

Paramagnetic Tagging of Oligonucleotides
for Structure Determination using
NMR-Spectroscopy

Dissertation
for the award of the degree
Doctor rerum naturalium
of the Georg-August University Göttingen

within the doctoral program *Physics of Biological and Complex Systems*
of the Georg-August University School of Science (GAUSS)

submitted by
Sebastian Täubert

from Rinteln

Göttingen 2014

Thesis Committee

Prof. Dr. Christian Griesinger

NMR based Structural Biology, Max Planck Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Ulf Diederichsen

Institute for Organic and Biomolecular Chemistry, University of Göttingen

Dr. Claudia Höbartner

Nucleic Acid Chemistry, Max Planck Institute for Biophysical Chemistry, Göttingen

Members of the Examination Board

Referee:

Prof. Dr. Christian Griesinger

NMR based Structural Biology, Max Planck Institute for Biophysical Chemistry, Göttingen

2nd Referee:

Prof. Dr. Ulf Diederichsen

Institute for Organic and Biomolecular Chemistry, University of Göttingen

Further Members of the Examination Board

Prof. Dr. Claudia Steinem

Institute for Organic and Biomolecular Chemistry, University of Göttingen

Prof. Dr. Bert de Groot

Computational Biomolecular Dynamics, Max Planck Institute for Biophysical Chemistry, Göttingen

Dr. Claudia Höbartner

Nucleic Acid Chemistry, Max Planck Institute for Biophysical Chemistry, Göttingen

Dr. Martin Kollmar

Systems Biology of Motor Proteins, Max Planck Institute for Biophysical Chemistry, Göttingen

Date of oral examination: 16.01.2015

Affidavit

I hereby declare that I wrote this thesis *Paramagnetic Tagging of Oligonucleotides for Structure Determination using NMR-Spectroscopy* on my own and without the use of any other than the cited sources and tools.

.....
Place, Date

.....
Signature

für meine Familie

Acknowledgments

First of all, I would like to thank Prof. Christian Griesinger for the opportunity to work in his department on this very interesting topic. I am very thankful for his ongoing support, the several constructive scientific discussions and the always open door.

I would like to thank Dr. Claudia Höbartner and Prof. Ulf Diederichsen for the participation in my thesis committee. These meetings always provided me with valuable new perspectives and ideas to find answers to the scientific questions.

Furthermore, I thank Prof. Claudia Steinem, Prof. Bert de Groot and Dr. Martin Kollmar for the participation in my oral examination.

I would particularly like to thank Dr. Andrei Leonov for the mentoring of the chemical synthesis. His constant support and his many useful suggestions were crucial for the success of my thesis. I would also like to thank Dr. Mitcheel Maestre Martinez for his support in the data analysis and the performed structure calculations. Thanks to Dr. Florian Siepel for the productive collaboration and the many useful and open discussions.

I thank Kerstin Overkamp and Gerhard Wolf for the HPLC purification and the mass spectrometry. I would like to thank Claudia Schwiegk for the RNA sample preparation and Dr. Dirk Bockelmann for the support with the computers and software.

I would like to thank all the proofreaders of this work and especially Eibe Dücker for his patience and the many useful (and necessary!) suggestions.

I thank Dr. Manuel Schmidt, Dr. Philip Lottmann, Eibe Dücker, Sergey Ryazanov, Benjamin Schomburg, Dr. Fabian Peters, Ruth Linder and Dr. Nina Schützenmeister for their help, the many discussions and the great atmosphere in the lab. I would also like to thank Nele Milsch for the trust to write her bachelor thesis with my supervision. Many thanks to all the members of the department for the friendly atmosphere.

I would especially like to thank all of my friends and, of course, the inner circle for their support and all the good times we have had together!

Meinen Eltern und Großeltern möchte ich von ganzem Herzen für ihre großartige Unterstützung während meiner ganzen Studienzeit danken. Ohne euer großes Vertrauen und euren starken Rückhalt wäre diese Arbeit nicht möglich gewesen.

Summary

Structure determination is one of the most fundamental fields of research, as it provides insights in biological systems and their mechanisms. Therein, NMR spectroscopy is an essential tool for investigations under physiological conditions and the development of new techniques, which enable the detection of more detailed information, is an ongoing topic of research.

By using paramagnetic tagging of biomolecules, highly valuable NMR parameters, e.g. residual dipolar couplings (RDCs) and pseudocontact shifts (PCSs), can be determined, which provide angle and distance information about the target molecule. Therefore, several lanthanide-binding tags have been developed and successfully attached to proteins and oligosaccharides. Due to the paramagnetic properties of the lanthanide ions, an alignment of the target molecule is induced, which facilitates the determination of RDCs. Additionally, the NMR signals are shifted due to a dipole-dipole interaction between the nuclei and the lanthanide ion (PCS). Both of these effects can not be detected in conventional NMR spectroscopy, due to the isotropic distribution of orientations of a molecule in solution.

A successful transfer of the protein tagging method to oligonucleotides has of yet not been reported, although DNA and RNA are two of the most important groups of biomolecules. Based on the recently developed Cys-Ph-TAHA tag, a reliable protocol for the determination of paramagnetic effects in a DNA molecule is presented in this work. To achieve this, a modified nucleobase was synthesized, which provides a suitable binding site for the Cys-Ph-TAHA tag. A tagging and purification protocol was established, by which two paramagnetic samples (thulium and terbium) and a diamagnetic reference sample (lutetium) were prepared (see Fig.0.1).

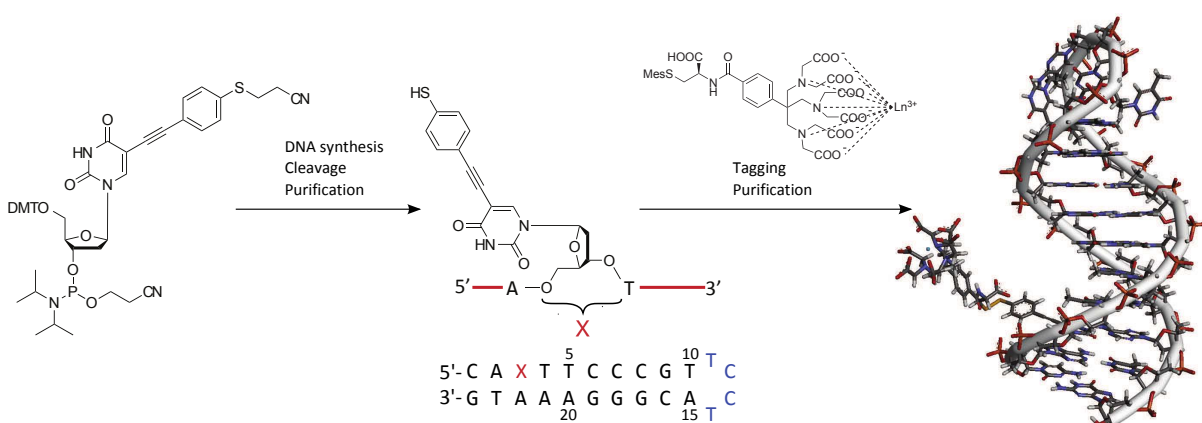


Figure 0.1: Established approach for the preparation of paramagnetically tagged DNA.

High resolution NMR spectroscopy revealed paramagnetically-induced PCSs and RDCs in the tagged oligonucleotide. The data evaluation proved a high quality of the determined

PCSs for both samples. As $^{13}\text{C}/^{15}\text{N}$ labeling of synthesized DNA strands is unreasonable, due to the expensive starting materials and the low overall yield of the DNA synthesizer, all NMR spectra were performed using samples with natural abundance. Consequently, the signal-to-noise ratio of the NMR signals was relatively low and together with the paramagnetic relaxation enhancement, the determined RDCs have a significant error. Nevertheless, a suitable ensemble model was created, by which the combined PCS and RDC data was satisfactorily described.

In conclusion, the technique of paramagnetic tagging was successfully transferred to an oligonucleotide. To do this, a reliable strategy for the incorporation of a binding site and sufficient tagging and purification protocols were established. This approach was verified by the determination and evaluation of PCSs and RDCs in a DNA strand.

Zusammenfassung

Strukturaufklärung gehört zu den wichtigsten Gebieten der Grundlagenforschung, da sie direkte Einblicke in biologische Systeme und ihre Mechanismen liefert. Der NMR Spektroskopie kommt dabei eine besondere Bedeutung zu, denn sie ermöglicht Forschung unter physiologischen Bedingungen. Dementsprechend ist die Entwicklung neuer Techniken zur Verbesserung dieser Methode weiterhin ein zentrales Forschungsgebiet.

Paramagnetische Markierung von Biomolekülen ermöglicht die Bestimmung von NMR Parametern, wie z.B. residuale dipolare Restkopplungen (RDCs) oder Pseudokontaktverschiebungen (PCSs), die für die Strukturaufklärung wertvolle Winkel- und Abstandsinformationen über das Zielmolekül beinhalten. In diesem Zusammenhang wurden Lanthanoidionen-koordinierende Tags entwickelt und erfolgreich an Proteinen angebracht. Durch die paramagnetischen Eigenschaften der Lanthanoidionen wird eine partielle Ausrichtung des Zielmoleküls im Magnetfeld des NMR Spektrometers induziert und somit das Messen residueller dipolarer Kopplungen ermöglicht. Zusätzlich werden die NMR Signale durch eine Dipol-Dipol-Wechselwirkung zwischen dem Lanthanoidion und den Kernen verschoben (PCS). In der konventionellen NMR Spektroskopie werden diese Effekte, aufgrund der Brownschen Molekularbewegung und dem Fehlen eines Metallions, nicht beobachtet.

In der Fachliteratur ist ein Transfer dieser Methode auf Oligonukleotide nicht bekannt, obwohl DNA und RNA zu den wichtigsten Biomolekülen überhaupt zählen. In dieser Arbeit wurde mit Hilfe des kürzlich entwickelten Cys-Ph-TAHA Tags ein Protokoll zur Bestimmung von paramagnetischen Effekten in der DNA entwickelt. Dafür wurde eine modifizierte Nukleobase synthetisiert, die eine passende Bindungsstelle für den Tag aufweist. Mit der neu entwickelten Methode wurden zwei paramagnetische und eine diamagnetische Referenzprobe hergestellt.

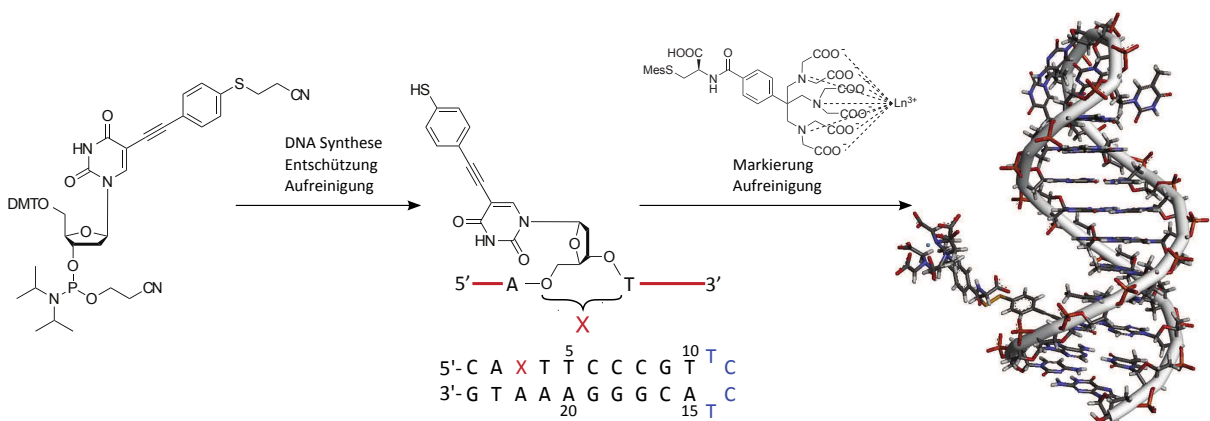


Figure 0.2: Entwickelte Methode zur Herstellung von paramagnetisch markierter DNA.

Mittels hochauflösender NMR Spektroskopie konnten paramagnetisch-induzierte PCSs und RDCs gemessen werden. Die Auswertung zeigte eine hohe Qualität der gemessenen

PCSs in beiden paramagnetischen Proben. Die RDCs wiesen einen signifikanten Fehler auf. Die in der NMR Spektroskopie übliche Isotopenmarkierung ($^{13}\text{C}/^{15}\text{N}$) ist bei im DNA-Synthesizer hergestellten Oligonukleotiden auf Grund der teuren Ausgangsmaterialien nicht möglich, sodass die hergestellten NMR Proben eine natürliche Isotopenhäufigkeit aufwiesen. In den NMR Spektren zur Bestimmung der RDCs ist damit das Verhältnis von Signal-zu-Rausch relativ niedrig, was zusammen mit der paramagnetischen Relaxationsverstärkung zu einem größeren Messfehler führt. Dennoch konnten die erhaltenen paramagnetischen Daten mit einem Ensemblemodell beschrieben werden.

In der vorliegenden Arbeit wurde die Methode der paramagnetischen Markierung erfolgreich auf die Stoffklasse der Oligonukleotide übertragen. Dabei wurde ein reproduzierbares Protokoll entwickelt, mit dem eine Bindungsstelle in einen DNA Strang eingebaut und das Zielmolekül anschließend mit dem Cys-Ph-TAHA Tag markiert wurde. Die erfolgreiche Anwendung der Methode konnte durch die erhaltenen paramagnetischen Messwerte von hoher Qualität verifiziert werden.

Contents

1	Introduction	1
1.1	Paramagnetic Tags	3
1.2	Modifications of Oligonucleotides	8
1.3	Objective of this Work	12
2	Theory	15
2.1	Residual Dipolar Couplings	16
2.2	Pseudocontact Shifts	20
2.3	Paramagnetic Relaxation Enhancement	21
3	Synthesis	23
3.1	Tagging Strategies of Oligonucleotides	23
3.2	Disulfide Protected Binding Site	27
3.3	Cyanoethyl Protected Binding Site	29
3.4	Cleavage of Sulfur Modified DNA Strand on Solid Support and Tagging with Cys-Ph-TAHA	31
3.5	Tagging on Phosphorothioate Oligonucleotides	32
3.6	Modification of Cys-Ph-TAHA	36
4	NMR Spectroscopy of Paramagnetically Tagged DNA	41
4.1	Sample Preparation	41
4.2	Measurement of Paramagnetically Induced NMR Parameters	43
4.2.1	Pseudocontact Shifts	45
4.2.2	Residual Dipolar Couplings	50
4.3	Data Evaluation and Discussion	60
4.3.1	Evaluation of the PCS and RDC Data Without a Structural Model of Cys-Ph-TAHA	60
4.3.2	Evaluation of the Combined PCS and RDC Data Using a Structural Model of Cys-Ph-TAHA	64
5	Summary and Outlook	69

6	Materials and Methods	73
6.1	Experimental Details	76
7	Bibliography	105
8	List of Figures	115
9	List of Abbreviations	119
10	Appendix	123

1 Introduction

Nuclear magnetic resonance spectroscopy (NMR) is one of the central tools to reveal structures and dynamics in biomolecular science. The large scope of the technique, which covers the range from conformational analysis of small chemical compounds up to investigation of domain dynamics, is unique. On the one hand, NMR spectroscopy has become an irreplaceable routine to, for example, monitor chemical reactions and on the other hand, new spectrometers with stronger magnetic fields are being developed, which will enable more accurate insights into the structures of biomolecules than currently possible. Compared to other structure determination techniques, e.g. crystallography, the opportunity to approach structural biology under native conditions is the major advantage of solution-state NMR spectroscopy.

The development of new methods, which enable the measurement of structural restraints, is an ongoing task to improve the accuracy of structure determination. Over the last 20 years, the determination of NMR parameters, which are usually not achieved in conventional experiments without either changing the solution to an anisotropic one (alignment media) or to modify the molecule by paramagnetic tagging, came into the focus of research [Lee1983] [LaMar1993] [Tolman1995] [Bertini2001] [Otting2008] [Bothe2011]. As one of these, residual dipolar couplings (RDCs) have shown to be an excellent tool for structure determination as they provide angle and distance information about a target molecule. RDCs can only be determined in samples in which the biomolecules are aligned within the magnetic field of the NMR spectrometer such that the isotropic distribution of orientations is changed to an anisotropic one. By comparing the coupling constants of an aligned molecule with a non aligned reference sample, RDCs can be measured as an additional contribution to the scalar couplings. For structure determination, the magnitude of the alignment, the distance between the regarded nuclei and the angle between its inter-nuclear vector and the external magnetic field are the important factors that alter the size of the measured RDC. Taking, for example, a protein with two domains that are connected by a flexible linker, in which one domain is strongly aligned, the average RDCs of this domain is expected to be significantly larger than the average RDCs of the less strongly aligned domain [Russo2013] (for RNA application [Zhang2006]). Therefore, RDCs are global parameters that yield information about the whole molecule. In 1995, cyanometmyoglobin was the first protein to be investigated with regard to RDCs,

as it aligns in the external magnetic field due to its inherent high magnetic susceptibility, resulting from the iron ion it contains [Tolman1995].

To investigate the conformations of small molecules using RDCs, different strategies were developed, which generate an alignment by spatial restriction using external media. For this purpose, phages [Hansen1998], gels [Tycko2000] [Schmidt2012] or liquid crystals [Bax1997] were employed in order to achieve an anisotropic orientation distribution and to reveal the RDCs.

Today, the artificial introduction of unpaired electrons into proteins has enormously advanced the number of accessible NMR parameters which can provide structural information. Due to their anisotropic magnetic susceptibility, an alignment of the target molecule can be induced, which then enables the measurement of RDCs. The direct interaction of an unpaired electron with a nuclear spin results in paramagnetic relaxation enhancement (PRE) and pseudocontact shifts (PCS). Both effects are highly distance dependent and therefore, the resonances close in space to the newly created paramagnetic center experience line-broadening or can not be observed at all (PRE). The unpaired electron also induces an additional magnetic field, resulting in different chemical shifts for the resonances close to the electron (PCS). Overall, NMR spectroscopy becomes more difficult by the influence of these two effects, but the benefits for structure determination have encouraged researchers to establish different strategies for the introduction of unpaired electrons.

Lanthanide ions have proven to be an excellent tool for this purpose, owing to their magnetic susceptibility properties and PRE radii [Otting2008]. Consequently, lanthanide ions were coordinated to modified metal binding proteins, e.g. calmodulin. To increase the affinity to lanthanide ions, additional negatively charged amino acids were introduced into one of the calcium binding sites [Bertini2003]. A more versatile strategy uses the attachment of small lanthanide coordinating tags, which is a key point of this work and therefore its development is summarized in the following section.

Furthermore, paramagnetic ions can be coordinated by specific protein sequences with a suitable binding affinity [Woehnert2003] [Nitz2003] [Barthelmes2010] or the zinc finger motif [Gaponenko2000], which can also be incorporated into the target molecule.

Even though the first description of the DNA double helix is more than 60 years old, the huge diversity and dynamics of oligonucleotides, especially RNA, is still a challenging task in structure determination [Watson1953]. Up until now, the measurement of RDCs in oligonucleotides remains difficult and a reliable method, which can be transferred to different RNA strands, has not yet been reported. In the following sections, the different paramagnetic tags (reviewed by Koehler *et al.*) and modifications of oligonucleotides that enable the introduction of tags into oligonucleotides are summarized [Koehler2011].

1.1 Paramagnetic Tags

In this work, the paramagnetic center was introduced into the molecules of interest using the Cys-Ph-TAHA (cysteine-phenyl-triaminohexaacetic acid) tag **1** shown in Fig.1.1, which was developed in our group and published in 2011 by Peters *et al.*. It was synthesized based on the TAHA complex [Viguier2001] and the previously published EDTA (ethylenediamine tetraacetic acid) tags. This tag features nine coordination sites, which saturate the coordination sphere of lanthanide ions, giving a metal complex with a high sterical requirement. A terbium loaded tag was successfully attached to ubiquitin, yielding PCS of 2 ppm and a max. NH RDC of 17.8 Hz for one specific mutant (T12C) at 900 MHz. In the NMR spectra, paramagnetically Cys-Ph-TAHA-tagged ubiquitin has shown one unambiguous signal set.

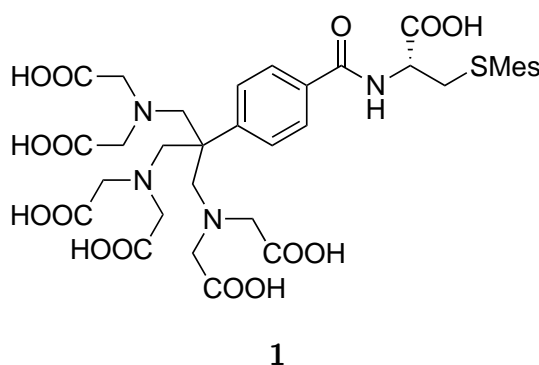


Figure 1.1: Cys-Ph-TAHA tag **1**.

The tag was also employed in studies of a lac repressor complex with DNA, allowing the conclusion that its lanthanide affinity is sufficiently high for an application on oligonucleotides, regarding the high cation affinity of the phosphate backbone [Peters2011]. Furthermore, its convenient synthesis enables tag modifications that discard the flexible cysteine linker. A promising approach is presented in section 3.6.

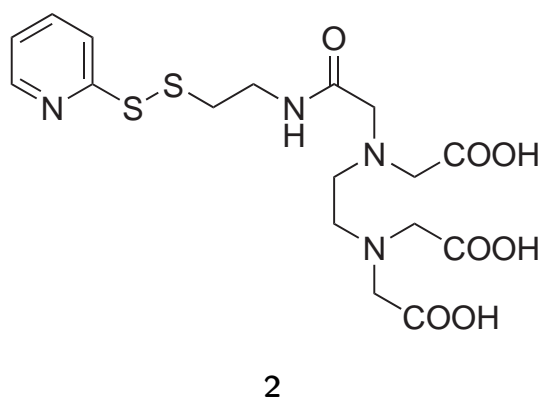


Figure 1.2: EDTA-based tag **2**.

In 2002, the (2-pyridylthio-)cysteaminy-EDTA tag **2** was published by Dvoretzky *et al.*, followed by a second generation of EDTA-based tags from 2004 to 2006 (see Fig.1.2 and 1.3) [Dvoretzky2002]. The disulfide linkage to the target molecule was established via an activated thiol moiety of the tag and a free thiol group of a cysteine residue in the protein. The first EDTA-tag **2** has already shown PCSs and RDCs, but its pseudo-asymmetric nitrogen atom, which connects the linker to the EDTA fragment, has produced two distinct signal sets in NMR spectra [Ikegami2004]. In the second EDTA tag generation, the linker is consequently connected to the C1 ethylene carbon with a defined stereochemistry [Ikegami2004] [Leonov2005]. Subsequently, an improvement of the flexible linker, which was replaced by a rigid phenyl ring, was published in 2006 (see Fig.1.3, **3** and **4**) [Haberz2006]. This generation of EDTA-tags was, for example, applied to Apo-Calmodulin, giving NH RDCs up to 8 Hz at a spectrometer frequency of 800 MHz. Furthermore, EDTA-tag **3** was successfully attached to DNA strands, yet its application on oligonucleotides was not further pursued [Woeltjen2009].

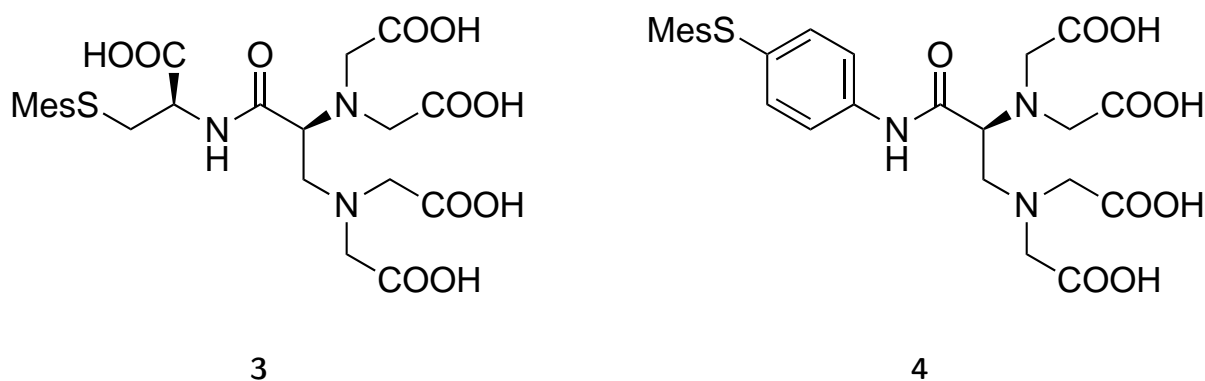
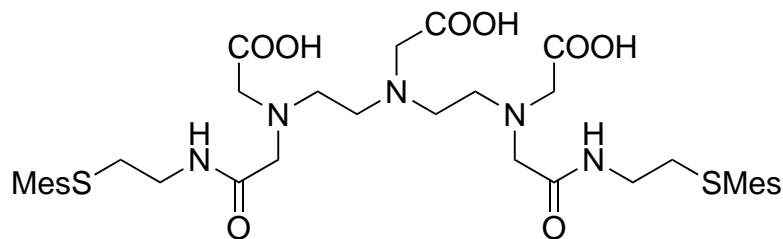


Figure 1.3: Second generation of EDTA-based tags with a flexible cysteine linker **3** or a more rigid phenyl ring modification **4**.

Apart from the EDTA or TAHA based tags, there has been a second group of tags, which was successfully used for paramagnetic tagging of proteins. In 2004, the *caged lanthanide NMR probe* (CLaNP-1) **5** was published by Prudencio *et al.* (see Fig.1.4), followed by several improvements up to 2008 [Prudencio2004] [Keizers2007] [Keizers2008]. In these tags, the lanthanide is either coordinated by a diethylene triaminepentaacetic acid (DTPA) or a 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) ligand with high stabilities ($K_D < 10^{-20}$ mol/L) and extensive sterical requirements.

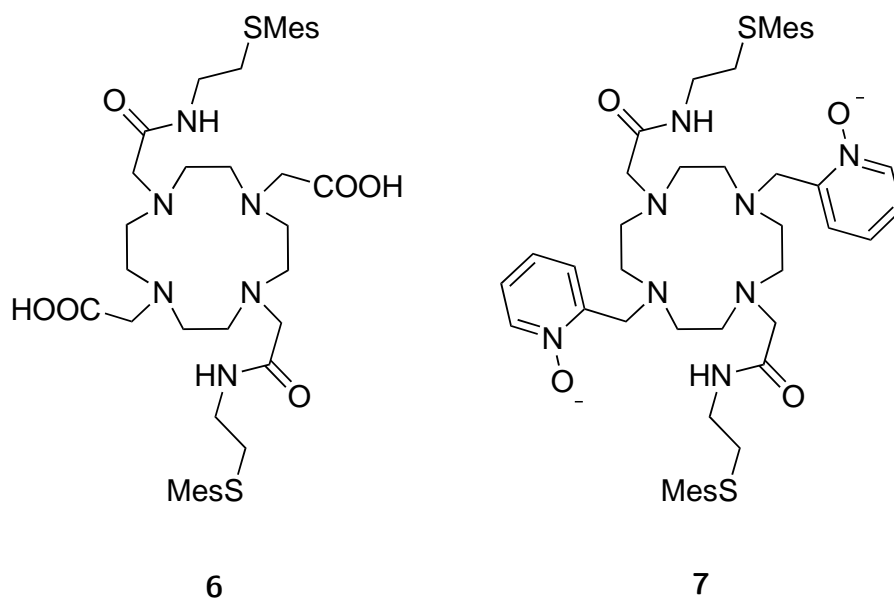
Several different stereoisomers were generated by tagging with CLaNP-1 and subsequent lanthanide coordination, resulting in five signal sets in the final NMR spectra. The DOTA tags exhibited the same stereochemical problem, which was subsequently solved by the introduction of different substituents (see Fig.1.5). Due to the sterical hindrance, the stereoisomers are not equally populated, resulting in two substantially populated isomers



5

Figure 1.4: CLaNP-1 5.

for CLaNP-3 **6** and only one significantly populated stereoisomer for CLaNP-5.1 **7**. These tags are connected to the target molecule via two disulfide bridges, giving large NH dipolar couplings of 25 Hz at 600 MHz [Keizers2008]. The disadvantages of their application lie in the two required solvent exposed binding sites in the protein and consequently, the necessity of the *a priori* knowledge of the target molecule. CLaNP-5.1 was successfully used for conformational space studies on calmodulin [Dasgupta2011] or on tagged reporter proteins that transfer a partial alignment on a target molecule [Camacho-Zarco2014]. Compared to CLaNP-5.1, the one-armed analogous of the tag was producing significantly smaller RDCs. In 2012, the group of M. Ubbink reported a further variation of the DOTA tag, the CLaNP-7 **8**. Therein, the amino side-chains were modified in order to reduce the charge by 2 units, which resulted in a different alignment tensor compared to its predecessor [Liu2012]. It was used to determine the structure of the putidaredoxin-cytochrome P450cam complex using PCSs, RDCs and PRE data [Hiruma2013].



6

7

Figure 1.5: CLaNP-3 **6** and CLaNP-5.1 **7**.

In 2009, Daniel Haeussinger *et al.* presented the one-armed, methyl-substituted DOTA-M8 tag **9** (see Fig.1.6), which enabled the measurement of large PCSs (>5 ppm) and RDCs (>20 Hz) for dysprosium-tagged ubiquitin at 800 MHz [Haeussinger2009]. Due to its sterical requirement and the short linker, an interaction with the protein surface is presumed, which causes the strong alignment. The synthesis of DOTA-M8 is, however, very sophisticated, as the overall yield of the methylated DOTA fragment is extremely low (9%). Furthermore, a second set of NMR signals with an intensity of 15-20% was observed at room temperature, which increased to about 50% at 50 °C.

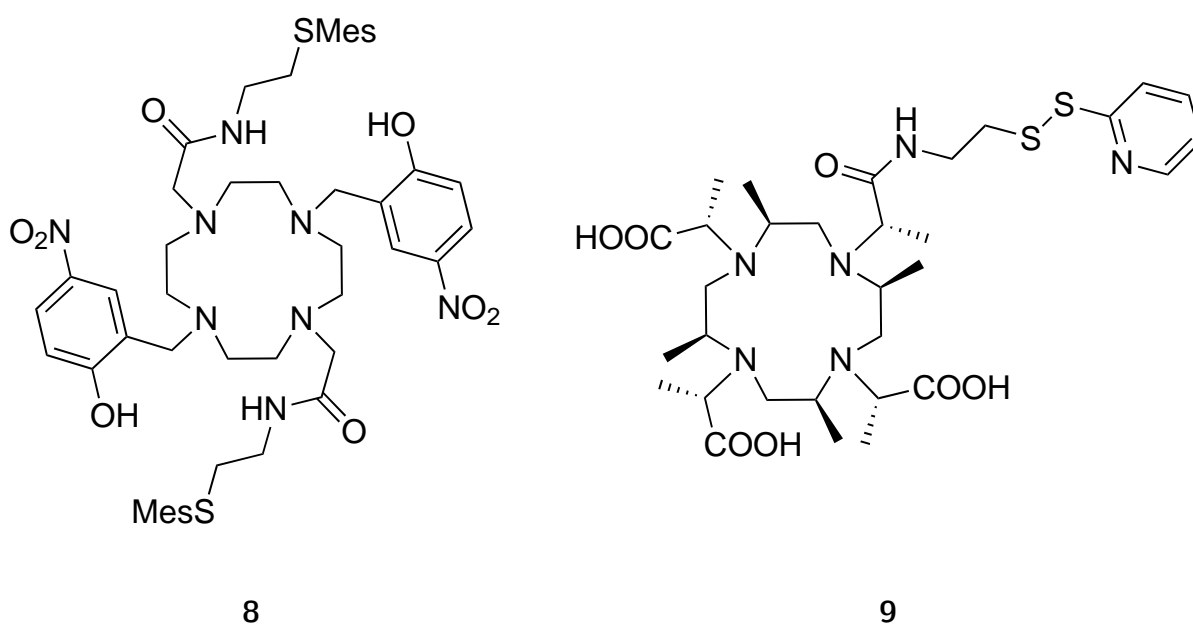


Figure 1.6: CLaNP-7 **8** and DOTA-M8 **9**.

Further variations of the DOTA tag were reported by Graham *et al.* and Loh *et al.* [Graham2011] [Loh2013]. Therein, not the DOTA fragment, but the amino side-chains were stereospecifically modified with bulky substituents. In comparison to the previously applied disulfide bridge, this tag can also be attached to proteins via 1,3-dipolar cycloaddition, using a linker with an acetylene moiety. To do this, the target molecule consequently requires a solvent exposed azido-phenylalanine residue.

Compared to the bulky TAHA or DOTA tag, a different approach was reported by the group of G. Otting in 2008. Based on dipicolinic acid (DPA) **10**, small lanthanide binding tags were attached to proteins. Interestingly, the disulfide bridge was generated using Ellman's reagent, which activates the thiol moiety of cysteine [Ellman1959] [Su2008]. The coordination sphere of the lanthanide ion is not completely saturated by one 4MMDPA **11** (see Fig.1.7) alone and therefore the ion is additionally coordinated by carboxy groups of neighboring amino acids. In further improved versions of the tag, the flexible methylene group was removed [Man2010] [Jia2011a] and different substituents of

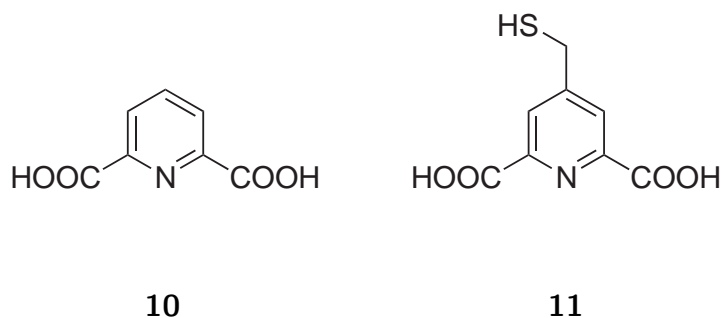


Figure 1.7: DPA **10** and 4MMDPA **11**.

the DPA ligand were investigated [Wei2013]. The iminodiacetic-based lanthanide tag **12** (IDA) was successfully attached to a protein helix, resulting in large PCSs up to 8 ppm (see Fig.1.8) [Swarbrick2011]. Furthermore, various different binding motifs of DPA tags were reported. First, a non-covalent variation was presented in 2010 by Yagi *et al.*, in which the $[\text{Ln}(\text{DPA})_3]^{3-}$ complex was coordinated by two positively charged amino acids of the target protein. Therein, the positive charge must not be compensated by neighboring residues in order to guarantee a sufficient level of tagging. This facilitated PCS and PRE measurements in different proteins [Yagi2010] [Jia2011b]. In a later approach, the connection to the protein was established via a thiol-ene reaction with the 4VPyMTA **13** tag, which possesses an EDTA-like structure, but features an even higher binding affinity for lanthanide ions (see Fig.1.8) [Yang2013].

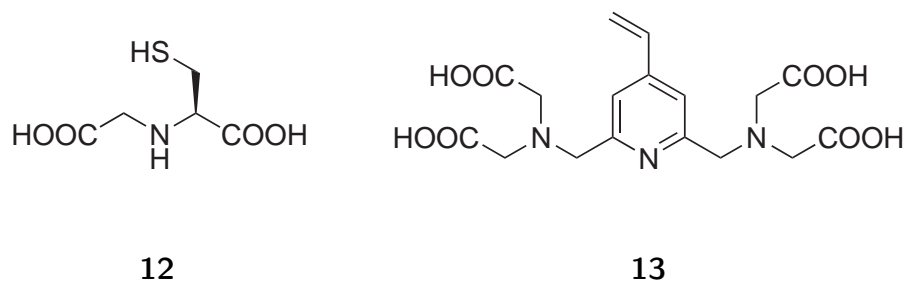
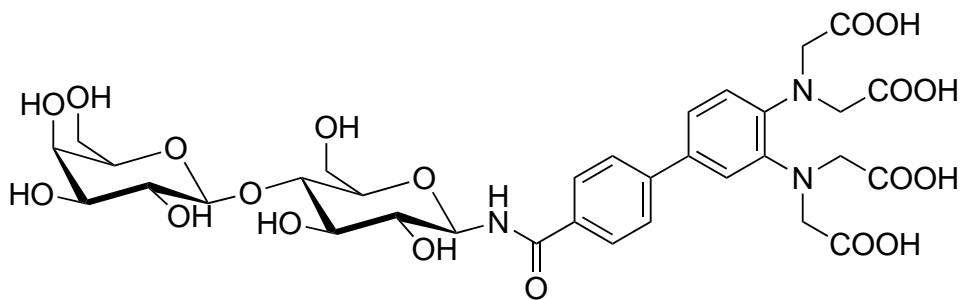


Figure 1.8: IDA-based tag **12** and 4VPyMTA **13**.

In contrast to the classic paramagnetic tags, which are small compared to the studied protein, lanthanide binding sites were also introduced into carbohydrate molecules [Erdelyi2011] [Yamamoto2011] [Canales2013] in order to investigate the conformations of lactose using PCSs and RDCs. In addition, Canales *et al.* employed this approach to measure small PCS on a carbohydrate-protein complex [Canales2014]. Therein, the lanthanide ion was coordinated by an EDTA-like modification that is covalently bound to the disaccharide (see Fig.1.9). This approach demonstrates a possible application for a paramagnetically tagged DNA strand that could be used for studying protein-DNA complexes.



14

Figure 1.9: Modified lactose with a EDTA-like lanthanide coordination site **14**.

Even though there are several paramagnetic tags available that have been used in structural biology of proteins, no application of a tagging method on oligonucleotides has been reported so far. In proteins, binding sites for the disulfide connection occur naturally in form of the free thiol moiety of cysteine. Moreover, they can be conveniently generated by *site-directed mutagenesis*. Regarding the introduction of a binding site, the successful tagging strategies of disaccharides can not be compared to the tagging of proteins or oligonucleotides, as the target molecule is significantly smaller.

1.2 Modifications of Oligonucleotides

In 1995, Kung *et al.* reported on the measurement of RDCs on DNA strands due to the anisotropic magnetic susceptibility of the molecule itself [Kung1995]. This form of self-alignment had already been described by Gayathri for small molecules [Gayathri1982]. Consequently, no modification was required yet the low level of self alignment did not result in a successful application for structure determination of DNA.

In oligonucleotides, however, the introduction of a suitable binding site for paramagnetic tags is more complex than in proteins. In our group, several different tagging strategies for oligonucleotides were investigated. Therein, paramagnetic tags were successfully attached to DNA or RNA strands but unfortunately, an efficient method has not been established so far [Woeltjen2009] [Siepel2013].

E. Woeltjen presented a 24-mer DNA hairpin structure for paramagnetic tagging of oligonucleotides, which is based on a STAT protein (*signal transducer and activator of transcription*) binding sequence [Chen1998] [Neculai2005] [Mao2005]. The DNA hairpin structure with its modification position is shown in Fig.1.10, in which the modification site points out of the major groove of the DNA strand. It was demonstrated that the introduced binding site did not change the native hairpin structure and thus proved to be a suitable target molecule to establish a new tagging strategy. Furthermore, this sequence

region was revealed [Zhang2006] [Zhang2007]. Therein, a twist of up to 110° and a bending of 94° was discovered. In the presented work, one of the helices was extensively elongated with either G-C or A-U base pairs and the corresponding other base pair was ^{13}C and ^{15}N labeled in order to enable efficient NMR measurements. The elongated helices were aligned using Pfl phages, giving large RDCs in the elongated helix and significantly smaller RDCs in the flexible domain (see Fig.1.12). Even though Zhang *et al.* did not present a tagging method to induce an alignment, this application proved the potential of RDCs in oligonucleotides, which can also be achieved by paramagnetic tags. For this purpose, E. Woeltjen has presented an elongated TAR-RNA construct using a DNA:RNA hybrid [Woeltjen2009]. Therein, the target RNA is elongated by only eleven nucleobases and the complementary, paramagnetically tagged oligonucleotide is connected via Watson-Crick base pairing (see Fig.1.13). The advantage of this approach is the possibility of using ^{13}C and ^{15}N labeled RNA samples, as only the attached oligonucleotide is chemically modified and produced by a synthesizer.

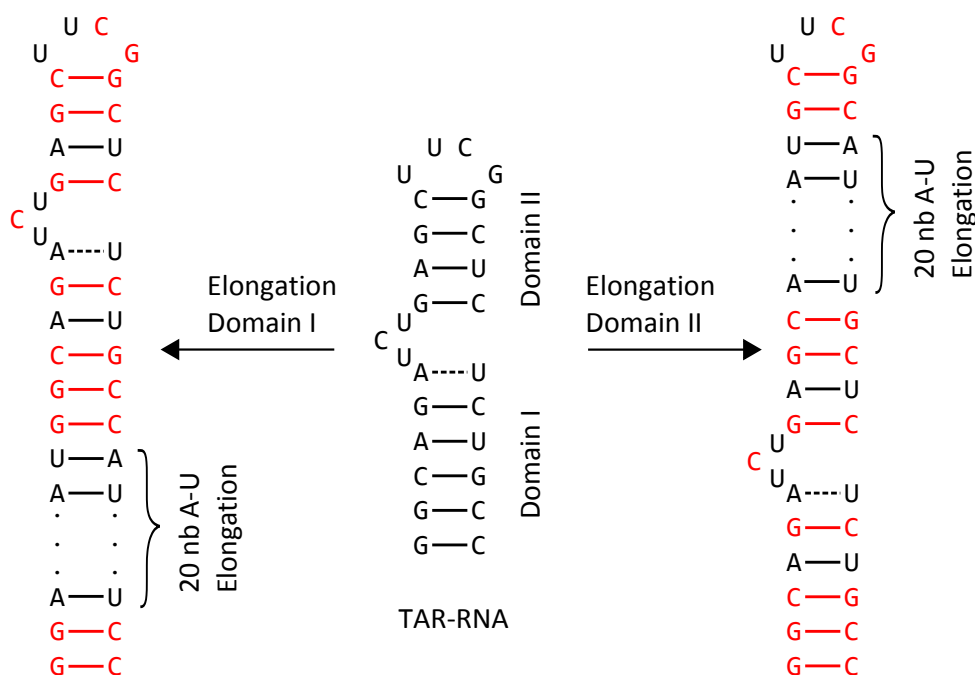
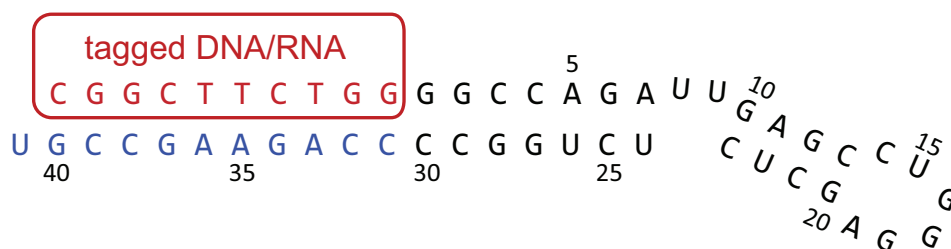


Figure 1.12: Elongations of TAR-RNA in the two different domains that induces alignment in Pfl phages presented by Zhang *et al.*. $^{13}\text{C}/^{15}\text{N}$ labeled nucleotides are shown in red.

Apart from tagging or modification strategies in order to induce a molecular alignment, oligonucleotides were also tagged, for example, with the 2,2,6,6-tetramethylpiperidine 1-oxyl (TEMPO) radical, its analogues or pyrene. In the literature, several successful strategies have been reported for the tagging of nucleobases, e.g. spin label modifications of amino groups [Edwards2001] [Sicoli2010], tagging of 4-thiouridine [Qin2003] or click-chemistry using acetylene modifications [Piton2007] [Ingale2014] (see Fig.1.14).



16

Figure 1.13: Elongated (*blue*) TAR-RNA hairpin structure (*black*) with complementary paramagnetically tagged oligonucleotide (*red*) presented by E. Woeltjen.

Furthermore, the TEMPO radical was introduced via phosphoramidite coupling at the 5' hydroxy group of TAR-RNA for the measurement of PRE in NMR spectroscopy [Wunderlich2013]. It was also applied to oligonucleotide phosphorothioate for EPR studies [Qin2001] [Qin2007] [Esquiaqui2014], a tagging strategy that was transferred to a paramagnetic application in this work (see section 3.5).

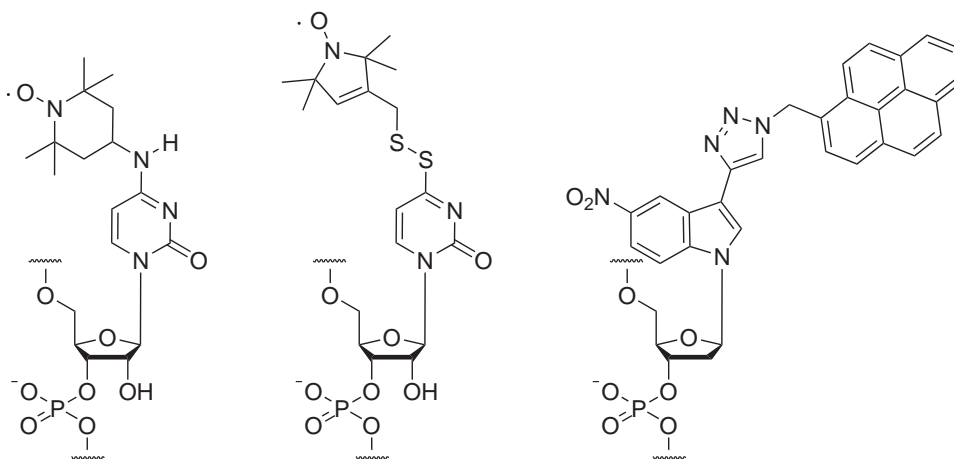


Figure 1.14: Different modifications of nucleobases in DNA or RNA strands.

