Southern Methodist University

SMU Scholar

Summer 8-4-2021

The Role of Hydrogen Bonding and Hydrogen Migration in Different Environments

Nassim Beiranvand
Southern Methodist University, nbeiranvand@smu.edu

THE ROLE OF HYDROGEN BONDING AND

HYDROGEN MIGRATION IN DIFFERENT ENVIRONMENTS

Approved by:

_____

Dr. Elfi Kraka
Professor of Chemistry

_____

Dr. Peng Tao
Professor of Chemistry

_____

Dr. Marek Freindorf
Research Associate Professor of
Chemistry

_____

Dr. Wenjie Su
Professor of Civil and Environmental
Engineering

THE ROLE OF HYDROGEN BONDING AND

HYDROGEN MIGRATION IN DIFFERENT ENVIRONMENTS

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Theoretical and Computational Chemistry

by

Nassim Beiranvand

MSc., Physics, Damghan University, Iran
BSc., Physics, Khoramabad University, Iran

August 4 , 2021

ACKNOWLEDGMENTS

Beiranvand, Nassim

MSc., Physics, Damghan University, Iran
BSc., Physics, Khoramabad University, Iran

The Role of Hydrogen Bonding And
Hydrogen Migration In Different Environments

Advisor: Dr. Elfi Kraka
Doctor of Philosophy degree conferred August 4 , 2021
Dissertation completed July 23, 2021

Hydrogen bonding is one of the most important topics in chemistry and has been a subject of many studies for years, using both experimental techniques and computational methods. My dissertation focuses on mechanistic studies of various types of intermolecular and intramolecular hydrogen bonding in different systems based on quantum chemical calculations.

In my work, I investigated the interplay of ring puckering and hydrogen bonding in DNA building blocks; hydrogen bonding in natural and unnatural base pairs and the nature of dihydrogen bonding by applying the Ring Puckering Analysis and the Local Vibrational Mode Theory (LVM). An additional aim of my research included a new approach to study chemical reactions and the formation of CN bonds in Titan's atmosphere by utilizing the Unified Reaction Valley Approach (URVA). New comprehensive insights into the reaction mechanisms were obtained via the analysis of the curving of the reaction path traced out on the potential energy surface.

TABLE OF CONTENTS

To My Loving Parents

(Tayebeh and Hossein)

Chapter 1

INTRODUCTION

1.1. Overview

This dissertation summarizes my projects during the four years of my study in the Computational and Theoretical Chemistry Group here at SMU. I have been exposed to various theoretical and computational tools in the CATCO research group to address a diverse array of problems in the chemistry research. My research focuses on: a) *Ring Puckering Analysis)* b) *Local Vibrational Mode Theory (LVM)*, and c) *Unified Reaction Valley Approach (URVA)*; all were developed by the CATCO group. Ring Puckering analysis provides puckering coordinates to describe geometry of ring puckering and pseudorotation. LVM provides the local stretching force constant which represents a physically meaningful measurement of chemical bond strength. URVA provides a comprehensive analysis of a reaction mechanism based on Reaction Path Hamiltonian. The details are described in Chapter 2.

This advancement have been used to cast light in the following projects: i) Interplay of Ring Puckering and Hydrogen Bonding in DNA Building Blocks, ii) Hydrogen Bonding in Natural and Unnatural Base Pairs, iii) A New Approach to Study Chemical Reactions in Titan Atmosphere, and iv) The Nature of Dihydrogen Bonds.

1.2. General Introduction

A Deoxyribonucleoside is a type of building blocks of DNA including a purine or pyrimidine base bonded to deoxyribose ring. They are important for the synthesis of anti cancer drugs, [1–3] and the development of polymerase chain reactions with deoxyribonucleotides [4–6]. The 2'-deoxycytidine (Cytosine), **dC**, (b) 2'-deoxyadenosine (Adenine), **dA**, (c) Thymidine (Thymine), **dT**, (d) 2'-deoxyguanosine (Guanine), **dG** are the different types

1

of deoxyribonucleoside. Most studies were focused on influence of the intramolecular hydrogen bonds on the conformational of deoxyribonucleosides [7–12]. But the possibility of studying ring puckering effects on the formation of internal hydrogen bonding and vise versa had not previously been fully explored because of difficulties in describing both the puckering coordinates of a deoxyribose ring [13]. In my work, I was focused on the relationship between ring puckering and hydrogen bonding. The details are described in Chapter 3.

Hydrogen bonds play crucial role in chemical interactions in biological systems [14–20], for example in the natural base pairs ($\mathbf{NBP}_s$) known as Watson-Crick base pairs [21–23] specific hydrogen-bonding patterns shows the specific genetic information. In the nature base pairs selectivity is very exclusive although it is one of the greatest mysteries. Also, long standing research has been devoted to creating and designing other base pairs to expand the genetic alphabet of DNA, by introducing "unnatural" base pairs ($\mathbf{UBP}_s$) to increase nucleic acid functionalities and creation living organism [24–33]. In this connection our work is a thorough quantum-chemical analysis of hydrogen bonding in a diverse set of 36 ($\mathbf{UBP}_s$) and 3 ($\mathbf{NBP}_s$) and also analysis of the strength of hydrogen bonds for four Watson–Crick base pairs embedded in DNA environment using the QM/MM approach. This study aims to cast light on importance of hydrogen bonding in base pairs selectively and different types of hydrogen bonds in $\mathbf{UBP}_s$ and $\mathbf{NBP}_s$ using a comprehensive quantum chemistry tools as it described in computational section.. The details are described in Chapter 4.

Deoxycytidine is one of the building blocks of DNA that is vital for all living organisms and even plants [4, 34–37]. It consist of three major components: a sugar ring, a phosphate group and a nitrogenous base and joined together by intermolecular hydrogen bonds. The study of the hydrogen bonds between DNA unit has long been a topic of significant investigation [12, 38–41]. However, only a few studies have been devoted to evaluate the intramolecular hydrogen bonds within a deoxyribonucleotide. The investigation of these intramolecular hydrogen bonds in the case of deoxyribonucleotides is very complicated, there are numerous intramolecular hydrogen bonds [N⋯H⋯O] and [C⋯H⋯O] between phosphate

group and nitrogenous base as well as sugar ring . In this situation this is quite reasonable that the orientation of the phosphate group and nitrogenous base with respect to the sugar ring can determine the the formation and existence of those hydrogen bonds which is also concerns the equilibrium conformation of deoxyribonucleotides.. The details are described in Chapter 7.

Ion-molecule reactions, are binary collisions of ions (positive or negative) with neutral molecules resulting in chemical reactions, in which at least one chemical bond is ruptured or one new chemical bond is formed. This is the definition of ion-molecule reactions in a narrow but rather conventional sense. The increasing interest in ion-molecule reactions during recent years was undoubtedly motivated by their importance in basic chemistry and physics, but a considerable part of the motivation also stemmed from the fact that detailed knowledge of ion-molecule reactions is required in many fields, such as the physics of the upper atmosphere of the earth and other planets, radiation chemistry, ion-lasers, flames and electrical discharges Understanding the mechanism of elementary chemical reactions that initiate ion processing in Titan's atmosphere, analyze the role of the charge transfer and associated changes in C-N bond formation are goals of this work. See Chapter 5 for details.

In recent years, another very intriguing intermolecular interaction has been discovered, where the two atoms directly involved in the interaction are hydrogens. This interaction was called by Richardson [42] a dihydrogen bond interaction, and it describes attraction between hydridic ($H_{\delta-}$) and protonic ($H_{\delta+}$) hydrogen atoms, where the hydridic atom acts as an electron donor, whereas the protonic atom represents a proton donor of the interaction, similar as in a conventional hydrogen bond. Interaction energies of dihydrogen bonds are generally between 1 and 7 kcal/mol, and because of similarity, dihydrogen bonds are often treated as a subclass of hydrogen bond interactions [43]. The nature of dihydrogen bonds is mostly of the electrostatic origin [43–46], although in stronger dihydrogen bonds a contribution from electronic interactions can be found [43, 47–49]. Therefore in my study I characterized the nature of different dihydrogen bonds, starting from weak interactions of the electrostatic

origin and ending with strong interactions showing a covalent character. See Chapter 6 for details.

Chapter 2

COMPUTATIONAL METHODS

2.1. Local Vibrational Modes

The normal vibrational modes (3N-L) of an N-atomic molecule has all important molecule's electronic structure and therefore, should be well suited as a measure of bond strength. However, in systems larger than diatomic molecules, the normal vibrational modes cannot assess bond strength because it is difficult to decode this information into individual atom-atom interactions. These modes are delocalized because of electronic and mass mode–mode coupling. [50–53] Dieter Cremer and Zoran Konkoli have developed an approach to localize vibrational motions to isolate the frequency of a given internal coordinate (q_n) by decoupling normal modes, including puckering coordinate, bond angle (θ), bond length (r), and dihedral (ϕ). [54–59] The fundamental aspect was to derive constants of local stretching force from local vibrational modes (LVMs), free of any mode-mode coupling, which would then serve as a direct measure of a chemical bond's intrinsic strength. In the Wilson equation's mass–decoupled analog, the mass–coupling was solved [54] that gives an LVM free of any mode–mode coupling. [58] LVM was used soon after its initial formulation in 1997 to measure the chemical bond's intrinsic strength. [60–63] Each LVM is associated with a q_n corresponded to a local mode force constant (k^a) and a local mode frequency (ω^a). There is a unique, direct, and vibrational spectroscopy-based relationship between the local mode force constant and a bond's intrinsic strength. [64] In 2012, Wenli Zou et al. demonstrated a one–to–one relationship between a non–redundant set of local vibrational modes and normal vibrational modes by an adiabatic connection scheme (ACS) known as the foundation for the decomposition of normal modes into local mode contributions (CNM), leading to a basic physical foundation for LVM. [65–67] As such, it resulted in a detailed analysis of a vibrational

spectrum from a computational and measurement point of view. [56,57] The new theory was off and running at this point, which is applied to several chemical systems with both non-covalent and covalent interactions (NCI).Thus, LVM analysis has advanced as a powerful analytical tool and ever since has been extensively applied to a broad range of chemical systems covering both covalent bonds [64, 68–79] and in particular, weak chemical interactions including hydrogen bond, [65, 80–87] halogen bond, [88–93] pnicogen bond, [67, 94, 95] chalcogen bond, [78, 96, 97] tetrel bond, [98] and atom-π interactions. [99, 100] LVM analysis was performed for the investigation of the hydrogen bonds strength in our projects. In my projects the bond strength assessment was simplified [73, 75] by converting k^a to bond strength order (BSO n) with the extended Badger [73, 75, 101] and two reference bonds:

$$BSO\ n \quad = \quad a\,(k^a)^{\,b} \tag{2.1}$$

The constants a and b in Eq. (2.1) were determined via Eq. (2.2) where for hydrogen bonds $n = 1$ for the FH bond and $n = 0.5$ for the FH bond in $[\text{F}\cdots\text{H}\cdots\text{F}]^-$ anion were used [83], leading to $a = 0.515$ and $b = 0.291$, and two reference values and the requirement that for a zero force constant the BSO n is zero.

$$\text{BSO}n(scaled) = 0.515(k^a)^{0.291} \tag{2.2}$$

According to Eq. (2.1) the OH bond in H_2O has a BSO n value of $= 0.966$. Therefore, we scaled the reference values, so that the BSO n of the OH bond in H_2O is 1.

2.2. Unified Reaction Valley Approach

Reaction path hamiltonian (RPH) approach [102] by Miller, Handy, and Adams describe a chemical reaction within a specific region on the potential energy surface (PES), i.e. *reaction valley*, which starts at the transition state (TS) and ends at the PES minima from forward (products) and reverse (reactants) directions. The majority applications of RPH has been applied in the calculation of the dynamics of a chemical reaction, rate constant and

tunneling coefficients. [103,104] Unified Reaction Valley Approach (URVA) provide in-depth mechanistic information about a chemical reaction by focusing on the curving of the reaction path. [105–109] URVA can measure electronic structural changes along the reaction path as shown by the normal vibrational changes, leading to changed mode–mode coupling between translational and vibrational motions of the reacting complex. These effects together highlight the reaction path scalar curvature $\kappa(s)$ changes (s refers to an arc length of reaction path. [110] Significant changes in $\kappa(s)$ relate to the most significant electronic and structural changes, including bond cleavage, formation, charge polarization and transfer, rehybridization, etc for curvature maxima. And curvature minima is related to minimal chemical change which are known as *hidden intermediates* [61, 75, 111, 112]. Also, *reaction phase* defines the region surrounding by two curvature minima containing a curvature peak. [107, 111] To determine which bonds break, form, stretch, and compress, the scalar curvature can also be decomposed into internal coordinate components q_n. [110] Positive contributions q_n to $\kappa(s)$ demonstrates that a particular chemical event is supported by a given component, whereas negative contributions to $(\kappa(s))$ demonstrates a resistant nature. URVA has revealed new mechanistic insight for hundreds of reactions, including the Claisen rearrangement of chorismate performed in the gas phase, water solution, and protein [85], cycloaddition of ethylene or acetylene to 1,3–dipoles; [113, 114] β–hydride elimination in Au(II) and Au(III) complexes; [115] and Au(I) assisted [3,3]-sigmatropic rearrangement of allyl acetate. [116] In my research, each point along the reaction path characterize by its distance from the transition state (TS) given by the arc length s (s = 0 at the TS, with negative values in the entrance channel, s < 0, and positive values in the exit channel, s > 0). The maxima and minima of the curvature profile represents the significant chemical information. Curvature minima could be a representative of hidden intermediates and hidden TSs, which can become real after suitable changes of the substitution patterns or environmental conditions and we use the minima to determine the different phases of reaction. The curvature maxima correspond to locations where chemical events take place such as bond formation and cleavage, charge transfer, charge polarization, and rehybridization. So that each chemical reaction has

a uniquely pattern of curvature maxima and minima with a different number of reaction phases.

2.3. Ring Puckering Analysis

Configurational space is defined by 3N-6 internal parameters that completely describe the geometry of the molecule. But it is difficult to describe the puckered N-membered ring motion via internal coordinate or Cartesian coordinate. In 1972, Altona and Sundaralingam [13] introduced a procedure describing both the puckering angle and puckering amplitude of a five-membered ring via torsional angles. However, the definition of the ring puckering amplitude expressed via an angle does not allow the appropriate description of the ring pseudo-rotation process. In this process, the maximum puckering amplitude is rotating around the edge of the ring without raising a substantial potential energy change. Such a movement can be best described with a puckering amplitude defined as displacement from a reference. Cremer and Pople [117] proposed to use N-3 puckering coordinates for extending a puckered N-membered ring conformational space. For odd-membered rings, it is possible to divide the N-3 puckering coordinates into pairs of pseudo-rotational coordinates $\{q_m, \phi_m\}$ (m = 2, 3, ... N-3) and for even-membered rings, an additional puckering amplitude $q_{N/2}$ that would describe ring inversion. The mth puckering mode's ring puckering degree is represented by the puckering amplitude q_m. The ring puckering mode is defined by the puckering angle ϕ_m. [118, 119] Proposed by Cremer and Pople, the plane of the planar ring can apply to functioning as a mean plane as reference. The N-3 puckering coordinate pairs q_m, ϕ_m can help determine the out-of-plane coordinates z_j that describe the displacement perpendicular to the mean plane [117, 118, 120–122]. Despite the Infinite number of ring conformations placed on a pseudorotation cycle for deoxyribonucleosides' sugar ring, it is sufficient to investigate a representative number of conformations. Referring to previous research on description of tetrahydrofuran, [123] which is an comparable model to the five-membered sugar ring in deoxyribose, a subset of 20 ring conformations were optimized using a mixed set of puckering and internal coordinates. Fig. 2.1, illustrates 10 envelope (E) forms

located at $\phi_2 = (0 + k \times 360)/10$ (for k = 0,1,2,...,9), and 10 twist forms located at $\phi_2 = (18 + k \times 360)/10$ (for k = 0,1,2,...,9). A dashed line is used to represent the internal H-bonds between the base and CH_2OH group of the sugar ring. The ring atom lying above the mean plane is superscripted in front of the conformational symbol E/T. On the other hand, if the ring atom lies below the mean plane, it is subscripted below the symbol E/T. Since the entire studied deoxyribonucleosides in this research have C_1 symmetry, it can be concluded that the conformational energy surface (CES) should also have C_1 symmetry.

The application of ring puckering coordinates yields two main advantages. [124, 125] At first, it is the optimization phase during which, for any given value of q_2 and ϕ_2, sugar ring's geometry is optimized, despite the possible lacking a stationary point on the CES. A Cartesian or internal coordinates based description of the ring would not make this possible. [126] A mixed set of puckering coordinates and internal/Cartesian coordinates for the substituents can apply to specifically identifying any conformer located on the CES for substituted ring systems such as deoxyribonucleosides. Besides, it is possible to gain physically meaningful pseudo-rotation path as a function of two. The expression of any property P of a puckered ring including geometry, magnetic properties, energy, dipole moment, charge distribution, and vibrational frequencies as a Fourier series of the puckering coordinates is the other advantage. In my research, in order to obtain the most stable and global minimum of each deoxyribonucleosides, the optimization calculations were performed using Cartizen coordinates and utilizing the ωB97X-D functional [127, 128] and Pople's 6-31++G(d,p) basis set [129–132]. To the best accuracy considerations, tight convergence criteria, (SCF iterations: 10^{-10} Hartree and geometry optimizations: 10^{-7} Hartree/Bohr) were used and an ultra-fine grid was applied for the DFT numerical integration. [133]. Then for finding the starting point of the pseudo-rotation path, and obtaining the related phase angle ϕ_2 and the puckering amplitude q_2 the RING Puckering program [134] for the puckering analysis was utilized. From this point 20 conformers were calculated by constrained optimizations with fixing phase angle ϕ_2 in increments of 18° to determine the pseudo-rotation path. For the constrained optimizations a mixture of internal coordinates for the substituents and puck-
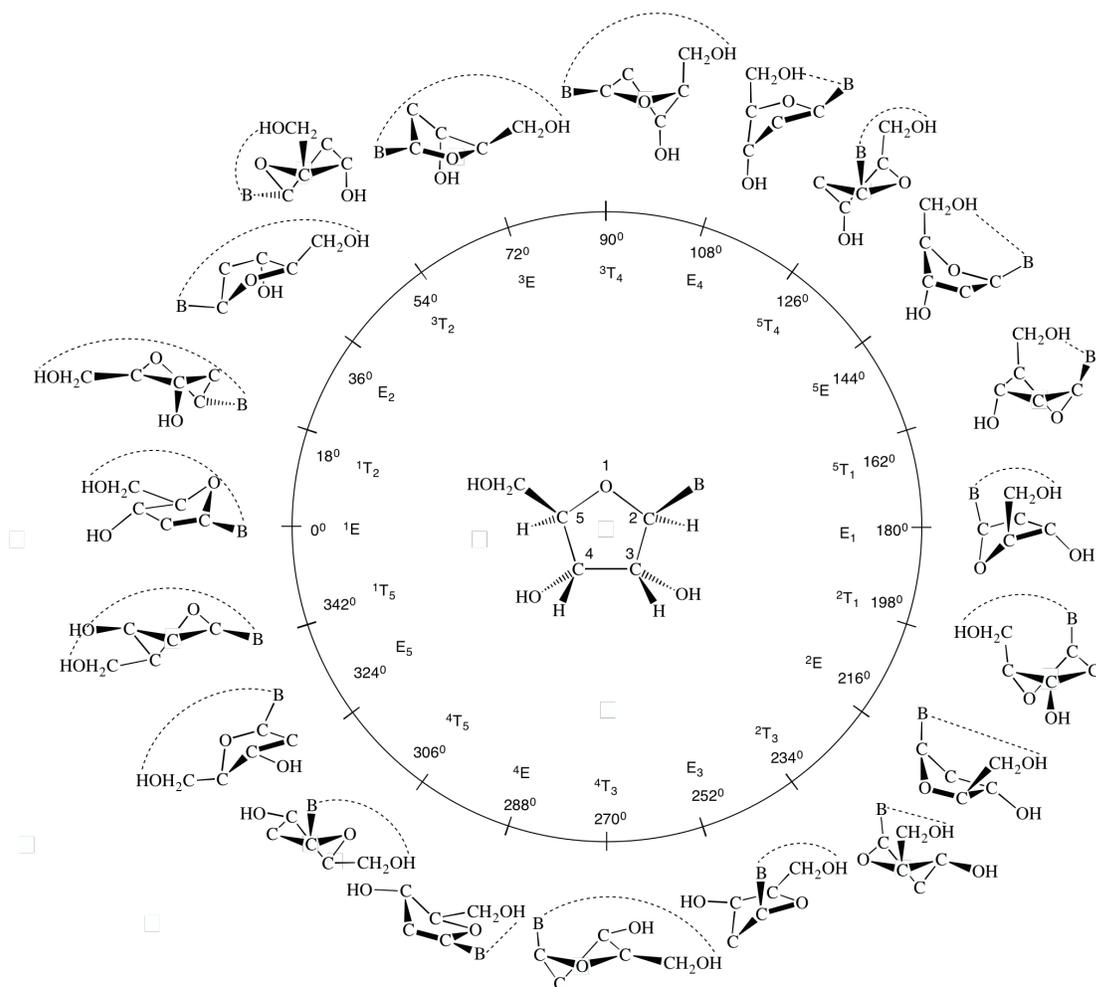
Figure 2.1: Pseudo-rotational cycle ($\phi_2 = 0° \rightarrow 360°$) of the deoxyribonucleoside sugar-ring represented by 20 conformers, (10 envelope (E) forms, $\phi_2 = (0 + k \cdot 360)/10$ and 10 twist (T) forms $\phi_2 = (18 + k \cdot 360)/10$, where $k = 0, 1, 2, ..., 9$). The planar form is located at the center, ($q_2 = 0$ Å). The symbol B represents the base of the deoxyribonucleoside. The dash line indicates the internal H-bond between the base and the CH_2OH substituent of the sugar ring.

ering coordinates for deoxyribose ring was used in the same computational model. In order to verify each structure as a minima (no imaginary frequencies) on the CES the frequency calculations were performed. To calculate the CESs, the same constrained optimization for each fixed ϕ_2 and specific q_2 (0.1Å, 0.2Å, 0.3Å and 0.4Å) which is leading to 80 data points on each CES.

Chapter 3

HYDROGEN BONDING IN DEOXYRIBONUCLEOSIDES

3.1. Background

The results of this work are published in Ref. 87 (Appendix A), therefore here a short overview is given.

The aim of this publication was to combine the Cremer-Pople *Ring Puckering Analysis* [117, 123, 126, 134–137] and the LVM of Konkoli and Cremer, [54–58] to investigate the relationship between ring puckering and hydrogen bonding. The conformational energy surface (CES) for the 2'-deoxycytidine (Cytosine), **dC**, 2'-deoxyadenosine (Adenine), **dA**, Thymidine (Thymine), **dT**, and 2'-deoxyguanosine (Guanine), **dG** (Figure 3.1) were mapped to determine the pseudo-rotation path, conformational energy, the puckering amplitude, and the hydrogen bond properties such as distance and bond strength change along the pseudo-rotation path of deoxyribose ring to investigate the relationship between these two essential structural features.
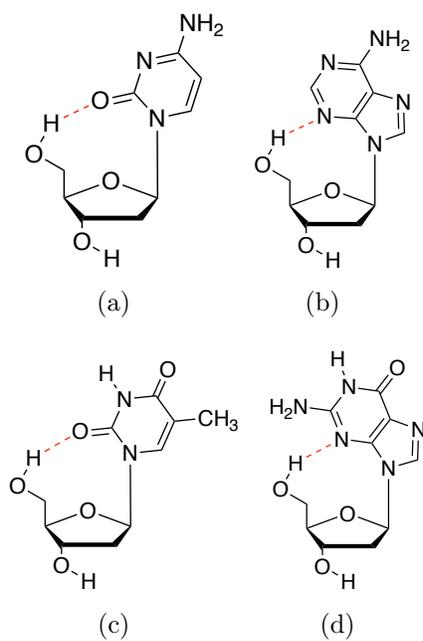
Figure 3.1: Deoxyribonucleosides investigated in this work, (a) 2'-deoxycytidine (Cytosine), **dC**, (b) 2'-deoxyadenosine (Adenine), **dA**, (c) Thymidine (Thymine), **dT**, (d) 2'-deoxyguanosine (Guanine), **dG**. The red dash line indicates the internal hydrogen bond between the base and the CH_2OH substituent of the sugar ring.

3.2. Results and Discussion

The findings show that due to the deoxyribose ring inversion the pseudo-rotation paths on the CESs is an incomplete open curve for all four deoxyribonucleosides. A global and a local minimum separated by a transition state were found on each pseudo-rotation path. Also, the elimination of the hydrogen bond doesn't poses the sugar ring's full conformational flexibility. The dominant role played by ring puckering determines the conformational flexibility of the deoxyribonucleosides. The largest puckering amplitude poses the lower conformational energy since the largest and smallest puckering amplitude occurs in the global minimum and the transition region respectively. While no direct correlation between conformational energy and hydrogen bond strength could be found.

The shortest and longest hydrogen bonds were found at the beginning and ending points of the pseudo-rotation paths while, the longest and weakest hydrogen bonds were located in the local minimum region. By applying the Cremer-Kraka criterion of covalent bonding [138–140], I found that the hydrogen bonds belong to the purine bases have more covalent character than pyrimidine bases.

Via two geometrical parameters defining the rotation of the substituents perpendicular to the sugar ring and their tilting relative to the ring center, I showed how the reciprocal orientation of the CH_2OH group and the base affects hydrogen bond formation. Rotation is more important for H-bond formation, as per these analysis and findings. Greater delocalization energies for the purine bases corresponded to stronger hydrogen bonds according evacuation of the effect of the lone pair's hydrogen bond acceptor through the energy of delocalization.

3.3. Summary

In summary, in this study the Cremer-Pople ring puckering analysis and the Konkoli-Cremer local mode analysis supported by the topological analysis of the electron density were applied to systematically analyze for the first time the interplay between deoxyribose ring puckering and the intramolecular H-bonding in 2'-deoxycytidine (Cytosine), **dC**, 2'-deoxyadenosine (Adenine), **dA**, Thymidine (Thymine), **dT**, and 2'-deoxyguanosine (Guanine), **dG**. Using Cremer-Pople puckering coordinates the CES for any substituted ring system such as a deoxyribonucleoside can be determined in analytical form and a physically meaningful pseudo-rotation path is obtained. Local mode force constants are a unique measure of bond strength and leads to the following significant outcome : We found for all four deoxyribonucleosides incomplete pseudorotation paths on the CESs caused by ring inversion. The pseudorotation path is not a closed circle but an open curve. There are two major factors determining the conformational flexibility of the deoxyribonucleosides, ring puckering and internal H-bonding. Our work showed that ring puckering plays the dominant role. We found a direct correlation between H-bond strength and H-bond length. Longest and weakest H-bonds were found in the local minimum region for all deoxyribonucleosides, whereas shortest and strongest H-bonds were found outside both the global minimum region. This analysis can be applied to the bigger system, the DNA building blocks deoxyribonucleotides, which possess a more complex internal H-bonding pattern caused by the phosphate group.

This work is published in The Journal of Physical Chemistry A (2019) 123:7087–7103 titled:

The Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleosides (see Appendix A).

Chapter 4

HYDROGEN BONDING IN NATURAL AND UNNATURAL BASE PAIRS

4.1. Background

The results of this work were published in Ref. 141 (Appendix B), therefore here a short overview is given.

Investigation of Hydrogen bond (HB) is one of the most rapidly growing areas in scientific research due to its significance in material science, organic and inorganic chemistry, biochemistry, and molecular medicine [142–145]. One of the most crucial chemical interactions in biological systems is the HB [14–20].

The storage of genetic information is based on the specific hydrogen-bonding patterns of the natural base pairs ($\mathbf{NBP}_s$) known as Watson-Crick base pairs. [21–23]. Despite the fact that base pairs were demonstrated such exclusive selectivity in the nature, it remains one of the greatest mysteries. Also, long standing research has been carried out to design other base pairs to expand the genetic alphabet of DNA, by introducing "unnatural" base pairs ($\mathbf{UBP}_s$) to increase nucleic acid functionalities and creation living organism [24–33]. Our work aims to analysis HBs in $\mathbf{NBP}_s$ and $\mathbf{UBP}_s$ systematically at the quantum chemical level to investigate if their HBs differences.

