Electronic Structure Studies of Noncovalent Interactions in B-DNA: Effect of Sugar-Phosphate Backbone

Sumit Mittal

A dissertation submitted for the partial fulfillment of BS-MS dual degree in Science

Indian Institute of Science Education and Research Mohali
April 2013

Certificate of Examination

This is to certify that the dissertation titled "**Electronic Structure Studies of Noncovalent Interactions in B-DNA: Effect of Sugar-Phosphate Backbone**" submitted by **Mr. Sumit Mittal** (Reg. No. MS08049) for the partial fulfillment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Prof. K.S. Viswanathan Dr. K. R. Shamasundar Prof. N. Sathyamurthy
(Supervisor)

Dated: April 26, 2013

DECLARATION

The work presented in this dissertation has been carried out by me under the guidance of Prof. N. Sathyamurthy at the Indian Institute of Science Education and Research, Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Sumit Mittal

(Candidate)

Dated: April 26, 2013

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Prof. N. Sathyamurthy

(Supervisor)

ACKNOWLEDGEMENT

I would like to express my sincere thanks and admiration to Professor N. Sathyamurthy for supervising my masters' project. He was flexible about choosing and working on a problem of my interest. He has been supportive and highly motivating at all times in his unique inspiring style. It was a great pleasure to be in his group. For creating one of the best work environments a student could ever ask for, thank you.

I thank Dr. Brijesh K. Mishra for helping with the calculations and the discussions at various times throughout the project. Thank you.

I would not be in the field of computational chemistry if it were not for Dr. Ramesh Ramachandran. He is a brilliant teacher and I had the luxury of being taught by him for 4 years. Thank you.

I would also like to thank Prof. K.S. Viswanathan who has been a steady source of knowledge. He is a wonderful teacher, and more than that, an amazing person. Thank you for giving inputs on all doubts, even the silly ones.

I thank my parents for their support during my whole life. I thank my sisters for everything they have done for me. They all have been amazing and always been there for me when required. I cannot thank them enough. Thank you for your everlasting love and support from hundreds of miles away.

Time flew by in the company of Manish, Manmeet, Nikhil, Gaurav, Asif, Ankit, Nilmani, Keshav and other classmates who are not mentioned by name. I had a great laugh around the campus with you people. You guys have been and always will be my dear friends and family away from home.

I am grateful to IISER Mohali for providing me the computing and library facilities. A special thanks to Dr. Paramdeep Singh for setting up and help up with the softwares used for the work reported in the thesis.

Last but not the least thank you, The Reader. You, if not interested by profession, must be a close friend and surely should have deserved a special mention.

LIST OF FIGURES

LIST OF FIGURES (Cont.)

LIST OF TABLES

ABBREVIATIONS

DNA	Deoxyribose Nucleic Acid
NMR	Nuclear Magnetic Resonance
A	Adenine
T	Thymine
Ad	Deoxyriboseadenosine
Th	Deoxyribosethymidine
dAMP	Deoxyriboseadenosinemonophosphate
dTMP	Deoxyribosethymidinemonophosphate
HF	Hartree-Fock
DFT	Density Functional Theory
GGA	Generalized Gradient Approximation
BLYP	Becke-Lee-Yang-Parr
M06	Minnesota functional
DFT-D	Dispersion corrected density functional theory
BSSE	Basis Set Superposition Error

Table of Contents

Sing the song of Crick and Watson
Wilkins, Franklin in Kings London
Tell of DNA chains spiral
And of ribose sugars (chiral).
Tell of Levene's theory, misbegotten
And three-chain models - best forgotten.
Tell of bases pairing
And our two heroes daring
To outthink the great L Pauling
Giving a structure calling
Loud and clear
The secret of all life, is here!

: J. Field

*Journal of Irreproducible results, **1968**, 17,53*

ABSTRACT

Hydrogen bonding and stacking interactions between nucleobases A and T and their base pairs in B-DNA have been studied by using density functional theory that includes dispersion interaction. The effect of the sugar-phosphate backbone on these interactions has also been investigated.

A comparison of the interaction energy values for the Watson-Crick base pairs of A-T, Ad-Th and dMAP-dTMP reveals that both the sugar and phosphate moieties have only a marginal influence on the hydrogen bond interactions between A and T.

A comparison of the interaction energy values for three different conformations of two stacked base pairs of A-T, Ad-Th and dAMP-dTMP shows that for a given conformation, the addition of the sugar moiety as well as the phosphate group does not influence the stacking interaction in a significant manner.

The role of the sugar-phosphate backbone in the formation of the helical structure of B- DNA was also investigated. It is shown that the base pairs themselves have an inherent tendency to form the double helical structure and the backbone does not significantly influence the helix formation.

Chapter 1

Introduction

1. 1 DNA

Nearly sixty years ago three papers were published in the journal, Nature, pertaining to one of the greatest discoveries of the 20th Century. Two of these papers were on X-ray diffraction studies by M. Wilkins and R. Franklin at King's College, London[1,2] and they demonstrated the helical nature of **deoxyribose nucleic acid or DNA**. In the third paper, J.D. Watson and F. Crick[3] proposed a double helical model as the structure of DNA. For the first time, the inheritance of genetic information could be understood in terms of a real chemical structure!

"This structure has novel features which are of considerable biological interest", wrote Watson and Crick in this historical paper and it aptly describes the importance of DNA. It is an indispuTable fact that all living organisms, from bacteria, flower, and fruit fly to man, depend upon DNA and ribonucleic acid (RNA) for their survival, function and for storage and transfer of the genetic information. It stores information for the synthesis of specific proteins. It might not be THE secret of life but it is "ONE" of the secrets of life, for sure.

Watson and Crick's discovery, as is often the case with a fundamental discovery, transformed the subject of DNA structure into an extremely active area of research. What followed was a plethora of structural studies on nucleic acids using X-ray and NMR[4]. As a result, a database of over a hundred nucleic acid crystal structures is now available[5]. Apart from experimental studies, computational studies have helped in getting more details on the structure and functioning of DNA.

1.1.1 Structure

DNA is a heteropolymer of monomeric units called 2'-deoxyribonucleotide[6]. A nucleotide consists of a nitrogenous base, a deoxyribose sugar, and a phosphate group (Figure 1.1).

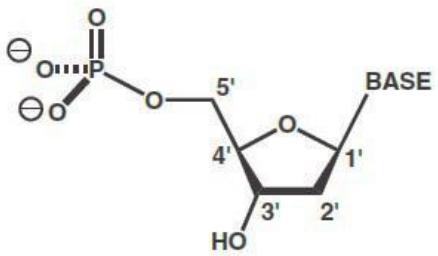

Figure 1.1: The basic structure of a 2'-deoxyribonucleotide.[6]

There are four types of nucleotides present in DNA, which differ by the attached nucleobase. These nitrogenous bases are adenine (A), guanine (G), thymine (T) and cytosine(C) (Figure 1.2). Two of these bases, Adenine and Guanine, are bicyclic and belong to the category of purine. Cytosine and Thymine, on the other hand, are monocyclic and fall into the category of pyrimidine.

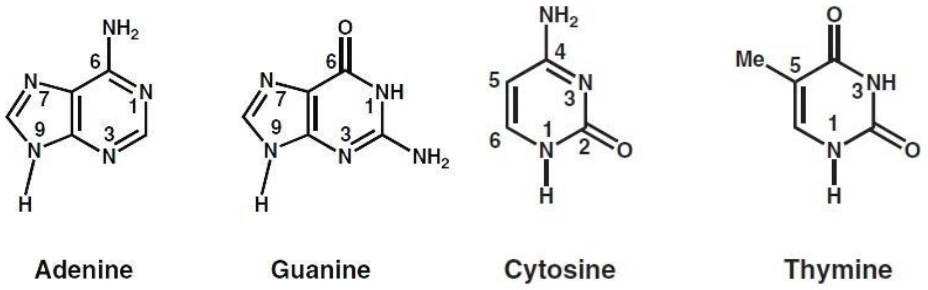

Figure 1.2: Structures of nitrogenous bases (purines and pyrimidines) present in DNA.[6]

Three of these bases have exocyclic oxygens in keto form, and three of them have exocyclic nitrogens in amino form. Exocyclic amino group of A, G and C make these bases non-planar. Nucleobases link with the deoxyribose sugar resulting in the formation of 2'-deoxyribonucleosides. This linkage between the base and the sugar is known as β-glycosidic or sometimes the nucleosidic bond. In the case of purines, this bond is formed between the N-9 and C-1' and in pyrimidines this occurs between the N-1 and C-1'. The resulting nucleosides are 2'-deoxyadenosine (dA), 2'-deoxyguanosine (dG), 2'-deoxycytidine (dC), 2'-deoxythymidine (dT) (Figure 1.3).

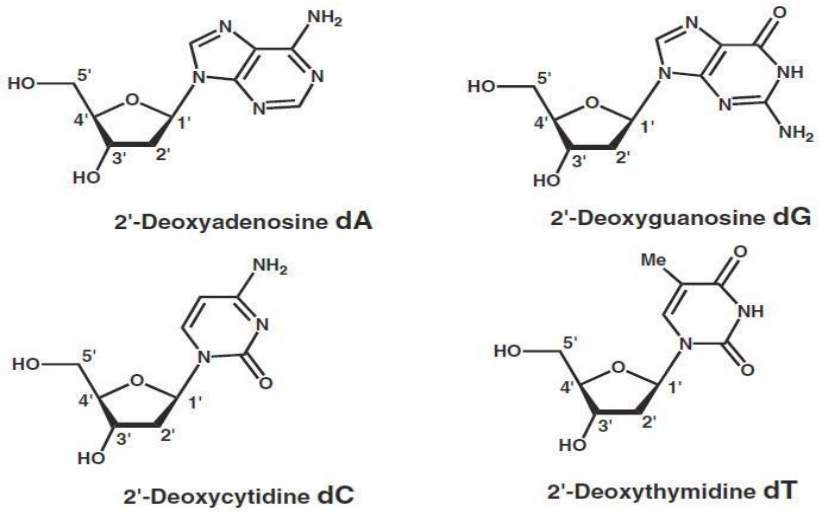

Figure 1.3: Structures of the four deoxyribonucleosides present in DNA.[6]

Deoxyribose sugars are finally connected to phosphate groups resulting in the formation of the basic building blocks of DNA: 2'-deoxyribose nucleotides. The bond between the deoxyribose sugar of the nucleoside and the phosphate group is a 3'-5' phosphodiester linkage. There are four nucleotides, differing by the type of base involved, known as 2'-deoxyadenosine 5'-monophosphate (dpA), 2'-deoxyguanosine 5'-monophosphate (dpG), 2'-deoxycytidine 5'-monophosphate (dpC) and 2'-deoxythymidine 5'-monophosphate (dpT) (Figure 1.4).

The deoxynucleotides are joined together by **phosphodiester links** forming oligodeoxynucleotides or polydeoxynucleotides (depending on the length of units combined) (Figure 1.5).

The chain of phosphodiester links and deoxyribose sugar constitutes one of the main components of DNA: the sugar-phosphate backbone. As the name suggests, this backbone provides stability to the secondary structure of DNA. Nucleobases are considered as 'side-chains' in the DNA structure.

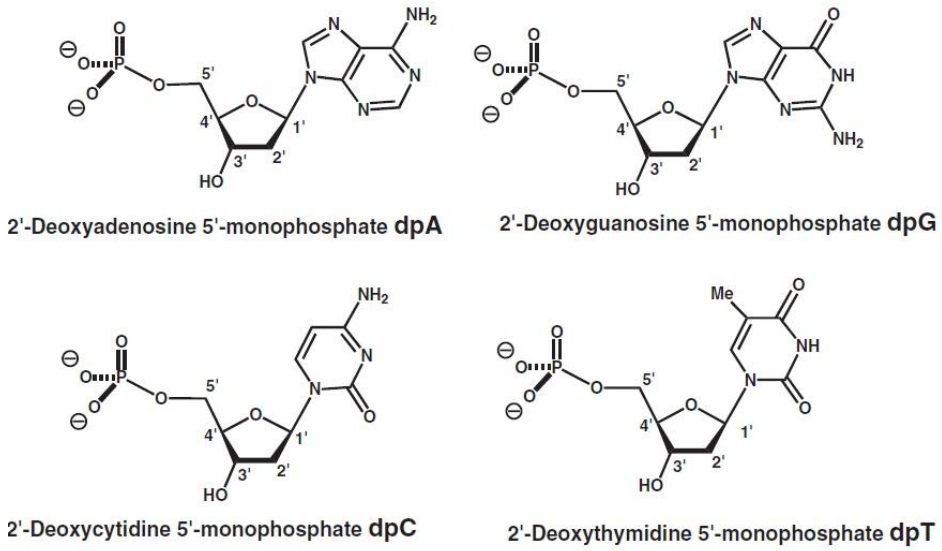

Figure 1.4: Structures of the four deoxyribonucleotides present in DNA.[6]

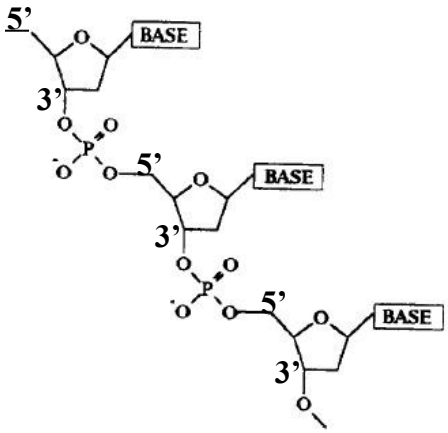

Figure 1.5: An oligodeoxynucleotide (3 units) formed by the phosphodiester link between three deoxynucletides.[6]

By convention, a DNA chain begins at the 5'- end (i.e., where carbon atom C-5' of the terminal residue is involved in a phosphodiester link) and terminate at the 3'- end (where carbon atom C-3' is not involved in a phosphodiester link). Each chain is therefore said to run by convention from5' to 3' (5'→3').