A different approach to site-directed spin labeling has recently been reported by Babaylova *et al.*, which enables tagging of long RNA sequences that exceed the scope of oligonucleotide synthesizers [Babaylova2014]. The crucial step lies in the introduction of the binding site after the RNA synthesis using a *complementary addressed reagent* at a sequence specified position of the target RNA. An aliphatic amino group is transferred from a modified decamer DNA strand to a specific RNA guanine nucleobase via alkylation of N9. The introduced amino group was then subsequently tagged with a spin label (see Fig.1.15). However, this strategy can not easily be transferred for the purpose of paramagnetic tagging, due to the flexible linker, a possible change of the native structure and,

most importantly, the different tagging reaction. The advantage of this method is the possibility of using ^{13}C and ^{15}N labeled RNA samples and the potential to attach the tag at different positions of large RNA molecules, using the specific complementary DNA sequence.

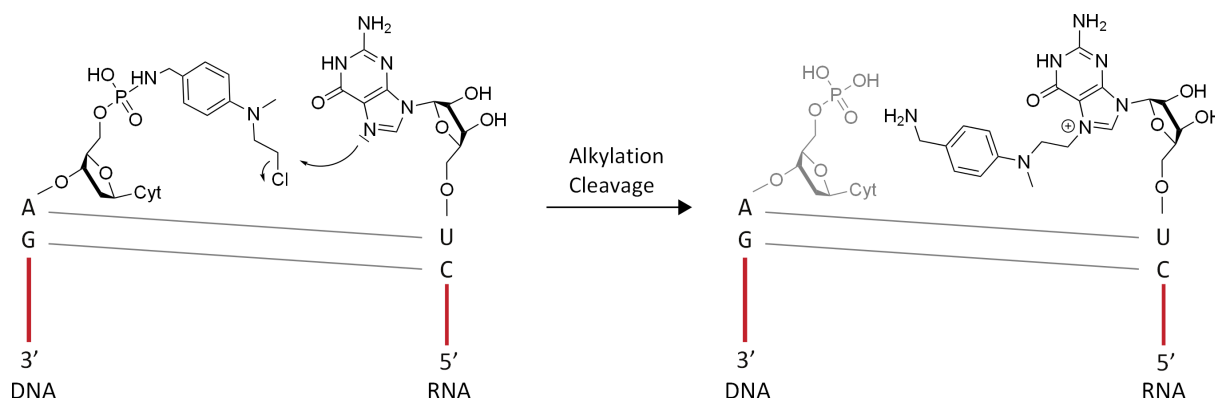


Figure 1.15: Introduction of a binding site using a *complementary addressed reagent* with 10 base pairs (only 2 are illustrated) presented by Babaylova *et al.*. At first, a guanine nucleobase of the target RNA is alkylated, followed by the cleavage of the modification from the reagent.

1.3 Objective of this Work

The development of an efficient strategy for paramagnetic tagging of oligonucleotides and the measurement of PCSs and RDCs in DNA are the major aims of this work. Due to the satisfying results that were obtained for the tagging of ubiquitin using the Cys-Ph-TAHA tag, it was a straightforward idea to employ this method on oligonucleotides. Even though paramagnetic tagging of proteins is an established approach for the measurement of PCSs and RDCs, up until now, no application for oligonucleotides has been published.

At first, the introduction of a suitable binding site for the Cys-Ph-TAHA tag needs to be established (see Fig.1.16). Therefore, it is crucial to develop a reliable protocol with reasonable yields and a high reproducibility. Regarding the different strategies which have been reviewed in the previous section for site-specific spin labeling and paramagnetic tagging, two different approaches are pursued in this work. On the one hand, a new modified nucleobase with a rigid sulfur binding site is to be incorporated into a DNA strand using the phosphoramidite method. To do so, an efficient protection group strategy needs to be found to guarantee a convenient synthetic pathway and satisfactory yields in DNA synthesis. Subsequently, the DNA purification and deprotection protocols, which generate a free thiol moiety, and the tagging reaction have to be established. To enable NMR measurements, traces of lanthanide ions that are coordinated to the backbone have

to be removed in order to avoid substantial line broadening caused by PRE [Siepel2013]. Finally, suitable sample conditions have to be found.

On the other hand, tagging of oligonucleotide phosphorothioate is a promising second strategy, as it does not require a complex chemical synthesis to generate a thiol moiety in DNA or RNA molecules. Compared to the conventional disulfide connection between Cys-Ph-TAHA and the target molecule, tagging at the backbone was performed using thioether linkages. Therefore, the tag needs to be modified to enable a substitution reaction that is comparable to the previously reported procedure [Qin2001] [Qin2007]. In both of the strategies, the Cys-Ph-TAHA tag needs to be coordinated with the lanthanide ion before tagging (*preloading*), as otherwise the tag can not be loaded quantitatively after it has been attached to the oligonucleotide [Peters2011].

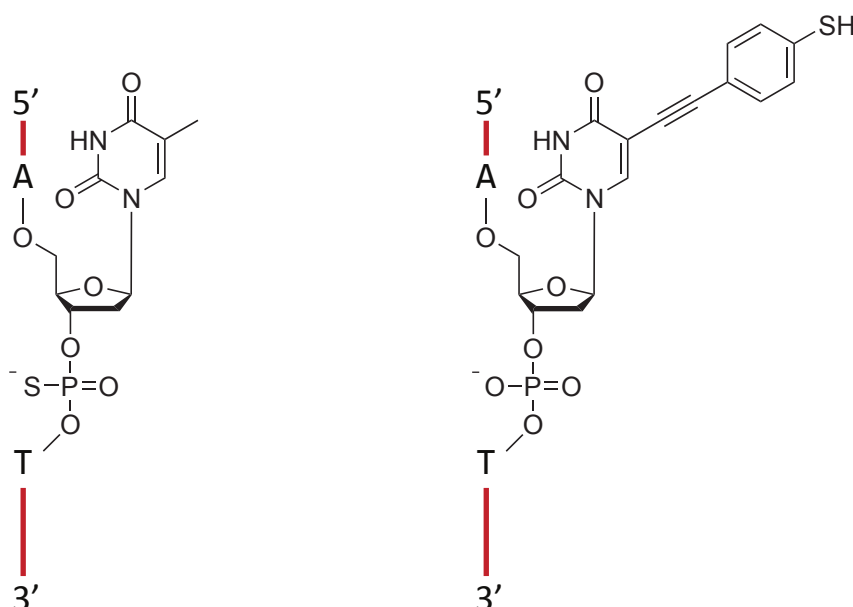


Figure 1.16: Sulfur binding sites in oligonucleotides resulting from the two pursued strategies.

For the new tagging method, the native structure of the target molecule has to be maintained. The modified nucleobase strategy will be applied to the 24-mer DNA strand that was presented by E. Woeltjen (see section 1.2, Fig.1.10), as the effect of a modification has already been investigated for a very similar linker at the same position. However, the influence of a new modification still needs to be examined. The strategy can be evaluated by the measurement and analysis of PCSs and RDCs in NMR spectroscopy.

In addition, the Cys-Ph-TAHA tag will be improved by discarding the cysteine linker in order to generate a more rigid version of the tag, which is supposed to result in the measurement of larger dipolar couplings.

Altogether, the different objectives of this work are summarized in the following list:

1. An efficient method to employ paramagnetic tagging should be established by
 - a) a sulfur binding site in a modified nucleobase.
 - b) tagging of oligonucleotide phosphorothioate.
2. The new strategy needs to be evaluated by the measurement of paramagnetically induced structural parameters like PCSs and RDCs on an oligonucleotide.
3. The Cys-Ph-TAHA tag should be improved by shortening of the cysteine linker.

2 Theory

In NMR spectroscopy, paramagnetic tags are attached to biomolecules to enable the determination of residual dipolar couplings, pseudocontact shifts and paramagnetic relaxation enhancement. These NMR parameters originate from different effects, induced by unpaired electrons. For the purpose of this work, the RDC is the most important parameter, as it provides global structural information about the target molecule, yet its determination is highly difficult in unlabeled NMR samples. The PCSs complement the RDCs and can be measured more easily, albeit they have a more local character and a lower structural sensitivity due to their distance dependence. Therefore, they should be combined with RDCs in order to exploit the full potential of the paramagnetic effects. In this approach, the PRE complicates NMR spectroscopy, as the measurement of unlabeled samples is already limited by a low signal-to-noise ratio, which is further reduced by the paramagnetic center.

Even though PCSs and PRE can exclusively be observed in close contact to the paramagnetic center and RDCs provide a more global structural information, all of these effects can be used to obtain long-range structural information. In contrast, other NMR parameters such as NOE restraints or scalar couplings provide short-range information, which are limited by a stronger distance dependency (r^{-6}) or few chemical bonds [Clare2009].

Lanthanide ions have become a commonly used source for the introduction of paramagnetic centers in NMR spectroscopy, as reviewed by Otting *et al.* [Otting2008] [Otting2010]. The unpaired electron is localized in an inner f-orbital, resulting in the different lanthanide ions having similar chemical properties and a presumed identical coordination by a tag. In contrast, their magnetic properties are highly individual, which increases the scope of their applications [Pintacuda2007]. For the purpose of this work, the lanthanide ions with the largest paramagnetic properties (Tb and Tm) were employed, in order to compensate the long linker and tag. For other applications, it might be advantageous to incorporate lanthanide ions that induce smaller paramagnetic effects (Ce or Sm), minimizing an interfering PRE. Lanthanide ions without paramagnetic properties (Lu or La) can be used for the preparation of reference samples as they provide similar ionic radii.

In this chapter, the concepts of RDCs, PCSs, PRE and their key equations are reviewed based on the publications by Bertini *et al.*, Kramer *et al.* and the comprehensive review of F. Siepel [Bertini2002a] [Kramer2004] [Siepel2013].

2.1 Residual Dipolar Couplings

Dipolar couplings are through-space interactions of nuclei that exhibit a nuclear spin. Compared to scalar couplings, the dipolar contribution can not be observed in conventional solution-state NMR spectra. The effect is dependent on the angle θ , which is spanned by the inter-nuclear vector \vec{R} and the external magnetic field vector \vec{B} (see Fig.2.1). Due to the isotropic tumbling of molecules in solution, the dipolar coupling is averaged to zero, rendering the effect invisible in conventional NMR studies. The tumbling can, however, be artificially disturbed by a partial alignment induced by spatial restrictions, making the RDCs observable. Therefore, external media, such as phages [Hansen1998], gels [Tycko2000] [Schmidt2012] or liquid crystals [Bax1997], have been employed in order to generate a strong alignment, which can even be adjusted by manipulations of the applied media. Paramagnetic tags usually generate a smaller level of alignment, resulting in residual dipolar couplings of similar magnitude as scalar couplings. The employed Cys-Ph-TAHA tag has been used to generate NH RDCs up to 17.8 Hz, corresponding to a theoretical CH RDC maximum up to 44 Hz. This expected magnitude of the CH RDCs is presumed to be sufficient, providing that the tag is attached to a rigid binding site [Peters2011].

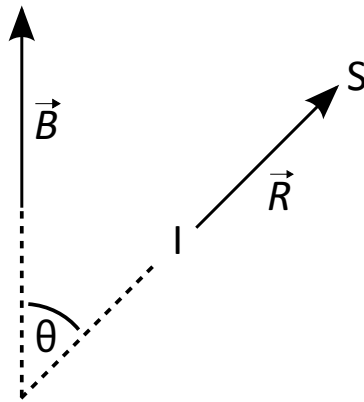


Figure 2.1: Illustration of the magnetic field vector \vec{B} , the inter-nuclear vector \vec{R} and their connecting angle θ .

The Hamiltonian of the dipolar coupling \hat{H}_D can be expressed in analogy to the scalar coupling. For the heteronuclear interaction between two coupled spins I_z and S_z it is given by

$$\hat{H}_D = 2\pi D I_z S_z, \quad (2.1)$$

with the dipolar coupling constant

$$D = -\frac{3\gamma_I\gamma_S\mu_0\hbar}{8\pi^2 R^3} \left(\cos^2 \theta - \frac{1}{3} \right) \quad (2.2)$$

in units of Hz. Therein, the gyromagnetic ratios of the coupled nuclei are given by γ_I and γ_S and $\hbar = h/2\pi$ and μ_0 refer to the Planck constant and the permeability of vacuum, respectively. The important structural dependent variables are given by the distance R between the spin pair and the angle θ , which is illustrated in Fig.2.1. Undisturbed, the $(\cos^2 \theta - \frac{1}{3})$ term is averaged to zero in conventional solution-state NMR spectra.

The molecular tumbling can be described as a time-dependency of the magnetic field vector $\vec{B}(t)$ in a molecular frame, in which the internuclear vector \vec{R} between I and S is fixed (see Fig.2.2). Further contributions to the flexibility of this vector, for example inter-domain dynamics, are neglected and the molecule assumed to be rigid. Consequently, the angle θ and the dipolar coupling constant D also become time-dependent in the molecular frame. The time-averaged dipolar coupling constant, which describes the residual dipolar coupling, is given as

$$\bar{D} = -\frac{3\gamma_I\gamma_S\mu_0\hbar}{8\pi^2R^3}\left(\overline{\cos^2\theta} - \frac{1}{3}\right). \quad (2.3)$$

The time-averaged $\overline{\cos^2\theta}$ term in this equation expresses the molecular tumbling, which can be either isotropic or, in the case of an aligned molecule, anisotropic. A useful description of the rotational motion, and therefore for the $\overline{\cos^2\theta}$ term, is given by the probability tensor P , which can be directly transformed into the more common alignment tensor A using

$$A = P - \frac{1}{3}1. \quad (2.4)$$

The probability tensor P , however, represents the likelihood of finding the magnetic field vector $\vec{B}(t)$ along one of the directions within its frame, which is defined by its principal axes $(\tilde{x}, \tilde{y}, \tilde{z})$. The overall probability of finding the magnetic field vector along the axes of the probability tensor is given as $P_{\tilde{x}} + P_{\tilde{y}} + P_{\tilde{z}} = 1$. Consequently, for an isotropically tumbling molecule, the values for these principal axes are $P_{\tilde{x}} = P_{\tilde{y}} = P_{\tilde{z}} = 1/3$. The probability tensor is symmetric and can be expressed as a matrix in the form of

$$P = \begin{pmatrix} \overline{b_{\tilde{x}}(t)b_{\tilde{x}}(t)} & \overline{b_{\tilde{x}}(t)b_{\tilde{y}}(t)} & \overline{b_{\tilde{x}}(t)b_{\tilde{z}}(t)} \\ \overline{b_{\tilde{y}}(t)b_{\tilde{x}}(t)} & \overline{b_{\tilde{y}}(t)b_{\tilde{y}}(t)} & \overline{b_{\tilde{y}}(t)b_{\tilde{z}}(t)} \\ \overline{b_{\tilde{z}}(t)b_{\tilde{x}}(t)} & \overline{b_{\tilde{z}}(t)b_{\tilde{y}}(t)} & \overline{b_{\tilde{z}}(t)b_{\tilde{z}}(t)} \end{pmatrix}, \quad (2.5)$$

using the individual components of the time-dependent magnetic field vector

$$\vec{B}(t) = \begin{pmatrix} b_{\tilde{x}}(t) \\ b_{\tilde{y}}(t) \\ b_{\tilde{z}}(t) \end{pmatrix}. \quad (2.6)$$

The probability tensor is visualized as an ellipsoid, which transforms into a ball for an isotropically tumbling molecule and reduces to a vector for a perfectly aligned molecule.

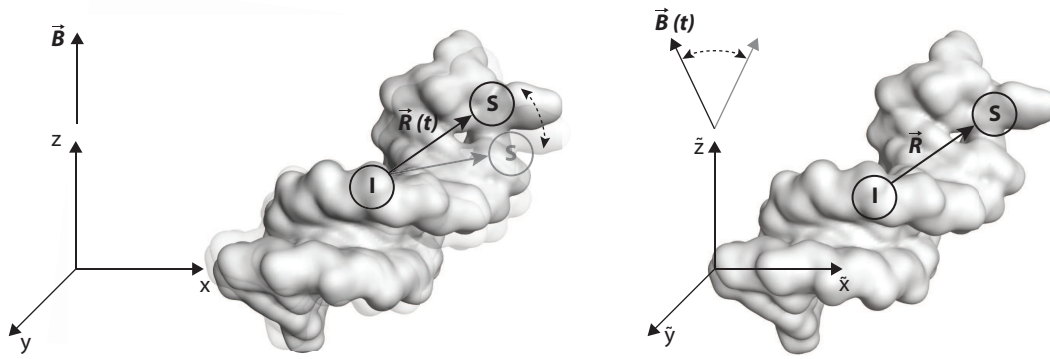


Figure 2.2: Molecular tumbling expressed in the molecular (*left*) or laboratory frame (*right*), resulting in a time-dependent movement of either the magnetic field vector $\vec{B}(t)$ or the inter-nuclear vector $\vec{R}(t)$, respectively. This figure was reproduced according to [Kramer2004].

Using equation 2.4, the corresponding alignment tensor can be determined, which is the traceless part of the probability tensor with $A_{\hat{x}} = A_{\hat{y}} = A_{\hat{z}} = 0$, giving a sphere with positive and negative values. Due to their symmetry and the respectively defined traces, both tensors contain five independent variables, so that five experimentally determined RDCs are required to calculate the molecular alignment, presuming that these RDCs describe the whole space. In a DNA helix, the internuclear vectors of the aromatic CH couplings lie approximately in one plane and consequently more than five RDCs are necessary to accurately calculate the alignment tensor.

The dipolar coupling constant is linked to the tensors P and A by

$$\overline{\cos^2 \theta} = P_{\hat{x}} r_{\hat{x}}^2 + P_{\hat{y}} r_{\hat{y}}^2 + P_{\hat{z}} r_{\hat{z}}^2 \quad (2.7)$$

and

$$\left(\overline{\cos^2 \theta} - \frac{1}{3} \right) = A_{\hat{x}} r_{\hat{x}}^2 + A_{\hat{y}} r_{\hat{y}}^2 + A_{\hat{z}} r_{\hat{z}}^2, \quad (2.8)$$

with \vec{r} being the internuclear unit vector in the molecular frame. Consequently, the residual dipolar couplings (equation 2.3) can be calculated by

$$\overline{D} = -\frac{3\gamma_I\gamma_S\mu_0\hbar}{8\pi^2 R^3} \left(P_{\hat{x}} r_{\hat{x}}^2 + P_{\hat{y}} r_{\hat{y}}^2 + P_{\hat{z}} r_{\hat{z}}^2 - \frac{1}{3} \right), \quad (2.9)$$

using the probability tensor P.

In paramagnetic tagging, the alignment of the target molecule is caused by the anisotropic magnetic susceptibility of the unpaired electron. It is common to describe the residual dipolar couplings obtained by this method with the χ -tensor, which is also used for the description of pseudocontact shifts. In a rigid molecule, the residual dipolar couplings

are dependent on the external magnetic field B_0 , the anisotropic magnetic susceptibility of the lanthanide ion, the gyromagnetic ratios of the coupled nuclei and their distance R :

$$\Delta\nu^{\text{RDC}} = -\frac{\gamma_I\gamma_S\hbar B_0^2}{8\pi^2 R^3 15kT} \left(\Delta\chi_{ax}(3\cos^2\vartheta - 1) + \frac{3}{2}\Delta\chi_{rh}\cos 2\varphi\sin^2\vartheta \right). \quad (2.10)$$

Here, the χ -tensor is separated into an axial (χ_{ax}) and a rhombic (χ_{rho}) component (see section 2.2) and the internuclear vector is defined by polar coordinates in the χ -tensor frame (see Fig.2.3). Therein, a unit vector is expressed in the form of

$$\vec{R} = \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} = \begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix}. \quad (2.11)$$

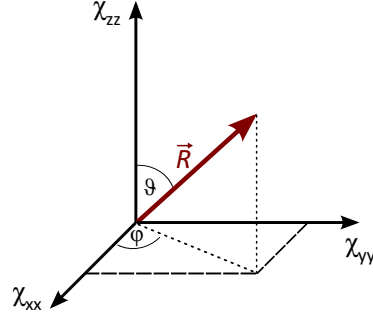


Figure 2.3: Definition of an inter-nuclear vector \vec{R} in the magnetic susceptibility frame.

For the data analysis, the alignment tensor A is calculated using the experimentally determined RDCs ν_{exp} . Based on this tensor, the theoretical couplings ν_{cal} (and their average $\overline{\nu_{cal}}$) are back-calculated and evaluated using the Q -factor and the R^2 -value:

$$Q = \sqrt{\frac{\sum (\nu_{exp} - \nu_{cal})^2}{\sum \nu_{exp}^2}} \quad R^2 = 1 - \frac{\sum (\nu_{exp} - \nu_{cal})^2}{\sum (\nu_{exp} - \overline{\nu_{cal}})^2} \quad (2.12)$$

A data set with the highest possible agreement results in a Q -factor of 0 and an R^2 -value of 1.

In a common molecular frame, the angle β between different alignment tensors A^1 and A^2 can be determined by

$$\cos\beta = \frac{\langle A^1 | A^2 \rangle}{|A^1| |A^2|}, \quad (2.13)$$

which is the normalized scalar product, using the definitions

$$\langle A^1 | A^2 \rangle = \sum_{i,j} A_{ij}^1 A_{ij}^2 \quad |A| = \sqrt{\sum_{i,j} (A_{i,j})^2}. \quad (2.14)$$

for the individual matrix elements A_{ij} . In this work, alignment tensors are calculated based on PCS and RDC data. The different tensors, and therefore the agreement of the different data sets, was analyzed by the angle spanned between them.

2.2 Pseudocontact Shifts

Pseudocontact shifts are caused by a direct interaction of a nucleus with the unpaired electron of a paramagnetic NMR sample. The electron induces an additional magnetic field that, compared to a diamagnetic reference sample, alters the chemical shifts of the observed nuclei. Just as residual dipolar couplings, the PCS is a dipolar interaction through space which, however, does not result in a contribution to the scalar coupling, but in a significant resonance shift. The difference between this effect and the dipolar interaction that yields in RDCs originates from the gyromagnetic ratio of the electron and its rapid longitudinal relaxation. Therefore, only an averaged interaction is observed, which is dominated by the higher populated β -state of the electron, that results in the pseudocontact shift

$$\delta^{\text{PCS}} = \frac{1}{12\pi r^3} \left(\Delta\chi_{ax}(3\cos^2\vartheta - 1) + \frac{3}{2}\Delta\chi_{rh}\cos 2\varphi\sin^2\vartheta \right). \quad (2.15)$$

The distance dependence of r^{-3} to the electron illustrates its local character, so that only nuclei in close proximity to the electron give rise to a significant PCS. For terbium and dysprosium, PCSs can be observed for distances up to 40 Å [Biekofsky1999] [Allegrozzi2000]. Similar to the RDCs, the angles ϑ and φ are defined within the magnetic susceptibility frame (see Fig.2.3). The χ tensor is separated into the axial component

$$\Delta\chi_{ax} = \chi_{zz} - \frac{\chi_{xx} - \chi_{yy}}{2} \quad (2.16)$$

and a rhombic component

$$\Delta\chi_{rh} = \chi_{xx} - \chi_{yy}. \quad (2.17)$$

The angle-dependent $(3\cos^2\vartheta - 1)$ term in the equation 2.15 can be illustrated as a pseudocontact sphere, depicting the nucleus-electron orientation [Pintacuda2007]. At the so-called magic angle of $\vartheta = 54.7^\circ$, the PCS is zero and its value changes its sign. Compared to RDCs, the pseudocontact shifts do not originate from a molecular alignment, but simply from the direct electron-nucleus interaction, which is also present in free-tumbling molecules.

The molecular alignment does, anyhow, also have an influence on the pseudocontact shifts, yet it is of negligible magnitude and given by an additional field-dependent term

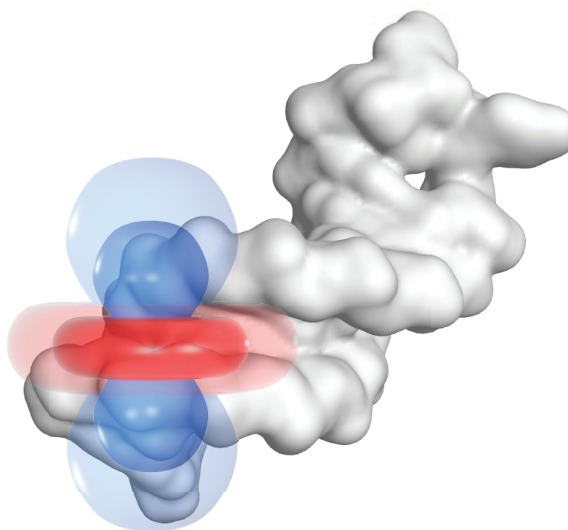


Figure 2.4: Illustration of pseudocontact shifts of an unpaired electron inducing either positive (*red*) or negative (*blue*) shifts in the oligonucleotide.

for an axial symmetric susceptibility tensor:

$$\delta^{\text{PCS}} = \frac{1}{12\pi r^3} \Delta\chi_{ax} (3 \cos^2 \vartheta - 1) \left(1 + \Delta\chi_{ax} \frac{B_0^2}{15\mu_0 kT} \right). \quad (2.18)$$

On a 900 MHz NMR spectrometer, the contribution of the alignment is only around 0.6%, based on the largest possible anisotropic χ -tensor of terbium [Otting2010].

Further effects that may alter the chemical shift, e.g. the residual chemical shift anisotropy (RCSA) or the Fermi contact interaction, do not influence NMR spectroscopy of paramagnetically tagged biomolecules. Compared to the PCS, a partial alignment does not give rise to a significant RCSA, yet this effect was successfully exploited for conformational studies using external alignment media, which enforce a higher level of alignment [Hallwass2011]. The Fermi contact interaction influences the chemical shifts of nuclei that are directly attached to the unpaired electron [Kurland1970], but even though this effect can also be transferred to adjacent nuclei, the target molecule is not affected due to the employed tag and the linker.

2.3 Paramagnetic Relaxation Enhancement

The paramagnetically induced relaxation enhancement provides an additional tool for structural biology, as it enables the detection of low populated states [Schmidt1984] [Kosen1986] [Iwahara2006] [Clare2009]. For this purpose, nitroxide spin-labels or Mn^{2+}

are attached to the target molecule in order to introduce an unpaired electron. For the determination of PCSs and RDCs, the relaxation enhancement is highly interfering, as it leads to significant line-broadening. In the worst case, PRE influences the resonances close to the paramagnetic center beyond detection, which is referred as *bleaching*. The dipolar interaction between the paramagnetic center and the nuclei results in the relaxation enhancement, which is given by

$$R_2^{\text{PRE}} \propto \frac{\gamma_I^2 B_0^2 (g_e S_e (S_e + 1))^2 \tau_r}{r^6 T^2} \quad (2.19)$$

for the Curie-spin relaxation [Gueron1975]. It is the major component of relaxation enhancement for most of the lanthanide ions, due to their short electron relaxation times. In equation 2.19, g_e represents the g -factor of the electron, S_e is the electron spin and τ_r is the rotational correlation time. The correlation of the PRE to the gyromagnetic ratio γ_I illustrates that protons are more strongly affected by PRE than ^{13}C or ^{15}N .

An interaction of the Curie-spin relaxation with other relaxation effects, such as the dipole-dipole relaxation, can result in a cross correlation on the transversal relaxation [Ghose1997] [Boisbouvier1999] [Bertini2002b] [Pintacuda2004]. In that case, the relaxation is modulated by an additional cross correlation relaxation term, leading to either higher or lower relaxation rates of the nuclear polarization. As a result, the individual linewidth of an affected CH doublet is increased or decreased, depending on the spin state of the coupled nuclei. The cross correlation effect, which is employed in *transversal relaxation optimized spectroscopy*, is highly efficient for the NH spin pair of the backbone of proteins at high magnetic fields and enables the examination of large biomolecules [Pervushin1997] [Fernandez2003].

In the context of this work, the distance dependence of the relaxation process ($\propto r^{-6}$) and its magnetic field dependence ($\propto B^2$) are of substantial importance. Higher magnetic fields are only necessary to induce a stronger alignment for RDC measurements, yet the determination of PCSs, however, is achieved at lower fields to minimize the influence of the PRE. Fortunately, the stronger distance dependence of the PRE compared to that of PCS (r^{-3}) facilitates its measurement.

3 Synthesis

3.1 Tagging Strategies of Oligonucleotides

Paramagnetic tagging of oligonucleotides requires the introduction of a binding site into the target molecule. There are several different approaches for attaching a tag to DNA or RNA molecules, for example thioether or disulfide as the linking functional group (see section 1). The Cys-Ph-TAHA tag **1**, which has been demonstrated to exhibit excellent properties in paramagnetic NMR spectroscopy, is connected via a disulfide bridge, providing a high stability and a satisfactory tagging yield [Peters2011]. Consequently in order to introduce a tag, a sulfur moiety needs to be implemented in the oligonucleotide. In section 1.2, several examples of oligonucleotide modification for paramagnetic tagging are presented. Regarding the synthetic approach, the following requirements need to be considered:

1. The modification needs to be solvent-exposed to enable efficient tagging.
2. The modification has to be rigid with an unambiguous stereochemistry to guarantee strong alignment and explicit NMR spectra.
3. The strategy requires a high reproducibility and a reasonable yield.

Furthermore, the native structure of the target molecule should remain unaltered. Considerations on the modification position are outlined in section 1.2 and were followed throughout the different strategies.

In general, two different approaches are presented in this chapter, e.g. nucleobase modification and oligonucleotide phosphorothioates [Eckstein1985]. Both strategies are based on a chemical oligonucleotide synthesis as it provides the necessary high tolerance towards manipulations as opposed to a biological synthesis. First attempts on a thymine nucleobase with a bulky EDTA based modification revealed that the introduction of a paramagnetic center has to be separated into two parts. Modifications, which directly provide a coordination site for the lanthanide ions were too bulky for efficient DNA synthesis [Woeltjen2005]. With an eye on previous efforts using 2 or 3 step approaches to

generate binding sites in DNA molecules, to which a paramagnetic tag has subsequently been attached after oligonucleotide synthesis [Woeltjen2009] [Siepel2013], a new synthesis strategy is presented in this chapter.

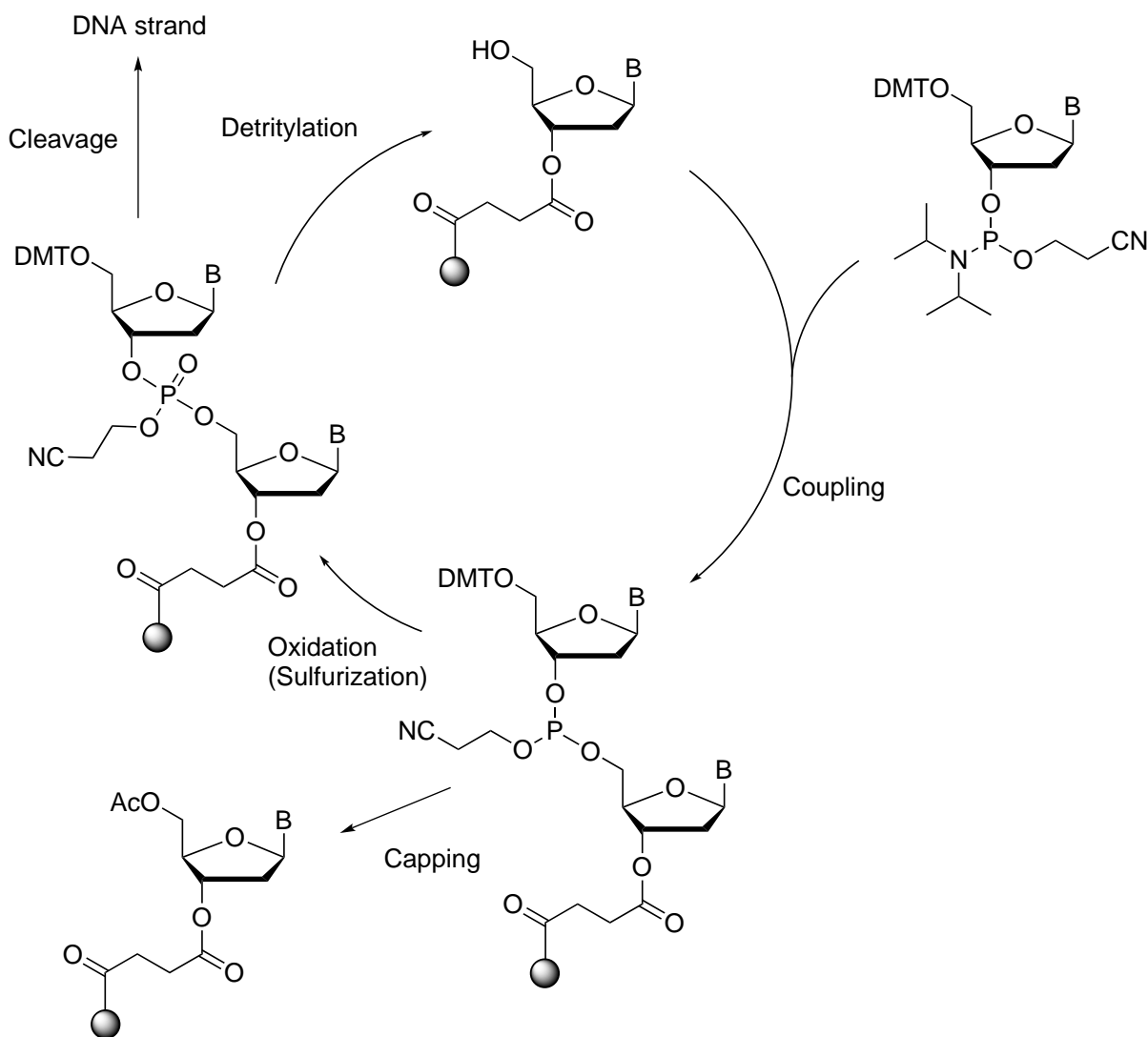


Figure 3.1: Synthesis of DNA by the phosphoramidite method using a succinyl linkage to a solid support. Shown are the 4 steps of the synthetic cycle (detritylation, coupling, capping, oxidation or sulfurization) and the final cleavage procedure.

The phosphoramidite method and the oligonucleotide synthesizer illustrated in Fig.3.1 play a key role in all of the different approaches. Therefore, the four steps of the synthetic cycle shall be reviewed briefly [Matteucci1981] [Beaucage1981] [Caruthers2001]:

- **Detritylation**

At the beginning of every synthetic cycle, the 4,4'-dimethoxytrityl protection group

(DMT) at the 5'-position is cleaved using 3% trichloroacetic acid solution (TCA) in dichloromethane. The coupling efficiency is monitored by controlling the characteristic UV-VIS absorption of DMT-cations at 498 nm [Gaur1989].

- **Coupling**

The coupling reaction is performed using a 2-(cyanoethyl)-diisopropylphosphoramidite activated nucleobase and tetrazole, generating a reactive tetrazolyl-phosphoramidite. The 5'-hydroxy group of the oligonucleotide forms a phosphite triester which elongates the sequence from the 3'- to the 5'-position by one nucleoside.

- **Capping**

Due to incomplete coupling, the remaining 5'-hydroxy groups are deactivated using acetic anhydride in order to avoid coupling in the next cycle at this position.

- **Oxidation**

The phosphite triester is oxidized to an acid stable phosphate triester using, for example, iodine, taking account of the conditions of the following reaction cycle. The resulting triester is also a precursor of the phosphate diester backbone of the final oligonucleotide. Pyridine or lutidine is used to neutralize the reaction mixture. By substituting the conventional oxidation agent with a sulfur source (e.g. DDTT or *Beaucage* reagent), it is possible to generate a phosphorothioate oligonucleotide [Beaucage1990] [Guzaev2011]. In this case, the capping procedure is carried out subsequently.

After synthesis, the oligonucleotide is cleaved from the solid support and all base-labile acyl protection groups attached to any nucleobases are removed using aqueous ammonium hydroxide. The oligonucleotide synthesis can be carried out for DNA and RNA molecules, using the same protocol, however, RNA synthesis requires an additional silyl protection group (TBDMS or TOM), which can be cleaved by fluoride ions [Usman1987] [Wu1998]. Altogether, acid or base labile protection groups provide no suitable protection for any functionalisation as they would be removed in either the synthetic cycle (by TCA or pyridine) or eventually during the cleavage procedure (by conc. ammonia solution).

Considering these limitations, two different protection group strategies were administered resulting in the modified phosphoramidites **17** and **18**, following either a disulfide or a cyanoethyl protection strategy (Fig.3.2). Both approaches are based on *Sonogashira* reactions, connecting a sulfur functionalized phenyl linker to a acetylene modified deoxythymidine [Sonogashira1975] [Sonogashira2002]. Details on the individual approaches are given in sections 3.2 and 3.3.

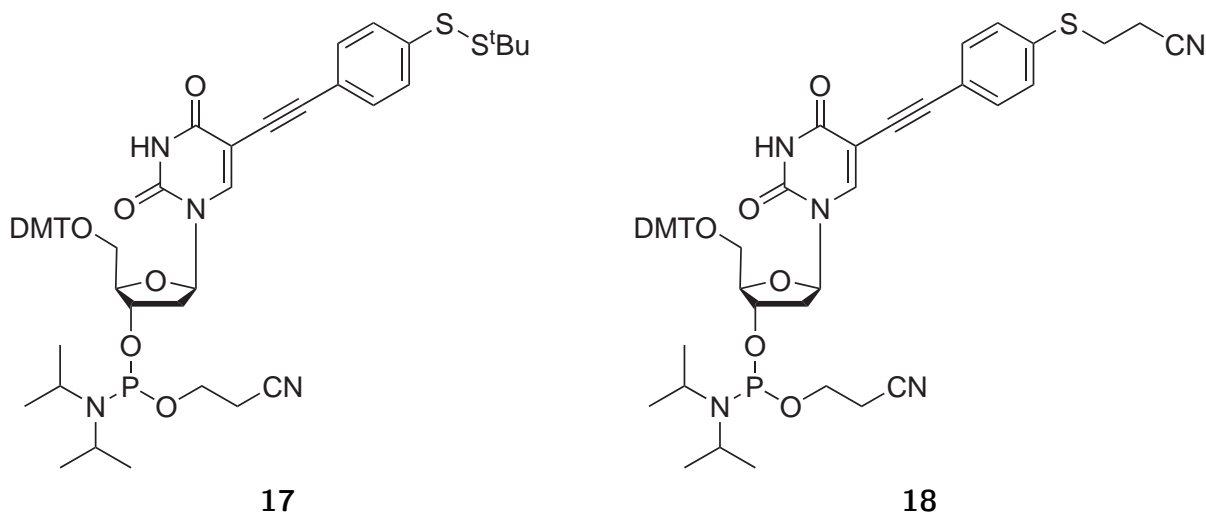


Figure 3.2: Modified and phosphoramidite-activated nucleobases **17** and **18** providing a sulfur binding site in DNA.

For paramagnetic tagging, the modified nucleobase will be implemented into a well studied DNA strand, which contains ten complementary base pairs and a four base pair loop region, forming the 24-mer hairpin structure **15** (Fig.3.3) [Woeltjen2009]. The indicated thymine nucleobase at position 3 will be replaced by the modified nucleobase, taking into account the conventional DNA synthesis from 3'- to the 5'-position. Due to steric hindrance, the coupling efficiency of a modified nucleobase might be smaller compared to conventional nucleosides and therefore a manipulation in a later step of the synthesis might be favorable. In section 4.2 the effect of the manipulation in comparison to the wild-type DNA is outlined.

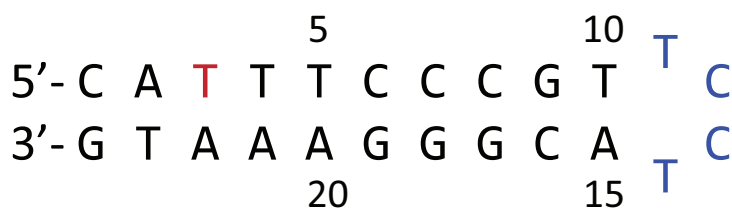


Figure 3.3: Wild-type 24-mer DNA hairpin structure **15**. Thymine at position 3 (*red*) was replaced by a modified nucleobase.

Additionally, an entirely different tagging strategy was attempted using a modification of the oligonucleotide's backbone. A non-bridging oxygen atom of the phosphate diester linkage is replaced by a sulfur atom at the oxidation step of the synthetic cycle, generating an oligonucleotide phosphorothioate (OPS) in DNA or RNA strands. A solvent exposed binding site at the oligonucleotide backbone is created in one specific position, which has already been used for tagging previously [Qin2001] [Qin2007]. Details on this strategy are given in section 3.5.

For all of the different approaches, the Cys-Ph-TAHA tag [Peters2011] was used due to its excellent properties in the resulting NMR spectra and its chemical availability. Its convenient synthesis requires 9 steps with good yields, which readily allows for modifications compared to the more sophisticated synthetic pathways of other tags (e.g. DOTA-M8) [Haeussinger2009]. In this chapter, different alterations of the Cys-Ph-TAHA tag are presented, to come into use in two different tagging strategies. On one hand, a modification was required for the OPS approach, as tagging could not be achieved via disulfide bond, but via thioether linkage [Qin2001] [Qin2007]. On the other hand, a promising reaction pathway for a much shorter version of the tag is presented. Although the synthesis was not completed, an encouraging intermediate was synthesized with good yields. The synthesis of the new tag requires only one additional step, followed by a conventional deprotection procedure.

3.2 Disulfide Protected Binding Site

The first approach for introducing a linker with a free thiol moiety into a DNA strand was based on a disulfide protection group. In respect of the synthesis and cleavage conditions in which no reduction agents are used, a disulfide bridge is a stable bond to protect the thiol towards electrophiles. Furthermore, deprotection can be achieved easily with various reduction agents, e.g. TCEP or DTT. The connection between linker and nucleobase was achieved by a conventional *Sonogashira* reaction, which unfortunately lacks tolerance towards a disulfide moiety. Therefore, the disulfide protection group could only be introduced after this key reaction step and the thiol moiety had to be protected in another form. Since in previous studies, sulfur moieties had shown to reduce the yield of palladium catalysed reactions significantly [Taeubert2010]. The bulky DMT protection group was a promising alternative as it was expected to be cleaved efficiently, afterwards. Consequently, the alcohol moieties needed to be protected in a different order to avoid interference with the DMT group attached to the sulfur. The corresponding reaction pathway is shown in Fig.3.4.

Starting with 5-iodo-2'-desoxyuridine **22**, the nucleobase was protected and the ethynyl group synthesized over 3 steps [Woeltjen2009]. The linker **21** was prepared over 3 steps according to the literature [Haiss2006], followed by DMT protection with a moderate yield of 61%. The *Sonogashira* reaction was performed under conventional conditions with a 68% yield. Unfortunately, straightforward deprotection of the thiol group with TFA or silver nitrate to produce compound **27**, followed by oxidation to generate the disulfide were not be achieved with reasonable yields. Therefore, these two reactions were combined by transforming compound **24** into a silver salt which was then directly converted into

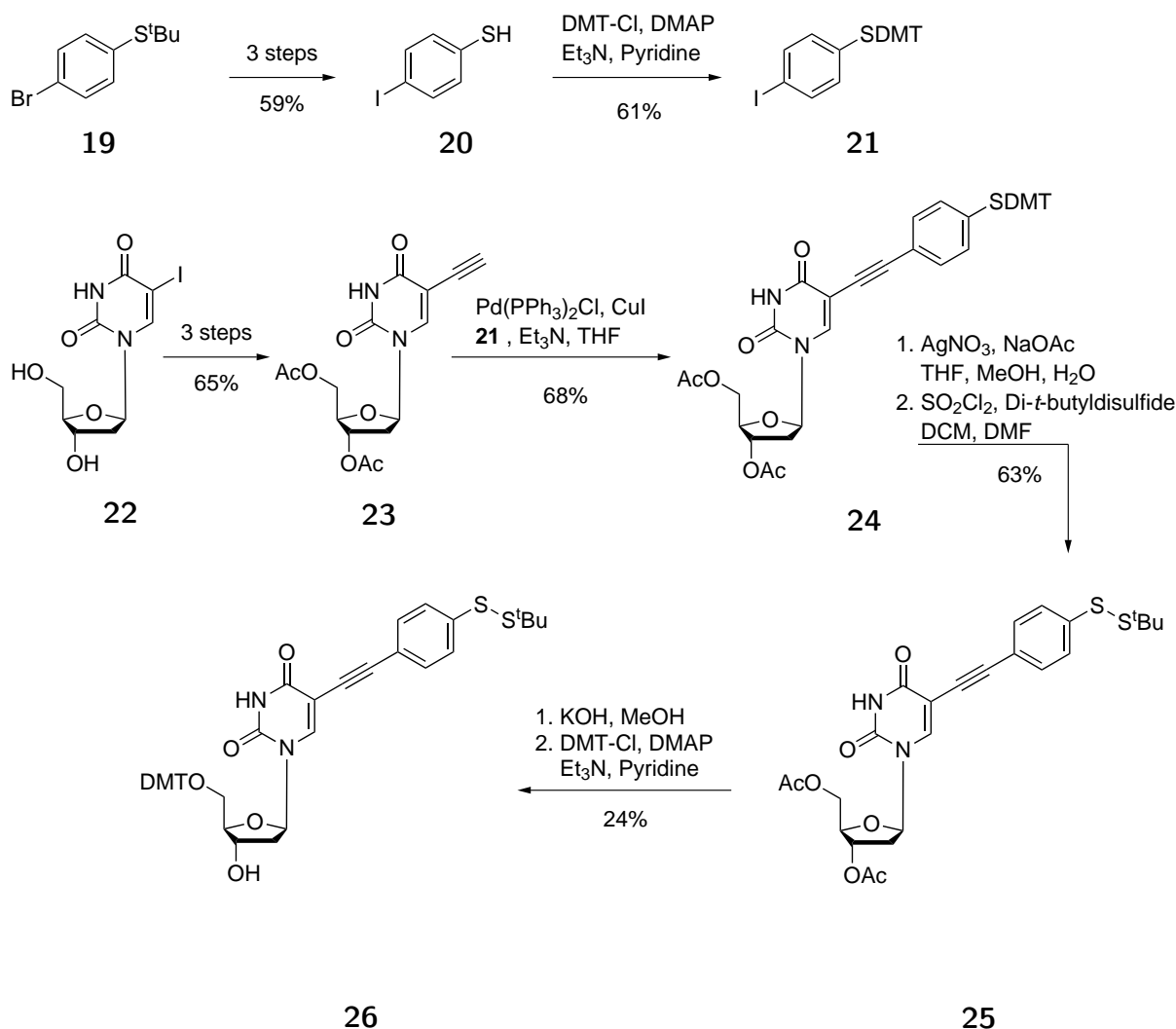


Figure 3.4: The reaction pathway of the disulfide protection group strategy. Intermediate **26** can be activated directly to give phosphoramidite-activated compound **17**.

the disulfide **25** using an *in situ* prepared *S*-chloro *tert*-butylthiolate, yielding 63% yield over two steps (see Fig.3.5). The acetyl protection groups were removed and subsequently the 5' position of the deoxyribose was protected with the DMT group. The low yield of 24% over these two steps illustrates the major drawback of this approach. Introducing the DMT group in a later step of the synthetic pathway is unfavorable due to the steric hindrance of the bulky DMT group.

