DFT Calculations on Metal-Nucleobase Complexes

Tushar van der Wijst

Technische Universität Dortmund

Voor mijn ouders

The work in this doctoral thesis has been carried out in the period of September 2005 till August 2009 at the Lehrstuhl für Bioanorganische Chemie, Fakultät Chemie, Technische Universität Dortmund and the Division of Theoretical Chemistry, Faculty of Sciences, Vrije Universiteit, Amsterdam, The Netherlands.

I am most grateful to my Ph. D. - supervisor

Prof. Dr. Bernhard Lippert

for giving me the opportunity to do research in his group and for the faith he had in me to complete my Ph.D. I am also very glad you gave me the freedom how to tailor my research. Apart from an excellent scientist you are also a person with a great engagement.

I am also most grateful to my Ph. D. co-supervisor

Prof. Dr. F. Matthias Bickelhaupt

for his hospitality in Amsterdam whenever I was there and the way he has helped me with interpreting the calculations and preparing the manuscripts. I also thank him for the faith he had in me to complete my Ph.D.

I am also very grateful to Dr. Célia Fonseca Guerra for her excellent daily supervision in Amsterdam. Without her large support this thesis would not have been so nice as it is now. I liked the way you improved my own research ideas and I thank you for the willingness to help me whenever you could. Fourth, I would like to thank Dr. Marcel Swart for helping me with more technical issues. Your recent program QUILD has proven to be mandatory for the calculations I had to do. I am also grateful for your maintaining the programs ADF and QUILD on the LISA so well.

Furthermore, I thank Dr. Pablo Sanz - Miguel and Dr. Barbara Müller for proof-reading parts of my thesis, thereby making the tough theoretical material presented in this thesis more readable for non-theoretical chemist. I also thank the secretaries of our group Birgit Thormann and Michaela Markert, who have guided me through the bureaucracy of Germany. To end, I would like to thank all group members of the Group of Biological Inorganic Chemistry at the Technische Unversität Dortmund and the Division of Theoretical Chemistry at the Vrije Universiteit in Amsterdam for the discussions I had with you individually, and for the nice atmosphere.

Moreover I am grateful to the examination committee for having critically read my doctoral thesis.

For the financial part I am most grateful that a stipend was awarded to me by the International Max-Planck Research School in Chemical Biology having its seat in Dortmund. I would like to thank in particular Dr. Jutta Rötter and Prof. Dr. Martin Engelhard, who coordinate the research school so well. I am also grateful to all the fellow-students with whom I have spent a very nice time.

Apart from the persons that have been involved in the scientific part I would like to thank some friends of mine. This will be done in their native language.

Beste Job, bedankt voor je vriendschap, gastvrijheid en het luisterende oor dat je geboden hebt tijdens mijn promotieperiode. Ik voel me altijd erg welkom bij jou en je vrouw Tosca. Ik hoop, dat je samen met haar en je dochters Marije en Lonneke een mooie toekomst tegemoet gaat. Tijdens mijn periode in Dortmund heb ik een, inmiddels, goede vriend leren kennen. Wesley, dear fellow, we hebben een erg gezellig periode gehad en ik bedank je voor je steun en voor je vriendschap. Onze discussies over muziek waren erg waardevol voor mij. Ik wens je alle goeds samen met Annemiek! Beste Erik, ik bedank je niet alleen voor je vriendschap maar ook voor de momenten dat je me hebt geholpen bij kleine problemen in mijn onderzoek. Ik heb onze discussies - met een lekker glas wijn of bier ! - altijd erg op prijs

gesteld. Ik wens je het allerbeste in je eigen loopbaan. Michel, jou ken ik het langst, en ook al raken we elkaar soms een tijd uit het oog, uit het hart ben je zeker niet! Bedankt voor je vriendschap, en heel veel geluk samen met Manon. Ik wens jullie beide heel veel geluk bij jullie aanstaand ouderschap.

Walter und Erika, ich danke Euch rechtherzlich für Eure Interesse in mir und großzügige Gastfreiheit und Freundschaft. Ich habe mich an jedem Sonntagabend, falls ich in Dortmund war, bei Euch in Eurem Gasthaus Wolf vergnügt und eine Menge Spaß gehabt. Darüber hinaus bedanke ich mich bei alle Stammkunden des Gasthauses Wolf für Eure Interesse in meiner Arbeit und die gemütlichen Stunden.

Lieve papa en mama, bedankt voor jullie onvoorwaardelijke steun en liefde tijdens deze promotie en alles wat daaraan vooraf ging. Zonder jullie hulp had ik me niet kunnen ontwikkelen tot degene die ik nu ben. Ik heb jullie altijd keihard nodig gehad (en dat zal altijd wel zo blijven) en jullie zijn er altijd voor me geweest, en daarom is dit proefschrift aan jullie opgedragen. Lakshmi, je bent de liefste, leukste en grappigste zus die ik me kan voorstellen en bedankt voor alles wat we hebben gedaan in Dortmund en in Nederland.

Al mijn andere familieleden bedank ik voor hun interesse in mijn promotie. Bedankt ook voor de gezellige uurtjes die ik samen met jullie heb doorgebracht als ik in Nederland was. In het bijzonder wil ik nog mijn oma noemen. Hoewel het allemaal abacadabra voor je zal zijn wat er in dit proefschift staat, heb je altijd gevraagd hoe het daar in Dortmund was. Ik bedank je voor je liefde die ik altijd van je heb mogen ontvangen.

I would like to end with a poem written by Franz Schobert, which was musically set by Franz Schubert. This poem is about my companion in hard times: the music.

An die Musik

Du holde Kunst, in wieviel grauen Stunden,
Wo mich des Lebens wilder Kreis umstrickt,
Hast Du mein Herz zu warmer Lieb' entzunden,
Hast mich in eine beßre Welt entrückt!

Oft hat ein Seufzer, Deiner Harf' entflossen,
Ein süßer, heiliger Akkord von Dir
Den Himmel beßrer Zeiten mir erschlossen,
Du holde Kunst, ich danke Dir dafür!

Contents

List of Publications

Parts of this doctoral thesis have already been published or are accepted for publication. A list is given below.

Chapter 3:
T. van der Wijst, C. Fonseca Guerra, M. Swart, F. M. Bickelhaupt.
Chem. Phys. Lett. **2006**, *426*, 415-421.

Chapter 4:
T. van der Wijst, C. Fonseca Guerra, M. Swart, F. M. Bickelhaupt, B. Lippert.
Chem. Eur. J. **2009**, *15*, 209-218.

Chapter 5:
C. Fonseca Guerra, T. van der Wijst, J. Poater, M. Swart, F. M. Bickelhaupt.
Theor. Chem. Acc. **2009**, on-line. DOI: 10.1007/s00214-009-0634-9.

Chapter 6:
T. van der Wijst, B. Lippert, C. Fonseca Guerra, M. Swart, F. M. Bickelhaupt.
J. Biol. Inorg. Chem. **2009**, on-line. DOI: 10.1007/s00775-009-0611-8.

Chapter 7:
T. van der Wijst, C. Fonseca Guerra, M. Swart, F. M. Bickelhaupt, B. Lippert.
Angew. Chem. **2009**, *121*, 3335-3337.
Angew. Chem. Int. Ed. **2009**, *48*, 3285-3287.

List of Abbreviations

1-MeUH	neutral 1-methyluracil
1-MeTH	neutral 1-methylthymine
1-MeU$^-$	deprotonated (N3) 1-methyluracil
9-MeAH	neutral 9-methyladenine
9-MeGH	neutral 9-methylguanine
A	Adenine base
A*	Adenine base in a rare tautomer form
A_4	Quartet of adenine bases
A_4*	Quartet of adenine base in a rare tautomer form.
A_4-N1	Quartet of adenine bases having N6 as donor, and N1 as acceptor
A_4-N3	Quartet of adenine bases having N6 as donor, and N3 as acceptor
A_4-N7	Quartet of adenine bases having N6 as donor, and N7 as acceptor
AT	Base pair consisting of an adenine and thymine bases, which can be hydrogen bonded or stacked
C	Cytosine base
C*	Cytosine base in a rare tautomer form
C_4	Quartet of cytosine bases
CC	Coupled Cluster
COSMO	COnductor-like Screening MOdel
DFT	Density Functional Theory
G	Guanine base
G*	Guanine base in a rare tautomer form
G_4	Quartet of guanine bases.
$(isoG)_4$	Quartet of isoguanine.
GC	Base pair consisting of a guanine and cytosine bases, which can be hydrogen bonded or stacked
GGA	Generalised Gradient Appoximation
GTO	Gaussian Type Orbital
HF	Hartree-Fock theory
HOMO	Highest Occupied Molecular Orbital
HOMO–3	Molecular Orbital that is three energy levels below the HOMO
K_{Taut}	Tautomerisation constant
LDA	Local Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
Me	Methylgroup (CH_3)
MP2	Second order Møller-Plesset perturbation theory
NMR	Nuclear Magnetic Resonance
pH	Negative decade logarithm of the hydrogen-ion concentration
pK_a	Negative decade logarithm of the acidity constant
STO	Slater Type Orbital
T	Thymine base
T*	Thymine base in a rare tautomer form
T_4	Quartet of thymine bases
U	Uracil base
U*	Uracil base in a rare tautomer form
U_4	Quartet of uracil bases
ZORA	Zeroth Order Regular Approximation

Chapter 1
General Introduction

1.1 Functions of Ions in Nucleic Acid Chemistry

In the human body "natural" ions occur at different concentrations: potassium (K^+: 0.2 M), magnesium (Mg^{2+}: 0.1 M) and sodium (Na^+: 0.01 M). The role of calcium (Ca^{2+}) is unclear, but possibly should be added to this list. These "natural" ions have several important functions. Firstly, as DNA is a polyanionic molecule, its charge should be neutralised. Instead of being a huge cloud around the DNA as proposed by Manning[1] distinct sites can be assigned where ions occupy the minor (Na^+, Mg^{2+}, Ca^{2+}) or major (Na^+, K^+, Mg^{2+}, Ca^{2+}) groove, thereby making DNA a possible ionophore.[2] An uneven distribution leads to DNA curvature.[3] Secondly, the "natural" metal ions are also needed to keep nucleic acid structures in their form. Examples include the structure of ribozymes (catalytic RNA, Mg^{2+})[4] and guanine quartets (Na^+, K^+).[5] Another need for Mg^{2+} becomes also clear in the formation or degradation of the phosphodiester bond.[6] The function of Mg^{2+} for these kind of reactions is to polarise the P-O bonds, to provide the nucleophile OH^-, or to stabilise one of the transition states or the departing group.[5]

The occurrence of heavy-metal ion species is also relevant because these are under the tight control of chaperone proteins. However, in some diseases these species may be become abundant. Examples are the pile-up of iron (Fe) in thalassemia or hematochromatosis or of copper (Cu) in Wilson's disease.[7] Other heavy metal ion species fulfil several roles. It has been found that manganese (Mn^V) is involved in oxidative damage of guanine which in the end may lead to DNA strand breakage.[8] The glycosidic bond (the bond between the nucleobase and sugar moiety) can break due to platination (Pt^{II}) of a nucleobase.[9] Sometimes, the function of heavy metal ions may resemble that of the "natural" ions in that they may contribute to the thermal stabilisation and destabilisation of the nucleic acid.[10] Furthermore, these species may also lead to local charge neutralisation, which may cause a collapse of the DNA, or eventual DNA condensation, or a simple distortion of DNA.[11] Heavy metals that are involved in here are, Cu^{II}, Mn^{II}, cobalt (Co^{II}), nickel (Ni^{II}), zinc (Zn^{II}) and cadmium (Cd^{II}). The kinking, due to platination, of DNA may have large effects. It is believed that through the kinking of DNA, cisplatin exerts its anti-cancer function.[12] More subtle changes induced by metal binding are the shift of tautomer equilibria that may induce mutations and influence base pairing properties.[13]

1.2 DNA

In this section, DNA is focussed on and the most common forms are A-DNA, B-DNA and Z-DNA, which are shown in Figure 1.1.

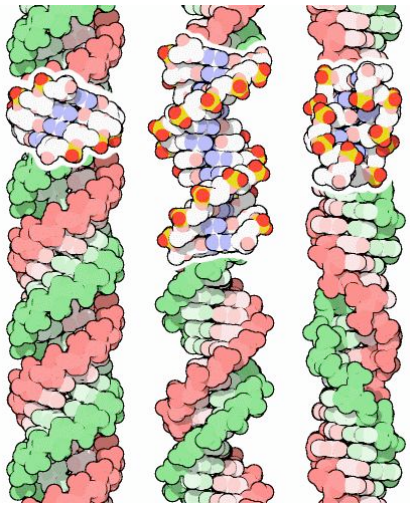
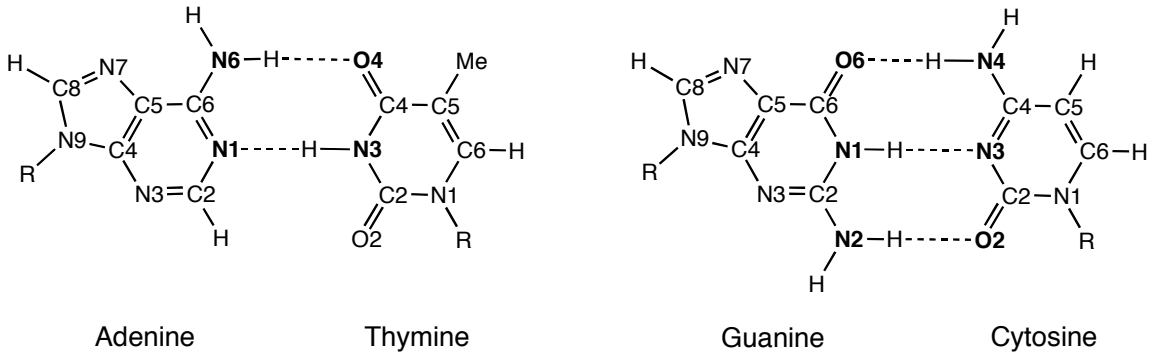


Figure 1.1. Schematic representation of A-DNA (left), B-DNA (middle), Z-DNA (right).

1.2.1 B-DNA

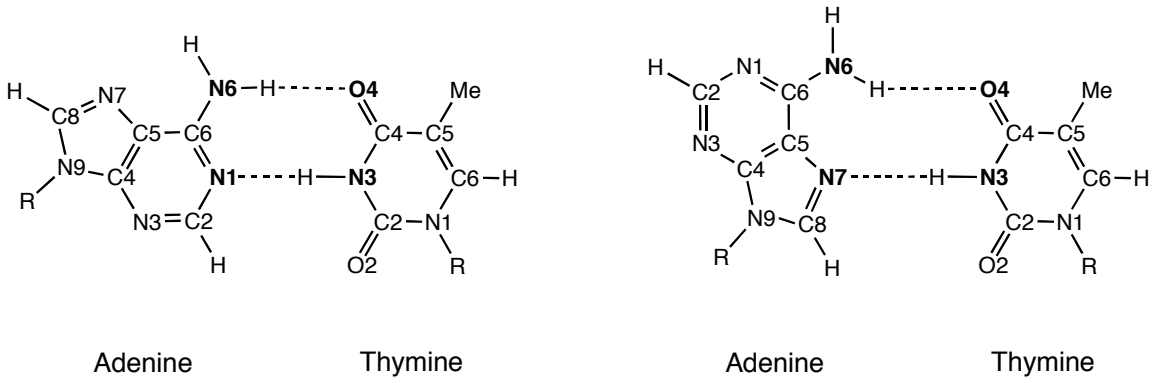
This form of DNA is the most common form of DNA and structural data about this molecule were proposed by James Watson and Francis Crick in 1953.[14] The main features they found will be discussed now. B-DNA, consists of two right handed helical chains existing of a central building block, the deoxyribonulecotides. These consist of one of the four nitrogenous bases, a sugar group and a phosphate group. The four nitrogenous bases are adenine (A), guanine (G), cytosine (C) and thymine (T). The nitrogenous bases can also be classified as a purine (A and G) or pyrimidine (C and T). The bases are at the inside of the helical chain, whereas the phosphate group are at the outside of the chain. Under normal conditions A pairs with T and G with C, which have two and three hydrogen bonds, respectively (Figure 1).[15a] In RNA thymine is substituted for uracil (U) and pairs with A. Because the two helical chains wind around each other, two gaps are formed, the major and minor groove. Restrictions of the actual sequence of AT and GC pairs do not occur and it is this unique sequence that contains the genetic information. In later crystallographic studies on small ribonucleic acid (RNA) fragments, such as AU and GC dimers, structural information became available and showed the base pairing directly.[16] Nowadays, crystallographic studies of larger fragments are known, but the observations that Watson and Crick made on the base pairing (*e.g.* A pairs T and G bairs C) in early days were recovered.[17] Furthermore, the base pairs AT and GC are widely used in supramolecular chemistry.[18]



Scheme 1.1. The base pairs adenine-thymine and guanine cytosine. Heavy atoms participating in hydrogen bonding are shown in bold.

1.2.2 Other Forms of DNA: A-DNA, Z-DNA and Multiple Stranded DNA

Although B-DNA is the best known form of double stranded DNA, other forms are known as well, namely, A- and Z-DNA.[15] There are some differences between these types of double stranded DNA. A- and B-DNA are right-handed but Z-DNA is left-handed. B-DNA is found in a relative high humid environment, whereas A-DNA is found in a relatively low salt environment. Apart from duplex DNA, multi-stranded DNA also exists. In this respect, triplex and quadruplex DNA will be discussed because of their biological role. Triplex DNA was discovered in 1957,[19a] only 4 years after the discovery of B-DNA. One of the first X-ray diffraction studies on triplex DNA appeared in 1974.[19b] The importance of triplex DNA stems from its potential use in gene therapy. Most illnesses occur due to a malfunctioning gene and, thus, leads to a malfunctional protein. First, the DNA is transcribed into RNA followed by the translation of the RNA into a protein. Due to the formation of triple helices, the DNA replication as well as transcription is inhibited, therefore, the malfunctioning protein will not be produced. In some case the illness is due to a single gene order (*e. g.* cystic fibrosis). The normal function of the gene can be restored by triple helix mediated mutagenesis, which in the end leads to a normal functioning protein.[20] The triple helices are formed in a site specific manner, according to the sequence of the malfunctioning gene, which makes the therapy very powerful. In analogy to duplex DNA, triplex DNA contains base triplets; these comprise C-G·C, T-A·A, T-A·T, C-G·G and C-C·(HC^+). Concerning the nomenclature of the triplets, a dash (-) indicates Watson-Crick base pairing and a dot (·) indicates Hoogsteen pairing, which was discovered in 1959 for the base pair AT (Scheme 1.2).[21]



Scheme 1.2. The base pair AT in Watson-Crick (left) and Hoogsteen (right) arrangement.

Apart from duplex and triplex DNA, quadruplex DNA or four stranded DNA, is also found. In quadruplex DNA, quartet arrangments of nucleobases are found. It has been found that the stabilisation of G-quartets inhibits telomerase, a crucial enzyme for the development of cancer. Under normal conditions, only double stranded DNA is target for telomerase, whereas quadruplex DNA is not. Stabilisation of G-quartets can be done via small aromatic molecules having an electron deficient ring, that can participate in stacking with the electron rich G-quartet.[22a] Other nucleobase quartets may also form, but it seems that the G-quartet is a scaffold for their formation, in the sense that it serves as a partner to stack with.[22b,c] G-quartets are also widely used in supramolecular chemistry.[23a] For example, by attaching them to calixarenes,[23b,c] recognition of ion-pairs is possible. The question now arises whether the role of the calixarene that recognises the anion, can be taken over by, for instance, an adenine quartet. Which implies that quadruplex DNA can be used as a salt extractor, as it binds an ion-pair.

1.3 Metal-Nucleic Acid Interactions: Binding Aspects and Implications

1.3.1 Metal Binding to Nucleic Acids

There is already ample information on metal-nucleic acids interactions,[24] but in this section a condensed overview about this issue will be given. Taking into account the structure of a nucleotide several binding parts are possible. First, binding to the phosphate backbone is possible for alkali and earth alkali metal ions and a large amount of X-ray structures is available.[25] Second, binding to the sugar entity is possible. Although it is the least frequent one, a chelate of O2' and O3' is formed with a variety of metal ions: Na^{+},[26] Cu^{II},[27] Sn^{IV},[28] and Os^{IV}.[29] Third, and by far the most frequent occurring one, is the interaction of a metal ion with the heterocyclic ring of the nucleobase. Almost every binding pattern is possible: a

single metal may bind but multiple metal binding is possible as well, and the metal species do not need to be the same.[24] With regard to binding to the heterocyclic ring of a nucleobase, there are two possibilities for metal binding. The first possibility is metal binding to a site that is not involved in Watson-Crick hydrogen bonding, thereby not changing the tautomeric structure of the nucleobase. A second possibility is that a metal replaces a proton that normally is involved in Watson-Crick hydrogen bonding. In this case, the metal "forces" the proton to reside at another place, thereby creating a different tautomer structure. So called "metal-stabilised-rare tautomers" are possible for all nucleobases, and occur for A-N6, G-N1, G-N2, C-N4 and T/U-N3. As a consequence of the changed tautomeric structure the normal base pairing scheme is affected and mutations may result.[24,30]

1.3.2 Implications for Metal Binding to Nucleic Acids

Generally, the binding of a metal ion will as a result in an *increase* of the acidity (decrease of the pK_a) of endocyclic NH groups and exocyclic NH_2 groups, or a *decrease* of the basicity (increase of the pK_a) of endocyclic N atoms.[24] To understand the changes in pK_a values one should have a clear view of the metal-nucleobase bond and some aspects should be discussed. First, the major contribution of the metal-nucleobase interaction is to be addressed to the electrostatic interaction.[31] This component is more pronounced in the gas phase than in solution. Secondly, stabilising interactions such as hydrogen bonding between co-ligands of the metal and the nucleobase will further improve the overall interaction.[32] The third and last aspect concerns non-electrostatic effects such as polarisation and charge-transfer.[33] These interactions refer to the redistribution of electrons in the nucleobase due to metal binding. The degree of the acidification depends both on the electronic effects brought about by the metal-ion, occurring though bonds as well as long range electrostatic repulsion between the positively charged metal and the proton under consideration. Of course, many factors will influence the latter factor namely: distance between the metal and the proton and net charge of the metal (take co-ligands into account). Solvent effects will affect both factors.[24]

1.3.3 Biological Applications

There are some applications that are important with regard to nucleic acid chemistry and several will be discussed shortly: acid-base chemistry, antisense therapy and the medicinal application of cisplatin. The pK_a-shift towards physiological values (7.0 - 7.4), due to metal

binding, is most relevant because acid-base catalysis seems to be a fruitful application. For instance protein synthesis (translation) is a process in which acid-base catalysis is involved. This is largely possible in ribozymes.[34] Binding patterns of nucleobases that will draw the pK_a to the physiological range are G-N7, G-N1, A-N6 and C-N4.[24] In this context, the pK_a value of 7.6 of an adenine residue should be mentioned.[34b-d]

Another application stems from the fact that within a base pair a proton that normally participates in hydrogen bonding can be substituted for a metal. One can take advantage of this property in antisense and antigene strategy.[35] In essence, one creates, apart from the reversible hydrogen bonding, an irreversible cross-link between the oligonucleotide and target RNA or DNA strand can be made. To this end one has managed to modify oligonucleotides with a *trans*-$(NH_3)_2Pt^{II}$ entity.[36] This Pt-modified oligonucleotide was successfully hybridised with single stranded oligonucleotides,[37] and double stranded DNA.[38]

In medicinal chemistry, especially cancer treatment, cisplatin is commonly used. The antitumour activity of cisplatin (*cis*-$[Pt^{II}Cl_2(NH_3)_2]$) was reported on in 1969 by Rosenberg.[39a] This was after the same author reported on the filamentous growth of *Escherichia Coli* bacteria.[39b] Rather than RNA or protein synthesis, DNA synthesis was affected which turned the investigations towards interactions between Pt and DNA.[40] One set out to investigate the reactivity towards DNA and several binding patterns were found or at least proposed: DNA interstrand cross-linking,[41a] DNA intrastrand cross-linking,[41b] and DNA protein cross-linking[41c] were identified. Around the 1980s the significance of the intrastrand binding of two adjacent guanines or adjacent guanine and adenine was recognised, and a few years later the major adducts of cisplatin were found.[42] It took some years before evidence based on X-ray crystallography became available but the structure of 1,2-GG adduct in a DNA fragment is noteworthy.[43] There also exist X-ray data on the 1,2-AG adduct, but this concerns a model nucleobase.[44]

The following intrastrand cross-link adducts occur at most: 60-65% (1,2-GG) and 20-25% (1,2-AG), less important is the 1,3-intrastrand crosslink adduct, which is formed up to a few percents.[42] With regard to the binding of Pt^{II}, the N7 position of A and G is crucial, and is readily accessible in the major groove of DNA.[45] Due to the metal binding the DNA helix is kinked such, that several complicated mechanisms in the cell are put to work which in the end will lead to apoptosis.[46] Although cisplatin is widely used, there are other transition metal complexes that have an anti-cancer activity, such of Fe, Co, gold (Au),

titanium (Ti), ruthenium (Ru), gallium (Ga) [47] Cancer can not only be cured through the interaction of metals with nucleic acids, it can also be detected alone by nucleic acid based methods.[48] In general, metals are widely used in medicinal chemistry, not only as therapeutics but also as diagnostics.[49]

1.4 Overview of This Thesis

In this thesis, some of the above mentioned topics in metal nucleic acid chemistry are discussed. The remainder of the thesis is organised as follows. Chapter 2 will give an overview of the theory and computational methods used in this thesis. Chapter 3 demonstrates the excellent performance of the BP86 functional for hydrogen bonds of the Watson-Crick base pairs AT and GC, whereas the popular B3LYP functional is shown to perform less satisfactory. Chapter 4 shows how one can influence, by means of Pt^{II} coordination, the relative stabilities of rare tautomers of 1-methylthymine and 1-methyluracil. From Chapter 5 on larger nucleobase aggregates will be treated, namely quartet structures. For this kind of systems dispersion should be treated well. Chapter 5 deals with this question: here, a performance study is carried out and shows the excellent performance of the dispersion-corrected BLYP-D functional. The latter performs very well for stacked AT and GC base pairs. Furthermore, it is demonstrated that the minima found with the BLYP-D functional do not coincide with those of the B3LYP functional. Thus, B3LYP fails particularly for stacking interactions. In Chapter 6 the bonding of cations and anions to diverse adenine quartets is demonstrated. As an extension of this work, chapter 7 demonstrates the possibility of a ditopic ion pair receptor for sodium chloride based on a stacking complex of A_4 and G_4.

1.5 References

[1] G. S. Manning, *Q. Rev. Biophys.* **1978**, *11*, 179-246.

[2] N. V. Hud, M. Polak, *Curr. Opin. Struct. Biol.* **2001**, *11*, 293-301.

[3] K. M. Kosikov, A. A. Gorin, X.-J. Lu, W. K. Olson, *J. Am. Chem. Soc.* **2002**, *124*, 4838-4847.

[4] (a) R. G. Kuimelis, L. W. McLaughlin, *Chem. Rev.* **1998**, *98*, 1027-1044.
(b) B. Lippert, *Chem. Biodiversity* **2008**, *5*, 1455-1474.

[5] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature (London)*, **2002**, *417*, 876-880.

[6] (a) E. Westhof, *Science* **1999**, *286*, 61-62.
(b) M. P. Roberston, W. G. Scott, *Science* **2007**, *315*, 1549-1553.

[7] B. Lippert in *Nucleic Acid-Metal Interactions* (Ed.: N. V. Hud), Royal Society of Chemistry, London, **2008**.

[8] C. Vialas, C. Claparols, G. Pratviel, B. Meunier, *J. Am. Chem. Soc.* **2000**, *122*, 2157-2167.

[9] (a) J. Arpalahti, A. Jokilammi, H. Hakala, H. Lönnberg, *J. Phys. Org. Chem.* **1991**, *4*, 301-309.
(b) A. Försti, P. Vodicka, K. Hemminki, *Chem.-Biol. Interact.* **1990**, *74*, 253-261.

[10] G. L. Eichhorn, Y. A. Shin, *J. Am. Chem. Soc.* **1968**, *90*, 7323-7328.

[11] V. A. Bloomfield, *Curr. Opin. Struct. Biol.* **1996**, *6*, 334-341.

[12] (a) D. B. Zamble, S. J. Lippard in *Cisplatin - Chemistry and Biochemistry of a Leading Anticancer Drug* (Ed.: B. Lippert), VHCA Zürich and Wiley-VCH, Weinheim, **1999**, pp. 73.
(b) E. R. Jamieson, S. J. Lippard. *Chem. Rev.* **1999**, *99*, 2467-2498.

[13] (a) M. A. Sirover, L.A Loeb, *Sciene* **1976**, *194*, 1434-1436.
(b) L. J. N. Bradley, K. J. Yarema, S. J. Lippard, J. M. Essigman, *Biochemistry* **1993**, *32*, 982-988.
(c) J. Müller, R. K. O. Sigel, B. Lippert, *J. Inorg. Biochem.* **2000**, *79*, 261-265.

[14] J. D. Watson, F. H. C. Crick, *Nature*, **1953**, *171*, 737-738.

[15] (a) L. Stryer, *Biochemistry*, W. H. Freeman, New York, **1995**.
(b) A. Ghosh, M. Bansal, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 620-626.

[16] a) N. C. Seeman, J. M. Rosenberg, F. L. Suddath, J. J. P. Kim, A. Rich, *J. Mol. Biol.* **1976**, *104*, 109-144.
b) J. M. Rosenberg, N. C. Seeman, R. O. Day, A. Rich, *J. Mol. Biol.* **1976**, *104*, 145-167.

[17] (a) X. Q. Shui, L. McFail-Isom, G. G. Hu, L. D. Williams, *Biochemistry*, **1998**, *37*, 8341-8355.
(b) G. Tereshko, M. Minasov, M. Egli, *J. Am. Chem. Soc.* **1999**, *121*, 470-471.

[18] J. L. Sessler, C. M. Lawrence, J. Jayawickramarajah, *Chem. Soc. Rev.* **2007**, 314-325.

[19] (a) G. Felsenfeld, D. R. Davies, A. Rich, *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.
(b) S. Arnott, E. Selsing, *J. Mol. Biol.* **1974**, *88*, 509-512.

[20] (a) P. P. Chan, P. M. Glazer, *J. Mol. Med.* **1997**, *75*, 267-282.
(b) K. M. Vasquez, J. H. Wilson, *Trends Biochem. Sci.* **1998**, *23*, 4-9.
(c) N. T. Thuong, C. Hélène, *Angew. Chem. Int. Ed.* **1993**, *32*, 666-669.

[21] K. Hoogsteen, *Acta Cryst.* **1959**, *12*, 822-823.

[22] (a) J. T. Davis, *Angew. Chem., Int. Ed.* **2004**, *43*, 668-698.
(b) S. Neidle, S. Balasubramanian (Eds.),*Quadruplex Nucleic Acids*, RSC Publishing, Cambridge, **2006**.

(c) J. Gros, F. Rosu, S. Amrane, A. De Cian, V. Gabelica, L. Lacroix, J.-L. Mergny, *Nucleic Acids Res*. **2007**, *35*, 3064-3075.

[23] (a) J. T. Davis, G. P. Spada, *Chem. Soc. Rev*. **2007**, *36*, 296-313.
(b) F. W. Kotch, V. Sidorov, K. Kayser, Y. F. Lam, M. S. Kaucher, H. Li,J. T. Davis, *J. Am. Chem. Soc*. **2003**, *125*, 15140-15150.
(c) F. W. B. van Leeuwen, J. T. Davis, W. Verboom, D. N. Reinhoudt, *Inorg. Chim. Acta* **2006**, *359*, 1779-1785.

[24] (a) B. Lippert, *Coord. Chem. Rev*. **2000**, *200-202*, 487-516.
(b) B. Lippert, *Prog. Inorg. Chem*. **2005**, *54*, 385-447.

[25] (a) K. Aoki, *Met. Ions. Biol. Syst*. **1996**, *32*, 91-120.
(b) K. Aoki in *Comprehensive Supramolecular Chemistry Vol. 5* (Ed. J. M. Lehn), Pergamon Press, Oxford, **1995**, pp. 249.

[26] T. J. Kistenmacher, C. C. Chiang, P. Chalilpoyil, L. G. Marzilli, *J. Am. Chem. Soc*. **1979**, *101*, 1143-1148.

[27] (a) Y.-Y. H. Chao, D. R. Kearns, *J. Am. Chem. Soc*. **1977**, *99*, 6425-6434.
(b) J. Galay, A. Mosset, I. Grenthe, I. Puigdomènech, B. Sjöberg, F. Hultén, *J. Am. Chem. Soc*. **1987**, *109*, 380-386.

[28] A. Janscó, L. Nagy, E. Moldrheim, E. Sletten, *J. Chem. Soc., Dalton Trans*. **1999**, 1587-1594.

[29] J. F. Conn, J. J. Kim, F. L. Suddath, P. Blattmann, A. Rich, *J. Am. Chem. Soc*. **1974**, *96*, 7152-7153.

[30] M. D. Topal, J. R. Fresco, *Nature* **1976**, *263*, 285-289.

[31] J. V. Bura, J. Sponer, P. Hobza, *J. Phys. Chem*. **1996**, *100*, 7250-7255.

[32] (a) C. B. Black, J. A. Cowan, *J. Am. Chem. Soc*. **1994**, *116*, 1174-1178.
(b) J. Sponer, J. V. Burda, M. Sabat, J. Leszczynki, P. Hobza, *J. Phys. Chem. A* **1998**, *102*, 5951-5957.
(c) J. Sponer, J. V. Burda, M. Sabat, J. Leszczynski, P. Hobza, *J. Biol. Struct. Dyn*. **1999**, *17*, 61-77.
(d) J. Sponer, M. Sabat, L. Gorb, J. Leszczynski, B. Lippert, P. Hobza, *J. Phys. Chem. B* **2000**, *104*, 7535-7544.

[33] J. V. Burda, J. Sponer, J. Leszczynski, P. Hobza, *J. Phys. Chem. B* **1997**, *101*, 9670-9677.

[34] (a) B. Lippert, *Chem. Biodiversity* **2008**, *5*, 1455-1474 (and refs. cited therein).
(b) P. Nissen, J. Hansen, N. Ban, M. Moore, T. A. Seitz, *Science* **2000**, *289*, 920-930.
(c) G. W. Muth, L. Ortoleva-Donnely, S. A. Strobel, *Science* **2000**, *289*, 947-950.
(d) P. B. Moore, T. A. Seitz, *RNA* **2003**, *9*, 155-159.

[35] (a) B. Lippert, M. Leng in *Topics Biol. Inorg. Chem. Vol 1*, (Eds.: M. J Clarke and P. J Sadler), Springer, Berlin, **1999**, pp. 117.
(b) P. C. Zamecnik, M. L. Stephenson, *Proc. Acad. Sci. USA* **1978**, *75*, 280-284.
(c) P. C. Zamecnik, M. L. Stephenson, *Proc. Acad. Sci. USA* **1978**, *75*, 285-288.

[36] (a) K. S. Schmidt, M. Boudvillain, A. Schwartz, G. A. van der Marel, J. H. van Boom, J. Reedijk, B. Lippert, *Chem. Eur. J*. **2002**, *8*, 5566-5570.
(b) K. S Schmidt, D. V. Filippov, N. J. Meeuwenoord, G. A. van der Marel, J. H. van Boom, B. Lippert, J. Reedijk, *Angew. Chem. Int. Ed*. **2000**, *39*, 375-377.
(c) U. Berghoff, K. S. Schmidt, M. B. L. Janik, G. Schröder, B. Lippert, *Inorg. Chim. Acta* **1998**, *269*, 135-142.

[37] (a) J. Müller, M. Drumm, M. Boudvillain, M. Leng, E. Sletten, B. Lippert, *J. Biol. Inorg. Chem*. **2000**, *5*, 603-611.
(b) M. B. L. Janik, B. Lippert, *J. Biol. Inorg. Chem*. **1999**, *4*, 645-653.

[38] C. Colombier, B. Lippert, M. Leng, *Nucleic Acids Res*. **1996**, *24*, 4519-4524.

[39] (a) B. Rosenberg, L. VanCamp, J. E. Trosko, V. H. Mansour, *Nature* **1969**, *222*, 385-386.
(b) B. Rosenberg, L. VanCamp, T. Krigas, *Nature* **1965**, *205*, 698-699.
[40] H. C Harder, B. Rosenberg, *Int. J. Cancer* **1970**, *6*, 207-216.
[41] (a) J. J. Roberts, J. M. Pascoe, *Nature* **1972**, *235*, 282-284.
(b) P. J. Stone, A. D. Kelman, F. M. Sinex, *Nature* **1974**, *251*, 736-737.
(c) L. A. Zwelling, K. W. Kohn, W. E. Ross, R. A. G. Ewig, T. Anderson, *Cancer Res.* **1978**, *38*, 1762-1768.
[42] (a) A. M. J. Fichtinger-Schepman, J. L. van der Veer, J. H. J. den Hartog, P. H. M. Lohman, J. Reedijk, *Biochemistry* **1985**, *24*, 707-713.
(b) A. Eastman, *Pharmac. Ther.* **1987**, *34*, 155-166 (and refs. cited therein).
[43] (a) P. M. Takahara, A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, *Nature* **1995**, *377*, 649-652.
(b) P. M. Takahara, C. A. Frederick, S. J. Lippard, *J. Am. Chem. Soc.* **1996**, *118*, 12309-12321.
[44] (a) G. Schröder, J. Kozelka, M. Sabat, M.-H. Fouchet, R. Beyerle-Pfnür, B. Lippert, *Inorg. Chem.* **1996**, *35*, 1647-1652.
(b) G. Schröder, M. Sabat, I. Baxter, J. Kozelka, B. Lippert, *Inorg. Chem.* **1997**, *36*, 490-493.
[45] B. Lippert, *Coord. Chem. Rev.* **2000**, *200-202*, 487-516.
[46] M. A. Fuertes, C. Alonso, J. M. Pérez, *Chem. Rev.* **2003**, *103*, 645-662.
[47] (a) I. Ott, R. Gust, *Arch. Pharm. Chem. Life Sci.* **2007**, *340*, 117-126.
(b) H. T. Chifotides, K. R. Dunbar, *Acc. Chem. Res.* **2005**, *38*, 146-156.
[48] D. Sidransky, *Science* **1997**, *278*, 1054-1058
[49] (a) S. M. Cohen, *Curr. Optin. Chem. Biol.* **2007**, *11*, 115-120.
(b) G. Orvig, in *Concepts and Models in Bioinorganic Chemistry* (Eds. H.-B. Kraatz, N. Metzler-Nolte), Wiley-VCH, Weinheim, **2006**.

Chapter 2
Theory and Computation

This chapter serves to give the reader some more insight into the theory that has been used in this thesis. It discusses the key concepts of quantum mechanics. For a more elaborate description, the interested reader is referred to various text books.[1] Moreover, a more detailed overview will be given on the computational settings being used in this thesis.

2.1 The Schrödinger Equation

The central problem in theoretical chemistry is to solve the Schrödinger equation, which can be written for all molecular systems.[1,2] The equation itself is named after Erwin Schrödinger, who proposed it in 1926.[2] The time-dependent Schrödinger equation can be written as follows:

$$\hat{H}\Psi = i\hbar\frac{\partial\Psi}{\partial t} \qquad \text{(Eq. 2.1)}$$

In equation 2.1 are several symbols: $\hat{H}$ is the Hamiltonian (or Hamilton operator). An operator is indicated by putting a circumflex on the letter representing the operator, Ψ represents the wavefunction, i is a complex number, $\hbar$ is the Planck constant (h = 6.626*10^{-34} m^{2}kg / s) divided by 2π and t represents the time. In this thesis, the molecular properties that are investigated (energy and geometry) are not dependent on time and the Schrödinger equation can be rewritten in a time-independent form.

$$\hat{H}\Psi = E\Psi \qquad \text{(Eq. 2.2)}$$

This equations says that when the Hamiltonian acts on the wavefunction, the total energy (E) is obtained. Since the state of a system is fully described by Ψ, solving this equation leads to information about all molecular properties of the system. For a general chemical system $\hat{H}$ contains terms due to kinetic (T) and potential energy (V) and one may write:

$$\hat{H} = \hat{T} + \hat{V} = \hat{T}_N + \hat{T}_n + \hat{V}_{NN} + \hat{V}_{nn} + \hat{V}_{Nn} \qquad \text{(Eq. 2.3)}$$

Equation 2.3 describes kinetic-energy effects due to electrons ($\hat{T}_N$) and nuclei ($\hat{T}_n$), as well as the potential energy effects due to mutual repulsion of electrons ($\hat{V}_{NN}$) and nuclei ($\hat{V}_{nn}$), as well as electron–nuclear attraction ($\hat{V}_{Nn}$). However, nuclei are heavier than electrons, thus the former will have a smaller velocity than the latter. We can now, as an approximation, treat the electrons in the field of fixed nuclei. This approximation is known as the Born-Oppenheimer approximation. Thus, Eq. 2.3 can be rewritten as:

$$\hat{H}_{elec} = \hat{T}_N + \hat{V}_{NN} + \hat{V}_{Nn} \qquad \text{(Eq. 2.4)}$$

The electronic Schrödinger equation reads as:

$$\hat{H}_{elec}\Psi_{elec} = E_{elec}\Psi_{elec} \qquad \text{(Eq. 2.5)}$$

E is now a sum of the electronic (E_{elec}) and the nuclear energy (E_{nuc}) and the energy E_{elec} itself is calculated as an expectation value of the electronic Hamiltonian.

$$E = E_{elec} + E_{nuc} \qquad \text{(Eq. 2.6a)}$$

$$E_{elec} = \left\langle \Psi_{elec} \middle| \hat{H} \middle| \Psi_{elec} \right\rangle \qquad \text{(Eq. 2.6b)}$$

According to the variational theorem, the expectation value of the exact Hamiltonian H for a (normalised and quadratically integrable) trial wavefunction (Ψ_{trial}) is always greater than or equal to the exact groundstate energy. Therefore, one seeks in practice for approximations to the exact wavefunction that yield the lowest possible energy $\left\langle \Psi_{trial} \middle| \hat{H} \middle| \Psi_{trial} \right\rangle$.