The three dimensional structure of DNA is that of a double helix (Figure 1.6), now immortalized by many books, articles, documentaries and films. It consists of two interwined helical strands in anti-parallel arrangement with one running in 3' to 5' direction and the other in a 5' to 3' direction. The two strands are twisted around a common central axis. The outer edges are formed by the sugar-phosphate backbone, while bases are projected, from each of the two strands, into the interior of the molecule, like the "rungs on a ladder".

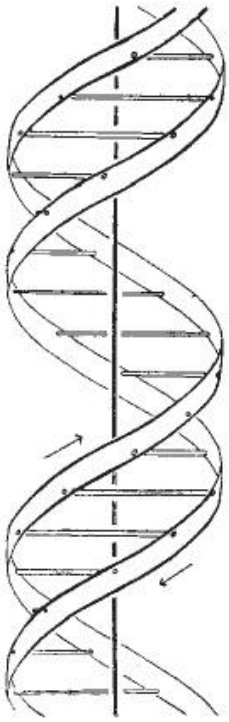

Figure 1.6: Odile Crick's most famous drawing. Francis Crick's wife was an artist, specializing mainly in Figure drawing. This sketch of the anti-parallel 10-fold double helix for the first Watson and Crick paper in *Nature*[3] has become the most famous image she created.

1.1.2 Double Helical Forms of DNA:

Depending on the environment present such as relative humidity, types of salt and solvent, several forms of DNA have been discovered. Three main forms of DNA double helix have been identified as A-, B-, and Z-DNA.

B-DNA, the classic structure first proposed by Watson and Crick, is the dominant form under physiological conditions (neutral pH, room temperature, about 200mM NaCl). This form of DNA is what is present in the genome and is the most important form by biological means. The helix is right-handed and makes one complete turn every 34 Å. The vertical distance between two neighboring base pairs is 3.4 Å along the double helix and hence there are about 10 base pairs per turn in B-DNA. It has two principal grooves, a wide major groove and a narrow minor groove. The grooves are very important for the functioning of DNA such as in DNA-Protein interactions.

The A-form of DNA is observed in environments with low water activity associated with low relative humidity or in solutions with high salt or high ethanol content. Like in B-DNA, in A-DNA the two complementary strands are anti-parallel to each other and form the right-handed helices. The vertical distance between the base pairs in the A-form is 2.6 Å and there are approximately 11 base pairs per turn. This shortness of A-DNA compared to B-DNA is because the base pairs in A-DNA are inclined with respect to the double helix axis. The actual physical separation between the base pairs has been measured to be the same as in B-DNA i.e. 3.4 Å. The only major difference from B-DNA is that the sugar conformation in the A-form is C3'-*endo* whereas that in the B-form is C2'-*endo (*Figure 1.7).

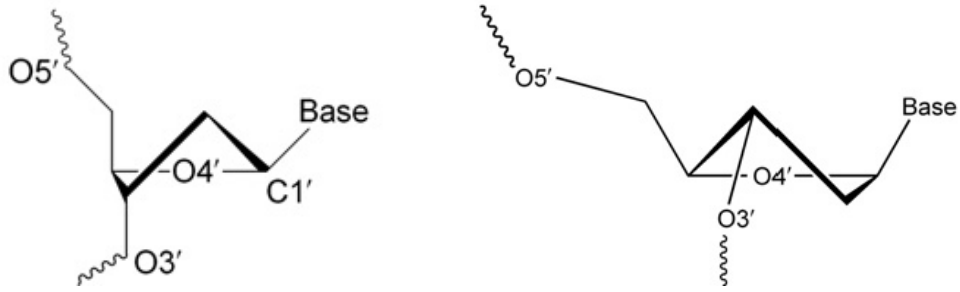

Figure 1.7: Two common pucker types: C2'– *endo* (left), and C3'– *endo* (right).[6]

The sugar conformations are commonly classified as *endo* and *exo*. *Endo* indicates that the given atom is displaced on the same side of the ring as the C5' atom while *exo* means that the given atom is displaced on the opposite side of the C5' atom.

While A- and B-DNA are right handed, another form of DNA double helix, Z-DNA is left-handed. The DNA molecule with alternating G–C sequences in alcohol or in high salt

solution tends to have such a structure. The vertical distance between base pairs is 3.8 Å and one complete turn comprises 12 base pairs. The DNA backbone takes on a zigzag appearance. One interesting feature of the Z-form is the sugar conformation. For pyrimidines, the sugar pucker conformation is C2'-*endo* and for purines, it is a C3'-*endo*. Unlike A- and B-DNA, which can be adopted by any arbitrary sequence of nucleotides, not any sequence can adopt the Z form. An alternate purine-pyridine sequence is strongly favored to form Z-DNA. So, in one sense, the repeating unit is a dinucleotide rather than a nucleotide as in A- and B-type conformations.

Usually these different forms of DNA double helix are described in terms of some parameters: handedness of helix, helix diameter, length of one complete turn (helix pitch), vertical distance between base pairs (rise), number of residues per turn, angle of rotation between adjacent base pair planes (twist). A summary of all the main differences between B-, A- and Z-forms of DNA in terms of dimensions, angles and conformations is presented in Table 1.1.

Table 1.1: Differences between the three helical forms of DNA.

	B-DNA	**A-DNA**	**Z-DNA**
Helical sense	Right handed	Right handed	Left handed
Helix pitch (Å)	34	28	45
Residues per turn	10	11	12
Helix rise (Å)	3.4	2.6	3.8
Twist (°)	36	33	-30
Sugar pucker conformation	C2'-*endo*	C3'-*endo*	C2'- *endo* for C; C3'-*endo* for G

Apart from the covalent interactions involved, the anti-parallel strands are held together by hydrogen bonds between complementary base pairs, the van der Waals base-stacking interactions, and the hydrophobic interactions. These non-covalent interactions are crucial both for structure and functioning of DNA.

1.2 Non-covalent Interactions

Atomic and molecular interactions are of two types: covalent and non-covalent. The first theoretical description of a covalent bond came from G.N. Lewis in 1916,[7] well before the development of quantum theory and now, after years of studies, bond formation and bond breaking processes are understood well experimentally as well as theoretically. There are various spectroscopic techniques to study molecules in gas phase. Computational chemistry methods successfully describe electronic structure of isolated molecules with chemical (1.0 kcal/mol) and often sub-chemical accuracy (0.1 kcal/mol).

Interactions between atoms and molecules, which do not involve formation or breaking of a covalent bond are known as non-covalent interactions. These interactions were first recognized by van der Waals in 1873.[8] While covalent interactions are of short range with covalent bonds generally being shorter than 2 Å, non-covalent interactions are known to act at distances of several angstroms.

Non-covalent interactions can be classified into different types, based on the fundamental forces involved in the interaction: between permanent electric multipoles (electrostatic interaction), between a permanent electric multipole moment and an induced multipole (induction or Debye force), and between an instantaneous multipole and an induced multipole moment (dispersion interaction or van der Waals interaction). To balance the effect of these attractive forces and to prevent interacting systems approaching too closely, there are repulsive forces known as exchange repulsion or Pauli repulsion. The relative contribution of these terms is system dependent.

Non-covalent interactions are considerably weak when compared to the covalent or pure ionic interactions, with a stabilization energy value of the order of a tenth to several kcal/ mol. And yet, non-covalent interactions are of utmost importance in almost every aspect of life. These are responsible for the very existence of liquids, particularly water, the elixir of life. Condensed-phase chemistry, solvation phenomena, molecular crystals, catalysis, polymer science are just some of the several fields of modern chemistry that are influenced by non-covalent interactions.[9] They are extensively involved in biological systems, molecular

recognition, structure and functioning of proteins, RNA, DNA, and drug designs. Easy formation and easy breaking property of non-covalent interactions is very important. What seems to be a weakness turns out to be the strength of such interactions! As Nobel Laureate Jean-Marie Lehn mentioned,[10] "Noncovalent interactions define the inter-component bond, the action and reaction, in brief, the behavior of molecular individuals and populations". In conclusion, a study of non-covalent interactions can reveal fundamental information about the chemistry that is happening in the world around us.

1.2.1 Hydrogen Bond:

Lying at the boundary of weak dispersion forces and the strong covalent bond, hydrogen bond is perhaps the most important and useful chemical structural concept in all of chemistry. Hydrogen bond plays a key role in various chemical and biological phenomena. It is responsible for the structure and properties of water. Further, it plays a key role in determining structures and, consequently, shapes, properties and functions of biomolecules.

It is difficult to define H-bonds in terms of all the features ascribed to them in different branches of science. The most recent and general definition states,[11] *"the H-bond is an attractive interaction between the hydrogen from a group X-H and an atom or group of atoms Y, in the same or different molecule(s), where there is evidence of a partial bond formation"*. At its core, the hydrogen bond is an attractive interaction between a proton donor D-H and an acceptor moiety A[12,13]

D-H … A

It is best considered as an electrostatic interaction but it can also have significant covalent and dispersion character. However, unlike other noncovalent interactions, it involves sharing of a hydrogen atom between two electronegative atoms (the most frequent example) that causes the linearity of the X-H…Y arrangement. This sharing results in a significant change in the properties of the X-H covalent bond. In most cases, the covalent bond becomes weaker upon formation of a H-bond and this results in a redshift of the corresponding vibrational stretching frequency of X-H. This shift is the most important and easily detecTable manifestation of the formation of the H-bond. However, it must be added that there are instances in which a hydrogen shift may exhibit no shift or blue shift in the stretching frequency.[14]

1.2.2 π - π Interactions:

π- π interactions are also an important type of noncovalent interactions that involve delocalized **π** systems. These interactions generally refer to stacking between aromatic moieties. An electron rich **π** system can also interact with a metal, an anion, or another molecule. The crystal structure of the complex between aniline and p-dinitrobenzene was the first structural report[15] showing aromatic rings in a stacked arrangement. These interactions play a key role in the structure of DNA and proteins, host-guest interactions, molecular recognition, supramolecular chemistry, crystal packing of organic molecules having aromatic moieties, etc.

"Aromatic stacking" commonly refers both to the geometry of face-to-face juxtaposition of two aromatic molecules, and to the forces that favor this geometry energetically. This face-to-face staking interaction is commonly observed when the subsystems have opposite quadrupole moment.[16]

Despite this enormous importance, noncovalent interactions are beginning to be understood only recently, in biological systems in particular. Their strength, directionality, and preferred geometric conformations, substituent effects are some of the issues that are discussed. Major efforts have gone into devising and employing experimental methods which provide insights into the nature of noncovalent interactions. Techniques such as vibrational spectroscopy, microwave spectroscopy, vibration-rotation tunneling spectroscopy, rotational electronic spectroscopy, mass spectrometry, zero electron kinetic energy (ZEKE) spectroscopy, etc., are used frequently for this purpose. These interactions can also be studied by standard quantum chemistry methods. As the stabilization energy values are usually of the order of a fraction to a few kcal/mol, most accurate methods of quantum chemistry are required for a reliable description of noncovalent interactions.

1.3 Non-covalent Interactions in DNA:

Hydrogen bonds are formed between base pairs of the anti-parallel strands in DNA. Nucleobases can form pairs via H-bonds between their polar groups and there are many such possible pair arrangements.[17] However, a base in the first strand forms a H-bond only with a specific base in the second strand, showing a remarkable pairwise complementarity. Guanine

is able to supply two hydrogen bond donor groups and one acceptor group to complement the two hydrogen bond acceptor groups and one donor group of cytosine, while adenine is able to provide one hydrogen bond donor and one acceptor group to complement the single hydrogen bond acceptor and one donor group of thymine (Figure 1.8), thus giving rise to the specific **Watson–Crick base pairs**, **G---C** and **A---T**. Watson and Crick found that these specific base pairs are those that best fit within the DNA structure having an isomorphous geometry (distance between N9 of purine base and C6 of pyrimidine base is same) and , ensuring that there are no 'bulges' or 'gaps' within the helix.

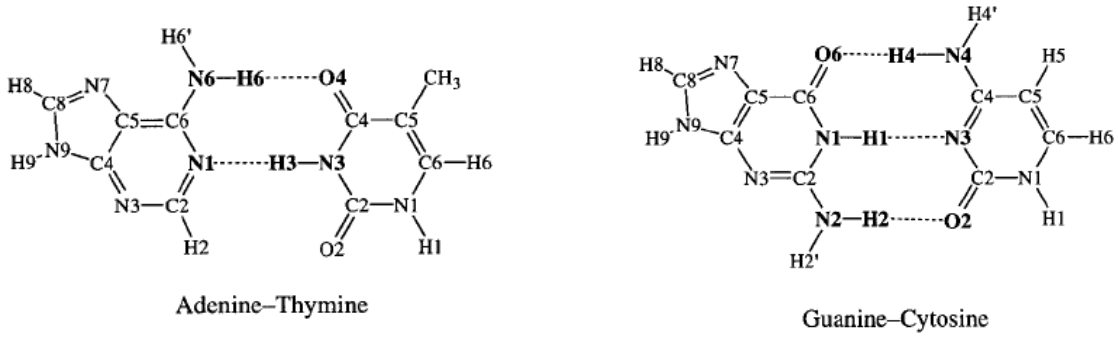

Figure 1.8: Watson-Crick base pairing of nucleobases by H-bonds: adenine with thymine, guanine with cytosine.[18]

Apart from Watson-Crick base pairing, some other non-Watson-Crick base pairs have also been found in several nucleic acid structures. Such base pairing includes reverse Watson-Crick H-bonding,[19] Hoogsteen base pairing,[20] and wobble base pairing.[21]

H-bonding is only one of the two most important interactions which stabilize the double helical structure of DNA. The second is the **stacking interaction** between neighboring bases along the vertical axis of the helix. However, because of the helical twist, there is also a contribution from inter-strand stacking. In the large majority of all known DNA structures, the bases are in face-to-face contact. They are generally not directly aligned but are rather offset. In general, the distance between two aromatic planes is 3.4 Å in a stacked structure, which corresponds to the rise of the B-DNA helix. The base stacking geometry is described using six degrees of freedom: twist, roll and tilt, which are rotational parameters, and slide, shift and rise, which are translational parameters (Figure 1.9).

