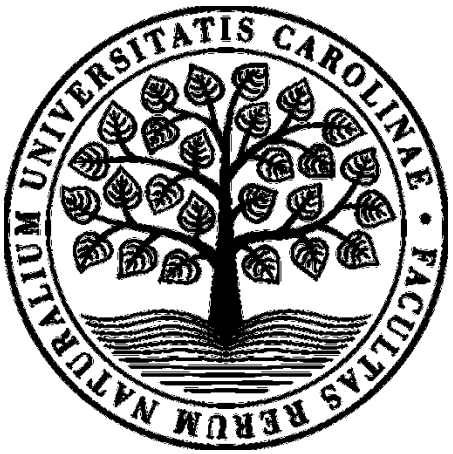
Charles University
Faculty of Science

Course of Study: Chemistry
Subject: Organic Chemistry



Bc. Zuzana Osifová

Mezimolekulové vodíkové vazby studované pomocí NMR spektroskopie

Intermolecular Hydrogen Bonds Studied by NMR Spectroscopy

Master Thesis

Supervisor: doc. RNDr. Martin Dračínský, Ph.D.

Praha, 2020

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité citační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 25. 5. 2020 ……

Bc. Zuzana Osifová

This work was supported by Czech Science Foundation (GA ČR), No. 18-11851S

Acknowledgment

I am very grateful to my supervisor doc. RNDr. Martin Dračínský, Ph.D. for the opportunity to work on this project, for his patience, willingness and constructive advice. I would like to thank to my colleagues at Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences – namely Jakub Radek Štoček, for his advice and support, Mgr. Ondřej Socha for the fitting of chemical shifts dependencies, and Ing. Lucie Čechová and RNDr. Michal Šála, Ph.D. for the synthesis of nucleobase derivatives.

Last but not least, I owe my thanks to my family, for their loving support during my study.

Abstrakt

N–methylace nukleových kyselin je jeden z nejdůležitějších epigenetických mechanismů v genomech organismů a virů. *N*–methylované nukleobáze jsou v lidském genomu registrovány ve spojitosti s fyziologickými i patologickými procesy. Například methylace adeninu vedoucí ke vzniku N^6-methyladeninu je spojována s rozvojem Alzheimerovy nemoci či obezity. Tyto systémy mohou být snadno studovány pomocí nízkoteplotní NMR spektroskopie, protože řád vazby mezi purinovým kruhem a jeho substituentem má částečně charakter násobné vazby. Díky omezené rotaci kolem této vazby lze jednotlivé rotamery odlišit v nízkoteplotních NMR spektrech jako dvě sady signálů.

Tato diplomová práce je zaměřena na výzkum rotamerních rovnováh *N*–methylovaných nukleobází a na výzkum změn volné Gibbsovy energie spojené se vznikem jejich páru s komplementárním partnerem za pomoci jednoduché a přímé metody NMR spektroskopie. Touto metodou jsme prokázali, že rotamerní rovnováhy *N*–methylovaných derivátů adeninu jsou závislé na teplotě, rozpouštědle a na koncentraci komplementárního partnera. S pomocí námi nově vyvinuté metodiky založené na pozorování změn chemických posunů závislých na koncentraci vazebného partnera jsme získali nejen geometrie vznikajících párů komplementárních derivátů adeninu s thyminem, ale i změny Gibbsovy energie spojené s jejich vznikem. Všechna experimentální data byla podpořena DFT výpočty.

Klíčová slova

NMR spektroskopie, *N*–methylace DNA, pár nukleobází, rotamerní rovnováha, Gibbsova energie, vodíková vazba, DFT výpočty

Abstract

N–methylation of nucleic acids is one of the most important epigenetic mechanisms in organisms and viruses and it is observed in connection with physiological or pathological processes in human body. For example, the *N*–methylation of adenine to give N^6-methyladenine is associated with development of obesity or Alzheimer disease. These systems can be successfully studied by low-temperature NMR spectroscopy because the bond between a purine ring and its substituent has a significant double-bond character and, due to restricted rotation around this bond, two sets of signals are observed in low-temperature NMR spectra.

This diploma thesis is focused on research of *N*–methylated nucleobases' rotamer equilibria and the free-energy changes associated with the hydrogen-bonded base pair formation via simple and straightforward method based on NMR spectroscopy monitoring. We proved that the rotamer equilibria of *N*–methylated adenine derivatives are dependent on temperature, solvent and bonding partner concentration. We also obtained the geometry of formed intermolecular complexes of *N*–methylated adenine derivatives with thymine and the free-energy changes associated with their formations via newly developed method based on chemical shift changes dependent on bonding partner concentration. All obtained data were supported by DFT calculations.

Keywords

NMR spectroscopy, methylation of DNA, base-pair, rotamer equilibrium, free-energy changes, hydrogen-bond, DFT calculations

Content

List of abbreviations

AIDS	*acquired immune deficiency syndrome*
CW	*continuous-wave*
DFT	*density functional theory*
DNA	*deoxyribonucleic acid*
DNMR	*dynamic nuclear magnetic resonance*
DOSY	*diffusion-ordered spectroscopy*
EWG	*electron withdrawing group*
FID	*free induction decay*
FT	*Fourier transform*
FTO	*fat mass and obesity-associated protein*
GTP	*guanosine triphosphate*
HMBC	*heteronuclear multiple bond correlation*
LC-MS/MS	*liquid chromatography with tandem mass spectrometry*
MRI	*magnetic resonance imaging*
mRNA	*messenger ribonucleic acid*
NMR	*nuclear magnetic resonance*
RNA	*ribonucleic acid*
SAM	*S-adenosyl methionine*
SARS	*severe acute respiratory syndrome*
SMRT	*single-molecule real-time sequencing*
S_N2	*bimolecular nucleophilic substitution*
tRNA	*transfer ribonucleic acid*
TROSY	*transverse relaxation-optimized spectroscopy*

1. Introduction

1.1 NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is one of the most important tools for determining the structure of (bio)molecules. In addition to laboratory applications, NMR is also used in medicine (MRI), food industry and customs. The following text is drawn up to summarize the basics of NMR spectroscopy. It relies on references.[1-7]

In December 1945 and in January 1946, the Purcell group at Harvard University and the Bloch group at Stanford University independently succeeded in observing the phenomenon of NMR. The basis of NMR spectroscopy is formed by the magnetic properties of atomic nuclei. These nuclei possess angular momentum and exhibit magnetic moment, μ.

$$\mu = \frac{\gamma h m}{2\pi} \tag{1.1}$$

where γ (in rad · T^{-1} · s^{-1}) is magnetogyric ratio, h is Planck constant (in J · s) and m is magnetic quantum number.

In the absence of a magnetic field, there are no differences in the behaviour of nuclei with a zero or non-zero spin quantum number. Only in a magnetic field B_0, the nuclear spin affects the behaviour of the nuclei and their magnetic moments´ vectors start to precess about the direction of the external field B_0. The frequency of precession is given by equation (1.2) and it is called Larmor frequency.

$$\nu_0 = \frac{\gamma \cdot B_0}{2\pi} \tag{1.2}$$

where γ (in rad · T^{-1} · s^{-1}) is magnetogyric ratio and B_0 is the magnitude of the applied magnetic field in T.

The magnetic moment, which precesses about the direction of the external field with Larmor frequency, can occupy *n* magnetic states according to quantum mechanics principles (equation 1.3).

$$n = 2l + 1 \tag{1.3}$$

where *l* is spin quantum number.

There are three types of nuclei; those with zero spin quantum number (^{12}C, ^{16}O) without spin; those with spin quantum number = ½ (^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P); and finally, those with spin quantum number higher than ½. Because nuclei with l = ½ are the most important in NMR spectroscopy, the following text is simplified for this case.

The eigenstates of these nuclei (α and β) are degenerate in the absence of magnetic field. Only in a static magnetic field is this degeneration lifted and both states have different energy. The energy difference, ΔE, is given by equation (1.4).

$$\Delta E = \gamma \cdot h \cdot B_0 \tag{1.4}$$

where γ (in rad · T^{-1} · s^{-1}) is magnetogyric ratio, *h* is Planck constant (in J · s) and B_0 is the magnitude of the applied magnetic field in T.

As shown in Figure 1.1, the energy difference is dependent on magnetic field strength. The distribution of nuclei between ground and excited state is given by the simplified Boltzmann relation.

$$\frac{N_\alpha}{N_\beta} = \frac{\Delta E}{k_b \cdot T} \tag{1.5}$$

where *N* is the number of nuclei, k_b is Boltzmann constant and *T* is thermodynamic temperature (in K).

The distribution of nuclei shown in Figure 1.1 is schematic. Due to the very small energy difference between the two spin states, the number of nuclei in the ground state at equilibrium is only slightly higher than the number of nuclei in the excited state. This small excess of nuclei builds up a macroscopic quantity – magnetization ***M***. The magnetization is oriented along the axis of B_0 (*z*-axis in Cartesian coordinate system).

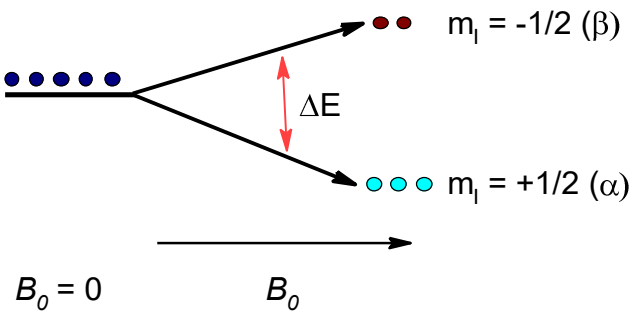


Figure 1.1 Energy separation between nuclear spin states – Zeeman splitting.

.

The fact that different bonded nuclei have slightly different electron distribution in their close vicinity is crucial for determining molecular structure. The motion of electrons creates additional magnetic field in the opposite direction to B_0.

$$B_{ef.} = B_0 \cdot (1 - \sigma) \tag{1.6}$$

where $B_{ef.}$ is effective magnetic field and σ is a shielding constant.

To determine the resonance frequency of a nuclei regardless of the magnetic field strength B_0, chemical shift δ is introduced. As equation (1.7) implies, the chemical shift δ is independent on the magnetic field strength.

$$\delta_x = \frac{(\nu_x - \nu_{ref})}{\nu_{ref}} \cdot 10^6 \tag{1.7}$$

where ν_x is absolute resonant frequency of nucleus X and ν_{ref} is the absolute resonant frequency of a reference nucleus.

In the infancy of NMR spectroscopy techniques, the continuous wave (CW) model was employed. The sample in a CW spectrometer is placed in a homogeneous magnetic field and irradiated by slowly sweeping the required frequency range. The spectrum is obtained by recording the energy absorption as a function of frequency. The commonly used terms “down-field” and “up-field” originate from this approach. The CW techniques have been completely superseded by pulsed Fourier transformation (FT) NMR techniques. In FT NMR, magnetization $\boldsymbol{M}$ is deflected from the z-axis by a radiofrequency pulse P. The precessing magnetization induces electric current in a

receiving coil. The receiver measures free induction decay (FID). The FID results from the superposition of all frequencies and the noise. The acquired intensity-time dependency is subjected to Fourier transform (FT) to provide intensity-frequency dependency (a spectrum). This principle is shown below in Figure 1.2. The scheme is simplified for nuclei with one resonant frequency. The FID obtained for a simple organic molecule is much more complicated.

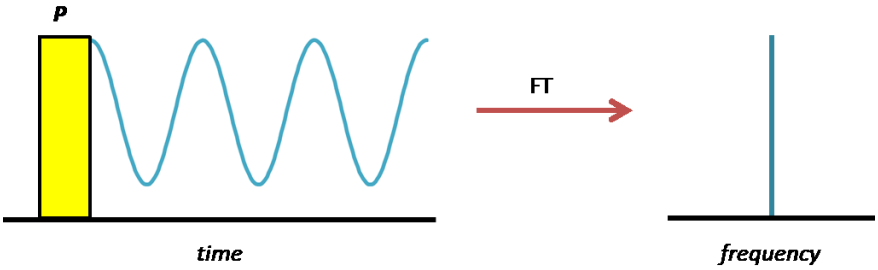


Figure 1.2 Simplified principle of the FT NMR experiment with one frequency

1.1.1 Dynamic NMR

The chemical shift and spin-spin coupling can be affected by temperature-dependent reversible processes like valence tautomerism, conformational changes or non-covalent bonding. The NMR spectra of some compounds therefore change with temperature. This phenomenon is known as dynamic nuclear magnetic resonance (DNMR) or chemical exchange.[1, 3]

The basis of chemical exchange is usually shown on *N,N*-dimethylformamide. The C-N bond between the nitrogen atom and the carbonyl group has a partially double bond character – the rotation around this bond is restricted. Therefore, the protons of the two methyl groups are in different magnetic environment. The differences in magnetic environment of the nuclei lead to different resonance frequencies, as depicted in Figure 1.3.

Because of the relatively high energy barrier of the rotation, two separate signals are observed in ^{1}H NMR spectrum at room temperature. The observation of two separate peaks originates in the slow exchange rate. The rate of chemical exchange is usually temperature-dependent (it increases with temperature) and one can observe one peak for both DMF methyl groups at the temperature of coalescence and higher.[1-4]

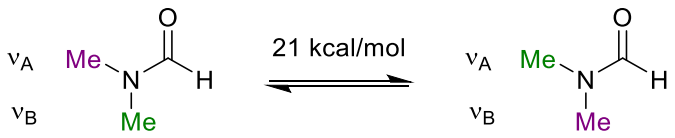


Figure 1.3 The rotation around the C-N bond in *N,N*-dimethylformamide. Created according to reference[1]

This observation is based on the fact, that the exchange rate constant k is much smaller than the frequency difference between inequivalent protons $\Delta\nu$ at low temperature. As shown in Figure 1.4 the spectrum consists of two sharp singlets at ν_A and ν_B. With increasing temperature, the lines start to broaden, and the spectrum consists of two overlapping lines. At the coalescence temperature, two separate peaks merge to one and exchange constant is defined by equation (1.8).

$$k = \frac{\pi \cdot \Delta\nu}{\sqrt{2}} \tag{1.8}$$

where $\Delta\nu$ is the separation (in Hertz) between the two sharp singlets.

Equation (1.8) is valid except for processes with other than first order kinetics, processes with coupled exchanging nuclei or with non-equal intensity of sharp singlets. At high temperatures, the spectrum consists of only one sharp singlet for the rate constant k is much bigger than the frequency difference. [3]

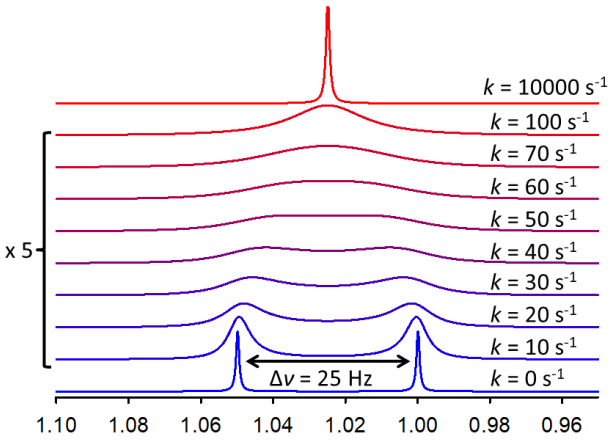


Figure 1.4 The simulated spectra illustrating the chemical exchange (500 MHz)[3]

The specific temperature of signal coalescence is dependent on the used magnetic field strength, as the frequency difference between signals scales with magnetic field, and the two peaks in stronger fields thus merge at a higher temperature.

As mentioned above, in the case of *N,N*-dimethylformamide, the coalescence temperature is relatively high because of a high energy barrier of rotation around the C–N bond. More often, one encounter cases where the rotation around a formally single bond is much faster due to a low energy barrier. The coalescence temperature of these systems can be deep below 273 K and it requires a cooled-down sample to observe two separate sets of signals. Other motivations of low-temperature NMR spectroscopy applications are immobilization of (bio)molecules, observation of meta-stable intermediates, study of processes that occur only at low temperatures and enhancement of the signal-to-noise ratio.[8-9]

A nice example to summarize this section is the study of rotational equilibria of 5-phenylazopyrimidines executed in our research group. The stability of the formed rotamers in these cases is affected by the intramolecular charge transfer through the pyrimidine ring and azo group (via push-pull mechanism). This effect is higher in the case of *para*-EWG substituted phenyl ring. The electron withdrawing groups (EWGs) increased the C5–N bond order and the energy of interconversion between two observed rotamers (the rotational barrier). The rotamer ratio is independent on R-substitution.[10]

The studied structures are depicted in Figure 1.5 together with their intramolecular six-membered pseudo rings. These pseudo rings are stabilized by N–H interactions. The results obtained from NMR line shapes are summarized in Table 1.1.[10]

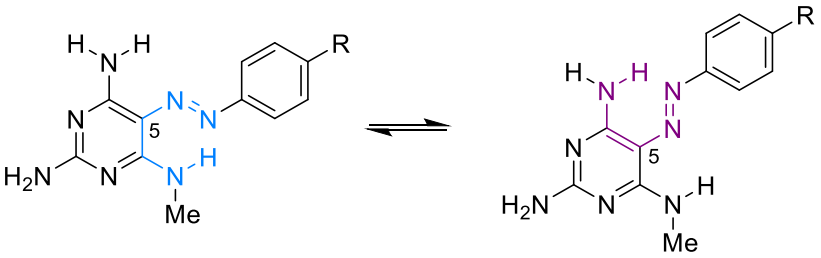


Figure 1.5 Two observed rotamers of 5-phenylazopyrimidines with different six-membered pseudo rings. Created according to reference[10].

Table 1.1 Experimental interconversion barrier between two rotamers of R-substituted 5-azopyrimidines determined on a 500 MHz spectrometer.[10]

R	Coalescence Temperature	ΔG (kcal/mol)
OMe	240 K	12.5 ± 0.3
H	260 K	13.4 ± 0.3
NO_2	320 K	16.2 ± 0.3

1.1.2 Hydrogen bonding in NMR spectroscopy

Hydrogen bonding (H-bonding) is a type of weak bonding and can vary in strength from weak to very strong. According to Grabowski there are three types of hydrogen bonds – weak (bond energy between 0.5–4 kcal/mol), moderate (4–15 kcal/mol) and strong (15–60 kcal/mol). [11] Most often, the binding energy is around 4.8 kcal/mol. [12] The formation of hydrogen bond can be explained by electrostatic model as an attraction between a hydrogen atom (partially positively charged) covalently bound to a more electronegative atom (O, S, N) and another electronegative atom bearing a lone electron pair (partially negatively charged). Another way of explaining hydrogen bond formation is by a quantum mechanical model. Three atomic orbitals of D (donor of hydrogen bond), H (hydrogen atom itself) and A (acceptor of hydrogen bond) combine into three molecular orbitals. The electrons occupy only the two molecular orbitals with the lowest energy. The energy of this system, D–H···A, is then lower than the energy of non-interacting molecules DH and A.[12]

Hydrogen bonds can be inter- or intramolecular and their formation affects compounds' lipophilicity, stability, melting and boiling point, tautomeric equilibria and at last, biological activity.

The interactions of bioactive compounds and drugs are mostly controlled by hydrogen bonds and other weak intermolecular forces. Nowadays, the binding mechanism is explained by the "induced-fit" model. This model postulates that the exposure of binding partners to the interacting (bio)molecules causes structural changes. These adaptations involve for example conformational or tautomeric changes.[13] Despite the importance of hydrogen bonds in life processes, the deep understanding of their energies, dynamics and geometric preferences is still limited.