To obtain these, we analyzed 3 $\mathbf{NBP}_s$ and 31 $\mathbf{UBP}_s$ pairs, suggested by Brovarets and co-workers [38] also, in order to have every possible pairing we chose 5 more $\mathbf{UBP}_s$ pairs [146] which do not include C–H$\cdots$O/N bonds, as shown in Figure 4.1. We combined the Konkoli - Cremer local mode analysis [54–58] and complemented this investigation with an analysis of the electron density.
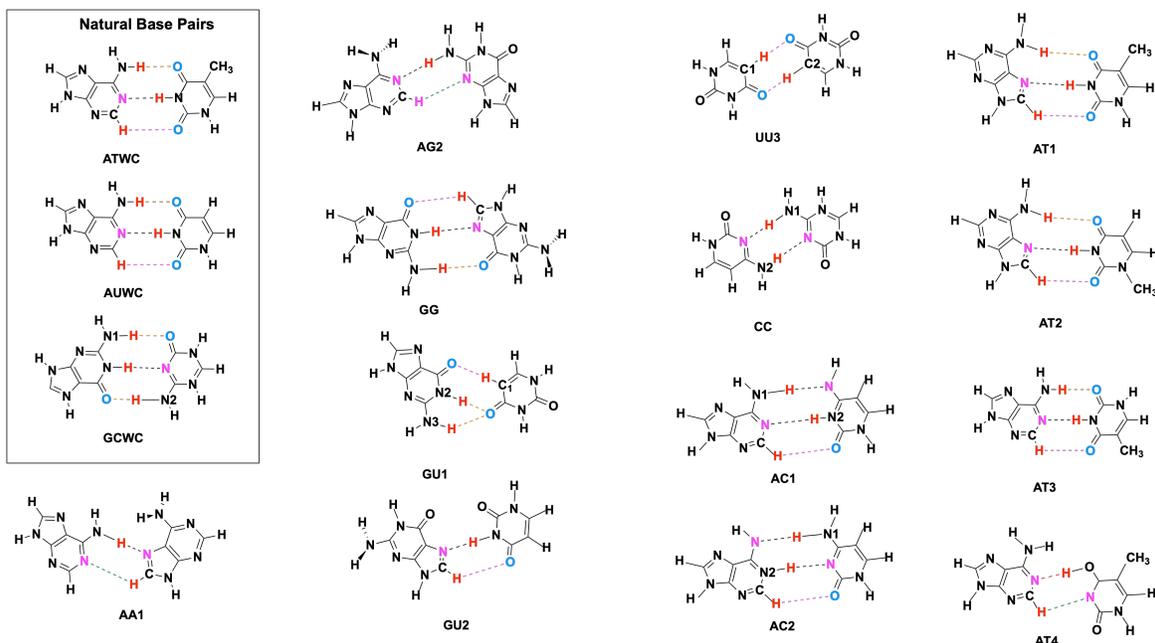
Figure 4.1: some of the Investigated Base Pairs. (Key: A = Adenine, T = Thymine, C = Cytosine, G = Guanine, U = Uracil , H = Hypoxanthine. Different HB was indicated by different color : N–H···N = Black, N–H···O = Orange, C–H···O = Pink, C–H···N = Green, O–H···O = Blue, O–H···N = Red)

4.2. Results and Discussion

Our finding shows a similar range of hydrogen bonds strength for both $\mathbf{UBP}_s$ and $\mathbf{NBP}_s$ based on our analysis, central [N-H···N] bond was found to be strongest hydrogen bond. However, the type of donor and acceptor atom has an important role in base selectivity. $\mathbf{NBP}_s$ are stabilized by [N-H···N] bond, [N-H···O] bond and [C-H···O] bond, while our results for $\mathbf{UBP}_s$ showed the in addition of those hydrogen bonds, there are possibility of [O-H···O] bond, [O-H···N] bond and [C-H···N] bond.

Also, hydrogen bonds for two important $\mathbf{NBP}_s$, $\mathbf{AT}$ and two $\mathbf{GC}$ embedded in DNA based on combined quantum mechanical - molecular mechanical (QM/MM) calculations and then compared with results form calculations of these base pairs obtained in the gas phase. The [N-H···N] bond is found to be strongest hydrogen bonds in both type of base pairs, and [C-H···O] bond has crucial role in structure stabilization. It was found that the DNA environment increases the strength of the central [N-H···N] bond and the [C-H···O] bond,

however it decreases the strength of the [N-H···O] bond in the **AT** base pairs which is one of the important bases to form DNA in nature.
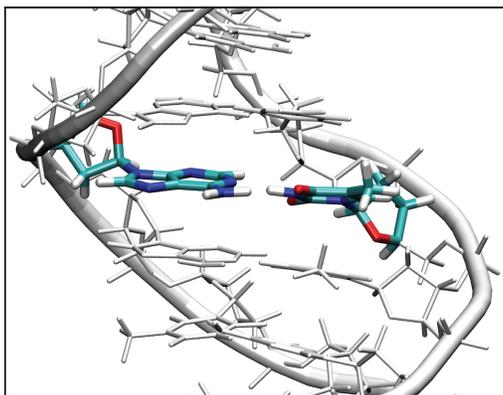
Figure 4.2: The QM/MM optimized geometry of the **AT** base pair in DNA

4.3. Summary

In summary, our study shows that the hydrogen bonding in Watson–Crick base pairs is not exceptionally strong and the hydrogen bond strength of unnatural base pairs fall in the same range as natural Watson–Crick base pairs. N-H⋯N bond, N-H⋯O bond and C-H⋯O bond are common between Both NBPs and UBPs while O-H⋯O bonds, O-H⋯N bonds and C-H⋯N bonds were observed only in UBPs. However, for base selectivity to recognize their match bases, the pattern of donor and acceptor in HB is important. According BSO n values, the N-H⋯N bond and C-H⋯O bond are the strongest and weakest bond in both NBPs and UBPs respectively. N-H⋯N is the most favorable hydrogen bonding and O-H⋯O bonds, O-H⋯N bonds are the less favorable. Non-classical weak hydrogen bonding found to have an important role in base pairs stabilization. Our research indicates that in natural base pairs, not only the intermolecular hydrogen bond strength, but also the combination of classical and non-classical hydrogen bonds plays a significant role, which should be replicated in the development of new promising unnatural base pair candidates. The bioengineering community now has an efficient design tool to analyze and evaluate the type and strength of hydrogen bonding in artificial base pairs, according to our local mode analysis, which was described and tested in this paper. Also, according our previous projects by applying Cremer-Pople ring puckering analysis we can expand this study to analysis the different structure of DNA, A and B, where there is a combination of phosphate group, sugar ring and nitrogenous bases. By applying theses significant quantum toolboxes (ring puckering and local mode analysis), such this project could cast light on the best form of sugar pucker (with respect to global minima) in both A and B conformation of DNA, different types of intramolecular hydrogen bonds involving in both A and B conformation of DNA and intermolecular hydrogen bonds in their Watson-Crick nucleic acid bases. This study was published in Molecules (2021) 26 titled:

Hydrogen Bonding in Natural and Unnatural Base Pairs - Explored with Vibrational Spectroscopy (see Appendix B).

Chapter 5

ON THE FORMATION OF CN BONDS IN TITAN'S ATMOSPHERE

5.1. Background

The results of this study has been submitted to the Journal of Molecular Modeling. Therefore here a short overview is given. (Appendix C). According to high-pressure mass spectrometry experiments combined with computational studies $N^+(^3P)$ then reacts with CH_4 to $HCNH^+$ [147]. The actual formation of $CH_2NH_2^+$ and the appearance of $HCNH^+$ and $CHNH_2^+$ along this path is not understood yet. Figure 5.1 shows the eight reactions which have been investigated in this work. And three main routes which lead to $CH_2NH_2^+$ and starting from CH_3NH^+. Reaction 1 (**R1**) shows the formation of the triplet CH_3NH^+ via a Van der Waals complex ($CH_4 + {}^3N^+$) tends to be triggered by the $N^+(^3P)$ into a C-H bond with a small energy barrier which will be discussed later. The first path is through hydrogen atom migration via reaction 2 (**R2**), reaction 3 (**R3**), reaction 4 (**R4**), and reaction 8 (**R8**). The second path is through hydrogen atom migration follows by reaction 5 (**R5**), reaction 6 (**R6**), then reaction 8 (**R8**). But the third path is an intersystem crossing from the triplet CH_3NH^+ to the unstable form of singlet CH_3NH^+ which is made possible by spin-orbit coupling. Then the singlet CH_3NH^+ through a downhill reaction 7 (**R7**) leads to the reaction 8 (**R8**).

From the reaction between CH_4 (1) and $N^+(^3P)$ (2) three different path-ways for the synthesis of $HCNH^+$ (7) and CH_2NH_2 (4S) result (see Figure 5.1). In this work we investigate these three pathways. Path 1 includes the re-actions R2, R3, R4, and R8, Path 2 includes reactions R5, R6, and R8. In addition to these two paths we also consider Path 3 which starts from the singlet state of CH_3NH^+ (3S), resulting from intersystem crossing of the triplet state CH_3NH^+ (3T) , and includes reactions R7 and R8. The main focus of this study is to explore

the detailed mechanism of the formation of HCNH$^+$ (7) and CH$_2$NH$_2$ (4S) using the Unified Reaction Valley Approach (URVA) [38] and to quantitatively assess and compare the strength of the CN bonds formed in these reactions through local vibrational mode analysis [39]. Our results will give new insights into these extraterrestrial reactions to help better understand prebiotic processes as well as develop guidelines for future investigations involving areas of complex interstellar chemistry.

Figure 5.1: Reaction pathways investigated in this work.

5.2. Results and Discussion

In this work we investigated the formation of protonated hydrogen cyanide $HCNH^+$ (7) and methylene amine cation $CH_2NH_2^+$ (4S) from three different pathways which stem from the interaction between CH_4 (1) and $N^+(^3P)$ (2). Our URVA studies could provide a comprehensive overview on bond formation/cleavage processes relevant to the specific mechanism of eight reactions R1 - R8 that occur across the three pathways. In addition we could explain the formation of $CH_2NH_2^+$ and the appearance of $HCNH^+$ and $CHNH_2$ along these paths. Although only smaller molecules are involved these in reactions including isomerzation, hydrogen atom abstraction and hydrogen molecule capture, we found a number of interesting features, such as roaming in reaction R3 or the primary interaction of H2 with the carbon atom in 7 in reaction R8 followed by migrating of one of the H2 hydrogen atoms to the nitrogen, which is more cost effective than breaking the HH bond first; a feature often found in catalysis. In all cases, charge transfer between carbon and nitrogen could be identified as a driving force for the CN bond formation. Among the three possible pathways, the path via the singlet potential energy surface is the shortest and therefore, most favorable path for the formation of 7 and 4.

5.3. Summary

In summary, our study shows the large potential of both, URVA and LMA to shed new light into these extraterrestrial reactions to help better understand prebiotic processes as well as develop guidelines for future investigations involving areas of complex interstellar chemistry. In particular the formation of CN bonds as a precursor to the extraterrestrial formation of amino acids will be the focus of future investigations. Our new analysis could help to predict the possibility of reactions in planets atmosphere which could lead to discovery of new species. Recently, NASA scientists discovered cyclopropenylidene or C_3H_2 [148] in Titan's atmosphere and they need to know the chemical reactions that lead to this complex to form.

This study has been submitted to the Journal of Molecular Modeling. (2021) titled: *On the formation of CN bonds in Titan's atmosphere - A Unified Reaction Valley Approach study* (see Appendix C).

Chapter 6

DIHYDROGEN BONDING - VIBRATIONAL SPECTROSCOPY

6.1. Background

The complete results of this study are drafted in the Appendix refapp:E, therefore here a short overview is given. Herein, we calculate the vibrational force constants of the adiabatic stretching modes of the molecular systems shown in Scheme 6.1. The focus of our study was to introduce a quantitative bond strength measure based on the local vibrational mode analysis (LMA), originally introduced by Konkoli and Cremer. [54–58] paired with Weinhold's Natural Bond Orbital (NBO) population analysis [149–151] and Bader's Quantum Theory of Atoms-In-Molecules (QTAIM) analysis [152–155]. Figure 6.1 shows some of the set of DHB complexes investigated in this complex (ordered into eight specific groups) providing a broad spectrum of DHBs, ranging from weak interactions of electrostatic origin and to strong interactions with covalent character. We also reassessed the proposed DHB interaction in aromatic hydrocarbons such as phenanthrene, dibenz[a]anthracene, and biphenyl [49,156] in which both hydrogen atoms involved are equally charged.
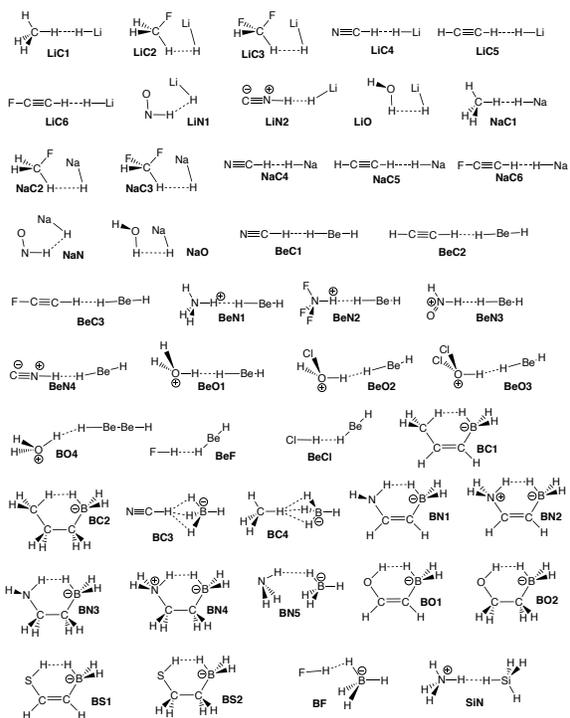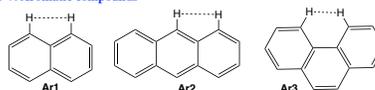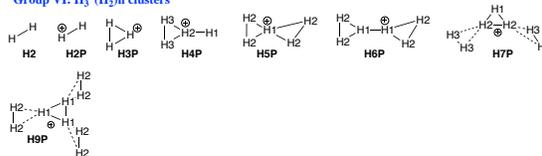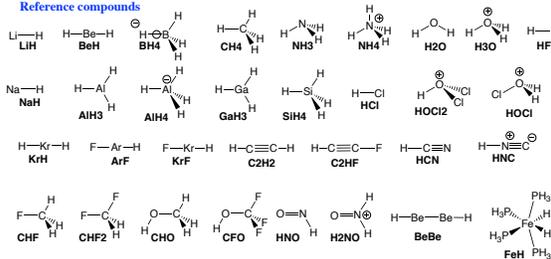
Figure 6.1: Sketches of molecular systems investigated in this study.

6.2. Results and Discussion

In this study we analyzed the strength of a dihydrogen bond by utilizing local vibration modes, local mode frequency shifts and the covalent character of the dihydrogen bond via the Cremer-Kraka criterion for covalent bonding. According the BSO n values for the DHBs, the protonic and hydridic H-bonds as well as the correlation between local mode force constants k^a(Protonic) and k^a(Hydridic), k^a(DHB) and k^a(Protonic) and k^a(DHB) and k^a(Hydridic), the values range from 0.14 (**FeF2**) to 0.41 (**BeO4**) which is in the same range as we previously found for HBs (BSO n=0.135, N-H···F; BSO n=0.33, N-H···N). The protonic H-bonds (i.e., donor part) ranging from 0.49 for **BeO3** to 0.99 for **CF**, close to the FH reference (BSO n=1) are stronger than their hydridic counter parts with BSO n=0.41 for **NaO** and 0.86 for **Ar3**, respectively. We did not find correlation between k^a(Protonic) and k^a(Hydridic) but here is a tendency that that weaker protonic H-bonds correspond to stronger DHBs, e.g. **BeO3** whereas stronger protonic H-bonds correspond to weaker DHBs, reflecting that the donor bond transfers electron density to the DHB. According to the Cremer-Kraka criterion for covalent bonding, the H_c/ρ_c values range from covalent, H_c/ρ_c= -0.6 [Hartree/electron] to electrostatic H_c/ρ_c= 0.2 (Hartree /electron) which matches the range of H_c/ρ_c values -0.6 to 0.3 (Hartree/electron) for the HBs investigated in our previous work, [81], disproving the assumption that DHBs are generally more electrostatic than their HB counterparts. Also, we did not find direct correlation between the difference Δq of protonic and hydridic H-atom NBO charges, which shows that differences in the atomic charges of protonic and hydridic H-atom cannot be used as direct measure for the strength of these complex interaction.

6.3. Summary

The bond orders of the dihydrogen interaction was obtained in our investigation, by using the theory of adiabatic internal vibrational modes, which provides a reliable characteristic of a bond strength. At the time of submitting this dissertation (July 23th, 2021) the manuscript is under preparation and will be submitted to the Molecules at the earliest. Draft is presented in Appendix E .

Chapter 7

PUCKERING AND HYDROGEN BONDING IN DEOXYRIBONUCLEOTIDES

7.1. Background

The complete results of this study are presented in (Appendix D), therefore here a short overview is given. We used in this work as efficient tool a combination of the Cremer-Pople *Ring Puckering Analysis* [117, 123, 126, 134–137] and the local mode analysis of Konkoli and Creme (LVM) of Konkoli and Cremer, [54–58] to determine the CES and pseudo-rotation path and the location of the global on the CES for the Deoxycytidine monophosphate (**dCMP** (Figure 7.1) which is a deoxyribonucleotides, and one of the four monomers that make up DNA [**?**]), in order to answering the following questions:

1. How the conformational energy, the puckering amplitude change along the pseudorotation path of sugar ring?

2. How the intramolecular H-bonds change along the pseudorotation path of sugar ring? Is there any connection?

3. What type of H-bonds contribution is possible between phosphate group and nitrogenous base and sugar ring?

4. Compare with our previous work [87], Deoxyribonucleosides to understand the role of phosphate group?
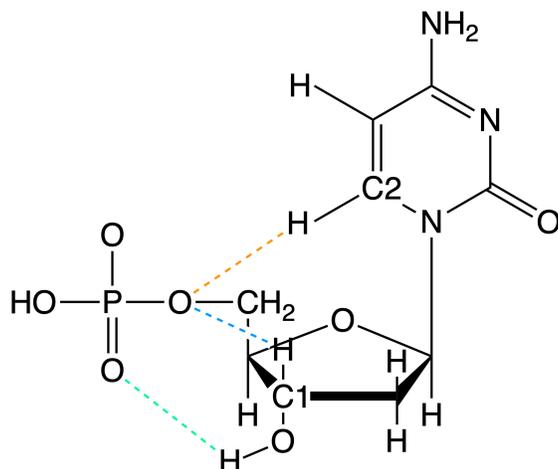
Figure 7.1: Deoxycytidine structure investigated in this work. Numbering in C1 and C2 is just for distinguishing between the C–H···O bond of phosphate group with ribose ring and nitrogenous base respectively.

7.2. CESs and pseudo-rotation (pseudo-libration) path

The pseudo-rotation paths on the corresponding CES of the **dCMP** is shown in Figure 7.2. We found negative puckering amplitudes on the pseudo-rotation paths in the range of ϕ_2 = 0 to 162° indicating a ring inversion. While in the case of deoxyribonucleosides we found negative puckering amplitudes in the range of ϕ_2 = 108 to 270° which is longer path. But similar to the deoxyribonucleosides the pseudo-rotation path is an incomplete open curve, which we made as pseudo-libration path. Global minimum was identified in the ϕ_2 = 270 °, ΔE = 0.28 kcal/mol on the pseudo-rotation path (Figure 7.2). At the beginning of path where ϕ_2 = 180.0° with the highest energy (ΔE = 11.96 kcal/mol) respect to the planar form, right after this point the conformer undergoes a complete rotation as shown in Figure 7.2.

Figure 7.2: Conformational Energy Surfaces (CESs) and pseudo-rotation path of **dCMP**. At the center of CES, the planar deoxyribose ring form is located. The corresponding energy is used as reference. The X axe shows the magnitude of puckering amplitude q_2 in Å. The labels around CES circle denote the value of the phase angle ϕ_2 in Degree. The solid blue line indicates the pseudo-rotation path. The color bar represents the energy on the CES in kcal/mol relative to the planar form; *yellow to red regions:* location of conformers higher in energy than the planar form; *green to blue regions:* location of conformers lower in energy than the planar form. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

7.3. Summary

For the first time, the Cremer-Pople ring puckering analysis and the Konkoli-Cremer local mode analysis, accompanied by topological analysis of the electron density, were used to systematically investigate the interplay between deoxyribose ring puckering and intramolecular H-bonding in Deoxycytidine (**dCMP**). Our work has led to the following conclusions:

We found incomplete pseudo-rotation paths on the CESs caused by ring inversion for **dCMP**, similar to deoxyribonucleosides in our previous study [87] where we we coined an open curve pseudo-rotation path as pseudo-libration. But due to the phosphate group, there are more H-bonds, which shortens the pseudo-libration path. Also, on the pseudo-libration paths a global minimum in the range of $\phi_2 = 270\text{-} 288°$ could be identified but not a local minimum or transition state. Our work showed that ring puckering plays the dominant role on determining the conformational flexibility of **dCMP** among two major factors (ring puckering and internal H-bonding). The largest puckering amplitude $q_2 = 0.408$ Åoccurs in the $\phi_2 = 270$ ° which is the global minimum region. This suggests that the lower the conformational energy, the larger the puckering amplitude, which is consistent with our previous work [87]. We found three H-bonds (O–H$\cdots$O bond and two C–H$\cdots$O bond which occurs between phosphate group with ribose ring and nitrogenous base of **dCMP**), then quantitatively assessed the H-bonds strength along the pseudo-rotation paths via bond strength orders BSO n derived from local vibrational force constants. The number of H-bonds decreases along the pseudo-rotation path, from three at the start to two at the global minimum and one at the end of the path. The O–H$\cdots$O bond between phosphate group and OH group of deoxyribose ring, constantly present with increasing trend of strength along the pseudo-rotation path but C1–H$\cdots$O bond just found in $\phi_2 = 180\text{-}234$ °. We found a direct correlation between H-bonds strength and H-bonds length for O–H$\cdots$O bond and C2–H$\cdots$O bond but we did not find it in C1–H$\cdots$O. We found the shortest and strongest O–H$\cdots$O bond and C2–H$\cdots$O bond at $\phi_2 = 324$ ° outside the global minimum region close to the end point and $\phi_2 = 252°$ respectively. These results clearly reveal that H-bonding determines the shape and length of the pseudo-libration paths.

This new analysis could be extended to different types of deoxyribonucleotides to understand the influence of puric and pyrimidic bases on formation internal H-bonding.

At the time of submitting this dissertation (July 23th, 2021) the manuscript is under preparation and will be submitted to the Journal of Physical Chemistry A at the earliest. Draft is presented in Appendix D.

BIBLIOGRAPHY

[1] Claudia Riccardi, Domenica Musumeci, Carlo Irace, Luigi Paduano, and Daniela Montesarchio. Ru(iii) complexes for anticancer therapy: The importance of being nucleolipidic. *Eur. J. Org. Chem.*, 2017(7):1100–1119, 2017. 1

[2] Vipin Kumar, Shyam Kishor, and Lavanya M Ramaniah. Chemical reactivity analysis of deoxyribonucleosides and deoxyribonucleoside analogues (nrtis): A first-principles density functional approach. *J. Mol. Model.*, 18(8):3969–3980, 2012. 1

[3] DC Baker and Chung Chu. *Nucleosides and Nucleotides as Antitumor and Antiviral Agents.* Springer, 2013. 1

[4] Michal Hocek. Synthesis of base-modified 2'-deoxyribonucleoside triphosphates and their use in enzymatic synthesis of modified dna for applications in bioanalysis and chemical biology. *J. Org. Chem.*, 79(21):9914–9921, 2014. 1, 2

[5] Pavel Kielkowski, Jindřich Fanfrlík, and Michal Hocek. 7-aryl-7-deazaadenine 2'-deoxyribonucleoside triphosphates (dntps): Better substrates for dna polymerases than datp in competitive incorporations. *Angew. Chem. Int. Ed.*, 53(29):7552–7555, 2014. 1

[6] Veronika Raindlová, Radek Pohl, Miloslav Šanda, and Michal Hocek. Direct polymerase synthesis of reactive aldehyde-functionalized dna and its conjugation and staining with hydrazines. *Angew. Chem. Int. Ed.*, 49(6):1064–1066, 2010. 1

[7] R R Wu, Bo Yang, C E Frieler, G Berden, J Oomens, and M T Rodgers. N3 and o2 protonated tautomeric conformations of 2'-deoxycytidine and cytidine coexist in the gas phase. *J. Phys. Chem. B*, 119(18):5773–5784, 2015. 2

[8] Piotr Storoniak, Janusz Rak, Yeon Jae Ko, Haopeng Wang, and Kit H Bowen. Photoelectron spectroscopic and density functional theoretical studies of the 2'-deoxycytidine homodimer radical anion. *J. Chem. Phys.*, 139(7):075101, 2013. 2

[9] Alla G Ponomareva, Yevgen P Yurenko, Roman O Zhurakivsky, Tanja Van Mourik, and Dmytro M Hovorun. Complete conformational space of the potential hiv-1 reverse transcriptase inhibitors d4u and d4c. a quantum chemical study. *Phys. Chem. Chem. Phys.*, 14(19):6787–6795, 2012. 2

[10] Zahra Aliakbar Tehrani, Marjan Jebeli Javan, Alireza Fattahi, and Mohammad Mahmoodi Hashemi. Structures, stabilities & conformational behaviors of hydrogen-atom abstractions of cytosine nucleosides: Aim & nbo analysis. *Comput. Theor. Chem.*, 971(1):19–29, 2011. 2

[11] Chérif F Matta, Norberto Castillo, and Russell J Boyd. Extended weak bonding interactions in dna: π-stacking (base-base), base-backbone, and backbone-backbone interactions. *J. Phys. Chem. B*, 110(1):563–578, 2006. 2

[12] Alexandre Hocquet and Mahmoud Ghomi. The peculiar role of cytosine in nucleoside conformational behavior: Hydrogen bond donor capacity of nucleic bases. *Phys. Chem. Chem. Phys.*, 2(23):5351–5353, 2000. 2

[13] C Altona and M Sundaralingam. Conformational analysis of the sugar ring in nucleosides and nucleotides. new description using the concept of pseudorotation. *J. Am. Chem. Soc.*, 94(23):8205–8212, 1972. 2, 8

[14] Gautam R Desiraju and Thomas Steiner. *The Weak Hydrogen Bond: In Structural Chemistry and Biology*, volume 9. International Union of Crystal, 2001. 2, 16

[15] Linus Pauling. *The Nature of the Chemical Bond*, volume 260. Cornell university press Ithaca, NY, 1960. 2, 16

[16] Thomas Steiner. The hydrogen bond in the solid state. *Angew. Chem. Int. Ed.*, 41(1):48–76, 2002. 2, 16

[17] Gopalakrishnan Bulusu and Gautam R Desiraju. Strong and weak hydrogen bonds in protein–ligand recognition. *J Indian Inst Sci*, pages 1–11, 2019. 2, 16

[18] C. Nick Pace, Hailong Fu, Katrina Lee Fryar, John Landua, Saul R. Trevino, David Schell, Richard L. Thurlkill, Satoshi Imura, J. Martin Scholtz, Ketan Gajiwala, Jozef Sevcik, Lubica Urbanikova, Jeffery K. Myers, Kazufumi Takano, Eric J. Hebert, Bret A. Shirley, and Gerald R. Grimsley. Contribution of hydrogen bonds to protein stability. *Protein Sci.*, 23(5):652–661, 2014. 2, 16

[19] C. Nick Pace, J. Martin Scholtz, and Gerald R. Grimsley. Forces stabilizing proteins. *FEBS Letters*, 588(14):2177–2184, 2014. 2, 16

[20] Petr Kolosov, Ludmila Frolova, Alim Seit-Nebi, Vera Dubovaya, Artem Kononenko, Nina Oparina, Just Justesen, Alexandr Efimov, and Lev Kisselev. Invariant amino acids essential for decoding function of polypeptide release factor erf1. *Nucleic Acids Res. Spec. Publ.*, 33(19):6418–6425, 2005. 2, 16

[21] James D Watson and Francis HC Crick. Molecular structure of nucleic acids. *Nature*, 171(4356):737–738, 1953. 2, 16

[22] James D Watson and Francis HC Crick. Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171(4361):964–967, 1953. 2, 16

[23] Célia Fonseca Guerra, F Matthias Bickelhaupt, Jaap G Snijders, and Evert Jan Baerends. The nature of the hydrogen bond in dna base pairs: The role of charge transfer and resonance assistance. *Chem.Eur. J.*, 5(12):3581–3594, 1999. 2, 16

[24] Armando R Hernandez, Yaming Shao, Shuichi Hoshika, Zunyi Yang, Sandip A Shelke, Julien Herrou, Hyo-Joong Kim, Myong-Jung Kim, Joseph A Piccirilli, and Steven A Benner. A crystal structure of a functional rna molecule containing an artificial nucleobase pair. *Angew. Chem. Int. Ed.*, 127(34):9991–9994, 2015. 2, 16

[25] Aaron W Feldman, Michael P Ledbetter, Yorke Zhang, and Floyd E Romesberg. Reply to hettinger: Hydrophobic unnatural base pairs and the expansion of the genetic alphabet. *PNAS. U.S.A.*, 114(32):E6478–E6479, 2017. 2, 16

[26] Janelle Weaver. Expanding the genetic alphabet. *BIOTECHNIQUES*, 62(6):252–253, 2017. 2, 16

[27] Elisa Biondi and Steven A Benner. Artificially expanded genetic information systems for new aptamer technologies. *Biomedicines*, 6(2):53– 66, 2018. 2, 16

[28] Kiyofumi Hamashima, Yun Ting Soong, Ken Ichiro Matsunaga, Michiko Kimoto, and Ichiro Hirao. Dna sequencing method including unnatural bases for dna aptamer generation by genetic alphabet expansion. *ACS Synth. Biol.*, 8(6):1401–1410, 2019. 2, 16

[29] Frank Eggert, Katharina Kurscheidt, Eva Hoffmann, and Stephanie Kath-Schorr. Towards reverse transcription with an expanded genetic alphabet. *ChemBioChem*, 20(13):1642–1645, 2019. 2, 16

[30] Aaron W. Feldman, Vivian T. Dien, Rebekah J. Karadeema, Emil C. Fischer, Yanbo You, Brooke A. Anderson, Ramanarayanan Krishnamurthy, Jason S. Chen, Lingjun Li, and Floyd E. Romesberg. Optimization of replication, transcription, and translation in a semi-synthetic organism. *J. Am. Chem. Soc.*, 141(27):10644–10653, 2019. 2, 16