Base stacking is basically determined by three contributions: dispersion attraction, short-range exchange repulsion, and electrostatic interaction. The stabilization of base stacking is dominated by the dispersion attraction, which is rather isotropic and proportional to the geometrical overlap of the bases. The vertical distance between stacked bases is determined by a balance between dispersion attraction and short-range exchange repulsion. Finally, the mutual orientation of bases and their displacements are determined primarily by electrostatic interactions.

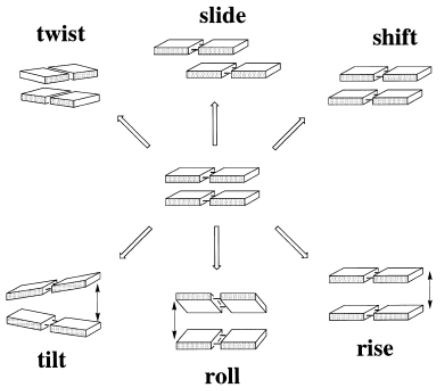

Figure 1.9: The six base step parameters which define the local geometry of a dinucleotide base step. The arrows indicate the motion of the upper base-pair.[22]

Dispersion interactions are crucial for base stacking. While individually weak, a large of number of them are present in DNA, which makes their cumulative contribution substantial. Recent studies on stacked DNA bases pairs showed that stabilization from stacking interactions can be surprisingly large and comparable to hydrogen- bonding.[23] Cryptolepine, a naturally occurring indoquinoline alkaloid used as an antimalarial drug intercalates into DNA and stacks between nucleic acid bases.[24] It has no hydrogen-bonding contacts either with nucleic acid bases or with solvents. The absence of such interactions suggests that stacking interactions provide the stabilizing mechanism for the complex.

The role of stacking in structure and hence functioning of DNA was shown recently by running molecular dynamics (MD) simulations with a modified energy function.[25] As the stacking energy is governed mainly by dispersion energy, MD simulations were performed on a DNA dodecamer first with standard energy functions and later with energy functions

that neglect dispersion interactions. Neglecting the dispersion energy leads to a dramatic change in the double-helical structure (Figure 1.10). The folded structure is transformed to the ladder like structure. The values of rise and twist for folded DNA amount to 3.4 Å and 31.5°, while neglecting dispersion energy changes these values to 5.1 Å and 2.4° respectively. Such drastic geometrical changes would have important biological consequences. Large base-base separation in one strand would lead to a loss of replication and transcription activity, proteins like transcription factors would not be able to bind to specific sequences of the bases exposed in the major groove, water can easily penetrate between the bases.

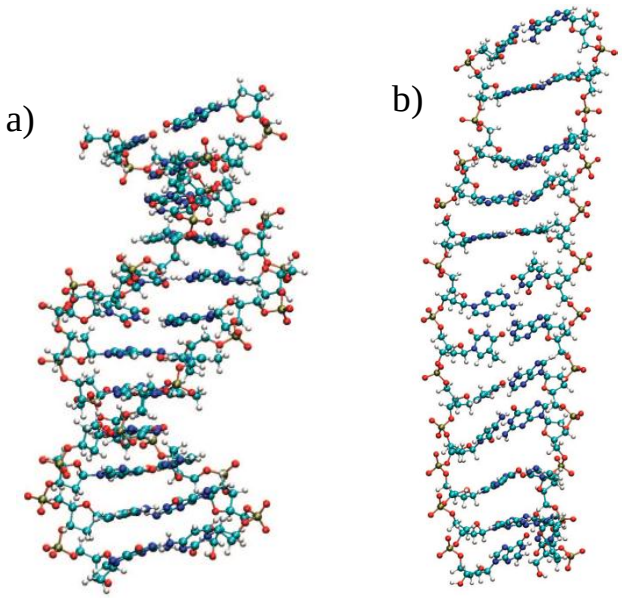

Figure 1.10: Snapshot Figure of (a) the initial crystal structure and (b) the final ladder-like structure of a DNA force field simulation that neglects dispersion interactions.[25]

1.3.1 Experimental Studies:

An experimental characterization of base pairs is extremely difficult and very few reliable studies are available. While no experimental data have been reported on the energetics of base stacking, there is one published gas-phase experiment for the H-bonded nucleic acid base pairs.[26] The "classic" mass field spectroscopy experiment reported stabilization enthalpies for 9-methyladenine…1-methylthymine and 9-methylguanine…1-methylcytosine

base pairs at high temperatures. While this experiment provided thermodynamic parameters for the base pairs, no information on the structure of the base pairs was obtained.

Atomic-resolution X-ray crystallography is a leading structural tool and it provides structural information of a nucleic acid within realistic environments. NMR experiments have also been used extensively to get structural information on nucleic acids. Though these techniques reveal the structure, neither of them provides insights into the energetics. Energy data for such systems can be obtained by condensed-phase thermodynamic experiments, but these experiments do not provide direct structural information.

1.3.2 Theoretical Studies:

In a gas-phase spectroscopic experiment, the geometry of a molecule is determined indirectly via the determination of the rotational constants. These rotational constants, however, do not yield an unambiguous structure of the system. Similarly, though the spectral shifts of X–H stretch vibrations upon formation of a complex indicates the formation of a H-bond, it is an indirect evidence and can be misleading sometimes. Even less information is obtained about the stabilization energy/enthalpy of noncovalent interactions by gas-phase experiments.[27] In contrast, quantum mechanical (QM) methods provide a direct structure–energy relationship and one can get information for any point on the potential energy surface (PES) of the system. Besides structures and energies, additional valuable information such as charge distributions, electrostatic potentials, polarizabilities, etc., can also be obtained by such studies.

Therefore, at present, high level quantum chemical calculations and computer simulations are the main tools to obtain some reference data on structure, energetics, vibrational frequencies, and other properties of nucleic acid base pairs. The role of individual interactions in the total stability of a system is often difficult to get by experimental techniques, but QM calculations can provide this information. For these reasons, many computational studies have appeared in the literature that examine the H-bonding and stacking interactions in B-DNA.

QM studies of DNA base pairs have been attempted for more than 30 years now. Prior to 1985, semi-empirical methods were in use but they are incapable of treating molecular complexes. Till 1994, *ab initio* Hartree-Fock (HF) calculations were carried out

approximation was applied. While such calculations provide a reliable description of H-bonding, they do not provide an accurate description of base stacking as the HF method neglects dynamic correlation effects. Modern specialized pair additive force fields such as that proposed by Cornell *et al.*[28] provide much better results than low-level QM data.[29] The first significant theoretical description[30] of base pairing involved medium quality calculations with the inclusion of electron correlation effects via MP2 theory.

Further extensive electron correlation studies[31] were performed on dozens of nucleic acid base pairs. A very accurate benchmark, S22, for hydrogen-bonding energies in a wide variety of nucleic acid base pairs was reported[32] in 2006 and it was revised very recently by Sherrill et al. [33] by using better basis sets for extrapolations. H-bonding interactions in DNA have also been understood well by high-level density functional theoretic (DFT) calculations.[18] Interaction energy value for the H-bond in Watson-Crick base pairs from some of the above mentioned studies are listed in the Table 1.2:

Table 1.2: Interaction energy values (kcal/mol) for the hydrogen bond in A---T Watson-Crick base pair reported in other theoretical studies.[32,33]

	ΔE	**ΔE (BSSE corrected)**
HF/6-31G**		-10.3
Experiment[a]	-12.1	------
B3LYP/6-31G**	-13.4	-12.3
MP2/DZP//HF/6-31G*		-14.7
MP2/cc-pVTZ		-14.92
RI-MP2/aug-cc-pVTZ		-16.4
CCSD(T)/CBS[b]		-16.37
CCSD(T)/CBS[c]		-16.66

[a]ΔH_{exp} from methyl substituted base pairs with corrections according to Brameld, K; Dasgupta, S; Goddard, W.A.; III. *J. Phys. Chem. B* **1997**, 101, 4851. [b]ΔCCSD(T) term was calculated[32] by extrapolation using the cc-pVDZ basis set. [c]ΔCCSD(T) term was calculated[33] by extrapolation using the aug-cc-pVTZ basis set.

In 1990s, extensive QM calculations were carried out to analyze the nature of the base pair stacking. Stacking energy values for a set of 11 base pairs were reported by Šponer *et al.*.[34]

This study involved the calculation of interaction energy values via geometry optimization at MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory and further extrapolation to get the MP2 complete basis set (CBS) limit. Similarly, CCSD(T)/CBS limit has been determined for 10 DNA base-pair steps at their gas phase minima with respect to the helical parameters rise, twist, and propeller twist.[35] A full PES with respect to the twist for all DNA base pair steps has been generated by using van der Waals density functional theory (vdW-DFT).[36] This study reported an average twist angle of 34±10° at the energetic minimum which is close to that of B-DNA as determined from multiple analyses of X-ray crystal structures deposited in the Nucleic Acids Database (36 ± 7°).[4]

1.4 Importance of the Sugar-Phosphate Backbone:

A common feature of the computational studies described above is the use of truncated models. That is, DNA nucleosides (nucleotides) are typically modeled as nucleobases, where the sugar-phosphate backbone is replaced with a hydrogen atom or a methyl group. As modeling of highly charged sugar phosphate is difficult, the use of such truncated systems enables the determination of π- π interactions. However, they do not take into account the role of deoxyribose sugar and the backbone effect.

The backbone can be considered as a substituent to the nucleobases and substituents have been shown to significantly affect stacking interactions.[37] This is because the London dispersion depends on the formation of instantaneous dipoles, which in turn depend on the polarizabilities of the nucleobases. The extent of polarizability of a π electron cloud depends on the electron withdrawing or donating potential of the substituent. For example, the potential energy surface of the benzene dimer possesses at least two energy minima, corresponding to the T-shaped and parallel displaced geometries. For the toluene dimer, it has been shown that the stacked configuration is preferred over the T-shaped one, and this preference is observed in both gas-phase and in aqueous solution.[38]

The backbone can have a significant control over the base pair stacking geometries as it restricts the conformational space accessible to the bases. Additionally, steric clashes arising from the backbones may prevent some stacked geometries from occurring in natural systems.

An experimental determination of the sugar-phosphate backbone is difficult. While it is possible to determine the position of the nucleobases and the centers of the phosphate groups reliably by X-ray diffraction, even at moderate resolution (~ 2.5 Å), it is much more difficult to determine the conformation of the sugar part.

The importance of accounting for the presence of the backbone has been shown by recent studies on QM description of DNA-protein interactions. These calculations revealed that the DNA backbone can increase or decrease the π-π stacking interaction.[39] Further analysis revealed that direct backbone-π interactions also significantly enhance the stability of the overall complex.[40]

Many theoretical studies have been performed to understand the importance of the backbone in shaping of DNA. A conformational energy analysis[41] of the dodecamer d-(CGCGAATTCGCG) suggested that the helical structure of DNA is mainly determined by the intrinsic interactions between base pairs. During the analysis, the backbone was not considered and yet the minimum energy geometry exhibited a helical conformation and it agreed remarkably well with the crystal structure. Based on these results, the authors suggested that the backbone may play a substantially passive role in determining the basic helical structure of DNA. Force field calculations carried out on base pair steps[22] showed that the experimentally observed values of roll, tilt and rise lie at the minima on the potential energy surface for the base stacking interactions and that they are essentially independent of the backbone.

The passive role of the backbone in controlling the helix formation has also been speculated by Sherrill in a recent work.[42] This study involved an analysis of the variation of energy with respect to slide, while keeping the values of rise and twist fixed at the optimal theoretical value. Values of the helical parameters rise and twist, corresponding to the minimum energy geometry were in accord with the corresponding values from the average B-DNA crystal structure geometry. As the backbone was not included in these calculations, the authors suggested that the backbone either does not significantly influence the rise and twist parameters in B-form DNA or that the backbone had a local (or global) conformational minimum that is an excellent match with the twist angles that provide optimal stacking energies.

However, as mentioned above, these studies did not explicitly consider the sugar-phosphate backbone. The passive role of backbone in controlling the helix formation has been speculated by observing that without considering the backbone, theoretical values of base step parameters are almost the same as the experimental values. An explicit consideration of the backbone might significantly improve our current understanding of the role of backbone in the helix formation in B-DNA.

1.5 Thesis Objectives:

The overall theme of the work reported in this thesis is to understand the effect of the sugar-phosphate backbone on noncovalent interactions present in B-DNA. To the best of our knowledge, no study, either experimental or theoretical, has explicitly considered the effect of sugar-phosphate backbone on stacking and H-bonding in B-DNA, and in the formation of the double helix. We have used DFT based electronic structure calculations for the purpose.

