2019•2020

Faculteit Industriële ingenieurswetenschappen

master in de industriële wetenschappen: chemie

Masterthesis

Base pairing analysis of nucleobase functionalized monomers

PROMOTOR :

Prof. dr. ir. Leen THOMASSEN

PROMOTOR :

Prof. dr. Tanja JUNKERS

BEGELEIDER :

De heer Lowie MAES

Pieterjan Broekx

Scriptie ingediend tot het behalen van de graad van master in de industriële wetenschappen: chemie

Gezamenlijke opleiding UHasselt en KU Leuven

2019•2020

Faculteit Industriële ingenieurswetenschappen

master in de industriële wetenschappen: chemie

Masterthesis

Base pairing analysis of nucleobase functionalized monomers

PROMOTOR :

Prof. dr. ir. Leen THOMASSEN

PROMOTOR :

Prof. dr. Tanja JUNKERS

BEGELEIDER :

De heer Lowie MAES

Pieterjan Broekx

Scriptie ingediend tot het behalen van de graad van master in de industriële wetenschappen: chemie

UHASSELT KU LEUVEN

Deze masterproef werd geschreven tijdens de COVID-19 crisis in 2020. Deze wereldwijde gezondheidscrisis heeft mogelijk een impact gehad op de opdracht, de onderzoekshandelingen en de onderzoeksresultaten.

Foreword and acknowledgements

Throughout my education in chemical engineering at Hasselt University, I learned a lot about the different branches within the chemical industry. I noticed quickly that my main interest lays in the organic chemistry and the field of R&D. Therefore, I was eager and motivated to get started on my master thesis as soon as I learned what my project entailed. I am grateful to have had the opportunity to execute this thesis within the Polymer Reaction Design Group (PRD).

First, I would like to express my gratitude towards those who made an effort in helping me during/with this project in the difficult times caused by the pandemic of the COVID-19 virus.

I would like to thank Prof. Dr. T. Junkers for giving me the opportunity to execute my master thesis within her research group and for proofreading my paper. I would also like to thank Prof. Dr. L. Thomassen for her guidance throughout my education, for providing me with this opportunity and for proofreading my master thesis.

Most importantly, special thanks go out to my supervisor Drs. Ing. L. Maes. His enthusiasm and tremendous guidance from the very beginning of this project caused me to be even more motivated. He went out of his way to teach me everything I needed to know from executing certain procedures, to the theoretical background of the project. He spent many hours proofreading my thesis, discussing data with me, answering my questions, synthesizing monomers, and providing me with NMR-analysis data. He has been a great mentor. As his own doctoral research reaches its final stages I wish him all the best in finishing his research and in his future career.

Furthermore, I would like to thank the members of the PRD and DSOS research groups for helping me in the lab whenever I could not find certain products or was experiencing troubles with certain procedures. Their never-ending help and kindness made me feel welcomed and caused every day in the lab to be enjoyable. Also, special thanks go out to Koen Van Vinckenroye for measuring the multitude of NMR-samples, during the lockdown, that I needed in this research.

Last but not least I would like to thank my family for their unconditional support throughout my entire education and during this final project. They helped me getting through difficult times and where always there for me in every step of the way. Finally, I would also like to thank my girlfriend. She did everything within her capabilities to get me through the final stages of my education despite everything she had going on herself.

Without everyone mentioned above, this thesis would not have been possible.

Table of Contents

List of tables

List of figures

Abbreviations

^{1}H-NMR	Proton nuclear magnetic resonance
A	Adenine
AAM	Adenine containing acrylate monomer
ACP	2-amino-6-chloropurine
AIBN	α,α-azoisobutyronitrile
ATRP	Atom transfer radical polymerization
BDDA	1,4-butanediol diacrylate
BDT	Benzene-1,4-dithiol
BEB	(1-bromoethyl)benzene
BHT	2,6-Di-tert-butyl-4-methylphenol
BMPH	N-[ß-maleimidopropionic acid] hydrazide
C	Cytosine
CAM	Cytosine containing acrylate monomer
CRP	Controlled radical polymerization
CS_2	Carbon disulfide
CTAW	2-(2-Carboxyethylsulfanylthiocarbonylsulfanyl)propionic acid
CTAW2	4-((((2-Carboxyethyl)thio)carbonothioyl)thio)-4-cyanopentanoic acid
DCT	1-dodecanethiol
DOSY	Diffusion-ordered spectroscopy
DPE-TTC	2-dodecyl-1-phenylethyl trithiocarbonate
EDC	N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide
EMCA	N-ε-maleimidocaproic acid
EWF	Evanescent wave field
FRP	Free radical polymerization
FT	Fourier transform
G	Guanine
GAM	Guanine containing acrylate monomer
HCOOH	Formic acid
K_2CO_3	Potassium carbonate
K_3PO_4	Potassium phosphate tribasic
KtBuO	Potassium tert-butoxide
LRP	Living radical polymerization
$MgSO_4$	Magnesium sulphate
MHB	Multiple hydrogen bond
NAAM	Nucleobase containing acrylamide monomers
NAM	Nucleobase containing acrylate monomer
PABH	Poly(3-aminobenzaldehyde)
PSB	Polystyrene beads

RAFT	Reversible addition-fragmentation chain transfer
RDRP	Reversible deactivation radical polymerization
RF	Radio frequency
SD	Sequence-defined
SPR	Surface plasmon resonance
SUMI	Single unit monomer insertion
T	Thymine
TAM	Thymine containing acrylate monomer
TEA	Tris(2-aminoethyl)amine
TMS	Trimethylsilane
U	Uracil
UAM	Uracil containing acrylate monomer
WC	Watson-Crick

Abstract

Nucleobases are key components in the storage of genetic information within the biopolymer called DNA. Besides their function in DNA, they have potential applications in synthetic polymer chemistry. A controlled radical polymerization technique enables the sequence-defined implementation of these nucleobases into synthetic polymers, which mimics the structure of DNA. Unfortunately, the low yield of this technique limits the length of the obtainable polymers, preventing them from being used in biomedical applications. Potential pathways towards increasing the yield rely on selective base pairing between the nucleobases, which is a phenomenon that still needs verification and characterization.

The objective of this research is the characterization of hydrogen bonding between the nucleobases implemented into polymerizable monomers. Two proposed analysis techniques, based on nuclear magnetic resonance, have been used to characterize this base pairing. The analysis revealed the presence of a complex stoichiometry and the influences of temperature, solvent polarity and concentration ratios on base pairing between the monomers.

Even though, this thesis revealed important effects and phenomena in base pairing, research confirming these conclusions is still required. Nevertheless, the results imply that the previously mentioned parameters could be optimized to obtain the desired selective base pairing between the nucleobase containing monomers. Therefore, the two potential pathways might indeed resolve the low yield problem.

Abstract in het Nederlands

Nucleobasen zijn essentieel voor de opslag van genetische informatie in DNA. Boven op hun functie in DNA zijn ze mogelijks bruikbaar in toepassingen in de synthetische polymeerchemie. Een gecontroleerde radicalaire polymerisatietechniek maakt sequentie-gedefinieerde implementatie van deze nucleobasen in synthetische polymeren mogelijk. Het lage rendement van deze techniek limiteert echter de ketenlengte en verhinderd daardoor het gebruik van deze polymeren in biomedische applicaties. Er zijn twee technieken die mogelijk een oplossing kunnen bieden. Helaas zijn deze gebaseerd op selectieve basen paring tussen nucleobasen, een fenomeen dat tot op heden nog onvoldoende gekarakteriseerd is.

Het doel van deze thesis is het karakteriseren van basen paring tussen de nucleobasen, geïmplementeerd in polymeriseerbare monomeren. Twee voorgestelde technieken, gebaseerd op nucleaire magnetisch resonantie, zijn gebruikt om dit te realiseren. De aanwezigheid van een complexe stoichiometrie van basen paring alsook de invloed van temperatuur, solvent polariteit en concentratie-ratio op basen paring werd met deze technieken bepaald.

Deze thesis onthuld belangrijke verschijnselen die plaats vinden in basen paring. Verder onderzoek ter verificatie van de bevindingen en ter karakterisatie van eerder vernoemde invloeden is echter nodig. Desondanks wordt hier aangetoond dat parameteroptimalisatie kan leiden tot selectieve basen paring, waardoor de potentiële oplossingen voor het lage rendementsprobleem realiseerbaar zouden kunnen zijn.

Chapter 1. Introduction

This study was executed to obtain the degree ‘Master of Industrial Engineering: Chemistry’ and was performed at the Polymer Reaction Design group. The research group is located at Hasselt University and Monash University and is led by Prof. Dr. T. Junkers. Their focus lies on obtaining specialized materials consisting of functional polymers obtained by examination of tailor-made synthesis procedures [1]. The researchers employ polymer conjugation techniques or controlled radical polymerization procedures to synthesize polymers with sequential and/or structural precision. Therefore, new methods for gaining control over the polymerization processes are developed. These processes are examined with kinetic studies to obtain insight in the reactions.

1.1 Context

1.1.1 Nucleobases

DNA and RNA are known as one of the most important biomolecules in living organisms. Since the isolation of DNA in 1869 an extensive amount of research has been done to reveal its structure and functions in organisms [2, 3]. Even though there is still a lot to uncover about these biomolecules, their structure is almost entirely known. DNA and RNA are similar biopolymers, both consisting of polynucleotide chains. These polynucleotides are biopolymers with nucleotides as their monomers, consisting of a cyclic pentose sugar, a nitrogenous base and a phosphate group [3]. Despite the similarities between DNA and RNA, there are also significant differences. Firstly, DNA consists of two polynucleotide chains while RNA consists of merely one. Additionally, DNA contains deoxyribonucleic acids in comparison to the ribonucleic acids from RNA. Finally, the nucleobase called ‘thymine’ is present in DNA, where ‘uracil’ is present in RNA. A schematic representation of the structure of DNA is given in Figure 1. The nitrogenous bases, also called nucleobases, execute important functions in DNA and RNA. There are 5 nucleobases called adenine (A), thymine (T), uracil (U), cytosine (C) and guanine (G). The colour of the nucleobases, as shown in the legend of Figure 1, also applies to the other figures in which these nucleobases are shown. The respective sequence of these nucleobases in the polynucleotides enable storage of the genetic information of the living organism, in which these biopolymers are present. Furthermore, selective hydrogen bonding or base pairing between the nucleotides from opposing polynucleotide strings causes two complementary strings to stay together in a double helix formation [4]. Selective hydrogen bonding is hydrogen bonding of a nucleobase with its complementary nucleobase. This results in the formation of base pairs (e.g. A-T, A-U and C-G). It is this capability of selective (reversible) hydrogen bonding and their biocompatibility that enables these nucleobases to not only find applications in nature (e.g. DNA and RNA), but also in the synthetic polymer

chemistry (e.g. hydrogels [5], self-healing materials [6, 7], DNA-delivery [8], adhesives [9], chemical data storage [10], etc.).

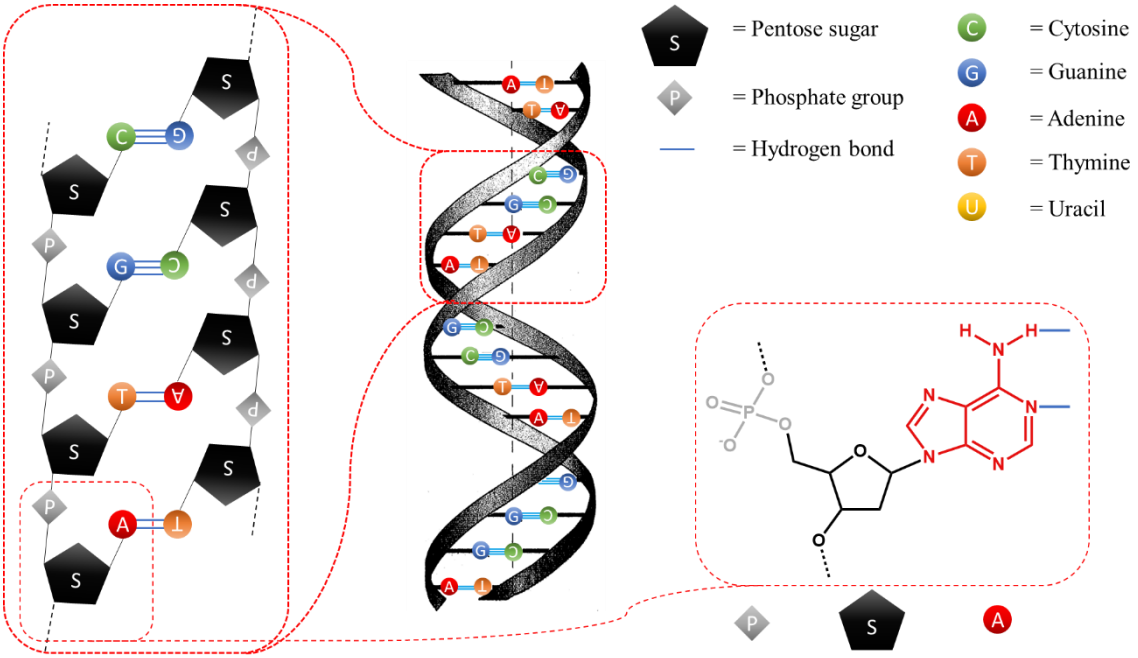


Figure 1. Molecular structure of DNA. The chemical structure of DNA (left) shows the molecular composition of a DNA double helix. The stylized diagram (right) shows the double helical structure of DNA. Hydrogen bonds between nucleobases are represented by the blue lines connecting the nucleobases. Note that no uracil nucleobases are present in the DNA structure, as uracil does not occur in DNA.

One of the methods to obtain such synthetic nucleobase containing polymers is radical polymerization. This method results in synthetic nucleobase containing polymers with a carbon-carbon backbone. A carbon-carbon backbone has certain advantages in comparison to the phosphate-sugar backbone of DNA. Some of the benefits are increased stability/inertia of the polymer backbone, possibility of implementation into classical polymer products and material development for potential medical applications. For certain applications (e.g. interaction with other biomolecules, data storage and recognition of specific DNA- or RNA-sequences) having a random sequence of nucleotides within such polymers is insufficient. Nucleobase containing monomers must be implemented, into the polymer, in a defined sequence. Therefore, a polymerization method must be found that enables the synthesis of such sequence-defined (SD) polymers.

1.1.2 Sequence-defined oligomers via controlled radical polymerization

One of the most frequently used polymerization methods in the industry is radical polymerization. This chain-growth polymerization method can be divided into two categories: free radical polymerization (FRP) and reversible-deactivation radical polymerization (RDRP). FRP is commonly used in the industry, but due to lacking control over the polymerization, FRP cannot be used for the synthesis of polymers with high structural control. Among polymers with a highly controlled structure, there are two classes: sequence-controlled and sequence-defined polymers. Sequence-controlled polymers are block copolymers. Depending on the control over

the polymerization, a limited polydispersity within these blocks might still be present, resulting in non-identical polymers ($1.5 > Đ > 1$). In SD polymers on the other hand, no polydispersity is observed ($Đ = 1$) and the exact sequence of the monomers implemented into the polymers is known and uniform for each synthesized polymer. This means that any possible sequence of monomers can be implemented into the polymer with precision. Unfortunately, the fast non-simultaneously propagation and unpredictable termination occurring in FRP, results in little control over the produced chain lengths and causes for a broad dispersity to occur in its produced polymers. Polymerization by means of RDRP, in contrary to FRP, does offer the necessary control for producing these desired SD polymers. RDRP-techniques have already been used to synthesize SD polymers such as atom transfer radical polymerization (ATRP) and reversible addition-fragmentation chain transfer polymerization (RAFT) [11, 12]. RAFT-polymerization shows many advantages over the other RDRP techniques for synthesizing the desired SD polymers. The mechanism of this technique followed by a detailed explanation is described in paragraph 2.1.

RAFT polymerization has already been used to synthesize sequence defined nucleobase containing polymers in recent research. Here, a sequence-defined tetramer containing adenine, cytosine, thymine and guanine was obtained [12]. The monomers used in this recent research employed acrylate derivatives of A (AAM), C (CAM), T (TAM), G (GAM) and U (UAM). Derivatizing the nucleobases is necessary as the bare molecules cannot be polymerized by means of RAFT polymerization. The resulting oligomer, after polymerization of these nucleobase containing acrylate monomers (NAM), is called a multiple hydrogen bond sequence-defined oligomer (MHB-SD oligomer).

1.2 Problem statement

An MHB-SD oligomer can be synthesized via the single unit monomer insertion approach (SUMI). The SUMI procedure consists of a combination of multiple RAFT polymerizations and subsequent intermediate purification steps resulting in only one monomer insertion in each cycle. Each polymerization results in a mixture of oligomers with a narrow molecular weight distribution. Thus, oligomers are obtained in which zero, one or multiple monomers are inserted as shown in in Figure 2. Therefore, the intermediate purification is required. Flash chromatography is currently used to purify these reaction mixtures, enabling isolation of the desired oligomers containing only one insertion. Unfortunately, the SUMI-cycle of polymerization and purification results in a milligram scale yield of the desired product, starting from grams of RAFT-agent [12]. This yield is obtained after four SUMI-cycles resulting in an MHB-SD tetramer. As each insertion has a yield between 10%-30%, increasing the amount of insertions means further decreasing the overall yield of the desired MHB-SD. This yield problem therefore limits the length of producible polymers. A consequence of the limited producible chain length is that these MHB-SD cannot be used in for example interaction with biopolymers. The oligomers should contain more nucleobases for them to be applicable in such bio-applications. To obtain a sufficient length of MHB-SD oligomers for bio-applications, a pentamer for example, the yield of the polymerization process must increase significantly.

Possible methods to increase the yield are template-assisted polymerization and purification by means of affinity separation. Both methods share the same start. They both require an MHB-SD oligomer, containing the complementary sequence of nucleobases relative to the desired compound. This MHB-SD is formed with the existing SUMI-process as described in Figure 2 [12]. In a following step these sequences could be placed on a substrate or particle.

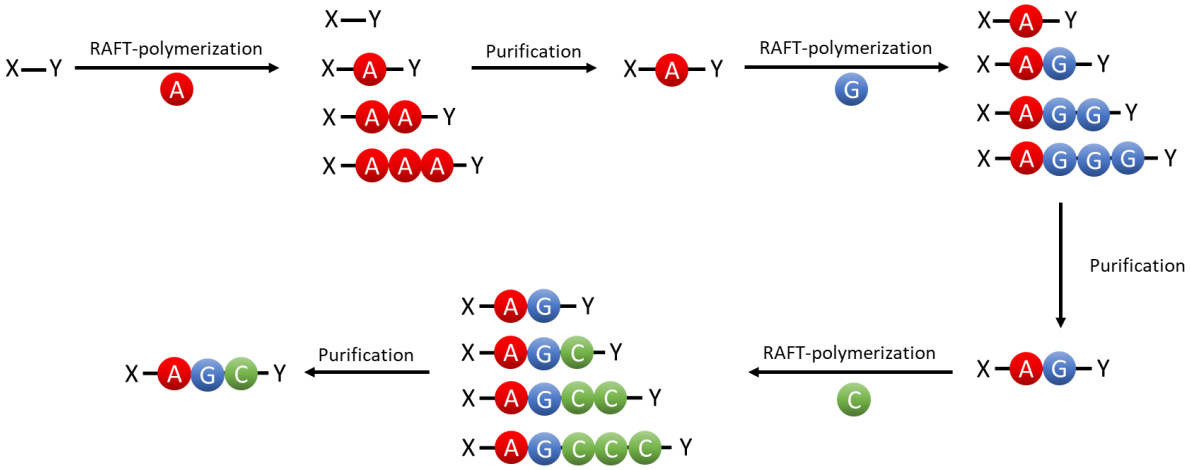


Figure 2. Existing SUMI-polymerization process. X-Y represents the end-groups of the RAFT-agent.

A first pathway that can be followed is template-assisted polymerization [13-15]. This technique has already been used in templated Sonogashira coupling of an adenine containing monomer and in radical polymerizations such as ATRP [16]. The principle of template-assisted polymerization might also be applied to MHB-SD oligomers as depicted in Figure 3. Here, the nucleobase containing monomers, that form the desired oligomer, can be mixed with the complementary MHB-SD attached onto a particle or substrate. Due to selective hydrogen bonding, the free nucleobases should bind to their complementary nucleobase within the template under the right circumstances. After binding, polymerization could be used to covalently link the hydrogen bonded monomers, which are now close to each other. Note that the solvent, temperature, pH and concentration(-ratios) have a significant influence on the hydrogen bonding and the polymerization [9, 17-19]. In a following step, side products should be removed by washing. Afterwards, the hydrogen bonds between the newly formed desired oligomer and its complementary sequence should be broken. Finally, particles or substrates and side products will be removed and a pure MHB-SD oligomer is isolated. The principle of bringing monomers close to each other before polymerization to obtain a sequence-defined polymer or oligomer has already been used in other methods. An example is DNA-templated synthesis [20]. Short DNA adapter strands, with a monomer attached to it, will bring the monomers in close proximity to the polymer chain. This is a result of the fact that the adapter strand containing the monomer will bind by means of hydrogen bonding to the adapter strand of the polymer. What follows is a click reaction coupling the monomer to the polymer chain.

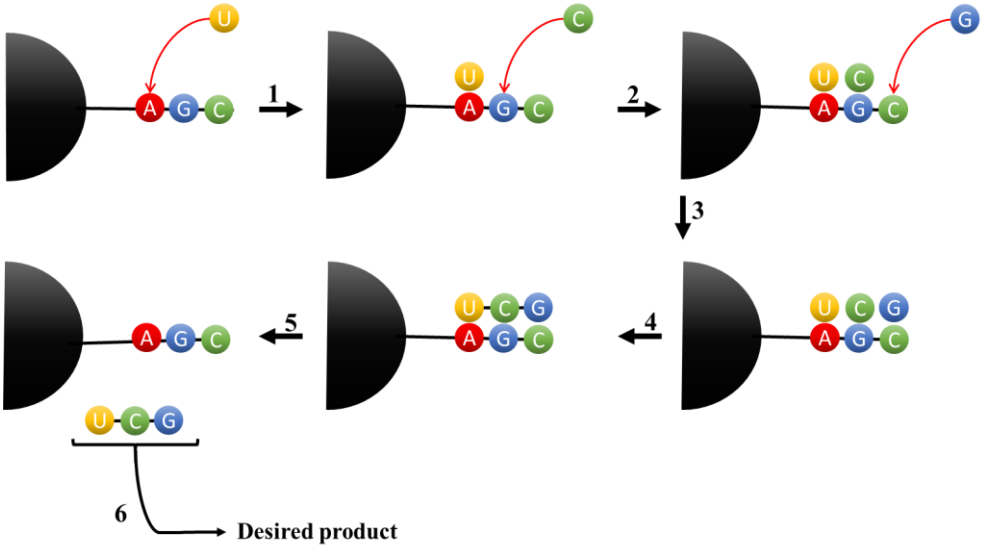


Figure 3. Schematic representation of template-assisted RAFT polymerization. In step 1 to 3 the monomers are attached to the templated by selective hydrogen bonding (red arrows). In step 4 RAFT polymerization causes these monomers to bond covalently to each other. In step 5 the desired MHB-SD is detached from the template and in step 6 the desired product is extracted.

A second pathway that can be followed, shown in Figure 4, is purification by means of affinity separation. This method resembles the existing SUMI-cycle, except for the fact that purification by means of flash column separation between every SUMI-cycle is not executed. Here, a particle or substrate containing the complementary sequence of the desired final compound ensures purification by forming hydrogen bonds with the desired compound at the end of the polymerization. The particle, with a complementary nucleobase sequence to the final desired product, is mixed with the reaction mixture. It will bind the desired product by means of selective hydrogen bonding. When the particles are separated from the mixture, the final desired product should be extracted as well. After breaking of the hydrogen bonds, pure product should be obtained. Note that a setup as used in affinity chromatography could also be used.

The yield obtained by this process is likely to show an increase in comparison to the yield obtained by the existing SUMI-process. It could be a result of eliminated intermediate flash chromatography. These purifications cause a decrease in yield in the current SUMI-process due to loss of product in the mix fractions after flash chromatography. Furthermore, the same principle can be implemented using a substrate or affinity chromatography.

As the two potential methods have been elucidated above, a short preliminary comparison of both techniques will be given. The assumption could be made that, if both methods work perfectly, template-assisted polymerization would render higher yields in the synthesis of the desired oligomers. In the affinity separation technique, presence of a polydispersity after each polymerization causes for a certain amount of produced polymers to consist of a deviating sequence from the desired MHB-SD. The deviation would decrease the yield in comparison to the template-assisted technique, as in this second technique no polydispersity would be present. This comparison between the techniques is under the assumption that both techniques work exactly as described in theory.

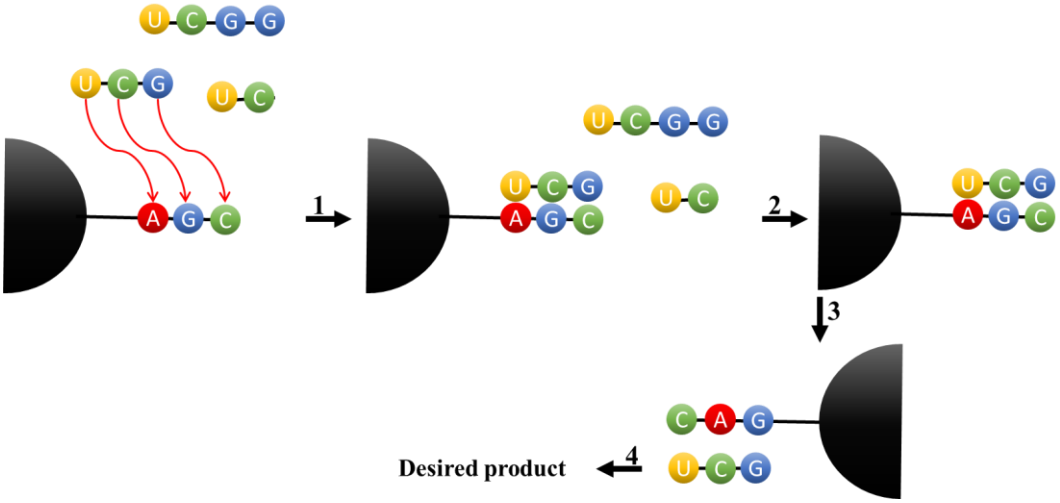


Figure 4. Schematic representation of affinity separation. In step 1 the desired product, obtained via multiple subsequent RAFT polymerizations, is hydrogen bonded (red arrows) to the complementary sequence of the desired product. In step 2 the residual sequences are removed from the mixture. In step 3 the desired product is detached from the template and in step 4 the desired product is extracted from the template mixture.

A prior requirement for both solutions to work is the fact that hydrogen bonding between the nucleobases must be selective. This means that a nucleobase should only form hydrogen bonds with its complement. Frequent formation of mismatches will render the previously mentioned solutions ineffective. Therefore, the selectivity of the hydrogen bonding between nucleobases in MHB-SD oligomers has to be examined. Only after reaching sufficient selectivity, the template-based polymerization or affinity purification can be studied. Additionally, an environment causing for breaking of the hydrogen bonds should be examined as well. It is important as the desired MHB-SD will have to be detached from the template at the end of the process.

This thesis focusses on analysis of the selectivity of hydrogen bonding between nucleobases implemented in NAM. Several analysis techniques will be handled to give an overview on how selectivity between nucleobases, implemented in MHB-SD, can be examined and characterized. Furthermore, influences of environmental parameters (e.g. solvents, temperature, etc.) on selectivity will be examined.

1.3 Objectives

1.3.1 Selectivity

The main objective is examining the chemical affinity and selectivity of the hydrogen bonding or base pairing between the complementary monomers and sequences. This objective consists of two sub-objectives.

A first sub-objective is finding methods which unambiguously detect hydrogen bonds between the (complementary) nucleobases and provides the characteristic affinity of these components. These methods preferably give information about the selectivity of the formed hydrogen bonds. This is important to detect formation of mismatches.

A second sub-objective is finding the optimal conditions for formation and breaking of selective hydrogen bonds. Conditions promoting selective hydrogen bonding are necessary for attachment of NAM or MHB-SD to their complement. The optimal environment is the environment in which the maximum amount of hydrogen bonds between complementary nucleobases take place, while maintaining a minimal percentage of mismatches. Secondly, optimal environment for breaking of the hydrogen bonds must be found to separate the desired MHB-SD oligomer from its complementary sequence. This is the environment in which weak or no hydrogen bonding between MHB-SD occurs.

1.4 Methodology

To determine selectivity of the MHB-SD there are several analysis methods:

1. Proton nuclear magnetic resonance (^{1}H-NMR)
2. 1D Nuclear Overhauser effect analysis (NOE-analysis)
3. 2D Nuclear Overhauser effect analysis (NOESY-analysis)
4. Diffusion-ordered spectroscopy NMR (DOSY-NMR)
5. Surface plasmon resonance (SPR)
6. Affinity extraction
7. Affinity chromatography

These methods can be used to verify whether hydrogen bonding between NAM and between MHB-SD takes place. They can also be used to give an indication whether this hydrogen bonding is selective. In this research, the first two methods will be executed and evaluated.

1.5 Outline of the dissertation

This dissertation will commence with an elaborate literature study in chapter two. The discussed topics are related to synthesis of nucleobase containing oligomers and interaction between these synthetic nucleobase containing structures. Firstly, the choice for RAFT-polymerization, as technique for synthesizing MHB-SD, is explained. Afterwards, the mechanism of this technique is elucidated followed by a section in which the potential and selected RAFT-agents and monomer structures are discussed. Subsequently, the theory behind the procedure for the synthesis of MHB-SD is given. A following sub-chapter commences with theory behind the formation and strength of hydrogen bonding. This part is followed by several sub-sections in which the different models of base pairing between complementary and non-complementary nucleobases is elaborated. These models are important as they give an insight in potential problems with the selectivity of base pairing. In the next sub-chapter, parameters such as temperature, solvent polarity and pH, affecting the strength of hydrogen bonds and base pair selectivity will be discussed. Finally, the literature study ends with the theory behind several analysis methods for examining the formation of hydrogen bonding and determination of selectivity in base pairing between NAM and MHB-SD. These methods will also be linked directly to the specific case of interaction between MHB-SD and/or NAM.

Chapter three contains laboratory results. Here, the results of synthesis of NAM will be given. Also, the procedures that were used to analyse the interaction between the NAM will be described in this section.

In the fourth chapter, all results will be listed and discussed in detail. Firstly, the stoichiometry of base pairing will be examined. Subsequently, an analysis of mismatching in base pairing between NAM is discussed.

Chapter five will contain a conclusion as well as an outlook in which the prospects will be discussed.

Chapter 2. Literature study

2.1 The radical synthesis of sequence-defined polymers

Synthesizing sequence-defined (SD) oligomers can be realized by executing multiple consecutive polymerizations. When each polymerization binds a controlled amount of monomers onto the growing chain, a defined sequence can be obtained. To bind such a controlled amount of monomers onto the growing chain, control of the polymerization is key. FRP is a commonly used radical polymerization technique in the polymer industry showing many advantages (e.g. simple mechanism, moderate temperatures, wide range of usable monomers and fast polymerization), but it lacks the necessary control over the polymerization [21]. Lacking control means that polydispersity of the obtained polymers is broad, as shown in Figure 5, and that the amount of added monomers is not controlled. In FRP a low and controlled molecular weight of the synthesized polymers can hardly/not be obtained. The technique instantaneously renders a high molecular weight within the polymers, even at low monomer conversions. The broad dispersity and high molecular weight of the synthesised polymers are a result of slow initiation, fast non-simultaneously propagation and uncontrolled termination [21, 22]. Reversible deactivation radical polymerization (RDRP) on the other hand, also known as living radical polymerization (LRP) or controlled radical polymerization (CRP), is an optimal polymerization method for the synthesis of (short) SD oligomers. RDRP introduces the desired narrow control in the molecular weight and ensures low dispersity, as depicted in Figure 5. Additionally, due to the slow and controlled propagation, lower molecular weights within the synthesized polymers is obtainable. Added to the advantage of low dispersity and controlled chain length of the resulting polymers, is the high end-group fidelity, the capacity of continued chain growth and the possibility to synthesize polymers of complex structure that make it the preferred method over FRP [23, 24].

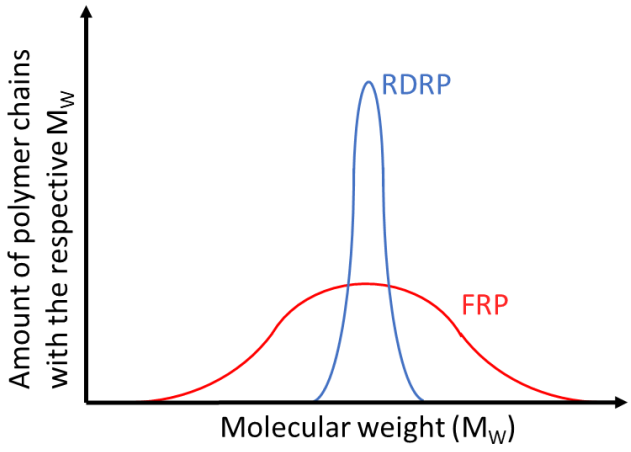


Figure 5. Difference in dispersity between RDRP and FRP.

There are two known forms of RDRP capable of synthesizing SD oligomers:

- Atom transfer radical polymerization (ATRP) [11]
- Reversible addition-fragmentation chain transfer (RAFT) [25]

For the synthesis of synthetic nucleobase containing polymers, requiring a defined sequence of nucleobases, RAFT polymerization is a promising technique. RAFT has several advantages over ATRP, that make it the most fitting polymerization technique. Firstly, RAFT polymerization has the capability of polymerizing a wide range of monomers [17, 24, 26]. Additionally there are no metal catalysts needed which is beneficial towards purification and eliminates potential problems with metal-nucleobase interactions [24, 27]. Furthermore, the technique provides synthesis of a wide range of polymers with a low dispersity and end or side chain functionality [26]. Additionally, these polymers can be synthesized without the need for protection or deprotection in a one step process [26]. Finally, RAFT polymerization is a technique that can be executed at mild temperatures in a wide range of solvents [17, 26].

2.1.1 Reversible addition-fragmentation chain transfer polymerization

2.1.1.1 Mechanism

The mechanism of RAFT polymerization is based on an equilibrium between dormant and active chains [28]. This equilibrium is achieved by a degenerative transfer via a chain transfer agent (CTA), in this research also referred to as the RAFT agent. An active radical chain becomes dormant (non-radical) by means of addition onto a CTA-molecule. The reversed process, where a dormant chain breaks from the RAFT agent and forms a radical, is called fragmentation. Degenerative transfer means that no overall change in the number of radicals occurs during the (activation) process. The presence of radicals is caused by a certain source, usually the initiator. Propagating radical (active) chains are reversibly deactivated by the degenerative transfer meaning that an equilibrium between dormant chains and propagating radicals is present. The addition/fragmentation-rate should be higher than the propagation rate, resulting in the addition of zero or one monomer onto the chain per fragmentation/addition-cycle. Control over this equilibrium and low radical or initiator concentration causes for the inhibition of termination and ensures a controlled and steady polymerization rate among all polymers during polymerization. It is this control that enables manipulation of the number of monomer insertions during polymerization by tuning the monomer/CTA-concentration ratio [23]. When the RAFT agent is absent in the process, control over the polymerization is lost resulting in an FRP-like polymerization.

The four main steps in thermal RAFT polymerization are shown in Figure 6 [23, 29]. In the first step, called initiation, an initiator (In) decomposes homolytically resulting in two radical molecules ($In^{\bullet}$). These radicals react with a monomer (M) and start the polymerization by producing an active polymerizing chain ($P_m{}^{\bullet}$). In the pre-equilibrium phase the active chain from initiation reacts with a CTA. This reaction causes a reversible release of the radical leaving group ($R^{\bullet}$). Re-initiation causes the leaving group to react with a monomer resulting in another growing and active polymer chain. In the main equilibrium phase, an equilibrium occurs

between the dormant and active (growing) chains. Dormant chains are bound to the RAFT-agent and cannot polymerise or undergo termination. Active chains are radical chains detached from the RAFT-agent. These active chains can react with monomers causing the chain to grow.

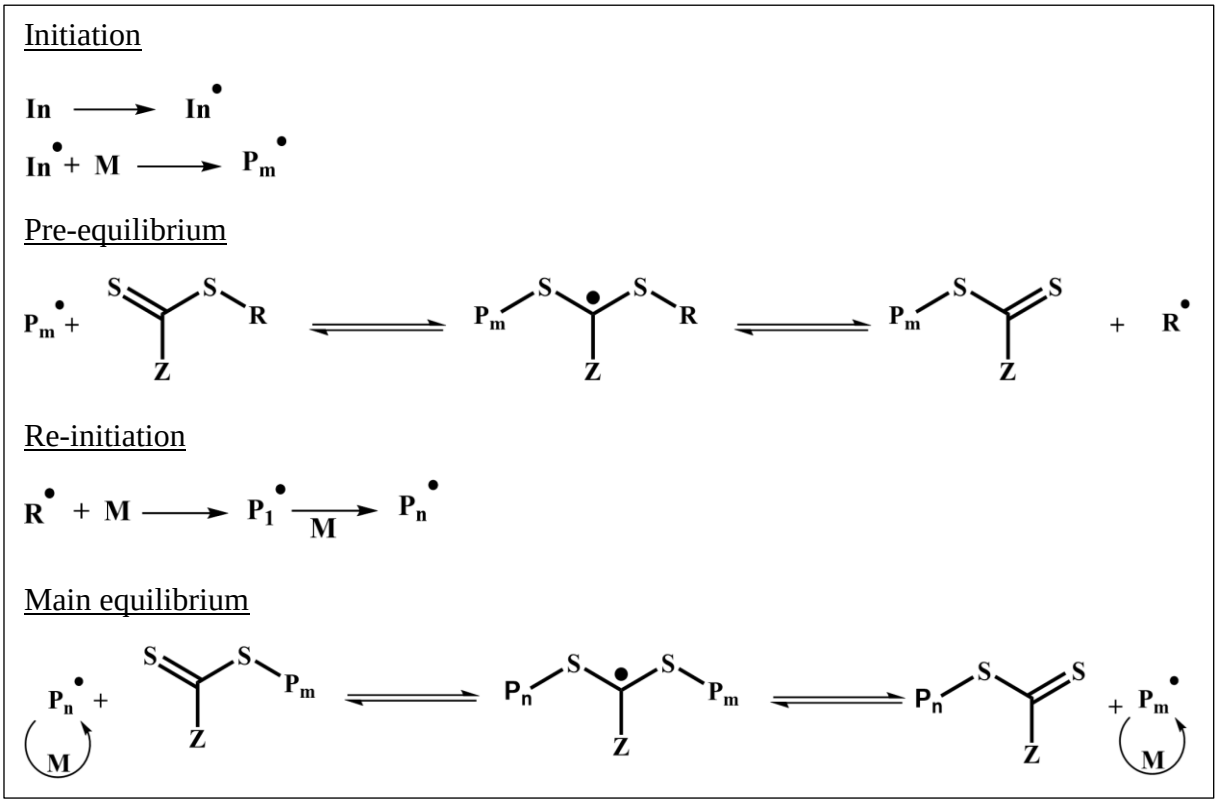


Figure 6. General RAFT polymerization mechanism.

2.1.1.2 RAFT-agent

RAFT-agents, also referred to as chain transfer agents, are generally molecules consisting of a R- and Z-group which are connected by a thiocarbonylthio-functional group [29]. The structure of such RAFT-agents is depicted in Figure 7 [29]. Every RAFT-agent differs from another by difference in R- and/or Z-group. Choosing the right groups is key for obtaining a working controlled polymerization with the selected monomers. In other words, a CTA is tuned depending on the monomers that are utilized.

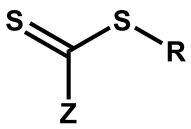


Figure 7. General structure of a RAFT agent.

The R-group in RAFT-agents primarily acts as homolytic leaving group when the RAFT-agent interacts with an active chain [29]. The more stable the resulting radical ($R^{\bullet}$), the more this R-group has the tendency to splice of homolytically. Therefore, radical-stabilizing effects such as electron-donating effects are important. If the requirement of good homolytic splicing is not

met within the R-group, inhibition[I] or retardation[II] of the polymerization can occur [30-32]. Secondly, the R-group acts as radical compound in the re-initiation phase [23]. Therefore, it must also be capable of reinitiating the polymerization, meaning that the radical should also be reactive or unstable enough for reinitiating to occur.

The Z-group, in the RAFT-agent, is responsible for the reactivity of the C=S and has an influence on the addition/fragmentation rates [29]. It is also responsible for the stability of the intermediate radical. The relation between this radical and a radical monomer is important. When a radical monomer is relatively stable, the Z-group will have to provide more stability to the intermediate RAFT-radical to favour the addition of the monomer onto the RAFT-agent. If this requirement is not met, inhibition, retardation and control problems could occur. When the monomer is not added onto the RAFT-agent, it will be able to react accordingly to the FRP mechanism. On the other hand, for a reactive and unstable monomer, the Z-group should not provide too much stability as it would result in slow fragmentation and again inhibition or retardation [29].

On top of the importance of the chosen CTA towards equilibria, is the importance of its polarity. The CTA should dissolve in the same solvent as the monomer. Otherwise, no RAFT polymerization can be executed. It means that the polarity of the monomer and CTA cannot differ too much. On the other hand, a difference in polarity is beneficial towards purification. When both compounds contain a different polarity, an increase in amount of monomer insertions causes for a change in polarity of the resulting polymer. The change in polarity is beneficial towards the production process as it enables separation of short polymers containing a different amount of monomers, based on polarity. Note that, for increasing chain lengths it is rather the monomer that will determine the polarity of the polymer instead of the RAFT agent. This means that also separation based on polarity will become less efficient as the chain length increases. Currently, for the polymerization of NAM, 2-dodecyl-1-phenylethyl trithiocarbonate (DPE-TTC) is used as RAFT-agent. DPE-TTC is shown in Figure 8. Due to the R- and Z-group, respectively the phenylethyl- and the dodecyl-group, this RAFT-agent is relatively non-polar. NAM are relatively polar causing for an increased the polarity of the structure upon each additional inserted monomer. The increasing polarity enables a good separation of the resulting polymers based on polarity for lower chain lengths. It is important to notice that for certain applications, where interaction of MHB-SD oligomers with biopolymers plays a role, the non-polar character of the DPE-TTC could cause problems. The problem concerning the polarity lays in the fact that DNA and RNA are polar (water soluble) biomolecules and for interaction to occur both have to dissolve in the same solvent.

[I] Inhibition period = a defined period in the beginning of the polymerization where (almost) no polymerization takes place.

[II] Rate retardation = a significant reduced polymerization rate compared to the polymerization rate in absence of the RAFT agent.

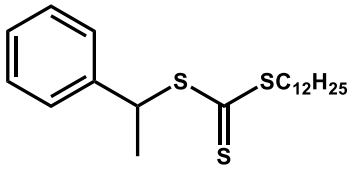


Figure 8. 2-dodecyl-1-phenylethyl trithiocarbonate (DPE-TTC).

For the interaction with biopolymers, possibly another RAFT-agent should be found. This RAFT-agent would then consist of a more polar R- and/or Z-group. Potential RAFT agents are 2-(2-Carboxyethylsulfanylthiocarbonylsulfanyl)propionic acid (CTAW) and 4-((((2-Carboxyethyl)thio)carbonothioyl)thio)-4-cyanopentanoic acid (CTAW2). Both are compatible with acrylates and soluble in water [33]. The structures of these RAFT-agents are shown in Figure 9.

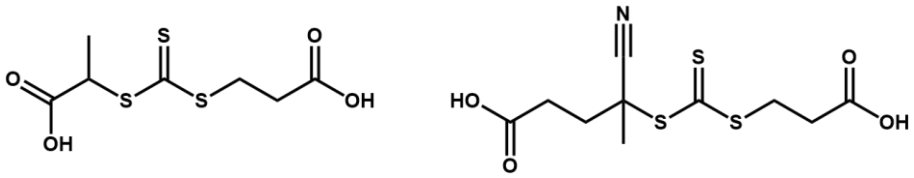


Figure 9. Structure of RAFT agents CTAW (left) and CTAW2 (right).

2.1.2 Structures for nucleobase functionalized monomers.

Nucleobases cannot readily be used in RAFT polymerization. Derivatization is needed for the molecules to obtain a functional group necessary for participation in radical polymerizations. The four types of nucleobase containing monomers, shown in Figure 10, are the most commonly used derivatives of nucleobases for RAFT polymerization [34].

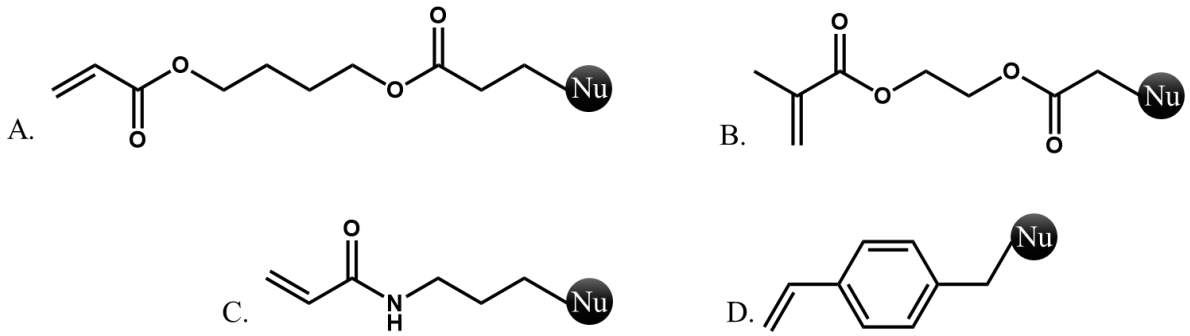


Figure 10. Four types of nucleobase containing monomers for RAFT polymerization. The respective monomers are nucleobase containing acrylate (A), methacrylate (B), acrylamide (C) and vinylbenzyl (D) monomers.

These four nucleobase containing monomers are compared in order select the optimal structure, here defined as the monomer that shows the most benefits towards synthesis and purification. For the synthesis this means that the procedure for synthesizing the nucleobase containing monomer should require the fewest amount of steps possible. Additionally, also the required conditions should be as such that the synthesis-setup is as simple as possible. For the separation, the monomer should cause for a sufficient increase in polarity when implemented into the monomer. When present, this property causes for an easy separation of the different polymer

lengths based on polarity. Based on polarity, nucleobase vinyl benzyl monomers (structure D, Figure 10) can be ruled out. Vinyl benzyl monomers are more non-polar than the other monomers and will therefore cause a smaller increase in polarity as the number of insertions in the MHB-SD increases. When comparing the number of steps and reaction conditions of the synthesis, NAM (structure A, Figure 10) prove to be the best choice. Their synthesis requires the fewest amount of steps and also synthesis conditions are milder, compared to the procedures used for all the other monomers [9, 17, 35, 36].

In the procedure, for synthesis of the desired NAM, a Michael-addition with 1,4-butanediol diacrylate (BDDA) on the desired nucleobase A, U, C or T is used to produce the corresponding monomers (AAM, UAM, CAM and TAM). The guanine acrylate monomer (GAM) requires a different procedure. Guanine reacts barely in the given Michael-addition and when it does, the resulting products are a mixture of two isomers. These isomers are difficult to separate. Therefore, for the synthesis of GAM, the procedure starts a Michael addition between 2-amino-6-chloropurine and BDDA. Subsequently a nucleophilic substitution is performed with $HCOOH/H_2O$, [12]. These monomers are shown in Figure 11.

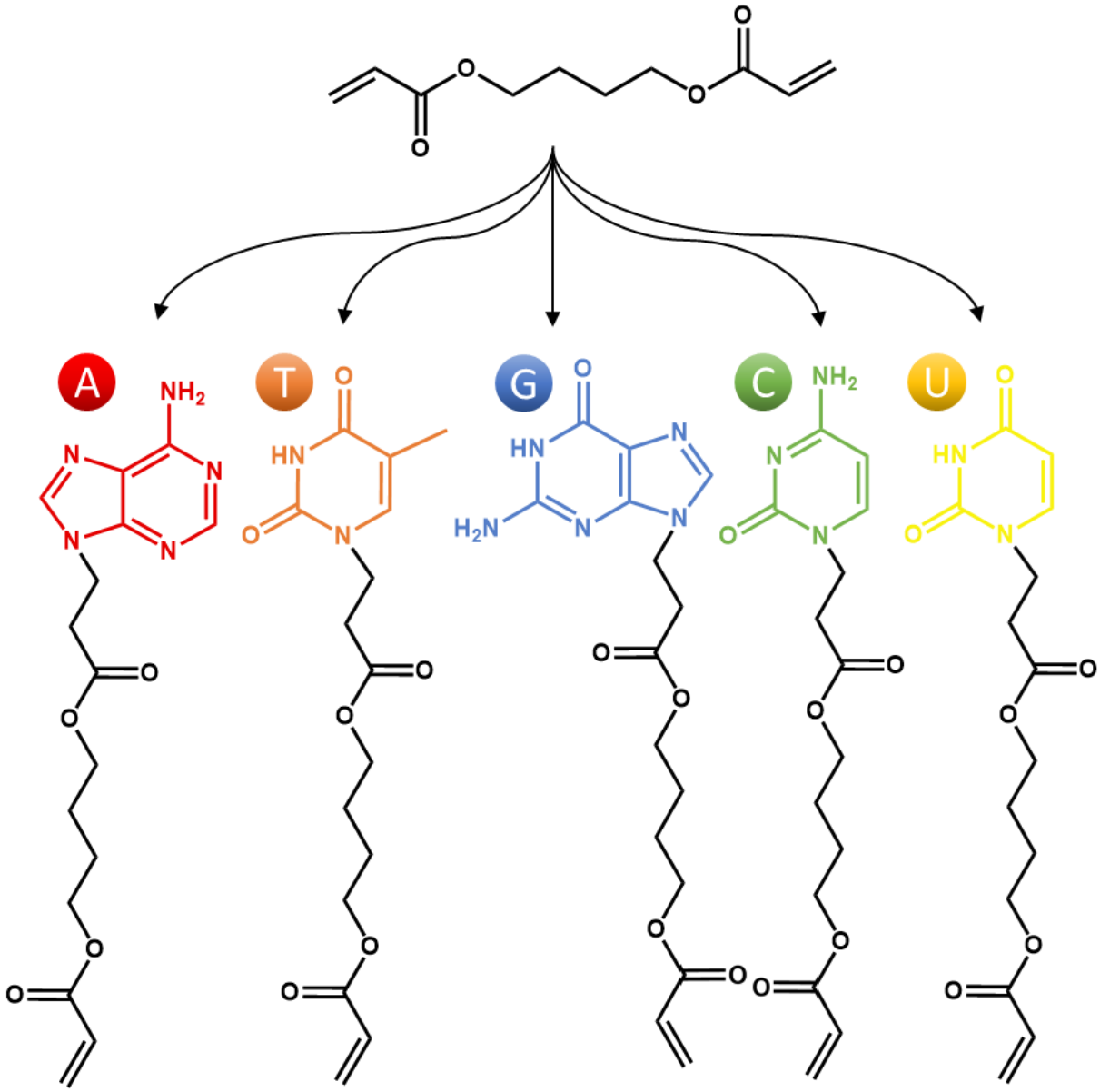


Figure 11. Molecular structures of the five nucleobase containing monomers. AAM, TAM, CAM and UAM are synthesized via an aza-Michael-addition between the corresponding nucleobase and BDDA. The reaction is executed at room temperature (50 °C for AAM) for 24 h (5 h for AAM). For GAM, a Michael addition between 2-amino-6-chloropurine and BDDA is performed at 50 °C for 24 h. Subsequently a nucleophilic substitution is performed with $HCOOH/H_2O$ for 2 h at 75 °C resulting in the GAM [10].

Even though the NAM are the better choice for RAFT polymerizations, an important note should be stated concerning the applications where the MHB-SD should interact with biopolymers. The polarity of the MHB-SD is increasingly dependent on the polarity of the NAM as the number of insertions increases. The limited solubility of longer MHB-SD is therefore mainly the result of the polarity of the NAM. The NAM are polar, but not enough to be soluble in water. By creating water soluble nucleobase containing monomers, in combination with a new and polar RAFT agent, a water soluble MHB-SD could be created. Suggestions for other monomers are given below.

Even though nucleobase containing acrylamide monomers (NAAM) require a complex synthesis, they are more polar than NAM. The increased polarity results in better solubility of the MHB-SD in water. Therefore, for applications relying on interaction between MHB-SD and biopolymers, NAAM might be chosen over NAM despite their synthesis complications. On a more critical note, there might be another problem with NAAM as monomer. Specifically, in the procedures template-assisted polymerization and affinity separation. The monomer contains another nitrogen in its structure, apart from the ones that were already present in the bare nucleobases. If the proton, on the additional nitrogen, forms hydrogen bonds with the nucleobases from other monomers, complications could occur in the mentioned procedures. Therefore, before using NAAM as monomer, interaction characterization is recommended.

Ethylene glycol diacrylates might also offer a solution to the polarity and water solubility problem. They should show increased polarity in comparison to BDDA, resulting in better solubility in water. General structure of the ethylene glycol diacrylates is shown in Figure 12.

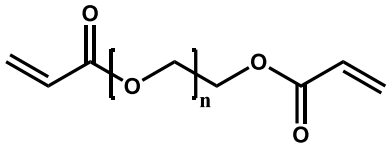


Figure 12. General structure of ethylene glycol diacrylates.

2.1.3 Single unit monomer insertion procedure for synthesising nucleobase containing oligomers

As mentioned before, RAFT polymerization is an RDRP-technique that gives great control over molecular weight and the dispersity of the polymer mixture. This is not the only requirement for synthesis of SD oligomers. To obtain SD oligomers, controlling the number of monomer insertions during polymerization is key. It can be calculated by means of equation (1) [23]. Note that this equation is a simplified equation but gives a good estimation on the number of insertions. The formula also does not give the exact amount of insertions obtained in all monomers but an average molar mass of the polymer mixture.

$$M_{n,th} = \frac{[M]_0 \cdot p \cdot M_M}{[CTA]_0} + M_{CTA} \quad (1)$$

Here $M_{n,th}$, M_M and M_{CTA} represent the molar masses of the resulting oligomer, the monomer and the RAFT-agent respectively. $[M]_0$ and $[CTA]_0$ represent the initial concentrations of the monomer and the RAFT-agent respectively. p is represents the monomer conversion and should be determined experimentally for the given conditions (e.g. initiator concentrations, reaction time, temperature, etc.).