[31] Vivian T. Dien, Matthew Holcomb, and Floyd E. Romesberg. Eight-letter dna. *Biochemistry*, 58(22):2581–2583, 2019. 2, 16

[32] Sk Jahiruddin and Ayan Datta. What sustains the unnatural base pairs (ubps) with no hydrogen bonds. *J. Phys. Chem. B*, 119(18):5839–5845, 2015. 2, 16

[33] Sk Jahiruddin, Nilangshu Mandal, and Ayan Datta. Structure and electronic properties of unnatural base pairs: The role of dispersion interactions. *ChemPhysChem*, 19(1):67–74, 2018. 2, 16

[34] George A. Jeffrey and Wolfram Saenger. *Hydrogen Bonding in Biological Structures.* Springer-Verlag Berlin Heidelberg, 1991. 2

[35] Wolfram Saenger. *Principles of Nucleic Acid Structure.* Springer, New York, NY, 1984. 2

[36] Anthony M Poole, Nobuyuki Horinouchi, Ryan J Catchpole, Dayong Si, Makoto Hibi, Koichi Tanaka, and Jun Ogawa. The case for an early biological origin of dna. *J. Mol. Evol.*, 79(5-6):204–212, 2014. 2

[37] DC Baker and Chung Chu. *Nucleosides and nucleotides as antitumor and antiviral agents.* Springer, 2013. 2

[38] Ol′ha O. Brovarets′, Yevgen P. Yurenko, and Dmytro M. Hovorun. Intermolecular ch⋯o/n h-bonds in the biologically important pairs of natural nucleobases: A thorough quantum-chemical study. *J. Biomol. Struct. Dyn.*, 32(6):993–1022, 2014. 2, 16

[39] Albert Rimola, Mariona Sodupe, and Piero Ugliengo. Role of mineral surfaces in prebiotic chemical evolution. in silico quantum mechanical studies. *Life*, 9(1), 2019. 2

[40] Iwona Dąbkowska, Petr Jurečka, and Pavel Hobza. On geometries of stacked and h-bonded nucleic acid base pairs determined at various dft, mp2, and ccsd(t) levels up to the ccsd(t)/complete basis set limit level. *J. Chem. Phys.*, 122(20):204322, 2005. 2

[41] Yosslen Aray, Ricardo Aguilera-García, and Daniel R. Izquierdo. Exploring the nature of the h-bonds between the human class ii mhc protein, hla-dr1 (drb*0101) and the influenza virus hemagglutinin peptide, ha306-318, using the quantum theory of atoms in molecules. *J. Biomol. Struct. Dyn.*, 37(1):48–64, 2018. 2

[42] Thomas Richardson, Susan de Gala, Robert H. Crabtree, and Per E. M. Siegbahn. Unconventional hydrogen bonds: intermolecular B-H⋯H-N interactions. *J. Am. Chem. Soc.*, 117(51):12875–12876, 1995. 3

[43] Radu Custelcean and James E. Jackson. Dihydrogen bonding: structures, energetics, and dynamics. *Chem. Rev.*, 101(7):1963–1980, 2001. 3

[44] Ibon Alkorta and Jose Elguero. Non-conventional hydrogen bonds. *Chem. Soc. Review.*, 27:163–170, 1998. 3

[45] Lina M. Epstein and Elena S. Shubina. New types of hydrogen bonding in organometallic chemistry. *Coord. Chem. Rev.*, 231(1-2):165 – 181, 2002. 3

[46] Mohammad Solimannejad, Leyla Mohammadi Amlashi, Ibon Alkorta, and Jose Elguero. XeH_2 as a proton-accepting molecule for dihydrogen bonded systems: A theoretical study. *Chemical Physics Letters*, 422(1-3):226 – 229, 2006. 3

[47] I. Alkorta, J. Elguero, Otilia Mo, Manuel Yanez, and Janet E. Del Bene. Ab initio study of the structural, energetic, bonding, and IR spectroscopic properties of complexes with dihydrogen bonds. *J. Phys. Chem. A*, 106(40):9325–9330, 2002. 3

[48] Ibon Alkorta, Krzysztof Zborowski, Jose Elguero, and Mohammad Solimannejad. Theoretical study of dihydrogen bonds between $(XH)_2$, X = Li, Na, BeH, and MgH, and weak hydrogen bond donors (HCN, HNC, and HCCH). *J. Phys. Chem. A*, 110(34):10279–10286, 2006. 3

[49] Slawomir J. Grabowski, W. Andrzej Sokalski, and Jerzy Leszczynski. Wide spectrum of H···H interactions: van der waals contacts, dihydrogen bonds and covalency. *Chem. Phys.*, 337(1-3):68–76, 2007. 3, 24

[50] E. B. Wilson, J. C. Decius, and P. C. Cross. *Molecular Vibrations : The Theory of Infrared and Raman Vibrational Spectra.* McGraw-Hill, New York, NY, USA, 1955. 5

[51] Leonard Ary Woodward. *Introduction to the Theory of Molecular Vibrations and Vibrational Spectroscopy.* Oxford University Press, Oxford, United Kingdom, 1972. 5

[52] G. Herzberg. *Molecular Spectra and Molecular Structure. Volume II: Infrared and Raman Spectra of Polyatomic Molecules.* Krieger Publishing Co., New York, NY, USA, 1991. 5

[53] G. Herzberg and K. P. Huber. *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules.* Springer, New York, NY, USA, 1979. 5

[54] Z. Konkoli and D. Cremer. A New Way of Analyzing Vibrational Spectra. I. Derivation of Adiabatic Internal Modes. *Int. J. Quant. Chem.*, 67:1–9, 1998. 5, 12, 16, 24, 27

[55] Z. Konkoli, J. A. Larsson, and D. Cremer. A New Way of Analyzing Vibrational Spectra. II. Comparison of Internal Mode Frequencies. *Int. J. Quant. Chem.*, 67:11–27, 1998. 5, 12, 16, 24, 27

[56] Z. Konkoli and D. Cremer. A New Way of Analyzing Vibrational Spectra. III. Characterization of Normal Vibrational Modes in terms of Internal Vibrational Modes. *Int. J. Quant. Chem.*, 67:29–40, 1998. 5, 6, 12, 16, 24, 27

[57] Z. Konkoli, J. A. Larsson, and D. Cremer. A New Way of Analyzing Vibrational Spectra. IV. Application and Testing of Adiabatic Modes within the Concept of the Characterization of Normal Modes. *Int. J. Quant. Chem.*, 67:41–55, 1998. 5, 6, 12, 16, 24, 27

[58] D. Cremer, J. A. Larsson, and E. Kraka. New Developments in the Analysis of Vibrational Spectra on the Use of Adiabatic Internal Vibrational Modes. In C. Parkanyi, editor, *Theoretical and Computational Chemistry*, pages 259–327. Elsevier, Amsterdam, 1998. 5, 12, 16, 24, 27

[59] Wenli Zou, Yunwen Tao, Dieter Cremer, and Elfi Kraka. Systematic Description of Molecular Deformations with Cremer-Pople Puckering and Deformation Coordinates Utilizing Analytic Derivatives: Applied to Cycloheptane, Cyclooctane, and Cyclo[18]carbon. *J. Chem. Phys.*, 152:154107–1–54107–15, 2020. 5

[60] W. Quapp, E. Kraka, and D. Cremer. Finding the Transition State of Quasi-Barrierless Reactions by a Growing String Method for Newton Trajectories: Application to the Dissociation of Methylenecyclopropene and Cyclopropane. *J. Chem. Phys. A*, 111:11287–11293, 2007. 5

[61] H. Joo, E. Kraka, W. Quapp, and D. Cremer. The Mechanism of a Barrierless Reaction: Hidden Transition State and Hidden Intermediates in the Reaction of Methylene with Ethene. *Mol. Phys.*, 105:2697–2717, 2007. 5, 7

[62] J. Gräfenstein and D. Cremer. Unusual Long-Range Spin-Spin Coupling in Fluorinated Polyenes: A Mechanistic Analysis. *J. Chem. Phys.*, 127:174704, 2007. 5

[63] Jürgen Gräfenstein and Dieter Cremer. Efficient Density–Functional Theory Integrations by Locally Augmented Radial Grids. *J. Chem. Phys.*, 127(16):164113, 2007. 5

[64] W. Zou and D. Cremer. C_2 in a Box: Determining its Intrinsic Bond Strength for the $X^1\Sigma^+_g$ Ground State. *Chem. Eur. J.*, 22:4087–4097, 2016. 5, 6

[65] R. Kalescky, E. Kraka, and D. Cremer. Local Vibrational Modes of the Formic Acid Dimer - The Strength of the Double H-Bond. *Mol. Phys.*, 111:1497–1510, 2013. 5, 6

[66] R. Kalescky, E. Kraka, and D. Cremer. New Approach to Tolman's Electronic Parameter Based on Local Vibrational Modes. *Inorg. Chem.*, 53:478–495, 2013. 5

[67] D. Setiawan, E. Kraka, and D. Cremer. Description of Pnicogen Bonding with the help of Vibrational Spectroscopy-The Missing Link Between Theory and Experiment. *Chem. Phys. Letters*, 614:136–142, 2014. 5, 6

[68] R. Kalescky, E. Kraka, and D. Cremer. Identification of the Strongest Bonds in Chemistry. *J. Phys. Chem. A*, 117:8981–8995, 2013. 6

[69] Dani Setiawan, Daniel Sethio, Dieter Cremer, and Elfi Kraka. From Strong to Weak NF Bonds: On the Design of a New Class of Fluorinating Agents. *Phys. Chem. Chem. Phys.*, 20:23913–23927, 2018. 6

[70] A. Humason, W. Zou, and D. Cremer. 11,11-Dimethyl-1,6-methano[10]annulene-An Annulene with an Ultralong CC Bond or a Fluxional Molecule? *J Phys Chem A.*, 119:1666–1682, 2014. 6

[71] R. Kalescky, E. Kraka, and D. Cremer. Are Carbon-Halogen Double and Triple Bonds Possible? *Int. J. Quant. Chem.*, 114:1060–1072, 2014. 6

[72] R. Kalescky, W. Zou, E. Kraka, and D. Cremer. Quantitative Assessment of the Multiplicity of Carbon-Halogen Bonds: Carbenium and Halonium Ions with F, Cl, Br, and I. *J. Phys. Chem. A*, 118:1948–1963, 2014. 6

[73] E. Kraka, J. A. Larsson, and D. Cremer. Generalization of the Badger Rule Based on the Use of Adiabatic Vibrational Modes. In J. Grunenberg, editor, *Computational Spectroscopy*, pages 105–149. Wiley, New York, 2010. 6

[74] J. A. Larsson and D. Cremer. Theoretical verification and extension of the Mckean relationship between bond lengths and stretching frequencies. *J. Mol. Struct.*, 485-486:385–407, 1999. 6

[75] D. Cremer and E. Kraka. From Molecular Vibrations to Bonding, Chemical Reactions, and Reaction Mechanism. *Curr. Org. Chem.*, 14:1524–1560, 2010. 6, 7

[76] D. Cremer, A. Wu, J. A. Larsson, and E. Kraka. Some Thoughts about Bond Energies, Bond Lengths, and Force Constants. *J. Mol. Model.*, 6:396–412, 2000. 6

[77] E. Kraka and D. Cremer. Weaker Bonds with Shorter Bond Lengths. *Rev. Proc. Quim.*, pages 39–42, 2012. 6

[78] D. Setiawan, E. Kraka, and D. Cremer. Hidden Bond Anomalies: The Peculiar Case of the Fluorinated Amine Chalcogenides. *J. Phys. Chem. A*, 119:9541–9556, 2015. 6

[79] E. Kraka, D. Setiawan, and D. Cremer. Re-Evaluation of the Bond Length-Bond Strength Rule: The Stronger Bond Is not Always the Shorter Bond. *J. Comp. Chem.*, 37:130–142, 2015. 6

[80] R. Kalescky, W. Zou, E. Kraka, and D. Cremer. Local Vibrational Modes of the Water Dimer - Comparison of Theory and Experiment. *Chem. Phys. Letters*, 554:243–247, 2012. 6

[81] M. Freindorf, E. Kraka, and D. Cremer. A Comprehensive Analysis of Hydrogen Bond Interactions Based on Local Vibrational Modes. *Int. J. Quant. Chem.*, 112:3174–3187, 2012. 6, 26

[82] R. Kalescky, W. Zou, E. Kraka, and D. Cremer. Vibrational Properties of the Isotopomers of the Water Dimer Derived from Experiment and Computations. *Aust. J. Chem.*, 67:426, 2014. 6

[83] Y. Tao, W. Zou, J. Jia, W. Li, and D. Cremer. Different Ways of Hydrogen Bonding in Water - Why Does Warm Water Freeze Faster than Cold Water? *J. Chem. Theory Comput.*, 13:55–76, 2017. 6

[84] Y. Tao, W. Zou, and E. Kraka. Strengthening of Hydrogen Bonding With the Push-Pull Effect. *Chem. Phys. Lett.*, 685:251–258, 2017. 6

[85] Marek Freindorf, Yunwen Tao, Daniel Sethio, Dieter Cremer, and Elfi Kraka. New Mechanistic Insights into the Claisen Rearrangement of Chorismate - A Unified Reaction Valley Approach Study. *Mol. Phys.*, 117:1172–1192, 2018. 6, 7

[86] Małgorzata Z. Makoś, Marek Freindorf, Daniel Sethio, and Elfi Kraka. New Insights into Fe–H_2 and Fe–H^- Bonding of a [NiFe] Hydrogenase Mimic – A Local Vibrational Mode Study. *Theor. Chem. Acc.*, 138:76, 2019. 6

[87] Siying Lyu, Nassim Beiranvand, Marek Freindorf, and Elfi Kraka. Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleosides. *J. Phys. Chem. A*, 123:7087–7103, 2019. 6, 12, 27, 30

[88] V. Oliveira, E. Kraka, and D. Cremer. The Intrinsic Strength of the Halogen Bond: Electrostatic and Covalent Contributions Described by Coupled Cluster Theory. *Phys. Chem. Chem. Phys.*, 18:33031–33046, 2016. 6

[89] V. Oliveira, E. Kraka, and D. Cremer. Quantitative Assessment of Halogen Bonding Utilizing Vibrational Spectroscopy. *Inorg. Chem.*, 56:488–502, 2016. 6

[90] V. Oliveira and D. Cremer. Transition from Metal-Ligand Bonding to Halogen Bonding Involving a Metal as Halogen Acceptor: A Study of Cu, Ag, Au, Pt, and Hg Complexes. *Chem. Phys. Letters*, 681:56–63, 2017. 6

[91] Seth Yannacone, Vytor Oliveira, Niraj Verma, and Elfi Kraka. A Continuum from Halogen Bonds to Covalent Bonds: Where Do λ^3 Iodanes Fit? *Inorganics*, 7:47, 2019. 6

[92] Vytor P. Oliveira, Elfi Kraka, and Francisco B. C. Machado. Pushing 3c-4e Bonds to the Limit: A Coupled Cluster Study of Stepwise Fluorination of First-Row Atoms. *Inorg. Chem.*, 58:14777–14789, 2019. 6

[93] Vytor P. Oliveira, Bruna L. Marcial, Francisco B. C. Machado, and Elfi Kraka. Metal-Halogen Bonding Seen through the Eyes of Vibrational Spectroscopy. *Materials*, 13:55, 2020. 6

[94] D. Setiawan, E. Kraka, and D. Cremer. Strength of the Pnicogen Bond in Complexes Involving Group VA Elements N, P, and As. *J. Phys. Chem. A*, 119:1642–1656, 2014. 6

[95] D. Setiawan and D. Cremer. Super-Pnicogen Bonding in the Radical Anion of the Fluorophosphine Dimer. *Chem. Phys. Letters*, 662:182–187, 2016. 6

[96] V. Oliveira, D. Cremer, and E. Kraka. The Many Facets of Chalcogen Bonding: Described by Vibrational Spectroscopy. *J. Phys. Chem. A*, 121:6845–6862, 2017. 6

[97] V. Oliveira and E. Kraka. Systematic Coupled Cluster Study of Noncovalent Interactions Involving Halogens, Chalcogens, and Pnicogens. *J. Phys. Chem. A*, 121:9544–9556, 2017. 6

[98] Daniel Sethio, Vytor Oliveira, and Elfi Kraka. Quantitative Assessment of Tetrel Bonding Utilizing Vibrational Spectroscopy. *Molecules*, 23:2763, 2018. 6

[99] X. Zhang, H. Dai, H. Yan, W. Zou, and D. Cremer. B-H π Interaction: A New Type of Nonclassical Hydrogen Bonding. *J. Am. Chem. Soc.*, 138:4334–4337, 2016. 6

[100] W. Zou, X. Zhang, H. Dai, H. Yan, D. Cremer, and E. Kraka. Description of an Unusual Hydrogen Bond Between Carborane and a Phenyl Group. *J. Organometal. Chem.*, 856:114–127, 2018. 6

[101] Richard M. Badger. A Relation Between Internuclear Distances and Bond Force Constants. *J. Chem. Phys.*, 2(3):128–131, 1934. 6

[102] William H. Miller, Nicholas C. Handy, and John E. Adams. Reaction Path Hamiltonian for Polyatomic Molecules. *J. Chem. Phys.*, 72(1):99–112, 1980. 6

[103] Bruce C Garrett and Donald G Truhlar. Variational transition state theory. In *Theory and Applications of Computational Chemistry*, pages 67–87. Elsevier, 2005. 7

[104] Javier González, Xavier Giménez, and Josep Maria Bofill. A reaction path–liouville approach to the rate constant for polyatomic chemical reactions. *Physical Chemistry Chemical Physics*, 4(13):2921–2926, 2002. 7

[105] Z. Konkoli, E. Kraka, and D. Cremer. Unified Reaction Valley Approach Mechanism of the Reaction $CH_3 + H_2 \rightarrow CH_4 +$ H. *J. Phys. Chem. A*, 101:1742–1757, 1997. 7

[106] Z. Konkoli, D. Cremer, and E. Kraka. Diabatic ordering of vibrational normal modes in reaction valley studies. *J. Comput. Chem.*, 18:1282–1294, 1997. 7

[107] E. Kraka. Reaction Path Hamiltonian and the Unified Reaction Valley Approach. *WIREs: Comput. Mol. Sci.*, 1:531–556, 2011. 7

[108] Elfi Kraka and Dieter Cremer. Dieter Cremer's Contribution to the Field of Theoretical Chemistry. *Int. J. Quantum Chem.*, 119:e25849, 2019. 7

[109] Elfi Kraka, Wenli Zou, Yunwen Tao, and Marek Freindorf. Exploring the Mechanism of Catalysis with the Unified Reaction Valley Approach (URVA) - A Review. *Catalysts*, 10:691, 2020. 7

[110] W. Zou, T. Sexton, E. Kraka, M. Freindorf, and D. Cremer. A New Method for Describing the Mechanism of a Chemical Reaction Based on the Unified Reaction Valley Approach. *J. Chem. Theory Comput.*, 12:650–663, 2016. 7

[111] E. Kraka and D. Cremer. Computational Analysis of the Mechanism of Chemical Reactions in Terms of Reaction Phases: Hidden Intermediates and Hidden Transition States. *Acc. Chem. Res.*, 43:591–601, 2010. 7

[112] D. Cremer, A. Wu, and E. Kraka. The Mechanism of the Reaction FH + $H_2C=CH_2$ $\rightarrow H_2C - CFH_3$. Investigation of Hidden Intermediates with the Unified Reaction Valley Approach. *Phys. Chem. Chem. Phys.*, 3:674–687, 2001. 7

[113] M. Freindorf, T. Sexton, E. Kraka, and D. Cremer. The Mechanism of the Cycloaddition Reaction of 1,3-Dipole Molecules with Acetylene - An Investigation with the Unified Reaction Valley Approach. *Theor. Chem. Acc.*, 133:1423, 2013. 7

[114] T. M. Sexton, M. Freindorf, E. Kraka, and D. Cremer. A Reaction Valley Investigation of the Cycloaddition of 1,3-Dipoles with the Dipolarophiles Ethene and Acetylene: Solution of a Mechanistic Puzzle. *J. Phys. Chem. A*, 120:8400–8418, 2016. 7

[115] M. Castiñeira Reis, C. S. López, E. Kraka, D. Cremer, and O. N. Faza. Rational Design in Catalysis: A Mechanistic Study of β-Hydride Eliminations in Gold(I) and Gold(III) Complexes Based on Features of the Reaction Valley. *Inorg. Chem.*, 55:8636–8645, 2016. 7

[116] M. Freindorf, D. Cremer, and E. Kraka. Gold(I)-Assisted Catalysis - A Comprehensive View on the [3,3]-Sigmatropic Rearrangement of Allyl Acetate. *Mol. Phys.*, 116:611–630, 2017. 7

[117] D. Cremer and J. A. Pople. General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.*, 97:1354–1358, 1975. 8, 12, 27

[118] H J Geise, W J Adams, and L S Bartell. Electron diffraction study of gaseous tetrahydrofuran. *Tetrahedron*, 25(15):3045–3052, 1969. 8

[119] Donald Allan McQuarrie. *Statistical Thermodynamics*. Harper & Row New York, 1973. 8

[120] Peter Luger and Jürgen Buschmann. Twist conformation of tetrahydrofuran in the crystal. *Angew. Chem. Int. Ed.*, 22(5):410–410, 1983. 8

[121] W I F David, R M Ibberson, T J S Dennis, J P Hare, and K Prassides. Structural phase transitions in the fullerene c_{60}. *Europhys. Lett.*, 18(3):219–225, 1992. 8

[122] Seong Jun Han and Young Kee Kang. Pseudorotation in heterocyclic five-membered rings: Tetrahydrofuran and pyrrolidine. *J. Mol. Struct.*, 369(1):157–165, 1996. 8

[123] A Wu and D Cremer. New approach for determining the conformational features of pseudorotating ring molecules utilizing calculated and measured nmr spin-spin coupling constants. *J. Phys. Chem. A*, 107(11):1797–1810, 2003. 8, 12, 27

[124] Joseph B Lambert, John J Papay, Shakil A Khan, Katharine A Kappauf, and Elaine S Magyar. Conformational analysis of five-membered rings. *J. Am. Chem. Soc.*, 96(19):6112–6118, 1974. 9

[125] Hans Otto Kalinowski, Stefan Berger, and Siegmar Braun. ^{13}*C-NMR-Spektroskopie [13 hoch C-NMR-Spektroskopie]: 200 Tabellen*. Thieme, 1984. 9

[126] D. Cremer. Calculation of puckered rings with analytical gradients. *J. Phys. Chem.*, 94:5502–5509, 1990. 9, 12, 27

[127] Jeng Da Chai and Martin Head-Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.*, 10(44):6615–6620, 2008. 9

[128] Jeng Da Chai and Martin Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128(8):084106, 2008. 9

[129] Warren J Hehre, Robert Ditchfield, and John A Pople. Self-consistent molecular orbital methods. xii. further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.*, 56(5):2257–2261, 1972. 9

[130] RHWJ Ditchfield, W J_ Hehre, and John A Pople. Self-consistent molecular-orbital methods. ix. an extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.*, 54(2):724–728, 1971. 9

[131] Timothy Clark, Jayaraman Chandrasekhar, Günther W Spitznagel, and Paul Von Ragué Schleyer. Efficient diffuse function-augmented basis sets for anion calculations. iii. the 3-21+g basis set for first-row elements, li–f. *J. Comput. Chem.*, 4(3):294–301, 1983. 9

[132] Michael J Frisch, John A Pople, and J Stephen Binkley. Self-consistent molecular orbital methods 25. supplementary functions for gaussian basis sets. *J. Chem. Phys.*, 80(7):3265–3269, 1984. 9

[133] Jürgen Gräfenstein and Dieter Cremer. Efficient density-functional theory integrations by locally augmented radial grids. *J. Chem. Phys.*, 127(16):164113, 2007. 9

[134] D. Cremer. RING - A Coordinate Transformation Program for Evaluating the Degree and Type of Puckering of a Ring Compound. *Quantum Chemical Program Exchange*, pages 1–8, 1975. 9, 12, 27

[135] D. Cremer and J. A. Pople. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. XXIII. Pseudorotation in Saturated Five-Membered Ring Compounds. *J. Am. Chem. Soc.*, 97:1358–1367, 1975. 12, 27

[136] D Cremer. Theoretical determination of molecular structure and conformation. xi. the puckering of oxolanes. *Isr. J. Chem.*, 23(1):72–84, 1983. 12, 27

[137] H Essén and D Cremer. On the relationship between the mean plane and the least-squares plane of an n-membered puckered ring. *Acta Crystallogr. B*, 40(4):418–420, 1984. 12, 27

[138] D. Cremer and E. Kraka. Chemical Bonds without Bonding Electron Density? Does the Difference Electron-Density Analysis Suffice for a Description of the Chemical Bond? *Angew. Chem. Int. Ed.*, 23:627–628, 1984. 14

[139] D. Cremer and E. Kraka. A Description of the Chemical Bond in Terms of Local Properties of Electron Density and Energy. *Croatica Chem. Acta*, 57:1259–1281, 1984. 14

[140] E. Kraka and D. Cremer. Chemical Implication of Local Features of the Electron Density Distribution. In *Theoretical Models of Chemical Bonding. The Concept of the Chemical Bond*, volume 2, page 453. Z.B. Maksic, ed., Springer Verlag, Heidelberg, 1990. 14

[141] Nassim Beiranvand, Marek Freindorf, and Elfi Kraka. Hydrogen Bonding in Natural and Unnatural Base Pairs - Explored with Vibrational Spectroscopy. *Molecules*, 26:2268–12268–22, 2021. 16

[142] Jens Thar and Barbara Kirchner. Hydrogen bond detection. *J. Phys. Chem. A*, 110(12):4229–4237, 2006. 16

[143] Christoph A. Bauer. How to model inter- and intramolecular hydrogen bond strengths with quantum chemistry. *J Chem Inf Model*, 59(9):3735–3743, 2019. 16

[144] Fatemeh Ghasemi, Ali Reza Rezvani, Khaled Ghasemi, Ibrahim Abdul Razak, and Mohd Mustaqim Rosli. Synthesis, characterization and crystal structure of a new organic salt of antidiabetic drug metformin resulting from a proton transfer reaction. *J. Mol. Struct.*, 1193:310 – 314, 2019. 16

[145] Natalia Rojas-Valencia, Sara Gómez, Sebastian Montillo, Marcela Manrique-Moreno, Chiara Cappelli, Cacier Hadad, and Albeiro Restrepo. Evolution of bonding during the insertion of anionic ibuprofen into model cell membranes. *J. Phys. Chem. B*, 124(1):79–90, 2020. 16

[146] Afshan Mohajeri and Fatemeh Fadaei Nobandegani. Detection and evaluation of hydrogen bond strength in nucleic acid base pairs. *J. Phys. Chem. A*, 112(2):281–295, 2008. 16

[147] Linsen Pei and James M. Farrar. Ion imaging study of reaction dynamics in the n^+ + ch_4 system. *Int. J. Chem. Phys.*, 137(15):154312, 2012. 20

[148] Conor A Nixon, Alexander E Thelen, Martin A Cordiner, Zbigniew Kisiel, Steven B Charnley, Edward M Molter, Joseph Serigano, Patrick GJ Irwin, Nicholas A Teanby, and Yi-Jehng Kuan. Detection of cyclopropenylidene on titan with alma. *The Astronomical Journal*, 160(5):205, 2020. 23

[149] A.E. Reed, L.A. Curtiss, and F. Weinhold. Intermolecular Interactions from A Natural Bond Orbital, Donor-Acceptor Viewpoint. *Chem. Rev.*, 88:899–926, 1988. 24

[150] F. Weinhold and C. R. Landis. *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective.* Cambridge University Press, 2003. 24

[151] E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, C. R. Landis, and F. Weinhold. NBO6, 2013. Theoretical Chemistry Institute, University of Wisconsin, Madison. 24

[152] R.F.W. Bader. *Atoms in Molecules: A Quantum Theory.* Clarendron Press, Oxford, 1995. 24

[153] Richard F. W. Bader. The quantum mechanical basis of conceptual chemistry. *Monatshefte fÃŒr Chemie*, 136(6):819–854, 2005. 24

[154] Paul L A Popelier. On quantum chemical topology. In *Applications of Topological Methods in Molecular Chemistry*, pages 23–52. Springer International Publishing, 2016. 24

[155] Chérif F. Matta and Russell J. Boyd. An introduction to the quantum theory of atoms in molecules. In *The Quantum Theory of Atoms in Molecules*, pages 1–34. Wiley-VCH Verlag GmbH & Co. KGaA, 2007. 24

[156] Chrif F. Matta, Jess Hernndez-Trujillo, Ting-Hua Tang, and Richard F. W. Bader. Hydrogen hydrogen bonding: a stabilizing interaction in molecules and crystals. *Chem. Europ. J.*, 9(9):1940–1951, 2003. 24

Appendix A

Paper on Hydrogen Bonding in Deoxyribonucleosides

Siying Lyu and Nassim Beiranvand and Marek Freindorf and Elfi Kraka, Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleosides,

J. Phys. Chem. A (2019) 7087–710

Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleosides

Siying Lyu,[‡] Nassim Beiranvand,[‡] Marek Freindorf, and Elfi Kraka*[iD]

Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, 3215 Daniel Ave, Dallas, Texas 75275-0314, United States