First we considered the effect of sugar-phosphate backbone on H-bonding interactions in A-T base pair. A detailed analysis of this effect was carried out by performing quality DFT calculations on the A-T base pair. It was found that H-bonding between nucleosides as well as nucleotides of adenine and thymine is almost the same as in nucleobases. This showed that the backbone has no or only a nominal effect on H-bonding between A-T base pairs.

Further, the thesis focuses on the examining weak interactions involved between stacked base pairs. Using DFT calculations, stacking interaction energy of the nucleosides and nucleotides of adenine and thymine was compared with the stacking energy of the corresponding nucleobases. Subsequently, using dispersion-corrected DFT methods (DFT-D3), we try to understand the role of sugar-phosphate backbone in the helix formation of the B-DNA.

Chapter 2

Computational Methods

This section will give an overview of the various computational/theoretical concepts used in the following studies. Detailed derivations and discussions of the mathematical aspects of the methods can be found in standard text book.[43]

2.1 The Schrödinger Equation

A chemical system can be described by the geometry and electronic structure of the atoms/molecules in the system. The electronic structure of a chemical system is described by a wave-function (Ψ), which is determined by solving the time-independent non-relativistic Schrödinger equation

$$\hat{H}\,\Psi = E\,\Psi \qquad \text{…….. } \mathbf{(2.1)}$$

where $\hat{H}$ denotes the Hamiltonian operator, Ψ is an eigenfunction of the Hamiltonian, and E is the eigenvalue. The Hamiltonian operator can be partitioned into the following parts:

$$\hat{H} = \check{T}_N(R) + \check{T}_e(r) + \hat{U}_{Ne}(R, r) + \hat{U}_{ee}(r_1, r_2) + \hat{U}_{NN}(R_1, R_2) \qquad \text{…….. } \mathbf{(2.2)}$$

Each term on the right hand side of eq. (2.2) has a physical interpretation: (a) $\check{T}_N(R)$ represnts the total kinetic energy of all the nuclei, (b) $\check{T}_e(r)$ corresponds to the kinetic energy of all the electrons, (c) $\hat{U}_{Ne}(R, r)$ represents the coulombic attraction between the electrons and the nuclei, (d) $\hat{U}_{ee}(r_1, r_2)$ represents the electron-electron repulsion, and (e) $\hat{U}_{NN}(R_1, R_2)$ corresponds to the nuclear-nuclear repulsion energy. Schrödinger equation can be solved exactly only for a handful of systems and a series of approximations are used to find the approximate solution for reasonable sized systems. Born-Oppenheimer approximation is one such approximation, which says that because the nuclei are much more massive than the

electrons and thus move much more slowly, the electrons can be considered to be moving in a field of fixed nuclei. As a result, $\check{T}_N$ becomes zero and $\hat{U}_{NN}$ is treated as a constant. The remaining three terms form the electronic Hamiltonian operator and the eigenvalue corresponds to the electronic energy.

For a many-electrons system, the eigenvalue equation is not exactly solvable and different quantum mechanical methods are used, to find the electronic energy. The primary difference in these methods is the manner in which the electron-electron repulsion $\hat{U}_{ee}$ is handled.

2.2 Hartree-Fock Theory

Hartree-Fock theory (HF) is based on an approximation within which each electron feels only the *average field* of all the other electrons in the system. That is, no explicit electron correlation is considered. The wave-function (Ψ_{HF}) in the HF method is represented as a Slater determinant in which each electron is assigned an orbital:

$$\Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_1(x_1) & \cdots & \chi_1(x_1) \\ \chi_1(x_1) & \chi_1(x_1) & \cdots & \chi_1(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_1) & \chi_1(x_1) & \cdots & \chi_1(x_1) \end{vmatrix}$$

Fock operator, which has the electron-electron repulsion terms, is then used variationally to find the lowest energy set of molecular orbitals (MOs) and of these MOs, occupied molecular orbitals (MOs) are used to write the Slater determinant. The Hartree-Fock energy, then, is the expectation value of the electronic Hamiltonian within this wave-function plus the nuclear potential energy.

The HF method can provide a reliable description of H-bonding, but fails completely when it comes to describe molecules having dispersion interactions. This is because it neglects electron correlation, which is important for dispersion and other weak interactions. To "recover" this correlation energy that the HF method overlooks, other methods, such as the nth order Møller-Plesset perturbation theory [MPn], coupled cluster theory (CCn), symmetry

adapted perturbation theory (SAPT), etc., have been developed, which refine the results from the HF method. These methods have been used extensively for the study of noncovalent interactions. However, each of them has certain limitations. For example, the second order Møller-Plesset perturbation theory (MP2) has been shown to strongly overestimate stacking interactions.[44] Though recently MP2 based methods such as local MP2 (LMP2),[45] spin component scaled MP2 (SCS-MP2),[46] density fitted MP2 (DF-MP2),[47] have been developed, they are of limited use. The golden standard of quantum chemical calculations is the singles and doubles coupled cluster theory with perturbative triples excitations (CCSD(T)). It provides standard data for stacking interactions, but it is prohibitively expensive for all but the smallest prototypes.

2.3 Density Functional Theory

Above mentioned *ab initio* methods involve determining the wave-function to obtain the desired information about the system. However, the wave-function contains a great deal of redundant information and generally lacks a physical meaning. Energy and other properties of a system can also be determined from the electron density and this observation makes the basis for the density functional theory (DFT). The electron density is physically observable and is a function of only three variables (x, y, z) regardless of the number of particles in the system. DFT methods have been shown to have an accuracy at par with many *ab initio* methods, but at a significantly lower computational cost.

The electron density is represented by a Slater determinant involving orbitals. Though this electron density corresponds to a system of noninteracting electrons, it represents the true electron density of the real system. Calculations of the orbitals involve solving Kohn-Sham equations, which include the noninteracting kinetic energy, the interaction of the electrons with the nuclei, the Coulomb interaction of the electron density with itself, and a potential arising from a functional derivative of the so called exchange–correlation energy functional. This energy functional accounts for the self-interaction correction, electron exchange, and electron correlation.

The exchange–correlation functional can only be approximated and DFT methods are classified based on the way they approximate this functional. The simplest such classification

involves two classes: local and nonlocal. Local functionals are further divided into local spin density approximation (LSDA), which depends on the local spin density, generalized gradient approximation (GGA) that also includes the gradient of these densities, and meta-GGA that also depends on local spin kinetic energy density. Nonlocality is introduced in these funtionals via the addition of Hartree-Fock exchange, involving occupied orbitals in a functional form. This addition results in nonlocal funtionals classified as hybrid GGAs and hybrid meta-GGAs.

These traditional density functional have shown remarkable accuracy in describing H-bond and some of them provide better results than the MP2. However, they fail to describe properly the dispersion interaction, stacking interaction in particular. Though the theory in itself is capable of providing a proper description, it is the approximations made in the DFT functionals which enforce this limitation. For example, B3LYP, a hybrid GGA very is an extensively used DFT functional, which determines the exchange-correlation energy based on the local spin density. Dispersion interactions require a nonlocal description and hence are not properly described by B3LYP.

Recently some new functionals, such as the M06 suite,[48] and DFT-D,[49] have been developed and they show remarkable success in describing dispersion interactions.

2.3.1 M06-2X

It is a hybrid meta-GGA functional developed by Zhao and Truhlar primarily to study noncovalent interactions.[48] Various types of reference data were used to find the optimum value of H-F exchange incorporated in this hybrid functional. The M06-2X functional involves simultaneous optimization of exchange and correlation functionals. This optimization accounts for medium-range correlation energy and as medium range includes the typical distance of interacting molecules in van der Waals systems, this functional performs excellently for noncovalent interaction studies. Gu et al.[50] employed the M06-2X functional to predict the structures and stacking interactions in uracil dimers and thymine dimers. Their study demonstrated that the M06-2X functional was able to predict stacked structures for the uracil and thymine dimers that were in better agreement than MP2 results when compared with the geometry solved at the CCSD(T)/CBS level of theory. However, in

the M06-2X, the long-range correlation is not treated in a correct manner and therefore, this functional is of limited to use for systems around equilibrium geometry and for H-bonding description.

2.3.2 DFT-D

Dispersion corrected density functional theory (DFT-D) is an improvement to other density functionals such as B3LYP, BLYP, PW91 etc.[49] DFT-D accounts for the noncovalent interactions by adding an empirical dispersion correction to the DFT results from traditional functionals which were developed without any regard for noncovalent interactions. This empirical correction is carried out by asymptotic formulas valid for long-range interactions. DFT provides reliable description of short range interactions. A damping function is introduced to control the amount of dispersion correction to DFT energy and it is determined mainly by the distance at which damping takes effect. At the equilibrium geometries, damping takes roughly around 15% of the dispersion contribution in the case of dispersion-bonded complexes while it is about 90% for the H-bonded complexes. This damping function is usually chosen by trial and error and it controls the accuracy of the DFT-D method.

Chapter 3

Effect of the Sugar-Phosphate Backbone on Non-covalent Interactions in B-DNA

3.1 Effect of Backbone on Watson-Crick A-T Base Pairing

This section describes the effect of the sugar-phosphate backbone on Watson-Crick hydrogen bonding between adenine and thymine nucleobases.

3.1.1 Methodology

Gaussian 09 suite of programs was used for the electronic structure calculations.[51] Geometries of adenine, thymine and the hydrogen bonded A-T base pair were optimized at the DFT/M06-2X level using Dunning's correlation consistent cc-pVTZ basis set.[52] Vibrational frequencies were also calculated for all the optimized geometries at the same level of theory to verify the energy minimum character of the optimized geometries. The effect of the backbone on base pairing in A-T is studied stepwise by going from Watson-Crick base pairing of adenine and thymine nucleic bases to their nucleosides and then the nucleotides. The charge distribution in isolated as well as paired bases was determined by performing the natural bond orbital analysis (NBO), which provides natural population atomic (NPA) charges in each system.[53]、

The interaction energies (ΔE_{int}) were computed using the supermolecule approach:

$$\Delta E_{int} = E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B) \quad \text{.....} \ \textbf{(3.1)}$$

where the energy of the molecule *M* in geometry *G* computed with basis set σ is denoted as $E_G^{\sigma}(\mathrm{M})$. Interaction energies were corrected for the basis set superposition error (BSSE), using the counterpoise correction (CP) method proposed by Boys and Bernardi.[54] The BSSE correction is given by

$$\text{BSSE} = E_{AB}^{A}(A) - E_{AB}^{AB}(A) + E_{AB}^{B}(B) - E_{AB}^{AB}(B). \quad \text{.....} \textbf{(3.2)}$$

Hence the BSSE-corrected interaction energy is calculated as

$$\Delta E_{int}^{CP} = \Delta E_{\text{int}} + \text{BSSE} \quad \text{.....} \textbf{(3.3)}$$

3.1.2 Results and Discussion:

3.1.2.1 Watson-Crick base pair of Adenine and Thymine (A-T):

Geometry and energetics of the A-T base pair

Optimized geometries of adenine, thymine, and Watson-Crick base pair obtained from DFT/M06-2X/cc-pVTZ calculations are shown in Figure 3.1. No imaginary frequencies were obtained in the vibrational frequency calculations, confirming that these geometries correspond to the true minima. The results of our studies are summarized and compared with the results from the literature in Tables 3.1 (geometries), and 3.2 (interaction energy values). Figure 1.8 defines the labels used throughout the studies.

The computed (DFT/M06-2X/cc-pVTZ) interaction energy value of, -14.89 kcal/mol for the hydrogen-bond in A-T agrees well with the reported value of -14.92 kcal/mol obtained from a MP2/cc-pVTZ calculation. It is worth adding that the DFT calculations are computationally less demanding and hence can be a good choice for studying hydrogen bond interaction in A-T base pair.

Table 3.1: Hydron bond parameters (bond distances in Å and bond angles in degrees) in A-T Watson-Crick base pair.

Level of theory	**H6 ---O4, R_1**	**N1 ---H3, R_2**	**<N6H6O4**	**<N1H3N3**
M06-2X/cc-pVTZ[a]	1.93	1.78	172.7	179.9
X-Ray data[b,55]	1.94	1.79	173.8	178.9
MP2/cc-pVTZ[56]	1.93	1.82	173.6	179.1
B3LYP/6-31G**[57]	1.92	1.80	173.1	178.9

[a]this work, [b]for methylated A-T bases,

Table 3.2: Values of the Hydron bond interaction (kcal/mol) obtained by different theoretical methods for the A-T Watson-Crick base pair.

Level of Theory	**ΔE_{int}**	**ΔE_{int} (BSSE-corrected)**
M06-2X/cc-pVTZ[a]	-15.19	-14.89
Experiment[26]	-12.1	------
MP2/cc-pVTZ[31]		-14.92
CCSD(T)/CBS[32]		-16.37
CCSD(T)/CBS[33]		-16.66

[a] this work.