The equation states that the $[M]_0$: $[CTA]_0$ ratio is the most decisive factor to determine the amount of insertions into the oligomer. To obtain oligomers in which mostly only one monomer has been inserted, the formula states that a 1:1-ratio should be used for the polymerization. Note that not only oligomers with the desired amount of insertions are obtained. There is still a polydisperse nature present in RAFT-polymerisation, meaning that oligomers with more or less insertions will also be obtained

The MHB-SD can now be obtained by inserting monomer after monomer into the polymer chain. Important is that the CTA/monomer-ratio for inserting one monomer is different for the first insertion in comparison to the subsequent insertions. For the first insertion a CTA/NAM-ratio of approximately 1/3 to 1/5 is necessary. This is a result of the fact that the first addition of a monomer radical onto the pure CTA is not efficient due to altered kinetics of the first insertion in comparison to further insertions. After this first insertion, a greater reactivity is present between the two structures. Here, also a lower CTA/NAM-ratio of approximately 1/1 is necessary to obtain as much as possible single monomer insertions. Isolation of the desired oligomer, obtaining one insertion, after each polymerization step is required as a dispersity will always occur. Note that, due to this statistical deviation, the maximum obtainable yield of desired product decreases after each monomer insertion. This results in a maximum number of monomer insertions, while still obtaining a sufficient amount of desired product.

Size exclusion chromatography has been used in research for the isolation of the desired oligomer [25]. Unfortunately, a limitation of this purification method is that it can only be used to purify small batches. Flash column chromatography has proven to be applicable as well and is capable of purifying larger batches making it the preferable method [12, 37, 38]. Even though flash column chromatography has proven to be useful for the separation of oligomers containing around four monomers, complications lay ahead when chain length is increased further. A decreased difference in polarity between the chains and base pairing between the nucleobases implemented into the chains will be disadvantageous towards the separation of the disperse mixture of polymers. This base pairing, impeding the column chromatography separation of the different oligomers, could be used as an advantage in template-assisted polymerization and affinity separation. However, this requires the occurrence of base pairing between solely complementary nucleobases and the absence of mismatches. These phenomena will be discussed in the following section.

2.2 Hydrogen bonding and base pairing

Synthetic nucleobase-containing polymers would use the characteristic of selective hydrogen bonding or selective base pairing in an application such as synthetic nucleobase containing biosensors. This selective bonding or pairing could also be used for improved synthesis of these MHB-SD as mentioned before. Hydrogen bonding is the central non-covalent molecular interaction between nucleobases. It is the phenomenon in which a proton, having a partial positive charge, interacts with two atoms that have a partially negative charge [39]. One of these atoms, with a partially negative charge, is covalently bonded to this proton. The other atom, containing a partially negative charge, is part of another molecule. The strength of a hydrogen bond between nucleobases is dependent on two main factors, besides the environmental factors. The first factor is electronegativity of the partially negative charged atoms. The greater this electronegativity, the greater the electromagnetic interaction that is the hydrogen bond. There is however a limit to the strengthening effect of increasing electronegativity. As the electronegativity increases, the tendency of an element to share its electrons will decrease. This means that the element will be less likely to take part in hydrogen bonding as an hydrogen bond acceptor [40]. This means that the electronegativity of an acceptor in a hydrogen bond is limited. The second factor is spatial arrangement and/or molecular structure. As the distance of opposing charges decreases, the electromagnetic forces pulling them towards each other increases. For hydrogen bonds it means that when the distance between two hydrogen bonding molecules decreases, due to reduced steric hindrance or kinetic energy, the hydrogen bonding strength increases. Increase in hydrogen bonding strength means that a larger amount of (thermal) energy is needed to disrupt such bonds.

Base pairing is the phenomenon in which (multiple) hydrogen bonds cause two nucleobases to form a complex. The stability of a base pair, existing due to hydrogen bonding, is dependent on the same factors as individual hydrogen bonds. There is however one additional factor. As the number of hydrogen bonds increases within a base pair, the (thermal) energy needed to cleave the complex into its two separate nucleobases increases as well.

Note that there are thermal limitations to the formation of hydrogen bonds or base pairs. The thermal limitation means that from a certain temperature onwards, hydrogen bonds will be disrupted and will not/rarely form. The reason for this disruption can be derived from the formula for change in Gibbs free energy:

$$\Delta G = \Delta H - T \cdot \Delta S \quad (2)$$

A molecule will always tend to take the form in which it obtains the lowest (most negative) Gibbs free energy. The change from one state to another, causes a change in Gibbs free energy noted as ΔG. When this difference in Gibbs free energy is lower than zero ($\Delta G < 0$) the molecule will have made a transition to a more stable state. These transitions occur spontaneously. The ΔG is dependent on the change in enthalpy (ΔH) and entropy (ΔS) and the temperature (T in Kelvin) of the environment. When a hydrogen bond or base pair is formed, two molecules will form one complex resulting in a decrease in entropy and thus a negative ΔS.

This causes an increase in ΔG and will, based on solely this criterium, prevent/reduce the formation of hydrogen bonds. The stabilization effect of the molecules due to hydrogen bonding, causing a decrease in enthalpy (ΔH<0), can however surpass the influence of this destabilizing entropy decrease ($|\Delta H| > |-T \cdot \Delta S|$). When this happens, hydrogen bonding is feasible in terms of energy and obtains the advantage over the state where molecules are not hydrogen bonded. The conditions for which this term is met, are limited. As the temperature increases, the influence of the change in entropy on the stability or Gibbs free energy, increases as well. This can cause for the ΔG of hydrogen bonding to become positive again. When the temperature is increased to this point or beyond, hydrogen bonding will become unfeasible again and will therefore not/rarely occur.

For hydrogen bonding, base pairing or complexation it is not that it either occurs or not. The complexation is rather described by an equilibrium that takes place and the parameters mentioned in previous paragraph influence this equilibrium. The equilibrium is described by:

$$H + G \underset{K_d}{\overset{K_a}{\rightleftharpoons}} HG \tag{3}$$

with HG as the complex, H as the free host molecule and G as the free guest molecule in the complexation. As the temperature increases, ΔG increases and the equilibrium will tend more to the left, meaning that the ratio K_a/K_d decreases. An increase in hydrogen bond or base pair strength on the other hand will cause a decrease in ΔG and will shift the equilibrium more to the right, meaning that the ratio K_a/K_d increases.

There are several ways for this base pairing to occur. All of them show different hydrogen bond strengths or enthalpies. An important distinction that should be made in this base pairing is that between correct matches and mismatches. Among these correct matches a further classification can be made into Watson-Crick (WC) base pairing and alternative base pairing models.

2.2.1 Base pairing between complementary nucleobases

2.2.1.1 Conventional Watson-Crick base pairing

One model describing these nucleobase interactions is the Watson-Crick (WC) base pairing model [39, 41]. It is the standard model for the formation of hydrogen bonds between nucleobases on which the applications of nucleobases in DNA, RNA and some applications of MHB-SD are based. The hydrogen bonding between nucleobases, according to this model, is shown in Figure 13.

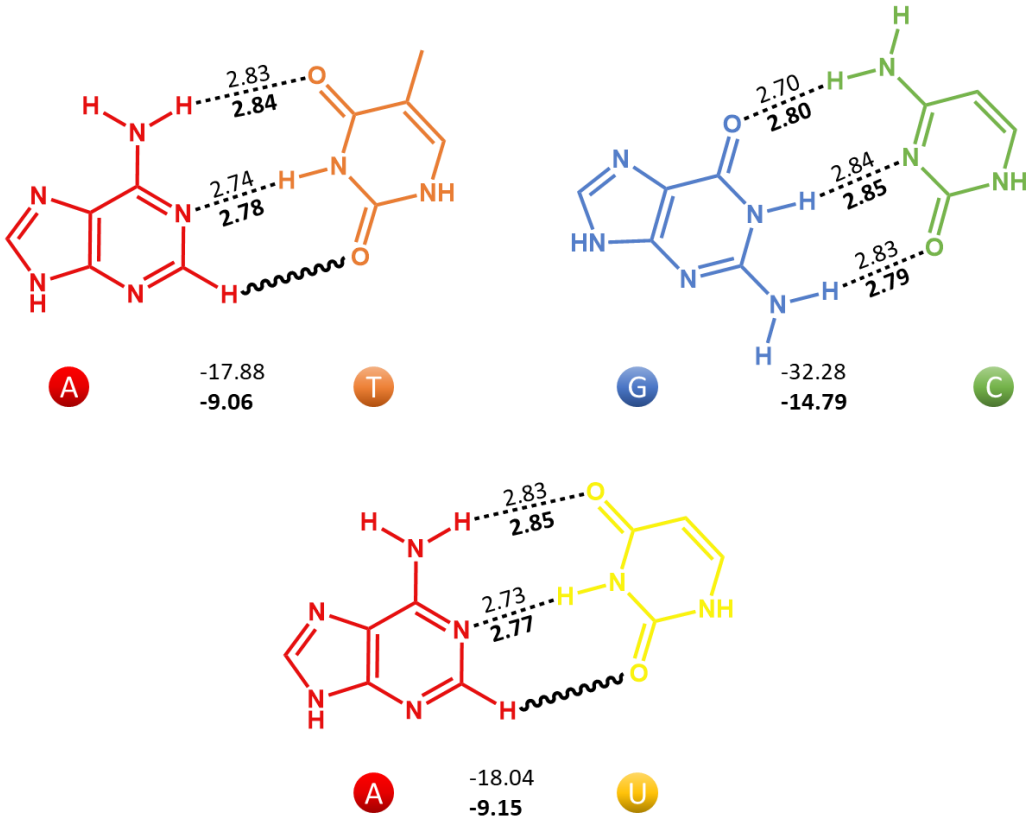


Figure 13. A-T (top left), G-C (top right) and A-U (bottom) Watson-Crick base pairs. The hydrogen bonds are represented by the dashed lines and the wavy lines represent a weak interaction between C-H and C=O. The lengths of the hydrogen bonds (in Å) are given in gas phase (plain) and in aqueous solution (bold) at the respective bonds. The hydrogen bond-energies are given below the respective base pair (in kcal/mol) [42].

As it can be seen in Figure 13, the A-T an A-U pair associate trough two hydrogen bonds and G-C pairs are held together by three of these interactions. Research also shows that G-C base pairs are more stable than A-T pairs which corresponds to the difference in number of hydrogen bonds, providing this stability [43, 44]. Additionally, a third stabilizing interaction can be observed in the A-T WC model. This is the interaction between the carbonyl-oxygen and a proton from the double bond of adenine [39, 45]. Also note that the length of hydrogen bonds are longer and the bonding energies increase in aqueous solution in comparison to gas phase. This difference is caused by solvation resulting in an increased stabilization of the charges of the free nucleobases in the aqueous solution [42].

2.2.1.2 Reversed Watson-Crick base pairing

Both WC and its reversed model seem similar. The most significant difference is present in the G-C reversed WC base pair. Here, one hydrogen bond is lost which translates into the drop in hydrogen bonding energies or base pair stability from -32.28 kcal/mol to -14.95 kcal/mol. Increase in the length of the hydrogen bonds also contributes to the lower stability of the base pairs. Because of the lower stability, this type of base pairing will occur less frequent than the WC base pairing. The drop in hydrogen bonding stability decreases, for all pairs, as they are present in an aqueous solution. This decreasing drop implies that selectivity towards the WC base pairing system would still be present but decreases as solvation of the monomers increases/occurs. For A-T and A-C base pairing, the drop in stability upon solvation is similar between WC and reversed WC. In the G-C WC base pairing, this drop is 12.81 kcal/mol greater compared to the reversed G-C WC base pairing. This is a result of the fact that in G-C WC base pairing, three hydrogen bonds are weakened due to the presence of a polar solvent. In the reversed G-C base pairing, only two hydrogen bonds are weakened upon solvation causing for a decreased drop in stability.

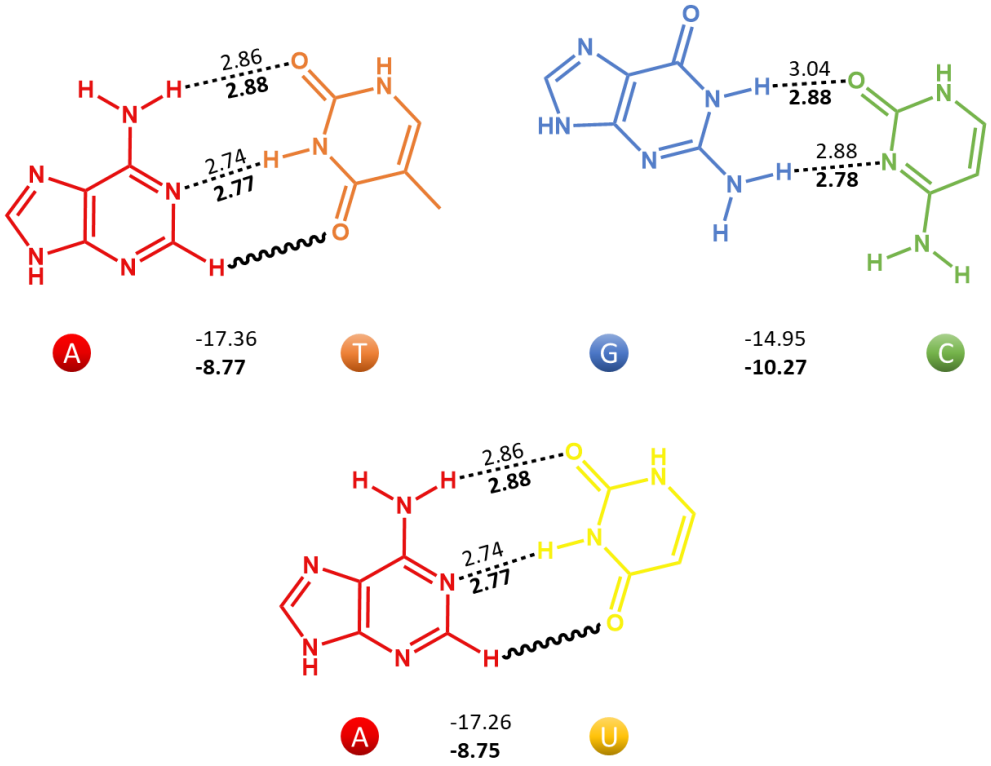


Figure 14. A-T (top left), G-C (top right) and A-U (bottom) reversed Watson-Crick base pairs. The hydrogen bonds are represented by the dashed lines and the wavy lines represent a weak interaction between C-H and C=O. The lengths of the hydrogen bonds (in Å) are given in gas phase (plain) and in aqueous solution (bold) at the respective bonds. The hydrogen bond-energies are given below the respective base pair (in kcal/mol) [42].

2.2.1.3 (Reversed) Hoogsteen base pairing

The Hoogsteen base pair model is depicted in Figure 15 and Figure 16. Decreased stability of the (reversed) Hoogsteen base pairing in A-T and A-U, in comparison to WC, is a result of disappearance or extreme weakening of the weak interaction between C-H and C=O due to the different orientation of the molecules in the complex [39, 41, 45]. Additionally, when comparing the complexes, a difference in stability/lengths of the hydrogen bond between C=O and C=NH_2 is observed. As the position of nitrogen, in the hexagon structure, alters relative to the hydrogen bonding carbonyl group, the strength of this hydrogen bond alters. This is a result of the electron withdrawing effect of the electronegative nitrogen. As this nitrogen is closer to the hydrogen bonding C=O group, it lowers the partial negative charge of the carbonyl-oxygen resulting in this weakening of its hydrogen bond. Note that the same phenomenon is observed when comparing WC and reversed WC. Because of the lower stability of the model, Hoogsteen base pairing is short lived (0.2 – 2.5 ms; ± 1 %) and will occur less frequent than the WC base pairing in the DNA double helix [46].

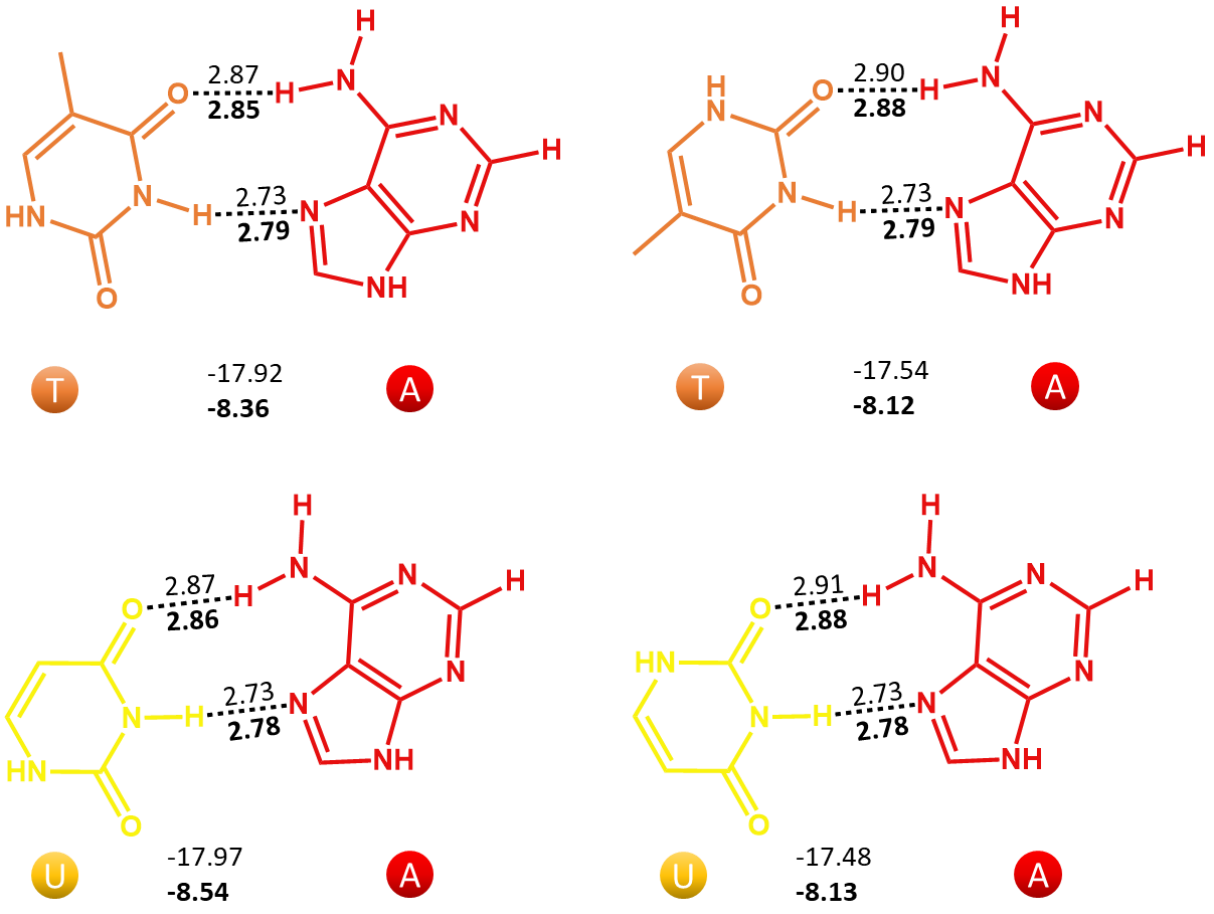


Figure 15. A-T Hoogsteen (top left), A-T Reversed Hoogsteen (top right), A-U Hoogsteen (bottom left) and A-U Reversed Hoogsteen (bottom right) base pairs. The hydrogen bonds are represented by the dashed lines. The lengths of the hydrogen bonds (in Å) are given in gas phase (plain) and in aqueous solution (bold) at the respective bonds. The hydrogen bond-energies are given below the respective base pair (in kcal/mol) [42].

Additionally a Hoogsteen and reversed Hoogsteen base pairing model exists for G-C base pairing [47, 48]. Unfortunately, no stability data or lengths of hydrogen bonds from this form of base pairing are available. Therefore, some theoretical assumptions will be made. Firstly, one can assume that these base pairs are a lot less stable in comparison to the G-C WC base pairs. This assumption can be made as these models only entail two hydrogen bonds and the G-C WC base pair entails three of these bonds. As observed for reversed WC base pairing, this causes for a significant decrease in the stability of the complex. Note that this would not be the only reason for reduced occurrence of these complexes. Additionally, protonation of a nitrogen on cytosine is required to enable this form of base pairing. Therefore, the pH should be low enough to cause this protonation. As base pairs are usually used in a pH-neutral environment, this form of base pairing is assumed to occur seldomly.

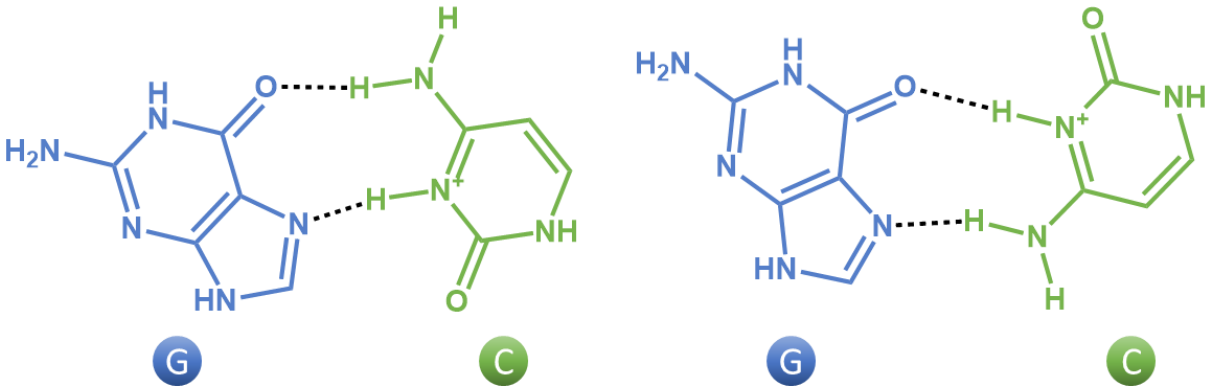


Figure 16. G-C Hoogsteen (left) and G-C reversed Hoogsteen (right) base pairs.

2.2.2 Base pairing between non-complementary nucleobases

Mismatching is the formation of hydrogen bonds between two non-complementary nucleobases. Occurrence of this phenomenon could have severe consequences. Mismatches have, for example, been related to 10%–30% of spontaneous cancers in a variety of tissues and they have also been associated with several hereditary cancers [49-51]. Furthermore, frequent occurrence of mismatches will cause for the templated polymerization and the affinity separation to be ineffective as solutions for the low yield problem in synthesis of MHB-SD.

2.2.2.1 Wobble mismatching

The Wobble base pair is a well-known model for base pair mismatching. This model describes mismatching between for example G and T or G and U [39, 52]. These forms of mismatching are shown in Figure 17. This form of base pairing occurs seldomly, in comparison with the WC base pairs, due to their lower stability [53, 54]. When both are compared to WC base pairs in aqueous solution, the WC model renders the most stable complexes. This is not the same for the gas phase. Here the A-T or A-U according to WC do not render a more stable complex than G-T or G-U. Therefore, T and U could be more likely to form a base pair with G. For G however, WC does give a much more stable base pair in comparison to the Wobble complexes. The higher stability of a WC base pair is caused by addition of a third hydrogen bond. G will now be more likely to form a base pair with C according to WC leaving the T and U naked. This will most likely happen as the increased stability of a G-C WC complex, compared to Wobble mismatching, is very high. The increased stability of G-T and G-U, compared to WC A-T and A-U, on the other hand is very small.

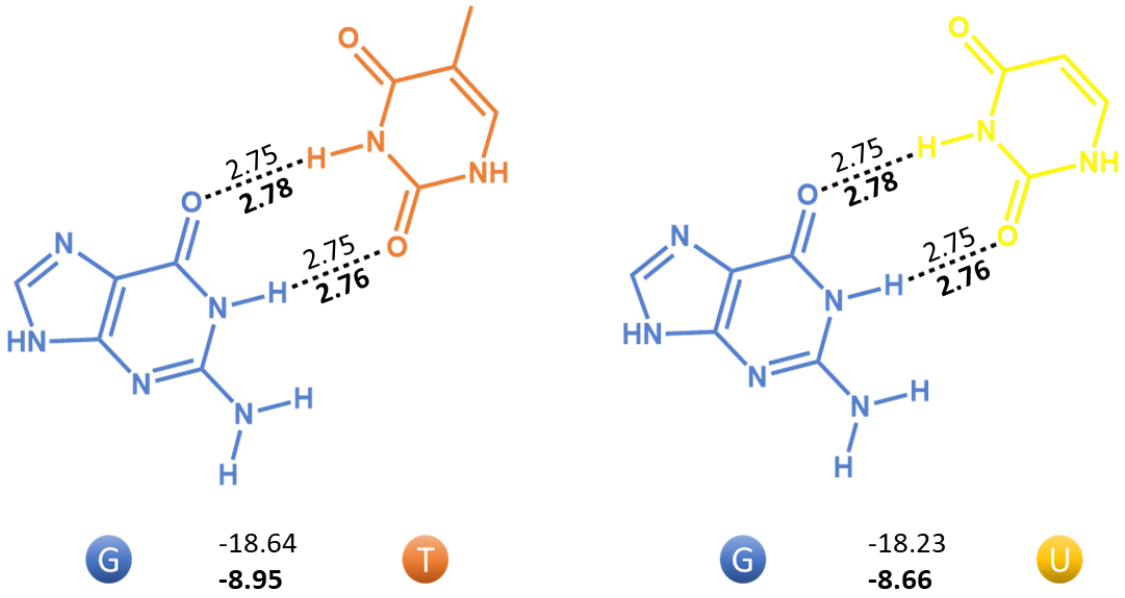


Figure 17. Wobble G.T (left) and G.U (right) base pair. The hydrogen bonds are represented by the dashed lines. The lengths of the hydrogen bonds (in Å) are given in gas phase (plain) and in aqueous solution (bold) at the respective bonds. The hydrogen bond-energies are given below the respective base pair (in kcal/mol) [42].

2.2.2.2 Anti-syn mismatching

Another form of mismatching is called anti-syn base pairing. This is a form of hydrogen bonding between two purines. Purine is a collective name for the group of nucleobases A and G. In this form of mismatching one purine occurs in the unusual syn configuration and the other one in the anti-configuration. An example with guanine in anti-configuration and adenine in the syn-configuration is given in Figure 18 [55]. Fortunately, the anti-configuration of a purine and therefore also this form mismatching occurs rarely [55]. Even if this anti configuration would occur frequently, the stability of this form of base pairing is less stable than both G-C and A-T WC base pairing.

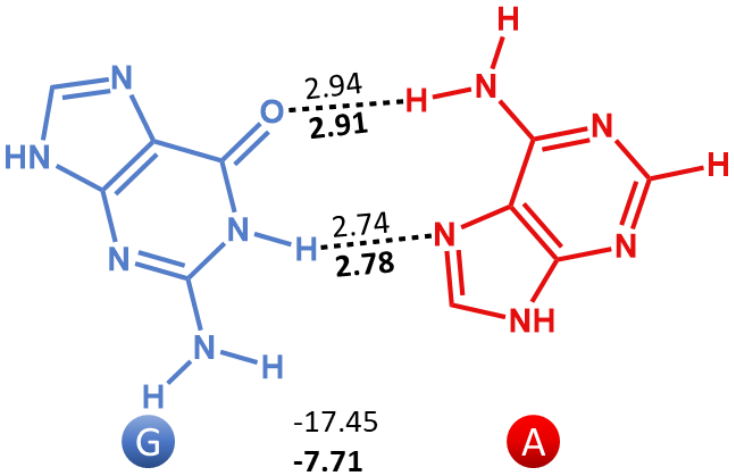


Figure 18. Anti-syn base pairing between guanine (anti) and adenine (syn). The lengths of the hydrogen bonds (in Å) are given in gas phase (plain) and in aqueous solution (bold) at the respective bonds. The hydrogen bond-energies are given below the respective base pair (in kcal/mol) [42].

2.2.2.2.1 Tautomerism mismatching

Mismatching can also be caused by tautomerization of the nucleobase as shown in Figure 19 [41, 55]. This model requires alternations in the structures of the molecule. The keto-enol or amino-imino tautomeric forms of the nucleobases can result in the A-C and G-T mismatching. Fortunately, enol forms of G and T and imino forms of A and C occur rarely [39].

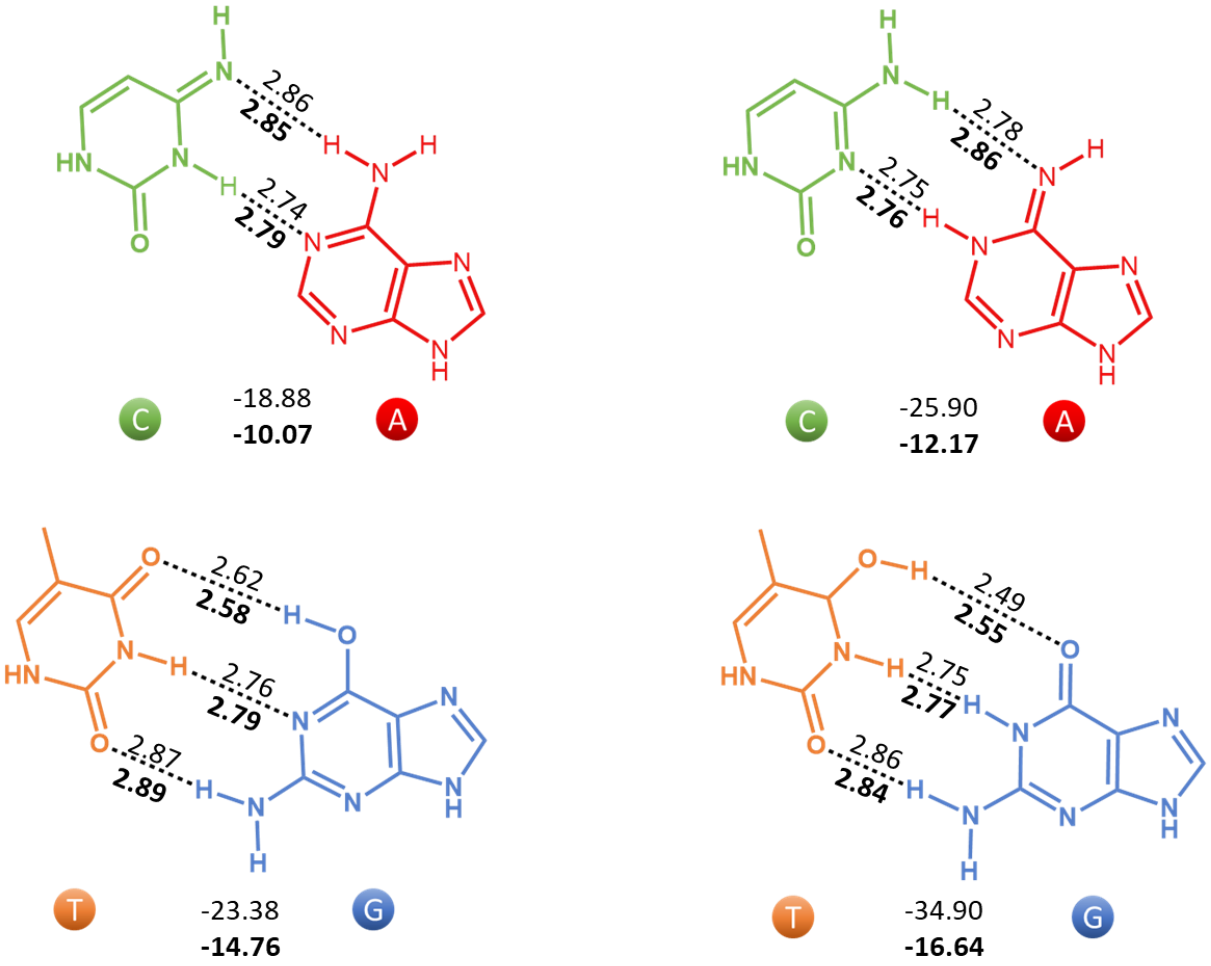


Figure 19. Base pairing of tautomeric isomers. The mismatches shown are mismatches between the imino form of cytosine and adenine (top left), cytosine and the imino form of adenine (top right), thymine and the enol form of guanine (bottom left) and between the enol form of thymine and guanine (bottom right). The lengths of the hydrogen bonds (in Å) are given in gas phase (plain) and in aqueous solution (bold) at the respective bonds. The hydrogen bond-energies are given below the respective base pair (in kcal/mol) [42].

2.2.2.2.2 Protonated mismatching

The last form of mismatching that will be discussed is an influence rather than a model. In literature, it is seen as a way to promote certain forms of base pair mismatching rather than being a model on its own [18]. This influence entails protonation of certain nitrogen-atoms in the nucleobase structure. Certain protonated forms of mismatching structures are given in Figure 20. Unfortunately, no data concerning energies or lengths of the hydrogen bonds, comparable to the data from the other model is available. However, research was executed determining stability by means of melting point and Gibbs-free energy analysis [54]. In lower pH ranges the mismatches showed lower Gibbs-free energy and higher melting temperatures and therefore increased stability of the mismatches, described above. This indicates that acidic environments do promote certain forms of mismatching.

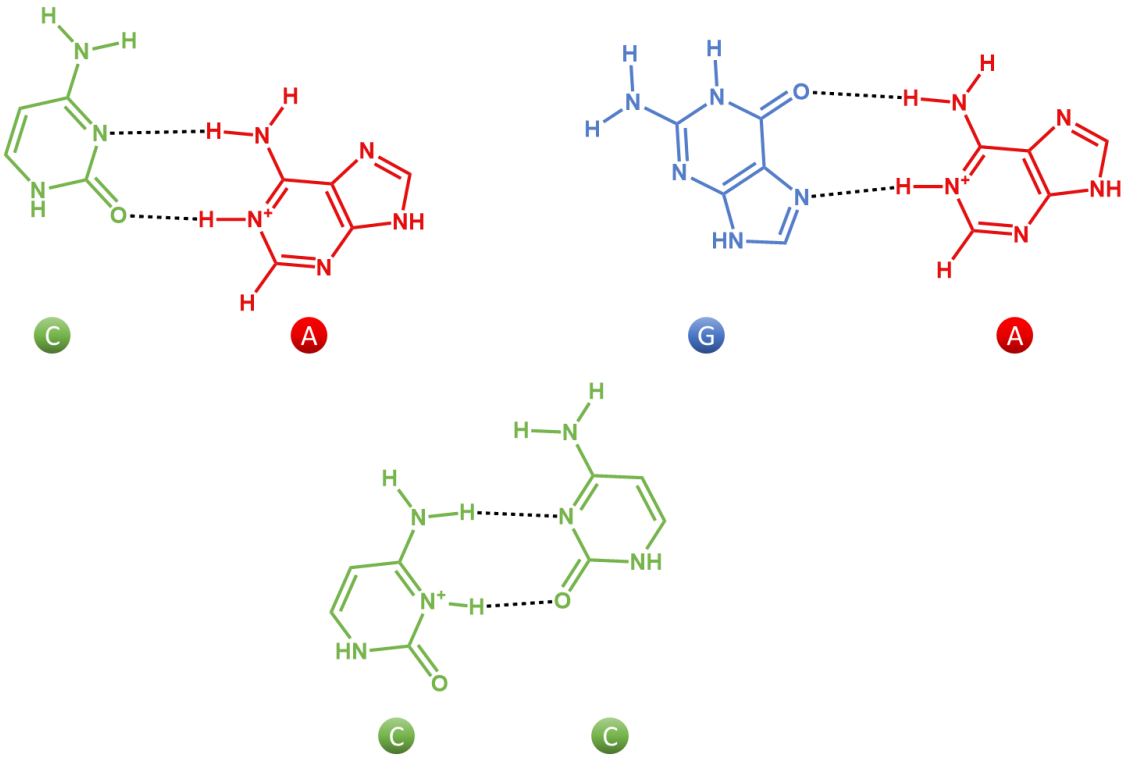


Figure 20. Mismatching due to protonation of the nucleobases. The mismatches that are shown are C-A^+ (left), G(syn)-A^+(anti) (right) and C-C^+ (bottom).

2.2.3 Optimal environment for selective hydrogen bonding

As proven by stability data, mismatching occurs seldomly due to rare occurrence of the configuration or tautomer, lower stability of the pairs or unlikely environment. Occurrence of mismatching could and should still be examined and minimized for the templated polymerization and the affinity separation to be reliable pathways for the synthesis of MHB-SD. This optimization could be possible by changing the environment in which the nucleobases are present. An optimal environment for selective hydrogen bonding or base pairing, maximises the formation of base pairs consisting of complementary nucleobases. This environment also minimizes, or even inhibits, the formation of mismatches. In this research, the assumption is made that such an environment might be obtained. This would be realized by setting effects, that are disadvantageously towards hydrogen bonding, at a sufficient high level that base pairs of non-complementary nucleobases are broken. Since base pairing between complementary nucleobases is more stable, the level of these effects should also be low enough for these base pairs to form. To obtain such an environment, following four main-effects on hydrogen bonding selectivity will be discussed: temperature, solvent polarity, pH and concentration ratio.

2.2.3.1 Temperature dependence

A first parameter affecting hydrogen bonding is the temperature. It is known that, as temperature increases, non-covalent bonds between molecules/atoms break. The reason is given in the introduction of 2.2. When hydrogen bonds between nucleobases break, base pairs are disrupted. Due to the difference in stability of the formed base pairs, every complex will be disrupted at a different temperature [41, 54, 56]. This difference in stability, as proven in the previous section, is a result of difference in molecular structure, electronegativity and number of hydrogen bonds that form the base pair. The stability of base pairs was extensively examined by examining the temperature at which base pairs were disrupted [54]. This research revealed that complementary nucleobase containing base pairs were more stable than mismatched base pairs, as shown in Table 1.

Table 1. Gibbs free energy of the different base pairs at a pH of 7.9 and a temperature of -273 K [54]. Correct matches (bold) prove to be more stable than mismatches (plain).

	A	T	G	C
C	-29.5 ± 0.9	-32.6 ± 0.7	**-46.3 ± 0.6**	-30.1 ± 0.7
G	-34.1 ± 0.8	-36.2 ± 0.5	-34.7 ± 0.4	
T	**-41.5 ± 0.6**	-35.0 ± 0.9		
A	-30.4 ± 1.9			

An important remark is that the difference in disruption-temperature between the WC model and models describing mismatches, at the given circumstances, is small [54]. Nevertheless, it implies that finetuning of the temperature of the environment might eliminate the number of mismatches formed. Also, note that the increase in temperature will also cause for the disruption of base pairs between complementary nucleobases. However, as these base pairs are more stable, their occurrence will most likely not be eliminated totally by this temperature increase.

An important factor to take into account, when examining the base pair disruption-temperature between nucleobase containing oligomers, is the influence of chain length. The strength of non-covalent binding between nucleobase containing polymers increases as their length increases [57, 58]. Partially due to increasing of the number of hydrogen bonds. This can be derived from the modified Marmur-Doty formula for the melting temperature of oligo-nucleotide complexes [59]:

$$T_m = 2(A + T) + 4(G + C) - \varepsilon \quad (4)$$

A, T, G and C represent the number the respective nucleotides in the chain and ε is a factor that takes the solvent into account. Furthermore, a higher C and/or G content in the polymers has proven to render more stable non-covalent binding of the chains [58]. This can be seen in Table 1 and is proven in the Marmur-Doty formula for the melting temperature of oligonucleotide complexes under specific circumstances (solvent containing 0.2 M Na^+) [60]:

$$T_m = 64.9 + 41.0\left(\frac{wG + xC - 16.4}{yA + zT + wG + xC}\right) \quad (5)$$

w, x, y and z are the number of the bases of G, C, A and T, respectively. This increased stability of chains with a higher C and/or G content, as depicted in Figure 21, is a result of higher stability of a C-G base pair compared to an A-T or A-U base pair. Base-stacking could also provide more stability to the complexes. This phenomenon is important towards the stability of the DNA-double helix [57, 58]. The presence or influence of the phenomenon in MHB-SD has not been proven yet. Nevertheless, this base-stacking should not be neglected.

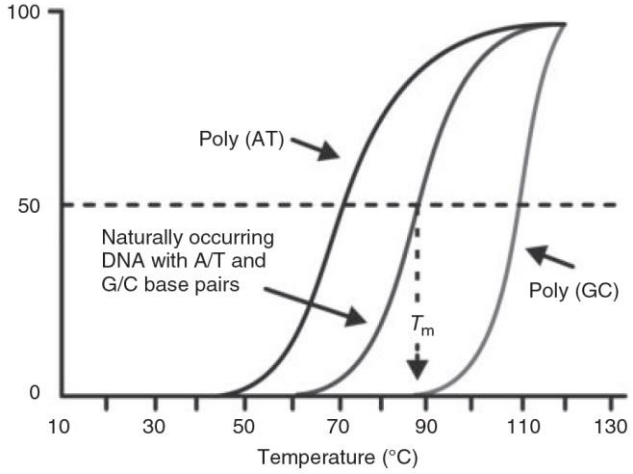


Figure 21. Melting curve of DNA [61]. The y-axis represents the percentage of the DNA-strand that has been disrupted.

2.2.3.2 Solvent polarity dependence

Another factor that influences the stability of hydrogen bonds is the polarity of the solvent. Studies have shown that solvent polarity alters the strength/stability of hydrogen bonds [42, 62]. The same is observed for nucleobases specifically. Research proves that stability of base

pairs and thus strength of hydrogen bonds weakens as nucleobases are dissolved in increasing polar solvent. The decreased stability is a result of lengthening of the hydrogen bonds due to stabilization of the lone pairs participating in these hydrogen bonds [42]. This stabilization of the lone pairs occurs due to a phenomenon called solvation. According to this theory, increasing base pair stability should be observed as polarity of the used solvent decreases. Research revealed that a decrease in solvent polarity, indeed causes a decrease in hydrogen bond distances and therefore increases hydrogen bond and base pair stability [63]. This means that choosing the right solvent could promote the formation of WC base pairs, by disrupting mismatches and leaving base pairs between complementary nucleobases intact.

2.2.3.3 pH dependence

The pH of the environment, in which the nucleobases are present, also influences the base pairing. The different nucleobases have different pK_a-values. This means that the nitrogen functional groups of the different nucleobases can be protonated at different acidic pH-values were a significant effect on the hydrogen bond stability is observed [18, 64, 65]. This protonation of nucleobases results in promotion of certain mismatches, such as described by the Wobble model [18, 64, 65]. Therefore, these acidic environments should be avoided at all times, for example when executing a templated polymerization or affinity separation with nucleobases.

2.2.3.4 Concentration ratio

A last important parameter for obtaining the optimal environment for selective base pairing to occur is the concentration in which the nucleobases are present in the mixture. This will be explained with a theoretical example. Assume a mixture consisting of G/C/T in a 5/1//1 molar concentration ratio. If now the assumption is made that all present C nucleobases form G-C base pairs, there would still be a significant amount of free G nucleobases present in the mixture. These nucleobases might now form base pairs with the free T nucleobases. Depending on the environmental parameters such as temperature and solvent polarity, formation of these mismatches might still provide more stability to the nucleobases than remaining in their free state or base paired to an identical nucleobase. Therefore, when examining the environmental effect, concentration ratios should not be neglected as it might affect the occurrence of mismatching significantly.

2.3 Characterizing base pairing in synthetic nucleobase containing molecules

The main objective of this research is the characterization of base pair formation between synthetic nucleobase containing molecules. To achieve this objective, analysis techniques have to be found that are capable of analysing the formation of these hydrogen bonds or base pairs. The following sections will discuss the different methods that could be used for obtaining the desired characterization.

2.3.1 Proton nuclear magnetic resonance spectroscopy

Proton nuclear magnetic resonance spectroscopy or ^{1}H-NMR spectroscopy is a first method capable of detecting secondary or non-covalent interactions [66-68]. This means that it can provide information about intra- and intermolecular hydrogen bonding interactions.

^{1}H-NMR analysis is based on the magnetic properties of protons present in organic molecules. In principle there are two forms of ^{1}H-NMR spectroscopy, continuous wave and Fourier transform (FT) NMR spectroscopy [68]. FT NMR spectroscopy is exclusively used nowadays and will also be used in this research. The principle behind the measurement will be briefly discussed. Proton NMR starts with applying a magnetic field B_0 to the analysis mixture. This field causes orientation of the spins of ^{1}H into two possible energy states, the ground (α) and excited (β) energy states as shown in Figure 22. The energy difference between the two states is referred to as ΔE. The number of ^{1}H in the α-spin state is slightly higher than the number of ^{1}H in the β-spin state, due to the preferred lower energy of ^{1}H in the α-spin state. In presence of B_0, a radio frequency (RF) signal is emitted by the NMR device. This RF-signal can excite the ^{1}H from the α-spin state into the β-spin state if it contains the wright frequency or the wright amount of energy ΔE. When excited, the ^{1}H will return to the ground state due to relaxation. This relaxation causes the ^{1}H to emit energy under the form of a frequency signal [68, 69]. After Fourier transformation of this signal, a frequency signal is obtained. The main advantage of this FT NMR analysis is that it can excite and therefore analyse the complete NMR spectrum with a single pulse.

The link between the ^{1}H-NMR analysis and the structure of a molecule will now be elucidated. The ΔE of a proton, therefore also the frequency obtained by the Fourier transformation, is strongly dependent on the environment or electrons surrounding the respective proton. The movement of the electrons causes the induction of a magnetic field ($B_{ind.}$). When this induced magnetic field opposes the orientation of B_0 at ^{1}H, the field experienced by the ^{1}H will be reduced to $B_{eff.}$. This is called the shielding effect. Due to the reduced experienced field, the energy difference between the two energy states ΔE, will also be reduced. This results in a reduction in the frequency obtained from the FT ^{1}H-NMR analysis according to the well-known formula: $\Delta E = h \cdot \upsilon$. Differences in electron density of different ^{1}H in an organic molecule, cause the structure of this molecule to be revealed by analysing the spectrum containing all the resonance frequencies.

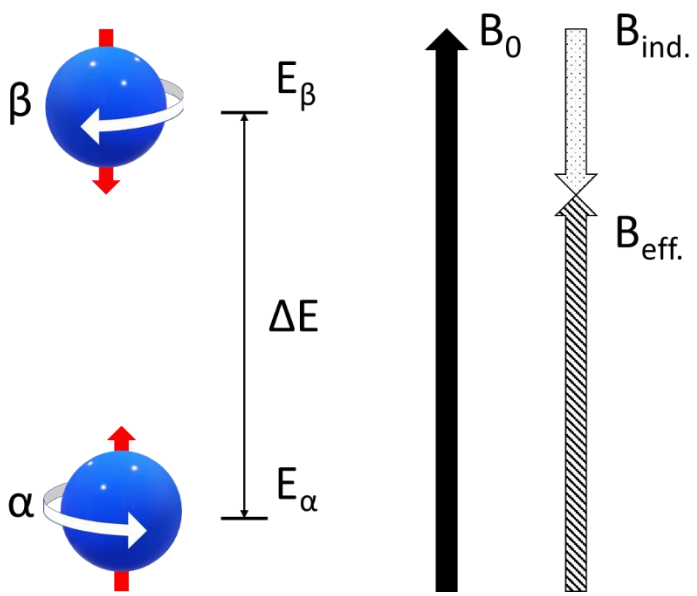


Figure 22. Principle of NMR-spectroscopy. The blue bowls represent the nuclei from ^{1}H-NMR and the white arrows represent their respective spins. The red arrows indicate the orientation of the field created due to the spins of the nuclei. E_α and E_β represent the energy states of the α and β spin-state respectively and ΔE represents the energy difference between the two spin states. B_0 represents the orientation of the external applied magnetic field.

In an NMR-spectrum it is not the resonance frequencies that are shown but the chemical shift. This chemical shift is described by the following formula:

$$\delta_{sample} = \frac{(\upsilon_{sample} - \upsilon_{reference})10^6}{\upsilon_{spectrometer}} \quad (6)$$

$\upsilon_{reference}$ is the frequency of the protons of the reference molecule. Tetramethylsilane is most often used as reference as its protons are more shielded than any other organic compound. This means that its δ is unlikely to coincide with the δ of a proton of the sample. $\upsilon_{spectrometer}$ is the frequency of the spectrometer and υ_{sample} is the frequency of the proton from the sample. The values of δ_{sample} are what is represented and examined in an NMR-spectrum.

2.3.1.1 ^{1}H-NMR and hydrogen bonding

The chemical shift (δ) of protons bonded to strongly electronegative atoms (e.g. nitrogen, oxygen, sulphur, etc.) is significantly affected by hydrogen bonding. As explained earlier, hydrogen bonding is a secondary interaction in which such a proton interacts with a second strong electronegative atom [39]. Due to the presence of an additional atom that draws electrons further away from the nucleus of the proton, referred to as the deshielding effect, the field experienced by this proton increases. The effect, resulting from the charge transfer and polarization, therefore causes an increase in the ΔE between the energy states [70]. This means that due to the deshielding effect of these hydrogen bonds, the resonance frequencies of these protons will shift upwards causing an increase in δ [67, 68, 71]. A ^{1}H-NMR spectrum concerning methanol, in which this deshielding effect is shown, is given in Figure 23.

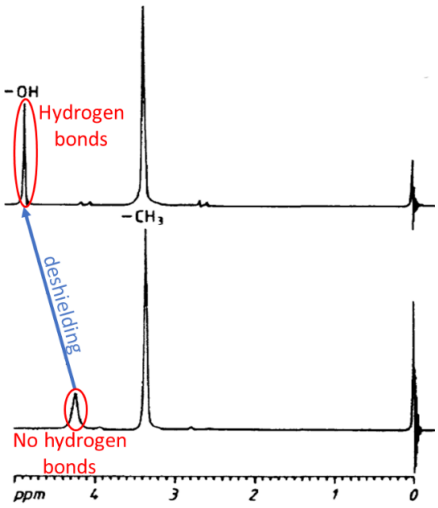


Figure 23. ^{1}H NMR spectra of methanol in the pure state (above) and diluted in tetrachloromethane solution (5%, below) [72]. Tetrachloromethane breaks down the hydrogen bridging meaning that the bottom spectrum represents a solution in which hydrogen bonding barely takes place. This breaking of hydrogen bonds is observed in the spectra as a decrease in the shift of the δ representing the proton on the hydroxyl group. This decrease in shift is a result of the increasing shielding effect due to absence of hydrogen bonds.

Due to the effect of hydrogen bonding on the δ of the proton taking part in this interaction, ^{1}H-NMR can be used to examine hydrogen bonding between nucleobases. Firstly, the ^{1}H-NMR spectrum of a solution containing only the respective nucleobase must be measured. By means of this measurement the δ of the proton, that could take part in hydrogen bonding, is determined when no hydrogen bonding with other nucleobases occurs. When the δ for the same proton is determined, now in presence of the complement of the respective nucleobase, hydrogen bonding between complementary nucleobases can be examined. For example, when no increase in δ for these protons is observed, absence of hydrogen bonding is assumed. When a downward shift is observed, occurrence of hydrogen bonding is observed due to the effect of deshielding on δ. The same experiments can be executed with nucleobases, other than the complement or mixtures, to examine selectivity.

The δ of these protons is strongly affected by solvent and temperature [68]. Therefore, all experiments of which data is compared should be executed in identical circumstances. Otherwise data comparison would not render valid conclusions. For the solvent, different interactions between the sample and the solvent can cause for a difference in observed δ of the reference sample. When DMSO is used for example, hydrogen bonding of the sample with the solvent can cause for an increasing δ of the proton participating in the interaction. When using the rather non-polar $CHCl_3$, polar nucleobases could be forced together by repulsive forces between the solvent and the sample. This could force the formation of base pairing between the identical nucleobases resulting in a different δ of the sample. Due to difference in strength of hydrogen bonds, the reference δ would not be the same and therefore the selected solvent is an important parameter to take into account. As temperature would also affect hydrogen bonding, also this parameter is important to take into account when measuring samples. Additionally, note that the parameters do not only affect the reference sample, but also the other samples. Also here, the same effects could cause for change in the measured δ.

2.3.1.2 Nuclear Overhauser Effect analysis

A second NMR-related analysis technique is based on the principle of the nuclear Overhauser effect (NOE). It is an important technique for the determination of distances between protons and is therefore key in structural analysis of molecules and interactions [68]. There are two forms of the NOE-analysis that will be handled in this literature study: 1D homonuclear NOE analysis (NOE-analysis) and 2D homonuclear NOE analysis (NOESY-analysis).

2.3.1.2.1 Nuclear Overhauser effect

Before explaining the NOE- and NOESY-analysis, an introduction to the NOE-effect is required. This will be done by means of the Solomon diagram for a two-spin system[III] consisting of two nuclear spins [68]. As mentioned in the beginning of 2.3.1, an outer magnetic field (B_0) is applied to the analysis mixture. This will cause a division of nuclei from protons over various energy levels (1 to 4) due to orientation of their spins, as displayed in Figure 24.

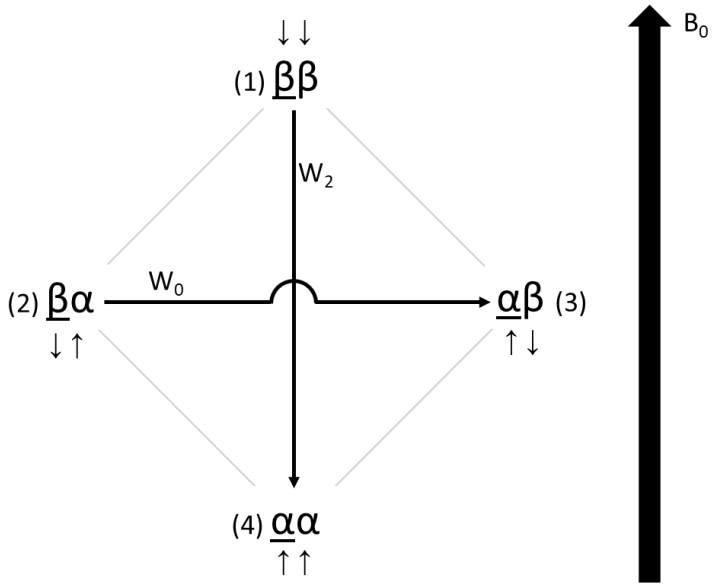


Figure 24. Solomon diagram for a two-spin consisting of two nuclear spins. The underlined state indicates the source spin or nucleus, while the other state indicates the interesting spin or nucleus.

A small difference in energy between the different spin states, causes for a slight difference in their population to occur. This means that more nuclei populate the ground spin state (α) in comparison to the excited or upper energy state (β). This distribution of nuclei is referred to as the Boltzmann distribution. When the nucleus of a source proton is irradiated by means of an RF-signal, at the resonance frequency of this proton, a redistribution of the population takes place resulting in an increased population of the upper spin state. At a certain point saturation will be reached, meaning that both upper and ground spin states contain equal populations of the source protons. Once the RF-pulse ends, the source nuclei will return to the equilibrium or Boltzmann distribution by means of relaxation to the ground state.

[III] Here, the term two-spin system is used to indicate a system consisting of two spin-active nuclei (of protons), both containing a spin that is coupled trough their magnetic properties the other nucleotide by trough space interactions.

