the C3'-endo and C2'-endo conformations. This agrees with the detailed relaxation studies found in literature using  $^{13}\text{C}$  labelling selectively at the C2' and C4' position. Both model free analysis and relaxation dispersion showed intense exchange involving the ribose moieties.

Preliminary cross correlated relaxation measurements showed the requirement of higher concentrated samples and / or additional accumulation of scans at optimized temperature. Alternatively, theoretical zero values can be also observed if the average of the cross correlated relaxation rates in the two conformations show different signs. Nevertheless, with possible dihedral angle constraints measured and the NOE information a structure calculation can be done.

For more detailed characterization of the dynamic behaviour or structural characterization of the alternative conformation(s) relaxation-based NMR methods, RDC measurements and the newly developed eNOEs could be used. If the slow exchange regime could be reached in supercooled water,<sup>[68,196–198]</sup> both conformers could be separately examined. GCAA tetraloop is a promising model structure to capture the highly dynamic state of RNAs by NMR spectroscopy and to aid the development of MD simulations on nucleic acids.



## REFERENCES

- [1] Jahn, T. R.; Radford, S. E. The Yin and Yang of Protein Folding. *FEBS J.* **2005**, *272* (23), 5962–5970.
- [2] Roder, H.; Shastry, M. R. Methods for Exploring Early Events in Protein Folding. *Curr. Opin. Struct. Biol.* **1999**, *9* (5), 620–626.
- [3] Nölting, B.; Golbik, R.; Fersht, A. R. Submillisecond Events in Protein Folding. *Proc. Natl. Acad. Sci.* **1995**, *92* (23), 10668–10672.
- [4] Gruebele, M.; Sabelko, J.; Ballew, R.; Ervin, J. Laser Temperature Jump Induced Protein Refolding. *Acc. Chem. Res.* **1998**, *31* (11), 699–707.
- [5] Jones, C. M.; Henry, E. R.; Hu, Y.; Chan, C. K.; Luck, S. D.; Bhuyan, A.; Roder, H.; Hofrichter, J.; Eaton, W. A. Fast Events in Protein Folding Initiated by Nanosecond Laser Photolysis. *Proc. Natl. Acad. Sci.* **1993**, *90* (24), 11860–11864.
- [6] Roder, H.; Maki, K.; Cheng, H. Early Events in Protein Folding Explored by Rapid Mixing Methods. *Chem. Rev.* **2006**, *106* (5), 1836–1861.
- [7] Agashe, V. R.; Shastry, M. C. R.; Udgaonkar, J. B. Initial Hydrophobic Collapse in the Folding of Barstar. *Nature* **1995**, *377* (6551), 754–757.
- [8] Karplus, M.; Weaver, D. L. Protein Folding Dynamics: The Diffusion-Collision Model and Experimental Data. *Protein Sci.* **1994**, *3* (4), 650–668.
- [9] Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. The Structure of the Transition State for Folding of Chymotrypsin Inhibitor 2 Analysed by Protein Engineering Methods: Evidence for a Nucleation-Condensation Mechanism for Protein Folding. *J. Mol. Biol.* **1995**, *254* (2), 260–288.
- [10] Kumar, A.; Balbach, J. Real-Time Protein NMR Spectroscopy and Investigation of Assisted Protein Folding. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2015**, *1850* (10), 1965–1972.
- [11] Korzhnev, D. M.; Kay, L. E. Probing Invisible, Low-Populated States of Protein Molecules by Relaxation Dispersion NMR Spectroscopy: An Application to Protein Folding. *Acc. Chem. Res.* **2008**, *41* (3), 442–451.
- [12] Zeeb, M.; Balbach, J. NMR Spectroscopic Characterization of Millisecond Protein Folding by Transverse Relaxation Dispersion Measurements. *J. Am. Chem. Soc.* **2005**, *127* (38), 13207–13212.
- [13] Neudecker, P.; Lundström, P.; Kay, L. E. Relaxation Dispersion NMR Spectroscopy as a Tool for Detailed Studies of Protein Folding. *Biophys. J.* **2009**, *96* (6), 2045–2054.
- [14] Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. Comparison of the Backbone Dynamics of a Folded and an Unfolded SH3 Domain Existing in Equilibrium in Aqueous Buffer. *Biochemistry* **1995**, *34* (3), 868–878.
- [15] Englander, S. W. Protein Folding Intermediates and Pathways Studied by Hydrogen Exchange. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29* (1), 213–238.
- [16] Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. A Heteronuclear Correlation Experiment for Simultaneous Determination of <sup>15</sup>N Longitudinal Decay and Chemical Exchange Rates of Systems in Slow Equilibrium. *J. Biomol. NMR* **1994**, *4* (5), 727–734.
- [17] Balbach, J.; Forge, V.; Lau, W. S.; Nuland, N. A. J. van; Brew, K.; Dobson, C. M. Protein Folding Monitored at Individual Residues During a Two-Dimensional NMR Experiment. *Science* **1996**, *274* (5290), 1161–1163.

- [18] Spraul, M. D.; Hofmann, M.; Schwalbe, H. D. NMR Measuring Cell Esp. for Flow through Sample Head for Protein Folding. DE19548977C1, February 27, 1997.
- [19] Mok, K. H.; Nagashima, T.; Day, I. J.; Jones, J. A.; Jones, C. J. V.; Dobson, C. M.; Hore, P. J. Rapid Sample-Mixing Technique for Transient NMR and Photo-CIDNP Spectroscopy: Applications to Real-Time Protein Folding. *J. Am. Chem. Soc.* **2003**, *125* (41), 12484–12492.
- [20] Kühn, T.; Schwalbe, H. Monitoring the Kinetics of Ion-Dependent Protein Folding by Time-Resolved NMR Spectroscopy at Atomic Resolution. *J. Am. Chem. Soc.* **2000**, *122* (26), 6169–6174.
- [21] Wirmer, J.; Kühn, T.; Schwalbe, H. Millisecond Time Resolved Photo-CIDNP NMR Reveals a Non-Native Folding Intermediate on the Ion-Induced Refolding Pathway of Bovine  $\alpha$ -Lactalbumin. *Angew. Chem.* **2001**, *113* (22), 4378–4381.
- [22] Schlepckow, K.; Wirmer, J.; Bachmann, A.; Kiefhaber, T.; Schwalbe, H. Conserved Folding Pathways of  $\alpha$ -Lactalbumin and Lysozyme Revealed by Kinetic CD, Fluorescence, NMR, and Interrupted Refolding Experiments. *J. Mol. Biol.* **2008**, *378* (3), 686–698.
- [23] Balbach, J.; Forge, V.; van Nuland, N. A. J.; Winder, S. L.; Hore, P. J.; Dobson, C. M. Following Protein Folding in Real Time Using NMR Spectroscopy. *Nat. Struct. Biol.* **1995**, *2* (10), 865–870.
- [24] Alderson, T. R.; Charlier, C.; Torchia, D. A.; Anfinrud, P.; Bax, A. Monitoring Hydrogen Exchange During Protein Folding by Fast Pressure Jump NMR Spectroscopy. *J. Am. Chem. Soc.* **2017**, *139* (32), 11036–11039.
- [25] Charlier, C.; Courtney, J. M.; Anfinrud, P.; Bax, A. Interrupted Pressure-Jump NMR Experiments Reveal Resonances of On-Pathway Protein Folding Intermediate. *J. Phys. Chem. B* **2018**, *122* (49), 11792–11799.
- [26] Charlier, C.; Courtney, J. M.; Alderson, T. R.; Anfinrud, P.; Bax, A. Monitoring  $^{15}\text{N}$  Chemical Shifts During Protein Folding by Pressure-Jump NMR. *J. Am. Chem. Soc.* **2018**, *140* (26), 8096–8099.
- [27] Charlier, C.; Alderson, T. R.; Courtney, J. M.; Ying, J.; Anfinrud, P.; Bax, A. Study of Protein Folding under Native Conditions by Rapidly Switching the Hydrostatic Pressure inside an NMR Sample Cell. *Proc. Natl. Acad. Sci.* **2018**, *115* (18), E4169–E4178.
- [28] Akasaka, K.; Naito, A.; Imanari, M. Novel Method for NMR Spectral Correlation between the Native and the Denatured States of a Protein. Application to Ribonuclease A. *J. Am. Chem. Soc.* **1991**, *113* (12), 4688–4689.
- [29] Akasaka, K.; Naito, A.; Nakatani, H. Temperature-Jump NMR Study of Protein Folding: Ribonuclease A at Low PH. *J. Biomol. NMR* **1991**, *1* (1), 65–70.
- [30] Akasaka, K.; Naito, A.; Nakatani, H.; Imanari, M. Construction and Performance of a Temperature-jump NMR Apparatus. *Rev. Sci. Instrum.* **1990**, *61* (1), 66–68.
- [31] Gullion, T.; Conradi, M. S. Temperature-Jump NMR: Molecular Twisting at the Phase Transition in Para-Terphenyl. *Phys. Rev. B* **1984**, *30* (3), 1133–1137.
- [32] Maresch, G. G.; Kendrick, R. D.; Yannoni, C. S. High-temperature NMR Using Inductive Heating. *Rev. Sci. Instrum.* **1990**, *61* (1), 77–80.
- [33] Massiot, D.; Bessada, C.; Echegut, P.; Coutures, J. P.; Taullele, F. High Temperature NMR Study of Lithium Sodium Sulfate. *Solid State Ion.* **1990**, *37* (2), 223–229.
- [34] Ferguson, D. B.; Krawietz, T. R.; Haw, J. F. Temperature-Jump MAS NMR with a Laser Heater. *J. Magn. Reson. A* **1994**, *109* (2), 273–275.

- [35] Ernst, H.; Freude, D.; Mildner, T.; Wolf, I. Laser-Supported High-Temperature MAS NMR for Time-Resolved in Situ Studies of Reaction Steps in Heterogeneous Catalysis. *Solid State Nucl. Magn. Reson.* **1996**, *6* (2), 147–156.
- [36] Ferguson, D. B.; Haw, J. F. Transient Methods for in Situ NMR of Reactions on Solid Catalysts Using Temperature Jumps. *Anal. Chem.* **1995**, *67* (18), 3342–3348.
- [37] Kawakami, M.; Akasaka, K. Microwave Temperature-Jump Nuclear Magnetic Resonance System for Aqueous Solutions. *Rev. Sci. Instrum.* **1998**, *69* (9), 3365–3369.
- [38] Naito, A.; Nakatani, H.; Imanari, M.; Akasaka, K. State-Correlated Two-Dimensional NMR Spectroscopy. *J. Magn. Reson.* **1990**, *87* (2), 429–432.
- [39] Yushmanov, P. V.; Furó, I. A Temperature-Jump Design for Conventional NMR Probes. *J. Magn. Reson.* **2006**, *181* (1), 148–153.
- [40] Frydman, L.; Gal, M.; Zibzener, K. A Capacitively Coupled Temperature-Jump Arrangement for High-Resolution Biomolecular NMR. *Magn. Reson. Chem.* **2010**, *48* (11), 842–847.
- [41] Andryieuski, A.; Kuznetsova, S. M.; Zhukovsky, S. V.; Kivshar, Y. S.; Lavrinenko, A. V. Water: Promising Opportunities For Tunable All-Dielectric Electromagnetic Metamaterials. *Sci. Rep.* **2015**, *5* (1), 1–9.
- [42] Rinnenthal, J.; Wagner, D.; Marquardsen, T.; Krahn, A.; Engelke, F.; Schwalbe, H. A Temperature-Jump NMR Probe Setup Using Rf Heating Optimized for the Analysis of Temperature-Induced Biomacromolecular Kinetic Processes. *J. Magn. Reson.* **2015**, *251*, 84–93.
- [43] Smeaton, J. R.; Elliott, W. H. Isolation and Properties of a Specific Bacterial Ribonuclease Inhibitor. *Biochim. Biophys. Acta BBA - Nucleic Acids Protein Synth.* **1967**, *145* (3), 547–560.
- [44] Hartley, R. W.; Smeaton, J. R. On the Reaction between the Extracellular Ribonuclease of *Bacillus Amylolyquefaciens* (Barnase) and Its Intracellular Inhibitor (Barstar). *J. Biol. Chem.* **1973**, *248* (16), 5624–5626.
- [45] Hartley, R. W. Barnase and Barstar: Expression of Its Cloned Inhibitor Permits Expression of a Cloned Ribonuclease. *J. Mol. Biol.* **1988**, *202* (4), 913–915.
- [46] Hartley, R. W. Directed Mutagenesis and Barnase-Barstar Recognition. *Biochemistry* **1993**, *32* (23), 5978–5984.
- [47] Matouschek, A.; Kellis, J. T.; Serrano, L.; Fersht, A. R. Mapping the Transition State and Pathway of Protein Folding by Protein Engineering. *Nature* **1989**, *340* (6229), 122–126.
- [48] Fersht, A. R. Protein Folding and Stability: The Pathway of Folding of Barnase. *FEBS Lett.* **1993**, *325* (1–2), 5–16.
- [49] Frisch, C.; Schreiber, G.; Johnson, C. M.; Fersht, A. R. Thermodynamics of the Interaction of Barnase and Barstar: Changes in Free Energy versus Changes in Enthalpy on Mutation. Edited by J. Karn. *J. Mol. Biol.* **1997**, *267* (3), 696–706.
- [50] Buckle, A. M.; Schreiber, G.; Fersht, A. R. Protein-Protein Recognition: Crystal Structural Analysis of a Barnase-Barstar Complex at 2.0-Å Resolution. *Biochemistry* **1994**, *33* (30), 8878–8889.
- [51] Wong, K.-B.; Fersht, A. R.; Freund, S. M. V. NMR <sup>15</sup>N Relaxation and Structural Studies Reveal Slow Conformational Exchange in Barstar C40/82A11. Edited by P. E. Wright. *J. Mol. Biol.* **1997**, *268* (2), 494–511.
- [52] Korchuganov, D. S.; Nolde, S. B.; Reibarkh, M. Ya.; Orekhov, V. Yu.; Schulga, A. A.; Ermolyuk, Y. S.; Kirpichnikov, M. P.; Arseniev, A. S. NMR Study of