There are two main ways of studying hydrogen bonds – X-ray crystallography and NMR spectroscopy, which is one of the most powerful tools for studying hydrogen

bonds, because all NMR observables are affected by the hydrogen bond formation. The list of NMR observable parameters for indirect studying of these bonds contains e.g. reduced hydrogen exchange rate with the solvent, isotropic proton chemical shift, primary isotope shifts by substitution of the hydrogen bonded proton by deuterium or tritium, size of the electric field gradient at the position of proton observed by the deuterium quadrupolar coupling constant, isotropic nitrogen chemical shift and the size of proton chemical shift anisotropy.[14]

Common way of hydrogen bond recognition by NMR spectroscopy is through evaluating the ^{1}H chemical shift in a non-aqueous solution, e.g. in chloroform, dimethylsulfoxide or *N,N*-dimethylformamide. The protons engaging in a hydrogen bond to electronegative atoms show slightly higher chemical shifts from those which are not hydrogen bonded. Moreover, strong hydrogen bonds lead to much higher chemical shift of interacting protons. The chemical shifts of strongly hydrogen bonded protons can be observed as high as 20 ppm.[15]

A direct evidence of hydrogen bonds can be obtained by observation of scalar couplings across the hydrogen bond (for example $^{2h}J_{N,N}$ scalar couplings in ^{15}N labelled compounds). The scalar couplings can be obtained by two dimensional ^{1}H-^{15}N-transverse optimized-relaxation spectroscopy (TROSY).[16] In the case of nucleic acids base pairs, Pervushin et al. found their values within the range of 6–7 Hz ($^{h}J_{N,N}$) and 2–3,6 Hz ($^{h}J_{H,N}$).[16] The biggest disadvantage of this method is its limitation to stable systems, such as Watson-Crick bound base pairs in nucleic acids.[14] The formed complex has to be stable on the NMR timescale. The values of $^{h}J_{N,N}$ coupling constants are dependent on interatomic distance. Therefore, these values are smaller for guanine-cytosine base pair than for adenine-thymine base pair, because of the longer N–N distance in guanine-cytosine pair.[16]

The assessment of the relative strength of hydrogen bonds between various components and ligands can be obtained by diffusion-ordered NMR spectroscopy (DOSY), the specific gradient NMR pulsed sequence. Using this method, Kapur et al. found out that phenol is forming stronger hydrogen bonds with deuterated dimethyl sulfoxide (d_6-DMSO) than cyclohexanol. This effect is due to the much higher acidity of phenol compared to cyclohexanol.[17]

The NMR spectra are a fingerprint of molecular geometry and structure. For a deeper understanding of the system under study, they are often supported by theoretical calculations e.g. DFT calculations.

1.2 Hydrogen bonding in life systems – DNA structure

As outlined above, hydrogen bonds play a key role in many life systems. They determine the tertiary structure of proteins, the structure and reactivity of enzymes, molecular recognition and the structure of nucleic acids.

The fascinating journey to reveal the structure of nucleic acids began in 1952, when Rosalind Franklin and Raymond Gosling obtained the X-ray diffraction image of crystalized DNA, called "Photo 51". In this picture (see Figure 1.6 below) regular arrangement of its structure is obvious. Photo 51 became a crucial data source for the determination of DNA structure.[18-19]

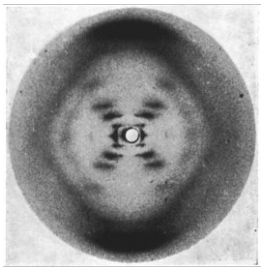


Figure 1.6 Photo 51[18]

In 1950, Erwin Chargaff published a paper about human DNA. He discovered that the amount of adenine is almost equal to thymine and the amount of guanine is almost equal to cytosine. In other words, the DNA from any cell of any organism has a 1:1 ratio of pyrimidine (thymine and cytosine) to purine bases (adenine and guanine). Chargaff also discovered that the composition of DNA varies from one species to another. His discovery attributed to the belief that DNA is of crucial importance for storing of genetic information. Nowadays, it is well known that the specific base pairing in double stranded DNA is responsible for the storing of genetic information.[20-22]

Based on the results of the mentioned researches, James D. Watson and Francis H. C. Crick concluded that DNA consists of a double helix with anti-parallel oriented strands. They described the B-form of DNA (see Figure 1.7), the most abundant form in cells.[19]

To clarify the Chargaff's rules, Watson and Crick described the canonical forms of base pairs, well known today as Watson-Crick base pairs. In this model, adenine (A)

is bound to thymine (T) via two hydrogen bonds and cytosine (C) to guanine (G) via three hydrogen bonds (Figure 1.8).[23-24]

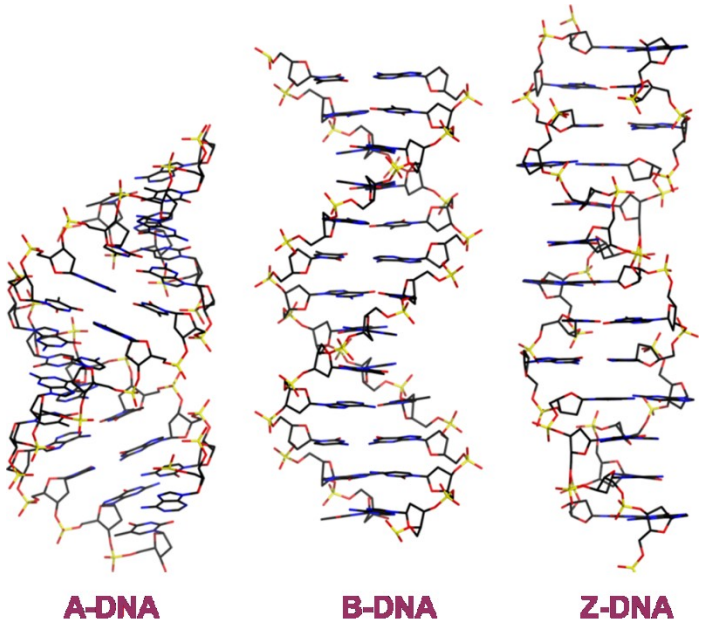


Figure 1.7 Three main conformations of DNA.[25] In living cells, the B-DNA form is the most abundant.

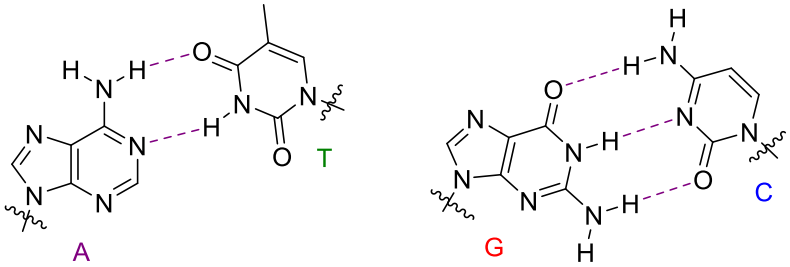


Figure 1.8 Watson-Crick base pairing. **A** is for adenine, **T** for thymine, **C** for cytosine and **G** for guanine.

Ten years later in 1963, Karst Hoogsteen reported another type of base pairing, known as the Hoogsteen type. In this model, the complementarity of bases is the same as in the Watson-Crick model, but the geometry differs (Figure 1.9). In this model, guanine and protonated cytosine are bound only via two hydrogen bonds instead of

three in the Watson-Crick model.[26-27] Hoogsteen type base pairing has been observed relatively rarely. It can be found in some specific structural motives described below.

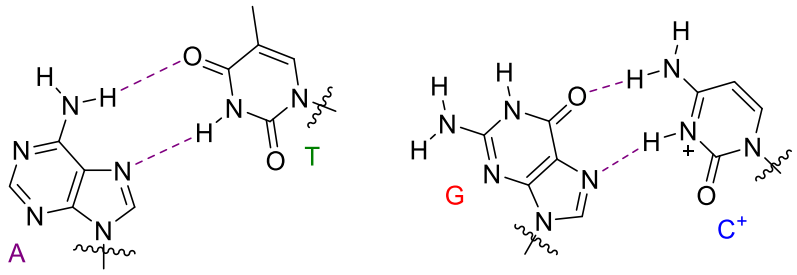


Figure 1.9 Hoogsteen base pairing

1.2.1 G-quadruplexes

In late 1980s, Dipankar Sen and Walter Gilbert found that oligonucleotides containing runs of three or four adjacent guanines formed four stranded DNA structures.[28] These specific structures are called G-quadruplexes or G-tetraplexes. Their building blocks are four guanines bound via Hoogsteen hydrogen bonding to form planar G-quartets (see Figure 1.10). The G-quartets are stacked on top of each other and give rise to a four stranded helical structure - G4-DNA (Figure 1.11).[29]

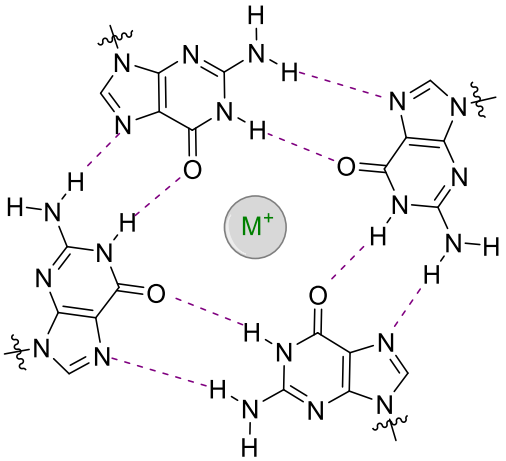


Figure 1.10 G-quartets

G-quadruplex DNA can be sub-grouped into families, e.g. inter- or intra-molecularly folded or parallel and antiparallel according to the relative strand orientation. These structures are highly polymorphic and the specific type formed depends on the number of G-quadruplexes in a strand.[30]

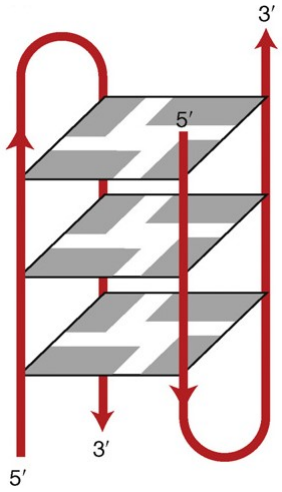


Figure 1.11 An example of a G-quadruplex[29]

Later in 1990, Sen and Gilbert found that G-quadruplexes are formed only in the presence of sodium or rubidium ions, but not in the presence of potassium ions. Potassium ions stabilize a fold-back intermediate structures and the G4-DNA is not formed.[31]

However, G-quadruplexes have not received scientific attention worldwide due to their structure. The most interesting phenomenon is their presence in genomic DNA and RNA, suggesting possible functional roles in various biological processes, e.g. regulation of expression of human oncogenes.[32] For this reason, a variety of small-molecule G-quadruplex ligands have been generated as potential anticancer agents.[33]

The functional properties of G-quadruplexes can be modulated by their multimerization.[34] It is hypothesized that multimeric DNA G-quadruplex structures bring together distant parts of chromosomes. This facilitates enhancer formation and homologous recombination.[28, 34] At our institute, GTP was found as an inhibitor for multimerization in the case of guanosine to adenosine mutation at position 11 of the central tetrad (GGGTGGGAAG**A**GTGGGA). Proton NMR experiments (Bruker Avance III, 850 MHz) proved that this mutation undergoes structural rearrangement in the presence of GTP.[34]

G-rich sequences are concentrated in telomeres. Human telomeres consist of non-coding DNA at the ends of chromosomes. These sequences include several tandem TTAGGG repeats. Every time a somatic human cell undergoes mitosis, about 50 – 200 base pairs are lost in telomeric DNA because it is not fully replicated. That means every time a cell divides, the telomeres are slightly shortened. Once telomeres shorten to a critical length, the cell division is stopped. The number of times a cell will divide before cell division stops is called the Hayflick limit.[30, 35-36]

Some cells, e.g. cancer cells or stem cells, do not have a Hayflick limit and can divide indefinitely. These types of cells have special enzymes called telomerases, which can elongate telomeres.

1.2.2 Other tetrads

Although most of the studied tetrad structures contains G-rich sequences, non-G-tetrads have also been found in DNA structures. Tetrads can be divided in two main groups; tetrads which are comprised of just one base type and mixed tetrads. In addition to G-tetrads, A-, T-, C- and U-tetrads have been observed in DNA (RNA in the case of U-tetrad) tetraplex structures. The uracil units in U-tetrad are linked by C–H···O contacts. These structures have been observed only in exceptionally stable RNA tetraplexes containing G- and U-tetrads.[37-38]

An interesting observation has been made for the G · C · G · C mixed tetrad. This structure exists in a Na^+ containing solution, but not in a K^+ containing solution. In a solution containing K^+ ions, the interbase H-bonds are disrupted because of the increasing distance between two GC pairs.[39] Planar, non-planar and twisted G · C · G · C mixed tetrads have been observed.[37]

By dimerization of two G · T pairs, a non-planar T · G · G · T tetrad is formed. This motif has been found in interacting loop regions.[40]

1.2.3 Triads

In 1978, the first base triad was observed at atomic resolution in yeast phenylalanine tRNA.[41] Triads occur in nucleic acid triplexes. These are formed by

interaction between a third strand and the major groove of a double helix. The third strand binds in a parallel or antiparallel orientation to one of the duplex strands.[37, 42]

There are two canonical triads with a parallel orientation of the third strand to one of the duplex strands. In 1994, the T · A · T triad was observed using NMR spectroscopy. Its structure is shown in Figure 1.12. The second C^+ · G · C canonical triad (see Figure 1.12) was observed via NMR spectroscopy in the same year. To form this triad, N3 protonation of one cytosine is required. That is why the formation of mixed C^+ · G · C triad is pH dependent.[43]

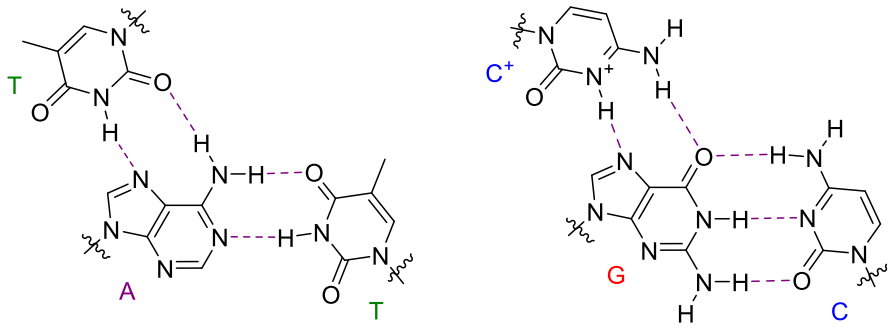


Figure 1.12 Parallel canonical triads

In 1993, two antiparallel oriented triads were observed using NMR spectroscopy. Namely, G · G · C and T · A · T triads.[44] The T · A · T triad with antiparallel oriented third strand structurally differs from the triad shown in Figure 1.12. The third canonical antiparallel triad is A · A · T.[37]

1.3 Methylation of nucleobases

In mammals, DNA methylation is associated with number of essential processes, e.g. genomic imprinting, aging or carcinogenesis. Methylation is performed as a post-replicative modification by specific enzymes – DNA methylases.[45]

Mainly two out of the four canonical nucleobases have been observed to undergo methylation – namely cytosine and adenine. Cytosine can be methylated in three ways giving N^3-methylcytosine (m3C), N^4-methylcytosine (m4C) or 5-methylcytosine (m5C). Adenine is methylated to N^6-methyladenine (m6A) or N^1-methyladenine (m1A).[46-47] All structures are summarized in Figure 1.13.

N^7-methylguanine can be generated by drugs (azinomycin, leinamycin) or alkylating mutagens (aflatoxin B_1). N^7-alkyl-dG has a half-life of several hours to days. These structures can undergo depurination. Generated abasic sites can lead to mutation due to G to T transversion.[48] In viruses, such as in the SARS coronavirus, N^7-methylation of guanine helps to RNA stability.[49] Guanine can be also methylated at oxygen to generate O^6-methylguanine. The methylation is mostly performed by reactive *N*-nitroso compounds or other alkylating agents such as SAM. This DNA damage leads to mutagenesis and cancer.[50] Described structures are shown in Figure 1.14.

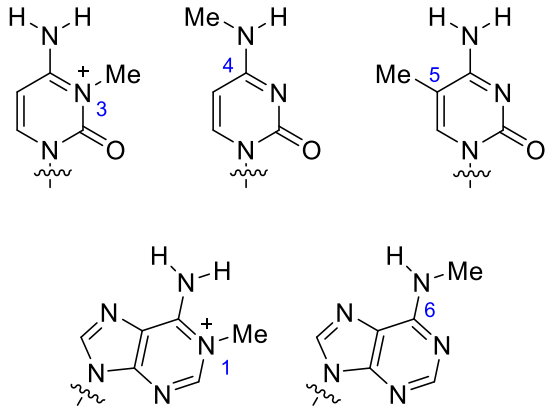


Figure 1.13 Methylated nucleobases

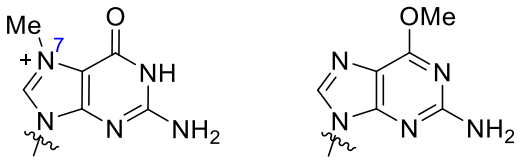


Figure 1.14 Two possibilities of guanine methylation

Thymine can be methylated to give O^2-methylthymine and O^4-methylthymine (Figure 1.15). These derivatives are proposed as pro-mutagenic bases.[51-52] According to Venkateswarlu, both *O*-methylated thymine derivatives can form a mutagenic base pair with guanine.[52]

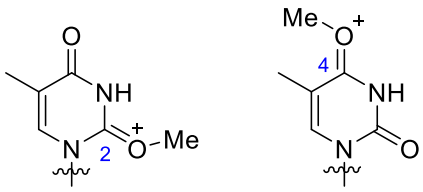


Figure 1.15 Methylated thymine structures in their protonated forms

Methylation is a way to change DNA structure and activity without changing its sequence. Nevertheless, methylation of cytosine to 5-methylcytosine can provide sequence changes. 5-methylcytosine can undergo spontaneous hydrolytic deamination to give thymine.[53] This leads to a change from GC base pair to AT pair and mutation. N^3-methylcytosine does not establish any interaction with guanine and it is a cytotoxic agent.[54] In bacterial DNA, the role of methylated cytosine (both types) is to protect against restriction enzymes.

N^1-methyladenine causes cytotoxic DNA lesions which can accumulate in the presence of S_N2 type methylation agents. It forms Hoogsteen type base pairs with complementary thymine in DNA structure.[54-55]

The role of N^6-methyladenine is multifunctional and includes repairing, replication, transposition and expression of bacterial DNA.[46] For example, 2 % of adenine is methylated in *Escherichia coli* $15T^-$.[45]

Based on previous research of the key role of DNA adenine methyltransferases in bacterial viability, the investigation of biological function of adenine methylation in eukaryotic cells is needed. Further research of methylated adenine status in human cells may help to find out new antiviral and other drugs.[46] Nowadays, some nucleoside derivatives are used in human medicine as active components in treatment of AIDS and hepatitis type B.[56]

In earlier studies, DNA m6A modifications in eukaryotic cells were considered absent.[46] Nowadays, deep sequencing and analytical methods are developed and m6A has been detected in the genome of *Chlamydomonas,*[57] mice[58] and pigs.[59] In 2018, Xiao et al. found out that m6A is extensively present in human genome.[60] Using SMRT sequencing and LC-MS/MS assay, they observed that 0.051% of the total adenines is methylated in human genome. Moreover, the motif (G/C)AGG(C/T) is significantly associated with adenine methylation.[60]

In cancer cells, m6A is the most regulated modification of DNA and Liang et al. found a dramatic decrease of this modification relative to its normal levels.[61] The methylation is regulated by methyltransferase N6AMT1 and demethylase ALKBH1. The decreased level of N6AMT1 and the increased level of ALKBH1 lead to downregulation of m6A modification in cells and promotes tumorigenesis.[60]

Apart from tumorigenesis, fat mass and obesity-associated protein (FTO) dependent N^6-methyladenine demethylation regulates mRNA splicing. This is required for adipogenesis.[62-63]

As 5-phenylazopyrimidines mentioned above, methylated adenine has its own rotational barrier between two observed rotamers (Figure 1.16). The rotamer equilibrium is affected by steric and electronic effects.

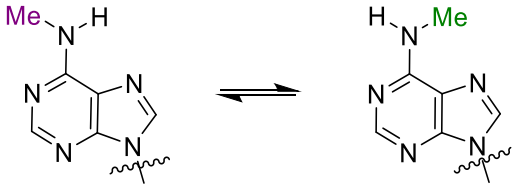


Figure 1.16 Observed rotamers of methylated adenine

Good knowledge of base pairing thermodynamic parameters is required for accurate predictions of nucleic acid structures and stability. Most experimental data of nucleic acid Gibbs energy have been obtained from NMR measurements,[64] thermal denaturation experiments of double-stranded DNA[65] and gas-phase mass spectrometry.[66-67] Other experimental methods are differential scanning calorimetry and isothermal titration calorimetry.[67-68] Apart from NMR spectroscopy, all techniques described above are experimentally challenging or destructive to the sample.