2.2 Electronic Structure Calculations

It is a pity that an analytical solution of the Schrödinger equation is only known for the hydrogen atom, containing only one electron. The solutions are the well known atomic orbitals. However, one is more interested in larger molecular systems, having 2 or many more electrons. But what should a many electron or N-electron wavefunction look like? A good example of how such a wavefunction can be constructed is offered by the Hartree-Fock (HF) approximation. In this approximation the N-electron wavefunction (Ψ_{elec}) is written as an

anti-symmetrised product of one-electron functions, spin-orbitals that constitute the so-called Slater determinant (Φ_{SD}), which is used in Hartree-Fock theory.

$$\Phi_{SD} = \begin{vmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & \dots & \chi_3(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & \dots & \chi_3(\vec{x}_2) \\ \dots & \dots & \dots & \dots \\ \chi_N(\vec{x}_1) & \chi_N(\vec{x}_2) & \dots & \chi_N(\vec{x}_N) \end{vmatrix} \qquad \text{(Eq. 2.7)}$$

By definition, the Slater determinant obeys the Pauli principle which states that the wavefunction must be anti-symmetric, *i.e.*, it changes sign under permutation of two fermions, in our case electrons. The spin-orbitals consists of a spatial part ($\phi(\vec{r})$) and a spin part ($\sigma(s)$).

$$\chi(\vec{x}) = \phi(\vec{r})\sigma(s) \qquad \text{(Eq. 2.8)}$$

For $\sigma(s)$ there are two options, spin up (α) and spin down (β). It has to be noted that anti-symmetry leads to a well-known phenomenon, the Pauli principle: two electrons can be placed in the same molecular orbital, only if these have opposite spins. The phenomenon that electrons of same spin, and same spatial part avoid each other is called exchange correlation.

One may now minimise the energy, by variation of the spin-orbitals. The fact that one uses in Hartree-Fock theory a one-determinant wavefunction corresponds to employing a meanfield approximation to the electron-electron repulsion. In this approximation, the electron "feels" the presence of the other electrons through an average field. This means that the instantaneous Coulomb correlation is neglected. Normally the movements of electrons are correlated, as they repel each other. There are several techniques to include this kind of correlation, which are designated post-Hartree-Fock methods for instance second order perturbation theory due to Møller and Plesset (MP2) or coupled cluster (CC) methods.[3] Highly correlated *ab inito* methods can serve to produce reference data which can be used to evaluate the performance of the various approximate approaches in density functional theory (DFT). This can guide one to choose the optimal the functional for a given type of chemical problems. This will for instance be demonstrated in chapters 3 and 5.

2.3 Density Functional Theory (DFT)

The wavefunction itself is a mere mathematical description, and it may be more interesting to work with a more physical quantity that is also able to give the total energy E. This can indeed be done by using the electron density, ρ, and this property is the fundament of DFT.[4] In 1964 Hohenberg and Kohn proved two theorems.[5] First, E is uniquely determined by the electron density ρ, *i.e.*, the energy is a functional of the density (Eq. 2.9):

$$E = E[\rho] \qquad \text{(Eq. 2.9)}$$

Of course the first question is: what is a functional? The reader will be more familiar with the mathematical function: a number is put in a function and another number is put out. A functional is a bit more complicated. It assigns a number to a function. Pictorially speaking the working of a functional can be made clear through Equation 2.10:

$$f(x) \xrightarrow{F[f(x)]} y \qquad \text{(Eq. 2.10)}$$

In the same 1964 paper it was proven that the variational principle holds also for DFT. However, the theory seems very interesting, but the functional form is not known, but a year later, in 1965, Kohn and Sham invented a scheme to put the theory to work. In this scheme one starts from a non-interacting system, but this is the starting point to come to the results of the fully interacting system. In the end, one may write

$$E[\rho] = T_S[\rho] + J[\rho] + E_{Nn}[\rho] + E_{\mathrm{XC}}[\rho] \qquad \text{(Eq. 2.11)}$$

In this Equation $T_S[\rho]$ is the kinetic energy of the non-interacting system, $J[\rho]$ is the classical Coulomb interaction, $V_{Nn}[\rho]$ the electron-nuclei attraction term. The term that is the most difficult is $E_{\mathrm{XC}}[\rho]$ and accounts for exchange correlation, Coulomb correlation, self-interaction as well as the differences in kinetic energy between the non-interacting and the interacting system. However the last term is completely unknown and has to be modelled. In the past years several approximations have become available.

A very rudimentary way to model $E_{\mathrm{XC}}[\rho]$ is the Local Density Approximation (LDA). In the LDA, the functional that holds for a homogeneous electron gas is applied to molecular systems. $E_{\mathrm{XC}}[\rho]$ is then written according to Eq. 2.12:

$$E_{XC}^{LDA} = \int \rho(r)\varepsilon_{XC}(\rho(r))dr \qquad \text{(Eq. 2.12)}$$

Here, ε_{XC} is the exchange correlation energy per particle of a uniform electron gas of density $\rho(r)$. This entity can be decomposed into individual contributions of exchange and correlation.

$$\varepsilon_{XC}(\rho(r)) = \varepsilon_X(\rho(r)) + \varepsilon_C(\rho(r)) \qquad \text{(Eq. 2.13)}$$

The LDA approximation will in general lead to an overestimation of binding energies, or differently stated in too negative energies. It is not surprising that one seeks now for better approximations. This can be done by the generalised gradient approximation (GGA), which contains not only the density itself but also contributions of its gradient ($\nabla\rho(r)$). After all, a homogenous electron density was assumed, which intuitively is definitely not physical for the highly inhomogeneous molecular systems. Gradient corrections can be made to the exchange and correlation individually. Another solution is found in the use of hybrid functionals in which the exchange is calculated exactly from the HF approximation, and the correlation is approximated with DFT.

2.4 Computational Settings

In this section detailed information is given about the computational settings used in this thesis. However, because it is a lot of data, the most important settings are repeated in each chapter. The reader should be aware that all important aspects are described here.

2.4.1 Programs and Their Basis Sets

As major DFT programs the Amsterdam Density Functional (ADF)[7] was used alone or in combination with the program QUantum-regions Interconnected by Local Descriptions (QUILD). The latter is a wrapper around ADF used for its superior geometry optimiser based on adapted delocalized coordinates coordinates.[8] Adapted delocalized coordinates are expressed in terms of internal coordinates, *i.e.*, bond length, bond angle and dihedral angle. For the geometry optimisations using B3LYP, Gaussian was used because this was originally not possible in ADF. In these instances, ADF was used to do an energy calculation on a single geometry obtained with another functional, which is done in chapter 5. A very important difference between ADF and Gaussian are the basis sets, which are the set of

functions, the atomic orbitals, from which the wavefunction is built. ADF and Gaussian use two different types of basis sets.

For ADF, the molecular orbitals (MOs) were expanded in a large un-contracted set of Slater type atomic orbitals (STOs) containing diffuse functions: TZ2P (no Gaussian functions are involved).[10a] The basis set is of triple-ζ quality for all atoms and has been augmented with two sets of polarization functions, *i.e.* 2*p* and 3*d* on H, 3*d* and 4*f* on C, N, O, Li, Na, K, F and Cl, 4*d* and 4*f* on Br, and *5f* and *6p* on Pt. The 1*s* core shell of carbon, nitrogen, oxygen, lithium, sodium and fluorine, up to and including the 2*p* core shell of chlorine and potassium, up to and including the 3*p* of bromine and up to and including the 4*d* core shell of the Pt atom were treated by the frozen-core approximation.[10b] A frozen-core means that core electrons are not treated in a calculation, but only the valence ones, which will make the calculation faster. Moreover, core electrons are less influenced by bond formation, than valence electrons are. An auxiliary set of *s*, *p*, *d*, *f* and *g* STOs was used to fit the molecular density and to represent the Coulomb and exchange potentials accurately in each self-consistent field cycle.[10c] The meaning of triple-ζ quality quality is that three basis functions per nl-shell of an atom have been used. One could have used single-ζ and double-ζ quality, but this is less accurate. In Gaussian a basis set of comparable accuracy was used, namely the cc-pVTZ basis set.[10d,e] The similarity of the STO-based TZ2P basis set and the GTO-based basis set is demonstrated in chapter 3.

The use of polarisation functions is to let the atomic orbitals have some more flexibility to describe the wavefunction better. Polarisation is indicated mostly by adding a "P" in the acronym for the basis set, as can be seen from the name TZ2P. Diffusion functions, at last, are needed to treat weak interactions.

2.4.2 Geometry Optimisations and Some of its Aspects

Throughout this thesis different density functionals have been used. They can be grouped into generalised gradient approximation (GGA) and hybrid functionals. The following functionals will occur in chapters 1-4: BLYP, BP86 mPW, OPBE, PBE and PW91 which are at the DFT level of GGA, but also at the hybrid level, B3LYP. For all GGA functionals exchange is described by Slater's Xα potential[11a] and correlation is treated in the Vosko-Wilk-Nusair (VWN) parameterisation[11b] (this is not true for BLYP, see later). The non-local GGA corrections are made as follows: (a) BLYP: exchange corrections due to Becke88,[12a] correlation is computed entirely by the Lee-Yang-Parr (LYP) scheme,[12b]

(b) BP86: exchange corrections due to Becke88,[12a] correlation corrections due to Perdew86,[13] (c) mPW: exchange corrections due to Adamo and Barone,[14a] correlation corrections due to Perdew and Wang (PW91c),[14b] (d) OPBE: exchange corrections from Cohen and Handy (OPTX),[15a] correlation corrections from the Perdew-Burke-Ernzerhof scheme (PBEc),[15b] (e) PBE: exchange and correlation corrections both from the Perdew-Burke-Ernzerhof scheme,[15b] (f) PW91: exchange and correlation corrections from Perdew and Wang (PW91),[14b, 16] (g) B3LYP: a hybrid functional formed by mixing a portion of exact exchange, local Slater exchange and non-local Becke88 exchange corrections, and for the correlation part a mix of local VWN and non-local LYP correlation.[12b, 17a, 17b]

Most GGA functionals can not deal with dispersion,[4a] although for some of the molecular systems that are treated a proper description of π-stacking, especially in chapter 5, 6 and 7 is mandatory to obtain good results. Dispersion corrected schemes were developed by S. Grimme for BLYP and PBE, and are because of their correction named BLYP-D,[18] BP86-D and PBE-D.[18a] The correction is done by augmenting these with an empirical correction for long-range dispersion effects, described by a sum of damped interatomic potentials of the form C_6R^{-6} added to the usual DFT energy.[18] Within this correction it is chosen as such that that normal bonds, which are well below the Van der Waals distance, are not affected by the correction. Another way to take into account dispersion is done using the M06-2X functional, which has been constructed to recover accurate *ab initio* (MP2 and CCSD(T)) data on dispersion complexes. This directness implies that no correction term is added to the energy calculated with the M06-2X functional.[19]

For atoms like Pt, velocities at the nucleus may reach the speed of light ($c = 3.0*10^8$ m/s), which requires to treat relativistic effects. In chapter 4, scalar relativistic effects are included by the zeroth-order regular (ZORA) approximation, which is implemented in ADF.[20]

To save time, one may apply symmetry to a molecule. For instance, as will be made clear in subsequent chapters, base pairs adenine-thymine (AT) and guanine-cytosine (GC) can be optimised in planar arrangement, which is associated with the C_S point group. Other symmetries, such as C_{4h}, C_4 and S_4 have also been used, especially where quartet arrangements are concerned, see chapters 5 and 6.

At the end of a geometry optimisation the respective geometry will fulfil several criteria. One of the most important ones is the convergence of the geometry. This means, a geometry does not, to a used-defined criteria, changes in geometry. It is now interesting to check whether this converged structure is a minimum. *i.e.*, an equilibrium structure. This can be

verified by means of a frequency analysis.[21] One can judge from the number of imaginary frequencies whether a structure is a minimum or not. A structure is a minimum in case the number of these imaginary frequencies is 0 (zero). Frequency analyses have been performed in chapter 4, 5 and 6. The frequency values are given in cm^{-1}. If a frequency is imaginary, it is preceded by the imaginary number *i*.

Let's take now the example in which the binding energy of AT is calculated as done in chapter 3, and assume that the geometry is converged, and is a minimum. The calculation of the binding energy is done by calculating the energy difference between the AT base pair and its bases. However, an oddity occurs: for a given basis set, the number of basis functions on each of the individual bases adenine and thymine is less than the number of basis functions used on the combined base pair. This difference may lead to a non-physical lowering of the energy, and is ascribed to the so-called basis super position error. This error can be calculated by use of the counterpoise method by Boys and Bernardi.[22] In effect, the monomers are calculated in the larger basis set of the dimer.

To end with some explanation on nomenclature: a geometry optimisation using functional A and basis set B is denoted as A/B. Thus BP86/TZ2P is a geometry optimisation using the BP86-functional and using a TZ2P basis set. This notation can be extended as follows: A/B // C/D. This means that first a geometry at the C/D level of theory is performed, this optimised geometry is then submitted to a so-called single point calculation, in which the used geometry is not optimised anymore, and the energy is evaluated at the A/B level.

2.4.3 Solvation

Most experiments are done in (aqueous) solution and that is the main reason that solvent effects have been treated throughout this thesis. One may do this explicitly and implicitly.

In chapter 3 solvent effects are treated explicitly, and up to five water molecules are part of a geometry optimisation. Of course one can extend the number of water molecules at one's own judgement, but the number that is tractable in a quantum chemical way is surely bounded. However, in molecular dynamics based on classical force fields one may take into account several thousands of molecules.

Another way to take solvent effects into account is by doing this implicitly by simulating the effect of the solvent by a dielectric continuum, as done in chapters 4-7. This has been done using the Conductor Like Screening Model (COSMO), which is implemented in ADF.[23] The settings for uncharged atoms have been published in the past.[24] From the

respective chapters it will become clear that the inclusion of solvation will influence both energetical and geometrical data. In this thesis water, having a dielectic constant of 78.4, interacts with the solvent. It has been shown that COSMO BP86/TZ2P solvation *free* energies of methylated DNA agree within ca. 1 kcal/mol with the solvation free energies obtained with AMBER/TI by Miller and Kollman[25], see Table 2.1.

The radii used for the atoms can be found in Table 2.2. They are based on an empirical modification of atomic radii used in a force-field method by Allinger and co-workers.[26] In COSMO the energy terms are calculated first for a conducting medium and then scaled by a factor (ε-1)/ε. As the solute is immersed in the solvent, a cavity has to be formed. The surface of this cavity is defined by following the path traced by a spherical solvent molecule rolling over the Van der Waals surface of the solvent molecule. One may think that concerning DFT-D the dispersion correction should be modified with respect to the gas phase, but this is not the case.[27]

Table 2.1. Solvation energies, ΔG_{Solv} and ΔG_{Solv} (in kcal/mol).

	ΔG_{Solv} [a]	ΔG_{Solv} [b]
9-methyladenine	–12.4	–12.0
9-methylguanine	–22.2	–22.4
1-methylcytosine	–18.6	–18.4
1-methylthymine	–11.7	–12.4
1-methyluracil	–12.7	–14.0

[a] Data taken from ref. 24.
[b] Data taken from ref. 25.

Table 2.2. Atomic radii (R, in Å) for several atoms.

Atom	R (in Å)[a]
H	1.350
C	1.700
N	1.608
O	1.517

[a] See ref. 26.

In Chapter 5 and onward dispersion corrected functionals are used which feature effective atomic radii as parameters that have been optimised for neutral atoms.[26] However, for the bare ions new parameters were derived based on experimental data on solvation free energies; they are presented in Table 2.3.[28] These new parameters are used in chapter 6 and 7.

Table 2.3. Experimental, ΔG(exp.) and calculated, ΔG(calcd.) hydration free energies (in kcal/mol) of several ions. Experimental data is given to compare with. The ionic radii (R in Å) are given as well. Calculations were done at the BLYP-D/TZ2P level of theory.

	ΔG(exp.)	ΔG(calcd.)	R (in Å)
Li^+	–118.1[a]	–118.7	1.364
Na^+	–90.6[a]	–90.5	1.775
K^+	–73.1[a]	–73.6	2.200
F^-	–110.7[a]	–103.1	1.250
Cl^-	–81.4[a]	–74.8	1.794
Br^-	–76.1[a]	–68.0	1.925

[a] See ref. 28.

2.5 References

[1] (a) R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press, London, **1992**.
(b) F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chicester, **1999**.
(c) Ch. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, Wiley, Chichester, **2002**.
(d) P. W. Atkins, R. S. Friedman, *Quantum Mechanics*, Oxford University Press, Oxford, **1999**.

[2] E. Schrödinger, *Phys. Rev.* **1926**, *28*, 1049-1070.

[3] R. J. Bartlett, J. F. Stanton, *Rev. Comput. Chem.* **1995**, *5*, 65-169.

[4] (a) W. Koch, M. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, Weinheim, **2001**.
(b) F. M. Bickelhaupt, E. J. Baerends, In: *Reviews in Computational Chemistry*; K. B. Lipkowitz, D. B. Boyd, Eds.; Wiley-VCH: New York, **2000**, Vol. 15, pp. 1-86.

[5] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864-B871.

[6] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133-A1138.

[7] (a) G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonseca Guerra, E. J. Baerends, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931-967.
(b) ADF, Scientific Computing & Modeling, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands. E.J. Baerends, J. Autschbach, A. Bérces, F.M. Bickelhaupt, C. Bo, P.M. Boerrigter, L. Cavallo, D.P. Chong, L. Deng, R.M. Dickson, D.E. Ellis, M. van Faassen, L. Fan, T.H. Fischer, C. Fonseca Guerra, S.J.A. van Gisbergen, A.W. Götz, J.A. Groeneveld, O.V. Gritsenko, M. Grüning, F.E. Harris, P. van den Hoek, C.R. Jacob, H. Jacobsen, L. Jensen, G. van Kessel, F. Kootstra, M.V. Krykunov, E. van Lenthe, D.A. McCormack, A. Michalak, J. Neugebauer, V.P. Nicu, V.P. Osinga, S. Patchkovskii, P.H.T. Philipsen, D. Post, C.C. Pye, W. Ravenek, J.I. Rodríguez, P. Ros, P.R.T. Schipper, G. Schreckenbach, J.G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T.A. Wesolowski, E.M. van Wezenbeek, G. Wiesenekker, S.K. Wolff, T.K. Woo, A.L. Yakovlev, and T. Ziegler

[8] a) M. Swart, F. M. Bickelhaupt, *Int. J. Quantum Chem.* **2006**, *106*, 2536-2544.
b) M. Swart, F. M. Bickelhaupt, *J. Comput. Chem.* **2008**, *29*, 724-734.

[9] Gaussian, Gaussian, Inc., Wallingford CT. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople.

[10] (a) J. G. Snijders, P. Vernooijs, E. J. Baerends, *At. Nucl. Data Tables* **1981**, *26*, 483-509.

(b) E. J. Baerends, D. E. Ellis, P. Ros, *Chem. Phys.* **1973**, *2*, 41-51.
(c) J. Krijn, E. J. Baerends, *Fit-Functions in the HFS-Method; Internal Report* (in Dutch), Vrije Universiteit, Amsterdam, **1984**.
(d) T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007-1023.
(e) D. E. Woon, T. H. Dunning, *J. Chem. Phys.* **1993**, *98*, 1358-1371.

[11] (a) J. C. Slater, *Quantum Theory of Molecules and Solids, Vol. 4*, McGraw-Hill, New York, **1974**.
(b) S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200-1211.

[12] (a) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098-3100.
(b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.

[13] J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822-8824.
(Erratum: *Phys. Rev. B* **1986**, *34*, 7406).

[14] (a) C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664-675.
(b) J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, C. Fiolhais, *Phys. Rev. B* **1992**, *46*, 6671-6687.

[15] (a) N. C. Handy, A.J. Cohen, *Mol. Phys.* **2001**, *99*, 403-412.
(b) J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865-3868.

[16] J. P. Perdew in *Electronic Structure of Solids* (Eds.: P. Ziesche, H. Eschrig) Akademie Verlag Berlin, **1991**, pp. 11.

[17] (a) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.
(b) P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *45*, 11623-11627.

[18] (a) S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463-1473.
(b) S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787-1799.
(c) J. Antony, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287-5293.
(d) M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, E. Kaxiras, *J. Chem. Phys.* **2001**, *114*, 5149-5155.
(e) X. Wu, M. C. Vargas, S. Nayak, V. Lotrich, G. Scoles, *J. Chem. Phys.* **2001**, *115*, 8748-8757.
(f) Q. Wu, W. Yang, *J. Chem. Phys.* **2002**, *116*, 515-524.
g) U. Zimmerli, M. Parrinello, P. Koumoutsakos, *J. Chem. Phys.* **2004**, *120*, 2693-2699.

[19] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215-241.

[20] E. van Lenthe, E. J. Baerends, J. G. Snijders, *J. Chem. Phys.* **1994**, *101*, 9783-9792.

[21] (a) A. Bérces, R. M. Dickson, L. Fan, H. Jacobsen, D. Swerhone, T. Ziegler, *Comput. Phys. Commun.* **1997**, *100*, 247-262.
(b) H. Jacobsen, A. Bérces, D. Swerhone, T. Ziegler, *Comput. Phys. Commun.* **1997**, *100*, 263-276.
(c) S. K. Wolff, *Int. J. Quantum Chem.* **2005**, *104*, 645-659.

[22] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, 19, 553-566.

[23] (a) A. Klamt, G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, 799-805.
(b) A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224-2235.
(c) C. C. Pye, T. Ziegler, *Theor. Chem. Acc.* **1999**, *101*, 396-408.

[24] M. Swart, E. Rösler, F. M. Bickelhaupt, *Eur. J. Inorg. Chem.* **2007**, 3646-3654.

[25] J. L. Miller, P. A. Kollman, *J. Phys. Chem.* **1996**, *100*, 8587-8594.

[26] N. L. Allinger, X. F. Zhou, J. Bergsma, *J. Mol. Struct. (THEOCHEM)* **1994**, *118*, 69-83.

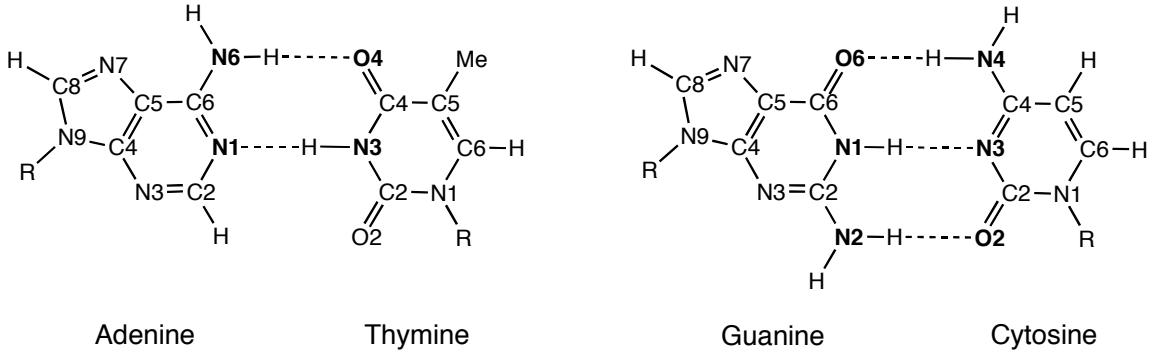
[27] K. E. Riley, J. Vondrasek, P. Hobza, *Phys. Chem. Chem. Phys.* **2007**, *9*, 5555-5560.

[28] R. Gomer, G. Tryson, *J. Chem. Phys.* **1977**, *66*, 4413-4424.

Chapter 3
Performance of Various Density Functionals for the Hydrogen Bonds in the DNA Base Pairs AT and GC

3.1 Introduction

Although the hydrogen bond is a weak interaction, it nevertheless plays an important role in biological systems.[1] As an example its role in the working of the genetic code should be mentioned. In B-DNA the double helical chains of nucleotides are stabilised through the hydrogen bonds that selectively occur between the base pairs adenine-thymine (AT) and guanine-cytosine (GC) as proposed by Watson and Crick in 1953 (see Scheme 3.1).[2]



Scheme 3.1. The Watson-Crick base pairs AT and GC. Heavy atoms participating in hydrogen bonding have been indicated in bold.

Many theoretical investigations were devoted to these hydrogen bonds in DNA at different levels of theory. MP2 is a very good method, but very time consuming. As a compromise, DFT is an excellent workhorse for treating hydrogen bonds. A very popular functional is B3LYP and is widely used for different chemical problems, and is a good alternative for MP2. This does not mean at all that B3LYP is a universal functional, for instance the functional OLYP performs much better when reaction-barriers are concerned,[3] and in the worst case B3LYP does not reproduce the reaction barrier. Another feature concerns a strange observation in the gas phase calculation on DNA base pairs.[4] Although the calculated hydrogen bond lengths of the base pair AT could be very well compared to crystallographic data,[5a] this comparison failed for the base pair GC.[5b]

Other studies showed that the functional BP86 gave base pair geometries of AT and GC that deviated much from the experimental data. Interestingly, this discrepancy could be taken

away by modelling the first-shell crystal environment.[6] Very interesting results were obtained for the base pair GC by going from the gas phase to the modelled crystal environment. The hydrogen bond length found in the upper (O6–N4), middle (N1–N3) and lower (N2–O2) hydrogen bond changed from short-long-long, respectively, for the gas phase (no water, no counter ions) to long-long-short, respectively, with inclusion of the first shell crystal environment. This latter pattern matches the one found in the X-ray structures measured for the crystal of sodium guanylyl-3',5'-cytidine nonahydrate.[5] The molecular environment consists of crystal water, ribose OH-groups of neighbouring base pairs (modelled also by water in the BP86 computations) and sodium (Na^+) ions.

The results that are presented in this chapter serve two different goals. First, it is investigated whether the geometrical changes due to the crystal-environment are an artefact of the method (*i. e.* BP86), or whether they are physically meaningful. Second, an evaluation of the performance of the different functionals for reproducing both gas phase and the above condensed phase Watson-Crick hydrogen bond structures is made.[6] As has already been mentioned in chapter 2, good benchmark results are needed for comparison with the DFT results. In this study we make, for the gas phase, use of very accurate *ab initio* MP2 data obtained recently by Sponer *et al.*,[7a] and coupled cluster (CC) energies obtained by Jurecka *et al.*.[7b] By comparing the MP2 and CC results with the results obtained with the different functionals, an indication is obtained to which extent a given functional is adequate for the description of the DNA base pairs AT and GC. It is expected that the different effects induced by the crystal environment are reproduced by all functionals. However, it is expected that B3LYP is not an appropriate functional to describe the DNA base pairs.

3.2. Summary of Computational Methods

Calculations were done with two different programs: an adapted version (QUILD) of the Amsterdam Density Functional Program (ADF), which uses Slater-type orbital (STO) basis sets, and the Gaussian03 program, which uses Gaussian-type orbital (GTO) basis sets. In this study an STO TZ2P basis set and a GTO cc-pVTZ were used. Geometries were optimised at the BLYP, BP86, mPW, OPBE, PBE and PW91 level of the generalised gradient approximation (GGA), but also at the hybrid level, B3LYP.

The hydrogen bond energy (ΔE_{HB}) is defined as follows:

$$\Delta E_{HB} = E_{\text{Base pair}} - E_{\text{Base1}} - E_{\text{Base2}} \qquad \text{(Eq. 3.1)}$$

The terms on the right hand side of Eq. 3.1 are the energies of the base pair ($E_{\text{Base pair}}$) and its constituting bases, base 1 and base 2 (E_{Base1} and E_{Base2}). The base pairs were optimised in C_S symmetry, its constituting bases in C_1 symmetry, to account for the pyramidalisation of the, if present, amino group.[8] Due to this pyramidalisation, a non-planar conformation of the amino group is obtained. The hydrogen bond energy is corrected for the basis superposition error (BSSE), through the counterpoise method using the DNA bases as fragments. The interested reader is referred to chapter 2 for more information about the theoretical aspects.

3.3 Results and Discussion

3.3.1 Base Pairs AT and GC in the Gas Phase

In Table 3.1 and 3.2 the hydrogen bond distances and their energy of the Watson-Crick base pairs AT and GC are given, respectively, along with the benchmark data of Sponer *et al.*,[7a] and Jurecka *et al.*[7b] Due to the different computational programs (ADF and Gaussian03) we first address a technical aspect concerning the basis sets. Previous studies showed a good performance of the STO TZ2P basis set.[6] One needs a GTO set of similar accuracy, which is obtained with the cc-pVTZ GTO basis set, as will be demonstrated. It can be taken from Table 3.1 and 3.2 that differences in hydrogen bond length are typically 0.02 Å, whilst the hydrogen bond energy differs by about a 0.5 kcal/mol. For instance the N6-O4 hydrogen bond distance at BP86/TZ2P is 2.92 Å and at BP86/cc-pVTZ 2.93 Å and the corresponding bonding energies are –11.0 and –11.6 kcal/mol, respectively.

To compare the energetical data one should also correct for the BSSE, because different values may be obtained for the STO and GTO basis set. Table 3.1 shows that this error is about 1 kcal/mol or less for the TZ2P STO basis set, but up to 2 kcal/mol for cc-pVTZ GTO basis set. By correcting for the BSSE one can compare, for instance, the B3LYP/cc-pVTZ data with the data of other functionals at the TZ2P level.

The results are now discussed in more detail. Eye-catching is how the gas phase results for the base pair AT obtained with B3LYP compare very well with the experimental hydrogen bond lengths in the crystal structure. In the gas phase B3LYP obtains hydrogen bond lengths for 2.93 and 2.88 Å (Table 3.1) for the upper and lower hydrogen bond,

respectively which is compared to the experimental values of 2.93 and 2.85 Å (Table 3.3), again, respectively.[5a] It can however be seen that this result emerges from coincidence and is, therefore, spurious as already found in previous results.[4] At variance, the results for the base pair GC for the gas phase are not in agreement with experimental results as judged from their hydrogen bonding pattern.[5b] The hydrogen bond pattern for the O6–N4, N1–N3 and N2–O2 hydrogen bond lengths is short-long-long for the gas phase and, on contrary, long-long-short in experiment, all respectively.

Table 3.1. Hydrogen bond distances (in Å) and bond energy, ΔE_{HB} (in kcal/mol), for AT computed at various levels of theory.

Method	N6–O4	N1–N3	MAD[a]	ΔE_{HB}	BSSE	ΔE_{HB}[b]
Ab initio data						
RI-MP2/aug-cc-pVQZ// RI-MP2/cc-pVTZ[c]	2.86	2.83				–15.1
"CCSD(T)/aug-cc-pVQZ"// RI-MP2/cc-pVTZ[c]	2.86	2.83				–15.4
STO basis: TZ2P						
BLYP	2.92	2.88	0.055	–11.6	0.6	–11.0
BP86	2.85	2.81	0.015	–13.0	0.7	–12.3
mPW	2.87	2.83	0.005	–12.7	0.7	–12.0
OPBE	3.00	2.93	0.120	–6.9	0.9	–6.0
PBE	2.87	2.80	0.020	–14.6	0.7	–13.9
PW91	2.85	2.79	0.025	–15.2	0.7	–14.5
GTO basis: cc-pVTZ						
B3LYP	2.93	2.88	0.060	–13.0	1.4	–11.6
BLYP	2.94	2.90	0.075	–12.5	1.7	–10.8
BP86	2.87	2.83	0.005	–13.7	1.5	–12.2

[a] Mean absolute deviation (MAD) in computed N6–O4 and N1–N3 distances between DFT and MP2.
[b] Bond energy with inclusion of BSSE correction, using the counterpoise method.
[c] Data from ref. 7a and 7b. The coupled-cluster energy has been obtained by adding a correction to the MP2 energies. This correction is calculated as a difference between the coupled-cluster energy and the MP2 energy obtained with smaller basis sets as explained in ref. 7a.

One may think that the above mentioned hydrogen bond length pattern difference is ascribed to the computational method (*i.e.*, DFT). This is however not true if one compares the various DFT results with the *ab initio* MP2/aug-cc-pVQZ results of Sponer and *et al.*,[7a] as seen from Tables 3.1 and 3.2. It appears that DFT compares much better to the MP2

results than to the experimental values, and the MP2 results show for GC the same, wrong, hydrogen bond pattern: short-long-long. Next, four density functionals show excellent agreement with MP2. These are BP86, mPW, PBE and PW91 and their mean absolute deviation (MAD) values are below 0.027Å (Table 3.2).

Table 3.2. Hydrogen bond distances (in Å) and bond energy, ΔE_{HB} (in kcal/mol), for GC computed at various levels of theory.

Method	O6-N4	N1-N3	N2-O2	MAD[a]	ΔE_{HB}	BSSE	ΔE_{HB}[b]
***Ab initio* data**							
RI-MP2/aug-cc-pVQZ// RI-MP2/cc-pVTZ[c]	2.75	2.90	2.89				–27.7
"CCSD(T)/aug-cc-pVQZ"// RI-MP2/cc-pVTZ[c]	2.75	2.90	2.89				–28.8
STO basis: TZ2P							
BLYP	2.79	2.94	2.93	0.040	–23.9	0.7	–23.2
BP86	2.73	2.88	2.87	0.020	–26.1	0.9	–25.2
mPW	2.74	2.90	2.89	0.003	–25.4	0.9	–24.5
OPBE	2.80	2.97	2.98	0.070	–17.1	1.1	–16.0
PBE	2.73	2.89	2.87	0.017	–27.8	0.9	–26.9
PW91	2.72	2.88	2.86	0.027	–28.5	0.9	–27.6
GTO basis: cc-pVTZ							
B3LYP	2.79	2.94	2.93	0.040	–26.1	1.7	–24.4
BLYP	2.80	2.96	2.96	0.060	–24.6	2.0	–22.6
BP86	2.73	2.90	2.89	0.007	–26.6	1.8	–24.8

[a] Mean absolute deviation (MAD) in computed O6–N4, N1–N3 and N2–O2 distances between DFT and MP2.
[b] Bond energy with inclusion of BSSE correction.
[c] Data from ref. 7 and 7b. The coupled-cluster energy has been obtained by adding a correction to the MP2 energies. This correction is calculated as a difference between the coupled-cluster energy and the MP2 energy obtained with smaller basis sets as explained in ref. 7a.

Summarising from Tables 3.1 and 3.2, one can compare the PW91 results of –14.5 and –27.6 kcal/mol for AT and GC to the corresponding MP2 results of –15.1 and –27.7 kcal/mol along with the CC-values of –15.1 and –28.2 kcal/mol. It should be mentioned that the difference between MP2 and CC is 0.5 kcal/mol for GC. BLYP and B3LYP underestimate the hydrogen bond energy by 3 to 5 kcal/mol and as a result thereof hydrogen bond lengths are longer than the ones at the MP2 level. The finding that PW91 outperforms B3LYP is not new as confirmed by other studies.[7a,9] A functional that performs quite bad is OPBE as it

underestimates the hydrogen bond energy severely and, as a result, gives hydrogen bond lengths that are too long.

To this end no comparison has been made with the complexation enthalpies, which are available for methylated base pairs AT and GC which amount to –13.0 and –21.0 kcal/mol, respectively.[10] These values are, however not useful to compare with DFT data, because the experimental enthalpies were calculated from the temperature dependence of the equilibrium constant at 323 and 381 K. Morever Kabelác and Hobza showed, based on molecular dynamics, that at elevated temperatures that the Watson-Crick hydrogen bonded base pairs are negligibly populated,[11] and that instead under these circumstances π-stacked species are predominant. In chapter π-stacked species are not dealt with, but this will be the case from chapter 5 on.

From the above discussion it is concluded that BP86 along with PW91 agree to large extent with the *ab initio* MP2 benchmark data of Sponer on Watson-Crick base pairs as derived from the comparison between their hydrogen bond distances as well as their binding energies. Therefore, both functionals are recommended to be used in case hydrogen bonding in biological systems is concerned.

3.3.2 AT and GC with Modeled Crystal-Environment

An aspect that is yet to be solved, is the difference between the calculated and the experimental geometry of the Watson-Crick base pairs. Is has already been mentioned that this difference is not an artefact of DFT as judged from its agreement with the benchmark data. Previously, it has been shown that the geometries agree better, at least for BP86,[6] when one takes into account the molecular environment that the base pairs face in their crystals of sodium adenylyl-3',5'-uridine hexahydrate (**1**),[5a] and sodium guanylyl-3',5'-cytidine nonahydrate (**2**),[5b] into the theoretical model systems.

The question that has to be answered is whether this trend is also found with other functionals. Therefore, two model systems per base pair were chosen from previous studies that resemble the full crystal environment at best.[6] These model systems are shown in Figure 3.1 and these are **1a** and **1b** for AT and **2a** and **2b** GC. The models comprise a sodium cation (Na^+) and up to five water molecules (in case of GC) that model the crystal water or hydroxy molecules of neighbouring ribose units. The sugar-phosphate backbone has been left out, as its inclusion does not have a large effect on the base pair hydrogen bonds. This is confirmed in a previous study,[6b] in which the backbone at the N1 of the pyrimidines and N9 of the

purine was modelled consecutively by a methyl group, sugar group and phosphate-sugar group. It was found that the effect of the backbone was in the order of 0.01 Å, which is negligible.

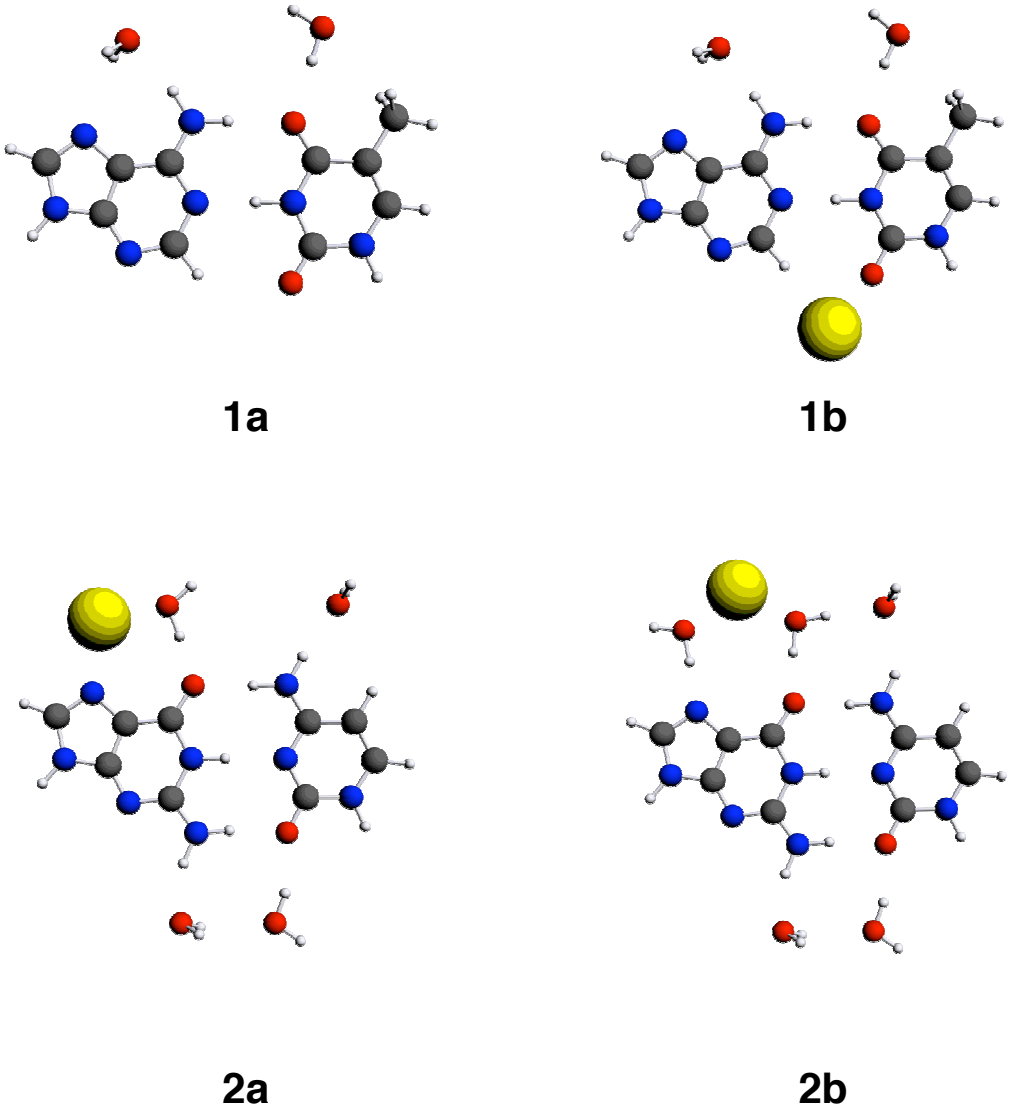


Figure 3.1. Model systems of AT and GC with water molecules and counter ions (Na^+) that simulate the crystal environment: **1a** = AT with two water molecules, **1b** = AT with two water molecules and sodium cation, **2a** = GC with four water molecules and a sodium cation, and **2b** = GC with five water molecules and a sodium cation. The yellow ball represents the Na^+ ion.

Geometry optimisations for **1a**, **1b**, **2a** and **2b** were performed for BLYP, BP86 and PW91 in combination with a TZ2P basis set, and for B3LYP and BP86 in a cc-pVTZ basis

set. The results of these optimisations are shown in Table 3.3 and comparison is made with experimental data. First, an aspect on the experimental data has to be discussed. For AT there are two different values for the hydrogen bond length because the two AU pairs in $(ApU)_2$ have different environments. The first value of N4–O6 and N1–N3 refers to the first pair and the second value of N4–O6 and N1–N3 refers to the second pair. Now, again, a technical aspect of the use of different basis sets in Gaussian and ADF is addressed based on the results obtained with BP86 for which both sets were used. The difference in hydrogen bond lengths obtained at either BP86/cc-pVTZ or BP86/TZ2P is in the order of 0.02 Å, which is quite small and comparable to the gas phase situation, which was discussed in the previous section. As the difference is only small it is justified to compare the B3LYP/cc-pVTZ results with the other results obtained at the TZ2P level. The central result of these computations is that the same effect is seen when going from the gas phase (Tables 3.1 and 3.2) to the situation in which the crystal environment is modelled (Table 3.3). In the base pair AT the upper hydrogen bond (N6–O4) becomes longer by about 0.1 Å, the lower hydrogen bond (N1–N3) does not change in length and remains effectively the same. In the base pair GC the upper (O6–N4) and middle hydrogen bond (N1–N3) become longer by about 0.1 Å, and the lower one (N2–O2) does not change effectively. As the effect that is induced by the modelled crystal environment on the base pairs is found for all functionals, it strongly advocates that these effects are brought about by physically meaningful events and are not an artefact of the specific method BP86/TZ2P.[6] Interestingly, for the GC pair, due to the inclusion of the crystal environment, the experimental order long-long-short is recovered with all functionals for **2a**, in **2b** this becomes short-long-short. Thus, it also depends on the modelling of the crystal environment to which extent the calculation will reproduce the experimental pattern.