- Monomer–Dimer Equilibrium of Barstar in Solution. *J. Am. Chem. Soc.* **2001**, *123* (9), 2068–2069.
- [53] Guillet, V.; Laphorn, A.; Hartley, R. W.; Mauguen, Y. Recognition between a Bacterial Ribonuclease, Barnase, and Its Natural Inhibitor, Barstar. *Structure* **1993**, *1* (3), 165–176.
- [54] Jones, D. N. M.; Bycroft, M.; Lubienski, M. J.; Fersht, A. R. Identification of the Barstar Binding Site of Barnase by NMR Spectroscopy and Hydrogen-Deuterium Exchange. *FEBS Lett.* **1993**, *331* (1), 165–172.
- [55] Zhuravleva, A.; Korzhnev, D. M.; Nolde, S. B.; Kay, L. E.; Arseniev, A. S.; Billeter, M.; Orekhov, V. Yu. Propagation of Dynamic Changes in Barnase Upon Binding of Barstar: An NMR and Computational Study. *J. Mol. Biol.* **2007**, *367* (4), 1079–1092.
- [56] Schreiber, G.; Fersht, A. R. Energetics of Protein-Protein Interactions: Analysis Ofthe Barnase-Barstar Interface by Single Mutations and Double Mutant Cycles. *J. Mol. Biol.* **1995**, *248* (2), 478–486.
- [57] Lee, L.-P.; Tidor, B. Barstar Is Electrostatically Optimized for Tight Binding to Barnase. *Nat. Struct. Mol. Biol.* **2001**, *8* (1), 73–76.
- [58] Schreiber, G.; Fersht, A. R. The Refolding of Cis- and Trans-Peptidylprolyl Isomers of Barstar. *Biochemistry* **1993**, *32* (41), 11195–11203.
- [59] Shastry, M. C. R.; Agashe, V. R.; Udgaonkar, J. B. Quantitative Analysis of the Kinetics of Denaturation and Renaturation of Barstar in the Folding Transition Zone. *Protein Sci.* **1994**, *3* (9), 1409–1417.
- [60] Nath, U.; Udgaonkar, J. B. Folding of Tryptophan Mutants of Barstar: Evidence for an Initial Hydrophobic Collapse on the Folding Pathway. *Biochemistry* **1997**, *36* (28), 8602–8610.
- [61] Shastry, M. C. R.; Udgaonkar, J. B. The Folding Mechanism of Barstar: Evidence for Multiple Pathways and Multiple Intermediates. *J. Mol. Biol.* **1995**, *247* (5), 1013–1027.
- [62] Bhuyan, A. K.; Udgaonkar, J. B. Observation of Multistate Kinetics during the Slow Folding and Unfolding of Barstar. *Biochemistry* **1999**, *38* (28), 9158–9168.
- [63] Killick, T. R.; Freund, S. M. V.; Fersht, A. R. Real-Time NMR Studies on a Transient Folding Intermediate of Barstar. *Protein Sci.* **1999**, *8* (6), 1286–1291.
- [64] Sridevi, K.; Juneja, J.; Bhuyan, A. K.; Krishnamoorthy, G.; Udgaonkar, J. B. The Slow Folding Reaction of Barstar: The Core Tryptophan Region Attains Tight Packing before Substantial Secondary and Tertiary Structure Formation and Final Compaction of the Polypeptide Chain11Edited by C. R. Matthews. *J. Mol. Biol.* **2000**, *302* (2), 479–495.
- [65] Li, H.; Frieden, C. Comparison of C40/82A and P27A C40/82A Barstar Mutants Using 19F NMR. *Biochemistry* **2007**, *46* (14), 4337–4347.
- [66] Privalov, P. L. Cold Denaturation of Protein. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25* (4), 281–306.
- [67] Lopez, C. F.; Darst, R. K.; Rossky, P. J. Mechanistic Elements of Protein Cold Denaturation. *J. Phys. Chem. B* **2008**, *112* (19), 5961–5967.
- [68] Babu, C. R.; Hilser, V. J.; Wand, A. J. Direct Access to the Cooperative Substructure of Proteins and the Protein Ensemble via Cold Denaturation. *Nat. Struct. Mol. Biol.* **2004**, *11* (4), 352–357.
- [69] Zweckstetter, M.; Jaremko, M.; Jaremko, Ł.; Kim, H.-Y.; Cho, M.-K.; Schwieters, C. D.; Giller, K.; Becker, S. Cold Denaturation of a Protein Dimer Monitored at Atomic Resolution. *Nat. Chem. Biol.* **2013**, *9* (4), 264–270.



- [70] Szyperski, T.; Mills, J. L.; Perl, D.; Balbach, J. Combined NMR-Observation of Cold Denaturation in Supercooled Water and Heat Denaturation Enables Accurate Measurement of  $\Delta C_p$  of Protein Unfolding. *Eur. Biophys. J.* **2006**, *35* (4), 363–366.
- [71] Agashe, V. R.; Udgaonkar, J. B. Thermodynamics of Denaturation of Barstar: Evidence for Cold Denaturation and Evaluation of the Interaction with Guanidine Hydrochloride. *Biochemistry* **1995**, *34* (10), 3286–3299.
- [72] Wong, K.-B.; Freund, S. M. V.; Fersht, A. R. Cold Denaturation of Barstar:  $^1H$ ,  $^{15}N$  and  $^{13}C$  NMR Assignment and Characterisation of Residual Structure. *J. Mol. Biol.* **1996**, *259* (4), 805–818.
- [73] Nölting, B.; Golbik, R.; Soler-González, A. S.; Fersht, A. R. Circular Dichroism of Denatured Barstar Suggests Residual Structure. *Biochemistry* **1997**, *36* (32), 9899–9905.
- [74] Privalov, P. L.; Griko, Yu. V.; Venyaminov, S. Yu.; Kutysenko, V. P. Cold Denaturation of Myoglobin. *J. Mol. Biol.* **1986**, *190* (3), 487–498.
- [75] Braun, D.; Wider, G.; Wüthrich, K. Sequence-Corrected  $^{15}N$  “Random Coil” Chemical Shifts. *J. Am. Chem. Soc.* **1994**, *116* (19), 8466–8469.
- [76] Kibbe, W. A. OligoCalc: An Online Oligonucleotide Properties Calculator. *Nucleic Acids Res.* **2007**, *35* (suppl\_2), W43–W46.
- [77] Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins Struct. Funct. Bioinforma.* **2005**, *59* (4), 687–696.
- [78] Lubienski, M. J.; Bycroft, M.; Freund, S. M.; Fersht, A. R. Three-Dimensional Solution Structure and  $^{13}C$  Assignments of Barstar Using Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **1994**, *33* (30), 8866–8877.
- [79] Solyom, Z.; Schwarten, M.; Geist, L.; Konrat, R.; Willbold, D.; Brutscher, B. BEST-TROSY Experiments for Time-Efficient Sequential Resonance Assignment of Large Disordered Proteins. *J. Biomol. NMR* **2013**, *55* (4), 311–321.
- [80] Gadian, D. G.; Robinson, F. N. H. Radiofrequency Losses in NMR Experiments on Electrically Conducting Samples. *J. Magn. Reson.* **1979**, *34* (2), 449–455.
- [81] Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. Rapid Recording of 2D NMR Spectra without Phase Cycling. Application to the Study of Hydrogen Exchange in Proteins. *J. Magn. Reson.* **1989**, *85* (2), 393–399.
- [82] Bermel, W.; Bertini, I.; Chill, J.; Felli, I. C.; Haba, N.; V, V. K. M.; Pierattelli, R. Exclusively Heteronuclear  $^{13}C$ -Detected Amino-Acid-Selective NMR Experiments for the Study of Intrinsically Disordered Proteins (IDPs). *ChemBioChem* **2012**, *13* (16), 2425–2432.
- [83] Golbik, R.; Fischer, G.; Fersht, A. R. Folding of Barstar C40A/C82A/P27A and Catalysis of the Peptidyl-Prolyl Cis/Trans Isomerization by Human Cytosolic Cyclophilin (Cyp18). *Protein Sci.* **1999**, *8* (7), 1505–1514.
- [84] Montelione, G. T.; Wagner, G. 2D Chemical Exchange NMR Spectroscopy by Proton-Detected Heteronuclear Correlation. *J. Am. Chem. Soc.* **1989**, *111* (8), 3096–3098.
- [85] Wider, G.; Neri, D.; Wüthrich, K. Studies of Slow Conformational Equilibria in Macromolecules by Exchange of Heteronuclear Longitudinal 2-Spin-Order in a 2D Difference Correlation Experiment. *J. Biomol. NMR* **1991**, *1* (1), 93–98.

- [86] Shaka, A. J.; Barker, P. B.; Freeman, R. Computer-Optimized Decoupling Scheme for Wideband Applications and Low-Level Operation. *J. Magn. Reson.* **1969**, *1985*, *64* (3), 547–552.
- [87] Bort, G.; Gallavardin, T.; Ogden, D.; Dalko, P. I. From One-Photon to Two-Photon Probes: “Caged” Compounds, Actuators, and Photoswitches. *Angew. Chem. Int. Ed.* **2013**, *17*, 4526–4537.
- [88] Klán, P.; Šolomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J. Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. *Chem. Rev.* **2013**, *113* (1), 119–191.
- [89] Luo, J.; Samanta, S.; Convertino, M.; Dokholyan, N. V.; Deiters, A. Reversible and Tunable Photoswitching of Protein Function through Genetic Encoding of Azobenzene Amino Acids in Mammalian Cells. *ChemBioChem* **2018**, *19* (20), 2178–2185.
- [90] Kathan, M.; Hecht, S. Photoswitchable Molecules as Key Ingredients to Drive Systems Away from the Global Thermodynamic Minimum. *Chem. Soc. Rev.* **2017**, *46* (18), 5536–5550.
- [91] Broichhagen, J.; Schönberger, M.; Cork, S. C.; Frank, J. A.; Marchetti, P.; Bugliani, M.; Shapiro, A. M. J.; Trapp, S.; Rutter, G. A.; Hodson, D. J.; Trauner, D. Optical Control of Insulin Release Using a Photoswitchable Sulfonylurea. *Nat. Commun.* **2014**, *5*, 5116.
- [92] Beharry, A. A.; Woolley, G. A. Azobenzene Photoswitches for Biomolecules. *Chem. Soc. Rev.* **2011**, *40* (8), 4422–4437.
- [93] Szymański, W.; Beierle, J. M.; Kistemaker, H. A. V.; Velema, W. A.; Feringa, B. L. Reversible Photocontrol of Biological Systems by the Incorporation of Molecular Photoswitches. *Chem. Rev.* **2013**, *113* (8), 6114–6178.
- [94] Frank, J. A.; Yushchenko, D. A.; Hodson, D. J.; Lipstein, N.; Nagpal, J.; Rutter, G. A.; Rhee, J.-S.; Gottschalk, A.; Brose, N.; Schultz, C.; Trauner, D. Photoswitchable Diacylglycerols Enable Optical Control of Protein Kinase C. *Nat. Chem. Biol.* **2016**, *12* (9), 755–762.
- [95] Ankenbruck Nicholas; Courtney Taylor; Naro Yuta; Deiters Alexander. Optochemical Control of Biological Processes in Cells and Animals. *Angew. Chem. Int. Ed.* **2017**, *57* (11), 2768–2798.
- [96] Gautier, A.; Gauron, C.; Volovitch, M.; Bensimon, D.; Jullien, L.; Vriza, S. How to Control Proteins with Light in Living Systems. *Nat. Chem. Biol.* **2014**, *10* (7), 533–541.
- [97] Hemphill, J.; Borchardt, E. K.; Brown, K.; Asokan, A.; Deiters, A. Optical Control of CRISPR/Cas9 Gene Editing. *J. Am. Chem. Soc.* **2015**, *137* (17), 5642–5645.
- [98] Baker, A. S.; Deiters, A. Optical Control of Protein Function through Unnatural Amino Acid Mutagenesis and Other Optogenetic Approaches. *ACS Chem. Biol.* **2014**, *9* (7), 1398–1407.
- [99] Ellis-Davies, G. C. R. Nitrophenyl-Based Caged Neurotransmitters. *Cold Spring Harb. Protoc.* **2014**, *2014* (6), pdb.top081794–pdb.top081794.
- [100] Kumar, P.; Shukhman, D.; Laughlin, S. T. A Photocaged, Cyclopropene-Containing Analog of the Amino Acid Neurotransmitter Glutamate. *Tetrahedron Lett.* **2016**, *57* (51), 5750–5752.
- [101] Kantevari, S.; Matsuzaki, M.; Kanemoto, Y.; Kasai, H.; Ellis-Davies, G. C. R. Two-Color, Two-Photon Uncaging of Glutamate and GABA. *Nat. Methods* **2010**, *7* (2), 123–125.

- [102] Heckman, L. M.; Grimm, J. B.; Schreiter, E. R.; Kim, C.; Verdecia, M. A.; Shields, B. C.; Lavis, L. D. Design and Synthesis of a Calcium-Sensitive Photocage. *Angew. Chem. Int. Ed.* **2016**, *55* (29), 8363–8366.
- [103] Ellis-Davies, G. C. R. Neurobiology with Caged Calcium. *Chem. Rev. Wash. DC U. S.* **2008**, *108* (5), 1603–1613.
- [104] Buhr, F.; Kohl-Landgraf, J.; tom Dieck, S.; Hanus, C.; Chatterjee, D.; Hegelein, A.; Schuman, E. M.; Wachtveitl, J.; Schwalbe, H. Design of Photocaged Puromycin for Nascent Polypeptide Release and Spatiotemporal Monitoring of Translation. *Angew. Chem. Int. Ed.* **2015**, n/a-n/a.
- [105] Inlay, M. A.; Choe, V.; Bharathi, S.; Fernhoff, N. B.; Baker, J. R.; Weissman, I. L.; Choi, S. K. Synthesis of a Photocaged Tamoxifen for Light-Dependent Activation of Cre-ER Recombinase-Driven Gene Modification. *Chem. Commun.* **2013**, *49* (43), 4971.
- [106] Binder, D.; Bier, C.; Grünberger, A.; Drobietz, D.; Hage-Hülsmann, J.; Wandrey, G.; Büchs, J.; Kohlheyer, D.; Loeschke, A.; Wiechert, W.; Jaeger, K.-E.; Pietruszka, J.; Drepper, T. Photocaged Arabinose: A Novel Optogenetic Switch for Rapid and Gradual Control of Microbial Gene Expression. *ChemBioChem* **2016**, *17* (4), 296–299.
- [107] Lubbe, A. S.; Szymanski, W.; Feringa, B. L. Recent Developments in Reversible Photoregulation of Oligonucleotide Structure and Function. *Chem Soc Rev* **2017**, *46* (4), 1052–1079.
- [108] Hansen, M. J.; Velema, W. A.; Lerch, M. M.; Szymanski, W.; Feringa, B. L. Wavelength-Selective Cleavage of Photoprotecting Groups: Strategies and Applications in Dynamic Systems. *Chem. Soc. Rev.* **2015**, *44* (11), 3358–3377.
- [109] Ji, Y.; Yang, J.; Wu, L.; Yu, L.; Tang, X. Photochemical Regulation of Gene Expression Using Caged SiRNAs with Single Terminal Vitamin E Modification. *Angew. Chem. Int. Ed.* **2016**, *55* (6), 2152–2156.
- [110] Jain, P. K.; Ramanan, V.; Schepers, A. G.; Dalvie, N. S.; Panda, A.; Fleming, H. E.; Bhatia, S. N. Development of Light-Activated CRISPR Using Guide RNAs with Photocleavable Protectors. *Angew. Chem. Int. Ed.* **2016**, *55* (40), 12440–12444.
- [111] Buck, J.; Fürtig, B.; Noeske, J.; Wöhnert, J.; Schwalbe, H. Time-Resolved NMR Methods Resolving Ligand-Induced RNA Folding at Atomic Resolution. *Proc. Natl. Acad. Sci.* **2007**, *104* (40), 15699–15704.
- [112] Helmling, C.; Klötzner, D.-P.; Sochor, F.; Mooney, R. A.; Wacker, A.; Landick, R.; Fürtig, B.; Heckel, A.; Schwalbe, H. Life Times of Metastable States Guide Regulatory Signaling in Transcriptional Riboswitches. *Nat. Commun.* **2018**, *9* (1).
- [113] Grün, J. T.; Hennecker, C.; Klötzner, D.-P.; Harkness, R. W.; Bessi, I.; Heckel, A.; Mittermaier, A. K.; Schwalbe, H. Conformational Dynamics of Strand Register Shifts in DNA G-Quadruplexes. *J. Am. Chem. Soc.* **2020**, *142* (1), 264–273.
- [114] Liu, Q.; Deiters, A. Optochemical Control of Deoxyoligonucleotide Function via a Nucleobase-Caging Approach. *Acc. Chem. Res.* **2014**, *47* (1), 45–55.
- [115] Rinnenthal, J.; Klinkert, B.; Narberhaus, F.; Schwalbe, H. Modulation of the Stability of the Salmonella FourU-Type RNA Thermometer. *Nucleic Acids Res.* **2011**, *39* (18), 8258–8270.
- [116] Wagner, D.; Rinnenthal, J.; Narberhaus, F.; Schwalbe, H. Mechanistic Insights into Temperature-Dependent Regulation of the Simple Cyanobacterial Hsp17