S *Supporting Information*

ABSTRACT: The Cremer−Pople ring puckering analysis and the Konkoli−Cremer local mode analysis supported by the topological analysis of the electron density were applied for the first comprehensive analysis of the interplay between deoxyribose ring puckering and intramolecular H-bonding in 2′-deoxycytidine, 2′-deoxyadenosine, 2′-deoxythymidine, and 2′-deoxyguanosine. We mapped for each deoxyribonucleoside the complete conformational energy surface and the corresponding pseudorotation path. We found only incomplete pseudorotation cycles, caused by ring inversion, which we coined as pseudolibration paths. On each pseudolibration path a global and a local minimum separated by a transition state were identified. The investigation of H-bond free deoxyribonucleoside analogs revealed that removal of the H-bond does not restore the full conformational flexibility of the sugar ring. Our work showed that ring puckering predominantly determines the conformational energy; the larger the puckering amplitude, the lower the conformational energy. In contrast no direct correlation between conformational energy and H-bond strength was found. The longest and weakest H-bonds are located in the local minimum region, whereas the shortest and strongest H-bonds are located outside the global and local minimum regions at the turning points of the pseudolibration paths, i.e., H-bonding determines the shape and length of the pseudolibration paths. In addition to the H-bond strength, we evaluated the covalent/electrostatic character of the H-bonds applying the Cremer−Kraka criterion of covalent bonding. H-bonding in the puric bases has a more covalent character whereas in the pyrimidic bases the H-bond character is more electrostatic. We investigated how the mutual orientation of the CH_2OH group and the base influences H-bond formation via two geometrical parameters describing the rotation of the substituents perpendicular to the sugar ring and their tilting relative to the ring center. According to our results, rotation is more important for H-bond formation. In addition we assessed the influence of the H-bond acceptor, the lone pair (N, respectively O), via the delocalization energy. We found larger delocalization energies corresponding to stronger H-bonds for the puric bases. The global minimum conformation of 2′-deoxyguanosine has the strongest H-bond of all conformers investigated in this work with a bond strength of 0.436 which is even stronger than the H-bond in the water dimer (0.360). The application of our new analysis to DNA deoxyribonucleotides and to unnatural base pairs, which have recently drawn a lot of attention, is in progress.
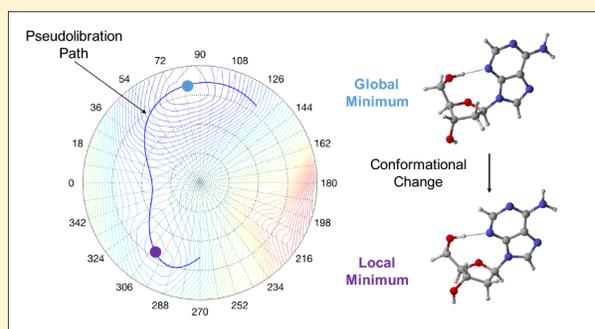
INTRODUCTION

Deoxyribonucleosides and their analogs are important building blocks for the synthesis of antiviral drugs and anticancer agents,[1−3] and the production of deoxyribonucleotides for polymerase chain reactions.[4−6] They can serve as biological probes.[7] Certain bacteria, e.g., the Gram-positive soil bacterium *Bacillus subtilis*, can utilize deoxyribonucleosides via the catabolic deoxyribonucleoside pathway as a sources of carbon and energy.[8] The reverse deoxyriboaldolase enzyme (DERA)[9] can drive the deoxyribonucleoside synthesis showing great promise for industrial-scale production of deoxyribonucleosides.[10,11] The DERA pathway has recently attracted a lot of attention as a possible alternative to the production of

deoxyribonucleotides via ribonucleotide reductases (RNRs).[9,12] Therefore, understanding more about deoxyribonucleosides and their synthesis via the DERA pathway offers an important gateway to new mechanistic insights into this pathway, which could help us to understand the suggested DERA production of deoxyribonucleotides shedding more light onto the origin of DNA.[12] One important ingredient along this route is to understand the conformational flexibility of deoxyribonucleosides, which is determined by a complex

interplay of ring puckering of the deoxyribose sugar ring and intramolecular hydrogen bonding (H-bonding). So far, conformational studies of deoxyribonucleosides have focused mostly on either the ring puckering or the H-bonding aspect. In 1972, Altona and Sundaralingam[13] introduced a procedure describing both the puckering angle and puckering amplitude of a five-membered ring via torsional angles. However, the definition of the ring puckering amplitude expressed via an angle does not allow the appropriate description of the ring pseudorotation process. In this process, the maximum puckering amplitude is rotating around the edge of the ring without raising a substantial potential energy change. Such a movement can be best described with a puckering amplitude defined as displacement from a reference. Ring puckering and pseudorotation of nucleosides based on the Altona and Sundaralingam approach were able to describe the molecular properties of specific deoxyribose sugar ring conformations; however, they failed to describe the changes of molecular properties along the pseudorotation path in a smooth form.[2,14−17] Other studies focused on the influence of the intramolecular hydrogen bonds (H-bond)s on the conformational of deoxyribonucleosides.[18−23] They investigated predominantly the internal H-bonds between the base and the CH_2OH group of the sugar ring. Other internal H-bonds resulting from a rotation of the base unit were not considered in a systematic fashion. Studies of the correlation between ring puckering and H-bonding have been rarely reported, even when the conformational energy surface (CES) was explored,[24] or when global reactivity parameters were derived.[2]

Therefore, the way that deoxyribose ring puckering affects the formation of internal H-bonding and vice versa is not fully understood yet. To understand the driving force of the conformational flexibility of deoxyribonucleosides it is important (i) to quantitatively assess both, ring puckering and H-bonding, and (ii) to investigate the relationship between these two essential structural features. We used in this work as an efficient tool a combination of the Cremer−Pople *ring puckering analysis*[25−30] and the *local mode analysis* of Konkoli and Cremer,[31−34] which describes the conformational process at the quantum mechanical level. In particular, the following objectives were included in our study:

- To exploit the CES, to determine the pseudorotation path and the location of the global and local minima on the CES for the deoxyribonucleosides **dC**, **dA**, **dT**, and **dG**; see Figure 1.

- To investigate how the conformational energy, the puckering amplitude, and the H-bond properties such as distance and bond strength change along the pseudorotation path of deoxyribose ring and how these changes are connected; as an important step to understand the interplay between sugar ring puckering and H-bonding.

- To monitor the change of the covalent character of the H-bonds along the pseudorotation path via the topological analysis of the electron density.

The paper is structured in the following way. In the second section the methods used in this work are described as well as computational details. The third section presents the results and discussion. Conclusions and an outlook are made in the final section.
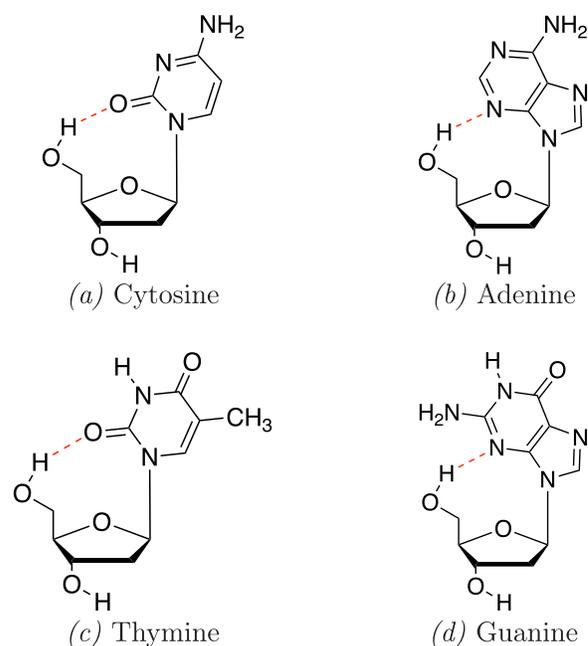
Figure 1. Deoxyribonucleosides investigated in this work: (a) 2′-deoxycytidine (cytosine), **dC**, (b) 2′-deoxyadenosine (adenine), **dA**, (c) thymidine (thymine), **dT**, (d) 2′-deoxyguanosine (guanine), **dG**. The red dashed line indicates the internal H-bond between the base and the CH_2OH substituent of the sugar ring.

■ METHODOLOGIES

In this section, the important tools applied in this work are introduced, the Cremer−Pople ring puckering analysis,[25−30] the local mode analysis of Konkoli and Cremer,[31−34] and the topological analysis of the electron density $\rho(r)$ via Bader's quantum theory of atoms in molecules (QTAIM) approach.[35,36] Then the computational details of this study are described.

Cremer−Pople Ring Puckering Analysis. Cremer and Pople[25] suggested to span the conformational space of a puckered N-membered ring with $N − 3$ puckering coordinates. The $N − 3$ puckering coordinates can be split up into pairs of pseudorotational coordinates $\{q_m, \phi_m\}$ ($m = 2, 3, ..., N − 3$) for odd-membered rings and an additional puckering amplitude $q_{N/2}$ describing ring inversion for even-membered rings. The puckering amplitude q_m describes the degree of ring puckering of the mth puckering mode. The puckering angle ϕ_m defines ring puckering mode.[37,38] Cremer and Pople[25] introduced a mean plane as reference, e.g. the plane of the planar ring. The out-of-plane coordinates z_j, describing the displacement perpendicular to the mean plane can be determined by the $N − 3$ puckering coordinate pairs $\{q_m, \phi_m\}$.[25,37,39−41] For a five-membered ring with ring atoms $j = 1, ..., 5$, the out-of-plane displacements z_j for any ring conformation are given by eq 1:[25,42]

$$z_j = \left(\frac{2}{5}\right)^{1/2} q_2 \cos\left[\frac{4\pi(j - 1)}{5} + \phi_2\right]$$

$$\text{for } j = 1, ..., 5 \text{ and } \phi_2 \in [0; 2\pi] \tag{1}$$

where the displacement z_j coordinates are normalized according to
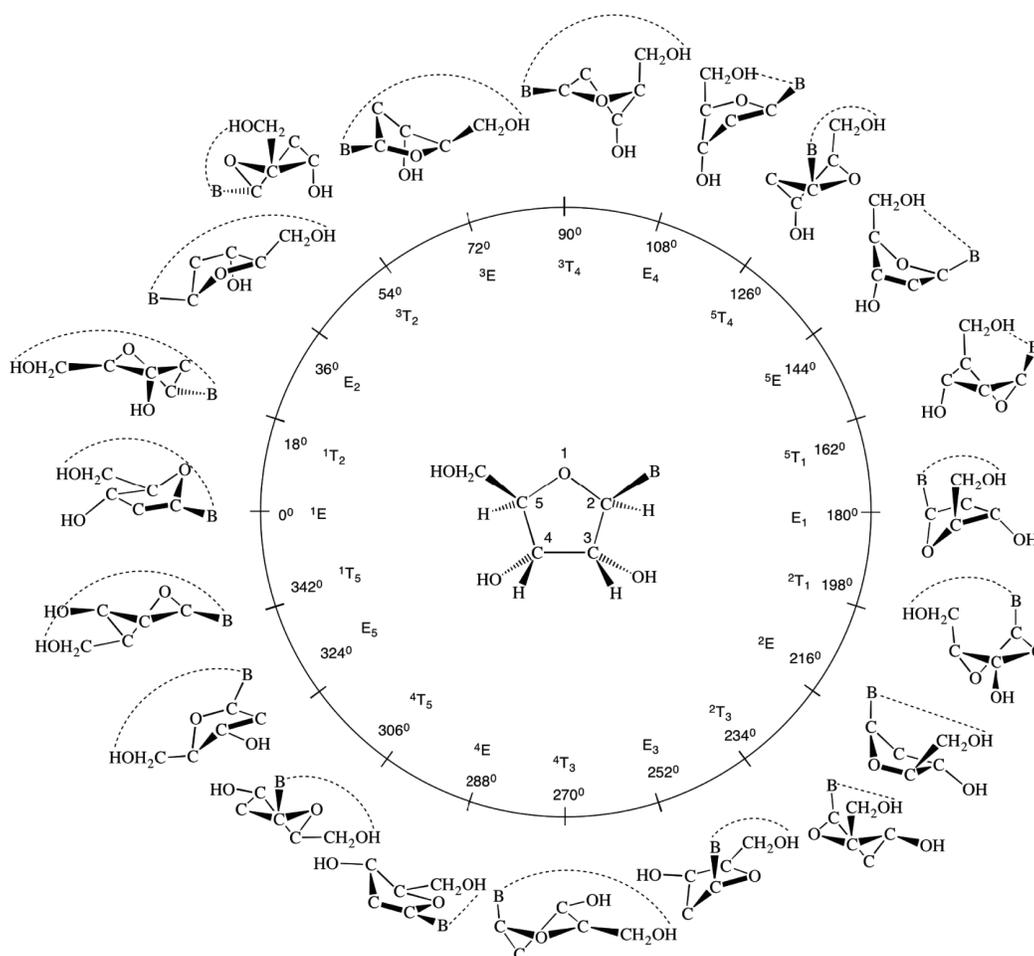
Figure 2. Pseudorotational cycle ($\phi_2 = 0° \rightarrow 360°$) of the deoxyribonucleoside sugar-ring represented by 20 conformers, (10 envelope (E) forms, $\phi_2 = (0 + 360k)/10$ and 10 twist (T) forms $\phi_2 = (18 + 360k)/10$, where $k = 0, 1, 2, ..., 9$). The planar form is located at the center ($q_2 = 0$ Å). The symbol B represents the base of the deoxyribonucleoside. The dashed line indicates the internal H-bond between the base and the CH_2OH substituent of the sugar ring.

$$\sum_{j=1}^{5} z_j^2 = q_2^2 \qquad (2)$$

Using puckering coordinates, the full set of $3N - 6$ independent Cartesian coordinates of any puckered N-membered ring can be uniquely determined via $N - 3$ puckering coordinates (q and ϕ), $N - 3$ bond angles, and N bond lengths. Besides the out-of-plane displacement z_j, the displacements in the perpendicular x_j and y_j directions can also be specified. First, the z_j coordinates are calculated according to eq 1 (or similar formulas for $N > 5$);[39,42] then, the N bond lengths and $N - 3$ bond angles are projected onto the mean plane of the ring. Finally, the projected ring is partitioned into segments, for which the displacements x_j and y_j are calculated according to a procedure described by Cremer.[25]

Although there is an infinite number of ring conformations located on a pseudorotation cycle for the sugar ring of deoxyribonucleosides, it is sufficient to investigate a representative number of conformations, as shown in Figure 2. Referring to previous work on description of tetrahydrofuran,[42] which is an analog to the five-membered sugar ring in deoxyribose, a subset of 20 ring conformations were optimized using a mixed set of puckering and internal coordinates. As shown in Figure 2, there

are 10 envelope (E) forms located at $\phi_2 = (0 + 360k)/10$ (for $k = 0, 1, 2, ..., 9$) and 10 twist forms located at $\phi_2 = (18 + 360k)/10$ (for $k = 0, 1, 2, ..., 9$). The internal H-bonds between the base and CH_2OH group of the sugar ring are indicated by a dashed line. The ring atom, which lies above the mean plane, is noted in front of the conformational symbol E/T as superscript, whereas a ring atom, which lies below the mean plane, is noted after the symbol E/T as subscript. All deoxyribonucleosides considered in this work possess C_1 symmetry; therefore, the CES must also have C_1 symmetry.

There are two major advantages of using ring puckering coordinates.[43,44] First, the geometry of the sugar ring for any given value of q_2 and ϕ_2 can be optimized, even if this conformation does not occupy a stationary point on the CES. This would not be feasible by a description of the ring with Cartesian or internal coordinates.[30] For substituted ring systems such as deoxyribonucleosides, any conformer located on the CES can be uniquely described by a mixed set of puckering coordinates for the ring and internal/Cartesian coordinates for the substituents, and a physically meaningful pseudorotation path as a function of ϕ_2 can be obtained. The second advantage is that any property P of a puckered ring, such as energy, geometry, dipole moment, charge distribution, vibrational frequencies, or magnetic properties, can be expressed as a

Fourier series of the puckering coordinates. In the case of a five-membered ring, P takes the following general form:[42]

$$P(q_2, \phi_2) = \sum_{k=0}^{\infty} [P_k^c(q_2)\cos(k\phi_2) + P_k^s(q_2)\sin(k\phi_2)] \quad (3)$$

where the Fourier coefficients P_k^c and P_k^s are in turn expressed as power series in the puckering amplitude q_2:[42]

$$P_k(q_2) = \sum_{l=0}^{\infty} P_{kl}q_2^l \quad (4)$$

Depending on the symmetry of the ring, eq 3 can be simplified by selecting trigonometric terms. Because the puckering amplitude as well as the coefficients P_{kl} change independent of the phase angle ϕ_2, the puckering amplitude q_2 the Fourier coefficients can be treated as normal coefficients for further simplification. Thus, the property P can be expressed as a pure function of ϕ_2, which also holds for the change of a property P along the pseudorotation path. For C_1 symmetric systems eq 3 can be simplified to eq 5, which was applied in this work:

$$P(\phi_2) = \sum_{k=0}^{\infty} (A_k\cos(k\phi_2) + B_k\sin(k\phi_2)) \quad (5)$$

which according to the C_1 symmetry of the deoxyribonucleosides **dC**, **dA**, **dT**, and **dG**, investigated in this work leads to

$$P(\phi_2) = A_0 + \sum_{i=1}^{\infty} \left(A_i\cos(i\phi_2) + \sum_{j=1}^{\infty} B_j\sin(j\phi_2) \right) \quad (6)$$

Experimentally, for a free or slightly hindered pseudorotor molecule, only a property $\langle P \rangle$ averaged over all pseudorotational modes can be measured. Focusing exclusively on the large amplitude pseudorotational mode of the ring, the calculated property $\langle P \rangle$ averaged over the pseudorotational motion can be determined, once the functional form of eq 6 is known, according to[42]

$$\langle P \rangle = \int_b^a \rho(\phi_2)P(\phi)\, d\phi_2 \quad (7)$$

where a and b are the boundary of pseudorotation. If $a \neq 2\pi$ and $b \neq 0$, the pseudorotation is incomplete, which we call pseudolibration in reference to physics calling an incomplete rotation a libration.

The conformational probability distribution $\rho(\phi_2)$ can be defined as a Boltzmann distribution:

$$\rho(\phi_2) = \frac{e^{-[V(\phi_2)-V(0)]/RT}}{\int_b^a e^{-[V(\phi_2)-V(0)]/RT}\, d\phi_2} \quad (8)$$

where $V(\phi_2)$ is the electronic potential energy of the conformer at ϕ_2 on the pseudorotation path and $V(0)$ is the electronic potential energy of planar form. The conformational probability distribution $\rho(\phi_2)$ can be treated as a property P and, therefore, can be defined using eq 9:

$$\rho(\phi_2) = C_0 + \sum_{i=1}^{\infty} \left(C_i\cos(i\phi_2) + \sum_{j=1}^{\infty} D_j\sin(j\phi_2) \right) \quad (9)$$

The conformational probability distribution $\rho(\phi_2)$ reflects the most likely conformers of a puckered ring on the CES.

Local Mode Analysis. The normal $(3N - L)$ vibrational modes of an N atomic molecule (L is the number of translational

and rotational motions of the molecule) contain important electronic structure information and, therefore, should be well suited as a measure of bond strength. However, it is difficult to decode this information into individual atom−atom interactions (e.g., those resulting in bonding) because normal vibrational modes are generally delocalized due to the coupling of the motions of the atoms within the molecule. There are two different coupling mechanisms between vibrational modes as a consequence of the fact that there is a kinetic and a potential contribution to the energy of a vibrational mode, as reflected in the corresponding Euler−Lagrange equations $L(\mathbf{q}, \dot{\mathbf{q}})$ for a vibrating molecule given by eq 10 in internal coordinates $\mathbf{q}$:[31−34]

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \underbrace{\frac{1}{2}\dot{\mathbf{q}}^{\dagger}\mathbf{G}^{-1}\dot{\mathbf{q}}}_{\text{mass coupling}} - \underbrace{\frac{1}{2}\dot{\mathbf{q}}^{\dagger}\mathbf{F}^q\mathbf{q}}_{\text{electronic coupling}} \quad (10)$$

The electronic coupling between the vibrational modes is reflected by the off-diagonal elements of the force constant matrix $\mathbf{F}^q$. By solving the Euler−Lagrange equations for a vibrating molecule, the basic equation of vibrational spectroscopy, e.g. the Wilson equation,[45] is obtained:

$$\mathbf{F}^q\mathbf{D} = \mathbf{G}^{-1}\mathbf{D}\Lambda \quad (11)$$

with the diagonal matrix Λ collecting the vibrational eigenvalues $\lambda_\mu = 4\pi^2c^2\omega_\mu$, where ω_μ represents the harmonic vibrational frequency of mode $\mathbf{d}_\mu$ given in reciprocal centimeters, c is the speed of light, and $\mu = (1\cdots N - L)$. Solution of eq 11, e.g. diagonalizing the Wilson equation leads to the diagonal force constant $\mathbf{K}$ given in normal coordinates $\mathbf{Q}$ which is free of electronic coupling:

$$\mathbf{K} = \mathbf{D}^{\dagger}\mathbf{F}^q\mathbf{D} \quad (12)$$

Kinematic coupling or mass coupling is still present when the electronic coupling is eliminated by solving the Wilson equation. In 1998, Konkoli and Cremer[31−34] determined for the first time local, mass-decoupled vibrational modes $\mathbf{a}_i$ directly from normal vibrational modes $\mathbf{d}_\mu$ by solving the mass-decoupled Euler−Lagrange equations. The subscript i specifies an internal coordinate q_i and the local mode is expressed in terms of normal coordinates $\mathbf{Q}$ associated with force constant matrix $\mathbf{K}$ of eq 12. Konkoli and Cremer showed that this is equivalent to requiring an adiabatic relaxation of the molecule after enforcing a local displacement of atoms by changing a specific internal coordinate as, e.g., a bond length (*leading parameter principle*).[31] The local modes are unique and the local counterparts of the normal vibrational modes. They can be based on either calculated or experimentally determined vibrational frequencies via[31−34,46−48]

$$\mathbf{a}_i = \frac{\mathbf{K}^{-1}\mathbf{d}_i^{\dagger}}{\mathbf{d}_i\mathbf{K}^{-1}\mathbf{d}_i^{\dagger}} \quad (13)$$

To each local mode $\mathbf{a}_i$ a corresponding local mode frequency ω_i^a, local mode mass $G_{i,i}^a$, and a local force constant k_i^a can be defined.[31] The local mode frequencies can be uniquely connected to the normal-mode frequencies via an adiabatic connection scheme.[33,34] The local mode frequency ω_i^a is defined by

$$(\omega_i^a)^2 = \frac{G_{i,i}^a k_i^a}{4\pi^2c^2} \quad (14)$$

and the corresponding local mode force constant k_i^a by

(a) Cytosine

(b) Adenine

(c) Thymine

(d) Guanine
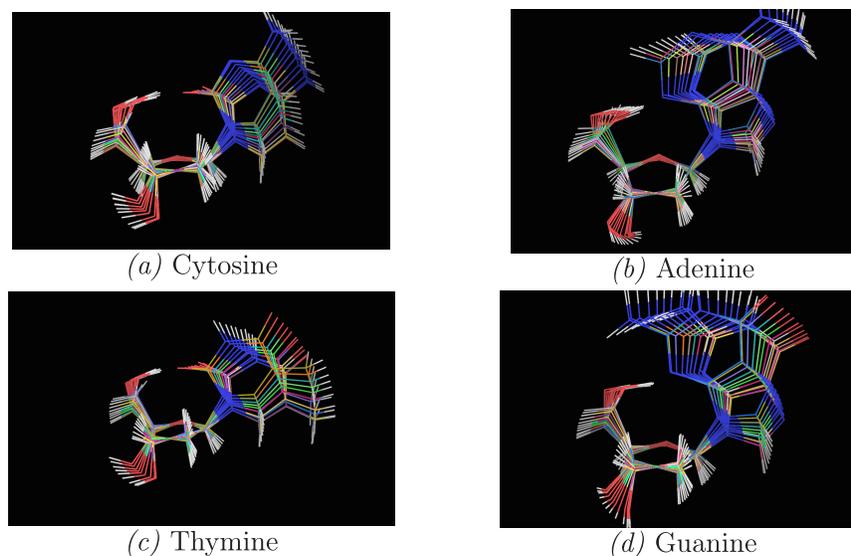
Figure 3. Illustration of the 20 representative path points for each of the four deoxyribonucleosides (a) **dC**, (b) **dA**, (c) **dT**, and (d) **dG** calculated by constrained optimizations using a fixed phase angle ϕ_2 and increments of 18°, e.g. ϕ_2 = 0, 18, 36, ..., 342°.

$$k_i^a = \mathbf{a}_i^\dagger \mathbf{K} \mathbf{a}_i \tag{15}$$

Local mode force constants, contrary to normal mode force constants, have the advantage of being independent of the choice of the coordinates used to describe the molecule in question and in contrast to vibrational frequencies they are independent of the atomic masses. They are extremely sensitive to differences in the electronic structure (e.g., caused by changing a substituent), and they capture only electronic effects directly associated with the intrinsic strength of the atom−atom interaction leading to a chemical bond as recently shown by Zou and Cremer.[49] Therefore, the local vibrational force constants provide a unique tool for assessing the strength of a chemical bond via vibrational spectroscopy, which has recently been applied to characterize strong and weak covalent bonds,[49−55] for the definition of new electronic parameters such as the generalized Tolman parameter[56−58] or a new aromaticity index.[48,59−61] In particular, weak chemical interactions including hydrogen, halogen, and pnicogen bonding were quantitatively assessed for the first time.[62−74]

In this work, we have used the local H-bond stretching frequencies ω^a and the local stretching force constants k^a for the characterization of the strength of the intramolecular H-bonds of the deoxyribonucleosides **dC**, **dA**, **dT**, and **dG**.

It is convenient to base the comparison of the bond strength of a series of molecules on a bond strength order (BSO) n rather than on a comparison for local force constant values. Both are connected via a power relationship according to the generalized Badger rule derived by Cremer and co-workers:[52,75]

$$\text{BSO } n = a(k^a)^b \tag{16}$$

The constants a and b in eq 16 can be determined via two reference values and the requirement that for a zero force constant the BSO n is zero. For H-bonds, $n = 1$ for the FH bond and $n = 0.5$ for the FH bond in the $[\text{F}\cdots\text{H}\cdots\text{F}]^-$ anion were used,[68] leading to $a = 0.515$ and $b = 0.291$, (calculated at the ωB97X-D/6-31++G(d,p) level of theory). According to eq 16 the OH bond in H_2O has a BSO n value of = 0.966. Therefore, we scaled the reference values, so that the BSO n of the OH bond in H_2O is 1.

QTAIM and NBO Analysis. Bader's quantum theory of atoms in molecules (QTAIM) presents the theoretical framework for identifying and characterizing chemical bonds and weak chemical interactions via the study of the topological features of the total electron density $\rho(r)$.[76,77] In this work we used QTAIM as a complementary tool to the local mode analysis to determine the covalent/electrostatic character of internal the H-bonds via the Cremer−Kraka criterion[78,79] of covalent bonding. According to this criterion a covalent bond between two atoms A and B is defined by the following two conditions: (i) *Necessary condition* The existence of a bond path and bond critical point $\mathbf{r}_c = c$ between A and B. (ii) *Sufficient condition* The energy density $H(\mathbf{r}_c) = H_c$ is smaller than zero. $H(\mathbf{r})$ is defined as

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) \tag{17}$$

where $G(\mathbf{r})$ is the kinetic energy density and $V(\mathbf{r})$ is the potential energy density. The negative $V(\mathbf{r})$ corresponds to a stabilizing accumulation of density whereas the positive $G(\mathbf{r})$ corresponds to depletion of electron density.[79] As a result, the sign of H_c indicates which term is dominant.[80] If $H_c < 0$, the interaction is considered covalent in nature, whereas $H_c > 0$ is indicative of electrostatic interactions.