For BP86 the effects induced by the modelled crystal environment lead to results that can be very well compared with the experimental results. This is best represented by the MAD values: for the systems **1b** and **2a** this value amount with respect to the experiment to less than 0.035 Å (see Table 3.3). This should be contrasted to the MAD values of the hydrogen bond lengths of the gas phase base pairs AT and GC with respect to the experiments and amounts to a maximum value of 0.086 Å (not shown in the tables). Again PW91 shows a good performance that is yet a bit worse than BP86. The MAD of the hydrogen bond lengths in **1a** and **2b** with respect to the experimental is less than 0.040 Å (see Table 3.3). Systems with modelled crystal environment the functionals B3LYP and BLYP give again hydrogen

bond lengths that are too long. The MAD values with respect to the experimental results are 0.043 and 0.060 Å for BLYP and B3YP, respectively.

Table 3.3. Hydrogen bond lengths (in Å) for model systems of AT (**1a**, **1b**) and GC (**2a**, **2b**) in the crystal environment computed with various DFT methods.

System	H-bond	Exp[a,b]	**GTO:** cc-pVTZ		**STO:** TZ2P		
			B3LYP	BP86	BLYP	BP86	PW91
1a	N4–O6	2.93	3.00	2.93	2.98	2.92	2.91
		2.95					
	N1–N3	2.85	2.88	2.82	2.87	2.80	2.79
		2.82					
	MAD[c,d]		0.050	0.015	0.035	0.030	0.040
			0.055	0.010	0.040	0.025	0.035
1b	N4–O6		3.01	2.96	2.98	2.93	2.91
	N1–N3		2.88	2.81	2.87	2.79	2.79
	MAD[c,d]		0.055	0.035	0.035	0.030	0.040
			0.060	0.010	0.040	0.025	0.035
2a	O6–N4	2.91	3.00	2.93	2.97	2.91	2.89
	N1–N3	2.95	3.00	2.95	2.99	2.94	2.93
	N2–O2	2.86	2.88	2.85	2.88	2.83	2.83
	MAD[c]		0.053	0.010	0.040	0.013	0.023
2b	O6–N4		2.93	2.87	2.93	2.86[e]	2.85
	N1–N3		3.01	2.97	3.01	2.95[e]	2.94
	N2–O2		2.93	2.89	2.91	2.86[e]	2.86
	MAD[c]		0.050	0.030	0.043	0.017	0.023

[a] X-ray crystallographic data taken from ref. 5a.
[b] X-ray crystallographic data taken from ref. 5b.
[c] Mean absolute deviation (MAD) between the calculated and the experimental hydrogen bond distances for each system.
[d] First and second MAD values refer to deviation with respect to the first and second AU pair in $(ApU)_2$, respectively.
[e] Geometry converged to $2.6*10^{-5}$ Hartree/Bohr.

It is concluded that BP86 and also PW91 agree very well with the experimental measured hydrogen bond distances. The fine agreement that is achieved by the B3LYP level between the geometry of the gas-phase AT pair and the crystallographic structure of this base pair (*i.e.*, sodium adenylyl-3',5'-uridine hexahydrate (**1**)),[5a] is the result of the a cancellation between a physical effect, the N6–O4 bond in the crystal **1** is longer than in the isolated gas-phase AT, and an error in B3LYP, which leads to an overestimation of the N6–O4 hydrogen bond and an underestimation of the hydrogen-bond strength in AT, and in general.

3.4 Conclusions

The performance of seven popular density functionals (B3LYP, BLYP, BP86, mPW, OPBE, PBE, PW91) on the hydrogen bond lengths and their energy of the Watson-Crick base pairs AT and GC is evaluated. The gas phase results were compared to very accurate *ab initio* results (MP2). To compare the results with the experiment the first crystal-environment shell was taken into account that consisted of several water molecules and a counter ion. The effects that were induced by the modelled crystal environment were observed for all DFT results. The fucntionals BP86 and PW91 are furthermore able to reproduce to high extent the good *ab initio* data as well as the experimental values. B3LYP consistently underestimates hydrogen bond strengths for both AT and GC and it gives significantly too long hydrogen-bond distances for the gas phase as well as for the systems with modelled crystal environment.

3.5 References

[1] (a) G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Systems*, Springer-Verlag, Berlin, New-York, Heidelberg, **1991**.
(b) G. A. Jeffery, *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, **1997** (Chapter 10).

[2] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737-738.

[3] (a) M. Swart, M. Solà, F. M. Bickelhaupt, *J. Comput. Chem.* **2007**, *28*, 1551-1560.
(b) G. Th. de Jong, F. M. Bickelhaupt, *J. Chem. Theory Comput.* **2006**, *2*, 322-335.
(c) G. Th. de Jong, D. P. Geerke, A. Diefenbach, M. Solà, F. M. Bickelhaupt, *J. Comput. Chem.* **2005**, *26*, 1006-1020.
(d) A. P. Bento, M. Solà, F. M. Bickelhaupt, *J. Comput. Chem.* **2005**, *26*, 1497-1504.
(e) J. Baker, P. Pulay, *J. Chem. Phys.* **2002**, *117*, 1441-1449.
(f) T. P. M. Goumans, A. W. Ehlers, K. Lammertsma, E. U. Würthwein, S. Grimme, *Chem. Eur. J.* **2004**, 10, 6468-6475.

[4] J. Bertran, A. Oliva, L. Rodríguez-Santiago, M. Sodupe, *J. Am. Chem. Soc.* **1998**, *120*, 8159-8167.

[5] (a) N. C. Seeman, J. M. Rosenberg, F. L. Suddath, J. J. P. Kim, A. Rich, *J. Mol. Biol.* **1976**, *104*, 109-144.
(b) J. M. Rosenberg, N. C. Seeman, R. O. Day, A. Rich, *J. Mol. Biol.* **1976**, *104*, 145-167.

[6] a) C. Fonseca Guerra, F. M. Bickelhaupt, *Angew. Chem.* **1999**, *111* 3120-3122; *Angew. Chem. Int. Ed.* **1999**, *38*, 2942-2945.
b) C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *J. Am. Chem. Soc.* **2000**, *122*, 4117-4128.

[7] (a) J. Sponer, P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2004**, *126*, 10142-10151.
(b) P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2003**, *125*, 150608-10893.

[8] (a) C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *Chem. Eur. J.* **1999**, *5*, 3581-3594.
(b) J. Sponer, P. Hobza, *J. Phys. Chem.* **1994**, *98*, 3161-3164.

[9] S. Tsuzuki, H. P. Lüthi, *J. Chem. Phys.* **2001**, *114*, 3949-3957.

[10] I. K. Yanson, A. B. Teplitsky, L. F. Sukhodub, *Biopolymers* **1979**, *18*, 1149-1170.

[11] M. Kabelác, P. Hobza, *J. Phys. Chem. B* **2001**, *105*, 5804-5817.

Chapter 4
Metal-Stabilised Tautomers of 1-Methyluracil and 1-Methylthymine

4.1 Introduction

Tautomer equilibria, those of heterocycles included, are influenced by factors such as temperature and bulk solvent properties (polarity and dielectric constant).[1] Trivial to state, *any* chemical modification may change the tautomer equilibrium. Due to the introduction of computational chemistry it has been possible to improve the knowledge about the dependence of tautomer equilibria on factors such as solvent,[2a-g] and metal binding.[2h-n] Important for nucleobases in this respect are the amino-imino and keto-enol tautomer equilibria. Amino and keto species are more present than the rare imino and enol species.

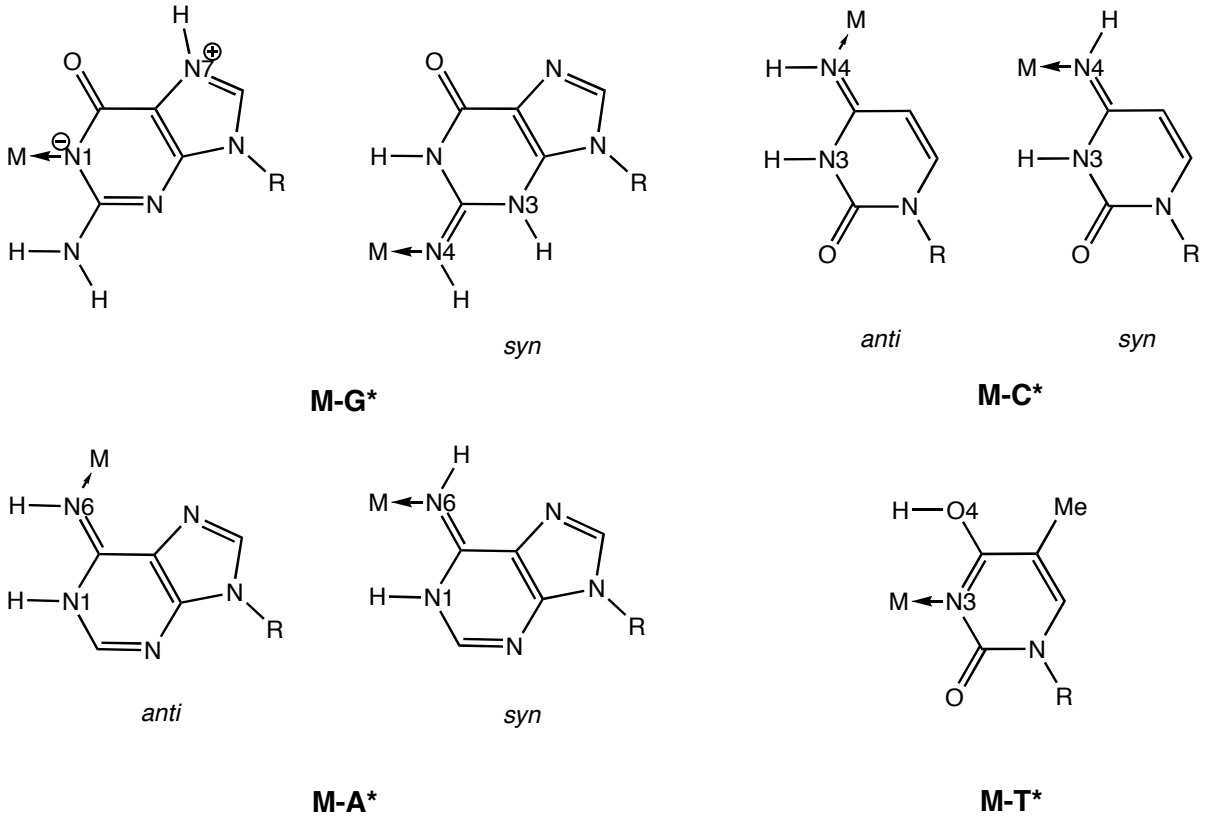
Shifts in tautomer equilibria towards the rarer form are important with regard to nucleobases, the constituents of DNA. Shifts in which a rare form is stabilised will in the end lead, if not repaired in an early stage, to a mutation.[3] Mutations may also result from more drastic changes in DNA composition such as hydrolytic or oxidative damage of nucleobases, deletion of bases, or addition of new ones. These pathways are not the focus of this chapter, however.

In the past there was a large interest in the possible roles of metal ions in mutagenic events. For this reason, scenarios were designed that could rationalise metal-mediated processes of nucleobases.[4] These can include effects on metal coordination to DNA (*e.g.* DNA kinking), strand cross-linking by metal ions, or changes in electronic structure of nucleobases with effects on acid-base equilibria, base-pairing patterns, and base-pair strength.[5]

A tautomeric change can be achieved via two pathways. The first one concerns metal binding to a site that does not participate in Watson-Crick hydrogen bonding. This situation occurs for adenine (A) in which a $[Pt(NH_3)_3]^{2+}$ fragment is bonded to N7 and in this sense through hydrogen bonding can stabilise a rare imino-form (A*), as demonstrated by a density functional theory (DFT) calculation.[6] The addition of an asterisk (*) to the name of the nucleobase means that this nucleobase is in a rare tautomer form. The stabilisation of a rare tautomer of guanine (G*) by a neutral Pt^{II} species considered unlikely, again based on DFT calculations.[7] Experimentally, efforts made to determine a shift in tautomeric equilibrium in

water for adenine were not sufficiently indicative to conclude a rare tautomer was present indeed.[8]

The second one concerns metal binding to a site that is normally blocked by a proton, at least in its canonical form, e.g. N1 or N2 of G, N6 of A, N4 of cytosine (C), and N3 of thymine (T) or, in RNA, uracil (U). In such, metal binding is only possible after deprotonation, and the liberated proton can potentionally bind at another site. The result is a so-called metal-stabilised rare tautomer and examples of these are shown in Scheme 4.1.



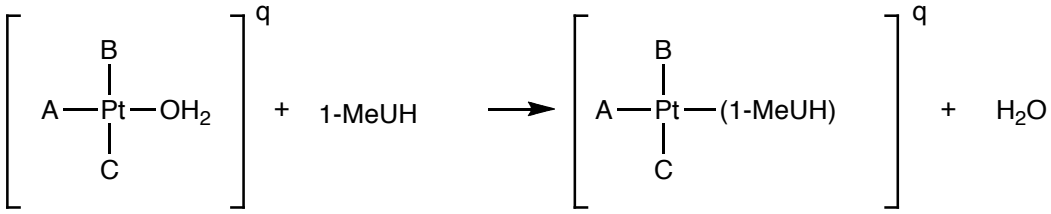
Scheme 4.1. Metal-stabilised rare tautomers of G, C, A and T bases.

Representations in which metal-binding occurs along the second pathway have been synthesised for the common model nucleobases 9-methyladenine,[9] 9-alkylguanine,[10] 1-methylcytosine,[11] 1-methyluracil,[12] and 1-methylthymine.[13] With un-substituted parent nucleobases such as cytosine,[14] or thymine/uracil,[15] or related iso-cytosine[16] metal complexes of individual tautomers were reported as well.

These metal-stabilised rare tautomers are relevant for several mutagenic events. For instance the *anti*-rotamers of M-A* or M-C* (M = Pt^{II}) may mispair via their Watson-Crick faces. This mispairing is not possible with M-T*, as the Pt^{II}-species blocks the Watson-Crick

face of the nucleobase. However the M-T* bond is prone to hydrolysis, at least in acidic medium, and in such the *free* rare tautomer is formed. Although this rare T* species will have a short lifetime, there is still the possibility that it may mispair with a complementary nucleobase, such as guanine.[12a]

By using density functional theory (DFT), in this chapter insight is first gained in the tautomerisation energy of uracil and thymine and their N1-methylated analoga. Thereafter, insight is gained on how the relative stabilities of tautomers of 1-methyluracil (1-MeUH) and 1-methylthymine (1-MeTH) can be modified through coordination to Pt^{II} complexes with a given set of ligands. Essentially, we have investigated the water/1-MeUH exchange reactions of $[Pt^{II}(A)(B)(C)(OH_2)]^q$ + 1-MeUH, shown in Scheme 4.2. A negative reaction energy indicates a given product is formed, and the more negative the reaction energy, the more stable a given product is. Consequently, the more plausible, at least on a thermodynamic basis, the exchange reaction may occur.



Scheme 4.2. Water-(1-MeUH) exchange reaction at platinum complex. An analogous reaction exists for 1-MeTH.

Another aspect of this study is to investigate what role the overall charge q in the complex plays, *i.e.* to what extent the reaction energy is affected. Ligands A, B and C are neutral amine (NH_3) or anionic chlorido (Cl^-) ligands and the composition of A, B and C is varied from all ammine (q = +2) to all chlorido, (q = –1) via consecutive substitution of an ammine for a chlorido ligand. The effect of q will be largest in the gas phase but due to the screening effect of, for instance water, the charge effect will be less pronounced. For this reason, solvent effects due to water have been taken into account as well.

The effects that Pt^{II} binding has on the tautomeric structure, has not only been investigated for 1-MeUH but also for 1-MeTH. In addition the effect on 1-MeTH tautomerisation by the anti-cancer drug cisplatin *cis*-$[Pt(NH_3)_2Cl_2]$, which is known to convert, in a first hydrolysation step, into *cis*-$[Pt^{II}(NH_3)_2Cl(OH_2)]^+$ (A, B = NH_3 and C = Cl^-) and, in a second hydrolisation step, into *cis*-$[Pt^{II}(NH_3)_2(OH_2)_2]^{2+}$ is also discussed.[17] Some more details of cisplatin are given in chapter 1.

4.2 Summary of Computational Methods

Geometry optimisations have been carried out with the program QUILD, a wrapper around ADF. The tautomerisation energies of uracil, thymine and its N1-methylated analoga are investigated by the BP86/TZ2P method and the bases were kept planar, C_S. The tautomerisation energies (ΔE_{Taut}) are computed according to Eq. 4.1.

$$\Delta E_{Taut} = E_{Rare\ tautomer} - E_{Canonical\ tautomer} \qquad \text{(Eq. 4.1)}$$

The terms on the right hand side are the energies of a given rare tautomer ($E_{Rare\ tautomer}$) and the canonical tautomer ($E_{Canonical\ tautomer}$). Because the total energy of the canonical tautomer is more negative than the one of the rare tautomer the tautomerisation is a positive entity.

The reaction energies were computed by optimising every component without any symmetry constraint, C_1, and by taking, for all components, relativistic effects into account. This method is referred to as ZORA-BP86/TZ2P. This latter method proved to be well working for several transition metals.[18] The reaction energy ($\Delta E_{Reaction}$) is computed according to Scheme 4.2 and can be represented as follows, Eq. 4.2.

$$\Delta E_{Reaction} = \Sigma E_{Products} - \Sigma E_{Reactants} \qquad \text{(Eq. 4.2)}$$

According to Eq. 4.2 the reaction energy is computed as the difference between the sum of the energies of the products ($\Sigma E_{Products}$) and the sum of the energies of the reactants ($\Sigma E_{Reactants}$).

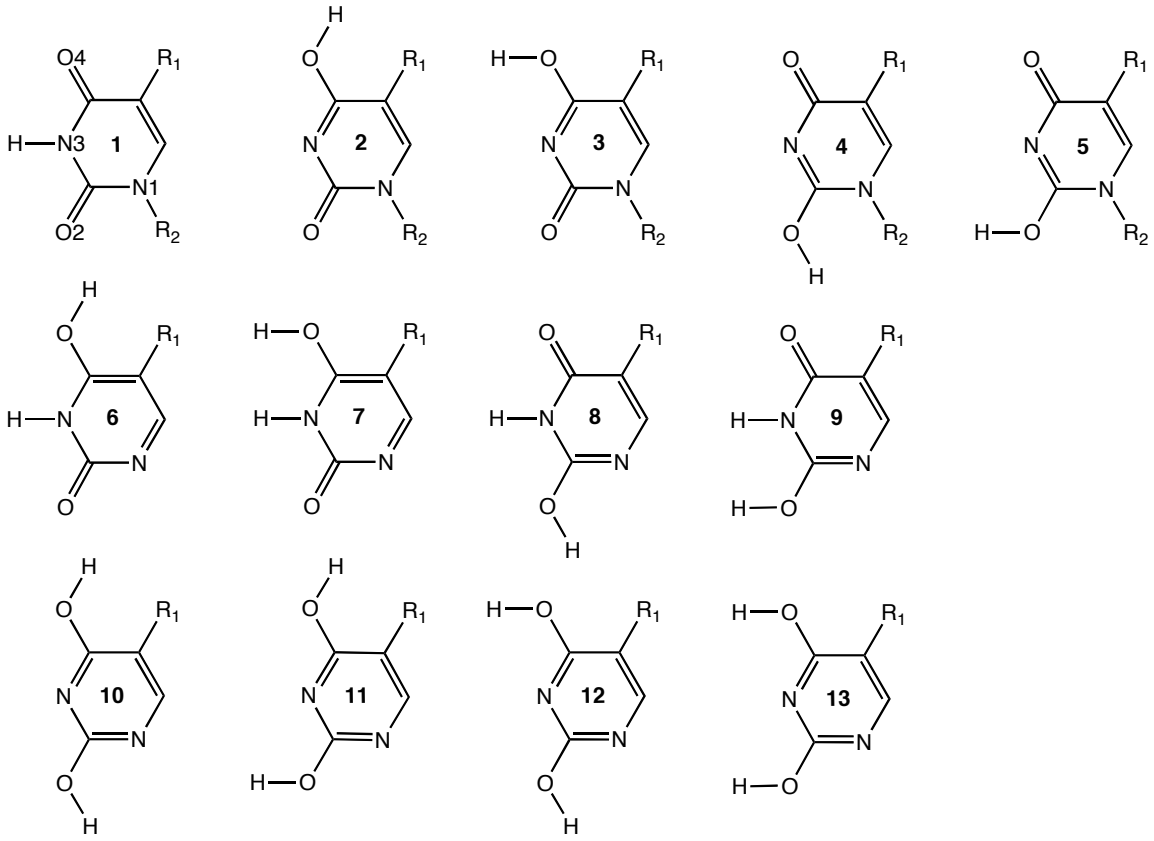
In the gas phase a frequency analysis was performed to check whether Pt^{II}-nucleobase complex is an equilibrium structures. In only a few cases an imaginary frequency of about i 20 cm^{-1} was found, but was judged to be spurious as it concerned an unhindered rotation of an ammine ligand, or a rotation of the nucleobase with respect to the $[Pt^{II}(A)(B)(C)]^q$ fragment. The solvent effects are treated implicitly by the COSMO method. The interested reader is referred to chapter 2 in which a complete overview of the computational settings is given.

4.3 Results and Discussion

The main feature of this chapter is to discuss which tautomer is stabilised by a given $[Pt^{II}(A)(B)(C)]^q$ moiety. However first results are presented of the tautomerisation of free uracil and thymine, followed by a short summary of their metal binding patterns.

4.3.1 Tautomer Energies of Uracil and Thymine

First, the results of tautomerisation on uracil (R_1 = H, R_2 = H), 1-MeUH (R_1 = H, R_2 = Me), thymine (R_1 = Me, R_2 = H) and 1-MeTH (R_1 = Me, R_2 = Me) are discussed, thus, without the binding of a Pt^{II}-complex. These results have been summarised in Table 4.1 and the respective structures can be found in Scheme 4.3.



Scheme 4.3. Different tautomeric forms of uracil (R_1, R_2 = H), 1-methyluracil (R_1 = H, R_2 = Me), thymine (R_1 = Me, R_2 = H) and 1-methylthymine (R_1, R_2 = Me).

There are a lot of tautomers, but to start with, the tautomers are discussed whose proton at N3 is shifted (**1** - **5**). In all cases the relative order of stability is **1** < **3** < **2** < **5** < **4**, in which "<" means the preceding species is more stable than the following species. This indicates that methylation at either N1 or C5 does not have any significance on tautomerisation. This order is observed in both the gas phase as well as in water. It is, however, unrealistic that a rare tautomer will be present in water, because the tautomerisation energy of the most stable tautomer (*i.e*, species **3**) is 10 kcal/mol. Rejnek *et al.*[2d] have investigated the tautomerisation of unsubstituted thymine and uracil as well at the MP2/TZVPP//RI-MP2/TZVPP level of theory. The relative order of stability that is found with this accurate MP2 method does not fully recover our order: **1** < **3** < **5** < **2** < **4**. Thus, there is a change of tautomers **2** and **5**, which

is not quite surprising in the sense that **2** and **5** lie quite close in energy at both levels of theory (BP86/TZ2P and RI-MP2/TZVPP).

Table 4.1. Tautomerisation energies (in kcal/mol) for the tautomers of uracil (R_1, R_2 = H), thymine (R_1 = Me, R_2 = H), 1-methyluracil (R_1 = H, R_2 = Me) and 1-methylthymine (R_1 = Me, R_2 = Me) in the gas phase and in water.

	R_1 = H			R_1 = Me		
	Gas phase		Water	Gas phase		Water
	BP86	MP2[a]	BP86	BP86	MP2[a]	BP86
R_2 = H						
1	0.0	0.0	0.0	0.0	0.0	0.0
2	17.8	20.3	11.6	18.9	21.8	13.0
3	11.6	12.5	10.3	12.5	13.4	11.0
4	29.2	31.3	18.0	28.3	32.1	17.4
5	19.7	19.8	15.5	18.9	19.2	14.9
6	20.9	22.3	16.7	22.3	25.2	18.5
7	23.5	25.7	18.5	25.0	27.7	19.6
8	11.5	11.1	14.6	11.3	10.7	14.3
9	18.7	19.9	16.9	18.5	19.5	16.6
10	18.4	17.4	20.4	19.1	18.5	21.5
11	18.5	17.4	20.6	19.2	18.5	21.7
12	13.6	11.6	19.7	14.2	13.2	20.0
13	14.7	12.8	20.0	15.3	14.3	20.2
R_2 = Me						
1	0.0		0.0	0.0		0.0
2	17.0		10.8	18.0		12.1
3	10.9		9.6	11.8		10.2
4	29.6		19.2	28.6		18.4
5	20.3		15.8	19.4		15.1

[a] Computed at RI-MP2/TZVPP under C_S symmetry constraint; data taken from ref. 2d.

With regard to the unsubstituted tautomer, species **8** is even more stable than species **3**. It should be noted, however, that in this case the shifted proton comes from N1, the site that is in DNA blocked by the sugar moiety. These species, including the di-hydroxo species will not be discussed further, when metal binding is concerned.

4.3.2 Reaction Energies for Exchange of Water by 1-MeUH and 1-MeTH in Pt^{II} Complexes in the Gas-Phase

We have investigated which 1-MeUH and 1-MeTH tautomers are stabilised by which $[Pt^{II}(A)(B)(C)]^q$ moiety by calculating reaction energies of the reaction shown in Scheme 4.3. In this reaction energy, when applicable, contributions owing to tautomerisation are included.

This is the case when due to Pt^{II} binding **1** tautomerises into **2**, **3**, **4** or **5** (Scheme 4.4). There are a lot ways by which a metal can coordinate to uracil and thymine,[19] which will be discussed in more detail. According to Scheme 4.4 tautomer **1** can coordinate in four possible ways to Pt^{II}: O2, O4 in either *anti* or *syn* arrangement. O4 binding is represented by **1a** and **1b**. This binding pattern was first settled for $HgCl_2$[20a] and later by other transition metals.[20b-e] Binding of alkali cations to O2 as represented by **1c** and **1d** is possible in B-DNA,[21] but has also been observed with model compounds.[20e-22] There are no examples in which Pt^{II} binds O2 or O4, because they are poor donors.[19] Stated differently, Pt^{II} is a soft acid and oxygen is a hard base, which does not interact strongly according to the theory of 'Hard and Soft Acid and Bases' (HSAB) as proposed by Pearson.[23] Good partners are those in which the acid and base are both soft or both hard. There are, however, examples of oxo compounds including platinum to be found in the literature, but instead of the soft Pt^{II} the much harder Pt^{IV} is used.[24]

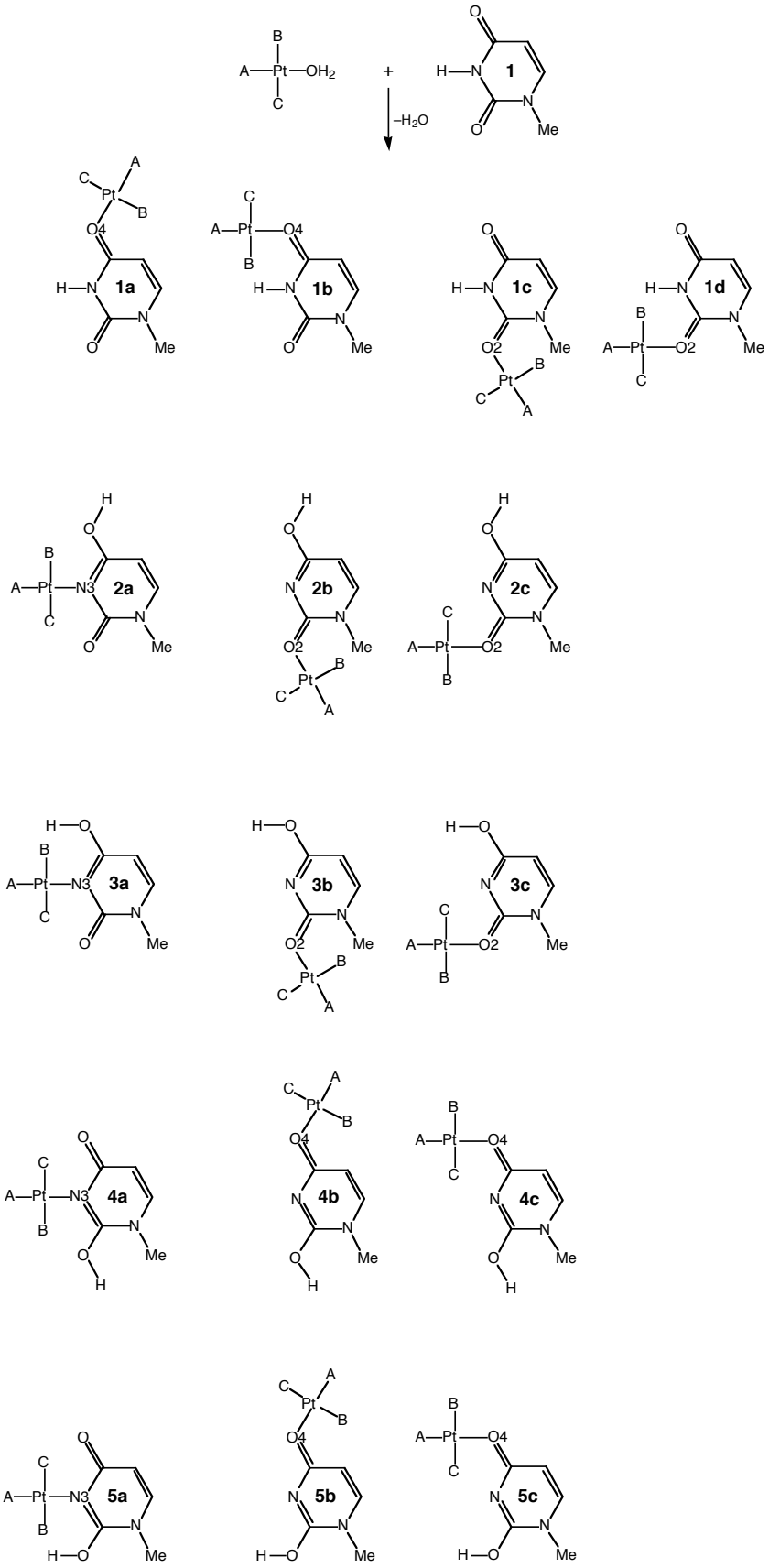
Tautomers **2-5** can coordinate in three ways to Pt^{II}: N3, or O2 (*anti* or *syn*) or O4 (*anti* or *syn*). For Pt^{II} a rare tautomer has been found experimentally and resembles products **2a** or **3a**.[12, 13] At least the rare tautomer was protonated at O4, but there was no evidence on how the proton was orientated, either *anti* or *syn*. Rare tautomer products **2b-5b** and **2c-5c** were not observed in combination with Pt^{II} binding. Note that **2b-5b** and **2c-5c** differ from **2a-5a** in the sense that Pt^{II} is bound to O2 or O4 instead of N3. In this study only the hetero-atoms were treated as possible binding site of a Pt^{II} atom, but there are examples of C5 binding for $Hg(CH_3COO)_2$,[25] Pt^{III},[26] Pt^{II}[27] and Au^{III}.[28] From Scheme 4.4 it is furthermore clear that only one Pt^{II} atom per nucleobase has been concerned, also termed as 1:1 complexes.

The reaction energies due to the formation of all possible products with different ligand compositions are presented in Table 4.2 and calculated according to Scheme 4.2. In Table 4.2, the net charge and the Pt^{II} binding sites (X) are also given. In the columns one can find the ligand composition, starting with the most positively charged species $[Pt^{II}(NH_3)(NH_3)(NH_3)]^{2+}$ and ending with the most negatively charged complex $[Pt^{II}(Cl)(Cl)(Cl)]^{-}$. The rows represent the products and a separation between normal tautomer and rare tautomer products is made. The set-up of Table 4.2 is used for most forthcoming tables. The more negative a reaction energy is, the more stabilised the respective product is, or differently stated: the higher the likelyhood it will form.

Taking a closer look at the reaction energies in Table 4.2, one can see that all products $[Pt^{II}(NH_3)_3(1\text{-MeUH})]^{2+}$ can form, because the associated reaction energy is negative. Upon binding of the dipositive species, containing three neutral NH_3 ligands, not only all products can form, but these are also the most stable, because the reaction energy is the most negative. The most and least favoured products are **2c** and **1d**. They have reaction energies of –39.0 and –24.5 kcal/mol, respectively. However, experimental data that corroborate this calculation are not available.

If the net charge is lowered from +2 to –1, by substituting an ammine ligand for a chlorido ligand, the reaction energy becomes more and more positive. In some cases products, either in rare or normal tautomer form will not be formed, as is the case when the charge of the Pt^{II}-complex is 0 or –1 and is mirrored by a positive value of the reaction energy. For instance, product **2c** is favoured in combination with $[Pt^{II}(NH_3)(NH_3)(NH_3)]^{+2}$ by –39.0 kcal/mol but as the charge is decreased, the reaction energy gradually rises and the product becomes least favoured with $[Pt^{II}(Cl)(Cl)(Cl)]^{-1}$ as indicated by the postive reaction energy of 29.9 kcal/mol.

Depending on their coordination mode, tautomers of **1** and **3** form stable reaction products, independent of the Pt^{II}-complex used. Tautomer **1** gives stable products through O4 coordination in a *syn* orientation (**1b**), and allowing for tautomerisation of **3** into **3a**, always a stable reaction product is formed. This is very interesting as this is a metal-stabilised rare tautomer and has even been observed experimentally in an X-ray structure.[12] However the orientation of the proton at O4 could not be determined from the experiment, and so product **2a** could also have been oberved but this product is only formed when, positively charged Pt^{II}-complexes are used.



Scheme 4.4. The reaction of exchanging water by 1-MeUH and 1-MeTH in the $[Pt^{II}(A)(B)(C)(OH_2)]^q$ complex. The methyl group at C5 in 1-MeTH and formal charges have been omitted for clarity (see Scheme 4.2).

The reaction energies were not only calculated for 1-MeUH but also for 1-MeTH. It can be seen from Table 4.3, that the substitution of H5 (1-MeUH) by a methyl group (1-MeTH) will lead to some changes. Especially the products, in which the Pt^{II}-species binds O4 in an *anti* manner, are affeced: **1a**, **4b**, **5b**. The reaction energies are increased by about 3 kcal/mol, wih respect to the non-methylated analogs. Obviously, there is a repulsive interaction between the methylgroup of thymine and the Pt^{II} complex.

Products having a 1-MeTH moiety, exothermic reaction energies (*i.e.*, negative reaction energy) occur for positively charged Pt^{II}-complexes and can, therefore, form. At contrast, in case of the negatively charged Pt^{II}-complexes, most reaction energies are or become endothermic (*i.e.*, positive reaction energy) and are, therefore, not feasible. Importantly, products **1b** and **3a** are formed independent of the used Pt^{II}-complex. The existence of **3a** has been settled experimentally,[13] and may participate in mutagenic events (see later). Again, because the position of the proton at O4 could not be determined from the experiments, the obseved product could also be **2a**, but is only formed when positively charged Pt^{II} complexes are applied.

Table 4.2. Reaction energies (in kcal/mol) in the gas phase for the exchange of H_2O by 1-MeUH in $[Pt^{II}(A)(B)(C)(OH_2)]^q$ forming $[Pt^{II}(A)(B)(C)(1\text{-MeUH})]^q$.

Ligands A		NH_3	NH_3	Cl	Cl	NH_3	Cl
B		NH_3	Cl	NH_3	Cl	Cl	Cl
C		NH_3	NH_3	NH_3	NH_3	Cl	Cl
Charge q		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	–35.7	–15.9	–17.5	–2.1	–2.3	3.8
1b	O4	–34.8	–19.1	–18.3	–7.2	–5.4	–1.0
1c	O2	**1d**[a]	**1d**[a]	–4.8	3.2	3.7	5.3
1d	O2	–24.5	–10.5	–11.7	–1.4	–2.4	0.3
Rare Tautomer							
2a	N3	–38.9	–14.7	–17.0	4.8	5.2	18.3
2b	O2	**2c**[a]	**2c**[a]	**2c**[a]	16.4	18.1	24.9
2c	O2	–39.0	–7.9	12.3	14.3	15.4	29.9
3a	N3	–35.1	–21.6	–20.6	–10.8	–7.6	–3.3
3b	O2	**3c**[a]	**3c**[a]	**3c**[a]	10.8	12.0	18.4
3c	O2	–36.4	–8.9	–13.0	10.0	8.9	22.0
4a	N3	–27.4	–4.8	–5.0	**5a**[a]	17.7	31.8
4b	O4	–24.6	0.9	0.8	21.5	22.9	34.3
4c	O4	–37.0	0.4	–6.8	24.7	24.7	41.5
5a	N3	–27.6	–16.6	–12.9	–6.1	0.4	4.1
5b	O4	–32.1	–6.9	–7.6	13.0	13.7	24.9
5c	O4	–38.3	–6.9	–10.9	15.4	14.8	29.3

[a] Expected species spontaneously transforms into indicated species.

Table 4.3. Reaction energies (in kcal/mol) in the gas phase for the exchange of H_2O by 1-MeTH in $[Pt^{II}(A)(B)(C)(OH_2)]^q$ forming $[Pt^{II}(A)(B)(C)(1\text{-MeTH})]^q$.[a]

Ligands A		NH_3	NH_3	Cl	Cl	NH_3	Cl
B		NH_3	Cl	NH_3	Cl	Cl	Cl
C		NH_3	NH_3	NH_3	NH_3	Cl	Cl
Charge q		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	–32.9	–13.0	–14.0	1.2	0.4	6.5
1b	O4	–35.9	–18.9	–18.8	–6.7	–5.6	–1.2
1c	O2	**1d**[a]	**1d**[a]	–8.0	2.5	3.1	5.7
1d	O2	–28.5	–12.6	–13.8	–2.1	–3.0	0.7
Rare Tautomer							
2a	N3	–41.4	–15.6	–17.8	5.1	5.9	19.7
2b	O2	**2c**[a]	**2c**[a]	**2c**[a]	16.7	18.4	26.3
2c	O2	–42.2	–9.0	–13.5	14.7	15.9	31.3
3a	N3	–37.5	–22.5	–21.4	–10.6	–7.2	–2.4
3b	O2	**3c**[a]	**3c**[a]	**3c**[a]	10.9	12.2	19.7
3c	O2	–39.8	–10.2	–14.3	10.2	9.1	23.2
4a	N3	–31.5	–7.0	–7.3	**5a**[a]	16.5	31.3
4b	O4	–22.7	4.4	3.7	25.4	25.7	37.6
4c	O4	–39.4	–2.1	–8.4	22.8	23.2	40.5
5a	N3	–31.5	–18.5	–15.1	–6.7	–0.7	3.8
5b	O4	–29.9	–3.6	–4.4	16.8	16.5	28.0
5c	O4	–40.5	–8.2	–12.3	14.6	13.5	28.4

[a] Expected species spontaneously transforms into indicated species.

4.3.3 Reaction Energies for Exchange of Water by 1-MeUH and 1-MeTH in Pt^{II} Complexes in Water

The solvent effects due to water on the reaction were also taken into account, and led to some interesting differences with respect to the gas phase. The results can be found in Table 4.4. In water, the charge effect has almost been eliminated, due to the charge screening effect of the solvent. This phenomenon describes the decrease in attraction between the electon and the nucleus of an atom. This will have as an effect that electrostatic contribution becomes less, and the reaction energies will be less dependent on the charge. Thus, for all products (**1a** to **5c**) there is only a small change in reaction energy in the order of a few kcal/mol when the composition of the Pt^{II}-complex, and thus the overall charge, q, is changed.

In accordance with the gas phase, tautomer **1** and **3** lead to stable reaction products **1b** and **3a**, irrespective of the Pt^{II} complex used, as all reaction energies are negative. However, in solution the products **1a**, **1d**, **2a** (and to lesser extent **5a)** become stable products as well. It should be noted that the reaction energy associated with the formation of **1d** is only small: –0.8 to –1.8 kcal/mol. The formation of **5a** is also doubtful as its reaction energy is, depending on the Pt^{II}-complex slightly positive or negative. Thus, under experimental conditions, it is expected that coordination of a Pt^{II} complex will occur through binding via O4 in either *anti* (**1a**) or *syn* (**1b**) orientation, or binding via the (deprotonated) N3 atom. However, the former possibility contradicts the expectations from the HSAB theory. In case the charge of the complex is +2 the preference of the O4 binding mode is favoured by 1.2 kcal/mol over the N3 binding mode by comparing products **1a** and **3a**. Clearly, all other Pt^{II} complexes with a charge of +1, 0 or –1 will prefer the N3 coordination, and in particular products **2a** and **3a**. The reaction energies of these rare tautomer products are all negative, irrespective of the Pt^{II} complex, and this species is also found experimentally.[12] Note, that from the experiment, the proton orientation (*anti* or *syn*) at O4 could not be determined. One can calculate that **1b** is preferred over **3a** by 0.8, 2.5, and 3.6 kcal/mol for a positively charged (q = +1), neutral (q = 0) and negatively charged (q = –1) Pt^{II} complex, respectively.

In water the results for 1-MeTH (see Table 4.5) are, again, quite similar to those obtained with 1-MeUH and the charge effect is almost eliminated. Binding modes that are most feasible are *syn* binding to O4 (**1b**), and N3 binding (**2a** and **3a**). Other products are only slightly stable (**1d**), or either stable or not (**5a**). With respect to 1-MeUH, reaction energies become more positive for **1a**, **4b** an **5b**, which have an *anti* binding to O4 in common. The increase of the reaction energy is due to the larger steric repulsion between the methyl group

and the Pt^{II}-complex. Interestingly the *syn* binding of a dipositive species (q = +2) to O4 is preferred by 0.5 kcal/mol over N3 binding, compare **1b** and **3a**. All other complexes prefer the N3 binding (**3a**) by 0.9, 2.2 and 3.3 kcal/mol for a positively (q = +1), neutral (q = 0) and negatively (q = –1) charged Pt^{II} complex over O4 binding (**1b**). Again, all reaction energies for the formation **2a** and **3a** are negative, and these rare tautomer products have been found experimentally.[13] However, the orientation of the proton at O4 could not be determined from the experiment.