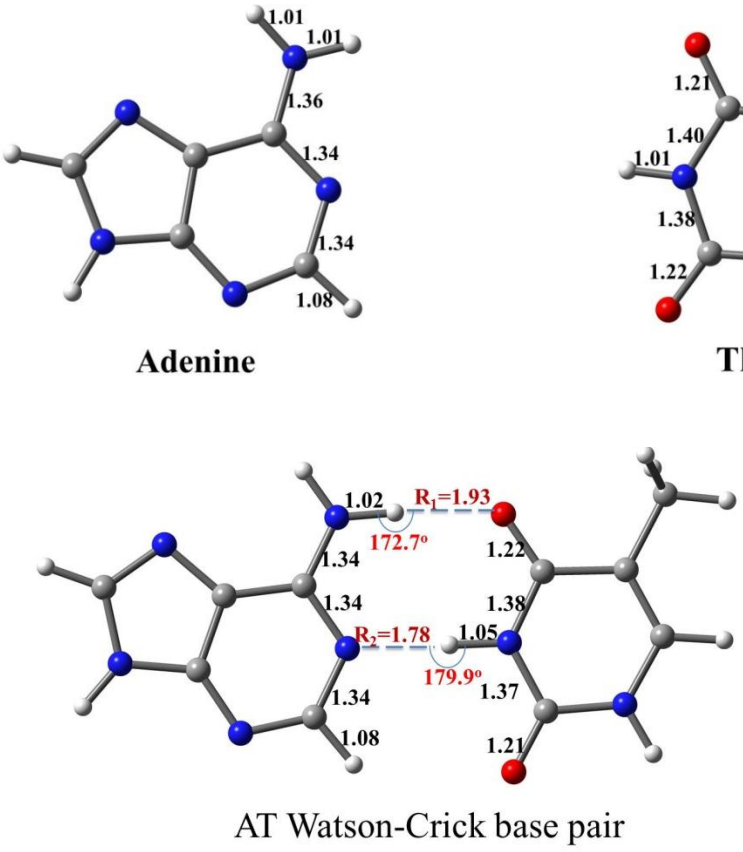

Figure 3.1: DFT/M06-2X/cc-pVTZ optimized geometries of adenine, thymine, and A-T Watson-Crick base pair. Selected bond distances (Å) and bond angles (degrees) are shown in the picture. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Bond angles between the atoms involved in the H-bonds are close to 180^{o} (i.e. 172.7 ^{o} and 179.9^{o}), which make them essentially linear in nature. The basis set superposition error of ~0.3 kcal/mol is quite small, which implies that the choice of level of theory and the basis set is good for the study. Hydrogen bond distances, R_1 and R_2, obtained from DFT/M06-2X/cc-pVTZ calculations agree well with the experiments and high level electronic structure calculations (Table 3.1). All the N-H bonds that participate in hydrogen bonding elongate by 0.01-0.04 Å, with the largest elongation (0.04 Å) observed for N3-H3 of thymine monomer (Figure 3.1).

Frontier Molecular Orbital (FMO) Analysis:

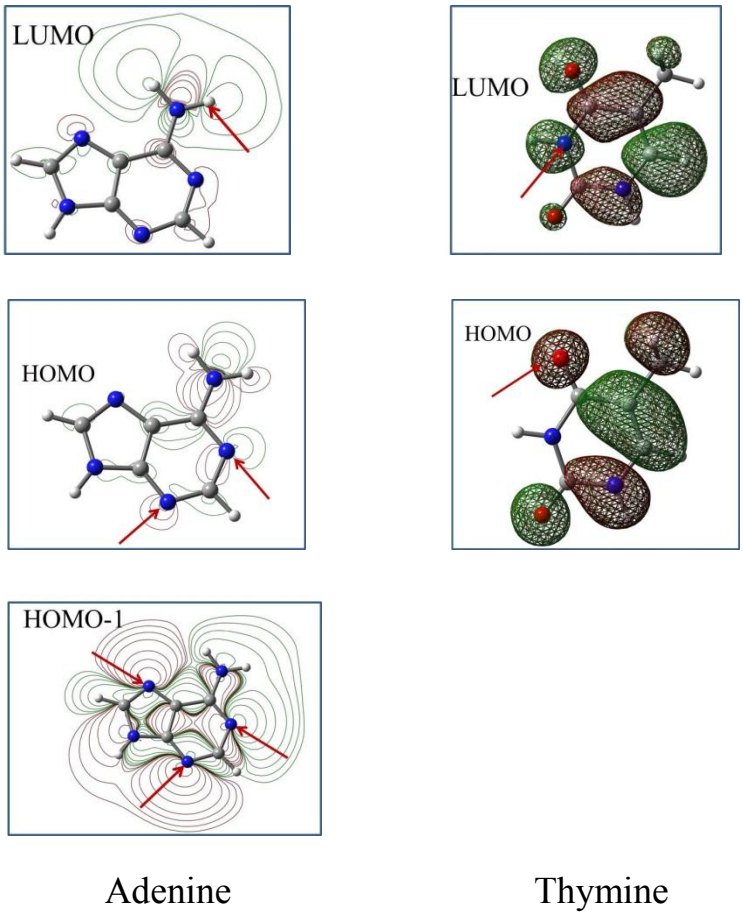

Figure 3.2: Contour plots of HOMO-1, HOMO, and LUMO of adenine (isovalue = 0.2) and HOMO, and LUMO of thymine (isovalue = 0.2) obtained at the DFT/M06-2X/cc-pVTZ level of theory. Red contours represent the positive values while green denotes the negative values. Molecular orbitals are arranged according to increasing energies.

Figure 3.2 depicts the contour plots of the frontier orbitals of the DNA bases and they turned out to be favorable for hydrogen bonding. The highest occupied molecular orbital (HOMO) and HOMO-1 of adenine have lone pair like lobes on the nitrogen atoms, N1, N3, and N7 (pointed by arrows in Figure 3.2), whereas the lowest unoccupied molecular orbital (LUMO) of thymine primarily has N3-H3 antibonding character. Through their lobes, adenine can overlap with and donate charge into the LUMO of thymine and form a hydrogen bond. Likewise, HOMO of thymine is primarily the lone pair on O4, while LUMO of adenine has N6-H6 antibonding character facilitating the formation of a hydrogen bond via overlapping and charge transferring. It is this donation of charge into the N-H antibonding orbitals of adenine and thymine, which is responsible for a slight elongation of the N-H bond involved in hydrogen bond formation.

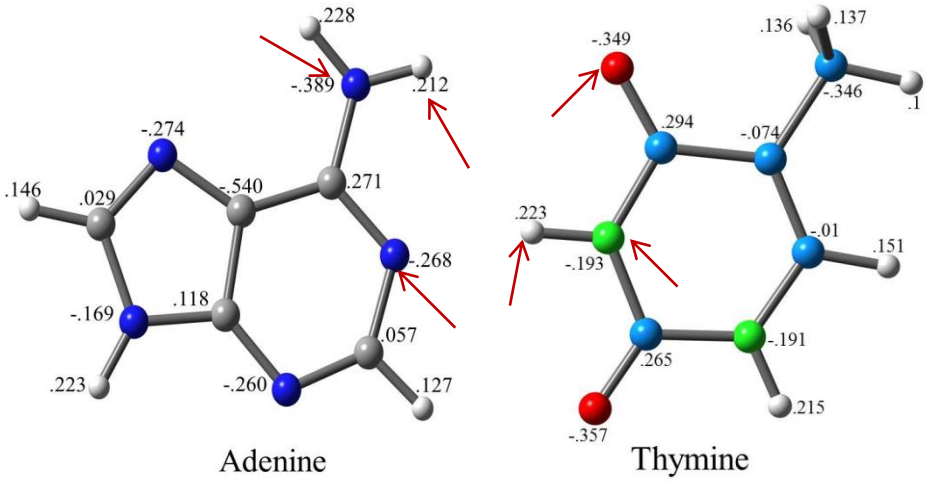

Figure 3.3: NPA charges (electrons) of adenine and thymine bases obtained at the DFT/M06-2X/cc-pVTZ level of theory. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Charge Analysis:

The formation of a hydrogen bond requires the right charge distribution in the bases. This turns out to be the case, as can be seen from Figure 3.3, which displays the NPA charges in adenine and thymine bases. All proton-acceptor atoms (N1, O4) have a negative charge, while the corresponding protons (H3, H6) are positively charged (highlighted by arrows in Figure 3.3). While adenine and thymine individually are neutral, in the base pair geometry

thymine has a total charge of -0.123, which clearly indicates the charge transfer from adenine to thymine during hydrogen bond formation. Though the charge transfer takes place from thymine to adenine as well but charge transfer is more in the opposite direction.

3.1.2.2 Watson-Crick base pair of deoxyadenosine and deoxythymidine (Ad-Th):

Geometry and energetics of the Ad-Th base pair

To study the effect of deoxyribose sugar group on hydrogen bonding in A-T base pair, we considered the nucleoside base pair: deoxyadenosine(Ad)-deoxythymidine (Th). Optimized geometries of the individual nucleosides and that of the base pair are shown in Figure 3.4.

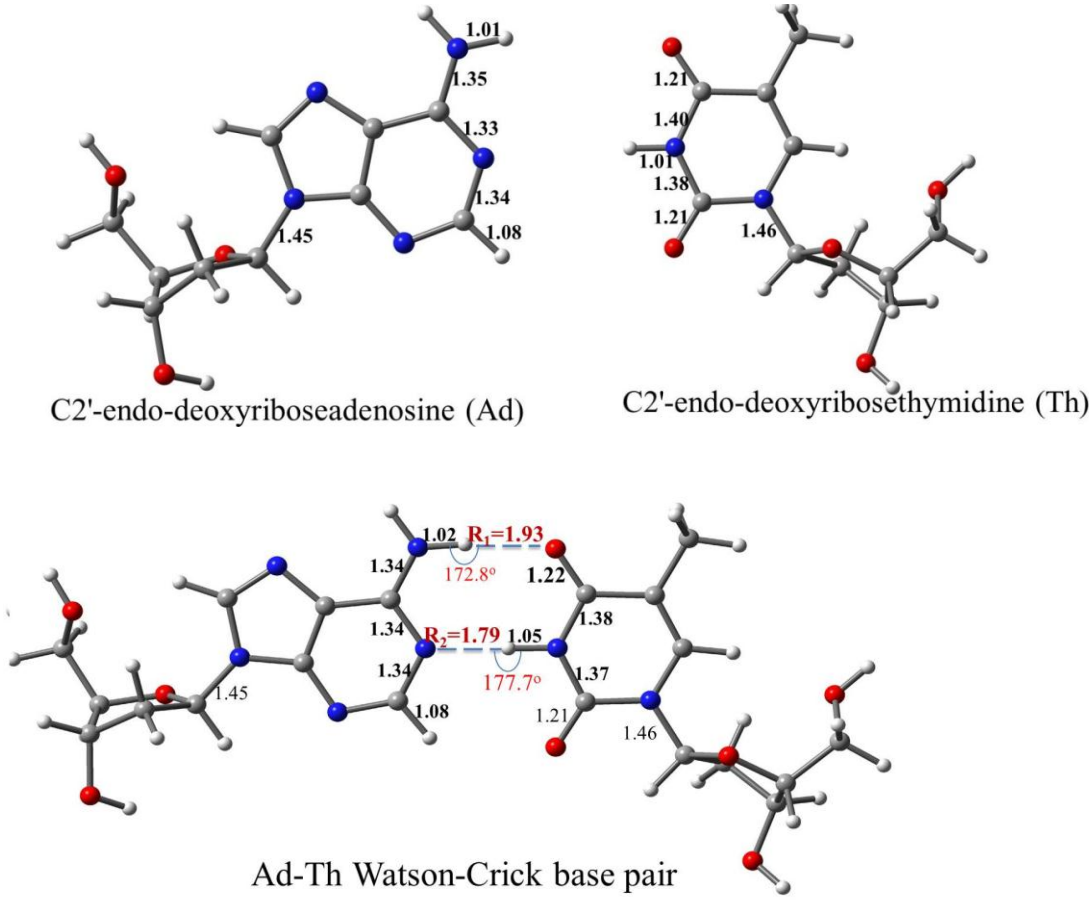

Figure 3.4: Optimized geometries of adenosine (Ad), thymidine (Th), and Ad-Th Watson-Crick base pair at the DFT/M06-2X/cc-pVTZ level of theory. Selected bond distances (Å) and bond angles (degrees) are shown in the picture. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

While the H6-O4 hydrogen bond distance remains constant in going from the A-T base pair to Ad-Th base pair, the N1-H3 hydrogen bond elongates nominally from 1.78 Å to 1.79 Å. The Ad-Th base pair remains linear as in the case of the A-T base pair. A comparison of the geometry of the nucleic bases with that of the corresponding nucleosides shows that deformation in the structure of the bases is negligible (≤ 0.01 Å) on addition of the sugar moiety.

At the DFT/M06-2X/cc-pVTZ level of theory, the interaction energy of Ad-Th base pair is found to be -14.42 kcal/mol. This means that the hydrogen bonding energy decreases by only 0.47 kcal/mol on going from the Watson-Crick base pair of nucleic bases to that of the nucleosides.

Charge Analysis:

An NPA charge distribution for Ad and Th are shown in Figure 3.5. Similar to adenine and thymine, adenosine and thymidine have charge distributions favorable for hydrogen bond formation (highlighted by arrows in Figure 3.5).