Compound **26** has ultimately not been activated and implemented into a DNA strand, as a second approach had coincidentally shown more favorable results (see section 3.3).

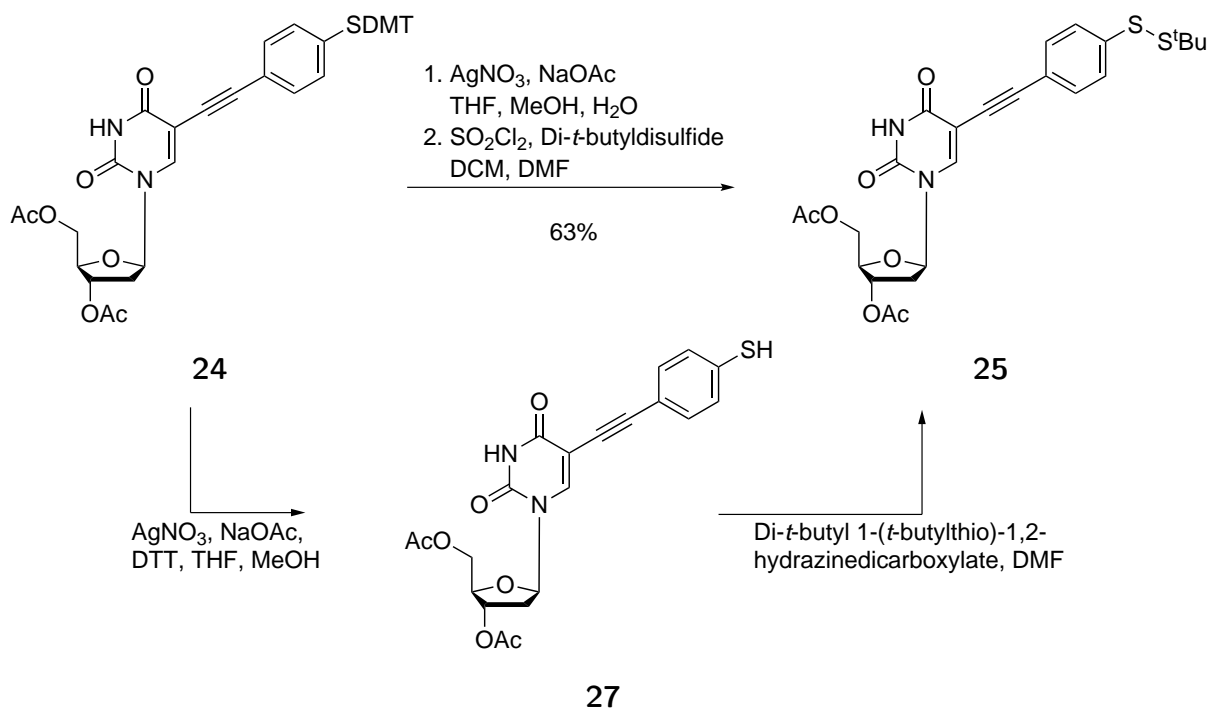


Figure 3.5: Transformation of protection groups at the sulfur moiety.

3.3 Cyanoethyl Protected Binding Site

The most convenient approach to introduce a Cys-Ph-TAHA binding site into an oligonucleotide was achieved using a cyanoethyl protection group for the sulfur moiety. The synthesis of the final phosphoramidite **18** requires a 9-step synthesis with a total yield of 20%, starting from commercially available compounds. The major advantages lie in the excellent yields of the individual reactions combined with the convergent synthesis strategy illustrated in Fig.3.6. Two individual fragments **28** and **29** were synthesized according to literature reported procedures [Haiss2006] [Woeltjen2009], followed by a connecting *Sonogashira* reaction and a straightforward phosphoramidite activation. In contrast to the previous approach, the protection group strategy was more convenient and led to the higher yield, yet a special deprotection procedure was required after DNA synthesis.

Over the course of the synthesis no change of protection groups takes place, whatsoever. The cyanoethyl protection group enabled the *Sonogashira* reaction and was directly applicable for the DNA synthesis as well. Furthermore, the DMT protection group was introduced in the first step of the synthesis which led to a significantly higher yield compared to an introduction at a later point. Due to the convergent strategy, its lability towards acids was no disadvantage. In contrast to the previously presented approaches (see section 3.2), the critical step was not the synthesis of the modified nucleobase **18** itself but the cleavage and deprotection procedure following the DNA synthesis, as it follows a different mechanism compared to that for a disulfide bond. However, deprotection of

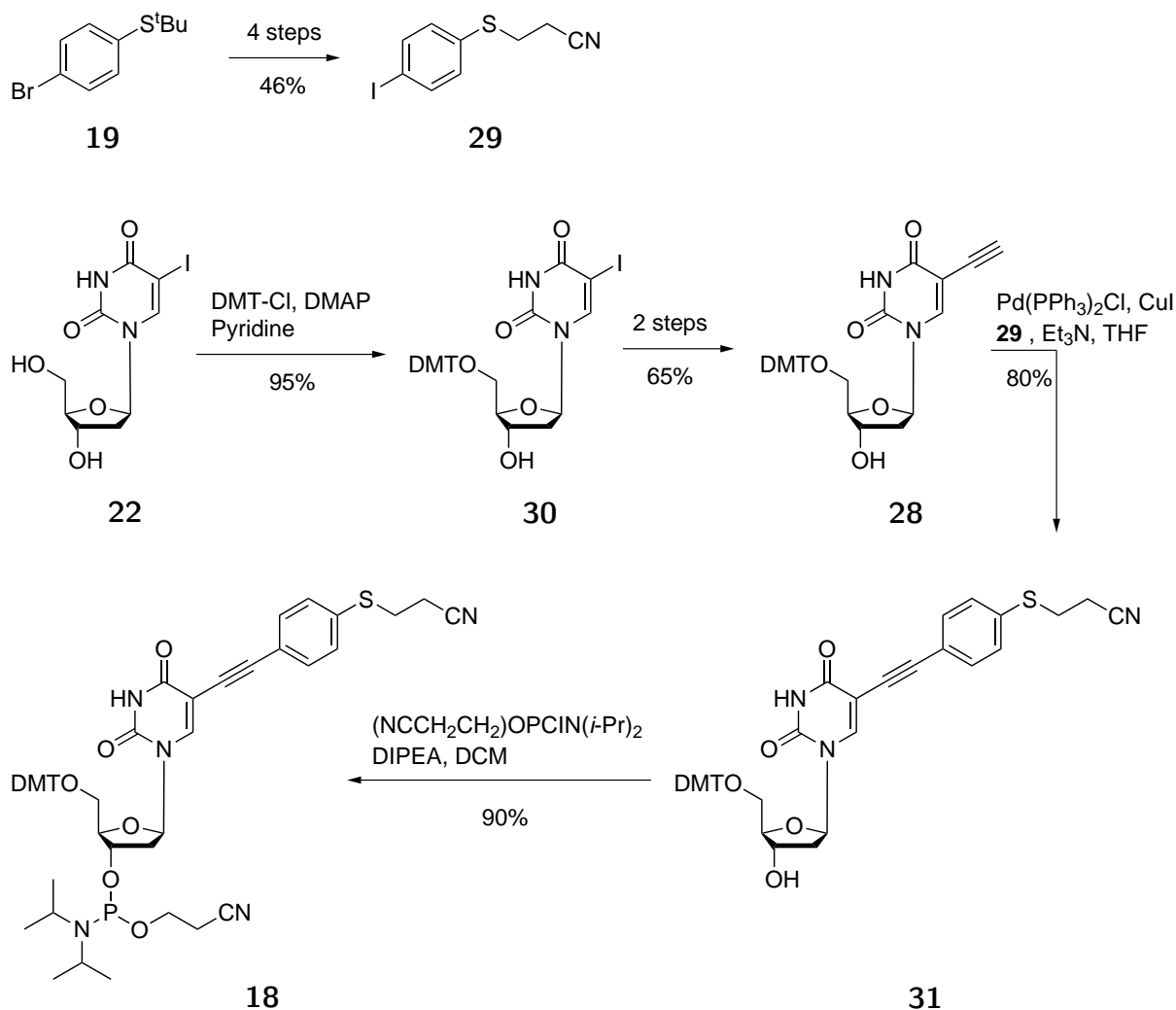


Figure 3.6: The reaction pathway to generate the phosphoramidite-activated compound **18**.

a cyanoethyl protection group had already been reported [Christopherson1991] and was successfully applied to fulfill this purpose. (see section 3.4).

A small variation of the presented approach is illustrated in Fig.3.7. Therein, the cyanoethyl protection group was substituted by an acetyl group. Within the conventional DNA cleavage procedure, the acetyl group would be removed leading to the free thiol. Under conventional conditions, the electrophilic acrylonitrile, an elimination product of the cleavage procedure, would react with the thiol, forming a cyanoethyl group [Taubert2010]. Thus, both strategies (i.e. acetyl or cyanoethyl protection) would lead to the same result. The required compound **32** was already available as it was an intermediate of the linker synthesis. The corresponding *Sonogashira* reaction was performed with a 53% yield. Even though, in this approach, the linker synthesis was reduced by 2 steps, the cyanoethyl protection group is still favorable due to the significantly higher reproducibility and yield.

Starting from 5-iodo-2'-desoxyuridine **22**, DMT protection and ethynyl functionalisation was achieved within 3 steps. In parallel, the linker **29** was synthesized over 4 steps, starting from 1-bromo-4-(*tert*-butylsulfanyl)benzene **19**. Within this sequence, the final cyanoethyl protection group was introduced, giving a good 80% yield in the following *Sonogashira* reaction. Finally, the two reaction pathways were combined to give compound **31**, which was subsequently transformed to the final product **18**.

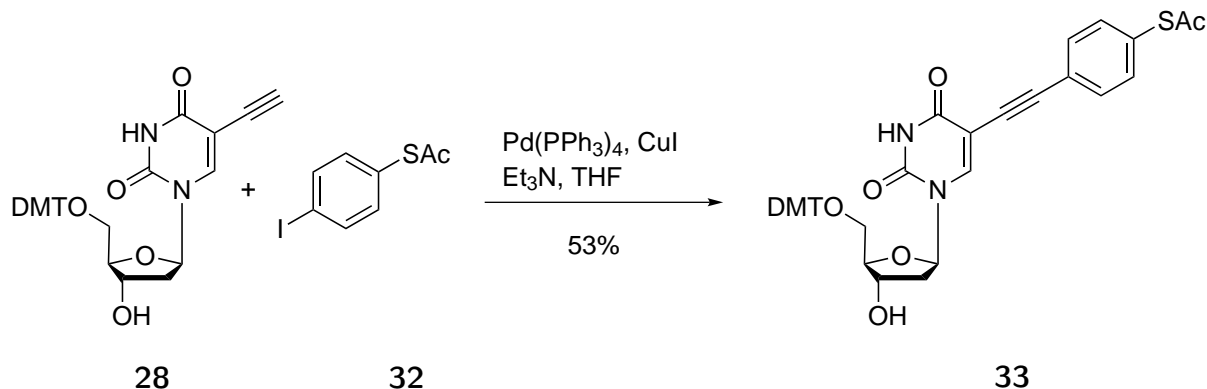


Figure 3.7: Alternative acetyl protection of thiol, resulting in a lower yield in the *Sonogashira* reaction.

The activated nucleobase **18** was implemented into the DNA strand **15** by the company *IBA*. The sufficient coupling efficiency (95-97%) for the modified nucleobase and the following coupling reactions allowed for a reasonable overall yield for the oligonucleotide synthesis. The modified DNA strand requires, however, a special workup procedure, which is described in section 3.4.

3.4 Cleavage of Sulfur Modified DNA Strand on Solid Support and Tagging with Cys-Ph-TAHA

The cyanoethyl protection group is not only applied in the modification, but also it is used for the protection of the backbone during the DNA synthesis. Therefore, its elimination product acrylonitrile does not interfere with the conventional synthesis and cleavage procedure. Furthermore, the CEM protection group, which is a cyanoethyl analogous of the MOM protection group, has improved the coupling efficiency of the *P*-stereodefined OPS synthesis, in comparison to a conventional silyl protection of RNA 2'-hydroxy group [Nukaga2012]. The high nucleophilicity of the free thiol and its reactivity towards the electrophilic acrylonitrile is the crucial factor of this approach. In the literature, deprotection

of a cyanoethyl protected 2'-deoxy-6-thioguanosine was reported using a 1 M NaOH / 0.1 M NaSH solution [Christopherson1991], but the direct application to the synthesized DNA proved unsuccessful. However, an excess of sodium sulfide, followed by the reduction agent TCEP lead to the desired free thiol. Therefore, sodium disulfide was added to the conventional deprotection solution, so that it acts as a *scavenger* for the unwanted acrylonitrile [Coleman1994]. In an additional step, TCEP reduced the sulfur moiety to give the free thiol (Fig.3.8). Fortunately, no purification was required between these two steps, leading to a convenient and fast protocol with a high reproducibility.

Oligonucleotides were purchased on solid support, in which one nucleoside was substituted by the modified nucleobase. Without any alterations, the cleavage and purification procedure was applied to two different DNA sequences with either 10 or 24 nucleobases. The purchased oligonucleotide was incubated at 55 °C with 0.2 M sodium sulfide in conc. ammonium hydroxide and ethanol (3:1). In the next step, TCEP was added to the mixture, giving the free thiol. Subsequent HPLC purification produced the isolated DNA strand, which can be stored for several weeks. To avoid oxidation of the free thiol to a disulfide, it was required to protect the DNA under argon atmosphere, following to the HPLC purification.

Tagging of the modified DNA follows in principle the previously reported procedure [Siepel2013]. However, two alterations were necessary to achieve the desired DNA. Following the preparation of the tag [Peters2011], the functionalized DNA was incubated with the preloaded tag solution for 12 h. Compared to the approaches described in literature, the employed linker is significantly shorter and thus, the binding site is less solvent exposed. This leads to an incomplete tagging reaction at room temperature. Therefore, the first significant alteration was the elevation of the incubation temperature to 55 °C. As described in the literature, approximately 10 washing steps with 1 M sodium chloride were necessary to remove traces of lanthanide ions coordinated to the DNA backbone. In addition, the DNA needs to be folded correctly as the 24-mer DNA strand tends to build a dimer rather than the desired hairpin structure (see section 4.2). For this purpose, the tagged DNA was melted at 70 °C and allowed to fold into the monomer by gradually lowering the temperature. Details on the NMR sample preparation and buffers are given in the Experimental Part 6.

3.5 Tagging on Phosphorothioate Oligonucleotides

As described in section 1.3, the backbone of oligonucleotides is a promising modification target, into which sulfur can be easily introduced. In the synthesizer, nucleosides are attached in a phosphorous III-form and oxidized subsequently. In contrast to a conventional synthetic cycle, a sulfur-carrying oxidation agent can be used to regioselectively

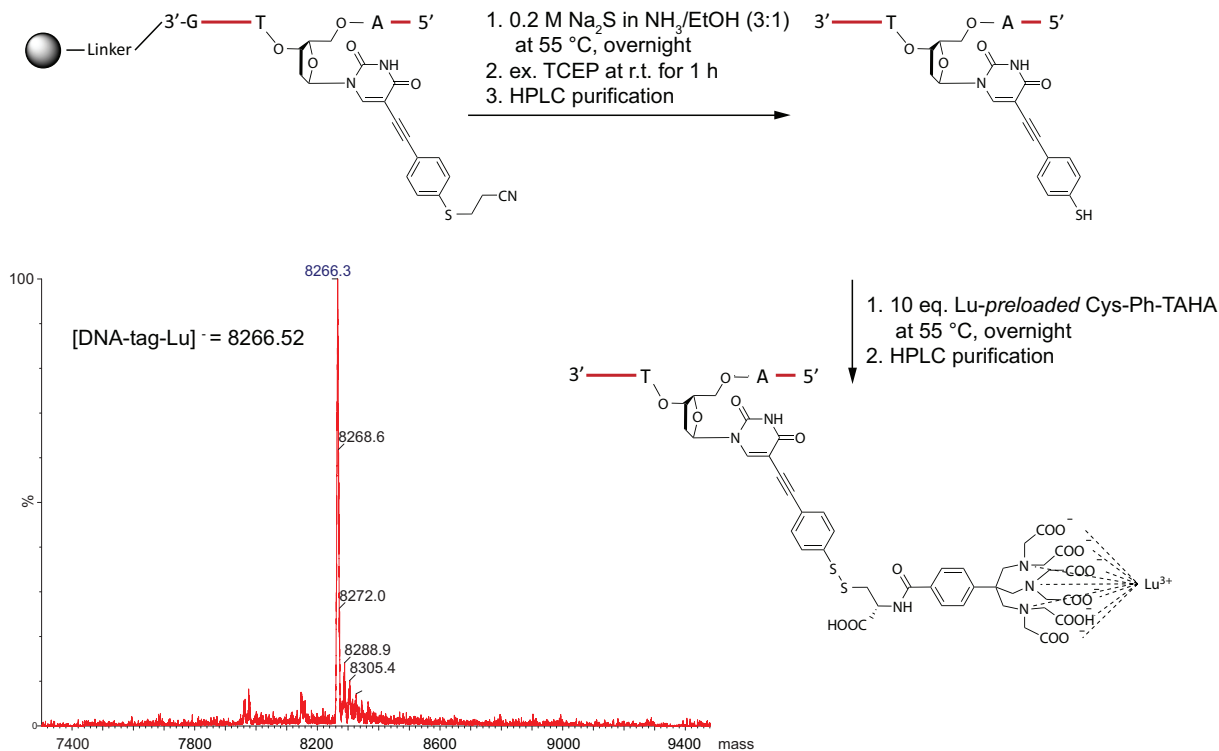


Figure 3.8: Cleavage and tagging scheme for 24-mer DNA strand with modified nucleobase at position 3 and *Lu-preloaded Cys-Ph-TAHA*. Mass spectrum of the purified DNA shows the successful tagging reaction.

generate oligonucleotide phosphorothioates (OPS) (see section 3.1). Compared to the previous approach, in which sulfur is introduced via a complex phosphoramidite synthesis, OPS synthesis is significantly faster. Unfortunately, the sulfur atom generates a new chiral center in the backbone, resulting in a doubling of NMR signals. Nevertheless, both stereoisomers can be used independently for paramagnetic tagging as they provide different information. Although first stereoselective strategies have been reported in the literature using phosphoramidites with a defined stereochemistry [Oka2008] [Nukaga2012], the two stereoisomers were separated conventionally via HPLC chromatography (see Fig.3.9).

In general, the feasibility of the separation highly depends on the sequence, the sulfur position and the length of the oligonucleotide [Frederiksen2009]. Considering the final purpose of paramagnetically tagging, the RNA:RNA or DNA:RNA construct **16** (illustrated in Fig.3.10), the sequence of the tagged oligonucleotide is irrelevant as long as it does not interfere with the target RNA.

A suitable 10-mer RNA, which can not build more than 3 consecutive base pairs with the target RNA, was separated according to a previously reported protocol [Slim1991] [Christian2000]. Fig.3.9 shows two HPLC chromatograms with significant differences in

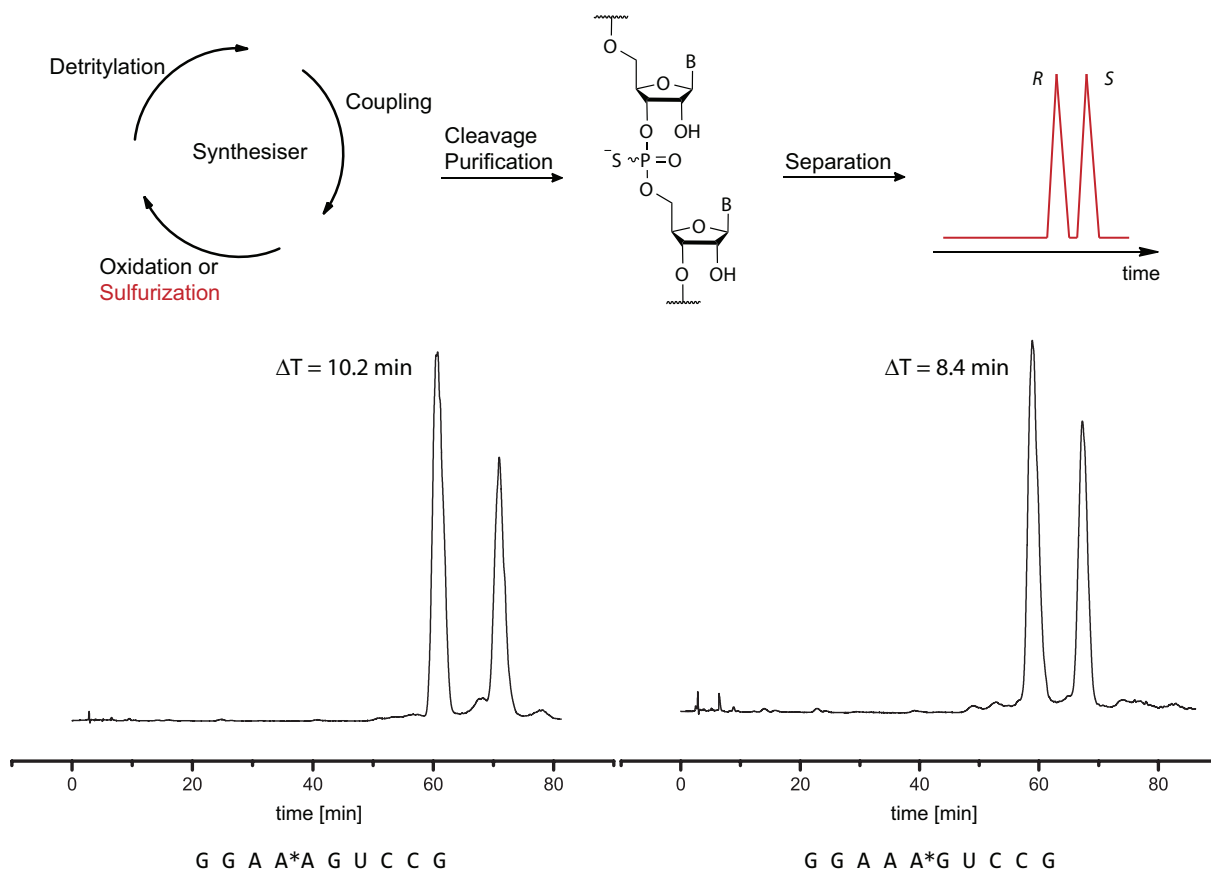
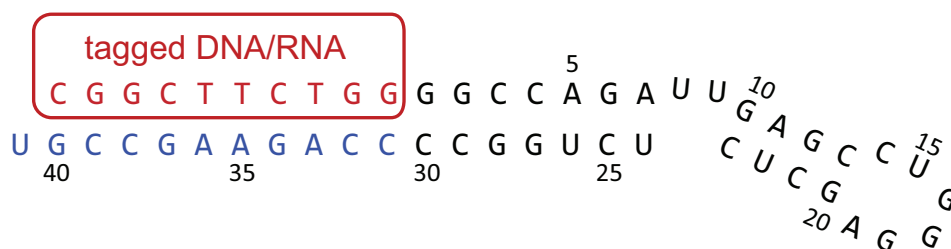


Figure 3.9: *Top:* Overview of synthesis and separation pathway for phosphorothioate RNA. *Bottom:* Sequence and HPLC chromatograms of two phosphorothioate RNA molecules modified at the labeled positions illustrate the successful separation procedure.

retention times of 10.2 min and 8.4 min between of the *R* and *S* conformer. In this approach, two HPLC runs were carried out. At first, the oligonucleotides were purified to remove protection groups and unwanted RNA fragments. Subsequently, the diastereomer separation was successfully performed for different sulfur positions.

For the tagging of OPS, the Cys-Ph-TAHA tag is not suitable as it is attached via a disulfide bond in contrast to the reported thioether linkage [Qin2001] [Qin2007]. Therefore, a significant modification of the tag was necessary, starting with an intermediate of the Cys-Ph-TAHA synthesis. In contrast to the original tag, the cysteine part was discarded, but the phenyl ring, which is crucial for the TAHAs synthesis, was retained [Peters2010]. The targeted tags each contain a halide atom for a nucleophilic substitution in benzylic position (see Fig.3.11).

Starting with a reduction of the intermediate **34** under mild conditions using isobutyl chloroformate and sodium borohydride, alcohol **35** was synthesized, which served as a starting material for various functionalisation. The most promising result was achieved



16

Figure 3.10: Elongated (*blue*) TAR RNA hairpin structure (*black*) with complementary paramagnetically tagged OPS (*red*).

by substituting the hydroxy group with a chloride moiety using thionyl chloride, followed by a straightforward deprotection using trifluoroacetic acid, giving compound **36** (Cl-Bn-TAHA) with a 39% yield over 3 steps. Furthermore, the iodine analogue **37** (I-Bn-TAHA) was synthesized with a 40% yield. Due to the significantly higher electrophilicity of iodine compared to chlorine, the corresponding I-Bn-TAHA tag was hydrolyzed in water within 1 h. The stability of the new Cl-Bn-TAHA was monitored using LCMS chromatography over the time range of a tagging reaction. At 7 °C, no hydrolysis product was observed and at room temperature, 80-90% of the tag was preserved. However, at 55 °C, the Cl-Bn-TAHA was hydrolyzed completely. Therefore, Cl-Bn-TAHA **36** was used for tagging reactions at room temperature on OPS.

Tagging reactions using the new *unloaded* Cl-Bn-TAHA tag and a 10-mer phosphorothioate DNA strand **38** were performed successfully and confirmed using HPLC chromatography and mass spectrometry. The chromatograms revealed, however, a low turnover rate (see Fig.3.12). In this approach, it was essential to maintain a high tag concentration (100 mM) and therefore an extensive excess (52 fold) was employed. No tagging product was found when lowering the reaction temperature to 7 °C. Ethanol and dimethylformamide were used as co-solvents in a 1 M MES buffer with pH=6. Considering the necessary purification procedure and the final amount of tagged oligonucleotide **39** that is required for one NMR sample, the excessive amount of tag (approx. 3 mg per 100 nmol DNA) posed a fundamental problem.

In the next step, the tagging reaction was performed using a *preloaded* tag. Unfortunately, this proved unsuccessful as the solubility of the compounds was found to be too low with respect to the crucial high tag concentration. Various solvents (DMF, EtOH, 1,4-dioxane) were tested without any success. Even though, phosphate buffer usually provides excellent solubility for oligonucleotides, its high affinity towards lanthanide ions resulted in an *unloaded* tag, so that its application was not an option. As described in section 1.3, *preloading* of the paramagnetic tags is inevitable as coordination with para-

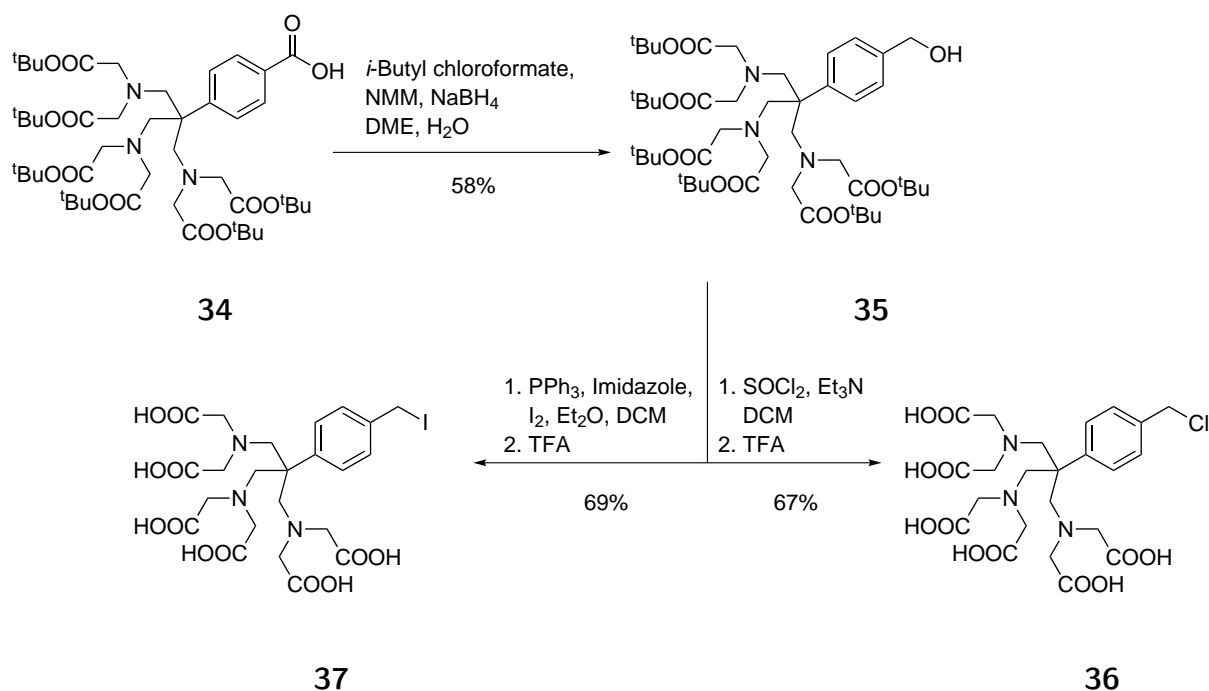


Figure 3.11: Reaction scheme for new paramagnetic tags **36** and **37**

magnetic lanthanide after the tagging reaction, leads to an incomplete loading of the tag [Peters2011]. In that case, the corresponding NMR spectra show two sets of signals, one for the diamagnetic and one for paramagnetic component.

At this point, efforts of tagging OPS had to be suspended due to the following reasons:

1. Incomplete tagging reaction despite a vast excess of tag.
2. Insufficient solubility of the *preloaded* tag while maintaining the required high concentration.
3. First promising tagging results using a modified nucleobase (see section 3.4).

In summary, a suitable decamer phosphorothioate RNA strand was found with convenient separation conditions. The original Cys-Ph-TAHA tag was successfully modified and a new Cl-Bn-TAHA tag was synthesized which was stable under tagging conditions. Phosphorothioate DNA was tagged successfully even though this approach did not lead to paramagnetically tagged oligonucleotides.

3.6 Modification of Cys-Ph-TAHA

The excellent properties of the Cys-Ph-TAHA tag for paramagnetic tagging originate from the triaminohexaacetate (TAHA) construct with its symmetry properties and its high

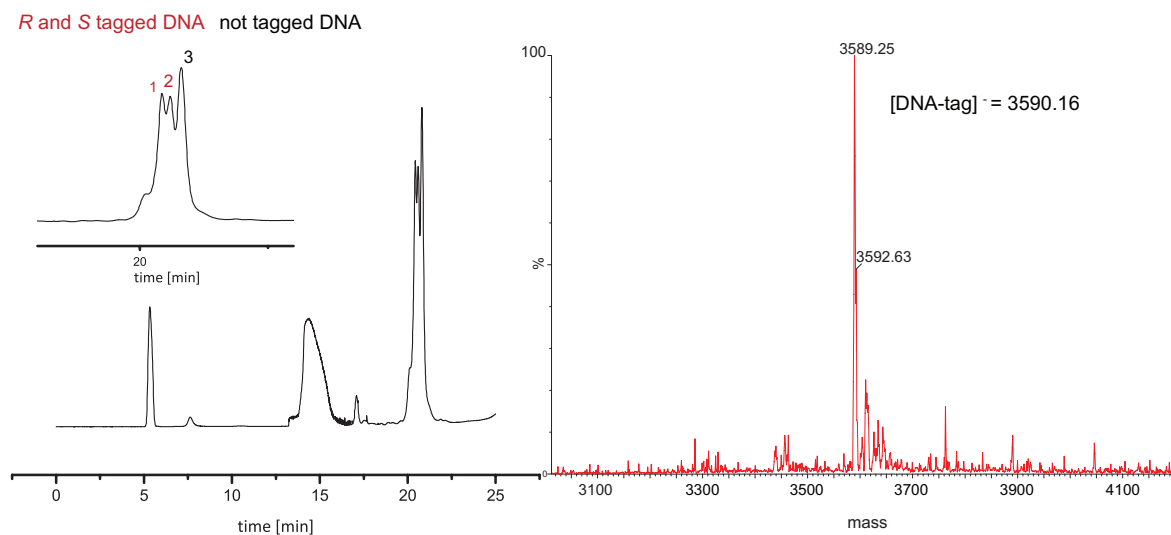
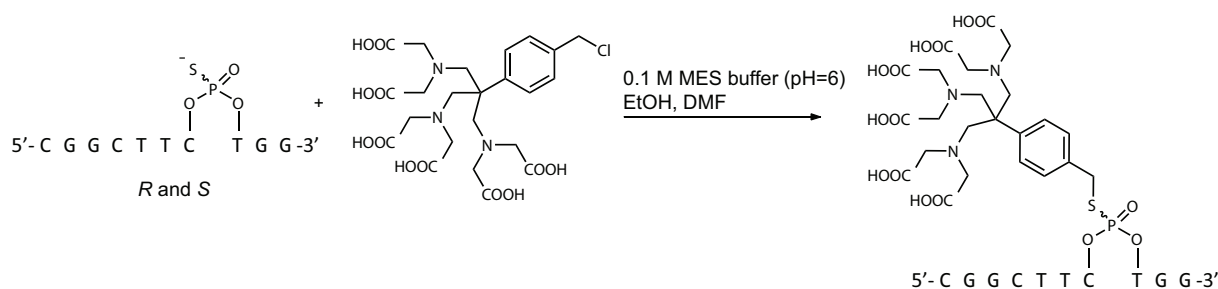


Figure 3.12: *Top:* Tagging reaction of phosphorothioate DNA **38** and Cl-Bn-TAHA **37**. *Bottom:* HPLC chromatogram showing 3 DNA peaks and mass spectrum of DNA peak 2 illustrating successful tagging and a low turnover rate.

affinity towards lanthanide ions [Viguiet2001]. As described in the literature, the phenyl ring is essential for the TAHA synthesis and in addition, it provides a rigid and axially symmetric linker [Peters2010]. The cysteine part of the tag, however, can be substituted by a more symmetric and less flexible linker or can even be entirely omitted. The shortest derivatives of the Cys-Ph-TAHA tag is shown in Fig.3.13. Therein, the cysteine has been removed and the sulfur binding site has directly attached to the phenyl moiety, leading to the highly rigid MesS-Ph-TAHA **40** tag. Furthermore, this variant does not include any functional groups that could interact with the target molecule. Even though the following

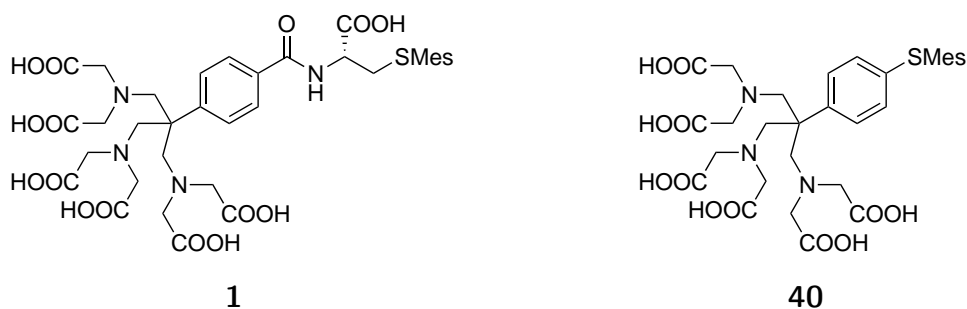


Figure 3.13: Cys-Ph-TAHA **1** and a shorter analogues MesS-Ph-TAHA **40**.

synthesis could not be completed, first results show that the presented synthetic pathway (see Fig.3.14) is very promising. The synthesis strategy of the new tag was based on the Cys-Ph-TAHA tag, while the essential TAHA fragment synthesis was maintained. The sulfur moiety was introduced during the TAHA synthesis, as previous attempts of a later introduction were unsuccessful [Siepel2013].

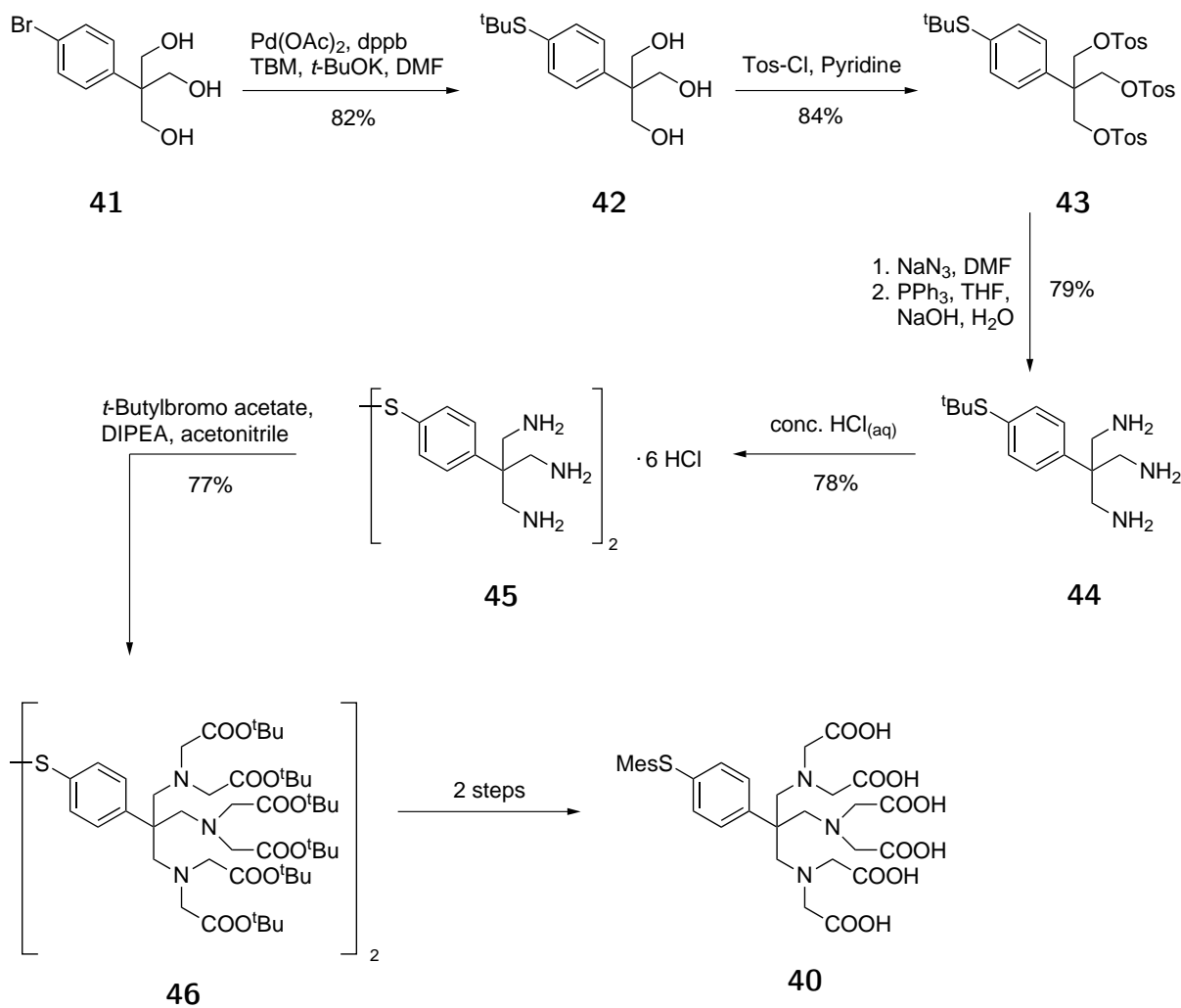


Figure 3.14: Scheme for modification of Cys-Ph-TAHA.

Starting from intermediate **41** of the Cys-Ph-TAHA synthesis, *tert*-butyl thiol was introduced via a palladium catalysed reaction with 1,4-bis(diphenylphosphino) butane, giving a good yield of 82% [Kawabuchi2008]. In the two following steps, the triamino compound **44** was synthesized. At first, the alcohol **42** was functionalized with tosyl groups, followed by a straightforward *Staudinger* reduction to afford compound **44** with a good yield of 79% over two steps. Subsequently, the *tert*-butyl protection group was removed using conc. hydrochloric acid. The TAHA synthesis was completed by a 12-fold alkylation with a yield of 77%, affording compound **46**.

Up until now, the activation of the disulfide **46** has not yet been achieved, according to literature reported procedures [Fujiki2002] [Haberz2006], but the synthesis of the TAHA fragment in presence of a sulfur moiety was established successfully. Compared to the Cys-Ph-TAHA synthesis, the reaction times of the individual steps were prolonged due to steric hindrance of the bulky *tert*-butyl group, but the yields of at least 77% for the alkylation are satisfactory. In addition, the purification procedures of this synthetic pathway are very convenient. No chromatography on silica was required prior to the tosylation reaction and compound **45** was afforded in its hydrochloride form, which was easily separated from apolar impurities. In the following steps, compound **46** shall be activated, followed by conventional cleavage of the *tert*-butyl groups and HPLC purification.

4 NMR Spectroscopy of Paramagnetically Tagged DNA

4.1 Sample Preparation

The paramagnetically loaded Cys-Ph-TAHA tag **1** was attached to the 24-mer DNA strand **15** based on the routine for the introduction of a binding site at a modified nucleobase, which is described in detail in section 3.4. The deoxythymidine in position 3 of the oligonucleotide was replaced by the modified nucleobase, yielding a free thiol moiety after DNA synthesis, which was subsequently connected to the preloaded tag. Figure 4.1 summarizes the synthetic approach that was employed for all NMR studies.

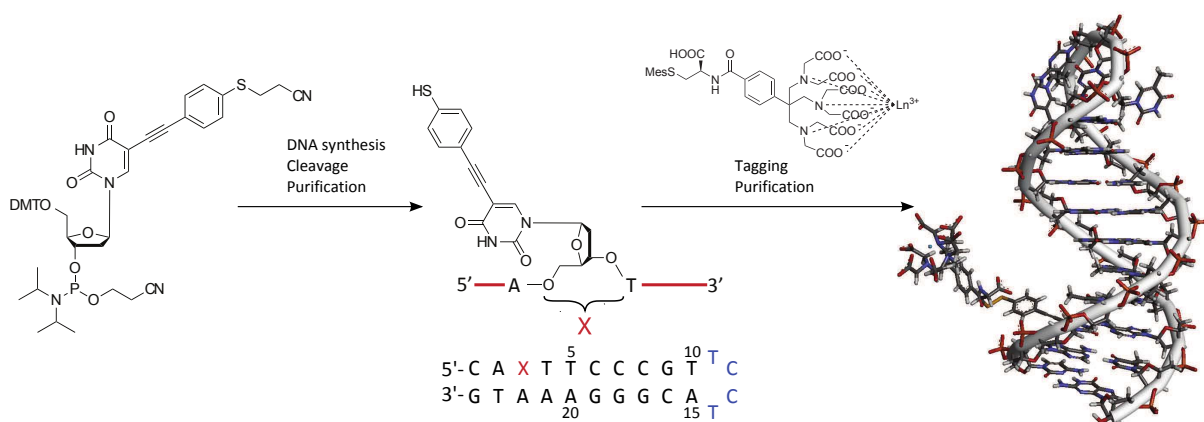


Figure 4.1: Established synthetic approach for the preparation of a paramagnetically tagged DNA strand, which is illustrated as a structure model (see section 4.2).

After HPLC purification, two additional steps were necessary to facilitate successful NMR spectroscopy. Due to the high affinity of lanthanide ions to the phosphate backbone, traces of the ions remain coordinated to the DNA molecule, which have to be removed completely. Randomly distributed paramagnetic ions at the backbone cause excessive line-broadening (see section 2.3), rendering NMR spectroscopy impossible, but also diamagnetic lanthanide ions need to be removed in order to guarantee natural sample conditions. Therefore, the DNA samples were repeatedly diluted with 1 M sodium chloride

solution in MOPS buffer at pH = 7.6, followed by concentration of the solution to a small volume using a *Millipore* Amicon concentration device (12 times) [Siepel2013]. Thereby, the excess of sodium chloride slowly displaced the lanthanide ions. During this step it is absolutely critical to avoid acidic conditions, as these would lead to a protonation of the tag's carboxy groups. Consequently, the lanthanide ions would also be removed from the Cys-Ph-TAHA tag, resulting in NMR signals from diamagnetic molecules in an originally paramagnetically tagged DNA sample. Unfortunately, the loss of lanthanide ions from the tag is not entirely preventable, yet the otherwise unwanted, smaller diamagnetic signals can still be used for calibration of the spectra.