Apart from preferred binding modes, there are other aspects that are of interest. Some years ago Deubel also investigated the same water exchange reaction with the $[Pt(NH_3)_3(OH_2)]^{2+}$ fragment, albeit with an other nucleobase ligand, namely guanine.[29] It was essentially shown that the reaction energy became less exothermic, and even endothermic in one case, as the dielectric constant was increased from 1 (vacuum, gas phase) to 78.4 (water). This result is confirmed in our study for $[Pt^{II}(NH_3)_3(OH_2)]^{2+}$ (q = +2) and is reflected by the formation of product **3a**. The formation of this product in the gas phase goes along with –35.1 and –37.5 kcal/mol using 1-MeUH and 1-MeTH, respectively. However, in solution, these reaction energies amount to –4.6 and –3.7 kcal/mol, again by using 1-MeUH and 1-MeTH. Obviously, the reaction energy has become *less* exothermic, for a doubly charged Pt^{II} complex. This feature changes however, as the charge of the species is decreased and it most pronounced in case of the tri-chlorido complex (q = –1). The reaction energy in the gas phase forming **3a** is –3.3 and –2.4 kcal/mol using 1-MeUH or 1-MeTH respectively. In water these reaction energies amount to –7.8 and –7.9 kcal/mol, again respectively for using 1-MeUH and 1-MeTH. This shows that the reaction energy has become *more* exothermic.
This means that the exchange reaction using positively charged Pt^{II} complexes become less feasible and those of negative complexes are more feasible, due to the treatment of solvent. Thus, as was already alluded on in chapter 2, solvent treatment is important.

Table 4.4. Reaction energies (in kcal/mol) in water for the exchange of H_2O by 1-MeUH in $[Pt^{II}(A)(B)(C)(OH_2)]^q$ forming $[Pt^{II}(A)(B)(C)(1\text{-MeUH})]^q$.

Ligands A		NH_3	NH_3	Cl	Cl	NH_3	Cl
B		NH_3	Cl	NH_3	Cl	Cl	Cl
C		NH_3	NH_3	NH_3	NH_3	Cl	Cl
Charge q		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	–5.8	–5.2	–5.2	–3.4	–3.8	–2.3
1b	O4	–5.1	–5.6	–5.3	–4.6	–4.9	–4.2
1c	O2	4.4	2.3	3.9	3.2	3.5	4.3
1d	O2	–0.8	–1.8	–1.0	–0.9	–1.6	–1.2
Rare Tautomer							
2a	N3	–3.8	–3.9	–4.8	–3.1	–3.6	–3.3
2b	O2	10.7	12.3	11.5	10.6	10.9	12.0
2c	O2	2.3	4.8	3.7	6.4	6.2	7.2
3a	N3	–4.6	–6.4	–5.9	–6.4	–7.4	–7.8
3b	O2	10.3	8.1	10.4	9.7	10.0	10.9
3c	O2	3.1	4.6	3.7	6.0	5.6	6.5
4a	N3	4.8	4.4	4.2	5.5	5.7	6.0
4b	O4	8.7	10.0	10.0	12.0	11.5	13.3
4c	O4	7.1	9.5	8.3	11.4	11.3	12.3
5a	N3	1.8	–0.7	0.2	–0.2	–1.3	–1.7
5b	O4	5.5	6.7	6.7	8.9	8.3	10.1
5c	O4	4.8	7.1	5.7	8.6	8.3	9.5

Table 4.5. Reaction energies (in kcal/mol) in water for the exchange of H_2O by 1-MeTH in $[Pt^{II}(A)(B)(C)(OH_2)]^q$ forming $[Pt^{II}(A)(B)(C)(1\text{-MeTH})]^q$.

Ligands A		NH_3	NH_3	Cl	Cl	NH_3	Cl
B		NH_3	Cl	NH_3	Cl	Cl	Cl
C		NH_3	NH_3	NH_3	NH_3	Cl	Cl
Charge q		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	–1.3	–1.7	–0.8	0.5	–0.5	0.9
1b	O4	–4.2	–5.1	–4.5	–4.3	–5.0	–4.6
1c	O2	3.3	2.7	3.6	2.5	2.8	3.7
1d	O2	–1.0	–2.2	–1.6	–1.4	–2.0	–1.6
Rare Tautomer							
2a	N3	–2.3	–2.3	–2.9	–1.7	–2.1	–2.0
2b	O2	11.7	9.8	12.3	11.2	11.8	12.9
2c	O2	3.6	6.1	5.0	7.8	7.4	8.4
3a	N3	–3.7	–5.8	–5.4	–6.1	–7.2	–7.9
3b	O2	10.5	7.9	10.4	9.3	10.1	11.0
3c	O2	3.2	4.7	3.8	6.3	5.8	6.8
4a	N3	4.2	3.8	3.4	5.1	4.6	4.9
4b	O4	13.0	13.4	14.1	16.1	14.8	16.6
4c	O4	6.5	9.3	8.0	11.0	10.6	11.6
5a	N3	0.9	–1.3	–0.4	–0.3	–2.0	–2.6
5b	O4	9.7	10.6	10.9	12.9	11.7	13.4
5c	O4	4.2	7.4	5.6	8.1	7.8	8.8

4.3.4 Geometrical Aspects

Apart from the energetics, some geometrical features shall be discussed in more detail. With B = Cl^- and the possibility of interacting with the proton at N3, the complex becomes flat as is the case for **1b** and **1d**, *i.e.*, a dihedral angle of approximately 0°. This dihedral angle is defined as the angle between the coordination plane of the Pt^{II} metal and that of 1-MeUH, connected through the coordination bond. The dihedral angle becomes for the same structures, in contrast, approximately 90° in case B and C are NH_3, because the propensity of a favourable interaction has vanished. In complex **2c**, **3c**, **4c**, and **5c** in which B and C are both Cl^- the dihedral angle between the planes is roughly 90°. The explanation lies in the fact that repulsion occurs between Cl^- and the lone pair at N3 of 1-MeUH.

In Table 4.2 and 4.3 it can be seen that some *anti* structures (**1c**, **2b** and **3b**) were not located on the potential energy surface (PES), but spontaneously turned into their *syn* analoga (**1d**, **2c** and **3c**). The products in which this *anti* to *syn* transition takes place have a Pt^{II} binding to O2. In contrast, the *anti* structures in which Pt^{II} is bound to O4 (**1a**, **4b**, **5b**) could be located on the PES. Thus, an *anti* to *syn* transition is not induced by the substitution of a proton by a methyl group at C5. In case the *anti* to *syn* transition takes place, though, it can be explained as follows. The PES has normally two wells: the deeper well is the *syn* structure and the higher well is the *anti* structure and the two wells correspond to the global and local minimum, respectively. There are two processes that may happen now, that have as a consequence that the *anti* structure can not be located anymore. The minimum of the *anti* structure can raise in energy due to repulsive interactions and/or the barrier between the minima can drop due to attractive interaction leading to the *anti* structure becoming a shoulder (but no longer a stationary point) on the PES.

Furthermore, a relation between the reaction energies and the Pt^{II}-X distance (X is the Pt^{II} binding site). From Tables 4.6 - 4.9 it becomes clear that such a relationship does not exist, as there is no change in the Pt^{II}-X distance as the charge is varied. Rather the bond length is detemined by the ligand that is *trans* to X, thus ligand A. The Pt^{II}-X distance is in general longer in case A = Cl^-, than in case A = NH_3. This observation can be explained by the "structural *trans*-effect" (or "*trans*-influence"). This is termed as the ability of a ligand to lengthen the bond *trans* to it, in preference to those that are *cis*,[30] and this effect is indeed larger for Cl^- than for NH_3, as supported by the results. The experimental distance, measured in a crystal-structure, of the Pt^{II}-N3 bond is 2.04 Å (O4 protonated, product **3a**), and has standard deviation of 0.023 Å.[12a] This distance is better reproduced in water, than in the gas

phase. In the gas phase there is a deviation of up to 0.05 Å, depending on the Pt^{II} complex. In solution, the deviation is always 0.01 Å, and there is thus a much better agreement between experiment and theory. The effect of the identity of the nucleobase, either 1-MeUH or 1-MeTH, is negligible. In almost all cases the bond distance is the same. If not, the absolute difference amounts to only 0.02 Å, which coincides with the experimentally found standard deviation.

Table 4.6. Bond distances (in Å) in the gas phase between Pt and its binding atom (X) of 1-MeUH.

Ligands	A	NH_3	NH_3	Cl	Cl	NH_3	Cl
	B	NH_3	Cl	NH_3	Cl	Cl	Cl
	C	NH_3	NH_3	NH_3	NH_3	Cl	Cl
Charge		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	2.03	2.04	2.09	2.09	2.04	2.07
1b	O4	2.03	2.05	2.09	2.09	2.04	2.07
1c	O2	**1d**[a]	**1d**[a]	2.12	2.13	2.06	2.13
1d	O2	2.04	2.06	2.11	2.11	2.05	2.09
Rare Tautomer							
2a	N3	2.07	2.06	2.09	2.06	2.02	2.02
2b	O2	**2c**[a]	**2c**[a]	**2c**[a]	2.12	2.05	2.12
2c	O2	2.05	2.05	2.10	2.08	2.05	2.11
3a	N3	2.05	2.07	2.08	2.10	2.03	2.05
3b	O2	**3c**[a]	**3c**[a]	2.10	2.12	2.05	2.11
3c	O2	2.04	2.05	2.10	2.11	2.05	2.08
4a	N3	2.07	2.06	2.10	**5a**[a]	2.03	2.03
4b	O4	2.01	2.03	2.07	2.07	2.03	2.06
4c	O4	2.04	2.04	2.08	2.06	2.03	2.04
5a	N3	2.05	2.08	2.08	2.11	2.03	2.06
5b	O4	2.01	2.03	2.07	2.07	2.03	2.07
5c	O4	2.03	2.04	2.07	2.08	2.03	2.05

[a] Expected species spontaneously transforms into indicated species.

Table 4.7. Bond distances (in Å) in the gas phase between Pt and its binding atom (X) of 1-MeTH.

Ligands	A	NH_3	NH_3	Cl	Cl	NH_3	Cl
	B	NH_3	NH_3	NH_3	Cl	Cl	Cl
	C	NH_3	Cl	NH_3	NH_3	Cl	Cl
Charge		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	2.03	2.04	2.09	2.09	2.04	2.07
1b	O4	2.03	2.05	2.09	2.09	2.04	2.07
1c	O2	**1d**[a]	**1d**[a]	2.10	2.13	2.05	2.13
1d	O2	2.04	2.06	2.10	2.11	2.06	2.10
Rare Tautomer							
2a	N3	2.07	2.06	2.10	2.06	2.02	2.03
2b	O2	**2c**[a]	**2c**[a]	**2c**[a]	2.13	2.05	2.12
2c	O2	2.05	2.05	2.10	2.09	2.04	2.11
3a	N3	2.05	2.07	2.08	2.10	2.03	2.05
3b	O2	**3c**[a]	**3c**[a]	**3c**[a]	2.12	2.05	2.11
3c	O2	2.04	2.05	2.10	2.11	2.05	2.08
4a	N3	2.07	2.06	2.10	**5a**[a]	2.03	2.03
4b	O4	2.01	2.03	2.07	2.06	2.02	2.05
4c	O4	2.04	2.04	2.08	2.06	2.03	2.04
5a	N3	2.05	2.08	2.08	2.11	2.03	2.06
5b	O4	2.01	2.03	2.07	2.07	2.03	2.06
5c	O4	2.03	2.04	2.07	2.06	2.03	2.05

[a] Expected species spontaneously transforms into indicated species.

Table 4.8. Bond distances (in Å) in water between Pt and its binding atom (X) of 1-MeUH.

Ligands	A	NH_3	NH_3	Cl	Cl	NH_3	Cl
	B	NH_3	NH_3	NH_3	Cl	Cl	Cl
	C	NH_3	Cl	NH_3	NH_3	Cl	Cl
Charge		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	2.04	2.05	2.07	2.07	2.06	2.07
1b	O4	2.05	2.06	2.07	2.07	2.06	2.07
1c	O2	2.07	2.08	2.09	2.09	2.09	2.10
1d	O2	2.06	2.07	2.08	2.08	2.07	2.08
Rare Tautomer							
2a	N3	2.06	2.06	2.06	2.05	2.05	2.05
2b	O2	2.06	2.07	2.08	2.08	2.07	2.09
2c	O2	2.06	2.06	2.07	2.07	2.06	2.07
3a	N3	2.05	2.05	2.05	2.05	2.05	2.05
3b	O2	2.06	2.07	2.08	2.08	2.08	2.09
3c	O2	2.05	2.06	2.07	2.07	2.06	2.07
4a	N3	2.06	2.06	2.06	2.06	2.05	2.05
4b	O4	2.04	2.05	2.05	2.06	2.05	2.06
4c	O4	2.05	2.05	2.06	2.06	2.05	2.06
5a	N3	2.05	2.05	2.06	2.07	2.05	2.05
5b	O4	2.04	2.05	2.06	2.06	2.05	2.06
5c	O4	2.04	2.05	2.06	2.06	2.05	2.06

Table 4.9. Bond distances (in Å) in water between Pt and its binding atom (X) of 1-MeTH.

Ligands	A	NH_3	NH_3	Cl	Cl	NH_3	Cl
	B	NH_3	NH_3	NH_3	Cl	Cl	Cl
	C	NH_3	Cl	NH_3	NH_3	Cl	Cl
Charge		2	1	1	0	0	–1
Normal Tautomer	X						
1a	O4	2.04	2.05	2.06	2.06	2.06	2.07
1b	O4	2.05	2.06	2.07	2.07	2.06	2.07
1c	O2	2.06	2.06	2.09	2.09	2.08	2.10
1d	O2	2.06	2.07	2.08	2.08	2.07	2.08
Rare Tautomer							
2a	N3	2.06	2.06	2.06	2.05	2.05	2.05
2b	O2	2.05	2.07	2.08	2.09	2.07	2.09
2c	O2	2.06	2.06	2.07	2.07	2.06	2.07
3a	N3	2.05	2.05	2.05	2.05	2.05	2.05
3b	O2	2.05	2.07	2.08	2.09	2.07	2.09
3c	O2	2.05	2.06	2.07	2.07	2.06	2.07
4a	N3	2.06	2.06	2.06	2.06	2.05	2.05
4b	O4	2.04	2.04	2.05	2.05	2.05	2.05
4c	O4	2.05	2.05	2.06	2.06	2.05	2.05
5a	N3	2.05	2.05	2.06	2.07	2.05	2.05
5b	O4	2.03	2.04	2.05	2.05	2.05	2.05
5c	O4	2.04	2.05	2.05	2.06	2.05	2.05

The existence of a crystal structure in which a Pt^{II} ion bridges two uracil molecules (one neutral and one deprotonated), of which the neutral one is in its rare tautomer form, has been mentioned several times by now. The rare tautomer was platinum bound via N3 and present in its 2-oxo-4-hydroxo-structure. In this very study,[12a] predictions were made on the geometry of the free rare tautomer. According to these, the largest changes were observed for the internal ring angles at C2, N3 and C4. With respect to the platinum-bound rare tautomer, the free rare tautomer should have internal ring angles at C2 and C4 that are increased by 1-2.5°, the internal ring angle at N3 should be decreased by the same amount. This has been checked computationally, and the ring angles have been measured and shown in Table 4.10.

In Table 4.10 internal ring angles values in solution are shown for a *cis*-$[Pt^{II}(NH_3)_2(Cl)]^+$ moiety for products **2a** and **3a** and the free rare tautomer **3**. Note that in the experimental structure a uracil and a uracilate are present, for the computation the uracilate has been substituted for a chlorido ligand.

Table 4.10. Internal ring angles (in degrees) at C4, N3 and C2 in solution.

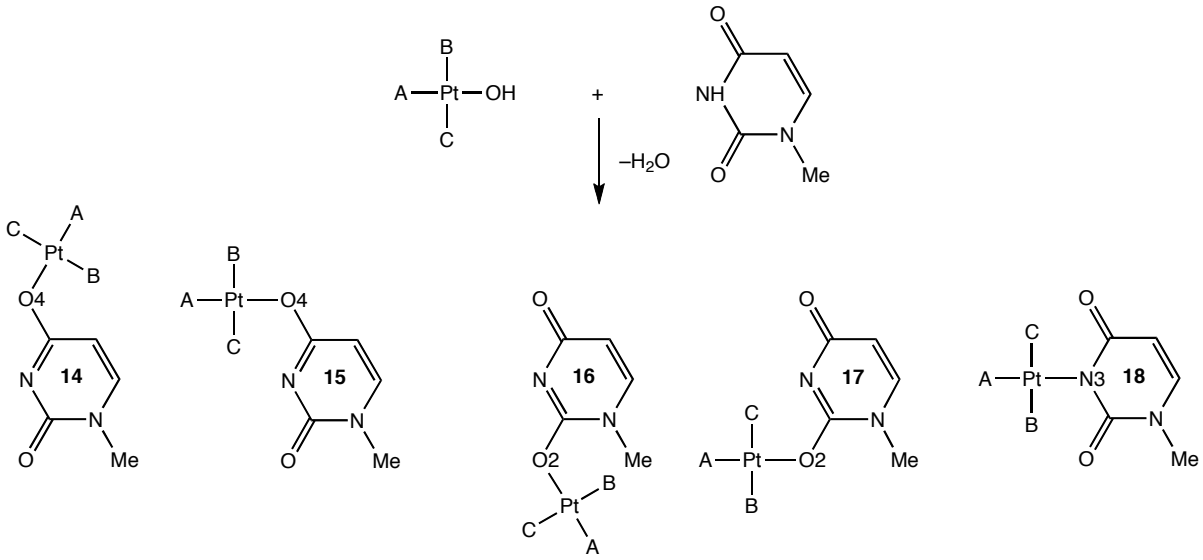
		2a	**3a**	**3**
Ligands	A	NH_3	NH_3	--
	B	Cl	Cl	--
	C	NH_3	NH_3	--
q		1	1	0
N3-C4-C5		122.2	121.9	124.3
C2-N3-C4		120.5	121.3	119.7
N1-C2-N3		117.0	116.3	117.9

The internal ring angles for product **2a** and **3a** do differ a little bit, and the largest absolute difference is found with the N3 ring angles and amounts to 0.8°. Much importantly, with respect to the platinated rare tautomer, the C4 internal ring angle increases by 2.3°, the C2 internal ring angle increases by 1.3°, and the N3 internal ring angle decreases by 1.2°. The latter ring angle changes are averages over **2a** and **3a**. These changes confirm the predictions based on experimental data, which were discussed above.

4.3.5 Complexes Containing the 1-MeU$^-$ Anion

In the literature there are several X-ray structures of Pt^{II} complexes known having either 1-methyluracil or 1-methylthymine in their N3 deprotonated form (see structure **18** in Scheme 4.7), also denoted as 1-MeU$^-$ and 1-MeT$^-$.[31] This metal binding pattern has been found with Hg^{II},[32] Pt^{II},[33] Pt^{IV},[34] Rh^{I},[35] Au^{I},[36] and Au^{III}[37] We, therefore, have also investigated the formation of these products via an exchange reaction, as shown in Scheme 4.5, with some slight modifications. Instead of using $[Pt^{II}(A)(B)(C)(OH_2)]^q$ as done previously, $[Pt^{II}(A)(B)(C)(OH)]^{q-1}$ has been used because the aqua ligand is more acidic than 1-MeUH by a 4-5 log units.[38]

In the reaction the nucleobase can be bound to Pt^{II} via O4 or O2 in either *anti* or *syn* orientation (**14** – **18**), or can be bound via N3 (**18**). The reaction energies in water are shown in Table 4.11. The reaction energies for products **14** and **17** are all endothermic, irrespective of the ligand composition, indicating these species will not be present in water. The reaction energy for **15** is slighly exothermic by up to –0.2 kcal/mol. Consequently, its existence in water is very questionable. On the contrary, the reaction energy for product **18** is always exothermic, and will be present, as corroborated by several experimental findings.



Scheme 4.5. The reaction of exchanging OH^- by 1-MeU$^-$ in the $[Pt^{II}(A)(B)(C)(OH)]^{q-1}$ complex.

Structure **18** is a building block for other structures, because the basicity of O4 is increased.[19] First, **18** can be reprotonted at O4 or O2, by using the free proton of N3 to form a rare tautomer structure. The reprotonation can be observed by vibrational spectroscopy,[39] or by determination of the pK_a values.[12,13] Second, another metal can bind to the oxygen atoms, leading to multiple metal binding motifs.[19] The maximum number is four metals, that do not need to be the same.[40]

Table 4.11. Reaction energy (in kcal/mol) in water for the exchange of OH^- by 1-MeUH in $[Pt^{II}(A)(B)(C)(OH)]^{q-1}$ forming H_2O and $[Pt^{II}(A)(B)(C)(1\text{-MeU})]^{q-1}$.

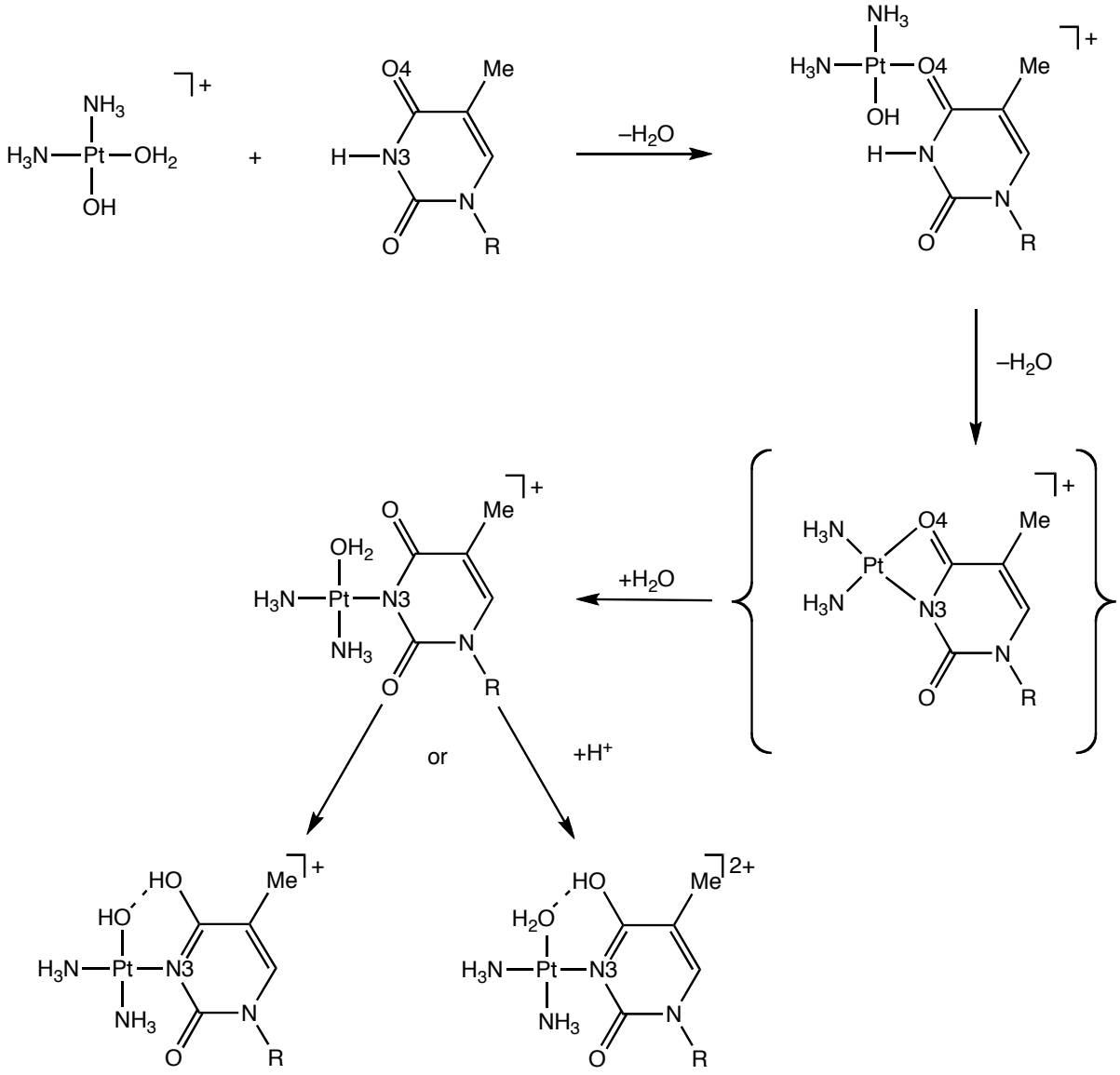
Ligands A			NH_3	NH_3	Cl	Cl	NH_3	Cl
B			NH_3	Cl	NH_3	Cl	Cl	Cl
C			NH_3	NH_3	NH_3	NH_3	Cl	Cl
Charge (q–1)			1	0	0	–1	–1	–2
14	O4	*anti*	3.3	2.7	2.6	3.4	1.2	2.2
15	O4	*syn*	–0.2	1.3	–0.1	1.9	1.1	1.4
16	O2	*anti*	**17**[a]	9.7	**17**[a]	9.5	8.9	9.1
17	O2	*syn*	4.5	5.5	4.1	5.8	4.9	5.1
18	N3		–7.5	–6.8	–9.2	–7.3	–8.5	–9.0

[a] Expected species **16** spontaneously transforms into indicated species **17**.

4.4 Pt^{II} Coordination and Implications for Biology

It is well known that thymine in DNA and uracil in RNA are not the kinetically preferred targets for cisplatin, which are rather guanine, adenine or cytosine. However, the binding of

Pt^{II} to N3 of uracil or thymine in either their free state or in DNA or RNA was reported.[41] Judging form the pK_a of the N(3)H proton which is around 9 – 10,[42] one would expect that one needs alkaline conditions to accomplish this binding. However this view is not quite true as binding has also been observed at very low pH values (3.5 – 4.0).[41a,43] There are, however, two explanations that rationalise why Pt^{II} binding takes place even at these low pH values, see also Scheme 4.6.



Scheme 4.6. Proposed tautomerisation pathway of 1-MeTH due to the binding of cisplatin.

First, before reaching its target, N3, initial binding to either O4 or O2 or both simultaneously, can be proposed, and this will acidify the proton at N3.[44] Secondly, the aquated species has is own base included, Pt-OH, (pK_{a1} of *cis*-$[Pt(NH_3)_2(H_2O)]^{2+}$ is ca. 5.5[45]). Thus, a change in

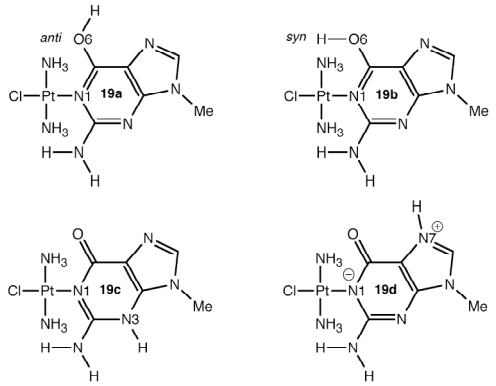
the tautomer structure can be explained in terms of initial N3 deprotonation by Pt^{II}-OH, followed by metal migration from O4/O2 to N3,[46] and reprotonation of either O4 or O2. The substitution of the proton at N3 by an Pt^{II} moiety, will increase the basicity of the exocyclic oxygen atoms, notably O4, which facilitates protonation of these and thus permits formation of a neutral rare tautomer.[5,12,13] The pK_a values of these protons are rather acidic, at least 3 – 4 in model systems, but its value depends, amongst others, on the co-ligands of the complex and complex charge.[47] In the respect of "shifted pK_a values", which are the result in shift of the parent pK_a of a nucleobase (metalated or not) in either direction (upward or downward) depending on the microenvironment,[48] the existence of a platinated rare tautomer of thymine or uracil at physiological pH is not completely excluded. Stated differently, species **2a** or **3a** might occur in a pH range in which Pt-N3 binding can take place (weakly acidic pH), and at a pH approaching 7. Under alkaline pH the nucleobase is deprotonated without doubt, making any discussions about changes in tautomer structure irrelevant.

Another issue is the mispairing possibility with G, C or T*, that may result from Pt-T*, but on steric arguments only it is not realistic that platinated complexes **2a** or **3a** will do so. The metal as well as its co-ligands prevent this mispairing as they completely block the Watson-Crick face. From a steric point of view complex **3b** could mispair with G, but the formation of the corresponding complex with 1-MeTH is endothermic (7.9 kcal for formation of *cis*-$[Pt^{II}(NH_3)_2Cl(1\text{-MeTH})]^+$, 8.8 kcal/mol for formation of *cis*-$[Pt^{II}((NH_3)_2(H_2O)(1\text{-MeTH})]^+$), thus such pathway is ruled out. The species T* can only participate in mispairing if the Pt-N3 bond is cleaved.[12a,13a] However, in that case the rare tautomer that is produced thereby, should be stable as long as the next base pairing step takes place. Although this kind of decomposition takes place, at least under acidic condition, it must be questioned if such happens at a biological relevant (physiological) pH value.

All in all, mispairing that stems from a platinated rare tautomer of thymine, in the end leading to a mutation, is not a likely pathway. Although the mispairing event due to a free rare tautomer generation though Pt binding will not be completely ruled out, there may be other mechanisms at work by which Pt^{II} complexes, and notably cisplatin, could cause mutagenesis.[4,49]

4.5 Metal-Stabilised Rare Tautomers of Guanine

In the preceding sections the focus was put on uracil and thymine. However, for other nucleobases the stabilisation of rare tautomer can also be investigated. Complexes of 9-methylguanine (9-MeGH) and Pt^{II} (bound via N1) are known,[10] and the question arises which site (N7, N3 or O6) could be protonated to form a rare tautomer complex. The proton comes from N1, and is liberated by Pt^{II}-binding. To answer the question, exactly the same procedure has been applied as with 1-MeUH and 1-MeTH, but the study has been restrained to one set of ligands, A = Cl^-, B = NH_3 and C = NH_3, that are *trans*-oriented. Complexes under investigation are termed **19a-d** and are shown in Scheme 4.7.



Scheme 4.7. Possible structures of *trans*-$[Pt(NH_3)_2(Cl)(9\text{-MeGH-}N1)]^+$ complex.

The reaction energies are shown in Table 4.12 and it can be seen that the formation of all products is possible in both the gas phase and water solution. In both phases the reaction energy for **19d** is the most exothermic, and the proton will thus reside at N7. Note, however, that the reaction energy has decreased by more than 50%. It is not very surprsing that the proton binds N7 of G, because it is also the best metal-binding site.[19] The rare tautomer form in which G occurs has also been found by NMR.[10a]

Table 4.12. Reaction energy (in kcal/mol) in water for the exchange of H_2O by 9-MeGH in $[Pt(NH_3)_2(OH_2)]^+$ forming *trans*-$[Pt^{II}(Cl)(NH_3)(NH_3)(9\text{-MeGH-}N1)]^+$.

	Gas phase	Solution
19a	–27.0	–2.5
19b	–23.9	–3.4
19c	–16.9	–6.1
19d	–31.5	–12.1

4.6 Conclusions

In this chapter light was shed on how rare tautomers of 1-MeUH and 1-MeTH can be stabilised. To this end, reaction energies of the water exchange of $[Pt^{II}(A)(B)(C)(OH_2)]^q$ (A, B, C = NH_3, Cl^-) by 1-MeUH have been calculated using DFT at the ZORA-BP86/TZ2P level. It has been found that in the gas phase the order of the stabilisation of a given reaction product is governed by the charge q of the Pt^{II} complex. So, all tautomers were stabilised best by reaction with the dipositive species $[Pt^{II}(NH_3)_3(OH_2)]^{2+}$, and this number diminished by decreasing the charge stepwise from +2 to –1. Treatment of solvent effects (water) changed the gas phase result completely. The charge becomes almost irrelevant, as variations in the reaction energy are only a few kcal/mol. Moreover, only few tautomers are stabilised, notably the N3 coordinated systems.

As a major result it has been found, as an extension of an earlier result by Deubel, that with regard to the gas phase the exchange reaction in water is less favourable for positively charged complexes, but more favourable when negatively charged are applied. This strongly advocates the inclusion of solvent effect when tautomerisation of DNA nucleobases is concerned, as these may strongly alter trends in relative stability observed in the gas phase.

The computations indicate that N3 coordination of Pt^{II} to either 1-MeUH or 1-MeTH is preferred over O2 and O4, be it in either *anti* or *syn* orientation. Although this result is not surprising for its N3 deprotonated analoga, it is more important to stress that this stabilisation is also possible when the neutral nucleobase is concerned. The latter thus implies that a rare tautomer of either 1-MeUH or 1-MeTH is present in its 4-hydroxo-2-oxo species.

For 9-MeGH the tautomerisation has been investigated as well. In case a Pt^{II}-complex (*trans*-$[Pt^{II}(Cl)(NH_3)(NH_3)(OH_2)]^+$) binds to N1, the proton that was initially bound there becomes liberated. Calculations shown that the liberated proton will be bound to N7, thereby generating a rare form that was also experimentally observed.[10]

When not coordinated to Pt^{II}, the rare tautomer species are very rare ($K_{Taut} \sim 10^{-5}$).[8] So, the earlier coined term “metal-stabilised rare tautomer” is thus very appropriate, and despite their low abundance, rare tautomers are chemically relevant.[50]

4.7 References

[1] J. Elguero, C. Marzin, A. R. Katritzky, P. Linda, *The Tautomerism of Heterocycles*, Acad. Press, New York, **1976**.

[2] (a) M. Hanus, M. Kabelác, J. Rejnek, F. Ryjácek, P. Hobza, *J. Phys. Chem. B* **2004**, *108*, 2087-2097.
(b) S. A. Trygubenko, T. V. Bogdan, M. Rueda, M. Orozco, F. J. Luque, J. Sponer, P. Slavícek, P. Hobza, *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192-4203.
(c) J. Rejnek, M. Hanus, M. Kabelác, F. Ryjácek, P. Hobza, *Phys. Chem. Chem. Phys.* **2005**, *7*, 2006-2017.
d) M. Hanus, M. Kabelác, D. Nachtigallová, P. Hobza, *Biochemistry* **2005**, *44*, 1701-1707.
(e) C. Colominas, F. J. Luque, M. Orozco, *J. Am. Chem. Soc.* **1996**, *118*, 6811-6821.
(f) C. Alhambra, F. J. Luque, J. Estelrich, M. Orozco, *J. Org. Chem.* **1995**, *60*, 969-976.
(g) J. R. Blas, F. J. Luque, M. Orozco, *J. Am. Chem. Soc.* **2004**, *126*, 154-164.
(h) N. Russo, M. Toscano, A. Grand, *J. Phys. Chem. A* **2003**, *107*, 11533-11538.
(i) N. Russo, M. Toscano, A. Grand, *J. Phys. Chem. B* **2001**, *105*, 4735-4741.
(j) N. Russo, M. Toscano, A. Grand, *J. Am. Chem. Soc.* **2001**, *123*, 10272-10279.
(k) A. M. Lamsabhi, M. Alcamí, O. Mó, M. Yáñez, *Chem. Phys. Chem.* **2003**, *4*, 1011-1016.
(l) A. Martínez, *J. Chem. Phys.* **2005**, *123*, 024311.
(m) M. Kabelác, P. Hobza, *J. Phys. Chem. B* **2006**, *110*, 14515-14523.
(n) D. Kosenkov, L. Gorb, O. V. Shishkin, J. Sponer, J. Leszczynski, *J. Phys. Chem. B* **2008**, *112*, 150-157.

[3] M. D. Topal, J. R. Fresco, *Nature* **1976**, *263*, 285-289.

[4] J. Müller, R. K. O. Sigel, B. Lippert, *J. Inorg. Biochem.* **2000**, *79*, 261-265.

[5] B. Lippert, *Prog. Inorg. Chem.* **2005**, *54*, 385-447, and references therein.

[6] J. V. Burda, J. Sponer, J. Leszczynski, *J. Biol. Inorg. Chem.* **2000**, *5*, 178-188.

[7] I. L. Zilberberg, V. I. Avdeev, G. M. Zhidomirov, *J. Mol. Struct. (Theochem)* **1997**, *418*, 73-81.

[8] B. Lippert, H. Schöllhorn, U. Thewalt, *Inorg. Chim. Acta* **1992**, *198-200*, 723-732.

[9] (a) F. Zamora, M. Kunsman, M. Sabat, B. Lippert, *Inorg. Chem.* **1997**, *36*, 1583-1587.
(b) J. Sponer, J. E. Sponer, L. Gorb, J. Leszczynski, B. Lippert, *J. Phys. Chem. A* **1999**, *103*, 11406-11413.
(c) D. V. Deubel, *J. Am. Chem. Soc.* **2008**, *130*, 665-675.
(d) M. Kang, H. T. Chifotides, K. R. Dunbar, *Biochemistry* **2008**, *47*, 2265-2276.
(e) A. C. G. Hotze, M. E. T. Broekhuisen, A. H. Velders, K. van der Schilden, J. G. Haasnoot, J. Reedijk, *Eur. J. Inorg. Chem.* **2002**, 369-376.

[10] (a) G. Frommer, I. Mutikainen, F. J. Pesch, E. C. Hillgeris, H. Preut, B. Lippert, *Inorg. Chem.* **1992**, *31*, 2429-2434.
(b) B. Müller, Dissertation, TU Dortmund, **2009**.
(c) G. Admiraal, J. L. van der Veer, R. A. G. de Graaff, J. H. J. den Hartog, J. Reedijk, *J. Am. Chem. Soc.* **1987**, *109*, 592-594.
(d) D. B. Pedersen, B. Simard, A. Martinez, A. Moussatova, *J. Phys. Chem. A* **2003**, *107*, 6464-6469.

[11] (a) B. Lippert, H. Schöllhorn, U. Thewalt, *J. Am. Chem. Soc.* **1986**, *108*, 6616-6621.
(b) F. Pichierri, D. Holthenrich, E. Zangrando, B. Lippert, L. Randaccio, *J. Biol. Inorg. Chem.* **1996**, *1*, 439-445.

(c) J. Müller, E. Zangrando, N. Pahlke, E. Freisinger, L. Randaccio, B. Lippert, *Chem. Eur. J.* **1998**, *4*, 397-405,
(d) P. J. Sanz Miguel, P. Lax, M. Willermann, B. Lippert, *Inorg. Chim. Acta* **2004**, *357*, 4552-4561.

[12] (a) H. Schöllhorn, U. Thewalt, B. Lippert, *J. Am. Chem. Soc.* **1989**, *111*, 7213-7221.
(b) F. Zamora, H. Witkowski, E. Freisinger, J. Müller, B. Thormann, A. Albinati, B. Lippert, *J. Chem. Soc., Dalton Trans.* **1999**, 175-182.

[13] (a) B. Lippert, *Inorg. Chim. Acta* **1981**, *55*, 5-14.
(b) O. Renn, B. Lippert, A. Albinati, *Inorg. Chim. Acta* **1991**, *190*, 285-289,
(c) I. Mutikainen, O. Renn, B Lippert, unpublished results.

[14] (a) W. Brüning, E. Freisinger, M. Sabat, R. K. O. Sigel, B. Lippert, *Chem. Eur. J.* **2002**, *8*, 4681-4692.
(b) W. Brüning, I. Ascaso, E. Freisinger, M. Sabat, B. Lippert, *Inorg. Chim. Acta* **2002**, *339*, 400-410.

[15] (a) B. Lippert, *J. Raman Spectrosc.* **1980**, *9*, 324-333.
(b) R. Pfab, P. Jandik, B. Lippert, *Inorg. Chim. Acta* **1982**, *66*, 193-204.
(c) I. Escorihuela, L. R. Falvello, M. Tomás, E. P. Urriolabeitia, *Cryst. Growth Des.* **2004**, *4*, 655-657.

[16] D. Gupta, M. Huelsekopf, M. Morell Cerdà, R. Ludwig, B. Lippert, *Inorg. Chem.* **2004**, *43*, 3386-3393.

[17] (a) J. K.-C. Lau, D. V. Deubel, *J. Chem. Theory Comput.* **2006**, *2*, 103-106.
(b) J. V. Burda, M. Zeizinger, J. Leszczynski, *J. Comput. Chem.* **2005**, *26*, 907-914.

[18] G. Th. de Jong, F. M. Bickelhaupt, *J. Chem. Theory Comput.* **2006**, *2*, 322-335.

[19] B. Lippert, *Coord. Rev. Chem.* **2000**, *200-202*, 487-516.

[20] (a) J. A. Carrabine, M. Sundaralingam, *Biochemistry* **1971**, *10*, 292-299.
(b) S. Mansy, R. S. Tobias, *Inorg. Chem.* **1975**, *14*, 287-291.
(c) M. Goodgame, K. W. Johns, *Inog. Chim. Acta* **1978**, *30*, L335-L337.
(d) B. A. Cartwright, M. Goodgame, K. W. Johns, A. C. Skapsi, *Biochem. J.* **1978**, *175*, 337-339.
(e) B. Fischer, H. Preut, B. Lippert, H. Schöllhorn, U. Thewalt, *Polyhedron* **1990**, *9*, 2199-2204.

[21] (a) M. A. Young, B. Jayaram, D. L. Beveridge, *J. Am. Chem. Soc.* **1997**, *119*, 59-69
(b) X. Shui, L. McFail-Isom, G. G. Hu, L. D. Williams, *Biochemistry* **1998**, 37, 8341-8355.
(c) V. Tereshko, G. Minasov, M. Egli, *J. Am. Chem. Soc.* **1999**, *121*, 3590-3595.

[22] B. L. Kindberg, E. H. Griffith, E. L. Amma, *J. Chem. Soc., Chem. Commun.* **1975**, 195-196.

[23] (a) R. G. Pearson, *J. Am. Chem. Soc.* **1963**, *85*, 3533-3539.
(b) R. G. Pearson, *J. Chem. Educ.* **1968**, *45*, 581-586.
(c) R. G. Pearson, *J. Chem. Educ.* **1968**, *45*, 643-648.

[24] (a) E. Poverenov, I. Efremenko, A. I. Frenkel, Y. Ben-David1, L. J. W. Shimon, G. Leitus, L. Konstantinovski, J. M. L. Martin, David Milstein, *Nature* **2008**, *455*, 1093-1096.
(b) C. Limberg, *Angew. Chem. Int. Ed.* **2008**, *48*, 3357-3361.

[25] R. M. K. Dale, D. C. Livingston, D. C. Ward, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 2238-2242.

[26] H. Schöllhorn, U. Thewalt, B. Lippert, *J. Chem. Soc., Chem. Commun.* **1986**, 258-260.