In Figure 24, zero order quantum transition (W_0) and double quantum transition (W_2) cross relaxation are shown. W_0 is the phenomenon where two anti-parallel spins flip simultaneously due to dipolar spin-spin coupling (DSSC) between the interesting and source nucleus. For the example depicted in Figure 24, the transition from state (2) to (3) represents the W_0-transition. This transition causes the nucleus of the interesting proton to transfer into the upper spin state. For the interesting proton, it means a deviation from the Boltzmann distribution is introduced. More specifically, an increase in population of the upper spin state or a decrease of the population difference between the upper and ground spin state will occur. The decreased population difference between the spin states results in a lowered intensity of the signal in the NMR-measurement and is referred to as the negative NOE-effect. For W_2, a similar but reversed effect is observed. Here, two aligned spins flip simultaneously meaning that, for the example depicted in Figure 24, a transition from state (1) to (4) occurs. W_2 therefore causes the nucleus of the interesting proton to transfer into the ground spin state, increasing the population at the ground state, therefore also increasing the population difference, resulting in an amplified NMR-signal. This phenomenon is referred to as the positive NOE-effect. W_0 and W_2 are the only relaxation effects that are elucidated and depicted as they are the only ones that contribute to the NOE-effect. It is important to notice that both transitions are a result of DSSC between the source and interesting nuclei and not by scalar spin-spin coupling between nuclei.

There are two main characteristics of this nuclear DSSC that enable the NOE-effect to be used in analysis of molecular structures and interactions. A first one is the distance restriction of the effect. The effect of DSSC between two protons is limited to an interatomic distance of approximately 6 Å [76]. It means that a potential positive or negative NOE-effect of the source proton on an interesting proton, when the distance between the two is more than 6 Å, is not usable or measurable by means of NOE- or NOESY-analysis. This enables determination of the protons in close proximity to the source protons, by means of NMR-analysis, as only the near protons will be affected by this source proton. A second important characteristic is the fact that DSSC is a trough space effect, meaning that also protons, not covalently attached to the same molecule, can be affected by the source proton. Therefore, secondary intermolecular interactions can also be examined by means of NOE- and NOESY-analysis. G-C base pairing can be used as an example. When the base pair is formed, the two molecules are in close proximity to each other. Especially the three protons participating in hydrogen bonding, during this base pairing, are in very close proximity to each. When a G-C mixture is analysed by means of NOE- and/or NOESY-analysis and base pairs are formed, one should clearly see the DSSC or NOE-effect between these protons.

2.3.1.2.2 1D NOE-analysis

For 1D NOE-analysis, the NMR-spectrum is measured with and without irradiation of the interesting proton [68]. The spectrum without this irradiation is subtracted from the spectrum with irradiation of this interesting proton resulting in a spectrum displaying solely the NOE-effect. The irradiated proton and protons for which $W_0 > W_2$ are observed as negative signals in the NMR-spectrum. The protons for which $W_0 < W_2$ are observed as positive signals. By means of this spectrum the protons close to the interesting proton can be uncovered.

2.3.1.2.3 2D NOESY-analysis

The 2D technique provides similar info compared to the 1D technique. The measuring techniques are however very different. 2D NOESY-analysis does not rely on the irradiation of specific spins or nuclei. Instead it uses the following pulse sequence [71]:

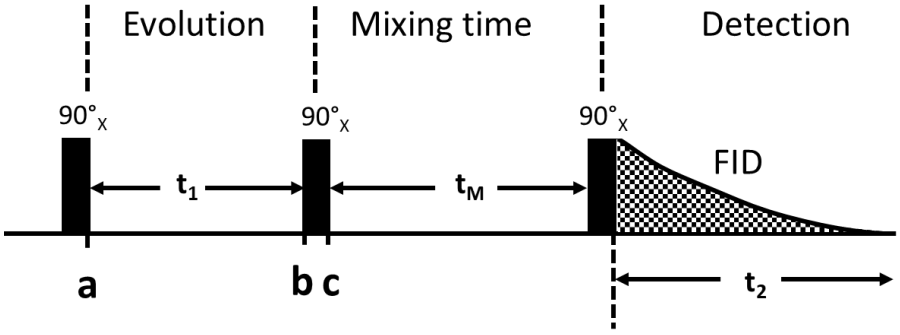


Figure 25. Pulse sequence for 2D NOESY-NMR.

A first pulse ($90°_x$) causes for a transverse magnetisation of the nuclei of all protons present in the mixture [71]. During evolution time (t_1) these magnetisations will develop accordingly to their individual Larmor frequencies[IV] [68, 77]. This causes each proton to be labelled with their respective Larmor frequency at the end of t_1. Subsequently, a second pulse ($90°_x$) is introduced causing for a longitudinal orientation of the magnetization relative to the external field B_0. During the mixing time (t_M), after the second pulse, DSSC will cause the magnetization transfer or NOE-effect to take place. Finally, a third pulse ($90°_x$) is introduced causing again a transverse magnetization of the nuclei. The resulting free induction decay (FID) signal after this last pulse contains the frequency-related information concerning the individual protons as well as the influences of the NOE-effect. It is also this signal that will be captured and processed to obtain the 2D NOE-spectrum. It is important to note that the amplitude of the last magnetisation is dependent on the duration of t_1 and the efficiency of the magnetisation transfer. This means that the NOE-effect is related to the duration of t_1 trough the FID-signal observed in t_2.

In order to obtain the 2D NOESY-spectrum, previous described process (Figure 25) is executed numerous times whilst incrementing the duration of t_1 after each measurement. This multitude of experiments can be seen as a data matrix with t_1 and t_2 as its two time axes. After two Fourier transformations, one with respect to t_1 and one with respect to t_2, the 2D NOESY-spectrum can be obtained. This spectrum has two frequency domains F_1 and F_2. A frequency domain F_n contains frequencies of mechanisms present during the corresponding t_n. This mean that F_1 contains the DSSC influences as the increments on t_1 carry the info concerning this coupling. F_2 then contains info concerning the chemical shift. Both axes represent the data in terms of δ [ppm].

[IV] In presence of an external magnetic field (B_0) nuclei can, in addition to alignment with this magnetic moment, produce a secondary spin or precession around this field. The Larmor frequency refers to this precession. This Larmor frequency is related to the strength of B_0 and characteristic for each specific nucleus.

In Figure 26 an example of a NOESY-spectrum is given. The signals on the red diagonal, seen as black marks, are called cross peaks and form the 1D ^{1}H-NMR spectrum. The off-diagonal points, seen as grey marks render information about the DSSC. A possible procedure, used to determine which protons are in close proximity to the proton of interest (H_a), starts with following the vertical dashed line of the δ from proton (H_a) to the diagonal. From this diagonal a horizontal line should be drawn (blue). At each point where the horizontal line crosses a grey mark/signal, a vertical line should be drawn upwards (dashed blue). The protons corresponding to the signals where these lines end, are the protons close enough to the interesting proton (H^a) to be affected by its NOE-effect during NOESY-NMR analysis.

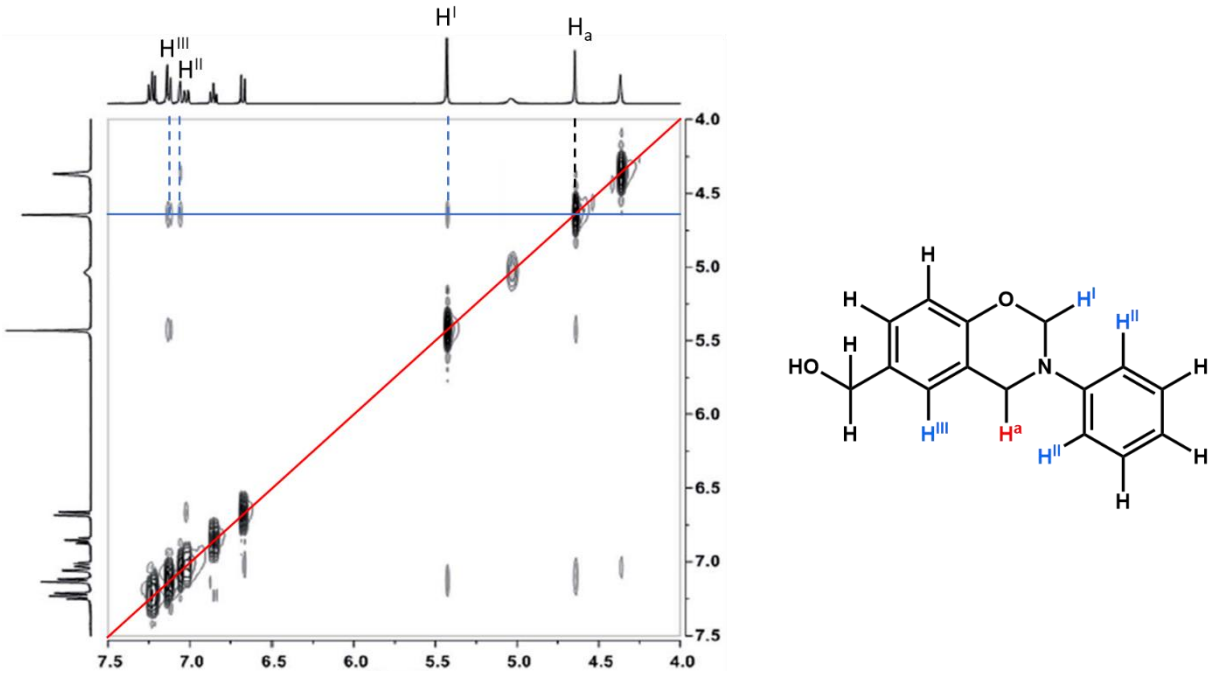


Figure 26. 2D ^{1}H-^{1}H NOESY NMR spectra of (3-phenyl-3,4-dihydro-2H-benzo[e][1,3]oxazin-6-yl)methanol [78].

2.3.1.3 Diffusion ordered spectroscopy

The final NMR-based analysis technique is diffusion ordered spectroscopy (DOSY). More specifically, it is an NMR-based diffusion analysis technique. It relies on the use of field gradients to determine physical location of a molecule or complex in a sample. By means of this type of measurement, the diffusion along the direction of the applied field gradient is characterized [68, 73]. The most basic form of DOSY is based on spin-echo [74]. Based on this method the principle of DOSY will be further explained.

2.3.1.3.1 Principle

As mentioned before, DOSY employs field gradients. The principle of DOSY lays in the use of two opposing field gradients [74]. A first field gradient imposes a phase on the magnetisation vectors dependant on the physical location of the molecule. A second field gradient is introduced afterwards. This field gradient has the same duration and magnitude but a 180° phase shift in comparison with first one. When the time between these field gradients is zero, the second one will realign the magnetization to the state it had before introduction of the first field gradient. However, in DOSY, a time interval between the two field gradients is introduced and is referred to as Δ. This gives molecules the time to diffuse inside the volume. Note that magnetisation of the spins is dependent on the physical location of the molecule in the examined volume. Therefore, complete realignment due to the second field gradient will not take place. This results in an incomplete recovery of magnetization of the molecule and an NMR-signal that is dependent on the physical location and thus the diffusion of the molecule. The principle of this measurement is given in Figure 27 [74].

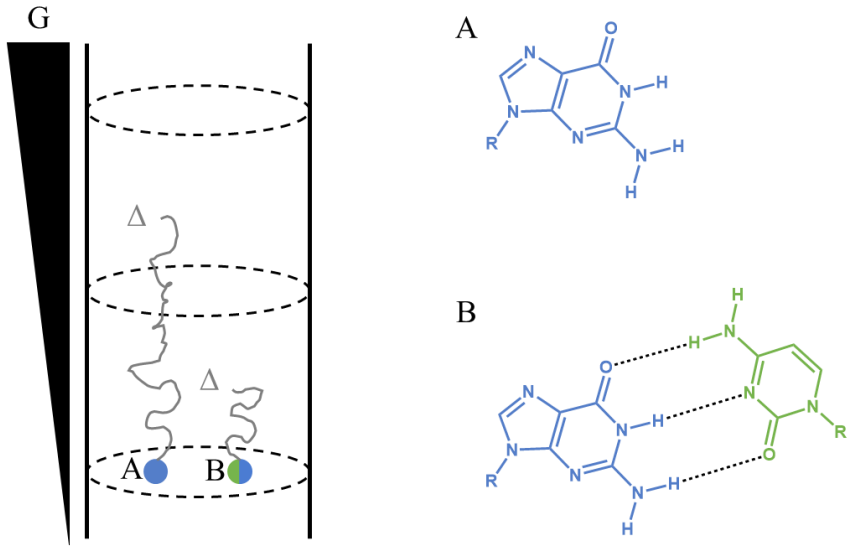


Figure 27. Schematic representation of a hypothetical DOSY-experiment with a GAM (A) and a base pair consisting of CAM and GAM (B). R in the molecule represents the acrylate containing part of the monomer. G represents the applied field gradient. The first field gradient is applied when A and B are in the same position. The second field gradient is applied a specific amount of time (Δ) after the first field gradient, in which A and B have moved to the end of their path (grey line). Because B has a larger hydrodynamic volume, it diffuses slower and will have travelled a smaller distance than A. After Δ, A will perceive another field than B and therefore the signal detected from A will differ to that from B.

To characterize diffusion rates, most often the strength of the gradient pulses G is progressively increased or decreased and the corresponding change in NMR signal intensity I is measured [73]. This results in a multitude of 1D ^{1}H NMR experiments as displayed in Figure 28. Plotting the maximum of each peak, at the same δ, in an I vs. G graph results in decay curves as displayed in Figure 29. Note that one peak should correspond to the proton(s) of a single molecule.

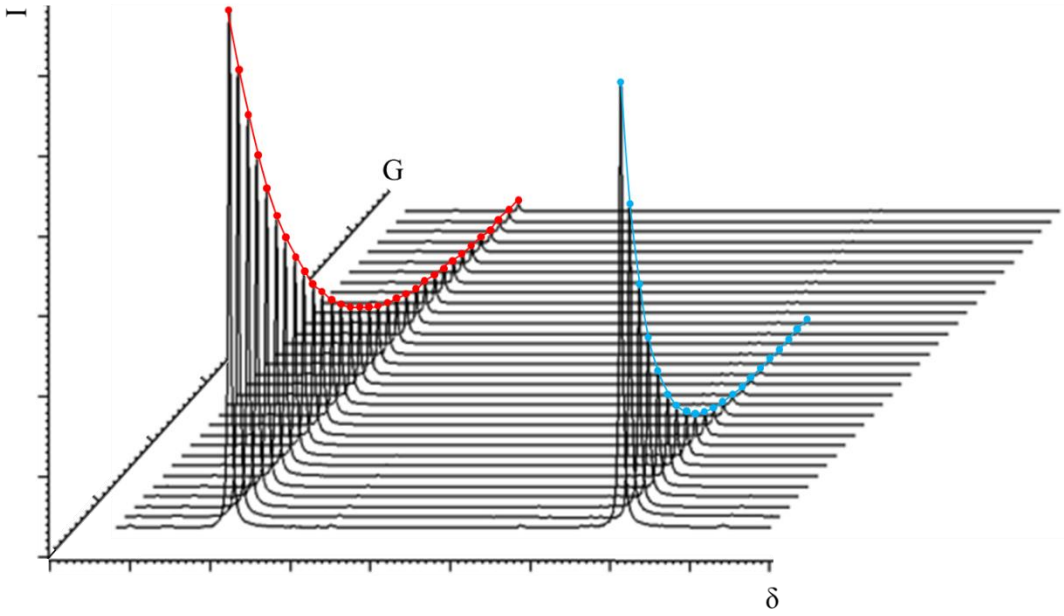


Figure 28. Example of a multitude of standard 1D ^{1}H-NMR experiments (I vs. δ graph) at different field gradient strengths (G) [75].

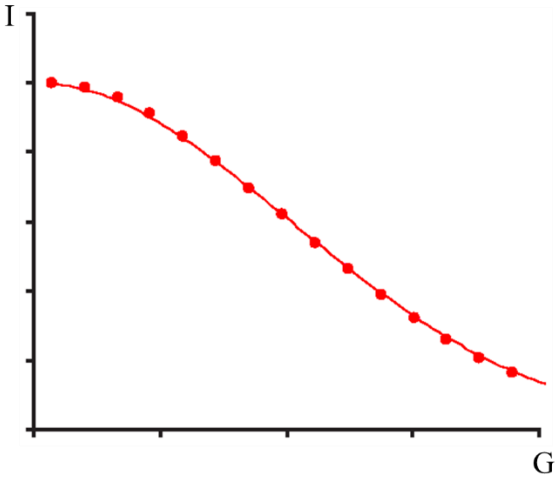


Figure 29. Regression analysis of the I(G)-curve for the determination of the diffusion coefficient [74]. Note that this curve represents the intensity curve of the red line from Figure 28. These are merely sketches and therefore not exact matches.

From these last curves, the diffusion constant can be derived using regression analysis according to the following equation [74]:

$$I = I_{G=0} \exp\left(-(\gamma\delta G)^2 D\left(\Delta - \frac{\delta}{3}\right)\right) \quad (7)$$

Now the signal intensity in absence of the gradient spin-echo ($I_{G=0}$), the length of the gradient pulse (δ), the magnetogyric ratio of the observed nuclide (γ) and the diffusion period(Δ) are known constants and the diffusion coefficient and other data can be determined.

As stated by the Stokes-Einstein equation for a sphere, diffusion is directly related to the size of the diffusing molecule according to:

$$D = \frac{k_b\, T}{6\, \pi\, \eta\, r_s} \quad (8)$$

Here k_b is the Boltzmann constant, T is the absolute temperature, η is the solution viscosity and r_s is the radius of the diffusing molecule [74]. Therefore, due to dependency of diffusion on molecule radius and dependency of NMR-signal on diffusion, molecular size can be examined with DOSY-NMR. Small molecules for example diffuse quickly, meaning that recovered magnetization due to the second field gradient will decrease.

The result of DOSY-NMR is a 2D-spectra with chemical shift on one axis (usually x-axis) and diffusion coefficient on the other axis (thus usually the y-axis) as shown in Figure 30. Due to this 2D-spectrum, different molecules/complexes can be examined simultaneously in one mixture.

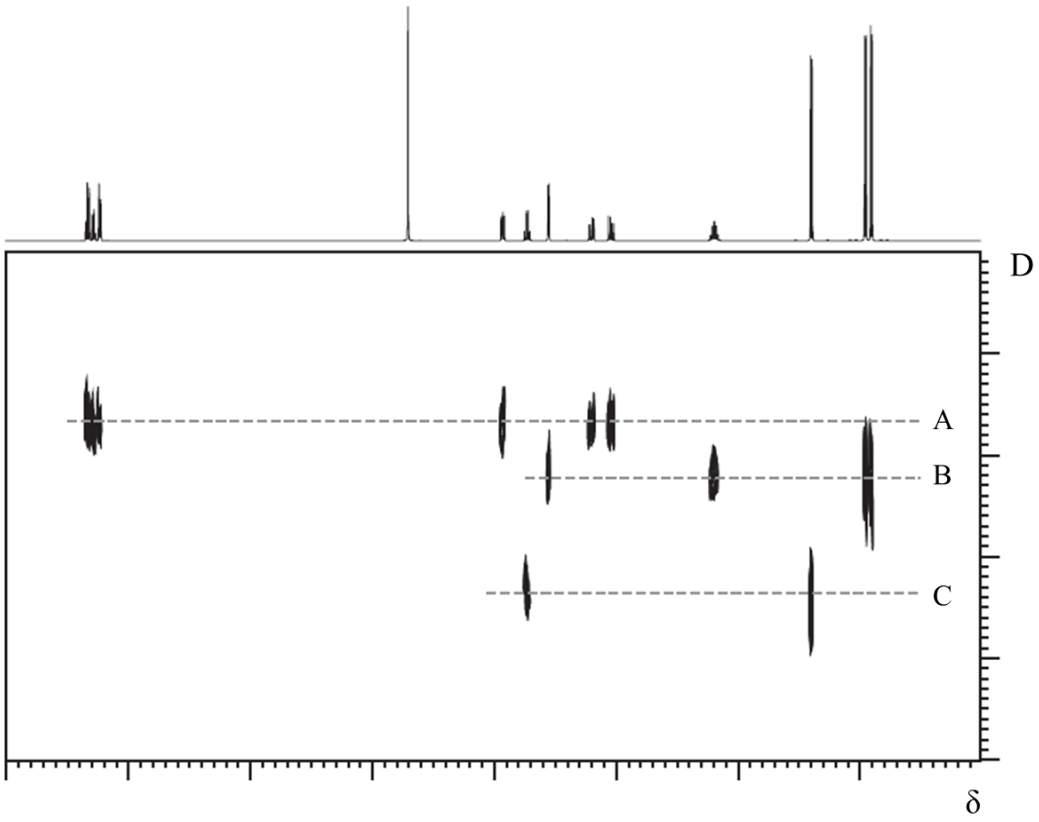


Figure 30. Illustration of a 2D DOSY spectrum in which separation of resonances from a mixture of three molecules (A, B and C) is observed [74].

2.3.1.3.2 Diffusion ordered spectroscopy as applied to nucleobase containing molecules

DOSY-NMR can be used for analysis of numerous phenomena and properties (e.g. molecular size, molecular weight distribution, hydrogen bonding, signal suppression, host guest complexes, etc.) [74]. As mentioned before, hydrogen bonding is of main interest in this research. Hydrogen bonding between two molecules can hypothetically be examined accordingly to the following reasoning. Three monomers (GAM, CAM and AAM) can be diluted into the same solvent, but in separate samples. DOSY-NMR on each sample will then give the diffusion coefficients for all three monomers individually (D_{GAM}, D_{CAM} and D_{AAM}). In a following step, the same experiment is executed, but now with all three monomers present in the same sample in equimolar concentrations. In this mixture the two monomers, containing complementary nucleobases, should form intermolecular hydrogen bonds resulting in a complex. This can be confirmed by DOSY-NMR. If the complementary molecules show a new shared diffusion coefficient (D_{GAM} and D_{CAM} for which $D_{GAM} \approx D_{CAM}$) it indicates that the monomers diffuse together. The diffusion coefficient of AAM should not (or barely) change compared to the individual coefficient. If, for example CAM and AAM, now show a new and similar diffusion coefficient, mismatching occurs in this mixture between CAM and AAM. The principle of this analysis is shown in Figure 31. Note that not all nucleobases will be present in a base pair complex as this complexation entails an equilibrium. The diffusion constants as observed when the nucleobases were analysed separately, will most likely still be observed in the resulting spectrum.

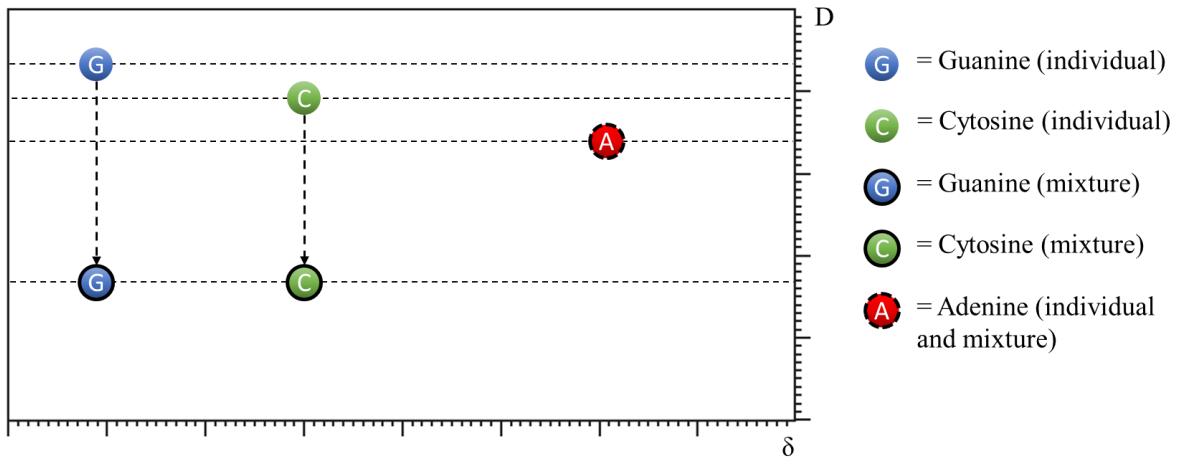


Figure 31. Principle of base pairing and selectivity analysis with DOSY-NMR. The circles without a black edge indicate a signal of the respective monomer individually. The circles with a black edge indicate the signal of the respective monomer in a mixture of the three monomers. A dashed black edge indicates that the diffusion coefficient for this monomer remains the same in mixture as individually.

2.3.2 Surface plasmon resonance analysis

A last method that will be discussed, capable of analysing base pairing, is surface plasmon resonance (SPR). This is an analysis method frequently used to characterize biomolecular interactions [79, 80]. Furthermore, SPR has already been used to examine DNA, nucleobase interactions and mismatching [80-82]. In this research, the analysis technique would be used differently as it would have to give information about affinity, binding kinetics and selectivity of base pairing between free NAM or short MHB-SD oligomers.

2.3.2.1 Principle

A fundamental concept of SPR-analysis is the fact that an electrical field intensity, or evanescent wave field (EWF), is created when total internal reflection of a light beam occurs at an interface [80]. If this EWF interacts with a metal, the p-polarized component of the EWF can penetrate the metal layer as displayed in Figure 32 [83].

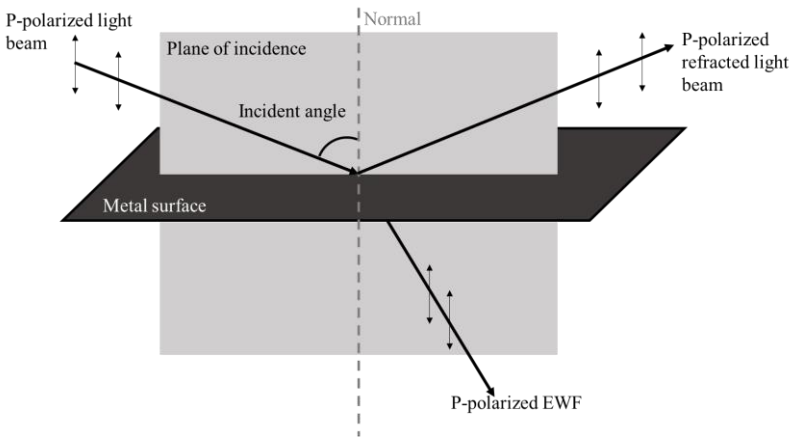


Figure 32. P-polarized EWF penetrating a metal surface (e.g. gold, silver or aluminium).

This can create electromagnetic waves propagating within the metal surface at the interface with the sample solution [80]. These waves are called surface plasmons. Note that the plasmons are bound to the metal layer in the plane of the interface. The EWF on the other hand penetrates (± 700 nm) into the analyte medium. Furthermore, formation of the surface plasmons only occurs if momentum and energy of the surface plasmons equal that of the incident light vector in the plane of the metal, as shown in Figure 33. This is called the resonance condition. For a light beam to correspond to this resonance condition, the wavelength or the angle of the incident light beam can be adjusted. The energy from incident light, corresponding to the resonance condition, is then transferred to surface plasmons. This results in reduced intensity of the reflected light beam at this incident angle or wavelength as shown in Figure 33.

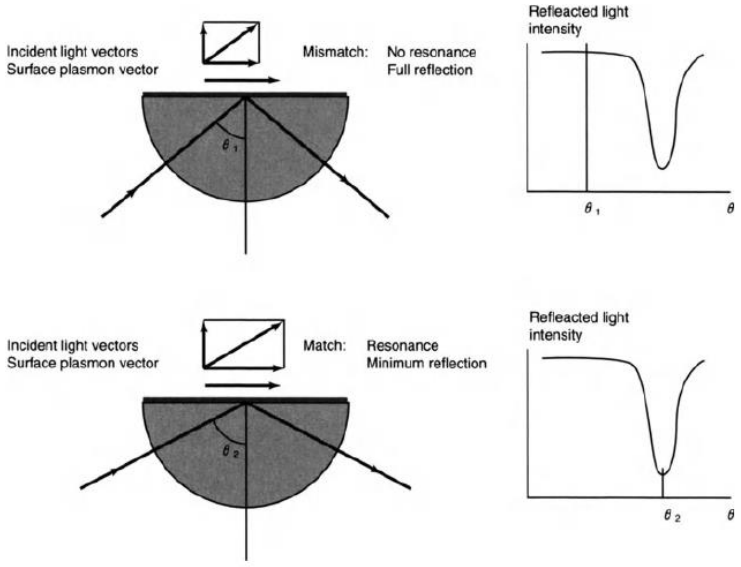


Figure 33. A visualisation of plasmon resonance [80]. The momentum of the incident light photons consists of two vectors parallel and perpendicular to the surface. Resonance takes place when the momentum of surface plasmons, within the metal film, matches the component of the incident light photons parallel to the surface.

For the metal substrate there are only three options concerning the material. It has to be gold, silver or aluminium. They are the only metals generating surface plasmons under resonance conditions using light in the visible or near IR spectrum [80].

As mentioned before, plasmons are confined to the metal. However, their propagation energy and momentum depend partially on the interaction between the EWF and the refractive index of the analyte medium [80]. In this property lays the key to the functioning of SRP. When the refractive index of the medium changes, the incident angle or wavelength to obtain resonance conditions changes. In following paragraphs, the link between the refractive index, incident angle and non-covalent interactions will be elucidated.

2.3.2.2 Example of SPR-analysis

As the principle of SPR has been elucidated above, the SPR-analysis will now be handled. This method analyses the non-covalent interactions between molecules. One of the interacting molecules required to be attached onto the metal substrate. Assume GAM will be attached onto the metal substrate by means of a coupling arm, called the ligand. One side of this ligand should be capable of binding to the metal substrate and on the other side to the monomer. This metal, with monomer attached onto it via the ligand, will now encounter a small current of solvent passing the functionalized surface as displayed in Figure 34. This creates the initial environment on the analyt-side of the substrate. This environment has a certain refractive index corresponding to a certain incident angle that meets the resonance condition as shown in graph A in Figure 34. The intensity measured at this angle will now be set as the baseline for the measurement as shown in graph B in Figure 34.

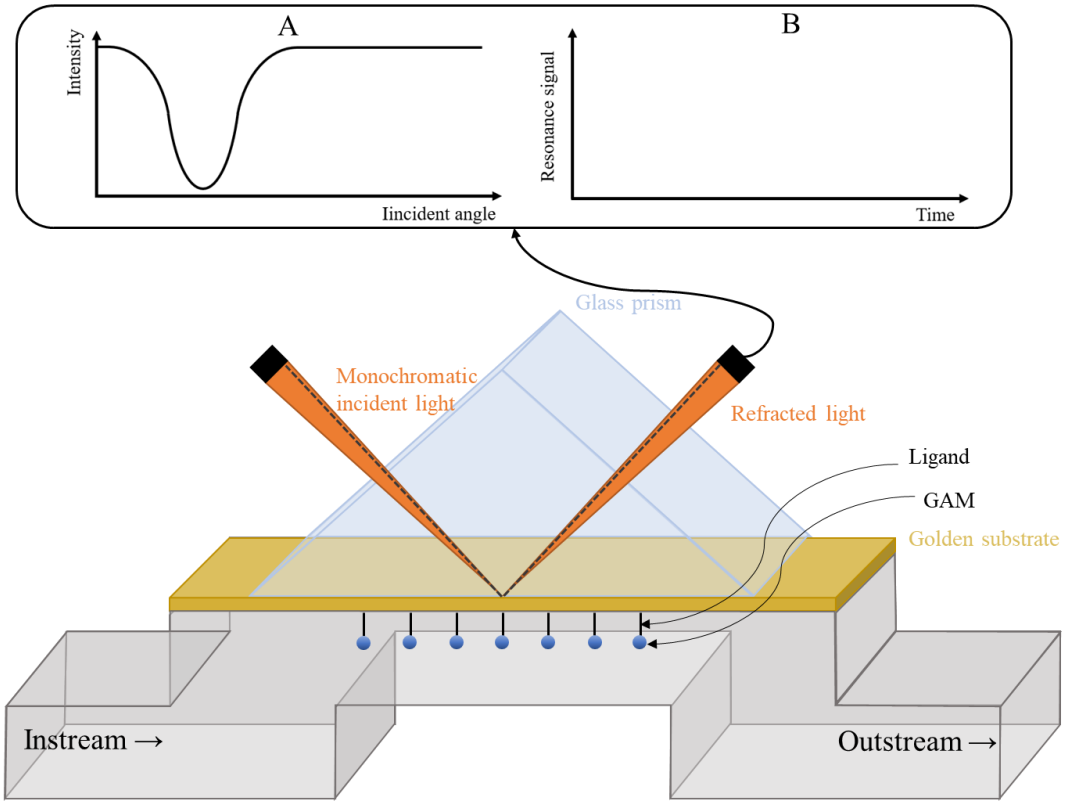


Figure 34. SPR-analysis standard set-up without product stream. The dashed dark grey line in the refracted light indicates the refracted beam showing reduced intensity due to surface plasmon resonance.

In the following step, another monomer will be added to the solvent stream. Assume that the monomer in the solvent stream is CAM. When this monomer passes the substrate, with AAM attached onto it, CAM will form base pairs with GAM. This causes for a change in the refractive index in the environment close to the gold substrate to occur. Therefore, the incident angle that corresponds to the resonance condition will change (graph A of Figure 35) and the intensity at the current incident angle will increase (graph B of Figure 35). This increase in intensity will now be measured and therefore implies the occurrence of base pair formation. This is shown in Figure 35.

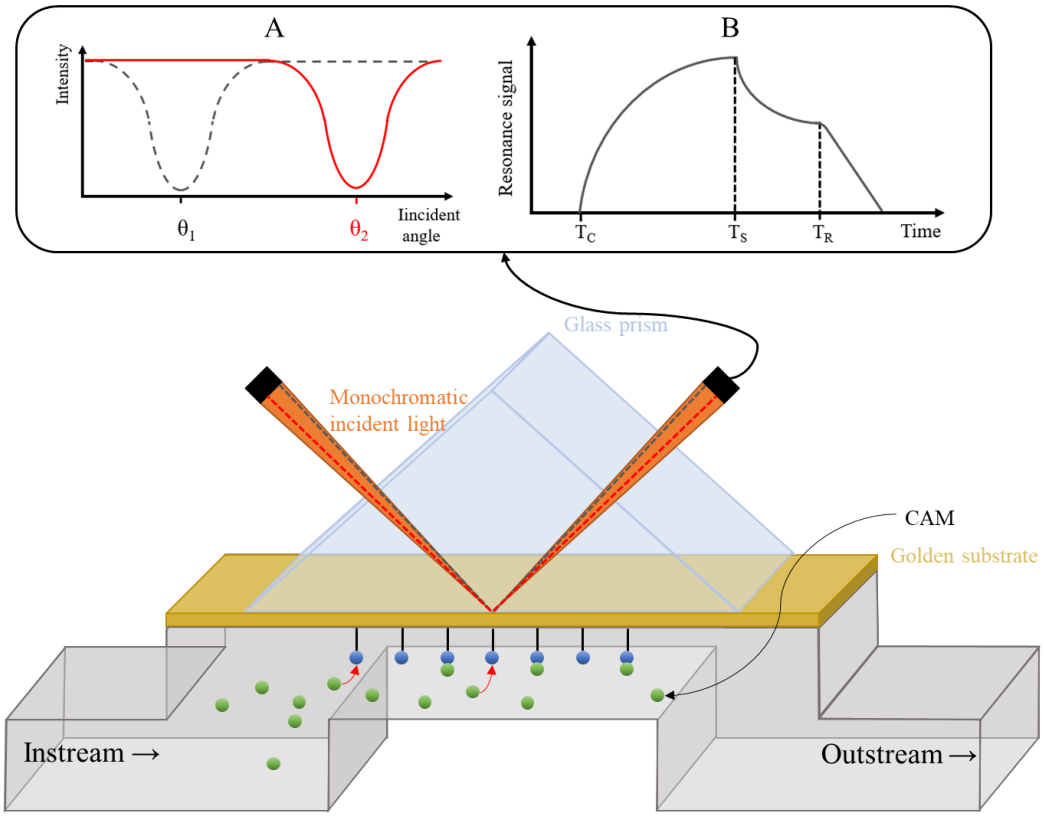


Figure 35. SPR-analysis standard set-up with product stream. The grey dashed line indicates the previous light beam for which incident angle θ_1 corresponded to the resonance conditions. The red dashed line indicates the new monochromatic light beam, having a different incident angle θ_2, that corresponds to the resonance conditions. These incident angles are shown in graph A. Graph B shows the signal during the experiment. T_C indicates the moment from which CAM is added to the stream. T_S indicates the moment from which again pure solvent is put through the system. T_R indicates the moment from which regeneration takes place, meaning that conditions will be set as such that base pairs will be broken.

The selectivity can now also be examined by repeating the same experiment for the other monomers. However, for non-complementary monomers no signal or a weak signal, compared to analysis between complementary nucleobases, should be measured when these non-complementary nucleobases are added to the solvent stream. Occurrence of a strong signal would imply the formation of fairly stable mismatches and thus poor selectivity in the current environment. Also, instead of using merely monomers, the same experiment can also be executed for MHB-SD sequences.

As mentioned, SPR-analysis can be used to examine selectivity and presence of base pairing. What has not been mentioned, is that SPR-analysis also renders information concerning the binding characteristics of the base pairs. Meaning that an equilibrium between formation and separation of base pairs, as well as association and dissociation rates and conditions can be examined.

A steeper slope of the curve between T_C and T_S indicates faster formation of base pairs between the two respective nucleobases. This implies the presence of a quickly reached equilibrium promoting base pair formation. A steeper slope of the curve between T_S and T_R or the end of the experiment on the contrary, indicates the presence of an equilibrium preferring no formation

of base pairs or indicates that the base pairs are weak and therefore break easily. This is shown in Figure 36. It is desired that the first slope is steep, meaning that base pairs are formed quickly. The second slope should be rather flat, indicating the formation of strong base pairs. This would render the best results in templated polymerization and affinity separation.

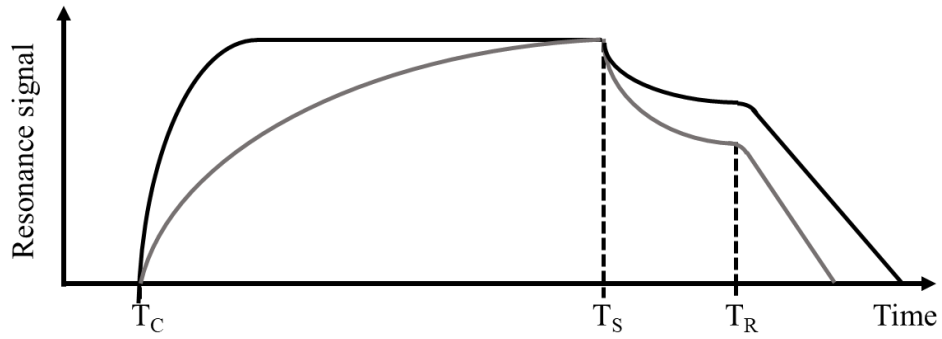


Figure 36. Signal intensity vs. time graph of a SPR-analysis. The black curve implies strong tendency towards complex formation. The grey curve implies strong tendency towards free molecules.

2.3.2.3 Preparation of substrate containing MHB-SD or NAM

To execute an SPR-analysis on MHB-SD or NAM, a method must be found to attach the molecules onto the gold substrate. A first method is shown in Figure 37. In the first step an end group modification of the trithiocarbonate functional group should be performed [84]. The trithiocarbonate functional group is converted into a thiol functional group by means of aminolysis [84]. Afterwards a simple coupling of the sulfhydryl functional group onto the gold substrate can be executed.

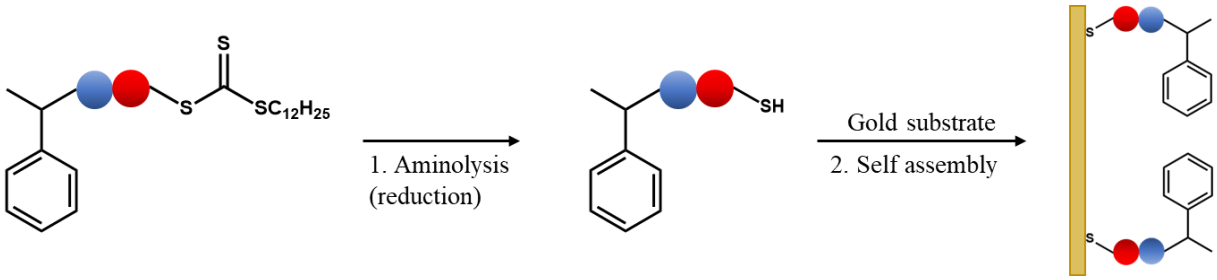


Figure 37. Process for functionalizing gold substrate with MHB-SD via direct self-assembly of the thiol functionalized MHB-SD onto the gold substrate. The blue and red circle represent two different NAM.

Self-assembly of a thiol-functionalized structure is a commonly used technique for creating a monolayer onto gold substrates in for example SPR-analysis [80, 85, 86]. There is however one major concern for the use of this process, described in Figure 37. The method causes the nucleobases to be in close proximity to the gold substrate. Due to steric hindrance, this could prevent the free MHB-SD from obtaining the right orientation towards the bonded MHB-SD for base pairing to occur. This means that a spacer arm between the gold substrate and the MHB-SD could be needed to create the necessary distance between the MHB-SD and the gold substrate. A process for coupling the MHB-SD onto the gold substrate, by means of a spacer arm, is given in Figure 38.

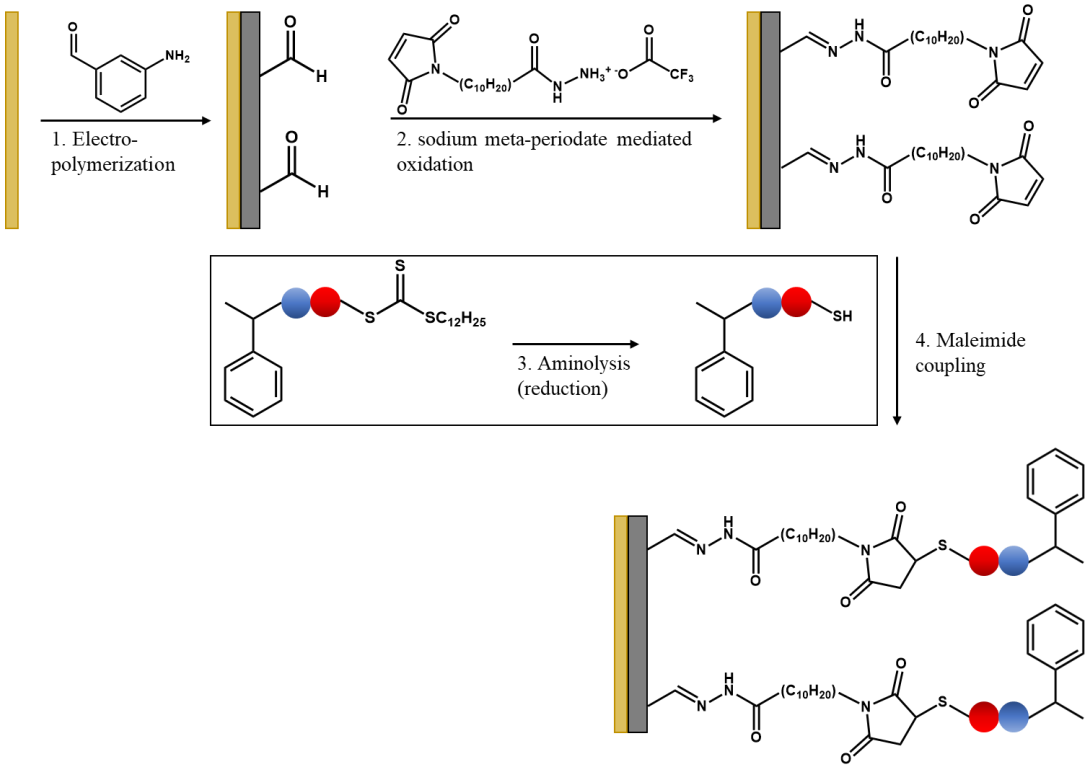


Figure 38. Process for functionalizing gold substrate with MHB-SD by means of a spacer arm. The blue and red circle represent two different NAM.

In the first step an electro-polymerization with 3-amino benzaldehyde is executed. This causes the formation of a poly(3-aminobenzaldehyde) (PABH) film [87]. In step two, an oxidation reaction between the aldehyde group on this film and the spacer arm N-[ß-maleimidopropionic acid] hydrazide (BMPH) takes place [88]. This results in a ligand on the gold substrate containing a sulfhydryl reactive maleimide functional group. In order to place the MHB-SD onto the surface, an end group modification of the trithiocarbonate functional group should be performed [84]. The trithiocarbonate functional group is converted into a sulfhydryl functional group by means of aminolysis. Finally the coupling reaction can be executed in which the sulfhydryl of the MHB-SD reacts with the maleimide functional group on the ligand [88].

Ensuring the presence of enough space between the substrate and the MHB-SD oligomers could also be obtained by another method. Instead of using a spacer arm, additional monomer insertions by means of RAFT polymerization could also be used. This method would require even fewer steps than the process described in Figure 38. Note that the monomers used to create this spacing will not be NAM. Preferably, the additional inserted monomers will be small acrylate molecules, such as methyl or ethyl acrylate. As stated before, the CTA and monomers are chosen as such, to obtain a controlled polymerization. When strongly deviating structures from acrylates are chosen, polymerization problems could occur. By tuning the CTA/monomer-ratio, the proper amount of monomer insertions can be optimized to obtain the desired spacing. An example of this process is depicted in Figure 39.

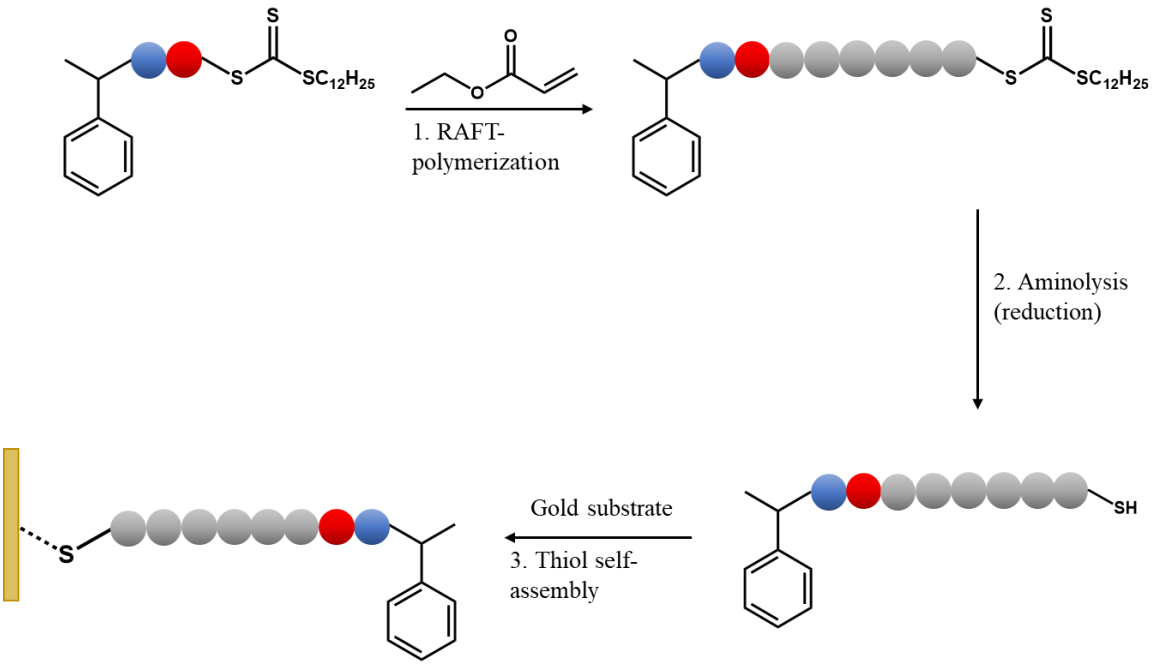


Figure 39. Insertion of small acrylate monomers into MHB-SD, creating a 'spacer arm' between the nucleobases and the gold substrate.

A process for coupling the NAM onto the gold substrate is described in Figure 40. In the first step, a NAM is coupled to benzene-1,4-dithiol (BDT) by means of a Michael-addition. Note that two monomers could be coupled onto each other in this process when both react with the same BDT-molecule. Therefore, it is important for the molar BDT/monomer-ratio to be two or higher, in the reaction mixture. In the final step self-assembly will takes place were the free sulfhydryl functional group of the molecule will be adsorbed onto the gold substrate [79, 80, 89, 90].

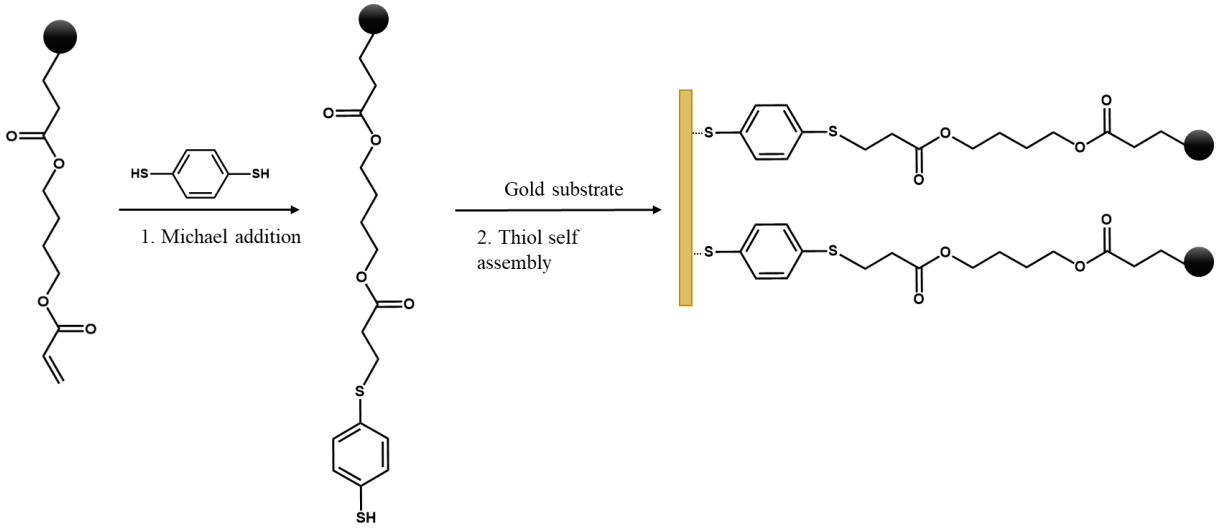


Figure 40. Process for functionalizing gold substrate with NAM.

2.4 Yield improving techniques

When research concerning the selectivity of nucleobases in MHB-SD oligomers and NAM has been executed, the results can be implemented in RAFT polymerization processes producing these MHB-SD oligomers. There are two main pathways possible for obtaining higher yields than the current SUMI-procedure. Both are theoretically elucidated in 1.2. Here it was mentioned that both methods require the use of particles. Onto these particles, MHB-SD oligomers containing the complementary sequence of nucleobases should be covalently attached. In affinity separation this particle is used for isolation of the desired counterpart. In template-assisted polymerization, this particle is used as template and afterwards ensures an easy isolation of the formed MHB-SD oligomer. As stated before, if both methods work perfectly, the template assisted polymerization would be the preferred process. This method would not show the problem of a small dispersity present in each polymerization step causing for a reduction in the yield in desired product.

For the particles, a choice must be made between polystyrene beads (PSB) or silica particles. Since the nucleobase containing oligomers might have bio-related applications, polystyrene beads are preferred. It is known that silica-particles contain certain health risks dependant on the size and crystallinity [91, 92]. Larger and amorphous silica-nanoparticles show fewer to no health hazards. Even though not all silica-particles show toxicity towards organisms, no health hazards have been reported for polystyrene beads. These beads are considered generally non-toxic and will therefore be favoured in this application [93].

Two methods for coupling the MHB-SD onto amine functionalized PSB were considered. A first method employed the copper(I)-catalysed alkyne-azide cycloaddition (CuAAC) click reaction [94]. This process is described in Figure 41.

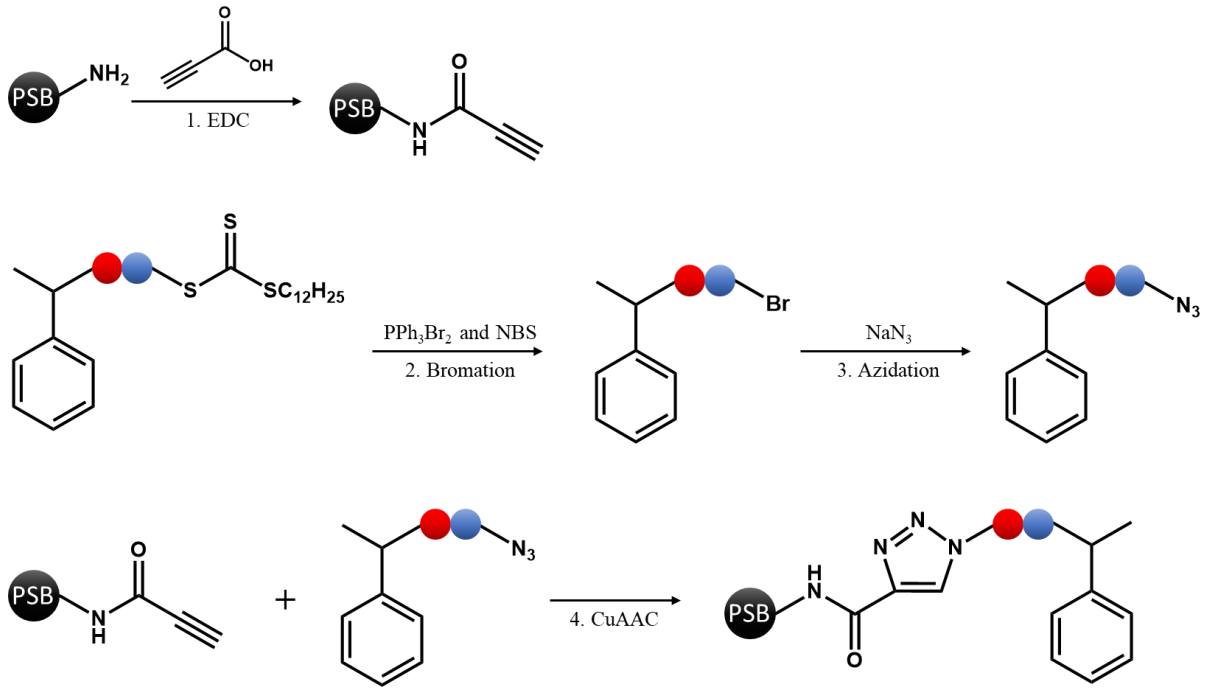


Figure 41. Process for coupling MHB-SD onto an amine functionalized PSB via a CuAAC click reaction. In step one, EDC coupling between the amino functionalized PSB is executed [95]. The trithiocarbonate functional group of the MHB-SD is converted into an azide functional group trough bromation followed by azidation [84, 96]. Finally the azide functionalized MHB-SD is coupled onto the alkyne functionalized PSB via a CuAAC click reaction [97].