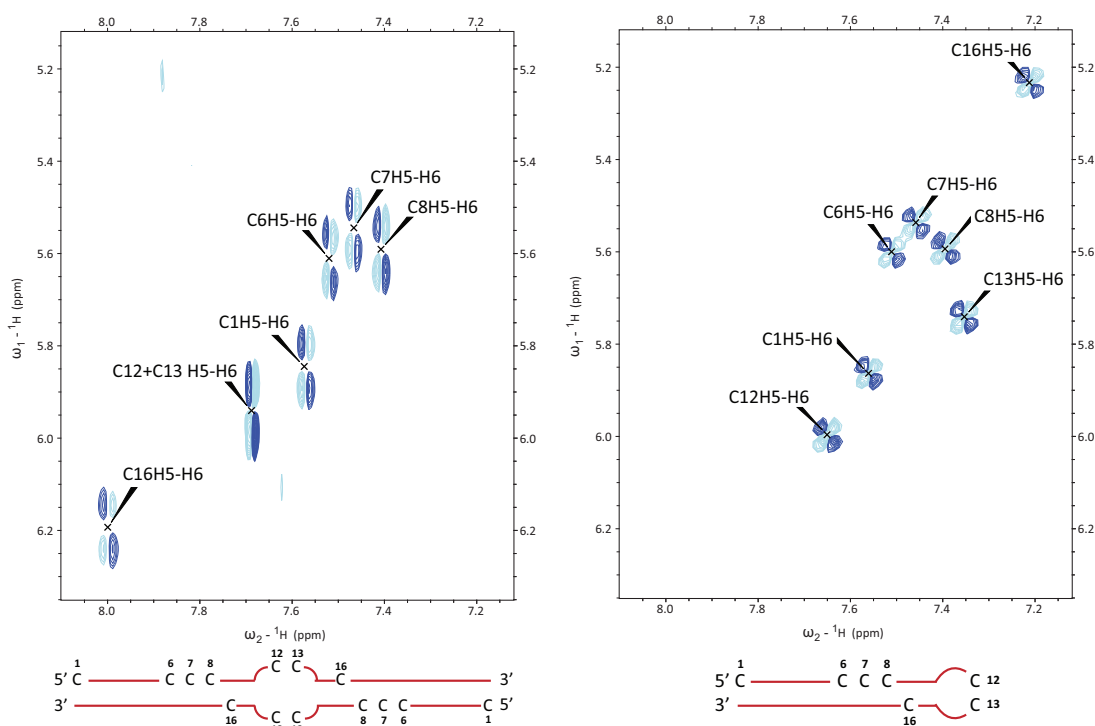


Figure 4.2: Folding of the DNA structure at low concentration leads to the expected chemical shifts for diamagnetically tagged DNA. *Left:* COSY spectrum of cytosine H5-H6 correlations folded at a high concentration with proposed assignment of a dimer structure. *Right:* After folding at a low concentration, the expected chemical shifts, which originate from the hairpin structure, were observed.

Compared to previous approaches (see section 1.2), the Cys-Ph-TAHA tag with its bulky coordination site is closer in space to the DNA strand. Therefore, an additional folding procedure was necessary in order to generate the hairpin structure of the tagged DNA strand. In COSY spectra, the characteristic *fingerprint* region shows the H5-H6 correlations in the aromatic cytosine system (see Fig.4.2). In the first attempts, only six peaks were observed for the seven cytosine residues due to a dimerisation of the hairpin structure, which presumably results in an overlap of the peaks from C12 and

C13. Chemical shift differences before and after correct folding are especially pronounced for C13 and C16, indicating significant structural changes between the monomer and the dimer in this region. Furthermore, Watson-Crick base pairing may have been broken for C16, as the chemical shift differences were considerably smaller for the other cytosine residues.

In order to generate the monomer hairpin structure, the DNA strand was re-folded at a low concentration by heating the diluted sample to 70 °C for 5 min, followed by gradually lowering the temperature. Finally, the seven expected NMR cross peaks for the diamagnetically tagged sample were observed (see Fig.4.2 (*right*)).

The sample preparation was completed by replacing the washing buffer with the final NMR buffer (20 mmol MOPS, 180 mM sodium chloride in deuterium oxide, pD = 8.0). The high salt concentration and the pD-value ensure the stability of the NMR sample, which can be used for several months. For the NMR studies, three differently tagged DNA samples were prepared using lanthanide ions with paramagnetic (thulium and terbium) or diamagnetic (lutetium) properties, generating paramagnetic centers or serving as a reference sample, respectively.

4.2 Measurement of Paramagnetically Induced NMR Parameters

Modification of biomolecules can lead to significant changes of their native structure. Even though it has been previously demonstrated that the introduced linker points out of the major groove [Woeltjen2009], the wild-type sample was compared to the diamagnetic reference, as a different linker might lead to a disruption of the DNA strand's original structure. Therefore, the chemical shifts of the different samples were analyzed and the differences are listed in Table 4.1. As expected, the largest chemical shift differences were determined for the residues directly attached to modified nucleobase and the adjacent cytosine 1. Overall, the agreement of the modified and the wild-type sample is satisfactory, so that significant changes of the native structure can be excluded.

For data analysis, two different structural models were employed. On the one hand, a set of five structures was used, which had previously been created based on NOE, angle, planarity and base pairing restraints [Taeubert2010]. On the other hand, an idealized B-DNA hairpin was generated without NOE information. To do so, a structure calculation was performed using restraints for planarity, base pairing and the characteristic dihedral angle of B-DNA [Roberts1993]. An overlay of the five NOE derived structures and the B-DNA hairpin is depicted in Fig.4.3.

Table 4.1: Chemical shift differences between the diamagnetic and the wild-type sample in ppm.

Residue	H1'	H2'	H2''	H3'	H2	H41	H42	H5	H6/H8	H7
C1	-0.05	0.26	0.17	0.09	x			0.05	0.10	x
A2	0.07	0.10	-0.20	0.07	0.22	x	x	x	0.34	x
X3	-0.16	-0.17	0.16	-0.06	x	x	x	x	-0.66	x
T4	-0.08	-0.06	-0.04	-0.04	x	x	x	x	-0.07	-0.06
T5	-0.02	0.02	-0.01	0.00	x	x	x	x	0.00	-0.01
C6	0.00	-0.02	0.01	-0.02	x	-0.02	0.00	0.00	-0.02	x
C7	-0.01	0.00	0.01	-0.01	x	0.00	0.01	-0.02	-0.01	x
C8	0.02	0.00	0.00	-0.01	x	-0.01	0.00	0.00	0.00	x
G9	0.00	0.00	0.00	0.00	x	x	x	x	0.00	x
T10	-0.01	overlap	overlap	0.00	x	x	x	x	0.00	-0.01
T11	overlap	overlap	overlap	0.00	x	x	x	x	0.00	0.00
C12	0.00	0.01	0.00	0.00	x			0.01	0.00	x
C13	0.00	0.00	0.00	0.00	x			0.02	0.01	x
T14	0.00	0.01	0.01	0.00	x	x	x	x	0.00	0.00
A15	0.01	0.00	0.00	-0.01	0.00	x	x	x	0.01	x
C16	0.01	0.01	0.01	overlap	x	-0.01	0.01	0.01	0.00	x
G17	0.04	0.02	0.02	0.01	x	x	x	x	0.01	x
G18	0.02	0.04	0.03	0.01	x	x	x	x	0.02	x
G19	0.08	0.03	0.04	0.01	x	x	x	x	0.02	x
A20	0.03	0.01	0.01	0.01	0.00	x	x	x	0.01	x
A21	-0.02	0.02	-0.02	-0.01	0.00	x	x	x	0.01	x
A22	0.00	0.01	-0.01	0.00		x	x	x	0.00	x
T23	0.03	overlap	overlap	0.02	x	x	x	x	0.04	-0.04
G24	0.01	-0.01	-0.03	-0.01	x	x	x	x	0.00	x

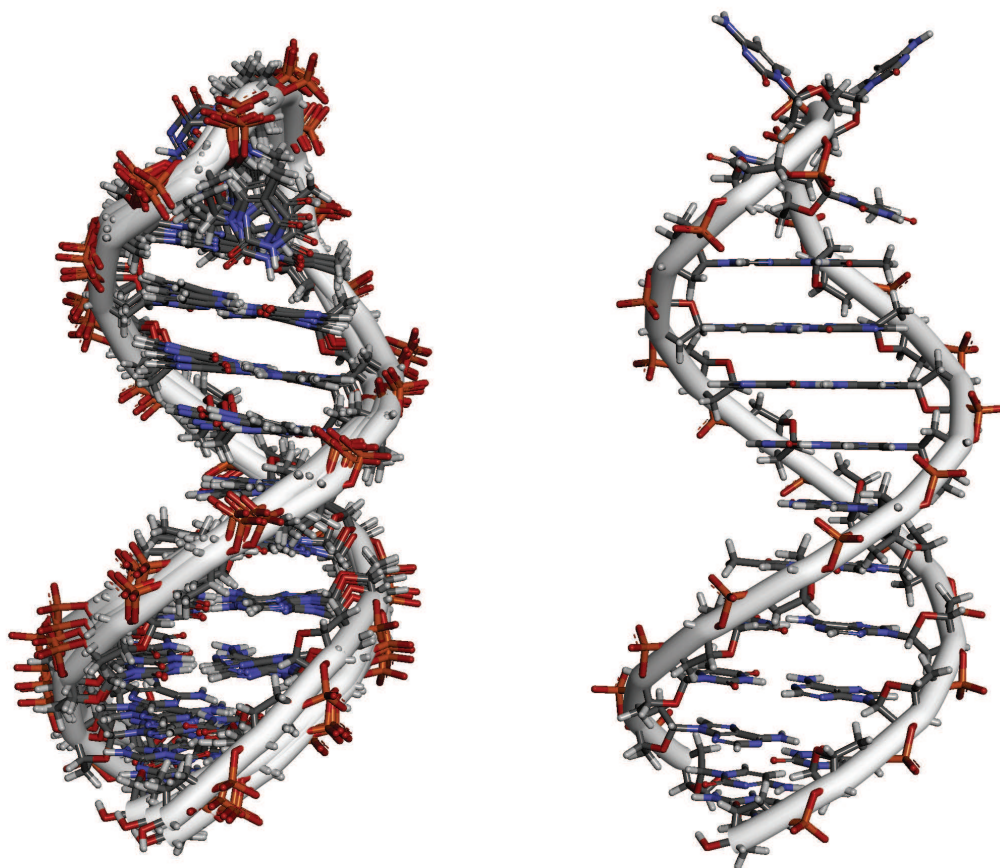


Figure 4.3: Different DNA models used for the structure calculation. *Left:* Overlay of the 5 NOE derived structures. *Right:* An idealized B-DNA model.

4.2.1 Pseudocontact Shifts

Depending on their individual properties, paramagnetic lanthanide ions generate a distinct additional magnetic field, resulting in different chemical shifts compared to a diamagnetic sample. The PCSs are highly distance dependent so that resonances in spatial proximity to the metal ion experience a stronger effect from the paramagnetic center.

The residues that are very close to the paramagnetic center can either not be observed or appear significantly smaller (*bleaching*), due to paramagnetic relaxation enhancement (PRE) (see section 2.3). These effects can be illustrated by a COSY spectra overlay of the *fingerprint* regions of the three differently tagged NMR samples (see Fig.4.4). Compared to the reference sample, the signals are shifted either upfield (thulium) or downfield (terbium), depending on the different signs of their respective susceptibility tensor [Pintacuda2007]. The PCS is of almost identical magnitude in both of the monitored dimensions, resulting in diagonal shifts of the H5-H6 cross peaks. The distance dependence of the PCSs is easily recognizable, as the shifts become smaller from C6 to C8,

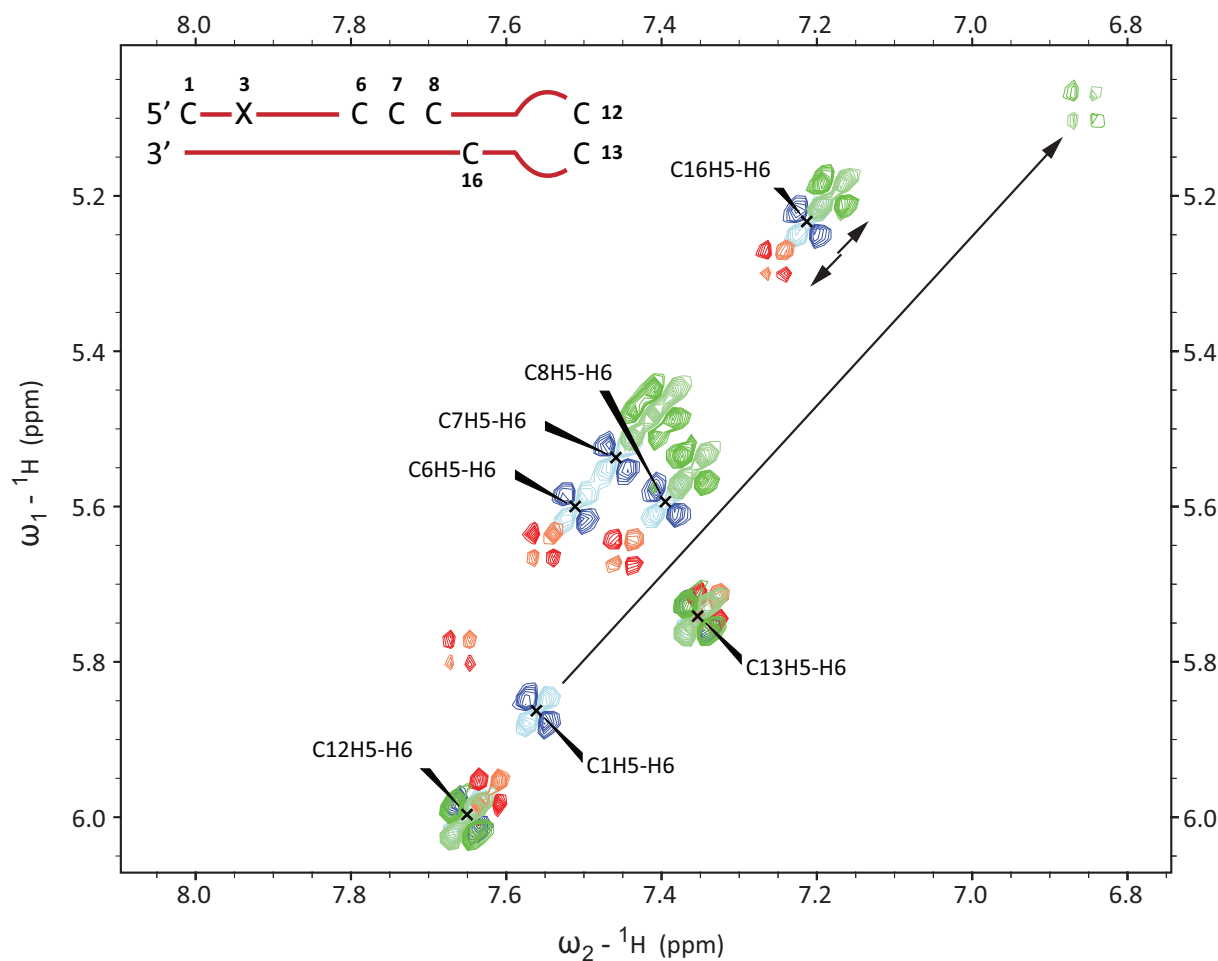


Figure 4.4: Overlay of the COSY *fingerprint*, showing PCSs in opposite directions for terbium (*red*) and thulium (*green*), compared to the reference sample (*blue*). The largest PCS measured for thulium (0.78 ppm for H5) is highlighted. The position of the modified nucleobase is labeled.

attributing to an increasing distance from the tag. In the loop region, hardly any PCSs were observable.

The PRE significantly complicates the correct assignment of the paramagnetic spectra, which is conventionally performed by a so-called NOESY-walk in diamagnetic DNA spectra. Correlations between aromatic base protons (H8 for G and A, H6 for C and T) and H1' sugar protons attached to the same and the (n+1) nucleoside give rise to strong signals, leading to a full *walk-through* along the DNA sequence. In addition, the intra-nucleoside peak is usually larger than the inter-nucleoside signals due to the distance dependence of the NOE. Unfortunately, paramagnetic DNA samples lack this features as many signals can not be observed at all or their intensity is manipulated. Therefore, the assignment was achieved by comparing the chemical shifts of the Tm and Tb samples to the reference sample. This procedure is illustrated by the NOESY overlay in Fig.4.5, in which the paramagnetic and diamagnetic peaks form a characteristic diagonal line for

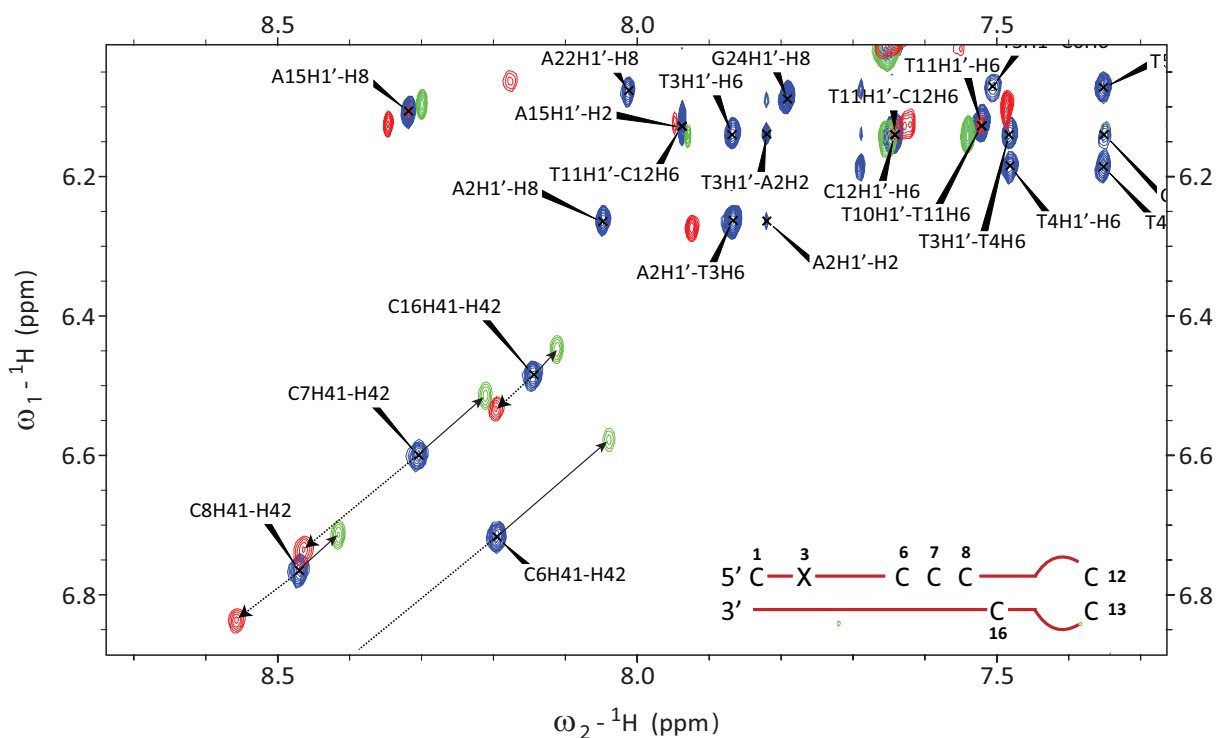


Figure 4.5: NOESY spectra overlay of the terbium (*red*), thulium (*green*) and lutetium sample (*blue*) illustrating the diagonal line that can be observed for the three corresponding signals.

the cytosine amine protons (C6, C7, C8 and C16). The more crowded NOESY-walk region (H1'-H6(H8) correlations) at around 6.2 ppm shows that no paramagnetically shifted peaks were observed for some of the cross peaks, e.g. A2H1'-H8 or A22H1'-H8. Figure 4.6 shows the NOESY spectra overlay (Tm- and Lu-sample) of the sugar protons (H1'-H2'(H2'')), by which most of the PCSs have been determined. Vital information for the peak assignment was obtained by comparing the chemical shift differences between H2' and H2'' protons for two residues, e.g. A15 and A22. Due to the chemical shift perturbation of sugar protons H3', H4' and H5' (H5''), no PCSs were determined for these residues.

In total, 123 PCSs were determined (56 for terbium, 67 for thulium) using COSY and NOESY spectra (see Table 4.2 and 4.3). Due to the stronger PRE caused by terbium, more signals are expected to be observed in the Tm than in the Tb spectra. Terbium induced the largest PCS of 1.21 ppm for cytosine 1 (0.78 ppm for thulium) and consequently a stronger PRE for the closest nuclei. Apart from C1, the residue G18 also experience a large PCS and G19 was not even observable in the thulium spectra, indicating a close distance to the metal ion. Due to the higher PRE, the resonances of these residues were not found in the Tb spectra. For residues G9 and A15, which are located in the helix, no significant shifts were determined, underlining the local character of the PCS. The loop

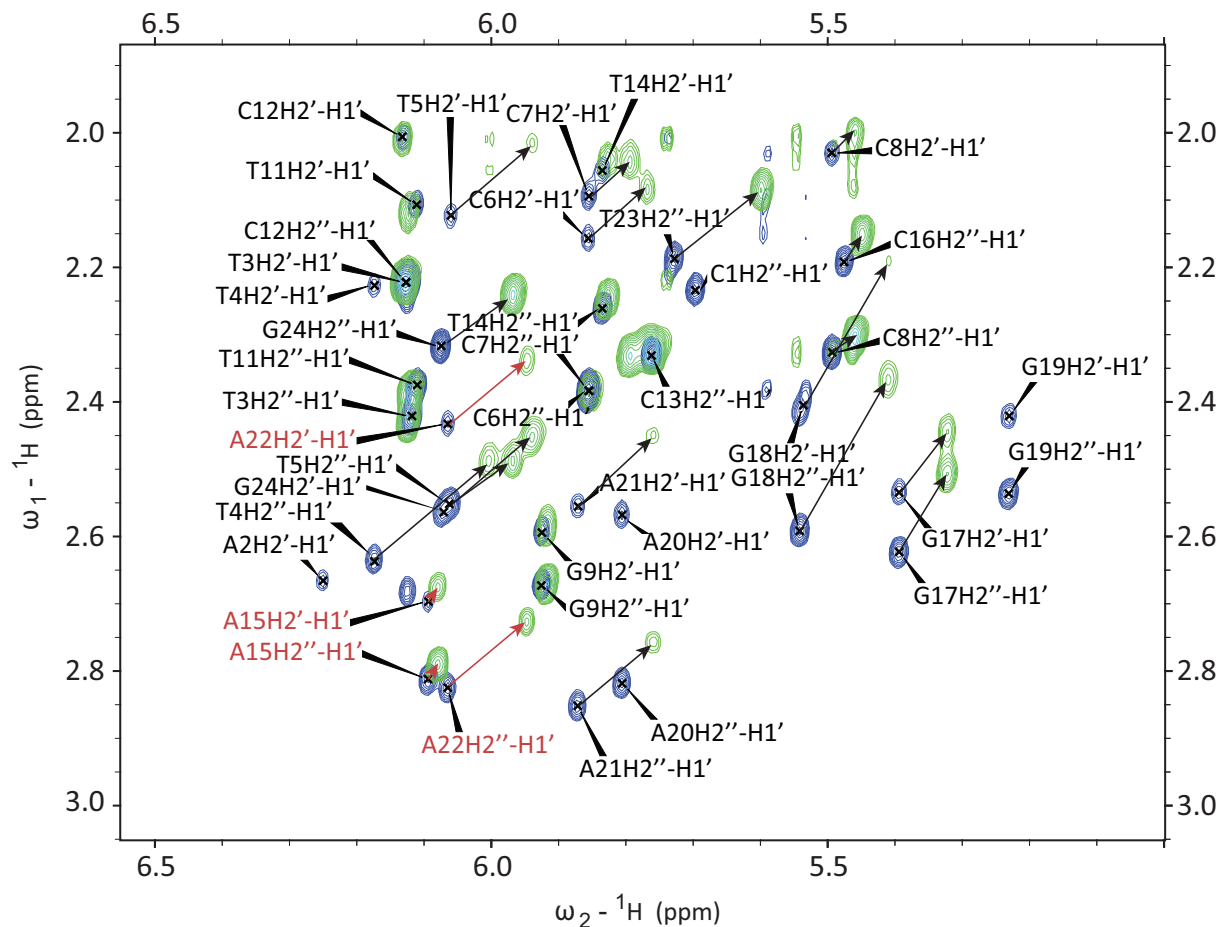


Figure 4.6: NOESY spectra overlay of the thulium (*green*) and lutetium sample (*blue*) showing PCSs of the H1'-H2'(H2'') correlation. The highlighted resonances illustrate the assignment strategy based on the differences of the chemical shifts of H2' and H2''.

region was disregarded due to its flexibility and, moreover, no PCSs were observed for these residues.

For the PCS analysis, only unambiguously assigned PCSs were used. Therefore, it was particularly important to perform a correct assignment, even though not all signals for the paramagnetically tagged samples were found. In Fig.4.7, the distribution of the measured PCSs in the hairpin is highlighted.

The PCS data was analyzed together with the RDC data, which is presented in section 4.3.

Table 4.2: Determined thulium-induced PCSs in ppm.

Residue	H5	H6/H8	H41	H42	H1'	H2'	H2''
C1	-0.78	-0.71			-0.50	-0.73	-0.65
T5	x		x	x	-0.12	-0.11	-0.10
C6	-0.10	-0.09	-0.14	-0.16	-0.09	-0.07	
C7	-0.06	-0.06	-0.09	-0.09	-0.06	-0.05	-0.05
C8	-0.04	-0.04	-0.05	-0.05	-0.03	-0.03	-0.02
G9	x	-0.01	x	x	-0.01	-0.01	0.00
A15	x	-0.02	x	x	-0.02	-0.02	-0.02
C16	-0.03	-0.03	-0.04	-0.03	-0.03	-0.04	-0.04
G17	x	-0.07	x	x	-0.07	-0.09	-0.12
G18	x		x	x	-0.13	-0.22	-0.23
A20	x		x	x	-0.13	-0.15	-0.14
A21	x		x	x	-0.11	-0.11	-0.10
A22	x		x	x	-0.12	-0.09	-0.10
T23	x	-0.10	x	x	-0.13	-0.09	-0.10
G24	x	-0.08	x	x	-0.10	-0.08	-0.08

Table 4.3: Determined terbium-induced PCSs in ppm.

Residue	H5	H6/H8	H41	H42	H1'	H2'	H2''
C1	1.21	1.09					
T5	x		x	x	0.21	0.18	
C6	0.19	0.16	0.24	0.28	0.15	0.12	0.12
C7	0.12	0.10	0.14	0.16	0.10	0.08	0.07
C8	0.07	0.06	0.07	0.09	0.06	0.04	0.04
G9	x	0.02	x	x	0.01	0.00	0.00
A15	x	0.03	x	x	0.02		0.02
C16	0.05	0.05	0.05	0.05	0.03	0.05	0.05
G17	x	0.12	x	x	0.12	0.13	0.21
A21	x		x	x	0.18		0.13
A22	x		x	x	0.20	0.14	0.15
T23	x	0.16	x	x	0.23	0.16	0.17
G24	x	0.13	x	x	0.19	0.14	0.14



Figure 4.7: Distribution of the determined PCSs for the Tm (*green*) and Tb (*red*) sample, which is illustrated in the hairpin structure.

4.2.2 Residual Dipolar Couplings

Residual dipolar couplings (RDCs) have previously demonstrated excellent properties for structure determination, as they provide global structural information about the target molecule [Zhang2006] [Bothe2011] [Russo2013]. Lanthanide ions with paramagnetic properties align along the magnetic field of the NMR spectrometer, causing a disruption of the free tumbling of the molecules and thereby, enabling the measurement of RDCs. Even though paramagnetically induced RDCs were successfully used for structure determination and refinement in proteins, a transfer of this application to oligonucleotides has not yet been possible.

Due to the distance dependency ($\propto r^{-3}$), only nuclei that are close in space can give rise to a significant RDC. Therefore, the proton-proton couplings of cytosine H5 and H6, which are 2.4 Å apart, were analyzed by a combination of NOESY and COSY spectra. The *Differences and Sums of Traces within COSY spectra* (DISCO) method is an effective tool to measure RDCs of these aromatic protons [Kessler1985]. In aligned DNA strands, two separable parts of the proton-proton couplings are present in the NOESY spectra, i.e. the scalar 3J coupling and the dipolar coupling D through space. In the COSY spectra, however, only the scalar coupling is determined, which facilitates the extraction of the dipolar contribution by addition or subtraction of the individual COSY traces to or from

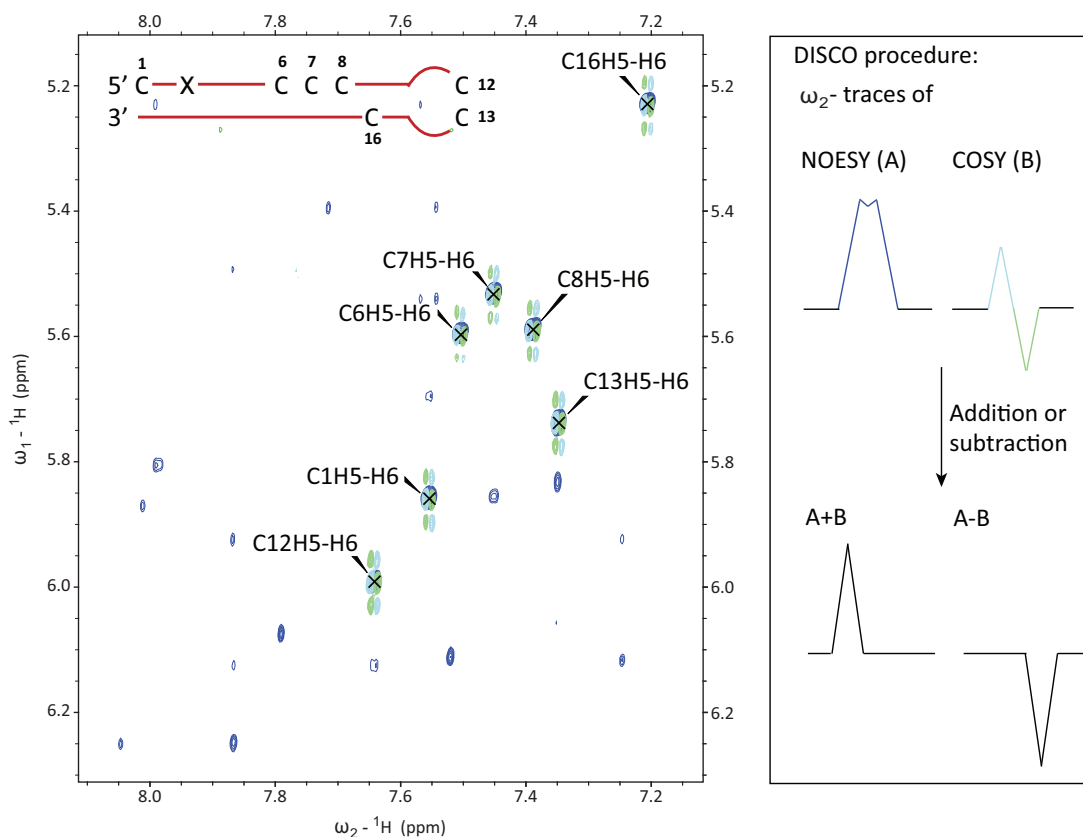


Figure 4.8: *Left:* Overlay of NOESY and COSY spectra of the diamagnetic sample to determine dipolar couplings of cytosine H5-H6 cross peaks. *Right:* Schematic of DISCO procedure showing the NOESY (A) and COSY (B) ω_2 -traces of C7 in order to extract RDCs by addition (A+B) and subtraction (A-B).

the traces of the NOESY peaks (see Fig.4.8 (*right*)).

The DNA hairpin molecule contains seven cytosine residues, which were analyzed for the thulium and the reference sample at a magnetic field of 900 MHz, using the DISCO method. An overlay of the corresponding NOESY and COSY spectra is shown in Fig.4.8 (*left*).

The determined RDCs of up to 2.2 Hz were too small for a valid evaluation (see Table 4.4). For the reference sample, which exhibits a considerably higher signal-to-noise ratio and superior line-shapes of the NMR signals, the DISCO protocol was performed for a second time, resulting in significant errors of 0.2 - 0.9 Hz. Based on the PCS measurement, RDCs of max. 2.8 Hz were predicted, which is generally in agreement with the data.

The hetero-nuclear proton-carbon RDCs were also analyzed. Compared to the cytosine H5-H6 couplings, these nuclei are closer in space (1.1 Å) and can be found throughout the entire DNA strand. At a magnetic field of 800 MHz, a max. RDC of 6.2 Hz was predicted based on the PCS data. The RDCs were measured using HSQC IPAP (*in-phase anti-phase*) spectroscopy, in which the hetero-nuclear couplings have been evolved into the

Table 4.4: Determined RDCs of cytosine H5-H6 correlations using the Tm-tagged DNA sample.

Residue	RDC [Hz]
C6	1.6
C7	0.8
C8	1.3
C12	1.8
C13	1.5
C16	2.2

indirect dimension. Compared to a conventionally decoupled HSQC spectrum, HSQC IPAP is a combination of two separate spectra, in which the peaks are either in-phase (A) or anti-phase (B). Addition (A+B) or subtraction (A-B) results in two different spectra, in which the originally decoupled HSQC signals are shifted $\pm 1/2 J$ [Ottiger1998].

The measurement of RDCs in ^{13}C labeled protein samples using the Cys-Ph-TAHA tag has demonstrated the potential of this method [Peters2011], but a successful application on paramagnetically tagged oligonucleotides has of yet been unsuccessful. Due to expensive ^{13}C labeled phosphoramidites and the low overall yield of synthetic oligonucleotides, labeling of synthetic oligonucleotides is unreasonable. In contrast, the recorded spectra of samples that only exhibit natural abundance of ^{13}C show a low signal-to-noise ratio. In Fig.4.9 (*right*), the in-phase HSQC of the cytosine C6-H6 couplings and the trace of

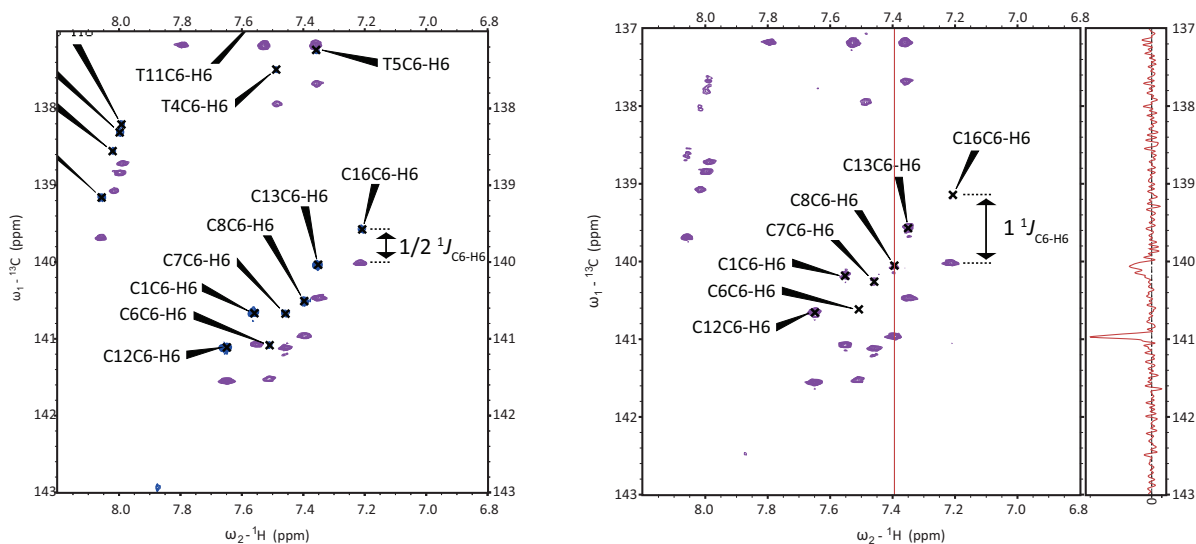


Figure 4.9: *Left:* Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Lu sample to determine $1/2$ of the coupling. *Right:* In phase part of the HSQC IPAP spectrum with the highlighted ω_1 -trace of C8 that demonstrates the different peak intensities in the indirect dimension.

C8 are shown. The intensity difference between the two in-phase peaks originates from different relaxation rates of ^{13}C when coupled to a proton that is either in the α or the β state. Therefore, RDCs have been determined by comparing the decoupled HSQC peaks at the center and the stronger peaks from the A+B part of the IPAP spectra, illustrated in Fig.4.9 (*left*), yielding $1/2$ of the coupling constant.

Using this method, the aromatic (C6-H6 or C8-H8) and the sugar (C1'-H1') couplings of both the Tm and the reference sample were determined on a 800 MHz spectrometer. To do so, one HSQC IPAP and one decoupled HSQC were recorded for the two different regions of the spectrum, requiring one week of measurement time for each sample. Overlays of the HSQC spectra are shown in Fig.4.11 and 4.12 for the reference sample and in Fig.4.13 and 4.14 for the Tm sample. Generally, the sugar and the aromatic region can be recorded together in one spectrum, using an average delay for the INEPT transfer, yet the individual measurements have proven to give superior signal-to-noise ratio. To optimize the RDC measurement, the traces of the peak maxima have been added to the adjacent traces of approximately 60 % of the peak intensity in order to increase the signal-to-noise ratio. The final NMR data set was independently analyzed by three different persons, resulting in the RDCs, which are shown in Table 4.5.

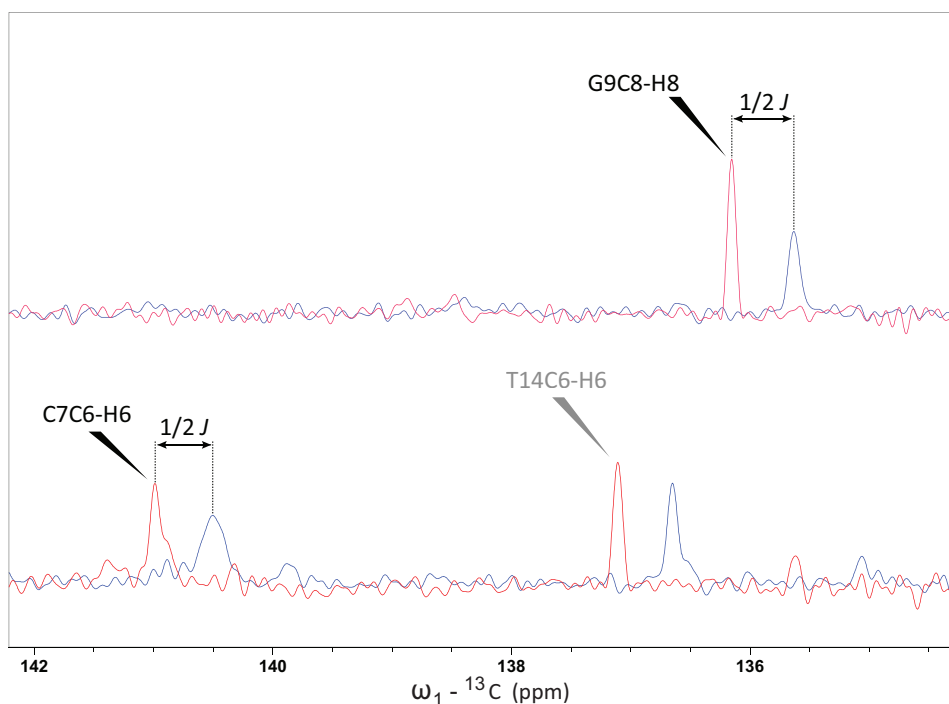


Figure 4.10: Overlay taken from ω_1 -traces of decoupled HSQC (*blue*) and HSQC IPAP (*red*) spectra of the Tm sample to determine the coupling constants of residues G9 (*top*) and C7 (*bottom*). Due to the asymmetric lineshape, the coupling constant of C7 was discarded in the structure calculations.

In total, 29 RDCs were determined for the Tm sample, but only 15 RDCs were employed in the final data analysis, resulting in a Q -factor of 0.28 over the three individually performed peak picking procedures. Therefore, the standard deviations of the diamagnetic and the paramagnetic coupling constant determination were multiplied by two and summed up to give an overall error, which was compared to the RDC values according to equation 2.12. The 14 remaining RDCs were discarded for different reasons. As only 1/2 the value of the coupling constants was determined and subsequently multiplied by two in order to calculate the actual coupling constant, this procedure is highly sensitive towards small errors. The RDCs of the loop region were rejected, as the quality of the structure models might not be sufficient for this flexible region. Other peaks were discarded due to their low signal-to-noise ratio (G17C1'-H1', G18C8-H8 and G24C8-H8) or lineshape errors in the paramagnetic (C6C1'-H1', C7C6-H6) or diamagnetic spectra (A15C8-H8) (e.g. see Fig.4.10). A maximum RDC of $D_{C1'-H1'} = -6.6$ Hz was found for the sugar coupling of G18 and the highest RDC for the aromatic coupling was $D_{C8-H8} = -4.4$ Hz for residue G9. The data analysis is presented in the following section 4.3.

Table 4.5: Residual dipolar coupling constants determined for the Tm-tagged DNA strand. The values highlighted in grey were omitted in the structure calculation due to low signal-to-noise ratios or lineshape errors. The loop region was disregarded due to its deviation from the structure models.

Residue	Resonance	RDC
C1	C1'-H1'	-1.3
C6	C1'-H1'	0.2
C7	C1'-H1'	1.6
C7	C6 -H6	13.8
C8	C1'-H1'	4.2
C8	C6 -H6	-2.0
G9	C1'-H1'	-2.8
G9	C8 -H8	-4.4
T10	C1'-H1'	-4.1
T10	C6 -H6	-0.1
T11	C1'-H1'	-1.8
T11	C6 -H6	2.7
C12	C1'-H1'	0.3
C12	C6 -H6	-1.5
C13	C1'-H1'	3.4
C13	C6 -H6	-3.7
T14	C1'-H1'	0.8
T14	C6 -H6	5.1
A15	C1'-H1'	-0.2
A15	C8 -H8	-3.5
C16	C1'-H1'	-0.5
C16	C6 -H6	-4.1
G17	C1'-H1'	-9.6
G17	C8 -H8	-2.4
G18	C1'-H1'	-6.6
G18	C8 -H8	4.0
A22	C1'-H1'	-2.0
T23	C1'-H1'	-6.0
G24	C8 -H8	1.7

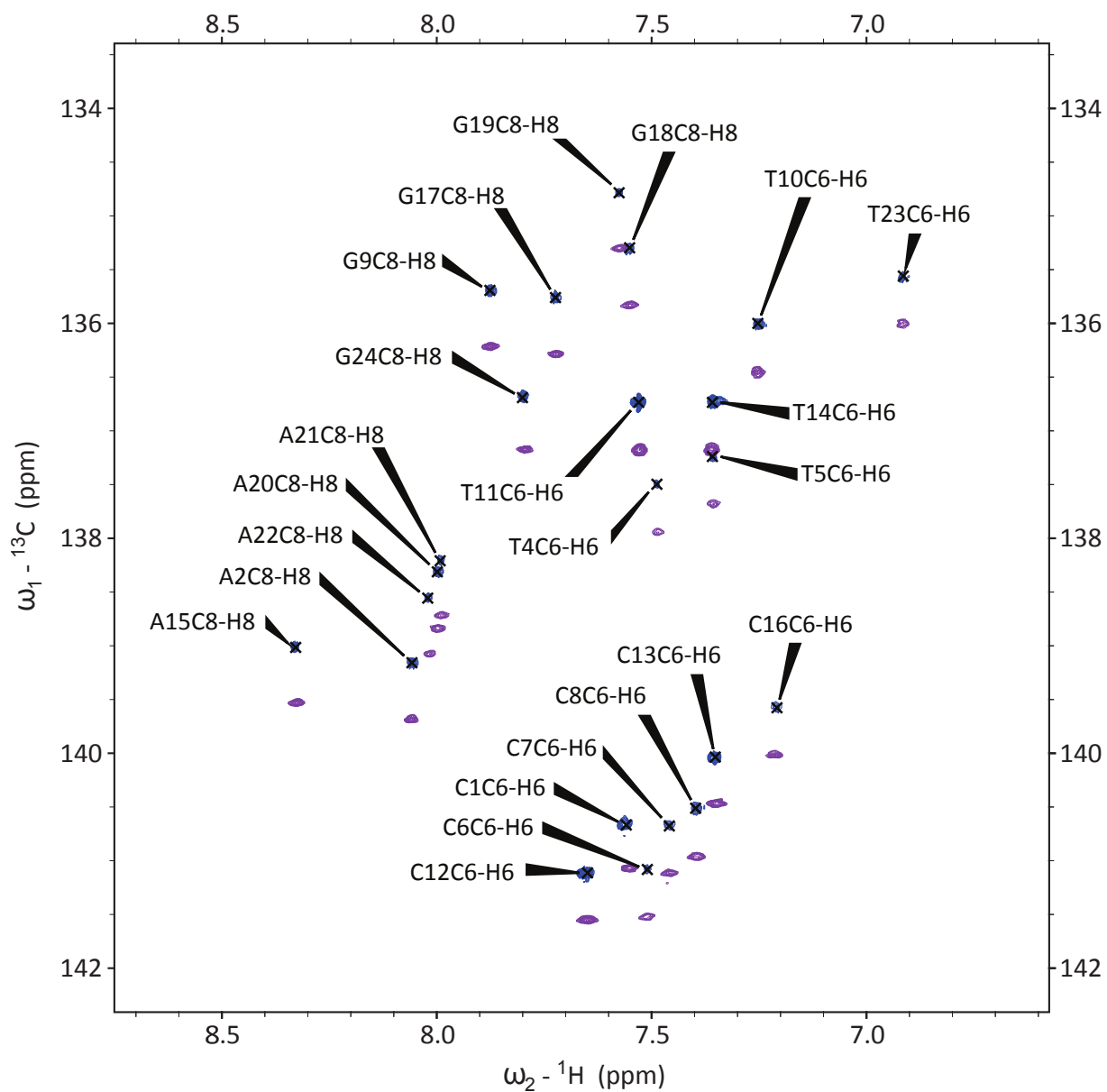


Figure 4.11: Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Lu sample to determine the coupling constants for the C6-H6 (C8-H8) resonances.