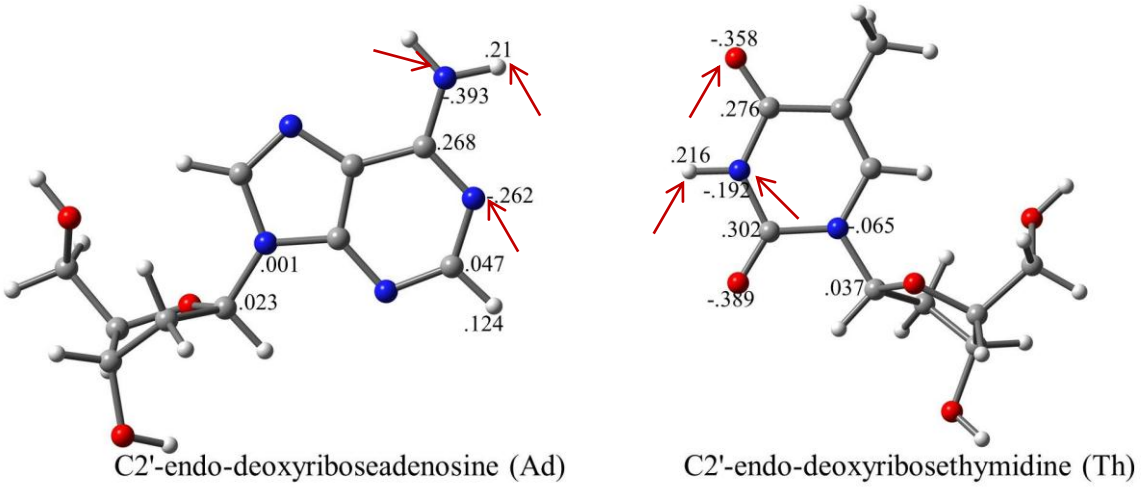

Figure 3.5: Electron density values around selected set of atoms of adenosine (Ad), and thymidine (Th) obtained at the DFT/M06-2X/cc-pVTZ level of theory. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Comparison of charge distribution in nucleosides and nucleic bases (Figure 3.3) indicates that addition of sugar does not influence the charge on atoms involved in hydrogen bond formation by a significant amount. As the hydrogen bond parameters (bond lengths and bond angles), charge distribution and interaction energy values for Ad-Th base pair differ only slightly from those of A-T, we can conclude that the sugar molecule has only a marginal influence on Watson-Crick hydrogen bonding in A-T base pair.

3.1.2.3 Watson-Crick base pair of deoxyadenosinemonophosphate and deoxythymine-monophosphate (dAMP-dTMP):

Geometry and energetics of dAMP-dTMP base pair

To study the effect of sugar-phosphate moiety on hydrogen bonding in Watson-Crick base pair of adenine and thymine, the base pair of 2'-deoxyadenosine 5'-monophosphate (dAMP) and 2'-deoxythymidine 5'-monophosphate (dTMP) was considered. Optimized geometries of the individual nucleotides of adenine and thymine and that of the base pair are shown in Figure 3.6. The H6-O4 and N1-H3 hydrogen bond distances in dAMP-dTMP base pair are 1.94 Å and 1.79 Å, respectively and the corresponding values are 1.93 Å and 1.78 Å in the case of A-T and 1.93 Å and 1.79 Å for Ad-Th Watson-Crick base pair. The dAMP-dTMP base pair remains linear as A-T and Ad-Th base pairs do. The addition of the sugar-phosphate group causes only slight change ($\leq$ 0.02 Å) in the geometry of nucleobases.

At the DFT/M06-2X/cc-pVTZ level of theory, the interaction energy of dAMP-dTMP base pair is found to be -14.38 kcal/mol. That is, the hydrogen bonding energy decreases by only 0.51 kcal/mol in going from the Watson-Crick base pair of nucleic bases to that of the nucleotides. The interaction energy of the nucleotide base pair of A-T differs by only 0.04 kcal/mol from that of the nucleoside base pair and therefore, we can conclude that the sugar and phosphate groups have only a marginal influence on the hydrogen bonds in adenine and thymine.

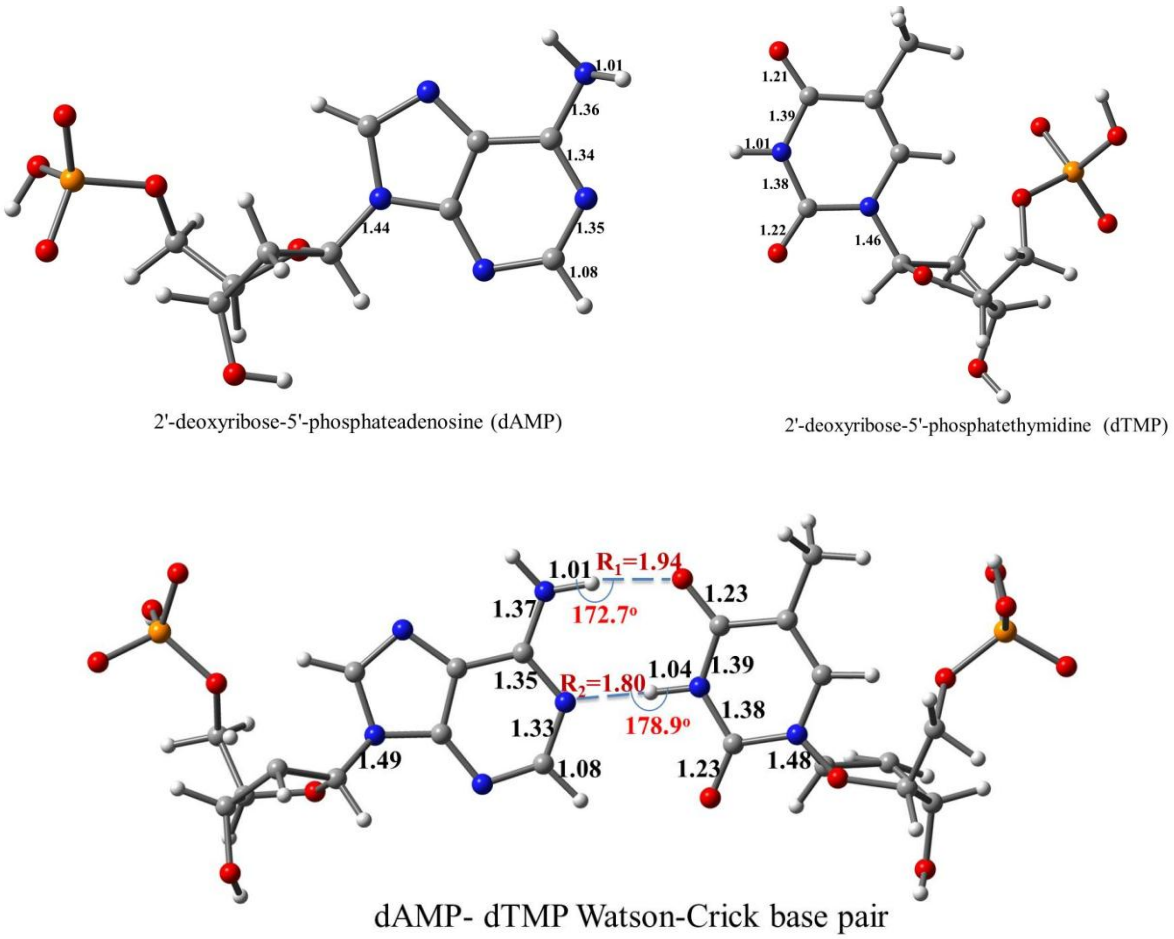

Figure 3.6: Optimized geometries of 2'-deoxyribose 5'- phosphateadenosine (dAMP), 2'-deoxyribose-5'-phosphatethymidine (dTMP), and dAMP-dTMP Watson-Crick base pair at the DFT/M06-2X/cc-pVTZ level of theory. Selected bond distances (Å) and bond angles (degrees) are shown in the picture. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

3.2 The Effect of Backbone on Stacking Interactions Between Adenine-Thymine Base Pairs

This section describes the results of electronic structure calculations performed to study the effect of the sugar-phosphate backbone on stacking interaction in base pair steps.

3.2.1 Methodology:

All the electronic structure calculations were carried out by using Gaussian 09 suite of programs[51] at the DFT/M06-2X level using the polarized double-ζ basis set 6-31G*. The systems involved in these calculations is larger compared to the base pairing studies and available computational resources necessitates the use of a different and smaller level of theory. Geometries for A, T and A-T were optimized. The geometry of the dimer of $(A\text{-}T)_2$ was constructed using the parameters obtained from the optimized geometry of the base pair. The vertical separation between two stacked adenine bases corresponding to the minimum was determined. Keeping this vertical separation fixed, optimal displacement of one adenine compared to the other along X- and Y- axes were found. A set of displacement values along X- , Y- and Z-axes were found for thymine dimer as well. Different geometries for $(A\text{-}T)_2$ were then constructed by a mutual displacement of the monomeric base pair along X-, Y-, and Z-axes. These displacement values were the same as those obtained for adenine and thymine dimers. Only single point energies were computed for these geometries.

The same values for the displacements were used to construct A-T nucleoside dimers and single point energy calculations were carried for these dimer geometries also. The effect of the sugar group on the stacking interaction was investigated by comparing the BSSE-corrected interaction energy values for the nucleoside dimer in a particular conformation with the interaction energy values of base pair dimer having the corresponding conformation. This procedure was extended to nucleotides and the effect of backbone on stacking interaction was examined.

The stabilization energy values for base pair dimers in various conformations were obtained by using the supermolecule approach:

$$\Delta E_{\text{dim}} = E \text{ (base pair dimer)} - 2\,E \text{ (base pair monomer)}. \quad \ldots\ldots \textbf{(3.4)}$$

The results were corrected for basis set superposition error (BSSE) using the counterpoise correction method proposed by Boys and Bernardi.[54] The corrected interaction energy values were calculated using the equation:

$$\Delta E_{\text{dim}}\,(\text{corrected}) = E^{\text{AB}}\,(\text{AA}) - 2\,E^{\text{AB}}\,(\text{A}) \quad \ldots\ldots \textbf{(3.5)}$$

where E^{AB} (*AA*) is the energy of the dimer and E^{AB} (*A*) is the energy of the monomer in the dimer geometry using the dimer basis set.

3.2.2 Results and Discussion:

The co-ordinate system and the different displacement notations are shown in Figure 3.7. R_X, R_Y and Rz correspond to displacement along the X, Y and Z-axis respectively.

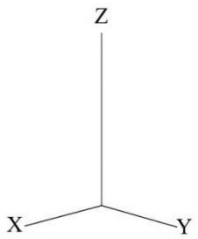

Figure 3.7: Directions for displacement between base pairs.

Stacked dimers of adenine and thymine:

Optimized geometries of stacked dimers of A and T along different axes are shown in Figure (3.8a). the vertical separation corresponding to an energy minimum was found to be 3.6 Å. The energy minimum was verified by vibrational frequency calculations, which showed no imaginary frequency values for this geometry. Keeping two adenine bases at this vertical separation, the geometry representing an energy minimum along the Y-axis has the bases separated by 1.6 Å. Likewise, keeping Rz fixed at 3.6 Å, the bases were found to be separated by 1.4 Å along the X-axis.

In a similar manner, values of displacements in geometries (Figure 3.8b) corresponding to the minima along X-, Y- and Z-axes were determined for the stacked thymine dimer and these values were the same as for the stacked adenine dimer, *i.e.* Rz = 3.6 Å, Ry =1.6 Å, and Rx = 1.4 Å.

BSSE-corrected interaction energy values for stacked adenine and thymine dimers corresponding to the energy minima along different axes are reported in Table 3.3.

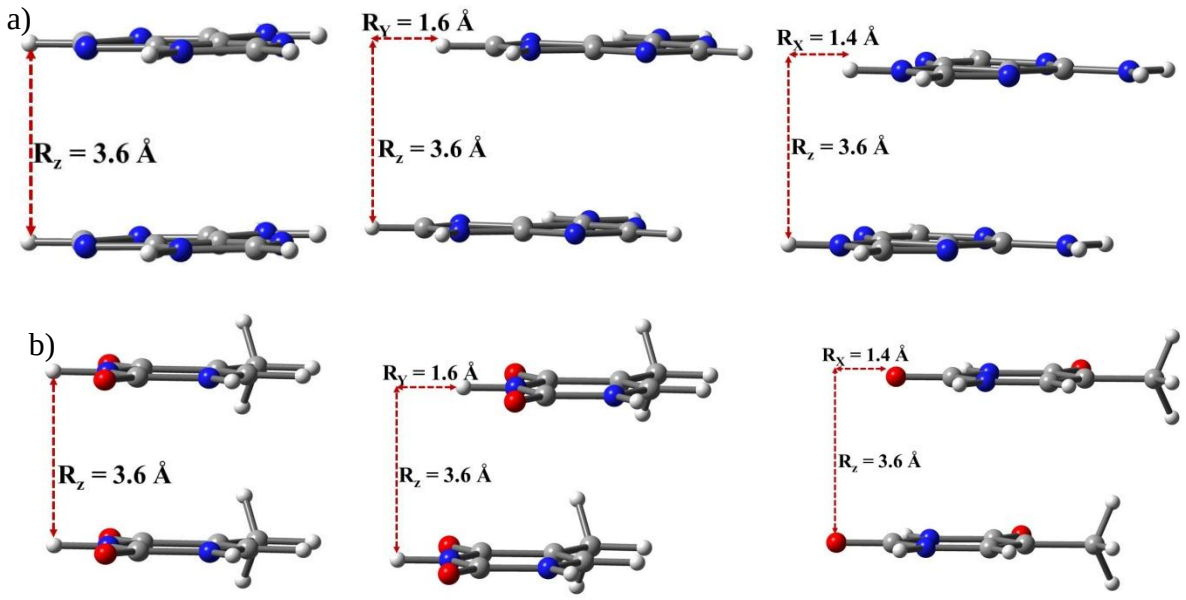

Figure 3.8: Geometries representing minima on the potential energy curve along X-, Y- and Z-axes for a) stacked adenine dimer, b) stacked thymine dimer. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Table 3.3: BSSE-corrected interaction energy values (kcal/mol) at the M06-2X/cc-pVTZ level of theory for stacked adenine and thymine dimers for different values of separation along X-, Y- and Z-axes.