- RNA Thermometer at Base-Pair Resolution. *Nucleic Acids Res.* **2015**, *43* (11), 5572–5585.
- [117] Rinnenthal, J.; Klinkert, B.; Narberhaus, F.; Schwalbe, H. Direct Observation of the Temperature-Induced Melting Process of the Salmonella FourU RNA Thermometer at Base-Pair Resolution. *Nucleic Acids Res.* **2010**, *38* (11), 3834–3847.
- [118] Noeske, J.; Buck, J.; Fürtig, B.; Nasiri, H. R.; Schwalbe, H.; Wöhnert, J. Interplay of ‘Induced Fit’ and Preorganization in the Ligand Induced Folding of the Aptamer Domain of the Guanine Binding Riboswitch. *Nucleic Acids Res.* **2007**, *35* (2), 572–583.
- [119] Fürtig, B.; Nozinovic, S.; Reining, A.; Schwalbe, H. Multiple Conformational States of Riboswitches Fine-Tune Gene Regulation. *Curr. Opin. Struct. Biol.* **2015**, *30*, 112–124.
- [120] Fürtig, B.; Buck, J.; Richter, C.; Schwalbe, H. Functional Dynamics of RNA Ribozymes Studied by NMR Spectroscopy. In *Ribozymes: Methods and Protocols*; Hartig, J. S., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; pp 185–199.
- [121] Hoogstraten, C. G.; Legault, P.; Pardi, A. NMR Solution Structure of the Lead-Dependent Ribozyme: Evidence for Dynamics in RNA Catalysis I Edited by I. Tinoco. *J. Mol. Biol.* **1998**, *284* (2), 337–350.
- [122] Mergny, J.-L.; Lacroix, L. Analysis of Thermal Melting Curves. *Oligonucleotides* **2003**, *13* (6), 515–537.
- [123] Spink, C. H. Differential Scanning Calorimetry. In *Methods in Cell Biology*; Elsevier, 2008; Vol. 84, pp 115–141.
- [124] Russu, I. M. Probing Site-Specific Energetics in Proteins and Nucleic Acids by Hydrogen Exchange and Nuclear Magnetic Resonance Spectroscopy. In *Methods in Enzymology*; Energetics of Biological Macromolecules, Part D; Academic Press, 2004; Vol. 379, pp 152–175.
- [125] Leroy, J. L.; Kochoyan, M.; Huynh-Dinh, T.; Guéron, M. Characterization of Base-Pair Opening in Deoxynucleotide Duplexes Using Catalyzed Exchange of the Imino Proton. *J. Mol. Biol.* **1988**, *200* (2), 223–238.
- [126] Guéron, M.; Leroy, J.-L. [16] Studies of Base Pair Kinetics by NMR Measurement of Proton Exchange. In *Methods in Enzymology*; Nuclear Magnetic Resonance and Nucleic Acids; Academic Press, 1995; Vol. 261, pp 383–413.
- [127] Steinert, H. S.; Rinnenthal, J.; Schwalbe, H. Individual Basepair Stability of DNA and RNA Studied by NMR-Detected Solvent Exchange. *Biophys. J.* **2012**, *102* (11), 2564–2574.
- [128] Huang, Y.; Weng, X.; Russu, I. M. Enhanced Base-Pair Opening in the Adenine Tract of a RNA Double Helix. *Biochemistry* **2011**, *50* (11), 1857–1863.
- [129] Harada, N. Determination of Absolute Configurations by X-Ray Crystallography and <sup>1</sup>H NMR Anisotropy. *Chirality* **2008**, *20* (5), 691–723.
- [130] Harada, N. Chiral Auxiliaries Powerful for Both Enantiomer Resolution and Determination of Absolute Configuration by X-Ray Crystallography. In *Topics in Stereochemistry*; John Wiley & Sons, Ltd, 2006; pp 177–203.
- [131] Freedman, T. B.; Cao, X.; Dukor, R. K.; Nafie, L. A. Absolute Configuration Determination of Chiral Molecules in the Solution State Using Vibrational Circular Dichroism. *Chirality* **2003**, *15* (9), 743–758.
- [132] Seco, J. M.; Quiñoá, E.; Riguera, R. The Assignment of Absolute Configuration by NMR. *Chem. Rev.* **2004**, *104* (1), 17–118.

- [133] Hoye, T. R.; Jeffrey, C. S.; Shao, F. Mosher Ester Analysis for the Determination of Absolute Configuration of Stereogenic (Chiral) Carbinol Carbons. *Nat. Protoc.* **2007**, *2* (10), 2451–2458.
- [134] Wenzel, T. J.; Chisholm, C. D. Using NMR Spectroscopic Methods to Determine Enantiomeric Purity and Assign Absolute Stereochemistry. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59* (1), 1–63.
- [135] Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. Determination of Relative Configuration in Organic Compounds by NMR Spectroscopy and Computational Methods. *Chem. Rev.* **2007**, *107* (9), 3744–3779.
- [136] Seyfried, P.; Heinz, M.; Pintér, G.; Klötzner, D.-P.; Becker, Y.; Bolte, M.; Jonker, H. R. A.; Stelzl, L. S.; Hummer, G.; Schwalbe, H.; Heckel, A. Optimal Destabilization of DNA Double Strands by Single-Nucleobase Caging. *Chem. – Eur. J.* **2018**, *24* (66), 17568–17576.
- [137] Steinert, H. S.; Schäfer, F.; Jonker, H. R. A.; Heckel, A.; Schwalbe, H. Influence of the Absolute Configuration of NPE-Caged Cytosine on DNA Single Base Pair Stability. *Angew. Chem. Int. Ed.* **2014**, *53* (4), 1072–1075.
- [138] Heus, H. A.; Pardi, A. Novel Proton NMR Assignment Procedure for RNA Duplexes. *J. Am. Chem. Soc.* **1991**, *113* (11), 4360–4361.
- [139] Liang, X.; Asanuma, H.; Kashida, H.; Takasu, A.; Sakamoto, T.; Kawai, G.; Komiyama, M. NMR Study on the Photoresponsive DNA Tethering an Azobenzene. Assignment of the Absolute Configuration of Two Diastereomers and Structure Determination of Their Duplexes in the *Trans* -Form. *J. Am. Chem. Soc.* **2003**, *125* (52), 16408–16415.
- [140] Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. D Biol. Crystallogr.* **1998**, *54* (5), 905–921.
- [141] Kumar, R.; Grubmüller, H. Do\_x3dna: A Tool to Analyze Structural Fluctuations of DsDNA or DsRNA from Molecular Dynamics Simulations. *Bioinformatics* **2015**, *31* (15), 2583–2585.
- [142] Linge, J. P.; Williams, M. A.; Spronk, C. A. E. M.; Bonvin, A. M. J. J.; Nilges, M. Refinement of Protein Structures in Explicit Solvent. *Proteins Struct. Funct. Bioinforma.* **2003**, *50* (3), 496–506.
- [143] Nilges, M.; Macias, M. J.; O’Donoghue, S. I.; Oschkinat, H. Automated NOESY Interpretation with Ambiguous Distance Restraints: The Refined NMR Solution Structure of the Pleckstrin Homology Domain from  $\beta$ -Spectrin1 Edited by P. E. Wright. *J. Mol. Biol.* **1997**, *269* (3), 408–422.
- [144] Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. The Three Roles of RNA in Protein Synthesis. *Mol. Cell Biol. 4th Ed.* **2000**.
- [145] Yao, R.-W.; Wang, Y.; Chen, L.-L. Cellular Functions of Long Noncoding RNAs. *Nat. Cell Biol.* **2019**, *21* (5), 542–551.
- [146] Gebert, L. F. R.; MacRae, I. J. Regulation of MicroRNA Function in Animals. *Nat. Rev. Mol. Cell Biol.* **2019**, *20* (1), 21–37.
- [147] Serganov, A.; Nudler, E. A Decade of Riboswitches. *Cell* **2013**, *152* (1), 17–24.
- [148] Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71–73.
- [149] Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and



- Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.
- [150] Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science* **2011**, *334* (6055), 517–520.
- [151] Vangaveti, S.; Ranganathan, S. V.; Chen, A. A. Advances in RNA Molecular Dynamics: A Simulator’s Guide to RNA Force Fields. *WIREs RNA* **2017**, *8* (2), e1396.
- [152] Šponer, J.; Bussi, G.; Krepl, M.; Banáš, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurečka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118* (8), 4177–4338.
- [153] Condon, D. E.; Kennedy, S. D.; Mort, B. C.; Kierzek, R.; Yildirim, I.; Turner, D. H. Stacking in RNA: NMR of Four Tetramers Benchmark Molecular Dynamics. *J. Chem. Theory Comput.* **2015**, *11* (6), 2729–2742.
- [154] Tubbs, J. D.; Condon, D. E.; Kennedy, S. D.; Hauser, M.; Bevilacqua, P. C.; Turner, D. H. The Nuclear Magnetic Resonance of CCCC RNA Reveals a Right-Handed Helix, and Revised Parameters for AMBER Force Field Torsions Improve Structural Predictions from Molecular Dynamics. *Biochemistry* **2013**, *52* (6), 996–1010.
- [155] Yildirim, I.; Stern, H. A.; Tubbs, J. D.; Kennedy, S. D.; Turner, D. H. Benchmarking AMBER Force Fields for RNA: Comparisons to NMR Spectra for Single-Stranded r(GACC) Are Improved by Revised  $\chi$  Torsions. *J. Phys. Chem. B* **2011**, *115* (29), 9261–9270.
- [156] Nozinovic, S.; Fürtig, B.; Jonker, H. R. A.; Richter, C.; Schwalbe, H. High-Resolution NMR Structure of an RNA Model System: The 14-Mer CUUCG Tetraloop Hairpin RNA. *Nucleic Acids Res.* **2010**, *38* (2), 683–694.
- [157] Banáš, P.; Hollas, D.; Zgarbová, M.; Jurečka, P.; Orozco, M.; Cheatham, T. E.; Šponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6* (12), 3836–3849.
- [158] Kührová, P.; Best, R. B.; Bottaro, S.; Bussi, G.; Šponer, J.; Otyepka, M.; Banáš, P. Computer Folding of RNA Tetraloops: Identification of Key Force Field Deficiencies. *J. Chem. Theory Comput.* **2016**, *12* (9), 4534–4548.
- [159] Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Cheatham, T. E. Highly Sampled Tetranucleotide and Tetraloop Motifs Enable Evaluation of Common RNA Force Fields. *RNA* **2015**, *21* (9), 1578–1590.
- [160] Jucker, F. M.; Heus, H. A.; Yip, P. F.; Moors, E. H. M.; Pardi, A. A Network of Heterogeneous Hydrogen Bonds in GNRA Tetraloops. *J. Mol. Biol.* **1996**, *264* (5), 968–980.
- [161] Heus, H.; Pardi, A. Structural Features That Give Rise to the Unusual Stability of RNA Hairpins Containing GNRA Loops. *Science* **1991**, *253* (5016), 191–194.
- [162] Wolters, J. The Nature of Preferred Hairpin Structures in 16S-like RRNA Variable Regions. *Nucleic Acids Res.* **1992**, *20* (8), 1843–1850.
- [163] Andronescu, M.; Bereg, V.; Hoos, H. H.; Condon, A. RNA STRAND: The RNA Secondary Structure and Statistical Analysis Database. *BMC Bioinformatics* **2008**, *9* (1), 340.
- [164] Bottaro, S.; Lindorff-Larsen, K. Mapping the Universe of RNA Tetraloop Folds. *Biophys. J.* **2017**, *113* (2), 257–267.
- [165] Rinnenthal, J.; Richter, C.; Ferner, J.; Duchardt, E.; Schwalbe, H. Quantitative  $\Gamma$ -HCNCH: Determination of the Glycosidic Torsion Angle  $\chi$  in RNA

- Oligonucleotides from the Analysis of CH Dipolar Cross-Correlated Relaxation by Solution NMR Spectroscopy. *J. Biomol. NMR* **2007**, *39* (1), 17–29.
- [166] Richter, C.; Griesinger, C.; Felli, I.; Cole, P. T.; Varani, G.; Schwalbe, H. Determination of Sugar Conformation in Large RNA Oligonucleotides from Analysis of Dipole–Dipole Cross Correlated Relaxation by Solution NMR Spectroscopy. *J. Biomol. NMR* **1999**, *15* (3), 241–250.
- [167] Nozinovic, S.; Richter, C.; Rinnenthal, J.; Fürtig, B.; Duchardt-Ferner, E.; Weigand, J. E.; Schwalbe, H. Quantitative 2D and 3D  $\Gamma$ -HCP Experiments for the Determination of the Angles  $\alpha$  and  $\zeta$  in the Phosphodiester Backbone of Oligonucleotides. *J. Am. Chem. Soc.* **2010**, *132* (30), 10318–10329.
- [168] Schwalbe, H.; Samstag, W.; Engels, J. W.; Bermel, W.; Griesinger, C. Determination of  $3J(C,P)$  and  $3J(H,P)$  Coupling Constants in Nucleotide Oligomers with FIDS-HSQC. *J. Biomol. NMR* **1993**, *3*, 479–486.
- [169] Marino, J. P.; Schwalbe, H.; Glaser, S. J.; Griesinger, C. Determination of  $\gamma$  and Stereospecific Assignment of H5' Protons by Measurement of  $^2J$  and  $^3J$  Coupling Constants in Uniformly  $^{13}C$  Labeled RNA. *J. Am. Chem. Soc.* **1996**, *118* (18), 4388–4395.
- [170] Nichols, P. J.; Henen, M. A.; Born, A.; Strotz, D.; Güntert, P.; Vögeli, B. High-Resolution Small RNA Structures from Exact Nuclear Overhauser Enhancement Measurements without Additional Restraints. *Commun. Biol.* **2018**, *1* (1), 1–11.
- [171] Bottaro, S.; Nichols, P. J.; Vögeli, B.; Parrinello, M.; Lindorff-Larsen, K. Integrating NMR and Simulations Reveals Motions in the UUCG Tetraloop. *Nucleic Acids Res.* **2020**, *48* (11), 5839–5848.
- [172] Nivón, L. G.; Shakhnovich, E. I. All-Atom Monte Carlo Simulation of GCAA RNA Folding. *J. Mol. Biol.* **2004**, *344* (1), 29–45.
- [173] Zhang, Y.; Zhao, X.; Mu, Y. Conformational Transition Map of an RNA GCAA Tetraloop Explored by Replica-Exchange Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2009**, *5* (4), 1146–1154.
- [174] Johnson, J. E.; Hoogstraten, C. G. Extensive Backbone Dynamics in the GCAA RNA Tetraloop Analyzed Using  $^{13}C$  NMR Spin Relaxation and Specific Isotope Labeling. *J. Am. Chem. Soc.* **2008**, *130* (49), 16757–16769.
- [175] Menger, M.; Eckstein, F.; Porschke, D. Dynamics of the RNA Hairpin GNRA Tetraloop. *Biochemistry* **2000**, *39* (15), 4500–4507.
- [176] Zhao, L.; Xia, T. Direct Revelation of Multiple Conformations in RNA by Femtosecond Dynamics. *J. Am. Chem. Soc.* **2007**, *129* (14), 4118–4119.
- [177] Schürer, H.; Lang, K.; Schuster, J.; Mörl, M. A Universal Method to Produce in Vitro Transcripts with Homogeneous 3' Ends. *Nucleic Acids Res.* **2002**, *30* (12), e56.
- [178] Milligan, J. F.; Uhlenbeck, O. C. [5] Synthesis of Small RNAs Using T7 RNA Polymerase. In *Methods in Enzymology*; RNA Processing Part A: General Methods; Academic Press, 1989; Vol. 180, pp 51–62.
- [179] Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D.  $^1H$ ,  $^{13}C$  and  $^{15}N$  Chemical Shift Referencing in Biomolecular NMR. *J. Biomol. NMR* **1995**, *6* (2), 135–140.
- [180] Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics* **2015**, *31* (8), 1325–1327.
- [181] Liebecq, C. *Biochemical Nomenclature and Related Documents*; Portland Press, 1992.
- [182] Sklenář, V.; Peterson, R. D.; Rejante, M. R.; Feigon, J. Two- and Three-Dimensional HCN Experiments for Correlating Base and Sugar Resonances in