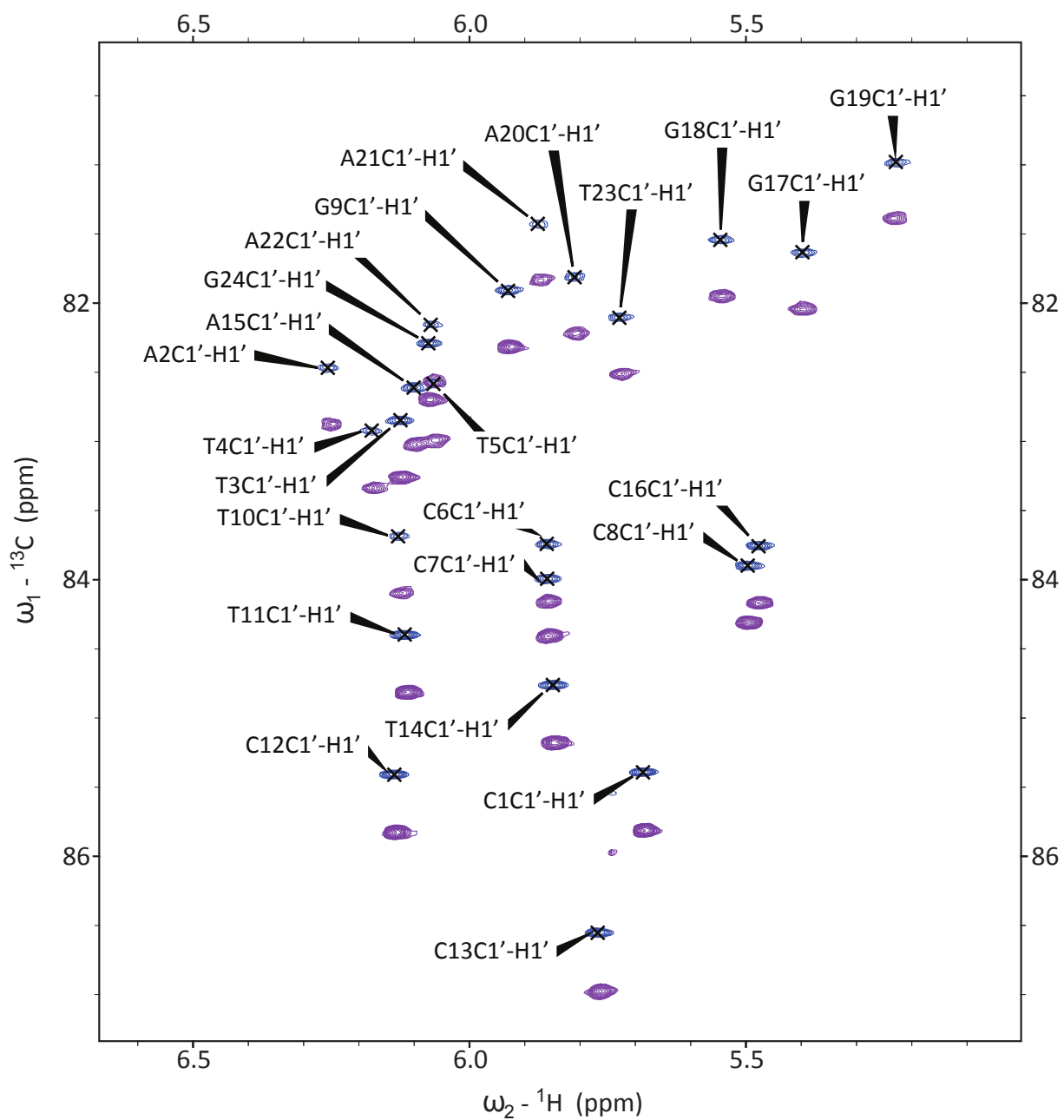


Figure 4.12: Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Lu sample to determine the coupling constants for the C1'-H1' resonances.

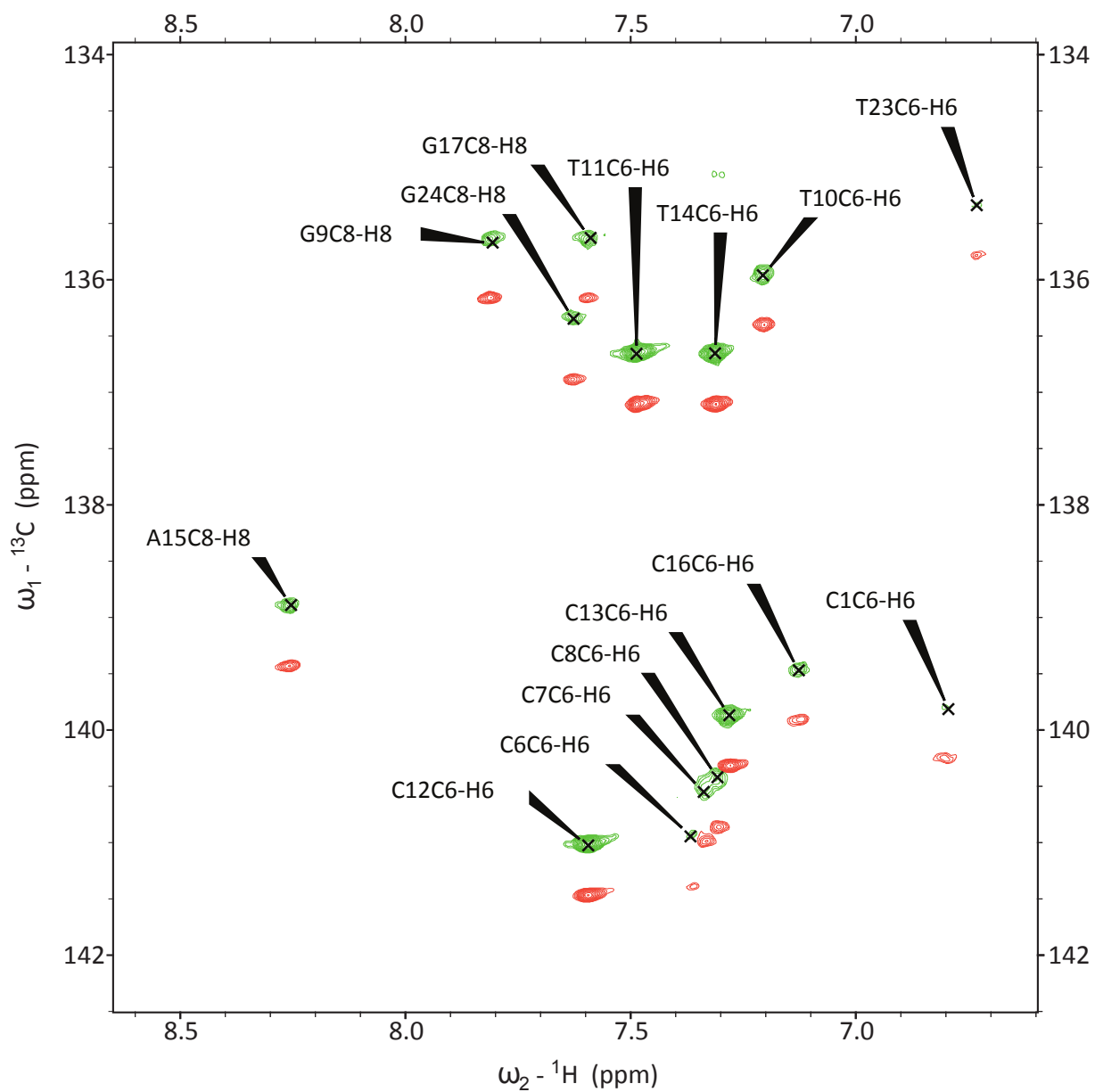


Figure 4.13: Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Tm sample to determine the coupling constants for the C6-H6 (C8-H8) resonances.

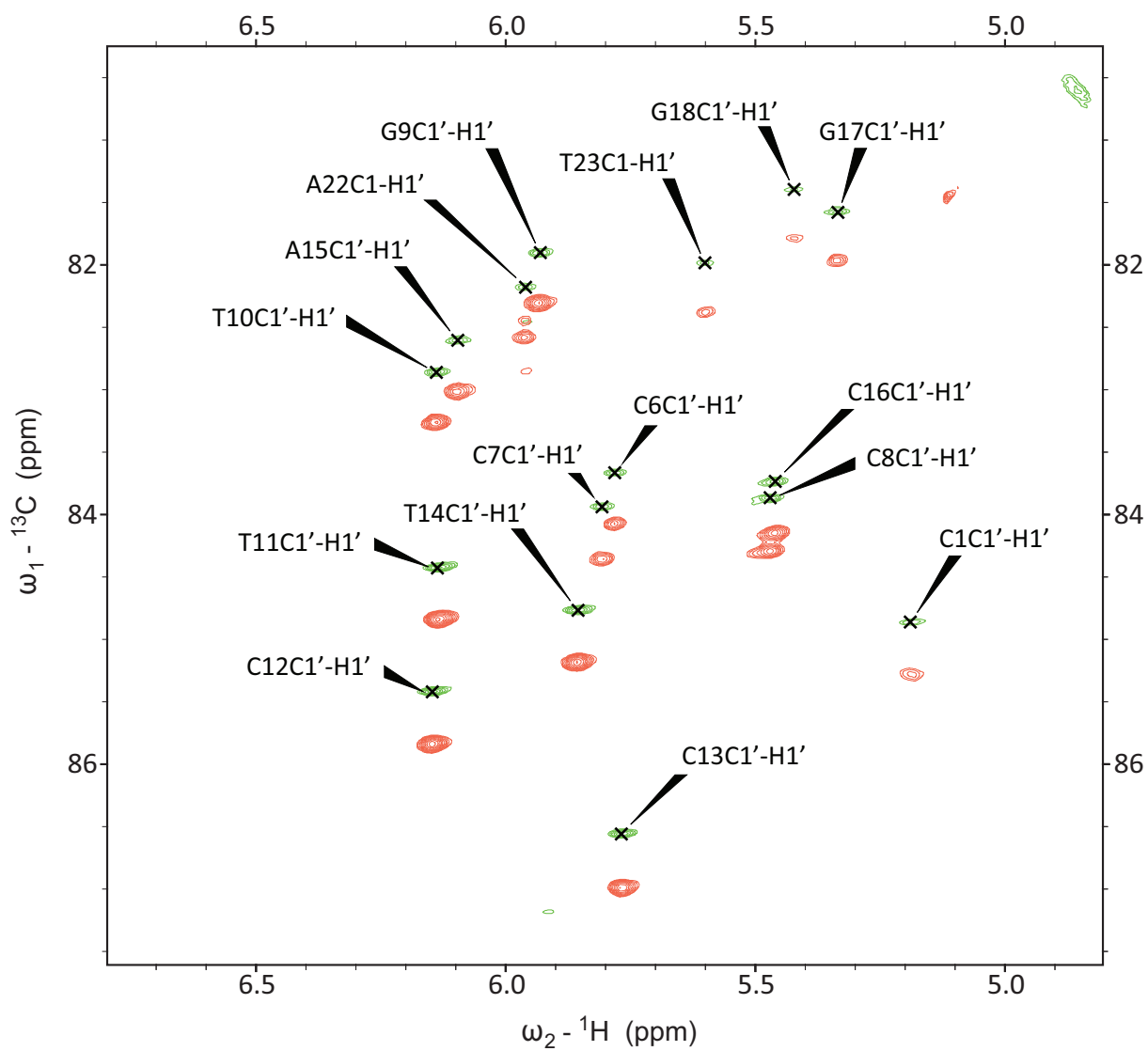


Figure 4.14: Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Tm sample to determine the coupling constants for the C1'-H1' resonances.

4.3 Data Evaluation and Discussion

The collected data was analyzed by Dr. Mitchell Maestre Martinez using in-house written Python scripts (see section 6). At first, the NOE structures (NOE 1-5) and the idealized B-DNA strand (see Fig.4.3) were used for an evaluation without a structural model of the Cys-Ph-TAHA tag, the binding site or their conformations (see section 4.3.1). A common metal position and a tensor were calculated for the different starting structure coordinates. Therein, the lanthanide ion was shifted in a grid search around the DNA strand, to find a suitable metal position. For every position, theoretical values were predicted and compared to the experimental data using equation 2.12. The resulting Q -factors were analyzed and the metal position changed until a minimum was found.

In a second calculation approach for the combined PCS and RDC data of the Tm-tagged DNA molecule, a model of the tag, the binding site and its conformations were implemented (section 4.3.2).

4.3.1 Evaluation of the PCS and RDC Data Without a Structural Model of Cys-Ph-TAHA

For the terbium- and thulium-induced PCSs satisfactorily low Q -factors (0.08 for Tb and 0.12 for Tm) were calculated, which illustrates the correct assignment of the paramagnetic spectra (see Table 4.6). The highest deviation between experimental and back-calculated PCSs was found for the C1H5 and C1H6 resonances of the Tm sample (see Appendix, Table 6.6). These resonances were, however, unambiguously assigned using the COSY *fingerprint* region (see Fig.4.4). A strong mismatch of the starting structures can be ruled out, as the same resonances do not give rise to significant errors in the Tb sample. Apart from small variations for the Tm sample, the results are identical for the different models. As this analysis was performed with an unrestrained search for the metal position, the resulting Q -factors represent the best fitting values achievable from the PCS data. The alignment tensors and the deviation between experimental and back-calculated PCSs are given in the Appendix.

In the next step, the RDC data was analyzed independently. The alignment tensors and the Q -factors were calculated for the different DNA models (see Table 4.7 and Appendix, Table 6.10 and 6.11), yet the resulting Q -factors of 0.36 for the best NOE structure and 0.41 for the idealized B-DNA are significantly higher compared to the literature (Q -factor of 0.18) [Peters2011]. This increased error originates from the coupling constant determination using unlabeled NMR samples. The conventional examination procedure is performed using both parts of the HSQC IPAP spectra, which had to be modified so that

Table 4.6: PCS-derived metal positions in the molecular frame and Q -factors of Tm and Tb.

Structure	Metal	PCSs	Lanthanide position			Q -factor
NOE 1	Tb	56	-3.11	13.75	11.50	0.08
NOE 2	Tb	56	-3.60	14.19	11.18	0.08
NOE 3	Tb	56	-4.52	14.23	11.32	0.08
NOE 4	Tb	56	-4.76	14.59	10.28	0.08
NOE 5	Tb	56	-5.23	16.10	7.23	0.08
B-DNA	Tb	56	-3.64	18.11	0.42	0.08
NOE 1	Tm	67	-3.45	15.42	16.59	0.13
NOE 2	Tm	67	-1.77	22.07	16.70	0.12
NOE 3	Tm	67	-2.86	20.31	16.88	0.12
NOE 4	Tm	67	-3.07	17.67	15.58	0.12
NOE 5	Tm	67	-2.43	18.63	13.17	0.13
B-DNA	Tm	67	-1.55	20.29	-1.43	0.11

only 1/2 of the coupling constant was actually determined. This results, however, in an increased sensitivity towards small phase errors or peaks with unsymmetrical lineshapes, due to the subsequent multiplication by two, which is performed in order to determine the actual coupling constant. Unfortunately, no alternative to this procedure was applicable, due to the very low peak intensity of one of the CH doublet lines (see Fig. 4.9). The RDC determination results in a high error of the peak picking (Q -factor = 0.28). A further error was caused by the spectral resolution of 0.7 Hz, which was partly caused by the relaxation enhancement in paramagnetic samples, which leads to fast signal decay in the indirect dimension and consequently to a relatively low FID resolution.

Table 4.7: Q -factors of RDC analysis.

Structure	Metal	RDCs	Q -factor
NOE 1	Tm	15	0.36
NOE 2	Tm	15	0.41
NOE 3	Tm	15	0.42
NOE 4	Tm	15	0.39
NOE 5	Tm	15	0.41
B-DNA	Tm	15	0.41

In general, there are two major problems that prevent a more accurate determination of dipolar couplings in this approach. On the one hand, the measurement of unlabeled samples requires an extensive amount of measurement time, yet the signal-to-noise ratio is still not always sufficient. When conducting the experiments, the NMR spectrometer

and especially the shims needs to be stable in order to guarantee a spectrum with an adequate quality. This appears to be highly challenging under paramagnetic conditions for longer measurements. As a result, the spectral resolution and low signal-to-noise ratio that is exacerbated by PRE, give rise to a significant error. On the other hand, the determined RDCs have proven to be smaller than expected (see section 2.1), indicating a higher flexibility of the tag compared to the binding sites in previous studies [Peters2011]. Consequently, the errors affect the accuracy of the RDC measurement more significantly.

The individual analyses of the PCSs and RDCs show that the different starting models are leading to very similar results, allowing the conclusion that the data is sufficiently described by the NOE structures and therefore, the B-DNA model was neglected in the following calculations.

Table 4.8: The Q -factors of the combined PCS and RDC analysis of Tm sample.

Structure	Q -factor (PCS)	Q -factor (RDC)	Q -factor (PCS+RDC)
NOE 1	0.27	0.50	0.32
NOE 2	0.34	0.43	0.36
NOE 3	0.34	0.44	0.36
NOE 4	0.26	0.52	0.32
NOE 5	0.32	0.43	0.34

The combined data of the Tm sample (PCSs and RDCs) was analyzed in the next step (see Table 4.8), resulting in significantly higher Q -factors than for the individual analyses. The combined analysis and the analysis for the PCSs alone result in significantly deviating lanthanide positions, which are 14 Å apart (see Fig.4.15). For both approaches, the distance from the DNA strand is reasonable, in relation to the length of the modification (approximately 20 Å). However, the back-calculated magnetic susceptibility tensors of the PCS data and the corresponding tensors that originate from the combined data set, point in different directions, illustrated by deviating values for the angle β (see Table 4.9). These relative orientations were calculated using equation 2.13 [Kramer2004]

Table 4.9: Angle β between the back-calculated PCS and RDC tensors.

Structure	angle β [°]
NOE 1	81
NOE 2	119
NOE 3	113
NOE 4	86
NOE 5	83

[Russo2013]. Furthermore, the magnitude of the calculated alignment tensors is smaller for the combined PCS and RDC analysis ($\Delta A_{ax} = -1.56$) compared to the individual PCS analysis ($\Delta A_{ax} = -3.60$). The angle β between the tensors, their magnitude, the different metal positions and the high Q -factors exhibit an overall unsatisfactory agreement of the combined PCS and RDC data analysis compared to the PCS analysis alone. On the one hand, this mismatch is caused by the errors of the determined RDCs. On the other hand, the dynamics of the tag have not been considered in the calculation, which was done by restricting the metal position to structurally reasonable positions in relation to the tag's conformation.

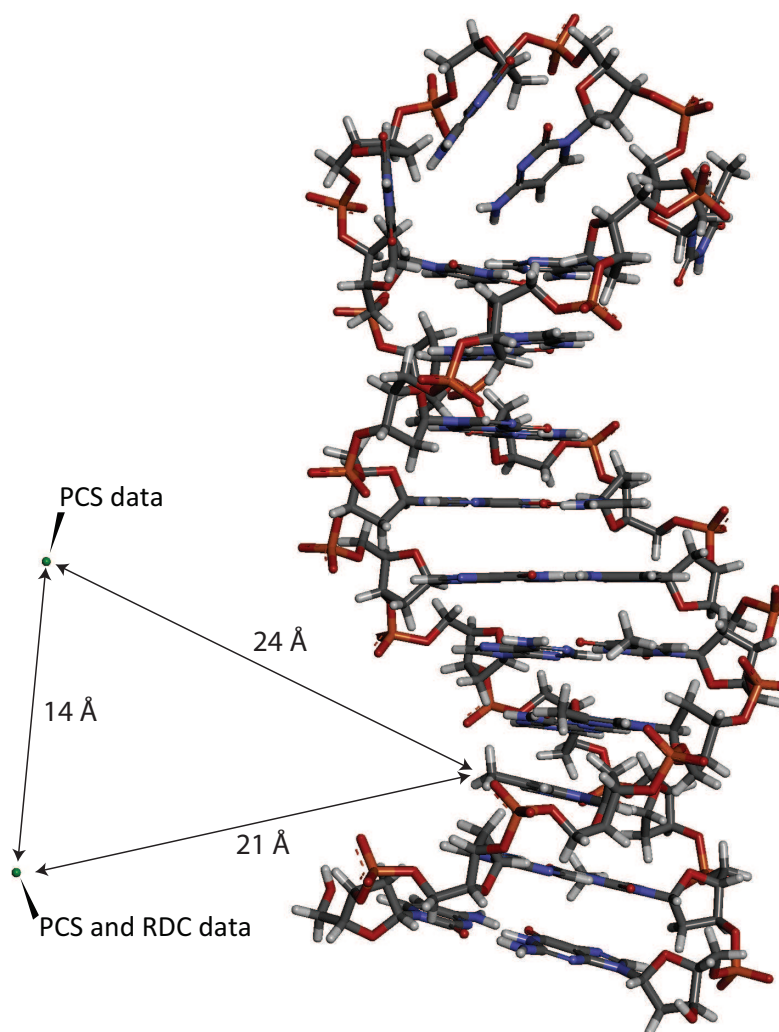


Figure 4.15: Two different T_m positions calculated from the PCS data alone or from the combined PCS and RDC data using the NOE model 1.

4.3.2 Evaluation of the Combined PCS and RDC Data Using a Structural Model of Cys-Ph-TAHA

In order to account for the tag flexibility, a model based approach was performed in a second data evaluation. For this purpose, the free rotation of the triple bond and the two conformations of the disulfide bridge were incorporated, as these are the major contributions to the flexibility (see Fig.4.16). Compared to these parameters, smaller deviations of the tag conformation have little effect on the final metal position.

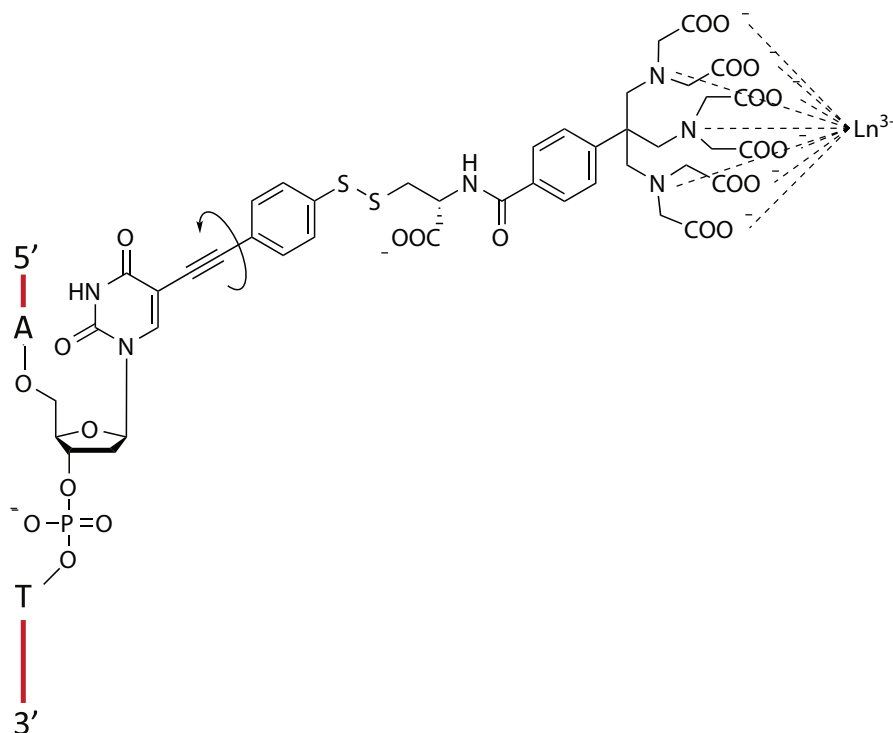


Figure 4.16: 52 models were created by rotating the triple bond in steps of 10° for both of the disulfide bridge conformations.

In the calculations, the Cys-Ph-TAHA tag and the incorporated binding site were attached to the NOE model 1. Different models were generated by rotating the triple bond in 10° steps for each of the two possible disulfide bridge conformers (dihedral angle $\phi = \pm 90^\circ$). In total, 72 structures of the tagged DNA strand were generated, of which only 52 structures were employed in the analysis, as conformers that result in sterical clashes between the tag and the DNA strand were discarded.

Ensembles consisting of two to five conformers were created using any possible combination of these 52 models. For the calculation, the individual structures of one ensemble were aligned along the phenyl ring of the tag and the PCS and RDC data was analyzed for this defined metal position, yet the tensor was calculated freely (see Fig.4.17). The

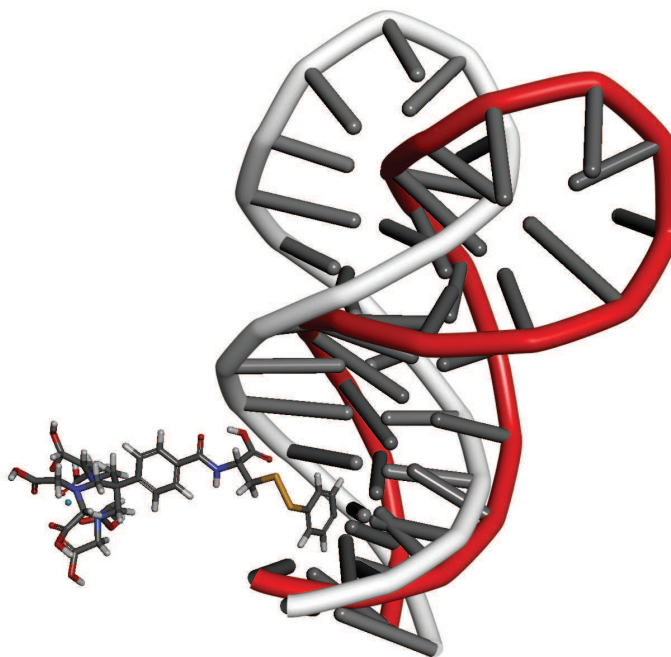


Figure 4.17: Two-structure ensemble aligned along the Cys-Ph-TAHA tag.

best Q -factors for the individual ensembles are given in Table 4.10.

By including more structures into the ensembles, lower resulting Q -factors are expected, as the actual metal position can be described more accurately. This was, however, not the case for the ensembles that were created from only two or three conformers, but for ensembles consisting of more structures. The influence of an additional structure is significantly stronger for smaller ensembles, which may have led to this deviation.

Table 4.10: Best Q -factors of the different data sets for ensembles consisting of two to five structures.

Number of Structures	PCSs	RDCs	Q -factor (PCS)	Q -factor (RDC)	Q -factor (PCS+RDC)
2	67	15	0.141	0.380	0.207
3	67	15	0.133	0.438	0.222
4	67	15	0.133	0.373	0.200
5	67	15	0.136	0.366	0.199

The best fit was found for an ensemble consisting of 5 conformers, which is shown in an overlay in Fig.4.18. Therein, one metal position and one alignment tensor are used for the description of the PCS and RDC data. Compared to the previous analysis (see section 4.3.1, Table 4.8), the calculated Q -factors are significantly smaller for the PCS, RDC and the combined PCS+RDC data. Only a slight improvement of the Q -factors was achieved

by increasing the number of structures in an ensemble so that the experimental data is accurately described by an ensemble consisting of only two structures.

The overlay of the best fitting ensembles illustrates the dynamic of the tag due to its rotation around the triple bond and the disulfide conformations. The ensemble shows a high degree of conformational freedom, in which the tag is turned away from the DNA's loop region, explaining the larger PCSs of the terminal cytosine 1, compared to the other residues.

Table 4.11: Euler angles and alignment tensors for the ensembles.

Number of Structures	Euler angles [°]			Alignment tensor				
	α	β	γ	ΔA_{ax}	ΔA_{rh}	A_{xx}	A_{yy}	A_{zz}
2	76.46	49.83	-176.34	3.29	0.44	-1.11	-5.48	6.58
3	79.90	52.87	-170.95	3.00	0.37	-1.36	-4.64	6.00
4	76.67	50.21	-176.52	3.65	0.45	-1.21	-6.08	7.29
5	125.25	62.98	-163.98	4.69	0.50	-1.14	-8.23	9.37

The calculated tensors and Euler angles are given in Table 4.11. In relation to the C_3 -symmetry of the Cys-Ph-TAHA tag, the tensor is expected to be axial-symmetric and oriented along the tag's C_3 -axis, presuming an unhindered rotation around this axis. So far, however, the calculated tensor is highly rhombohedral ($\Delta A_{ax} = 0.5$) and it is not aligned along the tag's C_3 -axis. As this axis is aligned along the z -axis of the molecular frame, the relative orientation between the tag's C_3 -axis and the tensor is directly given by the Euler angles in Table 4.11. This mismatch can be explained by a hindered rotation around the C_3 -axis of the tag, which could be induced by the lanthanide coordination, or by the errors of the determined data, in particular the RDCs (see section 4.3.1). Furthermore, the actual metal position might not be sufficiently represented by the structure models (undersampling), which could be improved by the addition of more conformations. Nevertheless, the description of the combined paramagnetic data using the structure and conformation of the tag results in a common metal position and alignment tensor with reasonable Q -factors.

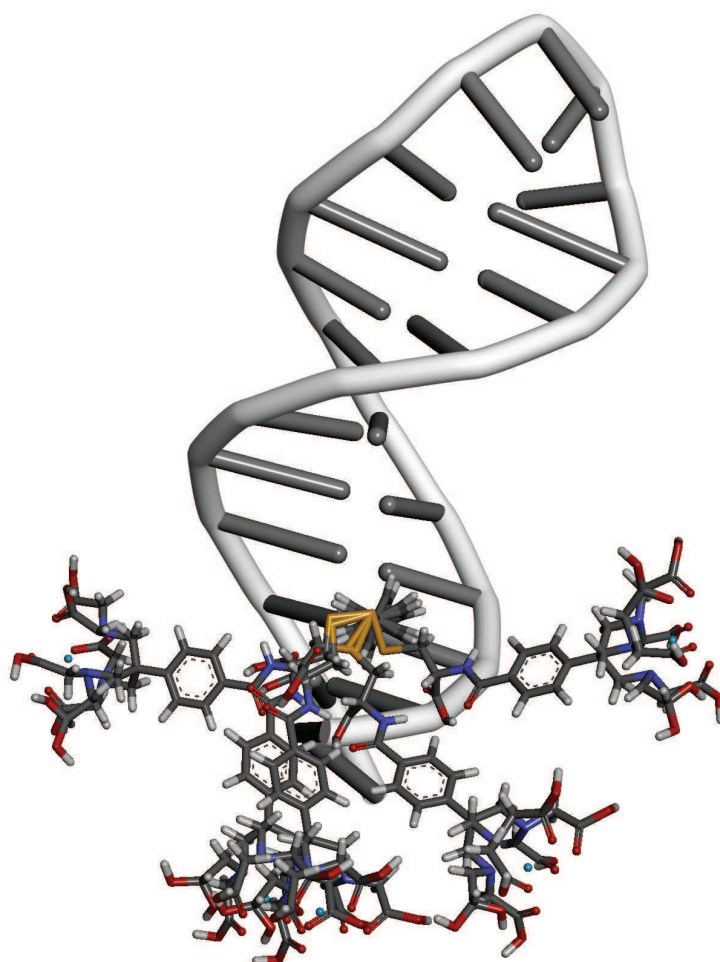


Figure 4.18: Five-structure ensemble for the NOE model 1 resulting in the lowest Q -factor.

5 Summary and Outlook

A reliable tagging strategy for the determination of paramagnetic NMR restraints is presented in this work. A sulfur binding site was successfully incorporated into a DNA strand, which was then utilized to attach the Cys-Ph-TAHA tag to the oligonucleotide. Two conceptually different approaches were pursued for the introduction of the free thiol moiety, i.e., the modification of a nucleobase and an oligonucleotide phosphorothioate strategy. In both cases, a paramagnetic tag was attached to a DNA molecule, with tagging on a modified nucleobase turning out to be superior with regard to the lanthanide ion coordination, reproducibility and overall yield.

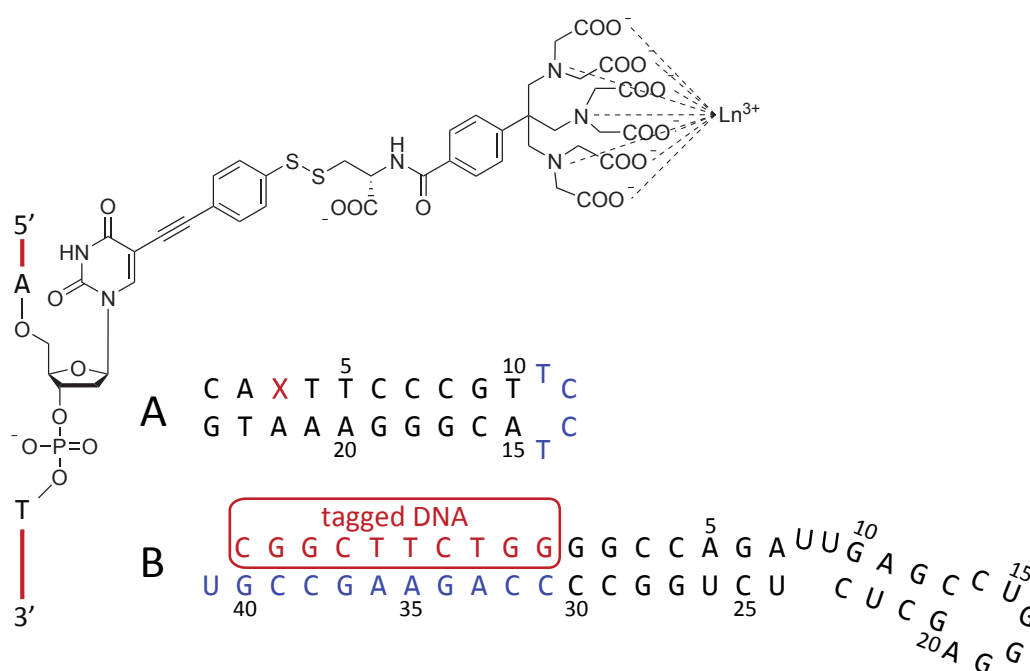


Figure 5.1: Scheme of a tagged DNA strand using a modified nucleobase and the Cys-Ph-TAHA tag, which was attached to the DNA hairpin molecule (A) and can be transferred to a DNA:RNA hybrid (B).

After the tagging reaction, three additional steps were necessary in order to prepare a stable NMR sample. Lanthanide ions, which were randomly coordinated by the DNA's backbone, had to be removed and sufficient sample conditions were found to guarantee a lanthanide coordination explicitly limited to the tag. Furthermore, the DNA strand

needs to be folded correctly in order to generate the desired hairpin structure, which was performed without significant removal of the lanthanide ion. The prepared NMR samples are stable for several months using a sodium chloride concentration of at least 180 mM.

Three different samples (Lu, Tm, Tb) were prepared for the measurement of high resolution ^1H - ^1H COSY, ^1H - ^1H NOESY, and ^1H - ^{13}C HSQC spectra, which revealed paramagnetically induced PCSs and RDCs. In total, 123 PCSs and 15 RDCs were recorded and evaluated in a structure calculation.

The data evaluation verified a satisfactorily high quality of the determined PCSs, which is expressed by the obtained low Q -factors. The errors from the RDC evaluation were significantly higher, which essentially originates from two different aspects. Firstly, the coupling constant determination of unlabeled NMR samples requires long measurement times of the NMR experiments in order to achieve a sufficient signal-to-noise ratio, which has proven to be especially difficult in paramagnetic samples. Moreover, the determined RDCs were smaller than expected, based on previous studies using the Cys-Ph-TAHA tag on ubiquitin [Peters2011] and, therefore, the relative errors are larger given the small size of the RDCs. Consequently, the flexibility of the tag relative to the DNA is higher in this approach compared to previous studies on proteins.

Nonetheless, by introducing a structural model of the modification and the disulfide conformations, a suitable ensemble model was found that sufficiently describes the combined paramagnetic data. The calculated alignment tensors show, however, an unexpected orientation in relation to the tag's symmetry. Therefore, further calculations will be performed, in which the tensor is defined to be axial symmetric.

In conclusion, a tagging protocol, starting with the synthesis of a modified nucleobase and resulting in the measurement of paramagnetic effects in a DNA strand, was established. The presented method provides a very convenient tool for the implementation of a sulfur binding site in oligonucleotides. As the overall yield of the phosphoramidite synthesis is satisfactorily high and can be performed on a gram scale, a large amount of starting material is accessible by this method for paramagnetic tagging.

Two major aspects can be derived from the presented work in order to improve the determination of RDCs in oligonucleotides.

Firstly, the flexibility of the modification and tag could be restricted further, in order to generate a stronger alignment. This suggests to be a promising approach for the synthesis of a more rigid Cys-Ph-TAHA derivative as described in section 3.6. A significant contribution to its flexibility originates from the rotational freedom of the triple bond in combination with the two different disulfide bridge conformations. A refinement of the developed binding site is expected to result in the determination of larger RDCs.

Secondly, a transfer of this protocol to a $^{13}\text{C}/^{15}\text{N}$ labeled target molecule is highly

recommendable, as it enables a more accurate RDC determination. Therefore, the method should be applied to a DNA:RNA hybrid, as presented by E. Woeltjen (see section 3.5, Fig.3.10) [Woeltjen2009]. Paramagnetic effects could also be transferred to DNA-binding proteins, as was recently reported for a protein-protein application [Camacho-Zarco2014]. In both examples, a paramagnetically tagged DNA molecule can be synthesized according to the presented protocol and the paramagnetic effects subsequently transferred to a $^{13}\text{C}/^{15}\text{N}$ labeled target molecule. Thereby, the accessible toolkit for the detection of long-range interactions in NMR spectroscopy can be expanded in order to cover one of the most important groups of biomolecules, the oligonucleotides.

6 Materials and Methods

Materials

All solvents were purchased in *pro analysis* quality from *Merck*, *Fluka* and *Acros*. The chemicals were purchased from *Sigma Aldrich*, *Merck*, *Fluka*, *Alfa Aesar*, *Acros*, and *Lancaster*. Sensitive reactions were shielded with argon gas which had been dried using phosphorus pentoxide.

All oligonucleotides were purchased from *IBA*. Cleavage and purification of the modified DNA was performed according to the protocol shown below. Wildtype DNA and phosphorothioate DNA were purified by *IBA*.

Chromatography

Purification using silica was performed via *flash column chromatography* with a pressure of 1.0 and 1.5 bar. Silica gel 60 was purchased from *Merck* with a particle size of either 63-100 μm or 15-40 μm . The amount of silica was 50-100 times that of the crude product, which was applied as a concentrated solution. Thin layer chromatography using either phosphomolybdic solution in ethanol (10%) or iodine as a dye, was used for reaction control on SIL G/UV 254 plates from *Merck*.

High performance liquid chromatography was performed on a *Jasco* system with a multi-wavelength detector. Reverse phase *Knauer* Eurosphere C18 columns (250 x 8 mm, 250 x 16 mm) were used with a flow rates of 3 ml/min and 7 mL/min, respectively. The applied gradients are given in the experimental details. Phosphorothioate oligonucleotides was separated according to literature reported procedure [Slim1991].

Liquid chromatography-mass spectrometry was performed on a *Waters Alliance 2295* system with a reverse phase *Knauer* Eurosphere C18 (250 x 4.6 mm) and the following gradient: 0.1% TFA in water : 0 - 30 min 0 \rightarrow 60% acetonitrile with a flowrate of 1 mL/min. Details on the mass spectrometer are given below.

NMR Spectroscopy

All synthetic molecules were analysed using a *Bruker* Avance I Ultrashielded Spectrometer (400 MHz) at 298 K. The employed solvents are given in the experimental details. Chemical shifts are listed in *parts per million* (ppm) in respect to the residual solvent signals of chloroform ($\delta=7.26$ ppm for proton, $\delta=77.0$ ppm for carbon), dimethyl sulfoxide ($\delta=2.49$ ppm for proton, $\delta=39.5$ ppm for carbon), methanol ($\delta=3.31$ ppm for proton, $\delta=49.0$ ppm for carbon) or water ($\delta=4.79$ ppm for proton). The signals are listed in the following notation: chemical shift δ in ppm, multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet, br = broadened), scalar coupling constant J in Hz, intensity and nucleus.

NMR spectra of oligonucleotides were recorded using *Shigemi* NMR tubes with a total volume of 250 μL at 298 K. Sample preparation and NMR buffer conditions are given in the experimental details.

NOESY and COSY spectra of oligonucleotides were recorded on a *Bruker* Avance I Ultrashielded Spectrometer at 900 MHz or 400 MHz using conventional pulse programs. A mixing time of 200 ms for the diamagnetic and 150 ms for the paramagnetic sample was applied.

$[^1\text{H}, ^{13}\text{C}]$ -HSQC and $[^1\text{H}, ^{13}\text{C}]$ -HSQC IPAP spectra of oligonucleotides were recorded on a *Bruker* Avance III Ultrashielded Spectrometer at 800 MHz. The pulse sequence is given in the appendix 6.1 [Ottiger1998] [Siepel2013]. The paramagnetic spectra were recorded with 128 scans. The spectra were processed using the Bruker *TopSpin* with a size of 8k (F1), which results in a spectral resolution of at least 0.7 Hz. A Gaussian window function was applied with a broadening factor of -7 Hz and a position factor of 0.3.

Mass Spectrometry

Mass spectra were recorded on a *Waters* electron spray ionization mass spectrometer (Micromass ZQ) with a quadrupole detector or on a *Thermo* DSQII/Focus GC. All values are given in *mass per charge* (m/z). The used solvents are given in the experimental details.

Software

The DNA models were created using the program *Accelrys Discovery Studio v.3.1.1.11157*. The Cys-Ph-TAHA model was energy minimized using a CHRAMm forcefield and the so called Smart Minimizer with a maximum of 200 steps and an RMS gradient of 0.1 [Brooks1983].

The idealized B-DNA model was created in accordance with the NOE-models

[Taeubert2010] using the program *Crystallography and NMR Systems* [Brunger1998] [Brunger2007].

The two-dimensional NMR spectra were analyzed using the program *Sparky* [Goddard2008].

Analysis of Paramagnetic Data

The PCS and RDC data analysis was performed according to the literature reported procedure [Russo2013]. The metal position was optimized using a grid search around the DNA models until the lowest Q -factors were found. The tensors were calculated by least-squared fits of the experimental PCS and RDC data to the different model coordinates. As starting structures either the idealized B-DNA and the NOE models were used. The calculation was performed according to the equation

$$\delta_i^{\text{PCS}}, \Delta\nu_i^{\text{RDC}} = k \cdot \text{trace} \cdot \left[\begin{pmatrix} (3x_i^2 - r_i^2) & 3x_i y_i & 3x_i z_i \\ 3x_i y_i & (3y_i^2 - r_i^2) & 3y_i z_i \\ 3x_i z_i & 3y_i z_i & (3z_i^2 - r_i^2) \end{pmatrix} \begin{pmatrix} \chi_{xx} & \chi_{xy} & \chi_{xz} \\ \chi_{xy} & \chi_{yy} & \chi_{yz} \\ \chi_{xz} & \chi_{yz} & \chi_{zz} \end{pmatrix} \right]$$

with

$$k^{\text{PCS}} = \frac{1}{12\pi r_i^5}, \quad k^{\text{RDC}} = \frac{B_0^2 \gamma_H \gamma_C \hbar}{120\pi^2 k_B T r_i^5}.$$

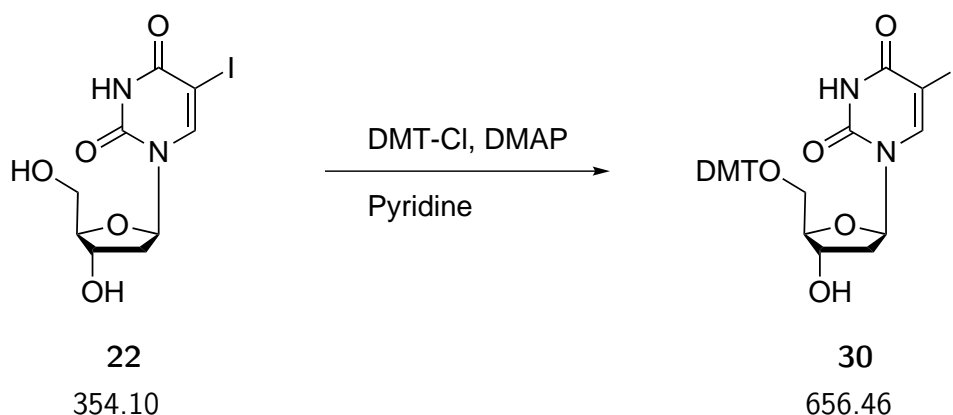
Therein, r_i is either defined as the distance from the nucleus to the paramagnetic center (PCSs) or as the inter-nuclear vector (RDCs) [Bertini2002a]. In the combined PCS and RDC analysis, the values of the data sets were scaled to the largest experimental values in order to avoid overestimation of any of the two parameters. Composite Q -factors were calculated for evaluation, taking into account the number of respective values N used to determine the Q -factors:

$$Q_{\text{PCS+RDC}} = \sqrt{\frac{(N_{\text{PCS}} Q_{\text{PCS}}^2 + N_{\text{RDC}} Q_{\text{RDC}}^2)}{(N_{\text{PCS}} + N_{\text{RDC}})}}.$$

The Q -factors were calculated according to equation 2.12 (see section 2.1).

6.1 Experimental Details

5'-*O*-Dimethoxytrityl-5-iodo-2'-desoxyuridine [Woeltjen2009]



To a solution of 5-iodo-2'-desoxyuridine **22** (1.50 g, 4.24 mmol, 1.00 eq) in pyridine (15 ml), triethylamine (587 μ L, 4.24 mmol, 1.00 eq), 4-(dimethylamino)pyridine (52.0 mg, 0.42 mmol, 0.10 eq) and 4,4'-dimethoxytrityl chloride (1.72 g, 5.08 mmol, 1.20 eq) were added. The reaction mixture was stirred for 16 h at ambient temperature.

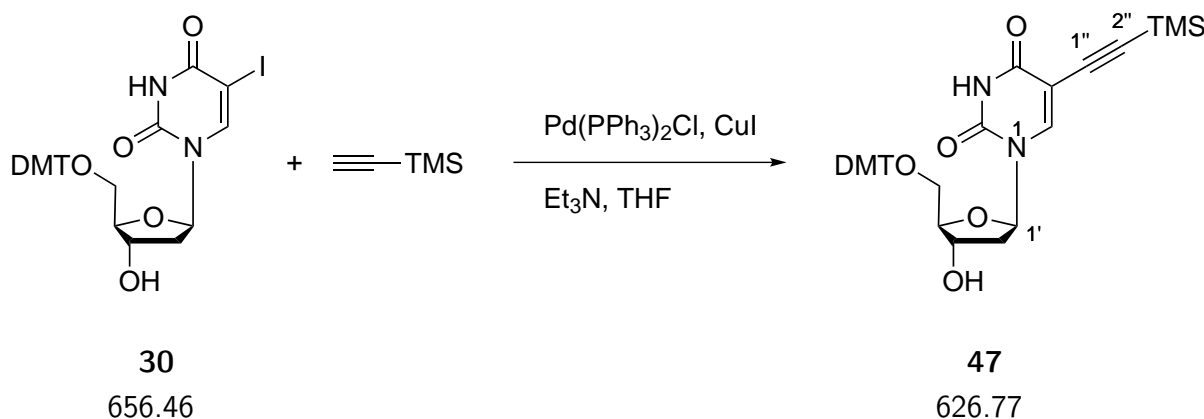
The solvent was removed under reduced pressure and coevaporated with methanol (15 mL). The crude product was purified by chromatography on silica using chloroform/methanol/triethylamine (100:1:1 \rightarrow 100:2:1) to afford product **30** (2.65 g, 4.04 mmol, 95%) as a white foam.