We complemented the QTAIM analysis with the natural bond orbital (NBO) of Weinhold and co-workers[81,82] in order to obtain further insights into the intermolecular orbital interactions, particularly the charge transfer from the nitrogen/oxygen lone pair to the adjacent σ^* of the O−H moiety stabilizing the H-bond. Within the NBO framework the delocalization energy $\Delta E_{del}(\text{HB})$ associated with the electron delocalization between the nitrogen/oxygen lone pair (donor) and the σ^* of O−H moiety (acceptor) can be estimated via the charge transfer from the localized (donor) Lewis-type orbital (L) into the empty (acceptor) non-Lewis orbital (NL)[83,84]

$$E^2 = \Delta E_{del}(\text{HB}) = \frac{q_L F(\text{L, NL})^2}{\epsilon_{NL} - \epsilon_L} \tag{18}$$

where q_L is the orbital and ϵ_{NL} are the diagonal elements of the NBO Fock matrix corresponding to orbital energies of the donor orbital L and acceptor orbital NL, and $F(\text{L, NL})$ is the off-

Table 1. Energies, Geometries, Vibrational, and Electron Density Data for Deoxyribonucleosides dA, dG, dC, and dT along Their Pseudolibration Paths[a]

ϕ_2 [deg]	q_2 [Å]	ΔE [kcal/mol]	ΔG [kcal/mol]	R(HB) [Å]	k^a(HB) [mdyn/Å]	ω^a(HB) [1/cm]	BSO n (HB)	ΔE_{del}(HB) [kcal/mol]	$\omega_\mu(\omega^a$, no., %) [1/cm]	ρ_c [e/Å³]	H_c [Hartree/Å³]	H_c/ρ_c [Hartree/e]
						dA						
270	0.256	−0.11(4.64)	−0.99(3.73)	1.866	0.226	638	0.372	20.4	130(5, 51.8)	0.232	−0.003	−0.011
288	0.304	−1.31(3.44)	−1.91(2.81)	1.909	0.152	524	0.353	18.7	130(5, 30.0), 114(4, 29.6)	0.211	−0.001	−0.006
306	0.272	−1.49(3.26)	−2.33(2.40)	1.969	0.138	499	0.330	15.4	112(4, 72.8)	0.185	−0.001	−0.005
324	0.214	−1.24(3.51)	−2.27(2.45)	1.971	0.144	510	0.329	15.6	106(4, 83.7)	0.184	−0.001	−0.004
342	0.173	−1.17(3.58)	−1.04(3.68)	1.951	0.156	531	0.337	17.2	100(4, 89.9)	0.192	−0.001	−0.004
0	0.165	−1.31(3.43)	−1.17(3.56)	1.941	0.151	523	0.341	18.2	97(4, 90.5)	0.197	−0.001	−0.004
18	0.186	−1.67(3.08)	−1.41(3.31)	1.948	0.159	536	0.338	18.2	99(4, 92.2)	0.194	−0.001	−0.003
36	0.234	−2.39(2.36)	−1.97(2.75)	1.959	0.164	544	0.333	17.9	100(4, 89.8)	0.190	−0.001	−0.004
54	0.291	−3.53(1.22)	−3.96(0.76)	1.945	0.135	493	0.339	19	98(4, 47.9), 76.3 (3, 27.9)	0.196	−0.001	−0.005
72	0.325	−4.52(0.15)	−4.64(0.09)	1.885	0.216	624	0.363	23.3	111(4, 38.7), 179(7, 28.5)	0.225	−0.002	−0.009
90	0.342	−4.68(0.07)	−4.77(−0.05)	1.826	0.249	670	0.390	28.3	182(7, 64.7)	0.258	−0.005	−0.021
108	0.343	−3.69(1.05)	−3.61(1.12)	1.784	0.326	767	0.409	32	194(7, 93.5)	0.285	−0.010	−0.034
126	0.326	−1.69(3.06)	−2.66(2.06)	1.799	0.163	543	0.411	31	6(1, 74.4)	0.288	−0.011	−0.037
						dG						
270	0.267	−0.76(3.76)	−1.40(3.49)	1.819	0.238	655	0.377	22.7	160(7, 45.6), 128(5, 23.8)	0.262	−0.007	−0.028
288	0.308	−2.22(2.31)	−2.80(2.09)	1.862	0.187	581	0.351	20.7	162(7, 38.1), 118(4, 27.0)	0.239	−0.005	−0.020
306	0.28	−2.45(2.08)	−3.53(1.36)	1.914	0.118	462	0.323	17.5	105(4, 46.0), 69 (3, 36.9)	0.212	−0.003	−0.014
324	0.23	−2.14(2.39)	−4.94(−0.04)	1.917	0.138	498	0.321	17.8	101(4, 54.4)	0.209	−0.002	−0.011
342	0.191	−1.96(2.57)	−1.73(3.17)	1.898	0.171	556	0.331	19.8	101(4, 83.0)	0.218	−0.002	−0.011
0	0.181	−2.00(2.53)	−1.80(3.09)	1.882	0.155	528	0.340	21.6	94(4, 78.1)	0.226	−0.003	−0.011
18	0.199	−2.28(2.24)	−1.90(3.00)	1.881	0.207	611	0.340	22.4	101(4, 85.4)	0.226	−0.002	−0.010
36	0.249	−2.95(1.58)	−2.56(2.34)	1.895	0.162	541	0.333	22	6(4, 65.7)	0.219	−0.002	−0.009
54	0.306	−3.98(0.55)	−4.09(0.81)	1.879	0.223	635	0.341	23.7	117(4, 47.2), 76 (3, 41.1)	0.228	−0.002	−0.011
72	0.336	−4.53(0.00)	−4.91(−0.01)	1.827	0.229	644	0.372	28.3	75(3, 50.6)	0.257	−0.005	−0.020
90	0.334	−3.78(0.74)	−3.99(0.91)	1.768	0.327	768	0.410	34.5	186(8, 48.3), 79 (3, 35.4)	0.296	−0.011	−0.039
108	0.304	−1.79(2.73)	−2.09(2.81)	1.732	0.38	828	0.436	38.4	194(8, 41.1), 199(9, 27.3)	0.323	−0.019	−0.057
						dC						
252	0.112	−0.17(2.95)	−1.73(1.36)	1.794	0.174	559	0.385	11.6	16(1, 56.9)	0.218	0.008	0.038
270	0.177	−0.85(2.27)	−1.71(1.38)	1.824	0.193	588	0.368	10	122(5, 65.6)	0.205	0.006	0.032
288	0.228	−1.61(1.51)	−1.96(1.13)	1.878	0.179	566	0.339	8	123(5, 73.3)	0.181	0.004	0.021
306	0.234	−1.85(1.27)	−2.38(0.70)	1.934	0.117	457	0.312	6.5	108(4, 83.4)	0.160	0.002	0.011
324	0.201	−1.51(1.61)	−2.21(0.88)	1.937	0.114	452	0.310	6.7	106(4, 93.4)	0.158	0.002	0.012
342	0.158	−1.12(2.00)	−1.10(1.99)	1.895	0.167	546	0.331	8.2	106(4, 73.2)	0.172	0.004	0.023
0	0.139	−0.91(2.21)	−0.63(2.46)	1.873	0.172	555	0.342	9.2	104(4, 79.0)	0.180	0.005	0.030
18	0.152	−0.87(2.26)	−0.63(2.46)	1.869	0.171	553	0.344	9.5	104(4, 82.1)	0.182	0.006	0.031
36	0.215	−1.15(1.98)	−0.79(2.30)	1.889	0.169	551	0.333	9	104(4, 89.9)	0.174	0.005	0.026
54	0.297	−2.13(0.99)	−2.47(0.62)	1.887	0.166	545	0.334	9.2	106(4, 84.6)	0.175	0.004	0.025
72	0.34	−3.06(0.07)	−3.08(0.01)	1.825	0.213	618	0.367	11.8	159(6, 40.6), 116(4, 33.9)	0.202	0.008	0.039
90	0.348	−2.72(0.40)	−2.64(0.45)	1.764	0.298	731	0.402	15.3	185(8, 50.5)	0.231	0.012	0.050
108	0.319	−0.88(2.24)	−1.35(1.74)	1.74	0.289	719	0.418	17.1	189(8, 91.1)	0.245	0.013	0.053
						dT						
270	0.19	−0.79 (3.10)	−1.44 (2.31)	1.868	0.277	576	0.348	8.7	111(5, 34.4), 191(9, 21.1)	0.214	0.009	0.043
288	0.25	−1.80(2.08)	−2.29(1.46)	1.933	0.113	451	0.321	7.1	110(5, 38.4), 94 (3, 34.5)	0.162	0.002	0.011
306	0.26	−2.24(1.65)	−2.67(1.08)	2.016	0.092	405	0.290	5.4	106(4, 88.1)	0.135	0.000	0.001
324	0.228	−1.96(1.93)	−2.62(1.13)	2.028	0.089	400	0.286	5.4	102(4, 89.8)	0.131	0.000	0.001
342	0.186	−1.58(2.31)	−1.05(2.69)	1.982	0.111	446	0.303	7	102(4, 92.9)	0.143	0.001	0.007
0	0.17	−1.41(2.47)	−1.05(2.69)	1.953	0.113	450	0.313	8.4	97(4, 92.3)	0.152	0.002	0.012
18	0.186	−1.48(2.41)	−0.93(2.81)	1.948	0.127	476	0.316	9.3	100(4, 95.5)	0.153	0.002	0.014
36	0.244	−1.96(1.93)	−1.74(2.01)	1.964	0.105	433	0.309	9.7	90(4, 94.5)	0.148	0.001	0.010
54	0.313	−3.05(0.84)	−3.04(0.71)	1.947	0.137	496	0.316	11.1	104(4, 93.0)	0.154	0.002	0.011

Table 1. continued

ϕ_2 [deg]	q_2 [Å]	ΔE [kcal/mol]	ΔG [kcal/mol]	R(HB) [Å]	k^a(HB) [mdyn/Å]	ω^a(HB) [1/cm]	BSO n (HB)	ΔE_{del}(HB) [kcal/mol]	$\omega_\mu(\omega^a$, no., %) [1/cm]	ρ_c [e/ Å^3]	H_c [Hartree/ Å^3]	H_c/ρ_c [Hartree/e]
						dT						
72	0.35	−3.86(0.02)	−3.82(−0.08)	1.873	0.188	580	0.346	14.8	108(4, 44.4), 151(6, 23.0)	0.182	0.005	0.026
90	0.353	−3.30(0.58)	−3.17(0.58)	1.8	0.277	704	0.379	19.2	174(9, 93.1)	0.214	0.009	0.043
108	0.323	−1.24(2.65)	−1.49(2.26)	1.77	0.314	750	0.394	20.5	191(9, 93.2)	0.230	0.011	0.048

^{a}Phase angles ϕ_2 are in degrees; puckering amplitudes q_2 are in angstroms; ΔE and ΔG indicate relative energy and Gibbs free energy in kilocalories per mole with respect to the planar form as reference, while the values in parentheses are relative to the global minimum as reference; bond length values R(HB) are in angstroms; force constants k^a are in millidyne per angstroms; vibrational frequencies ω^a(HB) are in inverse centimeters; ΔE_{del}(HB) gives the delocalization energy of lone-pair electrons from H-bond acceptor to H-bond donor in kilocalories per mole; $\omega_\mu(\omega^a$, no., %) represents the decomposition of H-bond normal modes to local stretching mode, ω^a is the vibrational frequency, no. is the name of normal mode given in number, and the contribution of normal mode to H-bond local stretching mode is given in percentage; electron densities ρ_c at the bond critical points are in electrons per cubic angstrom; energy densities H_c at the bond critical points are in Hartree per cubic angstrom; energy densities per electron H_c/ρ_c are in Hartree per electron. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

diagonal NBO Fock matrix element between the L and NL orbitals. The NBO analysis was applied to all conformers located on the pseudorotation paths to determine ΔE_{del}(HB) for all internal H-bonds.

Computational Methods. All calculations were performed with the ωB97X-D functional[85,86] and Pople's 6-31++G(d,p) basis set.[87−90] Tight convergence criteria were applied, (SCF iterations $10^{−10}$ Hartree and geometry optimizations $10^{−7}$ Hartree/Bohr) and an ultrafine grid was used for the DFT numerical integration.[91] First, the most stable conformer for each deoxyribonucleoside, e.g. the global minimum, was determined by fully optimizing the structures derived from all possible H-bond combinations using Cartesian coordinates. The pseudorotation path was then calculated in the following way. A puckering analysis was performed for the global minimum to obtain the corresponding phase angle ϕ_2 and the puckering amplitude q_2. Starting from this phase angle 20 representative path points were calculated by constrained optimizations with a fixed phase angle ϕ_2, modified in increments of 18° for e.g. ϕ_2 = 0, 18, 36, ..., 342°, as shown in Figure 3 to determine the pseudorotation path. For the constrained optimizations a mixed set of internal coordinates for the substituents and puckering coordinates for deoxyribose sugar ring was used.

Harmonic frequency calculations were carried out in each case to confirm the optimized conformers located on the pseudorotation path as minima (no imaginary frequency) or transition states (one imaginary frequency). In this regard, the pseudorotation path can be considered as a special reaction path with the phase angle ϕ_2 as reaction coordinate. For the lowest energy conformers being located on the opposite site of global minima, e.g. phase angles ϕ_2 in the range of ≈300°, a full geometry optimization was performed, to identify possible local minima, which were found for all deoxyribonucleosides in this range. For the highest energy conformers, e.g. phase angles ϕ_2 in the range of 330−18°, a full geometry optimization was performed to identify the transition state (TS) for the pseudorotation. The CESs were also mapped via constrained optimizations. For each fixed phase angle ϕ_2, the puckering amplitudes q_2 were constrained to four values, namely 0.1, 0.2, 0.3, and 0.4 Å, and the remainder of the molecule was optimized, leading to 80 data points for each CES. For each conformer on the pseudorotation path, a local mode analysis was performed for the investigation of the H-bond strength. Delocalization energies of the internal H-bonds were calculated for further characterization.[84] These energies reflect the delocalization of lone-pair electrons from H-bond accept or to H-bond donor

The RING Puckering program[26] was used for analysis of the deoxyribose sugar ring conformations. For the geometry optimizations and frequency calculations, a combination of the RING Puckering program and the program package Gaussian09[92] was used. All local mode analyses calculations were carried out with the program package COLOGNE2019.[93] The NBO calculations were carried out with NBO 6.[94] The QTAIM analysis was performed with the AIMALL software.[95]

■ RESULTS AND DISCUSSION

In Table 1, properties of the deoxyribonucleosides **dC**, **dA**, **dT**, and **dG** calculated at different phase angles ϕ_2 are listed. These comprise the puckering amplitude q_2, relative energy ΔE, and free energy ΔG with regard to the planar form and the global minimum, H-bond distance R(HB), force constant k^a(HB), local mode frequency ω^a(HB), and bond strength order BSO n(HB) as well the H-bond delocalization energy ΔE_{del}(HB) and the normal vibration mode(s) ω_μ with the highest H-bond contribution(s). Table 1 also contains the electron density ρ_c (e/ Å^3) and the energy density H_c (Hartree/Å^3) at the bond critical point c. Table 2 summarizes H-bond properties of some

Table 2. Properties of Hydrogen Bonded Reference Complexes R1−R5^a

molecule	k^a	BSO n	R(HB)	ρ_c	H_c	H_c/ρ_c
R1	0.209	0.360	1.905	0.183	−0.001	−0.005
R2	0.214	0.362	1.895	0.176	0.002	0.012
R3	0.044	0.241	2.457	0.071	0.003	0.048
R4	0.116	0.309	1.981	0.162	−0.004	−0.027
R5	0.205	0.358	1.922	0.202	0.000	0.002

^{a}Force constants k^a are in millidyne per angstrom; H-bond lengths R(HB) are in angstroms, ρ_c values are in electrons per cubic angstrom; H_c values are in Hartree per cubic angstrom; energy densities per electron H_c/ρ_c are in Hartree per electron. Reference complexes **R1−R5** are shown in Figure 9. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

reference molecules. In Table 3, geometric features, including R(HB), angles α_1 and α_2, and dihedral angles τ_1 and τ_2 between the base and the CH$_2$OH group forming the intramolecular H-bond of the conformer with the strongest H-bond (ϕ_2 = 108°), and that with the weakest H-bond (ϕ_2 = 324°) are summarized for puckering amplitude values of 0.1, 0.2, 0.3, and 0.4 Å.

CESs and Pseudorotation (Pseudolibration) Paths. The pseudorotation paths on the corresponding CESs of the

Article

Table 3. Geometric Features of Conformers at $\phi_2 = 108°$ (Region of Strongest H-Bonds) and $\phi_2 = 324°$ (Region of Weakest H-Bonds)^a

	q_2 [Å]	r(HB) [Å]	α_1 [deg]	α_2 [deg]	τ_1 [deg]	τ_2 [deg]	$\sum(\alpha_1, \alpha_2)$ [deg]	$\Delta(\tau_1, \tau_2)$ [deg]	ΔE [kcal/mol]
					dA				
$\phi_2 = 108°$	0.1	1.800	23.9	32.2	54.8	−93.5	56.1	38.6	−0.76
	0.2	1.784	21.9	36.4	55.7	−93.4	58.4	37.7	−2.25
	0.3	1.783	20.3	40.0	56.8	−93.4	60.3	36.7	−3.52
	0.4	1.787	18.7	42.9	57.0	−92.2	61.6	35.2	−3.34
$\phi_2 = 324°$	0.1	1.872	25.7	23.7	51.3	−89.1	49.4	37.8	−0.74
	0.2	1.929	25.4	20.0	49.0	−86.5	45.4	37.5	−1.22
	0.3	1.988	24.9	16.8	47.6	−85.3	41.8	37.7	−0.78
	0.4	2.073	23.4	13.6	56.4	−85.6	36.9	29.2	1.21
					dG				
$\phi_2 = 108°$	0.1	1.747	23.2	30.9	64.3	−81.7	54.2	17.4	−0.23
	0.2	1.732	21.4	35.4	64.6	−83.0	56.7	18.4	−1.15
	0.3	1.732	19.7	39.2	65.2	−83.2	58.9	18.0	−1.79
	0.4	1.735	18.0	42.3	65.4	−82.0	60.3	16.6	−0.85
$\phi_2 = 324°$	0.1	1.825	25.1	21.9	60.8	−77.1	47.0	16.3	−1.29
	0.2	1.893	24.7	18.0	57.9	−75.7	42.7	17.8	−2.09
	0.3	2.000	24.3	14.3	55.4	−76.0	38.6	20.6	−1.72
	0.4	2.201	23.6	10.8	54.4	−77.9	34.3	23.5	0.50
					dC				
$\phi_2 = 108°$	0.1	1.759	23.6	30.9	63.6	−97.7	54.5	34.1	0.34
	0.2	1.742	21.1	35.2	64.8	−97.9	56.4	33.1	−0.19
	0.3	1.738	19.1	39.1	64.7	−97.8	58.2	33.1	−0.85
	0.4	1.750	17.3	42.3	63.4	−97.1	59.5	33.6	−0.25
$\phi_2 = 324°$	0.1	1.837	25.5	22.9	59.9	−91.1	48.4	31.2	−1.04
	0.2	1.937	24.8	19.5	58.6	−87.8	44.3	29.2	−1.51
	0.3	2.128	23.8	16.4	58.2	−86.7	40.2	28.6	−0.91
	0.4	2.672	22.1	12.9	62.7	−88.5	35.0	25.8	0.90
					dT				
$\phi_2 = 108°$	0.1	1.789	23.6	31.4	60.8	−98.3	55.0	37.5	0.20
	0.2	1.773	21.2	35.6	62.1	−98.6	56.7	36.5	−0.45
	0.3	1.768	19.2	39.3	61.5	−98.9	58.5	37.4	−1.20
	0.4	1.779	17.5	42.4	60.3	−98.5	59.8	38.2	−0.65
$\phi_2 = 324°$	0.1	1.878	25.5	23.5	56.6	−92.2	49.1	35.6	−1.20
	0.2	1.986	24.9	20.2	54.7	−89.2	45.1	34.5	−1.91
	0.3	2.198	24.1	17.3	54.5	−88.0	41.3	33.5	−1.65
	0.4	2.811	22.4	14.1	59.6	−89.7	36.5	30.1	−0.30

$^a\phi_2 = 108°$ corresponds to the envelope form E$_4$, and $\phi_2 = 324°$ corresponds to the envelope form E$_5$; see Figure 2. $\sum(\alpha_1, \alpha_2)$ is the sum of angles α_1 and α_2. $\Delta(\tau_1, \tau_2)$ is the difference of $|\tau_2| - \tau_1$. ΔE is the conformational energy with regard to the planar form. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

deoxyribonucleosides **dC**, **dA**, **dT**, and **dG** are shown in Figure 4. We found on the pseudorotation paths of all deoxyribonucleosides negative puckering amplitudes in the range of $\phi_2 = 108$ to 270° indicating a ring inversion. As an example, the ring inversion for **dT** is shown in Figure 5 for conformer 1 (green color) with $\phi_2 = 108°$ and $q_2 = +0.323$ Å and conformer 2 (red color) with $\phi_2 = 144°$ and $q_2 = −0.228$ Å. This inversion leads to an incomplete pseudorotation, i.e. the pseudorotation path is not a closed circle as expected (see Figure 2) but an open curve, which we coined as pseudolibration path. As discussed in the Supporting Information for **dA** as an example, the energies beyond the ring inversion point are higher than the energy of the planar form, and as such there is no longer a minimum energy path on the CES. This suggests that when the deoxyribonucleoside reaches the end-point of the path (e.g., ring inversion point) during a thermal motion, it prefers to return back to the energetically lower part of CES via the pseudolibration path instead of entering the high energy regions.

On each of the pseudolibration paths a global minimum in the range of $\phi_2 = 72−90°$ and a local minimum in the range of $\phi_2 = 306°$ separated by a transition state in the range of $\phi_2 = 18−342°$ was identified, as shown in Figure 4. Energy differences ΔE between the global and local minima are 3.24, 2.05, 1.26, and 1.65 kcal/mol, respectively for **dA**, **dG**, **dC** and **dT**; see also Table 1. The puric bases (**dA** and **dG**) process somewhat larger ΔE values compared to the pyrimidic bases (**dC** and **dT**) which is also reflected by the fact that the former possess longer pseudolibration paths, i.e., they are conformationally more flexible. The transition states for the pseudolibration are located at $\phi_2 = 337.7°$, 347.5°, 13.5°, and 4.7° for **dA**, **dG**, **dC**, and **dT**. The corresponding pseudolibration barriers from the global minimum to the local minimum are 3.59, 2.58, 2.26, and 2.48 kcal/mol for **dA**, **dG**, **dC**, and **dT**. The reverse barriers, e.g. from local minimum to global minimum, are only 0.34, 0.53, 1, and 0.84 kcal/mol; see Table 1.

(a) Cytosine

(b) Adenine

(c) Thymine

(d) Guanine
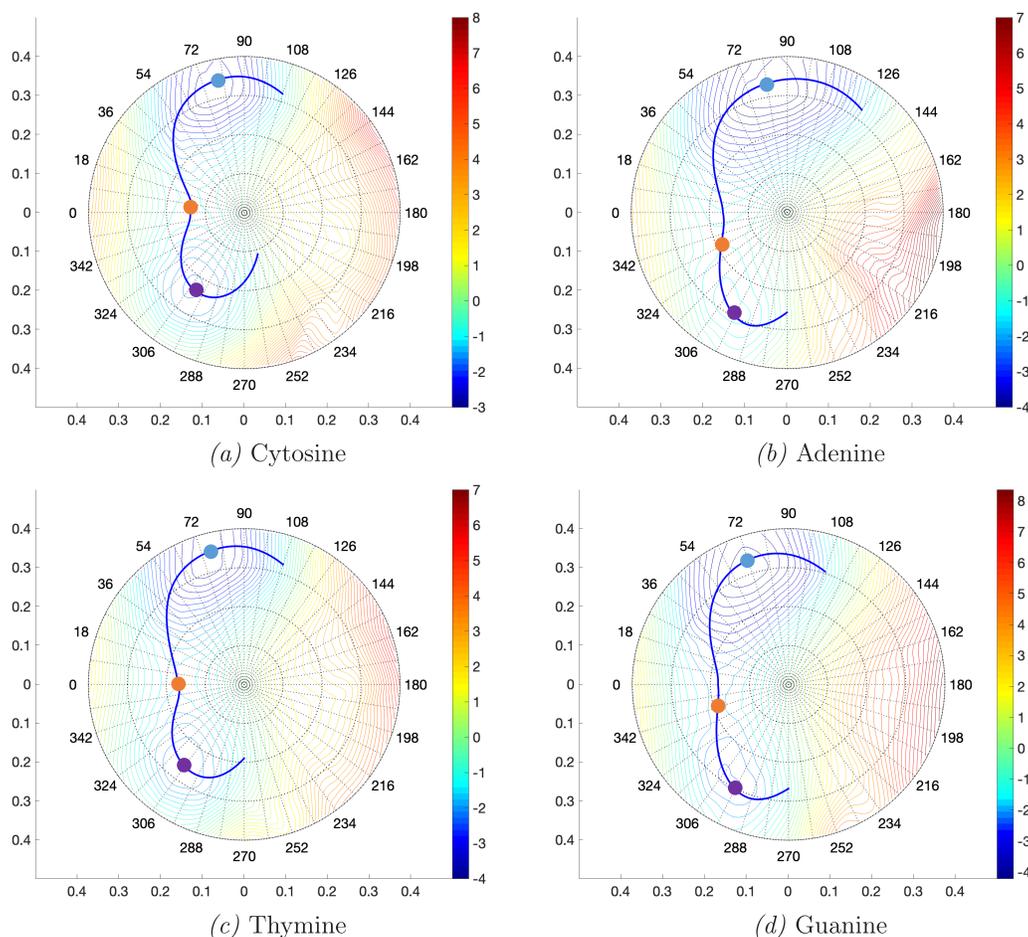
Figure 4. Conformational energy surfaces (CESs) and pseudolibration paths of deoxyribonucleosides **dC** (a); **dA** (b); **dT** (c); and **dG** (d). At the center of each CES, the planar deoxyribose ring form is located. The corresponding energy is used as reference. Both x and y axes show the magnitude of puckering amplitude q_2 in angstroms. The outer radius of the CES represents the maximum puckering amplitude of 0.4 Å. The labels around the CES circle denote the value of the phase angle ϕ_2 in degrees. The solid blue line indicates the pseudolibration path. The color bar represents the energy on the CES in kilocalories per mole relative to the planar form: (*yellow to red regions*) location of conformers higher in energy than the planar form; (*green to blue regions*) location of conformers lower in energy than the planar form. The blue dot in the range 72−90° shows the location of the global, the purple dot near 306° shows the location of the local minimum, and the yellow dot in the range 18−342° shows the transition state between the two forms. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.
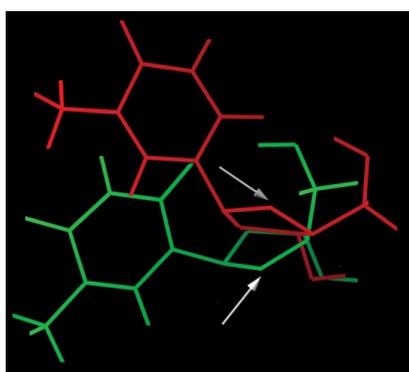
Figure 5. Ring inversion of the sugar ring (marked by silver arrows) of **dT**; conformer 1 is shown in green, $\phi_2 = 108°$ and $q_2 = +0.323$ Å, and conformer 2 is shown in red, $\phi_2 = 144°$ and $q_2 = −0.228$ Å.

More detailed information is obtained by plotting changes of ΔE, q_2, $R(HB)$, and BSO $n(HB)$ as a function of ϕ_2 as shown in Figure 6. The pseudolibration paths can be divided into three

distinct regions; (1) *Global minimum region*: including all conformers n with an energy difference $\Delta E = (E_n − E_{global})$ less than 1.0 kcal/mol. For all deoxyribonucleosides, this region covers ϕ_2 angles in the range of 54−90°. (2) *Local minimum region*: including all conformers m with an energy difference $\Delta E = (E_m − E_{local})$ less than 0.5 kcal/mol. For all deoxyribonucleosides, this region covers ϕ_2 angles in the range of 288−324°. (3) *Transition region*: Including all conformers in between, e.g. ϕ_2 angles in the range of 342−36°; see Figure 6. All conformers located either in the global or minimum region were identified to have no imaginary frequencies. All other conformers in the transition region and the regions outside the global and local minima (ϕ_2 angles in the range of 234−270° and 108−144°) were identified to have one imaginary frequency.

A comparison of Figure 6a and b reveals how the puckering amplitude and the conformational energy are related. All deoxyribonucleosides feature the same deoxyribose ring puckering pattern: the strongest puckering occurs in the global minimum region ($\phi_2 = 72−90°$, q_2 values up to 0.35 Å), followed by the puckering in the local minimum region ($\phi_2 = $

(a) ΔE as function of ϕ_2

(b) q_2 as function of ϕ_2

(c) H-bond distance as function of ϕ_2

(d) H-bond strength as function of ϕ_2
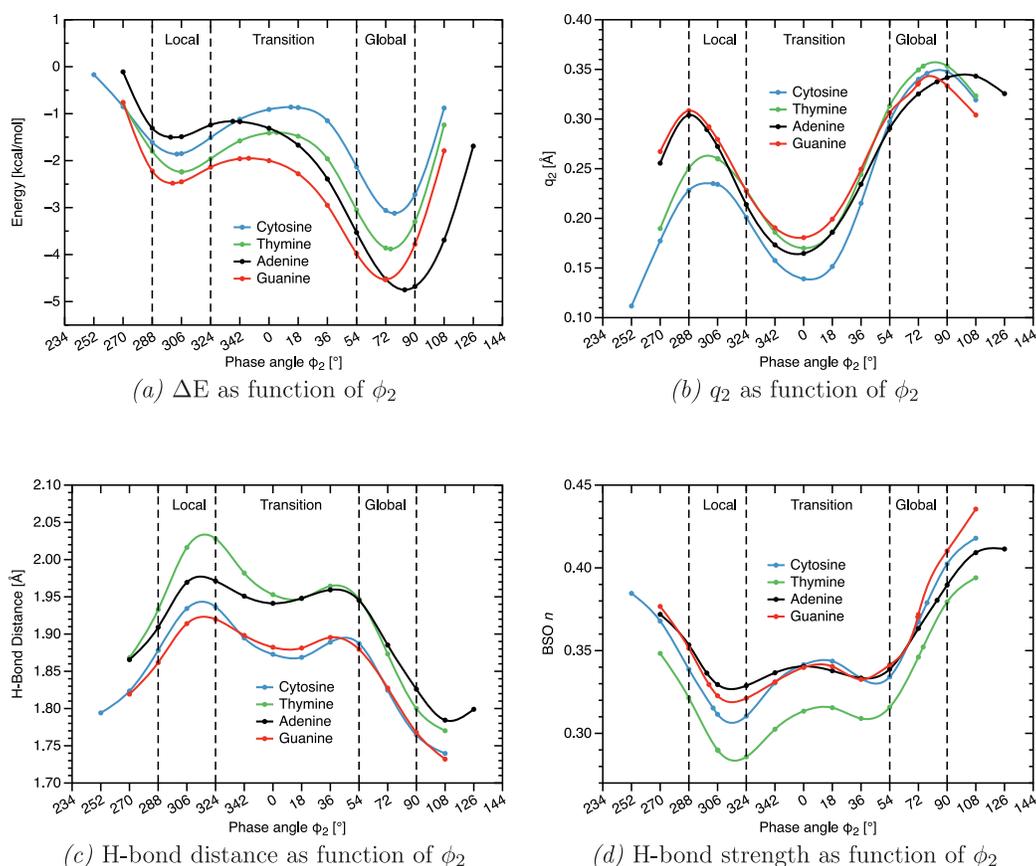
Figure 6. (a) Changes of the conformational energy ΔE with regard to the planar form, (b) puckering amplitude q_2, (c) H-bond distance $R(HB)$, and (d) H-bond strength order BSO n along the pseudolibration path described as a function of the puckering angle ϕ_2. The local minimum, transition, and global minimum regions are denoted by vertical dashed lines. For a definition of these regions, see the text. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

288−306°, q_2 values up to 0.31 Å), while the smallest puckering ($q_2 = 0.15−0.18$ Å) occurs in the transition region in the range of $\phi_2 = 324−54°$, suggesting that the larger the puckering amplitude, the lower the conformational energy. This holds for the pyrimidic bases **dC** and **dT** for which the energy and amplitude curves are perfect mirror images, e.g. smallest amplitude values correspond to highest energy values. The puckering amplitude curves of the puric bases **dA** and **dG** are almost identical in the local energy region, whereas the difference in the conformational energy is up to 1 kcal/mol. This shows that there are additional factors determining the conformational energy, such as constrained flexibility via H-bonding which in turn can be different for the different bases.