- 15 N, 13 C-Labeled RNA Oligonucleotides. *J. Biomol. NMR* **1993**, 3 (6), 721–727.
- [183] Schwalbe, H.; Marino, J. P.; Glaser, S. J.; Griesinger, C. Measurement of H,H-Coupling Constants Associated with .Nu.1, .Nu.2, and .Nu.3 in Uniformly <sup>13</sup>C-Labeled RNA by HCC-TOCSY-CCH-E.COSY. *J. Am. Chem. Soc.* **1995**, 117 (27), 7251–7252.
- [184] Marino, J. P.; Schwalbe, H.; Anklin, C.; Bermel, W.; Crothers, D. M.; Griesinger, C. Sequential Correlation of Anomeric Ribose Protons and Intervening Phosphorus in RNA Oligonucleotides by a <sup>1</sup>H,<sup>13</sup>C,<sup>31</sup>P Triple Resonance Experiment: HCP-CCH-TOCSY. *J. Biomol. NMR* **1995**, 5 (1), 87–92.
- [185] Marino, J. P.; Schwalbe, H.; Anklin, C.; Bermel, W.; Crothers, D. M.; Griesinger, C. Three-Dimensional Triple-Resonance <sup>1</sup>H, <sup>13</sup>C, <sup>31</sup>P Experiment: Sequential Through-Bond Correlation of Ribose Protons and Intervening Phosphorus along the RNA Oligonucleotide Backbone. *J. Am. Chem. Soc.* **1994**, 116 (14), 6472–6473.
- [186] Cherepanov, A. V.; Glaubitz, C.; Schwalbe, H. High-Resolution Studies of Uniformly <sup>13</sup>C,<sup>15</sup>N-Labeled RNA by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2010**, 49 (28), 4747–4750.
- [187] Ebrahimi, M.; Rossi, P.; Rogers, C.; Harbison, G. S. Dependence of <sup>13</sup>C NMR Chemical Shifts on Conformations of RNA Nucleosides and Nucleotides. *J. Magn. Reson.* **2001**, 150 (1), 1–9.
- [188] Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. New Description Using the Concept of Pseudorotation. *J. Am. Chem. Soc.* **1972**, 94 (23), 8205–8212.
- [189] Li, L.; Szostak, J. W. The Free Energy Landscape of Pseudorotation in 3'–5' and 2'–5' Linked Nucleic Acids. *J. Am. Chem. Soc.* **2014**, 136 (7), 2858–2865.
- [190] Jucker, F. M.; Pardi, A. Solution Structure of the CUUG Hairpin Loop: A Novel RNA Tetraloop Motif. *Biochemistry* **1995**, 34 (44), 14416–14427.
- [191] Richter, C.; Reif, B.; Wörner, K.; Quant, S.; Marino, J. P.; Engels, W.; Griesinger, C.; Schwalbe, H. A New Experiment for the Measurement of  $n J(C, P)$  Coupling Constants. *J. Biomol. NMR* **1998**, 2, 223–230.
- [192] Richter, C.; Reif, B.; Wörner, K.; Quant, S.; Marino, J. P.; Engels, J. W.; Griesinger, C.; Schwalbe, H. A New Experiment for the Measurement of  $NJ(C,P)$  Coupling Constants Including  $3J(C4'i,Pi)$  and  $3J(C4'i,Pi+1)$  in Oligonucleotides. *J. Biomol. NMR* **1998**, 12 (2), 223–230.
- [193] Fürtig, B.; Reining, A.; Sochor, F.; Oberhauser, E. M.; Heckel, A.; Schwalbe, H. Characterization of Conformational Dynamics of Bistable RNA by Equilibrium and Non-Equilibrium NMR: Conformational Dynamics of Bistable RNA by NMR. *Curr. Protoc. Nucleic Acid Chem.* **2013**, 55 (1), 11.13.1–11.13.16.
- [194] Xue, Y.; Kellogg, D.; Kimsey, I. J.; Sathyamoorthy, B.; Stein, Z. W.; McBairty, M.; Al-Hashimi, H. M. Characterizing RNA Excited States Using NMR Relaxation Dispersion. In *Methods in Enzymology*; Elsevier, 2015; Vol. 558, pp 39–73.
- [195] Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **1982**, 104 (17), 4546–4559.
- [196] Schroeder, K. T.; Skalicky, J. J.; Greenbaum, N. L. NMR Spectroscopy of RNA Duplexes Containing Pseudouridine in Supercooled Water. *RNA* **2005**, 11 (7), 1012–1016.



- [197] Skalicky, J. J.; Mills, J. L.; Sharma, S.; Szyperski, T. Aromatic Ring-Flipping in Supercooled Water: Implications for NMR-Based Structural Biology of Proteins. *J. Am. Chem. Soc.* **2001**, *123* (3), 388–397.
- [198] Skalicky, J. J.; Sukumaran, D. K.; Mills, J. L.; Szyperski, T. Toward Structural Biology in Supercooled Water. *J. Am. Chem. Soc.* **2000**, *122* (13), 3230–3231.



## ACKNOWLEDGEMENT

First of all, I would like to thank **Prof. Dr. Harald Schwalbe** – my supervisor – for the opportunity to join His group as part of the first cohort in the CliC graduate program. Thank You very much, Harald, for Your support and guidance all along this journey and the freedom to develop my projects and fulfil my goals in NMR research. I would also like to thank You for the opportunities to visit numerous conferences and other occasions to present my work.

It was a great time in the Horizon room with **Linda Schulte, Gerd Hanspach, Sven Trucks, Andreas Schmidt, Vijay Kumar Nimbarte**. Besides work we also had some great time together – ranging from funny – however sometimes distracting – chats to serious political discussions! Special thanks to **Linda** for the scientific discussions and the shared lunches in the kitchen. You always encouraged me to speak more in German.

Many thanks to **Dr. Christian Richter** for his guidance and technical help in all kinds of NMR spectroscopic topics. Christian, You guided me through the pulse programming mysteries in Topspin and I have learnt a lot about the electronics behind the magic of NMR spectroscopy. I am deeply grateful for Your help!

I would like to thank **Dr. Nina Kubatova** and **Robbin Schnieders** being the best company one can wish while visiting conferences. It was a great time visiting Warsaw together and repeating it in Berlin. I will always remember our daily gossiping during lunchtime in the Minerva Bistro, Nina!

I would like to thank **Manfred Strupf**, the best technician, who is always willing to have a look at broken equipment. To put it simply, the spectrometers would not run without you. Thank you for helping me out with Your technical expertise.

I am also grateful to **Prof. Dr. Alexander Heckel, Dr. Dean-Paulos Klötzner, Marcel Heinz** and **Dr. Patrick Seyfried** for sharing their knowledge about DNA caging and the great collaboration in our common project.

Thank you, **Kerstin Dathe**, for always being there to cheer up every one of us! Thank you for your patience with all the organisational issues – which can be especially challenging for a foreign student – and for the small surprises for Christmas holidays, making a festive and cheerful atmosphere.

It was a pleasure to share “The tetraloop” project with **Robbin Schnieders** and **Andreas Oxenfarth** and learning together the tricks of *in vitro* transcription as well as sharing the seemingly never ending problems of cross-correlated relaxation rate measurements.

Special thanks to all the “**CLiCies**” with whom we spent several summer schools and lunches together discussing problems of light sensitive systems. CliC provides a great international and interdisciplinary environment that promotes great research. Special thanks to **Dr. Christian Grünewald** for taking care of the initiative.

Thank you, **Dr. Boris Fürtig**, for being part of my mentoring committee, for your continuous support and for helping me – also outside the mentoring meetings – find the answer for some challenging scientific questions.

Thank you, **Gerd Hanspach**, for taking care of the skiing skills of the group! I always tried to keep up at least with the beginners of the group, however this was sometimes really challenging. **Islam Alshalem**, You would definitely deserve a speeding ticket, crazy guy!

I would like to thank the **whole AK Schwalbe** for all the support, especially in helping me and showing me around in the wet lab! The great atmosphere in the labs, all the fun during group trip, the Christmas-parties and the barbecues we spent together helped me a lot to feel me home in Germany. Simply the best AK to do a PhD in!

I am deeply grateful for **Dr. Julian de Mos**, we started together in CliC and you became my best friend and ally in my PhD study. You were always there for a coffee break, to help with German bureaucracy, etc. We shared lot of stimulating discussions about NMR spectroscopy and you took the time to explain biochemistry to a “simple” chemist.

Külön köszönet barátomnak, **Lacinak**, aki velem tartott Németországba és mindig türelemmel és odaadással támogatott. Köszönöm, hogy mindig bátorítottál és elviselted, ha frusztrált voltam.

Végül köszönöm a családomnak, **szüleimnek** és **testvéremnek**, hogy mindig bátorítottak és támogattak céljaim elérésében. Köszönöm, hogy mindig számíthatok Rátok!

## APPENDIX

## 1. CHEMICAL SHIFT ASSIGNMENT

## 1.1. BARSTAR AT 303 K FOLDED AND AT 270 K UNFOLDED STATE

Table 9 Chemical shift list of C82A single mutant barstar and C40A/C82A double mutant barstar in 10 mM  $K_2H_2PO_4$ , 115 mM KCl, 2.2 M urea, pH 6.8 buffer with 100  $\mu$ M DSS added as internal reference at 303 K and 270 K, respectively.  $^1H$  chemical shifts have been referenced to the methyl signal of DSS and  $^{13}C$  and  $^{15}N$  chemical shifts have been indirectly referenced according to Wishart et al.<sup>[179]</sup>

| Residue | Atom 303 K /<br>$\delta$ - chemical shift (ppm) of<br>C82A mutant |       |       |       | Atom 270 K /<br>$\delta$ - chemical shift (ppm) of<br>C40A/C82A mutant |        |        |       |       |
|---------|---|-------|-------|-------|--|--------|--------|-------|-------|
|         | H   | N     | Ca    | Cb    | H  | N      | CO     | Ca    | Cb    |
| Lys 1   | -   | -     | 55.49 | 34.3  | 8.53   | 125.65 | 173.17 | -     | -     |
| Lys 2   | 8.615   | 125.7 | 54.95 | 34.49 | 8.32   | 125.15 | 173.24 | 53.37 | 30.3  |
| Ala 3   | 9.043   | 129.5 | 50.18 | 22.16 | 8.25   | 127.34 | 174.57 | 49.37 | 16.24 |
| Val 4   | 8.878   | 123.6 | 60.98 | 34.39 | 8.07   | 121.71 | 173.2  | 59.29 | 29.91 |
| Ile 5   | 8.831   | 125.3 | 59.03 | 39.58 | 8.17   | 126.61 | 172.95 | 57.71 | 35.67 |
| Asn 6   | 8.969   | 126.3 | 52.62 | 37.64 | 8.45   | 124.53 | 172.73 | -     | -     |
| Gly 7   | 8.482   | 113.8 | 48.23 | -     | 8.21   | 110.18 | 171.14 | -     | -     |
| Glu 8   | 9.772   | 120.5 | 57.61 | 28.61 | 7.99   | 120.9  | 173.59 | 53.57 | 27.29 |
| Gln 9   | 7.864   | 117   | 55.14 | 28.93 | 8.24   | 122.41 | 173.01 | 52.84 | 26.34 |
| Ile 10  | 7.023   | 119.8 | 59.11 | 36.84 | 8.09   | 124.09 | 173.42 | 58.25 | 35.36 |
| Arg 11  | 9.248   | 123.6 | 55.86 | 30.93 | 8.31   | 126.92 | 173.26 | 52.99 | 28.03 |
| Ser 12  | 7.679   | 109.6 | 57.01 | 65.96 | 8.27   | 118.54 | 171.88 | 55.13 | 60.93 |
| Ile 13  | 9.474   | 122.3 | 62.46 | 38.13 | 8.13   | 123.42 | 173.64 | -     | -     |
| Ser 14  | 8.175   | 118.3 | 62.46 | -     | 8.05   | 119.34 | 171.47 | 55.72 | 60.73 |
| Asp 15  | 7.982   | 121.6 | 57.4  | 41.88 | 8.01   | 123.18 | 173.48 | -     | -     |
| Leu 16  | 7.775   | 121.7 | 58.36 | 41.12 | 7.87   | 122.62 | 174.6  | 52.63 | 38.99 |
| His 17  | 8.176   | 118.1 | 62.03 | 31.77 | 8.21   | 119.29 | 171.85 | 52.84 | 26.02 |
| Gln 18  | 8.618   | 119.1 | 59.46 | 28.1  | 8.08   | 121.65 | 173.14 | 52.92 | 26.65 |
| Thr 19  | 8.274   | 118.7 | 67.49 | 68.41 | 8.12   | 117.21 | 171.52 | 59.27 | 66.79 |
| Leu 20  | 8.447   | 121.3 | 58.46 | 41.69 | 8.04   | 126.56 | 174.25 | 52.13 | 39.29 |
| Lys 21  | 8.223   | 118.3 | 59.67 | 33.19 | 8.19   | 123.94 | 173.58 | 53.39 | 30.04 |
| Lys 22  | 7.416   | 117   | 58.96 | 33.22 | 8.21   | 124.47 | 173.61 | 53.49 | 30.11 |
| Glu 23  | 9.019   | 118   | 58.88 | 30.71 | 8.32   | 123.26 | 173.36 | 53.48 | 27.17 |
| Leu 24  | 8.242   | 113.6 | 54.31 | 40.85 | 8.07   | 123.94 | 173.93 | 51.91 | 39.44 |
| Ala 25  | 7.438   | 122   | 52.84 | 16.57 | 8.05   | 125.68 | 174.38 | 49.11 | 16.21 |
| Leu 26  | 7.975   | 116.4 | 53.63 | 37.88 | 8.02   | 123.46 | 172.33 | 50.01 | 38.41 |
| Pro 27  | -   | -     | 62.22 | 31.47 | -  | 136.1  | 173.65 | -     | -     |
| Glu 28  | 8.62  | 119.5 | 59.03 | 29.53 | 8.2  | 121.34 | 172.85 | 53.69 | 27.37 |
| Tyr 29  | 6.401   | 110.2 | 55.35 | 36.5  | 7.88   | 121.29 | 172.25 | 54.67 | 36.12 |
| Tyr 30  | 7.373   | 122.4 | 57.6  | 38.92 | 7.95   | 123.63 | 172.9  | 54.89 | 36.01 |
| Gly 31  | 8.19  | 115   | 46.03 | -     | 7.26   | 110.04 | 170.61 | 42.01 | -     |
| Glu 32  | -   | -     | 56.65 | 26.3  | 8.01   | 120.29 | 173.19 | 53.6  | 27.48 |
| Asn 33  | 6.408   | 109.5 | 51.92 | 39.87 | 8.32   | 119.99 | 172.46 | 50.16 | 35.45 |
| Leu 34  | 8.673   | 118.8 | 58.72 | 40.97 | 8.11   | 123.79 | 174.48 | 52.63 | 38.89 |
| Asp 35  | 8.204   | 121.3 | 57.69 | 40.36 | 8  | 120.5  | 173.24 | 51.86 | 37.95 |
| Ala 36  | 8.468   | 123.1 | 54.38 | 19.09 | 7.73   | 123.76 | 174.79 | 49.85 | 16.28 |
| Leu 37  | 8.375   | 120.2 | 58.65 | 40.58 | 7.89   | 120.98 | 174.16 | 52.19 | 39.4  |
| Trp 38  | 8.428   | 120.7 | 60.95 | 29.3  | 7.85   | 122.36 | -      | 54.64 | 26.53 |
| Asp 39  | 8.429   | 119.6 | 57.67 | 40.69 | -  | -      | 172.76 | 50.87 | 38.97 |