In our laboratory, Štoček et al. developed simple and straightforward method for determination of free energies of binding of natural nucleobases and their derivatives. The methodology is based on NMR monitoring of the conformational equilibrium of 2-(methylamino)pyrimidines. This compound exists in a mixture of two rotamers differing in the hydrogen-bonding pattern (see Figure 1.17). Because of the higher order of bond between carbon atom and nitrogen of methylamino group, two rotamers can be observed as two sets of signals at low temperature. Our group demonstrated that

interactions with hydrogen-bonding partners change the rotamer ratio, which can be used for determination of binding Gibbs energy.[67]

Gibbs energy was calculated from well-known equation (1.9) and the association constant from equation (1.10).

$$\Delta G = -R \cdot T \cdot lnK_2 \tag{1.9}$$

where R is universal gas constant, T is thermodynamic temperature and K_2 is association constant for binding process.

$$K_2 = \frac{K_1 \cdot [B_{TOT}] - [A_{TOT}]}{[A_{TOT}] \cdot \left([L_{TOT}] - [B_{TOT}] + \frac{[A_{TOT}]}{K_1}\right)} \tag{1.10}$$

where K_1 is the equilibrium constant describing the rotamer ratio in the absence of intermolecular interactions, $[A_{TOT}]$, $[B_{TOT}]$ and $[L_{TOT}]$ are total concentrations of rotamer A, rotamer B and interacting partner.

Applying equations (1.9) and (1.10) to variable temperature measurements of 2-(methylamino)pyrimidine derivatives with addition of 1–7 equivalents of binding partners enables determination of hydrogen-bonded complex formation free energies and corresponding enthalpic and entropic contributions.[67]

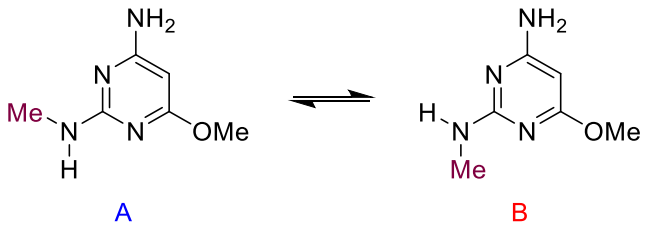


Figure 1.17 Two observed rotamers of 2-(methylamino)pyrimidine derivatives. Created according to reference[67].

This thesis is linked to research in our group described above and deals with rotational equilibria and determination of hydrogen-bound complex formation free energies of adenine derivatives.

2. Aims of work

The main aim of this thesis is the development of new methodology for determination of free energy changes associated with hydrogen-bonded complex formation of methylated nucleobase analogues based on monitoring by NMR spectroscopy as a simple and nondestructive analytical method. Another goal is description of the effect that methylation of nucleobases has on their bonding abilities and rotational equilibria.

3. Results and discussion

3.1 Rotamer equilibrium

Substituted purine, such as compound **1** in Figure 3.1, can exist as a mixture of two rotamers differing in the orientation of the substituent as the bond between the methylamino group and the aromatic ring is of order higher than one. These two rotamers of the methylamino group can be observed by low-temperature NMR as two sets of signals. At higher temperatures, these two sets of signals merge to one (Figure 3.4). The rotamer-ratio is given by equilibrium constant, K_1.

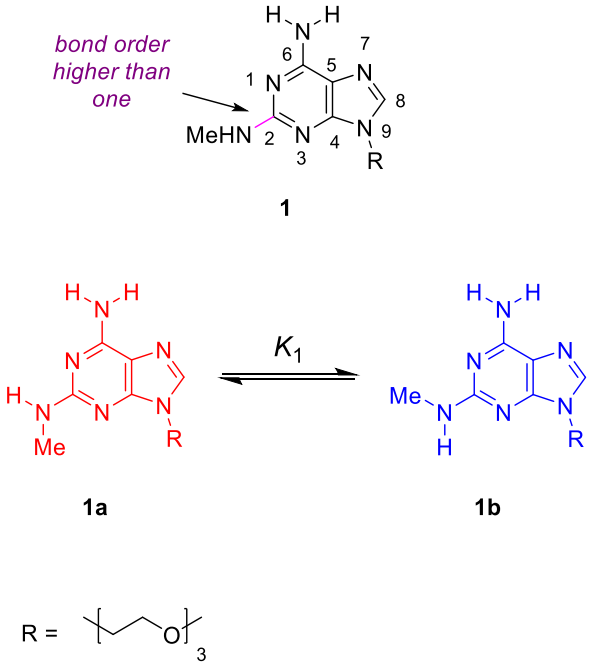


Figure 3.1 Two observed rotamers of compound **1**

The equilibrium constant K_1 is given by equation 3.1. The rotational-equilibrium is dependent on solvent[67] and temperature.

$$K_1 = \frac{[B]}{[A]} \tag{3.1}$$

where $[A]$ and $[B]$ are relative concentrations of rotamer A and B, respectively.

The relative concentrations $[A]$ and $[B]$ were obtained experimentally from the integral intensities in ^{1}H NMR spectra. The methylamino group orientation for each rotamer was established by long-range heteronuclear H–C coupling observed in the low-temperature ^{1}H,^{13}C-HMBC as four-bond correlations. Four-bond correlations are usually observed when the bonds between the coupled nuclei have the shape of "W". In the case of rotamer **1a**, a long-range cross-peak of methylamino proton and quaternary carbon C4 is observed. In rotamer **1b**, the long-range interaction of methylamino proton and carbon C6 is observed.

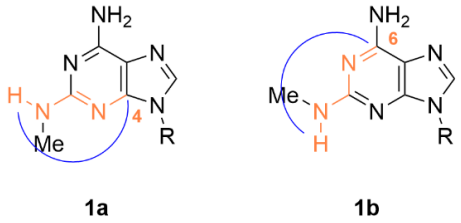


Figure 3.2 C,H long-range correlation observed in C,H-HMBC spectra at low temperature

The relative concentration of minor rotamer **1b** slightly increases with decreasing temperature – the free-energy difference between rotamer **1a** and rotamer **1b** grows smaller with decreasing temperature (Figure 3.4). The free-energy changes were obtained from equation 3.2, already introduced in Chapter 1.

$$\Delta G = -R \cdot T \cdot lnK \qquad (3.2)$$

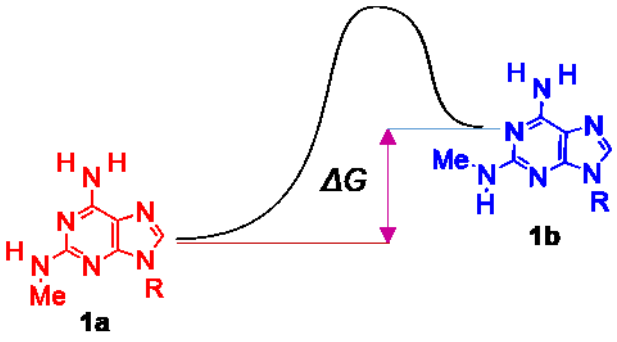


Figure 3.3 Schematic energy diagram of conversion between two rotamers

The relative rotamer concentrations were obtained experimentally from integral intensities of the imidazole ring proton (H8, see Figure 3.4). At low temperature, two well separated signals of H8 can be observed corresponding to the two rotamers.

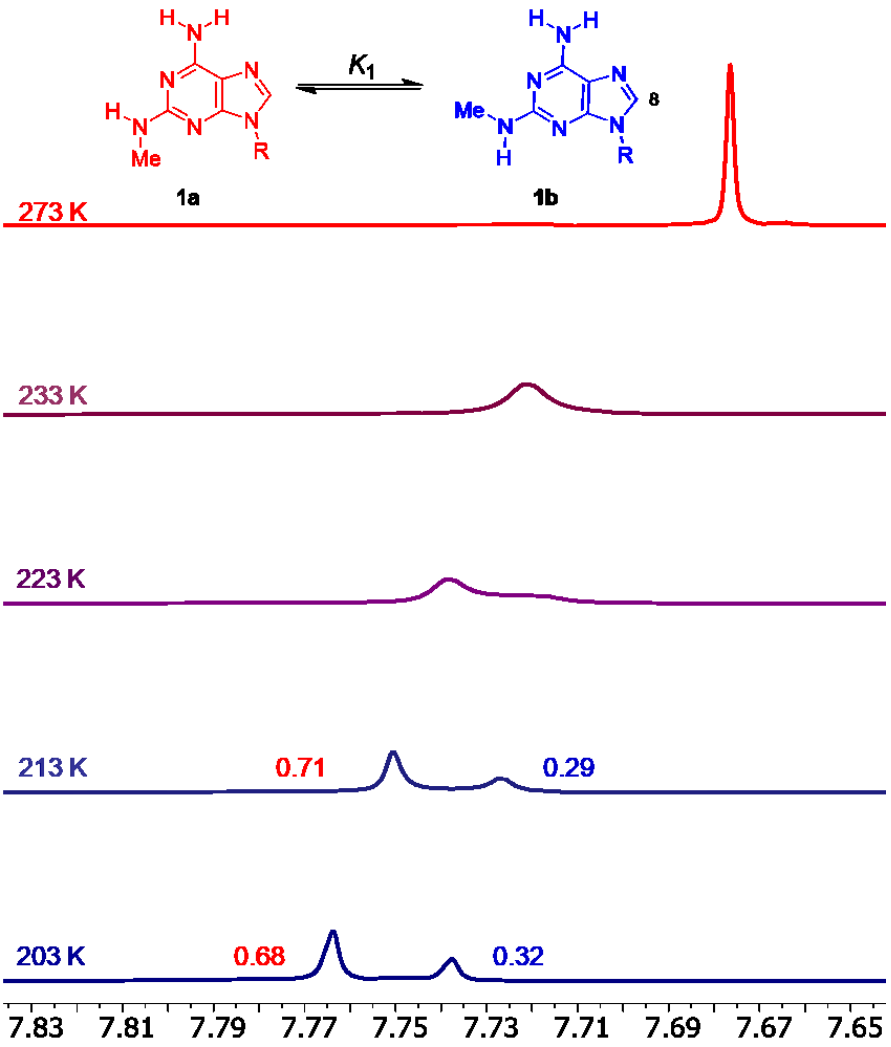


Figure 3.4 Part of variable-temperature ^{1}H NMR (500 MHz) spectra of imidazole proton (H8) of model compound **1** (concentration of 10 mmol/dm^{3}) in a DMF:DCM mixture (volumetric ratio 1:1) at variable temperatures

Only three values of ΔG in a DMF:DCM mixture were obtained because of the relatively low coalescence temperature (223 K at 500 MHz), significant line broadening and signal overlapping at lower temperatures that complicate the integration. For this reason, a similar set of experiments was performed also in d_4-methanol. Notice that the

free-energy difference between rotamer **1a** and rotamer **1b** is slightly smaller in methanol (Figure 3.5).

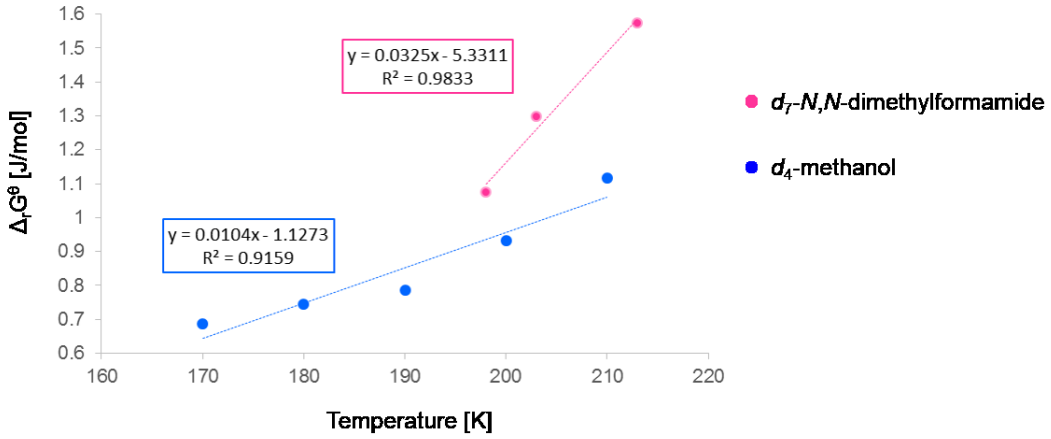


Figure 3.5 The free-energy difference between two rotamers of compound **1** in d_4-methanol and d_7-DMF

Other (methylamino)adenine derivatives were also investigated (see Figure 3.6). Similar experiments to those mentioned above were performed only in a DMF:DCM mixture in respect to further focus of the diploma thesis.

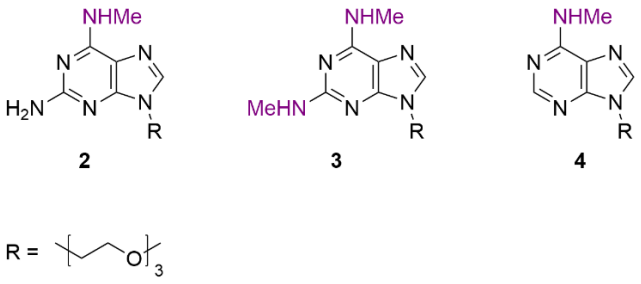


Figure 3.6 The structure of investigated compounds **2**, **3** and **4**

The relative rotamer concentrations of compound **2** cannot be obtained experimentally from integral intensities in ^{1}H NMR spectra (as shown above for the case of model compound **1**) because of a signal overlapping.

In the case of compound **3,** the experimental energy difference between two major rotamers **3c** and **3d** can be obtained from integral intensities but it is disturbed by the rotational equilibrium in the C6–NH region (there are four possible rotamers in

total). The free-energy difference between the major rotamer **3c** and the rotamer **3d** is 1.6 kJ/mol, which is slightly higher than the free-energy changes associated with the interconversion of rotamer **1a** to **1b** in the same solvent mixture.

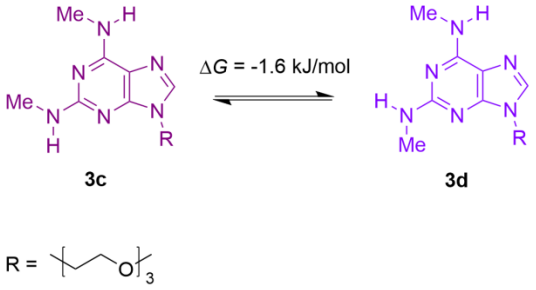


Figure 3.7 The structures of the rotamers **3c** and **3d**

Next, the N^6-methyladenine derivative **4** was investigated. The free-energy changes associated with conversion of rotamer **4a** to rotamer **4b** were 5.18±0.12 kJ/mol at 213 K and 5.01±0.06 kJ/mol at 193 K. This relatively high energy difference can be caused by large steric hindrance between the methyl group and the imidazole ring in rotamer **4b**.

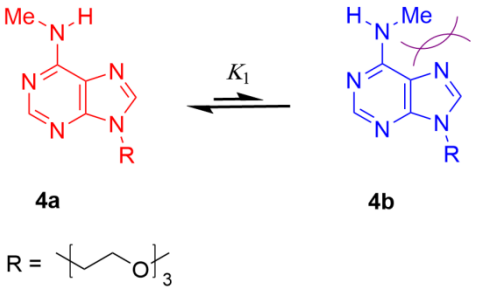


Figure 3.8 The rotamer equilibrium of N^6-methyladenine derivative **4**. Rotamer **4b** is energetically unfavorable due to steric hindrance of the methyl group and the imidazole ring.

3.2 Complex formation abilities

Model compound **1** exists as a mixture of two rotamers while each of them has different hydrogen bonding pattern and can thus form different types of

hydrogen-bonded complexes. The rotamer **1a** has a donor-acceptor-donor (DAD) hydrogen-bonding pattern and can form Watson-Crick type complex via three hydrogen bonds with thymine derivative **5** (acceptor-donor-acceptor hydrogen-bonding pattern). This intermolecular interaction is favorable as there is no steric hindrance and three hydrogen bonds are formed instead of two in the canonical adenine-thymine Watson-Crick type base pair.

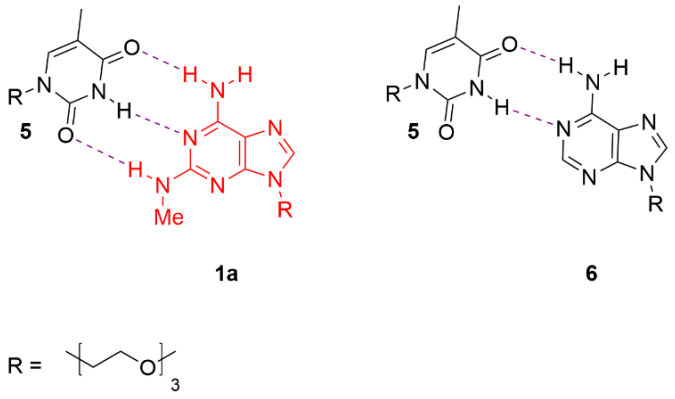


Figure 3.9 The comparison of compound **1** and natural adenine **6** complexes

On the other hand, rotamer **1b** cannot form the Watson-Crick type complex due to steric hindrance of the methyl group. The methyl group completely blocks one side of the molecule and the bonding partner cannot approach from this side.

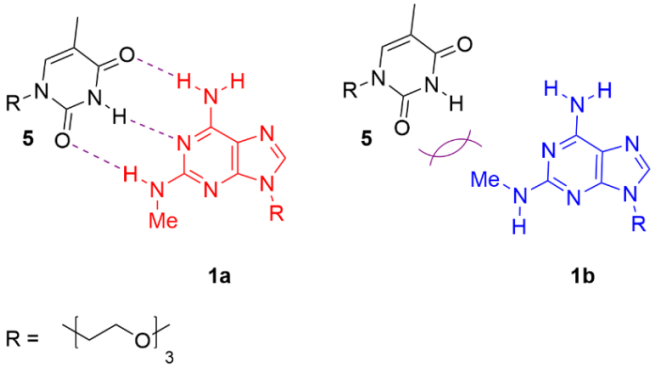


Figure 3.10 The structure of presumed Watson-Crick complex of compound **1**. The Watson-Crick type complex of structure **1b** cannot be formed due to steric hindrance of the methyl group.

Previous research in our laboratory has proven that these rotational equilibria are dependent on the binding partner concentration and an intermolecular interaction of one rotamer and the binding partner leads to an increase of relative concentration of the interacting rotamer.[67] Using equation 3.3 and the already introduced equation 3.2, the association constant K_2 and thus the free-energy changes associated with the hydrogen-bonded complex formation can be calculated from experimental data using ^{1}H NMR integral intensities and known concentrations of studied compounds.

$$K_2 = \frac{[A_{TOT}] - K_1 \cdot [B_{TOT}]}{K_1 \cdot [B_{TOT}] \cdot ([L_{TOT}] - [A_{TOT}] + K_1 \cdot [B_{TOT}])} \qquad (3.3)$$

where K_1 is the equilibrium constant describing the native rotamer ratio, $[A_{TOT}]$, $[B_{TOT}]$ and $[L_{TOT}]$ are total concentrations of rotamer A, rotamer B and ligand.

The addition of a suitable hydrogen-bonding partner shifts the equilibrium to the right – in the direction of complex formation according to Le Chatelier's principle. According to this principle, the increase of the concentration of thymine derivative **5** leads to an increase of rotamer **1a** concentration because of intermolecular hydrogen-bonded complex formation.

The free-energy changes associated with hydrogen-bonded complex formation can be calculated using equations 3.2 and 3.3. The total concentrations of adenine or thymine derivatives were calculated from their weighted masses, the rotamer-ratio was obtained experimentally from integral intensities of each set of signals in the ^{1}H NMR spectra. Apart from complexes shown in Figure 3.10, reverse-Watson-Crick type complex can also be formed. The reverse base pairing differs in the relative thymine orientation, but it is supposed to be similar from the NMR spectroscopy point of view.

The free energy changes associated with hydrogen-bonded complex formation of rotamer **1a** and thymine derivative **5** are temperature dependent as shown in Figure 3.11. Because of relatively low coalescence temperature (223 K at 500 MHz) only a narrow range of temperatures was measured. For this reason, the enthalpy and entropy contributions cannot be obtained. The obtained free energy is close to -5 kJ/mol.