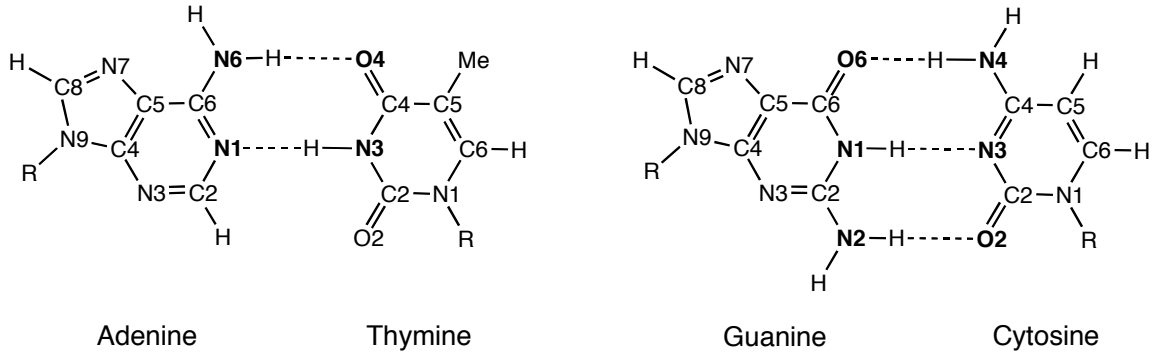
[27] (a) M. Höpp, A. Erxleben, I. Rombeck, B. Lippert, *Inorg. Chem.* **1996**, *35*, 397-403.
(b) L. Holland, Dissteration TU Dortmund, **2008**.

[28] F. Zamora, E. Zangrando, M. Furlan, L. Randaccio, B. Lippert, *J. Organomet. Chem.* **1998**, *552*, 127-134.
[29] D. V. Deubel, *J. Am. Chem. Soc.* **2002**, *124*, 5834-5842.
[30] K. M. Anderson, A. H. Orpen, *Chem. Commun.* **2001**, 2682-2683.
[31] B. Lippert, *Prog. Inorg. Chem.* **1989**, *37*, 1-97.
[32] L. D. Kosturko, C. Folzer, R. F. Steward, *Biochemistry* **1974**, *13*, 3949-3952.
[33] (a) G. Bandoli, G. Trovò, A. Dolmella. B. Longato, *Inorg. Chem.* **1992**, *31*, 45-51.
(b) M. Grehl, B. Krebs, *Inorg. Chem.* **1994**, *33*, 3877-3885.
[34] F. Lianza, A. Albinati, B. Lippert, *Inorg. Chim. Acta* **1997**, *255*, 313-318.
[35] H. Chen, M. M. Olmstead, M. F. Meastre, R. H. Fish, *J. Am. Chem. Soc.* **1995**, *117*, 9097-9098.
[36] R. Faggiani, H. E. Howard-Lock, C. J. L. Lock, M. A. Turner, *Can. J. Chem.* **1987**, *65*, 1568-1575.
[37] W. Micklitz, B. Lippert, G. Müller, P. Mikulcik, J. Riede, *Inorg. Chim. Acta* **1989**, *165*, 57-64.
[38] (a) S. J. Berners-Price, T. A. Frenkiel, U. Frey, J. D. Ranford, P. D. Sadler, *J. Chem. Soc., Chem. Commun.* **1992**, 789-791.
[39] B. Lippert, D. Neugebauer, *Inorg. Chim. Acta* **1980**, *46*, 171-179.
[40] H. Schöllhorn, U. Thewalt, B. Lippert, *J. Chem. Soc., Chem. Commun.* **1984**, 769-770.
[41] (a) B. Lippert in *Cisplatin-Chemistry and Biochemistry of a Leading Anticancer Drug* (Ed.: B. Lippert), VHCA Zürich and Wiley-VCH, Weinheim, **1999**, pp. 379-403;
(b) M.-A. Elizondo-Riojas, F. Gonnet, J.-C. Chottard, J.-P. Girault, J. Kozelka, *J. Biol. Inorg. Chem.* **1998**, *3*, 30-43.
(c) J. Vinje, E. Sletten, J. Kozelka, *Chem. Eur. J.* **2005**, *11*, 3863-3871.
[42] C. F. Moreno-Luque, E. Freisinger, B. Costisella, R. Griesser, J. Ochocki, B. Lippert, H. Sigel, *J. Chem. Soc., Perkin Trans. 2* **2001**, 2005-2011.
[43] H. Rauter, E. C. Hillgeris, A. Erxleben, B. Lippert, *J. Am. Chem. Soc.* **1994**, *116*, 616-624.
[44] B. Lippert, *Chem. Biodiversity* **2008**, *5*, 1455-1474.
[45] R. B. Martin, in *Cisplatin-Chemistry and Biochemistry of a Leading Anticancer Drug* (Ed.: B. Lippert), VHCA Zürich and Wiley-VCH, Weinheim, **1999**, pp. 183-205.
[46] Migration may proceed via an N,O-chelate. See for a appropriate structure:
J. Ruiz, M. D. Villa, V. Rodríguez, N. Cutillas, C. Vicente, G. López, D. Bautista, *Inorg. Chem.* **2007**, *46*, 5448-5449.
[47] L. Holland, W.-Z. Shen, W. Micklitz, B. Lippert, *Inorg. Chem.* **2007**, *46*, 11356-11365.
[48] (a) G. J. Narlikar, D. Herschlag, *Annu. Rev. Biochem.* **1997**, *66*, 19-59.
(b) P. C. Bevilacqua, T. S. Brown, S.-i. Nakano, R. Yajima, *Biopolymers* **2004**, *73*, 90-109.
(c) E. M. Moody, J. T. J. Lecomte, P. C. Bevilacqua, *RNA* **2005**, *11*, 157-172.
[49] (a) H. Louro, M. J. Silva, M. G. Boavida, *Environ. Mol. Mutagen.* **2002**, *40*, 283-291.
(b) L. J. N. Bradley, K. J. Yarema, S. J. Lippard, J. M. Essigmann, *Biochemistry* **1993**, *32*, 982-988.
(c) M. J. Pillaire, J. S. Hoffmann, M. Defais, G. Villani, *Biochimie* **1995**, *77*, 803-807.
[50] B. Lippert, D. Gupta, *Dalton Trans.* **2009**, *24*, 4619-4634.

Chapter 5
Dispersion-Corrected DFT Study on the Differences in π-Stacking and Hydrogen-Bonding Behaviour between the Guanine and Adenine Quartet

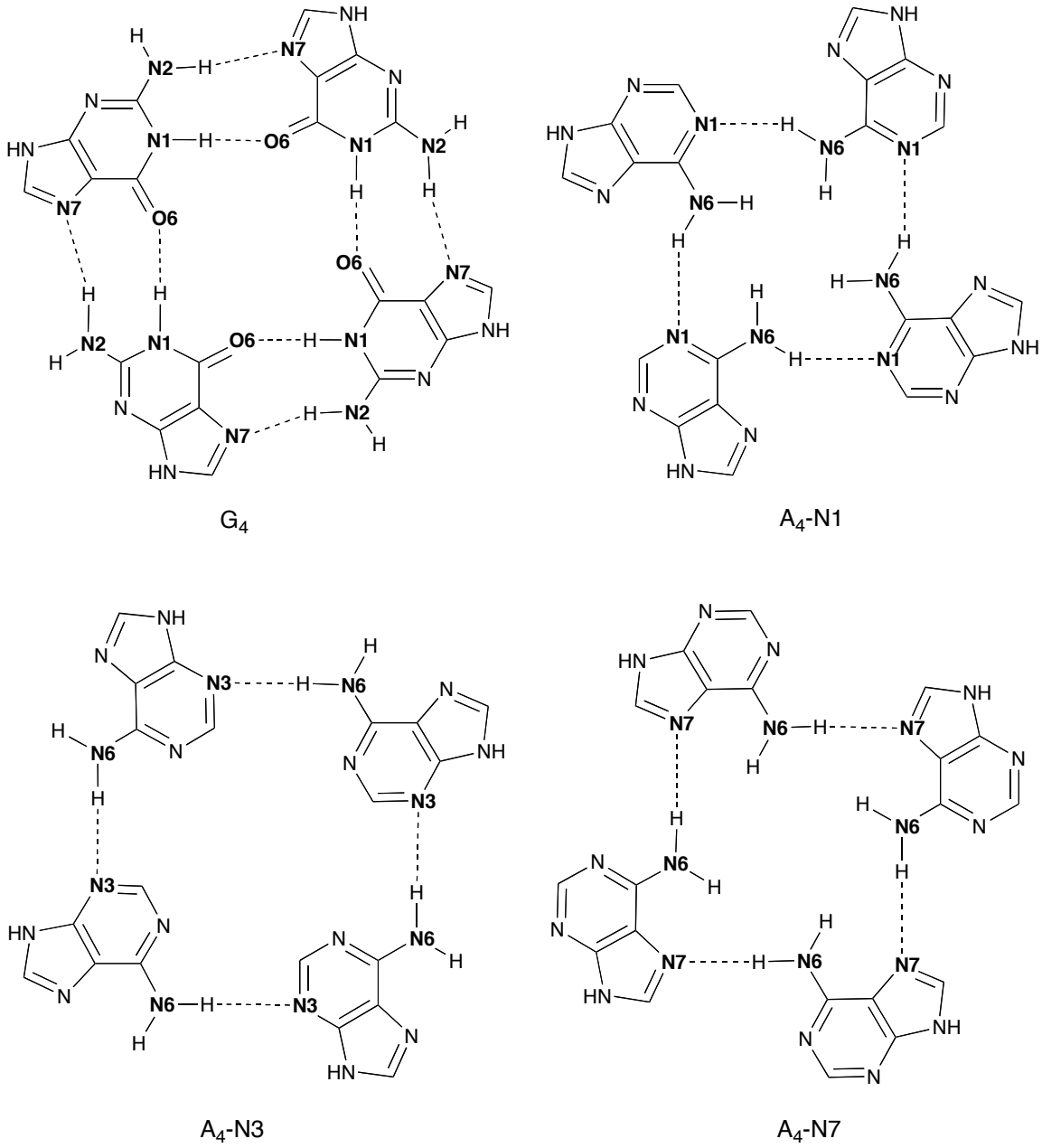
5.1 Introduction

The Watson-Crick base pairs adenine-thymine (AT) and guanine-cytosine (GC) are by far the most famous arrangements (Scheme 5.1) for nucleobases. However, larger aggregates are also found, among which a quartet arrangement is the most common. These quartets can stack together through π-stacking.[1-5] The biological relevance of these quartets is increasingly recognised.



Scheme 5.1. Watson-Crick base pairs AT and GC with atom numbering. Heavy atoms participating in hydrogen bonds are indicated in bold.

The most famous quartet structure is the guanine quartet (G_4), and has an exceptional stability because two hydrogen bonds occur between neighbouring G bases, *i.e.* the N1–O6 and N2–N7 hydrogen bond (see Scheme 5.2). The arrangement of the neighbouring bases is Hoogsteen-like. Stabilisation of G_4 quartets is brought about notably by the alkali cations Na^+ and K^+.[2] Because of the relative high stability of the G_4 quartet, it can serve as a scaffold to stack other quartets on, for instance an adenine (A_4) quartet.[4] There is experimental evidence for three kinds of A_4 quartets, which are sandwiched between two G_4 quartets.[4] Contrary to the G_4 quartet, two neighbouring bases of an A_4 quartet are not connected via *two* but via only *one* hydrogen bond between the bases. The three A_4 quartets have a common proton donor, namely N6–H of the amino-group, but different acceptor sites: the N1 atom (A_4-N1),[4a,b] the N3 atom (A_4-N3),[4c] or the N7 atom (A_4-N7),[4d] see Scheme 5.2.



Scheme 5.2. The quartets used in this study. Heavy atoms participating in hydrogen bonding are shown in bold.

Apart from the experimental work on the G_4 and A_4 quartet, computational studies are available for both quartets.[6,7] This chapter deals with the stability of these four quartets in both the gas phase as well as in aqueous solution using dispersion-corrected density functional theory (DFT-D), as developed by Grimme,[8] see also chapter 2. First the performance of several DFT-D functionals is investigated for several molecular sets which are grouped in hydrogen bonded, dispersion bonded and other weakly bonded molecules. Doing the same for the hydrogen and dispersion bonded base pairs AT and GC, this

evaluation is sharpened. Results of the hydrogen bonded AT and GC have already been discussed in detail in chapter 3, but only for dispersion-*un*corrected DFT. For the dispersion bonded molecules, it is very interesting what the performance of the dispersion corrected functionals will be, because most uncorrected density functionals fail for π-π interactions.[9] The best dispersion-corrected functional will then be used to investigate the G_4 and A_4 quartets in several symmetrical arrangements, namely C_{4h} (planar), C_4 (bowl-shaped) and S_4, under the constraint that the hydrogen bond pattern as shown in Scheme 5.2 remains the same. Experimentally it is found, that A_4 quartets are planar, by stacking with a G_4 quartet. The question that arises, is what may happen if this stacking partner is taken away.

Finally the performance of the widely used functional B3LYP is also assessed, and it is reconfirmed that this functional will lead to wrong results, because it does not deal properly with π-stacking.

5.2 Summary of Computational Methods

All calculations were done with QUILD, a wrapper around ADF. Geometry optimisations were performed at the BLYP-D/TZ2P, BP86-D/TZ2P, PBE-D/TZ2P and M06-2X/TZ2P level of theory. These functionals should be able to treat dispersion in the DFT framework.

In the first part of this chapter, a lot of bond energies (ΔE_{Bond}) of several complexes are computed according to Eq. 5.1.

$$\Delta E_{Bond} = E_{Complex} - \Sigma E_{Fragments} \qquad \text{(Eq 5.1)}$$

Thus, the bond energy is defined as the difference between the energy of the complex ($E_{Complex}$) and the sum of the energies of its constituting fragments ($\Sigma E_{Fragments}$). In some cases an interaction energy (ΔE_{Int}) is calculated, which is done in the same way as done for ΔE_{Bond}, but the energy of the fragments is calculated by a single point calculation and not by an geometry optimisation, thus, a given geometry is used. The geometry of the fragments, in this case, is the one they have in the complex. The hydrogen bond energy (ΔE_{HB}) or stacking energy (ΔE_{Stack}) of the base pairs AT and GC is computed according Eq. 5.2.

$$\Delta E_{HB} \text{ or } \Delta E_{Stack} = E_{Complex} - E_{Base\ 1} - E_{Base\ 2} \qquad \text{(Eq. 5.2)}$$

The terms in this equation have the same meaning as in Eq. 3.1 and the complex is either a hydrogen bonded or a stacked base pair. Hydrogen bonded base pairs were optimised in C_S symmetry. The stacked base pairs as well as the constituting bases were optimised without any symmetry constraint, C_1.

The hydrogen bond energy of G_4 or A_4 is defined as follows:

$$\Delta E_{HB} = E_{Quartet} - 4*E_{Base} \qquad \text{(Eq. 5.3)}$$

The terms on the right hand side of the equation are the energy of A_4 and G_4 ($E_{Quartet}$) in a given symmetry (C_{4h}, C_{4h}, S_4) and the energy of the constitiuting bases (E_{Base}), either adenine or guanine. At last a planarisation energy (ΔE_{Plan}) is calculated, which is defined as

$$\Delta E_{Plan} = E_{Quartet,\ C4h} - E_{Quartet,\ global\ minimum} \qquad \text{(Eq. 5.4)}$$

This planarisation energy shows how much energy is needed to make a quartet planar with respect to its global minimum. $E_{Quartet,\ C4h}$ is the energy of a fully planar quartet, and $E_{Quartet,\ global\ minimum}$ is the energy of the quartet in its global minimum. The global minimum corresponds to the structure of A_4 or G_4 that is the most stable, *i.e.*, has the most negative hydrogen bond energy.

All energy minima of hydrogen-bonded AT and GC pairs, stacked AT and GC dimers, and the global energy minima of DNA-base quartets, were verified in the gas phase and in water to be equilibrium structures through vibrational analysis, albeit only at the BLYP-D/TZ2P level of theory.

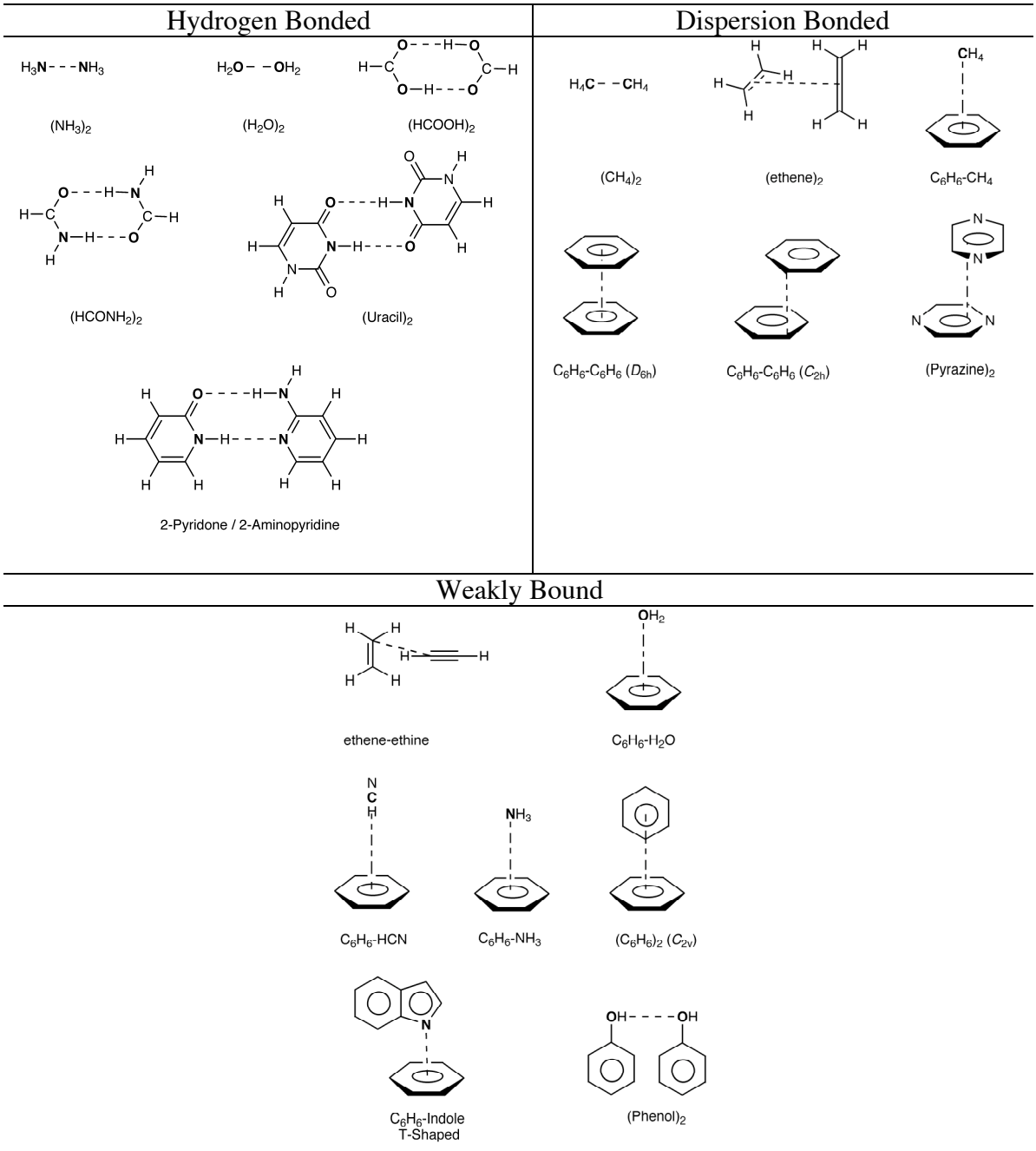
The basis superposition error (BSSE) is already absorbed in the dispersion-corrected functionals. This is not true for the M06-2X functional. The BSSE has been included afterwards, in case of the base pairs AT and GC, the bases were used as fragments.

Solvent effects were treated implicitly using the COSMO method, and the dispersion correction should not be modified when sovent effects are treated. Ample details on the computational settings can be found in chapter 2.

5.3 Results and Discussion

5.3.1 Test Sets

In Tables 5.1-5.3 DFT-D results of binding and interaction energies are discussed for the different molecular sets and compared to accurate *ab initio* data.[10] The sets contain hydrogen bonded (Table 5.1), dispersion bonded (Table 5.2) and weakly bound (Table 5.3) dimers. To judge which dispsersion corrected functional performs best, the Mean Absolute Error (MAE) with respect to the *ab initio* data was calculated. Interestingly, the MAE for the first set is an order of magnitude larger than for the other two. This is not surprising because the DFT-D functionals are not tailored for hydrogen bonded systems, the lowest MAE is found for BLYP-D and amounts to 1.9 kcal/mol. For the set of dispersion and weakly bonded molecules, lowest MAE values are obtained with BLYP-D and BP86-D and amount each 0.2 kcal/mol, all respectively. However, the performance on the base pairs AT and GC is, especially for the stacked ones, decisive for a proper choice.



Scheme 5.3 Schematic repsresentation of the molecules that have been used to test the performance of BLYP-D, BP86-D and PBE-D. If needed, atoms between which the distance is determined, are indicated in bold.

Table 5.1. Bond, ΔE_{Bond}, and interaction energy, ΔE_{Int}, (in kcal/mol) and hydrogen-bond distances[a1] (in Å) for various hydrogen-bonded complexes.

	DFT-D			
	BLYP-D	BP86-D	PBE-D	*ab initio*
$(NH_3)_2$ (C_{2h})				
ΔE_{Bond}	–3.6	–3.4	–4.0	–3.0[c]
R(N-N)	3.17	3.13	3.13	3.17[c]
$(H_2O)_2$ (C_S)				
ΔE_{Bond}	–5.4	–5.3	–5.9	–4.8[b]
R(O-O)	2.90	2.86	2.87	2.92[d]
$(HCOOH)_2$ (C_{2h})				
ΔE_{Bond}	–17.7	–19.2	–19.5	–13.9[b]
R(O-O)	2.67	2.61	2.61	2.66[d]
$(HCONH_2)_2$ (C_{2h})				
ΔE_{Int}	–17.3	–18.6	–18.8	–16.0[e]
R(N-O)	2.85	2.81	2.82	2.86[e]
$(Uracil)_2$ (C_{2h})				
ΔE_{Int}	–21.8	–23.3	–23.4	–20.7[e]
R(N-O)	2.78	2.74	2.74	2.80[e]
2-Aminopyridone 2-Aminopyridine (C_1)				
ΔE_{Int}	–20.0	–21.9	–21.9	–16.7[e]
R(N-O)	2.83	2.78	2.79	2.90[e]
R(N-N)	2.83	2.79	2.80	2.90[e]
MAE (ΔE)[f]	1.9	2.9	3.2	

[a] Definition of distances can be found in Scheme 5.3.
[b] Reference data taken from ref. 8a.
[c] Reference data taken from ref. 10a. Energy calculated at the MP4 level of theory.
[d] Reference data taken from ref. 10b. Energy calculated at the CCSD(T) level of theory.
[e] Reference data taken from ref. 10c. Geometries and energies are calculated at the MP2 or CCSD(T) level of theory. See reference for the level of theory.
[f] Mean Absolute Error (in kcal/mol) between DFT-D and MP2.

Table 5.2. Bond, ΔE_{Bond}, and interaction, ΔE_{Int}, energies (in kcal/mol) and intermolecular distances[a] (in Å) for various van der Waals complexes.

	DFT-D			
	BLYP-D	BP86-D	PBE-D	*ab initio*
$(CH_4)_2$ (D_{3d})				
ΔE_{Bond}	–0.5	–0.2	–0.9	–0.5[c]
R(C-C)	3.52	3.54	3.53	3.60[c]
$(C_2H_4)_2$ (D_{2d})				
ΔE_{Bond}	–1.8	–1.8	–2.2	–1.3[c]
R(C-C)	3.62	3.57	3.66	3.80[c]
C_6H_6•CH_4 (C_{3v})				
ΔE_{Bond}	–1.6	–1.5	–1.9	–1.6[b]
R	3.70	3.65	3.80	3.62[b]
$(C_6H_6)_2$ (C_{2h})				
ΔE_{Bond}	–2.3	–2.4	–2.6	–2.8[d]
R	3.56	3.41	3.63	3.40[d]
$(C_6H_6)_2$ (D_{6h})				
ΔE_{Bond}	–1.4	–1.1	–0.6	–1.8[d]
R	4.17	4.10	5.28	3.70[d]
$(Pyrazine)_2$ (C_S)				
ΔE_{Int}	–4.2	–4.3	–4.4	–4.4[f]
R[e]	3.32	3.20	3.31	3.26[f]
MAE[g] (ΔE)	0.3	0.4	0.5	

[a] Definition of distances can be found in Scheme 5.3
[b] Reference data is taken from ref. 8a.
[c] Reference taken from ref. 10b. Energies are at the CCSD(T) level of theory.
[d] Reference taken from ref. 10d. Energy calculated at the CCSD(T) level of theory.
[e] Between molecular planes.
[f] Reference data taken from ref. 10c. Geometries and energies are calculated at the MP2 CCSD(T) level of theory. See reference for the level of theory.
[g] Mean Absolute Error (in kcal/mol) between DFT-D and MP2.

Table 5.3. Bond, ΔE_{Bond}, and interaction, ΔE_{Int}, energies (in kcal/mol) and intermolecular distances[a] (in Å) for other weakly bound complexes.

	DFT-D			
	BLYP-D	BP86-D	PBE-D	*ab initio*
Ethene•Ethine(C_{2v})				
ΔE_{Bond}	–1.8	–1.8	–2.3	–1.5[b]
R	3.74	3.71	3.71	3.82[b]
Benzene•H_2O (C_S)				
ΔE_{Bond}	–3.4	–3.5	–3.6	–3.9[c]
R	3.58	3.36	3.82	3.21[c]
Benzene•NH_3 (C_S)				
ΔE_{Bond}	–2.4	–2.4	–2.8	–2.4[b]
R	3.60	3.50	3.62	3.45[b]
Benzene•HCN (C_S)				
ΔE_{Int}	–4.9	–5.1	–5.4	–4.5[d]
R	3.40	3.35	3.38	3.39[d]
$(Benzene)_2$ (C_{2v})				
ΔE_{Bond}	–2.9	–2.9	–3.2	–2.7[e]
R	4.90	4.84	4.93	4.90[e]
Indole•Benzene				
T-Shaped (C_1)				
ΔE_{Bond}	–6.0	–6.2	–6.3	–6.2[f]
R	3.29	3.24	3.28	3.16[f]
$(Phenol)_2$ (C_1)				
ΔE_{Int}	–7.5	–7.6	–7.7	–7.1[d]
R (O–O)	2.84	2.79	2.82	2.89[d]
MAE (ΔE)[g]	0.3	0.2	0.5	

[a] Definition of distances can be found in Scheme 5.3.
[b] Reference data is taken from ref. 8a.
[c] Reference data is taken from ref. 10e.
[d] Reference data is taken from ref. 10c. Geometries and energies are calculated at the MP2 or CCSD(T) level of theory. See reference for the level of theory.
[e] Reference data is taken from ref. 10d. Energy calculated at the CCSD(T) level of theory.
[f] Reference data is taken from ref. 10f.
[g] Mean Absolute Error (in kcal/mol) between DFT-D and MP2.

5.3.2 Hydrogen Bonded base pairs AT and GC

In Table 5.4 and 5.5 the hydrogen bond distances and their energies of the Watson-Crick base pairs AT and GC, respectively, in both the gas phase as well as solution are presented at several different levels of DFT-D together with accurate *ab initio* data.[11] Their structures can be found in Scheme 5.1. The hydrogen bond energy obtained with all dispersion-corrected and the M06-2X functionals is a bit weaker than the CCSD(T) reference value. Of the three dispersion-corrected functionals, the best agreement is found with BLYP-D, although BP86-D and PBE-D both perform reasonably well. The hydrogen bond energies for AT and GC using BLYP-D amount to –16.7 kcal/mol and –30.1 kcal/mol, respectively, which are both 1.3 kcal/mol more stable than the CCSD(T)/aug-cc-pVQZ//RI-MP2/cc-pVTZ reference value.[11] This close agreement between the *ab initio* and disperson-corrected DFT results has been found also by Grimme using a triple or quadruple zeta basis.[8c] In case solvent is treated as well, the hydrogen bond energies decreased to about 50% of their gas phase value. Because the BLYP-D functional performs relatively well, a frequency analysis has been performed to check whether these structures are minima.

Table 5.4. Hydrogen bond distances (in Å) and bond energies (in kcal/mol) for AT computed at various levels of theory.[a]

Method	N6–O4	N1–N3	ΔE_{HB}
Ab initio			
RI-MP2/aug-cc-pVQZ//RI-MP2/cc-pVTZ[b]	2.86	2.83	–15.1
"CCSD(T)/aug-cc-pVQZ"//RI-MP2/cc-pVTZ[b]	2.86	2.83	–15.4
DFT			
BLYP-D	2.89	2.78	–16.7
BP86-D	2.83	2.74	–17.9
PBE-D	2.84	2.75	–18.0
M06-2X[c]	2.91	2.79	–13.5
Inclusion of water			
BLYP-D	2.91	2.82	–9.8
BP86-D	2.85	2.78	–10.8
PBE-D	2.86	2.80	–11.1

[a] Calculations were done in C_s symmetry with a TZ2P basis set.
[b] Data from take from references 11a and 11b. The coupled-cluster energy has been obtained by adding a correction to the MP2 energies. This correction is calculated as a difference between the coupled-cluster energy and the MP2 energy obtained with smaller basis sets.
[c] Minimum was not checked by vibrational analysis. BSSE is calculated afterwards.

Table 5.5. Hydrogen bond distances (in Å) and bond energies (in kcal/mol) for GC computed at various levels of theory.[a]

Method	O6-N4	N1-N3	N2-O2	ΔE
Ab initio				
RI-MP2/aug-cc-pVQZ//RI-MP2/cc-pVTZ[b]	2.75	2.90	2.89	–27.7
"CCSD(T)/aug-cc-pVQZ"//RI-MP2/cc-pVTZ[b]	2.75	2.90	2.89	–28.8
DFT with dispersion				
BLYP-D	2.74	2.89	2.88	–30.1
BP86-D	2.70	2.84	2.83	–31.9
PBE-D	2.70	2.86	2.85	–31.9
M06-2X[c]	2.74	2.89	2.88	–26.5
DFT with dispersion in water				
BLYP-D	2.85	2.90	2.84	–13.6
BP86-D	2.80	2.85	2.79	–14.9
PBE-D	2.82	2.87	2.80	–15.0

[a] Calculations were done in C_S symmetry with a TZ2P basis set.
[b] Data taken from references 11a and 11b. The coupled-cluster energy has been obtained by adding a correction to the MP2 energies. This correction is calculated as a difference between the coupled-cluster energy and the MP2 energy obtained with smaller basis sets.
[c] Minimum was not checked by vibrational analysis, BSSE is calculated afterwards.

5.3.3 Stacked Base Pairs AT and GC

In Table 5.6 the stacking distance and the energies of the stacked base pairs AT and GC, respectively, in both the gas phase as well as solution are presented at several different levels of DFT-D together with accurate *ab initio* data of Jurecka *et al.*.[10c,11b] The DFT-D functionals are in very good agreement with this data, but BLYP-D performs the best. The BLYP-D energies –11.7 kcal/mol and –16.9 kcal/mol for AT and GC, respectively, are in excellent agreement with the energies of –11.6 and –16.9 kcal/mol, again respectively, at the CCSD(T)/aug-cc-pVQZ//RI-MP2/TZVPP level of theory. The M06-2X functional gives a bond energy that is 2 kcal/mol below the CCSD(T) reference value.

As has already been said in chapter 3, the B3LYP functional is very popular but is not working at all for stacking.[9] This is corroborated once again here, evaluating the B3LYP/TZ2P stacking energies with the use of the BLYP-D minima (B3LYP/TZ2P//BLYP-D/TZ2P). Notice that no geometry optimalisation has been done, because this is for B3LYP not possible in ADF, as stated in chapter 2. However, a single point calculation using B3LYP is possible, and is performed on the BLYP-D/TZ2P minimum. This reveals that the stacking energy of AT in the gas phase is 2.1 kcal/mol, for GC this equals –7.5 kcal/mol. The latter energy implies a big attraction, which can be

explained by its geometry, as shown in Figure 5.1. It can be very well seen, that AT is a more or less a parallel stack but GC is not. There is an angle between G and C due to the fact that two hydrogen bonds may be formed. In this respect the AT stack is governed by a pure π–π interaction, in GC this is not the case. Due to the small repulsive binding energy value for AT, it is clear that B3LYP can not deal with stacking. Things get even worse in solution where the stacking energy for AT and GC is 4.2 kcal/mol and 6.0 kcal/mol, respectively. This would imply that in DNA this kind of stacks are not present, which is obviously not the case. In water, the stacking energies are attractive, but are decreased by about 50%. Once again, the BLYP-D functional is performing best, and therefore it has been checked whether the gas phase as well as the solution structure are equilibrium structures, which they are both.

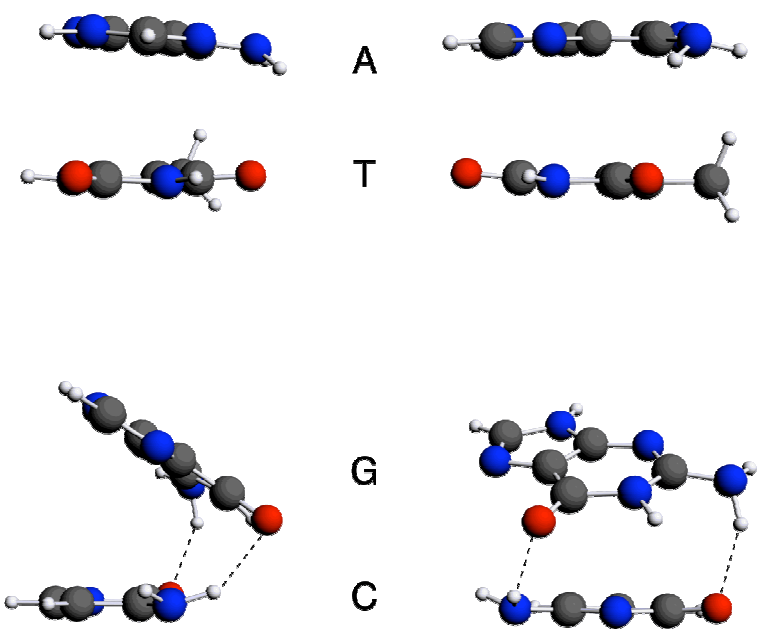


Figure 5.1. Stacked AT and GC dimers in the gas phase shown from two side views (BLYP-D/TZ2P).

Table 5.6. Stacking energies, ΔE_{Stack}, (in kcal/mol) and distances between the bases (in Å) for AT and GC computed at various levels of theory.[a]

	AT		GC	
Method	R(C4–N1)[b]	ΔE_{Stack}	R(N1–C2)[b]	ΔE_{Stack}
Ab initio				
"CCSD(T)/aug-cc-pVQZ"//RI-MP2/cc-pVTZ[c]	3.31	–11.6	3.34	–16.9
DFT-D				
BLYP-D	3.39	–11.7	3.36	–16.9
BP86-D	3.26	–12.3	3.30	–17.4
PBE-D	3.40	–11.2	3.37	–17.0
M06-2X[d]	3.29	–9.6	3.17	–14.9
B3LYP//BLYP-D		2.3		–7.5
DFT-D with inclusion of water				
BLYP-D	3.35	–8.0	3.32	–7.9
BP86-D	3.25	–8.5	3.27	–5.8
PBE-D	3.36	–7.7	3.33	–7.9
B3LYP//BLYP-D		4.2		6.0

[a] Calculations were done in C_1 symmetry with a TZ2P basis set.
[b] For the definition of the distances see reference 8a.
[c] Data taken from refs. 11b. The coupled-cluster energy has been obtained by adding a correction to the MP2 energies. This correction is calculated as a difference between the coupled-cluster energy and the MP2 energy obtained with smaller basis sets.
[d] Minima not checked by vibrational analysis, BSSE corrected afterwards.

5.4 Adenine and Guanine Quartets in the Gas Phase

In Table 5.7 the results of our BLYP-D/TZ2P study on the formation of the G_4, A_4-N1, A_4-N3 and A_4-N7 quartets in the point group symmetries C_{4h}, C_4 and S_4 in the gas phase and in water are summarised.

Table 5.7. Hydrogen bond energies (in kcal/mol) for DNA quartets in the gas phase and in water.

		BLYP-D				B3LYP//BLYP-D			
		C_{4h}	C_4	S_4	S_4 local	C_{4h}	C_4	S_4	S_4 local
Gas phase	G_4	–79.2	C_{4h}	–79.8		–66.4	C_{4h}	–66.3	
	A_4-N1	–25.5	–30.2	–46.0	–31.2	–9.5	–12.1	–9.7	–14.4
	A_4-N3	–33.1	–32.7	–32.8	–25.1	–19.9	–19.1	–10.0	–9.0
	A_4-N7	–31.9	–33.1	–45.5		–19.4	–19.1	–7.4	
Water	G_4	–33.6	C_{4h}	–33.8		–20.0	C_{4h}	–20.3	
	A_4-N1	–10.9	–17.0	–33.2	–15.3	3.0	–1.3	1.3	–0.4
	A_4-N3	–16.8	–16.3	–17.0	–15.3	–6.3	–6.2	–5.7	6.7
	A_4-N7	–16.2	–16.3	–27.2		–6.9	–6.5	6.0	

Pictures of the quartets can be found in Figures 5.2-5.4 for respectively C_{4h}, C_4 and S_4 symmetries. There is a similarity between the geometries in the gas phase and in condensed phase, however not for A_4-N3 in S_4 symmetry, see later. For C_4 symmetry all adenine quartets adopt a nice bowl shape but the G_4 quartet spontaneously converges to a C_{4h} symmetric arrangement. The S_4 structures can be grouped in two classes: G_4 and A_4-N3 adopt a distorted or twisted C_{4h} geometry, but A_4-N1 and A_4-N7 are described as a stack of adenine dimers (Figure 5.4) with one cyclic hydrogen bond according to the pattern in Scheme 5.2. Figure 5.5 shows the gas phase structure of A_4-N3, and its large difference with respect to the other structures can be seen. The number of imaginary frequencies can be found in Table 5.8. Importantly, all global minima, *i.e.* the structure with the most negative binding energy, have no imaginary frequencies and are, therefore, equilibrium structures, which is true for both the gas phase as well as for solution.

The group of Meyer and co-workers have already published data on the G_4 quartet, albeit at the B3LYP level.[7b] Our BLYP-D data give an energy difference of only 0.6 kcal/mol between the C_{4h} and the S_4 structures, which is in line with the data of Meyer. Also for the A_4-N3 quartet good agreement is found with previous work of Meyer.[7a] Our BLYP-D data predict that the C_4 geometry is 0.4 kcal/mol and the S_4 geometry 8.0 kcal/mol higher in energy than the C_{4h} structure. However, our S_4 structure differs from theirs.

For the S_4 structures of A_4-N1 and A_4-N7, it is found that these are 15.8 and 13.6 kcal/mol more stable than the C_{4h} symmetric system. This differs from the B3LYP results of Meyer.[7b] In the work by Meyer, the S_4 arrangement of A_4-N1 structure is a few kcal/mol lower in energy than the C_{4h} or C_4 structure, and for the A_4-N7 quartet all symmetries lead to bond energies that are within a 0.5 kcal/mol equal. However, this difference between this work and the work by Meyer can be explained by means of geometry. The A_4-N1 and A_4-N7 S_4 BLYP-D minima are like stacked adenine dimers, yet the structures that Meyer found are distorted C_{4h} structures. However, an S_4 local minima of A_4-N1 could be localised at the BLYP-D level of theory and was similar to the S_4 structure of Meyer, see Figure 5.6. This local S_4 structure has a binding energy of only –31.2 kcal/mol, and is thus 14.8 kcal/mol less stable than its global minimum. This brought up the question, whether B3LYP could reproduce this result. Therefore, B3LYP/TZ2P bonding energies were calculated with the use of the geometries of the global BLYP-D/TZ2P minima. The stacked S_4 structure has a binding energy of –9.7 kcal/mol and the distorted C_{4h} geometry has a binding energy of –14.4 kcal/mol. Thus, in the first place an underestimation is found. Moreover the stacked S_4 structure is not the global minimum anymore at the B3LYP level, this is now the distorted C_{4h} structure. This means that the B3LYP functional leads to an erroneous energy ordering and to wrong chemical conclusions.

Table 5.8. Number of imaginary frequencies for optimised DNA quartets in the gas phase and in water.[a]

		BLYP-D			
		C_{4h}	C_4	S_4	S_4 local
Gas phase	G_4	b	c	0	
	A_4-N1	b	1	0	0
	A_4-N3	0	0	0	0
	A_4-N7	b	0	0	
Water	G_4	d	c	d	
	A_4-N1	d	d	0	0
	A_4-N3	0	d	0	0
	A_4-N7	d	d	0	

[a] All calculations were done at the BLYP-D/TZ2P level of theory.
[b] These small imaginary frequencies found were shown to be spurious using an explicit potential-energy scan.
[c] C_4 starting geometry collapses spontaneously to C_{4h} structure.
[d] Vibrational analysis not carried out for these species.

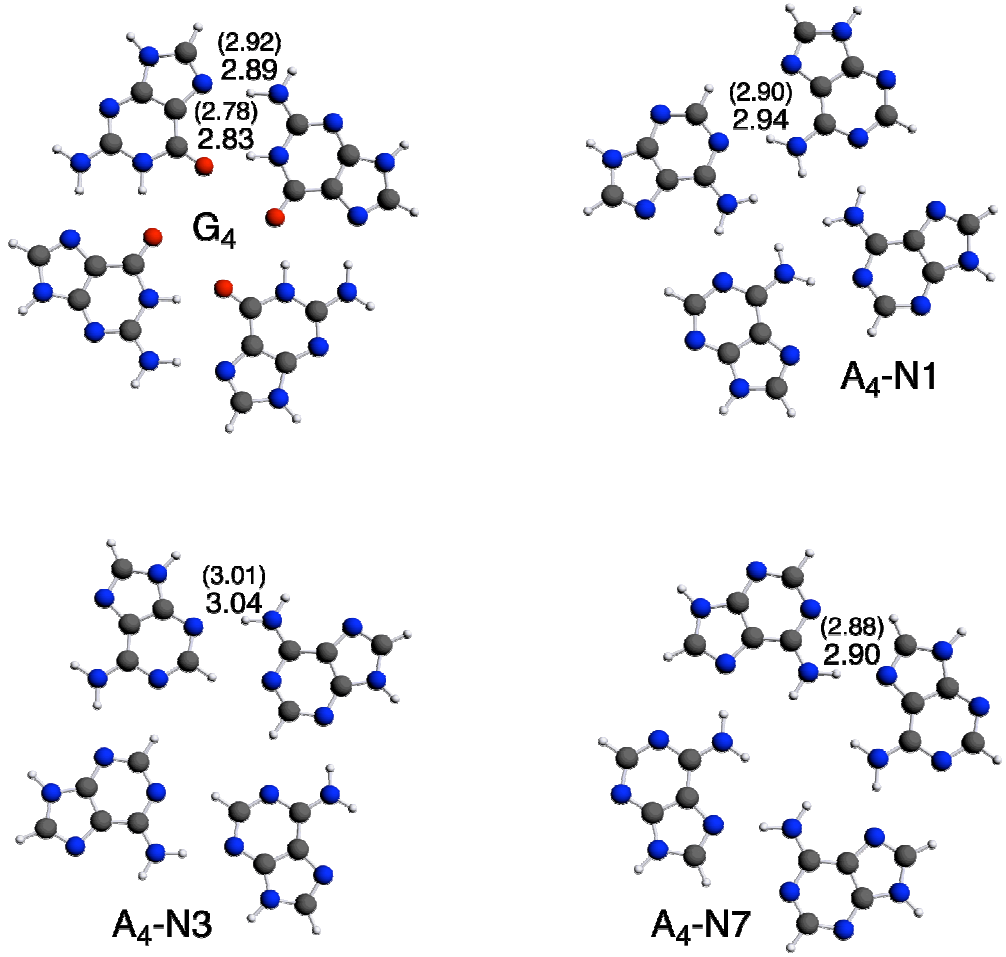


Figure 5.2. The C_{4h} structures in water of G_4, A_4-N1, A_4-N3 and A_4-N7 at the BLYP-D/TZ2P level of theory. Hydrogen-bond distances (in Å) are given for aqueous solution (gas-phase values in parentheses).