Appendix

|               |       |       |       |       |      |        |        |       |       |
|---------------|-------|-------|-------|-------|------|--------|--------|-------|-------|
| <b>Ala 40</b> | 8.322 | 117.1 | 63.51 | 27.46 | 7.79 | 124.99 | 175.42 | 50.6  | 15.94 |
| <b>Leu 41</b> | 8.757 | 120.8 | 57.12 | 40.48 | 7.93 | 118.98 | 175.34 | -     | -     |
| <b>Thr 42</b> | 7.836 | 110.5 | 62.89 | 69.88 | 7.6  | 112.51 | 172.31 | 59.3  | 66.86 |
| <b>Gly 43</b> | 7.606 | 108.7 | 46.49 | -     | 7.95 | 112.01 | 170.67 | 42.13 | -     |
| <b>Trp 44</b> | 8.365 | 121.2 | 60.24 | 32.23 | 7.87 | 122.57 | -      | -     | -     |
| <b>Val 45</b> | 7.942 | 118   | 63.25 | 33.34 | -    | -      | 171.83 | 58.58 | -     |
| <b>Glu 46</b> | 7.749 | 124.7 | 55.36 | 31.16 | 7.9  | 125.13 | -      | 53.42 | 27.54 |
| <b>Tyr 47</b> | 7.914 | 121.9 | 56.55 | 39.19 | 7.98 | 122.24 | 170.76 | 52.96 | 35.14 |
| <b>Pro 48</b> | -     | -     | 62.55 | 35.83 | -    | 136.73 | -      | -     | -     |
| <b>Leu 49</b> | 8.72  | 122.8 | 53.8  | 47.36 | 8.16 | 122.56 | -      | 53.51 | 27.33 |
| <b>Val 50</b> | 8.991 | 128.6 | 61.62 | 33.76 | -    | -      | -      | -     | -     |
| <b>Leu 51</b> | 9.118 | 130   | 52.83 | 44.71 | -    | -      | -      | -     | -     |
| <b>Glu 52</b> | 9.083 | 128.6 | 54.45 | 31.37 | -    | -      | -      | -     | -     |
| <b>Trp 53</b> | -     | -     | -     | -     | 8.02 | 123.69 | -      | 54.71 | 26.45 |
| <b>Arg 54</b> | -     | -     | -     | -     | 7.78 | 124.18 | 172.69 | 53.02 | 28    |
| <b>Gln 55</b> | -     | -     | -     | -     | 7.9  | 121.57 | 172.93 | -     | -     |
| <b>Phe 56</b> | -     | -     | 62.19 | 39.37 | 8.01 | 121.85 | 172.84 | -     | -     |
| <b>Glu 57</b> | 9.173 | 118.2 | 59.71 | 28.57 | 8.04 | 123.14 | 173.33 | -     | -     |
| <b>Gln 58</b> | 7.917 | 119.1 | 59.36 | 27.83 | 8.2  | 122.46 | 173.35 | 53.15 | 26.42 |
| <b>Ser 59</b> | 7.862 | 114   | 61.61 | 63.35 | 8.21 | 118.02 | 171.79 | 55.73 | 60.66 |
| <b>Lys 60</b> | 8.001 | 123.7 | 60.87 | 31.77 | 8.12 | 124.03 | 174.64 | 53.45 | 30.03 |
| <b>Gln 61</b> | 7.729 | 118.7 | 58.72 | 28.36 | 8.15 | 122.14 | 173.09 | 52.92 | 26.35 |
| <b>Leu 62</b> | 8.136 | 119.3 | 56.69 | 42.23 | 8.18 | 124.52 | 174.76 | 52.3  | 39.25 |
| <b>Thr 63</b> | 7.36  | 104.4 | 60.97 | 70.48 | 7.96 | 115.28 | 171.72 | -     | -     |
| <b>Glu 64</b> | 8.306 | 123.7 | 58.16 | 27.49 | 8.3  | 123.57 | 173.38 | -     | -     |
| <b>Asn 65</b> | 8.894 | 111.1 | 55.65 | 38.24 | 8.31 | 120.22 | 172.95 | 50.42 | 35.85 |
| <b>Gly 66</b> | 8.723 | 112.1 | 47.7  | -     | 8.19 | 110.23 | 171.07 | 42.4  | -     |
| <b>Ala 67</b> | 8.396 | 120.3 | 55.37 | 19.19 | 7.9  | 124    | 175.03 | -     | -     |
| <b>Glu 68</b> | 7.763 | 115.8 | 59.77 | 29.52 | 8.24 | 120.23 | 173.77 | 53.7  | 27.12 |
| <b>Ser 69</b> | 7.634 | 115.1 | 61.8  | 62.86 | 8.1  | 117.81 | 171.59 | 55.45 | 60.78 |
| <b>Val 70</b> | 7.505 | 121.3 | 66.76 | 31.92 | 7.94 | 122.85 | 173.21 | 59.38 | 29.64 |
| <b>Leu 71</b> | 8.304 | 120.4 | 58.8  | 40.9  | 8.21 | 123.2  | 172.61 | 52.65 | 26.49 |
| <b>Gln 72</b> | 8.378 | 117.1 | 58.89 | 27.67 | -    | -      | -      | -     | -     |
| <b>Val 73</b> | 7.582 | 119.5 | 67.17 | 31.43 | -    | -      | 172.77 | 59.23 | 29.99 |
| <b>Phe 74</b> | 7.507 | 116.4 | 62.96 | 39.42 | 8.27 | 125.92 | 172.42 | 54.92 | 36.7  |
| <b>Arg 75</b> | 8.45  | 117.7 | 60.69 | 29.78 | 7.99 | 125.14 | 172.46 | 52.61 | 28.32 |
| <b>Glu 76</b> | 8.899 | 122.7 | 59.24 | 28.95 | -    | -      | 173.08 | 53.3  | 27.35 |
| <b>Ala 77</b> | 8.513 | 123.5 | 55.64 | 17.34 | 8.24 | 126.78 | 173.61 | 49.39 | 16.12 |
| <b>Lys 78</b> | 8.246 | 120.4 | 58.73 | 33.49 | 8.15 | 122.14 | 173.35 | 53.06 | 30.28 |
| <b>Ala 79</b> | 8.332 | 123.7 | 55.1  | 18.07 | 8.24 | 127.08 | 174.79 | 49.42 | 16.11 |
| <b>Glu 80</b> | 7.662 | 115.6 | 56.53 | 29.62 | 8.29 | 121.54 | 174.11 | 53.64 | 27.36 |
| <b>Gly 81</b> | 7.769 | 105.8 | 45.16 | -     | 8.25 | 110.9  | 170.84 | 42.2  | -     |
| <b>Ala 82</b> | 7.895 | 124.8 | 52.85 | 18.78 | 7.91 | 123.87 | 174.5  | 49.58 | 16.52 |
| <b>Asp 83</b> | 7.564 | 122.4 | 52.98 | 39.19 | 8.18 | 120.05 | 173.43 | 51.16 | 37.86 |
| <b>Ile 84</b> | 7.408 | 124.9 | 59.64 | 42.6  | 7.86 | 121.77 | 173.49 | 58.26 | 35.72 |
| <b>Thr 85</b> | 8.783 | 126.1 | 62.27 | 70.59 | 8.14 | 120.23 | 171.28 | 59.51 | 66.72 |
| <b>Ile 86</b> | 9.069 | 128.1 | 60.9  | 39.69 | 7.95 | 125.85 | 172.69 | 57.96 | 35.61 |
| <b>Ile 87</b> | -     | -     | -     | -     | 8.18 | 128.15 | 173.17 | 57.7  | 35.3  |
| <b>Leu 88</b> | 8.59  | 128.7 | -     | -     | 8.31 | 129.44 | 173.53 | 51.91 | 39.22 |
| <b>Ser 89</b> | -     | -     | -     | -     | 7.75 | 123.2  | -      | 57    | 61.81 |

## 1.2. (R)-NPPM AND (S)-NPPM MODIFIED DNA DOUBLE STRAND

Table 10 Chemical shift list of (R)-NPPMdC9 and (S)-NPPMdC9 caged 15 basepair long DNA double strand. See sequence in Figure 23. All samples were in 25 mM K<sub>y</sub>H<sub>y</sub>PO<sub>4</sub>, 50 mM KCl, pH 6.4 buffer with 100 mM DSS added as internal reference. <sup>1</sup>H chemical shifts have been referenced to the methyl signal of DSS.

| <b>(R)-NPPM</b> |             |                       | <b>(S)-NPPM</b> |             |                       |
|-----------------|-------------|-----------------------|-----------------|-------------|-----------------------|
| <b>Nuclei</b>   | <b>Atom</b> | <b>Chemical shift</b> | <b>Nuclei</b>   | <b>Atom</b> | <b>Chemical shift</b> |
| G1              | H1'         | 5.97                  | G1              | H1'         | 5.972                 |
| G1              | H2'1        | 2.61                  | G1              | H2'1        | 2.61                  |
| G1              | H2'2        | 2.78                  | G1              | H2'2        | 2.779                 |
| G1              | H3'         | 4.836                 | G1              | H3'         | 4.834                 |
| G1              | H4'         | 4.227                 | G1              | H4'         | 4.225                 |
| G1              | H5'#        | 3.702                 | G1              | H5'#        | 3.702                 |
| G1              | H8          | 7.949                 | G1              | H8          | 7.95                  |
| C2              | H1'         | 5.628                 | C2              | H1'         | 5.623                 |
| C2              | H2'1        | 2.162                 | C2              | H2'1        | 2.156                 |
| C2              | H2'2        | 2.458                 | C2              | H2'2        | 2.455                 |
| C2              | H3'         | 4.875                 | C2              | H41         | 8.416                 |
| C2              | H41         | 8.418                 | C2              | H42         | 6.565                 |
| C2              | H42         | 6.57                  | C2              | H5          | 5.414                 |
| C2              | H5          | 5.412                 | C2              | H6          | 7.458                 |
| C2              | H6          | 7.459                 | A3              | H1'         | 6.287                 |
| A3              | H1'         | 6.289                 | A3              | H2          | 7.627                 |
| A3              | H2          | 7.626                 | A3              | H2'1        | 2.712                 |
| A3              | H2'1        | 2.712                 | A3              | H2'2        | 2.97                  |
| A3              | H2'2        | 2.974                 | A3              | H3'         | 5.034                 |
| A3              | H3'         | 5.035                 | A3              | H4'         | 4.446                 |
| A3              | H4'         | 4.448                 | A3              | H62         | 6.319                 |
| A3              | H8          | 8.354                 | A3              | H8          | 8.353                 |
| T4              | H1'         | 5.641                 | T4              | H1'         | 5.639                 |
| T4              | H2'1        | 2.092                 | T4              | H2'1        | 2.097                 |
| T4              | H2'2        | 2.481                 | T4              | H2'2        | 2.484                 |
| T4              | H3          | 13.187                | T4              | H3          | 13.175                |
| T4              | H3'         | 4.889                 | T4              | H6          | 7.164                 |
| T4              | H6          | 7.167                 | T4              | H7#         | 1.459                 |
| T4              | H7#         | 1.461                 | A5              | H1'         | 6.234                 |
| A5              | H1'         | 6.258                 | A5              | H2          | 7.023                 |
| A5              | H2          | 7.063                 | A5              | H2'1        | 2.65                  |
| A5              | H2'1        | 2.655                 | A5              | H2'2        | 2.961                 |
| A5              | H2'2        | 2.983                 | A5              | H3'         | 5.055                 |
| A5              | H3'         | 5.062                 | A5              | H4'         | 4.45                  |
| A5              | H4'         | 4.458                 | A5              | H8          | 8.289                 |
| A5              | H61         | 6.261                 | T6              | H1'         | 5.65                  |
| A5              | H62         | 7.432                 | T6              | H2'1        | 2.281                 |
| A5              | H8          | 8.306                 | T6              | H2'2        | 2.569                 |
| T6              | H1'         | 5.695                 | T6              | H3          | 13.131                |
| T6              | H2'1        | 2.343                 | T6              | H3'         | 4.957                 |
| T6              | H2'2        | 2.717                 | T6              | H4'         | 4.234                 |

Appendix

---

|            |      |        |            |      |        |
|------------|------|--------|------------|------|--------|
| <b>T6</b>  | H3   | 13.236 | <b>T6</b>  | H6   | 7.292  |
| <b>T6</b>  | H3'  | 4.975  | <b>T6</b>  | H7#  | 1.44   |
| <b>T6</b>  | H6   | 7.393  | <b>A7</b>  | H1'  | 6.172  |
| <b>T6</b>  | H7#  | 1.472  | <b>A7</b>  | H2   | 7.085  |
| <b>A7</b>  | H1'  | 6.195  | <b>A7</b>  | H2'1 | 2.572  |
| <b>A7</b>  | H2   | 7.157  | <b>A7</b>  | H2'2 | 2.819  |
| <b>A7</b>  | H2'1 | 2.673  | <b>A7</b>  | H3'  | 4.95   |
| <b>A7</b>  | H2'2 | 2.83   | <b>A7</b>  | H4'  | 4.427  |
| <b>A7</b>  | H3'  | 4.972  | <b>A7</b>  | H8   | 7.893  |
| <b>A7</b>  | H4'  | 4.458  | <b>X8</b>  | H1'  | 5.574  |
| <b>A7</b>  | H62  | 7.427  | <b>X8</b>  | H10  | 7.852  |
| <b>A7</b>  | H8   | 8.369  | <b>X8</b>  | H11  | 7.735  |
| <b>X8</b>  | H1'  | 5.668  | <b>X8</b>  | H12  | 7.577  |
| <b>X8</b>  | H10  | 7.997  | <b>X8</b>  | H13  | 8.058  |
| <b>X8</b>  | H11  | 7.924  | <b>X8</b>  | H2'1 | 1.852  |
| <b>X8</b>  | H12  | 7.511  | <b>X8</b>  | H2'2 | 2.454  |
| <b>X8</b>  | H2'1 | 1.857  | <b>X8</b>  | H3'  | 4.589  |
| <b>X8</b>  | H2'2 | 2.452  | <b>X8</b>  | H4   | 8.677  |
| <b>X8</b>  | H3'  | 4.577  | <b>X8</b>  | H4'  | 4.323  |
| <b>X8</b>  | H4   | 8.567  | <b>X8</b>  | H5'# | 4.18   |
| <b>X8</b>  | H4'  | 4.316  | <b>X8</b>  | H6   | 7.204  |
| <b>X8</b>  | H5   | 5.131  | <b>X8</b>  | H7   | 5.368  |
| <b>X8</b>  | H6   | 7.258  | <b>X8</b>  | Hd#  | 7.298  |
| <b>X8</b>  | H7   | 6.12   | <b>X8</b>  | He#  | 7.36   |
| <b>X8</b>  | Hd#  | 7.419  | <b>X8</b>  | Hz   | 7.298  |
| <b>X8</b>  | He#  | 7.36   | <b>T9</b>  | H1'  | 5.637  |
| <b>X8</b>  | Hz   | 7.418  | <b>T9</b>  | H2'1 | 2.054  |
| <b>T9</b>  | H1'  | 5.708  | <b>T9</b>  | H2'2 | 2.423  |
| <b>T9</b>  | H2'1 | 2.013  | <b>T9</b>  | H3   | 13.294 |
| <b>T9</b>  | H2'2 | 2.432  | <b>T9</b>  | H3'  | 4.854  |
| <b>T9</b>  | H3   | 13.118 | <b>T9</b>  | H6   | 7.29   |
| <b>T9</b>  | H3'  | 4.852  | <b>T9</b>  | H7#  | 1.379  |
| <b>T9</b>  | H6   | 7.233  | <b>A10</b> | H1'  | 5.886  |
| <b>T9</b>  | H7#  | 1.22   | <b>A10</b> | H2   | 6.954  |
| <b>A10</b> | H1'  | 5.839  | <b>A10</b> | H2'1 | 2.75   |
| <b>A10</b> | H2   | 6.884  | <b>A10</b> | H2'2 | 2.913  |
| <b>A10</b> | H2'1 | 2.728  | <b>A10</b> | H3'  | 5.049  |
| <b>A10</b> | H2'2 | 2.883  | <b>A10</b> | H4'  | 4.397  |
| <b>A10</b> | H3'  | 5.04   | <b>A10</b> | H8   | 8.253  |
| <b>A10</b> | H4'  | 4.374  | <b>A11</b> | H1'  | 5.934  |
| <b>A10</b> | H61  | 7.946  | <b>A11</b> | H2   | 7.411  |
| <b>A10</b> | H62  | 6.323  | <b>A11</b> | H2'1 | 2.593  |
| <b>A10</b> | H8   | 8.255  | <b>A11</b> | H2'2 | 2.837  |
| <b>A11</b> | H1'  | 5.935  | <b>A11</b> | H3'  | 5.045  |
| <b>A11</b> | H2   | 7.42   | <b>A11</b> | H4'  | 4.424  |
| <b>A11</b> | H2'1 | 2.592  | <b>A11</b> | H61  | 7.599  |
| <b>A11</b> | H2'2 | 2.832  | <b>A11</b> | H62  | 5.98   |
| <b>A11</b> | H3'  | 5.037  | <b>A11</b> | H8   | 8.041  |
| <b>A11</b> | H4'  | 4.42   | <b>G12</b> | H1   | 12.825 |