^1H NMR: (400 MHz, CDCl_3): δ = 8.73 (br, 1H, NH), 8.10 (s, 1H, H6), 7.43-7.16 (m, 9H, DMT), 6.86-6.84 (m, 4H, DMT), 6.33 (t, J = 6.2 Hz, 1H, H1'), 4.59 (br, 1 H, OH), 4.53-4.52 (m, 1H, H3'), 4.08-4.06 (m, 1H, H4'), 3.78 (s, 6H, 2 x OMe), 3.43-3.33 (m, 2H, H5'), 2.51-2.41 (m, 1H, H2'a), 2.30-2.20 (m, 1H, H2'b) ppm.

^{13}C NMR: (100.6 MHz, CDCl_3): δ = 161.1 (1C, C4), 158.4 (2C, DMT), 150.7 (1C, C2), 144.7 (1C, C6) 144.2-113.2 (16C, DMT), 86.8 (1C, DMT), 86.2 (1C, C4'), 85.1 (1C, C1'), 72.1 (1C, C3'), 69.0 (1C, C5), 63.2 (1C, C5'), 54.9 (2C, 2 x OMe), 41.3 (1C, C2') ppm.

ESI-MS m/z (acetonitrile, negative mode): calc. for $\text{C}_{30}\text{H}_{28}\text{IN}_2\text{O}_7$ $[\text{M}-\text{H}]^-$: 655.46, found: 655.06.

5'-*O*-Dimethoxytrityl-5-trimethylsilylethynyl-2'-desoxyuridine [Woeltjen2009]



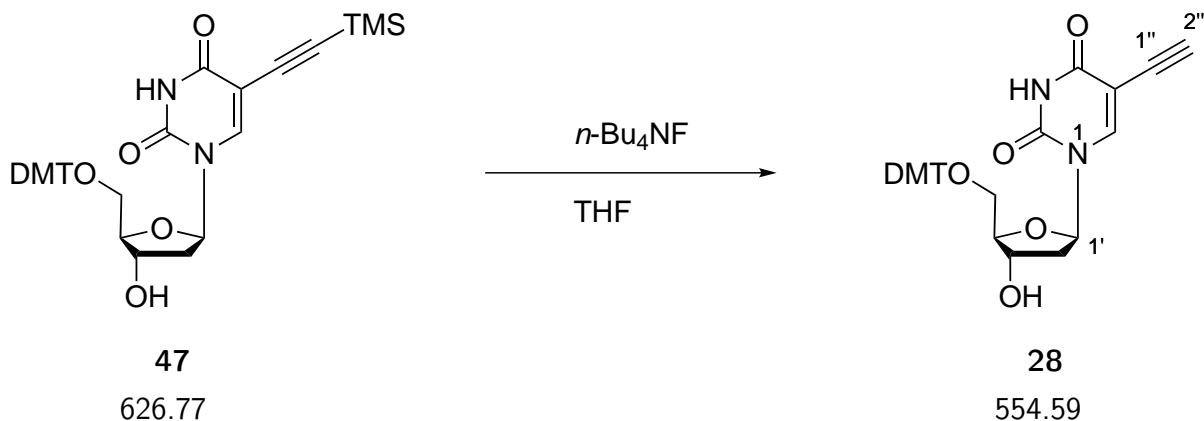
To a solution of 5'-*O*-dimethoxytrityl-5-iodo-2'-desoxyuridine **30** (2.64 g, 4.02 mmol, 1.00 eq) and copper(I) iodide (61 mg, 0.32 mmol, 0.08 eq) in triethylamine/tetrahydrofuran (1:1, 60 mL), trimethylsilylethyne (1.71 mL, 12.1 mmol, 3.00 eq) was added, followed by bis-(triphenylphosphino)-palladium(II) chloride (85.0 mg, 0.12 mmol, 0.03 eq). The reaction mixture was stirred for 4 h at 50 °C under argon atmosphere.

The solvent was removed under reduced pressure and the crude product was purified by chromatography on silica using chloroform/methanol/triethylamine (100:1:1 → 100:2:1) to afford the product **47** (1.79 g, 2.86 mmol, 71%) as a pale-yellow foam.

¹H NMR: (400 MHz, CDCl₃): δ = 7.91 (s, 1H, H6), 7.87 (br, 1H, NH), 7.36 - 7.05 (m, 9H, DMT), 6.75 - 6.69 (m, 4H, DMT), 6.25 - 6.19 (m, 1H, H1'), 4.37 - 4.33 (m, 1H, H3'), 4.04 - 4.00 (m, 1H, H4'), 3.64 (s, 6H, 2 x OMe), 3.32 - 3.13 (m, 2H, H5'), 2.42 - 2.33 (m, 1H, H2'a), 2.11 - 2.00 (m, 1H, H2'b), -0.02 (s, 9H, TMS) ppm.

¹³C NMR: (100.6 MHz, CDCl₃): δ = 163.8 (1C, C4), 158.0 (2C, DMT), 150.8 (1C, C2), 144.3 (1C, C6) 142.0-112.1 (16C, DMT), 99.9 (1C, C5), 98.2 (1C, C2''), 95.8 (1C, C1''), 86.3 (1C, DMT), 86.2 (1C, C4'), 85.3 (1C, C1'), 71.5 (1C, C3'), 63.4 (1C, C5'), 54.9 (2C, 2 x OMe), 41.2 (1C, C2') ppm, 0.0 (3C, TMS).

ESI-MS m/z (acetonitrile, negative mode): calc. for C₃₅H₃₇N₂O₇Si [M-H]⁻: 625.76, found: 625.09.

5'-*O*-Dimethoxytrityl-5-ethynyl-2'-desoxyuridine [Woeltjen2009]

To a solution of 5'-*O*-dimethoxytrityl-5-trimethylsilylethynyl-2'-desoxy-uridine **47** (1.79 g, 2.86 mmol, 1.00 eq) and triethylamine (396 μL , 2.86 mmol, 1.00 eq) in tetrahydrofuran (35 mL), a 1 M tetra-*n*-butylammoniumfluoride solution in tetrahydrofuran (5.71 mL, 5.71 mmol, 2.00 eq) added dropwise via syringe. The mixture was stirred for 6 h at ambient temperature under argon atmosphere.

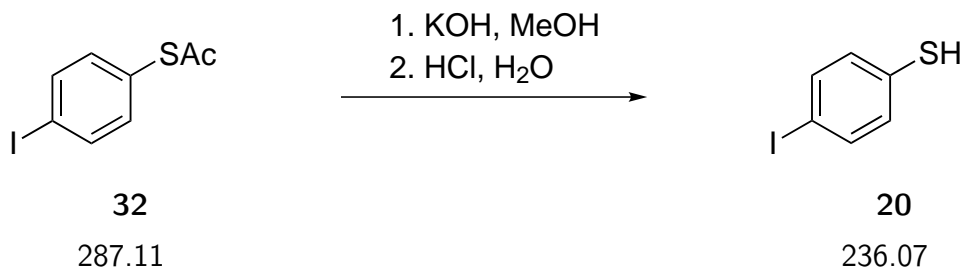
The solvent was removed under reduced pressure and the crude product was purified by chromatography on silica using chloroform/methanol/triethylamine (100:0:1 \rightarrow 100:10:1) to afford the product **28** (1.46 g, 2.63 mmol, 92%) as an off-white foam.

^1H NMR: (400 MHz, CDCl_3): δ = 7.95 (s, 1H, H6), 7.78 (br, 1H, NH) 7.38 - 7.04 (m, 9H, DMT), 6.76 - 6.68 (m, 4H, DMT), 6.28 - 6.20 (m, 1H, H1'), 4.48 - 4.42 (m, 1H, H3'), 4.07 - 4.01 (m, 1H, H4'), 3.63 (s, 6H, 2 x OMe), 3.36 - 3.15 (m, 2H, H5'), 2.77 (s, 1H, H2''), 2.45 - 2.38 (m, 1H, H2'a), 2.17 - 2.07 (m, 1H, H2'b) ppm.

^{13}C NMR: (100.6 MHz, CDCl_3): δ = 164.5 (1C, C4), 158.3 (2C, DMT), 151.0 (1C, C2), 144.4 (1C, C6) 143.0-112.0 (16C, DMT), 98.6 (1C, C5), 86.3 (1C, DMT), 86.0 (1C, C4'), 85.3 (1C, C1'), 80.9 (1C, C1''), 75.4 (1C, C2''), 71.4 (1C, C3'), 63.5 (1C, C5'), 54.0 (2C, 2 x OMe), 41.4 (1C, C2') ppm.

ESI-MS m/z (acetonitrile, negative mode): calc. for $\text{C}_{32}\text{H}_{29}\text{N}_2\text{O}_7$ $[\text{M}-\text{H}]^-$: 553.58, found: 553.29.

4-Iodothiophenol [Haiss2006]



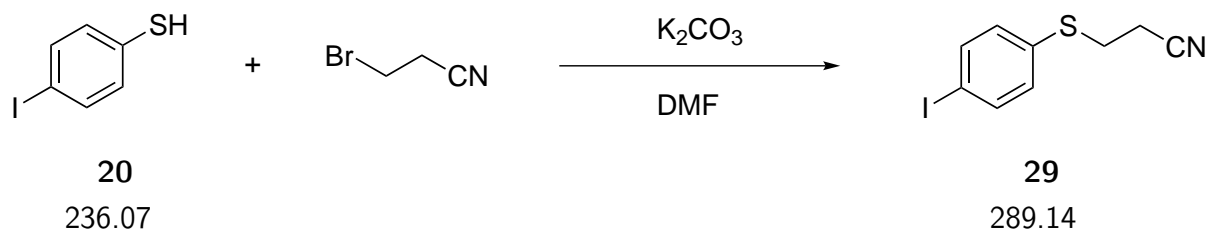
To a solution of *S*-(4-iodophenyl) thioacetate **32** (2.02 g, 7.04 mmol, 1.00 eq) in methanol (20 mL), potassium hydroxide (454 mg, 8.09 mmol, 1.15 eq) was added. The mixture was stirred under argon atmosphere for 1 h at ambient temperature, followed by 0.5 h at 50 °C. Concentrated hydrochloric acid was added to adjust the pH to approx. 1, followed by addition of ice-water (10 mL) with constant stirring.

The precipitated yellow crystals were collected by filtration and dried under vacuum to yield 4-iodothiophenol **20** (1.36 g, 5.76 mmol, 82%).

¹H NMR: (400 MHz, CDCl₃): δ = 7.54 (d, *J* = 8.5 Hz, 2H, Ph), 7.02 (d, *J* = 8.5 Hz, 2H, Ph), 3.43 (s, 1H, SH) ppm.

¹³C NMR: (100.6 MHz, CDCl₃): δ = 138.0 (2C, Ph), 131.1 (2C, Ph), 130.9 (1C, Ph), 90.2 (1C, Ph) ppm.

ESI-MS *m/z* (acetonitrile, negative mode): calc. for C₆H₄IS [M-H]⁻: 234.91, found: 235.00.

3-[(4-Iodophenyl)sulfanyl]propanenitrile [Haiss2006]

A solution of 4-iodothiophenol **20** (1.00 g, 4.24 mmol, 1.00 eq) and 3-bromopropionitrile (914 μ L, 11.0 mmol, 2.60 eq) in anhydrous dimethylformamide (6 mL) was degassed by bubbling argon through for 0.5 h. Potassium carbonate (586 mg, 4.24 mmol, 1.00 eq) was added and the mixture stirred at 105 °C for 3.5 h under argon atmosphere.

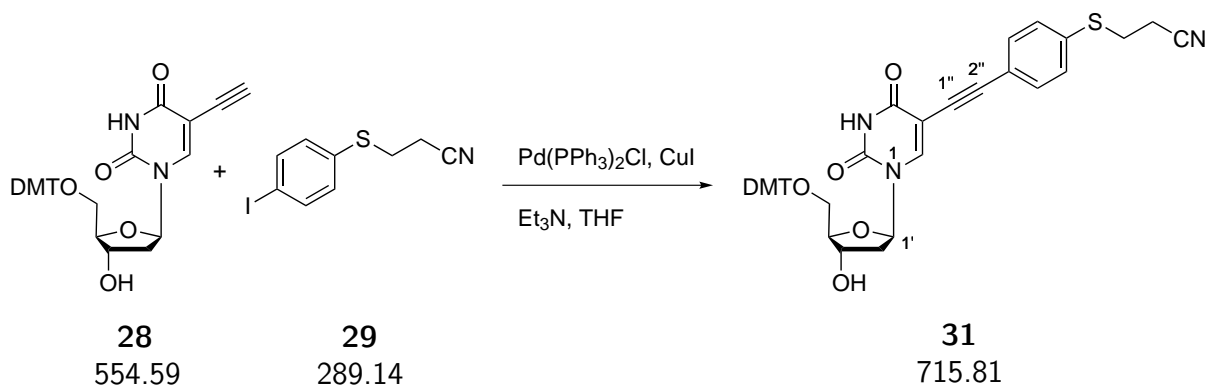
The cooled mixture was filtered and the solvent was evaporated under reduced pressure. The residual yellow oil was dissolved in methanol (8 mL), followed by slow addition of water (29 mL) with constant shaking. Suction filtration and washing with a methanol-water mixture (20 mL, 1:2) afforded 3-[(4-Iodophenyl)sulfanyl]propanenitrile **29** (951 mg, 3.29 mmol, 78%) as a white solid.

1H NMR: (400 MHz, $CDCl_3$): δ = 7.59 (d, J = 8.4 Hz, 2H, Ph), 7.08 (d, J = 8.4 Hz, 2H, Ph), 3.07 (t, 2H, CH_2), 2.56 (t, 2H, CH_2) ppm.

^{13}C NMR: (100.6 MHz, $CDCl_3$): δ = 128.3 (2C, Ph), 133.6 (1C, Ph), 132.6 (2C, Ph), 118.0 (1C, CN), 93.1 (1C, Ph), 29.9 (2C, CH_2), 18.3 (2C, CH_2) ppm.

GC-MS m/z (acetonitrile, negative mode): calc. for C_9H_7INS $[M-H]^-$: 287.93, found: 288.90.

5'-*O*-Dimethoxytrityl-5-[4-(2-cyanoethylsulfanyl)phenyl]ethynyl-2'-desoxyuridine



A solution of 5'-*O*-dimethoxytrityl-5-ethynyl-2'-desoxyuridine **28** (600 mg, 1.08 mmol, 1.00 eq) and 1-(2-cyanoethylsulfanyl)-4-iodobenzene **29** (313 mg, 1.08 mmol, 1.00 eq) in a tetrahydrofuran/triethylamine mixture (40 mL, 1:1) was degassed by bubbling argon through for 0.5 h. Copper(I) iodide (21.0 mg, 0.11 mmol, 0.10 eq) and bis-(triphenylphosphino)-palladium(II) chloride (38.0 mg, 0.06 mmol, 0.05 eq) were added and the mixture was stirred for 2 h at 40 °C under argon atmosphere.

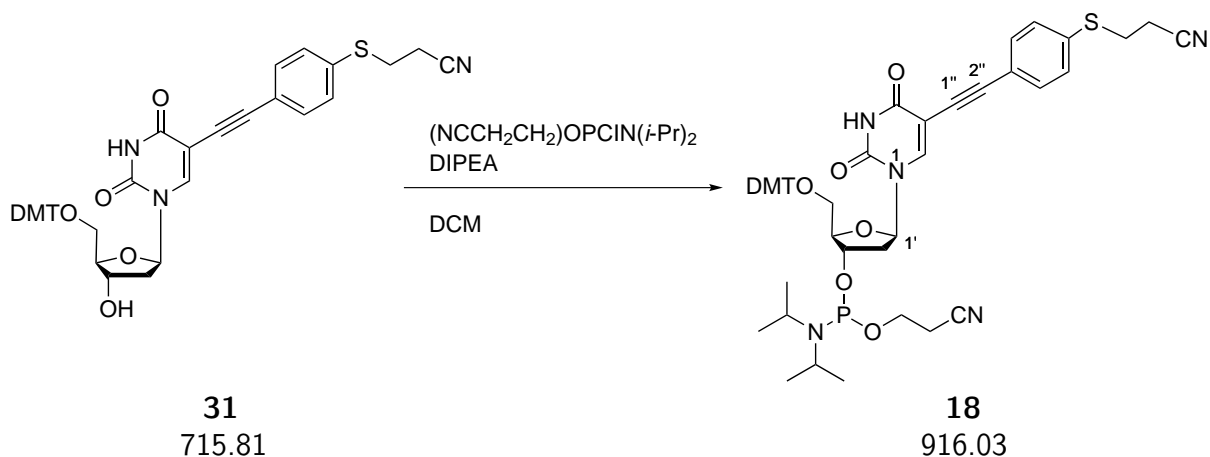
The solvents were removed under reduced pressure. The resulting oil was dissolved in ethyl acetate (50 mL), washed with saturated NaHCO₃ solution (3 x 50 mL), dried over anhydrous MgSO₄ and filtered. The solvent was removed under reduced pressure and the crude product was purified by chromatography on silica using chloroform/methanol/triethylamine (100:1:1) to yield product **31** (624 mg, 0.87 mmol, 81%) as an off-white foam.

¹H NMR: (400 MHz, CDCl₃): δ = 8.27 (s, 1H, H6), 7.49-7.23 (m, 8H, DMT), 7.17-7.09 (m, 3H, 2 x Ph, DMT), 6.90 (d, *J* = 8.4 Hz, 2H, 2 x Ph), 6.82-6.75 (m, 4H, DMT) 6.42-6.37 (m, 1H, H1'), 5.87 (br, 1H, NH), 4.60-4.55 (m, 1H, H3'), 4.20-4.15 (m, 1H, H4'), 3.71-3.66 (m, 6H, 2 x OMe), 3.49-3.43 (m, 1H, H5'a), 3.33-3.27 (m, 1H, H5'b), 3.09 (t, *J* = 7.2 Hz, 2H, CH₂), 2.62-2.54 (m, 3H, CH₂, H2'a), 2.36-2.27 (m, 1H, H2'b) ppm.

¹³C NMR: (100.6 MHz, CDCl₃): δ = 162.1 (1C, C4), 158.4 (2C, DMT), 149.7 (1C, C2), 144.4 (1C, C6), 142.5-112.9 (16C, DMT), 133.8 (1C, Ph), 132.1 (2C, Ph), 129.6 (2C, Ph), 121.5 (1C, Ph), 117.8 (1C, CN), 100.2 (1C, C5), 92.8 (1C, C2''), 86.9 (1C, DMT), 86.7 (1C, C4'), 85.8 (1C, C1'), 81.4 (1C, C1''), 72.1 (1C, C3'), 63.5 (1C, C5'), 55.1 (2C, 2 x OMe), 41.4 (1C, C2'), 29.4 (1C, CH₂), 17.9 (1C, CH₂) ppm.

ESI-MS m/z (acetonitrile, negative mode): calc. for $C_{41}H_{36}N_3O_7S$ $[M-H]^-$: 714.23, found: 714.30.

5'-*O*-Dimethoxytrityl-5-[4-(2-cyanoethylsulfanyl)phenyl]ethynyl-2'-desoxyuridin-3'-[2-cyanoethyl-*N,N*-bis(diisopropyl)]-phosphoramidite



To a solution of compound **31** (553 mg, 0.77 mmol, 1.00 eq) and diisopropylethylamine (202 μ L, 1.16 mmol, 1.50 eq) in anhydrous dichloromethane (20 mL), 2-cyanoethyl diisopropylchlorophosphoramidite (256 μ L, 1.16 mmol, 1.50 eq) was added under argon atmosphere. The mixture was stirred for 2 h at ambient temperature.

The solvent was removed under reduced pressure and the crude product was purified by chromatography on silica using chloroform/methanol/triethylamine (100:1:1) to yield product **18** (634 mg, 0.69 mmol, 90%) as a colorless oil.

^1H NMR: (400 MHz, CDCl_3): δ = 10.77 (br, 1H, NH), 8.27* (s, 1H, H6), 7.55-7.49 (m, 2H, DMT), 7.44-7.37 (m, 4H, DMT), 7.32-7.25 (m, 2H, DMT), 7.19-7.07 (m, 3H, 2 x Ph, DMT), 6.92-6.77 (m, 6H, 2 x Ph, DMT), 6.47-6.40 (m, 1H, H1'), 4.67-4.60 (m, 1H, H3'), 4.26-4.17 (m, 1H, H4'), 3.88-3.73 (m, 2H, $\text{OCH}_2\text{CH}_2\text{CN}$), 3.69 (s, 3H, OMe), 3.68 (s, 3H, OMe), 3.67-3.47 (m, 3H, H5'a, 2 x $\text{CH}(\text{CH}_3)_2$), 3.33-3.27 (m, 1H, H5'a), 3.09* (t, J = 7.2 Hz, 2H, CH_2), 2.60-2.54 (m, 3H, CH_2 , H2'a), 2.45-2.23 (m, 3H, H2'b, $\text{OCH}_2\text{CH}_2\text{CN}$), 1.21-1.16 (m, 12H, 3 x CH_3) ppm.

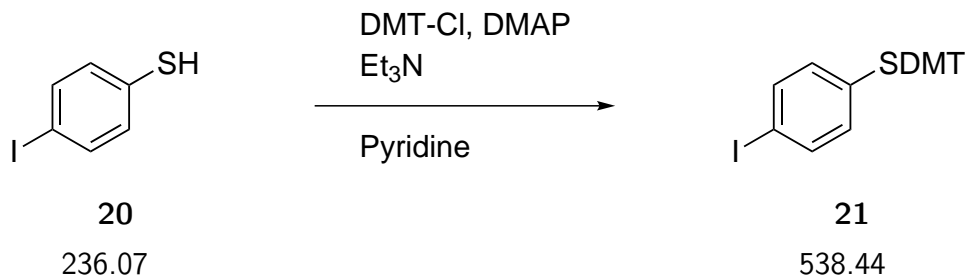
^{13}C NMR: (100.6 MHz, CDCl_3): δ = 164.7 (1C, C4), 158.5 (2C, DMT), 151.7 (1C, C2), 144.5 (1C, C6), 142.1-113.2 (DMT), 132.8 (1C, Ph), 132.1 (2C, Ph), 129.8 (2C, Ph), 126.9 (1C, DMT), 122.2 (1C, Ph), 117.8 (1C, CN), 117.6* (1C, $\text{OCH}_2\text{CH}_2\text{CN}$), 113.2 (4C, DMT), 100.3 (1C, C5), 92.0 (1C, C2''), 86.9 (1C, DMT), 85.9* (1C, C4'),

85.4/85.3* (1C, C1'), 82.9 (1C, C1''), 73.9-73.3* (1C, C3'), 63.4/63.2* (1C, C5'), 58.4* (1C, OCH₂CH₂CN), 55.2 (2C, 2 x OMe), 45.3/45.2* (1C, C2'), 43.2* (1C, NCH(CH₃)₂), 40.8* (1C, C2'), 29.5 (1C, CH₂), 24.7-24.4*(4C, NCH(CH₃)₂), 20.4-20.2 (1C, OCH₂CH₂CN), 18.0 (1C, CH₂) ppm.

Signals marked with * indicate two diastereomers due to the phosphor atom.

ESI-MS m/z (acetonitrile, positive mode): calc. for C₅₀H₅₄N₅O₈PSNa [M+Na]⁺: 938.33, found: 938.33.

1-(Dimethoxytritylsulfanyl)-4-iodobenzene



To a solution of 4-iodothiophenol **20** (1.00 g, 4.24 mmol, 1.00 eq) in pyridine (30 ml), triethylamine (531 μ L, 3.81 mmol, 0.90 eq), 4-(dimethylamino)pyridine (41.0 mg, 0.34 mmol, 0.08 eq) and 4,4'-dimethoxytrityl chloride (1.72 g, 5.08 mmol, 1.20 eq) were added. The reaction mixture was stirred for 12 h at ambient temperature.

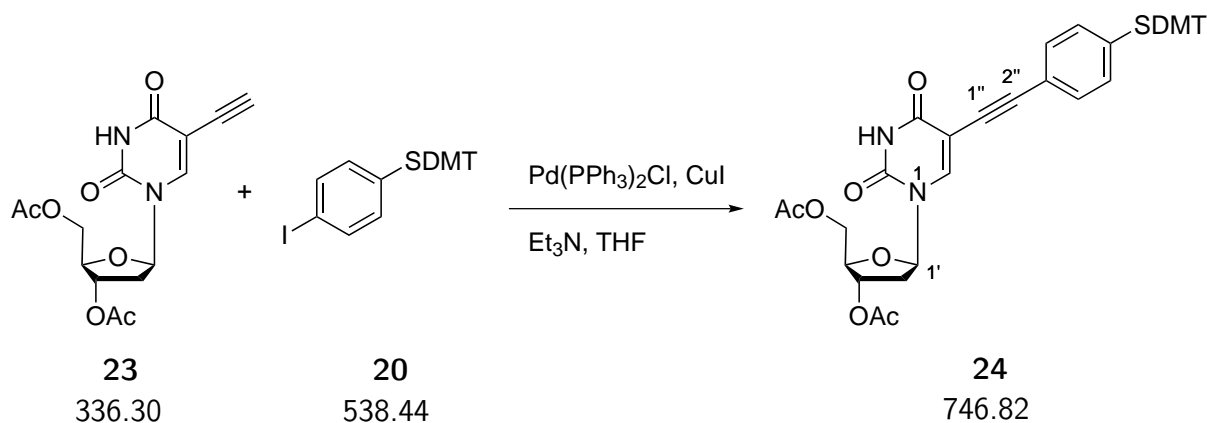
The solvent was removed under reduced pressure. Methanol (30 mL) was added to the residue, followed by suction filtration. The white residue was washed with methanol (2 x 30 mL) and dried *in vacuo* to yield compound **21** (1.39 g, 2.58 mmol, 61%) as a white solid.

¹H NMR: (400 MHz, CDCl₃): δ = 7.38-7.30 (m, 4H, DMT), 7.28-7.16 (m, 7H, DMT, Ph), 6.79-6.74 (m, 4H, DMT), 6.69-6.65 (m, 2H, Ph), 3.79 (s, 6H, 2 x OMe) ppm.

¹³C NMR: (100.6 MHz, CDCl₃): δ = 158.2 (2C, DMT), 144.7 (1C, DMT), 137.1 (2C, DMT), 136.5 (1C, Ph), 135.4 (2C, Ph), 135.2 (2C, DMT), 131.1 (4C, DMT, 2 x Ph), 129.8 (2C, DMT), 127.7 (2C, DMT), 126.8 (1C, DMT), 112.9 (4C, DMT), 93.3 (1C, Ph), 69.9 (1C, DMT) 55.2 (2C, OMe) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for C₂₇H₂₄IO₂S [M+H]⁺ 561.04, found: 561.09.

5-[4-(Dimethoxytritylsulfanyl)phenyl]ethynyl-2'-desoxyuridine-3',5'-di-*O*-acetate



A solution of 5-ethynyl-2'-desoxyuridine-3'-5'-di-*O*-acetate **49** (1.00 g, 2.97 mmol, 1.00 eq) in triethylamine/tetrahydrofuran (1:1, 160 mL) was degassed by bubbling argon through for 0.5 h. Copper(I) iodide (57 mg, 0.3 mmol, 0.1 eq), 1-(2-dimethoxytritylsulfanyl)-4-iodobenzene **20** (1.60 g, 2.97 mmol, 1.0 eq) and bis-(triphenylphosphino)-palladium(II) chloride (104 mg, 0.15 mmol, 0.05 eq) were added and the reaction mixture was stirred under argon atmosphere for 4 h at 45 °C.

The solvents were removed under reduced pressure. The residue was dissolved in ethyl acetate (100 mL), followed by an aqueous workup using a 5 % Na₃EDTA solution (3 x 100 mL). The combined organic layers were dried over anhydrous MgSO₄ and filtered. The crude product was purified by chromatography on silica using chloroform/methanol (100:0 → 100:1) to afford the product **24** (1.51 g, 2.02 mmol, 68%) as a yellow foam.

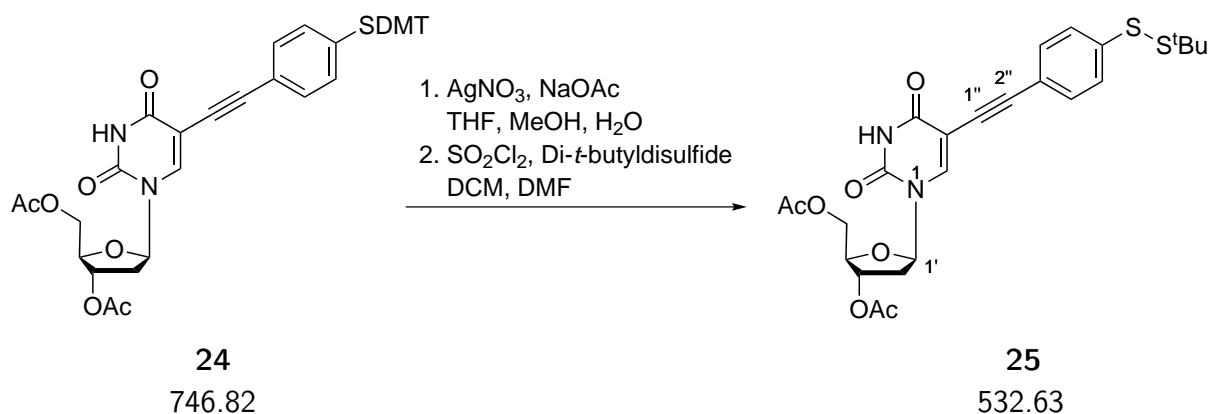
¹H NMR: (400 MHz, CDCl₃): δ = 9.55 (br, 1H, NH), 7.85 (s, 1H, H6), 7.39-7.17 (m, 9H, DMT), 7.12 (d, 2H, *J* = 8.4 Hz, 2H, Ph), 6.92 (d, *J* = 8.4 Hz, 2H, Ph), 6.80-6.75 (m, 4H, DMT) 6.35-6.30 (m, 1H, H1'), 5.27-5.23 (m, 1H, H3'), 4.38-4.35 (m, 2H, H5'), 4.32-4.28 (m, 1H, H4') 3.79 (s, 6H, 2 x OMe), 2.60-2.52 (m, 1H, H2'a), 2.27-2.18 (m, 1H, H2'b), 2.14-2.10 (m, 6H, 2 x OAc) ppm.

¹³C NMR: (100.6 MHz, CDCl₃): δ = 170.3 (1C, OAc), 170.0 (1C, OAc), 161.4 (1C, C4), 158.2 (2C, DMT), 149.4 (1C, C2), 144.6 (1C, C6), 141.1-112.9 (16C, DMT), 136.8 (1C, Ph), 132.9 (2C, Ph), 131.0 (2C, Ph), 121.2 (1C, Ph), 100.8 (1C, C5), 93.4 (1C, C2''), 85.4 (1C, C1'), 82.5 (1C, C4'), 81.1 (1C, C1''), 78.7 (1C, DMT), 73.9 (1C, C3'), 63.7 (1C,

C5'), 55.1 (2C, 2 x OMe), 38.2 (1C, C2'), 20.8 (1C, CH₃), 20.7 (1C, CH₃) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for C₄₂H₃₇N₂O₉S [M-H]³⁺: 186.30, found: 186.07.

5-[4-(*tert*-Butyldisulfanyl)phenyl]ethynyl-2'-desoxyuridine-3',5'-di-*O*-acetate



To a mixture of silver nitrate (641 mg, 3.78 mmol, 2.00 eq) in water/methanol (1:1, 1 mL), a solution of compound **24** (1.41 g, 1.89 mmol, 1.00 eq) in tetrahydrofuran/methanol (3:1, 173 mL) was added. A 3 M aqueous sodium acetate solution (1.61 mL, 4.81 mmol, 2.55 eq) was syringed to the mixture and stirred for 20 min at ambient temperature. Following filtration, methanol (10 mL) was added to the residue and the mixture sonicated. Filtration afforded a yellow silver salt.

To a solution of Di-*t*-butyldisulfide (1.64 mL, 8.50 mmol, 4.50 eq) in anhydrous dichloromethane (30 mL), sulfuryl chloride (689 μ L, 8.50 mmol, 4.50 eq) was added dropwise under argon atmosphere. The mixture was stirred for 2 h at ambient temperature and added to a suspension of the silver salt in anhydrous dimethylformamide (60 mL). The mixture was stirred for 16 h at ambient temperature under argon atmosphere.

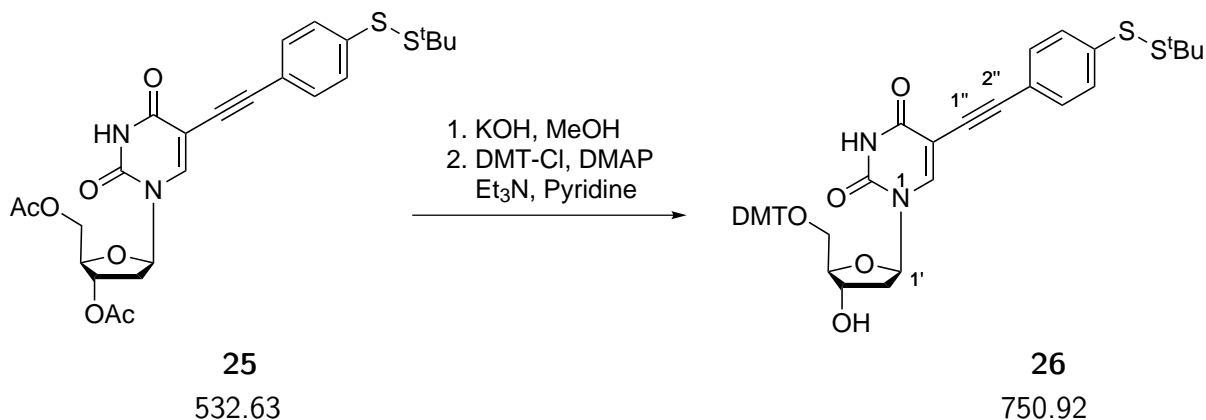
The solvents were removed under reduced pressure and the residue dissolved in dichloromethane (100 mL) and filtered. The filtrate was washed with water (2 x 50 mL), dried over anhydrous MgSO₄ and filtered. The crude product was purified by chromatography on silica using chloroform/methanol (400:1 \rightarrow 100:1) to afford product **25** (631 mg, 1.19 mmol, 63%) as a yellow oil.

¹H NMR: (400 MHz, CDCl₃): δ = 7.83 (s, 1H, H6), 7.44 (d, 2H, J = 8.0 Hz, 2H, Ph), 7.34 (d, J = 8.0 Hz, 2H, Ph), 6.38-6.22 (m, 1H, H1'), 5.21-5.16 (m, 1H, H3'), 4.36-4.28 (m, 2H, H5'), 4.24-4.21 (m, 1H, H4'), 2.55-2.46 (m, 1H, H2'a), 2.29-2.16 (m, 1H, H2'b), 2.11 (s, 3H, OAc), 2.06 (s, 3H, OAc), 1.29 (s, 9H, 3 x CH₃) ppm.

^{13}C NMR: (100.6 MHz, CDCl_3): $\delta = 170.1$ (1C, OAc), 169.9 (1C, OAc), 161.1 (1C, C4), 149.2 (1C, C2), 141.4 (1C, C6), 138.0 (1C, Ph), 131.7 (2C, Ph), 128.8 (2C, Ph), 121.4 (1C, Ph), 100.3 (1C, C5), 92.7 (1C, C2''), 85.3 (1C, C1'), 82.3 (1C, C4'), 81.3 (1C, C1''), 73.8 (1C, C3'), 63.5 (1C, C5'), 49.1 (1C, $\text{SC}(\text{CH}_3)_3$), 37.8 (1C, C2'), 29.7 (3C, $\text{SC}(\text{CH}_3)_3$), 20.6 (1C, CH_3), 20.5 (1C, CH_3) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $\text{C}_{25}\text{H}_{29}\text{N}_2\text{O}_7\text{S}_2$ $[\text{M}+\text{H}]^{3+}$: 177.71, found: 176.97.

5'-*O*-Dimethoxytrityl-5-[4-(*tert*-butyldisulfanyl)phenyl]ethynyl-2'-desoxyuridine



To a solution of potassium hydroxide (103 mg, 1.84 mmol, 4.00 eq) in methanol (25 mL), compound **25** (245 mg, 0.46 mmol, 1.00 eq) was added and the mixture stirred at ambient temperature for 2 h. The solution was neutralised using Amberlite IR-120 H⁺-form, filtered and concentrated under reduced pressure.

The residue was dissolved in pyridine (30 mL), followed by addition of triethylamine (64 μ L, 0.46 mmol, 1.00 eq), 4-(dimethylamino)pyridine (7.0 mg, 0.05 mmol, 0.10 eq) and 4,4'-dimethoxytrityl chloride (164 mg, 0.48 mmol, 1.50 eq). The reaction mixture was stirred for 16 h at ambient temperature under argon atmosphere.

The solvent was removed under reduced pressure and the crude product purified by chromatography on silica using chloroform/methanol/ triethylamine (100:0:1 \rightarrow 100:1:1) to afford product **26** (85.0 mg, 0.11 mmol, 25%) as a white oil.

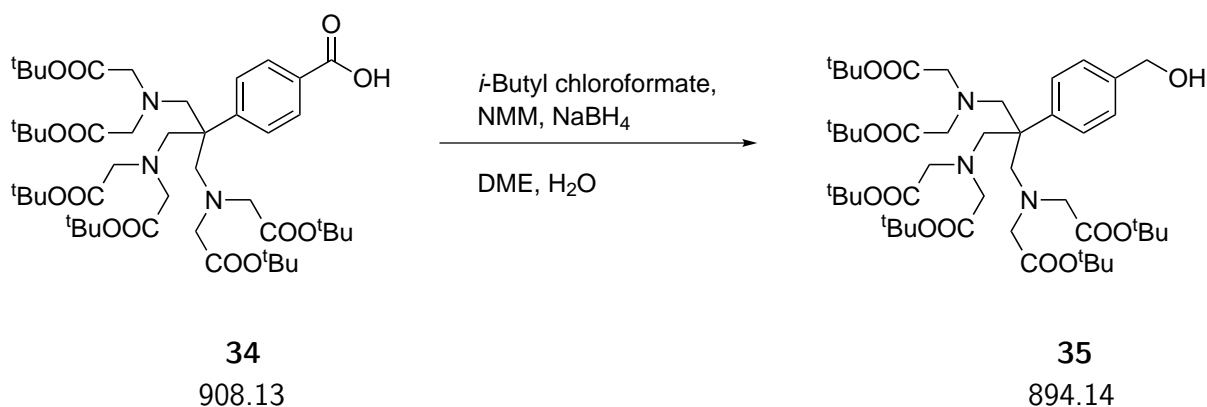
¹H NMR: (400 MHz, CDCl₃): δ = 8.17 (s, 1H, H6), 7.47-7.22 (m, 9H, DMT, 2 x Ph), 7.16-7.11 (m, 2H, DMT), 6.94 (d, J = 8.5 Hz, 2H, 2 x Ph), 6.81-6.76 (m, 4H, DMT), 6.41-6.36 (m, 1H, H1'), 6.10 (br, 1H, NH), 4.58-4.54 (m, 1H, H3'), 4.16-4.11 (m, 1H, H4'), 3.71-3.65 (m, 6H, 2 x OMe), 3.47-3.41 (m, 1H, H5'a), 3.36-3.30 (m, 1H, H5'b), 2.57-2.50 (m, 1H, H2'a), 2.34-2.25 (m, 1H, H2'b), 1.29 (s, 9H, 3 x CH₃) ppm.

¹³C NMR: (100.6 MHz, CDCl₃): δ = 162.8 (1C, C4), 158.5 (2C, DMT), 150.2 (1C, C2), 144.4 (1C, C6), 142.0-113.2 (16C, DMT), 139.0 (1C, Ph), 131.8 (2C, Ph), 126.9 (2C, Ph), 120.4 (1C, Ph), 100.3 (1C, C5), 92.9 (1C, C2''), 86.9 (1C, DMT), 86.6 (1C, C4'), 85.6 (1C, C1'), 81.0 (1C, C1''), 72.0 (1C, C3'), 63.5 (1C, C5'), 55.1 (2C, 2 x OMe), 49.4

(1C, SC(CH₃)₃), 41.7 (1C, C2'), 29.7 (3C, SC(CH₃)₃) ppm.

ESI-MS m/z (acetonitrile, negative mode): calc. for C₄₈H₅₇N₃O₇S₂ [M+TEA]⁻: 851.36, found: 852.42.

4-Hydroxymethyl- α,α,α -tris[[*N,N*-di(*tert*-butoxycarbonylmethyl)amino]methyl]toluene



Compound **34** (390 mg, 0.43 mmol, 1.00 eq) was dissolved in dimethoxyethane (5 mL) and cooled to $-15\text{ }^{\circ}\text{C}$ under argon atmosphere. *N*-methylemorpholine (52 μL , 0.47 mmol, 1.10 eq) and *i*-butyl chloroformate (61.0 μL , 0.47 mmol, 1.10 eq) were added and the mixture stirred for 2 h. The mixture was filtered and the residue washed with dimethoxyethane (5 x 1 mL) under argon atmosphere. The combined filtrate was cooled to $-15\text{ }^{\circ}\text{C}$ and sodium borohydride (24.0 mg, 0.64 mmol, 1.50 eq), dissolved in water (500 μL), was added. The mixture was stirred for 1 h at ambient temperature, followed by addition of water (20 mL).

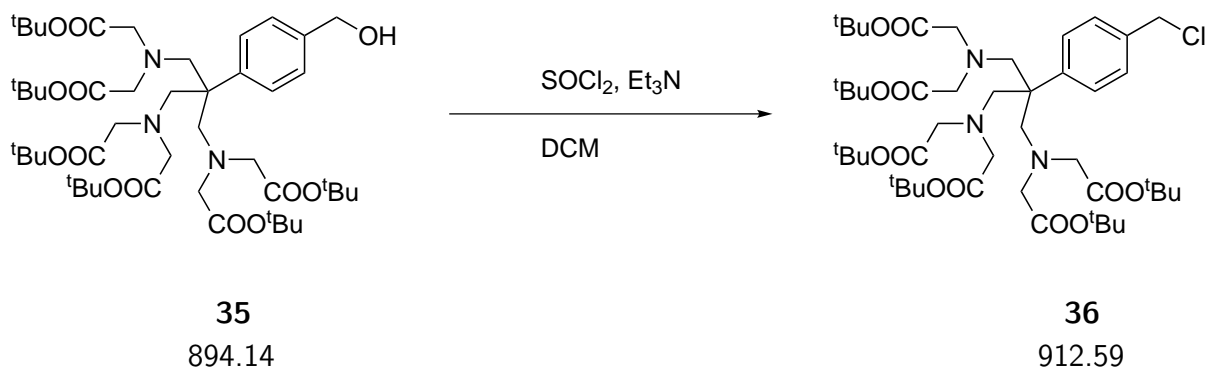
The solvents were removed under reduced pressure and the crude product was purified by chromatography on silica using chloroform/methanol (400:1) to yield product **35** (224 mg, 0.25 mmol, 58%) as a colorless oil.

^1H NMR: (400 MHz, CDCl_3): δ = 7.38 (d, J = 8.2 Hz, 2H, Ph), 7.24 (d, J = 8.2 Hz, 2H, Ph), 4.62 (s, 2 H, CH_2), 3.25 (s, 6H, 3 x CCH_2N), 3.21 (s, 12H, 6 x NCH_2CO), 1.41 (s, 54H, 6 x $\text{C}(\text{CH}_3)_3$) ppm.

^{13}C NMR: (100.6 MHz, CDCl_3): δ = 171.2 (6C, CO), 144.2 (1C, Ph), 138.5 (1C, PH), 127.2 (2C, Ph), 126.8 (2C, Ph), 80.5 (6C, 6 x $\text{C}(\text{CH}_3)_3$), 65.0 (1C, CH_2), 59.4 (3C, 3 x $\text{C}-\alpha\text{CH}_2$), 56.6 (6C, 6 x $\text{NCH}_2\text{CO}_2t\text{-Bu}$), 48.5 (1C, $\text{C}-\alpha$), 28.2 (18C, 18 x $\text{C}(\text{CH}_3)_3$) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $\text{C}_{47}\text{H}_{80}\text{N}_3\text{O}_{13}$ $[\text{M}+\text{H}]^+$: 894.57, found: 894.60.

4-Chloromethyl- α,α,α -tris[[*N,N*-di(*tert*-butoxycarbonylmethyl)amino]methyl]toluene



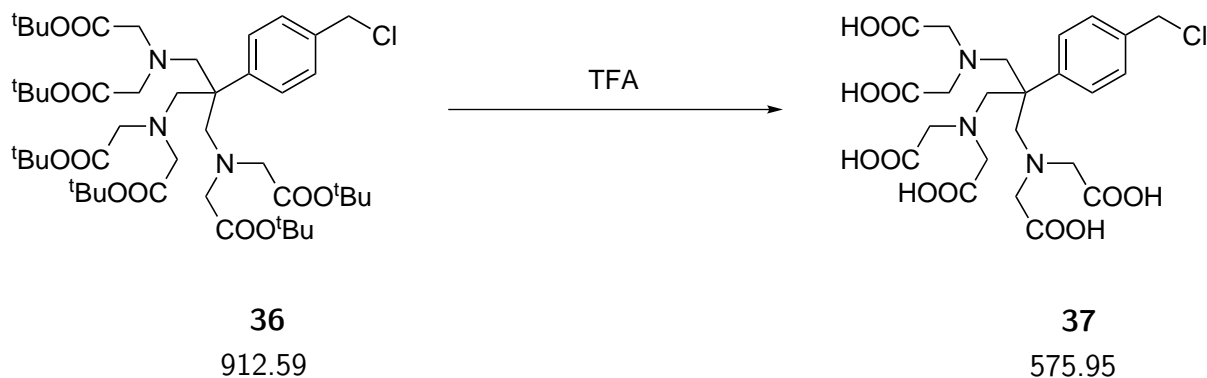
A solution of compound **35** (380 mg, 0.43 mmol, 1.00 eq) in dichloromethane (15 mL) was cooled to 0 °C under argon atmosphere. Triethylamine (71.0 μL , 0.51 mmol, 1.20 eq) and thionyl chloride (46.0 μL , 0.64 mmol, 1.50 eq) were added via syringe. The solution was stirred for 0.5 h at low temperature, followed by heating to reflux for 24 h.