The complex geometry can be estimated directly from ^{1}H NMR spectra at low temperatures. The proton participating in hydrogen bonding moves towards higher chemical shift with increasing bonding partner concentration. Should the formed

complex be stable on the NMR time scale, a new set of signals will be observed. If the complex is not stable on the NMR time scale, the observed signals represent a weighted average of the complex and its participating components.

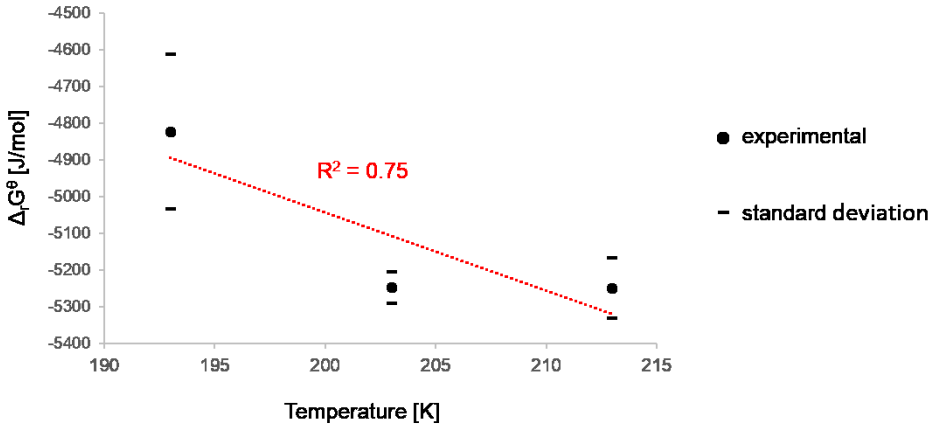


Figure 3.11 The free-energy changes associated with hydrogen-bonded complex formation of compound **1a** and **5** at variable temperatures

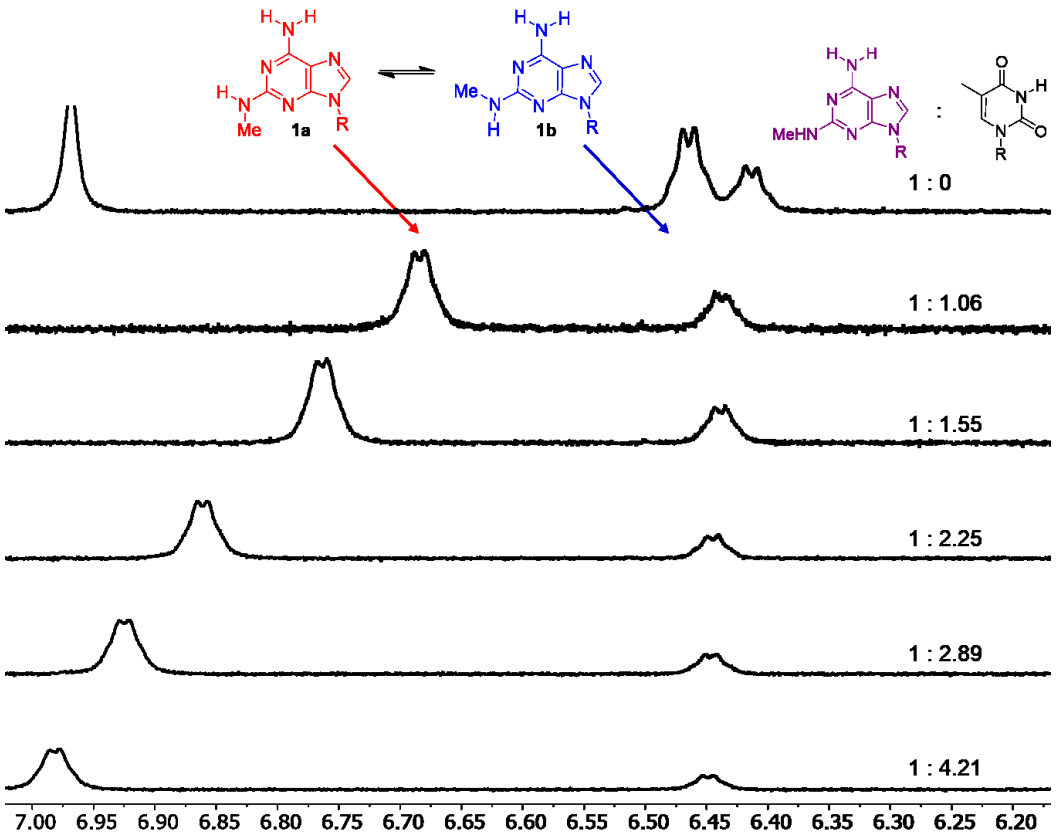


Figure 3.12 ^{1}H NMR (500 MHz) spectra of model compound **1** in DMF:DCM mixture (volumetric ratio 1:1) at 213 K in the presence of variable amounts of compound **5**.

Chemical shift of the methylamino proton of rotamer **1a** is dependent on the binding partner concentration and moves to higher chemical shifts with increasing concentration of thymine (Figure 3.12). This indicates the intermolecular interaction and formation of a Watson-Crick type complex. The methylamino proton chemical shift of rotamer **1b** is independent on the binding partner concentration.

This observation corresponds with presumed complex structure shown in Figure 3.10. The complex formed is not stable on the NMR time scale and the complex signals were not observed, even at lower temperatures. Therefore, averaged chemical shifts of the free and bound molecules were observed. Surprisingly, the imidazole ring proton signals of both rotamers show downfield effect dependent on thymine **5** concentration (see Figure 3.13). This indicates an intermolecular interaction between **1b** and **5**, but without participation of the methylamino proton.

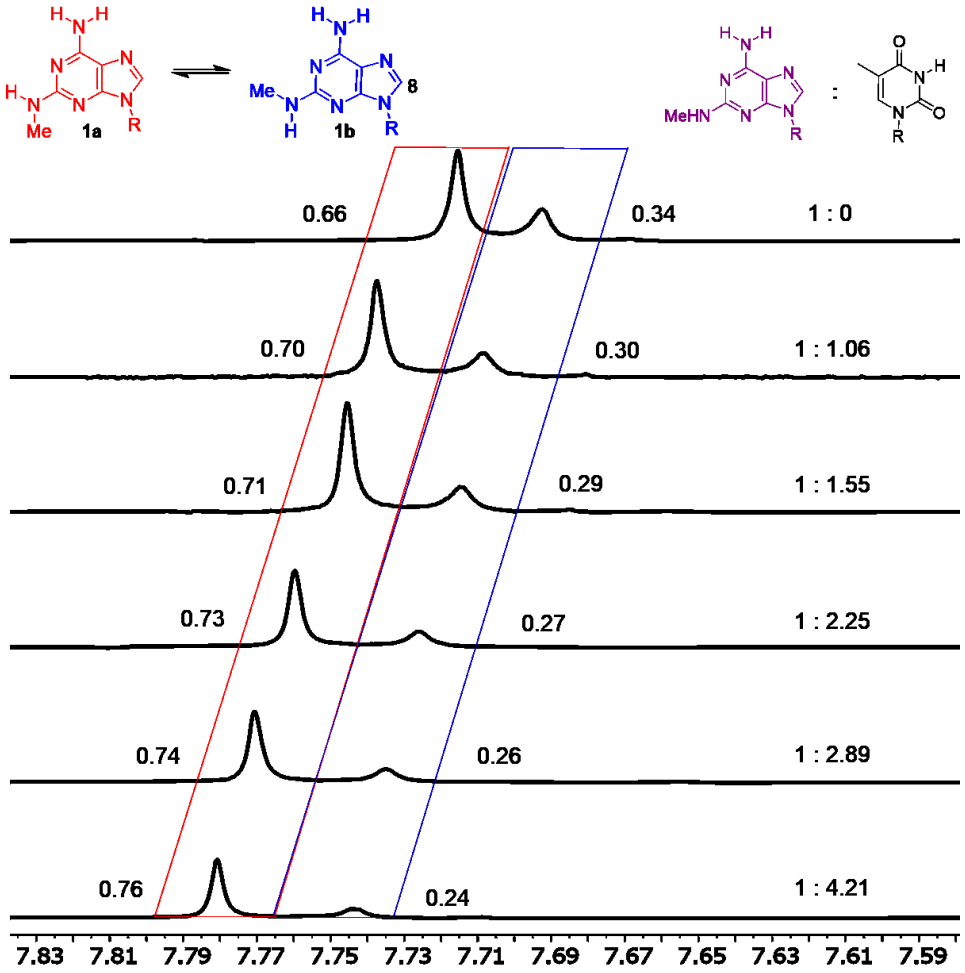


Figure 3.13 ^{1}H NMR (500 MHz) spectra of model compound **1** in DMF:DCM mixture (volumetric ratio 1:1) at 213 K. The selected area shows chemical shift changes of H8 protons of two rotamers **1a** and **1b** and the relative concentration increase of rotamer **1a**.

These observations indicate formation of another type of base pair – Hoogsteen type. The spatial vicinity of the hydrogen-bonding pattern in this type of complex and the imidazole ring proton H8 (see Figure 3.14 for more details) affects the magnetic environment of this proton and its chemical shift.

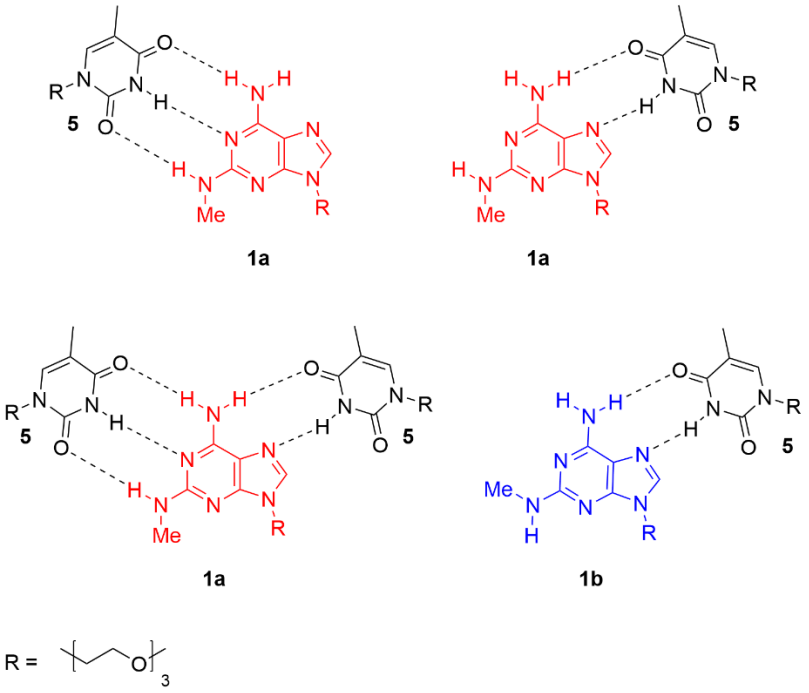


Figure 3.14 The structures of presumed complexes of compound **1** with thymine derivative **5**

Rotamer **1a** can form three types of complexes, summarized in Figure 3.14 – the Watson-Crick type complex, the Hoogsteen type complex and one bidentate complex with thymine molecules attached from both sides of the adenine derivative moiety. The equation (3.3) cannot be applied because it concludes only one equilibrium process instead of four processes shown in Figure 3.14. For this reason, we cannot obtain the equilibrium concentration of the thymine derivative $[L_{TOT}]$ and the equation (3.3) cannot be applied. The derivatives **2**, **3** and **4** can also form two types of complexes and $[L_{TOT}]$ cannot be obtained.

A new methodology based on chemical shift changes was developed to describe these multi-equilibrium processes via NMR spectroscopy. This methodology is discussed further in text.

3.3 Free-energy changes obtained from chemical shifts

The chemical shift is independent on the magnetic field strength, but it shows interesting dependency on inter- or intramolecular interactions, e.g. hydrogen bond formation and solvation. Compounds in inert solvents, such as chloroform, without intermolecular interactions show one set of signals. The chemical shift of signals is given by equation (1.7) shown in Chapter 1. The chemical exchange and the intermolecular interactions, e.g. solvation, aggregation and hydrogen-bonding formation affects the chemical shift. For example, the chemical shift observed in NMR spectra of adenine derivative is then a weighted arithmetic mean of non-bonded adenine derivative *Free* and its hydrogen-bonded complex *Bound*. It is given by equation (3.4).

$$\delta = x \cdot \delta_{Free} + (1 - x) \cdot \delta_{Bound} \qquad (3.4)$$

where x is the mole fraction of the free molecule and δ_{Free} and δ_{Bound} are chemical shifts of non-bonded adenine derivative *Free* and its hydrogen-bonded complex *Bound*.

3.3.1 Concentration dependence of ^{1}H NMR spectra of adenine and thymine derivatives alone

We found out that the chemical shift of our adenine derivatives is slightly concentration dependent and we presume some intermolecular interactions of adenine molecules with each other. For example, the derivative **1a** has DAD and DA hydrogen-bonding patterns from the Watson-Crick and Hoogsteen side respectively, and we can easily imagine hydrogen-bonding intermolecular interactions depicted in Figure 3.15. The formation of hydrogen-bonded complexes results in a higher chemical shift of the proton participating in the hydrogen-bond. Although concentration dependent chemical shift changes of compound **1a** are relatively small (just around 0.1 ppm), they are a significant sign of hydrogen-bonded complex formation.

The effect of intermolecular hydrogen-bonded complex formation of the adenine derivative itself was eliminated by calibration curves (see Figure 3.16). These calibrations describe the concentration dependence of chemical shifts of the nucleobase derivative in the DMF:DCM mixture. The chemical shift values can be determined for any concentration from linear trend line equations. I made these calibrations for

temperatures 213, 203 and 193 K. At these temperatures, the signals of each rotamer can be observed as separated set. The chemical shift changes between the lowest and highest concentration are around 0.09 ppm for the methylamino proton of adenine derivative **1a**.

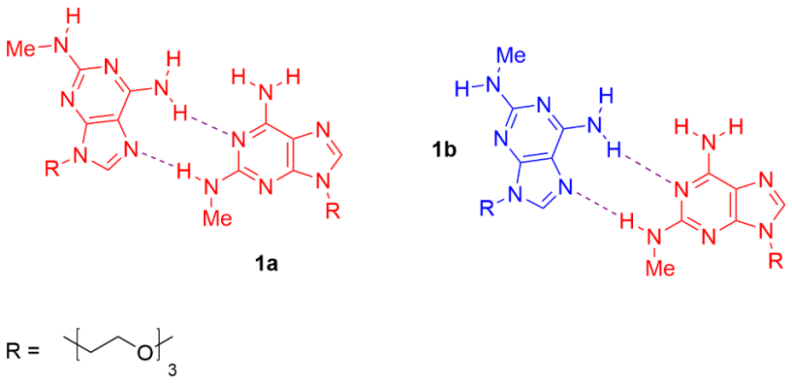


Figure 3.15 The structures of supposed self-interacting complexes of compound **1**

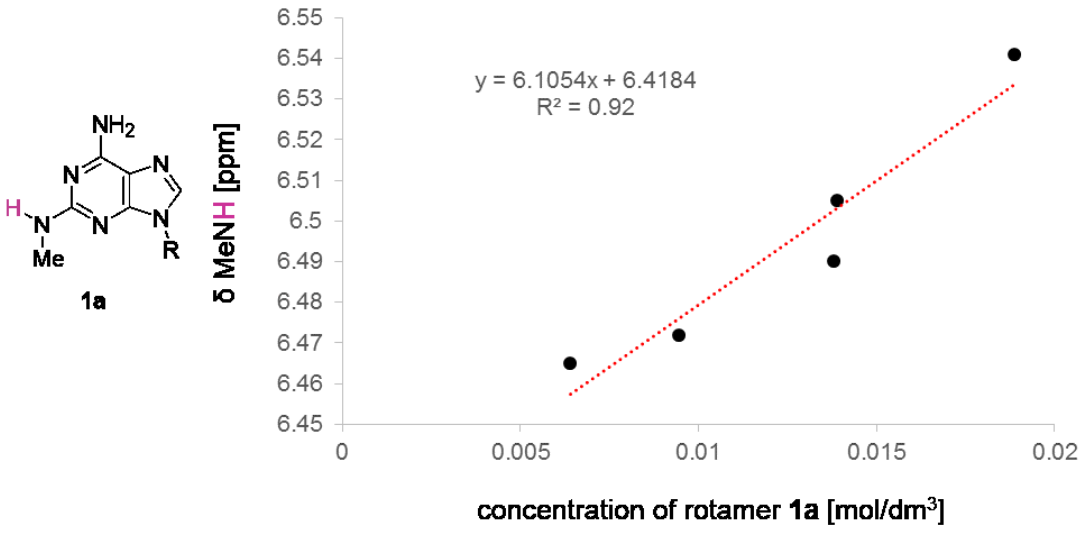


Figure 3.16 Concentration dependence of the methylamino proton chemical shift of rotamer **1a** on its concentration at 213 K.

The same set of experiments was done with thymine derivative **5**. For data evaluating, the NH proton chemical shift was used. The chemical shift changes are around 0.012 ppm in dependency on concentrations of thymine derivative **5** (Figure 3.17).

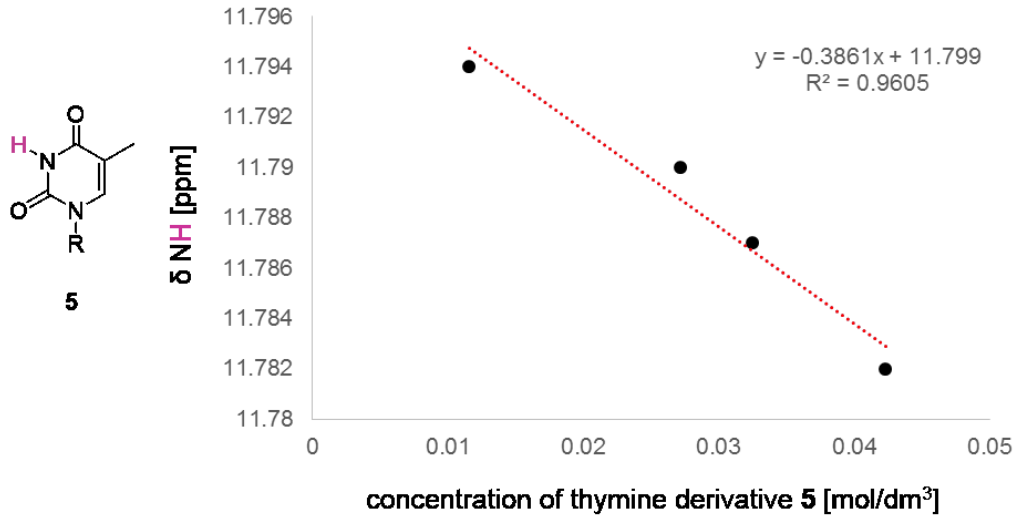


Figure 3.17 Concentration dependence of the NH chemical shift of derivative **5** obtained at 213 K

3.3.2 Dependence of ^{1}H NMR spectra of adenine derivatives on bonding partner concentration

The chemical shift changes dependent on nucleobase concentration are relatively small. On the other hand, the chemical shift changes of adenine derivatives dependent on concentrations of thymine derivative **5** are much higher. The spectra of the nucleobase mixtures with bonding partner stoichiometric ratios of 1:1, 1:1.5, 1:2, 1:2.5 and 1:3 were evaluated. Using simple equation (3.5), the chemical shift difference Δδ, induced by the presence of the bonding partner, was obtained. This chemical shift difference describes the contribution of the bonding partner to chemical shift changes.

Equation (3.5) serves for evaluation of chemical shift changes of the thymine derivative **5** induced by the presence (and excess) of an adenine derivative and vice versa.

$$\Delta\delta = \delta_{\text{experimental}} - \delta_{\text{calibration}} \qquad (3.5)$$

where $\delta_{\text{experimental}}$ is the experimental chemical shift of thymine derivative **5** of concentration x in the mixture with adenine derivative and $\delta_{\text{calibration}}$ is the calculated

chemical shift of thymine derivative **5** of concentration *x* obtained from a regression equation.