An important downside of the procedure mentioned above is the use of copper(I). Copper can cause severe health consequences called copper toxicosis [98]. Therefore, the procedure is not fitting for bio-applications. Additionally, copper can cause complexation of nucleobases, making it hard to remove completely [99]. Due to this complexation, copper could also interfere with the desired base pairing of the nucleobases implemented in MHB-SD.

The process for coupling the MHB-SD onto the PSB, by means of a cross linker or spacer arm, is shown in Figure 42. This process has two benefits in comparison to the previously mentioned procedure. The first one is that it does not require the use of Cu(I) in its process. The second advantage is that this process requires one less step in its process.

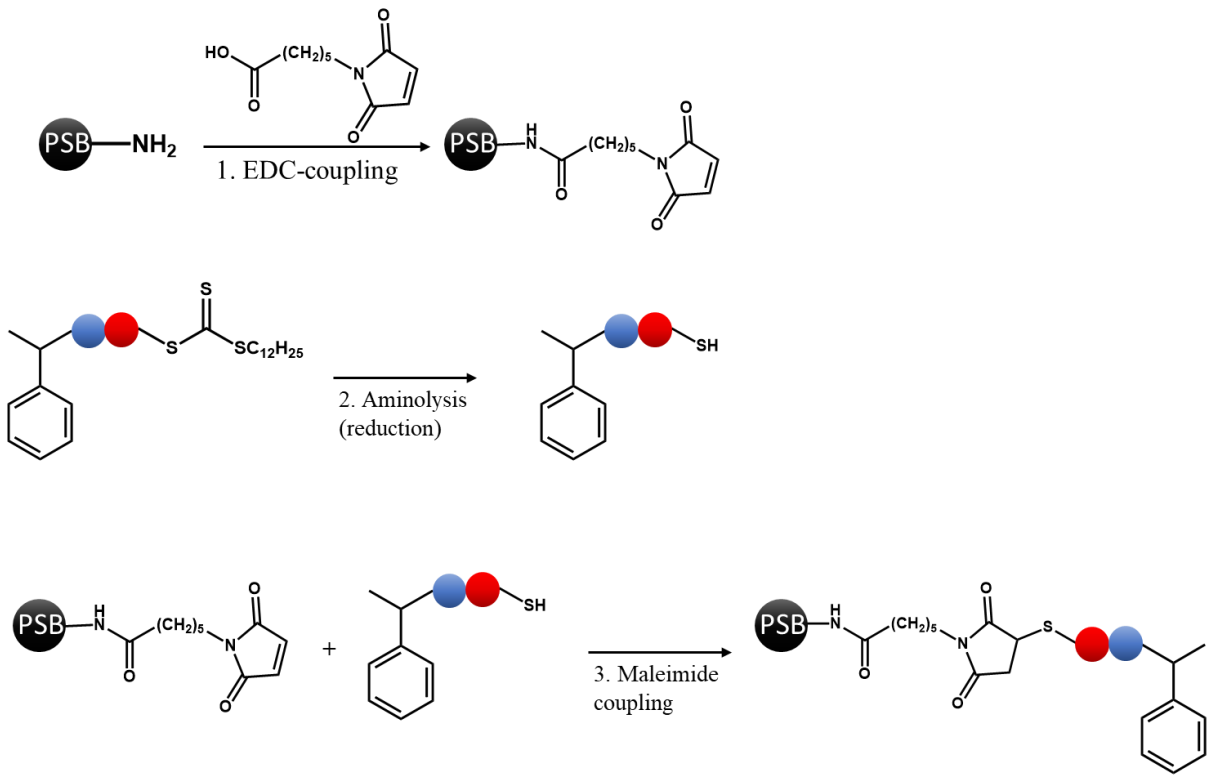


Figure 42. Process of attaching MHB-SD onto amine functionalized PSB via a spacer arm.

In the first step, the spacer arm N-ε-maleimidocaproic acid (EMCA) is coupled onto the PSB [95]. This is accomplished by means of EDC-coupling between the amine function of the PSB and the carboxylic acid function of the spacer arm [84, 100]. For coupling of the MHB-SD onto the functionalized particle, the MHB-SD should have a sulfhydryl functional group. This derivatization can be accomplished by means of aminolysis. Finally the two molecules can be covalently attached to each other by means of a maleimide coupling [95]. This method is chosen over other possible processes due to the few number of reactions required and the expected absence of side reactions [95].

2.4.1 Template polymerization

Templated polymerization requires selective base pairing between the nucleobase containing monomers and the nucleobases in the MHB-SD template. For this process, not all free NAM should be mixed with the templates simultaneously. When all NAM are poured into the mixture simultaneously, hydrogen bonding between these free NAM could take place. This could prevent the NAM from forming sufficient hydrogen bonds with the template. A first solution to this problem could be, adding the monomers in a certain sequence. Even though this seems to be a solution, still some free NAM (hydrogen bonded to each other) should be removed. A potential and hypothetical solution to this problem might lay in the principle of SPR-analysis. As in SPR-analysis, the templates should be coupled onto a solid substrate. A stream of NAM, dissolved in a solvent promoting selective hydrogen bonding, should pass these immobilized templates. This should be done until all nucleobases in the template formed hydrogen bonds with their complementary NAM. Due to the stream no free NAM (hydrogen bonded to each other) should be left in the mixture and polymerization could easily be initiated without further intermediate steps.

Furthermore, temperature is an important factor in template assisted polymerization. During polymerization, hydrogen bonds between NAM and the template should remain intact. However, certain polymerizations require elevated temperatures. This might cause the base pairs to separate as mentioned in 2.2.3.1. Therefore, the initiator should be chosen based on results of the temperature dependence of hydrogen bonding between NAM and MHB-SD oligomers, investigated in this research. The temperature should be high enough for the initiator to decompose homolytically and fast enough, but also be low enough for hydrogen bonds between nucleobases to remain intact. Light-induced initiation could resolve this problem. Here, an UV signal enables homolytic decomposition of the RAFT agent. The R-group is cleaved from the RAFT agent by means of the UV signal, meaning that no initiator is necessary. This process has already been used for trithiocarbonates and could therefore be a solution to the temperature issue [23]. Research showed that in solvents, showing a strong non-polar character, hydrogen bonding still occurred at approximately 60 °C [17]. Therefore, switching to light-induced initiators might not be necessary. This thesis will give an indication towards the environmental influences and examination methods to optimize the temperature for this process.

Potential thermal initiators are cumyl peroxyneodecanoate, V-65 and V-70 [101, 102]. Both show similar decomposition rates as AIBN at lower temperatures. These initiators can be found in Figure 43.

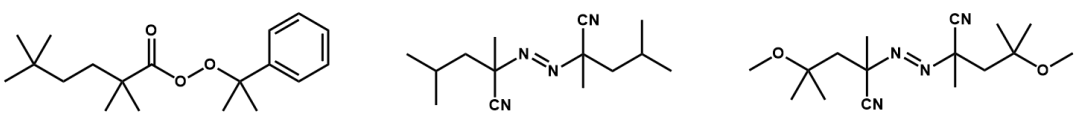


Figure 43. Thermal initiators cumyl peroxyneodecanoate (left), V-65 (middle) and V-70 (right) with their 10 hours half-life temperature.

2.4.2 Affinity separation

The temperature, solvent and chain length are important parameters in the affinity separation of MHB-SD oligomers. The effects of temperature and solvent have been discussed sufficiently in 2.2.3. A more elaborate discussion concerning the influence of the chain length will be given in this section. Longer nucleobase containing chains show an increased pairing stability with their complementary sequence as an increased amount of base pairs occurs. However, there might be a downside to this increased stability in the application of affinity separation. Due to the stability provided by a multitude of base pairs between complementary nucleobases, a slightly differing sequence could bind to the template anyway, as shown in Figure 44 [57]. The tremendous stability provided by the other base pairs could force the mismatch to occur. Additionally, due to the little influence of a single mismatch in a strain of correctly matched base pairs, the lowered stability of the non-complementary paired chains could differ barely from complementary paired chains. This could cause for insufficient isolation of the desired sequence if, by tuning temperature of solvent, this problem could not be resolved. This is however merely an assumption and requires verification.

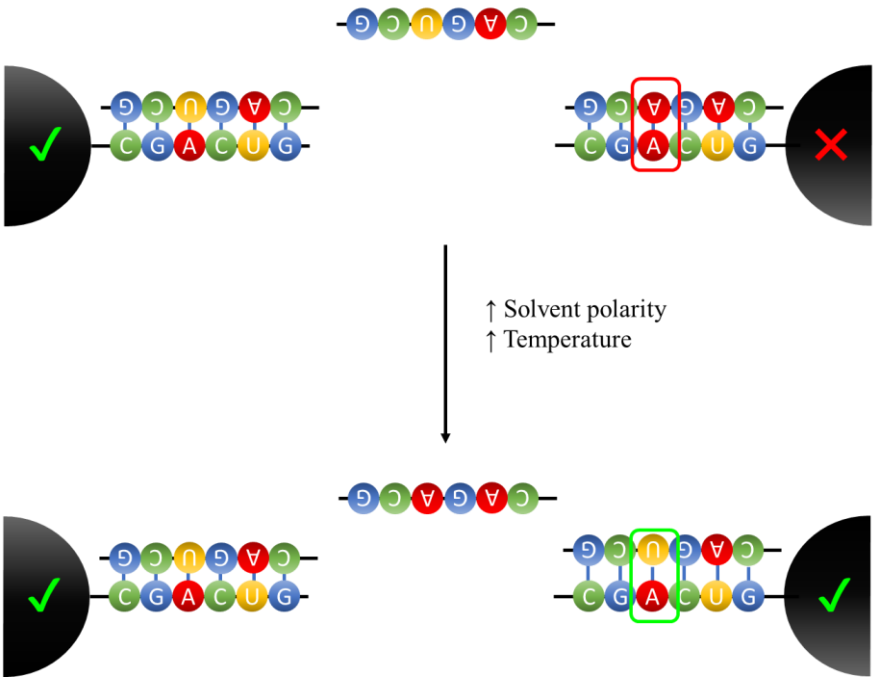


Figure 44. The influence of temperature and solvent polarity on mismatching in affinity separation. The blue lines represent hydrogen bonds.

Chapter 3. Materials and methods

3.1 Materials

Following products and solvents were used as received: adenine (A, TCI, >99.0%), thymine (A, TCI, >98.0%), uracil (U, ABCR, 99%), cytosine (C, TCI, >98%), 2-amino-6-chloropurine (ACP, ABCR, 98%), potassium carbonate (K_2CO_3, Janssen, anhydrous), magnesium sulphate ($MgSO_4$, Acros Organics, 99%), tris(2-aminoethyl)amine (TEA, Alfa Aesar, 97%), potassium tert-butoxide (KtBuO, Acros Organics, >98%), formic acid (HCOOH, Acros Organics, >98%), 2,6-Di-tert-butyl-4-methylphenol (BHT, Alfa Aesar, 99%), 1,4-butanediol diacrylate (BDDA, Sigma-Aldrich, 90%), potassium phosphate tribasic (K_3PO_4, Sigma-Aldrich, anhydrous), 1-dodecanethiol (DCT, TCI, >98%), carbon disulfide (CS_2, Acros Organics, 99.9%), (1-bromoethyl)benzene (BEB, TCI, >95%).

α,α-azoisobutyronitrile (AIBN, Glentham Life Sciences, 98%) was recrystallized twice from methanol before use.

Solvents are obtained from Sigma-Aldrich or Fisher Scientific and were used as received.

3.2 Characterization

Proton Nuclear Magnetic Resonance (^{1}H NMR) spectra were recorded on an Agilent/Varian 400 MHz Inova spectrometer using a 5 mm four-nucleus PFG probe [12]. The chemical shift scale (δ) in ppm was calibrated with TMS (0 ppm). Free induction decays were collected with a 90° pulse of approximately 6.0 μs, a spectral width of 6 kHz, an acquisition time of approximately 2 s. Spectra were analysed in Mestrenova software.

3.3 Methods

3.3.1 Synthesis of monomers and oligomers

3.3.1.1 Synthesis of adenine acrylate monomer (AAM)

AAM was synthesized according to literature [9, 12]. A mixture of adenine (10.00g, 1 eq.), K_2CO_3 (0.46 g, 0.04 eq.) and BHT (0.69 g, 0.04 eq.) in 200 ml DMSO was heated to 50°C and stirred for 1 h. Subsequently, BDDA (28.00 ml, 2 eq.) was added to the mixture and the reaction continued for 5 h. In the following step, the mixture was diluted with water (1500 ml). Washing with hexane (350 ml) was then executed to remove excess of BDDA. This step was followed by an extraction with DCM (3 x 200 ml). Removal of residual water was done by adding $MgSO_4$ and removing the precipitate with filtration. The dried mixture was then concentrated under reduced pressure. Column chromatography with $CHCl_3$/MeOH (90/10 vol%) was performed to purify the mixture. Finally, the solvent was removed by means of evaporation under reduced pressure and high vacuum using a cold trap vacuum installation with liquid nitrogen. A mass of 14.80 g of the pure product was obtained, resulting in a 60 % yield.

3.3.1.2 Synthesis of thymine acrylate monomer (TAM)

TAM was synthesized according to literature [9, 12]. A mixture of thymine (1.00g, 1 eq.), TEA (0.22 ml, 0.20 eq.) and BHT (0.06 g, 0.04 eq.) in 20 ml DMSO was stirred at room temperature for 1 h. Subsequently, BDDA (3.00 ml, 2 eq.) was added to the mixture and the reaction continued for 24 h. In the following step, the mixture was diluted with water (150 ml). Washing with hexane (35 ml) was then executed to remove excess of BDDA. This step was followed by an extraction with DCM (3 x 20 ml). Removal of residual water was done by adding $MgSO_4$ and removing the precipitate with filtration. The dried mixture was then concentrated under reduced pressure. Column chromatography with $CHCl_3$/MeOH (95/5 vol%) was performed to purify the mixture. Finally, the solvent was removed by means of evaporation under reduced pressure and high vacuum using a cold trap vacuum installation with liquid nitrogen. A mass of 1.75 g of the pure product was obtained, resulting in a 68 % yield.

3.3.1.3 Synthesis of uracil acrylate monomer (UAM)

UAM was synthesized according to literature [9, 12]. A mixture of uracil (6.00g, 1 eq.), TEA (1.50 ml, 0.20 eq.) and BHT (0.40 g, 0.03 eq.) in 120 ml DMSO was stirred at room temperature for 1 h. Subsequently, BDDA (20.00 ml, 2 eq.) was added to the mixture and the reaction continued for 24 h. In the following step, the mixture was diluted with water (900 ml). Washing with hexane (150 ml) was then executed to remove excess of BDDA. This step was followed by an extraction with DCM (4 x 100 ml). Removal of residual water was done by adding $MgSO_4$ and removing the precipitate by filtration. The dried mixture was then concentrated under reduced pressure. Column chromatography with $CHCl_3$/MeOH (97/3 vol%) was performed to purify the mixture. Finally, the solvent was removed by means of evaporation under reduced pressure and high vacuum using a cold trap vacuum installation with liquid nitrogen. A mass of 9.63 g of the pure product was obtained, resulting in a 58 % yield.

3.3.1.4 Synthesis of cytosine acrylate monomer (CAM)

CAM was synthesized according to literature [9, 12]. A mixture of cytosine (4.00g, 1 eq.), KtBuO (0.16 g, 0.04 eq.) and BHT (0.32 g, 0.04 eq.) in 80 ml DMSO was stirred at room temperature for 1 h. Subsequently, BDDA (14.00 ml, 2 eq.) was added to the mixture and the reaction continued for 24 h. In the following step, the mixture was diluted with water (600 ml). Washing with hexane (100 ml) was then executed to remove excess of BDDA. This step was followed by an extraction with DCM (3 x 80 ml). Removal of residual water was done by adding $MgSO_4$ and removing the precipitate with filtration. The dried mixture was then concentrated under reduced pressure. Column chromatography with $CHCl_3$/MeOH (97/3 vol%) was performed to purify the mixture. Finally, the solvent was removed by means of evaporation under reduced pressure and high vacuum using a cold trap vacuum installation with liquid nitrogen. A mass of 5.09 g of the pure product was obtained, resulting in a 45 % yield.

3.3.1.5 Synthesis of guanine acrylate monomer (GAM)

GAM was synthesized according to literature [12]. A mixture of ACP (1.00 g, 1 eq.), KtBuO (0.03 g, 0.04 eq.), BHT (0.06 g, 0.04 eq.) in 20 ml DMSO was stirred for 1 h at 50 °C. Subsequently, BDDA (2.39 ml, 2 eq.) was added and the reaction continued for 5 h at 50 °C. Afterwards, the mixture was diluted with water (600 ml). In the following step, washing with hexane (75 ml) was done to remove excess of BDDA. An extraction with DCM (75 ml) was repeated four times. Removal of residual water was done by adding $MgSO_4$ and removing the precipitate with filtration. The dried mixture was then concentrated under reduced pressure resulting in a crude mixture. 1.29 g of this mixture was dissolved in a $HCOOH/H_20$-mixture (80/20 vol%) and stirred for 2 h at 75 °C. This final mixture was dried under reduced pressure. Column chromatography with $CHCl_3$/MeOH (97/3 vol%) was then performed to purify the concentrated mixture. Finally, the solvent was removed by means of evaporation under reduced pressure and high vacuum using a cold trap vacuum installation with liquid nitrogen. A mass of 0.68 g of the pure product was obtained, resulting in a 33 % yield.

3.3.1.6 Synthesis of synthesis of RAFT-agent (DPE-TTC)

DPE-TTC was synthesized according to literature [12, 103]. K_3PO_4 (6.23 g, 1 eq.) and 1-dodecanethiol (7.0 ml, 1 eq.) were suspended in acetone (166 ml) and stirred for 10 minutes at room temperature. During this time, the mixture was also purged with nitrogen. Afterwards, CS_2 (5.31 ml, 3 eq.) is added to the suspension and stirring continued for 1 hour. Subsequently, 1-bromoethyl benzene (4.00 ml, 1 eq.) was added. After 5 hours of stirring the solvent was evaporated under reduced pressure and the crude mixture was purified using column chromatography (100% petroleum ether). After removing of the solvent by means of a rotavapor and drying under vacuum by means of a cold trap vacuum installation with liquid nitrogen, 7.95 g pure product was obtained (71% yield).

3.3.2 Nuclear magnetic resonance experiments

3.3.2.1 Job plot of AAM-TAM base pairing

For characterization of the base pairing interaction between NAM and TAM one sample set was analysed. For these samples, AAM and TAM were dissolved in 65 ml deuterated chloroform. Mole fractions of the components were varied whilst keeping the overall monomer concentration constant at 40 mM. This is called the continuous variation method. The ^{1}H-NMR spectra were measured at 25 °C and 50 °C. In Table 2 the mole fractions of these samples are given.

Table 2. Sample data from characterization of AAM-TAM base pairing

Sample	γ_{AAM}	γ_{TAM}
1	1.0000	0.0000
2	0.8917	0.1083
3	0.7798	0.2202
4	0.6889	0.3111
5	0.5869	0.4131
6	0.4902	0.5098
7	0.4050	0.5950
8	0.3016	0.6984
9	0.2036	0.7964
10	0.1117	0.8883
11	0.0000	1.0000

3.3.3 NOE-NMR experiments

Various mixtures of base pairs were analysed by means of ^{1}H-NMR, followed by NOE-NMR analysis of the same samples. Data of the samples is given in Table 3.

Table 3. Sample data from NOE-experiments for examination of base pairing selectivity. The shift given in the last column represents the shift that was radiated during the NOE-analysis.

Sample	γ_{AAM}	γ_{TAM}	γ_{UAM}	γ_{GAM}	γ_{CAM}	$CHCl_3$	$DMSO$	$MeOH$	T [°C]	δ [ppm]
1 nucleobase										
A1	1	0	0	0	0	X	/	/	25.0	/
A2	1	0	0	0	0	/	X	/	23.0	/
T1	0	1	0	0	0	X	/	/	25.0	/
T2	0	1	0	0	0	/	X	/	22.8	/
U	0	0	1	0	0	/	X	/	23.0	/
G	0	0	0	1	0	/	X	/	20.7	/
C1	0	0	0	0	1	X	/	/	25.0	/
C2	0	0	0	0	1	/	X	/	23.3	/
2 nucleobases										
AT1	0.2539	0.7461	0	0	0	/	X	/	22.2	11.23456
AT2	0.2506	0.7494	0	0	0	/	X	X	22.2	11.06203
AT3	0.2545	0.7455	0	0	0	/	X	X	50.0	11.06203
AT4	0.2798	0.7202	0	0	0	X	/	/	20.2	11.45158
AU1	0.3525	0	0.6475	0	0	X	/	/	20.8	11.99660
GC1	0	0	0	0.3558	0.6442	/	X	/	20.2	11.32000
GC2	0	0	0	0.4378	0.5622	X	/	/	21.2	13.67219
CT1	0	0.7327	0	0	0.2673	X	/	/	20.1	/
CT2									50.0	/
AC1	0.4946	0	0	0	0.5054	X	/	/	21.0	6.4346
3 nucleobases										
ACG1	0.3724	0	0	0.2985	0.3291	X	/	/	20.2	6.2676 13.5504
ACG2	0.4084	0	0	0.2120	0.3796	/	X	/	20.3	7.1152 7.3237 8.1020 11.3351
ACG3	0.3232	0	0	0.1908	0.4860	/	X	/	80.0	/
ATC1	0.3441	0.5271	0	0	0.1431	X	/	/	20.3	6.9303
ATC2	0.3415	0.5128	0	0	0.3293	/	X	/	21.6	11.2918
ATC3	0.3580	0.3160	0	0	0.3259	/	X	/	50.0	/
TCG1	0	0.4367	0	0.2833	0.2800	X	/	/	20.0	11.4730
TCG2	0	0.4286	0	0.1837	0.3878	/	X	/	21.7	7.0934 11.2059 11.9652
4 nucleobases										
ATCG	0.1553	0.3439	0	0.0952	0.4057	X	/	/	20.1	6.6563 6.9663 13.6189

Chapter 4. Results and discussion

The experiments executed and analysed in this research were conducted to test the NMR-based analysis techniques and gain insight into the occurring phenomena in base pairing between NAM. The results discussed in this chapter will mainly provide pointers for future characterization of base pairing between synthetic nucleobase containing structures. Additionally, template-assisted polymerization and affinity separation of synthetic nucleobase containing polymers will be evaluated with the results from the experiments.

4.1 NMR-analysis of base pairing between complementary NAM

In this first section of Chapter 3 the mechanism of base pairing between complementary nucleobases will be analysed. It is important to obtain or verify the presence of a 1:1-stoichiometry in this base pairing as any deviation can cause problems in applications of the synthetic nucleobase containing polymers and systems such as templated polymerization and affinity separation as described in 1.2.

4.1.1 NMR-analysis of base pairing between AAM and TAM

A-T base pairing in chloroform is examined by using the continuous variation method as described in 3.3.2.1. Here, the change in shift of certain protons is determined and used to characterize base pairing. The examined protons are marked in Figure 45.

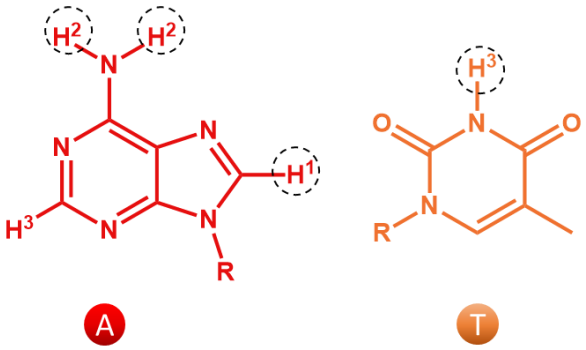


Figure 45. Examined protons in characterization of A-T base pairing.

As described in 2.3.1.1, protons participating in hydrogen bonding show a change in their respective chemical shift due to the deshielding effect. This phenomenon is however not exclusively observed in hydrogen bonding as deshielding of a proton can be caused by presence of any atom/group with a strong electronegativity. Therefore, other (similar) interactions can cause a deshielding effect and increase the chemical shift of a proton as well. The changing δ is also observed for proton H^1 and H^3 of adenine displayed in Figure 45 and is likely to be the

result of the interaction mentioned in the beginning of this paragraph. Examination of the shift proton H^1 of AAM will be used to characterize the A-T base pairing.

The change chemicals shift, caused by the secondary interactions, is summarized in Job plots. This type of plot is commonly used to determine the stoichiometry of complexation between host and guest molecules. This means that it could also be used to prove or disprove the expected 1:1-stoichiometry of base pairing between nucleobases, when both are implemented into monomers [17]. The x-axis of a Job plot represents the mole fraction of the guest molecule and the y-axis represents the change in the interaction parameter, observed in the host molecule, due to complexation. In this research the interaction parameter will be the chemical shift that changes due to complexation by means of hydrogen bonding or base pairing. Determination of the stoichiometry can be done by determining the mole fractions at which the interaction parameter from the Job plot reaches its maximum value [104]. When the interaction parameter reaches a maximum value, the stoichiometric point is reached [105]. The stoichiometry of complexation between a host and guest molecule, as calculated by a Job plot, is then given as:

$$\chi_{host\ (max)} : \chi_{guest\ (max)} \tag{9}$$

$\chi_{(max)}$ represents the mole fraction of the respective compound at which the interaction parameter reaches its maximum value. In hydrogen bonding there is an electronegative atom that is not covalently attached to the proton. In this research, the molecule containing this atom, will be referred to as the guest molecule. The molecule containing the electronegative atom, covalently bound to the hydrogen, is called the host molecule.

Examination of base pairing between AAM and TAM rendered the plots shown in Figure 46, Figure 47 and Figure 48. The shift difference used in the Job plot is calculated as the difference between the shift of the proton in hydrogen bonded state and the shift of the proton in a free state. After determination of the polynomial describing the measured data, by means of the polynomial fitting program of excel or the MATLAB-code from A1, the maximum value of the interaction parameter is calculated. This maximum value is calculated as the root of the first order derivative of the polynomial. Its corresponding mole fraction of AAM and TAM then give the stoichiometry of the complexation.

An example of determination of the stoichiometry is given below. It concerns the ^{1}H-NMR analysis from Figure 46 at 25 °C.

$$\begin{aligned} \chi_{host\ (max)} : \chi_{guest\ (max)} \rightarrow \chi_{TAM\ (max)} : \chi_{AAM\ (max)} \ &\rightarrow 0.5003 : 0.4997 \\ &\approx 1 : 1 \end{aligned} \tag{10}$$

The results for the other measurements are given in Table 4. Note that all experiments have been executed at both 25 °C and 50 °C to examine potential temperature effects.

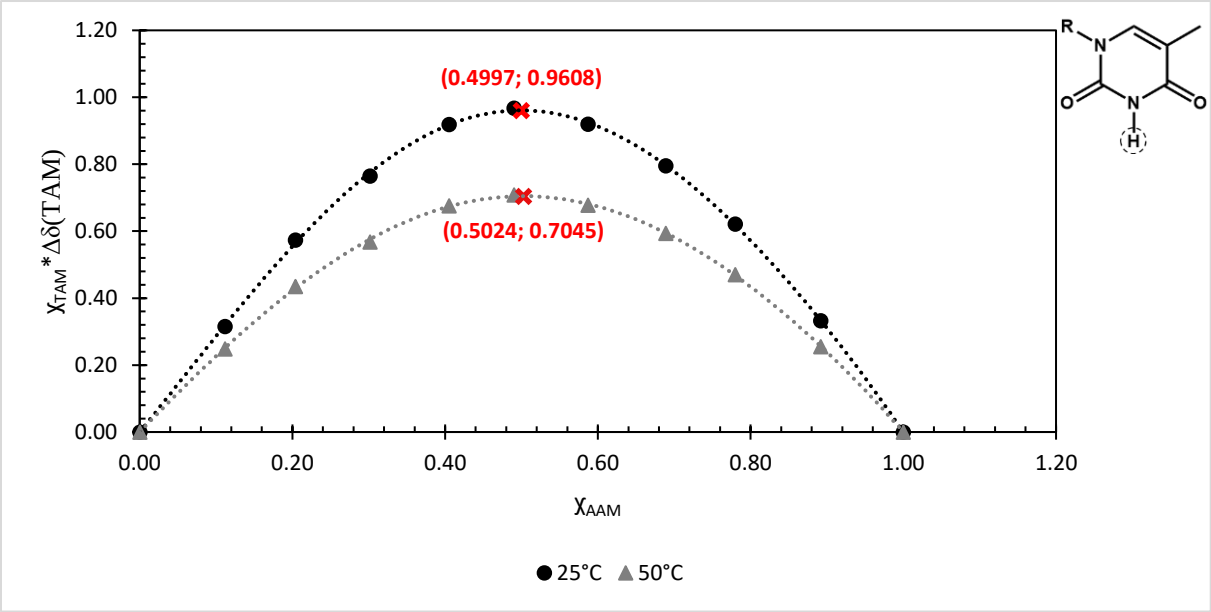


Figure 46. Job plot of ^{1}H-NMR analysis concerning hydrogen bonding between AAM (guest) and TAM (host). The shift used in the y-axis is the change in shift of the proton H^3 of thymine from TAM participating in hydrogen bonding with adenine from the AAM. This proton is indicated (dashed circle) in the thymine structure.

Table 4. Stoichiometry of base pairing between AAM and TAM. The protons of TAM and AAM to which is referred can be found in Figure 45.

Proton	Host	Guest	25 °C	50 °C
H^3 (TAM)	TAM	AAM	1:1 (0.5003:0.4997)	1:1 (0.4976:0.5024)
H^2 (AAM)	AAM	TAM	1:2 (0.3334:0.6666)	1:2 (0.3168:0.6832)
H^1 (AAM)	AAM	TAM	1:1.5 (0.4194:0.5806)	1:1.5 (0.4250:0.5750)

A 1:1-stoichiometry is observed, while examining H^3 from TAM in Figure 42 at both 25 °C and 50 °C, meaning that indeed one TAM-molecule would form a base pair with one AAM-molecule. As can be seen in Table 4, this is in contradiction with the results coming from examination of H^2 from AAM in Figure 45. Here, at both 25 °C and 50 °C, a 1:2-stoichiometry is observed. This means that one molecule AAM would form a base pair complex with two TAM molecules. Furthermore, analysis of H^1 from AAM renders a 1:1.5-stoichiometry at both temperatures. This implies that a combination of 1:1- and 1:2-stoichiometry occurs during base pairing of AAM and TAM. The results from examination of H^1 from AAM could also be interpreted as a 2:3-stoichiometry. This seems unlikely as there are no possibilities for stable base pairing to occur between the NAM resulting in a A_2T_3-complex. For this 2:3-stoichiometry to occur, a proton of a NAM would have to form hydrogen bonds with a carboxylic oxygen from the acrylate tail of the monomer. Formation of this hydrogen bonding renders less stability than the formation of a base pair and is therefore assumed to be unlikely.

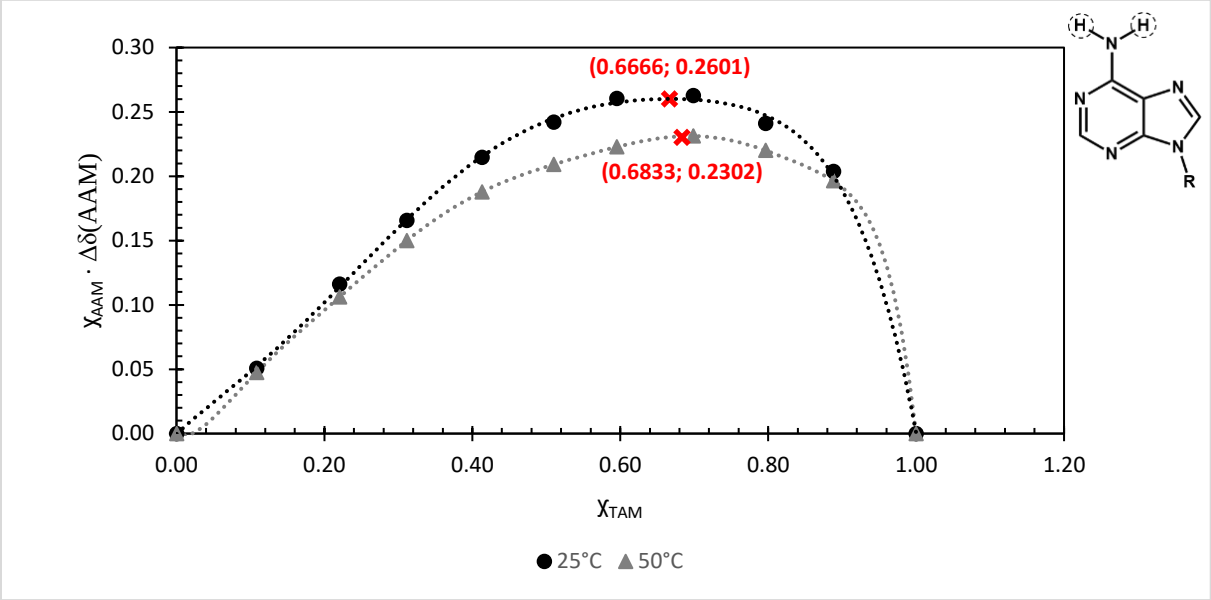


Figure 47. Job plot of ^{1}H-NMR analysis concerning hydrogen bonding between AAM (host) and TAM (guest). The shift used in the y-axis is the change in shift of the proton H^{2} of adenine from AAM participating in hydrogen bonding with thymine from the TAM. These protons from adenine are indicated (dashed circle) in the molecular structure of AAM.

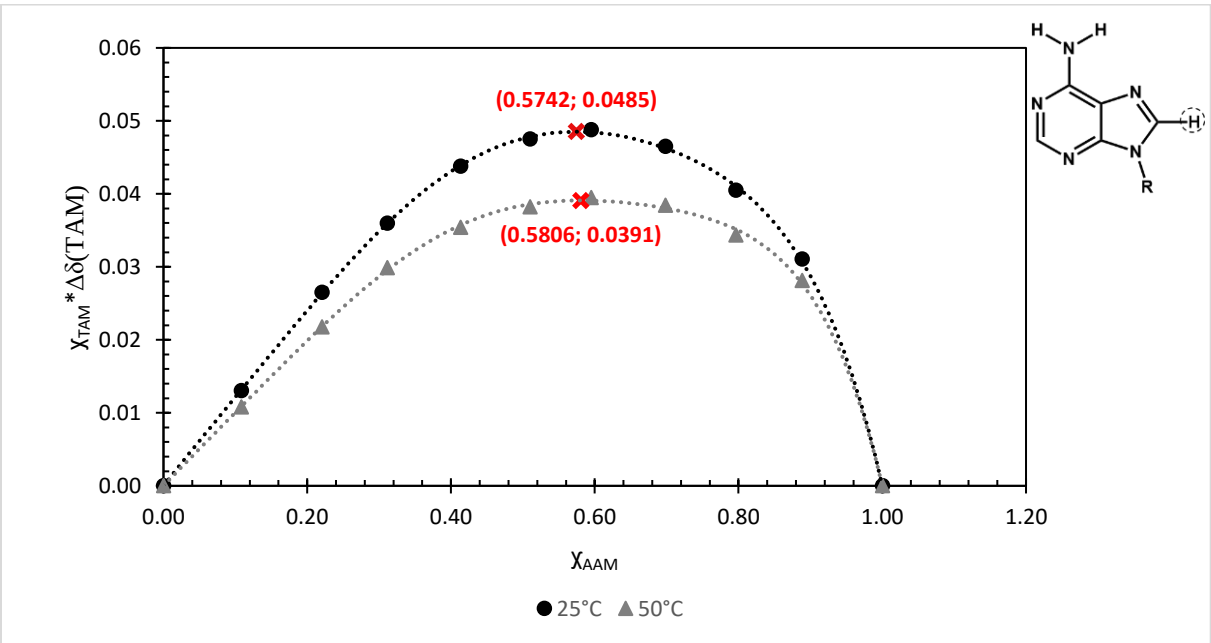


Figure 48. Job plot of ^{1}H-NMR analysis concerning hydrogen bonding between AAM (host) and TAM (guest). The shift used in the y-axis is the change in shift of the proton H^{1} on adenine participating in a weak interaction with thymine from the TAM. The proton on adenine is indicated (dashed circle) in the molecular structure of AAM.

Even though the Job plot analysis implies different stoichiometries, an attempt to elucidate the mechanism of the complexation occurring between AAM and TAM will be made. This theory will be based on the existing base pairing models, the Job plots and NOE-experiments.

The Job plot, also called the continuous variation method, is used in this research to examine the stoichiometry of A-T base pairing. It is known that for a 1:1-stoichiometry, as expected in this situation, this method is fairly reliable [17, 106]. In more complicated forms of complexation, these Job plots need to be interpreted with great caution as one can easily be guided to false conclusions. Research showed that Job plots, showing 1:1-stoichiometries, can often be the result of more complex complexation models and therefore lead to incorrect conclusions [106]. An indicator for the potentially false observation of a 1:1-stoichiometry is the presence of a flattened peak at the 0.5 mole fraction mark on the x-axis. True 1:1 complexation should result in a sharp peak at this point. Due to the fact that the Job plot in Figure 46 shows a flattened peak at the 0.5 mark, combined with the contradicting stoichiometries from the Job plots, suggest that the expected 1:1-stoichiometry does not occur as such. Instead, a more complex base pairing would occur. Here, it is assumed that the 1:1-, 1:2- and 1:1.5-stoichiometries imply a combined 1:1- and 1:2-stoichiometry resulting in both AT and AT_2 base pairs. This assumption is supported by the fact that both H^3 and H^1 show a change in their respective δ within the continuous variation method. This simultaneous shifting, shown in Figure 49**,** of the δ from both protons cannot be caused by the formation of an AT base pair. It is most likely that it is caused by the formation of an AT_2 base pair.

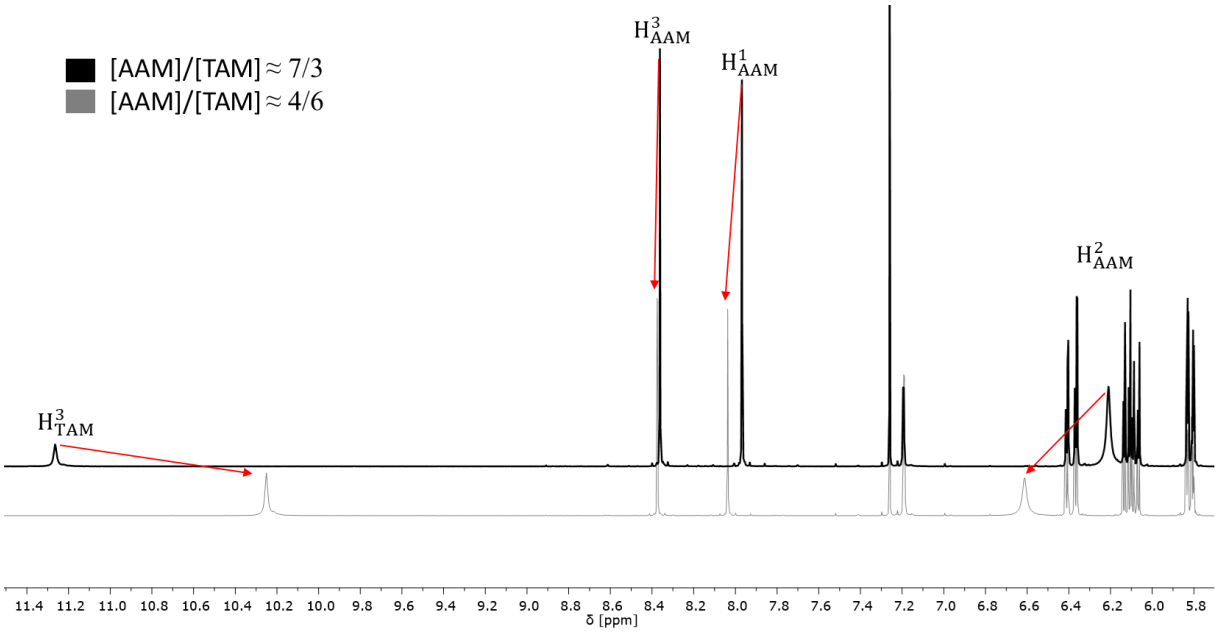


Figure 49. Shifting δ of protons observed during the continuous variation method used for the analysis of AAM-TAM complexation in chloroform.

For the AT_2 complex, the most likely and most stable configuration is a combination of the (reversed) WC and (reversed) Hoogsteen models. An example employing the WC and Hoogsteen model is depicted in Figure 50. This configuration entails the highest stability obtainable with an AT_2 base pair. The weak interactions of H^1 and H^3 from AAM with carbonyl-oxygens from TAM also correspond to their observed change in δ in the continuous variation method. This is the first experimental result confirming the theoretical AT_2-model.

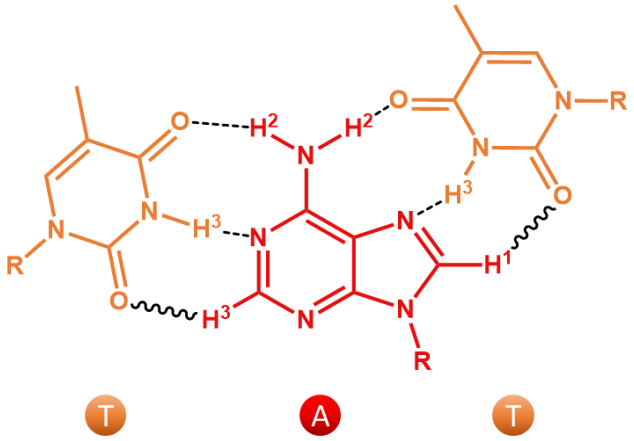


Figure 50. Combined model of Watson-Crick and Hoogsteen likely to explain the 1:2-stoichiometry observed in A-T base pairing.

To further confirm this form of base pairing, NOE-analysis is employed. The spectrum from this analysis can be found in appendix A2 and A3A1 In this analysis the H^3 of TAM is irradiated. The NOE-spectrum, revealing all protons in close proximity to H^3, reveals δ from protons H^1, H^2 and H^3 belonging to AAM as shown in Figure 51. This result implies that, for an AAM/TAM molar concentration ratio of 3/1, an AT_2-complex is formed combining (reversed) WC and (reversed) Hoogsteen base pairing. Sole occurrence of (reversed) WC or (reversed) Hoogsteen cannot cause the occurrence of both H^1 and H^3 of AAM in the NOE-spectrum. The distance between H^3 from TAM and H^1 from AAM in the (reversed) WC configuration is around 8 Å, exceeding the 6 Å limit of the NOE-effect. The distance between H^3 from TAM and H^3 from AAM in (reversed) Hoogsteen base pairing is around 7.5 Å, also exceeding the 6 Å limit of the NOE-effect. This confirms that the occurrence of the δ, belonging to H^1 and H^3 in the NOE-spectrum, should indeed be the result of a combined (reversed) WC and (reversed) Hoogsteen complexation.

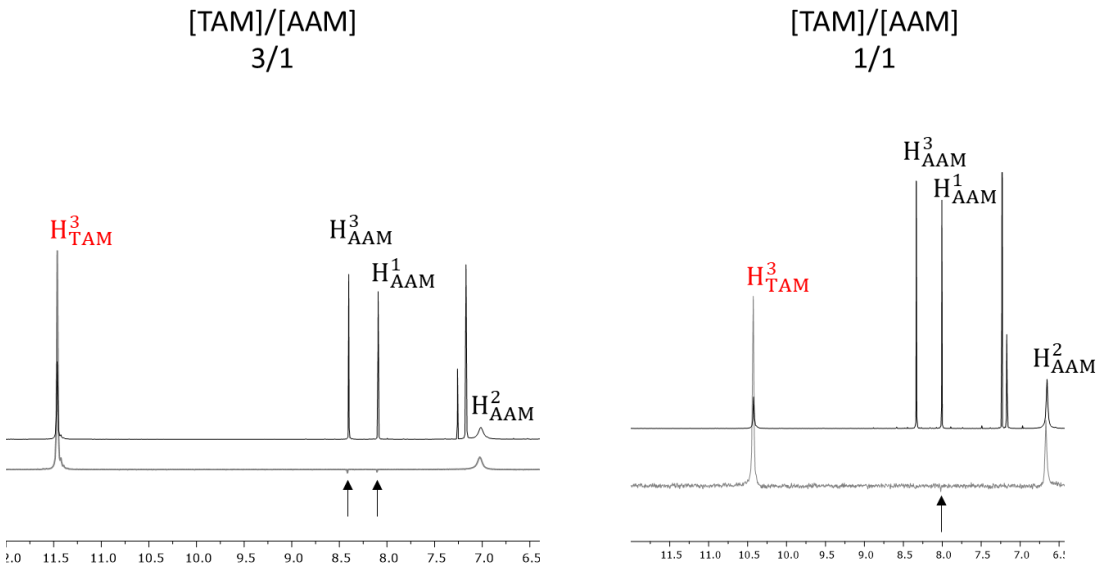


Figure 51. NOE-analysis of TAM-AAM base pairing in chloroform by means of irradiation of proton H^3 of TAM. The ratios above each partial spectrum are the respective molar concentration ratios of the NAM in the sample.

As mentioned in the previous paragraph, NOE-analysis implies the occurrence of the model described in Figure 50, when the concentration of TAM is much higher than that of AAM. To determine whether the occurrence of this form of complexation is dependent on the concentration ratio, another NOE-analysis was executed. Here, AAM and TAM are present in the mixture in near equimolar concentrations. The resulting NOE-spectrum, shown in Figure 51 or appendix A3, reveals the absence of the δ belonging to H^3 from AAM. In the same spectrum, a δ belonging to H^1 from AAM is present as shown by the arrow in Figure 51. The presence of the δ from this proton could imply that, as the [TAM]/[AAM]-ratio increases from 0/1, mainly (reversed) Hoogsteen base pairs are formed until the point of equimolar concentration resulting in mainly a 1:1-stoichiometry of complexation. This would be in contradiction with the data concerning the stability of the two base pairing models as given in 2.2.

Note that the data of the base pairs is based on interaction between nucleobases and not NAM. Therefore, a stabilizing effect of the structure, added onto each nucleobase upon derivatization, could alter the stabilities of the different base pairing models. A possible result is an increased stability of the (reversed) Hoogsteen base pairing. This would explain the thermodynamic preference for the (reversed) Hoogsteen model. Additionally, note that increased formation of Hoogsteen base pairs, based on the NOE-analysis is merely an assumption and not a proven fact. The absence of the δ of H^3 from AAM in the NOE-spectrum does not necessarily mean that the (reversed) WC base pairing does not occur at this near equimolar ratio. The presence of a δ proves that the proton belonging to this δ is in close proximity to the irradiated proton. However, absence of the δ of a proton in the NOE-spectrum is not conclusive evidence that this proton is not in close proximity to the irradiated proton. Further research is required if the base pairing mechanism is to be elucidated completely.

Note that in the continuous variation method, protons likely to represent the formation of (reversed) WC and (reversed) Hoogsteen base pairing both shift as the monomer-ratio increases from 0/1 onwards. This implies the formation of both (reversed) WC and (reversed) Hoogsteen base pairs at lower than equimolar TAM/AAM-ratios meaning that the presence of an equilibrium between the two types of base pairing is more likely. Potentially, based on the NOE-analysis results, this equilibrium tends slightly more to the side of (reversed) Hoogsteen base pairing. The ^{1}H-NMR data also implies that potentially AT_2-complexes could be formed at TAM/AAM ratios below 1. However, the formation of AT_2-complexes below the equimolar ratio is thermodynamically unlikely. When the ratio is increased beyond the equimolar ratio on the other hand, additional base pairing onto AT-complexes could occur, resulting in the formation of the AT_2 complex. Under these conditions, the formation of AT_2-complexes is likely as most AAM will already have formed an AT-base pair with one other TAM. Therefore, additional base pairing of TAM onto AAM could only result in the formation of the AT_2-complex as described in Figure 50. This concentration-ratio dependency of the base pairing stoichiometry is an important factor to take into account when examining optimal conditions for templated polymerization. Note that this is merely a theory, which still requires further verification/research.

An additional and remarkable fact is that both the δ of H^1 and H^3, belonging to AAM, disappear when the same NOE-experiment is executed in DMSO with and without MeOH. The corresponding spectra can be found in appendices A4, A5 and A6. This is troubling for the analysis of the base pairing stoichiometry in these solvents. Irradiation of H^3, belonging to TAM, cannot render information anymore concerning the occurrence of WC and/or Hoogsteen base pairing. The reason for disappearance of these signals in polar solvents might be found in interaction between the respective protons and the used solvent(s). When DMSO is used as a solvent, the δ from H^1 and H^3 of AAM disappear. A signal, representing the protons of water present in DMSO, appears instead. This could imply that both water and DMSO interact with protons H^1 and H^3 from AAM. The shift from DMSO is not observed as it is a deuterated solvent. To confirm the interaction with DMSO, the same experiment should be conducted with a sample spiked with non-deuterated DMSO. Water and DMSO could interact with these protons because the interaction of these protons with the carbonyl oxygen from thymine is weak [39]. This competition could then cause the respective protons to be shielded from the NOE-effect upon irradiation of H^3 from TAM. Instead, the water- and DMSO-molecules are now affected by the NOE-effect. The same reasoning can be followed when MeOH is added to the mixture. Also, the δ representing the protons of MeOH are present in the NOE-spectrum. The theory that interaction of a proton with a solvent causes disappearance of this proton in the NOE-spectrum is however merely an assumption, requiring further experimental verification. The δ of MeOH and water, appearing in the NOE-analysis, could also simply be caused by the interaction of these solvent molecules with the (temporarily) free TAM's. Note that this would not explain the disappearance of protons H^1 and H^3 of AAM in the NOE-spectra. Additionally, a combination of both theories is plausible.

A final remarkable phenomenon is observed in the NOE-spectra of the AT-samples, concerning occurrence of secondary interactions between TAM and the acrylate tail of the monomers. At a [TAM]/[AAM]-ratio of 3/1, δ representing the protons from the double bond of the acrylate monomer, are observed in the NOE-spectrum for which proton H^3 of TAM was irradiated. This is shown in Figure 52 and implies that previous mentioned interaction, with the acrylate tail, occurs. The likely mechanism to occur is the following. Firstly, AT and AT_2 base pairs will be formed upon increasing the [TAM]/[AAM]-ratio. When the ratio surpasses the 2/1-mark, TAM will not be able to form base pairs with AAM anymore as most binding sites will have been occupied. This means that two options remain. Either the remaining TAM molecule bind to each other or they form secondary interactions with the acrylate tail. Most likely a combination of the two system occurs.

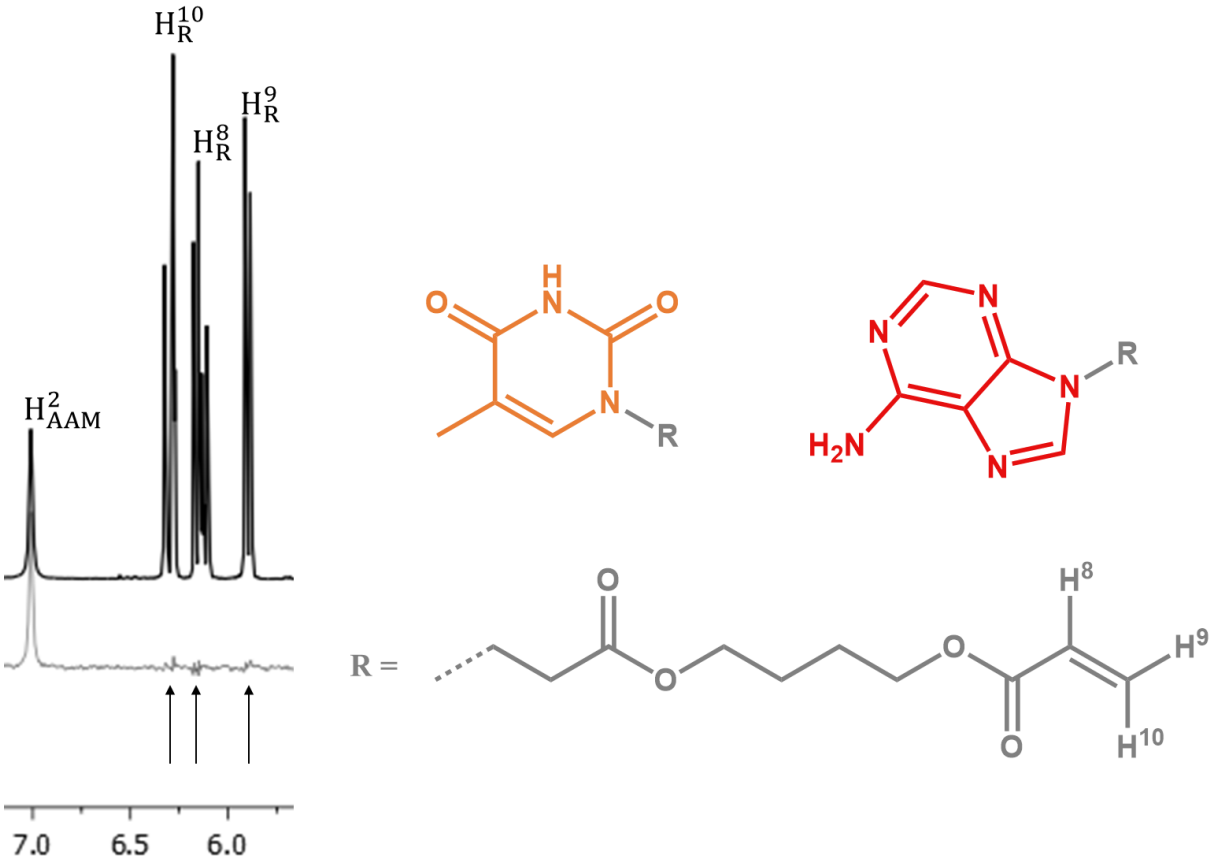


Figure 52. NOE-NMR example showing the occurrence of the δ representing the protons on the acrylate functional group on the tail of the monomer upon irradiation of H^3 from TAM in DMSO. The original spectrum can be found in appendix A6.

The Job plots, in combination with the NOE-spectra, have thus already revealed information concerning the base pairing stoichiometry. Additionally, a Job plot can render information about the stability of the formed complexes. As the peak of the Job plot is sharper and reaches a higher maximum value, a higher association constant (K_a) is present [107]. This is depicted in Figure 53 for a 1:1-complexation. Here, K_d represents the dissociation constant. An equilibrium of a 1:1 complexation between a host (H) and a guest (G) molecule can be described as:

$$H + G \underset{K_d}{\overset{K_a}{\rightleftharpoons}} HG \tag{11}$$

For which the equilibrium constant is described as:

$$K_a = \frac{[HG]}{[H] \cdot [G]} \tag{12}$$

[HG], [H] and [G] are the concentrations of the host-guest complex, the free host and the free guest respectively. This equation means, that as the value for K_a increases, the equilibrium will tend more to the side of the complex. The cause for this shift is an increase in complex-stability.

In all Job plot from the base pairing analysis between AAM and TAM an increase in K_a is observed as the temperature of the mixture decreases. This confirms that the formed base pairs are more stable as the temperature decreases.