The solvent was removed under reduced pressure and the crude product was purified by chromatography on silica using hexane/ethyl acetate (10:1) to yield product **36** (303 mg, 0.33 mmol, 78%) as a colorless oil while educt **35** was recovered (40 mg, 0.05 mmol, 11%).

^1H NMR: (400 MHz, CDCl_3): δ = 7.41 (d, J = 8.3 Hz, 2H, Ph), 7.27 (d, J = 8.3 Hz, 2H, Ph), 4.54 (s, 2 H, CH_2), 3.27 (s, 6H, 3 x CCH_2N), 3.22 (s, 12H, 6 x NCH_2CO), 1.42 (s, 54H, 6 x $\text{C}(\text{CH}_3)_3$) ppm.

^{13}C NMR: (100.6 MHz, CDCl_3): δ = 171.1 (6C, CO), 145.3 (1C, Ph), 134.9 (1C, Ph), 128.7 (2C, Ph), 127.0 (2C, Ph), 80.5 (6C, 6 x $\text{C}(\text{CH}_3)_3$), 59.4 (3C, 3 x $\text{C}-\alpha\text{CH}_2$), 56.5 (6C, 6 x $\text{NCH}_2\text{CO}_2t\text{-Bu}$), 48.7 (1C, $\text{C}-\alpha$), 45.9 (1C, CH_2), 28.2 (18C, 18 x $\text{C}(\text{CH}_3)_3$) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $\text{C}_{47}\text{H}_{79}\text{ClIN}_3\text{O}_{12}$ $[\text{M}+\text{H}]^+$: 912.53, found: 912.59.

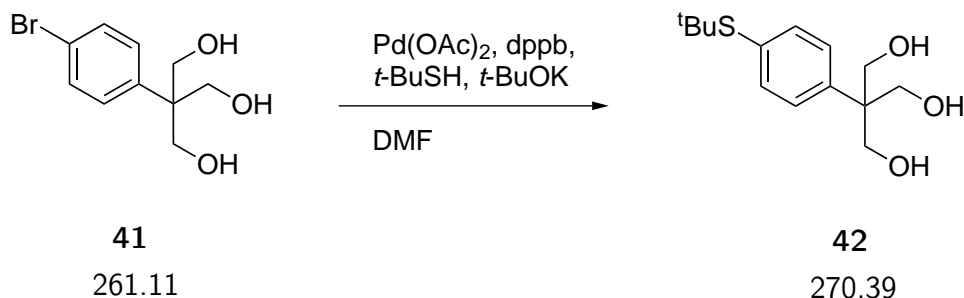
4-Chloromethyl- α,α,α -tris[[*N,N*-di(carboxymethyl)amino]methyl]-toluene, Cl-Bn-TAHA

Compound **36** (91.0 mg, 0.10 mmol, 1.00 eq) was dissolved in conc. trifluoroacetic acid (10 mL) and stirred at ambient temperature for 12 h, after which the solvent was removed under reduced pressure. Water (8 mL) was added to the residue and centrifuged. The supernatant was purified by HPLC (0.1% TFA in water : 0 - 30 min 0 \rightarrow 60% acetonitrile) to afford Cl-Bn-TAHA **37** as a white powder (48.8 mg, 0.09 mmol, 85%).

NMR: Due to the low stability of the molecule no NMR spectra were recorded.

HPLC (preparative): t_R : 18.8 min

ESI-MS m/z (acetonitrile, negative mode): calc. for $C_{23}H_{31}N_3O_{12}Cl$ $[M+H]^+$: 576.16, found: 576.13.

4-(*tert*-Butylsulfanyl)- α,α,α -tris(hydroxymethyl)toluene

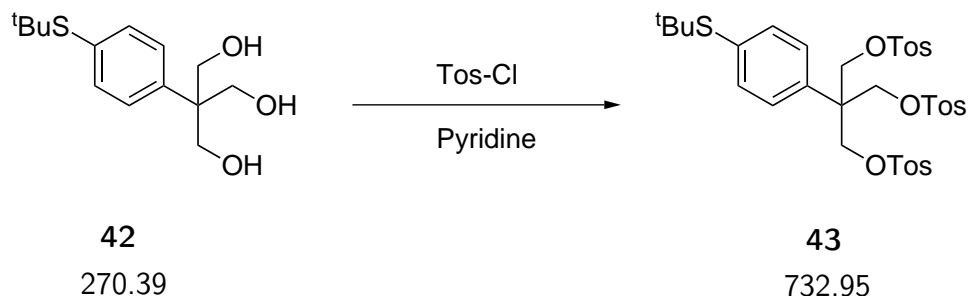
A solution of compound **41** (780 mg, 2.99 mmol, 1.00 eq) and potassium *tert*-butoxide (1.14 g, 10.2 mmol, 3.40 eq) in dimethylformamide (20 mL) was degassed by bubbling argon through for 0.5 h. Palladium(II) acetate (67.0 mg, 0.30 mmol, 0.10 eq), 1,4-bis(diphenylphosphino)butane (64.0 mg, 0.15 mmol, 0.05 eq) and 2-methyl-2-propanethiol (404 μ L, 3.59 mmol, 1.20 eq) were added. The suspension was stirred for 21 h at 90 °C under argon atmosphere.

The solvent was removed under reduced pressure and the crude product was dissolved in ethyl acetate and filtered. The resulting solution was evaporated under reduced pressure and purified by chromatography on silica using hexane/acetone (1:1) to yield product **42** (663 mg, 2.45 mmol, 82%) as a colorless oil.

^1H NMR: (400 MHz, MeOD): δ = 7.54-7.46 (m, 4H, Ph), 3.97 (s, 6H, 3 x CH_2), 1.30 (s, 9H, *t*-Bu) ppm.

^{13}C NMR: (100.6 MHz, MeOD): δ = 143.8 (1C, Ph), 138.3 (2C, Ph), 131.8 (1C, Ph), 128.9 (2C, Ph), 65.4 (3C, 3 x CH_2), 50.4 (1C, C- α), 46.3 (1C, C(CH_3) $_3$), 31.4 (3C, 3 x CH_3) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $\text{C}_{14}\text{H}_{23}\text{O}_3\text{S}$ $[\text{M}+\text{H}]^+$: 271.14, found: 271.12.

4-(*tert*-Butylsulfanyl)- α,α,α -tris(tosyloxymethyl)toluene

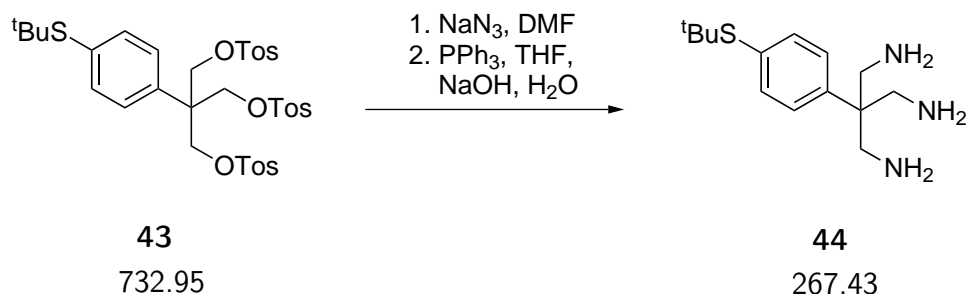
A solution of compound **42** (473 mg, 1.75 mmol, 1.00 eq) in pyridine (10 mL) was cooled to 0 °C. *p*-Toluenesulfonyl chloride (3.34 g, 17.5 mmol, 10.0 eq) was added under argon atmosphere and the solution was stirred for 60 h while the temperature was allowed to rise naturally to ambient temperature.

The resulting suspension was filtered and the solvent was removed under reduced pressure. Water (30 mL) was added to the residue and the mixture was extracted with dichloromethane (3 x 30 mL). The combined organic layers were dried over anhydrous MgSO₄ and filtered. The solvent was removed under reduced pressure and the crude product purified by chromatography on silica using hexane/dichloromethane (2:1 → 0:1) to yield product **43** (1.08 g, 1.47 mmol, 84%) as a yellow solid.

¹H NMR: (400 MHz, MeOD): δ = 7.65-7.61 (m, 6H, Tos), 7.34-7.28 (m, 8H, 6 x Tos, 2 x Ph), 6.88 (d, J = 8.3 Hz, 2H, Ph), 4.19 (s, 6H, 3 x CH₂), 2.46 (s, 9H, 3 x CH₃), 1.27 (s, 9H, *t*-Bu) ppm.

¹³C NMR: (100.6 MHz, MeOD): δ = 145.4 (3C, Tos), 137.5 (2C, Ph), 135.3 (1C, Ph), 133.3 (1C, Ph), 131.6 (3C, Tos), 130.1 (6C, Tos), 127.9 (6C, Tos), 126.2 (2C, Ph), 68.7 (3C, 3 x CH₂), 46.3 (1C, C- α), 46.0 (1C, C(CH₃)₃), 31.0 (3C, 3 x C(CH₃)₃), 21.7 (3C, 3 x CH₃) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for C₃₅H₄₀NaO₉S₄ [M+H]⁺: 755.14, found: 755.11.

4-(*tert*-Butylsulfanyl)- α,α,α -tris(aminomethyl)toluene

To a solution of compound **43** (826 mg, 1.13 mmol, 1.00 eq) in dimethylformamide (16 mL), sodium azide (440 mg, 6.76 mmol, 6.00 eq) was added under argon atmosphere and the mixture was stirred for 72 h at 100 °C. Water (50 mL) was added to the cooled mixture and extracted with dichloromethane (3 x 20 mL). The combined organic layers were dried over anhydrous MgSO₄ and filtered. The solvent was reduced to 3 mL, followed by the addition of tetrahydrofuran (10 mL). Again, the solvent was reduced to 3 mL and the procedure was repeated 2 times.

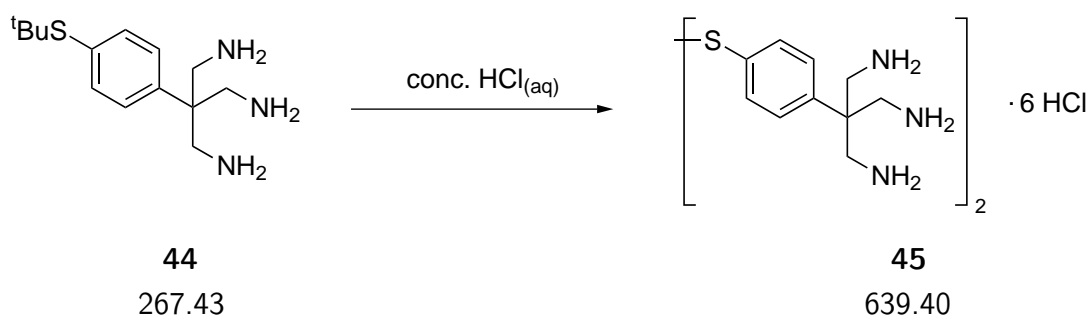
The resulting solution was diluted with tetrahydrofuran (10 mL), followed by the addition of triphenylphosphine (2.66 g, 10.1 mmol, 9.00 eq). The mixture was stirred at 65 °C for approx. 4 h until the gas evolution was finished. Subsequently, 2 M sodium hydroxide solution (10 mL) was added and the mixture was stirred at 65 °C for 2 h.

The mixture was cooled and the pH adjusted to 1 using hydrochloride solution. The aqueous layer was washed with dichloromethane (3 x 30 mL) and the pH of the aqueous layer adjusted to 12 using sodium hydroxide. The aqueous layer was extracted with dichloromethane (6 x 20 mL). The combined organic layers were dried over anhydrous MgSO₄ and filtered. The solvent was removed under reduced pressure to yield product **44** (239 mg, 0.89 mmol, 79%) as a yellow oil.

¹H NMR: (400 MHz, DMSO): δ = 8.35 (s, 6H, 3 x NH₂), 7.53 (d, J = 8.4 Hz, 2H, Ph), 7.46 (d, J = 8.4 Hz, 2H, Ph), 3.58 (s, 6H, 3 x CH₂), 1.29 (s, 9H, *t*-Bu) ppm.

¹³C NMR: (100.6 MHz, DMSO): δ = 137.7 (2C, Ph), 135.4 (1C, Ph), 132.5 (1C, Ph), 127.5 (2C, Ph), 65.4 (3C, 3 x CH₂), 43.0 (1C, C- α), 46.2 (1C, C(CH₃)₃), 30.7 (3C, 3 x CH₃) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $C_{14}H_{26}N_3S [M+H]^+$: 268.18, found: 268.03.

Bis[4-[α,α,α -tris(aminomethyl)methyl]phenyl] disulfide hexahydrochloride

To compound **44** (240 mg, 0.90 mmol, 1.00 eq) was added conc. hydrochloric acid (5 mL) and the mixture was stirred at 100 °C for 3 h.

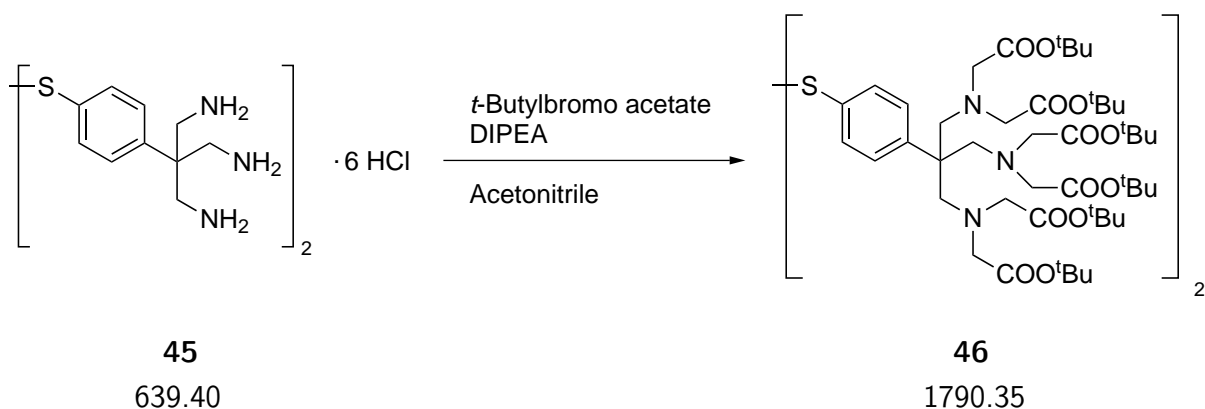
The solvent was removed under reduced pressure, followed by the addition of methanol (20 mL). Suction filtration and washing with methanol (20 mL) afforded compound **45** (225 mg, 0.35 mmol, 78%) as a brown solid.

^1H NMR: (400 MHz, D_2O): $\delta = 7.59$ (d, $J = 8.6$ Hz, 2H, Ph), 7.46 (d, $J = 8.6$ Hz, 2H, Ph), 3.69 (s, 6H, 3 x CH_2) ppm.

^{13}C NMR: (100.6 MHz, D_2O): $\delta = 134.5$ (1C, Ph), 130.4 (2C, Ph), 129.1 (1C, Ph), 127.3 (2C, Ph), 42.3 (1C, C- α), 40.8 (3C, 3 x CH_2) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $\text{C}_{20}\text{H}_{33}\text{N}_6\text{S}_2$ $[\text{M}+\text{H}]^+$: 421.22, found: 421.20.

Bis[4- α,α,α -tris[*N,N*-di(*tert*-butoxycarbonylmethyl)amino-methyl]methyl]phenyl] disulfide



To a solution of compound **45** (100 mg, 0.16 mmol, 1.00 eq) in acetonitrile (30 mL), diisopropylethylamine (665 μL , 3.82 mmol, 24.4 eq) and *t*-butyl bromoacetate (416 μL , 2.81 mmol, 18.0 eq) were added under argon atmosphere. The mixture was heated to reflux for 48 h.

The solvent was removed under reduced pressure, the residue dissolved in dichloromethane (30 mL) and washed with 5% citric acid solution (2 x 10 mL) and water (10 mL). The combined organic layers were dried over anhydrous MgSO_4 and filtered. The solvent was removed under reduced pressure and the crude product purified by chromatography on silica using hexane/ethyl acetate (6:1) to yield product **46** (221 mg, 0.12 mmol, 77%) as a colorless oil.

^1H NMR: (400 MHz, CDCl_3): δ = 7.34 (d, J = 8.3 Hz, 2H, Ph), 7.29 (d, J = 8.3 Hz, 2H, Ph), 3.21 (s, 6H, 3 x CH_2), 3.19 (s, 12H, 6 x CH_2), 1.38 (s, 54H, *t*-Bu) ppm.

^{13}C NMR: (100.6 MHz, CDCl_3): δ = 171.5 (6C, CO), 143.9 (1C, Ph), 132.6 (1C, Ph), 130.1 (2C, Ph), , 127.2 (2C, Ph), 80.4 (6C, 6 x $\text{C}(\text{CH}_3)_3$), 59.0 (3C, 3 x CH_2), 56.3 (6C, 6 x CH_2), 48.7 (1C, C- α), 28.2 (18C, 6 x $\text{CC}(\text{H}_3)_3$) ppm.

ESI-MS m/z (acetonitrile, positive mode): calc. for $\text{C}_{47}\text{H}_{80}\text{N}_3\text{O}_{15}$ $[\text{M}+\text{O}+\text{H}]^+$: 1806.03, found: 1806.28

Cleavage of Modified Oligonucleotide on Solid Support

The oligonucleotide (1 μmol scale) on solid support was incubated with a 0.2 M sodium sulfide solution in conc. aqueous ammonia/ethanol solution (4:1, 1 mL) for 12 h at 55 °C. The mixture was vortexed, centrifuged and the supernatant stored separately. The solid support was washed twice with a water/ethanol solution (3:1, 1 mL) and the combined aqueous solution was incubated with tris(2-carboxyethyl)phosphine (excess) at ambient temperature for 1 h under constant shaking. The solvent was removed under reduced pressure using a rotary evaporator. The crude product was dissolved in water (1 mL), centrifuged and the supernatant purified by HPLC (0.1 M triethylammonium acetate buffer/acetonitrile: 0 - 5 min 3% acetonitrile, 5 - 25 min 3 \rightarrow 25% acetonitrile) to afford one DNA peak with a retention time of 17.9 min, controlled by mass spectrometry. The combined product fractions were lyophilised and stored under argon at -20 °C.

Tagging Reaction and Sample Preparation of Modified Oligonucleotide

The modified free thiol oligonucleotide (1 μmol) was incubated with the solution of the preloaded tag (10 μmol , 10 eq) at 55 $^{\circ}\text{C}$ for 12 h under argon atmosphere. HPLC purification (0.1 M triethylammonium acetate buffer/acetonitrile: 0 - 5 min 3% acetonitrile, 5 - 25 min 3 \rightarrow 25% acetonitrile) afforded the product with a retention time of 18.6 min controlled by mass spectrometry. The combined product solutions were diluted with an aqueous buffer solution (1 M sodium chloride, 50 mM MOPS, pH = 7.6) to twice the volume and resulting solution was reduced to 200 μL using a vacuum concentrator at 45 $^{\circ}\text{C}$. To the DNA solution, the aqueous buffer (15 mL) was added and the resulting solution reduced to 500 μL with a *Millipore* Amicon concentration device with a MWCO of 3000 g/mol. The washing procedure was repeated 10 times, followed by addition of aqueous buffer (50 mL). The DNA solution was heated in an 70 $^{\circ}\text{C}$ water bath for 5 min and allowed to slowly cool to ambient temperature. The solution was reduced to 20 μL using Amicon concentration device with a MWCO of 3000 g/mol. The DNA containing solution was diluted with of the final NMR buffer (1 mL, 20 mmol MOPS, 180 mM sodium chloride in deuterium oxide, pD = 8.0), reduced again to 20 μL and transferred to a *Shigemi* NMR tube using 250 μL NMR buffer in total.

Preparation of the Tagging Solution [Siepel2013]

To a solution of Cys-Ph-TAHA **1** (752 μL , 10.0 μmol , 10.0 eq), lanthanide trichloride (226 μL , 12.0 μmol , 1.20 eq) was added. The solution was shaken at ambient temperature for 2 h and the pH adjusted to 7.8 with 0.1 M sodium hydroxide solution. After centrifugation, the supernatant was used for the tagging reaction.

Bibliography

- [Allegrozzi2000] M. Allegrozzi, I. Bertini, M. B. L. Janik, Y. M. Lee, G. Liu, C. Luchinat, *Journal of the American Chemical Society* **2000**, *122*, 4154–4161.
- [Babaylova2014] E. S. Babaylova, A. V. Ivanov, A. A. Malygin, M. A. Vorobjeva, A. G. Venyaminova, Y. F. Polienko, I. A. Kirilyuk, O. A. Krunkacheva, M. V. Fedin, G. G. Karpova, E. G. Bagryanskaya, *Organic & Biomolecular Chemistry* **2014**, *12*, 3129–3136.
- [Barthelmes2010] K. Barthelmes, A. M. Reynolds, E. Peisach, H. R. A. Jonker, N. J. DeNunzio, K. N. Allen, B. Imperiali, H. Schwalbe, *Journal of the American Chemical Society* **2010**, *133*, 808–819.
- [Bax1997] A. Bax, N. Tjandra, *Journal of Biomolecular NMR* **1997**, *10*, 289–292.
- [Beaucage1981] S. Beaucage, M. Caruthers, *Tetrahedron Letters* **1981**, *22*, 1859–1862.
- [Beaucage1990] R. P. Iyer, W. Egan, J. B. Regan, S. L. Beaucage, *Journal of the American Chemical Society* **1990**, *112*, 1253–1254.
- [Bertini2001] I. Bertini, M. B. L. Janik, Y. M. Lee, C. Luchinat, A. Rosato, *Journal of the American Chemical Society* **2001**, *123*, 4181–4188.
- [Bertini2002a] I. Bertini, C. Luchinat, G. Parigi, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2002**, *40*, 249–273.
- [Bertini2002b] I. Bertini, G. Cavallaro, M. Cosenza, R. Kümmerle, C. Luchinat, M. Piccioli, L. Poggi, *Journal of Biomolecular NMR* **2002**, *23*, 115–125.
- [Bertini2003] I. Bertini, I. Gelis, N. Katsaros, C. Luchinat, A. Provenzani, *Biochemistry* **2003**, *42*, 8011–8021.

- [Biekofsky1999] R. R. Biekofsky, F. W. Muskett, J. M. Schmidt, S. R. Martin, J. Browne, P. M. Bayley, J. Feeney, *FEBS Letters* **1999**, *460*, 519–526.
- [Boisbouvier1999] J. Boisbouvier, P. Gans, M. Blackledge, B. Brutscher, D. Marion, *Journal of the American Chemical Society* **1999**, *121*, 7700–7701.
- [Bothe2011] J. R. Bothe, E. N. Nikolova, C. D. Eichhorn, J. Chugh, A. L. Hansen, H. M. Al-Hashimi, *Nature Methods* **2011**, *8*, 919–931.
- [Brooks1983] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *Journal of Computational Chemistry* **1983**, *4*, 187–217.
- [Brunger1998] A. T. Brunger, P. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Cryst.* **1998**, *D54*, 905–921.
- [Brunger2007] A. T. Brunger, *Nat. Protocols* **2007**, *2*, 2728–2733.
- [Camacho-Zarco2014] A. R. Camacho-Zarco, F. Munari, M. Wegstroth, W. M. Liu, M. Ubbink, S. Becker, M. Zweckstetter, *Angewandte Chemie International Edition* **2014**, *53*, 1–5.
- [Canales2013] A. Canales, A. Mallagaray, J. Pérez-Castells, I. Boos, C. Unverzagt, S. André, H. J. Gabius, F. J. Canada, J. Jiménez-Barbero, *Angewandte Chemie International Edition* **2013**, *52*, 13789–13793.
- [Canales2014] A. Canales, A. Mallagaray, M. A. Berbis, A. Navarro-Vazquez, G. Dominguez, F. J. Canada, S. André, H. J. Gabius, J. Pérez-Castells, J. Jiménez-Barbero, *Journal of the American Chemical Society* **2014**, *136*, 8011–8017.
- [Caruthers2001] S. L. Beaucage, M. H. Caruthers, *Synthetic Strategies and Parameters Involved in the Synthesis of Oligodeoxyribonucleotides According to the Phosphoramidite Method*. John Wiley & Sons, Inc., **2001**.
- [Chen1998] X. Chen, U. Vinkemeier, Y. Zhao, D. Jeruzalmi, J. E. Darnell, Jr., J. Kuriyan, *Cell* **1998**, *93*, 827–839.

- [Christian2000] E. L. Christian, N. M. Kaye, M. E. Harris, *RNA* **2000**, *6*, 511–519.
- [Christopherson1991] M. S. Christopherson, A. D. Broom, *Nucleic Acids Research* **1991**, *19*, 5719–5724.
- [Clore2009] G. M. Clore, J. Iwahara, *Chemical Reviews* **2009**, *109*, 4108–4139.
- [Coleman1994] R. S. Coleman, E. A. Kesicki, *Journal of the American Chemical Society* **1994**, *116*, 11636–11642.
- [Dasgupta2011] S. Dasgupta, X. Hu, P. H. J. Keizers, W. M. Liu, C. Luchinat, M. Nagulapalli, M. Overhand, G. Parigi, L. Sgheri, M. Ubbink, *Journal of Biomolecular NMR* **2011**, *51*, 253–263.
- [Dvoretzky2002] A. Dvoretzky, V. Gaponenko, P. R. Rosevear, *FEBS Letters* **2002**, *528*, 189–192.
- [Eckstein1985] F. Eckstein, *Annual Review of Biochemistry* **1985**, *54*, 367–402.
- [Edwards2001] T. E. Edwards, T. M. Okonogi, B. H. Robinson, S. T. Sigurdsson, *Journal of the American Chemical Society* **2001**, *123*, 1527–1528.
- [Ellman1959] G. L. Ellman, *Archives of Biochemistry and Biophysics* **1959**, *82*, 70–77.
- [Erdelyi2011] M. Erdélyi, E. d’Auvergne, A. Navarro-Vazquez, A. Leonov, C. Griesinger, *Chemistry - A European Journal* **2011**, *17*, 9368–9376.
- [Esquiaqui2014] J. M. Esquiaqui, E. M. Sherman, S. A. Ionescu, J. D. Ye, G. E. Fanucci, *Biochemistry* **2014**, *53*, 3526–3528.
- [Fernandez2003] C. Fernandez, G. Wider, *Current Opinion in Structural Biology* **2003**, *13*, 570–580.
- [Frederiksen2009] J. K. Frederiksen, J. A. Piccirilli, *Chapter 14 - Separation of RNA Phosphorothioate Oligonucleotides by HPLC*, Academic Press, **2009**.
- [Fujiki2002] K. Fujiki, N. Tanifuji, Y. Sasaki, T. Yokoyama, *Synthesis* **2002**, *3*, 343–348.

- [Gaponenko2000] V. Gaponenko, A. Dvoretzky, C. Walsby, B. M. Hoffman, P. R. Rosevear, *Biochemistry* **2000**, *39*, 15217–15224.
- [Gaur1989] R. Gaur, S. Paliwal, P. Sharma, K. Gupta, *Journal of Biochemical and Biophysical Methods* **1989**, *18*, 323–329.
- [Gayathri1982] C. Gayathri, A. A. Bothner-By, P. C. M. van Zijl, C. Maclean, *Chemical Physics Letters* **1982**, *87*, 192–196.
- [Ghose1997] R. Ghose, J. H. Prestegard, *Journal of Magnetic Resonance* **1997**, *128*, 138–143.
- [Goddard2008] T. Goddard, D. Kneller, *Sparky 3*, University of California, San Francisco, **2008**.
- [Graham2011] B. Graham, C. T. Loh, J. D. Swarbrick, P. Ung, J. Shin, H. Yagi, X. Jia, S. Chhabra, N. Barlow, G. Pintacuda, T. Huber, G. Otting, *Bioconjugate Chemistry* **2011**, *22*, 2118–2125.
- [Gueron1975] M. Gueron, *Journal of Magnetic Resonance* **1975**, *19*, 58–66.
- [Guzaev2011] A. P. Guzaev, *Tetrahedron Letters* **2011**, *52*, 434 - 437.
- [Haberz2006] P. Haberz, F. Rodriguez-Castaneda, J. Junker, S. Becker, A. Leonov, C. Griesinger, *Organic Letters* **2006**, *8*, 1275–1278.
- [Haeussinger2009] D. Haeussinger, J. Huang, S. Grzesiek, *Journal of the American Chemical Society* **2009**, *131*, 14761–14767.
- [Haiss2006] W. Haiss, C. Wang, I. Grace, A. S. Batsanov, D. J. Schiffrin, S. J. Higgins, M. R. Bryce, C. J. Lambert, R. J. Nichols, *Nature Materials* **2006**, *5*, 995–1002.
- [Hallwass2011] F. Hallwass, M. Schmidt, H. Sun, A. Mazur, G. Kummerloewe, B. Luy, A. Navarro-Vazquez, C. Griesinger, U. M. Reinscheid, *Angewandte Chemie International Edition* **2011**, *50*, 9487–9490.
- [Hansen1998] M. R. Hansen, L. Mueller, A. Pardi, *Nature Structural & Molecular Biology* **1998**, *5*, 1065–1074.
- [Hiruma2013] Y. Hiruma, M. A. Hass, Y. Kikui, W. M. Liu, B. Oelmez, S. P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Loehr, H. Schwalbe, M. Nojiri, M. Ubbink, *Journal of Molecular Biology* **2013**, *425*, 4353–4365.

- [Ikegami2004] T. Ikegami, L. Verdier, P. Sakhaii, S. Grimme, B. Pescatore, K. Saxena, K. M. Fiebig, C. Griesinger, *Journal of Biomolecular NMR* **2004**, *29*, 339–349.
- [Ingale2014] S. A. Ingale, P. Leonard, H. Yang, F. Seela, *Organic & Biomolecular Chemistry* **2014**, 8519-8532.
- [Ippel1996] J. H. Ippel, S. S. Wijmenga, R. de Jong, H. A. Heus, C. W. Hilbers, E. de Vroom, G. A. van der Marel, J. H. van Boom, *Magnetic Resonance in Chemistry* **1996**, *34*, S156–S176.
- [Iwahara2006] J. Iwahara, G. M. Clore, *Nature* **2006**, *440*, 1227–1230.
- [Jia2011a] X. Jia, A. Maleckis, T. Huber, G. Otting, *Chemistry - A European Journal* **2011**, *17*, 6830–6836.
- [Jia2011b] X. Jia, H. Yagi, X. C. Su, M. Stanton-Cook, T. Huber, G. Otting, *Journal of Biomolecular NMR* **2011**, *50*, 411–420.
- [Kawabuchi2008] H. Kawabuchi, Jpn. Kokai Tokkyo Koho., No: JP2008110363, **2008**.
- [Keizers2007] P. H. J. Keizers, J. F. Desreux, M. Overhand, M. Ubbink, *Journal of the American Chemical Society* **2007**, *129*, 9292–9293.
- [Keizers2008] P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand, M. Ubbink, *Journal of the American Chemical Society* **2008**, *130*, 14802–14812.
- [Kessler1985] H. Kessler, A. Mueller, H. Oschkinat, *Magnetic Resonance in Chemistry* **1985**, *23*, 844–852.
- [Koehler2011] J. Koehler, J. Meiler, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2011**, *59*, 360–389.
- [Kosen1986] P. A. Kosen, R. M. Scheek, H. Naderi, V. J. Basus, S. Manogaran, P. G. Schmidt, N. J. Oppenheimer, I. D. Kuntz, *Biochemistry* **1986**, *25*, 2356–2364.
- [Kramer2004] F. Kramer, H. Deshmukh, H. Kessler, S. J. Glaser, *Concepts in Magnetic Resonance, Part A* **2004**, *21A*, 10–21.
- [Kung1995] H. C. Kung, K. Y. Wang, I. Goljer, P. H. Bolton, *Journal of Magnetic Resonance* **1995**, *109*, 323–325.

- [Kurland1970] R. J. Kurland, B. R. McGarvey, *Journal of Magnetic Resonance* **1970**, *2*, 286–301.
- [LaMar1993] G. N. LaMar, J. S. de Ropp, NMR Methodology for Paramagnetic Proteins. ; L. J. Berliner, J. Reuben, Eds., Springer US, **1993**.
- [Lee1983] L. Lee, B. D. Sykes, *Biochemistry* **1983**, *22*, 4366–4373.
- [Leonov2005] A. Leonov, B. Voigt, F. Rodriguez-Castaneda, P. Sakhaii, C. Griesinger, *Chemistry - A European Journal* **2005**, *11*, 3342–3348.
- [Liu2012] W. M. Liu, P. H. J. Keizers, M. A. S. Hass, A. Blok, M. Timmer, A. C. J. Sarris, M. Overhand, M. Ubbink, *Journal of the American Chemical Society* **2012**, *134*, 17306–17313.
- [Loh2013] C. T. Loh, K. Ozawa, K. L. Tuck, N. Barlow, T. Huber, G. Otting, B. Graham, *Bioconjugate Chemistry* **2013**, *24*, 260–268.
- [Man2010] B. Man, X. C. Su, H. Liang, S. Simonson, T. Huber, B. A. Messerle, G. Otting, *Chemistry - A European Journal* **2010**, *16*, 3827–3832.
- [Mao2005] X. Mao, Z. Ren, G. N. Parker, H. Sondermann, M. A. Pastorello, W. Wang, J. S. McMurray, B. Demeler, J. E. Darnell, Jr., X. Chen, *Molecular Cell* **2005**, *17*, 761–771.
- [Matteucci1981] M. D. Matteucci, M. H. Caruthers, *Journal of the American Chemical Society* **1981**, *103*, 3185–3191.
- [Neculai2005] D. Neculai, A. M. Neculai, S. Verrier, K. Straub, K. Klumpp, E. Pfizner, S. Becker, *The Journal of Biological Chemistry* **2005**, *280*, 40782–40787.
- [Nitz2003] M. Nitz, K. J. Franz, R. L. Maglathlin, B. Imperiali, *ChemBioChem* **2003**, *4*, 272–276.
- [Nukaga2012] Y. Nukaga, K. Yamada, T. Ogata, N. Oka, T. Wada, *Journal of Organic Chemistry* **2012**, *77*, 7913–7922.
- [Oka2008] N. Oka, T. Kondo, S. Fujiwara, Y. Maizuru, T. Wada, *Organic Letters* **2009**, *11*, 967–970.

- [Ottiger1998] M. Ottiger, F. Delaglio, A. Bax, *Journal of Magnetic Resonance* **1998**, *131*, 373–378.
- [Otting2008] G. Otting, *Journal of Biomolecular NMR* **2008**, *42*, 1–9.
- [Otting2010] G. Otting, *Annual Review of Biochemistry* **2010**, *39*, 387–405.
- [Pervushin1997] K. Pervushin, R. Riek, G. Wider, K. Wuethrich, *Proceedings of the National Academy of Sciences* **1997**, *94*, 12366–12371.
- [Peters2010] F. Peters, *Entwicklung von Lanthanoid-Tags fuer die biomolekulare NMR-Spektroskopie*, PhD thesis, Goettingen, **2010**.
- [Peters2011] F. Peters, A. Maestre-Martinez, M. Leonov, L. Kovacic, S. Becker, R. Boelens, C. Griesinger, *Journal of Biomolecular NMR* **2011**, *51*, 329–337.
- [Pintacuda2004] G. Pintacuda, A. Kaikkonen, G. Otting, *Journal of Magnetic Resonance* **2004**, *171*, 233–243.
- [Pintacuda2007] G. Pintacuda, M. John, X. C. Su, G. Otting, *Accounts of Chemical Research* **2007**, *40*, 206–212.
- [Piton2007] N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, J. W. Engels, *Nucleic Acids Research* **2007**, *35*, 3128–3143.
- [Prudencio2004] M. Prudencio, J. Rohovec, J. A. Peters, E. Tocheva, M. J. Boulanger, M. E. P. Murphy, H. J. Hupkes, W. Kusters, A. Impagliazzo, M. Ubbink, *Chemistry - A European Journal* **2004**, *10*, 3252–3260.
- [Qin2001] P. Z. Qin, S. E. Butcher, J. Feigon, W. L. Hubbell, *Biochemistry* **2001**, *40*, 6929–6936.
- [Qin2003] P. Z. Qin, K. Hideg, J. Feigon, W. L. Hubbell, *Biochemistry* **2003**, *42*, 6772–6783.
- [Qin2007] P. Z. Qin, I. S. Haworth, Q. Cai, A. K. Kusnetzow, G. P. G. Grant, E. A. Price, G. Z. Sowa, A. Popova, B. Herreros, H. He, *Nature Protocols* **2007**, *2*, 2354–2365.
- [Roberts1993] G. C. K. Roberts, *NMR of Macromolecules*, Oxford University Press, New York, **1993**.

- [Russo2013] L. Russo, M. Maestre-Martinez, S. Wolff, S. Becker, C. Griesinger, *Journal of the American Chemical Society* **2013**, *135*, 17111–17120.
- [Schmidt1984] P. G. Schmidt, I. D. Kuntz, *Biochemistry* **1984**, *23*, 4261–4266.
- [Schmidt2012] M. Schmidt, H. Sun, P. Rogne, G. K. E. Scriba, C. Griesinger, L. T. Kuhn, U. M. Reinscheid, *Journal of the American Chemical Society* **2012**, *134*, 3080–3083.
- [Sicoli2010] G. Sicoli, F. Wachowius, M. Bennati, C. Hoebartner, *Angewandte Chemie International Edition* **2010**, *49*, 6443–6447.
- [Siepel2013] F. Siepel, *The Advantages Of Paramagnetic NMR*, PhD thesis, Goettingen, **2013**.
- [Slim1991] G. Slim, M. J. Gait, *Nucleic Acids Research* **1991**, *19*, 1183–1188.
- [Sonogashira1975] K. Sonogashira, Y. Tohda, N. Hagihara, *Tetrahedron Letters* **1975**, *50*, 4467–4470.
- [Sonogashira2002] K. Sonogashira, *Journal of Organometallic Chemistry* **2002**, *653*, 46–49.
- [Su2008] X. C. Su, B. Man, S. Beeren, H. Liang, S. Simonson, C. Schmitz, T. Huber, B. A. Messerle, G. Otting, *Journal of the American Chemical Society* **2008**, *130*, 10486–10487.
- [Swarbrick2011] J. D. Swarbrick, P. Ung, S. Chhabra, B. Graham, *Angewandte Chemie International Edition* **2011**, *50*, 4403–4406.
- [Taeubert2010] S. Taeubert, *Paramagnetisch Markierung von DNA und Strukturaufklaerung mittels NMR-Spektroskopie*, Diplomarbeit, Goettingen, **2009**.
- [Tolman1995] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proceedings of the National Academy of Sciences* **1995**, *92*, 9279–9283.
- [Tycko2000] R. Tycko, F. J. Blanco, Y. Ishii, *Journal of the American Chemical Society* **2000**, *122*, 9340–9341.

- [Usman1987] N. Usman, K. K. Ogilvie, M. Y. Jiang, R. J. Cedergren, *Journal of the American Chemical Society* **1987**, *109*, 7845–7854.
- [Viguier2001] R. Viguier, G. Serratrice, A. Dupraz, C. Dupuy, *European Journal of Inorganic Chemistry* **2001**, *2001*, 1789–1795.
- [Watson1953] J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 737–738.
- [Wei2013] Z. Wei, Y. Yang, Q. F. Li, F. Huang, H. H. Zuo, X. C. Su, *Chemistry - A European Journal* **2013**, *19*, 5758–5764.
- [Woehnert2003] J. Woehnert, K. J. Franz, M. Nitz, B. Imperiali, H. Schwalbe, *Journal of the American Chemical Society* **2003**, *125*, 13338–13339.
- [Woeltjen2005] E. Woeltjen, *Synthese eines paramagnetischen Tags zur NMR-spektroskopischen Untersuchung von Oligonukleotiden*, Diplomarbeit, Göttingen, **2005**.
- [Woeltjen2009] E. Woeltjen, *Paramagnetisch markierte Oligonukleotide*, PhD thesis, Goettingen, **2009**.
- [Wu1998] X. Wu, S. Pitsch, *Nucleic Acids Research* **1998**, *26*, 4315–4323.
- [Wunderlich2013] C. H. Wunderlich, R. G. Huber, R. Spitzer, K. R. Liedl, K. Kloiber, C. Kreutz, *Chemical Biology* **2013**, *8*, 2697–2706.
- [Yagi2010] H. Yagi, K. V. Loscha, X. C. Su, M. Stanton-Cook, T. Huber, G. Otting, *Journal of Biomolecular NMR* **2010**, *47*, 143–153.
- [Yamamoto2011] S. Yamamoto, T. Yamaguchi, M. Erdélyi, C. Griesinger, K. Kato, *Chemistry - A European Journal* **2011**, *17*, 9280–9282.
- [Yang2013] Y. Yang, Q. F. Li, C. Cao, F. Huang, X. C. Su, *Chemistry - A European Journal* **2013**, *19*, 1097–1103.
- [Zhang2006] Q. Zhang, X. Sun, E. D. Watt, H. M. Al-Hashimi, *Science* **2006**, *311*, 653–656.
- [Zhang2007] Q. Zhang, A. C. Stelzer, C. K. Fisher, H. M. Al-Hashimi, *Nature* **2007**, *450*, 1263–1267.