Equation (3.5) enables a qualitative comparison of the binding abilities of the adenine derivatives with the thymine derivative **5** (see Figure 3.18). This qualitative comparison was obtained from chemical shift changes Δδ of the imido proton of the thymine derivative **5** calculated as the difference between experimental chemical shift of the thymine imido proton and its chemical shift obtained by linear extrapolation to zero concentration of the adenine derivative. This method was used because of slightly different water content in the samples, which affects the NH chemical shift. The difference Δδ corresponds to the ability of thymine derivative **5** to form hydrogen-bonded complexes with the adenine derivative.

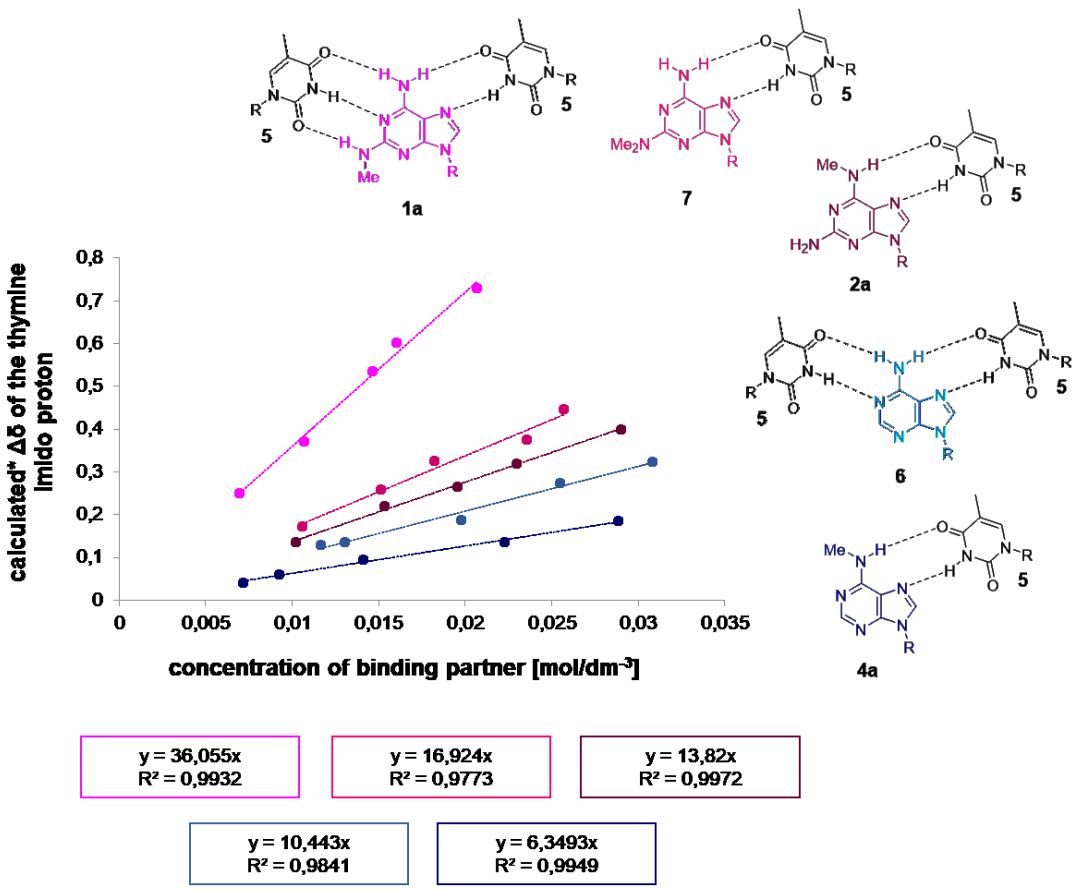


Figure 3.18 The dependence of induced chemical shift changes Δδ of the thymine derivative **5** NH proton on bonding partner concentration. * Δδ is the difference between experimental chemical shift of the thymine imido proton and its chemical shift obtained by linear extrapolation to zero concentration of the adenine derivative

The model rotamer **1a** has the steepest slope in Figure 3.18, which indicates the highest propensity of this compound to the formation of hydrogen-bonded complexes with thymine derivative **5**. Unfortunately, we cannot distinguish between the contribution of Watson-Crick and Hoogsteen types of complexes on the chemical shift changes of thymine derivative **5**.

^{15}N labelled derivatives of compounds **1** and **2**, derivatives **1-15N** and **2-15N** respectively, were synthetized to help us proving complex geometries. Their structures are depicted below. We expected that the ^{15}N chemical shift changes would distinguish between formations of Watson-Crick and Hoogsteen base pairs better than proton chemical shifts. The labelled methylamino group of each rotamer would move to higher chemical shifts in dependency to concentrations of thymine derivative **5**. Unfortunately, there were no significant chemical shifts changes dependent on binding partner **5** concentrations in the ^{15}N NMR spectra.

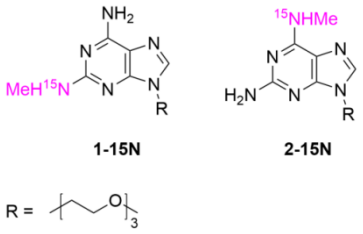


Figure 3.19 The ^{15}N-labelled compounds

For this reason, two sterically specific derivatives **7** and **8** were synthetized to help us distinguish between the contribution of Watson-Crick and Hoogsteen types of complexes on the proton chemical shift changes.

Adenine derivative **7** can only form Hoogsteen type complexes because of the steric hindrance of the dimethylamino group, which completely blocks the Watson-Crick side of the molecule (Figure 3.20). Derivative **8** is a reference compound (a negative control), because the steric hindrance blocks both sides of the molecule and hydrogen-bonded intermolecular complexes cannot be formed (Figure 3.20). This assumption was confirmed experimentally. Figure 3.21 shows ^{1}H NMR spectra of compound **8** with variable concentrations of thymine derivative **5**. There are no chemical shift changes dependent on thymine concentration.

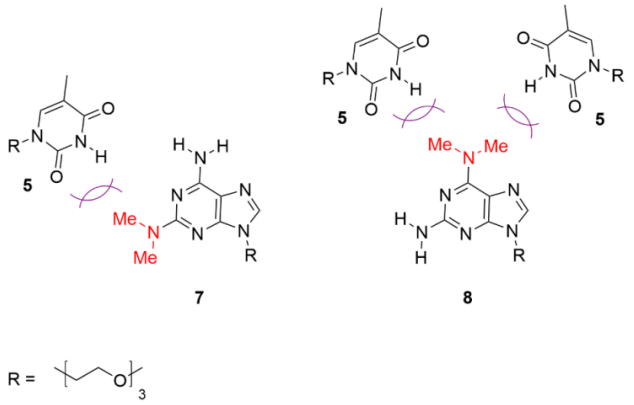


Figure 3.20 The structures of derivatives **7** and **8** and the steric clash with interacting thymine derivative **5**

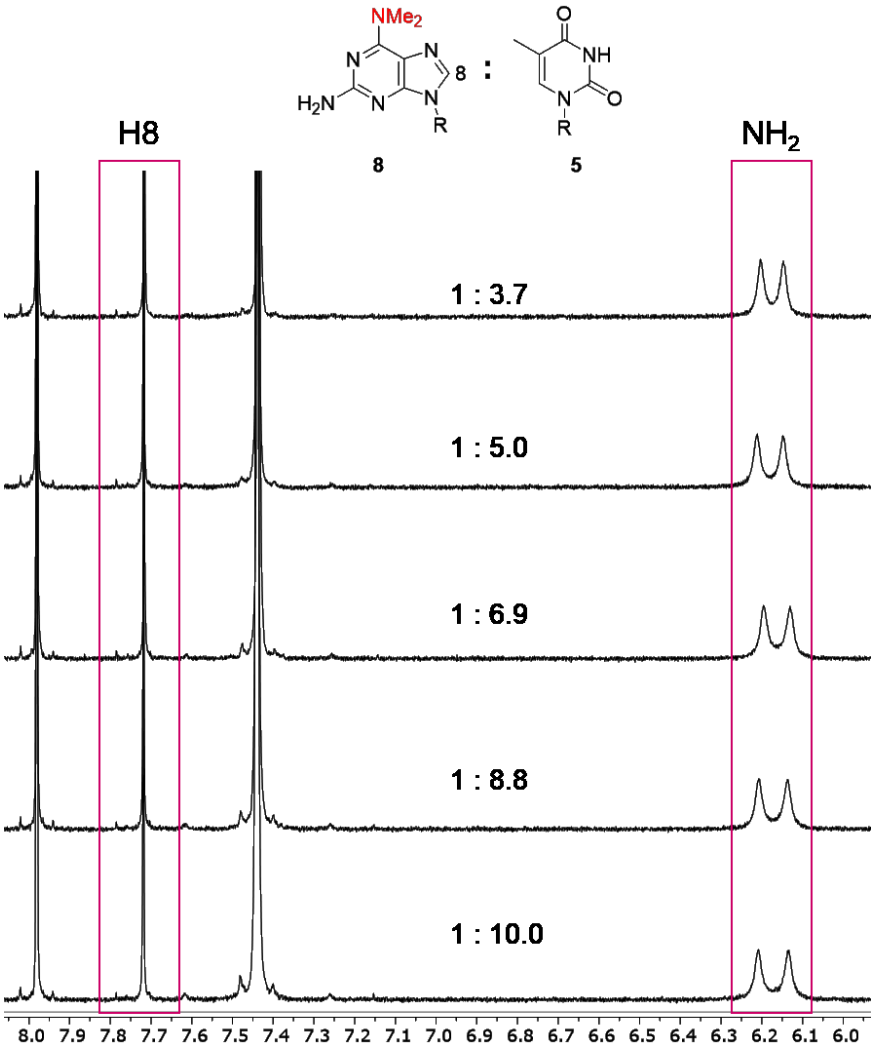


Figure 3.21 ^{1}H NMR (500 MHz) spectra of compound **8** in a DMF:DCM mixture (volumetric ratio 1:1) at 213 K with increasing concentration of thymine derivative **5**

Quantitative determination of the free energy changes associated with formation of hydrogen-bonded complexes was obtained by fitting of the chemical shift changes induced by the presence of an interacting partner to a simple complexation model described below.

$$[A] + [T] \rightleftharpoons [AT]$$

where [A] and [T] are equilibrium concentrations of adenine and thymine derivatives respectively and [AT] is equilibrium concentration of the bimolecular complex.

The resulting chemical shift depends on the relative concentration of individual species (according to equation 3.4). These data were fitted by my colleague Mgr. Ondřej Socha at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences. The obtained data are summarized in the following Table (3.1). The fitted signals are highlighted and the remarks *pro*-WC and *pro*-H correspond with the orientation of the amino proton in the molecule – e.g. *pro*-WC amino proton is involved in the hydrogen bond in the formed Watson-Crick type complex with thymine derivative **5**.

Table 3.1 The summarized Gibbs energies obtained by fitting data in temperature range of 193 – 213 K. The R2 and R6 represent substituents on carbon 2 and 6.

Adenine derivative		Cpd.	Rotamer	Fitted signals	Base pair type	$\Delta_r G_{203}$ [kJ/mol]
R2	R6					
$NHCH_3$	NH_2	**1**	**1a**	N**H**CH_3	WC-3	– 6.54
NH_2	$NHCH_3$	**2**	**2a**	N**H**CH_3 + **H**8	H-2	– 2.01
H	$NHCH_3$	**4**	**4a**	N**H**CH_3 + **H**8	H-2	– 3.61
			4b	**H**8 + **H**2	WC-2	– 7.46
H	NH_2	**6**		N**H** – *pro*-WC + **H**2	WC-2	– 6.64
				N**H** – *pro*-H + **H**8	H-2	– 6.54
$N(CH_3)_2$	NH_2	**7**		N**H**$_2$ + **H**8 + thymine imido proton	H-2	– 5.17

The atom numbering is shown in Figure 3.22 below. For more details see Appendix C – List of structures of all compounds and their possible rotamers.

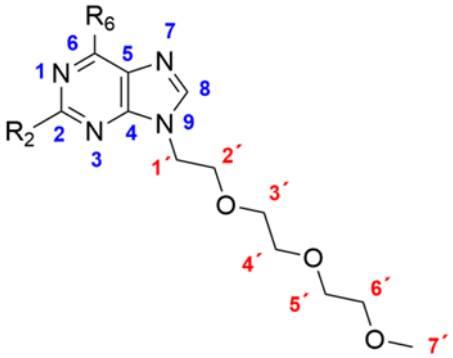


Figure 3.22 The atom numbering

Each base pair type is represented in Table 3.1 as unique code consisting from WC or H for Watson-Crick and Hoogsteen type respectively and number of hydrogen bonds forming the complex.

The enthalpy and entropy contributions cannot be obtained because of a narrow range of temperatures.

3.4 Energy changes obtained from DFT calculations

Computational chemistry is a powerful tool for studying of chemical problems. It deals, for example, with modelling short-lived intermediates and transition states, which cannot be observed experimentally. The computational chemistry in general supports experimental data and can help to elucidate unclear or ambiguous experimental results. In addition, NMR, IR and Raman spectra can be also predicted using computational methods. These predictions can support the assignment of signals.

I used computational chemistry for prediction of the intermolecular complex geometries, complexation energies and the NMR parameters of these complexes.

At first, the geometries of all studied compounds were optimized using the density functional theory method, utilizing the B3LYP functional (details in Chapter 4). We also performed geometry optimizations for one more adenine derivative (**9**), supporting our research. Then the geometries of all possible hydrogen-bonded complexes were optimized. The process for model compound **1** is simplified in Figure

3.23. Notice that the triethylene glycol chain was replaced by a methyl group to simplify the calculations.

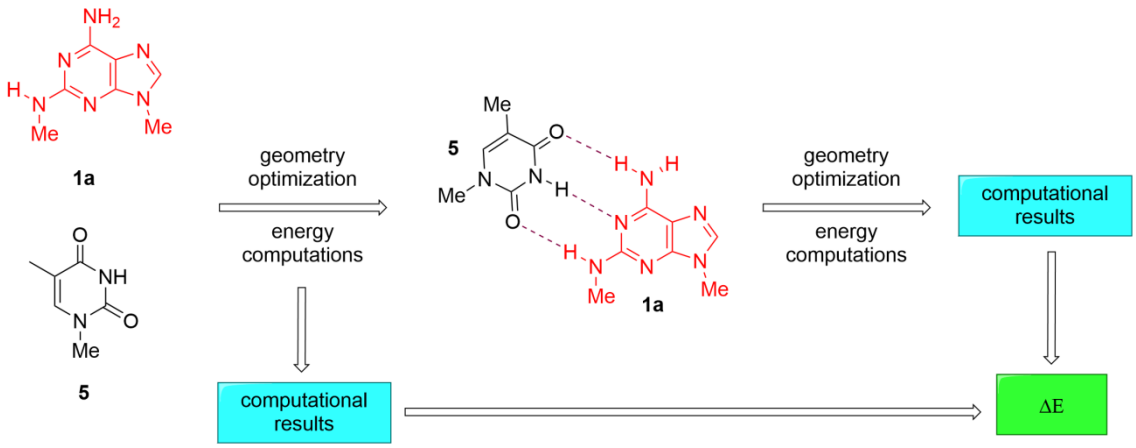


Figure 3.23 Scheme of the computational process

The complexation energy difference, ΔE_c, is given by equation (3.6). All obtained complexation energies are summarized in Table 3.2.

$$\Delta E_c = E_{\text{complex}} - E_{\text{adenine derivative}} - E_{\text{thymine derivative}} \qquad (3.6)$$

where E is the single point energy from Gaussian output file for complex, adenine derivative bound in the complex and thymine derivative, respectively.

The obtained complexation energies summarized in Table 3.2 are different from experimental data because computational methods did not involve explicit solvent molecules. In *N*,*N*-dimethylformamide solution, the monomers of both thymine and adenine derivatives will form hydrogen bonds with the solvent molecules that must be disrupted before the formation of the base pair complexes. Therefore, the solvation goes against the complex formation and the stabilisation of the complexes is much smaller than that predicted by DFT without explicit solvent molecules. Obtained complexation energies reflect the strength of intermolecular hydrogen-bonding interactions.

Apart from complexation energy, ΔE_c, also stabilisation energies, ΔE_{stabil}, were calculated. These energies are calculated from energy of formed complex and the energy of thymine derivative **5** and from the energy of the most stable rotamer of

adenine derivative (see equation 3.7). For this reason, stabilisation energy also includes the energy penalty needed for the formation of a less stable rotamer. Both energies, ΔE_c and ΔE_{stabil}, are presented together in Table 3.2.

$$\Delta E_{stabil} = E_{\text{complex}} - E_{\text{most stable rot. of A}} - E_{\text{thymine derivative}} \quad (3.7)$$

where E is the single point energy from Gaussian output file for complex, the most stable rotamer of the adenine derivative and thymine derivative, respectively.

Table 3.2 The complexation energies obtained from computational methods. R2 and R6 represent the substituents C2 and C6 in the adenine molecule.

Adenine derivative		Cpd.	Rotamer	Complex	ΔE_c	ΔE_{stabil}
R2	R6			type	[kJ/mol]	[kJ/mol]
$NHCH_3$	NH_2	**1**	**1a**	WC-3	– 59.59	– 59.59
			1a	H-2	– 49.32	– 49.32
			1b	H-2	– 49.61	– 47.78
NH_2	$NHCH_3$	**2**	**2a**	H-2	– 50.45	– 50.45
			2b	WC-3	– 58.18	– 51.38
$NHCH_3$	$NHCH_3$	**3**	**3a**	WC-3	– 59.12	– 51,57
			3c	H-2	– 50.57	– 49.03
			3d	H-2	– 50.41	– 50.41
H	$NHCH_3$	**4**	**4a**	H-2	– 49.92	– 49.92
			4b	WC-2	– 48.88	– 42.74
H	NH_2	**6**		WC-2	– 48.04	– 48.04
				H-2	– 49.16	– 49.16
$N(CH_3)_2$	NH_2	**7**		H-2	– 49.40	– 49.40
NH_2	NH_2	**9**		WC-3	– 58.68	– 58.68
				H-2	– 49.46	– 49.46

Even though the obtained complexation energies are far away from experimental results, the data show interesting trends. The Watson-Crick type complexes bonded via three hydrogen bonds are energetically the most favourable (ΔE_c around –59 kJ/mol). The Hoogsteen type complexes are less favourable (their ΔE_c is around –50 kJ/mol), yet still more favourable than Watson-Crick complexes bonded via two hydrogen bonds ($\Delta E_c \approx$ –49 kJ/mol). The energy difference is given by the type of complex, but it is

almost the same for different adenine derivatives bonded in the same complex type with thymine derivatives. Interestingly, the complexes with methylamino group participating in hydrogen bond are more stable than complexes bonded via amino protons. The results are depicted in Figure 3.24 below.