	Stacked adenine dimer	**Stacked thymine dimer**
R_z = 3.6 Å	-7.94	-4.74
R_z = 3.6 Å, R_y = 1.6 Å	-8.68	-5.47
R_z = 3.6 Å, R_x = 1.4 Å	-8.13	-5.12

Interaction energy values for optimized stacked adenine dimer have been reported to be -7.74 at the MP2/aug-cc-pVTZ level of theory and -8.59 at the CCSD(T)/CBS level of theory.[35] A recent study involved the calculation of the stacking interaction energy for the adenine dimer using DFT with LDA approximation for exchange-correlation functional.[58] Without the dispersion interaction the stacking interaction energy was reported to be -5.14 kcal/mol, and accounting for the dispersion lead to a value of -7.25 kcal/mol for the interaction energy. Our values are close to the reported values and the M06-2X/cc-pVTZ level of calculations prove

to be a good choice for including dispersion interaction. Our values of interaction energy values for stacked thymine dimer are also close to other reported values such as -4.85 kcal/mol at the MP2/aug-cc-pVTZ level and -5.14 at the CCSD(T)/CBS level of theory.[35]

The results clearly indicate that both adenine and thymine dimer structures tend to maximize bond to face interactions and avoid atom-atom repulsion similar to parallel-displaced benzene dimer or graphite sheets. The differences between our calculated interaction energy values and other reported values can be due to different optimization approaches and differences in the also proper consideration of twist angle, which was considered in our study.

Stacked dimer of Watson-Crick A-T base pair:

Three different conformations for the stacked A-T base pair dimers were constructed using the parameters obtained for the stacked adenine and thymine dimers, *i.e.*, 3.6 Å for the vertical separation, 1.6 Å for separation in the Y-direction, and 1.4 Å for mutual displacement of bases along the X-axis. These geometries are shown in Figure 3.9.

Interaction energy values for the three geometries are reported in the second column of Table 3.4. Previously reported stacking interaction energy values for the adenine-thymine base pair step (denoted as AA...TT) differ significantly from the calculated values. For example, MP2/6-31G*(0.25) level calculations reported -12.0 kcal/mol for AA…TT step of B-DNA[35] while CCSD(T)/CBS calculations reported it to be -13.1 kcal/mol.[35] However, these calculations differ from our calculations as these were calculated at the equilibrium geometries with respect to the rise and the twist. A considearation of propeller twist during optimization along with the rise and the twist provided an additional stability of 1.5 kcal/mol.

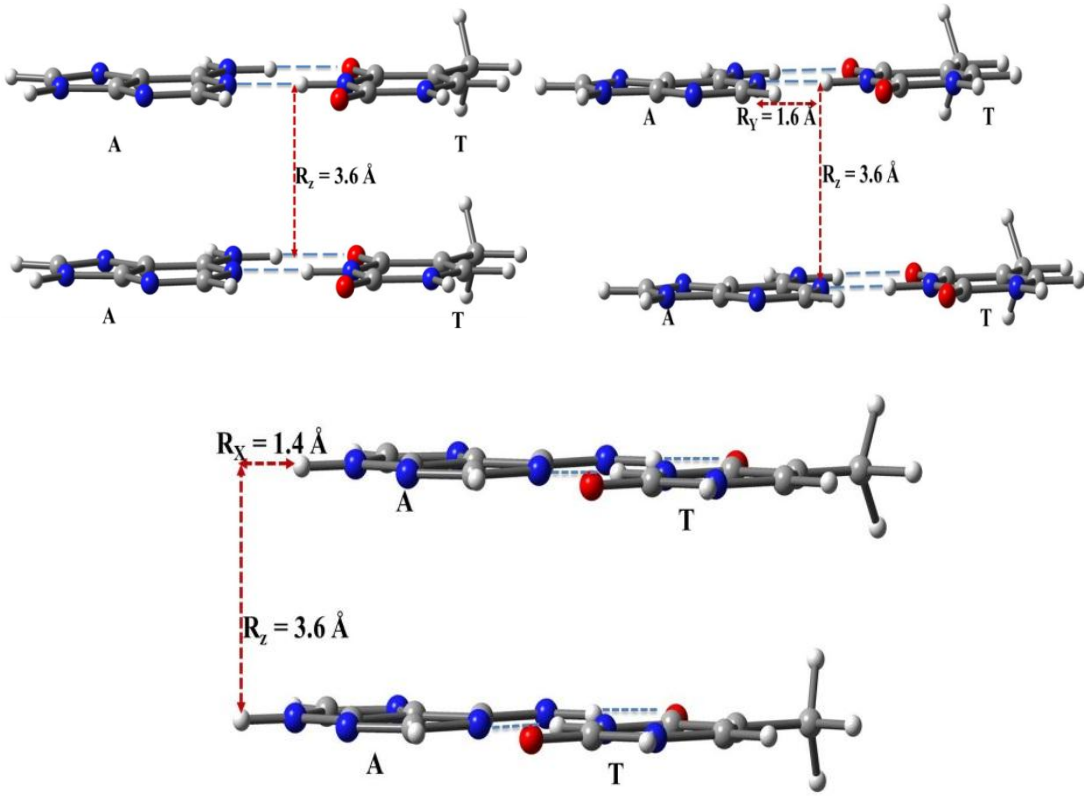

Figure 3.9: The geometries representing three different conformations obtained by mutual displacement of the bases along X-, Y- and Z-axes of the stacked dimer of Watson-Crick A-T base pair. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Table 3.4: BSSE-corrected interaction energy values (kcal/mol) at the M06-2X/cc-pVTZ level of theory for stacked Watson-Crick base pairs of A-T, Ad-Th, and dAMP-dTMP for different values of separation along X-, Y- and Z-axes.

	Stacked A-T	**Stacked Ad-Th**	**Stacked dAMP-dTMP**
R_z = 3.6 Å	-9.92	-10.44	-10.74
R_z = 3.6 Å, R_y = 1.6 Å	-14.34	-14.64	-14.96
R_z = 3.6 Å, R_x = 1.4 Å	-12.83	-13.17	-13.65

Stacked dimer of Watson-Crick Ad-Th base pair, dAMP-dTMP base pair:

To investigate the effect of the sugar-phosphate backbone on stacking interaction between base pair steps, dimers of Ad-Th and dAMP-dTMP base pairs were considered. Similar to the A-T base pair dimer, three dimers were considered for both nucleoside and nucleotide base pair step by mutual displacements of the monomers along X-, Y- and Z-axes. Values of these displacements were the same as for the AA…TT step. Geometries considered for Ad-Th base pair step and dAMP-dTMP base pair step are shown in Figure 3.10 and Figure 3.11, respectively and the interaction energy values obtained for each geometry are summarized in the third and fourth columns of Table 3.4, respectively.

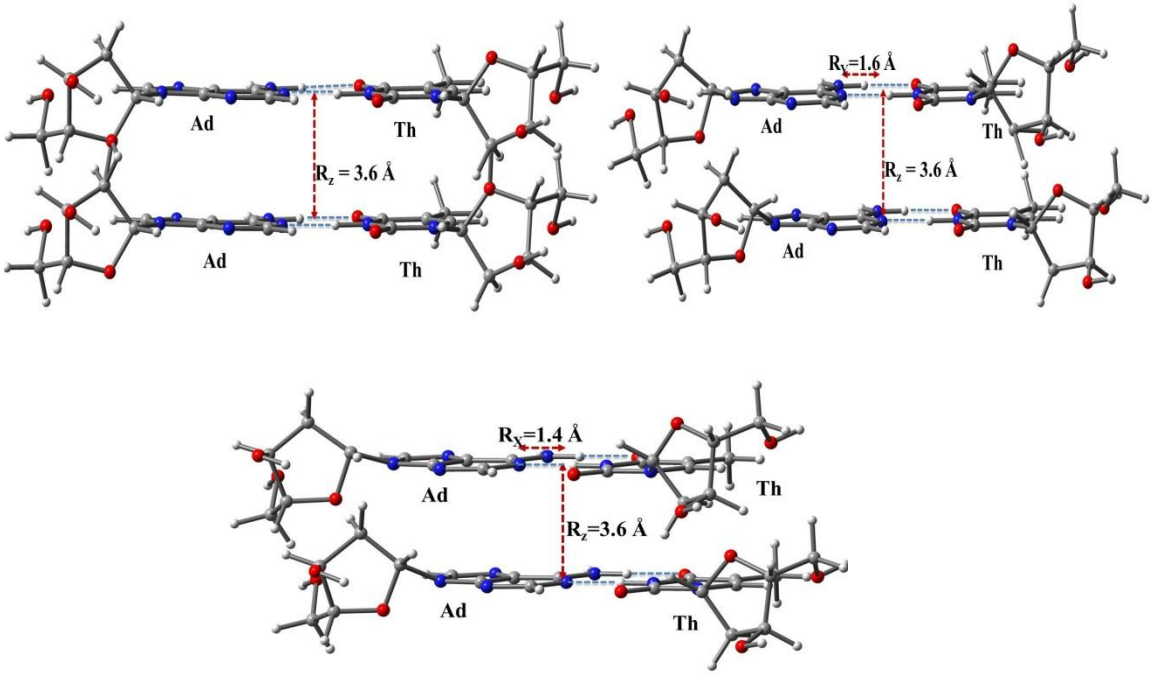

Figure 3.10: The geometries representing three conformations obtained by mutual displacement of the bases along X-, Y- and Z-axes of the stacked dimer of Watson-Crick Ad-Th base pair. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

A comparision of the stacking interaction energy values for the Ad-Th base pair step and A-T base pair step shows that for a particular conformation, the addition of sugar moiety provides a marginal stability to the system. Depending upon the conformation, this additional stability

by the sugar group ranges from 0.30 kcal/mol to 0.52 kcal/mol. However, this difference is not most significant and one can conculde that for a particular conformation, the addition of the sugar molecule does not have a significant influence on the base pair stacking interactions.

Likewise, the addition of the sugar-phosphate group provides an additional stability, ranging from 0.62 kcal/mol to 0.82 kcal/mol, to the A-T base pair step which is not much significant. A comparision of the sugar moiety, obtained by comparing the A-T base pair step with the Ad-Th base pair step, revelas that the phosphate group alone imparts a stability of approximately 0.30 kcal/mol.

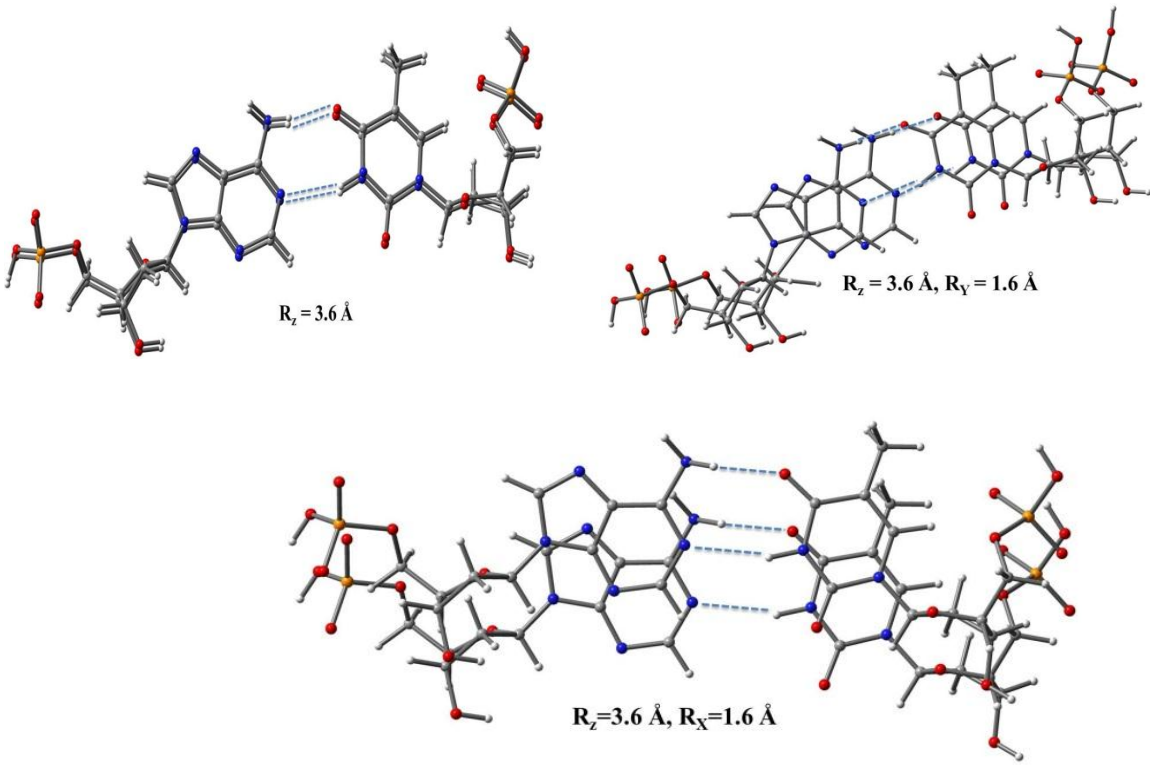

Figure 3.11: The geometries representing three conformations obtained by mutual displacement of bases along X-, Y- and Z-axes of the stacked dimer of Watson-Crick dAMP-dTMP base pair. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Therefore, we can conclude that, for a fixed point on the potential energy curve, stacking interaction energies are essentially same for the A-T base step, the Ad-Th base step and the dAMP-dTMp base pair step.