Appendix

---

|            |      |        |            |      |        |
|------------|------|--------|------------|------|--------|
| <b>A11</b> | H61  | 5.99   | <b>G12</b> | H1'  | 5.576  |
| <b>A11</b> | H62  | 7.615  | <b>G12</b> | H2'1 | 2.516  |
| <b>A11</b> | H8   | 8.046  | <b>G12</b> | H2'2 | 2.61   |
| <b>G12</b> | H1   | 12.835 | <b>G12</b> | H3'  | 4.933  |
| <b>G12</b> | H1'  | 5.58   | <b>G12</b> | H4'  | 4.357  |
| <b>G12</b> | H2'1 | 2.517  | <b>G12</b> | H8   | 7.543  |
| <b>G12</b> | H2'2 | 2.613  | <b>G13</b> | H1   | 12.767 |
| <b>G12</b> | H3'  | 4.931  | <b>G13</b> | H1'  | 5.912  |
| <b>G12</b> | H4'  | 4.355  | <b>G13</b> | H2'1 | 2.452  |
| <b>G12</b> | H5'# | 4.179  | <b>G13</b> | H2'2 | 2.702  |
| <b>G12</b> | H8   | 7.543  | <b>G13</b> | H4'  | 4.367  |
| <b>G13</b> | H1   | 12.775 | <b>G13</b> | H8   | 7.543  |
| <b>G13</b> | H1'  | 5.916  | <b>T14</b> | H1'  | 5.893  |
| <b>G13</b> | H2'1 | 2.454  | <b>T14</b> | H2'1 | 1.963  |
| <b>G13</b> | H2'2 | 2.705  | <b>T14</b> | H2'2 | 2.391  |
| <b>G13</b> | H3'  | 4.861  | <b>T14</b> | H3   | 13.904 |
| <b>G13</b> | H4'  | 4.371  | <b>T14</b> | H3'  | 4.854  |
| <b>G13</b> | H8   | 7.545  | <b>T14</b> | H6   | 7.196  |
| <b>T14</b> | H1'  | 5.896  | <b>T14</b> | H7#  | 1.353  |
| <b>T14</b> | H2'1 | 1.965  | <b>G15</b> | H1'  | 6.173  |
| <b>T14</b> | H2'2 | 2.393  | <b>G15</b> | H2'1 | 2.601  |
| <b>T14</b> | H3   | 13.911 | <b>G15</b> | H2'2 | 2.356  |
| <b>T14</b> | H3'  | 4.856  | <b>G15</b> | H3'  | 4.68   |
| <b>T14</b> | H4'  | 4.173  | <b>G15</b> | H4'  | 4.184  |
| <b>T14</b> | H6   | 7.2    | <b>G15</b> | H8   | 7.909  |
| <b>T14</b> | H7#  | 1.355  | <b>C16</b> | H1'  | 5.668  |
| <b>G15</b> | H1'  | 6.174  | <b>C16</b> | H2'1 | 1.942  |
| <b>G15</b> | H2'1 | 2.358  | <b>C16</b> | H2'2 | 2.412  |
| <b>G15</b> | H2'2 | 2.604  | <b>C16</b> | H3'  | 4.712  |
| <b>G15</b> | H3'  | 4.682  | <b>C16</b> | H4'  | 4.068  |
| <b>G15</b> | H8   | 7.909  | <b>C16</b> | H41  | 8.157  |
| <b>C16</b> | H1'  | 5.669  | <b>C16</b> | H42  | 7.045  |
| <b>C16</b> | H2'1 | 1.941  | <b>C16</b> | H5   | 5.969  |
| <b>C16</b> | H2'2 | 2.413  | <b>C16</b> | H5'# | 3.748  |
| <b>C16</b> | H3'  | 4.714  | <b>C16</b> | H6   | 7.714  |
| <b>C16</b> | H4'  | 4.067  | <b>A17</b> | H1'  | 6.285  |
| <b>C16</b> | H41  | 7.048  | <b>A17</b> | H2   | 7.939  |
| <b>C16</b> | H42  | 8.164  | <b>A17</b> | H2'1 | 2.786  |
| <b>C16</b> | H5   | 5.965  | <b>A17</b> | H2'2 | 2.941  |
| <b>C16</b> | H5'1 | 3.759  | <b>A17</b> | H3'  | 5.036  |
| <b>C16</b> | H5'2 | 3.74   | <b>A17</b> | H4'  | 4.437  |
| <b>C16</b> | H6   | 7.713  | <b>A17</b> | H8   | 8.375  |
| <b>A17</b> | H1'  | 6.287  | <b>C18</b> | H1'  | 5.878  |
| <b>A17</b> | H2   | 7.937  | <b>C18</b> | H2'1 | 2.118  |
| <b>A17</b> | H2'1 | 2.788  | <b>C18</b> | H2'2 | 2.474  |
| <b>A17</b> | H2'2 | 2.946  | <b>C18</b> | H3'  | 4.785  |
| <b>A17</b> | H3'  | 5.036  | <b>C18</b> | H4'  | 4.259  |
| <b>A17</b> | H4'  | 4.441  | <b>C18</b> | H41  | 8.079  |
| <b>A17</b> | H8   | 8.377  | <b>C18</b> | H42  | 6.636  |

Appendix

---

|            |      |        |            |      |        |
|------------|------|--------|------------|------|--------|
| <b>C18</b> | H1'  | 5.883  | <b>C18</b> | H5   | 5.303  |
| <b>C18</b> | H2'1 | 2.12   | <b>C18</b> | H6   | 7.326  |
| <b>C18</b> | H2'2 | 2.484  | <b>C19</b> | H1'  | 5.896  |
| <b>C18</b> | H3'  | 4.79   | <b>C19</b> | H2'1 | 2.107  |
| <b>C18</b> | H4'  | 4.262  | <b>C19</b> | H2'2 | 2.513  |
| <b>C18</b> | H41  | 8.084  | <b>C19</b> | H3'  | 4.749  |
| <b>C18</b> | H42  | 6.644  | <b>C19</b> | H4'  | 4.174  |
| <b>C18</b> | H5   | 5.305  | <b>C19</b> | H41  | 8.271  |
| <b>C18</b> | H6   | 7.331  | <b>C19</b> | H42  | 6.855  |
| <b>C19</b> | H1'  | 5.91   | <b>C19</b> | H5   | 5.469  |
| <b>C19</b> | H2'1 | 2.112  | <b>C19</b> | H6   | 7.501  |
| <b>C19</b> | H2'2 | 2.517  | <b>T20</b> | H1'  | 5.997  |
| <b>C19</b> | H3'  | 4.767  | <b>T20</b> | H2'1 | 2.09   |
| <b>C19</b> | H4'  | 4.179  | <b>T20</b> | H2'2 | 2.515  |
| <b>C19</b> | H41  | 8.288  | <b>T20</b> | H3   | 13.994 |
| <b>C19</b> | H42  | 6.873  | <b>T20</b> | H3'  | 4.843  |
| <b>C19</b> | H5   | 5.479  | <b>T20</b> | H6   | 7.444  |
| <b>C19</b> | H6   | 7.514  | <b>T20</b> | H7#  | 1.624  |
| <b>T20</b> | H1'  | 6.004  | <b>T21</b> | H1'  | 5.628  |
| <b>T20</b> | H2'1 | 2.068  | <b>T21</b> | H2'1 | 2.128  |
| <b>T20</b> | H2'2 | 2.513  | <b>T21</b> | H2'2 | 2.45   |
| <b>T20</b> | H3   | 14.017 | <b>T21</b> | H3   | 13.572 |
| <b>T20</b> | H3'  | 4.835  | <b>T21</b> | H6   | 7.355  |
| <b>T20</b> | H4'  | 4.166  | <b>T21</b> | H7#  | 1.638  |
| <b>T20</b> | H6   | 7.43   | <b>A22</b> | H1'  | 6.186  |
| <b>T20</b> | H7#  | 1.665  | <b>A22</b> | H2   | 7.134  |
| <b>T21</b> | H1'  | 5.627  | <b>A22</b> | H2'1 | 2.706  |
| <b>T21</b> | H2'1 | 2.168  | <b>A22</b> | H2'2 | 2.975  |
| <b>T21</b> | H2'2 | 2.451  | <b>A22</b> | H3'  | 5.017  |
| <b>T21</b> | H3   | 13.607 | <b>A22</b> | H4'  | 4.43   |
| <b>T21</b> | H3'  | 4.873  | <b>A22</b> | H61  | 7.832  |
| <b>T21</b> | H6   | 7.37   | <b>A22</b> | H62  | 5.439  |
| <b>T21</b> | H7#  | 1.51   | <b>A22</b> | H8   | 8.191  |
| <b>A22</b> | H1'  | 6.223  | <b>G23</b> | H1   | 12.245 |
| <b>A22</b> | H2   | 7.189  | <b>G23</b> | H1'  | 5.582  |
| <b>A22</b> | H2'1 | 2.73   | <b>G23</b> | H2'1 | 2.203  |
| <b>A22</b> | H2'2 | 2.986  | <b>G23</b> | H2'2 | 2.525  |
| <b>A22</b> | H3'  | 5.005  | <b>G23</b> | H3'  | 4.574  |
| <b>A22</b> | H4'  | 4.44   | <b>G23</b> | H4'  | 4.325  |
| <b>A22</b> | H61  | 6.104  | <b>G23</b> | H8   | 7.116  |
| <b>A22</b> | H62  | 8.395  | <b>T24</b> | H1'  | 5.897  |
| <b>A22</b> | H8   | 8.22   | <b>T24</b> | H2'1 | 2.177  |
| <b>G23</b> | H1   | 12.318 | <b>T24</b> | H2'2 | 2.574  |
| <b>G23</b> | H1'  | 5.537  | <b>T24</b> | H3   | 13.376 |
| <b>G23</b> | H2'1 | 2.133  | <b>T24</b> | H6   | 7.265  |
| <b>G23</b> | H2'2 | 2.504  | <b>T24</b> | H7#  | 1.037  |
| <b>G23</b> | H3'  | 4.573  | <b>A25</b> | H1'  | 6.22   |
| <b>G23</b> | H4'  | 4.284  | <b>A25</b> | H2   | 6.985  |
| <b>G23</b> | H8   | 7.029  | <b>A25</b> | H2'1 | 2.648  |

Appendix

---

|            |      |        |            |      |        |
|------------|------|--------|------------|------|--------|
| <b>T24</b> | H1'  | 5.819  | <b>A25</b> | H2'2 | 2.948  |
| <b>T24</b> | H2'1 | 2.137  | <b>A25</b> | H3'  | 4.995  |
| <b>T24</b> | H2'2 | 2.539  | <b>A25</b> | H4'  | 4.406  |
| <b>T24</b> | H3   | 13.319 | <b>A25</b> | H8   | 8.263  |
| <b>T24</b> | H3'  | 4.855  | <b>T26</b> | H1'  | 5.699  |
| <b>T24</b> | H4'  | 4.231  | <b>T26</b> | H2'1 | 2.149  |
| <b>T24</b> | H6   | 7.206  | <b>T26</b> | H2'2 | 2.515  |
| <b>T24</b> | H7#  | 1.014  | <b>T26</b> | H3   | 13.133 |
| <b>A25</b> | H1'  | 6.226  | <b>T26</b> | H6   | 7.183  |
| <b>A25</b> | H2   | 7.066  | <b>T26</b> | H7#  | 1.335  |
| <b>A25</b> | H2'1 | 2.644  | <b>A27</b> | H1'  | 6.228  |
| <b>A25</b> | H2'2 | 2.948  | <b>A27</b> | H2   | 7.152  |
| <b>A25</b> | H3'  | 4.992  | <b>A27</b> | H2'1 | 2.628  |
| <b>A25</b> | H4'  | 4.406  | <b>A27</b> | H2'2 | 2.931  |
| <b>A25</b> | H61  | 6.255  | <b>A27</b> | H3'  | 5.018  |
| <b>A25</b> | H62  | 7.388  | <b>A27</b> | H4'  | 4.444  |
| <b>A25</b> | H8   | 8.277  | <b>A27</b> | H61  | 7.451  |
| <b>T26</b> | H1'  | 5.693  | <b>A27</b> | H62  | 6.239  |
| <b>T26</b> | H2'1 | 2.147  | <b>A27</b> | H8   | 8.275  |
| <b>T26</b> | H2'2 | 2.512  | <b>T28</b> | H1'  | 5.783  |
| <b>T26</b> | H3   | 13.152 | <b>T28</b> | H2'1 | 1.975  |
| <b>T26</b> | H3'  | 4.897  | <b>T28</b> | H2'2 | 2.375  |
| <b>T26</b> | H6   | 7.186  | <b>T28</b> | H3   | 13.551 |
| <b>T26</b> | H7#  | 1.345  | <b>T28</b> | H4'  | 4.364  |
| <b>A27</b> | H1'  | 6.239  | <b>T28</b> | H6   | 7.096  |
| <b>A27</b> | H2   | 7.171  | <b>T28</b> | H7#  | 1.335  |
| <b>A27</b> | H2'1 | 2.633  | <b>G29</b> | H1   | 12.768 |
| <b>A27</b> | H2'2 | 2.931  | <b>G29</b> | H1'  | 5.892  |
| <b>A27</b> | H3'  | 5.021  | <b>G29</b> | H2'1 | 2.606  |
| <b>A27</b> | H4'  | 4.452  | <b>G29</b> | H2'2 | 2.693  |
| <b>A27</b> | H61  | 7.463  | <b>G29</b> | H3'  | 4.974  |
| <b>A27</b> | H62  | 6.264  | <b>G29</b> | H4'  | 4.364  |
| <b>A27</b> | H8   | 8.283  | <b>G29</b> | H8   | 7.87   |
| <b>T28</b> | H1'  | 5.787  | <b>C30</b> | H1'  | 6.195  |
| <b>T28</b> | H2'1 | 1.978  | <b>C30</b> | H2'# | 2.172  |
| <b>T28</b> | H2'2 | 2.378  | <b>C30</b> | H3'  | 4.49   |
| <b>T28</b> | H3   | 13.558 | <b>C30</b> | H42  | 6.698  |
| <b>T28</b> | H3'  | 4.875  | <b>C30</b> | H5   | 5.445  |
| <b>T28</b> | H6   | 7.103  | <b>C30</b> | H6   | 7.471  |
| <b>T28</b> | H7#  | 1.343  |            |      |        |
| <b>G29</b> | H1   | 12.769 |            |      |        |
| <b>G29</b> | H1'  | 5.894  |            |      |        |
| <b>G29</b> | H2'1 | 2.605  |            |      |        |
| <b>G29</b> | H2'2 | 2.695  |            |      |        |
| <b>G29</b> | H3'  | 4.973  |            |      |        |
| <b>G29</b> | H4'  | 4.367  |            |      |        |
| <b>G29</b> | H8   | 7.869  |            |      |        |
| <b>C30</b> | H1'  | 6.191  |            |      |        |
| <b>C30</b> | H2'1 | 2.16   |            |      |        |

|            |     |       |
|------------|-----|-------|
| <b>C30</b> | H2' | 2.184 |
| <b>C30</b> | H3' | 4.491 |
| <b>C30</b> | H41 | 8.184 |
| <b>C30</b> | H42 | 6.692 |
| <b>C30</b> | H5  | 5.43  |
| <b>C30</b> | H6  | 7.465 |