Figure 6c and d shows the change of the internal H-bond distance $R(HB)$ and the corresponding H-bond strength BSO n along the pseudolibration path. There is a direct relationship between H-bond length and strength, i.e. the strongest H-bonds are the shortest as also reflected by Figure 7. This is not always true as numerous examples have shown.[96,97] Figure 7 reveals that there is a considerable variation in the H-bond lengths (0.334 Å) leading to H-bond strength differences of 0.15.

There is no obvious correlation between the H-bond length/ strength and the conformational energy change. The longest and weakest H-bonds are found in the local minimum region for all deoxyribonuclesides, with the longest H-bond of 2.028 Å at $\phi_2 = 324°$ for **dT** compared with 1.917 Å at $\phi_2 = 324°$ for **dG**. If the H-bond strength would dominate the conformational energy,
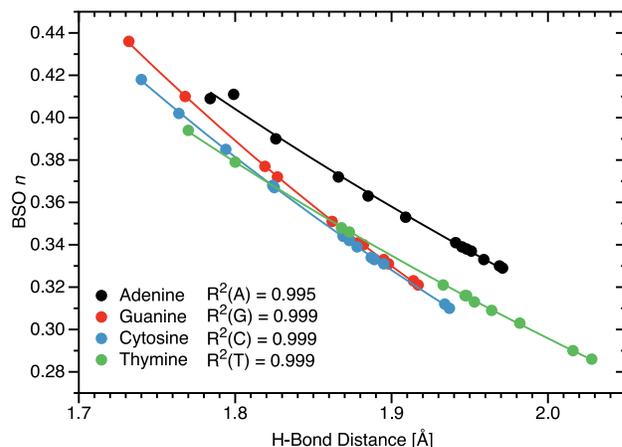
Figure 7. Correlation between H-bond strength and H-bond length. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

according to Figure 6a the longest H-bonds should be located in the transition region and at the end points of the pseudolibration paths. However, we find for all deoxyribonucleosides smallest BSO n values at the end of local minimum region ($\phi_2 = 306−324°$) which increase through the transition region to the global minimum region ($\phi_2 = 54−90°$). However, the shortest and strongest H-bonds are located outside both the global minimum region (1.732 Å at $\phi_2 = 108°$ for **dG**) and the local minimum

region $(1.794$ Å at $\phi_2 = 252°$ for **dC**), i.e. at the turning points of the pseudolibration path, where the deoxyribonuclesides bounce back into the opposite path direction. These results clearly reveal that H-bonding determines the shape and length of the pseudolibration paths.

Internal H-Bond Strength and Covalent/Electrostatic Character. The weakest H-bond (BSO $n = 0.286$) is found for the **dT** conformer at $\phi_2 = 324.0°$ which is still stronger than the H-bond in the formaldehyde dimer **R3**, (BSO $n = 0.241$, see Table 2). The **dG** conformer at $\phi_2 = 108.0°$ has the strongest H-bond (BSO $n = 0.436$)—see Figure 8—which is even stronger
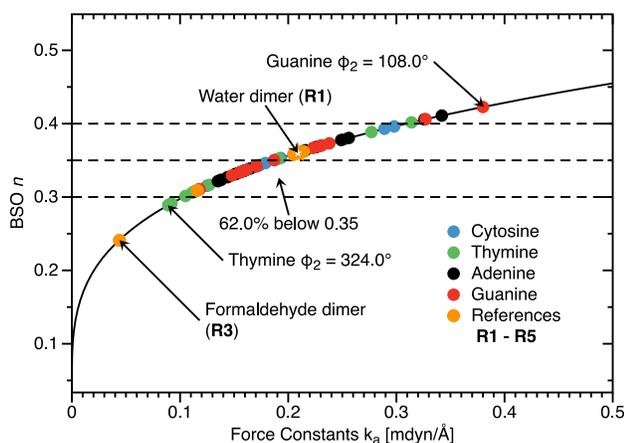
Figure 8. Bond strength BSO n of the H-bonds of all deoxyribonucleoside conformers as a function of the corresponding local stretching force constant k^a as determined via eq 16. For comparison H-bonded complexes **R1**−**R5** defined in Figure 9 are included. Dashed horizontal lines mark BSO n values of 0.3, 0.35, and 0.4, respectively. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

than the H-bond in the water dimer R1 with a BSO n value of 0.360; see Table 2. The percentage of conformers with BSO n values between 0.3 and 0.35 is 62.0%, in the range of the most common H-bond types in water clusters.[68] This clearly shows that the intermolecular H-bonds are of considerable strength and that they are an important feature of deoxyribonucleosides.

In addition to the H-bond strength we evaluated the covalent/electrostatic character of the H-bonds applying the Cremer−Kraka criterion of covalent bonding[78−80,98] described above. As reflected by data in Table 1 H_c/ρ_c values range from -0.057 Hartree/e (covalent character) for **dG** at the $\phi_2 = 108°$ turning point (strongest H-bond, BSO $n = 0.436$) to 0.053 Hartree/e (electrostatic character) for **dC** again at the $\phi_2 = 108°$ turning point (second strongest H-bond, BSO $n = 0.418$). Similar values were also found for the H-bonds of the reference compounds **R1**−**R5** shown in Figure 9 as revealed by the data in Table 2. These H_c/ρ_c magnitudes are typical of H-bonding[99] and small compared with the corresponding values for the OH bond in H_2O (-1.666 Hartree/e) and the H-donor bond in **R1** (-1.689 Hartree/e).[68,100,101]

In Figure 10a normalized energy density H_c/ρ_c values are correlated with the amplitudes q_2 for all conformers of the four deoxyribonucleosides. As expected there is no general relationship because the amplitude q_2 is a more global property reflecting complex geometry changes induced by ring puckering while H_c/ρ_c is taken only at a single point. However, one finds a clear separation into puric and pyrimidic bases caused by the different sign of the corresponding H_c/ρ_c values. In Figure 10b,
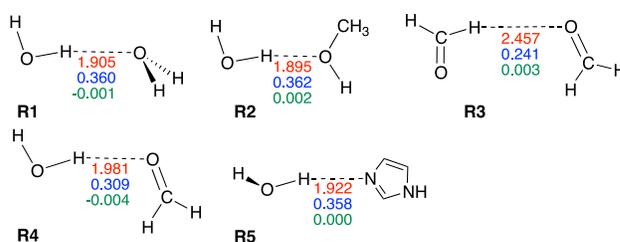

Figure 9. Reference complexes **R1**−**R5**. The intramolecular H-bond distance (Å) is given in red, the corresponding BSO n value is in blue, and energy density H_c (Hartree/Å^3) is in green. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

BSO n values are correlated with H_c/ρ_c. There is no overall relationship between BSO n and H_c/ρ_c for all conformers. However, there is a correlation between BSO n and H_c/ρ_c for the puric bases **dA** and **dG** as well as for the pyrimidic bases **dC** and **dT**. Larger BSO n values correspond to larger H_c/ρ_c magnitudes in both cases.

The finding that **dA** and **dG** conformers possess negative H_c/ρ_c values, i.e. being located on the covalent side of Figure 10, whereas all **dC** and **dT** conformers possess positive H_c/ρ_c, i.e. being located on the electrostatic side of Figure 10, which reflects the influence of the base determining (i) the different H-bond type, $[O−H\cdots N]$ for the puric bases and $[O−H\cdots O]$ bonds for the pyrimidic bases, and (ii) leading to a different electronic environment, i.e. in the more bulky puric bases the electron density can delocalize more than in the pyrimidic bases influencing the lone pair density of the H-bond acceptor atom. This will be discussed in more detail in the following.

Influence of the Orientation of the CH_2OH Group and the Base on H-Bonding. As discussed above (see Figure 7) there is a correlation between H-bond distance and H-bond strength for all four deoxyribonucleosides, but no obvious correlation between the conformational energy and H-bond distance, in particular in the local and transition region, (see Figure 6a and c), which needs further analysis. The H-bond distance depends on the mutual orientation of the CH_2OH substituent (*H-bond donor*) of the C5 sugar ring carbon and the base (*H-bond acceptor*) attached to the C2 sugar ring carbon, which in turn depends on the ring pucker. This mutual orientation is also critical for optimal orbital overlap, i.e. the O−H should point to the lone pair of the H-bond acceptor atom A (A = N for **dA**, **dG**; A = O for **dC**, **dT**) in a way that the angle (OHA) is close to 180°.[101]

For the quantification of this orientation we used two torsion angles, defined in Figure 11; torsion angle τ_1 describing the rotation of the base and torsion angle τ_2 describing the rotation of the CH_2OH group perpendicular to the sugar ring. The dihedral difference $\Delta(\tau_1, \tau_2) = |\tau_2| - \tau_1$ can serve as measure of the distance between H-bond donor and acceptor; i.e. smaller dihedral differences indicate shorter H-bonds and in this way stronger H-bonds. For the global minimum geometries the dihedral differences $\Delta(\tau_1, \tau_2)$ are 24.3°, 1.3°, 13.1°, and 16.3° for **dA**, **dG**, **dC**, and **dT**, respectively. For the local minimum geometries $\Delta(\tau_1, \tau_2)$ are larger with values of 43.1°, 25.1°, 33.8°, and 39.1° for **dA**, **dG**, **dC**, and **dT**, respectively. This reveals that for the global minimum geometries H-bond interactions are stronger. The largest difference between the global and the local minimum $\Delta(\tau_1, \tau_2)$ values are found for **dG** (23.8°), followed by **dT** (22.8°), **dC** (20.7°), and **dA** (18.8°). Obviously, in the local minimum region, i.e. for puckering angles ϕ_2 between 288° and
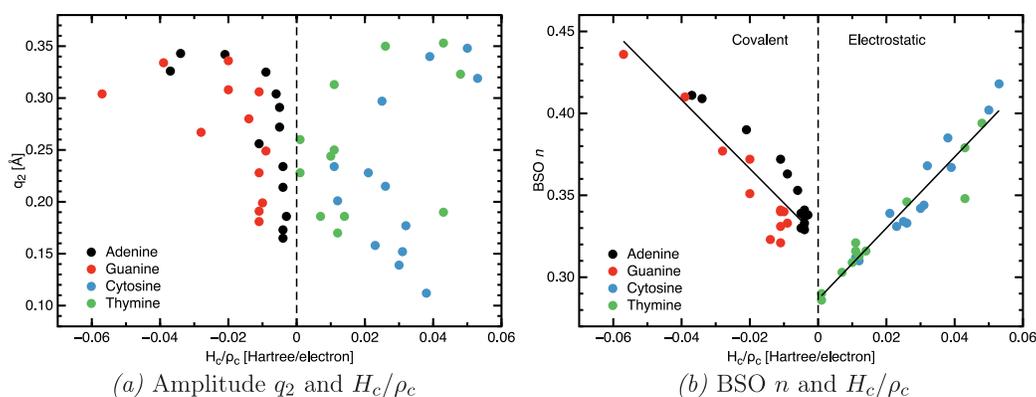
(a) Amplitude q_2 and H_c/ρ_c *(b)* BSO n and H_c/ρ_c

Figure 10. (a) Correlation between amplitude q_2 and normalized energy density H_c/ρ_c. (b) Correlation of BSO n and normalized energy density H_c/ρ_c. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.
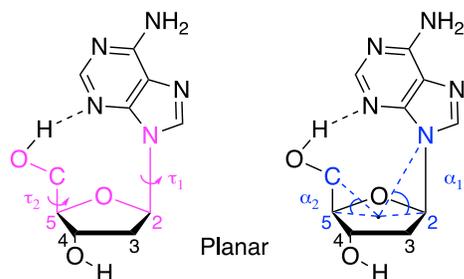

Figure 11. Definition of torsion angles τ_1 and τ_2 (pink color) and angles α_1 and α_2 (blue color) as listed in Table 3.

$324°$ the CH$_2$OH group and the base cannot orient in an optimal way.

Besides a rotational movement the substituents can tilt relative to the main plane of the deoxyribose ring described by angles α_1 and α_2 for the base and the CH$_2$OH group respectively, as shown in Figure 11. For the global minimum geometries the tilting angle α_2 of the CH$_2$OH adapts values of $37.1°$, $34.1°$, $35.6°$, $35.3°$ for **dA**, **dG**, **dC**, and **dT**, respectively. For the local minimum geometries α_2 values are smaller: $17.3°$, $14.6°$, $18.3°$, and $18.3°$ for **dA**, **dG**, **dC**, and **dT**, respectively. The tilting angles α_1 of the base are $15.2°$, $13.3°$, $12.4°$, and $12.1°$ for the global minimum geometries and $29.0°$, $29.1°$, $28.0°$, and $27.8°$ for the local minimum geometries of **dA**, **dG**, **dC**, and **dT**, respectively. The sum $\sum(\alpha_1, \alpha_2)$ of the tilting angles α_1 and α_2 reflects the distance between H-bond donor and acceptor. The larger the sum $\sum(\alpha_1, \alpha_2)$, the smaller the distance between the H-bond donor and acceptor and the stronger is the H-bond. For the puric bases, the differences of $\sum(\alpha_1, \alpha_2)$ between the global and local minimum geometries are $6.1°$ and $3.6°$ for **dA** and **dG**, respectively. The corresponding differences for the pyrimidic bases are somewhat smaller with $1.7°$ and $1.4°$ for **dC** and **dT**, respectively. This indicates that the tilting movement is more pronounced for the puric bases although they are more bulky than their pyrimidic counterparts. Overall, our analysis suggests that side group rotation is more important for the H-bond formation than tilting.

The strength of the H-bond also depends on the orientation and electron density distribution of the lone pair of the H-bond acceptor atom (N, respectively O). Both can be assessed via the delocalization energy ΔE_{del}(HB) as defined in eq 18. The larger ΔE_{del}(HB), the stronger the H-bond.[83,84] As revealed by the data in Table 1 larger ΔE_{del} values along the pseudolibration

paths are found for the puric bases **dA** and **dG** than for the pyrimidic bases **dC** and **dT**. ΔE_{del} values for the pyrimidic bases at the global minima are 32.0 kcal/mol ($\phi_2 = 108°$, $q_2 = 0.343$ Å) for **dA** and 38.4 kcal/mol ($\phi_2 = 108°$, $q_2 = 0.304$ Å) for **dG**, compared to the pyrimidic bases, 17.1 kcal/mol ($\phi_2 = 108°$, $q_2 = 0.319$ Å) for **dC** and 20.5 kcal/mol ($\phi_2 = 108°$, $q_2 = 0.323$ Å) for **dT**. This is in line with the finding that puric bases generally tend to have a more delocalized electron density than pyrimidic bases[102] inducing a more delocalized electron density at the lone pair of the H-bond acceptor atom and in this way leading to stronger H-bonding. This is in line with our finding that H-bonding in the puric bases is of more covalent nature and that the global minimum conformation of **dG** has the strongest H-bond of all conformers investigated in this work with a BSO n value of 0.436.

In order to gain additional insights into these important side group effects and for better comparison, we evaluated torsion and tilting angles for all deoxyribonucleosides for the same two puckering modes, $\phi_2 = 108°$ and $\phi_2 = 324°$ (envelope forms E$_4$ and E$_5$, respectively, see Figure 2), corresponding to strong and weak H-bond regions and for four different amplitudes 0.1, 0.2, 0.3, and 0.4 Å mapping the angle changes along the CES; see Table 3. As suggested by the pseudorotational cycle shown in Figure 2, ring puckering for the $\phi_2 = 108°$ conformation should allow a better interaction between the CH$_2$OH group and the base than that for $\phi_2 = 324°$.

According to the data in Table 3 for puckering mode E$_4$, α_1 is decreasing whereas α_2 is increasing for increasing puckering amplitudes for all deoxynucleosides. These two opposite movements balance each other, so that as a consequence, the $\sum(\alpha_1, \alpha_2)$ is increasing to $61.6°$, $60.3°$, $59.5°$, and $59.8°$ for **dA**, **dG**, **dC**, and **dT**, respectively, reflecting strong H-bonding. It is also noteworthy, that the effect is almost similar for all deoxyribonucleosides quantifying the trends shown in Figure 6. In contrast, for puckering mode E$_5$ both α_1 and α_2 are decreasing for increasing amplitudes, so that the $\sum(\alpha_1, \alpha_2)$ is also decreasing to $36.9°$, $34.3°$, $35°$, and $36.5°$ for **dA**, **dG**, **dC**, and **dT**, respectively, reflecting weak H-bonding. In addition, the decreasing values of both tilting angles indicate that the equatorial position is more favorable when the ring puckers with a larger amplitude. As shown in Table 3 overall changes in $\Delta(\tau_1, \tau_2)$ for increasing puckering amplitudes are somewhat less pronounced. Overall this analysis shows that both the puckering amplitude and the puckering mode have a substantial influence on H-bond formation.
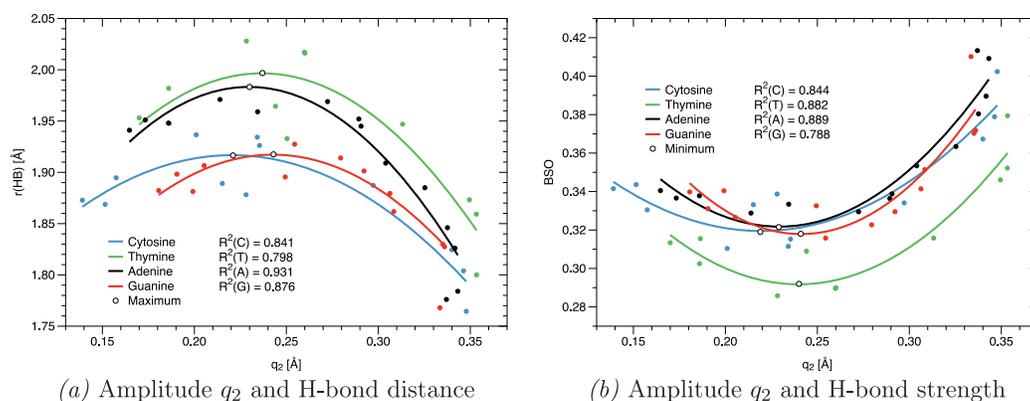
(a) Amplitude q_2 and H-bond distance   *(b)* Amplitude q_2 and H-bond strength

Figure 12. (a) Correlation of puckering amplitude q_2 and H-bond distance, (b) correlation of puckering amplitude q_2 and H-bond strength, expressed as a quadratic function of the puckering amplitude q_2. The color points indicate the original calculated data. The lines represent the corresponding H-bond property as a quadratic function of q_2. The curve maxima (a) and minima (b) are denoted by open cirles. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

Interplay of Ring Puckering and Internal H-Bonding. In the following it will be discussed how the puckering amplitude q_2 is related to the H-bond length/strength. As shown in Figure 12, there is a qualitative quadratic relationship between the two quantities.

An increase of the puckering amplitude q_2 first leads to an increase of the H-bond length (Figure 12a) reaching a maximum value around $q_2 = 0.25$ Å, followed by a decrease of the H-bond length for larger puckering amplitudes. This corresponds to a decrease of the H-bond strength, reaching a minimum value around 0.25 Å, followed by an increase of the bond strength (Figure 12b). This is in line with Figure 6c revealing that at the local minima on the CESs located around the puckering angle $\phi_2 = 306°$ with puckering amplitudes q_2 in the range of 0.23 and 0.27 Å (see Figure 6b) the longest H-bonds are found. After passing through the local minima, the H-bond distances decrease until reaching minimum values beyond the global minima at the turning points of the pseudolibration paths. This implies that the H-bond strength increases along the pseudolibration path from the local minimum, through the transition state, to the global minimum on CES and beyond as a result of ring puckering, clarifying that there is no direct relationship between the conformational energy and the H-bond length/strength, which would have implied longest/weakest H-bonds in the transition region.

Comparison with H-Bond Free Model Analogues. Finally we studied the four H-bond free deoxyribonucleosides analogs **dmC**, **dmA**, **dmT**, and **dmG** shown in Figure 13 in order to determine (i) if the pseudolibration path on the CES surface converts into a full pseudorotation path if the H-bonding is eliminated and (ii) if there is still a global and local minimum separated by a transition state.

The CESs of the H-bond free analogs are shown in Figure 14. For all analogs there exists still a pseudolibration path, which for the pyrimidic bases **dCm** and **dTm** is considerably longer compared with their original counterparts. That means that removal of the H-bond does not restore the full conformational flexibility of the sugar ring. We still find global and local minima separated a transition state on the CESs of **dCm** and **dTm** (see Figure 14a and c). However, without the influence of the internal H-bond, the transition region shrinks from $\phi_2 = 342°-36°$ to $\phi_2 = 342°-18°$ and the position of local and global minima are exchanged. In the case of the puric bases **dAm** and

(a) Cytosine, **dCm**   *(b)* Adenine, **dAm**

(c) Thymine, **dTm**   *(d)* Guanine, **dGm**
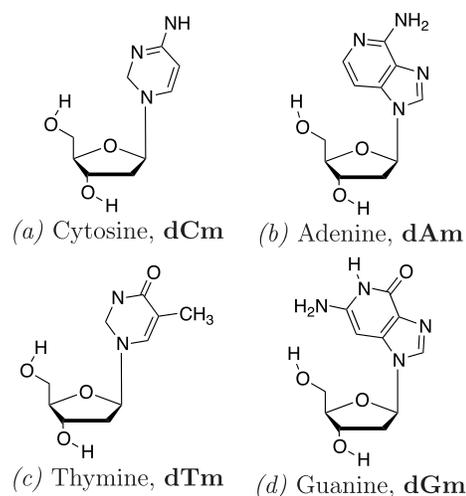
Figure 13. H-bond free deoxyribonucleoside analogs **dCm**, **dAm**, **dTm**, and **dGm**.

dGm only one minimum is found at $\phi_2 = 342°$ and the pseudolibration paths do not become considerably longer. This suggests that H-bonding invokes more strain on the pyrimidic bases, which is line with the above finding that pyrimidic H-bonding is more covalent in nature.

■ CONCLUSIONS AND OUTLOOK

In this study, the Cremer−Pople ring puckering analysis and the Konkoli−Cremer local mode analysis supported by the topological analysis of the electron density were applied to systematically analyze for the first time the interplay between deoxyribose ring puckering and the intramolecular H-bonding in 2′-deoxycytidine (cytosine), **dC**, 2′-deoxyadenosine (adenine), **dA**, thymidine (thymine), **dT**, and 2′-deoxyguanosine (guanine), **dG**. Using Cremer−Pople puckering coordinates, the CES for any substituted ring system such as a deoxyribonucleoside can be determined in analytical form and a physically meaningful pseudorotation path is obtained. Local mode force constants are a unique measure of bond strength. Our work has led to the following conclusions:

1. We found for all four deoxyribonucleosides **dC**, **dA**, **dT**, and **dG** incomplete pseudorotation paths on the CESs
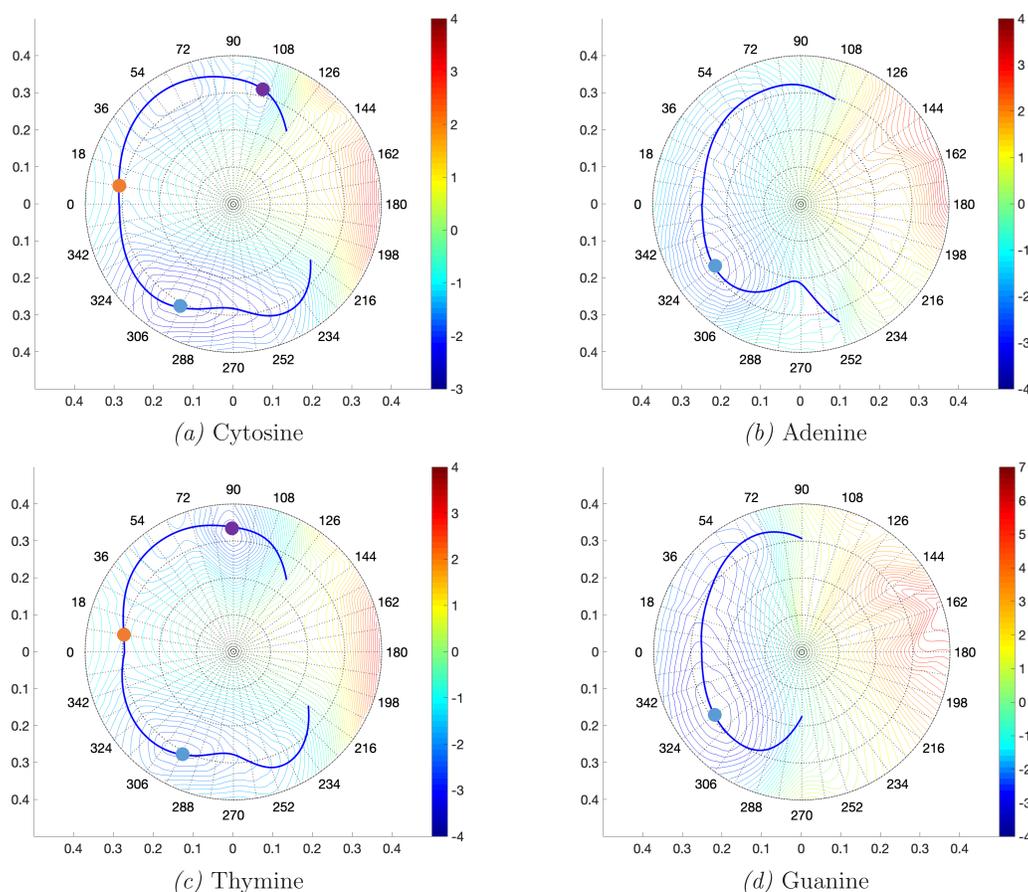
Figure 14. CESs and pseudolibration paths of the H-bond free deoxyribonucleosides **dCm** (a); **dAm** (b); **dTm** (c); and **dGm** (d). At the center of each CES, the planar deoxyribose ring form is located. The corresponding energy is used as reference. Both x and y axes show the magnitude of puckering amplitude q_2 in angstroms. The outer radius of the CES represents the maximum puckering amplitude of 0.4 Å. The labels around CES circle denote the value of the phase angle ϕ_2 in degrees. The solid blue line indicates the pseudolibration path. The color bar represents the energy on the CES in kilocalories per mole relative to the planar form: (*yellow to red regions*) location of conformers higher in energy than the planar form; (*green to blue regions*) location of conformers lower in energy than the planar form. The purple dot in the range 90−108° shows the location of the local, the blue dot near 306° the location of the global minimum, and the yellow dot near 9° the transition state between the two minima for **dCm** and **dTm**. The blue dot at 324° shows the single minimum for **dAm** and **dGm**. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

caused by ring inversion, i.e. the pseudorotation path is not a closed circle but an open curve, which we coined as a pseudolibration path. On each of the pseudolibration paths a global minimum in the range of $\phi_2 = 72°-90°$ and a local minimum in the range of $\phi_2 = 306°$ separated by a transition state in the range of $\phi_2 = 18°-342°$ could be identified. Harmonic frequency calculations identified all conformers in the global and local minimum regions as minimum structures (no imaginary frequency) and that in the transition region as a transition structure (one imaginary frequency). In this regard, the pseudorotation path can be considered as a special reaction path with the phase angle $\phi_2$ as a reaction coordinate. Studies of H-bond free deoxyribonucleoside analogs revealed that removal of the H-bond does not restore the full conformational flexibility of the sugar ring. For all analogs there exists still a pseudolibration path, which for the pyrimidic bases is considerably longer compared with their original counterparts. This reflects that H-bonding invokes more strain on the pyrimidic bases than on the puric bases.