3.3 Effect of the Backbone on Helix Formation in B-DNA

This section of the thesis describes the electronic structure calculations performed to understand the role of sugar-phosphate backbone on the helix formation in B-DNA.

3.3.1 Methodology:

All the electronic structure calculations were carried out using TurboMole 6.3.1 program suites.[59] Geometry optimization for all the geometries was carried out at the BLYP-D3 level of theory using the def2-SVP[60] and def2-TZVP[61] basis sets. Resolution of identity (RI) approximation is used for evaluating the two-electron integrals, which fastens the calculations without loss of accuracy. The def2-SVP basis set represents the split-valence polarization (SVP) functions. That is the inner shell atomic orbitals (AOs) are described by a single basis function, and the two basis functions are provided for each valence shell AO, augmented by a set of polarization functions. The triple-zeta valence plus polarization basis sets def2-TZVP represents a basis set with one set of d-functions added on the heavy atoms and one set of p-functions added on the hydrogen atoms. Initially the geometries were optimized using the def2-SVP basis set. Subsequently, geometry optimization was performed with the higher basis set (def2-TZVP) to obtain quantitative results. BSSE corrected stacking interaction energy values for base pair dimers were calculated using the equation (3.5).

3.3.2 Results and Discussions:

5’-(dApdA)-3’//5’-(dTpdT)-3’ step of B-DNA was considered to investigate the role of sugar-phosphate backbone in influencing the stacking interaction between the base pairs and also in the helical shape of B-DNA formation. High resolution X-ray crystallographic data of an oligonucleotide (PDB code 1ENN) was used to obtain the geometry of this base step[62]. However, the direct use of experimental geometry can introduce a bias in the energy calculations. A small error in the experimental interbase distances may cause a major bias in the energy calculations. This bias can lead to a false interpretation of the results. Therefore,

the base-pair step was fully optimized using the BLYP-D3/def2-TZVP method and the optimized geometry is shown in Figure 3.12. Twist and displacement in this geometry incorporates the effect of the backbone on the stacked base pairs.

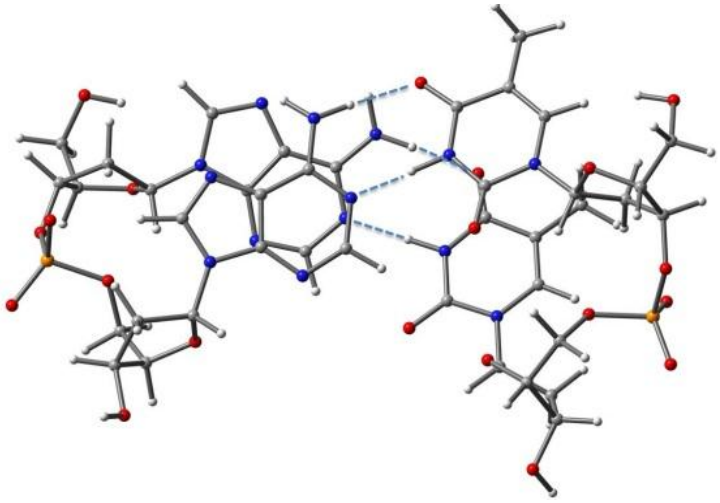

Figure 3.12: Structure of 5'-(dApdA)-3'//5'-(dTpdT)-3' step of B-DNA optimized at BLYP-D3/def2-TZVP level of theory. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

To determine the stacking interaction energy, sugar-phosphate backbone atoms were removed (Figure 3.13a). Single-point energy calculation was carried out using the BLYP-D3/def2-TZVP method and the stacking interaction energy was found to be -19.39 kcal/mol. Full geometry optimization results are shown in Figure 3.13b. The interaction energy was found to be -20.12 kcal/mol using the BLYP-D3/def2-TZVP method.

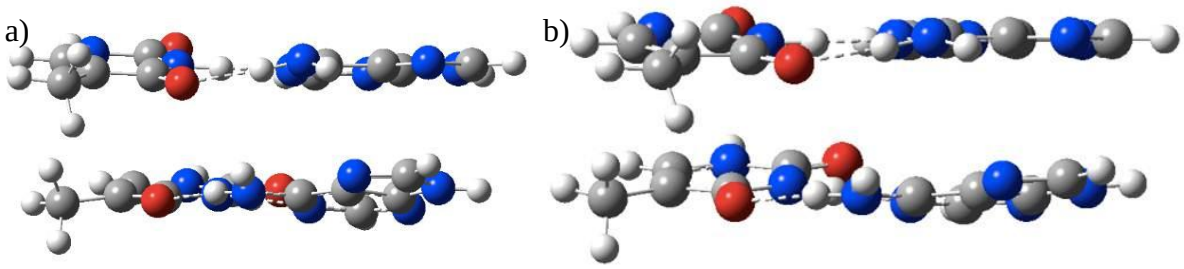

Figure 3.13: Structures of a) AA…TT base step which incorporates the effect of backbone, b) AA…TT base step optimized at the BLYP-D3/def2-TZVP level of theory. Color code: Black for carbon, blue for nitrogen, red for oxygen and white for hydrogen atoms.

Interestingly, fully optimized geometry is only slightly more sTable (by 0.73 kcal/mol) than the constrained geometry. The two geometries do not differ much in terms of their base step parameters (*i.e.* twist, rise, slide etc.) in any significant manner. Therefore, based on our calculations of the geometry and energy of the $(A\ldots T)_2$ with and without considering the sugar-phosphate backbone, it is inferred that, despite the constraints present by the backbone in terms of conformational space, the $(A\ldots T)_2$ retains its near optimal geometry and contributes significantly (~ 20 kcal/mol) to the overall stabilization of the B-DNA structure.

Chapter 4

Summary and Conclusion

Hydrogen bonding and stacking interactions between nucleobases A and T and their base pairs in B-DNA have been studied by using density functional theory that includes dispersion interaction. The effect of the sugar-phosphate backbone on these interactions has also been investigated.

A comparison of the interaction energy values for the Watson-Crick base pairs of A-T, Ad-Th and dMAP-dTMP reveals that both the sugar and phosphate moieties have only a marginal influence on the hydrogen bond interactions between A and T.

A comparison of the interaction energy values for three different conformations of two stacked base pairs of A-T, Ad-Th and dAMP-dTMP shows that for a given conformation, the addition of the sugar moiety as well as the phosphate group does not influence the stacking interaction in a significant manner.

The role of the sugar-phosphate backbone in the formation of the helical structure of B- DNA was also investigated. It is shown that the base pairs themselves have an inherent tendency to form the double helical structure and the backbone does not significantly influence the helix formation.

Clearly, this is a limited study. A more elaborate study considering more than two base pairs, preferably up to five base pairs would be needed to draw firm conclusions on the role of the sugar-phosphate moiety in the formation of the double helix. The inclusion of the solvent molecules and the counter ions (to the phosphate) are clearly beyond the scope of the present investigation.

Bibliography

[1] Wilkins, M.H.F.; Stokes, A.R.; Wilson, H.R. *Nature* **1953**, 171, 738.

[2] Franklin, R.E.; Gosling, R.G. *Nature* **1953**, 171, 740.

[3] Watson, J.D.; Crick, F H.C. *Nature* **1953**, 171, 737.

[4] Hassan, M.A. El; Calladine, C.R. *Phil. Trans. Ro. Soc. Lond. A* **1997,** 355, 43.

[5] Berman et al. *Biophys. J.* **1992**, 63, 751.

[6] Saenger, W. *Principles of Nucleic Acid Structure*, 1st ed.; Springer-Verlag: NY, **1983**.

[7] Lewis, G. N. *Valence and the Structure of Atoms and Molecules*; Chemical Catalog Co.: NY, **1923**; p 109.

[8] van der Waals, J. D. Doctoral Dissertation, Leiden, **1873**.

[9] C.D. Scherrill, *Reviews in Computational Chemistry*, Chapter 1: Computations of non-covalent interactions, pages 1–38. Jon Wiley & Sons, **2009**.

[10] Lehn, J.-M., *Supramolecular Chemistry: Concepts and Perspectives*. NY: VCH, **1995**.

[11] IUPAC Task Group 2004-026-2-100, August 24, **2007**.

[12] Desiraju, G.R., *Angew. Chem. Int. Ed*, **2011**, 50, 52.

[13] Pauling, L., *J. Am. Chem. Soc.* **1931**, 53, 1367.

[14] Hobza, P.; Spirko, V.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem. A* **1998**, 102, 2501.

[15] Landauer, J.; McConnell, H. *J. Am. Chem. Soc*. **1952**, *74*, 1221.

[16] Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc. Perkin Trans. 2* **2001**, 651.

[17] Leontis, N.B.; Westhof, E. *RNA* **2001**, **7**, 499.

[18] Guerra, C. F.; Bickelhaupt, F.M; Snijders, J.G; Baerends E.J; *Chem. Eur. J.* **1999**, 5, 3581.

[19] Crick, F. H. C., *J. Mol. Biol.* **1967**,*19*, 548.

[20] Hoogsteen, K. *Acta. Cryst.* 1963, 16, 907.

[21] Seela, F.; Wei, C.; Melenewski, A. *Nucl. Acids Res.* **1996**, *24*, 4940.

[22] Hunter, C. A.; Lu, X. J. *J. Mol. Biol.* **1997**, 265, 603.

[23] Jurecka, P.; Šponer, J.; Cerny, J.; Hobza, P. *Physical Chemistry Chemical Physics*, **2006**, *8*, 1985.

[24] Lisgatren, J. N.; Coll, M.; Portugal, J.; Wright, C. W.; Aymani, J. *Nature Structural Biology*, **2002**, *9*, 57.

[25] Cerny J, Kabelac P, Hobza P, *J Am Chem Soc* **2008**, 130,16055.

[26] Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.

[27] Muller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, 100,143.

[28] Cornell *et al*, *J. Am. Chem. Soc.*, **117**, 5179.

[29] Hobza, P., Kabelac, M., poner, J., Mejzlik, P. & Vondrasek, J. (1997). *J. Comput. Chem.* **18**, 1136.

[30] Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493.

[31] Hobza, P.; Sÿponer, J. *Chem. Re*v. **1999**, *99*, 3247.

[32] Šponer, J.; Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, 126, 10142.

[33] Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *J. Chem. Phys.* **2011**, 135, 194102.

[34] Hobza P, Šponer J, *J Am Chem Soc* **2002**, 124, 11802.

[35] Šponer, J.; Jurecka, P.; Marchan, I.; Luque, F. J.; Orozco, M.; Hobza, P. *Chem. Eur. J.* **2006**, 12, 2854.

[36] Cooper, V. R.; Thonhauser, T.; Puzder, A.; Schroder, E.; Lundqvist, B. I.; Langreth, D. C. *J. Am. Chem. Soc.* **2008**, 130, 1304.

[37] Raju, R.K.; Bloom, J.W.G.; An, Y.; Wheeler, S.E. *Chem. Phys. Chem* **2011**, 12, 3116.

[38] Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, 118, 11217.

[39] Churchill, C. D. M.; Navarro-Whyte, L.; Rutledge, L. R.; Wetmore, S. D. *Phys. Chem. Chem. Phys.* **2009**, 11, 10657.

[40] Churchill, C. D. M.; Rutledge, L. R.; Wetmore, S. D. *Phys. Chem. Chem. Phys.* **2010**, 12, 14515.

[41] Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, 149, 761.

[42] Parker, T. M.; Hohenstein,E.G.; Parrish, R. M.; Hud, N. V.; Sherrill, C. D. *J. Am. Chem. Soc.* **2013**, 135, 1306.

[43] Szabo, A.; Ostlund, N. S. Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. Dover: New York, **1996**.

[44] Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. J. *Am. Chem. Soc*. **2002**, 124, 10887.

[45] Pulay, P.. *Chem. Phys. Lett.*, **1983**, 100, 151.

[46] Grimme, S., *J. Chem. Phys.* **2003**, 118, 9095.

[47] Werner, H.J.; Manby, F.R.; Knowles, P.J., *J. Chem. Phys.* 2003, 118, 8149.

[48] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

[49] Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.

[50] Gu,J.; Wang,J.; Leszczynski, J.; Xie,Y; Schaefer H. F., *Chem. Phys. Lett.* **2008**, 459, 164.

[51] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, 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, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O¨ . Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Rev. A.02, Wallingford, CT, **2009**.

[52] Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. *J. Chem. Phys.*, **1992**, 96, 6796

[53] Reed, A. E.; Weinhold, F. J. *Chem. Phys.* **1985**, 83, 1736

[54] Boys, S. F.; Bernadi, F.; *Mol. Phys.* **1970**, *19*, 553.

[55] Seeman, N. C.; Rosenberg, J.M.; Suddath F. L.; Kim, J. J. P.; Rich, A. *J. Mol. Biol.* **1976**, 104, 109.

[56] Riley, E. K.; Hobza, P. *J. Phys. Chem. A* **2007**, 111, 8257.

[57] Bertran, J; Olivia, A; Rodriguez-Santiago, L.; Sodupe, M., *J. Am. Chem. Soc.* **1998**, 120, 8159.

[58] Samanta et al., *Int. J. Quant. Chem*, **2008**, 108, 1173.

[59] Electronic structure calculations on workstation computers: the program system TURBOMOLE. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett*. **1989**, 162, 165.

[60] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, 97, 2571.

[61] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1994**, 100, 5829.

[62] Soler-lopez, M.; Malinia, L.; Subirana, J.A.; *J. Bio. Chem.* **2000**, 275, 23034.