Values for the association constants cannot be determined from these Job plots as they do not imply the same stoichiometry. Due to the complex base pairing, deviating from the simple 1:1-stoichiometry, a complex methods is required to determine association constants for this system [108] .

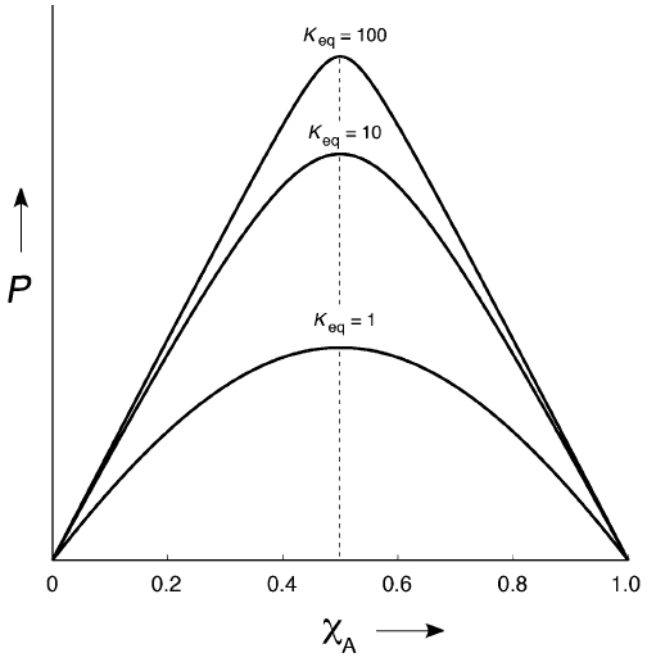


Figure 53. Job plots for 1:1-complexation according to equation 11 [107]. χ_A represents the mole fraction of the guest molecule taking part in the complexation. P represents the physical property of the complex. K_{eq} equals the K_a from formula (11) and (12).

4.1.2 NMR analysis of base pairing between AAM and UAM

No Job plots were constructed to analyse base pairing between AAM and UAM. Due to the high structural similarity between UAM and TAM, a similar base pairing mechanism is to be expected. A minor possible difference is a change in association constants due to absence of the methyl-group on the hexagonal cyclic structure. This might cause a slight reduction in the partial negative charge of the oxygen in the neighbouring carboxylic function due to disappearance of its weak +I-effect. This would weaken the hydrogen bond and potentially change or lower the K_a-constant of the complexation mechanism.

NOE-analysis of the AAM-UAM base pairing implies the same phenomenon to occur as described in 4.1.1 which verifies the similarity in base pairing as described for AAM-TAM base pairing.

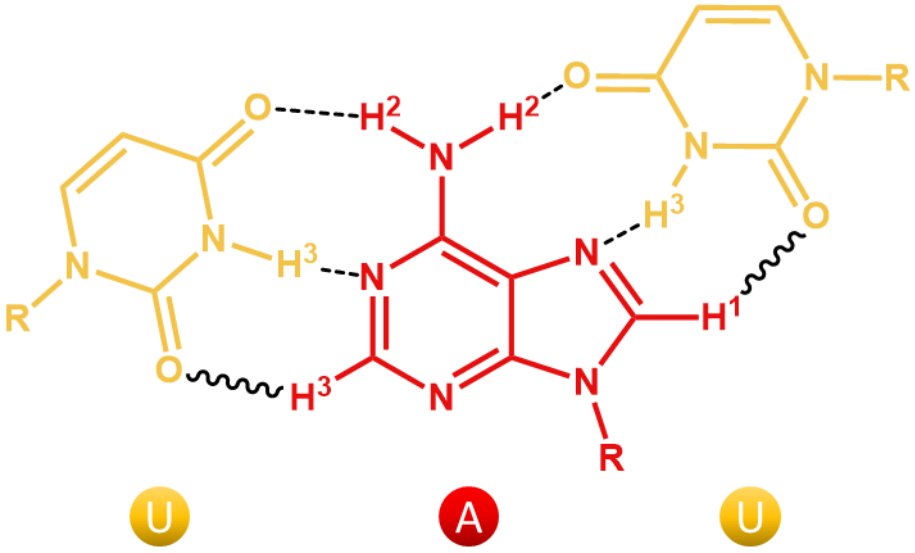


Figure 54. Combined model of Watson-Crick and Hoogsteen likely to explain the 1:2-stoichiometry observed in A-U base pairing.

Here, H^3 of UAM was irradiated. The NOE-spectrum, partially depicted in Figure 55, revealed signals from protons H^1, H^2 and H^3 belonging to AAM. As the structure of the molecules is nearly identical and hydrogen bonds have approximately the same length, in comparison to A-T base pairing, occurrence of the δ belonging to H^1 and H^3 in the NOE-spectrum, would indeed be the result of a combined (reversed) WC and (reversed) Hoogsteen complexation. Here, the condition from AAM-TAM base pairing, [UAM] >> [AAM], is also met. Therefore, the same phenomenon is observed as in AAM-TAM base pairing as expected. As mentioned before, further concentration-ration dependency is expected to be the same for UAM-AAM as for TAM-AAM base pairing but still requires further research.

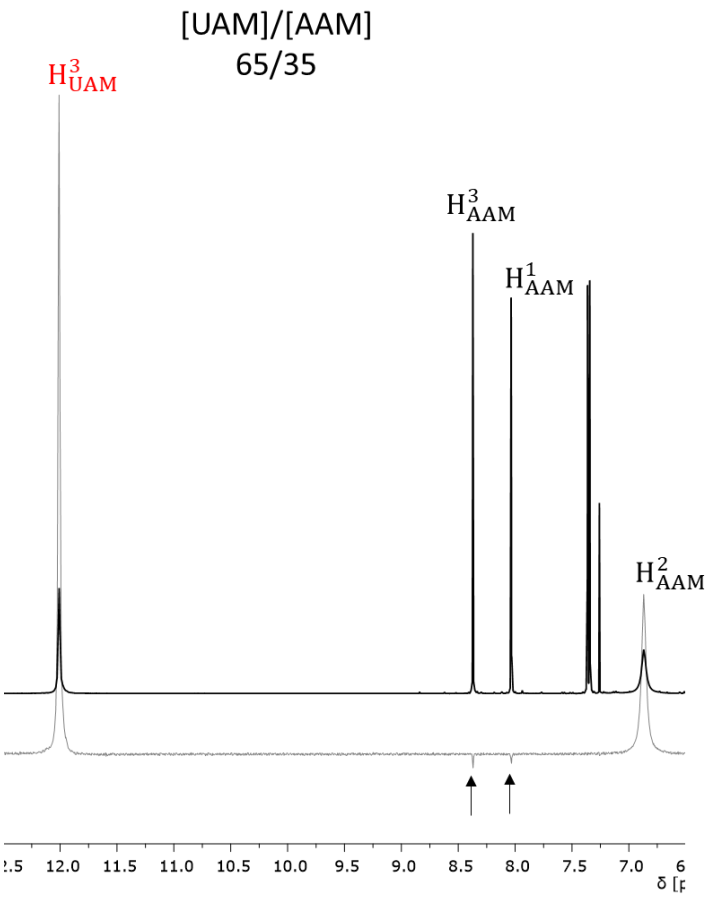


Figure 55. NOE-analysis of UAM-AAM base pairing in chloroform by means of irradiation of proton H^3 of UAM. The ratio given above the partial spectrum is the molar concentration ratios of the NAM present in the sample. The full spectrum can be found in appendix A7.

4.1.3 NMR analysis of base pairing between GAM and CAM

Since the base pairing stoichiometry of AAM-TAM and AAM-UAM rendered unexpected results, also GAM-CAM base pairing will be examined. Since GAM barely dissolves in chloroform, Job plots examining the GAM CAM base pairing could not be constructed. Measuring the NMR-spectra in DMSO would render a small change in δ resulting in even less reliable Job plots compared to analysis in chloroform. This means that the analysis will be based on theoretical models and NOE-analysis.

A possible GC_2 base pairing model is proposed in Figure 56. It combines the WC G-C base pairing with an alternative and merely theoretical form of base pairing. This model seems the most likely to cause a 1:2-stoichiometry as is entails the largest number of hydrogen bonds and is therefore the most stable obtainable GC_2-complex. The wavy bond in the figure represents a potential additional stabilizing interaction.

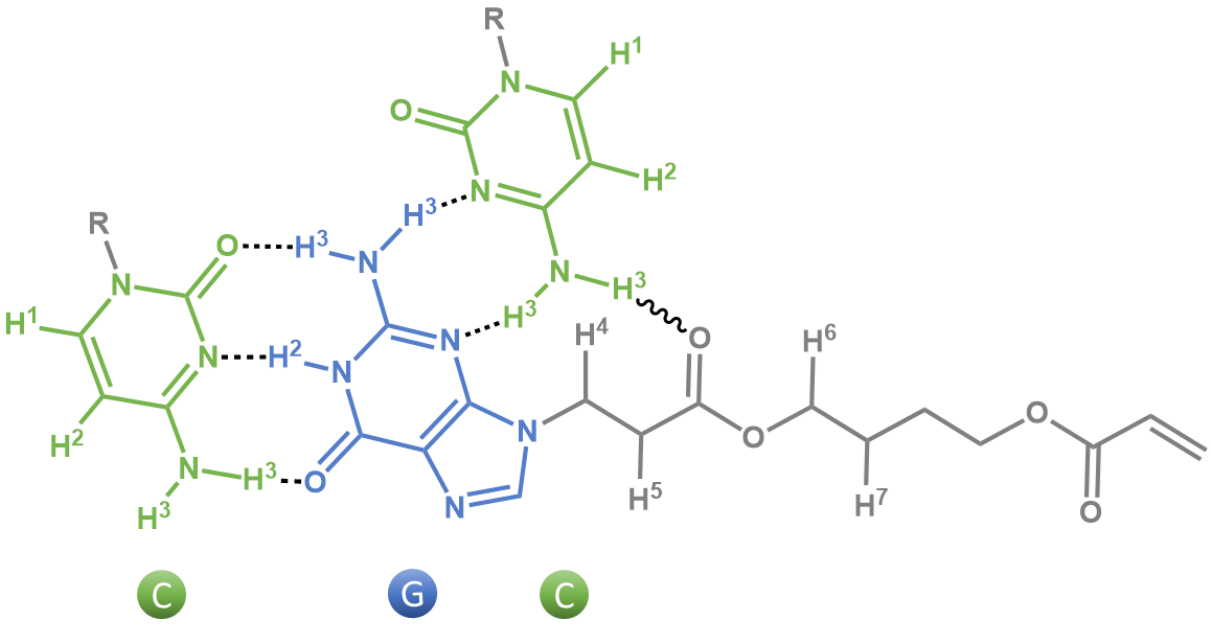


Figure 56. Combined model of Watson-Crick and an alternative G-C base pairing likely to explain the 1:2-stoichiometry observed in G-C base pairing.

To examine the stoichiometry of G-C base pairing, several NOE-experiments should be executed at two conditions. A first condition requires the concentration of CAM to be at least double of the molar concentration of GAM. The second condition requires near equimolar concentrations of GAM and CAM. Analysis of these conditions could reveal resembling concentration effects, compared to the A-T base pairing, on the complexation mechanism. A first NOE-analysis requires irradiation of H^2 and/or H^3 from CAM and/or irradiation of H^2 from GAM to observe WC base pairing at both conditions. This NOE-analysis has been executed, at the first condition, and indeed proved the presence of WC base pairing. The spectra can be found in appendices A8 and A9. Here, H^2 of GAM was irradiated and δ representing H^1 and H^2 from CAM appeared in the NOE-spectrum as can be seen in Figure 57. As H^1 and H^2 from the alternative base paired CAM are too far away from the irradiated proton, it can only be caused

by the protons of the WC base paired CAM. The same experiment should still be executed under the first condition to prove or deny the presence of this form of base pairing when concentration conditions are altered. Finally, the occurrence of the alternative base pairing structure should still be proven or denied. When H^1 and/or H^2 of CAM are irradiated and the NOE-spectrum shows the presence of protons H^4, H^5, H^6 and/or H^7 of GAM, the alternative base pairing model is plausible. According to the WC model, the distance between these protons would exceed the distance limitations of the NOE-effect meaning that the WC model cannot cause these signals to occur. The same conclusion could be drawn when the respective protons of GAM are irradiated and δ representing protons H^1 and/or H^2 are observed in the NOE-spectrum. Alternatively, a 2D-NOESY analysis would suffice as well, as it analyses the entire spectrum at once instead of irradiating each proton individually. The model, as described in Figure 56, remains fully theoretical until the mentioned NOE-experiments are executed..

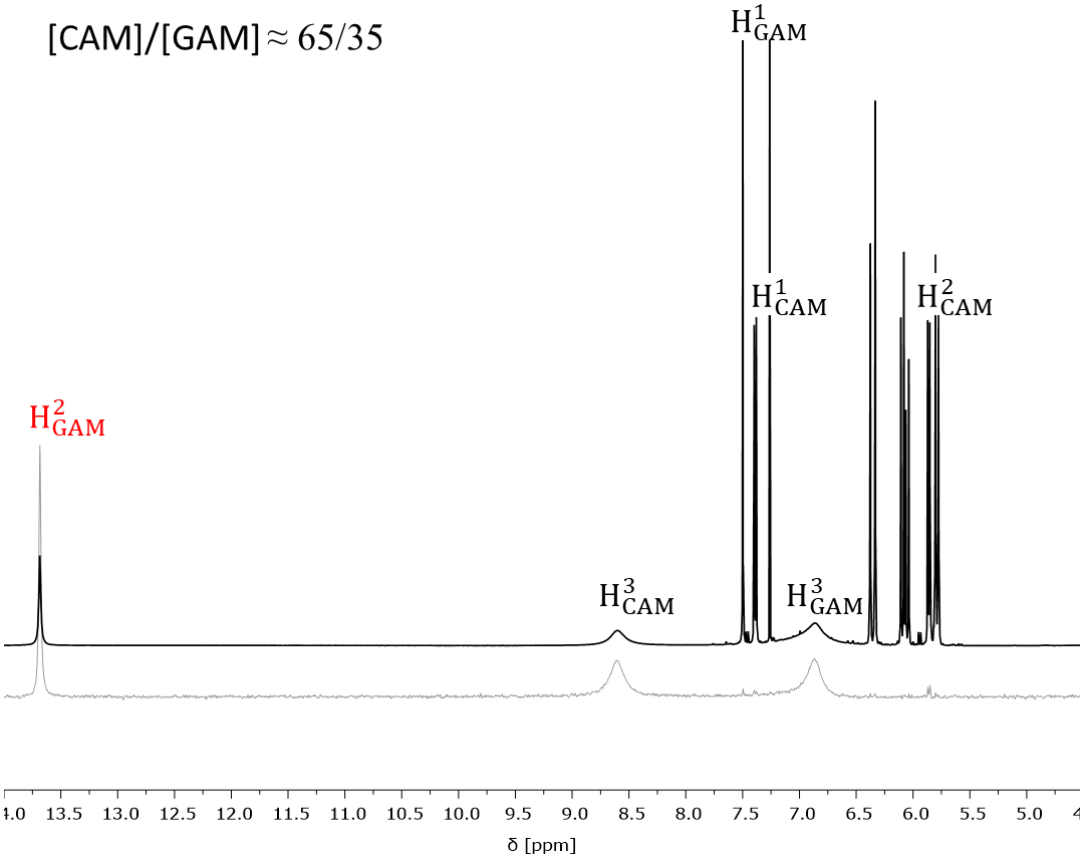


Figure 57. NOE-analysis of CAM-GAM base pairing in chloroform by means of irradiation of proton H^2 of GAM. The ratio given above the partial spectrum is the molar concentration ratios of the NAM present in the sample. The full spectrum can be found in appendix A9.

One could argue that the protons H^3 from CAM were not taken into account when determining ways to identify the WC and/or alternative base pairing model. These protons could also be used for the observation of the alternative base pairing model. There is however one major issue with these protons when an analysis is executed in chloroform. In this solvent, the δ corresponding to the H^3 protons of CAM are often weak or even absent. This phenomenon can be observed in the spectra given in appendices A10, A11, A12, A20 and A27. The phenomenon

could be caused by the exchange of protons between cytosine and the deuterated solvent. The protons on chloroform used for NMR-analysis are deuterated, meaning that they will not be visible in the NMR-spectrum. Additionally, chloroform has a fairly acidic character meaning that it is likely to donate its proton to a more partially negatively charged molecule/atom. When the nitrogen of CAM, onto which H^3 is bound, swaps its proton(s) for a deuterated proton from chloroform, present in great excess, the intensity of the δ corresponding to the H^3 will be reduced. It is most likely that this phenomenon causes the absence of the H^3-shift of CAM in NMR-analysis using chloroform as a solvent. This also means that this proton is not reliable for NOE-analysis.

4.1.4 Environmental effects on base pairing stoichiometry

The fact that the expected simple 1:1-stoichiometry of base pairing is not a certainty, as implied above, is problematic for future applications and procedures employing the synthetic nucleobase containing structures. Before, only mismatching was considered to be a potential problem towards these processes. Now, even the formation of base pairs between complementary NAM can cause problems. Note that, for a full understanding of the base pairing stoichiometry and mechanism, further research remains necessary. DOSY-NMR for example could be a useful analysis-tool to verify the presence of a 1:1 or 1:2 base pairing stoichiometry or even to verify the presence of an even more complex base pairing mechanism. Additionally, there are potentially several factors that might prevent the problem of formation of 1:2-base pairs. These parameters, given in explained in following paragraphs, should be examined as well.

A first factor, only relevant for templated polymerization, is the monomer concentration in the mixture. The results from this research imply that a high monomer concentration in the mixture could cause for the presence of a 1:2-stoichiometry in base pairing. This could result in faults in the produced MHB-SD by means of templated polymerization. Note that this assumption still has to be proven by DOSY-NMR for example. To prevent the stoichiometry problem from occurring, a characterization of base pairing between NAM and MHB-SD should be executed. Additionally, optimization of the monomer concentration in template-assisted polymerization will be key in synthesizing the desired sequence.

A second important factor is the temperature. As stated by the equation for Gibbs free energy, an increased temperature causes for a reduction in complexation. A decrease in entropy, caused by complexation, causes for an increase in the Gibbs free energy and therefore a decrease in complex-stability. The decrease in entropy is greater for 1:2-complexation, compared to 1:1-complexation, and an elevated temperature increases the influence of the entropy. Combination of these facts also state that 1:1-stoichiometric base pairs obtain an increased energetic preference over 1:2-stoichiometric base pairs and this energetic preference increases as the temperature increases. The influence of the temperature on the stoichiometry should still be examined by means of NOE-analysis to verify this theory.

A third factor is solvent polarity. Increasing the solvent polarity is assumed to have a similar effect as increasing the temperature. It should reduce the gained stability of base pairing, by stabilizing the bare NAM. This means that an increasing solvent polarity will cause for an increasing positive enthalpy difference. In other words, this means that again the decrease in entropy becomes increasingly important. Therefore, a polar solvent should also reduce the formation 1:2-complexation even further. To determine whether this theory is correct, NMR-experiments examining the solvent effects on the stoichiometry are conducted and analysed. Note that upon increasing the temperature or solvent polarity, equilibriums of both WC and Hoogsteen base pairing shift towards the side of free NAM. Tuning the environmental parameters will be important to eliminate the 1:2 base pairing but remain the presence of 1:1 base pairing.

A fourth factor that might be crucial to the phenomenon of 1:2-stoichiometry in base pairing is steric hindrance. Assume a template as depicted in Figure 58. When all free NAM have bound to their complement in the template, the most stable system is obtained. The neighbouring nucleobases from the thymine monomer, bound to the template trough hydrogen bonds, might cause for a steric hindrance to prevent the second TAM to form hydrogen bonds with the AAM from the template. Note that also this theory still requires further research. Examination of this theory could be performed by means of an affinity separation of a mixture of monomers or MHB-SD using a template attached onto a particle or substrate. Also, NOE-analysis could be of importance in examining this theory.

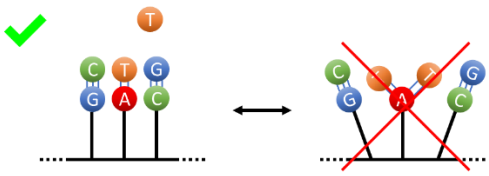


Figure 58. Template polymerization example, depicting the effect of steric hindrance inhibiting 1:2-base pairing

4.2 NMR-analysis of mismatching

In the second section of this chapter, the phenomenon of mismatching is examined. Characterization of influences on this phenomenon is key in determination of the optimal conditions for a templated polymerization or affinity separation to be executed. As mentioned previously, frequent mismatching causes for faults to occur in MHB-SD obtained by templated polymerization or affinity separation.

To examine occurrence of mismatching between the monomers, various mixtures of NAM's were examined by means of NOE-analysis and regular ^{1}H-NMR. Details concerning these mixtures of non-complementary NAM can be found in Table 3 and the corresponding ^{1}H-NMR and NOE-spectra can be found in appendices A10 to A28. Additional details concerning the observed δ of the different protons from ^{1}H-NMR analysis can be found in appendix A29.

The NOE-analysis of mismatching will provide a qualitative analysis of the phenomenon. Not all possible or occurring mismatching models will be handled as they are not of main importance in thesis. A multitude of reports containing extensive theoretical and practical research can already be found in literature. The main forms of mismatching, of importance in this research, are also reported in 2.2.2.

4.2.1 NOE-analysis of mismatching

The NOE-spectra from appendices A10 to A28 seem to show the absence of selectivity in base pairing between the different NAM. In almost every spectrum, irradiation of a proton belonging to a specific NAM gives rise to the δ of protons belonging to complementary and non-complementary NAM. An example is given in Figure 59. This figure depicts the NOE- and ^{1}H-NMR analysis of a sample containing AAM/CAM/GAM in a 4/4/2-ratio in DMSO. The NOE-spectrum was obtained by irradiation of H^{2} from GAM. As GAM is present in a much lower fraction, compared to the other two NAM, one might expect that only base pairing between the GAM and excessive amount of CAM would be observed upon irradiation of a proton from GAM. Certainly, as a GC-base pair entail the most stability and DMSO was used as solvent, which should weaken the mismatches. Instead, the δ of protons belonging to all three NAM appear in the NOE-spectrum, implying that hydrogen bonds are formed between GAM and CAM and between AAM and GAM.

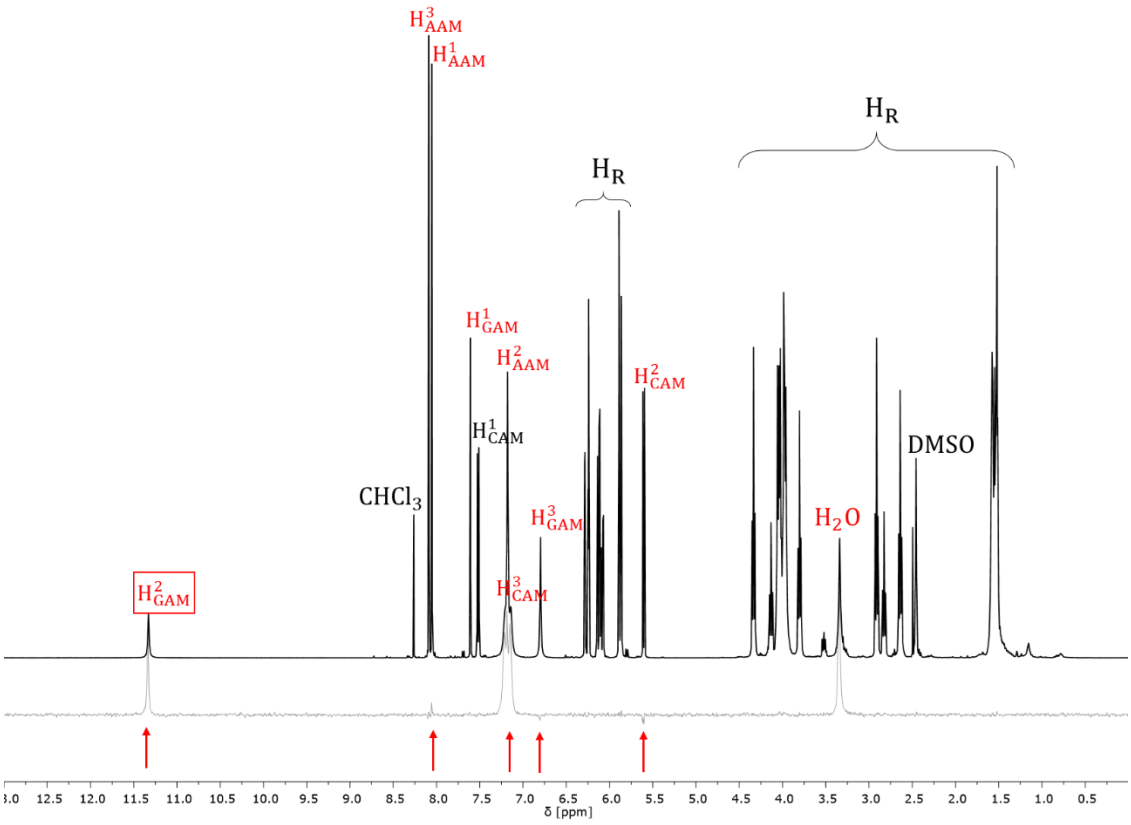


Figure 59. ^{1}H-NMR (black) and NOE- (grey) spectrum of a sample containing AAM/CAM/GAM in a 4/4/2-ratio in DMSO. The NOE-spectrum was obtained by irradiation of H^2 of GAM.

This might seem to rule out the idea of selectivity and therefore the use of templated polymerization and affinity separation to obtain the desired MHB-SD. However, note that the NOE-analysis is a qualitative analysis method. This means that it cannot be used to determine any distribution in formation of the different base pairs. When mismatching is observed in the NOE-spectrum, it is possible that these mismatches only represent a small fraction of the formed base pairs. This could still leave room for the presence of selectivity and condition optimization. For a more quantitative analysis of base pairing, the Δδ caused by the deshielding effect due to hydrogen bonding will be examined in mixtures containing different NAM in 0. NOE-analysis alone can only be used to determine presence of mismatching.

Despite the fact that the desired the desired selectivity, with complete absence of mismatching, is not observed in the NOE-spectra, two interesting phenomena can be observed in the spectra.

A first phenomenon is the strong interaction of the nucleobase with water. Even though water is barely present in the solvents, irradiation of the protons on host sites for hydrogen bonding, almost always gives rise to δ representing water in the NOE-spectra as shown in Figure 59. All the NOE-spectra showing this interaction with water can be found in appendices A1-A9, A12A12A16, A18, A21, A23-A26 and A28. This implies a strong interaction or hydrogen bonding to occur between the NAM and water. This strong interaction can potentially be used to obtain improved selectivity of base pairing between NAM. When the concentration of water in the mixture rises, interference with base pairing between the NAM might occur as it occupies binding sites of these molecules. Water should be added until mismatches do not or barely occur

anymore. Note that the concentration of water should also be small enough to prevent complete inhibition of base pairing between complementary NAM. It is important to know that the competition of hydrogen bonding with other NAM or water does not only affect the base pairing between non-complementary NAM. As mentioned before, increasing solvent polarity causes for a shift to occur in the base pairing equilibriums. Increasing the amount of water will cause for the equilibria between complementary and non-complementary NAM to shift towards the side of free NAM. The theory of adding the optimal amount of water merely relies on the potential presence of an increased stability or k_a in base pairing between complementary NAM. Therefore, a point could be reached where occurrence of mismatching is negligible and base pairing between complementary NAM still occurs. Additionally, an aqueous environment might also be ideal for the breaking hydrogen bonds. This can be useful when hydrogen bonds between the desired MHB-SD and its template have to be broken to isolate the desired product after affinity separation or templated polymerization.

A last phenomenon, observable in the NOE-spectra, is the occurrence of the δ belonging to the protons H^8, H^9 and H^{10} of the NAM. This can be observed in appendices A2, A4-A7, A9, A14-A16, A20, A21, A24, A26 and A28. This could imply the presence of an interaction between the acrylate tail and the nucleobases of the NAM as mentioned in 4.1.1. It is most likely that a proton, from a nucleobase, forms a hydrogen bond with the carbonyl oxygen as acceptor. Additionally, a weak interaction between the proton in the β-position of the respective carbonyl group and a nitrogen or carbonyl oxygen from a nucleobase, capable of forming hydrogen bonds, could occur. This type of pairing is depicted in Figure 60

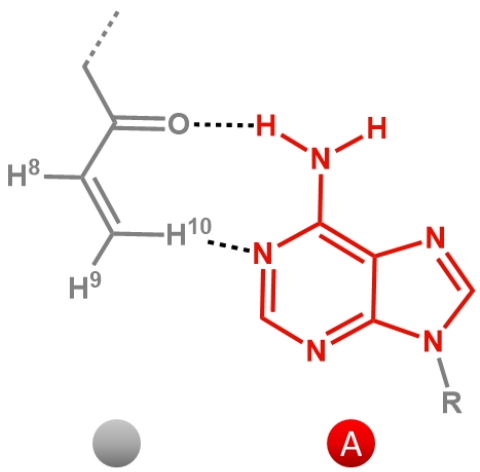


Figure 60. An example of interaction between a NAM (AAM, right) and the acrylate tail of another NAM-molecule (left).

4.2.2 ^{1}H-NMR analysis of mismatching

The Δδ, resulting from the ^{1}H-NMR analysis, will be analysed for each proton capable of hydrogen bonding. The shifts are depicted for AAM, TAM, GAM and CAM respectively in appendix A30, A31, A32 and A33. Note that the temperatures of most measurements, executed around room temperature, range from 20.1 °C to 22.2 °C. Additionally, the temperatures at which the shifts of the sole NAM were measured, range from 20.7 °C to 25.0 °C. The temperature, as mentioned in 2.2, affects the stability of hydrogen bonds and therefore also the observed Δδ. This causes for the need of caution when interpreting the results. Mainly, observed trends in hydrogen bonding strength can be discussed. Comparison of the exact Δδ of protons should be done with caution, while taking the potential effects of the temperature into account. Additionally, when preparing the samples, the effects of concentration ratio were not taken into account as it was not the intention of examining merely the influence of the concentration(ratios). The goals is to obtain an overall view of the effects of multiple parameters. When comparing exact Δδ of the samples, one should also take the potential effect of these ratios into account. Neglection of these ratios or the temperature influence could lead to false conclusions.

Before discussing the results, the importance and potential effect of the concentration-ratios on mismatching will be elucidated by means of two theoretical examples. Assume a mixture containing AAM/GAM/CAM in an 8/1/1-ratio. Due to the large excess of AAM present in the mixture, it is likely that a certain amount of this AAM will form mismatched base pairs with GAM and/or CAM. Additionally, due to the presence of very few GAM and CAM molecules, it is also likely that correct matches could occur less frequent as these molecules will be less likely to collide with each other in the mixture, resulting in a base pair. The formation of base pairs between GAM and CAM is lowered even further due to the occupation of binding sites on GAM and/or CAM due to base pairing with AAM, present in great excess. This would result in different Δδ for the hydrogen bonding protons from a situation where AAM/GAM/CAM are present in a 2/9/9-ratio for example. Here, the formation of mismatches will occur less frequent resulting in different Δδ, which would then lead to different conclusions. Another example is the following. Assume the presence of AAM/GAM/CAM in a 2/6/2-ratio. Due to an excess in GAM, with respect to CAM, a lot of free GAM molecules will be present even if all possible correct matches have been formed. Therefore, it is most likely that a certain amount of the remaining free GAM molecules will form base pairs with the free AAM molecules. This would again result in different Δδ and conclusions compared to a mixture containing AAM/GAM/CAM in a 1/1/1-ratio. These examples, even though focussing on the same monomer interactions, are likely to result in different conclusions due to the difference in monomer-ratios. Therefore, one should be cautious when interpreting the results from ^{1}H-NMR analysis. Analysis of the influence of concentration ratios on mismatching should therefore still be executed as this will not be handled in this research.

A first phenomenon, observed in the results of all NAM-mixtures depicted in Figure 61, is the dependency of the hydrogen bond stability on the solvent polarity. An increased Δδ is observed when the analysis is executed in $CDCl_3$, in comparison to analysis in DMSO, implying that the hydrogen bonds between NAM become more stable as the solvent polarity decreases. This means that a decrease in solvent polarity results in an equilibrium promoting base pairing, which is in line with the theory discussed in 2.2.3.2. This effect is underestimated in analysis of GAM. Here, the Δδ of all samples, regardless the used solvent, are determined by subtraction of the shift of the bare proton in DMSO. Analysis of the bare proton was not possible in chloroform, as GAM does not dissolve sufficiently for ^{1}H-NMR analysis. As the shift of the bare proton in chloroform would be higher, than in DMSO, the calculated Δδ is an underestimation. Unfortunately, mostly a strong increase in Δδ is observed for all kinds of base pairs, meaning that a strong non-polar solvent promotes the formation of base pairing between both complementary and non-complementary NAM.

Figure 61. Δδ of all protons determined by ^{1}H-NMR analysis at near room temperature. All Δδ have been calculated using the δ of the analysed proton in a pure solution containing solely the examined NAM in the respective solvent. Only for GAM also the mixtures analysed in chloroform are shown relative to their shift in DMSO as GAM does not dissolve in DMSO.

A second phenomenon, observed for the samples depicted in Figure 62, is the temperature dependency of the δ and therefore also temperature dependency of hydrogen bonding stability. Here, the Δδ is calculated as the difference between the shift of the proton in the sample mixture at elevated temperature and the shift of the proton in the same sample mixture at room temperature. The δ of the hydrogen bonding protons decreases as the temperature increases implying that an increased temperature will cause for a shift, in the previously mentioned equilibrium, towards the free NAM. This conclusion is in line with the theory discussed in 2.2.3.1 and the equation of Gibbs free energy.

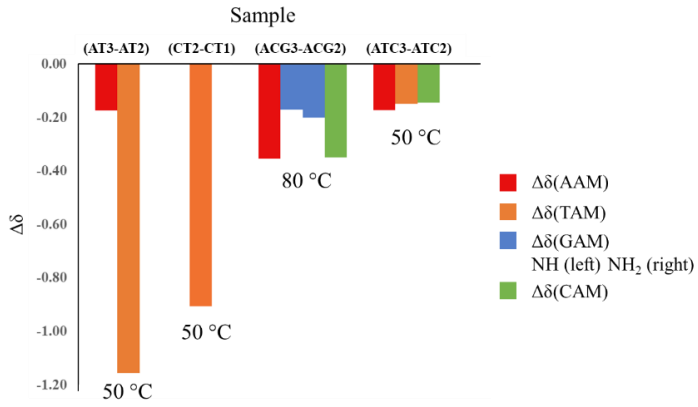


Figure 62. Δδ showing the influence of temperature on hydrogen bonding. The Δδ is calculated as the difference between the shift of the proton in the sample mixture at elevated temperature and the shift of the proton in the same sample at room temperature. For more detailed information concerning the samples, the reader is referred to appendices A30-A33. No shift in the CT sample is given for CAM as its proton was not observable in the ^{1}H-NMR spectrum.

Note that the data, given in this research, might exaggerate the effect of the temperature. The determined Δδ entails the temperature effects on the proton and temperature effects on its hydrogen bonding property. To get a more correct view on the effect of the temperature on solely hydrogen bonding another comparison should be made. A reference sample of a solution containing solely the respective NAM must be measured at room temperature and at the elevated temperature (δ_{RT} and δ_E). Subsequently, the proton of the respective NAM should be analysed in the desired mixture again at room and the same elevated temperature (δ'_{RT} and δ'_E). The Δδ, calculated by means of equation (13), is a result containing solely the influences of the temperature on hydrogen bonding.

$$\Delta\delta = (\delta_E - \delta_{RT}) - (\delta'_E - \delta'_{RT}) \qquad (13)$$

Finally, a third and important trend can be deduced from the Δδ-analysis. Even though, NOE-analysis revealed the occurrence of base pairing between both complementary and non-complementary NAM, analysis of Δδ shows the presence of a selectivity in base pairing is present, using the same solvent, are compared as shown in Figure 63. The first two samples showing this trend are samples AT5 and AC1. The proton of AAM, capable of hydrogen bonding, shows a higher Δδ in presence of TAM than in presence of CAM. This increased Δδ implies a greater affinity of AAM towards TAM compared to CAM. The temperature difference of 2.0 °C between the two samples and the slightly altering concentration-ratio only shows an underestimation of the selectivity. The same trend is observed when comparing the Δδ of the proton from TAM in sample AT4 and CT1. Their temperature and concentration ratio difference seems negligible and the respective Δδ show a greater affinity of TAM towards AAM than towards the non-complementary CAM. Again, the same conclusion could be formed when comparing the Δδ of the proton from CAM in sample GC2 and AC1. The observed Δδ of the proton from CAM is higher in presence of GAM in comparison to AAM. Finally, when comparing samples ACG1 and TCG1 another phenomenon can be observed. Adding AAM or

TAM into a mixture containing GAM and CAM in a near equimolar ratios, barely affects the shift of proton H^3 from GAM. This implies that GAM shows a significant affinity towards its complement CAM and would therefore barely be affected by adding non-complementary NAM into the mixture. Other data does not seem to be comparable as the concentration ratios of the NAM vary too much. Full comparison of the exact data can only be executed after examination of the effects of concentration ratios on base pairing. The importance and potential effects of these ratios were mentioned in the beginning of 4.2.2.

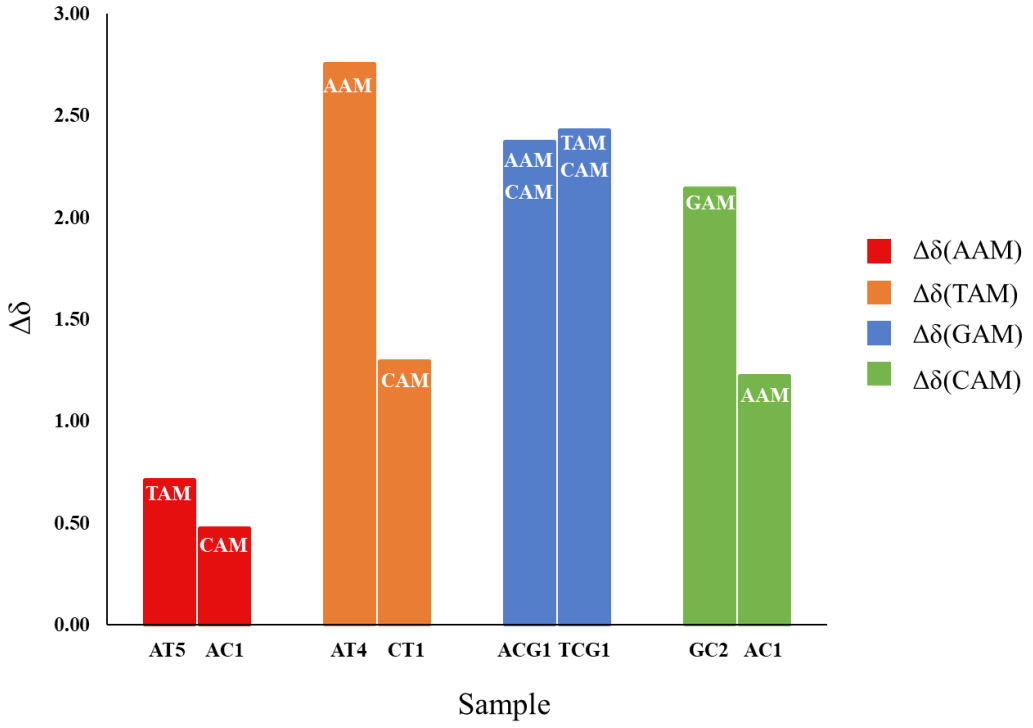


Figure 63. Δδ representing the selectivity of the NAM towards their complement.

4.2.3 Conclusion

Results from the Δδ-analysis of hydrogen bonding protons revealed that the expected trends caused by the temperature and solvent polarity are present in the base pairing mechanism. Additionally, a certain selectivity is observed for the base pair formation between complementary NAM. Presence of the selectivity implies that optimization of environmental parameters could cause for a more selective base pairing to occur between NAM. However, the acquired data is insufficient for the desired optimization. Therefore, a more elaborate characterization is required in future research for obtaining the fitting environmental conditions.

On a more critical note there is one form of mismatching, most likely present in the samples, that has not been discussed in this research. Mismatching between identical NAM was not handled due to the inability of the used analysis techniques to provide sufficient analysis of the phenomenon. Nevertheless, it is a form of mismatching that is most likely present in certain samples. The Δδ between free protons and protons hydrogen bonded to an identical NAM cannot be detected as analysis of a single free NAM by the NOE- or proton-NMR method is not possible. On the other hand, the Δδ was obtained using the δ of protons potentially hydrogen bonded to an identical NAM, as only one type of NAM was present in the reference samples. This means that the determined Δδ renders information about hydrogen bonding characteristics with non-identical NAM, potentially using the mismatching between identical NAM as reference. Therefore, almost all forms of base pairing seem to be more stable than base pairing between identical NAM. Only for analysis of protons from CAM in samples CT1, AGC3 and ATC3, TAM in samples AT3 and ATC3 and AAM in samples AT3, ACG3 and ATC3 the contrary seems to be true. Here, a negative Δδ is present for the examined protons. This could mean that hydrogen bonding of the respective NAM occurring in the sample, under the given conditions, is weaker/occurs less than hydrogen bonding between the identical NAM in the reference sample, under its given conditions. In the mentioned samples this potentially weaker or reduced formation of hydrogen bonding in the sample, is usually caused by the increased temperature. Only for analysis of CAM in sample CT1 it might effectively mean that hydrogen bonding between two CAM is more stable than hydrogen bonding between CAM and TAM under near identical conditions. An analysis technique, capable of detecting and characterizing base pairing between identical NAM is the SPR-analysis. In the SPR-analysis method a NAM could be bound to the surface of a substrate. When a stream containing the identical NAM, passes the surface, characterization of the interaction is possible. Unfortunately, use of SPR-analysis for this research was not possible. Note that theoretical work, for the SPR-analysis, is implemented in the literature study in section 2.3.2

Additionally, note that the analysis of temperature and solvent polarity on selective base pairing was executed on NAM. When these monomers are implemented into MHB-SD by means of RAFT-polymerization using DPE-TTC as CTA, similar effects/trends of the environmental parameters can be expected. However, the magnitude of these parameters on base pairing between the nucleobases implemented in MHB-SD is likely to differ from influence on base pairing between NAM. This can be the result of a lot of factors. A first factor is the influence

of the CTA-structure in the MHB-SD on the polarity of the molecule. Additionally, there is a possibility of interactions to occur between the nucleobases and the sulphur atoms present in the MHB-SD. Furthermore, the mobility of nucleobases in MHB-SD can also influence base pairing between the MHB-SD. Due to the fact that now multiple nucleobases will be locked into one MHB-SD, it might be possible that perfect alignment of the MHB-SD to form a complex will become much more difficult. On the other hand, the increased amount of hydrogen bonds in the resulting complexes could cause for an increased stability of these complexes. All these mentioned influences are merely theoretical but could be important. Therefore, exact characterization of base pairing between MHB-SD or between MHB-SD and NAM remains necessary before using template-assisted polymerization and affinity separation.

4.3 Template-assisted polymerization vs. affinity separation

Now that certain phenomena have been revealed and certain assumptions have been verified, a final comparison between the template-assisted polymerization and affinity separation will be made.

Firstly, there is the problem of the possible occurring 1:2-stoichiometry in base pairing between nucleobases, implemented in the synthetic structures. Even though, temperature, solvent polarity and concentration(-ratio) of the monomers might be tuned to reduce the occurrence of this undesired stoichiometry, it is yet another factor to take into account. This unwanted stoichiometry might cause for bigger problems in the template-assisted polymerization, compared to the affinity separation. Template-assisted polymerization relies on bringing the monomers into close presence of each other in the desired sequence and afterwards covalently attaching these monomers. Due to the potential 1:2-stoichiometry, errors in the synthesised MHB-SD could occur. As mentioned before, steric hindrance could reduce/prevent the occurrence of this undesired stoichiometry. To verify if the problem occurs, further research is required and a trial template-assisted polymerization could be executed to examine which faults occur. In the process of affinity separation the undesired stoichiometry could be less of a problem. Here, the steric hindrance could have an influence as well. Additionally, in affinity separation the nucleobases are ‘locked’ in a certain position into the MHB-SD. Therefore, they have a reduced freedom of motion which could prevent them from obtaining the right orientation for the occurrence of a second base pairing onto a nucleobase in the template. This is however merely an assumption and could easily be tested by means of a trial affinity separation. Finally, the concentration-ratio of monomers will not be a problem in the affinity separation as the monomers have already been polymerized into the MHB-SD.

Secondly, tuning the temperature and solvent polarity will be important in optimizing the selectivity of base pairing between the nucleobases as well. However, there is one major problem in again the template-assisted polymerization. Due to increasing the temperature or solvent polarity, for obtaining selective base pairing, also the association constant for base pairing between complementary nucleobases will be lowered. Therefore, increasing the temperature and solvent polarity will also decrease the amount of formed base pairs between

complementary nucleobases. This could result in incomplete base pairing of free NAM to the template which could then result in faults in the synthesized MHB-SD as depicted in Figure 64. This means that tuning of the environment will be very difficult if not impossible to obtain an operational template-assisted polymerization technique. When looking at the principle of the affinity separation one can clearly see that this problem is not present in this process.

In theory, when both methods would work perfectly as described in theory, the template-assisted polymerization would be the best option for increasing the yield of the process. However, when taking problems concerning hydrogen bonding into account, it is likely that the affinity separation will be the better option. Note that these assumptions are based on an unverified theory and on analysis of hydrogen bonding between NAM. Examination of hydrogen bonding between nucleobases implemented in MHB-SD and/or NAM should still be executed to come to the wright conclusions.

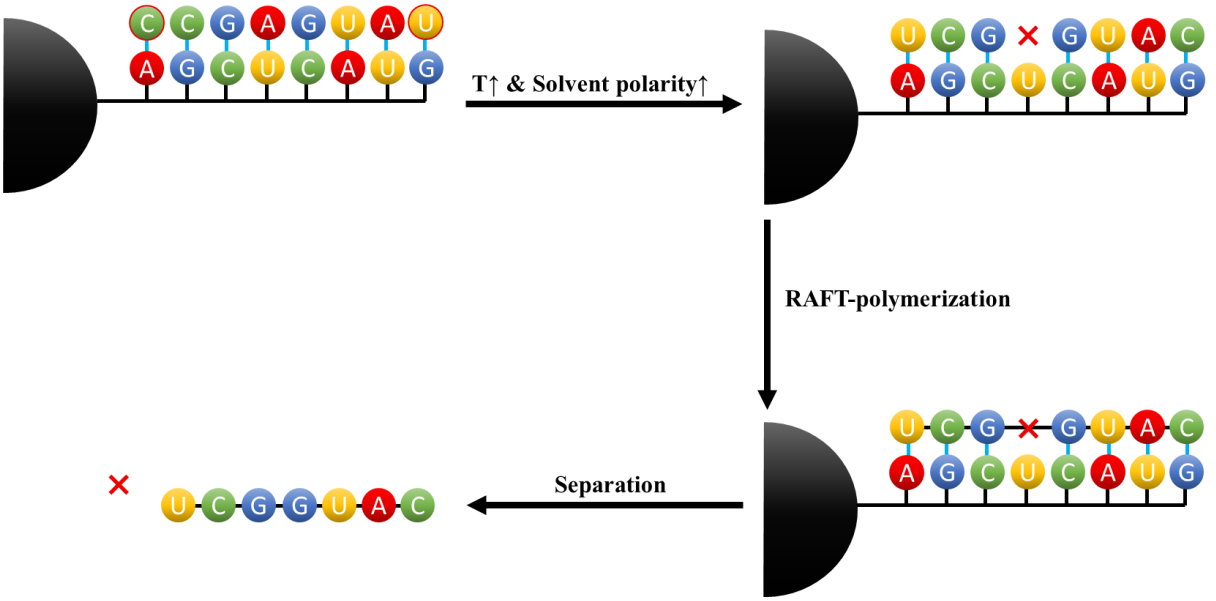


Figure 64. An example of breaking hydrogen bonds between complementary nucleobases upon increasing the temperature and/or solvent polarity to disrupt mismatches.

Chapter 5. Conclusion and outlook

The first objective, examined in this thesis, was finding analysis methods that could be used to detect and enable characterization of the formation of hydrogen bonding or base pairing between NAM. In the literature study, multiple methods have been listed and explained, that could be used for this examination. Two of these methods, regular ^{1}H-NMR and NOE-NMR, have successfully been used in this research to determine the occurrence of base pairing. NOE-analysis has proven to be useful in the qualitative examination. It enabled detection of base pairing and clarification of the occurring model of base pairing. Regular ^{1}H-NMR was used for quantitative analysis. The magnitude of the change in $\Delta\delta$ enabled examination of trends, present in base pairing. Even though these methods have proven to be useful, further verification and characterization remains necessary.

Examination of change in $\Delta\delta$ should be further verified by other methods to confirm whether conclusions based on this method are correct. Affinity separations of mixtures of NAM, might prove to be useful for this verification. Affinity separation can also be used to determine and characterize base pair formation between identical NAM. DOSY-NMR could be used for further verification of the occurring complexation model as it can render information concerning the hydrodynamic volume of a complex. Furthermore, SPR-analysis might be of great importance as well. The results of this analysis can render information about the thermodynamic behaviour of base pairing between identical, complementary and non-complementary NAM and MHB-SD. Additionally, this method can be used to confirm results acquired by means of $\Delta\delta$-analysis.

The second objective was the determination of optimal conditions for enabling selective base pairing on one hand and breaking of hydrogen bonds on the other hand. Even though full characterization was not accomplished, important phenomena have been revealed. A first phenomenon is the problematic and unexpected deviation from the 1:1-stoichiometry of base pairing. The potential 1:2 stoichiometry, combining the Watson-Cricks and Hoogsteen model in AAM-TAM and AAM-UAM base pairing, adds another problem into the determination of the ideal conditions for base pairing in applications such as templated polymerization and affinity separation. Further research on the influence of temperature, solvent polarity and concentration-ratio of the nucleobases on this the base pairing stoichiometry, should be executed to prevent occurrence of this unwanted 1:2-stoichiometry in base pairing. Furthermore, based on the results of this thesis, the expected influences of temperature and solvent polarity on base pairing, did seem to occur. An increase in temperature and solvent polarity seemed to weaken the formed hydrogen bonds between the nucleobases. Additionally, the importance of examination/characterization of the effects of the concentration-ratio of NAM

were mentioned as well. The ratio might have an influence on the stoichiometry of base pairing and on the formation of mismatches.

Important fundamental research, concerning base pairing between NAM, has been executed in this thesis. Additionally, important statements have been made for execution of further research. This future research is essential for full characterization of base pairing between NAM and MHB-SD. After the required characterization, templated polymerization and affinity separation can be examined as alternatives for the current SUMI-procedure for obtaining the desired MHB-SD. When the problem of the low yield of the current procedure is resolved, actual applications of the MHB-SD could be examined.

Bibliography

[1] UHasselt. "POLYMER REACTION DESIGN (PRD)." UHasselt. https://www.uhasselt.be/UH/IMO/Visit-the-groups/Polymer-reaction-design-(PRD).html (accessed 2020).

[2] Á. Anguera De Sojo, J. Ares, M. A. Martínez, J. Pazos, S. Rodríguez, and J. G. Zato, "Serendipity and the Discovery of DNA," *Foundations of Science*, vol. 19, no. 4, pp. 387-401, 2014, doi: 10.1007/s10699-014-9348-0.

[3] P. J. Russel, B. Wilbur, Ed. *iGenetics: A Molecular Approach*, third ed. Pearson (in English), 2010, p. 828.

[4] J. D. Watson and F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature*, vol. 171, no. 4356, pp. 737-738, Apr 25 1953, doi: 10.1038/171737a0.

[5] H. Tan, C. Xiao, J. Sun, D. Xiong, and X. Hu, "Biological self-assembly of injectable hydrogel as cell scaffold via specific nucleobase pairing," *Chem Commun (Camb)*, vol. 48, no. 83, p. 10289, Oct 25 2012, doi: 10.1039/c2cc35449g.

[6] C.-C. Cheng, F.-C. Chang, S. Dai, Y.-L. Lin, and D.-J. Lee, "Bio-complementary supramolecular polymers with effective self-healing functionality," *RSC Adv.*, vol. 5, no. 110, pp. 90466-90472, 10/16 2015, doi: 10.1039/C5RA19708B.

[7] X. Ye, X. Li, Y. Shen, G. Chang, J. Yang, and Z. Gu, "Self-healing pH-sensitive cytosine- and guanosine-modified hyaluronic acid hydrogels via hydrogen bonding," *Polymer*, vol. 108, pp. 348-360, 2017/01/13/ 2017, doi: https://doi.org/10.1016/j.polymer.2016.11.063.

[8] A. Harguindey, D. W. Domaille, B. D. Fairbanks, J. Wagner, C. N. Bowman, and J. N. Cha, "Synthesis and Assembly of Click-Nucleic-Acid-Containing PEG-PLGA Nanoparticles for DNA Delivery," *Advanced Materials*, vol. 29, no. 24, p. 1700743, Jun 2017, doi: 10.1002/adma.201700743.

[9] S. Cheng, M. Zhang, N. Dixit, R. B. Moore, and T. E. Long, "Nucleobase Self-Assembly in Supramolecular Adhesives," *Macromolecules*, vol. 45, no. 2, pp. 805-812, 2012, doi: 10.1021/ma202122r.

[10] Q. Ma, D. Lee, Y. Q. Tan, G. Wong, and Z. Gao, "Synthetic genetic polymers: advances and applications," *Polymer Chemistry*, 10.1039/C6PY01075J vol. 7, no. 33, pp. 5199-5216, 2016, doi: 10.1039/C6PY01075J.

[11] J. Vandenbergh, G. Reekmans, P. Adriaensens, and T. Junkers, "Synthesis of Sequence-Defined Acrylate Oligomers via Photoinduced Copper-Mediated Radical Monomer Insertions," *Chem. Sci.*, vol. 6, no. 10, pp. 5753-5761, 07/03 2015, doi: 10.1039/C5SC02035B.

[12] L. Maes, D. Massana Roquero, L. Pitet, P. Adriaensens, and T. Junkers, "Sequence-defined nucleobase containing oligomers via reversible addition-fragmentation chain transfer single monomer addition," *Polymer Chemistry*, vol. 11, no. 12, pp. 2027-2033, 2020, doi: 10.1039/c9py01853k.

[13] Y. Yong Tan, "The synthesis of polymers by template polymerization," *Progress in Polymer Science*, vol. 19, no. 4, pp. 561-588, 1994/01/01/ 1994, doi: https://doi.org/10.1016/0079-6700(94)90028-0.

[14] R. Saito, "Combination of template polymerization and atom transfer radical polymerization: Strategy for synthesis of specifically structural polymers," *Polymer*, vol. 49, no. 11, pp. 2625-2631, 2008, doi: 10.1016/j.polymer.2008.03.011.