2. There are two major factors determining the conformational flexibility of the deoxyribonucleosides, ring puckering and internal H-bonding. Our work showed that ring puckering plays the dominant role. All deoxyribonucleosides feature the same deoxyribose ring puckering pattern, and strongest puckering occurs in the global minimum region ($\phi_2 = 72°-90°$, q_2 values up to 0.35 Å), followed by the puckering in the local minimum region ($\phi_2 = 288°-306°$, q_2 values up to 0.31 Å), while smallest puckering ($q_2 = 0.15-0.18$ Å) occurs in the transition region in the range of $\phi_2 = 324°-54°$. This suggests that the larger the puckering amplitude, the lower the conformational energy. In contrast no direct correlation between conformational energy and H-bond strength could be found.

3. We quantitatively assessed the H-bond strength of all conformers along the pseudorotation paths via bond strength orders BSO n derived from local vibrational force constants. We found a direct correlation between H-bond strength and H-bond length. Longest and weakest H-bonds were found in the local minimum region for all deoxyribonuclesides, whereas shortest and strongest H-

bonds were found outside both the global minimum region at the turning points of the pseudolibration paths, where the deoxyribonuclesides bounce back into the opposite path direction. These results clearly reveal that H-bonding determines the shape and length of the pseudolibration paths. In addition to the H-bond strength we evaluated the covalent/electrostatic character of the H- bonds applying the Cremer—Kraka criterion of covalent bonding. H-bonding in the puric bases has more covalent character whereas for the pyrimidic bases the character of the H-bond is more electrostatic.

4. The formation of the internal H-bond is affected by the mutual orientation of the CH_2OH group and the base which in turn is depends on the puckering mode. The influence of this orientation on the H-bond formation was quantified via (i) two dihedral angles describing the rotation of the CH_2OH group and the base perpendicular to the sugar ring and (ii) two angles describing the tilting of to the he CH_2OH group and the base relative to the center of the deoxyribose ring. We found that the tilting movement is more pronounced for the puric bases although they are more bulky than their pyrimidic counterparts. Overall, our analysis suggests that side group rotation is more important for H-bond formation. Besides geometric factors the strength of the H-bond also depends on the orientation and electron density distribution of the lone pair of the H-bond acceptor atom (N, respectively O). We assessed both via the delocalization energy $\Delta E_{del}(HB)$. With larger $\Delta E_{del}(HB)$ values corresponding stronger H-bonds were found along the pseudolibration paths for puric bases in line with our finding that H-bonding in the puric bases is of more covalent nature and that the global minimum conformation of **dG** has the strongest H-bond of all conformers investigated in this work with a BSO n value of 0.436, which is even stronger than the H-bond in the water dimer (BSO n = 0.360).

Currently, we are applying our new analysis to the DNA building blocks deoxyribonucleotides, which possess a more complex internal H-bonding pattern caused by the phosphate group, and to the characterization of unnatural base pairs which have recently drawn a lot of attention.[103,104]

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.9b05452.

Cut through the CES of Adenine for a fixed value of the puckering phase angle ϕ_2. 3D plot of the CES for adenine. Geometries of **dA**, **dG**, **dC**, and **dT** at the global minimum, transition state, and local minimum (PDF)

AUTHOR INFORMATION

Corresponding Author
*E-mail: ekraka@smu.edu.

ORCID Ⓞ
Elfi Kraka: 0000-0002-9658-5626

Author Contributions
‡S.L. and N.B. contributed equally to this work.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Riccardi, C.; Musumeci, D.; Irace, C.; Paduano, L.; Montesarchio, D. Ru(III) Complexes for Anticancer Therapy: The Importance of Being Nucleolipidic. *Eur. J. Org. Chem.* **2017**, *2017*, 1100−1119.

(2) Kumar, V.; Kishor, S.; Ramaniah, L. M. Chemical Reactivity Analysis of Deoxyribonucleosides and Deoxyribonucleoside Analogues (NRTIs): A First-Principles Density Functional Approach. *J. Mol. Model.* **2012**, *18*, 3969−3980.

(3) Baker, D.; Chu, C. *Nucleosides and Nucleotides as Antitumor and Antiviral Agents*; Springer, 2013.

(4) Hocek, M. Synthesis of Base-Modified 2′-Deoxyribonucleoside Triphosphates and Their Use in Enzymatic Synthesis of Modified DNA for Applications in Bioanalysis and Chemical Biology. *J. Org. Chem.* **2014**, *79*, 9914−9921.

(5) Kielkowski, P.; Fanfrlík, J.; Hocek, M. 7-Aryl-7-deazaadenine 2′-Deoxyribonucleoside Triphosphates (dNTPs): Better Substrates for DNA Polymerases than dATP in Competitive Incorporations. *Angew. Chem., Int. Ed.* **2014**, *53*, 7552−7555.

(6) Raindlová, V.; Pohl, R.; Šanda, M.; Hocek, M. Direct Polymerase Synthesis of Reactive Aldehyde-Functionalized DNA and Its Conjugation and Staining with Hydrazines. *Angew. Chem., Int. Ed.* **2010**, *49*, 1064−1066.

(7) Morales, J. C.; Kool, E. T. Varied Molecular Interactions at the Active Sites of Several DNA Polymerases: Nonpolar Nucleoside Isosteres as Probes. *J. Am. Chem. Soc.* **2000**, *122*, 1001−1007.

(8) Škerlová, J.; Fábry, M.; Hubálek, M.; Otwinowski, Z.; Řezáčová, P. Structure of the Effector-Binding Domain of Deoxyribonucleoside Regulator DeoR from Bacillus Subtilis. *FEBS J.* **2014**, *281*, 4280−4292.

(9) Salleron, L.; Magistrelli, G.; Mary, C.; Fischer, N.; Bairoch, A.; Lane, L. DERA is the Human Deoxyribose Phosphate Aldolase and is Involved in Stress Response. *Biochim. Biophys. Acta, Mol. Cell Res.* **2014**, *1843*, 2913−2925.

(10) Horinouchi, N.; Ogawa, J.; Kawano, T.; Sakai, T.; Saito, K.; Matsumoto, S.; Sasaki, M.; Mikami, Y.; Shimizu, S. Biochemical Retrosynthesis of 2′-Deoxyribonucleosides from Glucose, Acetaldehyde, and a Nucleobase. *Appl. Microbiol. Biotechnol.* **2006**, *71*, 615−621.

(11) Honda, K.; Maya, S.; Omasa, T.; Hirota, R.; Kuroda, A.; Ohtake, H. Production of 2-deoxyribose 5-phosphate from Fructose to Demonstrate a Potential of Artificial Bio-Synthetic Pathway Using Thermophilic Enzymes. *J. Biotechnol.* **2010**, *148*, 204−207.

(12) Poole, A. M.; Horinouchi, N.; Catchpole, R. J.; Si, D.; Hibi, M.; Tanaka, K.; Ogawa, J. The Case for an Early Biological Origin of DNA. *J. Mol. Evol.* **2014**, *79*, 204−212.

(13) Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. New Description Using the Concept of Pseudorotation. *J. Am. Chem. Soc.* **1972**, *94*, 8205−8212.

(14) Szczepaniak, M.; Moc, J. Conformational Studies of Gas-Phase Ribose and 2-deoxyribose by Density Functional, Second Order PT and Multi-Level Method Calculations: The Pyranoses, Furanoses, and Open-Chain Structures. *Carbohydr. Res.* **2014**, *384*, 20−36.

(15) Taniguchi, T.; Nakano, K.; Baba, R.; Monde, K. Analysis of Configuration and Conformation of Furanose Ring in Carbohydrate and Nucleoside by Vibrational Circular Dichroism. *Org. Lett.* **2017**, *19*, 404−407.

(16) Li, L.; Szostak, J. W. The Free Energy Landscape of Pseudorotation in 3′-5′ and 2′-5′ Linked Nucleic Acids. *J. Am. Chem. Soc.* **2014**, *136*, 2858−2865.

(17) Shishkin, O. V.; Pelmenschikov, A.; Hovorun, D. M.; Leszczynski, J. Molecular Structure of Free Canonical 2′-Deoxyribo-

nucleosides: A Density Functional Study. *J. Mol. Struct.* **2000**, *526*, 329−341.

(18) Wu, R. R.; Yang, B.; Frieler, C. E.; Berden, G.; Oomens, J.; Rodgers, M. T. N3 and O2 Protonated Tautomeric Conformations of 2′-Deoxycytidine and Cytidine Coexist in the Gas Phase. *J. Phys. Chem. B* **2015**, *119*, 5773−5784.

(19) Storoniak, P.; Rak, J.; Ko, Y. J.; Wang, H.; Bowen, K. H. Photoelectron Spectroscopic and Density Functional Theoretical Studies of the 2′-deoxycytidine Homodimer Radical Anion. *J. Chem. Phys.* **2013**, *139*, 075101.

(20) Ponomareva, A. G.; Yurenko, Y. P.; Zhurakivsky, R. O.; Van Mourik, T.; Hovorun, D. M. Complete Conformational Space of the Potential HIV-1 Reverse Transcriptase Inhibitors d4U and d4C. A Quantum Chemical Study. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6787−6795.

(21) Tehrani, Z. A.; Javan, M. J.; Fattahi, A.; Hashemi, M. M. Structures, Stabilities & Conformational Behaviors of Hydrogen-Atom Abstractions of Cytosine Nucleosides: AIM & NBO Analysis. *Comput. Theor. Chem.* **2011**, *971*, 19−29.

(22) Matta, C. F.; Castillo, N.; Boyd, R. J. Extended Weak Bonding Interactions in DNA: π-Stacking (base-base), Base-Backbone, and Backbone-Backbone Interactions. *J. Phys. Chem. B* **2006**, *110*, 563−578.

(23) Hocquet, A.; Ghomi, M. The Peculiar Role of Cytosine in Nucleoside Conformational Behavior: Hydrogen Bond Donor Capacity of Nucleic Bases. *Phys. Chem. Chem. Phys.* **2000**, *2*, 5351−5353.

(24) Huang, M.; Giese, T. J.; Lee, T. S.; York, D. M. Improvement of DNA and RNA Sugar Pucker Profiles from Semiempirical Quantum Methods. *J. Chem. Theory Comput.* **2014**, *10*, 1538−1545.

(25) Cremer, D.; Pople, J. A. General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.

(26) Cremer, D. RING - A Coordinate Transformation Program for Evaluating the Degree and Type of Puckering of a Ring Compound. *Quantum Chem. Program Exchange* **1975**, *288*, 1−8.

(27) Cremer, D.; Pople, J. A. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. XXIII. Pseudorotation in Saturated Five-Membered Ring Compounds. *J. Am. Chem. Soc.* **1975**, *97*, 1358−1367.

(28) Cremer, D. Theoretical Determination of Molecular Structure and Conformation. XI. the Puckering of Oxolanes. *Isr. J. Chem.* **1983**, *23*, 72−84.

(29) Essén, H.; Cremer, D. On the Relationship Between the Mean Plane and the Least-Squares Plane of an N-membered Puckered Ring. *Acta Crystallogr., Sect. B: Struct. Sci.* **1984**, *40*, 418−420.

(30) Cremer, D. Calculation of Puckered Rings with Analytical Gradients. *J. Phys. Chem.* **1990**, *94*, 5502−5509.

(31) Konkoli, Z.; Cremer, D. A New Way of Analyzing Vibrational Spectra I. Derivation of Adiabatic Internal Modes. *Int. J. Quantum Chem.* **1998**, *67*, 1−11.

(32) Konkoli, Z.; Cremer, D. A New Way of Analyzing Vibrational Spectra III. Characterization of Normal Vibrational Modes in Terms of Internal Vibrational Modes. *Int. J. Quantum Chem.* **1998**, *67*, 29−41.

(33) Konkoli, Z.; Larsson, J. A.; Cremer, D. A New Way of Analyzing Vibrational Spectra II. Comparison of Internal Mode Frequencies. *Int. J. Quantum Chem.* **1998**, *67*, 11−29.

(34) Konkoli, Z.; Larsson, J. A.; Cremer, D. A New Way of Analyzing Vibrational Spectra IV. Application and Testing of Adiabatic Modes within the Concept of the Characterization of Normal Modes. *Int. J. Quantum Chem.* **1998**, *67*, 41−55.

(35) Bader, R. F. W.; Slee, T. S.; Cremer, D.; Kraka, E. Description of Conjugation and Hyperconjugation in Terms of Electron Distributions. *J. Am. Chem. Soc.* **1983**, *105*, 5061−5068.

(36) Cremer, D.; Kraka, E.; Slee, T. S.; Bader, R. F. W.; Lau, C. D. H.; Dang, T. T. N.; MacDougall, P. J. Description of Homoaromaticity in Terms of Electron Distributions. *J. Am. Chem. Soc.* **1983**, *105*, 5069−5075.

(37) Geise, H. J.; Adams, W. J.; Bartell, L. S. Electron Diffraction Study of Gaseous Tetrahydrofuran. *Tetrahedron* **1969**, *25*, 3045−3052.

(38) McQuarrie, D. A. *Statistical Thermodynamics*; Harper & Row: New York, 1973.

(39) Luger, P.; Buschmann, J. Twist Conformation of Tetrahydrofuran in the Crystal. *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 410−410.

(40) David, W. I. F.; Ibberson, R. M.; Dennis, T. J. S.; Hare, J. P.; Prassides, K. Structural Phase Transitions in the Fullerene C_{60}. *Europhys. Lett.* **1992**, *18*, 219−225.

(41) Han, S. J.; Kang, Y. K. Pseudorotation in Heterocyclic Five-membered Rings: Tetrahydrofuran and Pyrrolidine. *J. Mol. Struct.: THEOCHEM* **1996**, *369*, 157−165.

(42) Wu, A.; Cremer, D. New Approach for Determining the Conformational Features of Pseudorotating Ring Molecules Utilizing Calculated and Measured NMR Spin-Spin Coupling Constants. *J. Phys. Chem. A* **2003**, *107*, 1797−1810.

(43) Lambert, J. B.; Papay, J. J.; Khan, S. A.; Kappauf, K. A.; Magyar, E. S. Conformational Analysis of Five-membered Rings. *J. Am. Chem. Soc.* **1974**, *96*, 6112−6118.

(44) Kalinowski, H. O.; Berger, S.; Braun, S. *^{13}C-NMR-Spektroskopie [13 hoch C-NMR-Spektroskopie]: 200 Tabellen*; Thieme, 1984.

(45) Wilson, G. M. Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing. *J. Am. Chem. Soc.* **1964**, *86*, 127−130.

(46) Zou, W.; Kalescky, R.; Kraka, E.; Cremer, D. Relating Normal Vibrational Modes to Local Vibrational Modes With the Help of an Adiabatic Connection Scheme. *J. Chem. Phys.* **2012**, *137*, 084114.

(47) Zou, W.; Cremer, D. Properties of Local Vibrational Modes: The Infrared Intensity. *Theor. Chem. Acc.* **2014**, *133*, 1451−1466.

(48) Zou, W.; Kalescky, R.; Kraka, E.; Cremer, D. Relating Normal Vibrational Modes to Local Vibrational Modes: Benzene and Naphthalene. *J. Mol. Model.* **2013**, *19*, 2865−2877.

(49) Zou, W.; Cremer, D. C_2 in a Box: Determining its Intrinsic Bond Strength for the $X^1\Sigma_g^+$ Ground State. *Chem. - Eur. J.* **2016**, *22*, 4087−4089.

(50) Kraka, E.; Setiawan, D.; Cremer, D. Re-evaluation of the Bond Length-Bond Strength Rule: The Stronger Bond is Not Always the Shorter Bond. *J. Comput. Chem.* **2016**, *37*, 130−142.

(51) Setiawan, D.; Kraka, E.; Cremer, D. Hidden Bond Anomalies: The Peculiar Case of the Fluorinated Amine Chalcogenides. *J. Phys. Chem. A* **2015**, *119*, 9541−9556.

(52) Kalescky, R.; Kraka, E.; Cremer, D. Identification of the Strongest Bonds in Chemistry. *J. Phys. Chem. A* **2013**, *117*, 8981−8995.

(53) Humason, A.; Zou, W.; Cremer, D. 11, 11-Dimethyl-1, 6-methano[10]annulene - An Annulene with an Ultralong CC Bond or a Fluxional Molecule? *J. Phys. Chem. A* **2015**, *119*, 1666−1682.

(54) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Quantitative Assessment of the Multiplicity of Carbon-Halogen Bonds: Carbenium and Halonium Ions with F, Cl, Br, I. *J. Phys. Chem. A* **2014**, *118*, 1948−1963.

(55) Kraka, E.; Cremer, D. Characterization of CF Bonds with Multiple-Bond Character: Bond Lengths, Stretching Force Constants, and Bond Dissociation Energies. *ChemPhysChem* **2009**, *10*, 686−698.

(56) Kalescky, R.; Kraka, E.; Cremer, D. New Approach to Tolman's Electronic Parameter Based on Local Vibrational Modes. *Inorg. Chem.* **2014**, *53*, 478−495.

(57) Setiawan, D.; Kalescky, R.; Kraka, E.; Cremer, D. Direct Measure of Metal-Ligand Bonding Replacing the Tolman Electronic Parameter. *Inorg. Chem.* **2016**, *55*, 2332−2344.

(58) Cremer, D.; Kraka, E. Generalization of the Tolman Electronic Parameter: The Metal-Ligand Electronic Parameter and the Intrinsic Strength of the Metal-Ligand Bond. *Dalton T.* **2017**, *46*, 8323−8338.

(59) Li, Y.; Oliveira, V.; Tang, C.; Cremer, D.; Liu, C.; Ma, J. The Peculiar Role of the Au_3 Unit in Au_m Clusters: σ-Aromaticity of the Au_5Zn^+ Ion. *Inorg. Chem.* **2017**, *56*, 5793−5803.

(60) Kalescky, R.; Kraka, E.; Cremer, D. Description of Aromaticity with the Help of Vibrational Spectroscopy: Anthracene and Phenanthrene. *J. Phys. Chem. A* **2014**, *118*, 223−237.

(61) Setiawan, D.; Kraka, E.; Cremer, D. Quantitative Assessment of Aromaticity and Antiaromaticity Utilizing Vibrational Spectroscopy. *J. Org. Chem.* **2016**, *81*, 9669−9686.

(62) Oliveira, V.; Kraka, E.; Cremer, D. Quantitative Assessment of Halogen Bonding Utilizing Vibrational Spectroscopy. *Inorg. Chem.* **2017**, *56*, 488−502.

(63) Oliveira, V.; Kraka, E.; Cremer, D. The Intrinsic Strength of the Halogen Bond: Electrostatic and Covalent Contributions Described by Coupled Cluster Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 33031−33046.

(64) Oliveira, V.; Cremer, D. Transition from Metal-Ligand Bonding to Halogen Bonding Involving a Metal as Halogen Acceptor a Study of Cu, Ag, Au, Pt, and Hg Complexes. *Chem. Phys. Lett.* **2017**, *681*, 56−63.

(65) Setiawan, D.; Kraka, E.; Cremer, D. Strength of the Pnicogen Bond in Complexes Involving Group Va Elements N, P, and As. *J. Phys. Chem. A* **2015**, *119*, 1642−1656.

(66) Setiawan, D.; Kraka, E.; Cremer, D. Description of Pnicogen Bonding With the Help of Vibrational Spectroscopy - The Missing Link Between Theory and Experiment. *Chem. Phys. Lett.* **2014**, *614*, 136−142.

(67) Setiawan, D.; Cremer, D. Super-Pnicogen Bonding in the Radical Anion of the Fluorophosphine Dimer. *Chem. Phys. Lett.* **2016**, *662*, 182−187.

(68) Tao, Y.; Zou, W.; Jia, J.; Li, W.; Cremer, D. Different Ways of Hydrogen Bonding in Water - Why Does Warm Water Freeze Faster than Cold Water? *J. Chem. Theory Comput.* **2017**, *13*, 55−76.

(69) Freindorf, M.; Kraka, E.; Cremer, D. A Comprehensive Analysis of Hydrogen Bond Interactions Based on Local Vibrational Modes. *Int. J. Quantum Chem.* **2012**, *112*, 3174−3187.

(70) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Vibrational Properties of the Isotopomers of the Water Dimer Derived from Experiment and Computations. *Aust. J. Chem.* **2014**, *67*, 426−434.

(71) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Local Vibrational Modes of the Water Dimer - Comparison of Theory and Experiment. *Chem. Phys. Lett.* **2012**, *554*, 243−247.

(72) Kalescky, R.; Kraka, E.; Cremer, D. Local Vibrational Modes of the Formic Acid Dimer - The Strength of the Double Hydrogen Bond. *Mol. Phys.* **2013**, *111*, 1497−1510.

(73) Zhang, X.; Dai, H.; Yan, H.; Zou, W.; Cremer, D. B-H···π Interaction: A New Type of Nonclassical Hydrogen Bonding. *J. Am. Chem. Soc.* **2016**, *138*, 4334−4337.

(74) Kraka, E.; Freindorf, M.; Cremer, D. Chiral Discrimination by Vibrational Spectroscopy Utilizing Local Modes. *Chirality* **2013**, *25*, 185−196.

(75) Kraka, E.; Larsson, J. A.; Cremer, D. *Computational Spectroscopy: Methods, Experiments and Applications*; Wiley-VCH Verlag GmbH and Co. KGaA, 2010; pp 105−149.

(76) Bader, R. F. W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18*, 9−15.

(77) Bader, R. *Atoms in Molecules: A Quantum Theory*; International Series of Monographs on Chemistry; Clarendon Press: Oxford, 1990.

(78) Cremer, D.; Kraka, E. Chemical Bonds without Bonding Electron Density? Does the Difference Electron-Density Analysis Suffice for a Description of the Chemical Bond? *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 627−628.

(79) Cremer, D.; Kraka, E. A Description of the Chemical Bond in Terms of Local Properties of Electron Density and Energy. *Croat. Chem. Acta* **1984**, *57*, 1259−1281.

(80) Kraka, E.; Cremer, D. *Theoretical Models of Chemical Bonding. The Concept of the Chemical Bond*; Maksic, Z. B., Ed.; Springer Verlag: Heidelberg, 1990; Vol. 2, p 453.

(81) Reed, A.; Curtiss, L.; Weinhold, F. Intermolecular Interactions from A Natural Bond Orbital, Donor-Acceptor Viewpoint. *Chem. Rev.* **1988**, *88*, 899−926.

(82) Weinhold, F.; Landis, C. R. *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*; Cambridge University Press, 2003.

(83) Weinhold, F. Natural Bond Critical Point Analysis: Quantitative Relationships Between Natural Bond Orbital-Based and QTAIM-Based Topological Descriptors of Chemical Bonding. *J. Comput. Chem.* **2012**, *33*, 2440−2449.

(84) Zhang, Z.; Li, D.; Jiang, W.; Wang, Z. The Electron Density Delocalization of Hydrogen Bond Systems. *Adv. Phys.* **2018**, *3*, 298−315.

(85) Chai, J. D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.

(86) Chai, J. D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

(87) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257−2261.

(88) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724−728.

(89) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li−F. *J. Comput. Chem.* **1983**, *4*, 294−301.

(90) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-Consistent Molecular Orbital Methods 25. Supplementary Functions for Gaussian Basis Sets. *J. Chem. Phys.* **1984**, *80*, 3265−3269.

(91) Gräfenstein, J.; Cremer, D. Efficient Density-Functional Theory Integrations by Locally Augmented Radial Grids. *J. Chem. Phys.* **2007**, *127*, 164113.

(92) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. *Gaussian 09*, Revision B.1; Gaussian, Inc.: Wallingford, CT, 2010.

(93) Kraka, E.; Cremer, D.; Filatov, M.; Zou, W.; Grafenstein, J.; Izotov, D.; Gauss, J.; He, Y.; Wu, A.; Polo, V. et al. *COLOGNE2019*; Southern Methodist University: Dallas, TX, 2019.

(94) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Landis, C. R.; Weinhold, F. *NBO6*; Theoretical Chemistry Institute, University of Wisconsin: Madison, 2013.

(95) Keith, T. *TK Gristmill Software*; Overland Park, KS, USA, 2011; see http://aim.tkgristmill.com.

(96) Kraka, E.; Cremer, D. Weaker Bonds with Shorter Bond Lengths. *Rev. Proc. Quim.* **2012**, *6*, 31.

(97) Kraka, E.; Setiawan, D.; Cremer, D. Re-Evaluation of the Bond Length-Bond Strength Rule: The Stronger Bond is not Always the Shorter Bond. *J. Comput. Chem.* **2016**, *37*, 130−142.

(98) Cremer, D. *Modelling of Structure and Properties of Molecules*; Maksic, Z. B., Ed.; Ellis Horwood: Chichester, 1987; p 125.

(99) Freindorf, M.; Kraka, E.; Cremer, D. A Comprehensive Analysis of Hydrogen Bond Interactions Based on Local Vibrational Modes. *Int. J. Quantum Chem.* **2012**, *112*, 3174−3187.

(100) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Vibrational Properties of the Isotopomers of the Water Dimer Derived from Experiment and Computations. *Aust. J. Chem.* **2014**, *67*, 426−434.

(101) Tao, Y.; Zou, W.; Kraka, E. Strengthening of Hydrogen Bonding with the Push-Pull Effect. *Chem. Phys. Lett.* **2017**, *685*, 251−258.

(102) Herbert, H. E.; Halls, M. D.; Hratchian, H. P.; Raghavachari, K. Hydrogen-Bonding Interactions in Peptide Nucleic Acid and Deoxyribonuclei Acid: A Comparative Study. *J. Phys. Chem. B* **2006**, *110*, 3336−3343.

(103) Eremeeva, E.; Herdewijn, P. Non Canonical Genetic Material. *Curr. Opin. Biotechnol.* **2019**, *57*, 25−33.

(104) Hoshika, S.; Leal, N. A.; Kim, M.-J.; Kim, M.-S.; Karalkar, N. B.; Kim, H.-J.; Bates, A. M.; Watkins, N. E., Jr.; SantaLucia, H. A.; Meyer, A. J.; et al. Hachimoji DNA and RNA: A Genetic System with Eight Building Blocks. *Science* **2019**, *363*, 884−887.

Article

Hydrogen Bonding in Natural and Unnatural Base Pairs—A Local Vibrational Mode Study

Nassim Beiranvand †, Marek Freindorf † and Elfi Kraka *

Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, 3215 Daniel Ave, Dallas, TX 75275-0314, USA; nbeiranvand@smu.edu (N.B.); mfreindorf@smu.edu (M.F.)
* Correspondence: ekraka@smu.edu; Tel.: +1-214-76-1609
† These authors contributed equally to this work.

Abstract: In this work hydrogen bonding in a diverse set of 36 unnatural and the three natural Watson Crick base pairs adenine (A)–thymine (T), adenine (A)–uracil (U) and guanine (G)–cytosine (C) was assessed utilizing local vibrational force constants derived from the local mode analysis, originally introduced by Konkoli and Cremer as a unique bond strength measure based on vibrational spectroscopy. The local mode analysis was complemented by the topological analysis of the electronic density and the natural bond orbital analysis. The most interesting findings of our study are that (i) hydrogen bonding in Watson Crick base pairs is not exceptionally strong and (ii) the N–H···N is the most favorable hydrogen bond in both unnatural and natural base pairs while O–H···N/O bonds are the less favorable in unnatural base pairs and not found at all in natural base pairs. In addition, the important role of non-classical C–H···N/O bonds for the stabilization of base pairs was revealed, especially the role of C–H···O bonds in Watson Crick base pairs. Hydrogen bonding in Watson Crick base pairs modeled in the DNA via a QM/MM approach showed that the DNA environment increases the strength of the central N–H···N bond and the C–H···O bonds, and at the same time decreases the strength of the N–H···O bond. However, the general trends observed in the gas phase calculations remain unchanged. The new methodology presented and tested in this work provides the bioengineering community with an efficient design tool to assess and predict the type and strength of hydrogen bonding in artificial base pairs.

Keywords: natural base pairs; unnatural base pairs; hydrogen bonding; vibrational spectroscopy; local vibrational mode analysis

1. Introduction

Deoxyribonucleic acid (DNA) is one of the most intriguing biomolecules found in nature; it basically encodes all necessary information for the diverse functions of life [1–3]. In the early 1950s, a race started to determine the structure of this fascinating biomacro-molecule [4–7]. Linus Pauling and Robert B. Corey published an article in February of 1953 [8] proposing a triple helix DNA structure with the bases oriented at the outside. Although Pauling's and Corey's model was proven to be incorrect by Watson and Crick a couple of months later [9] they were one of the first scientists who came to the important conclusion that genes are segments of DNA that contain the code for a specific protein determining its function in different cells in the body, which can be considered as a first milestone for gene sequencing and gene cloning [10,11]. As described by Watson and Crick in their landmark paper [9], DNA forms a double strand helix, in which the four nucleobases guanine, cytosine, adenine and thymine of the DNA single strands form two pairs, guanine-cytosine (GC) and adenine-thymine (AT), also known as Watson-Crick or natural base pairs (**NBP**$_s$), bound to each other by intermolecular hydrogen bonds [12–14]. Why nature has decided to use just these two base pair combinations remains one of the greatest mysteries [15,16].

An increasing number of efforts have been made to use Nature's genius DNA concept in practical applications. For example, the physical and chemical properties of DNA have been exploited to create machines that are both encoded by and built from DNA molecules [17,18]. Utilizing DNA as a material building block in molecular and structural engineering has already led to the creation of numerous molecular-assembly systems and materials at the nanoscale [19]. Substantial efforts have been made to expand the genetic alphabet of DNA by introducing other base pair combinations, so-called unnatural base pairs (**UBP**$_s$) to increase nucleic acid functionalities [20–31]. Recently, modifications of DNA containing four **NBP**$_s$ and four additional **UBP**$_s$ which efficiently replicated [32,32] were reported. A range of **UBP**$_s$, termed xeno nucleic acids (**XNA**$_s$) were introduced [33,34]. **XNA**$_s$ are constructed by replacing natural bases, sugars, and phosphate linkages of DNA with artificial structures in order to synthesize potential alternative genetic materials, which may open new horizons of genetically modified organisms [18,35–39].