List of Figures

0.1	Established approach for the preparation of paramagnetically tagged DNA.	IX
0.2	Entwickelte Methode zur Herstellung von paramagnetisch markierter DNA.	XI
1.1	Cys-Ph-TAHA tag 1 .	3
1.2	EDTA-based tag 2 .	3
1.3	Second generation of EDTA-based tags with a flexible cysteine linker 3 or a more rigid phenyl ring modification 4 .	4
1.4	CLaNP-1 5 .	5
1.5	CLaNP-3 6 and CLaNP-5.1 7 .	5
1.6	CLaNP-7 8 and DOTA-M8 9 .	6
1.7	DPA 10 and 4MMDPA 11 .	7
1.8	IDA-based tag 12 and 4VPyMTA 13 .	7
1.9	Modified lactose with a EDTA-like lanthanide coordination site 14 .	8
1.10	Wild-type 24-mer DNA hairpin structure 15 . Thymine at position 3 (<i>red</i>) was replaced by a modified nucleobase.	9
1.11	Scheme of tagged DNA strand using a 1,3-cycloaddition to introduce a sulfur binding site for a disulfide bridge.	9
1.12	Elongations of TAR-RNA in the two different domains that induces alignment in Pfl phages presented by Zhang <i>et al.</i> . ¹³ C/ ¹⁵ N labeled nucleotides are shown in red.	10
1.13	Elongated (<i>blue</i>) TAR-RNA hairpin structure (<i>black</i>) with complementary paramagnetically tagged oligonucleotide (<i>red</i>) presented by E. Woeltjen.	11
1.14	Different modifications of nucleobases in DNA or RNA strands.	11
1.15	Introduction of a binding site using a <i>complementary addressed reagent</i> with 10 base pairs (only 2 are illustrated) presented by Babaylova <i>et al.</i> . At first, a guanine nucleobase of the target RNA is alkylated, followed by the cleavage of the modification from the reagent.	12
1.16	Sulfur binding sites in oligonucleotides resulting from the two pursued strategies.	13
2.1	Illustration of the magnetic field vector \vec{B} , the inter-nuclear vector \vec{R} and their connecting angle θ .	16

2.2	Molecular tumbling expressed in the molecular (<i>left</i>) or laboratory frame (<i>right</i>), resulting in a time-dependent movement of either the magnetic field vector $\vec{B}(t)$ or the inter-nuclear vector $\vec{R}(t)$, respectively. This figure was reproduced according to [Kramer2004].	18
2.3	Definition of an inter-nuclear vector \vec{R} in the magnetic susceptibility frame.	19
2.4	Illustration of pseudocontact shifts of an unpaired electron inducing either positive (<i>red</i>) or negative (<i>blue</i>) shifts in the oligonucleotide.	21
3.1	Synthesis of DNA by the phosphoramidite method using a succinyl linkage to a solid support. Shown are the 4 steps of the synthetic cycle (detritylation, coupling, capping, oxidation or sulfurization) and the final cleavage procedure.	24
3.2	Modified and phosphoramidite-activated nucleobases 17 and 18 providing a sulfur binding site in DNA.	26
3.3	Wild-type 24-mer DNA hairpin structure 15 . Thymine at position 3 (<i>red</i>) was replaced by a modified nucleobase.	26
3.4	The reaction pathway of the disulfide protection group strategy. Intermediate 26 can be activated directly to give phosphoramidite-activated compound 17	28
3.5	Transformation of protection groups at the sulfur moiety.	29
3.6	The reaction pathway to generate the phosphoramidite-activated compound 18	30
3.7	Alternative acetyl protection of thiol, resulting in a lower yield in the <i>Sonogashira</i> reaction.	31
3.8	Cleavage and tagging scheme for 24-mer DNA strand with modified nucleobase at position 3 and Lu- <i>preloaded</i> Cys-Ph-TAHA. Mass spectrum of the purified DNA shows the successful tagging reaction.	33
3.9	<i>Top</i> : Overview of synthesis and separation pathway for phosphorothioate RNA. <i>Bottom</i> : Sequence and HPLC chromatograms of two phosphorothioate RNA molecules modified at the labeled positions illustrate the successful separation procedure.	34
3.10	Elongated (<i>blue</i>) TAR RNA hairpin structure (<i>black</i>) with complementary paramagnetically tagged OPS (<i>red</i>).	35
3.11	Reaction scheme for new paramagnetic tags 36 and 37	36
3.12	<i>Top</i> : Tagging reaction of phosphorothioate DNA 38 and Cl-Bn-TAHA 37 . <i>Bottom</i> : HPLC chromatogram showing 3 DNA peaks and mass spectrum of DNA peak 2 illustrating successful tagging and a low turnover rate.	37
3.13	Cys-Ph-TAHA 1 and a shorter analogues MesS-Ph-TAHA 40	37

3.14	Scheme for modification of Cys-Ph-TAHA.	38
4.1	Established synthetic approach for the preparation of a paramagnetically tagged DNA strand, which is illustrated as a structure model (see section 4.2).	41
4.2	Folding of the DNA structure at low concentration leads to the expected chemical shifts for diamagnetically tagged DNA. <i>Left</i> : COSY spectrum of cytosine H5-H6 correlations folded at a high concentration with proposed assignment of a dimer structure. <i>Right</i> : After folding at a low concentration, the expected chemical shifts, which originate from the hairpin structure, were observed.	42
4.3	Different DNA models used for the structure calculation. <i>Left</i> : Overlay of the 5 NOE derived structures. <i>Right</i> : An idealized B-DNA model.	45
4.4	Overlay of the COSY <i>fingerprint</i> , showing PCSs in opposite directions for terbium (<i>red</i>) and thulium (<i>green</i>), compared to the reference sample (<i>blue</i>). The largest PCS measured for thulium (0.78 ppm for H5) is highlighted. The position of the modified nucleobase is labeled.	46
4.5	NOESY spectra overlay of the terbium (<i>red</i>), thulium (<i>green</i>) and lutetium sample (<i>blue</i>) illustrating the diagonal line that can be observed for the three corresponding signals.	47
4.6	NOESY spectra overlay of the thulium (<i>green</i>) and lutetium sample (<i>blue</i>) showing PCSs of the H1'-H2'(H2'') correlation. The highlighted resonances illustrate the assignment strategy based on the differences of the chemical shifts of H2' and H2''.	48
4.7	Distribution of the determined PCSs for the Tm (<i>green</i>) and Tb (<i>red</i>) sample, which is illustrated in the hairpin structure.	50
4.8	<i>Left</i> : Overlay of NOESY and COSY spectra of the diamagnetic sample to determine dipolar couplings of cytosine H5-H6 cross peaks. <i>Right</i> : Schematic of DISCO procedure showing the NOESY (A) and COSY (B) ω_2 -traces of C7 in order to extract RDCs by addition (A+B) and subtraction (A-B).	51
4.9	<i>Left</i> : Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Lu sample to determine 1/2 of the coupling. <i>Right</i> : In phase part of the HSQC IPAP spectrum with the highlighted ω_1 -trace of C8 that demonstrates the different peak intensities in the indirect dimension.	52

4.10	Overlay taken from ω_1 -traces of decoupled HSQC (<i>blue</i>) and HSQC IPAP (<i>red</i>) spectra of the Tm sample to determine the coupling constants of residues G9 (<i>top</i>) and C7 (<i>bottom</i>). Due to the asymmetric lineshape, the coupling constant of C7 was discarded in the structure calculations.	53
4.11	Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Lu sample to determine the coupling constants for the C6-H6 (C8-H8) resonances.	56
4.12	Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Lu sample to determine the coupling constants for the C1'-H1' resonances.	57
4.13	Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Tm sample to determine the coupling constants for the C6-H6 (C8-H8) resonances.	58
4.14	Overlay of decoupled HSQC and HSQC IPAP (A+B) spectra of the Tm sample to determine the coupling constants for the C1'-H1' resonances.	59
4.15	Two different Tm positions calculated from the PCS data alone or from the combined PCS and RDC data using the NOE model 1.	63
4.16	52 models were created by rotating the triple bond in steps of 10° for both of the disulfide bridge conformations.	64
4.17	Two-structure ensemble aligned along the Cys-Ph-TAHA tag.	65
4.18	Five-structure ensemble for the NOE model 1 resulting in the lowest Q -factor.	67
5.1	Scheme of a tagged DNA strand using a modified nucleobase and the Cys-Ph-TAHA tag, which was attached to the DNA hairpin molecule (A) and can be transferred to a DNA:RNA hybrid (B).	69

List of Abbreviations

4MMDPA	4-Merpactomethyl-dipicolinic acid
4VPyMTa	4-Vinyl(pyridine-2,6-diyl)bismethylenitrilo tetrakisacetid acid
Ac	Acetyl
Bu	Butyl
calc	calculated
CEM	2-Cyanoethoxymethyl
CLaNP	Caged lanthanide NMR probe
conc	concentrated
COSY	Correlation spectroscopy
Cys	Cysteine
DCM	Dichloromethane
DDTT	3-(Dimethylaminomethylidene)amino)-3H-1,2,4-dithiazole-3-thione
DIPEA	<i>N,N</i> -Diisopropylamine
DISCO	Differences and sums of traces within COSY spectra
DMAP	4-Dimethylaminopyridine
DME	Dimethoxyethane
DMSO	Dimethyl sulfoxide
DMT	4,4'-Dimethoxytrityl
DNA	Deoxyribonucleic acid
DOTA	1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetraacetic acid

DPA	Dipicolinic acid
dppb	1,4-Bis(diphenylphosphino)butane
DTPA	Diethylene triaminepentaacetic acid
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
EPR	Electron paramagnetic resonance
ESI	Electrospray ionization
Et	Ethyl
ex	excess
HPLC	High-performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
IDA	Iminodiacetic acid
INEPT	Insensitive nuclei enhanced by polarization transfer
IPAP	In phase - anti phase
Me	Methyl
Mes	Mesyl
MOM	Methoxymethyl
MOPS	3-(<i>N</i> -Morpholino)propanesulfonic acid
MS	Mass spectrometry
MWCO	Molecular weight cut off
nb	Nucleobases
NMM	<i>N</i> -Methylmorpholine
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
OPS	Oligonucleotide phosphorothioates

PCS	Pseudocontact shift
Ph	Phenyl
ppm	parts per million
PRE	Paramagnetic relaxation enhancement
RCSA	Residual chemical shift anisotropy
RDC	Residual dipolar coupling
RNA	Ribonucleic acid
STAT	Signal transducer and activator of transcription
TAHA	Triaminohexaacetic acid
TAR	Transactivation response
TBDMS	<i>t</i> -Butyldimethylsilyl
TCA	Trichloroacetic acid
TCEP	<i>tris</i> -(2-Carboxyethyl)phosphine
TEMPO	2,2,6,6-Tetramethyl-piperidine 1-oxyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TMS	Trimethylsilane
TOM	Tri- <i>i</i> -propylsilyloxymethyl
Tos	Tosyl

Appendix

Chemical shift tables

Table 6.1: Chemical shifts of Tm sample in ppm.

Residue	H1'	H2'	H2''	H41	H42	H5	H6/H8
C1	5.20	0.96	1.58			5.10	6.85
A2							
X3							
T4							
T5	5.94	2.01	2.45				
C6	5.77	2.09		6.58	8.04	5.51	7.42
C7	5.79	2.04	2.33	6.52	8.21	5.48	7.39
C8	5.46	2.00	2.31	6.71	8.42	5.56	7.35
G9	5.92	2.59	2.67				7.85
T10							
A15	6.08	2.68	2.79				8.30
C16	5.45	1.75	2.15	6.45	8.11	5.21	7.18
G17	5.32	2.45	2.51				7.64
G18	5.41	2.19	2.37				
G19							
A20	5.68	2.42	2.68				
A21	5.76	2.45	2.76				
A22	5.95	2.34	2.73				
T23	5.60	1.61	2.09				6.81
G24	5.97	2.24	2.49				7.71

Table 6.2: Chemical shifts of Tb sample in ppm.

Residue	H1'	H2'	H2''	H41	H42	H5	H6/H8
C1						7.07	8.64
A2							
X3							
T4							
T5	6.27	2.31					
C6	6.01	2.28	2.51	6.96	8.48	5.80	7.66
C7	5.96	2.17	2.45	6.74	8.46	5.66	7.55
C8	5.55	2.07	2.36	6.84	8.56	5.67	7.45
G9	5.94	2.60	2.67				7.89
A15	6.11		2.84				8.35
C16	5.51	1.84	2.24	6.53	8.20	5.30	7.25
G17	5.51	2.66	2.83				7.83
G18							
G19							
A20							
A21	6.05		2.98				
A22	6.26	2.57	2.97				
T23	5.96	1.86	2.35				7.08
G24	6.26	2.45	2.70				7.92

Table 6.3: Chemical shifts of diamagnetic sample in ppm.

Residue	H1'	H2'	H2''	H3'	H2	H41	H42	H5	H6/H8	H7
C1	5.70	1.69	2.23	4.57	X			5.86	7.55	X
A2	6.25	2.66	3.18	4.93	7.83	X	X	X	8.06	X
X3	6.12	2.23	2.42	4.88	X	X	X	X	7.86	X
T4	6.18	2.22	2.64	4.89	X	X	X	X	7.48	1.60
T5	6.06	2.12	2.55	4.86	X	X	X	X	7.35	1.60
C6	5.86	2.15	2.38	4.82	X	6.71	8.20	5.59	7.50	X
C7	5.86	2.08	2.38	4.81	X	6.59	8.30	5.55	7.45	X
C8	5.49	2.02	2.33	4.80	X	6.76	8.47	5.59	7.39	X
G9	5.93	2.59	2.67	4.93	X	X	X	X	7.87	X
T10	6.13	overlap	overlap	4.87	X	X	X	X	7.25	1.52
T11	overlap	overlap	overlap	4.79	X	X	X	X	7.52	1.80
C12	6.13	2.00	2.22	4.52	X			5.99	7.64	X
C13	5.76	1.82	2.33	4.56	X			5.74	7.35	X
T14	5.83	2.05	2.25	4.60	X	X	X	X	7.35	1.70
A15	6.09	2.70	2.81	4.92	7.94	X	X	X	8.32	X
C16	5.48	1.78	2.19		X	6.48	8.14	5.23	7.20	X
G17	5.39	2.53	2.62	4.90	X	X	X	X	7.71	X
G18	5.54	2.41	2.59	4.91	X	X	X	X	7.54	X
G19	5.23	2.42	2.54	4.90	X	X	X	X	7.57	X
A20	5.81	2.56	2.82	4.98	7.20	X	X	X	7.99	X
A21	5.87	2.55	2.85	4.99	7.14	X	X	X	7.98	X
A22	6.07	2.43	2.83	4.93	7.70	X	X	X	8.01	X
T23	5.73	1.70	2.19	4.77	X	X	X	X	6.92	1.33
G24	6.08	2.32	2.56	4.63	X	X	X	X	7.79	X

Coupling constants of diamagnetic sample.

Table 6.4: Aromatic coupling constants of the diamagnetic sample determined at 800 MHz spectrometer compared to nucleoside monophosphate [Ippel1996].

Residue	Coupling constant [Hz]	Difference [Hz]
C1C6-H6	183.9	-0.1
A2C8-H8	214.7	-1.3
T3C6-H6	X	X
T4C6-H6	185.7	0.7
T5C6-H6	181.5	-3.5
C6C6-H6	181.0	-3.0
C7C6-H6	179.9	-4.1
C8C6-H6	183.2	-0.8
G9C8-H8	214.1	-1.9
T10C6-H6	180.1	-4.9
T11C6-H6	181.2	-3.8
C12C6-H6	183.7	-0.3
C13C6-H6	182.5	-1.5
T14C6-H6	179.3	-5.7
A15C8-H8	216.2	0.2
C16C6-H6	185.5	1.5
G17C8-H8	212.9	-3.1
G18C8-H8	215.8	-0.2
G19C8-H8	211.5	-4.5
A20C8-H8	211.6	-4.4
A21C8-H8	215.9	-0.1
A22C8-H8	213.9	-2.1
T23C6-H6	184.2	-0.8
G24C8-H8	215.8	-0.2

Table 6.5: C1'-H1' coupling constants of the diamagnetic sample determined at 800 MHz spectrometer compared to nucleoside monophosphate [Ippel1996].

Residue	Coupling constant [Hz]	Difference [Hz]
C1C1'-H1'	169.9	0.4
C6C1'-H1'	165.9	-3.6
C7C1'-H1'	167.5	-2.0
C8C1'-H1'	167.4	-2.1
G9C1'-H1'	164.8	-1.2
T10C1'-H1'	166.4	-3.6
T11C1'-H1'	168.3	-1.7
C12C1'-H1'	168.5	-1.0
C13C1'-H1'	170.5	1.0
T14C1'-H1'	167.2	-2.8
A15C1'-H1'	165.6	-0.9
C16C1'-H1'	166.4	-3.1
G17C1'-H1'	166.3	0.3
G18C1'-H1'	163.0	-3.0
A22C1'-H1'	165.1	-1.4
T23C1'-H1'	164.5	-5.5

Results of PCS analysis

Table 6.6: Structure calculation of Tm PCS data. Differences between determined PCSs and back-calculated PCSs (PCScal) in ppm.

Structure	Residue	PCS	PCScal	Diff	Diff ²
NOE 1	C1H1'	-0.50	-0.47	0.03	0.00
NOE 1	C1H2'	-0.73	-0.76	-0.03	0.00
NOE 1	C1H2''	-0.65	-0.57	0.09	0.01
NOE 1	C1H5	-0.78	-0.68	0.10	0.01
NOE 1	C1H6	-0.71	-0.82	-0.11	0.01
NOE 1	C5H1'	-0.12	-0.09	0.03	0.00
NOE 1	C5H2'	-0.11	-0.09	0.02	0.00
NOE 1	C5H2''	-0.10	-0.08	0.02	0.00
NOE 1	C6H1'	-0.09	-0.07	0.02	0.00
NOE 1	C6H2'	-0.07	-0.06	0.01	0.00
NOE 1	C6H5	-0.10	-0.10	0.00	0.00
NOE 1	C6H6	-0.09	-0.08	0.01	0.00
NOE 1	C6H41	-0.14	-0.14	0.00	0.00
NOE 1	C6H42	-0.16	-0.14	0.01	0.00
NOE 1	C7H1'	-0.06	-0.04	0.02	0.00
NOE 1	C7H2'	-0.05	-0.04	0.01	0.00
NOE 1	C7H2''	-0.05	-0.03	0.02	0.00
NOE 1	C7H5	-0.06	-0.07	-0.01	0.00
NOE 1	C7H6	-0.06	-0.05	0.01	0.00
NOE 1	C7H41	-0.09	-0.09	-0.01	0.00
NOE 1	C7H42	-0.09	-0.10	-0.01	0.00
NOE 1	C8H1'	-0.03	-0.02	0.01	0.00
NOE 1	C8H2'	-0.03	-0.02	0.02	0.00
NOE 1	C8H2''	-0.02	-0.01	0.01	0.00
NOE 1	C8H5	-0.04	-0.04	0.00	0.00
NOE 1	C8H6	-0.04	-0.03	0.01	0.00
NOE 1	C8H41	-0.05	-0.06	-0.01	0.00
NOE 1	C8H42	-0.05	-0.06	-0.01	0.00
NOE 1	G9H1'	-0.01	0.00	0.00	0.00
NOE 1	G9H2'	-0.01	0.00	0.01	0.00
NOE 1	G9H2''	0.00	0.01	0.01	0.00
NOE 1	G9H8	-0.02	-0.01	0.01	0.00

NOE 1	A15H1'	-0.02	-0.03	-0.01	0.00
NOE 1	A15H2'	-0.02	-0.03	-0.01	0.00
NOE 1	A15H2''	-0.02	-0.04	-0.01	0.00
NOE 1	A15H8	-0.02	-0.03	-0.01	0.00
NOE 1	C16H1'	-0.03	-0.04	-0.02	0.00
NOE 1	C16H2'	-0.04	-0.05	-0.01	0.00
NOE 1	C16H2''	-0.04	-0.06	-0.02	0.00
NOE 1	C16H5	-0.03	-0.04	-0.01	0.00
NOE 1	C16H6	-0.03	-0.04	-0.01	0.00
NOE 1	C16H41	-0.04	-0.04	0.00	0.00
NOE 1	C16H42	-0.03	-0.04	0.00	0.00
NOE 1	G17H1'	-0.07	-0.07	0.00	0.00
NOE 1	G17H2'	-0.09	-0.09	0.00	0.00
NOE 1	G17H2''	-0.12	-0.11	0.01	0.00
NOE 1	G17H8	-0.07	-0.08	0.00	0.00
NOE 1	G18H1'	-0.13	-0.14	-0.01	0.00
NOE 1	G18H2'	-0.22	-0.20	0.02	0.00
NOE 1	G18H2''	-0.23	-0.24	-0.01	0.00
NOE 1	A20H1'	-0.13	-0.12	0.01	0.00
NOE 1	A20H2'	-0.15	-0.21	-0.06	0.00
NOE 1	A20H2''	-0.14	-0.14	0.00	0.00
NOE 1	A21H1'	-0.11	-0.10	0.01	0.00
NOE 1	A21H2'	-0.11	-0.10	0.00	0.00
NOE 1	A21H2''	-0.10	-0.09	0.00	0.00
NOE 1	A21H1'	-0.12	-0.11	0.01	0.00
NOE 1	A21H2'	-0.09	-0.09	0.01	0.00
NOE 1	A21H2''	-0.10	-0.09	0.01	0.00
NOE 1	T23H1'	-0.13	-0.12	0.01	0.00
NOE 1	T23H2'	-0.09	-0.10	0.00	0.00
NOE 1	T23H2''	-0.10	-0.10	0.00	0.00
NOE 1	T23H6	-0.10	-0.11	-0.01	0.00
NOE 1	G24H1'	-0.10	-0.15	-0.04	0.00
NOE 1	G24H2'	-0.08	-0.11	-0.03	0.00
NOE 1	G24H2''	-0.08	-0.12	-0.04	0.00
NOE 1	G24H8	-0.08	-0.12	-0.04	0.00
Structure	Residue	PCS	PCScal	Diff	Diff^2
B-DNA	C1H1'	-0.50	-0.39	0.11	0.01
B-DNA	C1H2'	-0.73	-0.77	-0.04	0.00

B-DNA	C1H2''	-0.65	-0.62	0.03	0.00
B-DNA	C1H5	-0.78	-0.73	0.05	0.00
B-DNA	C1H6	-0.71	-0.77	-0.06	0.00
B-DNA	C5H1'	-0.12	-0.11	0.02	0.00
B-DNA	C5H2'	-0.11	-0.11	0.00	0.00
B-DNA	C5H2''	-0.10	-0.09	0.01	0.00
B-DNA	C6H1'	-0.09	-0.06	0.03	0.00
B-DNA	C6H2'	-0.07	-0.07	0.01	0.00
B-DNA	C6H5	-0.10	-0.12	-0.01	0.00
B-DNA	C6H6	-0.09	-0.09	0.00	0.00
B-DNA	C6H41	-0.14	-0.16	-0.02	0.00
B-DNA	C6H42	-0.16	-0.16	-0.01	0.00
B-DNA	C7H1'	-0.06	-0.03	0.04	0.00
B-DNA	C7H2'	-0.05	-0.03	0.02	0.00
B-DNA	C7H2''	-0.05	-0.02	0.03	0.00
B-DNA	C7H5	-0.06	-0.07	-0.01	0.00
B-DNA	C7H6	-0.06	-0.05	0.01	0.00
B-DNA	C7H41	-0.09	-0.10	-0.01	0.00
B-DNA	C7H42	-0.09	-0.10	-0.01	0.00
B-DNA	C8H1'	-0.03	-0.01	0.03	0.00
B-DNA	C8H2'	-0.03	0.00	0.03	0.00
B-DNA	C8H2''	-0.02	0.00	0.02	0.00
B-DNA	C8H5	-0.04	-0.04	0.00	0.00
B-DNA	C8H6	-0.04	-0.02	0.02	0.00
B-DNA	C8H41	-0.05	-0.06	0.00	0.00
B-DNA	C8H42	-0.05	-0.06	0.00	0.00
B-DNA	G9H1'	-0.01	-0.01	0.00	0.00
B-DNA	G9H2'	-0.01	0.00	0.01	0.00
B-DNA	G9H2''	0.00	0.00	0.01	0.00
B-DNA	G9H8	-0.02	-0.01	0.01	0.00
B-DNA	A15H1'	-0.02	-0.03	-0.02	0.00
B-DNA	A15H2'	-0.02	-0.03	-0.01	0.00
B-DNA	A15H2''	-0.02	-0.04	-0.02	0.00
B-DNA	A15H8	-0.02	-0.03	-0.01	0.00
B-DNA	C16H1'	-0.03	-0.05	-0.03	0.00
B-DNA	C16H2'	-0.04	-0.05	-0.01	0.00
B-DNA	C16H2''	-0.04	-0.06	-0.02	0.00
B-DNA	C16H5	-0.03	-0.05	-0.01	0.00

B-DNA	C16H6	-0.03	-0.05	-0.02	0.00
B-DNA	C16H41	-0.04	-0.04	0.00	0.00
B-DNA	C16H42	-0.03	-0.04	-0.01	0.00
B-DNA	G17H1'	-0.07	-0.08	-0.01	0.00
B-DNA	G17H2'	-0.09	-0.09	0.00	0.00
B-DNA	G17H2''	-0.12	-0.11	0.01	0.00
B-DNA	G17H8	-0.07	-0.08	-0.01	0.00
B-DNA	G18H1'	-0.13	-0.13	-0.01	0.00
B-DNA	G18H2'	-0.22	-0.18	0.03	0.00
B-DNA	G18H2''	-0.23	-0.22	0.01	0.00
B-DNA	A20H1'	-0.13	-0.11	0.01	0.00
B-DNA	A20H2'	-0.15	-0.20	-0.05	0.00
B-DNA	A20H2''	-0.14	-0.12	0.02	0.00
B-DNA	A21H1'	-0.11	-0.12	-0.01	0.00
B-DNA	A21H2'	-0.11	-0.09	0.02	0.00
B-DNA	A21H2''	-0.10	-0.08	0.01	0.00
B-DNA	A21H1'	-0.12	-0.13	-0.01	0.00
B-DNA	A21H2'	-0.09	-0.10	0.00	0.00
B-DNA	A21H2''	-0.10	-0.10	0.00	0.00
B-DNA	T23H1'	-0.13	-0.13	0.00	0.00
B-DNA	T23H2'	-0.09	-0.10	0.00	0.00
B-DNA	T23H2''	-0.10	-0.10	0.01	0.00
B-DNA	T23H6	-0.10	-0.12	-0.02	0.00
B-DNA	G24H1'	-0.10	-0.13	-0.02	0.00
B-DNA	G24H2'	-0.08	-0.10	-0.02	0.00
B-DNA	G24H2''	-0.08	-0.09	-0.02	0.00
B-DNA	G24H8	-0.08	-0.12	-0.03	0.00

Table 6.7: PCS analysis of Tm sample, Euler angles and alignment tensors.

Structure	Euler angles in $^{\circ}$			Alignment tensors				
	α	β	γ	ΔA_{ax}	ΔA_{rh}	ΔA_{xx}	ΔA_{yy}	ΔA_{zz}
NOE 1	-82.92	149.90	-38.07	-3.60	0.39	1.52	5.69	-7.20
NOE 2	22.70	55.97	128.68	-6.08	0.22	4.05	8.11	-12.16
NOE 3	32.59	47.67	128.77	-5.44	0.07	4.87	6.01	-10.87
NOE 4	111.27	36.05	131.66	-3.78	0.30	2.09	5.47	-7.56
NOE 5	131.18	33.78	128.57	-3.45	0.58	0.45	6.44	-6.89
B-DNA	80.91	70.21	-65.08	-2.25	0.38	0.98	3.53	-4.51

Table 6.8: Structure calculation of Tb PCS data. Differences between determined PCSs and back-calculated PCSs (PCScal) in ppm.

Structure	Residue	PCS	PCScal	Diff	Diff ²
NOE 1	C1H5	1.21	1.20	0.00	0.00
NOE 1	C1H6	1.09	1.10	0.01	0.00
NOE 1	C5H1'	0.21	0.17	-0.04	0.00
NOE 1	C5H2'	0.18	0.15	-0.03	0.00
NOE 1	C6H1'	0.15	0.12	-0.03	0.00
NOE 1	C6H2'	0.12	0.11	-0.02	0.00
NOE 1	C6H2''	0.12	0.09	-0.03	0.00
NOE 1	C6H5	0.19	0.19	0.00	0.00
NOE 1	C6H6	0.16	0.15	-0.01	0.00
NOE 1	C6H41	0.24	0.27	0.03	0.00
NOE 1	C6H42	0.28	0.28	0.00	0.00
NOE 1	C7H1'	0.10	0.08	-0.03	0.00
NOE 1	C7H2'	0.08	0.07	-0.01	0.00
NOE 1	C7H2''	0.07	0.06	-0.01	0.00
NOE 1	C7H5	0.12	0.12	0.01	0.00
NOE 1	C7H6	0.10	0.10	0.00	0.00
NOE 1	C7H41	0.14	0.17	0.03	0.00
NOE 1	C7H42	0.16	0.18	0.02	0.00
NOE 1	C8H1'	0.06	0.04	-0.02	0.00
NOE 1	C8H2'	0.04	0.03	-0.01	0.00
NOE 1	C8H2''	0.04	0.02	-0.02	0.00
NOE 1	C8H5	0.07	0.08	0.01	0.00
NOE 1	C8H6	0.06	0.05	0.00	0.00
NOE 1	C8H41	0.07	0.11	0.04	0.00
NOE 1	C8H42	0.09	0.11	0.03	0.00
NOE 1	G9H1'	0.01	0.02	0.01	0.00
NOE 1	G9H2'	0.00	0.01	0.01	0.00
NOE 1	G9H2''	0.00	0.01	0.01	0.00
NOE 1	G9H8	0.02	0.02	0.00	0.00
NOE 1	A15H1'	0.02	0.04	0.02	0.00
NOE 1	A15H2''	0.02	0.05	0.03	0.00
NOE 1	A15H8	0.03	0.04	0.01	0.00
NOE 1	C16H1'	0.03	0.06	0.03	0.00
NOE 1	C16H2'	0.05	0.07	0.02	0.00

NOE 1	C16H2''	0.05	0.08	0.03	0.00
NOE 1	C16H5	0.06	0.06	0.01	0.00
NOE 1	C16H6	0.05	0.06	0.02	0.00
NOE 1	C16H41	0.05	0.06	0.01	0.00
NOE 1	C16H42	0.05	0.06	0.01	0.00
NOE 1	G17H1'	0.12	0.10	-0.02	0.00
NOE 1	G17H2'	0.13	0.12	-0.01	0.00
NOE 1	G17H2''	0.21	0.15	-0.06	0.00
NOE 1	G17H8	0.12	0.11	-0.01	0.00
NOE 1	A21H1'	0.18	0.19	0.01	0.00
NOE 1	A21H2''	0.13	0.16	0.03	0.00
NOE 1	A22H1'	0.20	0.19	-0.01	0.00
NOE 1	A22H2'	0.14	0.15	0.01	0.00
NOE 1	A22H2''	0.15	0.15	0.00	0.00
NOE 1	T23H1'	0.23	0.19	-0.04	0.00
NOE 1	T23H2'	0.16	0.14	-0.02	0.00
NOE 1	T23H2''	0.17	0.14	-0.02	0.00
NOE 1	T23H6	0.16	0.18	0.02	0.00
NOE 1	G24H1'	0.19	0.18	-0.01	0.00
NOE 1	G24H2'	0.14	0.13	0.00	0.00
NOE 1	G24H2''	0.14	0.13	-0.01	0.00
NOE 1	G24H8	0.13	0.17	0.04	0.00
Structure	Residue	PCS	PCScal	Diff	Diff^2
B-DNA	C1H5	1.21	1.20	0.00	0.00
B-DNA	C1H6	1.09	1.09	0.00	0.00
B-DNA	C5H1'	0.21	0.18	-0.02	0.00
B-DNA	C5H2'	0.18	0.17	-0.01	0.00
B-DNA	C6H1'	0.15	0.12	-0.03	0.00
B-DNA	C6H2'	0.12	0.12	-0.01	0.00
B-DNA	C6H2''	0.12	0.10	-0.02	0.00
B-DNA	C6H5	0.19	0.19	0.00	0.00
B-DNA	C6H6	0.16	0.15	0.00	0.00
B-DNA	C6H41	0.24	0.27	0.03	0.00
B-DNA	C6H42	0.28	0.27	-0.01	0.00
B-DNA	C7H1'	0.10	0.07	-0.03	0.00
B-DNA	C7H2'	0.08	0.07	-0.01	0.00
B-DNA	C7H2''	0.07	0.05	-0.01	0.00
B-DNA	C7H5	0.12	0.12	0.00	0.00

B-DNA	C7H6	0.10	0.09	-0.01	0.00
B-DNA	C7H41	0.14	0.16	0.02	0.00
B-DNA	C7H42	0.16	0.16	0.00	0.00
B-DNA	C8H1'	0.06	0.04	-0.02	0.00
B-DNA	C8H2'	0.04	0.03	-0.01	0.00
B-DNA	C8H2''	0.04	0.03	-0.01	0.00
B-DNA	C8H5	0.07	0.07	0.00	0.00
B-DNA	C8H6	0.06	0.05	0.00	0.00
B-DNA	C8H41	0.07	0.09	0.02	0.00
B-DNA	C8H42	0.09	0.10	0.01	0.00
B-DNA	G9H1'	0.01	0.03	0.02	0.00
B-DNA	G9H2'	0.00	0.02	0.02	0.00
B-DNA	G9H2''	0.00	0.02	0.02	0.00
B-DNA	G9H8	0.02	0.03	0.01	0.00
B-DNA	A15H1'	0.02	0.05	0.03	0.00
B-DNA	A15H2''	0.02	0.05	0.03	0.00
B-DNA	A15H8	0.03	0.04	0.02	0.00
B-DNA	C16H1'	0.03	0.07	0.03	0.00
B-DNA	C16H2'	0.05	0.07	0.02	0.00
B-DNA	C16H2''	0.05	0.08	0.03	0.00
B-DNA	C16H5	0.06	0.07	0.02	0.00
B-DNA	C16H6	0.05	0.07	0.02	0.00
B-DNA	C16H41	0.05	0.07	0.02	0.00
B-DNA	C16H42	0.05	0.07	0.02	0.00
B-DNA	G17H1'	0.12	0.11	-0.01	0.00
B-DNA	G17H2'	0.13	0.12	-0.01	0.00
B-DNA	G17H2''	0.21	0.14	-0.07	0.01
B-DNA	G17H8	0.12	0.11	-0.01	0.00
B-DNA	A21H1'	0.18	0.20	0.02	0.00
B-DNA	A21H2''	0.13	0.12	-0.01	0.00
B-DNA	A22H1'	0.20	0.22	0.02	0.00
B-DNA	A22H2'	0.14	0.15	0.01	0.00
B-DNA	A22H2''	0.15	0.15	0.01	0.00
B-DNA	T23H1'	0.23	0.21	-0.03	0.00
B-DNA	T23H2'	0.16	0.15	-0.01	0.00
B-DNA	T23H2''	0.17	0.15	-0.02	0.00
B-DNA	T23H6	0.16	0.19	0.03	0.00
B-DNA	G24H1'	0.19	0.17	-0.02	0.00

B-DNA	G24H2'	0.14	0.13	-0.01	0.00
B-DNA	G24H2''	0.14	0.12	-0.02	0.00
B-DNA	G24H8	0.13	0.17	0.03	0.00

Table 6.9: PCS analysis of Tb sample, Euler angles and alignment tensors.

Structure	Euler angles in °			Alignment tensors				
	α	β	γ	ΔA_{ax}	ΔA_{rh}	ΔA_{xx}	ΔA_{yy}	ΔA_{zz}
NOE 1	-130.27	152.12	-65.90	2.62	0.46	-0.80	-4.45	5.25
NOE 2	152.99	30.51	107.66	2.62	0.61	-0.21	-5.03	5.24
NOE 3	160.52	29.55	101.04	2.67	0.64	-0.12	-5.23	5.35
NOE 4	159.15	31.54	96.82	2.44	0.57	-0.35	-4.54	4.89
NOE 5	49.03	88.33	-6.27	-2.34	0.53	0.48	4.20	-4.68
B-DNA	85.06	83.43	-73.38	2.25	0.57	-0.34	-4.16	4.51

Results of RDC analysis

Table 6.10: Structure calculation of Tm RDC data. Differences between determined RDCs and back-calculated RDCs (RDCcal) in Hz.

Structure	Residue	RDC	RDCcal	Diff	Diff ²
NOE 1	C8C6-H6	-2.0	-3.3	-1.3	1.8
NOE 1	G9C8-H8	-4.4	-4.8	-0.4	0.1
NOE 1	T10C6-H6	-0.1	-1.0	-0.9	0.8
NOE 1	C16C6-H6	-4.1	-4.3	-0.2	0.0
NOE 1	G17C8-H8	-2.4	-4.9	-2.5	6.0
NOE 1	C1C1'-H1'	-1.3	-0.9	0.5	0.2
NOE 1	C7C1'-H1'	1.6	1.6	0.0	0.0
NOE 1	C8C1'-H1'	4.2	2.9	-1.3	1.8
NOE 1	G9C1'-H1'	-2.8	-0.6	2.2	4.9
NOE 1	T10C1'-H1'	-4.1	-4.5	-0.4	0.2
NOE 1	A15C1'-H1'	-0.2	0.1	0.3	0.1
NOE 1	C16C1'-H1'	-0.5	-0.1	0.4	0.1
NOE 1	G18C1'-H1'	-6.6	-4.6	2.0	4.0
NOE 1	A22C1'-H1'	-2.0	-1.0	1.0	0.9
NOE 1	T23C1'-H1'	-6.0	-4.7	1.3	1.7
NOE 2	C8C6-H6	-2.0	-3.6	-1.6	2.6
NOE 2	G9C8-H8	-4.4	-5.0	-0.6	0.3
NOE 2	T10C6-H6	-0.1	-1.5	-1.4	1.9
NOE 2	C16C6-H6	-4.1	-3.8	0.3	0.1
NOE 2	G17C8-H8	-2.4	-5.0	-2.6	6.6
NOE 2	C1C1'-H1'	-1.3	-0.9	0.4	0.1
NOE 2	C7C1'-H1'	1.6	1.1	-0.5	0.3
NOE 2	C8C1'-H1'	4.2	2.7	-1.5	2.4
NOE 2	G9C1'-H1'	-2.8	-0.6	2.2	5.0
NOE 2	T10C1'-H1'	-4.1	-2.3	1.8	3.1
NOE 2	A15C1'-H1'	-0.2	0.5	0.7	0.5
NOE 2	C16C1'-H1'	-0.5	1.0	1.5	2.1
NOE 2	G18C1'-H1'	-6.6	-4.9	1.7	2.9
NOE 2	A22C1'-H1'	-2.0	-0.8	1.2	1.6
NOE 2	T23C1'-H1'	-6.0	-5.2	0.8	0.7
NOE 3	C8C6-H6	-2.0	-3.6	-1.6	2.5
NOE 3	G9C8-H8	-4.4	-4.9	-0.5	0.2

NOE 3	T10C6-H6	-0.1	-1.5	-1.4	2.0
NOE 3	C16C6-H6	-4.1	-3.9	0.2	0.1
NOE 3	G17C8-H8	-2.4	-4.8	-2.4	5.9
NOE 3	C1C1'-H1'	-1.3	-0.9	0.4	0.1
NOE 3	C7C1'-H1'	1.6	0.8	-0.8	0.6
NOE 3	C8C1'-H1'	4.2	2.4	-1.8	3.2
NOE 3	G9C1'-H1'	-2.8	-0.3	2.5	6.1
NOE 3	T10C1'-H1'	-4.1	-2.5	1.6	2.6
NOE 3	A15C1'-H1'	-0.2	0.6	0.8	0.6
NOE 3	C16C1'-H1'	-0.5	1.1	1.6	2.5
NOE 3	G18C1'-H1'	-6.6	-4.9	1.7	2.8
NOE 3	A22C1'-H1'	-2.0	-0.7	1.3	1.6
NOE 3	T23C1'-H1'	-6.0	-5.3	0.7	0.5
NOE 4	C8C6-H6	-2.0	-3.6	-1.6	2.6
NOE 4	G9C8-H8	-4.4	-5.0	-0.6	0.3
NOE 4	T10C6-H6	-0.1	-1.0	-0.9	0.8
NOE 4	C16C6-H6	-4.1	-4.2	-0.1	0.0
NOE 4	G17C8-H8	-2.4	-5.0	-2.6	6.6
NOE 4	C1C1'-H1'	-1.3	-1.0	0.3	0.1
NOE 4	C7C1'-H1'	1.6	1.1	-0.5	0.3
NOE 4	C8C1'-H1'	4.2	2.5	-1.7	2.9
NOE 4	G9C1'-H1'	-2.8	-0.4	2.4	5.9
NOE 4	T10C1'-H1'	-4.1	-3.6	0.5	0.3
NOE 4	A15C1'-H1'	-0.2	0.5	0.7	0.4
NOE 4	C16C1'-H1'	-0.5	0.3	0.8	0.6
NOE 4	G18C1'-H1'	-6.6	-4.7	1.9	3.7
NOE 4	A22C1'-H1'	-2.0	-0.8	1.2	1.4
NOE 4	T23C1'-H1'	-6.0	-4.9	1.1	1.3
NOE 5	C8C6-H6	-2.0	-3.5	-1.5	2.1
NOE 5	G9C8-H8	-4.4	-5.1	-0.7	0.4
NOE 5	T10C6-H6	-0.1	-1.4	-1.3	1.8
NOE 5	C16C6-H6	-4.1	-4.0	0.1	0.0
NOE 5	G17C8-H8	-2.4	-4.9	-2.5	6.4
NOE 5	C1C1'-H1'	-1.3	-1.1	0.2	0.0
NOE 5	C7C1'-H1'	1.6	1.1	-0.5	0.2
NOE 5	C8C1'-H1'	4.2	2.6	-1.6	2.5
NOE 5	G9C1'-H1'	-2.8	-0.7	2.1	4.6
NOE 5	T10C1'-H1'	-4.1	-3.8	0.3	0.1

NOE 5	A15C1'-H1'	-0.2	0.1	0.3	0.1
NOE 5	C16C1'-H1'	-0.5	1.1	1.6	2.5
NOE 5	G18C1'-H1'	-6.6	-4.1	2.5	6.0
NOE 5	A22C1'-H1'	-2.0	-0.9	1.1	1.1
NOE 5	T23C1'-H1'	-6.0	-4.8	1.2	1.5
B-DNA	C8C6-H6	-2.0	-3.1	-1.1	1.1
B-DNA	G9C8-H8	-4.4	-4.9	-0.5	0.3
B-DNA	T10C6-H6	-0.1	-1.4	-1.3	1.8
B-DNA	C16C6-H6	-4.1	-3.7	0.4	0.2
B-DNA	G17C8-H8	-2.4	-4.9	-2.5	6.3
B-DNA	C1C1'-H1'	-1.3	-2.5	-1.2	1.5
B-DNA	C7C1'-H1'	1.6	2.2	0.6	0.4
B-DNA	C8C1'-H1'	4.2	2.9	-1.3	1.6
B-DNA	G9C1'-H1'	-2.8	-1.1	1.7	3.0
B-DNA	T10C1'-H1'	-4.1	-4.1	0.0	0.0
B-DNA	A15C1'-H1'	-0.2	0.8	1.0	0.9
B-DNA	C16C1'-H1'	-0.5	0.5	1.0	1.0
B-DNA	G18C1'-H1'	-6.6	-4.4	2.2	4.9
B-DNA	A22C1'-H1'	-2.0	-0.7	1.3	1.6
B-DNA	T23C1'-H1'	-6.0	-3.9	2.1	4.5

Table 6.11: RDC analysis of Tm sample, Euler angles and alignment tensors.

Structure	Euler angles in °			Alignment tensors				
	α	β	γ	ΔA_{ax}	ΔA_{rh}	ΔA_{xx}	ΔA_{yy}	ΔA_{zz}
NOE 1	-25.41	34.38	-100.01	-1.56	0.36	-0.72	2.40	-3.12
NOE 2	-10.27	37.81	-96.53	-1.38	0.64	0.05	2.71	-2.75
NOE 3	131.68	95.38	-13.26	1.42	0.54	-0.27	-2.58	2.84
NOE 4	-11.43	37.30	-98.29	-1.60	0.38	0.70	2.50	-3.19
NOE 5	-15.64	37.42	-90.80	-1.42	0.51	0.33	2.51	-2.85
B-DNA	-142.61	163.39	121.46	1.21	0.56	-0.19	-2.23	2.42

Results of combined RDC and PCS analysis

Table 6.12: Combined PCS-RDC analysis of Tm sample, Euler angles and alignment tensors.

Structure	Euler angles in $^{\circ}$			Alignment tensors				
	α	β	γ	ΔA_{ax}	ΔA_{rh}	ΔA_{xx}	ΔA_{yy}	ΔA_{zz}
NOE 1	161.77	90.55	133.81	1.52	0.52	-0.33	-2.70	3.03
NOE 2	-7.69	46.06	-98.66	-1.49	0.66	0.02	2.95	-2.98
NOE 3	6.74	133.09	80.90	-1.55	0.64	0.07	3.03	-3.10
NOE 4	0.65	57.74	54.54	-1.36	0.49	0.35	2.37	-2.73
NOE 5	-7.00	41.42	-88.85	-1.86	0.38	0.79	2.93	-3.72

HSQC IPAP

;15N and 13C=0 decoupling are not executed as long as d0 is too small

;G. Bodenhausen and D.J. Ruben, Chem. Phys. Lett. 69, 185 (1980)

;D.R. Muhandiram et al, JMR B102, 317-321 (1993)

;M. Ottinger et al, JMR, 373-378 (1998)

;p1 : power for 1H

;p12 : power for 13C hard

;p112 : power for 13C GARP decoupling

;p1 : 90 degree hard pulse 1H

;p3 : 90 degree hard pulse 13C

;p4 : 13C pulse, 225deg for 500/600, 180deg for 750/900

;pcpd2 : 90 deg cpd-pulse 13C (~85us)

;p20 : 1m (Gradient before first INEPT)

;p21 : 1m (Gradient in first INEPT)

;p22 : 800u (Gradient in first INEPT)

;p23 : 1m (Gradient for z-filter)

;p24 : 1m (Gradient for second INEPT)

;p25 : 500u (Gradient for IP AP conversion INEPT)

;gpz0 : 80%

;gpz1 : 19%

;gpz2 : 30%

;gpz3 : 65%

;gpz4 : 15%

;gpz5 : 17%

;d1 : relaxation delay

;d2 : 1H-13C INEPT delay (1.7m)

;in0 : 1/2*SW(in Hz)

#include <Avance_flos.incl>

define delay INEPT_1

```

define delay INEPT_2
define delay INEPT_3

#define GRADIENT0  10u p20:gp0 200u
#define GRADIENT1  10u p21:gp1 200u
#define GRADIENT2  10u p22:gp2 200u
#define GRADIENT3  10u p23:gp3 200u
#define GRADIENT4  10u p24:gp4 200u
#define GRADIENT5  10u p25:gp5 200u

"p2=p1*2"
"in0=inf1/2"
"d0=in0/2-p3*2/3.14159"
"INEPT_1=d2-(p21+210u)"
"INEPT_2=d2-(p24+210u)"
"INEPT_3=d2-(p25+210u)"
"l3=td1/4"

1 10u ze
;-----IP-part
18 40u
19 10u
20 1m
30 1m do:f2
    20u p19:f1
    10u LOCKH_OFF
    d1 cw:f1 ph29
    4u do:f1
    10u LOCKH_ON
    20u p11:f1
    20u p12:f2
    (p3 ph20):f2
    GRADIENT0
    10m
;-----first INEPT
    (p1 ph20):f1
    GRADIENT1

```

```

INEPT_1
(center (p2 ph20):f1 (p3 ph20 2u p4 ph21 2u p3 ph20):f2)
GRADIENT1
INEPT_1
(p1 ph2):f1
GRADIENT2
;-----13C evolution
(p3 ph1 d0 d0 p3 ph20):f2
GRADIENT3
(p1 ph20):f1
;-----second INEPT
GRADIENT4
INEPT_2
(center (p2 ph20):f1 (p3 ph20 2u p4 ph21 2u p3 ph20):f2)
GRADIENT4
INEPT_2 pl12:f2 LOCKH_OFF
;-----acquisition of IP-part
go=18 ph31 cpd2:f2
1m do:f2 wr #0 if #0 zd
;-----AP-part
40 goto 60
50 1m
60 50u
1m do:f2
20u pl9:f1
10u LOCKH_OFF
d1 cw:f1 ph29
4u do:f1
10u LOCKH_ON
20u pl1:f1
20u pl2:f2
(p3 ph20):f2
GRADIEN0
10m
;-----first INEPT
(p1 ph20):f1
GRADIENT1
INEPT_1

```

```
(center (p2 ph20):f1 (p3 ph20 2u p4 ph21 2u p3 ph20):f2)
GRADIENT1
INEPT_1
(p1 ph2):f1
GRADIENT2
;-----AP evolution INEPT
(p3 ph11):f2
GRADIENT5
INEPT_3
(center (p2 ph20):f1 (p3 ph12 2u p4 ph13 2u p3 ph12):f2)
GRADIENT5
INEPT_3
(p3 ph20):f1
;-----13C evolution
(d0 d0 p3 ph20):f2
GRADIENT3
(p1 ph20):f1
;-----second INEPT
GRADIENT4
INEPT_2
(center (p2 ph20):f1 (p3 ph20 2u p4 ph21 2u p3 ph20):f2)
GRADIENT4
INEPT_2 p112:f2 LOCKH_OFF
;-----acquisition of AP-part
go=50 ph31 cpd2:f2
1m do:f2 wr #0 if #0 zd
10u ip1
10u ip11
10u ip12
10u ip13
lo to 19 times 2
;-----for frequency discrimination
10u id0
lo to 20 times 13
;-----13=td1/2
10u LOCKH_OFF
exit
```

ph1=0 2
ph2=1 1 3 3
ph11=1 3
ph12=0
ph13=1
ph20=0
ph21=1
ph22=2
ph23=3
ph29=0
ph31=2 0 0 2

Curriculum Vitae

Personal data

Name Sebastian Täubert

Date of Birth 27.01.1985

Place of Birth Rinteln

Nationality German

Education

1991-1995 Primary school Domschule in Minden

1995-2004 Secondary school Besselgymnasium in Minden

2004-2005 Community Service at the children's home Salem-Köslin in Minden

2005-2010 Chemistry Studies at the Georg-August University of Göttingen

2010 Diploma thesis at the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany. Subject: *Paramagnetic Tagging of DNA and Structure Determination using NMR-spectroscopy*, supervisor: Prof. Dr. Christian Griesinger.

2010-2014 Ph.D. student at the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany. Subject: *Paramagnetic Tagging of Oligonucleotides for Structure Determination using NMR-spectroscopy*, supervisor: Prof. Dr. Christian Griesinger.

Göttingen, November 2014