[15] R. Saito and H. Kobayashi, "Synthesis of Polymers by Template Polymerization. 2. Effects of Solvent and Polymerization Temperature," *Macromolecules*, vol. 35, no. 19, pp. 7207-7213, 2002, doi: 10.1021/ma0121058.

[16] R. McHale and R. K. O'Reilly, "Nucleobase Containing Synthetic Polymers: Advancing Biomimicry via Controlled Synthesis and Self-Assembly," *Macromolecules,* vol. 45, no. 19, pp. 7665-7675, 2012, doi: 10.1021/ma300895u.
[17] Y. Kang, A. Lu, A. Ellington, M. C. Jewett, and R. K. O'Reilly, "Effect of Complementary Nucleobase Interactions on the Copolymer Composition of RAFT Copolymerizations," *ACS Macro Letters,* vol. 2, no. 7, pp. 581-586, 2013, doi: 10.1021/mz4001833.
[18] T. Brown, G. A. Leonard, E. D. Booth, and G. Kneale, "Influence of pH on the conformation and stability of mismatch base-pairs in DNA," *Journal of molecular biology,* vol. 212, no. 3, pp. 437-40, Apr 5 1990, doi: 10.1016/0022-2836(90)90320-L. 437-440.
[19] B. G. Alberding and B. J. Lear, "Concentration-Dependent Dynamics of Hydrogen Bonding between Acetonitrile and Methanol As Determined by 1D Vibrational Spectroscopy," *The Journal of Physical Chemistry A,* vol. 118, no. 25, pp. 4363-4371, 2014/06/26 2014, doi: 10.1021/jp4110147.
[20] R. K. O'Reilly, A. J. Turberfield, and T. R. Wilks, "The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery," *Accounts of Chemical Research,* vol. 50, no. 10, pp. 2496-2509, Oct 17 2017, doi: 10.1021/acs.accounts.7b00280.
[21] H. Bisht, "LIVING FREE-RADICAL POLYMERIZATION—A REVIEW," *Journal of Macromolecular Science Part C Polymer Reviews,* vol. 41, no. 3, p. 139, 07/31 2001, doi: 10.1081/MC-100107774.
[22] A. J. O'Lenick and T. G. O'Lenick, *Organic Chemistry for Cosmetic Chemists*. Allured Publishing, 2008.
[23] S. Perrier, "50th Anniversary Perspective: RAFT Polymerization—A User Guide," *Macromolecules,* vol. 50, no. 19, pp. 7433-7447, 2017, doi: 10.1021/acs.macromol.7b00767.
[24] D. A. Shipp, "Reversible-Deactivation Radical Polymerizations," *Polymer Reviews,* vol. 51, no. 2, pp. 99-103, 2011/04/22 2011, doi: 10.1080/15583724.2011.566406.
[25] J. Vandenbergh, G. Reekmans, P. Adriaensens, and T. Junkers, "Synthesis of sequence controlled acrylate oligomers via consecutive RAFT monomer additions," *Chem Commun (Camb),* vol. 49, no. 88, p. 10358, Nov 14 2013, doi: 10.1039/c3cc45994b.
[26] J. Chiefari *et al.*, "Living Free-Radical Polymerization by Reversible Addition−Fragmentation Chain Transfer: The RAFT Process," *Macromolecules,* vol. 31, no. 16, pp. 5559-5562, 1998, doi: 10.1021/ma9804951.
[27] J. Šponer and F. Lankas, "Nucleic acids in complexes," in *Computational studies of RNA and DNA*, no. 2). Netherlands: Springer, 2006, ch. 4, pp. 21-31.
[28] A. Goto and T. Fukuda, "Kinetics of living radical polymerization," *Progress in Polymer Science,* vol. 29, no. 4, pp. 329-385, 2004, doi: 10.1016/j.progpolymsci.2004.01.002.
[29] D. J. Keddie, G. Moad, E. Rizzardo, and S. H. Thang, "RAFT Agent Design and Synthesis," *Macromolecules,* vol. 45, no. 13, pp. 5321-5342, 2012, doi: 10.1021/ma300410v.
[30] M. Benaglia, J. Chiefari, Y. K. Chong, G. Moad, E. Rizzardo, and S. H. Thang, "Universal (Switchable) RAFT Agents," *Journal of the American Chemical Society,* vol. 131, no. 20, pp. 6914-6915, 2009/05/27 2009, doi: 10.1021/ja901955n.
[31] G. Moad and C. Barner-Kowollik, "The Mechanism and Kinetics of the RAFT Process: Overview, Rates, Stabilities, Side Reactions, Product Spectrum and Outstanding Challenges," *Handbook of RAFT Polymerization,* pp. 51-104, 09/16 2008, doi: 10.1002/9783527622757.ch3.
[32] A. Feldermann, M. L. Coote, M. H. Stenzel, T. P. Davis, and C. Barner-Kowollik, "Consistent Experimental and Theoretical Evidence for Long-Lived Intermediate Radicals in Living Free Radical Polymerization," *Journal of the American Chemical Society,* vol. 126, no. 48, pp. 15915-15923, Dec 8 2004, doi: 10.1021/ja046292b.
[33] L. G. wIlliams, K. Sullivan, and J. Tsanaktsidis. "Select RAFT Agents for Making Well-Defined Functionalized Polymers." Merck. (accessed.
[34] J. Li, Z. Wang, Z. Hua, and C. Tang, "Supramolecular nucleobase-functionalized polymers: synthesis and potential biological applications," *Journal of Materials Chemistry B,* vol. 8, no. 8, pp. 1576-1588, Feb 26 2020, doi: 10.1039/c9tb02393c.
[35] Z. Hua, A. Pitto-Barry, Y. Kang, N. Kirby, T. R. Wilks, and R. K. O'Reilly, "Micellar nanoparticles with tuneable morphologies through interactions between nucleobase-containing

synthetic polymers in aqueous solution," *Polymer Chemistry*, vol. 7, no. 25, pp. 4254-4262, 2016, doi: 10.1039/c6py00716c.

[36] M. Sedlák, P. ŠImůNek, and M. Antonietti, "Synthesis and15N NMR characterization of 4-vinylbenzyl substituted bases of nucleic acids," *Journal of Heterocyclic Chemistry*, vol. 40, no. 4, pp. 671-675, 2003, doi: 10.1002/jhet.5570400418.

[37] J. J. Haven, J. A. De Neve, and T. Junkers, "Versatile Approach for the Synthesis of Sequence-Defined Monodisperse 18- and 20-mer Oligoacrylates," *ACS Macro Letters*, vol. 6, no. 7, pp. 743-747, 2017/07/18 2017, doi: 10.1021/acsmacrolett.7b00430.

[38] J. Lawrence *et al.*, "A Versatile and Efficient Strategy to Discrete Conjugated Oligomers," *Journal of the American Chemical Society*, vol. 139, no. 39, pp. 13735-13739, 2017/10/04 2017, doi: 10.1021/jacs.7b05299.

[39] W. Olson, A. Colasanti, X.-J. Lu, and V. Zhurkin, "Watson–Crick Base Pairs: Character and Recognition," in *Wiley Encyclopedia of Chemical Biology*, 2008.

[40] E. Anslyn and D. Dougherty, "Solutions and non-covalent binding forces," in *Modern Physical Organic Chemistry*, J. Murdzek Ed. Sausalito, California: University Science Books, 2006, ch. 3, pp. 145-206.

[41] T. Chatake, A. Ono, Y. Ueno, A. Matsuda, and A. Takénaka, "Crystallographic studies on damaged DNAs. I. an N6-methoxyadenine residue forms a watson-crick pair with a cytosine residue in a B-DNA duplex," *J Mol Biol*, vol. 294, no. 5, pp. 1215-1222, Dec 17 1999, doi: 10.1006/jmbi.1999.3303.

[42] J. Poater, M. Swart, C. F. Guerra, and F. Matthias Bickelhaupt, "Solvent effects on hydrogen bonds in Watson–Crick, mismatched, and modified DNA base pairs," *Computational and Theoretical Chemistry*, vol. 998, pp. 57-63, 2012/10/15/ 2012, doi: https://doi.org/10.1016/j.comptc.2012.06.003.

[43] L. F. Sukhodub, "Interactions and hydration of nucleic acid bases in a vacuum. Experimental study," *Chemical Reviews*, vol. 87, no. 3, pp. 589-606, 1987/06/01 1987, doi: 10.1021/cr00079a006.

[44] I. K. Yanson, A. B. Teplitsky, and L. F. Sukhodub, "Experimental studies of molecular interactions between nitrogen bases of nucleic acids," *Biopolymers*, vol. 18, no. 5, pp. 1149-1170, 1979/05/01 1979, doi: 10.1002/bip.1979.360180510.

[45] M. C. Wahl and M. Sundaralingam, "C-H....O hydrogen bonding in biology," *Trends in Biochemical Sciences*, vol. 22, no. 3, pp. 97-102, Mar 1997, doi: 10.1016/S0968-0004(97)01004-9.

[46] B. Sathyamoorthy *et al.*, "Insights into Watson-Crick/Hoogsteen breathing dynamics and damage repair from the solution structure and dynamic ensemble of DNA duplexes containing m1A," (in eng), *Nucleic Acids Res*, vol. 45, no. 9, pp. 5586-5601, May 19 2017, doi: 10.1093/nar/gkx186.

[47] R. Johnson and L. Prakash, "Biochemical evidence for the requirement of Hoogsteen base pairing for replication by human DNA polymerase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 30, pp. 10466-71, 08/01 2005, doi: 10.1073/pnas.0503859102.

[48] E. N. Nikolova, E. Kim, A. A. Wise, P. J. O'Brien, I. Andricioaei, and H. M. Al-Hashimi, "Transient Hoogsteen base pairs in canonical duplex DNA," (in eng), *Nature*, vol. 470, no. 7335, pp. 498-502, Feb 24 2011, doi: 10.1038/nature09775.

[49] R. D. Kolodner, "Mismatch repair: mechanisms and relationship to cancer susceptibility," *Trends in Biochemical Sciences*, vol. 20, no. 10, pp. 397-401, 1995/10/01/ 1995, doi: https://doi.org/10.1016/S0968-0004(00)89087-8.

[50] E. Renkonen *et al.*, "Altered Expression of MLH1, MSH2, and MSH6 in Predisposition to Hereditary Nonpolyposis Colorectal Cancer," *Journal of Clinical Oncology*, vol. 21, no. 19, pp. 3629-3637, 2003/10/01 2003, doi: 10.1200/JCO.2003.03.181.

[51] G. Rossetti, P. D. Dans, I. Gomez-Pinto, I. Ivani, C. Gonzalez, and M. Orozco, "The structural impact of DNA mismatches," *Nucleic Acids Res*, vol. 43, no. 8, pp. 4309-4321, Apr 30 2015, doi: 10.1093/nar/gkv254.

[52] J. W. Szostak, "The eightfold path to non-enzymatic RNA replication," *Journal of Systems Chemistry*, vol. 3, no. 1, p. 2, 2012, doi: 10.1186/1759-2208-3-2.

[53] F. H. C. Crick, "Codon—anticodon pairing: The wobble hypothesis," *Journal of Molecular Biology*, vol. 19, no. 2, pp. 548-555, 1966/08/01/ 1966, doi: https://doi.org/10.1016/S0022-2836(66)80022-0.
[54] C. R. Geyer, T. R. Battersby, and S. A. Benner, "Nucleobase Pairing in Expanded Watson-Crick-like Genetic Information Systems," *Structure*, vol. 11, no. 12, pp. 1485-1498, Dec 2003, doi: 10.1016/j.str.2003.11.008.
[55] R. R. Sinden, "CHAPTER 1 - Introduction to the Structure, Properties, and Reactions of DNA," in *DNA Structure and Function*, R. R. Sinden Ed. San Diego: Academic Press, 1994, pp. 1-57.
[56] A. S. Gilbert, "Hydrogen Bonding and Other Physicochemical Interactions Studied By IR and Raman Spectroscopy," in *Encyclopedia of Spectroscopy and Spectrometry*, J. C. Lindon Ed. Oxford: Elsevier, 1999, pp. 837-843.
[57] B. D. Sattin, "DNA base pair resolution by single molecule force spectroscopy," *Nucleic Acids Res*, vol. 32, no. 16, pp. 4876-4883, 2004, doi: 10.1093/nar/gkh826.
[58] T.-B. Zhang, C.-L. Zhang, Z.-L. Dong, and Y.-F. Guan, "Determination of Base Binding Strength and Base Stacking Interaction of DNA Duplex Using Atomic Force Microscope," *Sci Rep*, vol. 5, p. 9143, Mar 16 2015, doi: 10.1038/srep09143.
[59] Merck. "Oligonucleotide Melting Temperature." Sigma-Aldrich. (accessed 17th April, 2020).
[60] J. Marmur and P. Doty, "Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature," *Journal of Molecular Biology*, vol. 5, no. 1, pp. 109-118, 1962/07/01/ 1962, doi: https://doi.org/10.1016/S0022-2836(62)80066-7.
[61] S. B. Prakash, D. Mousumi, and P. Godavarthi, "Identification and Classification of Microbes," in *Microbes: Concepts and Applications*, (Wiley Online Books. Hoboken, New Jersey: John Wiley & Sons, 2012, ch. 4, pp. 275-337.
[62] A. J. A. Aquino, D. Tunega, G. Haberhauer, M. H. Gerzabek, and H. Lischka, "Solvent Effects on Hydrogen BondsA Theoretical Study," *The Journal of Physical Chemistry A*, vol. 106, no. 9, pp. 1862-1871, 2002, doi: 10.1021/jp013677x.
[63] E. Romero and F. Hernandez, "Solvent effect on the Intermolecular Proton Transfer of the Watson and Crick Guanine-Cytosine and Adenine-Thymine base pairs: A Polarizable Continuum Model study," *Physical Chemistry Chemical Physics*, vol. 20, no. 2, pp. 1198-1209, 12/08 2017, doi: 10.1039/C7CP05356H.
[64] Y. Boulard, J. A. H. Cognet, J. Gabarro-Arpa, M. Le Bret, L. C. Sowers, and G. V. Fazakerley, "The pH dependent configurations of the C.A mispair in DNA," *Nucleic Acids Res*, vol. 20, no. 8, pp. 1933-1941, Apr 25 1992, doi: 10.1093/nar/20.8.1933.
[65] Y. Boulard, J. A. H. Cognet, and G. V. Fazakerley, "Solution structure as a function of pH of two central mismatches, C·T and C·C, in the 29 to 39 K-ras gene sequence, by nuclear magnetic resonance and molecular dynamics11Edited by T. Tinoco," *Journal of Molecular Biology*, vol. 268, no. 2, pp. 331-347, 1997/05/02/ 1997, doi: https://doi.org/10.1006/jmbi.1997.0975.
[66] P. Charisiadis, V. Kontogianni, C. Tsiafoulis, A. Tzakos, M. Siskos, and I. Gerothanassis, "1H-NMR as a Structural and Analytical Tool of Intra- and Intermolecular Hydrogen Bonds of Phenol-Containing Natural Products and Model Compounds," *Molecules*, vol. 19, no. 9, pp. 13643-13682, Sep 2 2014, doi: 10.3390/molecules190913643.
[67] J. Kroon, L. M. J. Kroon-Batenburg, B. R. Leeflang, and J. F. G. Vliegenthart, "Intramolecular versus intermolecular hydrogen bonding in solution," *Journal of Molecular Structure*, vol. 322, pp. 27-31, 1994/06/16/ 1994, doi: https://doi.org/10.1016/0022-2860(94)87018-7.
[68] H. Günther and H. Günther, *NMR Spectroscopy : Basic Principles, Concepts and Applications in Chemistry*. Somerset, GERMANY: John Wiley & Sons, Incorporated, 2013.
[69] G. Bydder and R. Steiner, "NMR imaging of the brain," *Neuroradiology*, vol. 23, no. 5, pp. 231-40, 02/01 1982, doi: 10.1007/BF00339389.
[70] S. Scheiner, "Assessment of the Presence and Strength of H-Bonds by Means of Corrected NMR," *Molecules*, vol. 21, no. 12, p. 1426, Oct 27 2016, doi: 10.3390/molecules21111426.
[71] S. Grzesiek and E. D. Becker, "Hydrogen Bonding," *eMagRes*, 2011/12/15 2011, doi: doi:10.1002/9780470034590.emrstm0216.pub2.
[72] E. Breitmaier and A. Sinnema, "Structure elucidation by NMR in organic chemistry," 1993.
[73] P. Groves, "Diffusion ordered spectroscopy (DOSY) as applied to polymers," *Polymer Chemistry*, vol. 8, no. 44, pp. 6700-6708, 2017, doi: 10.1039/c7py01577a.

[74] T. D. W. Claridge, "Chapter 9 - Diffusion NMR spectroscopy," in *Tetrahedron Organic Chemistry Series*, vol. 27, T. D. W. Claridge Ed., (Tetrahedron Organic Chemistry Series: Elsevier, 2009, pp. 303-334.

[75] R. Hoffman. "NMR Relaxation." Hebrew University of Jerusalem, Institute of chemistry. http://chem.ch.huji.ac.il/nmr/techniques/other/t1t2/t1t2.html (accessed 2020).

[76] B. Vögeli, *The Nuclear Overhauser Effect in NMR Structure and Dynamics Analysis*. Zürich, Switzerland: Swiss Federal Institute of Technology, 2015, p. 127.

[77] D. Kaseman. "Larmor Precession." LibreTexts. https://chem.libretexts.org/Courses/University_of_Illinois%2C_Springfield/Introduction_to_Organic_Spectroscopy/5%3A_Proton_Nuclear_Magnetic_Resonance_Spectroscopy_(NMR)/5.04%3A_Theory/Larmor_Precession (accessed 6th May, 2020).

[78] K. Zhang, L. Han, Y. Nie, M. L. Szigeti, and H. Ishida, "Examining the effect of hydroxyl groups on the thermal properties of polybenzoxazines: using molecular design and Monte Carlo simulation," *RSC Advances*, 10.1039/C8RA02033G vol. 8, no. 32, pp. 18038-18050, 2018, doi: 10.1039/C8RA02033G.

[79] N. de Mol and M. Fischer, *Surface Plasmon Resonance: Methods and Protocols* (Methods in Molecular Biology). 2010.

[80] K. Nagata and H. Handa, *Real-Time Analysis of Biomolecular Interactions: Applications of BIACORE*. 2000.

[81] K. Tawa, "Mismatching base-pair dependence of the kinetics of DNA-DNA hybridization studied by surface plasmon fluorescence spectroscopy," *Nucleic Acids Res*, vol. 32, no. 8, pp. 2372-2377, 2004, doi: 10.1093/nar/gkh572.

[82] E. R. Lacy, "Recognition of Tmiddle dotG mismatched base pairs in DNA by stacked imidazole-containing polyamides: surface plasmon resonance and circular dichroism studies," *Nucleic Acids Res*, vol. 30, no. 8, pp. 1834-1841, Apr 15 2002, doi: 10.1093/nar/30.8.1834.

[83] J. Masárová, F. Winquist, and B. Danielsson, "Chapter 5 - Bioanalytical Studies Based on Lectin–Carbohydrate Interactions Measured by Ellipsometry and Surface Plasmon Resonance Techniques," in *Lectins*, C. L. Nilsson Ed. Amsterdam: Elsevier Science B.V., 2007, pp. 103-128.

[84] I.-H. Lee *et al.*, "Desulfurization–bromination: direct chain-end modification of RAFT polymers," *Polymer Chemistry*, vol. 8, no. 46, pp. 7188-7194, Dec 14 2017, doi: 10.1039/c7py01702b.

[85] S. S. Hinman, K. S. McKeating, and Q. Cheng, "Surface Plasmon Resonance: Material and Interface Design for Universal Accessibility," *Analytical Chemistry*, vol. 90, no. 1, pp. 19-39, 2018/01/02 2018, doi: 10.1021/acs.analchem.7b04251.

[86] R. Debono, G. Loucks, D. Manna, and U. Krull, "Self-assembly of short and long-chain n-alkyl thiols onto gold surfaces: A real-time study using surface plasmon resonance techniques," *Canadian Journal of Chemistry*, vol. 74, no. 5, pp. 677-688, 02/05 2011, doi: 10.1139/v96-073.

[87] R. Janmanee, S. Chuekachang, S. Sriwichai, A. Baba, and S. Phanichphant, "Functional Conducting Polymers in the Application of SPR Biosensors," *Journal of Nanotechnology*, vol. 2012, pp. 1-7, 2012, doi: 10.1155/2012/620309.

[88] T. F. Scientific, "Instructions BMPH, EMCH, KMUH," Rockford, 2012. [Online]. Available: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0011363_BMPH_EMCH_KMUH_UG.pdf

[89] J. Kestell, R. Abuflaha, M. Garvey, and W. T. Tysoe, "Self-Assembled Oligomeric Structures from 1,4-Benzenedithiol on Au(111) and the Formation of Conductive Linkers between Gold Nanoparticles," *The Journal of Physical Chemistry C*, vol. 119, no. 40, pp. 23042-23051, 2015/10/08 2015, doi: 10.1021/acs.jpcc.5b07343.

[90] C. Vericat, M. Vela, G. Benitez, P. Carro, and R. Salvarezza, "Self-Assembled Monolayers of Thiols and Dithiols on Gold: New Challenges for a Well-Known System," *Chemical Society reviews*, vol. 39, no. 5, pp. 1805-34, 05/01 2010, doi: 10.1039/b907301a.

[91] S. Sharifi, S. Behzadi, S. Laurent, M. Laird Forrest, P. Stroeve, and M. Mahmoudi, "Toxicity of nanomaterials," *Chem. Soc. Rev.*, vol. 41, no. 6, pp. 2323-2343, Mar 21 2012, doi: 10.1039/c1cs15188f.

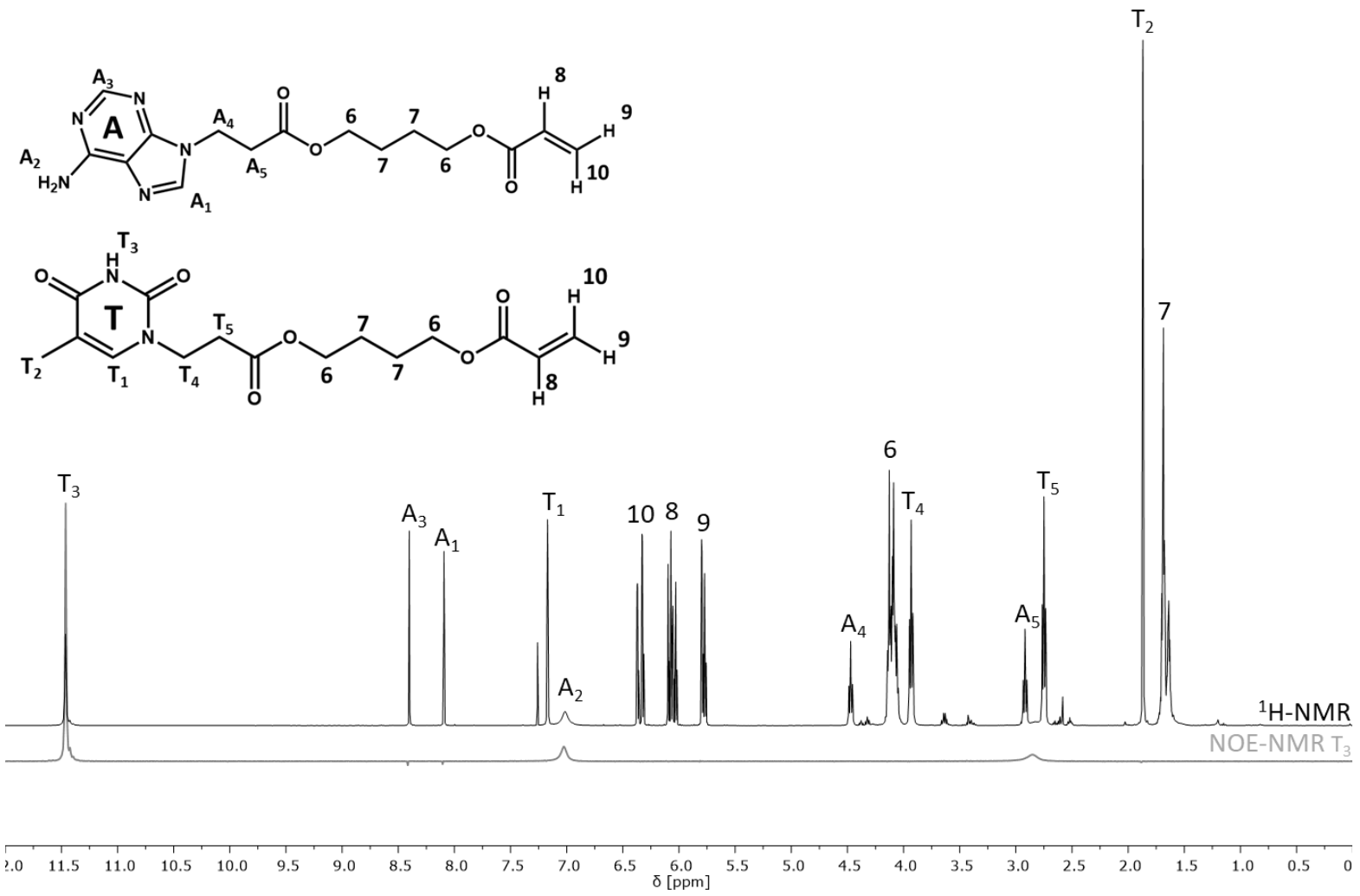
[92] L. Chen *et al.*, "The toxicity of silica nanoparticles to the immune system," *Nanomedicine*, vol. 13, no. 15, pp. 1939-1962, 2018/08/01 2018, doi: 10.2217/nnm-2018-0076.
[93] M. Hesler *et al.*, "Multi-endpoint toxicological assessment of polystyrene nano- and microparticles in different biological models in vitro," *Toxicology in Vitro*, vol. 61, p. 104610, Dec 2019, doi: 10.1016/j.tiv.2019.104610.
[94] L. Liang and D. Astruc, "The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) "click" reaction and its applications. An overview," *Coordination Chemistry Reviews*, vol. 255, no. 23, pp. 2933-2945, 2011/12/01/ 2011, doi: https://doi.org/10.1016/j.ccr.2011.06.028.
[95] G. Hermanson, "Bioconjugate Techniques: Third Edition," *Bioconjugate Techniques: Third Edition*, pp. 1-1146, 08/01 2013, doi: 10.1016/B978-0-12-382239-0.00005-4.
[96] T. M. V. D. Pinho e Melo, "Synthesis of Azides," *Organic Azides*, pp. 53-94, 2009/11/20 2009, doi: 10.1002/9780470682517.ch3.
[97] J. Vandenbergh, T. Tura, E. Baeten, and T. Junkers, "Polymer end group modifications and polymer conjugations via "click" chemistry employing microreactor technology," *Journal of Polymer Science Part A: Polymer Chemistry*, vol. 52, no. 9, pp. 1263-1274, 2014/05/01 2014, doi: 10.1002/pola.27112.
[98] C. National Research, *Copper in Drinking Water*. Washington, DC: The National Academies Press (in English), 2000.
[99] J. V. Burda and F. Šebesta, "Metal Interactions with Nucleobases, Base Pairs, and Oligomer Sequences; Computational Approach," in *Handbook of Computational Chemistry*, J. Leszczynski Ed. Dordrecht: Springer Netherlands, 2016, ch. Chapter 36-2, pp. 1-48.
[100] J. Xu, J. He, D. Fan, X. Wang, and Y. Yang, "Aminolysis of Polymers with Thiocarbonylthio Termini Prepared by RAFT Polymerization: The Difference between Polystyrene and Polymethacrylates," *Macromolecules*, vol. 39, no. 25, pp. 8616-8624, 2006/12/01 2006, doi: 10.1021/ma061961m.
[101] *Azo Polymerization Initiators : Comprehensive Catalog*, Fujifilm, Neuss, Germany. [Online]. Available: https://www.wako-chemicals.de/files/download/pdf/wako_azo_polymerization_initiators_catalog_25.pdf.
[102] M. Klussmann, B. Schweitzer-Chaput, J. Vandenbergh, and T. Junkers, "Low Temperature Radical Initiator System and Processes Making Use Thereof," Belguim Patent Appl. 15202134.1, 2015.
[103] A. Lu, T. P. Smart, T. H. Epps, D. A. Longbottom, and R. K. O'Reilly, "l -Proline Functionalized Polymers Prepared by RAFT Polymerization and Their Assemblies as Supported Organocatalysts," *Macromolecules*, vol. 44, no. 18, pp. 7233-7241, Sep 27 2011, doi: 10.1021/ma201256m.
[104] C. Y. Huang, "[27] Determination of binding stoichiometry by the continuous variation method: The job plot," in *Methods in Enzymology*, vol. 87, D. L. Purich Ed., (Methods in Enzymology: Academic Press, 1982, pp. 509-525.
[105] S. Upadhye *et al.*, "Preparation and Characterization of Inclusion Complexes of a Hemisuccinate Ester Prodrug of Δ9-Tetrahydrocannabinol with Modified Beta-Cyclodextrins," *AAPS PharmSciTech*, vol. 11, no. 2, pp. 509-17, 03/01 2010, doi: 10.1208/s12249-010-9401-4.
[106] F. Ulatowski, K. Dąbrowa, T. Bałakier, and J. Jurczak, "Recognizing the Limited Applicability of Job Plots in Studying Host–Guest Interactions in Supramolecular Chemistry," *The Journal of Organic Chemistry*, vol. 81, no. 5, pp. 1746-1756, 2016/03/04 2016, doi: 10.1021/acs.joc.5b02909.
[107] J. S. Renny, L. L. Tomasevich, E. H. Tallmadge, and D. B. Collum, "Method of Continuous Variations: Applications of Job Plots to the Study of Molecular Associations in Organometallic Chemistry," *Angewandte Chemie International Edition*, vol. 52, no. 46, pp. 11998-12013, 2013/11/11 2013, doi: 10.1002/anie.201304157.
[108] P. Thordarson, "Determining association constants from titration experiments in supramolecular chemistry," *Chemical Society Reviews*, 10.1039/C0CS00062K vol. 40, no. 3, pp. 1305-1323, Mar 2011, doi: 10.1039/C0CS00062K.

Appendices

A1. Matlab code for determination of polynomal fit trough Job plot data

```
%MA represents the matrix containing the mole fractions of the experimental data
%yT represents the interaction parameter for the Job plot, calculated with the experimental data
%P represents the order of the polynomial fit
%The code below is used to determine the equation describing the polynomial most fitting for the Job plot dataset
[PT_25,ST_25] = polyfit(MA,yT_25,P)
%The formula below calculates the R² of the polynomial fit
1 - (ST_25.normr/norm(yT_25 - mean(yT_25)))^2;
%The code below calculates all maximum values of the polynomial
rT_25=roots(polyder(PT_25));
%The code below extract all real maximum values from the maximum value dataset
rT_25=rT_25(imag(rT_25)==0);
```

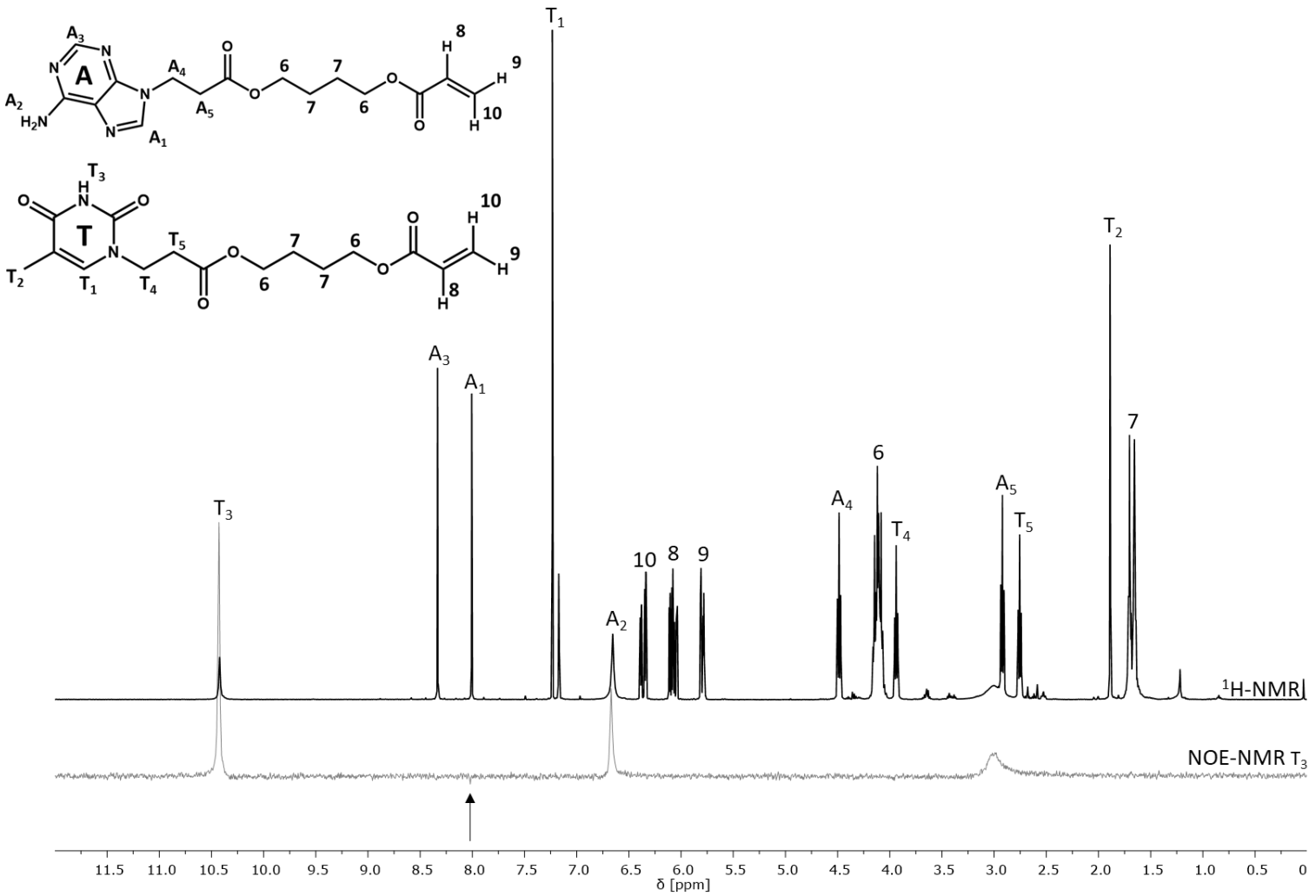
A2. ^{1}H-NMR analysis of sample AT4 and NOE-analysis of H^3 of TAM



Measuring specifications	
Sample	AT4
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.2 °C
T(NOE-NMR T_3)	20.4 °C
γ_A	0.2798
γ_T	0.7202
δ(NOE-NMR T_3)	11.4515 8 ppm

Impurities	
Compound	δ [ppm]
TEA	1.2002
TEA	2.5201
DMSO	2.5826
H_20	2.8610
MeOH	3.4257
$CDCl_3$	7.2600

A3. ^{1}H-NMR analysis of sample AT5 and NOE-analysis of H^3 of TAM

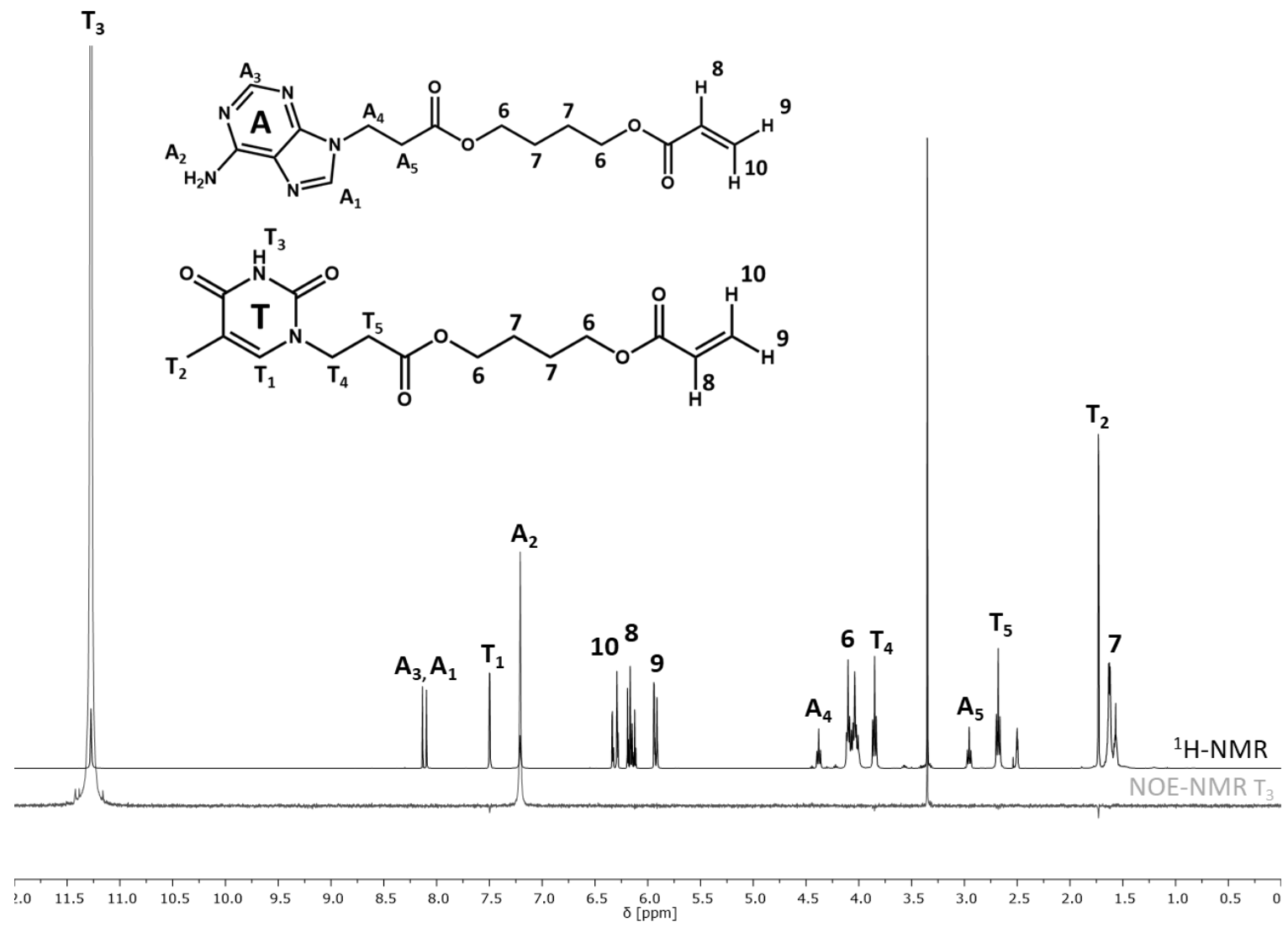


Measuring specifications	
Sample	AT5
Solvent	$CDCl_3$
T(^{1}H-NMR)	23.0 °C
T(NOE-NMR T_3)	23.3 °C
γ_A	0.5187
γ_T	0.4813
δ(NOE-NMR T_3)	11.4280 ppm

Impurities	
Compound	δ [ppm]
Si-grease	0.0627
n-hexane	0.8740
n-hexane	1.2454
DMSO	2.5509
DMSO	2.6134
H_2O	3.0341
$CHCl_3$	72603

Remark: the arrow indicates the weak signal in the NOE-spectrum, representing H^3 from AAM.

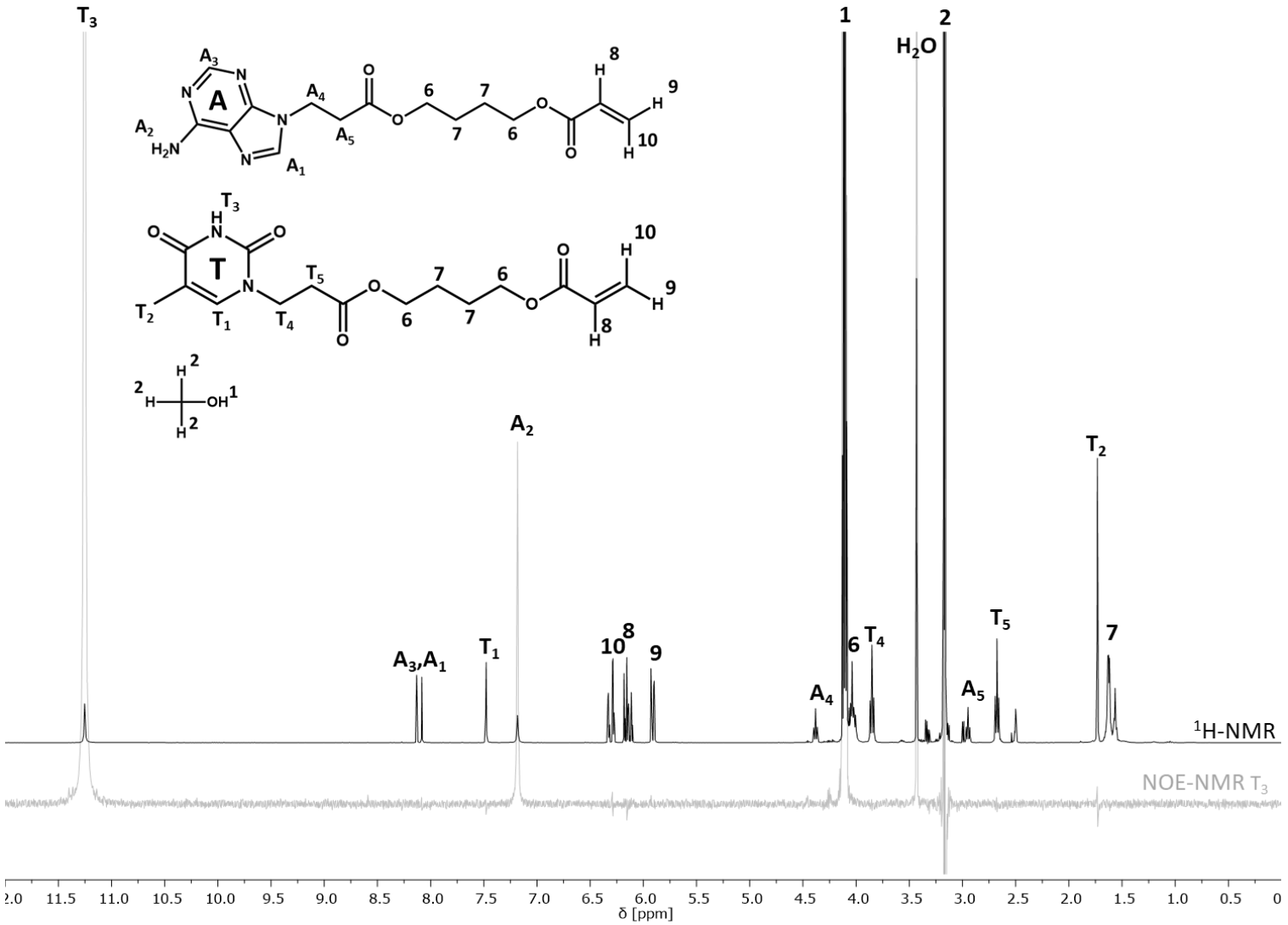
A4. ^{1}H-NMR analysis of sample AT1 and NOE-analysis of H^{3} of TAM



Measuring specifications	
Sample	AT1
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	22.2 °C
T(NOE-NMR T_3)	22.2 °C
γ_A	0.2539
γ_T	0.7461
δ(NOE-NMR T_3)	11.23465 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8312
TEA	1.0788
n-hexane	1.2045
DMSO	2.5000
TEA	2.5394
MeOH	4.2218
H_20	3.3505

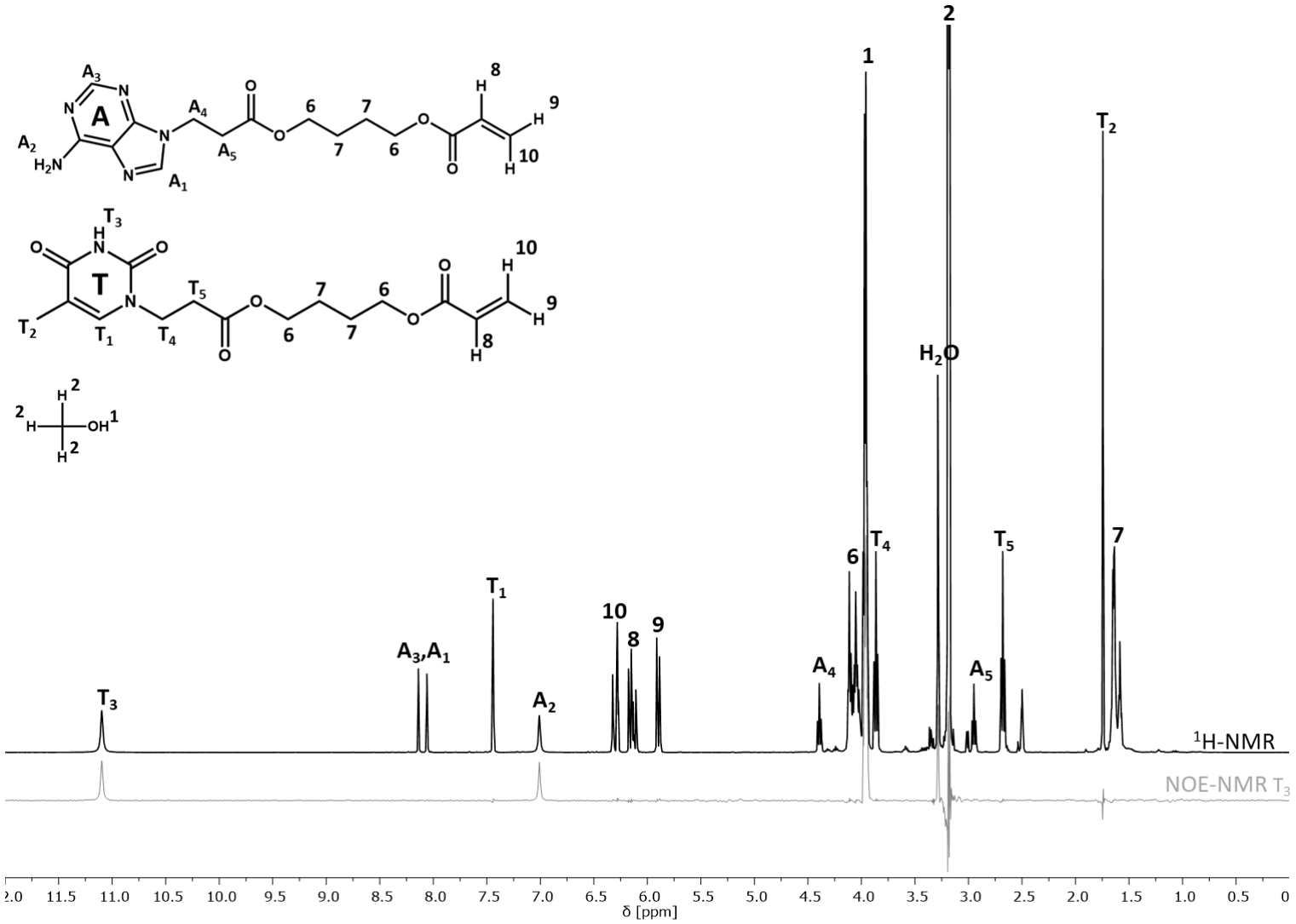
A5. ^{1}H-NMR analysis of sample AT2 and NOE-analysis of H^{3} of TAM



Measuring specifications	
Sample	AT2
Solvent	$(CD_3)_2SO$ + MeOH
T(^{1}H-NMR)	22.2 °C
T(NOE-NMR T_3)	22.3 °C
γ_A	0.2506
γ_T	0.7494
δ(NOE-NMR T_3)	11.06203 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8294
TEA	1.0487
n-hexane	1.2028
DMSO	2.5000
TEA	2.5393
MeOH	3.1764
H_20	3.4311
MeOH	4..1159

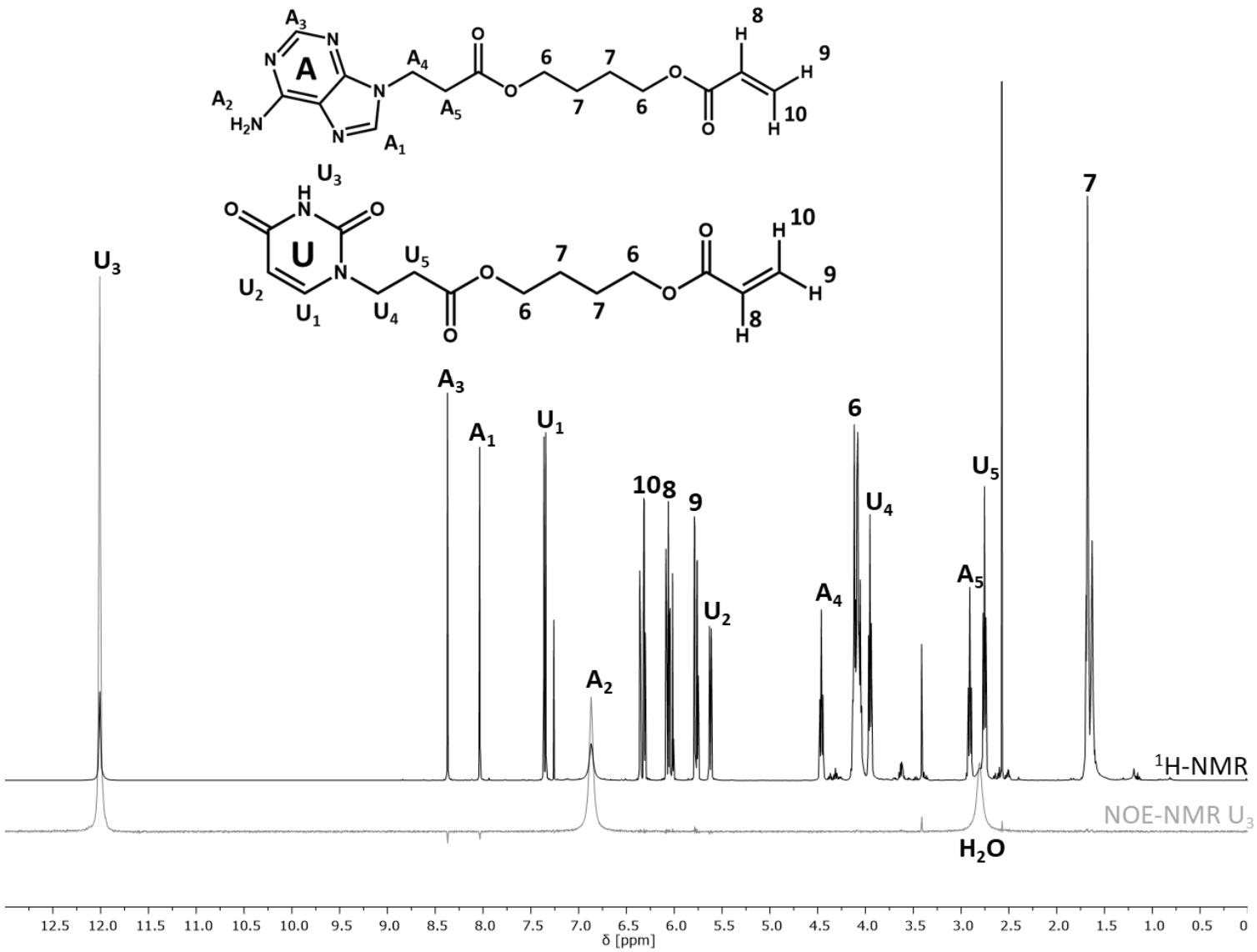
A6. ^{1}H-NMR analysis of sample AT3 and NOE-analysis of H^{3} of TAM



Measuring specifications	
Sample	AT3
Solvent	$(CD_3)_2SO$ + MeOH
T(^{1}H-NMR)	50.0 °C
T(NOE-NMR T_3)	50.0 °C
γ_A	0.2545
γ_T	0.7455
δ(NOE-NMR T_3)	11.06203 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8338
TEA	1.0629
n-hexane	1.2243
DMSO	2.5000
DMSO	2.5380
MeOH	3.1798
H_20	3.2863
MeOH	3.9585

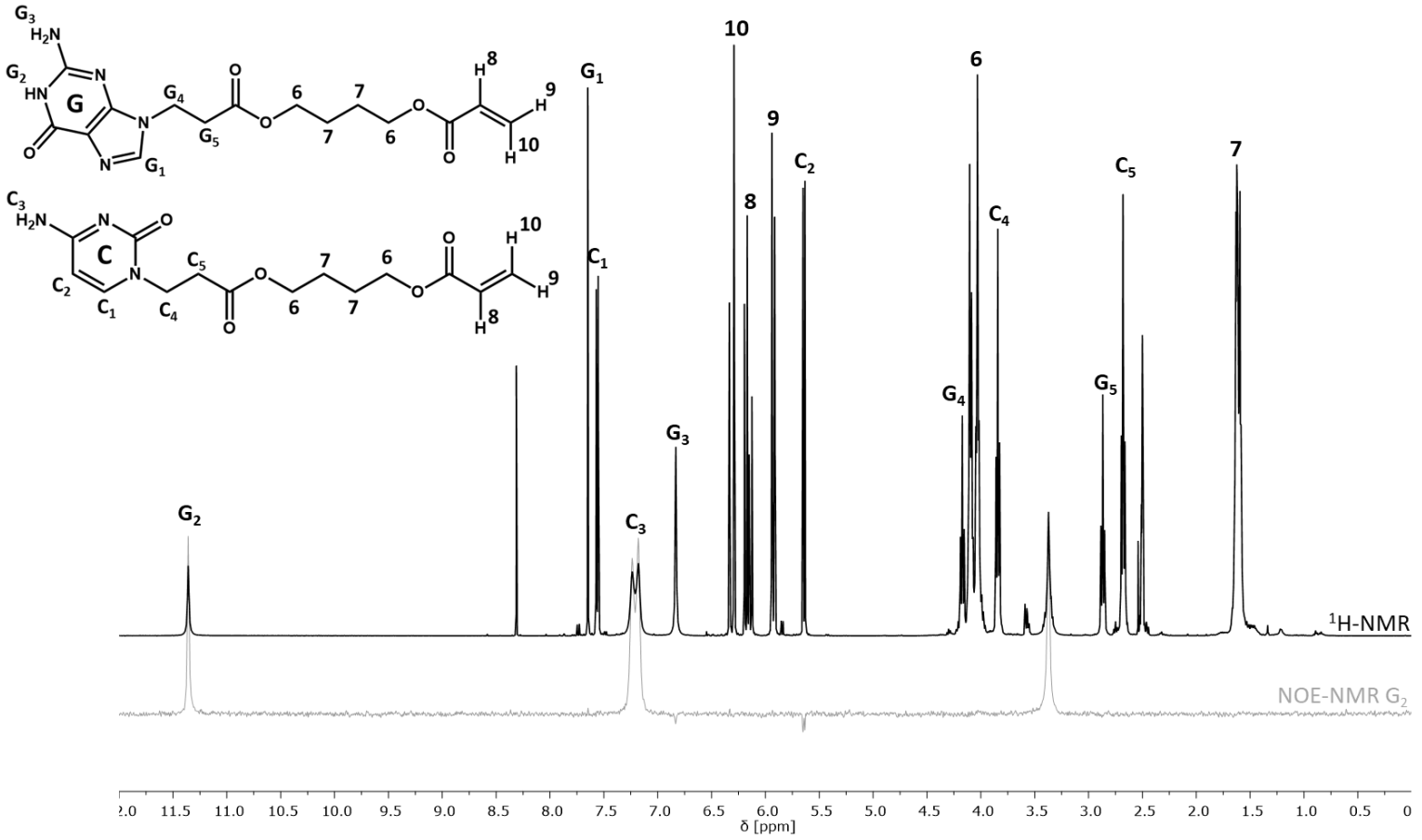
A7. ^{1}H-NMR analysis of sample AU1 and NOE-analysis of H^{3} of UAM