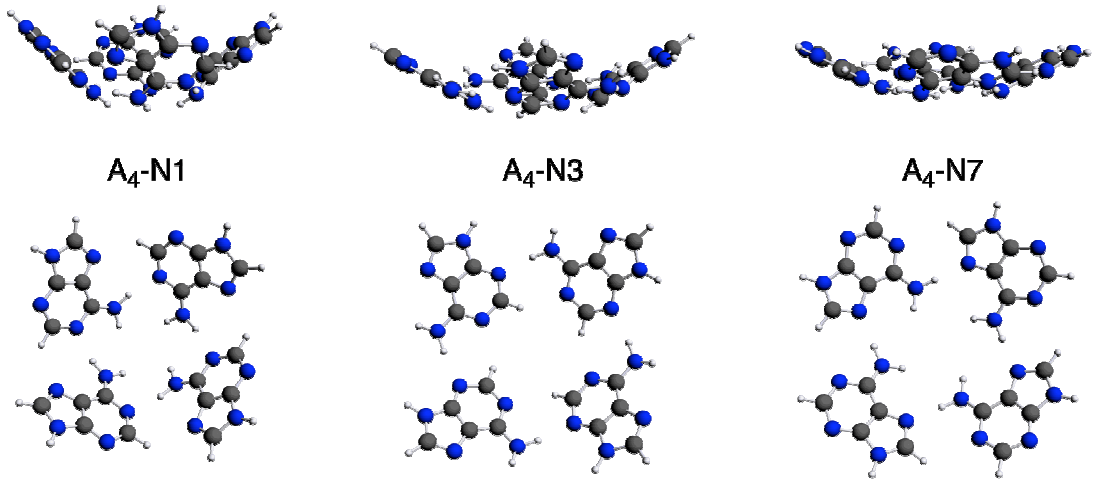


Figure 5.3. Side view and top view of the C_4 structures in water of A_4-N1, A_4-N3 and A_4-N7 at the BLYP-D/TZ2P level of theory.

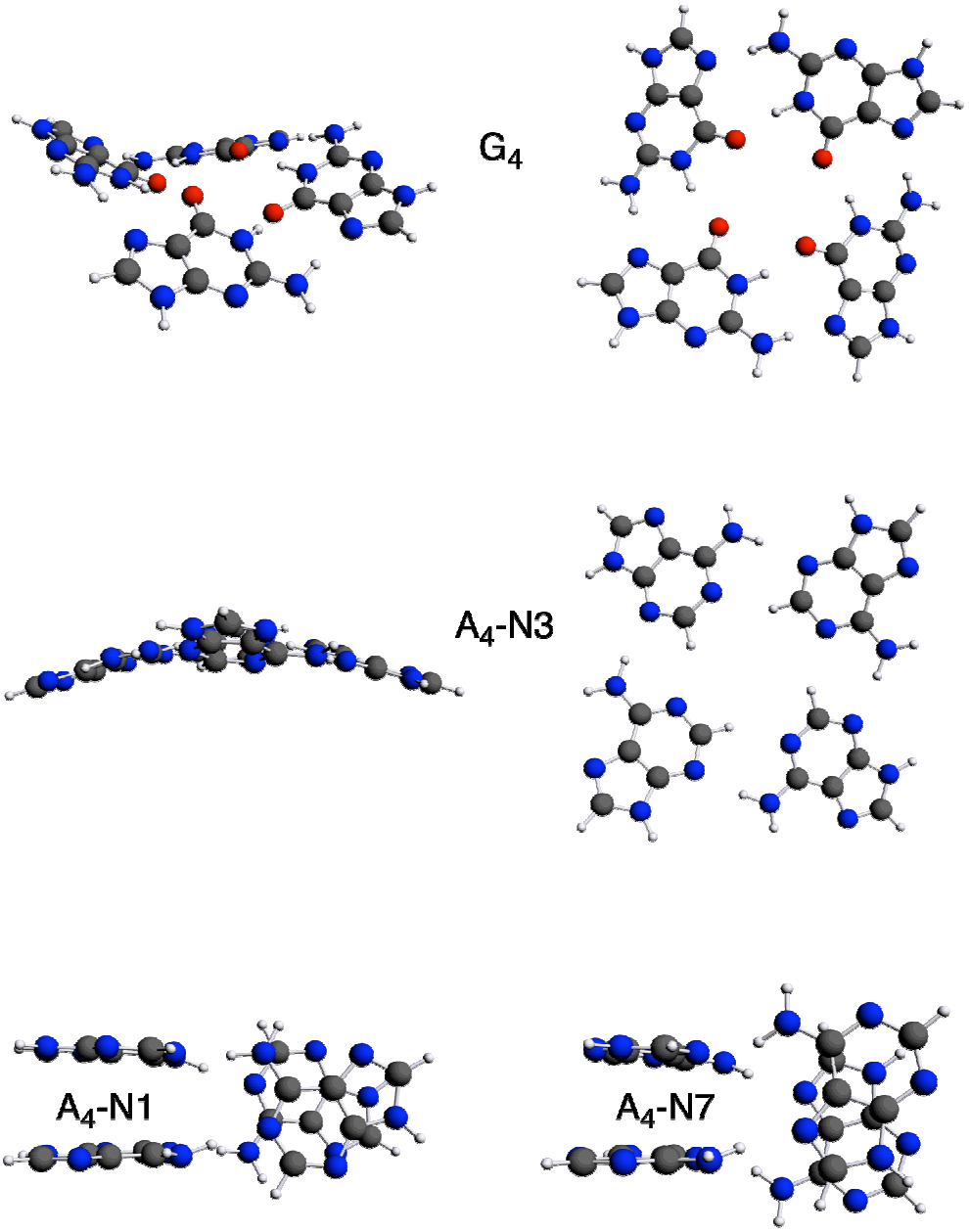


Figure 5.4. Structures of the S_4 global minima in water of G_4, A_4-N1, A_4-N3 and A_4-N7 at the BLYP-D/TZ2P level of theory (for G_4 and A_4-N3 a side and a top view are shown)

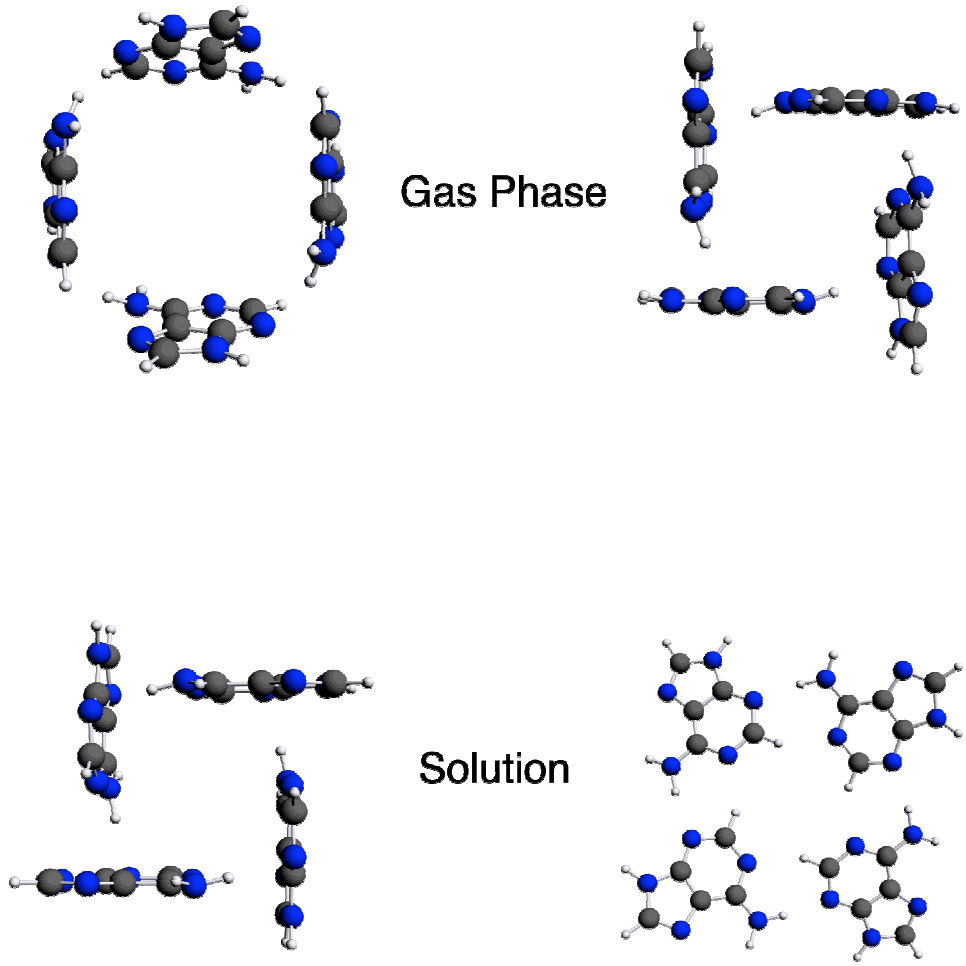
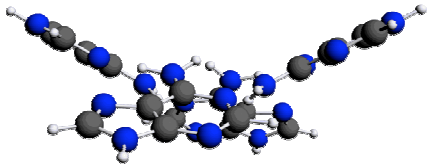


Figure 5.5. Top and side view of the gas phase S_4 structure of A_4-N3 at the BLYP-D/TZ2P level of theory



A_4-N1

Figure 5.6. Side view of the local S_4 minimum in the gas phase of A_4-N1 at the BLYP-D/TZ2P level of theory.

5.5 Adenine and Guanine Quartets Solution: To Stack or Not to Stack?

In water, the results for G_4 show a small energy difference, which amounts to 0.2 kcal/mol between the S_4 and C_{4h} structure. This 0.2 kcal/mol is needed to planarise the G_4 quartet, which is then able to act as stacking scaffold for other quartets. In water the bonding energy has decreased by 50% with respect to the gas phase. For A_4-N1 and A_4-N7 the bonding energy is of the same order of magnitude as found for the G_4 quartet. However the planarisation energies amount to 22.3 kcal/mol and 11.0 kcal/mol for A_4-N1 and A_4-N7 respectively. Due to the concomitant facts that the A_4 quartets have high planarisation energies and are in a planar configuration less stable than the planar G_4, insight is gained into why G_4 quartets were found earlier than A_4 quartets. The A_4 quartet has to be sandwiched between two G_4 quartets.[4] The stacking between these quartets may compensate the planarisation energy. The stacking occurs through the π-π interaction between the aromatic rings.

5.6 Conclusions

In this chapter the suitability of dispersion-corrected density functionals, particularly BLYP-D, for hydrogen bonded as well as stacked AT and GC pairs has been demonstrated. It has been shown that this dispersion-corrected functional is able to reproduce very accurate *ab initio* data of hydrogen bonded and stacked base pairs AT and GC, both bond energy as well as geometry.

Moreover, it has been shown that the B3LYP function is not able to describe π-stacking as it predicts A_4 quartets to adopt planar structures instead of the stacked dimers that were found with BLYP-D. This leads to the recommendation not to use B3LYP when DNA systems are under investigation in which π-stacking plays a large role.

The BLYP-D results predict in water G_4 to be the most stable quartet, which can easily adopt a planar conformation. The stabilities of the planar A_4 quartets in water are a factor 2-3 less stable than the planar G_4. The most stable A_4 quartet is surely not planar. To state it differently, G_4 adopts a geometry that is fine for stacking interactions, whereas the geometry of A_4 is not suited for stacking interactions. This result may be the reason why A_4 quartets are only found between two G_4 quartets. The A_4 quartet needs a lot of energy to become planar, which may be supplied through stacking with G_4 quartets

5.7 References

[1] *Quadruplex Nucleic Acids* (Eds.: S. Neidle, S. Balasubramanian), RSC Publishing, Cambridge, **2006**.

[2] J. T. Davis, *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698 (and references cited therein).

[3] J. Gros, F. Rosu, S. Amrane, A. de Cian, V. Gabelica, L. Lacroix, J. L. Mergny *Nucleic Acids Res.* **2007**, *35*, 3064-3075.

[4] (a) M. S. Searle, H. E. L. Williams, C. T. Gallagher, R. J. Grant, M. F. G. Stevens *Org. Biomol. Chem.* **2004**, *2*, 810-812.
(b) P. K. Patel, A. S. R. Koti, R. V. Hosur, *Nuceic. Acids Res.* **1999**, *27*, 3836-3843.
(c) B. Pan, Y. Xiong, K. Shi, M. Sundaralingam, *Structure* **2003**, *11*, 825-831.
(d) B. Pan, Y. Xiong, K. Shi, J. Deng, M. Sundaralingam, *Structure* **2003**, *11*, 815-823.

[5] (a) M. Roitzsch, B. Lippert, *Angew. Chem. Int. Ed.* **2006**, *45*, 147-150.
(b) M. Roitzsch, B. Lippert, *Inorg. Chem.* **2004**, *43*, 5483-5485.
(c) E. Freisinger, A. Schimanski, B. Lippert, *J. Biol. Inorg. Chem.* **2001**, *6*, 378-389.
(d) Metzger S, B. Lippert, *J. Am. Chem. Soc.* **1996**, *118*, 12467-12468.

[6] (a) H. Liu, J. W. Gauld, *Phys. Chem. Chem. Phys.* **2009**, *11*, 278-287.
(b) M. Meyer, A. Hocquet, J. Sühnel, *J. Comput. Chem.* **2005**, *26*, 352-364.
(c) T. van Mourik, A. J. Dingley, *Chem. Eur. J.* **2005**, *11*, 6064-6079.
(d) J. Gu, J. Leszczynski, *J. Phys. Chem. A* **2002**, *106*, 529-532.
(e) J. Sponer, N. Spacková, *Methods* **2007**, *43*, 278-290.

[7] (a) M. Meyer, J. Sühnel, *J. Phys. Chem. A* **2008**, *112*, 4336-4341.
(b) M. Meyer, C. Schneider, M. Brandl, J. Sühnel, *J. Phys. Chem. A* **2001**, *105*, 11560-11573.
(c) Gu J, J. Leszczynski, *Chem. Phys. Lett.* **2001**, *335*, 465-474.

[8] (a) S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463-1473.
(b) S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787-1799.
(c) J. Antony, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287-5293.
(d) M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, E. Kaxiras, *J. Chem. Phys.* **2001**, *114*, 5149-5155.
(e) X. Wu, M. C. Vargas, S. Nayak, V. Lotrich, G. Scoles, *J. Chem. Phys.* **2001**, *115*, 8748-8757.
(f) Q. Wu, W. Yang, *J. Chem. Phys.* **2002**, *116*, 515-524.
(g) U. Zimmerli, M. Parrinello, P. Koumoutsakos, *J. Chem. Phys.* **2004**, *120*, 2693-2699.

[9] (a) Y. Zhao, D. G. Truhlar, *Phys. Chem. Chem. Phys.* **2005**, *7*, 2701-2705.
(b) M. Swart, T. van der Wijst, C. Fonseca Guerra, F. M. Bickelhaupt, *J. Mol. Model.* **2007**, *13*, 1245-1257.

[10] (a) J. S. Lee, S. Y. Park, *J. Chem. Phys.* **2000**, *112*, 230-237.
(b) S. Tsuzuki, H. P. Lüthi, *J. Chem. Phys.* **2001**, *114*, 3949-3957.
(c) P. Jurecka , J. Sponer, J. Cerny, P. Hobza, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985-1993.
(d) M. O. Sinnokrot, E. F. Valeev, C. D. Sherrill, *J. Am. Chem. Soc.* **2002**, *124*, 10887-10893
(e) D. Feller, *J. Phys. Chem. A* **1999**, *103*, 7558-7561.
(f) J. Braun, H. J. Neusser, P. Hobza, *J. Phys. Chem. A* **2003**, *107*, 3918-3924.

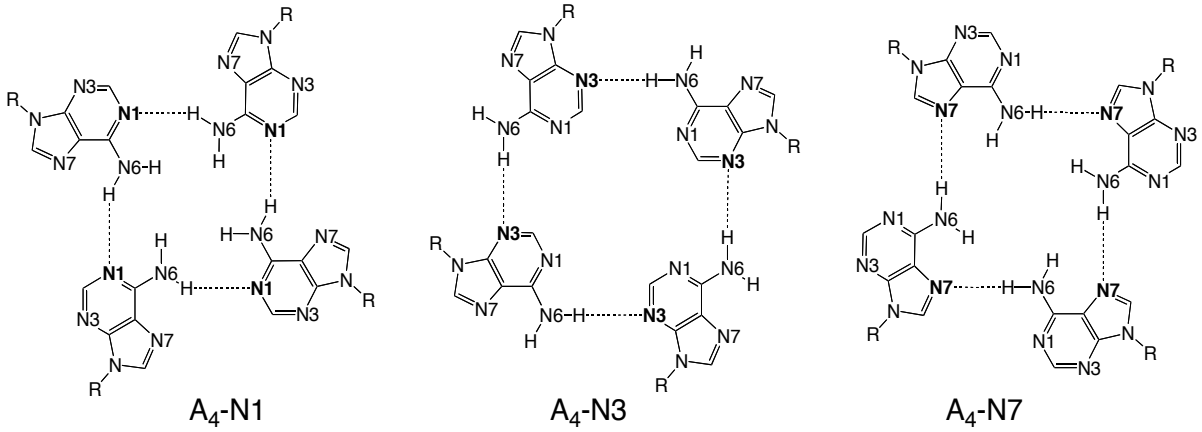
[11] (a) J. Sponer, P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2004**, *126*, 10142-10151.
(b) P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2003**, *125*, 150608-10893.

[12] (a) M. Kabelác, E. C. Sherer, C. J. Cramer, P. Hobza, *Chem. Eur. J.* **2007**, *13*, 2067-2077
(b) M. Kabelác, H. Valdes, E. C. Sherer, C. J. Cramer, P. Hobza, *Phys. Chem. Phys. Chem.* **2007**, *9*, 5000-5008.

Chapter 6
Differential Stabilisation of Adenine Quartets by Anions and Cations

6.1 Introduction

Tetrastrand structures of nucleic acids gain more and more biological significance,[1] and have numerous applications. These comprise, for example, protein recognition and denaturation,[2] recognition of potassium ions,[3] and bio-electronic devices.[4] DNA and RNA quadruplexes consist of nucleobase quartets, in which the bases are essentially coplanar and interact through hydrogen bonds. By far the most stable and at the same time longest known nucleobase quartet is the guanine quartet (G_4). It is arranged in a Hoogsteen fashion and two hydrogen bonds occur, see also chapter 5.[5] Nowadays, quartet structures of all DNA or RNA bases are known.[6] Compared to G_4 these quartets have lower stabilities and they have a larger structural diversity and often use G_4 as scaffold to stack on. For example, the U_4 quartet (U = uracil) may be planar,[7] or saddle-shaped,[8] it may involve different hydrogen bonding patterns,[6] or may even contain a water molecule in the hydrogen bonding scheme.[9] Other mono nucleobase quartets that have been observed are T_4 (T = thymine),[10] C_4 quartet (C = cytosine),[11] and $(isoG)_4$ (isoG = isoguanine).[12] The flexibility of quartets other than G_4 is demonstrated by T_4, which can have two different hydrogen bond patterns.[10]



Scheme 6.1. Representation of A_4-N1, A_4-N3 and A_4-N7. The different acceptor sited are shown in bold. R = H and R = Me refer to adenine and 9-methyladenine, respectively.

Hetero nucleobase quartets, having two different nucleobase species are also known and examples of these are ATAT,[13] (A = adenine) and GCGC.[14] The existence of the AGAG

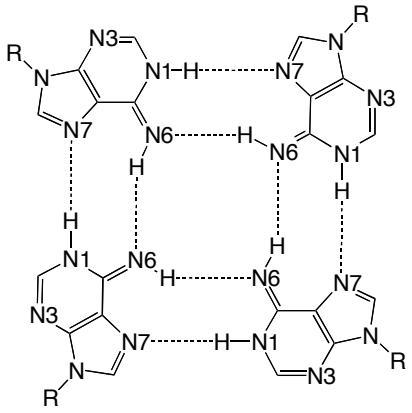
quartet remains surprisingly unclear,[15] although three mispairs of AG are known.[16] The use of two different nucleobase species can even be extented to all four DNA nucleobases: in fact, a GCAT quartet has been observed.[17]

Recently, three variants of adenine quartets (A_4) have been characterised experimentally and these have, as a common feature, four cyclic hydrogen bonds. There is a common donor site, the N-H moiety of the amino group of adenine, but the acceptor site is variable and can be N1 (A_4-N1), N3 (A_4-N3) and N1 (A_4-N7). A_4-N1 has been proven to exist by NMR experiments,[18] whereas A_4-N3,[19] and A_4-N7,[20] have been characterised by X-ray crystallography. Adenines can, instead of being arranged in a quartet, also form a ribbon of adenine pairs,[21] but this feature will not be discussed here. Apart from experiments, all quartets have been subjected to computational studies.[22]

G_4, U_4 and T_4 gain extra stability through the incorporation of metal cations such as Na^+, K^+, NH_4^+, Ca^{2+}, Sr^{2+}, which are located in the center of the quartet *or* can be found in the center of a sandwich of two adjacent quartets.[23] In these quartets, the inward pointing cabonyl groups can interact, through their lone pairs, with the metal cation. Moreover, this interaction will relieve the repulsion between these carbonyl groups.

In case of A_4 two options are possible. These two possibilities rely on the position of the amino groups, they can either point outward (A_4-N3) or inward (A_4-N3 and A_4-N7). To start with, for A_4-N3 an experimental study is known,[18] showing unambiguously the metal-binding possibility, which has also been corroborated by a computational study.[22c] Another study by Pan *et al.*,[20] showed A_4-N7 interacting with a sodium cation (Na^+). This feature prompted us also to look to the possibility whether this could in fact be an anion. The statement that anions can also bind to DNA or RNA structures is not new, and has also been observed experimentally.[24] Anions can also play a role in supramolecular chemistry. For instance, these are found in assemblies of metal-nucleobase assemblies, albeit having the anions at the pheriphery.[25] In the next chapter it will be shown that an ditopic ion-pair receptor based on nucleobase quartets, in theory, can exist.

This chapter deals with a computational study in which the interactions between monovalent anions (F^-, Cl^- and Br^-) and cations (Li^+, Na^+ and K^+) and adenine and 9-methyladenine (9-MeAH) quartets are investigated (A_4-N1, A_4-N3, A_4-N7). These interactions are investigated with dispersion corrected density functional theory (DFT-D), as descibed in chapter 2. Apart from the quartets in Scheme 6.1, a rare tautomer quartet (A_4*) is also taken into account, see Scheme 6.2.



A_4*

Scheme 6.2. Representation of an A_4 quartet having adenine in its most stable rare tautomer conformation denoted as A_4*.

Although there is no experimental evidence for such a rare tautomer quartet, the importance of rare tautomers should not be underestimated.[26] As a last feature the binding of ions at the periphery as opposed to the binding in the central cavity is addressed. This is a common feature for duplex DNA,[24,27] but has recently been found for quadruplex DNA as well.[28]

The chapter is organised as follows: to start with the adenine quartets without ions and followed by by the inclusion of cations and anions to investigate the influence of the ions on the structures and bond energies. The arrangement that the empty quartets and these including ions can adopt are fully planar (C_{4h}) bowl-shaped (C_4(Bowl) and box- or saddle-shaped (S_4). The chapter is ended by simulating the situation in a stack of quartets by keeping the A_4 quartet planar in C_{4h} symmetry with the exception of the amino groups which were allowed to pyramidalise. Furthermore the ion was allowed to move along the C_4-axis ($C_4(R_{vert})$). Calculations were done in both the gas phase as well as aqueous solution.

6.2 Summary of Computational Methods

This chapter can be regarded as a follow-up of the previous chapter. Again, stacking interactions will play an important role in the quartets. Normal DFT does not take into account stacking interactions properly, but dispersion corrected DFT (DFT-D) is very well able to do so. In the previous chapter it was shown that the BLYP-D functional can reproduce very accurate *ab initio* data of stacked base pairs of adenine-thymine and guanine-cytosine.

Geometry optimisations were done at the BLYP-D/TZ2P level of theory. Quartet without ions were optimised in a fully planar arrangement (C_{4h}), bowl-shaped arrangement (C_4(Bowl)), and a saddle-like arrangement (S_4). These arrangements including the ions can be made as well. Finally, a special arrangement was considered: A_4 was kept planar, but the aminogroups were allowed to pyramidalise, and the anion were allowed to optimise its vertical distance to the center of the quartet (R_{vert}). This option can be regarded as a simulation of A_4 in a stack, or oligonucleotide and is denoted as $C_4(R_{vert})$. Note, that in the previous chapter the addition "Bowl" has been kept away.

In these arrangements, the anion or cation binds in the cavity of the quartet, this is referred to as central cavity binding, or cavity binding along the C_4-axis. A cation can also bind to a site at the periphery of the quartet: N3 (A_4-N1, A_4-N7, A_4*) and N7 (A_4-N3), an anion can also interact with the aminogroup (A_4-N3). These latter binding modes are referred to as peripheral binding and such complexes are optimised in planar arrangement (C_S).

All energy values that are given in this chapter are based on structures that are minima. In some cases an imaginary frequency was observed, but was spurious, as found with a potential energy scan.

The hydrogen bond energy of A_4 is defined as in the previous chapter, but as a base adenine (R = H) or 9-methyladenine (R = Me) are used. The bonding for a quartet with an ion is defined as

$$\Delta E_{Bond} = E_{Complex} - 4*E_{Adenine} - E_{Ion} \qquad \text{(Eq. 6.1)}$$

In this equation $E_{Complex}$ is the energy of a quartet with ion in C_{4h}, C_4(Bowl), $C_4(R_{vert})$ or S_4 symetry, $E_{Adenine}$ is the energy of adenine or 9-methyladenine, and E_{Ion} is the energy of the ion.

For calculations of ΔE_{HB} and ΔE_{Bond} for the rare adenine quartet (A_4*), the canonical adenine base is used as a reference. In this way the tautomerisation energy of the four nucleobases is taken into account as well as the complexation between these in A_4*.

If the situation of A_4 with an ion in a stack, in which A_4 is kept planar apart from the amino groups, is simulated or peripheral binding of the ion to A_4 is treated, a stabilisation energy (ΔE_{Stab}) by the ion is calculated defined by:

$$\Delta E_{Stab} = E_{Complex,\ C4(Rvert)} - E_{Quartet,\ C4h} - E_{Ion} \quad \text{(Eq. 6.2a)}$$

$$\Delta E_{Stab} = E_{Complex,\ CS} - E_{Quartet,\ C4h} - E_{Ion} \quad \text{(Eq. 6.2b)}$$

The stabilisation energy indicates what the energy gain or loss is due to the binding of an ion with the respect to the empty A_4 in fully planar arrangement. The energy of the complex of A_4 with ion in the stack is defined by $E_{Complex,\ C4(Rvert)}$ and complex of A_4 with peripheral binding by $E_{Complex,\ CS}$. The energy of A_4 without an ion in a planar geometry is defined by $E_{Quartet,\ C4h}$.

At last a planarisation energy (ΔE_{Plan}) for empty A_4 quartets is calculated exactly as done in chapter 5.

All structures which are calculated for adenenine quartets (R = H) are checked to be minima by vibrational analysis. In case of R = Me this was not possible as the methylgroups can bias the outcome of result.

Solvent effects were treated by the COSMO model as implemented in ADF, but the parameters of the free ions were adapted to reproduce experimental solvation energies. Furthermore, upon treating solvent effects the dispersion correction should not be changed.

In the BLYP-D functional the basis super position error has been incorporate by its construction, and has, therefore, not been calculated separatly. The reader is referred to chapter 2 for a more detaillistic description of the used theoretical methods.

6.3 Results and Discussion

6.3.1 Adenine Quartets without Ions

The survey is started by investigating the bond energies of empty quartets (*i.e.*, A_4-N1, A_4-N3, A_4-N7 and A_4*) in both the gas phase and water solution. Some of the data has been presented in chapter 5 already, but is repeated here for clarity. It is investigated which geometrical shape is the most stable amongs the fully planar, bowl-shaped and

saddle-shaped arrangement which are defined by C_{4h}, C_4(bowl) and S_4, respectively. The bond energies are presented in Table 6.1. Therein, apart from a differentiation between phases, a differentiation between A_4 is made; one group has R = H, one group has R = Me.

To start with, it is mentioned that quartets for which R = H, all given structures are minima and representations of the structures in water can be found in Figure 6.1. The global minima of the normal quartets A_4-N1, A_4-N_3 and A_4-N7 are box- or saddle shaped (S_4) . On visualising these minima (Figurer 6.1) an interesting difference shows up. A_4-N1 and A_4-N7 can be characterised by two stacks of an adenine dimer, which are rotated with respect to each other by 90° but mutually bind through hydrogen bonding. For A_4-N3 the global minimum is box-shaped with the four bases positioned as the sides of a box. In this arrangement, the hydrogen bond between N6 and N3 still exists, but there will also be interactions based on dispersion that keep the structure together.

Table 6.1. Hydrogen bond energies (in kcal/mol) for the adenine (R = H) and 9-methyladenine (R = Me) quartets in the gas phase and in water.[a]

		R = H			R = Me		
		C_{4h}	C_4(Bowl)	S_4	C_{4h}	C_4(Bowl)	S_4
Gas phase	A_4-N1	–25.5	–30.2	–46.0	–25.5	–30.1	–49.0
	A_4-N3	–33.1	–32.7	–32.8	–33.3	–32.8	–39.9
	A_4-N7	–31.9	–33.1	–45.5	–32.5	–33.9	–52.3
	A_4*	–0.6	–1.7		–0.2	–1.5	
water	A_4-N1	–10.9	–17.0	–33.2	–10.3	–16.7	–36.0
	A_4-N3	–16.8	–16.3	–17.0	–17.1	–17.1	–19.3
	A_4-N7	–16.2	–16.3	–27.2	–16.0	–15.6	–31.2
	A_4* [b]	8.6	6.6		9.6	7.3	

[a] Computed at BLYP-D/TZ2P. Quartets with R = H were verified to be equilibrium structures through vibrational analysis.
[b] Destabilised complex, separated by a transition state from dissociation within C_{4h} symmetry into individual bases. This barrier occurs as it is unfavorable to break eight hydrogen bonds at the same time.

When A_4 is present in an oligonucleotide, it will adopt a more or less planar arangement. Therfore, it is interesting how much energy it takes to planarise the quartet starting from its global minimum. The global minma corresponds to the structure having the strongest hydrogen bonds, in other words have the most negative hydrogen bond energy. Focussed on an aqueous solution and R = Me, it costs 25.7 kcal/mol and 15.2 kcal/mol to planarise A_4-N1 and A_4-N7, respectively, with respect to their global minima (S_4). On contrary, it takes 2.2

kcal/mol to planarise A_4-N3 from its global minimum. In constrast, it takes in solution for R = H, only 0.2 kcal/mol to planarise A_4-N3.

Another feature that has to be noted is the instability of A_4* with respect to four *canonical* adenine bases. The term canonical refers to the most stable tautomer of adenine. A_4* is destabilised by about 7 kcal/mol in solution irrespective of A_4* having R = H or Me, based on the global minima (C_4(Bowl)). In the gas phase, the quartet is only slightly stabilised with respect to four canonical adenine bases, the bond energy amounts to up to –1.7 kcal/mol. The reason that the A_4* in water (C_{4h} symmetry, R = Me) is not stable can be explained as follows. The bonding energy with respect to four rare tautomers is –22.4 kcal/mol (data not shown). However, the total tautomerisation of four adenine bases into the most stable rare tautomer amounts to 32.0 kcal/mol. Consequently, the bonding energy can not compensate the larger tautomerisation energy, thus the quartet is destabilised. The rare tautomer of adenine, used to construct the quartet is the *most stable rare tautomer*, based on a computational study.[29]

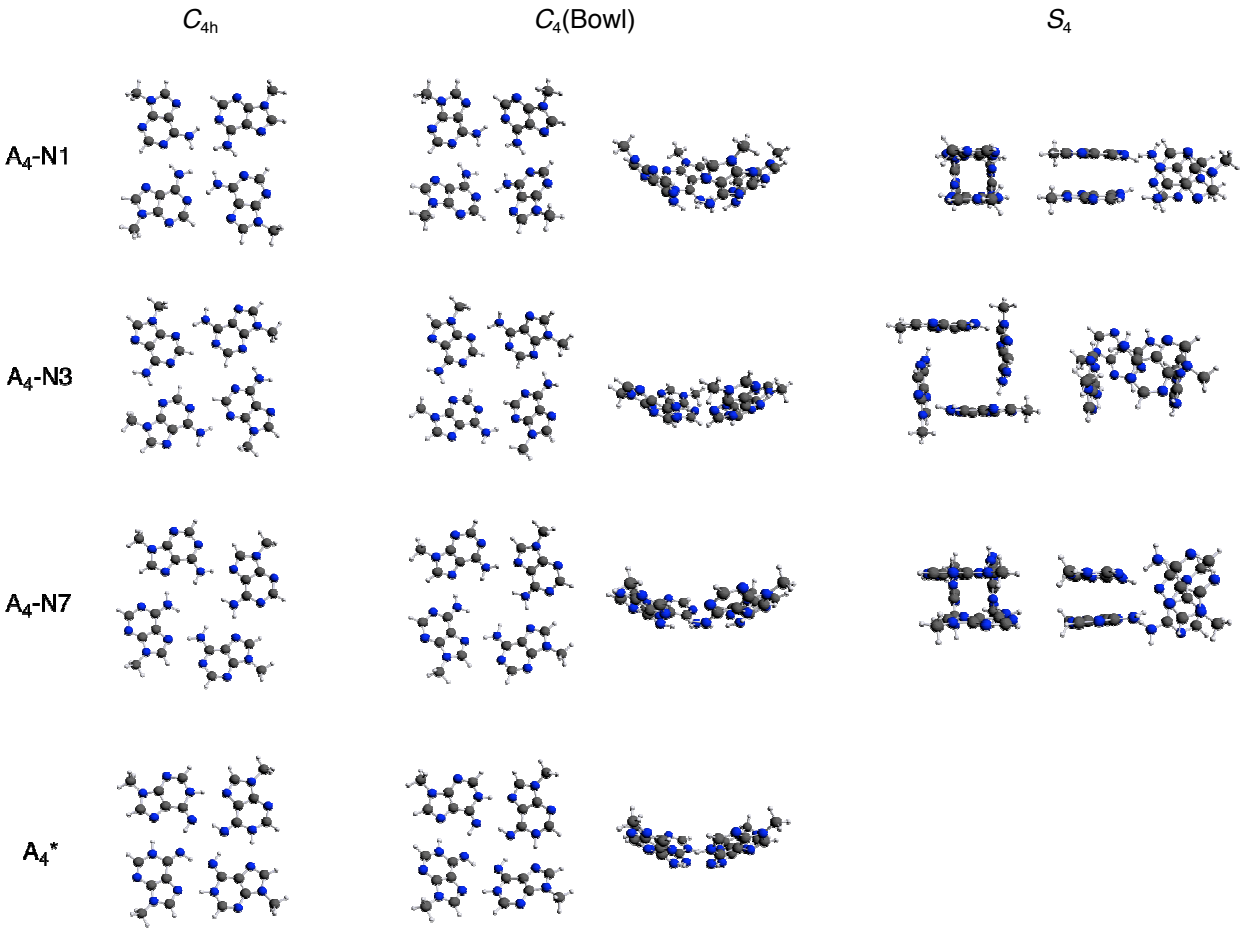


Figure 6.1. C_{4h}-, C_4(Bowl)-, and S_4-symmetric equilibrium structures of empty A_4 (R = Me) in water (for C_4 and S_4, top and side view are shown).

6.3.2 Adenine Quartets with Cations and Anions

This study is followed by investigating the possibility that adenine quartets can bind alkali cations or halide anions, and if this is true, how strong the bonding is. Especially, different complexes of adenine and ion are focussed on, *i.e.*, C_{4h}, C_4(Bowl) and S_4. The bond energies calculated from either adenine (R = H) or 9-methyladenine (R = Me) are shown in Table 6.2 and Table 6.3. Selected structures in water are shown in Figure 6.2 and Figure 6.3.

In the following discussion the quartets having R = H are focussed on, because for this systems a frequency analysis has been carried out. This is not computationally possible for systems having R = Me, because methyl groups can rotate which may bias results. The C_{4h} symmetric species may reflect the situation in a stack. However, it is in almost all cases *not* a stationay point, hence, a minimum. The structures are quite flexible, which has to do with the size of the ion, either it is too big, or too small. In the gas phase there is however a C_{4h} minimum possible for the canonical quartets with F^-, which is obviously small enough to fit in the quartet. The values for the bond energies are –111.7 kcal/mol, –91.5 kcal/mol and –113.2 kcal/mol for A_4-N1, A_4-N3 and A_4-N7, respectively. Another ion that is small enough is Li^+. It fits in A_4*, and has a bonding energy of –121.3 kcal/mol. A_4-N3 obviously has a cavity that is big enough to accomodate K^+, the A_4-N3-K^+ complex has a bonding energy of –91.5 kcal/mol.

In aqeuous solution, these C_{4h} symmetric complexes are even less stable, for instance A_4-N3-K^+ and A_4-N3-F^- are no minimum anymore, all other complexe that were discussed for the gas phase, are also minima in solution, but the binding energy has reduced substantially, by up to 90% (!) for the A_4*-Li^+ complex.

Table 6.2. Bonding energies (in kcal/mol) in the gas phase between four adenine bases (R = H) or four 9-methyladenine bases (R = Me) and an ion for equilibrium structures in C_{4h}, C_4(Bowl) and S_4 symmetry.[a]

		R = H			R = Me		
Quartet	Ion	C_{4h}	C_4(Bowl)	S_4	C_{4h}	C_4(Bowl)	S_4
A_4-N1	Li^+		–125.5	–151.0		–129.3	–158.1
	Na^+		–106.7	–119.7		–113.3	–126.3
	K^+		–90.2			–95.1	
	F^-	–111.7	–114.4		–109.1	–112.0	
	Cl^-		–77.8			–75.4	
	Br^-		–70.7			–68.4	
A_4-N3	Li^+			–139.4			–143.4
	Na^+			–116.8			–121.0
	K^+	–91.5	–92.0	–91.9	–94.2	–94.8	–97.7
	F^-	–53.3	–54.0		–52.7	–54.1	
	Cl^-		–35.6			–35.5	
	Br^-		–32.9			–32.7	
A_4-N7	Li^+		–129.3	–145.3		–134.8	–159.0
	Na^+		–117.1	–116.3		–122.4	–129.2
	K^+		–94.6			–101.0	
	F^-	–113.2			–110.7		
	Cl^-		–80.8			–78.7	
	Br^-		–74.2			–72.2	
A_4*	Li^+	–121.3	–123.6		–125.2	–128.0	
	Na^+		–95.4			–99.7	
	K^+		–65.5			–69.0	

[a] Computed at BLYP-D/TZ2P. Quartets with R = H were verified to be equilibrium structures through vibrational analysis.

Table 6.3. Bond energies (in kcal/mol) in water between four adenine bases (R = H) or four 9-methyladenine bases (R = Me) and an ion for equilibrium structures in C_{4h}, C_4 and S_4 symmetry.[a]

		R = H			R = Me		
Quartet	Ion	C_{4h}	C_4(Bowl)	S_4	C_{4h}	C_4(Bowl)	S_4
A_4-N1	Li^+		–25.6	–49.8		–26.6	–50.9
	Na^+		–37.5	–46.8		–38.4	–41.9
	K^+		–34.3	–32.3		–35.0	–34.5
	F^-	–27.1	–31.3		–26.8	–31.0	
	Cl^-		–23.9			–23.7	
	Br^-		–23.7			–23.6	
A_4-N3	Li^+			–33.6			–33.6
	Na^+			–37.7			–38.1
	K^+		–31.0	–29.5		–30.2	–29.5
	F^-		–6.6			–8.2	
	Cl^-		–10.2			–11.6	
	Br^-		–12.6			–13.9	
A_4-N7	Li^+		–29.2	–47.5		–29.8	–51.2
	Na^+		–34.7	–43.5		–36.3	–48.1
	K^+		–32.1	–32.6		–33.3	–36.0
	F^-	–26.7	–26.8		–26.5	–26.2	
	Cl^-		–21.7			–21.1	
	Br^-		–22.0			–21.6	
A_4*	Li^+	–13.2	–17.7		–12.3	–16.5	
	Na^+		–16.0			–15.3	
	K^+		–6.5			–5.2	

[a] Computed at COSMO-BLYP-D/TZ2P. Quartets with R = H were verified to be equilibrium structures through vibrational analysis.