### 1.3. RESONANCE ASSIGNMENT OF GCAA 14NT TETRALOOP

Table 11 Chemical shift assignment of the GCAA 14nt RNA tetraloop. Resonances marked with \* could not be unambiguously assigned. The resonances are from spectrum measured at 298 K.  $^{13}\text{C}$  and  $^1\text{H}$  chemical shift are from HSQC spectrum measured on non-isotope labelled sample.  $^{15}\text{N}$  values are from uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled samples.  $^1\text{H}$  chemical shifts have been referenced to DSS internal reference signal, and  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts have been indirectly referenced according to reference.<sup>[179]</sup>

| $\begin{matrix} \text{Nuc} \\ \text{Res} \end{matrix}$ | <b>H1'</b> | <b>H2'</b> | <b>H3'</b> | <b>H4'</b> | <b>H5'</b> | <b>H5''</b> |
|--|------------|------------|------------|------------|------------|-------------|
| <b>1</b>   | 5.82       | 4.94       | 4.72       | 4.57       | 4.40       | 4.26        |
| <b>2</b>   | 5.92       | 4.55       | 4.59       | 4.56       | 4.53       | 4.26        |
| <b>3</b>   | 5.53       | 4.50       | 4.58       | 4.45       | 4.59       | 4.13        |
| <b>4</b>   | 5.96       | 4.49       | 4.68       | 4.47       | 4.58       | 4.15        |
| <b>5</b>   | 5.38       | 4.35       | 4.29       | 4.36       | 4.47       | 4.05        |
| <b>6</b>   | 5.68       | 4.53       | 4.64       | 4.37       | 4.38       | 4.08        |
| <b>7</b>   | 5.56       | 4.38       | 4.26       | 4.10       | 4.14       | 3.94        |
| <b>8</b>   | 5.62       | 4.36       | 4.56       | 4.21       | 3.86       | 3.77        |
| <b>9</b>   | 6.06       | 4.63       | 4.19       | 4.49       | 4.44       | 4.43        |
| <b>10</b>  | 3.96       | 4.27       |            | 4.28       | 4.35       | 4.23        |
| <b>11</b>  | 5.55       | 4.63       | 4.55       | 4.41       | 4.45       | 4.06        |
| <b>12</b>  | 5.81       | 4.51       | 4.60       | 4.49       | 4.52       | 4.13        |
| <b>13</b>  | 5.46       | 4.18       | 4.42       | 4.37       | 4.54       | 4.04        |
| <b>14</b>  | 6.09       | 4.62       | 4.94       | 4.39       | 4.51       | 4.08        |

| $\begin{matrix} \text{Nuc} \\ \text{Res} \end{matrix}$ | <b>C1'</b> | <b>C2'</b> | <b>C3'</b> | <b>C4'</b> | <b>C5'</b> |
|--|------------|------------|------------|------------|------------|
| <b>1</b>   | 89.02      | 72.26      | 71.95      | 80.71      | 64.25      |
| <b>2</b>   | 90.30      | 72.62      | 70.25      | 79.72      | 63.19      |
| <b>3</b>   | 91.05      | 72.63      | 69.42      | 79.10      | 61.73      |
| <b>4</b>   | 90.22      | 73.01      | 69.93      | 79.21      | 62.16      |
| <b>5</b>   | 91.00      | 73.05      | 69.51      | 79.13      | 62.16      |
| <b>6</b>   | 89.72      | 73.39      | 71.00      | 79.84      | 62.46      |
| <b>7</b>   | 89.36      | 73.17      | 72.06      | 80.80      | 63.06      |
| <b>8</b>   | 89.02      | 74.21      | 69.41      | 81.09      | 63.46      |
| <b>9</b>   | 89.45      | 74.15      | 71.70      | 80.38      | 63.94      |
| <b>10</b>  | 90.14      | 72.01      |            | 80.33      | 66.85      |
| <b>11</b>  | 90.83      | 72.42      | 69.59      | 79.27      | 61.71      |
| <b>12</b>  | 90.06      | 72.62      | 70.09      | 79.21      | 62.54      |
| <b>13</b>  | 91.62      | 72.91      | 69.38      | 79.30      | 61.70      |
| <b>14</b>  | 89.27      | 82.51      | 75.09      | 82.79      | 62.89      |

Appendix

| Res \ Nuc | H1    | H2   | H3    | H5   | H6    | H8   |
|-----------|-------|------|-------|------|-------|------|
| 1         |       |      |       |      |       | 8.14 |
| 2         | 13.31 |      |       |      |       | 7.62 |
| 3         |       |      |       | 5.30 | 7.72* |      |
| 4         |       | 7.37 |       |      |       | 8.04 |
| 5         |       |      |       | 5.10 | 7.23  |      |
| 6         |       |      |       |      |       | 7.59 |
| 7         |       |      |       | 5.61 | 7.75  |      |
| 8         |       | 8.06 |       |      |       | 8.02 |
| 9         |       | 7.90 |       |      |       | 8.24 |
| 10        | 12.98 |      |       |      |       | 7.83 |
| 11        |       |      | 13.68 | 5.11 | 7.72* |      |
| 12        | 12.66 |      |       |      |       | 7.75 |
| 13        |       |      |       | 5.26 | 7.64  |      |
| 14        |       |      |       | 5.54 | 7.47  |      |

| Res \ Nuc | C2     | C5     | C6      | C8     | P     | N1     | N9     | N3     |
|-----------|--------|--------|---------|--------|-------|--------|--------|--------|
| 1         |        |        |         | 136.35 | -0.14 | 133.73 | 168.96 |        |
| 2         |        |        |         | 134.07 | -0.72 | 148.74 | 169.49 |        |
| 3         |        | 94.71  | 138.85* |        | -1.14 | 150.76 |        | 196.56 |
| 4         | 150.44 |        |         | 136.63 | -0.81 | 221.88 | 170.87 |        |
| 5         |        | 94.92  | 136.93  |        | -0.91 | 150.29 |        | 195.16 |
| 6         |        |        |         | 133.75 | -0.84 |        | 169.61 |        |
| 7         |        | 95.24  | 140.64  |        | 0.17  | 151.99 |        |        |
| 8         | 152.46 |        |         | 137.97 | -0.44 |        | 169.95 |        |
| 9         | 152.03 |        |         | 139.07 | -1.38 |        | 171.98 |        |
| 10        |        |        |         | 134.87 | 0.02  | 147.34 | 168.97 |        |
| 11        |        | 100.26 | 138.03* |        | -1.48 | 146.00 |        | 161.40 |
| 12        |        |        |         | 133.44 | -0.91 | 147.34 | 169.68 |        |
| 13        |        | 94.34  | 138.32  |        | -1.32 | 151.23 |        | 197.97 |
| 14        |        | 96.47  | 138.91  |        | -1.17 | 115.01 | 150.24 |        |

## 2. PULSE SEQUENCES AND AU PROGRAMS

### 2.1. AU PROGRAM TO SET UP CYCLING TWO DIMENSIONAL T-JUMP EXPERIMENT

The script has to be started on the desired T-jump experiment with optimized heating and pulses. Afterwards the measurement can be started with a python script from the first expno.

```
#include <inc/exptUtil>

int numExpt;
float T1counter;
int i;
GETINT("Enter number of experiments:", numExpt)

/*setting the right experiment and T1 delay*/
for( i = 0 ; i <= numExpt - 1 ; ++i )
{
  SETCURDATA;
  /*STOREPAR ("AUNM", aunm );*/
  STOREPAR ("CNST 16", T1counter);
  if(i % 4 == 3)
    T1counter = T1counter + 1.0 ;

  if(i % 8 == 0)
    STOREPAR ("ZGOPTNS", "-DFIRSTone");
  if(i % 8 == 1)
    STOREPAR ("ZGOPTNS", "-DFIRSTtwo");

  if(i % 8 == 2)
    STOREPAR ("ZGOPTNS", "-DSECONDone");
  if(i % 8 == 3)
    STOREPAR ("ZGOPTNS", "-DSECONDto");

  if(i % 8 == 4)
    STOREPAR ("ZGOPTNS", "-DTHIRDone");
  if(i % 8 == 5)
    STOREPAR ("ZGOPTNS", "-DTHIRDto");

  if(i % 8 == 6)
    STOREPAR ("ZGOPTNS", "-DFOURTHone");
  if(i % 8 == 7)
    STOREPAR ("ZGOPTNS", "-DFOURHTwo");

  IEXPNO
}
/*start of Tjump AU program with required wait time*/

QUIT
```

## 2.2. PULSE SEQUENCE FOR STATE-CORRELATED T-JUMP EXPERIMENT

T-jump state correlated experiment with control measurement at low and high temperature. Setting up the required number of indirect points are done with AU program shown in the Appendix chapter 2.1.

```
;TJ.hsqcfpf3gpplhwgex-F1-tjump-F2.T1.gp
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple>

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

;for Tjump and RF heating part
;-----

define loopcounter td1m
"td1m=td1-3"
"d63=(p30+d30)*16*3"
"d25=p25"
"l0=0"

;-----

"p2=p1*2"
"p22=p21*2"
"d11=30m"
"d12=20u"
"d26=1s/(cnst4*4)"
"in0=inf1"
"d0=in0/2-p21*4/3.1415"
"d24=2.7m"

"DELTA1=d26-p16-d16"
"DELTA2=d26-p16-d16-p11-12u"
"d9=d8-d7"

1 ze
  d11 pl16:f3
  d63
2 d1 do:f3
3 d12 pl1:f1 pl3:f3
  50u ;UNBLKGRAD
  "d0=in0*cnst16+in0/2-p21*4/3.1415"
  "d9=d8-d7-d0"
  (p1 ph1)
  p16:gp3
  d16
```

DELTA1  
(center (p2 ph2) (p22 ph6):f3 )  
DELTA1  
p16:gp3  
d16  
(p1 ph2)

4u pl0:f1  
(p11:sp1 ph8:r):f1  
4u  
p16:gp4  
d16 pl1:f1

(p21 ph1):f3

;~~~~~

d24  
(center (p2 ph1) (p22 ph1):f3)  
d24  
(p21 ph2):f3

d7  
p16:gp6  
d16

```
# ifdef LABEL_CN
  (center (p2 ph5):f1 (p8:sp13 ph1):f2 (DELTA1 p21 ph3 d0 p21 ph4 DELTA1):f3 )
# else
  (center (p2 ph5):f1 (p21 ph3 d0 p21 ph4):f3 )
# endif /*LABEL_CN*/
```

;~~~~~

d63 ;T-jump time  
d9  
p16:gp7  
d16  
(p21 ph2):f3

;~~~~~

d24  
(center (p2 ph1) (p22 ph4):f3)  
d24  
(p21 ph1):f3  
4u pl0:f1  
(p11:sp1 ph7:r):f1  
4u  
4u pl1:f1  
(p1 ph5)  
4u  
p16:gp5  
d16  
DELTA2 pl0:f1  
(p11:sp1 ph7:r):f1  
4u  
4u pl1:f1  
(center (p2 ph1) (p22 ph1):f3 )  
4u pl0:f1  
(p11:sp1 ph7:r):f1  
4u



```
p16:gp5
d16
DELTA2 pl16:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
30m do:f3 wr #0 if #0 zd
lo to 2 times 2

;Required to communicate with AU program, work only on Topspin 2.1
;-----
1m print "file: synchZg start experiment ready"
autosuspend

; Exchange experiment with T-jump part
;-----

12 d12 pl1:f1 pl3:f3
50u ;UNBLKGRAD
"d0=in0*cnst16+in0/2-p21*4/3.1415"
"d9=d8-d7-d0"
(p1 ph1)
p16:gp3
d16
DELTA1
(center (p2 ph2) (p22 ph6):f3 )
DELTA1
p16:gp3
d16
(p1 ph2)
4u pl0:f1
(p11:sp1 ph8:r):f1
4u
p16:gp4
d16 pl1:f1
(p21 ph1):f3
;~~~~~
d24
(center (p2 ph1) (p22 ph1):f3)
d24
(p21 ph2):f3
d7 ;compensation time to keep TOTAL time
p16:gp6
d16
# ifdef LABEL_CN
(center (p2 ph5):f1 (p8:sp13 ph1):f2 (DELTA1 p21 ph3 d0 p21 ph4 DELTA1):f3 )
# else
(center (p2 ph5):f1 (p21 ph3 d0 p21 ph4):f3 )
# endif /*LABEL_CN*/
; ---Heating loop 1---
20u fq=cnst18 (bf ppm):f2
9 p30:f2
d30
lo to 9 times l6

; ---Heating loop 2---
20u fq=cnst19 (bf ppm):f2
10 p30:f2
d30
lo to 10 times l6
```

```

; ---Heating loop 3---
 20u fq=cnst20 (bf ppm):f2
11 p30:f2
  d30
  lo to 11 times l6

;~~~~~
 d9 ;folding time before magnetization transferred to proton
 p16:gp7
  d16
  (p21 ph2):f3
;~~~~~
 d24
 (center (p2 ph1) (p22 ph4):f3)
 d24
 (p21 ph1):f3
 4u pl0:f1
 (p11:sp1 ph7:r):f1
 4u
 4u pl1:f1
 (p1 ph5)
 4u
 p16:gp5
  d16
 DELTA2 pl0:f1
 (p11:sp1 ph7:r):f1
 4u
 4u pl1:f1
 (center (p2 ph1) (p22 ph1):f3 )
 4u pl0:f1
 (p11:sp1 ph7:r):f1
 4u
 p16:gp5
  d16
 DELTA2 pl16:f3
 4u BLKGRAD
 go=12 ph31 cpd3:f3
 30m do:f3 wr #0 if #0 zd

;----- rf heating -----
5 d29 do:f3
  if "l0 < l31"                ;l31 l31-1=rf heating part
  {
6 p25:f2
  d30
  lo to 6 times l7
  4u
  3m iu0
  }
  else
  {
7 d25
  d30
  lo to 7 times l7
  4u
  3m iu0
  }
;-----Control experiment at high temperature-----

```

```

8 d12 pl1:f1 pl3:f3
50u ;UNBLKGRAD
"d0=in0*cnst16+in0/2-p21*4/3.1415"
"d9=d8-d7-d0"
(p1 ph1)
p16:gp3
d16
DELTA1
(center (p2 ph2) (p22 ph6):f3 )
DELTA1
p16:gp3
d16
(p1 ph2)
4u pl0:f1
(p11:sp1 ph8:r):f1
4u
p16:gp4
d16 pl1:f1
(p21 ph1):f3
;~~~~~
d24
(center (p2 ph1) (p22 ph1):f3)
d24
(p21 ph2):f3
d7 ;refolding time
p16:gp6
d16
;~~~~~

# ifdef LABEL_CN
  (center (p2 ph5):f1 (p8:sp13 ph1):f2 (DELTA1 p21 ph3 d0 p21 ph4 DELTA1):f3 )
# else
  (center (p2 ph5):f1 (p21 ph3 d0 p21 ph4):f3 )
# endif /*LABEL_CN*/
;~~~~~
d9
d63
p16:gp7
d16
(p21 ph2):f3
;~~~~~
d24
(center (p2 ph1) (p22 ph4):f3)
d24
(p21 ph1):f3
4u pl0:f1
(p11:sp1 ph7:r):f1
4u
4u pl1:f1
(p1 ph5)
4u
p16:gp5
d16
DELTA2 pl0:f1
(p11:sp1 ph7:r):f1
4u
4u pl1:f1
(center (p2 ph1) (p22 ph1):f3 )

```

```

4u pl0:f1
(p11:sp1 ph7:r):f1
4u
p16:gp5
d16
DELTA2 pl16:f3
4u BLKGRAD

```

```

go=12 ph31 cpd3:f3
30m do:f3 wr #0 if #0 zd
lo to 5 times td1m
1s
exit

```

;~~~~~Phase cycle to be used for each kinetic experiment~~~~~

```
# ifdef FIRSTone
```

```

ph1=0
ph2=1
ph3=0 ;2
ph4=0
ph5=0 ;0 ;2 2
ph6=0
ph7=2
ph8=0
ph31=0; 2 ;2 0

```

```
# endif
```

```
# ifdef FIRSTtwo
```

```

ph1=0
ph2=1
ph3=2 ;0 2
ph4=0
ph5=0 ;0 0 ;2 2
ph6=0
ph7=2
ph8=0
ph31=2;0 2 ;2 0

```

```
# endif
```

```
# ifdef SECONDone
```

```

ph1=0
ph2=1
ph3=1
ph4=0
ph5=0 ;0 ;2 2
ph6=1
ph7=2
ph8=0
ph31=0; 2 ;2 0