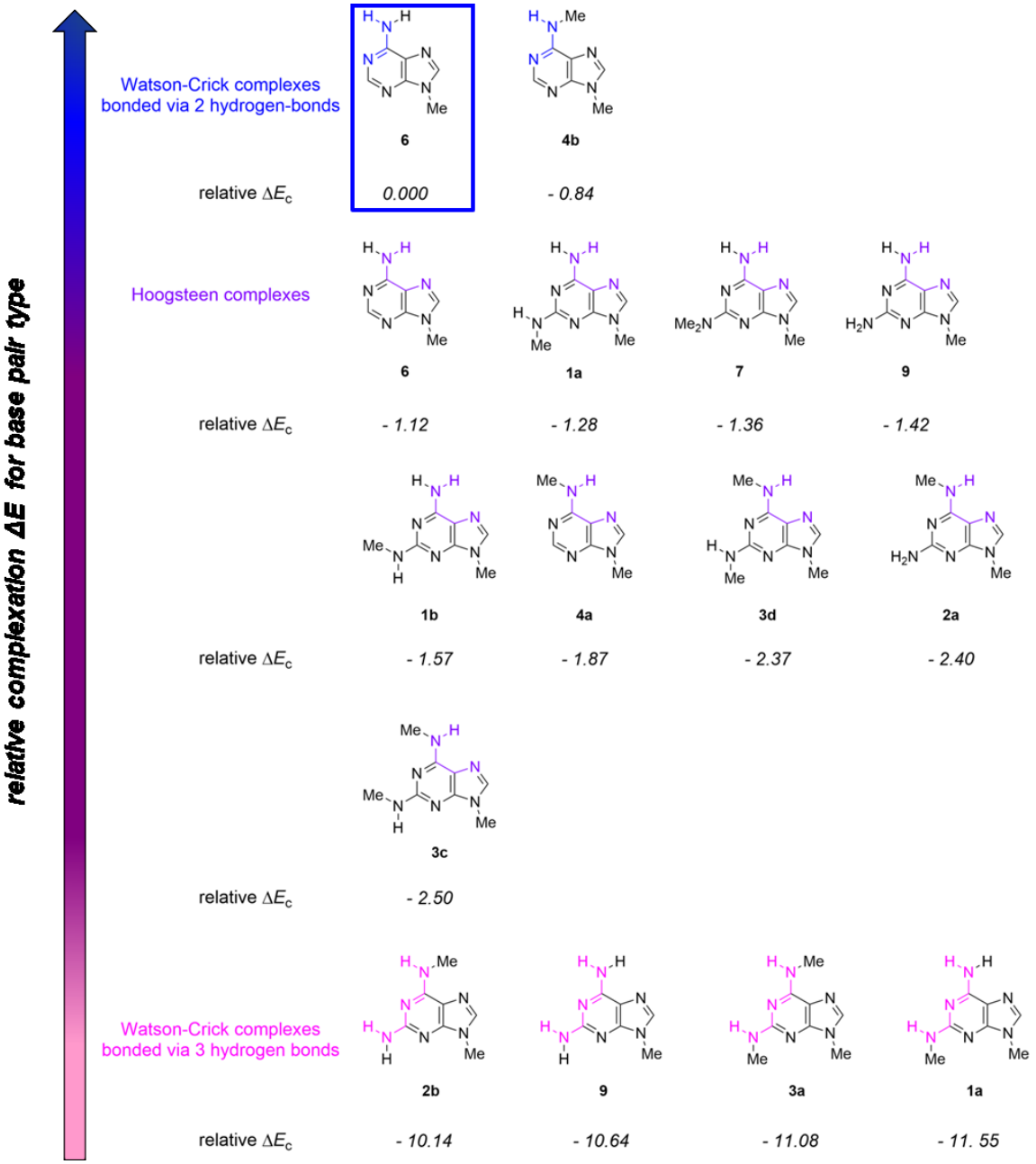


Figure 3.24 Comparison of ΔE_c for different base types normalized to the "natural" Watson-Crick base pair of derivative **6**

In Figure 3.24, the energy differences for Watson-Crick type complex of the "natural" adenine derivative **6** are normalized to 0.0 and the horizontal axis describes the hydrogen-bonded complex formation energy difference relative to adenine derivative **6** calculated from equation (3.6) for each adenine derivative. The hydrogen bonding patterns are highlighted for each compound and the derivative **6** is highlighted by a blue frame.

According to the computational study, N^6-methyladenine derivative **4** forms more stable complexes (Watson-Crick type and Hoogsteen type) with thymine derivative **5** than the "natural" adenine derivative **6**. On the other hand, the formation of Watson-Crick base pair of minor rotamer **4b** and thymine derivative **5** is limited by the unfavourable rotational equilibrium.

This unfavourable rotational equilibrium plays a key role in the Watson-Crick hydrogen-bonded complex formation between N^6-methyladenine derivative and thymine derivative **5**. The stabilisation energy, ΔE_{stabil} given by equation (3.7), includes also the energy penalty needed for the formation of a less stable rotamer **4b**. The counting with this energy penalty changes the order of the energies associated with the hydrogen-bonded complex formation depicted in Figure 3.24 above.

New order of the calculated energies associated with the hydrogen-bonded complex formation based on ΔE_{stabil} is shown in Figure 3.25 below. In Figure 3.25, the energy differences for Watson-Crick type complex of the adenine derivative **6** are normalized to 0.0 and the horizontal axis describes the hydrogen-bonded complex formation energy difference relative to adenine derivative **6** calculated from equation (3.7) for each adenine derivative. The hydrogen bonding patterns are highlighted for each compound and the derivative **6** is highlighted by a blue frame, as well as in Figure 3.24.

In the case of ΔE_{stabil}, there is no significant dependency of obtained energy on base-pair type as in the case of ΔE_c. The formation of Watson-Crick hydrogen-bonded complex between N^6-methyladenine derivative and thymine derivative **5** is less favourable than the formation of canonical Watson-Crick complex between thymine derivative **5** and adenine derivative **6**. It is caused by unfavourable rotational equilibrium between more stable *pro*-Hoogsteen rotamer **4a** and less stable *pro*-Watson-Crick rotamer **4b**. The Hoogsteen complex of N^6-methyladenine derivative **4** is more favourable than the complex of "natural" adenine derivative **6** according to both ΔE.

There are other interesting changes between the order of hydrogen-bonded complexes in dependency on ΔE_c and ΔE_{stabil}. For example, the formation of hydrogen-bonded complex between rotamer **1b** and thymine derivative **5** is the most unfavourable Hoogsteen complex in this study due to unfavourable rotational equilibrium. Also the relative energies associated with formation of the hydrogen-bonded complex formation between rotamers **2b**, **3a** and **3c** and thymine derivative **5** were shifted.

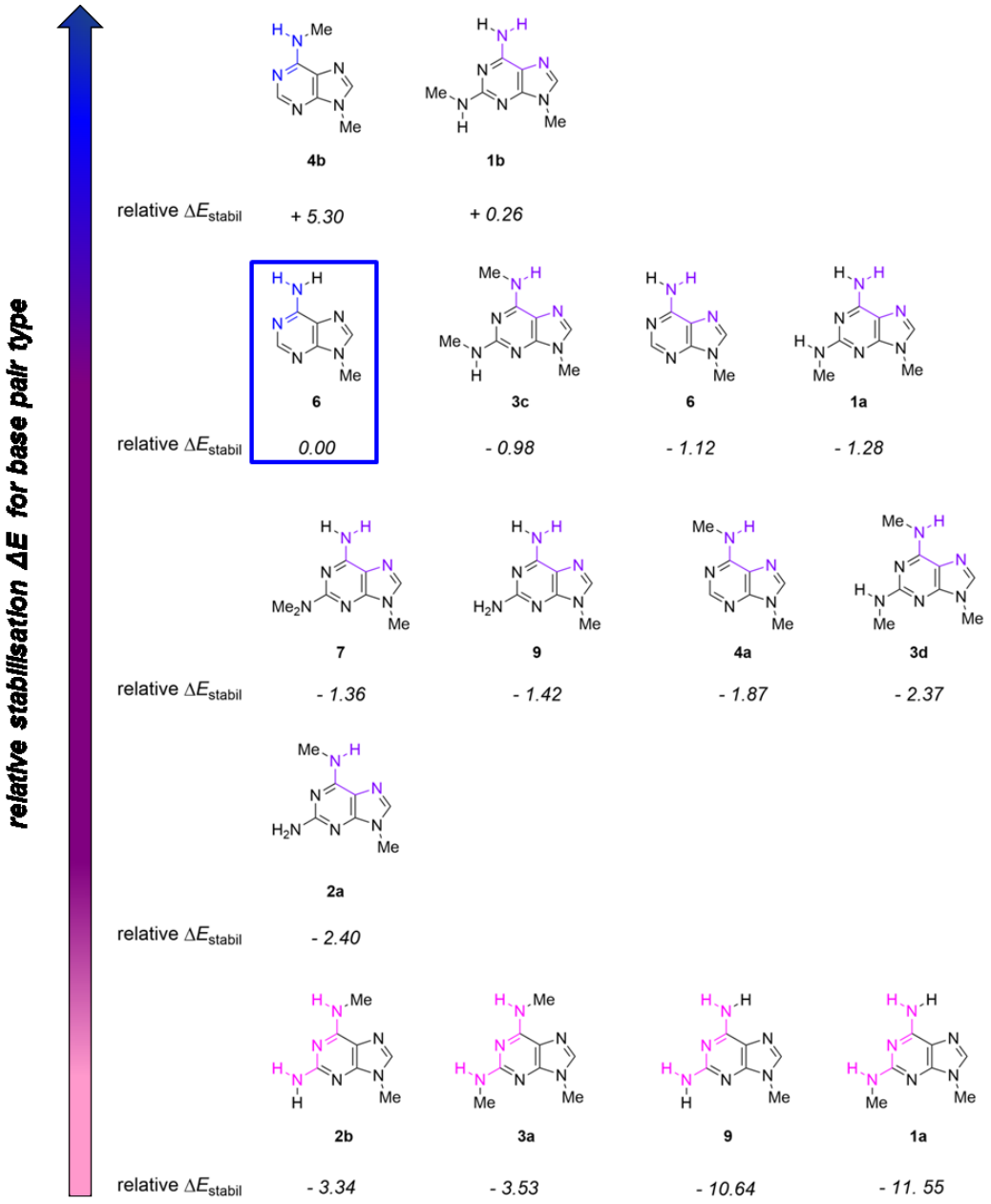


Figure 3.25 Comparison of ΔE_{stabil} for different base types normalized to the "natural" Watson-Crick base pair of derivative **6**

As depicted in the right corner of Figure 3.25 (compound **1a** and **9**), the trend of the stabilisation of the hydrogen bond by methylamino group is preserved.

NMR spectra can be predicted using computational chemistry by calculating the NMR shielding tensor. The shielding constant σ origins in the motion of electrons which create additional magnetic field in the opposite direction to B_0. These shielding values obtained from computations can be easily transformed into chemical shift values and compared with experimental data. The reported values represent the chemical shift relative to a standard compound, tetramethylsilane. The shielding values themselves help us to understand the electronic density in the molecule. Higher shielding values correspond to higher electronic density in vicinity of the observed nucleus and lower chemical shift.

For model compound **1,** the shielding tensors for all atoms were calculated at the same computational level as the geometry optimizations. Here, I only present the data obtained for amino- and methylamino group and imido proton of thymine derivative **5** because these protons are mostly affected by the hydrogen-bonded complex formation. We calculated the shielding values for isolated molecules of adenine and thymine derivatives investigated in this thesis and their complexes, utilizing the process described in Figure 3.23.

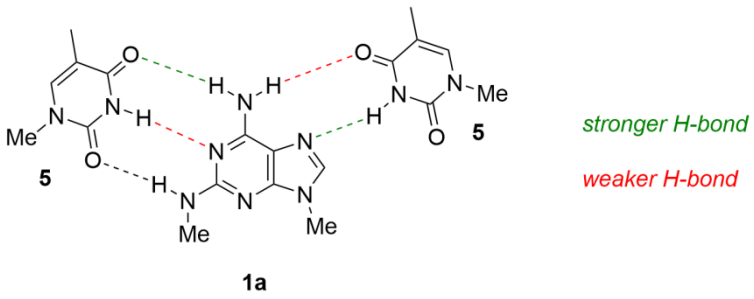


Figure 3.26 Comparison of the hydrogen-bond strength

The protons of adenine derivatives bonded in Watson-Crick type complexes are generally less shielded than the protons in Hoogsteen type complexes. This, according to the principles of general chemistry, indicates that the adenine hydrogen is involved in a stronger hydrogen bonds in Watson-Crick complexes than the thymine hydrogen. On the other hand, the thymine hydrogen is involved in stronger hydrogen bonds in Hoogsteen complexes. Bonding of the thymine derivative (H-bond donor) is stronger in

Hoogsteen type complexes because this proton is, in average, less shielded than in Watson-Crick complexes. These observations are depicted in Figure 3.26 for the case of bidentate complex of rotamer **1a**.

The shielding value differences (Δσ given by equation 3.8) of the imido proton in isolated thymine derivative **5** and in hydrogen-bonded complexes are summarized in Table 3.3. The increasing shielding value differences correlate with increasing bond strength, as it is an absolute value of Δσ presented in Table 3.3.

$$\Delta\sigma = \sigma_{\text{complex}} - \sigma_{\text{thymine derivative}} \tag{3.8}$$

where σ_{complex} is the shielding of the imido proton of thymine derivatives participating in a hydrogen-bonded complex with the adenine derivative and $\sigma_{\text{thymine derivative}}$ is the shielding of the imido proton of free thymine derivative **5**.

Table 3.3 The comparison of Δσ of thymine derivative **5** amino proton in hydrogen-bonded complexes

Adenine derivative		Cpd.	Rotamer	Complex	$\lvert\Delta\sigma\rvert$
R2	R6			type	[ppm]
$NHCH_3$	NH_2	**1**	**1a**	WC-3	7.08
			1a	H-2	7.97
			1b	H-2	8.01
NH_2	$NHCH_3$	**2**	**2a**	H-2	7.92
			2b	WC-3	6.92
$NHCH_3$	$NHCH_3$	**3**	**3a**	WC-3	6.82
			3c	H-2	8.03
			3d	H-2	7.98
H	$NHCH_3$	**4**	**4a**	H-2	7.77
			4b	WC-2	7.97
H	NH_2	**6**		WC-2	7.94
				H-2	7.74
$N(CH_3)_2$	NH_2	**7**		H-2	8.02
NH_2	NH_2	**9**		WC-3	7.12
				H-2	7.91

The shielding analysis for imidazole ring protons H8 also proved that the formation of Hoogsteen type complex affects the shielding value (in average 0.5 ppm). This is in good agreement with the experiment shown in Figure 3.13.

The similar analysis of Δσ can be obtained also for adenine derivatives. The following tables 3.4 and 3.5 show the shielding values differences calculated from analogous equation 3.9.

$$\Delta\sigma = \sigma_{\text{complex}} - \sigma_{\text{adenine derivative}} \qquad (3.9)$$

where σ_{complex} is the shielding of the amino and methylamino protons participating in a hydrogen-bonded complex with the thymine derivative **5** and $\sigma_{\text{adenine derivative}}$ is the shielding of these protons of free adenine derivatives.

Table 3.4 shows the shielding values differences obtained from equation (3.9) for the amino protons of the 6-aminoadenine derivatives.

Table 3.4 The comparison of Δσ of 6-N**H₂** protons in adenine derivatives in Watson-Crick and Hoogsteen hydrogen-bonded complexes

Adenine derivative		Cpd.	Rotamer	Complex	$\lvert\Delta\sigma\rvert$
R2	R6			type	[ppm]
$NHCH_3$	NH_2	**1**	**1a**	WC-3	4.84
			1a	H-2	3.81
			1b	H-2	3.81
H	NH_2	**6**		WC-2	4.56
				H-2	3.88
$N(CH_3)_2$	NH_2	**7**		H-2	3.81
NH_2	NH_2	**9**		WC-3	4.72
				H-2	3.86

The increasing shielding value differences correlate with increasing bond strength, as it is an absolute value of Δσ presented in Table 3.4. The amino protons bonded in the Watson-Crick complex with thymine are involved in a stronger hydrogen-bond than the protons participating on Hoogsteen type complex formation. This is in good agreement with Figure 3.26 above. The substituent on C2 has a small effect on the

hydrogen bond strength in Hoogsteen type complexes. The hydrogen bond strength in 2-methylaminoadenine derivative **1** is slightly higher than in 2-aminoadenine derivative **9** bonded in Watson-Crick complexes. This observation probably originates in electronic effects.

Table 3.5 shows the shielding values differences obtained from equation (3.9) for the methylamino protons of the N^6-methyladenine derivatives.

Table 3.5 The comparison of Δσ of 6-N**H**CH_3 protons in adenine derivatives in Watson-Crick and Hoogsteen hydrogen-bonded complexes

Adenine derivative		Cpd.	Rotamer	Complex	$\lvert\Delta\sigma\rvert$
R2	R6			type	[ppm]
NH_2	$NHCH_3$	**2**	**2a**	H-2	3.58
			2b	WC-3	4.27
$NHCH_3$	$NHCH_3$	**3**	**3a**	WC-3	4.31
			3c	H-2	3.51
			3d	H-2	3.50
H	$NHCH_3$	**4**	**4a**	H-2	3.60
			4b	WC-2	5.72

The substituents on C2 have only small effect on hydrogen bond strength in both types (WC-3) and (H-2) of complexes. Notice that the hydrogen bond involved in WC-2 of N^6-methyladenine derivative **4a** is significantly stronger than hydrogen bonds of other types (H-2 and WC-3).

Although the *N*-methylation leads to the stabilisation of hydrogen-bonded complexes formed, the hydrogen bonds involving methylamino protons are mostly weaker than the hydrogen-bonds involving only amino protons according to the Δσ study.

4. Technical details and methods

4.1 Synthesis

The nucleobase analogues were synthetized by Ing. Lucie Čechová and RNDr. Michal Šála, Ph. D. at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences. All synthetic details are summarized in Appendix A- Synthetic details.

4.2 Sample preparation and handling

The derivatives were dissolved in a mixture of d_7-*N,N*-dimethylformamide and d_2-dichloromethane (volume ratio 1:1) to give solutions with concentrations of 10, 15, 20, 25 or 30 mmol/dm^3. In the case of nucleobase mixtures, 3.6 ml of DMF-DCM solution was prepared at first. Weighted amount of a sample was dissolved in this solution to give concentration of 10 mmol/dm^3. Samples with binding partner stoichiometric ratio of 1:1, 1:1.5, 1:2, 1:2.5 or 1:3 were prepared by dissolving the stoichiometric weighted amounts of nucleobase analogue in excess in the prepared solution containing binding partner of concentration of 10 mmol/dm^3. Used solvents were purchased from Eurisotop (d_7-*N,N*-dimethylformamide) and Merck (d_2-dichloromethane).

4.3 NMR experiments

The NMR spectra were recorded on an Oxford Industries 500MHz superconducting magnet equipped with a triple resonance broad-band probe with ATM (5 mm PATBO BB-^{1}H/^{19}F/D Z-GRD) from Bruker. The probe was connected to temperature control units (B-VT 3200 for low temperature measurements and BCU 05 for room and higher temperature measurements). It is accompanied by a Bruker Avance II^{TM} 500 (two channels BOSS II, 34 magnetic field homogeneity corrections, BSMS 2, amplifiers BLAXH 300/100, BLAXH 300 and BLAX 500) NMR console.

The characterization experiments were performed on an UltraShield^{TM} 500 MHz superconducting magnet equipped with a broad-band cryo-probe with ATM module (5 mm CPBBO BB-^{1}H/^{13}C/^{15}N/ D Z/GRD) connected to temperature units BSVT and

BCU 05 accompanied by a Bruker Avance III^{TM} HD 500 console (three channels, BOSS III, 36 magnetic field homogeneity corrections, BSMS 2, amplifiers BLAX2H 300/100 and BLAX 300) and on an UltraShield^{TM} Plus 600 MHz superconducting magnet equipped with inverse triple resonance cryo-probe with ATM module (5 mm CPTCI $^{1}H/^{13}C/^{15}N/D$ Z-GRD) connected to temperature units BSVT and BCU 05 accompanied by a Bruker Avance III^{TM} HD 600 console (three channels, BOSS III, 36 magnetic field homogeneity corrections, BSMS 2, amplifiers BLAXH2H 300/100 and BLAX 500).

All chemical shifts are referenced to d_7-*N,N*-dimethylformamide (δ = 2.75 ppm). The characterization spectra are referenced to d_6-dimethylsulfoxide (δ = 2.50 ppm). The spectra were typically measured in the range of -1–13 ppm and 0–17 ppm.

4.4 Software and DFT parameters

Bruker Topspin 3.2 was used for operating the spectrometers. The spectra were processed by MestReNova 12.03 and the data analysis was performed in MS Excel. The inputs for DFT calculations were prepared with HyperChem 6 software and GaussView 6 software. The DFT calculations were performed by Gaussian 16 software[69] with the B3LYP functional[70] 6-311++G(3df, 2pd) basis set[71] polarizable continuum model of *N*,*N*-dimethylformamide solvation and empirical correction of dispersion GD3.[72] The triethylene glycol chain was replaced by a methyl group to simplify the calculations.

5. Conclusion

This thesis was focused on the study of the intermolecular hydrogen bonds between *N*-methylated adenine derivatives and thymine derivative via NMR spectroscopy. The first part of the thesis summarized the findings of dynamic NMR spectroscopy, hydrogen bonds and DNA structure published previously. This part can be easily used as a review for further research performed in these areas.

The aims of work were completed successfully. We used low temperature NMR spectroscopy to describe rotational equilibria of different *N*-methylated nucleobases, especially N^6-methyladenine derivative. The free-energy changes associated with the rotamer conversion is around 5 kJ/mol in the temperature range of 193 – 213 K. We proved that the rotamer with N^6-methyl oriented to the imidazole ring is less stable in all cases.