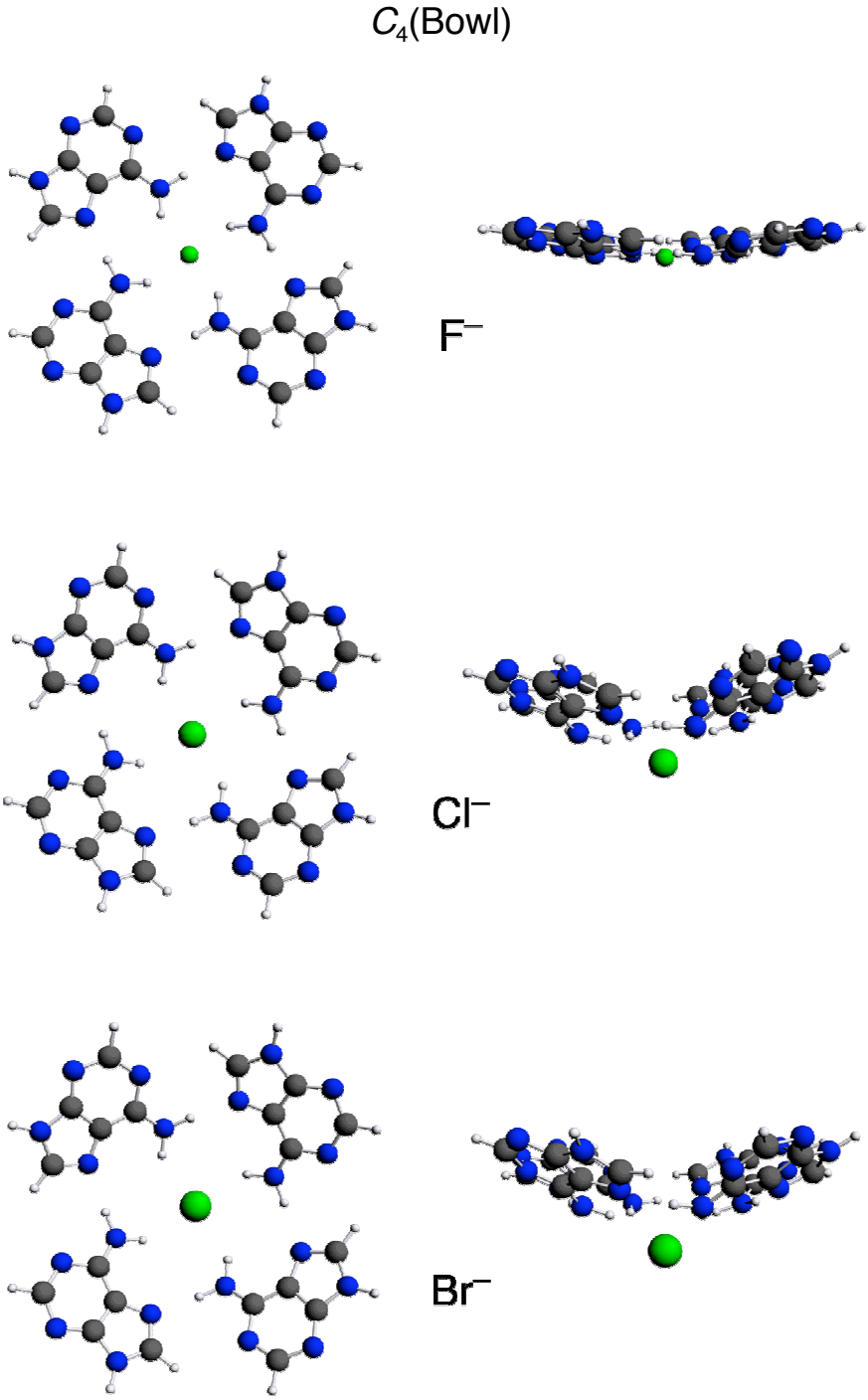


Figure 6.2. Top and side view of C_4(Bowl)-symmetric equilibrium structure of the A_4-N7 quartet binding halide anions (F^-, Cl^- and Br^-) in water. The anions are indicated by the green spheres.

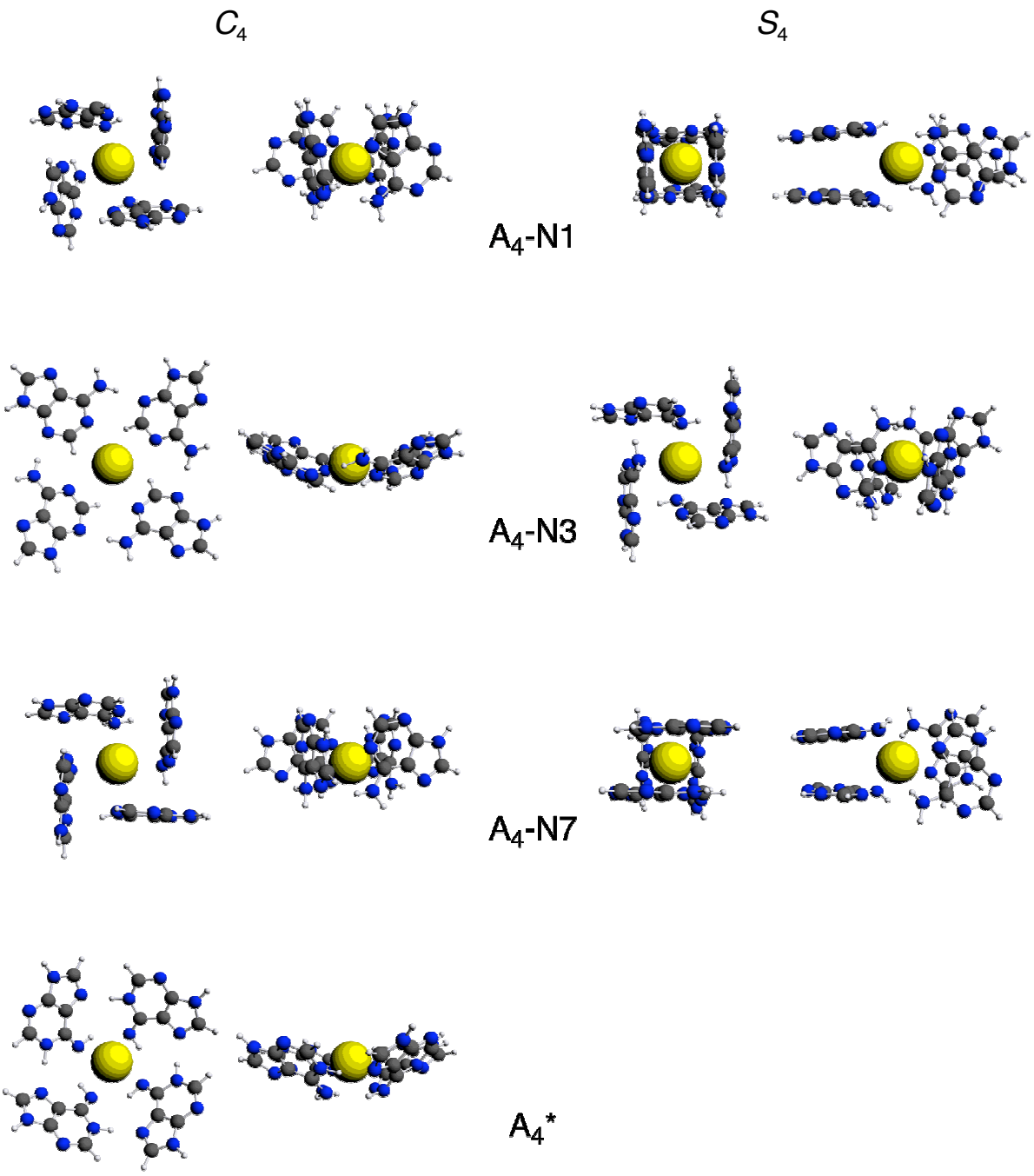


Figure 6.3. Top and side view of C_4(Bowl)- and S_4-symmetric equilibrium structures of A_4 quartets binding a potassium cation in water. The potassium is indicated by the yellow sphere.

With regard to the analogs with R = Me, there are no real changes. If present, these translate into differences in bond energies. In the gas phase the absolute difference amounts to 4 kcal/mol for the A_4*-Li^+ complex. In solution this difference is even smaller and amounts to up to 0.9 kcal/mol for the same complex.

The global minima are surely not planar, which is in line with findings with the empty quartets. Although the latter are all box- or saddle-shaped (S_4), the quartets with an ion have global minima, which are either bowl-shaped (C_4(Bowl)) or box- or saddle shaped (S_4). The ion plays a decisive role with regard to this aspect.

In both the gas phase as well as in solution, anions prefer to stabilise a C_4(Bowl) arrangement for A_4-N1, A_4-N3 and A_4-N7. In solution the bond energies of these systems amount to –26.8 (A_4-N7-F^-), –21.7 (A_4-N7-Cl^-) and –22.0 (A_4-N7-Br^-) kcal/mol. Structures of these can be found in Figure 6.2 and one can see that the larger the ion is (along F^-, Cl^- and Br^-) the more pronounced the bowl-shapedness is. There are three aspects that help to explain this feature. First, by increasing the size of the ion, the ion can no longer reside in the central cavity of the quartet. One can see clearly in Figure 6.2 that the complex with F^- is near to planar, but this situation changes as Cl^- and Br^- are applied. Second, one can cleary see that a hydrogen bond exists between the N-H moiety of the amino group of adenine and the anion. To retain this hydrogen bond, the quartet aligns itself towards the anion. The third and last aspect emerges as a result of the previous two ones. By aligning the hydrogen bond, an overlap with anion and the π-rich region of the quartet is avoided. There are no geometrical differences that arise due to the fact that either adenine or 9-methyladenine is used, but it is translated again in energy differences. In the gas phase the biggest absolute difference is 2.4 kcal/mol for A_4-N1-X^- and in solution 1.6 kcal/mol for the A_4-N3-F^- complex.

Cations prefer to stabilise box- or saddle-shaped arrangements (S_4) instead of bowl-shaped arrangements (C_4(Bowl)). However, the preference of S_4 over C_4(Bowl) depends on the size of the cation. In the gasphase the smaller ions Li^+ and Na^+ do all prefer the S_4 symmetry, apart form one case. A_4-N3-Na^+ (R = H) prefers a bowl-shaped over a box- or saddle-shaped arrangement, but the difference between these two arrangements is less than 1 kcal/mol. The large K^+ ion has no stationary points when it should adopt a box- or saddle-shaped arrangement with the A_4-N1 or A_4-N3 quartet. For A_4-N3 the preference is quite complicated and depends on whether R = H or Me. In case of R = H, a bowl-shaped arrangment is preferred (–92.0 kcal/mol) over a box-arrangement (–91.9 kcal/mol), but the

difference is negligible. For R = Me, the fact is inversed, hence, a box-shaped arrangments is preferred (–129.2 kcal/mol) over a bowl-shaped arrangement (–122.4 kcal/mol).

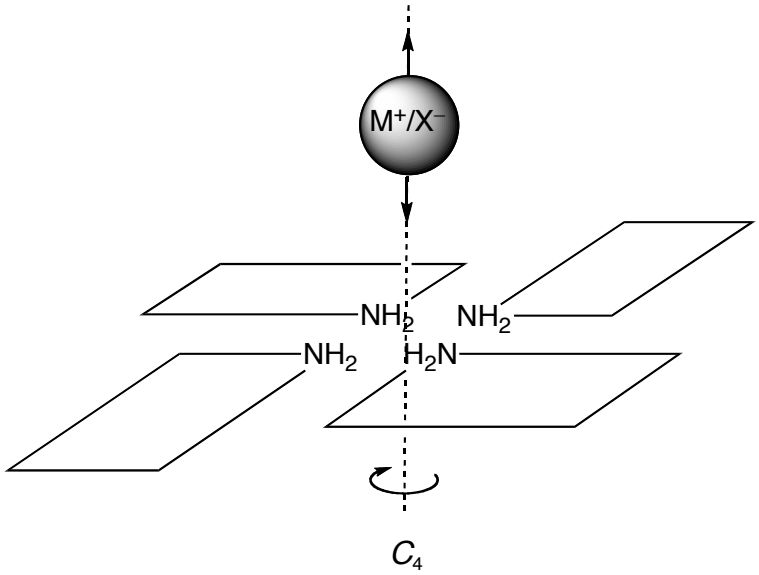
In solution the smaller ions Li^+ and Na^+ always prefer a box- or saddle-arrangement. The larger K^+ ion prefers, on contrary, a bowl-shaped arrangement in combination with A_4-N1 and A_4-N3. With A_4-N7 a box-shaped arrangement is preferred. This means that a switch in the preferential stabilisation exists in the case of A_4-N1 and A_4-N3, but not for A_4-N7. In Figure 6.3 the minima in water for K^+ with the different quartets are shown. They resemble the empty structures of Figure 6.1 a lot. There are, however, changes with respect to the quartets A_4-N1 and A_4-N7, because these fold around the cation. However, with respect to the empty quartets, the arrangement is changed such, that the quartets engage an extra nitrogen atom from the aromatic ring (*i. e.*, N7 for A_4-N1 and N1 for A_4-N7) in bonding with the alkali cation.

To conclude this discussion, it is interesting to note, that A_4* in combination with metal-cations is stable in both the gas phase as well as solution. In solution, the stabilisation energy due to the binding of a metal is up to –24.3 kcal/mol (R = H) and –23.8 kcal/mol (R = Me) when a Li^+ ion is concerned.

6.3.3 Models of Adenine Quartets in Stacks Binding Cations and Anions

In the last section of this chapter, the situation is simulated as if the quartet were part of a fourstranded-oligonucleotide in which the A_4 preserves a planar geometry. It has, however, already been mentioned that planar structures (C_{4h}) are in general no global minima. To make the quartets planar it takes a few tenths of a kcal (A_4-N3) to about 22 kcal/mol (A_4-N1). This planarisation energy should be compensated by the stacks that A_4 may form with G_4: one above and one below, as observed experimentally.[19,20] This compensation will be gained through the π - π interaction between the aromatic rings of the two quartets. In this section the emphasis is laid on A_4 in a geometry it may have in the stack. Therefore, quartet-ion complexes were optimised in which the adenine quartet was kept planar, apart from the amino group that was allowed to optimise in A_4-N1 and A_4-N7. Moreover, in all quartet-ion complexes the vertical distance between the center of the quartet and the ion (R_{vert}) was optimised as well (Scheme 6.3). The reader should be aware of the fact that these models are no equilibrium structures on their own. The quartets that are described by this models have quite large bonding energies in both the gas phase and solution, see Table 6.4 and 6.5, respectively. However, when going from the gas phase to solution, the bonding energy is

lowered. In the following discussion the solution (Table 6.5) shall be dealt with in more detail. For these model systems, only the case of R = Me is treated.



Scheme 6.3. The $C_4(R_{vert})$-symmetric systems in Tables 6.4 and 6.5 were optimised under the constraint that all atoms of the bases are kept in one plane, except for the nitrogen and hydrogen atoms of the amino groups of A_4-N1 and A_4-N7, which were allowed to pyramidalise; the cation or anion was allowed to move along the C_4-symmetry axes.

In case an alkali cation coordinates in the central cavity or along the C_4 axis, a bond energy of up to –29 kcal/mol is found. This corresponds to a stabilisation energy of –12 kcal/mol for A_4-N3-Na^+ and A_4-N3-K^+ between alkali cations and a C_{4h} quartet. The inner coordination complexes are not stable for A_4-N7, the bond energy is positive. However, A_4-N7 is able to coordinate a cation at the periphery, through N3, and a stabilisation energy is obtained by up to –10 kcal/mol in the case of Na^+. This peripheral coordination is also the preferred coordination for A_4-N1, also by using N3. However, coordination in the central cavity is possible too for A_4-N1 and the amino group can adapt itself such that the lone pair, situated on the N-atom (*i.e.*, the HOMO–3) can interact with, for instance Na^+, leading to a stabilisation enery of –6 kcal/mol. This situation is shown in Figure 6.4, and because an analysis was made thereafter, the gas phase structure is shown in which the A_4-N1-Na^+ complex is used as an example. Based on the left panel, the HOMO–3 on A_4-N1, donates 0.1 electron into the 3*s* orbital of the sodium cation.

Interestingly, A_4* a stabilisation energy of –22 kcal/mol is observed for Na^+ in the cavity, which is even larger than the stabilisation energies for the canonical quartets. Owing to this large stabilisation, the rare imino A_4 quartet becomes even stable with respect to four canonical adenine bases, in the case of Na^+, the bond energy is –12.7 kcal/mol. The reader is

reminded that empty A_4* is not stable. The fact that A_4* gets a negative bonding energy is a nice feature but all in all the bond energy in which canonical quartets play a role are more negative, thus an even better bonding exists. The values for Na^+ are –16.3 kcal/mol and –28.9 kcal/mol in combination with A_4-N1 and A_4-N3, respectively. In general, the value of the vertical distance (R_{Vert}) indicates the cation resides above the quartet. It varies between 0.1 Å (A_4-N3-K^+) and 2.9 Å (A_4-N1-K^+). There is one case in which the ion prefers to reside in the central cavity: A_4*-Li^+.

Apart from cation binding, canoncial quartets are also able to bind anions, even in a quite firm manner. For instance the bond energy between F^- and A_4-N1 amounts to –27 kcal/mol, and yield a stabilisation of about –17 kcal/mol, see Table 6.5. In contrast to cations, peripheral coordination is not possible for anions in case of A_4-N7 and A_4-N1 just because of a lacking amino group having an N-H bond exposed to the outside (See Scheme 6.1). A_4-N7, for instance, binds with stabilisation energies of –10.5, –5.0 and –5.7 kcal/mol, F^-, Cl^- and Br^-, respectively. In the complexes A_4-N7-Cl^- and A_4-N7-Br^- the ions have an R_{Vert} of 2.19 and 2.51 Å, respectively, and the amino groups of the quartet are pyramidalised such that the N-H bond that does not participate in hydrogen bonding with an adjacent adenine base, lines up in the direction of the anion. In the right panel of Figure 6.4 a picture is shown for A_4-N1-Cl^-. Again, for bonding analyses, the gas phase structure is shown. The LUMO of the quartet accepts 0.09 electrons from the chloride 3*p* atomic orbitals. Thus, in case of A_4-N1-Na^+ there is charge transfer from quartet to ion, in case of A_4-N1-Cl^- this is reversed, charge flows from ion to quartet.

The A_4-N3 quartet has different properties than A_4-N1 and A_4-N7, which has to do with the orientation of the N-H bond of the amino group that does not participate in hydrogen bonding with an adjacent adenine base. In the center of the quartet there are no such N-H groups and this is the reason why no stable complexes between anions and A_4-N3 are found. However, the N-H bond is exposed to the outside, thus, pheripheral coordination should be more favorable as is the case. The complex A_4-N3-F^- has a maximum bond energy of –3.1 kcal/mol. However, in the center of A_4-N3 the lone pairs of the N1 atoms point into the center of the quartet and can interact favourably with a cation, and a maximum stabilisation energy is obtained with Na^+ and amounts to –8.7 kcal/mol.

Table 6.4. Bond energies, ΔE_{Bond}, and stabilisation energies, ΔE_{Stab}, (in kcal/mol), and hydrogen bond distances, r_{HB}, and vertical separation, R_{Vert}, (in Å) in the gas phase for the complex of four 9-MeAH bases and an ion. [a]

		C_{4h}			$C_4(R_{vert})$				C_S	
Quartet	Ion	ΔE_{Bond}	ΔE_{Stab}	r_{HB} [b]	ΔE_{Bond}	ΔE_{Stab}	r_{HB} [b]	R_{Vert} [c]	ΔE_{Bond}	ΔE_{Stab}
A_4-N1	none	–25.5		2.91	–25.5		2.91		–25.5	
	Li^+	–32.4	–6.9	3.12	–105.5	–80.0	2.85	0.73	–83.2	–57.7
	Na^+	–48.8[d]	–23.3[d]	3.28[d]	–88.3	–62.8	2.88	1.30	–66.4	–40.9
	K^+	–47.1[d]	–21.6[d]	3.30[d]	–70.2	–44.7	2.90	1.89	–54.5	–29.0
	F^-	–109.1	–83.6	3.08	–109.1	–83.6	3.08	0.00		
	Cl^-	–60.3	–34.8	3.49	–64.0	–38.5	3.04	1.93		
	Br^-	–49.4	–23.9	3.75	–56.1	–30.6	3.03	2.24		
A_4-N3	none	–33.3		3.00	–33.3		3.00		–33.3	
	Li^+	–101.4	–68.1	2.88	–101.4	–68.1	2.88	0.00	–86.5	–53.2
	Na^+	–103.9	–70.6	2.89	–103.9	–70.6	2.89	0.00	–70.4	–37.1
	K^+	–94.2	–60.9	2.99	–94.2	–60.9	2.99	0.00	–58.3	–25.0
	F^-	–52.7	–19.4	3.06	–52.7	–19.4	3.06	0.00	[e]	
	Cl^-	–22.1	11.2	3.32	[f]	[f]	[f]	[f]	–54.8	–21.5
	Br^-	–14.0	19.3	3.47	[f]	[f]	[f]	[f]	–51.1	–17.8
A_4-N7	none	–32.5		2.88	–32.5		2.88		–32.5	
	Li^+	–50.4	–17.9	2.77	–91.1	–58.6	2.82	0.61	–90.2	–57.7
	Na^+	–44.3	–11.8	2.86	–82.5	–50.0	2.83	1.00	–73.4	–40.9
	K^+	–33.1	–0.6	2.97	–69.2	–36.7	2.85	1.53	–60.4	–27.9
	F^-	–110.7	–78.2	2.86	–110.7	–78.2	2.86	0.00		
	Cl^-	–76.5	–44.0	3.00	–76.7	–44.2	2.98	0.57		
	Br^-	–66.7	–34.2	3.07	–68.8	–36.3	2.97	1.28		
A_4*	none	–0.2		2.87(3.07)	–0.2		2.87(3.07)		–0.2	
	Li^+	–125.2	–125.0	2.92(3.19)	–125.2	–125.0	2.92(3.19)	0.00	–58.4	–58.2
	Na^+	–92.5	–92.3	3.24(3.53)	–93.0	–92.8	2.97(3.21)	1.22	–40.5	–40.3
	K^+	–54.0	–53.8	3.76(4.11)	–68.7	–68.5	2.92(3.13)	1.92	–26.8	–26.6

[a] Computed at BLYP-D/TZ2P. In case of empty quartets the bond energy is equal to the hydrogen bond energy.
[b] Hydrogen bond distances defined for each quartet as the distance between the proton donor and the proton acceptor. N7–N1 hydrogen bond distance in A_4* in parenthesis.
[c] Vertical separation between ion and plane of quartet.
[d] A_4-N1 quartet switches connectivity: N1 hydrogen bonds with the other N-H moiety of the aminogroup.
[e] F^- abstracts a proton from the amino group.
[f] Ion does not bind.

Table 6.5. Bond energies, ΔE_{Bond}, and stabilisation energies, ΔE_{Stab}, (in kcal/mol), and hydrogen bond distances, r_{HB}, and vertical separation, R_{Vert}, (in Å) in water for the complex of four 9-MeAH bases and an ion.[a]

		$C_4(R_{Vert})$				C_S	
Quartet	Ion	ΔE_{Bond}	ΔE_{Stab}	r_{HB}[b]	R_{Vert}[c]	ΔE_{Bond}	ΔE_{Stab}
A_4-N1	none	–10.3		2.95			
	Li^+	[d]	[d]	[d]	[d]	–17.8	–7.5
	Na^+	–16.3	–6.0	2.92	1.47	–20.5	–10.2
	K^+	–13.5	–3.2	2.94	2.05	–17.8	–7.5
	F^-	–26.8	–16.5	3.11	0.00		
	Cl^-	–15.8	–5.5	3.02	2.19		
	Br^-	–16.0	–5.7	3.00	2.51		
A_4-N3	none	–17.1		3.03			
	Li^+	[e]	[e]	[e]	[e]	–25.6	–8.5
	Na^+	–28.9	–11.8	2.93	0.1	–25.8	–8.7
	K^+	–28.9	–11.8	3.02	0.1	–23.7	–6.6
	F^-	[e]	[e]	[e]	[e]	–20.2	–3.1
	Cl^-	[e]	[e]	[e]	[e]	–17.8	–0.7
	Br^-	[e]	[e]	[e]	[e]	–18.5	–1.4
A_4-N7	none	–16.0		2.91			
	Li^+	5.4	21.4[f]	2.87	0.87	–22.2	–6.2
	Na^+	–9.0	7.0[f]	2.87	1.19	–26.1	–10.1
	K^+	–11.7	4.3[f]	2.89	1.75	–23.3	–7.3
	F^-	–26.5	–10.5	2.89	0.00		
	Cl^-	–21.0	–5.0	2.88	1.55		
	Br^-	–21.7	–5.7	2.95	1.91		
A_4*	none	9.6		2.91(3.07)			
	Li^+	–12.3	–21.9	2.95(3.20)	0.00	3.8	–5.8
	Na^+	–12.7	–22.3	2.91(3.11)	1.50	–0.1	–9.7
	K^+	–5.5	–15.1	2.91(3.09)	2.09	2.6	–7.0

[a] Computed at COSMO-BLYP-D/TZ2P. In the case of empty quartets, the bond energy equals the hydrogen bond energy.
[b] Hydrogen bond distances defined for each quartet as the distance between the proton donorand the proton acceptor. N7–N1 hydrogen bond distance in A_4* in parenthesis.
[c] Vertical separation between ion and plane of quartet.
[d] Ion does not bind.
[e] Ion binds very weakly at about 5.0 Å, which is not relevant for situation in stack of quartets.
[f] Destabilised complex, separated by a transition state from dissociation.

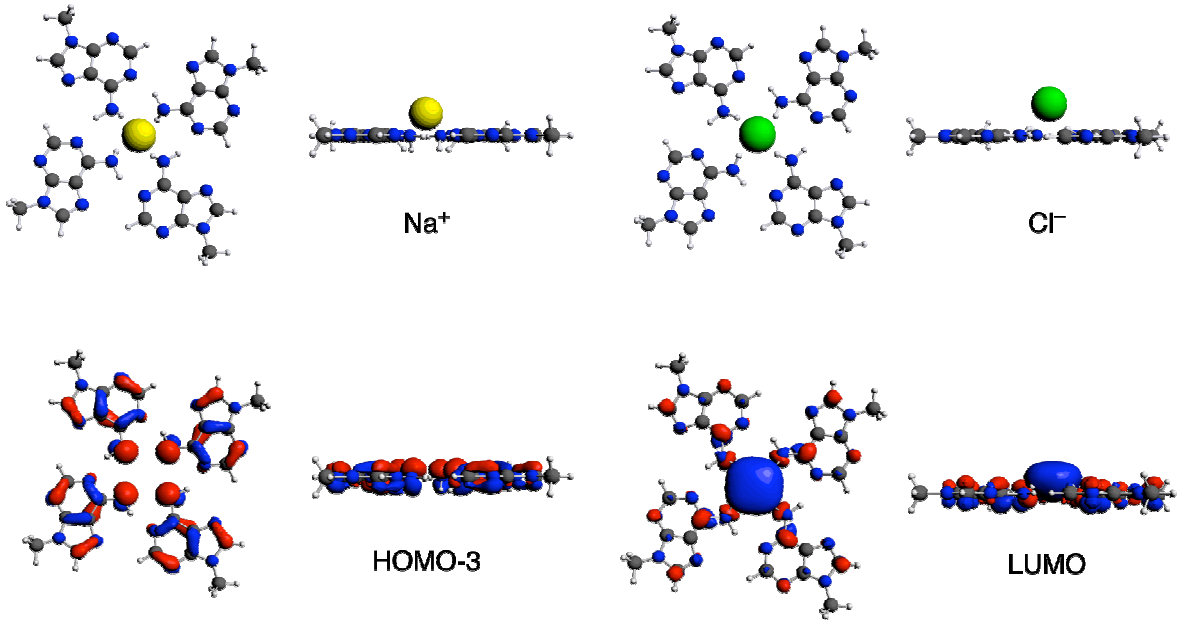


Figure 6.4. Upper: top and side view of a planar A_4-N1 quartet interacting in the gas phase with a sodium cation (left) or chloride ion (right). Lower: corresponding A_4-N1 HOMO-3 that donates charge into sodium 3*s* AO (left) and A_4-N1 LUMO that accepts charge from chloride 3*p* AO (right). The ion sizes in this picture do not scale.

6.4 Conclusions

This chapter demonstrate, by the use of dispersion-corrected DFT, that adenine quartets can bind both cations (Li^+, Na^+ and K^+) and anions (F^-, Cl^- and Br^-) in or above their central cavity, in the gas phase as well as in aqueous solution. In this chapter three A_4 quartets with a canonical structure were presented (A_4-N1, A_4-N3 and A_4-N7) as well as a rare imino-tautomer quartet (A_4*). In general global minima are bowl-, box- or saddle-shaped, but in combination with Li^+ or F^- planar structures are local minima and are found for A_4-N1-F^-, A_4-N7-F^- and A_4*-Li^+. Empty quartes prefer a non-planar geometry in both the gas phase as well as in water, and this feature does not change due to the addition of an ion.

The ion-quartet interaction has also been simulated by keeping the quartet aritficially planar, and the inner amino groups (A_4-N1 and A_4-N7) was allowed to be optimised, furthermore the distance between the ion and the central cavity of the quartet (R_{Vert}) was optimised as well. This model system leads to good stabilisations as well, and there is an interesting feature. The cations prefer to bind to the outside in case of A_4-N1 and, notably, A_4-N7, by interaction with the N lone pair of N3. Inner coordination of cations with A_4-N1, is also possible, though, via the amino group. In case of A_4-N3, central cavity coordination is the preferred binding mode of the cations, as it interacts with the N1 atoms having its lone pair in the center. In constrast, A_4-N1 and A_4-N7 prefer to coordinate anions in the central cavity (F^-), or above it (Cl^- and Br^-). In the latter case, the amino groups pyramidalise such that the N-H bond that does not participate in hydrogen bonding point toward the anion, and the geometry that is obtained hereby guarantees an optimised electrostatic as well as donor-acceptor interactions. The anions prefer to interact with the N-H bonds of the amino groups at the outside. Consequently, it is the orientation of the N-H bond of the amino group that does not participate in hydrogen bonding with an adjacent adenine base that decide which ions are preferred and how they are bonded. However, one can also assume, that the ion directs the geometry of the amino group. When this possibility is also taken into account, the "hen-and-egg" problem arises: does the amino group decides, by its geometry, what ion is bound or does the ion force the amino group in a given geometry. The answer is difficult to give, and a mutuality between these possibilities may exist.

Finally, the results of anion binding to adenine quartets is relevant when these are present in a tetrastranded structure. In case stacking interactions force an adenine quartet in a planar geometry having the amino groups in the interior, halide ions, especially the smallest one,

fluoride (F^-), is a good choice for stabilising such an arrangement. A_4 can occur in negatively charged tetratranded oligonucleotides as well as artificial neutral analogues.

6.5 References

[1] *Quadruplex nucleic acids* (Eds. S. Neidsle, S. Balasubramanian), RSC Publishing, Cambridge, **2006**.

[2] D. M. Tagore, K. I. Sprinz, S. Fletcher, J. Jayawickramarajah, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2006**, *46*, 223-225.

[3] C.-C. Huang, H. T. Chang, *Chem. Commun.* **2008**, 1461-1463.

[4] M. Mir, I. Katakis, *Mol. BioSyst.* **2007**, *3*, 620-622.

[5] J. T. Davis, *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.

[6] D. J. Patel, S. Bouaziz, A. Kettani, Y. Wang in *Oxford Handbook of Nucleic Acid Structure* (Ed. S. Neidle), Oxford University Press, Oxford, **1999**, pp. 389-453.

[7] (a) J. Deng, Y. Xiong, M. Sundaralingam, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13665-13670.
(b) B. Fischer, H. Preut, B. Lippert, H. Schöllhorn, U. Thewalt, *Polyhedron* **1990**, *9*, 2199-2204.
(c) C. Cheong, P. B. Moore, *Biochemistry* **1992**, 31, 8406-8414.

[8] H. Witkowski, E. Freisinger, B. Lippert, *Chem. Commun.* **1997**, 1315-1316.

[9] B. Pan, K. Shi, M. Sundaraligam, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3130-3134.

[10] P. Patel, R. V. Hosur, *Nucleic Acids Res.* **1999**, *27*, 2457-2464.

[11] (a) P. K. Patel, N. S. Bhavesh, R. V. Hosur, *Biochem. Biophys. Res. Commun.* **2000**, *270*, 967-971.
(b) P. K. Patel, S. Neel, S. Bhavesh, R. V. Hosur, *Biochem. Biophys. Res. Commun.* **2000**, *278*, 833-838.

[12] F. Seela, C. Wei, A. Melenewski, *Nucleic Acids Res.* **1996**, *24*, 4940-4945.

[13] G. N. Parkinson, M. P. H. Lee, Neidle, S. *Nature* **2002**, *417*, 876-880.

[14] A. Kettani, A. R. Kumar, D. J. Patel, *J. Mol. Biol.* **1995**, *254*, 638-656.

[15] (a) A. I. H. Murchie, D. M. J. Lilley, *EMBO J.* **1994**, *13*, 993-1001.
(b) D. Mohanty, M. Bansal, *Biophys. J.* **1995**, *69*, 1046-1067.
(c) M. C. Shiber, E. H. Braswell, H. Klump, J. R. Fresco, *Nucleic Acids Res.* **1996**,
24, 5004-5012.
(d) P. Amo-Ochoa, P. J. Sanz-Miguel, P. Lax, I. Alonso, M. Roitzsch, F. Zamora, B. Lippert, *Angew. Chem. Int. Ed.* **2005**, *44*, 5670-5674.

[16] W. N. Hunter, T. Brown in *Oxford Handbook of Nucleic Acid Structure* (Ed.: S. Neidle), Oxford University Press, Oxford, **1999**, p. 389.

[17] N. Escaja, J. L. Gelpí, M. Orozco, M. Rico, E. Pedroso, C. González, *J. Am. Chem. Soc.* **2003**, *125*, 5654-5662.

[18] (a) M. S. Searle, H. E. L Williams, C. T. Gallagher, R. J. Grant, M. F. G. Stevens. *Org. Biomol. Chem.* **2004**, *2*, 810-812.
(b) P. K. Patel, A. S. R. Koti, R. V. Hosur, *Nucl. Acids Res.* **1999**, *27*, 3836-3843.
(c) E. Gavathiotis, M. S. Searle, *Org. Biomol. Chem.* **2003**, *1*, 1650-1656.

[19] B. Pan, Y. Xiong, K. Shi, J. Deng, M. Sundaralingam, *Structure* **2003**, *11*, 823-831.

[20] B. Pan, Y. Xiong, K. Shi, J. Deng, M. Sundaralingam, *Structure* **2003**, *11*, 815-823.

[21] (a) J. Kumar, C. S. Purohit, S. Verma, *Chem. Commun.* **2008**, 2526-2528.
(b) C. S. Purohit, A. K. Mishra, S. Verma, *Inorg. Chem.* **2007**, *46*, 8493-8495.

[22] (a) J. Gu, J. Leszczynski, *Chem Phys. Lett.* **2001**, *335*, 465-474.
(b) M. Meyer, C. Schneider, M. Brandl, J. Sühnel, *J. Phys. Chem. A* **2001**, *105*, 11560-11573.

(c) M. Meyer, J. Sühnel, *J. Phys. Chem. A* **2008**, *112*, 4336-4341.
[23] A. Wong, G. Wu, *J. Am. Chem. Soc.* **2003**, *125*,13895-13905 (and refs. cited therein).
[24] P. Auffinger, L. Bielecki, E. Westhof, *Structure* **2004**, *12*, 379-388.
[25] T. Evan-Salem. L. Frish, F. W. B. van Leeuwen, D. N. Reinhoudt, W. Verboom, M. S. Kaucher, J. T. Davis, Y. Cohen, *Chem. Eur. J.* **2007**, *13*, 1969-1977.
[26] B. Lippert, D. Gupta, *Dalton Trans.* **2009**, 1419-1434.
[27] *Nucleic Acid-Metal Interactions* (Ed. N. V. Hud), RSC Publishing, Cambridge, **2009**.
[28] M. P. H. Lee, G. N. Parkinson, P. Hazel, S. Neidle, *J. Am. Chem. Soc.* **2007**, *129*, 10106-10107.
[29] C. Fonseca Guerra, F. M. Bickelhaupt, S. Saha, F. Wang, *J. Phys. Chem. A* **2006**, *110*, 4012-4020.

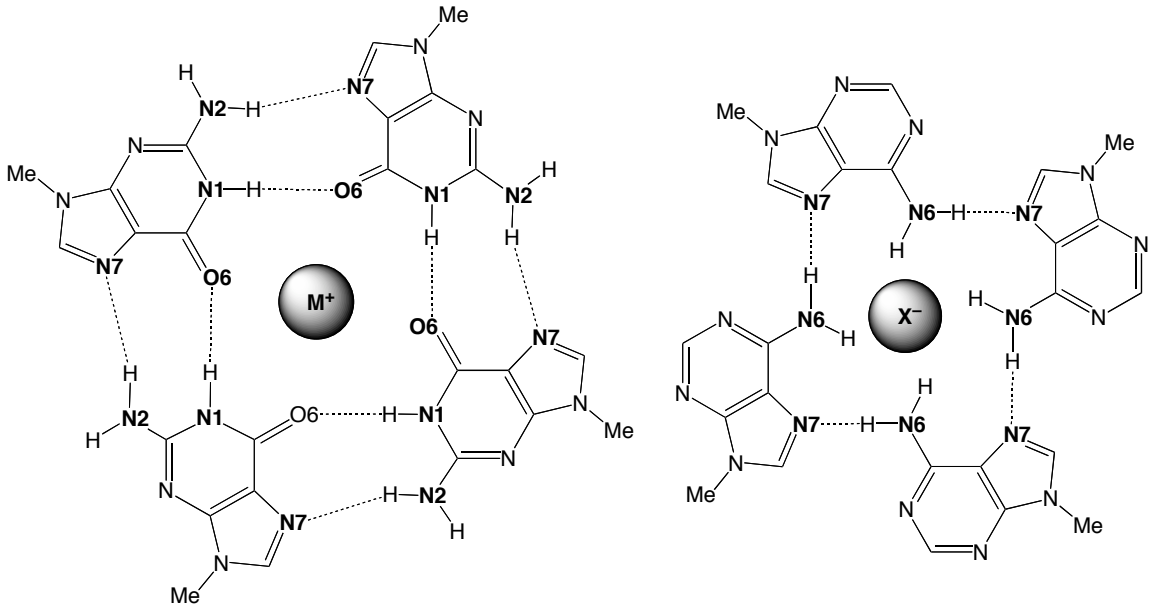
Chapter 7
A Ditopic Ion-Pair Receptor Based on Stacked Nucleobase Quartets

7.1 Introduction

The recognition of ionic species is an ongoing important topic in supramolecular chemistry.[1] Historically, cation receptors were studied first,[2] which have lead to the design of anion receptors.[3] However, one may also bind an anionic species and cationic species at the same time,[4] and this is termed as a ditopic receptor. These molecules can, for example, be chalix-like,[5] or have a crown ether.[6] Ditopic receptors are interesting because of the salt-extracting and membrane transport properties.[5,6]

In this study we show that an ion-pair receptor for sodium chloride (NaCl) can be built using DNA quartet structures: namely a guanine (G_4) and adenine (A_4) quartet. G_4 is very stable because of its ability to form two hydrogen bonds between neighbouring guanine bases. Furthermore, the formation of G_4 is brought about by cations, notably Na^+ and K^+.[7,8] G_4 will now be used to stack the anion binding partner on it, in this case the adenine quartet. It must however be noticed, that a stack of G_4's on its own can act as ditopic receptor, a cation coordinates in the central cavity and an anion interacts with the periphery.[9] The choice of A_4 is not straightforward, because throughout all possible quartet structures,[10] no anion binding to these has been evidenced. There exists however a structure in which an adenine quartet is stabilised by a sodium cation, which could feasibly have been an anion.[11] Moreover, a platinum-modified purine quartet was synthesised, which showed a preference for anion.[12]

In this chapter an NaCl ion pair receptor is presented. G_4 and A_4 bind the sodium cation (Na^+) and chloride anion (Cl^-), respectively, see Scheme 7.1. The cation can favourably interact with the cation through four coordinative bonds with the oxygen atom of the four carbonyl groups (Scheme 7.1, left), the chloride anion may be able to form four hydrogen bonds with the N-H moiety of the amino groups (Scheme, right).



Scheme 7.1. G_4–cation and A_4–anion receptor.

In total a stack of two purine quartets is obtained, G_4A_4, for which will be shown that it can act as a NaCl ion pair binder in water (Figure 7.1) according to Eq. 7.1.

$$G_4A_4(aq) + Na^+(aq) + Cl^-(aq) \rightarrow G_4NaClA_4(aq) \qquad \text{(Eq. 7.1)}$$

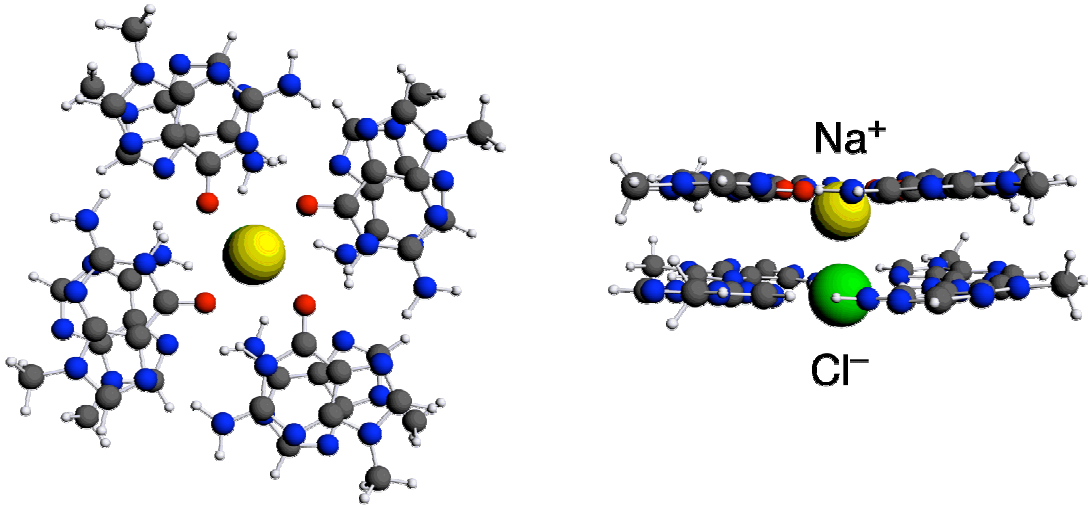


Figure 7.1. Ditopic receptor G_4NaClA_4 in aqueous solution: top view and side view.

7.2 Summary of Computational Methods

Calculations were performed in QUILD, a subprogram of ADF, and done at the BLYP-D/TZ2P level of density functional theory. As in several reaction products π-stacking plays a large role, it must be treated properly, which is done by the BLYP-D functional. In chapter 5, the outstanding performance of BLYP-D has been demonstrated for stacked dimers of adenine-thymine and guanine-cytosine, as it could reproduce very accurate *ab initio* binding energies. The reaction as defined by Eq. 7.1, energy is calculated as follows

$$\Delta E_{\text{Reaction}} = \Sigma E_{\text{Products}} - \Sigma E_{\text{Reactants}} \qquad \text{(Eq. 7.2)}$$

which means the reaction energy $\Delta E_{\text{Reaction}}$ is the difference between the sum of the energies of the products and the sum of the energies of the reactants. In some cases the difference between two reaction energies is given to indicate the extra stabilisation. Other parameters as bond energy and stabilisation energy can be analoguously derived from previous chapters.