Measuring specifications	
Sample	AU1
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.8 °C
T(NOE-NMR U_3)	20.6 °C
γ_A	0.3525
γ_U	0.6475
δ(NOE-NMR U_3)	11.99660 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8146
n-hexane	1.1916
DMSO	2.5734
H_2O	2.8052
MeOH	3.4128
DCM	3.6211
$CDCl_3$	7.2600

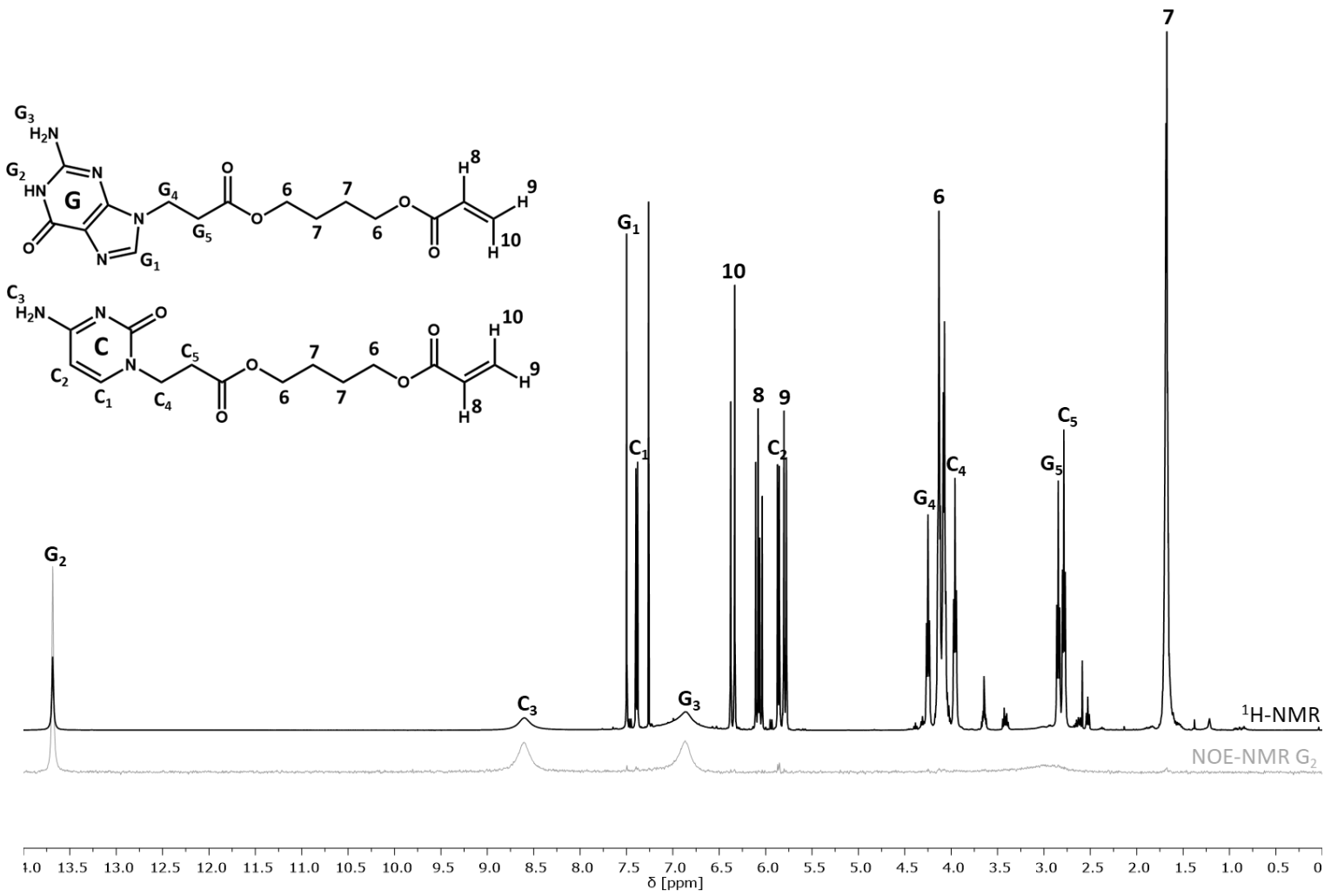
A8. ^{1}H-NMR analysis of sample GC1 and NOE-analysis of H^{2} of GAM



Measuring specifications	
Sample	GC1
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	20.2 °C
T(NOE-NMR G_2)	20.2 °C
γ_G	0.3558
γ_C	0.6442
δ(NOE-NMR G_2)	11.32000 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8920
n-hexane	1.2215
BHT	1.3373
BHT	2.3126
H_20	3.3722
MeOH	4.2996
DCM	5.8348
BHT	6.5479
$CHCl_3$	8.3122

A9. ^{1}H-NMR analysis of sample GC2 and NOE-analysis of H^{2} of GAM



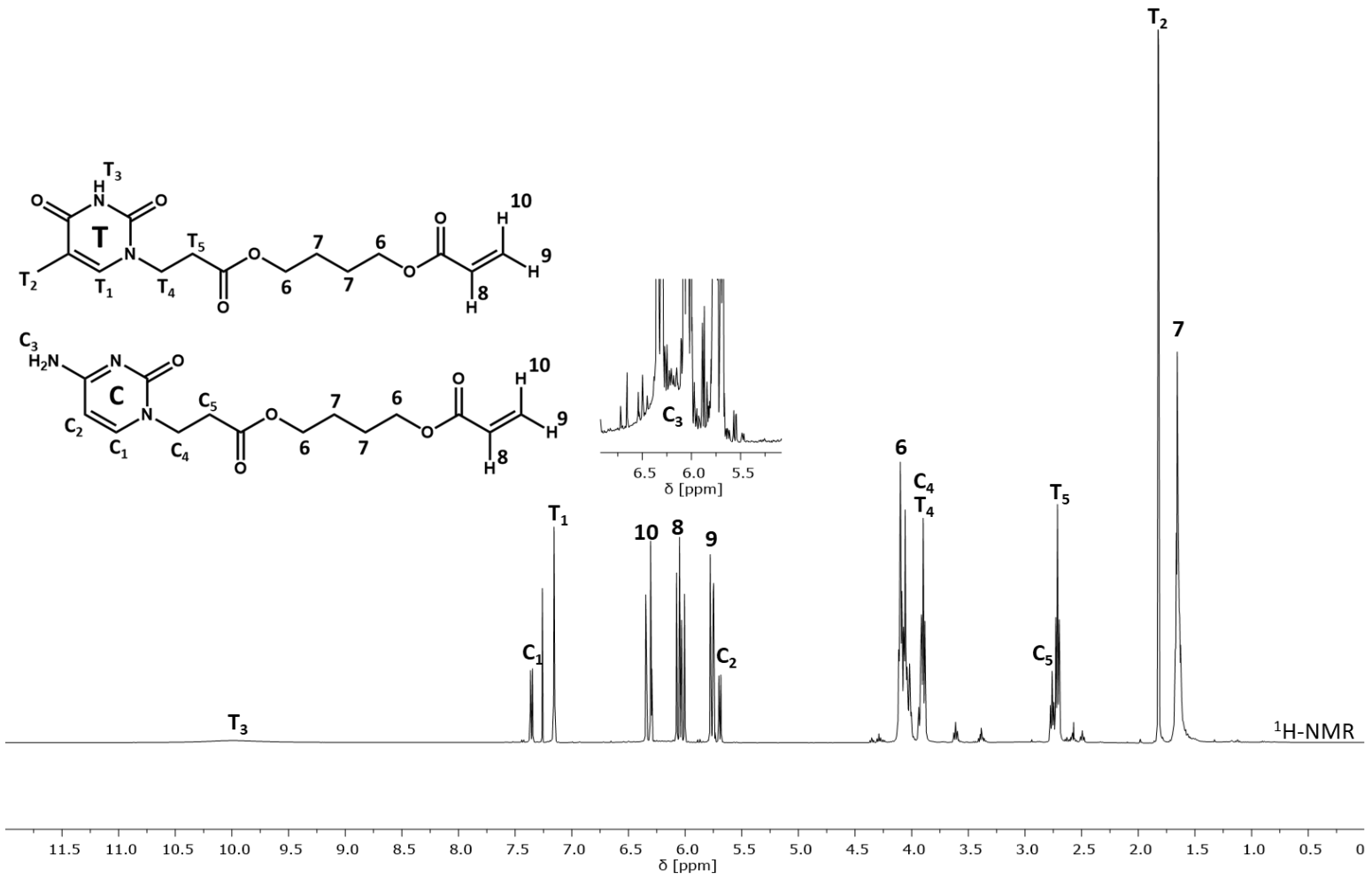
Measuring specifications	
Sample	GC2
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.2 °C
T(NOE-NMR G_2)	20.2 °C
γ_G	0.3558
γ_C	0.6442
δ(NOE-NMR G_2)	11.32000 ppm

Impurities	
Compound	δ [ppm]
Si-grease	0.0378
MeOH	1.2161
BHT	1.3779
DMSO	2.5865
H_20	2.8910
DCM	3.4280
MeOH	3.6468
$CHCl_3$	7.2600

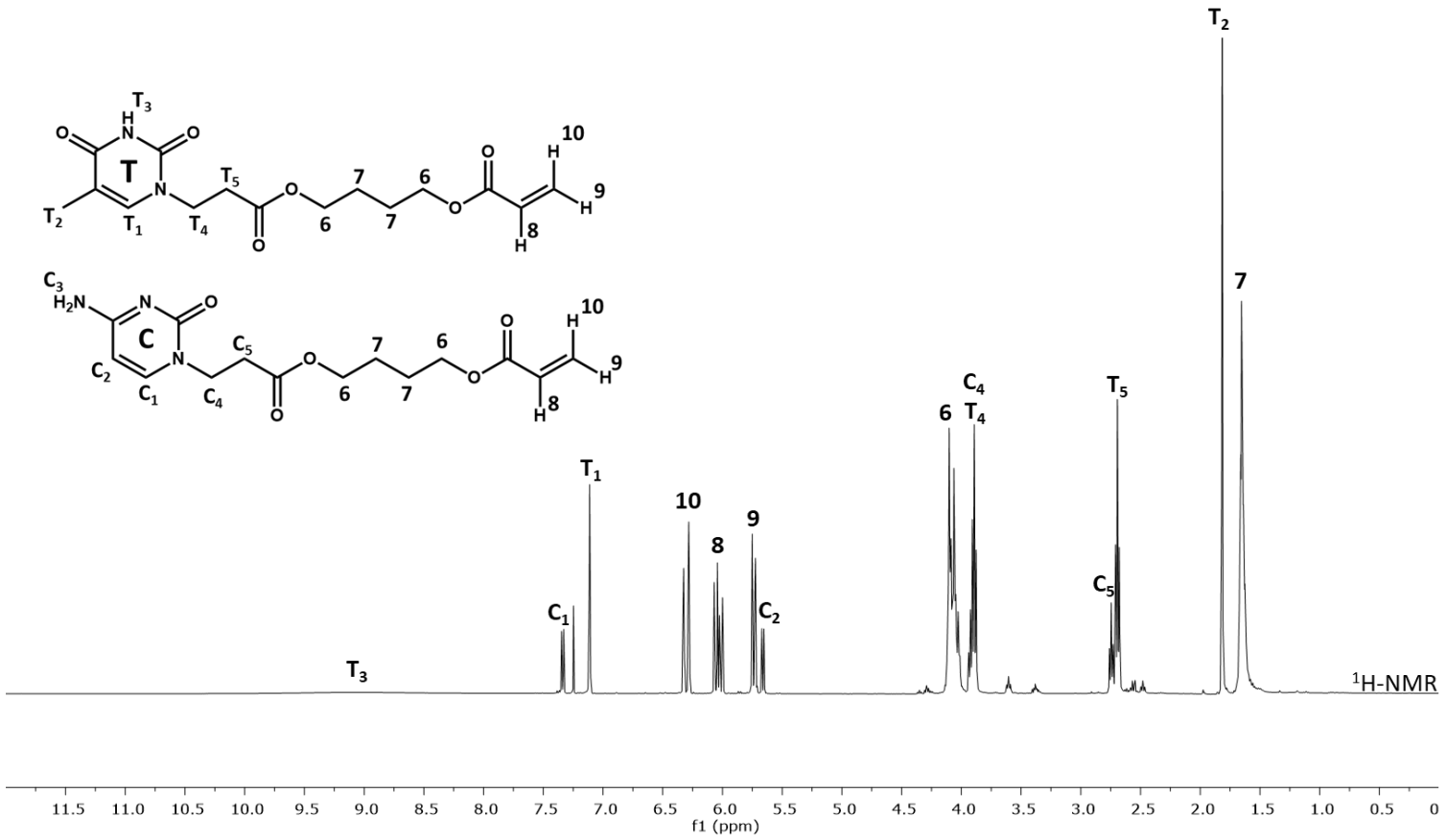
A10. ^{1}H-NMR analysis of sample CT1

Measuring specifications	
Sample	CT1
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.1 °C
γ_C	0.2673
γ_T	0.7327

Impurities	
Compound	δ [ppm]
MeOH	1.1748
Tert-butyl alcohol	1.3278
DMSO	2.5708
DCM	3.3868
MeOH	3.6122
$CHCl_3$	7.2600



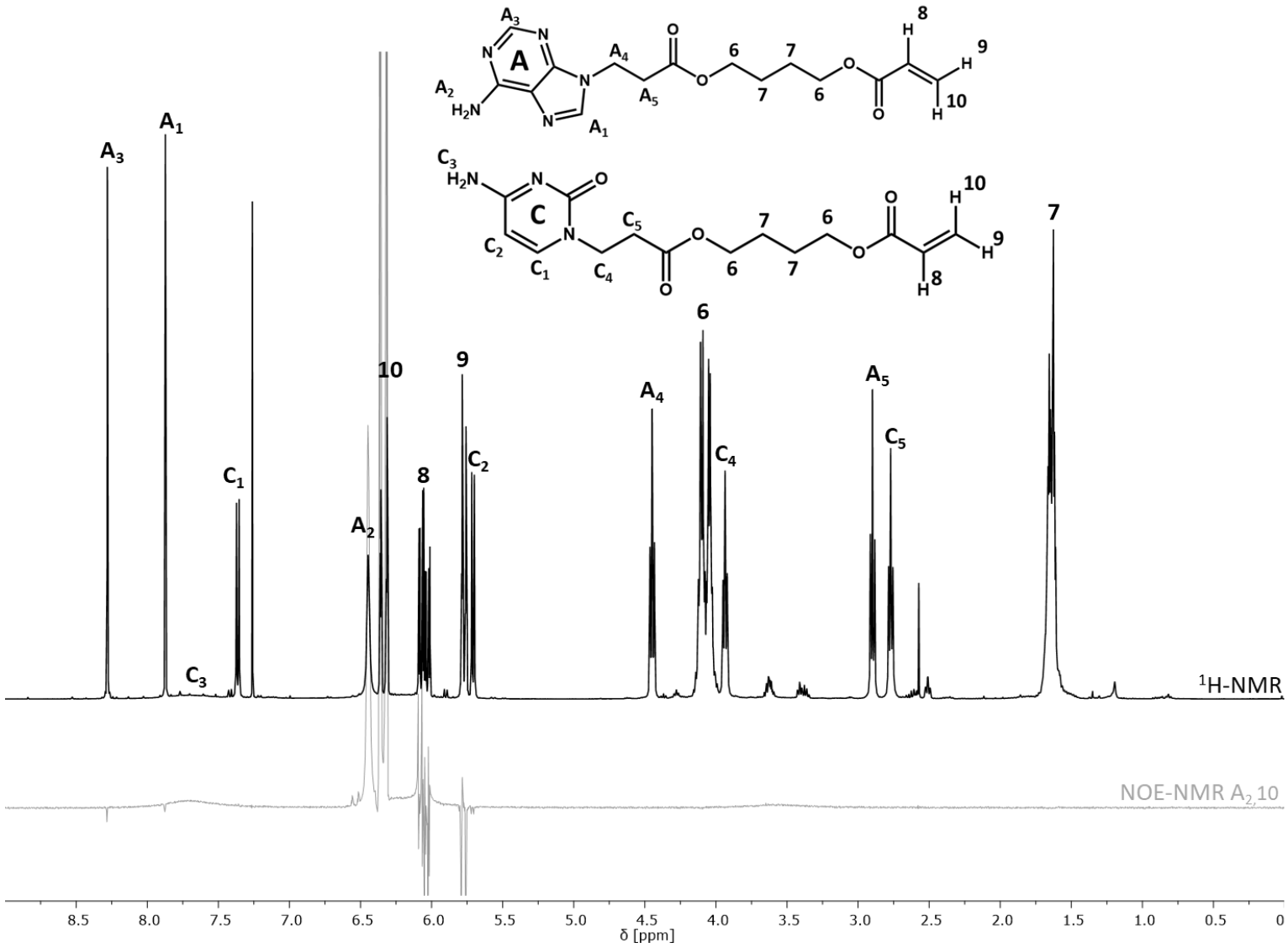
A11. ^{1}H-NMR analysis of sample CT2



Measuring specifications	
Sample	CT2
Solvent	$CDCl_3$
T(^{1}H-NMR)	50.0 °C
γ_C	0.2673
γ_T	0.7327

Impurities	
Compound	δ [ppm]
MeOH	1.1873
Tert-butyl alcohol	1.3343
DMSO	2.4806
DCM	3.3812
MeOH	3.6041
$CHCl_3$	7.2600

A12. ^{1}H-NMR analysis of sample AC1 and NOE-analysis of H^{2} of AAM and H^{10}

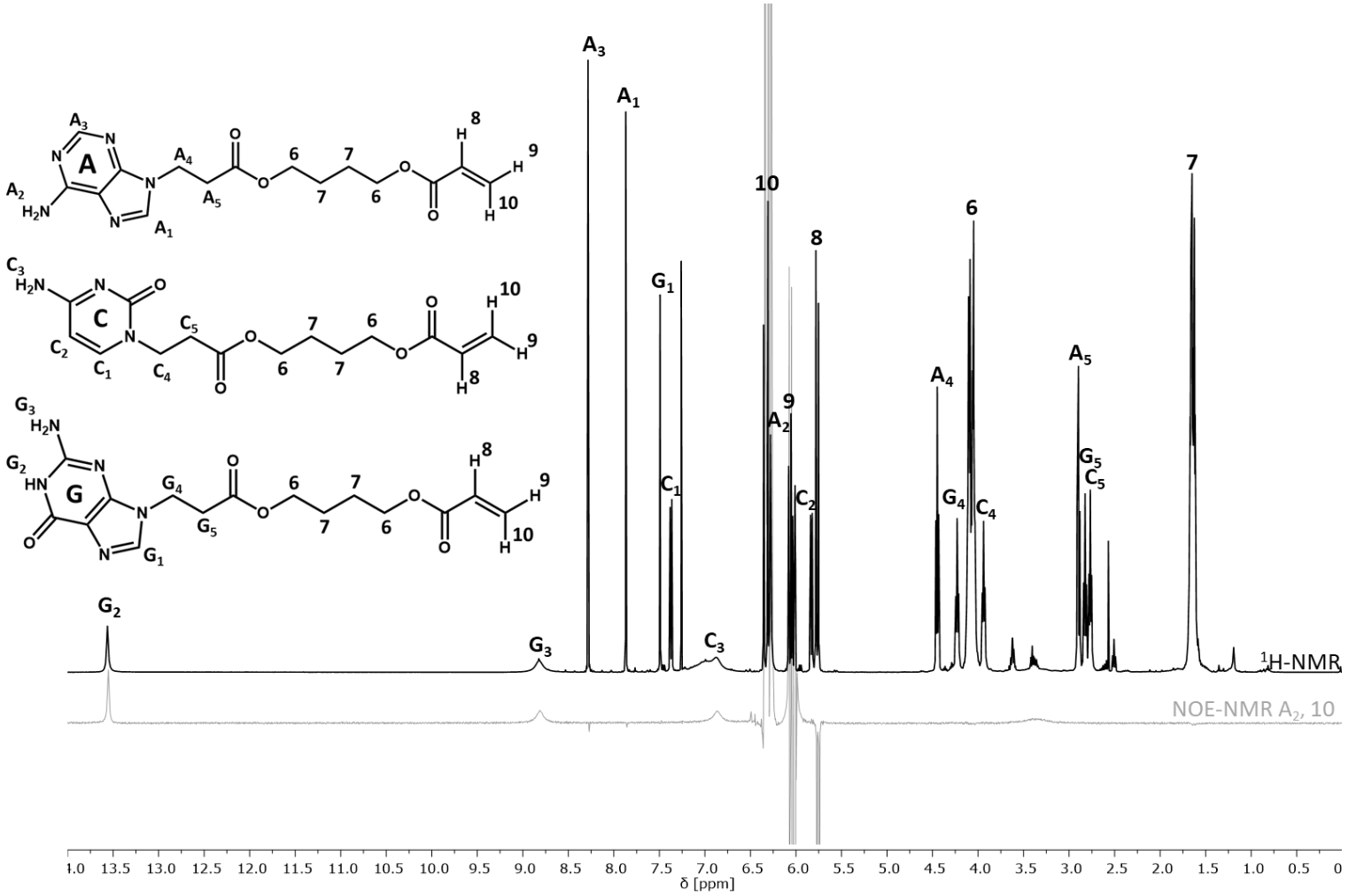


Measuring specifications	
Sample	AC1
Solvent	$CDCl_3$
T(^{1}H-NMR)	21.0 °C
T(NOE-NMR A_2, 10)	20.9 °C
γ_A	0.3558
γ_C	0.6442
δ(NOE-NMR A_2, 10)	6.43460 ppm

Impurities	
Compound	δ [ppm]
MeOH	1.1945
DMSO	2.5086
DMSO	2.5728
TEA	3.4118
H_2O	3.5936
MeOH	3.6309
$CHCl_3$	7.2600

Remark: Due to the irradiation of a frequency-range instead of one specific frequency and the small difference between the respective δ, also H^{10} was irradiated instead of merely A_2.

A13. ^{1}H-NMR analysis of sample AGC1 and NOE-analysis of H^{2} of AAM and H^{10}

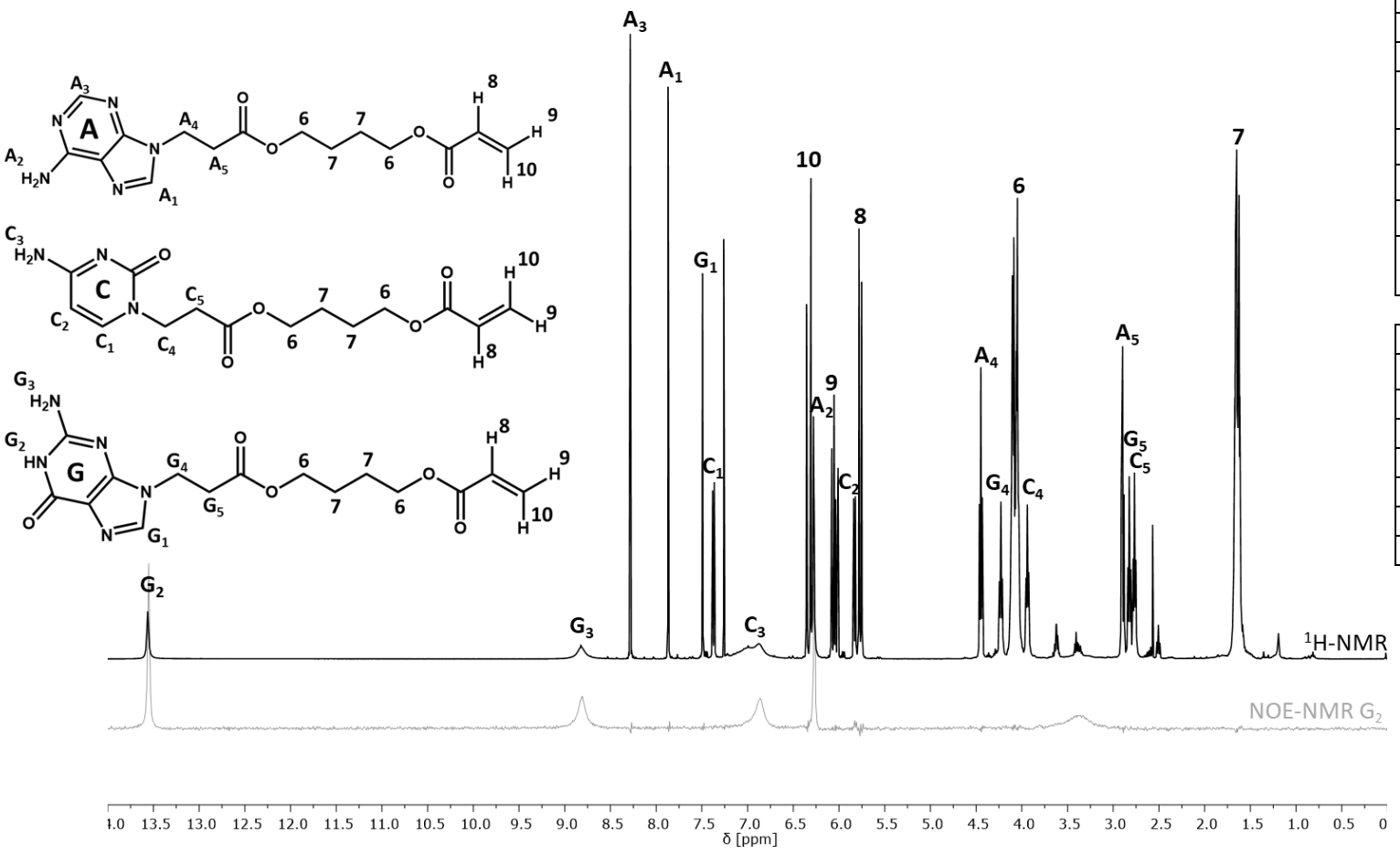


Measuring specifications	
Sample	ACG1
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.2 °C
T(NOE-NMR A_2,10)	20.1 °C
γ_A	0.3724
γ_C	0.3291
γ_G	0.2985
δ(NOE-NMR A_2, 10)	6.2676 ppm

Impurities	
Compound	δ [ppm]
Si-grease	0.0194
n-hexane	0.8165
n-hexane	1.3554
DMSO	2.5076
H_2O	3.4054
$CHCl_3$	7.2600

Remark: Due to the irradiation of a frequency-range instead of one specific frequency and small difference between the respective δ, also H^{10} was irradiated instead of merely A_2.

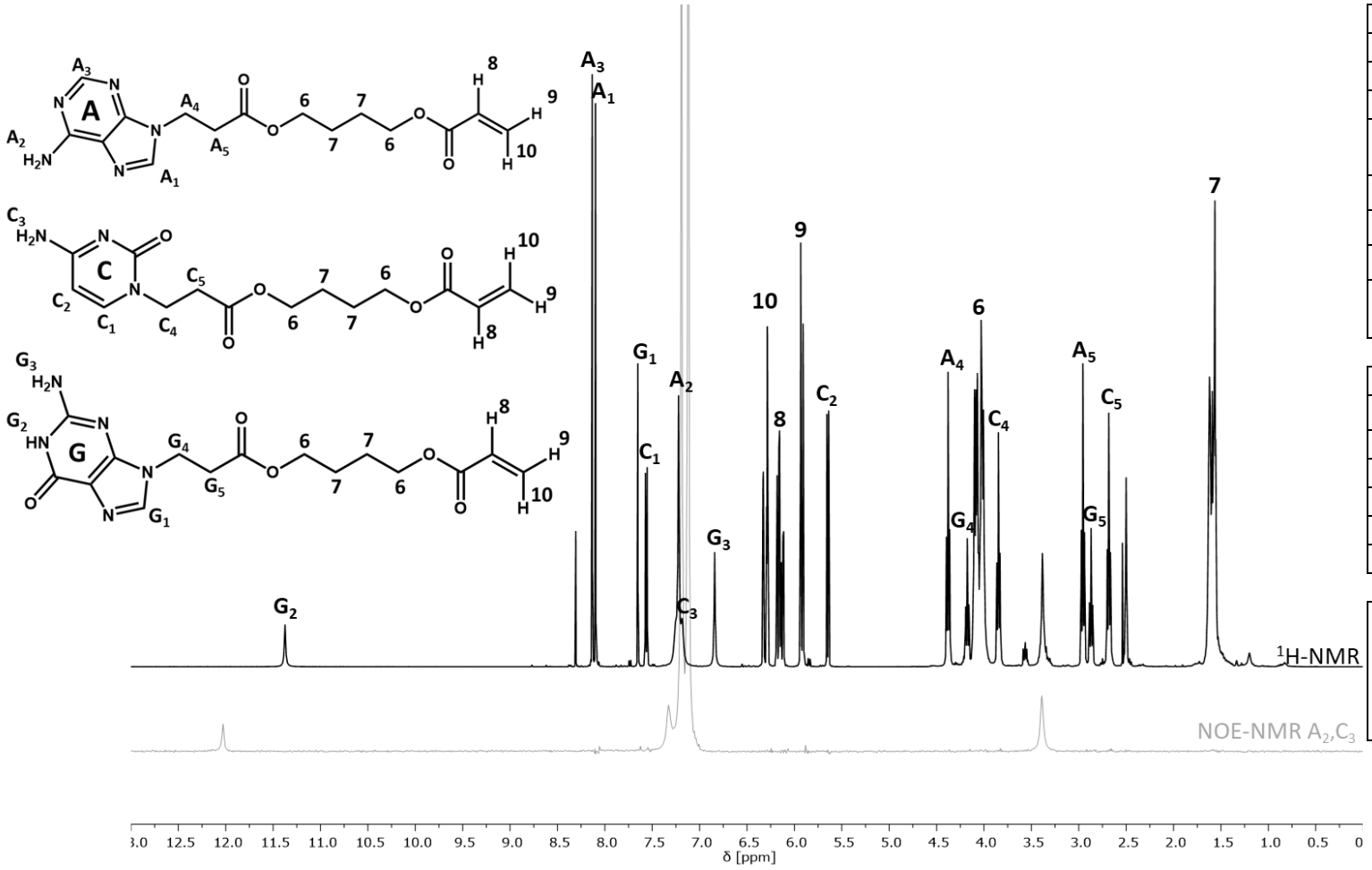
A14. ^{1}H-NMR analysis of sample AGC1 and NOE-analysis of H^{2} of GAM



Measuring specifications	
Sample	ACG1
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.2 °C
T(NOE-NMR G_2)	20.2 °C
γ_A	0.3724
γ_C	0.3291
γ_G	0.2985
δ(NOE-NMR G_2)	13.5504 ppm

Impurities	
Compound	δ [ppm]
Si-grease	0.0194
n-hexane	0.8165
n-hexane	1.3554
DMSO	2.5076
H_2O	3.4054
$CHCl_3$	7.2600

A15. ^{1}H-NMR analysis of sample AGC2 and NOE-analysis of H^2 of AAM and H^3 of CAM

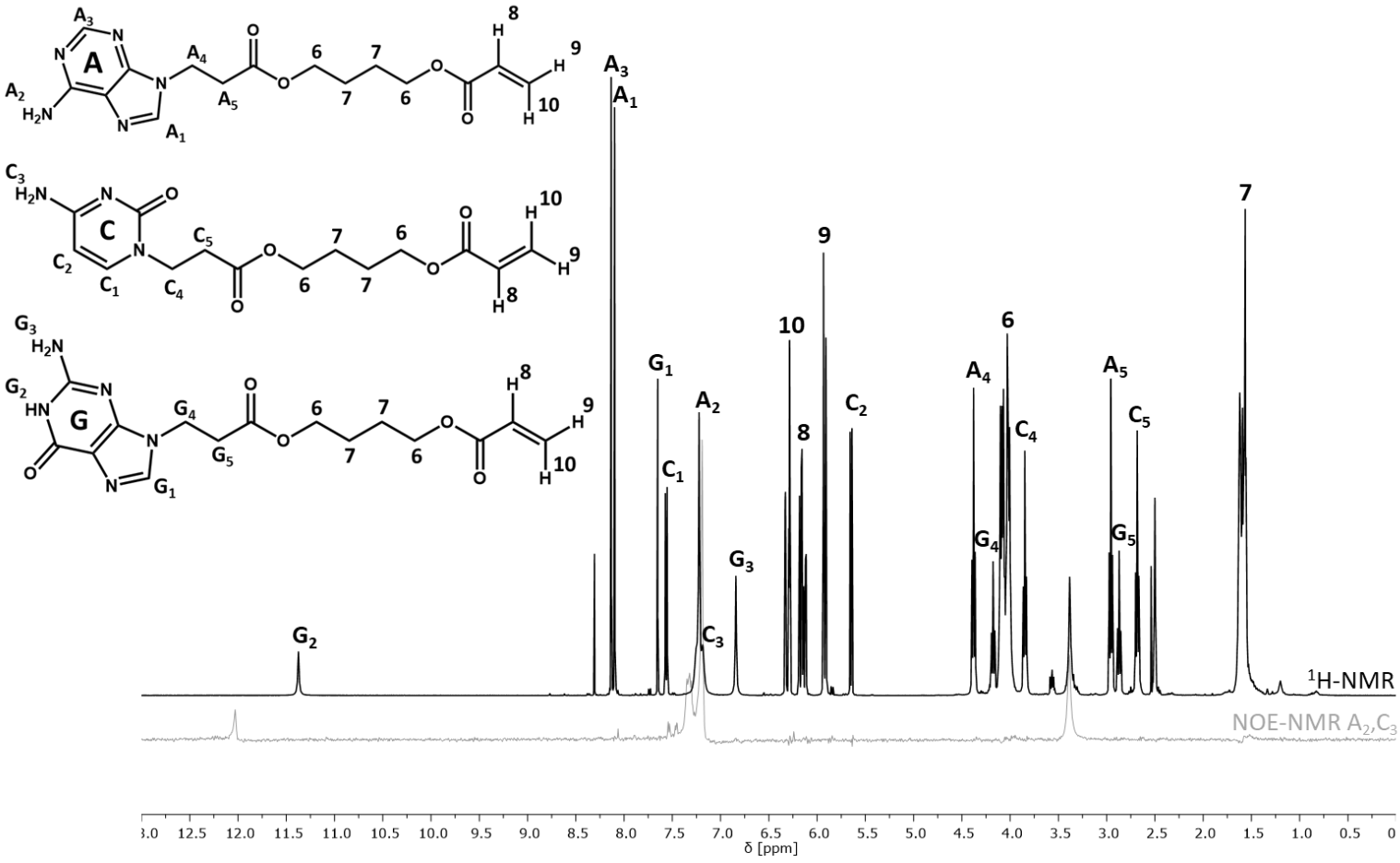


Measuring specifications	
Sample	ACG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	20.3 °C
T(NOE-NMR A_2,C_3)	22.9 °C
γ_A	0.4048
γ_C	0.3796
γ_G	0.2120
δ(NOE-NMR A_2,C_3)	7.1152 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8275
n-hexane	1.2001
DMSO	2.5002
H_2O	3.3840
$CHCl_3$	8.3067

Remark: Due to overlapping signals and irradiation of a frequency band, A_2 and C_3 cannot be irradiated separately.

A16. ^{1}H-NMR analysis of sample AGC2 and NOE-analysis of H^2 of AAM and H^3 of CAM

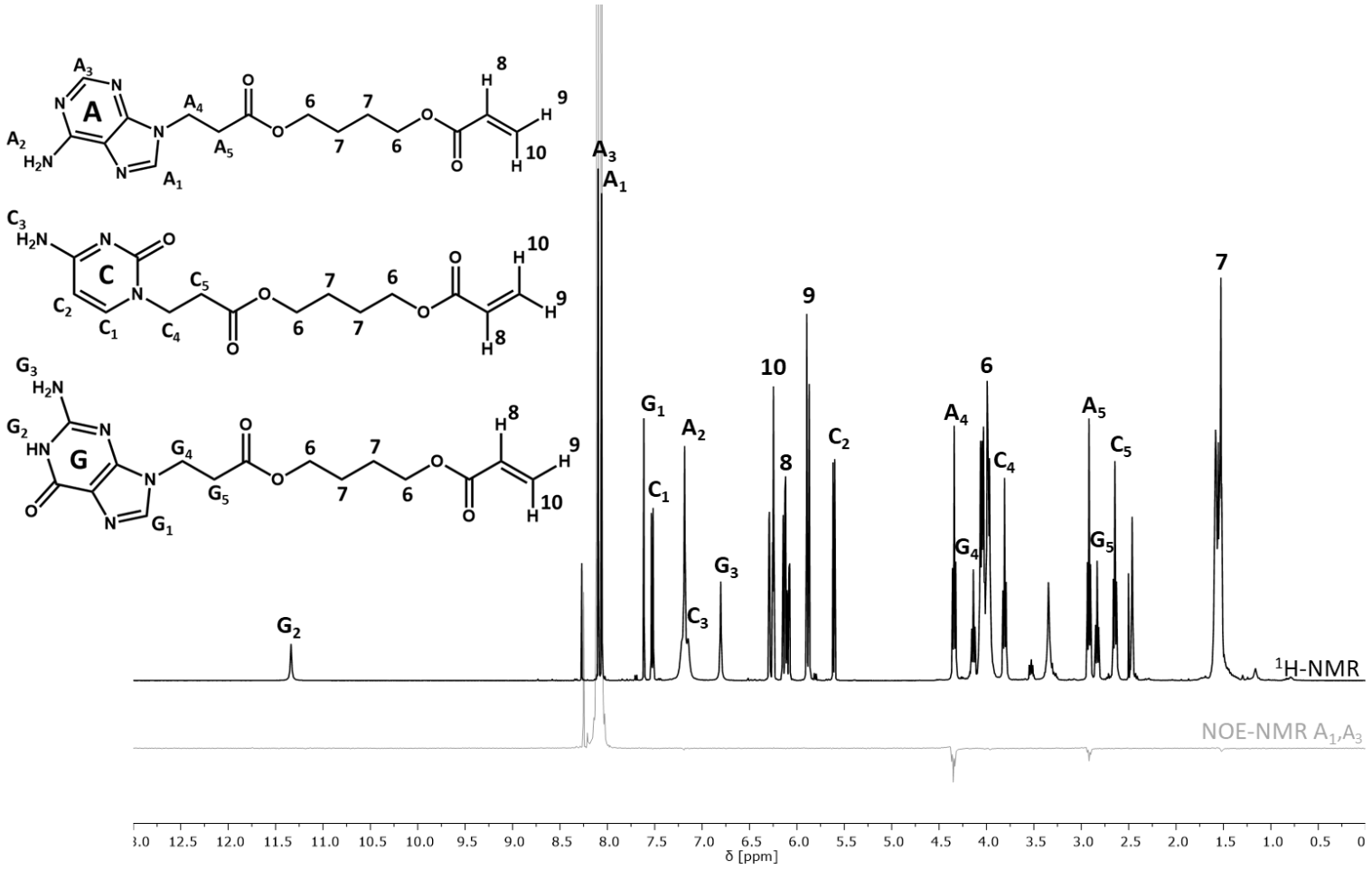


Measuring specifications	
Sample	ACG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	20.3 °C
T(NOE-NMR A_2, C_3)	22.9 °C
γ_A	0.4048
γ_C	0.3796
γ_G	0.2120
δ(NOE-NMR A_2, C_3)	7.3237 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8275
n-hexane	1.2001
DMSO	2.5002
H_2O	3.3840
$CHCl_3$	8.3067

Remark: Due to overlapping signals and irradiation of a frequency band, A_2 and C_3 cannot be irradiated separately.

A17. ^{1}H-NMR analysis of sample AGC2 and NOE-analysis of H^{1} and H^{3} of AAM

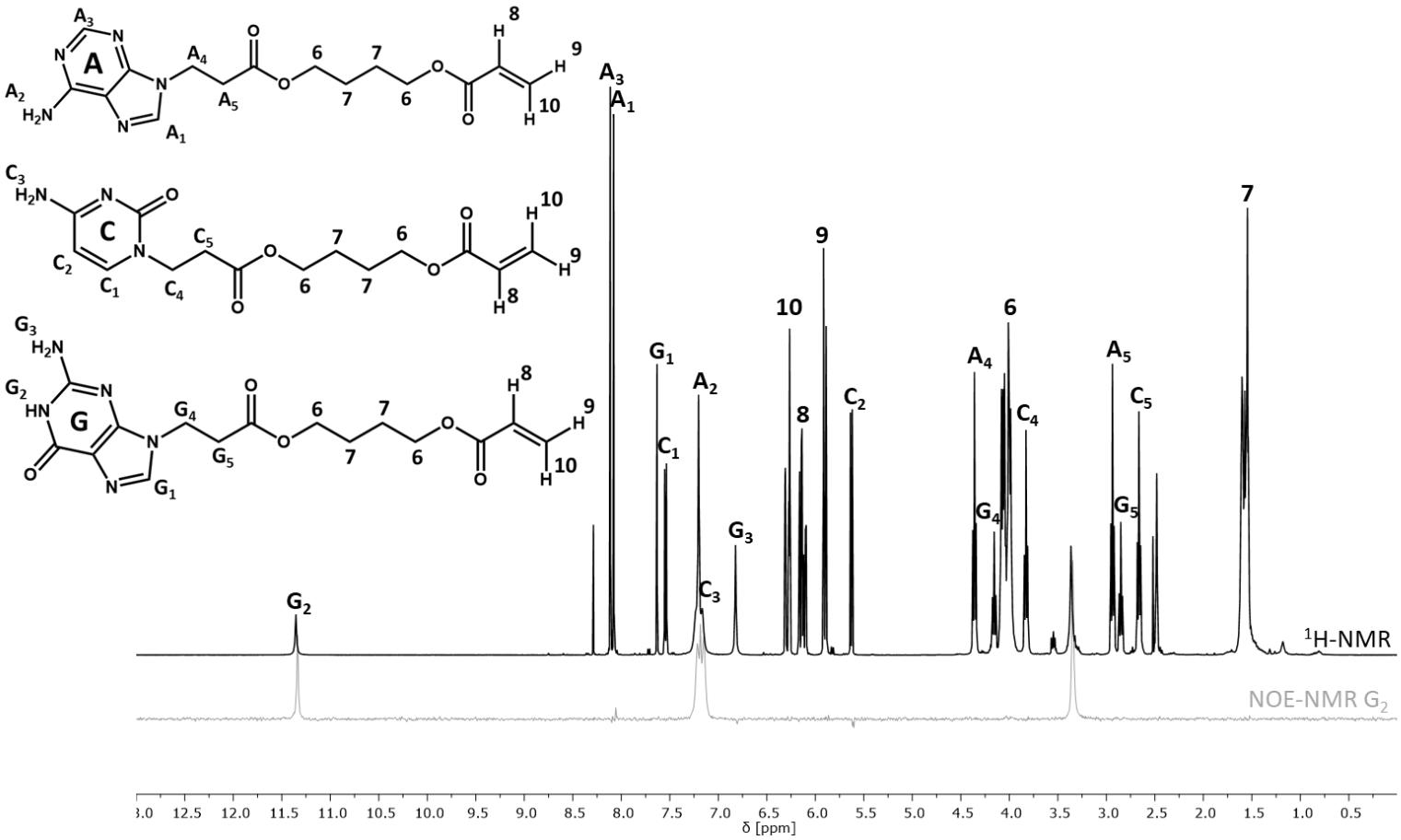


Measuring specifications	
Sample	ACG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	20.3 °C
T(NOE-NMR A_1, A_3)	22.9 °C
γ_A	0.4048
γ_C	0.3796
γ_G	0.2120
δ(NOE-NMR A_1, A_3)	8.1020 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8275
n-hexane	1.2001
DMSO	2.5002
H_2O	3.3840
$CHCl_3$	8.3067

Remark: Due to the irradiation of a frequency-range instead of one specific frequency and the small difference between the respective δ, A_1 and A_3 were both irradiated.

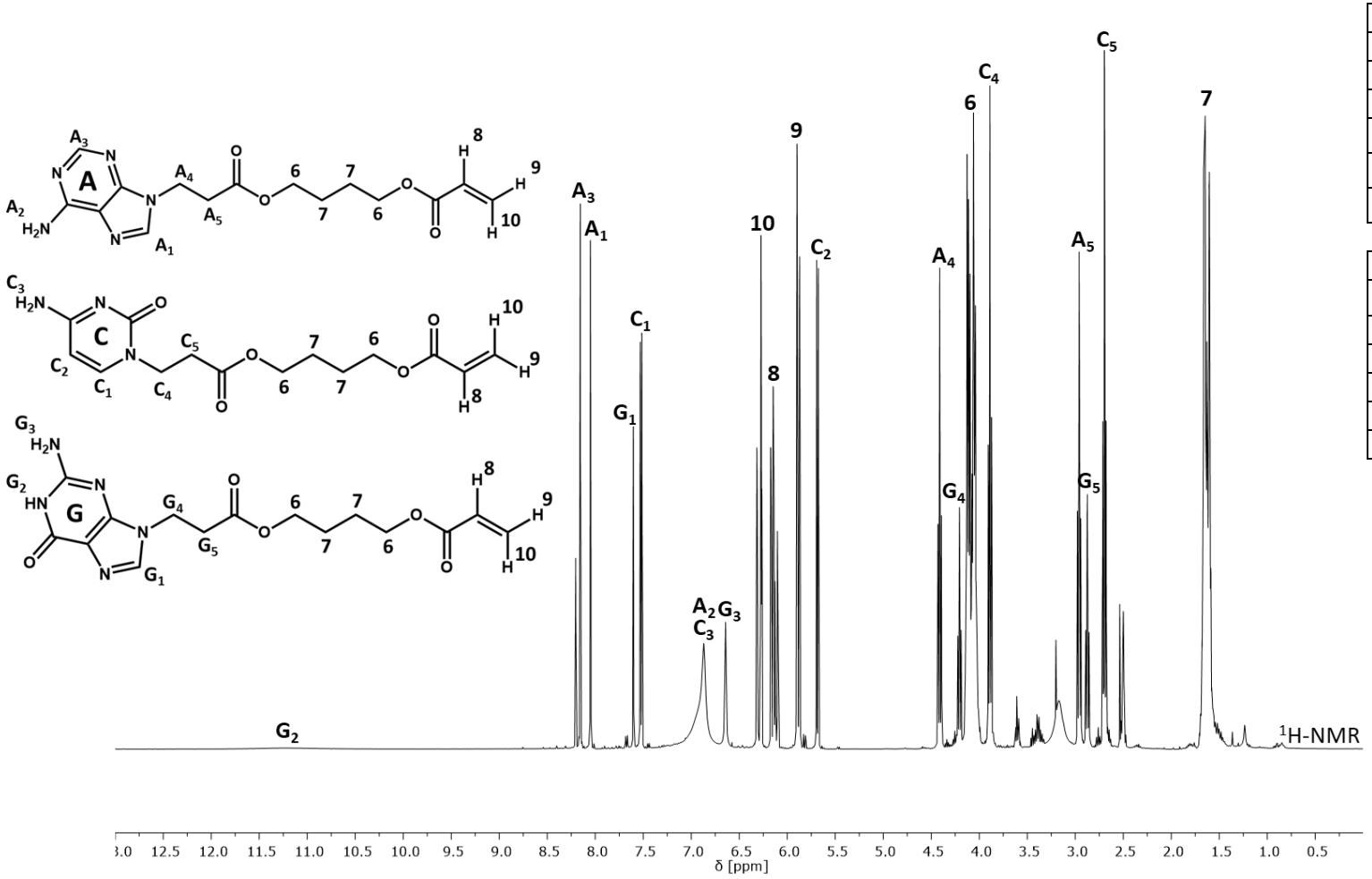
A18. ^{1}H-NMR analysis of sample AGC2 and NOE-analysis of H^{2} of GAM



Measuring specifications	
Sample	ACG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	20.3 °C
T(NOE-NMR G_2)	20.3 °C
γ_A	0.4048
γ_C	0.3796
γ_G	0.2120
δ(NOE-NMR G_2)	11.3351 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8275
n-hexane	1.2001
DMSO	2.5002
H_2O	3.3840
$CHCl_3$	8.3067

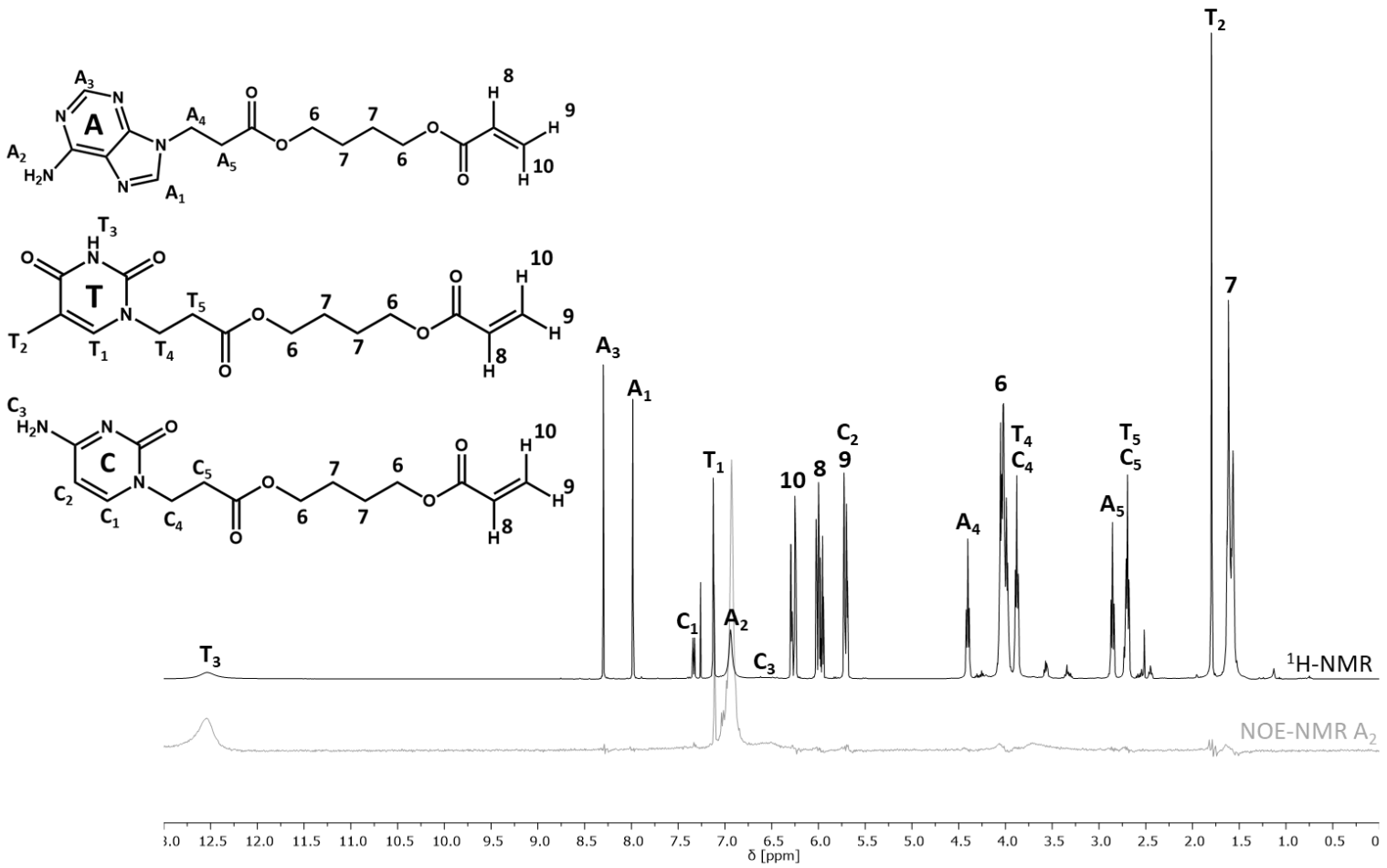
A19. ^{1}H-NMR analysis of sample AGC3



Measuring specifications	
Sample	ACG3
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	80.0 °C
γ_A	0.3232
γ_C	0.4860
γ_G	0.1908

Impurities	
Compound	δ [ppm]
n-hexane	0.8490
n-hexane	1.2360
DMSO	2.5002
H_2O	3.1696
$CHCl_3$	8.2037