```

```
# endif
```

```
# ifdef SECONDt看
```

```

ph1=0
ph2=1
ph3=3 ;0 2
ph4=0
ph5=0 ;0 0 ;2 2
ph6=1
ph7=2
ph8=0

```

```
ph31=2;0 2 ;2 0
# endif
# ifdef THIRDone
ph1=0
ph2=1
ph3=2
ph4=0
ph5=0 ;0 ;2 2
ph6=2
ph7=2
ph8=0
ph31=0; 2 ;2 0
# endif
# ifdef THIRDto
ph1=0
ph2=1
ph3=0 ;0 2
ph4=0
ph5=0 ;0 0 ;2 2
ph6=2
ph7=2
ph8=0
ph31=2;0 2 ;2 0
# endif
# ifdef FOURTHone
ph1=0
ph2=1
ph3=3
ph4=0
ph5=0 ;0 ;2 2
ph6=3
ph7=2
ph8=0
ph31=0; 2 ;2 0
# endif
# ifdef FOURTHtwo
ph1=0
ph2=1
ph3=1 ;0 2
ph4=0
ph5=0 ;0 0 ;2 2
ph6=3
ph7=2
ph8=0
ph31=2;0 2 ;2 0
# endif

;p0 : 0W
;p1 : f1 channel - power level for pulse (default)
;p3 : f3 channel - power level for pulse (default)
;p16: f3 channel - power level for CPD/BB decoupling
;sp1: f1 channel - shaped pulse 90 degree
;sp13: f2 channel - shaped pulse 180 degree (adiabatic)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p8 : f2 channel - 180 degree shaped pulse for inversion (adiabatic)
;p11: f1 channel - 90 degree shaped pulse
;p16: homospoil/gradient pulse
;p21: f3 channel - 90 degree high power pulse
```

```
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (2D)           [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O               [30 msec]
;d12: delay for power switching        [20 usec]
;d16: delay for homospoil/gradient recovery
;d26 : 1/(4)YH
;cnst4: = J(YH)
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;ns: 4 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient ratio:  gp 3 : gp 4 : gp 5 : gp6 : gp7
;                    50 : 80 : 30 : 20 : 40

;for z-only gradients:
;gpz3: 50%
;gpz4: 80%
;gpz5: 30%

;use gradient files:
;gpnam3: SMSQ10.100
;gpnam4: SMSQ10.100
;gpnam5: SMSQ10.100

;preprocessor-flags-start
;LABEL_CN: for C-13 and N-15 labelled samples start experiment with
; option -DLABEL_CN (eda: ZGOPTNS)
;1st_increment:
; option -DFIRSTone (eda: ZGOPTNS)
;1st_increment:
; option -DFIRSTtwo (eda: ZGOPTNS)
;2nd_increment:
; option -DSECONDone (eda: ZGOPTNS)
;2nd_increment:
; option -DSECONDto (eda: ZGOPTNS)
;3rd_increment:
; option -DTHIRDone (eda: ZGOPTNS)
;3rd_increment:
; option -DTHIRDto (eda: ZGOPTNS)
;4th_increment:
; option -DFOURTHone (eda: ZGOPTNS)
;4th_increment:
; option -DFOURTHto (eda: ZGOPTNS)
;preprocessor-flags-end

;Processing
;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1
;$Id: hsqc3gpphwg.v 1.12 2012/01/31 17:49:26 ber Exp $
```

### 2.3. AU PROGRAM TO CONCATENATE TOGETHER TWO-DIMENSIONAL SPECTRUMS FROM SERIES OF CYCLED T-JUMP EXPERIMENTS

The AU program must be started under a new experiment (expno) generated in the same folder as the kinetic experiments are stored.

```

int kinetic_step;
char nm1[PATH_MAX];
int ne = 1;
int firstTjump; /*tell where is first Tjump exp*/
int nTjump;
int Tjump_row = 1;
int i, k;
int m = 0;
int expHSQC;
strcpy(nm1, name);
int Tjump_position;

/*FETCHPAR1S("PseudoTD",&td)*/
GETINT("How many of the 2Ds do you want to reconstruct?", kinetic_step);
GETINT("How many T1 increment was measured in Tjumps? real Tjump expno number not
reconsutrcted TD", nTjump);
GETINT("Expno of the first Tjump", firstTjump);
GETINT("Expno of the first HSQC", expHSQC);
GETSTRING("Enter name of 1D series:", nm1);

for( i = 0; i <= kinetic_step - 1; i = i + 1) /*how many 2Ds will be generated*/
{
    for ( k = 0; k <= nTjump - 1 ; k = k + 2) /*step of 2, because of number of scans*/
    {
        Tjump_position = k + firstTjump;
        REXPNO(Tjump_position); /*go to the next Tjump experiment*/
        SETCURDATA;
        /******Generating from two scans an increment***** /

        REXPNO(Tjump_position); /*Reads the first scan*/
        SETCURDATA;
        RSER(Tjump_row,9997,1);
        REXPNO(9997);
        SETCURDATA;

        STOREPAR("DC",1.0);
        STOREPAR("ALPHA",1.0);
        STOREPAR("GAMMA",1.0);
        FT;
        REXPNO(Tjump_position+1); /*Reads the second scan*/
        SETCURDATA;
        RSER(Tjump_row,9998,1);
        REXPNO(9998);
        SETCURDATA;

        STOREPAR("DC",1.0);
        STOREPAR("ALPHA",1.0);
        STOREPAR("GAMMA",1.0);
        FT;
        REXPNO(9998); /*Generates a placeholder expno for add*/
    }
}

```

```

        SETCURDATA;
        IEXPNO;
        REXPNO(9999);
        SETCURDATA;
    STOREPAR("DC",1.0);
        STOREPAR("ALPHA",1.0);
        STOREPAR("GAMMA",1.0);
        DATASET2(name, expno-1, procno, disk, user); /*Adds together the two
FID read from 2D, generates a copy and clears placeholder*/
        DATASET3(name, expno-2, procno, disk, user);
        ADDFID;
        REXPNO(9999);
        SETCURDATA;
        FT;
        WRA(10000);
        DELETEEXPNO(name, expno, disk, user);
        /*REXPNO(10000);
        SETCURDATA;*/

        /*****Copying FID to SER*****/
        REXPNO(expHSQC); /*go to the 2D into which WSER will write the FID*/
        SETCURDATA;
        m = k / 2 ;
        WSER(m + 1, nm1, 10000, 1, disk, user); /*puts the previously generated FID into
the proper HSQC SER file*/

        /*****/
    }

    REXPNO(expHSQC);
    SETCURDATA;
    IEXPNO;
    expHSQC = expHSQC + 1;
    REXPNO(expHSQC);
    SETCURDATA;
    Tjump_row = Tjump_row + 1; /*basically this will set RSER to the next kinetic step*/
}

QUIT

```



## DECLARATIONS

Except where stated otherwise by reference or acknowledgment, the work presented was generated by myself under the supervision of my advisors during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research:

### **Declaration of contributions**

Chapter III section 2.1 Dr. Dean-Paulos Klötzner and Yvonne Becker, Institute for Organic Chemistry and Chemical Biology, developed a synthesis route and synthesized the caged DNA strand used in the NMR experimental studies. I designed, executed and analysed shown experiments.

Chapter III section 3.6 Dr. Patrick Seyfried, Institute for Organic Chemistry and Chemical Biology, developed a synthesis route and synthesized the caged DNA strand for medium throughput screening. Patrick Seyfried executed and analysed shown experiments. The data was used for the discussion.

Chapter III section 3.7, Figure 32, Figure 33, Figure 34, Figure 35 Marcel Heinz, MPI of Biophysics Department of Theoretical Biophysics, performed MD simulations and made comparison between simulation results to NMR spectroscopic data. I measured and analysed experimental NMR data.

Chapter III section 3.5 Setting up structure calculation was done by Dr. Hendrik R. A. Jonker, Institute for Organic Chemistry and Chemical Biology. I provided NOE constraint list and analysed the results of the structure calculation.

### **Declaration of published data**

The following parts of the thesis have been previously published:

Figure 7, Figure 10, Figure 11, Figure 12, Figure 13, Figure 14, Figure 15, Figure 16, Figure 17, Figure 19, Figure 21 have been published in Pintér, G.; Schwalbe, H. Refolding of Cold-Denatured Barstar Induced by Radio-Frequency Heating: A New Method to Study Protein Folding by Real-Time NMR Spectroscopy. *Angewandte Chemie International Edition* n/a (n/a). doi: 10.1002/anie.202006945.

Figure 24, Figure 25, Figure 27, Figure 31, Figure 32, Figure 33, Figure 34, Figure 35 have been published in Seyfried, P.; Heinz, M.; Pintér, G.; Klötzner, D.-P.; Becker, Y.; Bolte, M.; Jonker, H. R. A.; Stelzl, L. S.; Hummer, G.; Schwalbe, H.; Heckel, A. Optimal Destabilization of DNA Double Strands by Single-Nucleobase Caging. *Chemistry – A European Journal* **2018**, *24* (66), 17568–17576.

### **Declaration of figures obtained from published work**

The following figures or data part of this thesis have been obtained, are redrawn, adapted or modified from published work in accordance with copyright permissions or granted licenses by the respective copyright holders (licensors):

Figure 1 Adapted from Figure 2 in Kumar, A.; Balbach, J. Real-Time Protein NMR Spectroscopy and Investigation of Assisted Protein Folding. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2015**, *1850* (10), 1965–1972. (Permission from Elsevier with licence number 4914081272499)

Figure 2 Reproduced from Figure 2 in Rinnenthal, J.; Wagner, D.; Marquardsen, T.; Krahn, A.; Engelke, F.; Schwalbe, H. A Temperature-Jump NMR Probe Setup Using Rf Heating Optimized for the Analysis of Temperature-Induced Biomacromolecular Kinetic Processes. *Journal of Magnetic Resonance* **2015**, *251*, 84–93. (Permission from Elsevier with licence number 4914090937285)

Figure 4 Adapted with permission from Scheme I Schreiber, G.; Fersht, A. R. The Refolding of Cis- and Trans-Peptidylprolyl Isomers of Barstar. *Biochemistry* **1993**, *32* (41), 11195–11203. Copyright (1993) American Chemical Society

Figure 5 Redrawn based on Scheme I in Shastry, M. C. R.; Udgaonkar, J. B. The Folding Mechanism of Barstar: Evidence for Multiple Pathways and Multiple Intermediates. *Journal of Molecular Biology* **1995**, *247* (5), 1013–1027. (Permission from Elsevier with licence number 4914091167305)

Figure 6 a) Adapted from Figure 4 in Rinnenthal, J.; Wagner, D.; Marquardsen, T.; Krahn, A.; Engelke, F.; Schwalbe, H. A Temperature-Jump NMR Probe Setup Using Rf Heating Optimized for the Analysis of Temperature-Induced Biomacromolecular Kinetic Processes. *Journal of Magnetic Resonance* **2015**, *251*, 84–93. (Permission from Elsevier with licence number 4914090937285)

Figure 22 Adapted from Scheme I in Liu, Q.; Deiters, A. Optochemical Control of Deoxyoligonucleotide Function via a Nucleobase-Caging Approach. *Acc. Chem. Res.* **2014**, *47* (1), 45–55. Copyright (2013) American Chemical Society



## PUBLICATIONS

**G. Pintér**, H. Schwalbe, Refolding of cold denatured barstar induced by radio frequency heating – new method to study protein folding by real-time NMR spectroscopy, *Angewandte Chemie International Edition* n.d., n/a, DOI 10.1002/anie.202006945.

M. Novakovic, G. L. Olsen, **G. Pintér**, D. Hymon, B. Fürtig, H. Schwalbe, L. Frydman, A 300-fold enhancement of imino nucleic acid resonances by hyperpolarized water provides a new window for probing RNA refolding by 1D and 2D NMR, *PNAS* 2020, 117, 2449–2455.

P. Seyfried, M. Heinz, **G. Pintér**, D.-P. Klötzner, Y. Becker, M. Bolte, H. R. A. Jonker, L. S. Stelzl, G. Hummer, H. Schwalbe, A. Heckel, Optimal Destabilization of DNA Double Strands by Single-Nucleobase Caging, *Chemistry – A European Journal* 2018, 24, 17568–17576.

**G. Pintér**, H. Schwalbe, Unprecedented Carbon Signal Enhancement in Liquid-State NMR Spectroscopy, *Angewandte Chemie International Edition* 2017, 56, 8332–8334.

E. Balint, E. Fazekas, **G. Pinter**, A. Szollosy, T. Holczbauer, M. Czugler, L. Drahos, T. Kortve, G. Keglevich, Synthesis and Utilization of the Bis(>P(O)CH<sub>2</sub>)amine Derivatives Obtained by the Double Kabachnik–Fields Reaction with Cyclohexylamine; Quantum Chemical and X-Ray Study of the Related Bidentate Chelate Platinum Complexes, *Current Organic Chemistry* 2012, 16, 547–554.

## CONFERENCE CONTRIBUTIONS

**Oral presentation** at the BMRZ Workshop, *Rational Design of New Photo-Cages*, Bad Homburg, June 2017

**Poster presentation** at the EUROMAR conference, *NMR characterization of new modified nitro-phenyl type photocaged DNA*, Warsaw, July 2017

**Poster presentation** at the Design and Light Control conference, *NMR characterization of new modified nitro-phenyl type photocaged DNA*, Niedernberg, August 2017

**Poster presentation** at the 39th FGMR Annual Discussion Meeting, *The use of organic solvents in temperature jump probehead*, Bayreuth, September 2017

**Oral presentation** at the BMRZ Workshop, *Study of the folding of the cold denatured barstar by temperature jump NMR spectroscopy*, Bad Homburg, August 2019

**Poster presentation** at the EUROMAR conference, *Study of the folding of the cold denatured barstar by temperature jump NMR spectroscopy*, Berlin, August 2019



# CURRICULUM VITAE

## Personal information

Name György Pintér  
Date of birth October 23<sup>rd</sup>, 1987  
Place of birth Szeged, HUNGARY  
Nationality Hungarian



## Education

11/2014-10/2020 PhD research in chemistry  
Goethe University Frankfurt am Main, Germany  
**Supervisor:** Prof. Dr. Harald Schwalbe  
**Thesis title:** Optimisation and development of NMR spectroscopic methods for the investigation of biomolecular dynamics

08/2013-08/2014 Research Assistant, NMR Research Laboratory  
Research Centre for Natural Sciences, Hungary

09/2011-07/2013 Budapest University of Technology and Economics,  
Hungary  
MSc in Chemical Engineering  
Specialization: analytical chemistry

09/2006-07/2011 Budapest University of Technology and Economics,  
Hungary  
BSc in Chemical Engineering  
Specialization: analytical chemistry

09/2002-06/2006 SZTE Gyakorló Gimnázium,  
Hungary (Abitur 2006, Szeged)

## List of academic teachers:

Dr. Edit Székely, Dr. Zoltán Hórvölgyi, Dr. Miklós Kubinyi, Dr. Béla Pukánszky, Dr. András Szilágyi, Dr. András György Deák, Dr. Péter Mizsey, Dr. György Keglevich, Dr. Ferenc Faigl, Dr. Gábor Hornyánszky, Dr. Péter Huszthy, Dr. József Nagy, Dr. Ervin Róbert Gyurcsányi, Dr. József Balla, Dr. László Bezur, Dr. Béla Koczka, Dr. Ilona Kovács, Dr. János Madarász, Dr. György Pokol, Dr. András Simon, Dr. Áron Szöllősy, Dr. Janisz Sztatisz, Dr. Tamás Veszprémi, Dr. Béla Nagy, Dr. Lukenics Jánosné. Dr. Zsolt Meszéna, Dr. György Pátzay, Dr. László Nyulászi, Dr. Márta Lángné Lázi, Dr. Mihály Szabó, Dr. Mihály Kállay, Dr. Ferenc Márkus, Dr. György Horvai. Dr. Béla Pukánszky, Dr. Gábor Bajnóczy, Dr. János Tóth, Dr. Livia Sarkadi, Dr. Zoltán Noszticzius, Dr. László Poppe, Dr. Antal Tungler, Dr. Jenő Fekete, Dr. Sándor Kemény