Then we discussed the effect of binding partner, the thymine derivative, to the rotational equilibria. The concentration of that rotamer, which is able to form hydrogen-bonded complex with the thymine derivative, increases with the increasing concentration of binding partner. We also found out, that adenine derivatives molecules can form intermolecular hydrogen-bonded complexes with each other. For this reason, we build-up a whole new methodology for determination of free-energy changes associated with the formation of hydrogen-bonded complex consisting of adenine and thymine derivative. This methodology is based on the chemical shift changes of the nucleobases alone and in the mixture with binding partner on its concentrations. The chemical shift changes of nucleobase alone were less concentration dependent than the chemical shift changes associated with the increasing concentration of their natural binding partner – thymine derivative. The obtained complexation energies were similar for systems bonded via three or two hydrogen bonds.

Our work was intensively supported by DFT calculations. Using these calculations, we established that *N*-methylation stabilize the Watson-Crick and Hoogsteen base pairs. We also sorted fifteen unique hydrogen-bonded complexes of adenine derivative with thymine derivative by their stability.

N^6-methyladenine, often discussed in connection to Alzheimer's disease, forms more stable Hoogsteen complexes with thymine than the "natural" adenine derivative. The formation of Watson-Crick complex between N^6-methyladenine derivative and

thymine derivative is much less favorable than the formation of canonical Watson-Crick base pair, due to unfavorable rotational equilibrium. I presume, that N^6-methyladenine can form Hoogsteen complex preferably.

References

1. Günther, H., *NMR spectroscopy : basic principles, concepts, and applications in chemistry*. Third ed.; Wiley-VCH: Weinheim, 2013.
2. Sanders, J. K. M.; Hunter, B. K., *Modern NMR spectroscopy : a guide for chemists*. Second ed.; Oxford University Press: Oxford, 1994.
3. Dračínský, M., NMR spektroskopie pro chemiky. https://nmr.group.uochb.cz/cs/kurz-nmr-i
4. Berger, S.; Braun, S., *200 and more NMR experiments: a practical course*. Wiley-VCH Verlag: Weinheim, 2004.
5. McMurry, J., *Organická chemie*. sixth ed.; VŠCHT Praha: Prague, 2007.
6. Goldman, M., *Quantum description of high-resolution NMR in liquids*. Oxford University Press: New York, 2002.
7. Silverstein, R. M.; Webster, F. X.; Kiemle, D. J., *Spectrometric Identification of Organic Compounds*. seventh ed.; Wiley: 2005.
8. Sečkářová, P.; Marek, R.; Maliňáková, K.; Kolehmainen, E.; Hocková, D.; Hocek, M.; Sklenář, V., Direct determination of tautomerism in purine derivatives by low-temperature NMR spectroscopy. *Tetrahedron Lett* **2004,** *45* (33), 6259-6263.
9. Tycko, R., NMR at Low and Ultra low Temperatures. *Accounts Chem Res* **2013,** *46* (9), 1923-1932.
10. Bártová, K.; Čechová, L.; Procházková, E.; Socha, O.; Janeba, Z.; Dračínský, M., Influence of Intramolecular Charge Transfer and Nuclear Quantum Effects on Intramolecular Hydrogen Bonds in Azopyrimidines. *J Org Chem* **2017,** *82* (19), 10350-10359.
11. Grabowski, S. J., Hydrogen bonding strenghts - measures based on geometrical and topological parameters. *J. Phys. Org. Chem.* **2004,** *17*, 18-31.
12. Atkins, P.; de Paula, J., *Fyzikální chemie*. VŠCHT Praha: Prague, 2013.
13. Pohl, R.; Socha, O.; Šála, M.; Rejman, D.; Dračínský, M., The Control of the Tautomeric Equilibrium of Isocytosine by Intermolecular Interactions. *Eur J Org Chem* **2018,** (37), 5128-5135.
14. Dingley, A. J.; Grzesiek, S., Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide (2)J(NN) couplings. *J Am Chem Soc* **1998,** *120* (33), 8293-8297.
15. Emsley, J., Very Strong Hydrogen-Bonding. *Chem Soc Rev* **1980,** *9* (1), 91-124.
16. Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K., NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Biophysics* **1998,** *95*, 14147-14151.
17. Kapur, G. S.; Cabrita, E. J.; Berger, S., The qualitative probing of hydrogen bond strength by diffusion-ordered NMR spectroscopy. *Tetrahedron Lett* **2000,** *41* (37), 7181-7185.
18. Braun, G.; Tierney, D.; Schmitzer, H., How Rosalind Franklin discovered the helical structure of DNA: Experiments in diffraction. *The Physics Teacher* **2011,** *49*, 140-143.
19. Kodíček, M.; Valentová, O.; Hynek, R., *Biochemie: chemický pohled na biologický svět*. VŠCHT Praha: Prague, 2015.
20. Noguera, M.; Sodupe, M.; Bertrán, J., Effects of protonation on proton transfer processes in Watson-Crick adenine-thymine base pair. *Theor Chem Acc* **2007,** *118* (1), 113-121.
21. Manchester, K. L., Historical Opinion: Erwin Chargaff and his 'rules' for the base composition of DNA: why did he fail to see the possibility of complementarity? *Trends Biochem Sci* **2008,** *33* (2), 65-70.
22. Rudner, R.; Karkas, J. D.; Chargaff, E., Separation of B Subtilis DNA into Complementary Strands . 3. Direct Analysis. *P Natl Acad Sci USA* **1968,** *60* (3), 921.
23. Crick, F. H. C.; Watson, J. D., The Complementary Structure of Deoxyribonucleic Acid. *Proc R Soc Lon Ser-A* **1954,** *223* (1152), 80.

24. Watson, J. D.; Crick, F. H., A structure for deoxyribose nucleic acid. *Ann Intern Med* **2003,** *138* (7), 581-582.
25. Mauroesguerroto, Side and top view of A-, B-, and Z-DNA conformations. Dnaconformation.png, Ed. 2014.
26. Hoogsteen, K., The Structure of Crystals Containing a Hydrogen-Bonded Complex of 1-Methylthymine and 9-Methyladenine. *Acta Crystallogr* **1959,** *12* (10), 822-823.
27. Hoogsteen, K., Crystal and Molecular Structure of a Hydrogen-Bonded Complex between 1-Methylthymine and 9-Methyladenine. *Acta Crystallogr* **1963,** *16* (9), 907.
28. Sen, D.; Gilbert, W., Formation of Parallel 4-Stranded Complexes by Guanine-Rich Motifs in DNA and Its Implications for Meiosis. *Nature* **1988,** *334* (6180), 364-366.
29. Lipps, H. J.; Rhodes, D., G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol* **2009,** *19* (8), 414-422.
30. Patel, D. J.; Phan, A. T.; Kuryavyi, V., Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res* **2007,** *35* (22), 7429-7455.
31. Sen, D.; Gilbert, W., A Sodium-Potassium Switch in the Formation of 4-Stranded G4-DNA. *Nature* **1990,** *344* (6265), 410-414.
32. Balasubramanian, S.; Hurley, L. H.; Neidle, S., Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov* **2011,** *10* (4), 261-275.
33. Monchaud, D.; Teulade-Fichou, M. P., A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem* **2008,** *6* (4), 627-636.
34. Kolesnikova, S.; Srb, P.; Vrzal, L.; Lawrence, M. S.; Veverka, V.; Curtis, E. A., GTP-Dependent Formation of Multimeric G-Quadruplexes. *ACS Chem Biol* **2019,** *14* (9), 1951-1963.
35. Raynaud, C. M.; Sabatier, L.; Philipot, O.; Olaussen, K. A.; Soria, J., Telomere lenght, telomeric proteins and genomic instability during the multistep carcinogenic process. *Critical Review in Oncology/Hematology* **2008,** *66*, 99-117.
36. Blackburn, E. H., Structure and Function of Telomeres. *Nature* **1991,** *350* (6319), 569-573.
37. Sühnel, J., Beyond nucleic acid base pairs: From triads to heptads. *Biopolymers* **2001,** *61* (1), 32-51.
38. Cheong, C. J.; Moore, P. B., Solution Structure of an Unusually Stable Rna Tetraplex Containing G-Quartet and U-Quartet Structures. *Biochemistry-Us* **1992,** *31* (36), 8406-8414.
39. Bouaziz, S.; Kettani, A.; Patel, D. J., A K cation-induced conformational switch within a loop spanning segment of a DNA quadruplex containing G-G-G-C repeats. *J Mol Biol* **1998,** *282* (3), 637-652.
40. Gallego, J.; Chou, S. H.; Reid, B. R., Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad: Solution structure of the d(TCCCGTTTCCA) dimer. *J Mol Biol* **1997,** *273* (4), 840-856.
41. Hingerty, B.; Brown, R. S.; Jack, A., Further Refinement of Structure of Yeast Transfer-Rna Phe. *J Mol Biol* **1978,** *124* (3), 523-534.
42. Gilbert, D. E.; Feigon, J., Multistranded DNA structures. *Curr Opin Struc Biol* **1999,** *9* (3), 305-314.
43. Radhakrishnan, I.; Patel, D. J., Solution structure of a pyrimidine-purine-pyrimidine DNA triplex containing TAT, C+GC and GTA triples. **1994,** *2*.
44. Radhakrishnan, I.; Patel, D. J., Solution structure of purine-purine-pyrimidine DNA triplex containing GGC and TAT triples. *Structure* **1993,** *1*.
45. Dunn, D. B.; Smith, J. D., The occurrence of 6-methylaminopurine in deoxyribonucleic acids. *Biochem J* **1958,** *68* (4), 627-36.
46. Ratel, D.; Ravanat, J. L.; Berger, F.; Wion, D., N6-methyladenine: the other methylated base of DNA. *Bioessays* **2006,** *28* (3), 309-315.
47. Dominissini, D.; Nachtergaele, S.; Moshitch-Moshkovitz, S.; Peer, E.; Kol, N.; Ben-Haim, M. S.; Dai, Q.; Di Segni, A.; Salmon-Divon, M.; Clark, W. C.; Zheng, G. Q.; Pan, T.; Solomon, O.;

Eyal, E.; Hershkovitz, V.; Han, D.; Dore, L. C.; Amariglio, N.; Rechavi, G.; He, C., The dynamic N-1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* **2016,** *530* (7591), 441-+.
48. Kou, Y.; Koag, M. C.; Lee, S., N7 Methylation Alters Hydrogen-Bonding Patterns of Guanine in Duplex DNA. *J Am Chem Soc* **2015,** *137* (44), 14067-14070.
49. Chen, Y.; Su, C.; Ke, M.; Jin, X.; Xu, L.; Zhang, Z.; Wu, A.; Sun, Y.; Yang, Z.; Tien, P.; Ahola, T.; Liang, Y.; Liu, X.; Guo, D., Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog* **2011,** *7* (10), e1002294.
50. De Bont, R.; van Larebeke, N., Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis* **2004,** *19* (3), 169-85.
51. Abbott, P. J.; Saffhill, R., DNA-Synthesis with Methylated Poly(dA-dT) Templates - Possible Role of O4-Methylthymine as a Pro-Mutagenic Base. *Nucleic Acids Res* **1977,** *4* (3), 761-769.
52. Venkateswarlu, D.; Lyngdoh, R. H. D., Structural, Steric and Energetic Requirements for Induction of Base Substitutional Mutations by Methylated Guanines and Thymines. *J Chem Soc Perk T 2* **1995,** (4), 839-846.
53. Yebra, M. J.; Bhagwat, A. S., A Cytosine Methyltransferase Converts 5-Methylcytosine in DNA to Thymine. *Biochemistry-Us* **1995,** *34* (45), 14752-14757.
54. Lu, L. H.; Yi, C. Q.; Jian, X.; Zheng, G. Q.; He, C. A., Structure determination of DNA methylation lesions N-1-meA and N-3-meC in duplex DNA using a cross-linked protein-DNA system. *Nucleic Acids Res* **2010,** *38* (13), 4415-4425.
55. Yang, H.; Zhan, Y. Q.; Fenn, D.; Chi, L. M.; Lam, S. L., Effect of 1-methyladenine on double-helical DNA structures. *Febs Lett* **2008,** *582* (11), 1629-1633.
56. Baszczyňski, O.; Jansa, P.; Dračínský, M.; Klepetářová, B.; Holý, A.; Votruba, I.; de Clercq, E.; Balzarini, J.; Janeba, Z., Synthesis and antiviral activity of N-9-[3-fluoro-2-(phosphonomethoxy)propyl] analogues derived from N-6-substituted adenines and 2,6-diaminopurines. *Bioorgan Med Chem* **2011,** *19* (7), 2114-2124.
57. Fu, Y.; Luo, G. Z.; Chen, K.; Deng, X.; Yu, M.; Han, D.; Hao, Z.; Liu, J.; Lu, X.; Dore, L. C.; Weng, X.; Ji, Q.; Mets, L.; He, C., N6-methyldeoxyadenosine marks active transcription start sites in Chlamydomonas. *Cell* **2015,** *161* (4), 879-892.
58. Wu, T. P.; Wang, T.; Seetin, M. G.; Lai, Y.; Zhu, S.; Lin, K.; Liu, Y.; Byrum, S. D.; Mackintosh, S. G.; Zhong, M.; Tackett, A.; Wang, G.; Hon, L. S.; Fang, G.; Swenberg, J. A.; Xiao, A. Z., DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature* **2016,** *532* (7599), 329-33.
59. Liu, J.; Zhu, Y.; Luo, G. Z.; Wang, X.; Yue, Y.; Wang, X.; Zong, X.; Chen, K.; Yin, H.; Fu, Y.; Han, D.; Wang, Y.; Chen, D.; He, C., Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. *Nat Commun* **2016,** *7*, 13052.
60. Xiao, C. L.; Zhu, S.; He, M. H.; Chen, D.; Zhang, Q.; Chen, Y.; Yu, G. L.; Liu, J. B.; Xie, S. Q.; Luo, F.; Liang, Z.; Wang, D. P.; Bo, X. C.; Gu, X. F.; Wang, K.; Yan, G. R., N-6-Methyladenine DNA Modification in the Human Genome. *Mol Cell* **2018,** *71* (2), 306-+.
61. Liang, D.; Wang, H.; Song, W.; Xiong, X.; Zhang, X. H.; Hu, Z. P.; Guo, H. H.; Yang, Z. J.; Zhai, S. D.; Zhang, L. H.; Ye, M.; Du, Q., The decreased N-6-methyladenine DNA modification in cancer cells. *Biochem Bioph Res Co* **2016,** *480* (1), 120-125.
62. Huajie, L.; Ren, Y.; Mao, K.; Hua, F.; ZYang, Y.; Wei, N.; Yue, C.; Li, D.; Zhang, H., FTO is involved in Alzheimer's disease by targeting TSC1-mTOR-Tausignaling. *Biochem Bioph Res Co* **2018,** *498*, 234-239.
63. Zhao, X.; Yang, Y.; Sun, B. F.; Shi, Y.; Yang, X.; Xiao, W.; Hao, Y. J.; Ping, X. L.; Chen, Y. S.; Wang, W. J.; Jin, K. X.; Wang, X.; Huang, C. M.; Fu, Y.; Ge, X. M.; Song, S. H.; Jeong, H. S.; Yanagisawa, H.; Niu, Y. M.; Jia, G. F.; Wu, W.; Tong, W. M.; Okamoto, A.; He, C.; Danielsen, J. M. R.; Wang, X. J.; Yang, Y. G., FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res* **2014,** *24* (12), 1403-1419.

64. Janke, E. M. B.; Dunger, A.; Limbach, H. H.; Weisz, K., Hydrogen bonding in complexes of adenosine and 4-thiouridine: a low-temperature NMR study. *Magn Reson Chem* **2001,** *39*, S177-S182.
65. Oertell, K.; Harcourt, E. M.; Mohsen, M. G.; Petruska, J.; Kool, E. T.; Goodman, M. F., Kinetic selection vs. free energy of DNA base pairing in control of polymerase fidelity. *P Natl Acad Sci USA* **2016,** *113* (16), E2277-E2285.
66. Dey, M.; Moritz, F.; Grotemeyer, J.; Schlag, E. W., Base-Pair Formation of Free Nucleobases and Mononucleosides in the Gas-Phase. *J Am Chem Soc* **1994,** *116* (20), 9211-9215.
67. Štoček, J. R.; Bártová, K.; Čechová, L.; Šála, M.; Socha, O.; Janeba, Z.; Dračínský, M., Determination of nucleobase-pairing free energies from rotamer equilibria of 2-(methylamino)pyrimidines. *Chem. Commun.* **2019,** *55* (74), 11075-11078.
68. Vaitiekunas, P.; Crane-Robinson, C.; Privalov, P. L., The energetic basis of the DNA double helix: a combined microcalorimetric approach. *Nucleic Acids Res* **2015,** *43* (17), 8577-8589.
69. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, X.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; T. Henderson; D. Ranasinghe; V. G. Zakrzewski; J. Gao; N. Rega; G. Zheng; W. Liang; M. Hada; M. Ehara; K. Toyota; R. Fukuda; J. Hasegawa; M. Ishida; T. Nakajima; Y. Honda; O. Kitao; H. Nakai; T. Vreven; K. Throssell; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision A.03*, Gaussian, Inc.: Wallingford CT, 2016.
70. Becke, A. D., Density-Functional Thermochemistry. 3. The Role of Exact Exchange. *J Chem Phys* **1993,** *98* (7), 5648-5652.
71. Mclean, A. D.; Chandler, G. S., Contracted Gaussian-Basis Sets for Molecular Calculations .1. 2nd Row Atoms, Z=11-18. *J Chem Phys* **1980,** *72* (10), 5639-5648.
72. Foresman, J. B.; Frisch, A., *Exploring Chemistry with Electronic Structure Methods*. third ed.; Gaussian, Inc.: Wallingford CT, 2015.

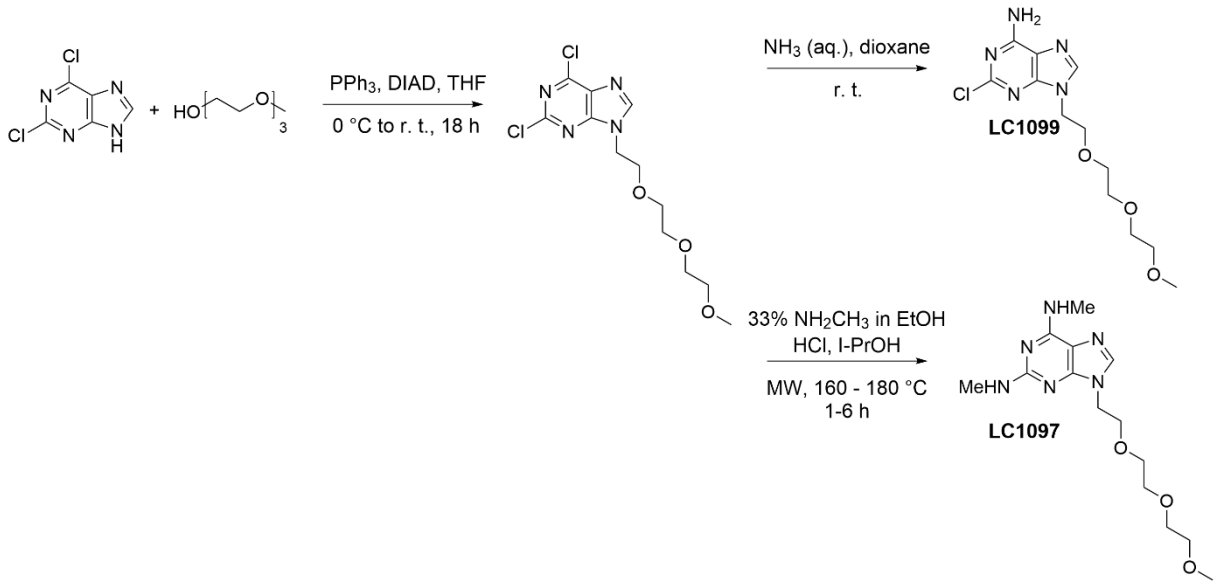
Appendix A – Synthetic details

All compounds investigated in this thesis were synthetized in the laboratories of Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences by Ing. Lucie Čechová and RNDr. Michal Šála, Ph.D.