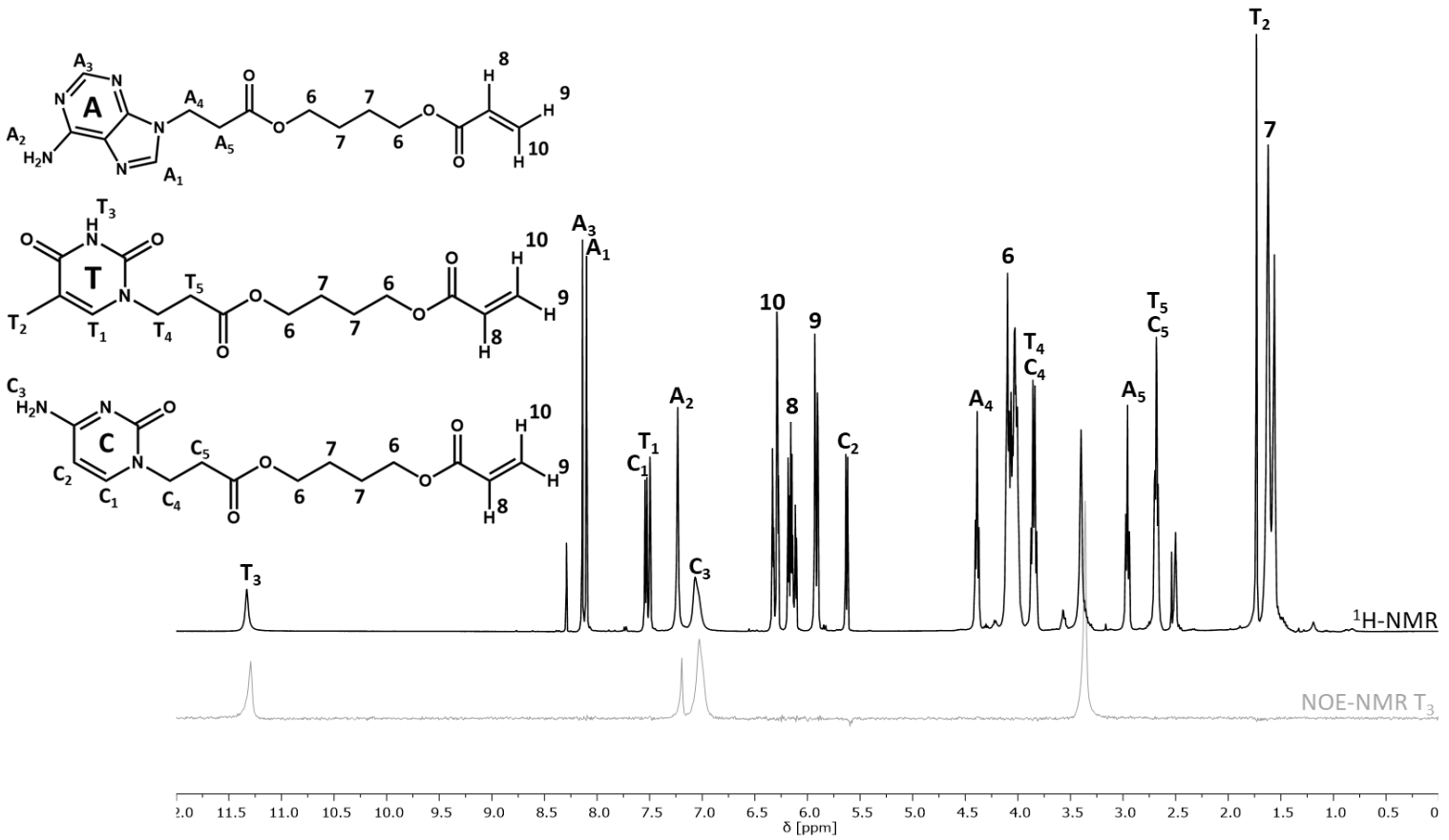
A20. ^{1}H-NMR analysis of sample ATC1 and NOE-analysis of H^{2} of AAM



Measuring specifications	
Sample	ATC1
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.3 °C
T(NOE-NMR A_2)	20.2 °C
γ_A	0.3441
γ_T	0.5128
γ_C	0.1431
δ(NOE-NMR A_2)	6.9303 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.7500
MeOH	1.1301
DMSO	2.5168
DMSO	3.5712
$CHCl_3$	7.2600

A21. ^{1}H-NMR analysis of sample ATC2 and NOE-analysis of H^{3} of TAM



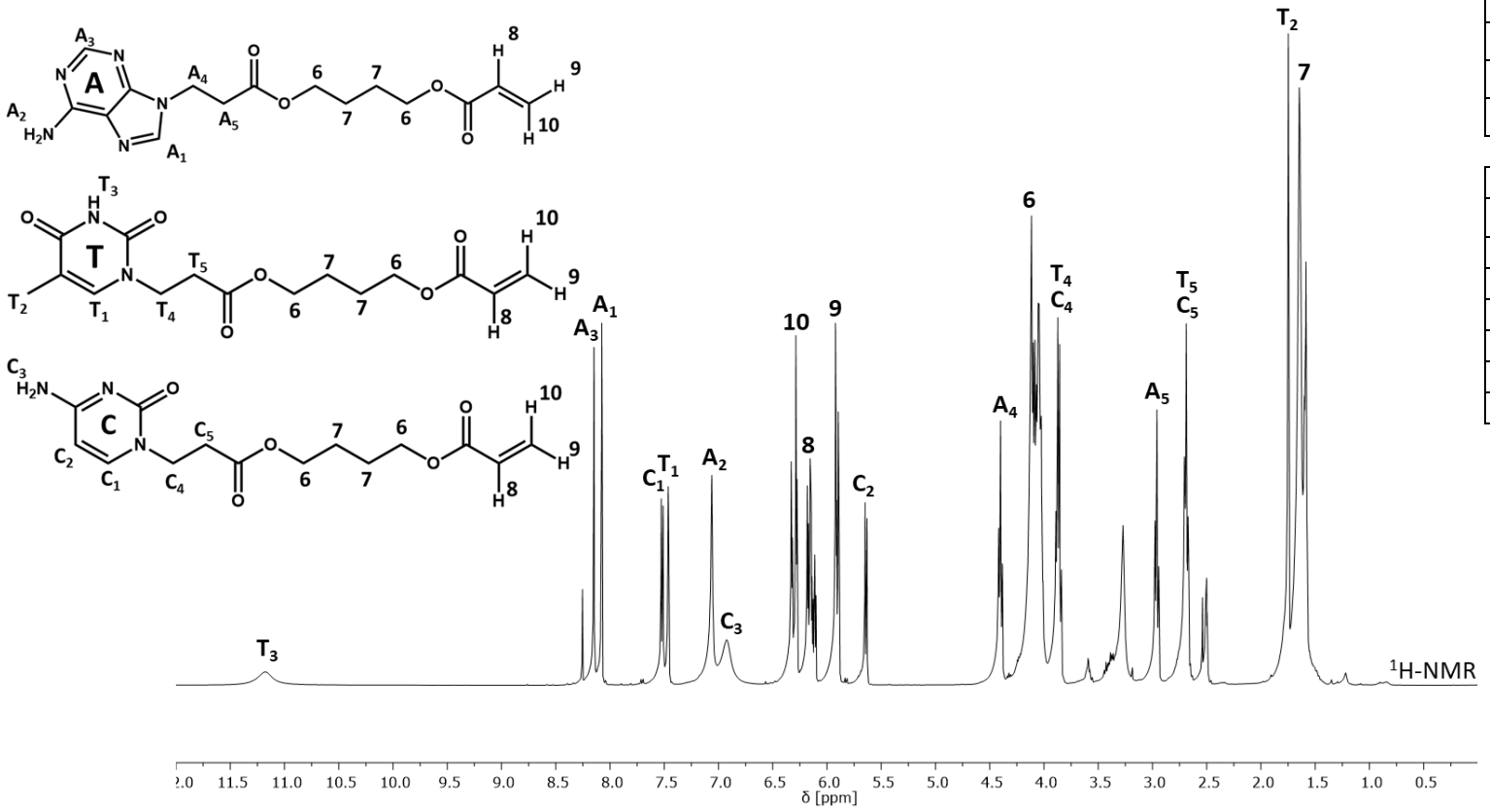
Measuring specifications	
Sample	ATC2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	21.6 °C
T(NOE-NMR T_3)	21.6 °C
γ_A	0.3415
γ_T	0.3293
γ_C	0.3293
δ(NOE-NMR T_3)	11.2918 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8244
n-hexane	1.1913
DMSO	2.5004
H_2O	3.1662
MeOH	3.5719
$CHCl_3$	8.2910

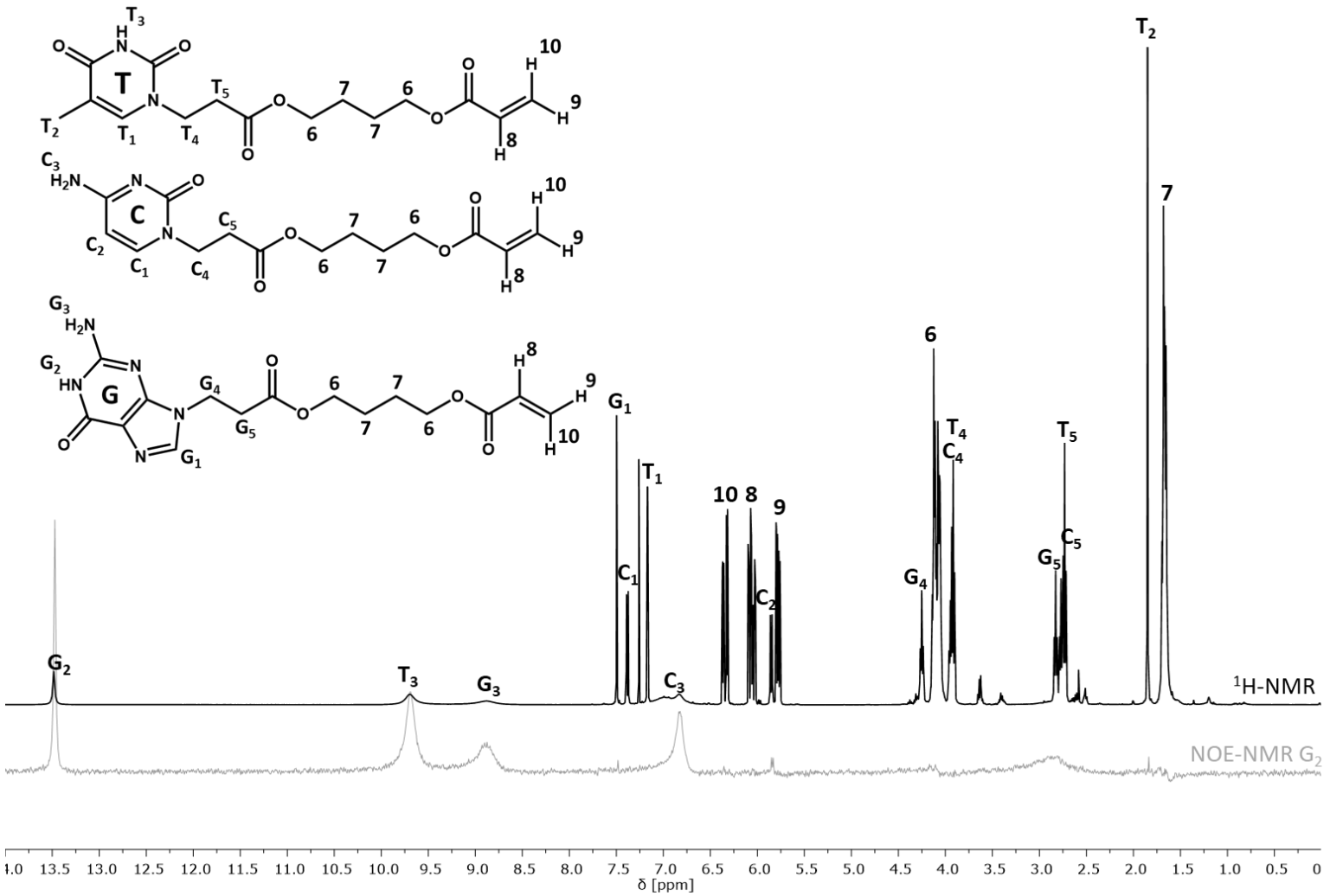
A22. ^{1}H-NMR analysis of sample ATC3

Measuring specifications	
Sample	ATC3
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	50.0 °C
γ_A	0.3580
γ_T	0.3160
γ_C	0.3259

Impurities	
Compound	δ [ppm]
n-hexane	0.8335
n-hexane	1.2188
DMSO	2.5005
H_2O	3.2713
MeOH	3.5938
$CHCl_3$	8.2538



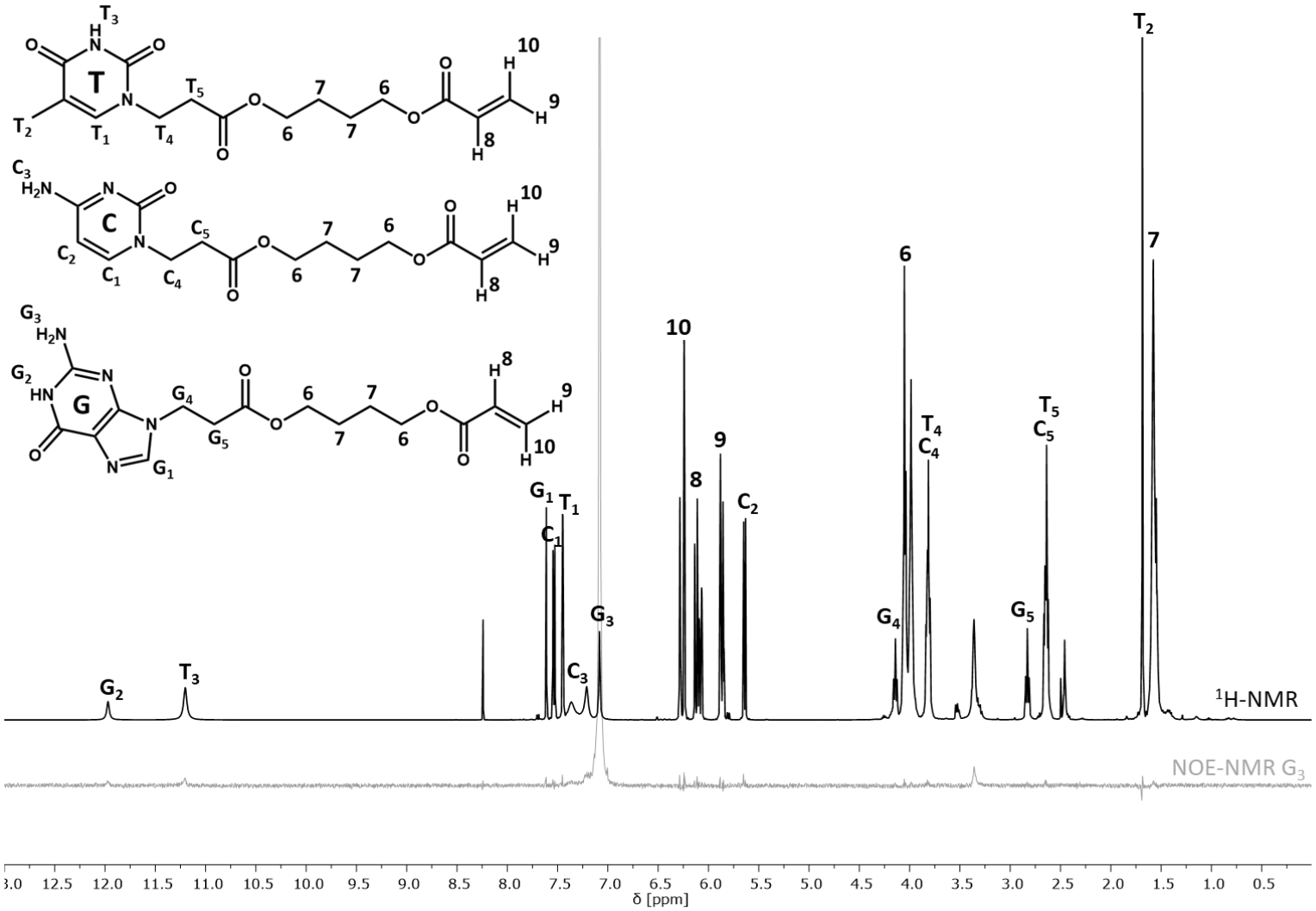
A23. ^{1}H-NMR analysis of sample TCG1 and NOE-analysis of H^{2} of GAM



Measuring specifications	
Sample	TCG1
Solvent	$CDCl_3$
T(^{1}H-NMR)	20.0 °C
T(NOE-NMR G_2)	20.0 °C
γ_T	0.4367
γ_C	0.2800
γ_G	0.2833
δ(NOE-NMR G_2)	13.4730 ppm

Impurities	
Compound	δ [ppm]
Si-grease	0.0208
n-hexane	0.8240
n-hexane	1.2002
BHT	1.3603
BHT	2.0075
TEA	2.5117
H_2O	2.8367
MeOH	3.4109
DCM	3.6236
$CHCl_3$	7.2599

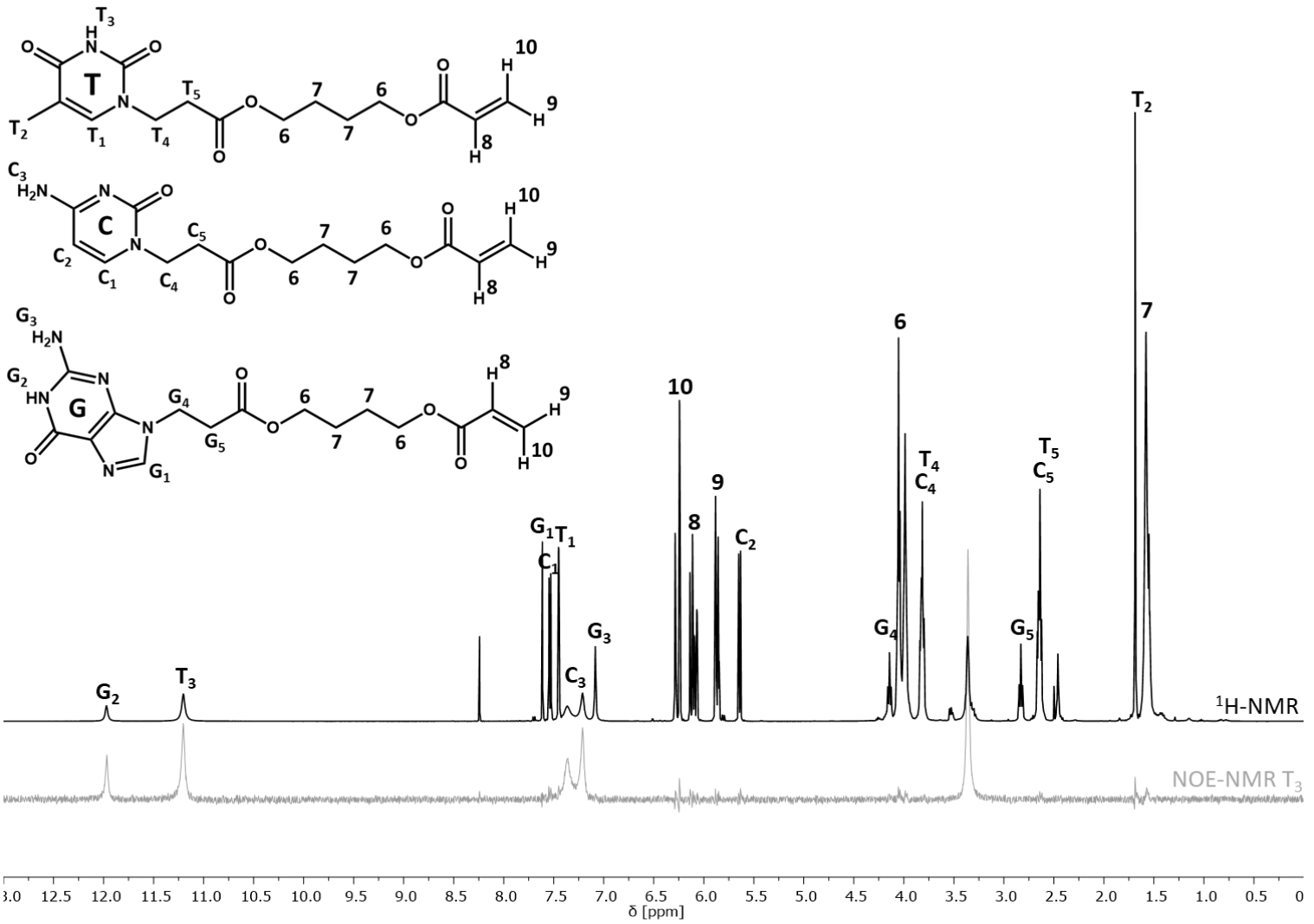
A24. ^{1}H-NMR analysis of sample TCG2 and NOE-analysis of H^{23} of GAM



Measuring specifications	
Sample	TCG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	21.7 °C
T(NOE-NMR G_3)	22.1 °C
γ_T	0.4286
γ_C	0.3878
γ_G	0.1837
δ(NOE-NMR G_3)	7.0934 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8179
TEA	1.0675
n-hexane	1.1832
DMSO	2.5002
H_20	3.4014
DCM	5.8348
$CHCl_3$	8.2815

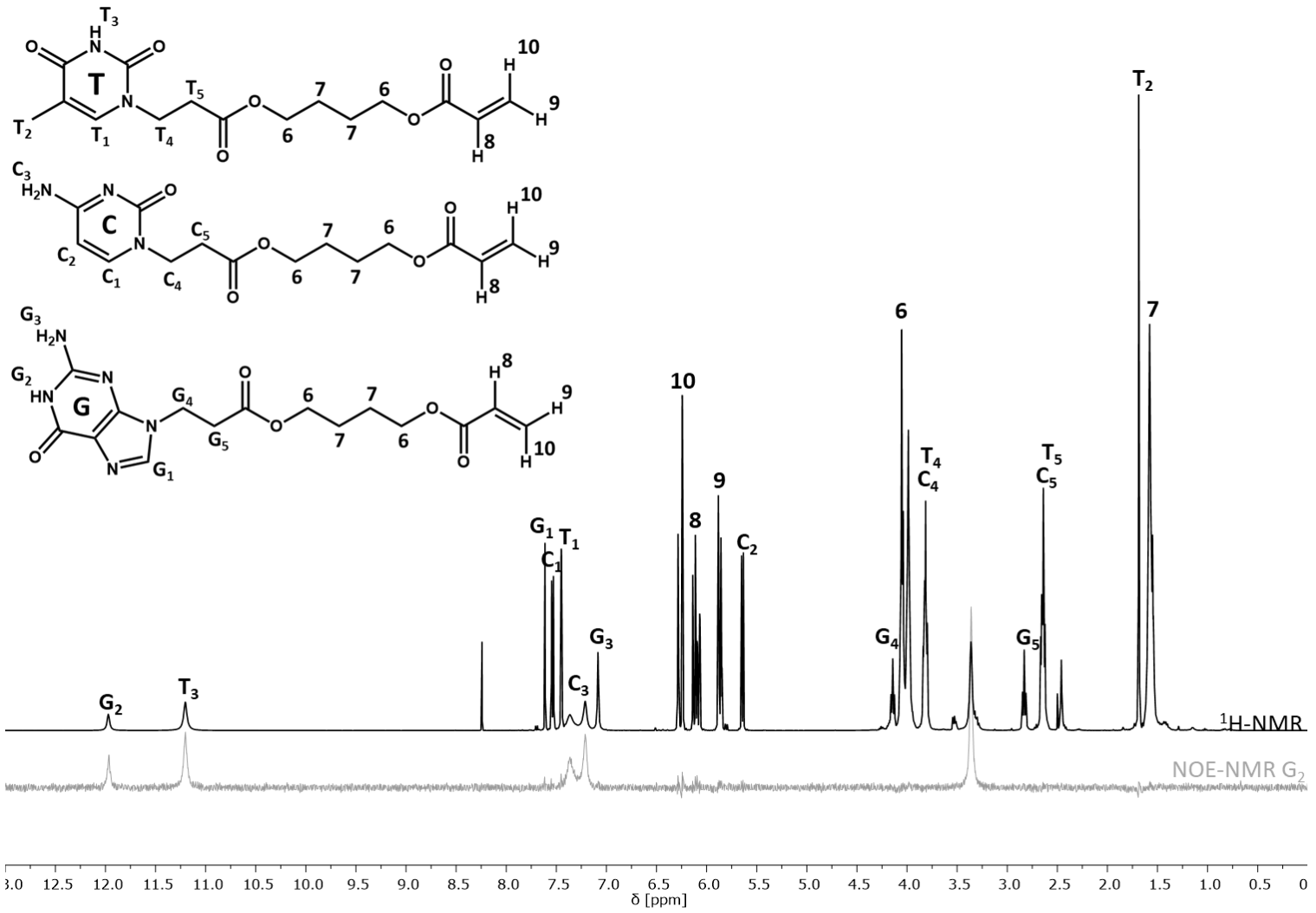
A25. ^{1}H-NMR analysis of sample TCG2 and NOE-analysis of H^{3} of TAM



Measuring specifications	
Sample	TCG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	21.7 °C
T(NOE-NMR T_3)	22.1 °C
γ_T	0.4286
γ_C	0.3878
γ_G	0.1837
δ(NOE-NMR T_3)	11.2059 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8179
TEA	1.0675
n-hexane	1.1832
DMSO	2.5002
H_20	3.4014
DCM	5.8348
$CHCl_3$	8.2815

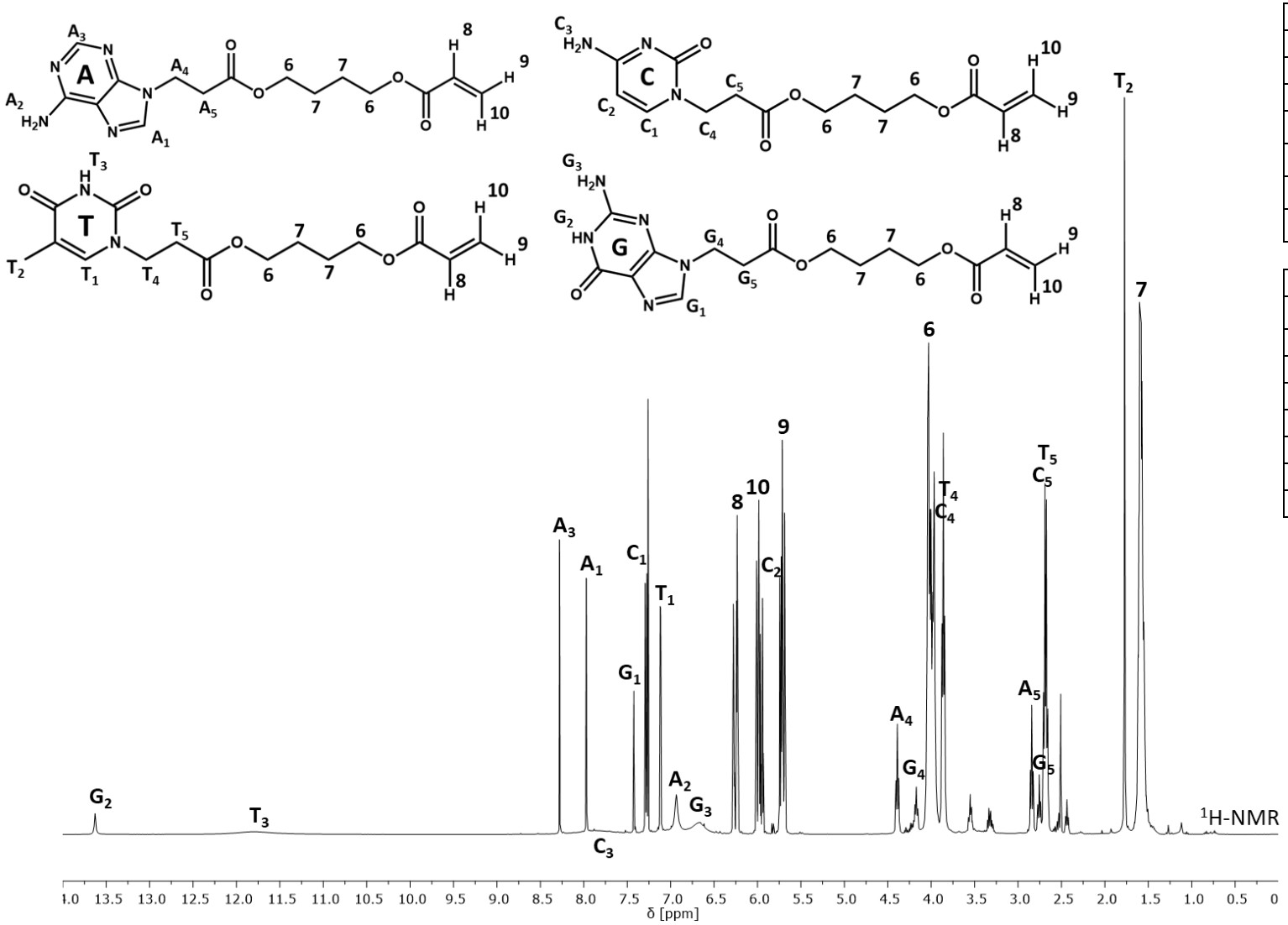
A26. ^{1}H-NMR analysis of sample TCG2 and NOE-analysis of H^{2} of GAM



Measuring specifications	
Sample	TCG2
Solvent	$(CD_3)_2SO$
T(^{1}H-NMR)	21.7 °C
T(NOE-NMR G_2)	22.1 °C
γ_T	0.4286
γ_C	0.3878
γ_G	0.1837
δ(NOE-NMR G_2)	11.9652 ppm

Impurities	
Compound	δ [ppm]
n-hexane	0.8179
TEA	1.0675
n-hexane	1.1832
DMSO	2.5002
H_20	3.4014
DCM	5.8348
$CHCl_3$	8.2815

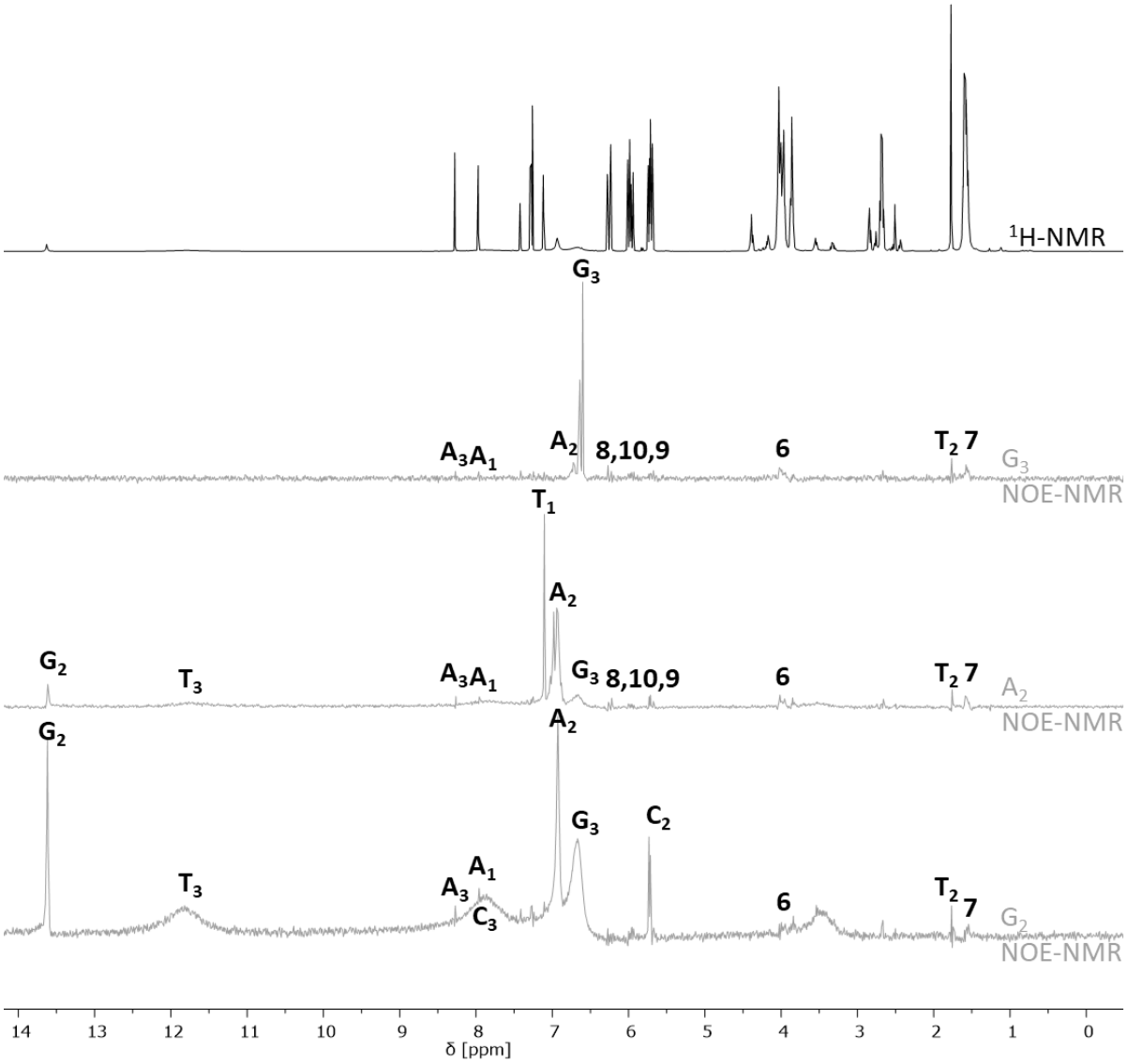
A27. ^{1}H-NMR analysis of sample ATCG



Measuring specifications	
Sample	ATCG
Solvent	$CDCl_4$
T(^{1}H-NMR)	20.1 °C
γ_A	0.1553
γ_T	0.3439
γ_C	0.4075
γ_G	0.0952

Impurities	
Compound	δ [ppm]
n-hexane	0.7821
MeOH	1.1176
n-hexane	1.2701
DMSO	2.4367
DMSO	2.5081
H_2O	3.5498
$CHCl_3$	7.2603

A28. ^{1}H-NMR analysis of sample ATCG and NOE-analysis of H^3 of GAM, H^2 of AAM and H^2 of GAM



Measuring specifications	
Sample	ATCG
Solvent	$CDCl_4$
T(^{1}H-NMR)	20.1 °C
T(NOE-NMR G_3)	20.0 °C
T(NOE-NMR A_2)	20.0 °C
T(NOE-NMR G_2)	19.9 °C
γ_A	0.1553
γ_T	0.3439
γ_C	0.4075
γ_G	0.0952
δ(NOE-NMR G‘’)	6.6563 ppm
δ(NOE-NMR A_2)	6.9266 ppm
δ(NOE-NMR G_2)	13.6189 ppm

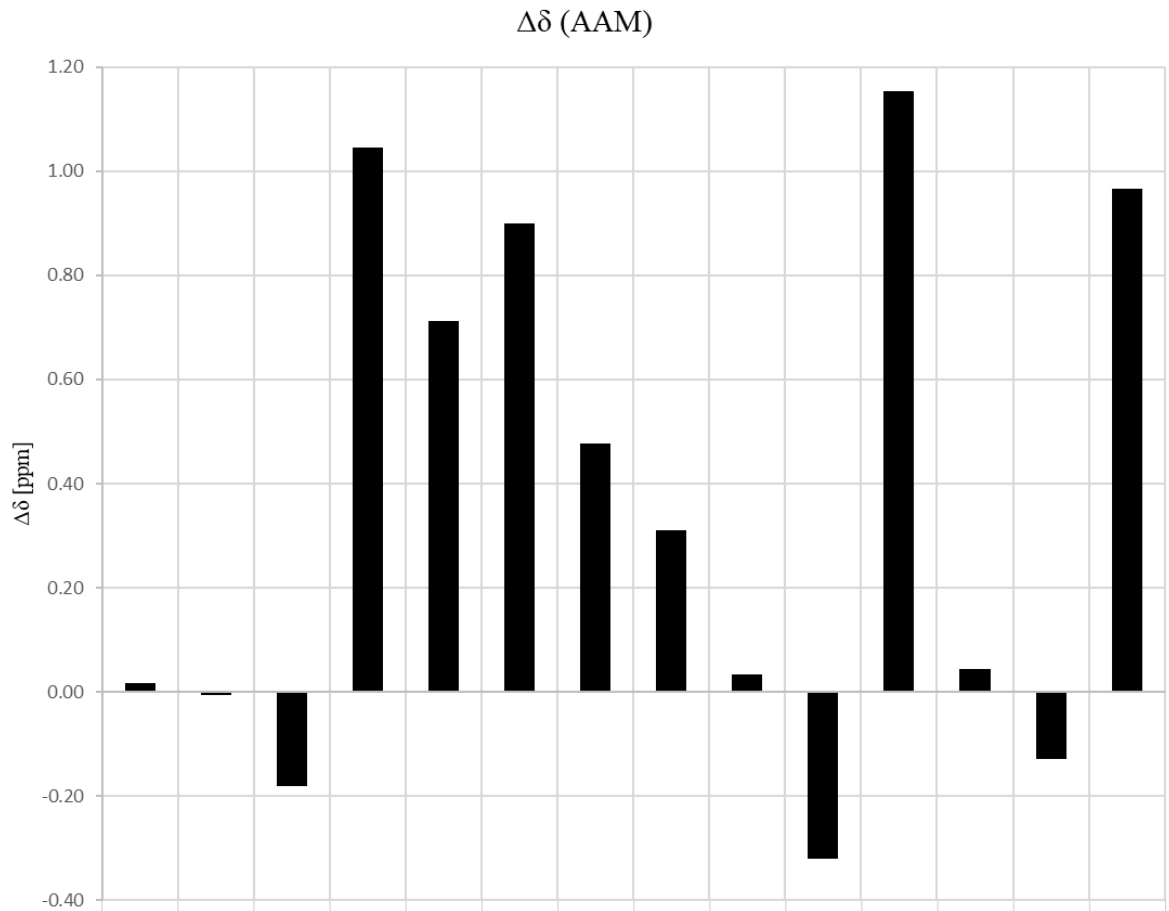
Impurities	
Compound	δ [ppm]
n-hexane	0.7821
MeOH	1.1176
n-hexane	1.2701
DMSO	2.4367
DMSO	2.5081
H_2O	3.5498
$CHCl_3$	7.2603

A29. Data concerning the NMR- and NOE-experiments.

Sample	Adenine			Thymine			Uracil			Guanine					Cytosine					CHCl3	MeOH			DMSO	T [°C]	δ* [ppm]
	γ	δ [ppm]	Δδ [ppm]	γ	δ [ppm]	Δδ [ppm]	γ	δ [ppm]	Δδ [ppm]	γ	δ (NH) [ppm]	Δδ [ppm]	δ (NH2) [ppm]	Δδ [ppm]	γ	δ [ppm]	Δδ [ppm]	δ [ppm]	Δδ [ppm]		S	δ(OH) [ppm]	δ(CH3) [ppm]			
A1	1	5.9697	/	0	/	/	0	/	/	0	/	/	/	/	0	/	/	/	/	X	/	/	/	/	25.0	/
A2	1	7.1905	/	0	/	/	0	/	/	0	/	/	/	/	0	/	/	/	/	/	/	/	/	X	23.0	/
T1	0	/	/	1	8.7069	/	0	/	/	0	/	/	/	/	0	/	/	/	/	X	/	/	/	/	25.0	/
T2	0	/	/	1	11.2461	/	0	/	/	0	/	/	/	/	0	/	/	/	/	/	/	/	/	X	22.8	/
U1	0	/	/	0	/	/	1	11.2264	/	0	/	/	/	/	0	/	/	/	/	/	/	/	/	X	23.0	/
G1	0	/	/	0	/	/	0	/	/	1	10.5385	/	6.4514	/	0	/	/	/	/	/	/	/	/	X	20.7	/
C1	0	/	/	0	/	/	0	/	/	0	/	/	/	/	1	6.4566	/	6.4566	/	X	/	/	/	0	25.0	/
C2	0	/	/	0	/	/	0	/	/	0	/	/	/	/	1	7.0264	/	6.9427	/	/	/	/	/	X	23.3	/
AT1	0.2539	7.2080	0.0175	0.7461	11.2741	0.0280	0	/	/	0	/	/	/	/	0	/	/	/	/	/	/	/	/	X	22.2	11.2347
AT2	0.2506	7.1836	-0.0069	0.7494	11.2533	0.0072	0	/	/	0	/	/	/	/	0	/	/	/	/	/	X	4.1094	3.1698	X	22.2	11.0620
AT3	0.2545	7.0100	-0.1805	0.7455	10.0977	-1.1484	0	/	/	0	/	/	/	/	0	/	/	/	/	/	X	3.9654	3.1863	X	50.0	11.0620
AT4	0.2798	7.0150	1.0453	0.7202	11.4631	2.7562	0	/	/	0	/	/	/	/	0	/	/	/	/	X	/	/	/	/	20.2	11.4516
AT5	0.5187	6.6823	0.7126	0.4813	10.4491	1.7422	0	/	/	0	/	/	/	/	0	/	/	/	/	X	/	/	/	/	23.00	10.4280
AU1	0.3525	6.8694	0.8997	0	/	/	0.6475	12.0090	**0.7826**	0	/	/	/	/	0	/	/	/	/	X	/	/	/	/	20.8	11.9966
GC1	0	/	/	0	/	/	0	/	/	0.3558	11.3588	0.8203	6.8321	0.3807	0.6442	7.2370	0.2106	7.1793	0.2366	/	/	/	/	X	20.2	11.3200
GC2	0	/	/	0	/	/	0	/	/	0.4378	13.8651	**3.3266**	6.8619	**0.4105**	0.5622	8.6018	2.1452	/	0	X	/	/	/	/	21.2	13.6722
CT1	0	/	/	0.7327	10.0034	1.2965	0	/	/	0	/	/	/	/	0.2673	6.1668	-0.2898	/	0	X	/	/	/	/	20.1	/
CT2	0	/	/	0.7327	9.0968	0.3899	0	/	/	0	/	/	/	/	0.2673	/	0	/	0	X	/	/	/	/	50.0	/
AC1	0.4946	6.4472	0.4775	0	/	/	0	/	/	0	/	/	/	/	0.5054	7.6794	1.2228	/	0	X	/	/	/	/	21.0	6.4346
ACG1	0.3724	6.2799	0.3102	0	/	/	0	/	/	0.2985	13.5629	**3.0244**	8.8265	**2.3751**	0.3291	6.9898	0.5332	6.8763	0.4197	X	/	/	/	/	20.2	6.2676 13.5504
ACG2	0.4084	7.223	0.0325	0	/	/	0	/	/	0.2120	11.3747	0.8362	6.8419	0.3905	0.3796	7.2430	0.2166	7.1865	0.2438	/	/	/	/	X	20.3	7.1152 7.3237 8.1020 11.3351
ACG3	0.3232	6.8695	-0.3210	0	/	/	0	/	/	0.1908	11.2051	0.6666	6.6427	0.1913	0.4860	6.8950	-0.1314	/	0	/	/	/	/	X	80.0	/
ATC1	0.3441	7.1238	1.1541	0.5128	12.5345	3.8276	0	/	/	0	/	/	/	/	0.1431	7.0514	0.5948	6.5811	0.1245	X	/	/	/	/	20.3	6.9303
ATC2	0.3415	7.2335	0.0430	0.3293	11.3296	0.0835	0	/	/	0	/	/	/	/	0.3293	7.0678	0.0414	/	0	/	/	/	/	X	21.6	11.2918
ATC3	0.3580	7.0612	-0.1293	0.3160	11.1819	-0.0642	0	/	/	0	/	/	/	/	0.3259	6.9238	-0.1026	/	0	/	/	/	/	X	50.0	/
TCG1	0	/	/	0.4367	9.6982	0.9913	0	/	/	0.2833	13.4840	**2.9455**	8.8825	**2.4311**	0.2800	6.9871	0.5305	6.8344	0.3778	X	/	/	/	/	20.0	13.4730
TCG2	0	/	/	0.4286	11.2433	-0.0028	0	/	/	0.1837	12.0113	1.4728	7.1241	0.6727	0.3878	7.4049	0.3785	7.2516	0.3089	/	/	/	/	X	21.7	7.0934 11.2059 11.9652
ATCG	0.1553	6.9352	0.9655	0.3439	11.8010	3.0941	0	/	/	0.0952	13.6282	**3.0897**	6.6665	**0.2151**	0.4057	7.8785	1.4219	6.8240	0.3674	X	/	/	/	/	20.1	6.6563 6.9663 13.6189

δ* represents the shift(s) at which irradiation of a proton(s) was executed for the NOE-analysis.
The bold numbers represent relative shifts that are obtained by subtraction of shifts measured in DMSO from shifts measured in $CHCl_3$ since the single NAM does not dissolve in $CHCl_3$.
X indicates the use of the corresponding solvent to prepare the analyte.

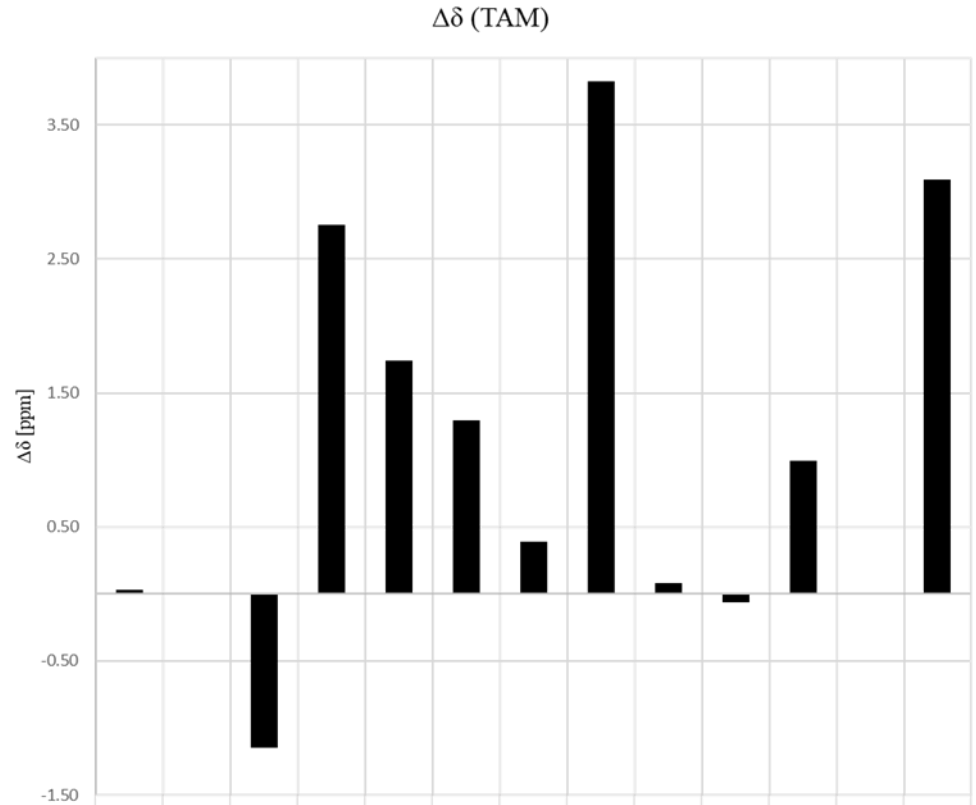
A30. Δδ of proton H^2 belonging to AAM



	AT1	AT2	AT3	AT4	AT5	AU1	AC1	ACG1	ACG2	ACG3	ATC1	ATC2	ATC3	ATCG
$CDCl_3$	/	/	/	X	X	X	X	X	/	/	X	/	/	X
DMSO	X	X	X	/	/	/	/	/	X	X	/	X	X	/
MeOH	/	X	X	/	/	/	/	/	/	/	/	/	/	/
T [°C]	22.2	22.2	50.0	20.2	23.0	20.8	21.0	20.2	20.3	80.0	20.3	21.6	50.0	20.1
γ_{AAM}	0.25	0.25	0.25	0.28	0.52	0.35	0.49	0.37	0.41	0.32	0.34	0.34	0.36	0.16
γ_{TAM}	0.75	0.75	0.75	0.72	0.48	/	/	/	/	/	0.51	0.33	0.32	0.34
γ_{UAM}	/	/	/	/	/	0.65	/	/	/	/	/	/	/	/
γ_{GAM}	/	/	/	/	/	/	/	0.30	0.21	0.19	/	/	/	0.10
γ_{CAM}	/	/	/	/	/	/	0.51	0.33	0.38	0.49	0.14	0.33	0.33	0.41

The Δδ is calculated as the difference between the measured shift in the sample and the shift of H^2 from pure AAM measured in the same solvent (DMSO or $CDCl_3$) at room temperature. Additional experimental data is given below each result.

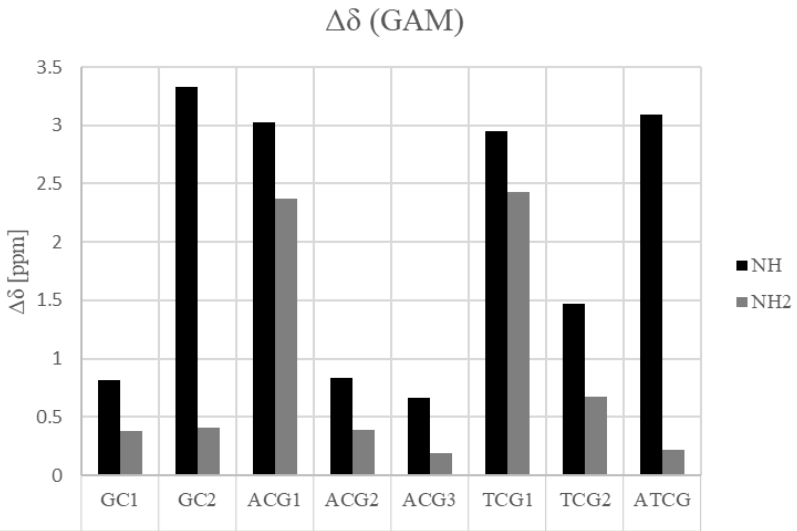
A31. Δδ of H^3 belonging to TAM



	AT1	AT2	AT3	AT4	AT5	CT1	CT2	ATC1	ATC2	ATC3	TCG1	TCG2	ATCG
$CDCl_3$	/	/	/	X	X	X	X	X	/	/	X	/	X
DMSO	X	X	X	/	/	/	/	/	X	X	/	X	/
MeOH	/	X	X	/	/	/	/	/	/	/	/	/	/
T [°C]	22.2	22.2	50.0	20.0	23.0	20.1	50.0	20.3	21.6	50.0	20.0	21.7	20.1
γ_{AAM}	0.25	0.25	0.25	0.28	0.52	/	/	0.34	0.34	0.36	/	/	0.16
γ_{TAM}	0.75	0.75	0.75	0.72	0.48	0.73	0.73	0.51	0.33	0.32	0.44	0.43	0.34
γ_{GAM}	/	/	/	/	/	/	/	/	/	/	0.28	0.18	0.10
γ_{CAM}	/	/	/	/	/	0.27	0.27	0.14	0.33	0.33	0.28	0.39	0.41

The Δδ is calculated as the difference between the measured shift in the sample and the shift of H^3 from pure AAM measured in the same solvent (DMSO or $CDCl_3$) at room temperature. Additional experimental data is given below each result.

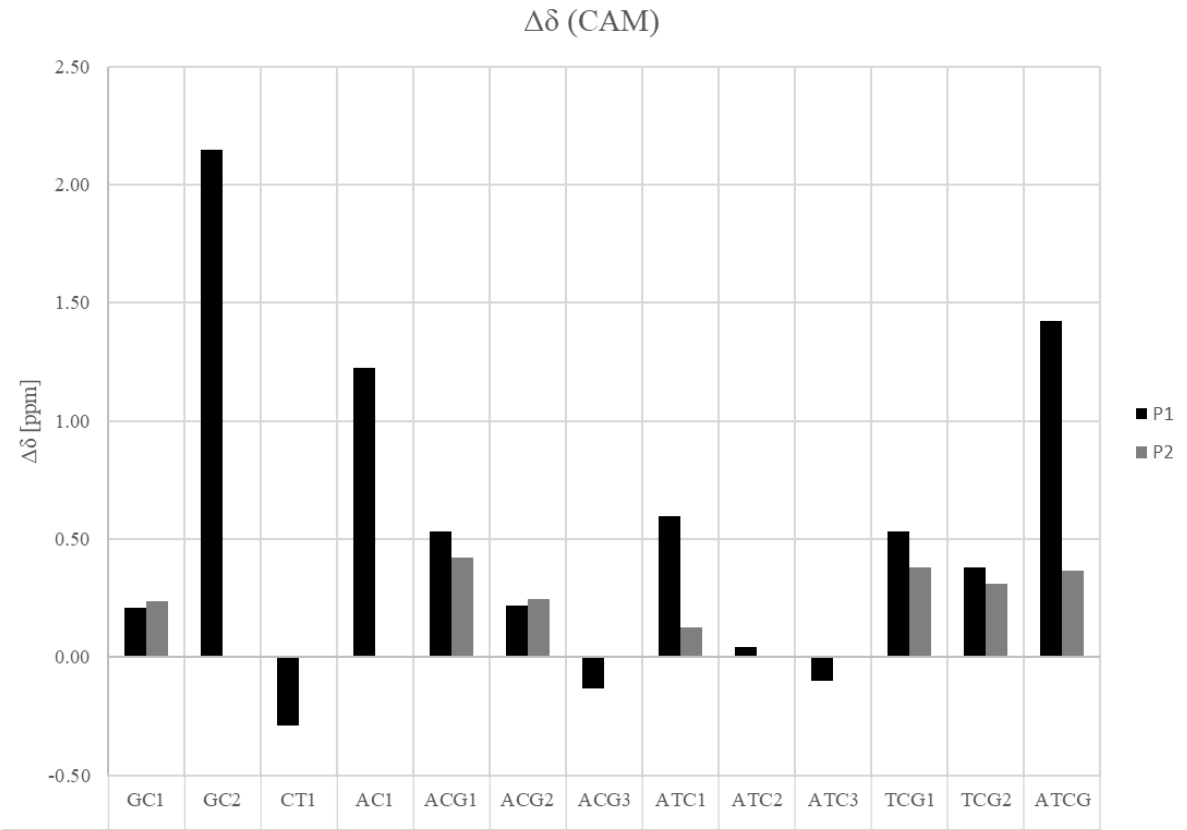
A32. Δδ of H^2 (black) and H^3 (grey) belonging to TAM



	GC1	GC2	ACG1	ACG2	ACG3	TCG1	TCG2	ATCG
$CDCl_3$	/	X	X	/	/	X	/	X
DMSO	X	/	/	X	X	/	X	/
T [°C]	20.2	21.2	20.2	20.3	80.0	20.0	21.7	20.1
γ_{AAM}	/	/	0.37	0.41	0.32	/	/	0.16
γ_{TAM}	/	/	/	/	/	0.44	0.43	0.34
γ_{GAM}	0.36	0.44	0.30	0.21	0.19	0.28	0.18	0.10
γ_{CAM}	0.64	0.56	0.33	0.38	0.49	0.28	0.39	0.41

The Δδ is calculated as the difference between the measured shift in the sample and the shift from pure TAM measured in DMSO at room temperature. Additional experimental data is given below each result.

A33. Δδ of proton H^3 belonging to CAM



	GC1	GC2	CT1	AC1	ACG1	ACG2	ACG3	ATC1	ATC2	ATC3	TCG1	TCG2	ATCG
$CDCl_3$	/	X	X	X	X	/	/	X	/	/	X	/	X
DMSO	X	/	/	/	/	X	X	/	X	X	/	X	/
T [°C]	20.2	21.2	20.1	21.0	20.2	20.3	80.0	20.3	21.3	50.0	20.0	21.7	20.1
γ_{AAM}	/	/	/	0.49	0.37	0.41	0.32	0.34	0.34	0.36	/	/	0.16
γ_{TAM}	/	/	0.73	/	/	/	/	0.51	0.33	0.32	0.44	0.43	0.34
γ_{GAM}	0.36	0.44	/	/	0.30	0.21	0.19	/	/	/	0.28	0.18	0.10
γ_{CAM}	0.64	0.56	0.27	0.51	0.33	0.38	0.49	0.14	0.33	0.33	0.28	0.39	0.41

The Δδ is calculated as the difference between the measured shift in the sample and the shift of proton H^3 from pure CAM measured in the same solvent (DMSO or $CDCl_3$) at room temperature. Additional experimental data is given below each result. Note that the respective proton renders a doublet when dissolved in DMSO. This means that the proton is identified by a doublet consisting of two δ (P1 and P2). The absence of data for P2 does not mean that its Δδ equals zero. It merely means that the doublet is observed as a singlet in the respective sample. Therefore, no Δδ can be determined for this signal.