Individual bases were optimised without any symmetry restrictions. A_4 and G_4 were optimised in C_{4h} symmetry, fully planar arrangements, and stacks were optimised in C_4(Bowl) symmetry. Solvent effects due to water were treated implicitly by the COSMO method. However, the parameters for the free ions were adapted such that experimental solvation energies could be reproduced. Furthermore upon treating solvent effects, the dispersion correction should not be modified. The interested reader is referred to chapter 2, for more details about the computational settings.

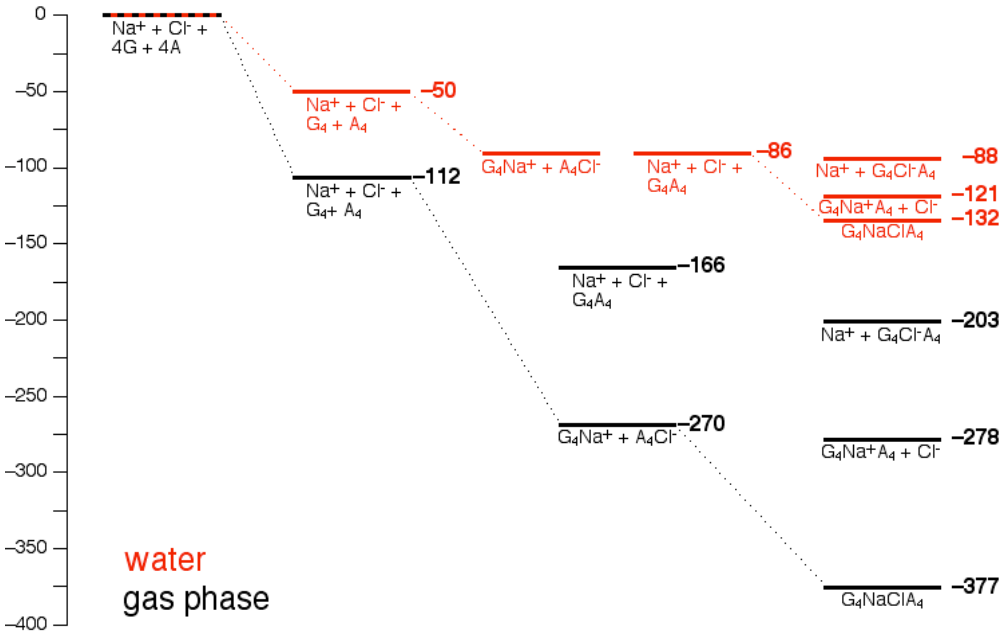
7.3 Results and Discussion

In Scheme 7.2 several conceivable steps are shown that lead to the formation of the ditopic receptor, in both the gas phase as well as solution. The numbers that are presented are relative reaction energies and are, as explained above, done at the BLYP-D/TZ2P level. As a common starting point the combination of the separate 9-methylated bases with the ions is used: 4G + 4A + Na^+ + Cl^-. It can be easily seen that in water reaction energies are about two- or threefold lower than in the gas phase, but the trends remain the same. Some features will be discussed in more detail.

In the first step, the two empty quartets G_4 and A_4 are formed, which go along with a reaction energy of of –112 and –50 kcal/mol in the gas phase and water solution, respectively. The bonding in the G_4 quartet is about as twice as strong in comparison with the

bonding in of the A_4 quartet, –79.2 kcal/mol and –32.5 kcal/mol in the gas phase and –33.8 kcal/mol and –16.0 kcal/mol, both respectively (see also previous chapters). This goes nicely along with the fact that the G_4 quartet has twice as many hydrogen bonds, eight, than the A_4 quartet, namely four.

In the second step, out of the quartets two species can form i) a G_4 quartet with Na^+ and an A_4 quartet with Cl^- or ii) the formation of a G_4A_4 stack. With respect to the first step the formation of G_4Na^+ and A_4Cl^- goes along with an extra stabilisation of –158 kcal/mol and the formation of the G_4A_4 stack yields an extra stabilisation of –54 kcal/mol. These two events yield in water the same energy stabilisation of –36 kcal/mol. Individually the insertion of Na^+ in G_4 and of Cl^- in A_4 gains –114.5 kcal/mol and –44.1 kcal/mol, respectively in the gas phase, and –33.9 and –2.9 kcal/mol, respectively, in water solution.



Scheme 7.2. Reaction energies (in kcal/mol) of various stages of complexation in the formation of the ditopic receptor G_4NaClA_4, starting from four 9-methylguanines (G), four 9-methyladenines (A) and the ions Na^+ and Cl^- in the gas phase (in black) and in aqueous solution (in red), computed at BLYP-D/TZ2P.

Finally, in the third step the behaviour of the total ditopic receptor is described. There are two possible reaction pathways to reach the final step. The ditopic receptor can be formed by two pathways: i) inclusion of Na^+ and Cl^- in the G_4A_4 stack and ii) the stack of the earlier discussed fragments G_4Na^+ and A_4Cl^-. In the gas phase both pathways are possible in the sense that these go along with stabilisation of –211 and –107 kcal/mol, in solution the stabilisation energy of the two pathways amount to both –46 kcal/mol, respectively. The

overall formaton of the ditopic receptor from four separate 9-methylguanines, 9-methyladenine and the ions Na^+ and Cl^- in water is –132 kcal/mol. Moreover, it is the most stable reaction product because $G_4Cl^-A_4$ and $G_4Na^+A_4$ form with –203 and –278 kcal/mol, respectively. All in all, it is likely that the ditopic receptor can exist.

With respect to the geometry there are other interesting features that are important to note, and for this reason G_4A_4 and G_4NaClA_4 will be compared, thus geometrical changes due to the inclusion of the ions pair NaCl are are given in Table 7.1

Table 7.1. Selected distances (in Å) of G_4A_4 and $G_4Na^+Cl^-A_4$ in both the gas phase as well as solution. Distances refer in most cases to separations from plane, except for Na-Cl.

	G_4A_4	$G_4Na^+Cl^-A_4$
Gas phase		
O6-N1	2.79	2.81
N2-N7	2.88	2.86
N6-N7	2.88	2.98
G_4-A_4	3.08	3.14
G_4(O6)-A_4(N6)	2.84	3.16
G_4-Na^+	--	1.08
G_4(O6)-Na^+	--	0.71
A_4-Cl^-	--	0.70
A_4(N6)-Cl^-	--	0.31
Na-Cl	--	2.75
COSMO		
O6-N1	2.82	2.82
N2-N7	2.88	2.86
N6-N7	2.90	3.01
G_4-A_4	3.09	3.20
G_4(O6)-A_4(N6)	2.85	3.33
G_4-Na^+	--	0.84
G_4(O6)-Na^+	--	0.70
A_4-Cl^-	--	0.45
A_4(N6)-Cl^-	--	0.16
Na-Cl	--	2.80

In both the gas phase as well as solution, the hydrogen bonds of G_4 are less affected than the ones of A_4 due to the presence of either Na^+ or Cl^-, respectively. The latter goes along with an elongation of about 0.1 Å. It can also be seen that the stack becomes larger in size, based on the distance between the planes that are determined by $(O6)_4$ of G_4 and $(N6)_4$ of A_4, an increase of up to 0.5 Å can be observed. The maximal distance of 3.3 Å is a distance that is more often observed in both B-DNA as well as G-rich quadruplex DNA. As induced by the

insertion of the NaCl, the empty G_4A_4 cage will be deformed in these deformations are 12.8 and 6.9 kcal/mol in the gas phase and water solution, respectively.

7.4 Conclusion

In summary the G_4A_4 stack is a stable complex and able to accept NaCl in both the gas phase as well as aqueous solution. However it is questionable whether this feature will also present in real oligonucleotides, which are negatively charged. It should however be mentioned that anion binding to polyanionic DNA has indeed been found.[13] Consequently, the binding of a chloride to an A_4 quartet is very well conceivable. This process may even be facilitated by the fact that the negative charges of the backbone are screened by cations, a feature has been mentioned in chapter 1.

7.5 References

[1] (a) J. W. Steed, J. J. Atwood, *Supramolecular Chemistry: An Introduction*, Wiley, Chichester, **2000**.
(b) P. D. Beer, P. A. Gale, D. K. Smith, *Supramolecular Chemisty*, Oxford University Press, Oxford, **1999**.
(c) J.-M. Lehn, *Supramolecular Chemistry Concepts and Perspectives*, VCH, Weinheim, **1995**.

[2] G. W. Gokel in, *Comprehensive Supramolecular Chemisty Vol. 1*, Eds.: J. L. Atwood, J. E. D. Davies, D. D. MacNicol, F. Vögtle, Pergamon, Oxford, **1996**.

[3] P. D. Beer, P. A. Gale, *Angew. Chem., Intl. Ed.* **2001**, *40*, 486-516.

[4] G. J. Kirkovits, J. A. Shriver, P. A. Gale, J. L. Sessler, *J. Inclusion Phenom. Macrocyclic Chem.* **2001**, *41*, 69-75.

[5] (a) M. D. Lankshear, I. M. Dudley, K.-M. Chan, A. R. Cowley, S. M. Santos, V. Felix, P. D. Beer, *Chem. Eur. J.* **2008**, *14*, 2248-2263.
(b) M. D. Lankshear, A. R. Cowley, P. D. Beer, *Chem. Commun.* **2006**, 612-614.
(c) J. B. Cooper, M. G. B. Drew, P. D. Beer, *J. Chem. Soc., Dalton Trans.,* **2001**, 392-401
(d) M. D. Lankshear, I. M. Dudley, K.-M. Chan, P. D. Beer, *New J. Chem.* **2007**, *31*, 684-690.
(e) P. R. A. Webber, P. D. Beer, *Dalton Trans.* **2003**, 2249-2252.
(f) R. Custelcean, L. H. Delmau, B. A. Moyer, J. L. Sessler, W.-S. Cho, D. Gross, G. W. Bates, S. J. Brooks, M. E. Light, P. A. Gale, *Angew. Chem. Int. Ed.* **2005**, *44*, 2537-2542.

[6] (a) M. K. Chae, J.-I. Lee, N.-K. Kim, K.-S. Jeong, *Tetrahedron Lett.* **2007**, *48*, 6624-6627.
(b) H. Miyaji, D.-S. Kim, B.-Y. Chang, E. Park, S.-M. Park, K. H. Ahn, *Chem. Commun.* **2008**, 753-755.
(c) J. M. Mahoney, K. A. Stucker, H. Jiang, I. Carmichael, N. R. Brinkmann, A. M. Beatty, B. C. Noll, B. D. Smith, *J. Am. Chem. Soc.* **2005** *127*, 2922-2928.
(d) C. Suksai, P. Leeladee, D. Jainuknan, T. Tuntulani, N. Muangsin, O. Chailapakul, P. Kongsaeree, C. Pakavatchai, *Tetrahedron Lett.* **2005**, *46*, 2765-2769.
(e) P. Gunning, A. C. Benniston, R. D. Peacock, *Chem. Commun.* **2004**, 2226-2227.

[7] *Quadruplex Nucleic Acids*, (Eds.: S. Neidle, S. Balasubramanian), RSC Publishing, Cambridge, **2006**.

[8] (a) J. T. Davis, *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698 (and references cited therein).
(b) J. Gros, F. Rosu, S. Amrane, A. De Cian, V. Gabelica, L. Lacroix, J.-L. Mergny, *Nucleic Acids Res.* **2007**, *35*, 3064-3075.

[9] (a) X. Shi, K. M. Mullaugh, J. C. Fettinger, Y. Jiang, S. A. Hofstadler, J. T. Davis, *J. Am. Chem. Soc.* **2003**, *125*, 10830-10841.
(b) (b) X. Shi, J. C. Fettinger, J. T. Davis, *Angew. Chem. Int. Ed.* **2001**, *40*, 2827-2831.
(c) F. W. Kotch, V. Sidorov, Y.-F. Lam, K. J. Kayser, H. Li, M. S. Kaucher, J. T. Davis, *J. Am. Chem. Soc.* **2003**, *125*, 15140-15150.

[10] (a) D. J. Patel, S. Bouaziz, A. Kettani, Y. Wang in *Oxford Handbook of Nucleic Acid Structure* (Ed.: S. Neidle), Oxford University Press, Oxford, **1999**.

(b) P. Amo-Ochoa, P. J. Sanz-Miguel, P. Lax, I Alonso, M. Roitzsch, *Angew. Chem. Int. Ed.* **2005**, *44*, 5670-5674 (and references cited therein).
[11] B. Pan, Y. Xiong, K. Shi, J. Deng, M. Sundaralingam, *Structure* **2003**, *11*, 815-823.
[12] M. Roitzsch, B. Lippert, *Angew. Chem. Int. Ed.* **2006**, *45*, 147-150.
[13] (a) P. Auffinger, L. Bielecki, E. Westhof, *Structure* **2004**, *12*, 379-388.
(b) B. Spingler, *Inorg. Chem.* **2005**, *44*, 831-833.
(c) J. Muñoz, J. L. Gelpí, M. Soler-López, J. A. Subirana, M. Orozco, F. J. Luque, *J. Phys. Chem. B* **2002**, *106*, 8849-8857.

Summary in English, Dutch and German

Summary

In this thesis, studies are presented of computations on metal-nucleobase complexes, based on density functional theory (DFT). This research field is very wide,[1] and that is why a selection of research topics has been made. The following systems of nucleobases are described. First, the Watson-Crick base pairs between adenine and thymine (AT) and between guanine and cytosine (GC) without and with consideration of the crystal environment. Second, complexes of platinum (Pt^{II}) and 1-methyluracil (1-MeUH) or 1-methylthymine (1-MeTH), in which the stabilisation of a rare tautomer is put central. In addition, it is also investigated in which rare-tautomer form a guanine base is present, if Pt^{II} coordinates to N1. Third, larger aggregates of nucleobases are described, in particular quartets. For instance, the cation and anion binding properties of adenine quartets (A_4) are demonstrated. Finally, a stack of one guanine (G_4) and one A_4 quartet is presented, which can function as a receptor for sodium chloride (NaCl). Now a summary per chapter is given.

Chapter 1 and 2 are, respectively, an introducion to metal-nucleobase chemistry and an overview of the theoretical methods that are used in this thesis.

In chapter 3 accurate *ab initio* reference data[2] on hydrogen bond lengths and their binding energy of the Watson-Crick pairs AT and GC are compared with DFT results. Amongst a series of seven popular density functionals (a.o. BLYP, B3LYP, BP86 and PW91), BP86 gave the best agreement. This result is based on the gas phase in which no crystal environment (water molecule, Na^+-ion) is taken into account. If this is done, though, structural changes that are induced hereby are mirrored by all used functionals. Promising functionals are BP86 and PW91, which reproduce both the *ab initio* as well as experimental values of the hydrogen bond lengths.[3] At contrast, B3LYP gives for AT in the gas phase hydrogen bond lengths that match these of the experiment, but this is not true for GC. It is however more important to notice in case of B3LYP, that when the crystal environment is taken into account, the hydrogen bonds become too long with respect to the exprimental data.[3] This shows that the widely used and popular functional B3LYP contains a deficiency.

Chapter 4 shows how rare tautomers of, among others, 1-MeUH or 1-MeTH can be stabilised through coordination to Pt^{II}. This is visualised by calculating the reaction energy of the exchange of the water ligand in the complex $[Pt^{II}(A)(B)(C)(OH_2)]^q$ (A, B, C = NH_3, Cl^-) by 1-MeUH or 1-MeTH. The reaction energy was calculated using the BP86 functional, with

relativistic corrections. The gas phase results are governed by the net charge q, which is determined by the composition of the ligands A, B and C. The used complexes had a net charge varying between +2 and –1. In case the complex has a charge of +2 (A, B, C are all NH_3), all products have a negative reaction energy, which means all of these can be formed. In the case q = +2, the reaction energy is also the most negative, meaning the products are the best stabilised. Decreasing the charge will make as a consequence, the reaction energy more positive, thus the products are less stabilised. In some cases, the reaction energy becomes positive, which means the respective product can not be formed. Taking into account solvation significantly affects the results as compared to the gas phase. The reaction energies become less exothermic, in some cases even positive. However, the dependence on the charge is less pronounced than in the gas phase. Moreover, solvation makes the exchange reaction of the positively charged Pt^{II} complexes less favourable, but it promotes those of the negatively charged complexes. In solution a rare tautomer of both 1-MeUH as well as 1-MeTH is stabilised via Pt^{II} binding at N3. This rare tautomer species is present in its 4-hydroxo-2-oxo form, which is also found experimentally.[4] The results on Pt^{II} bonding to N1 of guanine show that a rare tautomer is stabilised, in which the proton binds to N7. This is also experimentally observed.[5] Without such coordination, this rare tautomer would be present in only very small amounts ($K_{Taut} \sim 10^{-5}$).[6] The term "metal-stabilised rare nucleobase tautomer", which has been proposed about two decades ago, is thus very adequate.

Starting from chapter 5 larger aggregates of nucleobases are considered, in particular quartets. In these cases, dispersion can play an important role and one has to use a functional that is able to cope with dispersion. This is not possible with the usual GGA functionals but it is with the so-called dispersion-corrected functionals such as BLYP-D, BP86-D and PBE-D. The addition '-D' indicates the dispersion correction. First, for a group of molecules in which dispersion plays a big role, *ab initio* data[7] were compared to the DFT-D results, and BLYP-D gave the best agreement. Moreover, the BLYP-D functional could to excellent agreement reproduce *ab initio* binding energies of stacked and hydrogen bonded AT and GC.[2] Next, by use of the BLYP-D functional, minima of several A_4 quartets were localised. These quartets differ in hydrogen bond patterns. They have a common donor, the N-H bond of the amino group, but the acceptor site differs and can be N1 (A_4-N1), N3 (A_4-N1) or N7 (A_4-N7). In this chapter, it is clearly shown that a correct treatment of dispersion is decisive for finding the correct minima. More drastically, the minima obtained with BLYP-D, are no minima for

B3LYP. This illustrates the inability of B3LYP to treat dispersion. By choosing the wrong functional, the order of stabilisation of different minima of one and the same A_4 quartet is predicted wrongly, which leads to chemically wrong conclusions.

Chapter 6 is an extension of chapter 5 in the sense that the ion binding properties of the adenine quartets presented in the latter are investigated here. The minima of A_4 associated with an ion are in general non-planar structures. An important feature in this chapter is the simulation of an A_4 as were it in a stack of quartets. Therefore, the quartet was optimised in a planar manner, but the amino groups were allowed to pyramidalise. Moreover, the distance of the ion with respect to the central cavity was optimised as well. These simulations lead to an interesting view and the quartets can be grouped into two classes. The ion binding pattern of A_4-N1 and A_4-N7 resemble each other but A_4-N3 shows a reversed behaviour. Essentially, the orientation of the N-H bond that does not participate in hydrogen bonding with an adjacent adenine base is decisive for what ion is bound and for how it is bound. In the case of A_4-N1 and A_4-N7, this bond points into the center which makes a favourable interaction possible with the anion. The N3 atoms of these quartets are at the outside of the quartets, and can interact through their lone pairs with cations. In A_4-N3 this N-H bond is exposed to the outside which favours an interaction with an anion, its N1 atoms point to the center of the quartets where it can interact with a cation through its lone pair. In summary, through the orientation of the N-H bond that does particiapte in hydrogen bonding with an adjacent base, anions prefer to bind at the outside and cations at the outside (A_4-N1 and A_4-N7) and *vice versa* (A_4-N3).

In chapter 7, it is investigated whether a stack of G_4 and A_4-N7 quartets can be used as a receptor for NaCl. Hereby, advantage has been taken of two preceding chapters. From chapter 5 it is taken that the BLYP-D functional treats dispersion very well. This is also needed here, because a number of stacks will be described. From chapter 6 it has been taken, that an A_4 quartet can accept an anion. G_4 is known from experiment to be a cation acceptor. Starting from four adenine and four guanine bases, and the ions Na^+ and Cl^-, it is shown that such a receptor can be made, in both the gasphase and solution. The stack of a guanine and an adenine quartet having Na^+ and Cl^- inside is quite stable with respect to the same stack with only Na^+ or only Cl^-.

Samenvatting

In dit proefschrift worden studies gepresenteerd van berekeningen op basis van dichtheidsfunctionaaltheorie (DFT) aan metaal-nucleobase complexen. Dit onderzoeksveld is zeer breed,[1] en om die reden is er een keuze gemaakt tussen diverse onderzoeksonderwerpen. De volgende systemen van nucleobasen worden beschreven. Ten eerste de Watson-Crick basenparen adenine-thymine (AT) en guanine-cytosine (GC) met en zonder kristalomgeving. Ten tweede complexen van platina (Pt^{II}) en 1-methyluracil (1-MeUH) en 1-methylthymine (1-MeTH), waarbij de stabilisering van zeldzame tautomeren centraal staat. Als toevoeging word ook van guanine bepaald, welke tautomere vorm deze base aanneemt, wanneer Pt^{II} aan N1 bindt. Ten derde worden grotere aggregaten van nucleobasen beschreven. Er wordt gedemonsteerd dat adenine (A_4) kwartten de mogelijkheid bezitten anionen en cationen te binden. Tot stot wordt een stapel van een guanine (G_4) en een adenine kwartet gepresenteerd die kan functioneren als een receptor voor natriumchloride (NaCl). Nu volgt een korte samenvating per hoofdstuk.

In hoofstuk 1 en 2 wordt een inleiding over metaal-nukleobase chemie respectievelijk een overzicht van de in dit proefschrift gebruikte theoretische methoden gegeven.

In hoofdstuk 3 worden nauwkeurige *ab initio* literatuurwaarden,[2] van waterstofbruglengten en hun bindingsenergie van de Watson-Crick basenparen AT en GC vergeleken met DFT resultaten. In een serie van populaire funktionalen, o.a. BLYP, B3LYP, BP86 en PW91, gaf BP86 de beste overeenkomst met de *ab initio* data. Dit resultaat is op basis van de gasfase, waarbij de kristal omgeving (water moleculen en Na^+-ion) niet zijn meegenomen. Als dit wel gedaan wordt, worden de geometrische veranderingen, die hierdoor geïnduceerd worden, weerspiegeld door alle gebruikte functionalen. Veel belovende fuctionalen zijn BP86 en PW91, die zowel de *ab initio* als de experimentele waarden voor de waterstofbruglengten reproduceren.[3] Daarentegen geeft B3LYP in de gasfase voor het basenpaar AT goede watersofbruglengten die vergelijkbaar zijn met het experiment, maar dat geldt niet voor GC. Het is echter veel belangrijker te noemen dat wanneer voor B3LYP de kristalomgeving wordt meegenomen, de waterstofbruglengten te lang zijn, met betrekking tot de experimentele data.[3] Dit toont aan dat de veel gebruikte, populaire B3LYP functionaal een deficiëntie heeft.

Hoofdstuk 4 laat zien hoe zeldzame tautomeren van 1-MeUH en 1-MeTH kunnen worden gestabiliseerd door binding aan Pt^{II}. Dit wordt aanschouwelijk gemaakt door de reaktie-energie te berekenen van de uitwisseling van het water ligand in het complex

$[Pt^{II}(A)(B)(C)(OH_2)]^q$ (A, B, C = NH_3, Cl^-) tegen 1-MeUH of 1-MeTH. Reaktie-energieën werden berekend met de BP86 funktionaal, waarbij ook rekening is gehouden met relativistische effekten. De gasfase resultaten worden overheerst door de netto lading q en worden bepaald door de compositie van de liganden. De gebruikte complexen hadden een lading variërend van +2 tot –1. In het geval dat de lading van het complex +2 is (A, B, C zijn allen NH_3), is de reactie energie negatief voor alle producten, wat inhoudt dat deze allen gevormd kunnen worden. De reaktie energie is bij q = +2 het meest negatief. Dat betekent dat de producten het beste worden gestabiliseerd. Het verlagen van de lading heeft tot gevolg, dat de producten minder goed worden gestabiliseerd, dus dat de reaktie-energie hoger wordt. In sommige gevallen wordt een positieve reaktie-energie waargenomen en kan het bijbehorende product niet gevormd worden. Het meenemen van oplosmiddeleffecten (water) beïnvloedt de resultaten in vergelijking met de gasfase, zeer. De reaktie-energie is veel minder exotherm, en in sommige gevallen zelfs positief. Echter, de ladingsafhankelijkheid van de reactie-energie is veel minder uitgesproken dan in de gasfase. Verder wordt door oplosmiddeleffecten de uitwisselingsreaktie minder gunstig als het gaat om positief geladen Pt^{II}-complexen, maar gunstiger als als het gaat om negatief geladen Pt^{II}-complexen. In oplossing wordt een zeldzame tautomeer van zowel 1-MeUH als 1-MeTH gevormd, door Pt^{II} binding aan N3. Deze zeldzame tautomeer komt voor in zijn 4-hydroxo-2-oxo vorm, welke ook experimenteel waargenomen is.[4] De resultaten van Pt^{II}-binding aan N1 van guanine laten zien, dat guanine in oplossing als een zeldzame tautomeer voorkomt, waabij het proton aan N7 gebonden is. Dit is ook experimenteel geobserveerd.[5] Zonder deze platinacoördinatie zouden deze zeldzame tautomeren slechts in zeer kleine hoeveelheden voorkomen ($K_{Taut} \sim 10^{-5}$).[6] De term “metaal-gestabiliseerde zeldzame tautomeer”, die ongeveer twintig jaar geleden werd geopperd, is dus zeer passend.

Vanaf hoofdstuk 5 worden grotere systemen van nucleobasen, in het bijzonder kwartetten, beschreven. Hierbij kan dispersie een grote rol spelen, en het is dan ook nodig een functionaal te kiezen die dispersie goed beschrijft. De gebruikelijke GGA-funktionalen kunnen dit niet. Dit kan echter wel met zogenaamde dispersie-gecorrigeerde funktionalen zoals BLYP-D, BP86-D en PBE-D. De toevoeging ‘-D’ geeft de dispersie correctie aan. Eerst werden van een groep molekulen waar dispersie een grote rol speelt, *ab initio* data[7] vergeleken met de DFT-D resultaten, waarbij BLYP-D de beste overeenkomst gaf. Tevens is de BLYP-D funktionaal in staat om zeer nauwkeurig de *ab initio* bindingsenergieën te reproduceren van de basenparen AT en GC in zowel gestapelde als waterstofbruggebonden

vorm. Daarna worden met de BLYP-D funktionaal minima van A_4 kwartetten bepaald. Deze kwartetten hebben een gemeenschappellijke donor, namelijk de N-H binding van de aminogroep maar de acceptor is verschillend en kan N1 (A_4-N1), N3 (A_4-N3) of N7 (A_4-N7) zijn. In dit hoofdstuk komt zeer duidelijk naar voren, dat het correct beschrijven van dispersie doorslaggevend is voor het vinden van de juiste minima. Sterker, er wordt gevonden dat de BLYP-D minima niet overeenkomen met B3LYP minima. Dit illustreert dat de B3LYP funktionaal niet in staat is tot het juist beschrijven van dispersie. Door het verkeerd kiezen van de funktionaal wordt de stabilisatievolgorde van diverse minima van één en hetzelfde kwartet verkeerd voorspeld, wat leidt tot chemisch foutieve conclusies.

Hoofstuk 6 is een uitbreiding op hoofdstuk 5 omdat van de daar voorgestelde adenine kwartten wordt onderzocht of zij anionen of cationen kunnen binden. Een belanrijk punt in dit hoofdstuk is het bescrhijven van deze kwartetten als zaten zij in een oligonucleotide. Hiertoe werd het kwartet in een vlakke geometrie geoptimaliseerd, met uitzondering van de aminogroep, die mocht pyramidaliseren. Tevens werd de afstand van het ion tot het centrum van het kwartet geoptimaliseerd. De resultaten van deze optimalisaties leidden tot een interessant beeld waarbij de kwartetten in twee groepen kunnen worden verdeeld. De ion-bindingspatronen van A_4-N1 en A_4-N7 lijken opelkaar, terwijl dat van A_4-N3 een tegengesteld gedrag vertoont. In essentie is het de orientatie van de N-H binding van de aminogroep die niet deelneemt aan de waterstofbrug binding met aan aangrenzende adenine base, die doorslaggevend is voor welk ion er wordt gebonden en ook voor de manier waarop het ion gebonden wordt. In A_4-N1 en A_4-N7 wijst deze binding naar het centrum van het ion en kan er een gunstige donor-acceptor-wisselwerking onstaan tussen kwartet en ion. Het N3 atoom van deze kwartetten ligt aan de buitenkant en kan interacteren via zijn vrije elektronpaar met een cation. In kwartet A_4-N3 wijst die N-H binding naar buiten, zodat het kan interacteren met een anion. De N1 atomen liggen in het centrum en zorgen, door middel van een vrij elektronpaar, voor een gunstige interactie met een cation. Kortom, door de orientatie van de binding van de aminogroep die niet deelneemt aan een waterstofbrug met een aangrenzende adenine base, binden anionen bij voorkeur aan de binnenkant en kationen bij voorkeur aan de buitenkant van het kwartet (A_4-N1 en A_4-N7) en *vice versa* (A_4-N3).

In hoofdstuk 7 wordt onderzocht of een systeem van een gestapel van een G_4 en een A_4-N7 kwartet kan worden gebruikt als een receptor voor NaCl. Hierbij wordt gebruik gemaakt van twee voorgaande hoofdstukken. Uit hoofdstuk 5 het resultaat, dat BLYP-D de dispersie goed beschijft. Dit is ook hier noodzakelijk omdat er diverse stapels worden

gepresenteerd. Uit hoofdstuk 6 het gevonden resultaat dat A_4 anionen kan accepteren, van G_4 is het experimenteel bekend dat dit kationen kan accepteren. Uitgaand van de bouwstenen, 4 adenine en 4 guanine basen, en de losse ionen Na^+ en Cl^- blijkt het inderdaad mogelijk een receptor te construeren, in zowel de gasfase als oplossing. De stapel van een guanine en een adenine kwartet met daarin Na^+ en Cl^- is in vergelijk met eenzelfde stapel met of alleen Na^+ of alleen Cl^-, bijzonder stabiel.

Zusammenfassung

In dieser Doktorarbeit werden Studien von Berechnungen auf Basis von Dichtefunktionaltheorie (DFT) an Metall-Nukleobase-Komplexen vorgestellt. Dieser Forschungsbereich ist sehr breit gefächert.[1] Daher wurden verschiedene Forschungsthemen ausgewählt. Die folgenden Systeme von Nukleobasen werden beschrieben: (i) Watson-Crick-Basenpaare Adenin-Thymin (AT) und Guanin-Cytosin (GC) mit und ohne Kristallumgebung; (ii) Komplexe von Platin (Pt^{II}) mit 1-Methyluracil (1-MeUH) und 1-Methylthymine (1-MeTH), wobei die Stabilisierung seltener Tautomere im Mittelpunkt steht. Außerdem wird an Guanin untersucht, welches Tautomer dieser Base vorliegt, wenn Pt^{II} über N1 gebunden ist; (iii) größere Anordnungen von Nukleobasen, insbesondere Quartette. Es wird unter anderem gezeigt, dass A_4-Quartette in der Lage sind, Anionen zu akzeptieren. Zum Schluss wird ein Stapel bestehend aus einem Guanin- (G_4) und einem A_4-Quartett vorgestellt, der als Rezeptor für Natriumchlorid (NaCl) dienen kann. Es folgt jetzt eine Kurzfassung der Kapitel, in denen die Forschung besprochen wurde.

Bei Kapitel 1 und 2 handelt es sich um eine Einleitung in Metall-Nukleobasen-Chemie und eine Übersicht der in dieser Doktorarbeit verwendeten theoretischen Methoden.

In Kapitel 3 werden genaue *ab-initio*-Daten[2] von Wasserstoffbrückenlängen und deren Bindungsenergien in den Watson-Crick Basenpaaren AT und GC mit denen von DFT-Berechnung verglichen. In einer Reihe bekannter Funktionale (u.a. BLYP, B3LYP, BP86 and PW91) entsprachen die Ergebnisse von BP86 den *ab-initio*-Daten am besten. Dieses Ergebnis gilt in der Gasphase, wobei die Kristallumgebung der Basenpaare nicht berücksichtigt wurde. Im Fall der Berücksichtigung werden die geometrischen Änderungen, die dadurch induziert werden, von allen gebrauchten Funktionalen wiedergespiegelt. Viel versprechende Funktionale sind BP86 und PW91, da sie sowohl die *ab-initio*- als auch die experimentellen Daten für die Wasserstoffbrückenlänge wiedergeben.[3] Dagegen ergibt B3LYP für das Basenpaar AT in der Gasphase gute Werte für die Wasserstoffbrückenlängen,

die vergleichbar sind mit experimentellen Daten, jedoch nicht für das Basenpaar GC. Ein wichtigeres Ergebnis für B3LYP ist, dass mit der Berücksichtigung der Kristallumgebung, die Wasserstoffbrücken bezogen auf experimentelle Daten[3] zu lang werden. Dieses zeigt, dass das oft angewendete B3LYP Funktional eine Schwachstelle hat.

Kapitel 4 zeigt, wie seltene Tautomere von 1-MeUH und 1-MeTH durch Pt^{II}-Koordinierung stabilisiert werden können. Dies wird durch die Berechnung der Reaktionsenergie untersucht, wobei der Aqua-Ligand des Komplexes $[Pt^{II}(A)(B)(C)(OH_2)]^q$ (A, B, C = NH_3, Cl^-) durch 1-MeUH oder 1-MeTH ausgetauscht wird. Die Reaktionsenergie wird mithilfe des BP86-Funktionals berechnet, wobei relativistische Effekte berücksichtigt werden. Die Ergebnisse in der Gasphase zeigen, wie die Reaktionsenergie von der Nettoladung q beherrscht wird, die durch die Identität der Liganden bestimmt wird. Die verwendeten Komplexe haben eine zwischen +2 und –1 variierende Ladung. Falls q = +2 ist, ist die Reaktionsenergie für alle Produkte am negativsten, was zeigt, dass diese Produkte am besten stabilisiert werden. Durch Absenkung der Ladung wird die Reaktionsenergie positiver, also sind die Produkte weniger stabilisiert. In manchen Fällen ist die Reaktionsenergie sogar positiv, weshalb das entsprechende Produkt nicht gebildet werden kann. Die Berücksichtigung von Lösungsmitteleffekten (Wasser) beeinflusst die Ergebnisse im Vergleich zur Gasphase stark. Die Reaktionsenergien sind bedeutend weniger exotherm, und in manchen Fällen sogar positiv. Allerdings ist die Ladungsabhängigkeit weniger ausgeprägt als in der Gasphase. Darüber hinaus wird die Austauschreaktion durch Lösungsmitteleffekte nicht begünstigt, wenn positiv geladene Komplexen involviert sind. Eine solche Begünstigung tritt jedoch wohl auf, wenn es sich um negativ geladene Komplexe handelt. In Lösung wird ein seltenes Tautomer sowohl von 1-MeUH als auch von 1-MeTH durch Pt^{II}-Bindung an N3 stabilisiert. Dabei handelt es sich um die 4-Hydroxo-2-oxo-Form, die auch experimentell beobachtet werden konnte.[4] Die Ergebnisse für über N1 am Pt gebundenes 9-Methylguanin ergaben, dass ein seltenes Tautomer von Guanin in Lösung mit N7-Protonierung vorliegt. Dieses seltene Tautomer wurde auch experimentell beobachtet.[5] Ohne Pt^{II}-Koordinierung kommen diese seltenen Tautomere nur in geringen Mengen vor ($K_{Taut} \sim 10^{-5}$).[6] Der Ausdruck “metallstabilisiertes seltenes Tautomer”, der vor etwa zwanzig Jahren vorgeschlagen wurde, ist darum sehr passend.

Ab Kapitel 5 werden größere Anordnungen von Nukleobasen, insbesondere Quartette, beschrieben. Hierbei kann Dispersion eine große Rolle spielen, deswegen ist es notwendig ein Funktional anzuwenden, das in der Lage ist, Dispersion zu beschreiben; die üblichen

GGA-Funktionale sind hierzu nicht in der Lage. Möglich wird eine solche Beschreibung jedoch durch die Anwendung von dispersions-korrigierter Funktionalen wie BLYP-D, BP86-D und PBE-D möglich. Der Zusatz '-D' zeigt diese Dispersionskorrektur an. Zuerst werden von einer Gruppe Moleküle, wobei Dispersion eine große Rolle spielt *ab-initio*-Daten[7] verglichen mit DFT-D Daten, wobei BLYP-D die beste Übereinstimmung hatte. Weiterhin ist BLYP-D in der Lage sehr genaue *ab-initio*-Daten der Bindungsenergie von AT und GC, sowohl in gestapelter als auch in wasserstoffbrückengebundener Anordnung, wiederzugeben. Anschließend wurde von einigen A_4-Quartetten die Gleichgewichts-Struktur bestimmt. Die Quartette unterscheiden sich zwar in ihrem Wasserstoffbrückenbindungsmuster, haben jedoch miteinander einen gemeinschaftlichen Donor, eine der N-H-Bindungen der Aminogruppe, wobei der Akzeptor variabel ist. Dabei handelt es sich um N1 (A_4-N1), N3 (A_4-N3) oder N7 (A_4-N7). In diesem Kapitel ist eine Berücksichtigung von Dispersion erforderlich, damit das korrekte Minimum gefunden werden kann. Die Gleichgewichts-Struktur für BLYP-D entspricht nicht der von B3LYP. Damit ist noch eine weitere Schwachstelle für B3LYP illustriert: es ist nicht in der Lage Dispersion zu beschreiben. Die falsche Wahl eines Funktionals kann dazu führen, dass die energetische Folge der Stabilisierung von mehreren Minima von ein und demselben A_4-Quartet falsch vorhersagt wird. Dieses kann zu chemisch falschen Schussfolgerungen führen.

Kapitel 6 ist eine Erweiterung von Kapitel 5, weil dort untersucht wird, ob die dort vorgestellten Adenin-Quartette Anionen oder Kationen binden können. Ein wichtiger Punkt dieses Kapitels ist die Simulation eines A_4 als sei es in einem vierstrang-Oligonukleotid. Daher wird eine planare A_4-Struktur optimiert, wobei die Aminogruppe pyramidalisieren darf. Auch wird der Abstand zwischen Ion und Zentrum des Quartetts optimiert. Die Ergebnisse dieser Berechnungen zeigen ein interessantes Bild, wobei sich die Quartette in zwei Gruppen aufteilen lassen. Die Ionenbindungsmuster von A_4-N1 und A_4-N7 sind vergleichbar, das von A_4-N3 verhält sich dagegen gegenteilig. Im Wesentlichen ist die Orientierung der N-H Bindung der Aminogruppe, die nicht an Wasserstoffbrückenbindung mit einer angrenzenden Adenin Base beteiligt ist, ausschlaggebend dafür, ob und wie ein Anion oder Kation gebunden wird. In A_4-N1 und A_4-N7 ist diese Bindung nach innen orientiert, was eine günstige Donor-Akzeptor-Interaktion mit einem Anion hervorruft. Das N3-Atom liegt an der Außenseite und kann mittels eines freien Elektronenpaars mit einem Kation interagieren. In A_4-N3 ist diese Bindung nach außen orientiert, und kann entsprechend mit einem Anion wechselwirken. Die N1 Atome liegen im Zentrum dieses Quartetts, und

können mittels eines freien Elektronenpaars mit einem Kation interagieren. Somit werden durch die Orientierung der N-H-Bindung der Aminogruppe, die nicht an Wasserstoffbrückenbindung mit einer angrenzenden Adenin Base beteiligt ist, Anionen in der Mitte von A_4-N1 oder A_4-N7 gebunden und Kationen an der Außenseite, oder umgekehrt (A_4-N3).

In Kapitel 7 wird gezeigt, dass ein Stapel von einem G_4- und A_4-N7 Quartett in der Lage ist, als Wirt für Natriumchlorid zu dienen. Hierbei werden Ergebnissen aus vorherigen Kapiteln benutzt: Erstens aus Kapitel 5 wird die Tatsache benutzt, dass BLYP-D in der Lage ist, Dispersion am besten zu beschreiben. Das ist hier auch sehr wichtig, weil in diesem Kapitel gestapelte Komplexe vorkommen. Aus Kapitel 6 wurde die Tatsache benutzt, dass A_4-Quartette in der Lage sind, Anionen zu binden. Aus Experimenten ist bekannt, dass G_4-Quartette Kationen akzeptieren können. Ausgehend von ihren Bausteinen: vier Guanin- und vier Adenin-Basen und den Ionen Na^+ und Cl^- ist es in der Tat möglich, solch einen Rezeptor sowohl in der Gasphase als in Lösung zu konstruieren. Der Stapel aus einem Guanin- und Adeninquartett mit Na^+ und Cl^- im Innern, ist im Vergleich zu den gestapelten Quartetten mit nur einem Ion (Na^+ oder Cl^-) besonders stabil.

References

[1] B. Lippert, B. Lippert, *Prog. Inorg. Chem.* **2005**, *54*, 385-447.

[2] (a) J. Sponer, P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2004**, *126*, 10142-10151.
(b) P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2003**, *125*, 150608-10893.

[3] (a) N. C. Seeman, J. M. Rosenberg, F. L. Suddath, J. J. P. Kim, A. Rich, *J. Mol. Biol.* **1976**, *104*, 109-144.
(b) J. M. Rosenberg, N. C. Seeman, R. O. Day, A. Rich, *J. Mol. Biol.* **1976**, *104*, 145-167.

[4] (a) H. Schöllhorn, U. Thewalt, B. Lippert, *J. Am. Chem. Soc.* **1989**, *111*, 7213-7221.
(b) B. Lippert, *Inorg. Chim. Acta* **1981**, *55*, 5-14.

[5] (a) G. Frommer, I. Mutikainen, F. J. Pesch, E. C. Hillgeris, H. Preut, B. Lippert, *Inorg. Chem.* **1992**, *31*, 2429-2434.
(b) B. Müller, Dissertation, TU Dortmund, **2009**.

[6] B. Lippert, H. Schöllhorn, U. Thewalt, *Inorg. Chim. Acta* **1992**, *198-200*, 723-732.

[7] (a) S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463-1473.
(b) S. Tsuzuki, H. P. Lüthi, *J. Chem. Phys.* **2001**, *114*, 3949-3957.
(c) P. Jurecka , J. Sponer, J. Cerny, P. Hobza, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985-1993.
(d) M. O. Sinnokrot, E. F. Valeev, C. D. Sherrill, *J. Am. Chem. Soc.* **2002**, *124*,