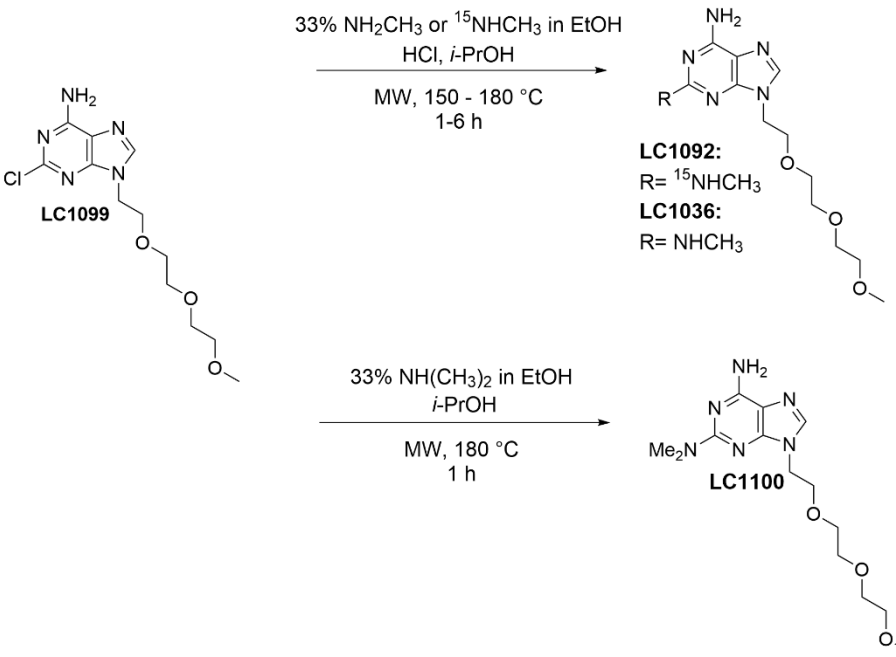
Each compound synthetized at our institute has a unique code which represents the compound in a register. Table A.1 below shows a transcript for each compound numbered in this thesis to its original unique code.

Table A.1 The synthetic codes of the investigated compounds

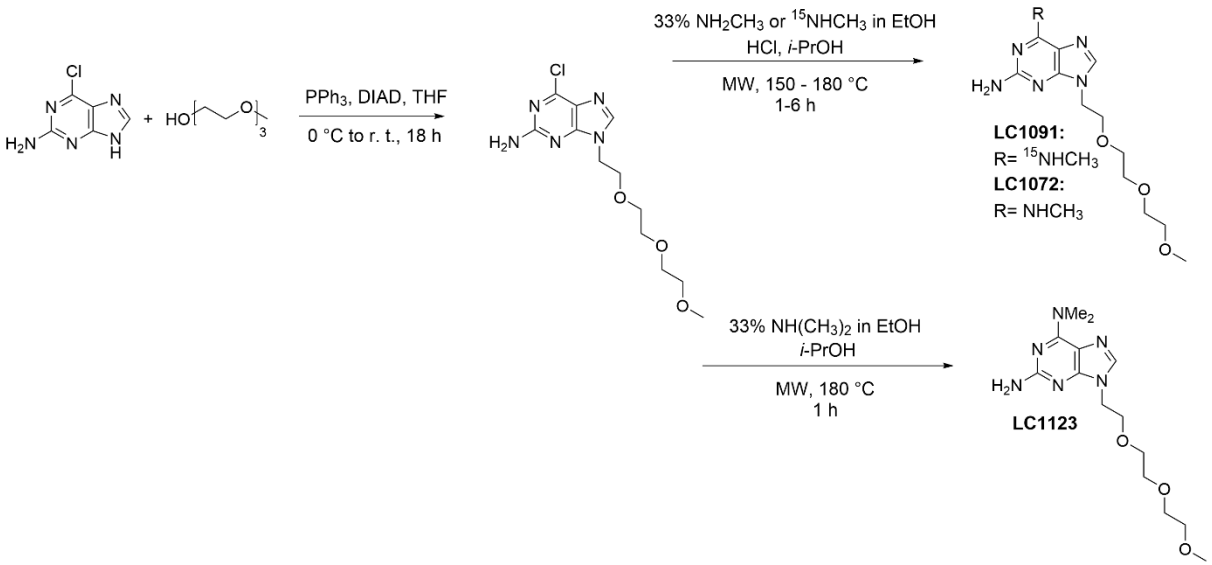
Number	Synthetic code
1	LC1036
1-15N	LC1092
2	LC1072
2-15N	LC1091
3	LC1097
4	LC1104
5	MS1471
6	MS991
7	LC1100
8	LC1123



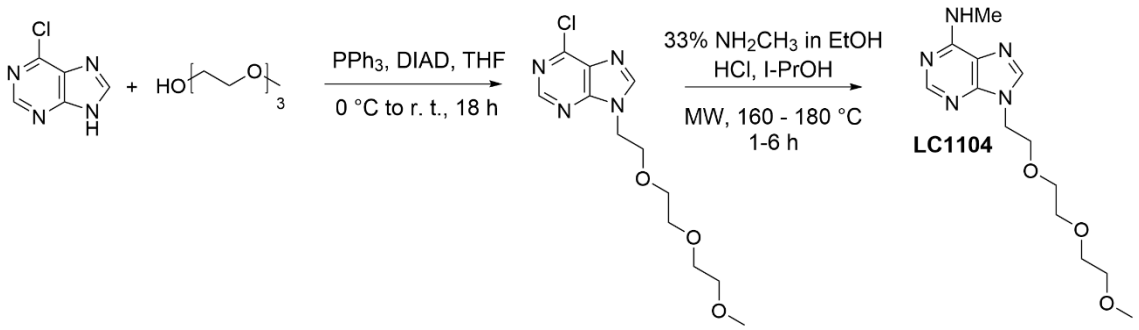
Scheme A.1 Synthesis of an intermediate LC1099 and compound LC1097



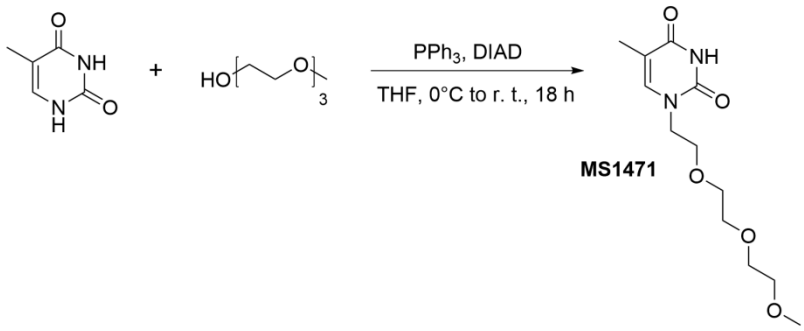
Scheme A.2 Synthesis of compounds LC1036, LC1092 and LC1100



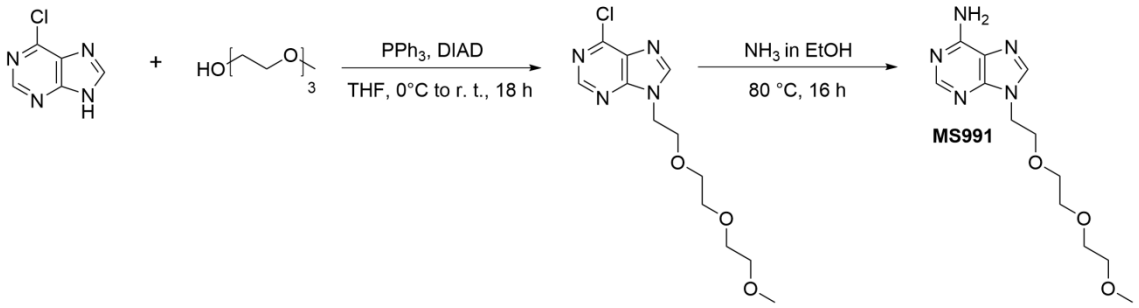
Scheme A.3 Synthesis of compound LC1091, LC1072 and LC1123



Scheme A.4 Synthesis of compound LC1104



Scheme A.5 Synthesis of compound MS1471



Scheme A.6 Synthesis of compound MS991

Appendix B – Characterization of compounds

All compounds investigated in this thesis were characterized via NMR spectroscopy and the ^{1}H and ^{13}C spectra transcripts are shown below. The atom numbering is depicted in Figure B.1.

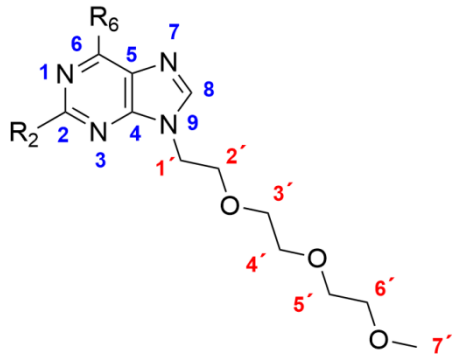


Figure B.1 The general atom numbering. R_2 and R_6 stay for substituents on C2 and C6 respectively.

Compound 1 (synthetic code LC1036)

^{1}H NMR (600 MHz, d_6-DMSO): δ = 2.76 (d, *J* = 4.82 Hz, NHC**H$_3$**), 3.22 (s, 3H, H7′), 3.37 – 3.54 (m, 8H, H3′, H4′, H5′, H6′), 3.74 (t, *J* = 5.5 Hz, 2H, H2′), 4.11 (t, *J* = 5.5 Hz, 2H, H1′), 6.16 (q, *J* = 4.7 Hz, N**H**CH$_3$), 6.62 (s, N**H$_2$**), 7.66 (s, 1H, H8)

^{13}C NMR (151 MHz, d_6-DMSO): δ = 28.36 (CH_3NH); δ = 42.11 (C1′), 58.04 (C7′), 68.34, 69.56, 69.65, 71.24 (C2′, C3′, C4′, C5′, C6′), 112.93 (C5), 137.60 (C8), 151.70 (C4), 155.89 (C6), 160.08 (C2)

Compound 1-15N (synthetic code LC1092)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 2.75 (d, *J* = 4.8 Hz, ^{15}NHC**H$_3$**), 3.21 (s, 3H, H7′), 3.33 – 3.57 (m, 8H, H3′, H4′, H5′, H6′), 3.73 (t, *J* = 5.5 Hz, 2H, H2′), 4.11 (t, *J* = 5.4 Hz, 2H, H1′), 6.19 (dq, $J_{C,N.}$ = 91.6 Hz, $J_{vic.}$ = 4.9 Hz, ^{15}N**H**CH$_3$), 6.65 (s, N**H$_2$**), 7.66 (s, 1H, H8)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 28.40 (d, J = 11.0, $CH_3{}^{15}NH$), 42.15 (C1′), 58.09 (C7′), 68.37, 69.60, 69.69, 71.28 (C2′, C3′, C4′, C5′, C6′), 112.94 (C5), 137.67 (C8), 151.73 (d, $J_{vic.}$ = 2.5 Hz, C4), 155.91 (d, $J_{vic.}$ = 2.9 Hz, C6), 160.09 (d, $J_{gem.}$ = 24.8 Hz, C2)

Compound 2 (synthetic code LC1072)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 2.86 (s, NHC**H$_3$**), 3.21 (s, 3H, H7′), 3.31 –3.52 (m, 8H, H3′, H4′, H5′, H6′), 3.70 (t, *J* = 5.6 Hz, 2H. H2′), 4.09 (t, *J* = 5.5 Hz, 2H, H1′), 5.85 (s, N**H$_2$**), 7.16 (s, N**H**CH_3), δ = 7.64 (s, 1H, H8)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 26.94 (CH_3NH), 42.20 (C1′), 58.09 (C7′), 68.45, 69.60, 69.65, 69.67, 71.27 (C2′, C3′, C4′, C5′, C6′), 113.33 (C5), 137.40 (C8), 150.82 (C4), 155.44 (C6), 160.29 (C2)

Compound 2-15N (synthetic code LC1091)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 2.86 (s, ^{15}NHC**H$_3$**), 3.21 (s, 3H, H7′), 3.34 -3.52 (m, 8H, H3′, H4′, H5′, H6′), 3.70 (t, *J* = 5.46 Hz, 2H, H2′), 4.09 (t, *J* = 5.4 Hz, 2H, H1′), 5.86 (s, N**H$_2$**), 7.16 (d, J = 94.2 Hz, ^{15}N**H**CH_3), 7.64 (s, 1H, H8)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 26.96 (CH_3NH), 42.22 (C1′), 58.10 (C7′), 68.46, 69.61, 69.66, 69.69, 71.28 (C2′, C3′, C4′, C5′, C6′), 113.32 (C5), 137.41 (C8), 150.74 (C4), 155.44 (d, *J* = 20.4 Hz, C6), 160.26 (C2)

Compound 3 (synthetic code LC1097)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 2.77 (d, *J* = 4.8 Hz, C2NHC**H$_3$**), 2.88 (s, C6NHC**H$_3$**), 3.21 (s, 3H, H7′), 3.31 – 3.54 (m, 8H, H3′, H4′, H5′, H6′), 3.73 (t, *J* = 5.4 Hz, 2H, H2′), 4.12 (t, *J* = 5.4 Hz, 2H, H1′), 6.28 (s, N**H**CH_3), 7.15 (s, N**H**CH_3), 7.64 (s, 1H, H8)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 27.20 (**C**H_3NHC6), 28.33 (**C**H_3NHC2), 42.12 (C1′), 58.08 (C7′), 68.38, 69.58, 69.59, 69.67, 71.26 (C2′, C3′, C4′, C5′, C6′), 113.58 (C5), 137.23 (C8), 151.39 (C4), 160.04 (C2)

Compound 4 (synthetic code LC1104)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 2.94 (s, 3H, NHC**H$_3$**), 3.20 (s, 3H, H7′), 3.34 – 3.53 (m, 8H, H3′, H4′, H5′, H6′), 3.77 (t, *J* = 5.3 Hz, 2H, H2′), 4.30 (t, *J* = 5.3 Hz, 2H, H1′), 7.68 (s, N**H$_2$**), 8.08 (s, 1H. H8), 8.22 (s, 1H. H2)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 27.20 (CH_3NH), 42.92 (C1′), 58.25 (C7′), 68.50, 69.75, 69.81, 71. 42 (C2′, C3′, C4′, C5′, C6′), 119.29 (C5), 141.08 (C8), 148.74 (C4), 152.59 (C2), 155.13 (C6)

Compound 5 (synthetic code MS1471)

See supporting information of reference 67 for more details.

Compound 6 (synthetic code MS991)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 3.21 (s, 3H, H7′), 3.32 – 3.53 (m, 8H, H3′, H4′, H5′, H6′), 3.77 (t, *J* = 5.4 Hz, 2H, H2′), 4.29 (t, *J* = 5.3 Hz, 2H, H1′), 7.20 (s, N**H$_2$**), 8.09 (s, 1H, H8), 8.13 (s, 1H, H2)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 42.73 (C1′), 58.08 (C7′), 68.30, 69.56, 69.57, 69.64 (C2′, C3′, C4′, C5′, C6′), 118.60 (C5), 141.19 (C8), 149.54 (C4), 152.37 (C2), 155.96 (C6)

Compound 7 (synthetic code LC1100)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 3.06 (s, N(C**H$_3$**)$_2$), 3.21 (s, 3H, H7′), 3.34 - 3.55 (m, 8H, H3′, H4′, H5′, H6′), 3.75 (t, *J* = 5.5 Hz, 2H, H2′), 4.13 (t, *J* = 5.5 Hz, 2H, H1′), 6.70 (s, N**H$_2$**), 7.69 (s, 1H, H8)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 37.11 ($(CH3)_2$NH), 42.11 (C1′), 58.08 (C7′), 68.30, 69.54, 69.60, 69.67, 71. 26 (C2′, C3′, C4′, C5′, C6′), 112.40 (C5), 137.98 (C8), 151.72 (C4), 155.56 (C6), 159.45 (C2)

Compound 8 (synthetic code LC1123)

^{1}H NMR (500 MHz, d_6-DMSO): δ = 3.22 (s, 3H, H7′), 3.32 – 3.52 (m, 8H, H3′, H4′, H5′, H6′, N(C**H$_3$**)$_2$), 3.68 (t, *J* = 5.4 Hz, 2H, H2′), 4.10 (t, *J* = 5.4 Hz, 2H, H1′), 5.81 (s, N**H$_2$**), 7.68 (s, 1H, H8)

^{13}C NMR (126 MHz, d_6-DMSO): δ = 37.87 ((CH3)$_2$NH), 42.38 (C1′), 58.23 (C7′), 68.52, 69.76, 69.83, 69.84, 71. 43 (C2′, C3′, C4′, C5′, C6′), 113.66 (C5), 137.03 (C8), 152.85 (C4), 154.90 (C4), 159.63 (C6)

Appendix C – List of structures of all compounds and their possible rotamers

All discussed structures and their possible rotamers are summarized in Figure C.1.

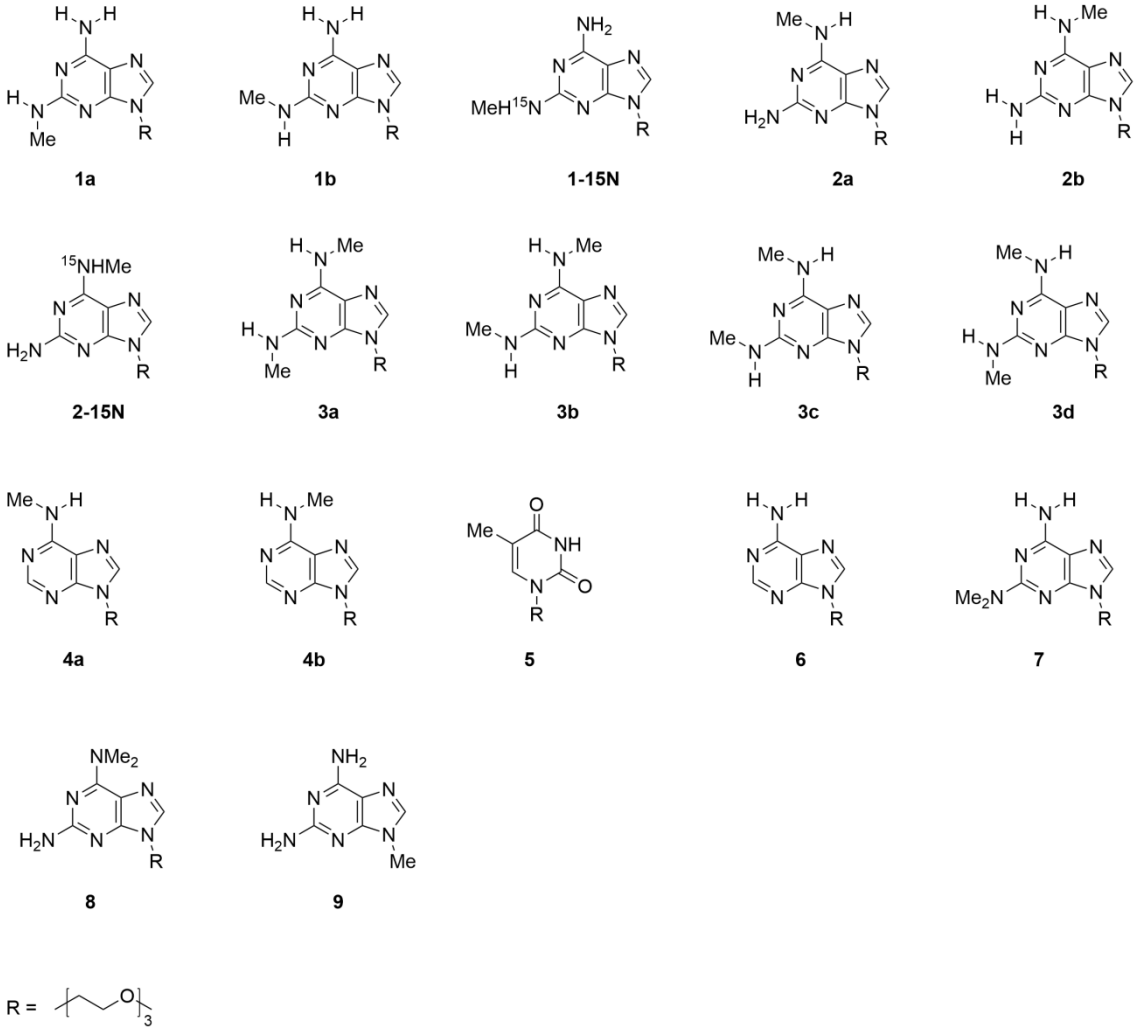


Figure C.1 All investigated structures