A key feature of the base pairs is their link via hydrogen bonds (HB) [40–45]. Therefore, a comprehensive study of the hydrogen bonds formed between the base pairs is imperative (i) for the deeper understanding of the structure and biological function of DNA, and (ii) to assess the qualification of designer **UBP**$_s$. HBs are one of the most important interactions found in biochemical molecules. Already in the 1950s Linus Pauling explored together with Robert B. Corey the importance of hydrogen bonding in proteins [46–48], work which contributed to his Nobel Price in Chemistry, awarded to him in 1954 for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances [49,50]. Up to now HBs have been the object of numerous experimental and theoretical investigations [45,51–53], and because of the complex interplay between different components, their nature is still subject of an ongoing debate [41,42,54]. Using intermolecular HBs as the key feature of base pairs selectivity GC base pairs with different bonding patterns and atomic organization were suggested [55,56], as well as different **UBP**$_s$ estimating the HBs via calculated interaction energies [57]. Brovarets and co-workers [58] discussed the formation of C–H$\cdots$O/N bonds in **UBP**$_s$ and showed that these HBs incorporate equally well into the structure of DNA. The effect of alkali metal cations on length and strength of HBs in DNA base pairs has been recently discussed [59] concluding that metal cations may help the base pairs stabilize to varying degrees depending on their position. DNA mutates as a result of proton transfer reactions. External electrostatic fields can modulate these reactions in DNA. Cerón-Carrasco and co-workers showed that mutagenic effects of high intensity electric fields on DNA have an impact on hydrogen bonding [60]. For example, during a pulse, HBs are elongated by widening the DNA strands, a reversible change once the electric field is removed. They also showed that in the guanine-cytosine base pair, the rate constants of proton transfer reactions can be changed by an electric field which is also in control of the mechanism of those reactions [61]. Through the design of a nanofluidic system that incorporates a number of synergistic functionalities displayed by both DNA molecules and the device itself, Kounovsky-Shafer and co-worker developed [62] an electrostatically inspired method for genome analysis. Molecular loading into nano-slits is aided by low ionic strength pressures, which are dynamically combined for efficient transport and temporal regulation. Proton transfer along DNA's hydrogen bonds can lead to gene mutation and, possibly, cancer. Slocombe and co-workers investigated energy barriers and tunneling rates of hydrogen transfer of canonical and tautomeric Watson-Crick DNA base pairs [63]. They showed that the guanine-cytosine structure plays a role in spontaneous point mutations, if it survives long enough to pass via the RNA polymerase. However, they found a slightly reverse reaction barrier for adenine-thymine, suggesting that the adenine-thymine tautomer is unstable [63]. As a result, populating the tautomeric state through double proton transfer from the canonical state is unlikely a biologically important mechanism for spontaneous point mutations. The tautomeric state seems to be more likely to occur through proton tunneling in the hydrogen-bonded conformation than in the single-stranded conformation, according to combined studies of the hydrogen-bonded and dissociated forms of the DNA bases. Florián and Leszczyński [64]

studied the energetic provisions for mutational DNA mechanisms. They showed that the guanine-cytosine base pair has more structural variability than previously thought. This base pair's ion-pair and imino-keto/amino-enol forms are energetically available, though the probability of their formation is less than 10^{-6} while they are significantly nonplanar. While all these efforts have definitely increased our knowledge about DNA they lack one important ingredient, a quantitative measure of the intrinsic HB strength, which we introduce in this work. Based on the local vibrational mode analysis, originally introduced by Konkoli and Cremer [65–69] we assessed the intrinsic strength of the HBs of the **NBP**$_s$ of DNA and the adenine-uracil (AU) base pair found in RNA, and a set of 36 **UBP**$_s$, shown in Figure 1 via local mode force constants, complemented with electron density and molecular orbital analyses. The **UBP**$_s$ where chosen from Brovarets' set which was designed to span over a large variety of different **UBP**$_s$ with both classical and non-classical HBs [58], complemented with several **UBP**$_s$ purely connected by classical HBs, (CU, TC, TT, TG, TU; see Figure 1) [70]. The main focus of our work was to shed some new light into the question why Nature chose the **NBP**$_s$ from the perspective of hydrogen bonding, exploring if hydrogen bonding in **UBP**$_s$ differ substantially from that in **NBP**$_s$. Based on our results we developed a roadmap for the design of **UBP**$_s$. The paper is structured in the following way. In the methodology section methods used in this work are described and computational details are provided. The results and discussion presents our finding, and conclusion and an outlook are made in the final section.
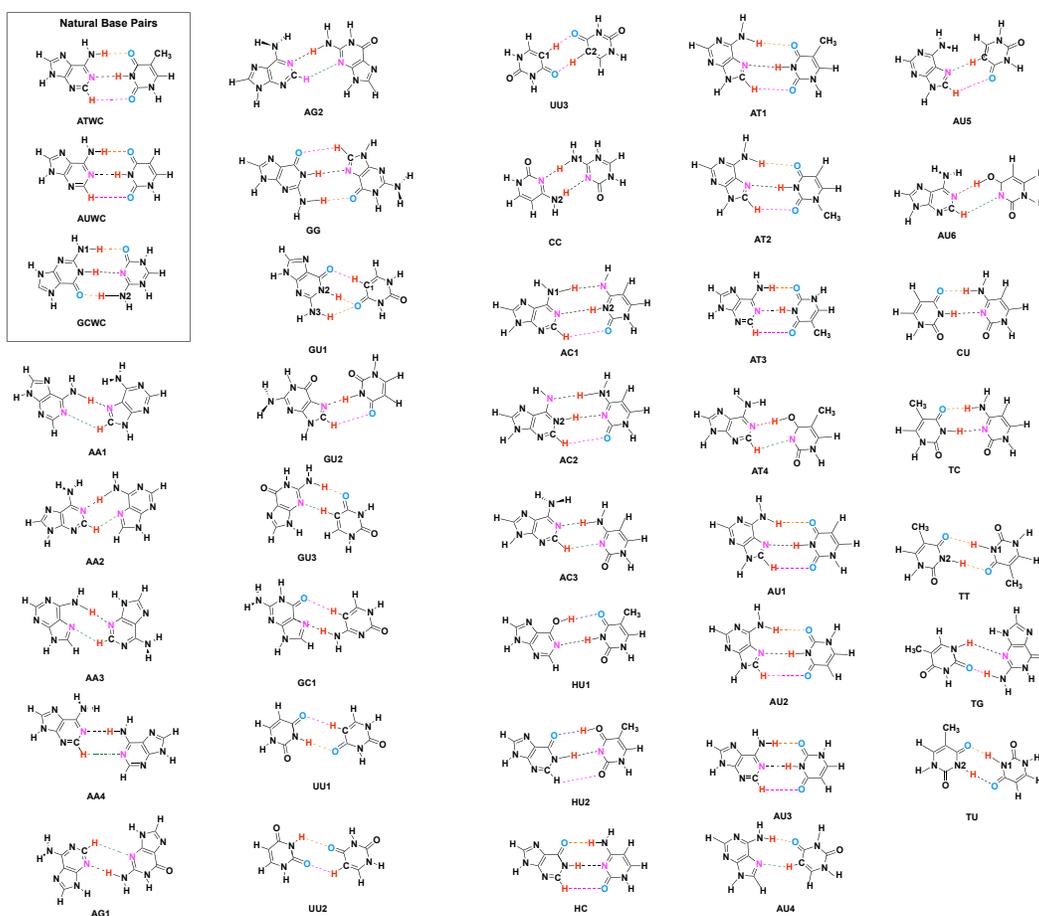

Figure 1. Investigated Base Pairs. (Key: A = Adenine, T = Thymine, C = Cytosine, G = Guanine, U = Uracil, H = Hypoxanthine. Different HBs are indicated by different color: N–H⋯N = Black, N–H⋯O = Orange, C–H⋯O = Pink, C–H⋯N = Green, O–H⋯O = Blue, O–H⋯N = Red). The selection of **UBP**$_s$ is described in the text.

2. Methodologies

In this section, first the tools applied in this work are introduced, i.e., the local vibrational mode analysis (LMA), originally introduced by Konkoli and Cremer [65–69] and the topological analysis of the electron density $\rho(\mathbf{r})$ based on Bader's quantum theory of atoms in molecules (QTAIM) [71,72]. Then the computational details are described.

2.1. The Local Vibrational Mode Analysis

A comprehensive mathematical derivation of the *local vibrational mode theory* can be found in a recent review article [73]. Therefore, in the following the most important essentials are summarized. The (3N-L) normal vibrational modes of a molecule being composed of N atoms (L = 5 for a linear and 6 for a nonlinear molecule) contain important information about the electronic structure and bonding. However, they tend to delocalize over the molecule due to the coupling of the atomic motions [74,75]. Therefore, one cannot directly derive an intrinsic bond strength measure from the normal modes. There are two coupling mechanisms, electronic coupling associated with the potential energy content of the vibrational mode and mass coupling associated with the kinetic energy. The electronic coupling between normal vibrational modes is promoted by the off-diagonal elements of the force constant matrix $\mathbf{F}^q$ in internal coordinates $\mathbf{q}$ and it is eliminated by diagonalizing $\mathbf{F}^q$, i.e., solving the fundamental equation of vibrational spectroscopy and transforming to normal mode coordinates [74,75]:

$$\mathbf{F}^q\,\mathbf{D} \quad = \quad \mathbf{G}^{-1}\,\mathbf{D}\,\mathbf{\Lambda} \tag{1}$$

where $\mathbf{G}$ represents the Wilson mass-matrix. Normal mode eigenvectors d_μ in internal coordinates $\mathbf{q}$ are collected in matrix $\mathbf{D}$, and the corresponding vibrational eigenvalues $\lambda_\mu = 4\pi^2 c^2 \omega_\mu$ in diagonal matrix $\mathbf{\Lambda}$, where ω_μ represents the harmonic vibrational frequency of mode d_μ in reciprocal cm, c is the speed of light, and $\mu = (1, 2 \dots N - L)$. Solving of Equation (1), e.g., diagonalizing the Wilson equation leads to the diagonal force constant matrix $\mathbf{K}$ in normal coordinates $\mathbf{Q}$, which is free of electronic coupling Equation (2) [74,75]:

$$\mathbf{K}^Q \quad = \quad \mathbf{D}^\dagger \mathbf{F}^q\,\mathbf{D} \tag{2}$$

However, this procedure does not resolve the mass-coupling which often has been overlooked. Konkoli and Cremer [65–69] solved this problem by introducing a mass-decoupled equivalent to the Wilson equation to derive mass-decoupled local vibrational modes $\mathbf{a}_i$ directly from normal vibrational modes $\mathbf{d}_i$ and the $\mathbf{K}$ matrix via Equation (3):

$$\mathbf{a}_i \quad = \quad \frac{\mathbf{K}^{-1}\mathbf{d}_i^\dagger}{\mathbf{d}_i\,\mathbf{K}^{-1}\,\mathbf{d}_i^\dagger} \tag{3}$$

For each local mode i, one can define a corresponding local model frequency ω_i^a, a local force constant k_i^a, and a local mode mass $G_{i,i}^a$ [65]. The local mode frequency ω_i^a is defined by:

$$(\omega_i^a)^2 \quad = \quad \frac{G_{i,i}^a\,k_i^a}{4\pi^2 c^2} \tag{4}$$

and the corresponding local mode force constant k_i^a by:

$$k_i^a \quad = \quad \mathbf{a}_i^\dagger\,\mathbf{K}\,\mathbf{a}_i \tag{5}$$

Local vibrational modes have a number of unique properties. Zou, Kraka and Cremer [67,68] verified the uniqueness of the local vibrational modes via an adiabatic connection scheme between local and normal vibrational modes. In contrast to normal mode force constants, local mode force constants have the advantage of not being dependent of the choice of the coordinates used to describe the target molecule and in contrast to vibrational frequencies they are independent of the atomic masses. They are of high sensitivity to elec-

tronic structure differences (e.g., caused by changing a substituent) and directly reflect the intrinsic strength of a bond or weak chemical interaction as shown by Zou and Cremer [76]. Thus, local vibration stretching force constants have been utilized as a unique measure of the intrinsic strength of a chemical bond [69,77–90] or weak chemical interaction [91–114] based on vibration spectroscopy.

In this work, we used local stretching frequencies ω^a and stretching force constants k^a for the characterization and comparison of the intrinsic strength of the intermolecular HBs of **NBP**$_s$ and their **UBP**$_s$ counter parts.

It is convenient to base the comparison of the bond strength of a series of molecules on a bond strength order (BSO) n rather than on a comparison for local force constant values. Both are connected via a power relationship according to the generalized Badger rule derived by Cremer and co-workers [79]:

$$BSO\ n \quad = \quad a\left(k^a\right)^b \tag{6}$$

The constants a and b in Equation (6) can be determined from two reference with known (BSO) n values and the requirement that for a zero force constant the BSO n is zero. For HBs, we generally use as references the FH bond in the FH molecule with BSO $n = 1$ and the FH bond in the $[\mathrm{H\cdots F\cdots H}]^-$ anion with BSO $n = 0.5$ [91,103,109]. For an ωB97X-D/6-31++G(d,p) model chemistry, applied in this study, this leads to $k^a(\mathrm{FH})$ = 9.782 mdyn/Å, $k^a(\mathrm{F\cdots H})$ = 0.901 mdyn/Å, $a = 0.515$ and $b = 0.291$. According to Equation (6) the OH bond in H_2O has a BSO n value of = 0.966. We scaled the reference values [109], so that the BSO n of the OH bond in H_2O is 1.

2.2. QTAIM and NBO Analysis

The Quantum Theory of Atoms-In-Molecules developed by Bader [115,116] presents a theoretical scheme for identifying, analyzing and characterizing chemical bonds and interactions via the topological features of the total electron density $\rho(\mathbf{r})$. In this work we used QTAIM as a complementary tool to the local mode analysis to determine the covalent/electrostatic character of internal HBs via the Cremer-Kraka criterion [117–119] of covalent bonding.

The Cremer-Kraka criterion is composed of two conditions; necessary condition: (i) existence of a bond path and bond critical point $\mathbf{r_c} = c$ between the two atoms under consideration; (ii) sufficient condition: the energy density $H(\mathbf{r_c}) = H_c$ is smaller than zero. $H(\mathbf{r})$ is defined as:

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) \tag{7}$$

where $G(\mathbf{r})$ is the kinetic energy density and $V(\mathbf{r})$ is the potential energy density. A negative $V(\mathbf{r})$ corresponds to a stabilizing accumulation of density whereas the positive $G(\mathbf{r})$ corresponds to depletion of electron density [118]. As a result, the sign of H_c indicates which term is dominant [119]. If $H_c < 0$, the interaction is considered covalent in nature, whereas $H_c > 0$ is indicative of electrostatic interactions.

In addition to the QTAIM analysis we used the Natural Bond Orbital (NBO) population analysis of Weinhold and co-workers [120–122] in order to obtain atomic charges and the charge transfer between the two monomers forming the base pair.

2.3. Computational Methods

Geometry optimizations and harmonic frequency calculations were performed with the Gaussian 16 program [123] using the ωB97X-D functional [124,125] in combination with Pople's 6-31++G(d,p) basis set [126–129]. An ultra-fine grid was used for the numerical DFT integration [130]. All local mode analysis calculations were carried out with the program package LModeA [73,131]. The NBO calculations were carried out with NBO 6 [122]. The QTAIM analysis was performed with the AIM2000 [132] software for calculating the bonds critical points and visualizing the bonds path. Binding energies (BE) were calculated

also at the ωB97X-D/6-31++G(d,p) level of theory, where the counterpoise correction of Boys and Bernardi [133] was used to correct for basis set superposition errors.

In order to assess the validity of our gas phase study we also analyzed the intrinsic HB strength for four Watson–Crick base pairs, two **AT**, and two **GC** in a DNA environment. The analysis was based on combined Quantum mechanics/Molecular mechanic (QM/MM) calculations [134], which were performed for each base pair using initial DNA coordinates from an X-ray structure of a synthetically constructed DNA dodecamer, PDB entry 6CQ3 [135] . The QM/MM geometry optimization were performed using ONIOM [136] with electronic embedding without constraints, followed by vibrational frequency calculations utilizing the ωB97X-D/6-31++G(d,p)/AMBER level of theory. For comparison we also calculated optimal geometries and vibrational frequencies of the **AT** base pair (labeled as **AT3**$_{gas}$) and the **GC** base pair (labeled as **GC3**$_{gas}$) in the gas phase, based on a starting geometries of the corresponding QM part of the QM/MM calculations at the ωB97X-D/6-31++G(d,p) level of theory. Calculations were performed with Gaussian16 [123]. Further details are given in the supporting information.

3. Results and Discussion

3.1. Internal HB Strength

Figure 2 shows the bond strength order BSO n of the HBs for the **NBP**$_s$ and **UBP**$_s$ as a function of the corresponding local stretching force constant k^a derived from Equation (6). The obtained BSO n values for **NBP**$_s$ range from 0.256 to 0.455 while BSO n values for **UBP**$_s$ range from 0.247 to 0.426, which leads to the important observation that the HBs of the **UBP**$_s$ and **NBP**$_s$ investigated in this work fall in to the same range, or in other words, the HB strength of **NBP**$_s$ does not stand out in any particular way. The N–H$\cdots$N bond shows the strongest BSO n values of either **NBP**$_s$ or **UBP**$_s$, and the weakest HB in both **NBP**$_s$ and **UBP**$_s$ is the C–H$\cdots$O bond. Also, **NBP**$_s$ are stabilized by three HBs while the majority of the **UBP**$_s$ are stabilized by two HBs. We found N–H$\cdots$N, N–H$\cdots$O and C–H$\cdots$O in **ATWC**, **AUWC** and **GCWC** are joined by one N–H$\cdots$N and two N–H$\cdots$O, but in case of **UBP**$_s$ in addition of these HBs, O–H$\cdots$O, C–H$\cdots$N and O–H$\cdots$N bonds were found. Different combinations of hydrogen donor and acceptor atoms appear to be the main difference between natural base pairs and unnatural base pairs (see Table S1 in the supporting information). In the following each individual HB type is discussed in more detail.

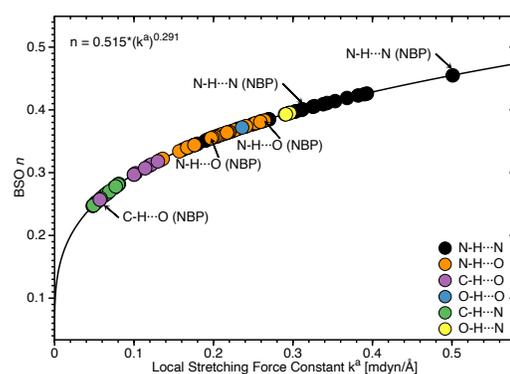**Figure 2.** Bond strength order BSO n of the HBs for the **NBP**$_s$ and **UBP**$_s$ as a function of the corresponding local stretching force constant k^a determined via Equation (6). Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

Table 1. Properties of hydrogen bonded reference complexes **R1–R5**.

Molecule	k^a	BSO n	R(HB)	ρ_c	H_c	H_c/ρ_c
R1	0.133	0.320	2.186	0.130	−0.002	−0.014
R2	0.113	0.307	2.122	0.115	0.000	0.002
R2*	0.026	0.212	2.705	0.051	0.007	0.143
R3	0.205	0.358	1.922	0.202	0.000	0.002
R4	0.209	0.360	1.905	0.183	−0.001	−0.005
R5	0.044	0.241	2.457	0.071	0.003	0.048

Force constant k^a in mdyn/Å, R(HB) in Å, ρ_c in e/Å^3 and H_c in Hartree/Å^3, normalized energy density H_c/ρ_c in Hartree/electron. Reference complexes **R1–R5** are shown in Figure 3. Calculated at the ωB97X-D/6-31++G(d,p) level of theory. R2* represents C–H···N bond.

Figure 3 shows compounds **R1–R5** used to compare HBs properties in base pairs and representative reference molecules. In Figure 4a, we compare BSO n values and force constant k^a for N–H···N bonds in base pairs and reference molecules. N–H···N bonds were found in 28 base pairs qualifying the N–H···N bond as the most favorable HB. BSO n values for N–H···N bonds range from 0.337 to 0.455, with **UBP**$_s$ values ranging from 0.337 to 0.426. It is interesting to note that all base pair N–H···N bonds are stronger than the HB bond of the reference molecule **R1** (BSO n = 0.320, see Table 1). The strongest N–H···N bond was found for **GCWC** which is stronger than the same bond in other **NBP**$_s$ (0.400 and 0.401 in **ATWC** and **AUWC**, respectively). The weakest N–H···N bond was found for **AA3**. Two gaps were observed in the Figure 4a; none of the **NBP**$_s$ and **UBP**$_s$ has a N–H···N bond in the BSO n range between 0.385 and 0.397 and between 0.426 and 0.455. One of the **UBP**$_s$ with the strongest N–H···N bonds, the **GG** base pair with a BSO n value of 0.426 has been discussed in so-called mismatched DNA causing genetic diseases [137,138]. The **GG** base pair has also the capability to form a H-bonding pattern close to that found in **NBP**$_s$, i.e., being stabilized by three HBs, N–H···N, N–H···O and C–H···O (see Figure 2) making this pair an interesting candidate for xenobiology [33,34,139].
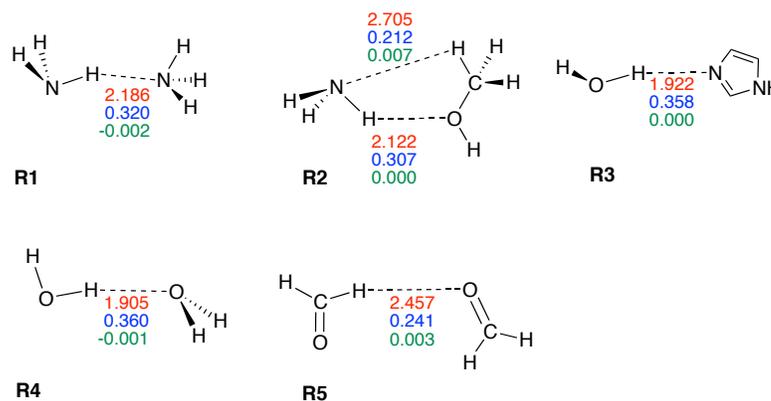
Figure 3. Reference complexes **R1–R5**. The intramolecular HB distance (Å) is given in red color, the corresponding BSO n value in blue color, and energy density $H(c)$ (Hartree/Å^3) in green color. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

In Figure 4b, we compared BSO n values and force constant k^a for N–H···O bond in base pairs and reference molecules. We found this HB in 21 base pairs. The BSO n values for N–H···O bonds range from 0.263 to 0.395. With the exception of **GG** (BSO n = 0.281) and **GU1** (BSO n = 0.263), all base pairs have stronger N–H···O bonds than reference molecule **R2** (BSO n value of 0.307, see Table 1). It should be noted that **GU1** makes two N–H···O bonds, two nitrogen atoms belong to guanine are donating electrons to an oxygen atom of uracil. The N–H···O bond in the center of this base pair has a BSO n = 0.368 which is stronger than the same bond in **R2**. The strongest N–H···O bond was found for **GCWC**, BSO n = 0.395 which is stronger than the same bond in other **NBP**$_s$ (0.356

and 0.355 in **ATWC** and **AUWC**, respectively). We observed three gaps in the Figure 4b, for BSO n between 0.322 and 0.334, between 0.344 and 0.355 and between 0.383 and 0.395. As the discussion above attests **GCWC** has the strongest HBs between the $\mathbf{NBP}_s$ and $\mathbf{UBP}_s$ and central N–H$\cdots$N bond is stronger than N–H$\cdots$O bond.

Figure 4c,d displays the BSO n values and force constant k^a for O–H$\cdots$N and O–H$\cdots$O bonds. These HBs are less favorable and occur only in $\mathbf{UBP}_s$. The O–H$\cdots$N bond was found in **AT4** and **AU6**, BSO n = 0.393 and 0.395, respectively, i.e., these HBs are stronger than the reference molecule **R3** (BSO n value of 0.358, see Table 1). The O–H$\cdots$O bond was found in **HU1** and **HU2** with BSO n 0.372 and 0.393, respectively, i.e., both HBs are even stronger than the HB in the water dimer (BSO n = 0.360, see reference molecule **R4** in Table 1). The O–H$\cdots$O bond in **HU2** is stronger than in **HU1**. The presence of a weak HB like C–H$\cdots$O bond and a strong central N–H$\cdots$N bond along with an O–H$\cdots$O bond in **HU2** obviously increases the overall stability of this base pair.

(a) BSO n and k^a for N–H$\cdots$N

(b) BSO n and k^a for N–H$\cdots$O

(c) BSO n and k^a for O–H$\cdots$N

(d) BSO n and k^a for O–H$\cdots$O

Figure 4. Bond strength order BSO n of the HBs for the $\mathbf{NBP}_s$ and $\mathbf{UBP}_s$ as a function of the corresponding local stretching force constant k^a as determined via Equation (6). For comparison the HBs of the reference complexes are included. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

Whereas base pair **HU1** is stabilized by O–H$\cdots$O and N–H$\cdots$N bonds, **HU2** exhibits three H-bond (O–H$\cdots$O, N–H$\cdots$N bonds and C–H$\cdots$O). The C–H$\cdots$O bond in **HU2** obviously makes a difference. Another weak HB, a C–H$\cdots$N bond along with a O–H$\cdots$N bond was found in **AT4** and **AU6** which appears to be stronger than the C–H$\cdots$O bond. In summary, weak HBs play an important role for the stabilization of base pairs, which will be discussed in more detail in the following section.

3.2. Significance of Non-Classical HBs

A classical HB is defined as the interaction between a hydrogen atom bonded to a highly electronegative atom such as oxygen, nitrogen and fluorine and the lone pair of

another such atom nearby [140,141]. Carbon is generally not considered as an electron donor which has led to this narrow definition of hydrogen bonding. In 2011, a new definition was introduced by the International Union of Pure and Applied Chemistry (IUPAC) [142,143], that emphasized the hydrogen donor does not always need to be one of the most electronegative atoms (oxygen, nitrogen and fluorine). One atom with a higher electronegativity than hydrogen is sufficient (non-classical HB). According to this new definition many interactions including less electronegative atoms such as carbon, chlorine, sulfur, phosphorus to act as the proton donor have been considered as HBs. In particular non-classical HBs were found to play a critical role for the structure and stability of biological systems, including DNA [144,145]. C–H$\cdots$O bonding between phosphate groups and nitrogenous bases was identified as a stabilizing part in DNA stability.

In the case of **AT** and **AU**, the absence of a C–H$\cdots$O bond causes their instability because the remaining oxygen in the minor groove can not be fully utilized as HB acceptor translating into a whole DNA structure [146]. Results presented in Figure 5 reveal the important stabilizing role of non classical HBs. In Figure 5a, the BSO n values and force constant k^a for the C–H$\cdots$N bond in base pairs are compared. This non-classical HB was found in 12 base pairs and it is less favorable for base pairs specially, we did not find it in any of the **NBP**$_s$. BSO n values for C–H$\cdots$N bond ranges from 0.248 to 0.297 which is much stronger than this HB in the reference molecule **R2***(BSO n value of 0.212, see Table 1). The C–H$\cdots$N bonds belong to the **AU6** and **AT4** with BSO n values 0.276 and 0.278, respectively, and are located in the middle of the Figure 5a. Thus they are strong enough to make the base pairs stable. In Figure 5b, the BSO n values and force constant k^a for C–H$\cdots$O bond in base pairs are compared. The non-classical C–H$\cdots$O bond is much more favorable than C–H$\cdots$N bond and it was found in 21 base pairs including both **NBP**$_s$ and **UBP**$_s$. BSO n values range from 0.247 to 0.318, which is stronger than this HB in the reference molecule **R5**(BSO n value of 0.241, see Table 1). None of the base pairs shows a C–H$\cdots$O bond in the gap between BSO n 0.281 to 0.297. This HB is naturally favored and it was found in **ATWC** and **AUWC** with BSO n value 0.257 and 0.256, receptively. However, we did not find any non-classical HBs in **GCWC**. **UU3** is stabilized just by two C–H$\cdots$O bonds and **AU5** is formed by two non classical weak HBs (C–H$\cdots$O bond and C–H$\cdots$N bond). Our results show that the C–H$\cdots$O bond is stronger than the C–H$\cdots$N bond (where, the percentage of C–H$\cdots$O bond with BSO n values ranges between 0.297 and 0.318 is 35.0%). This range is the most common HB type in water clusters [103]. In most cases, the base pairs are joined by three HBs, one of them is a weak non-classical C–H$\cdots$O bond, which is more naturally favored and stronger. The C–H$\cdots$N bonds were found always in **UBP**$_s$ which are stabilized by two HBs. We did not observe non classical HBs in 8 base pairs. These base pairs are stabilized by strong N–H$\cdots$N bond from BSO $n = 0.359$ to BSO $n = 0.455$ and N–H$\cdots$O bond from BSO $n = 0.364$ to BSO $n = 0.395$, and, C–H$\cdots$O bond with BSO $n = 0.372$.

(a) BSO n and k^a for C–H$\cdots$N
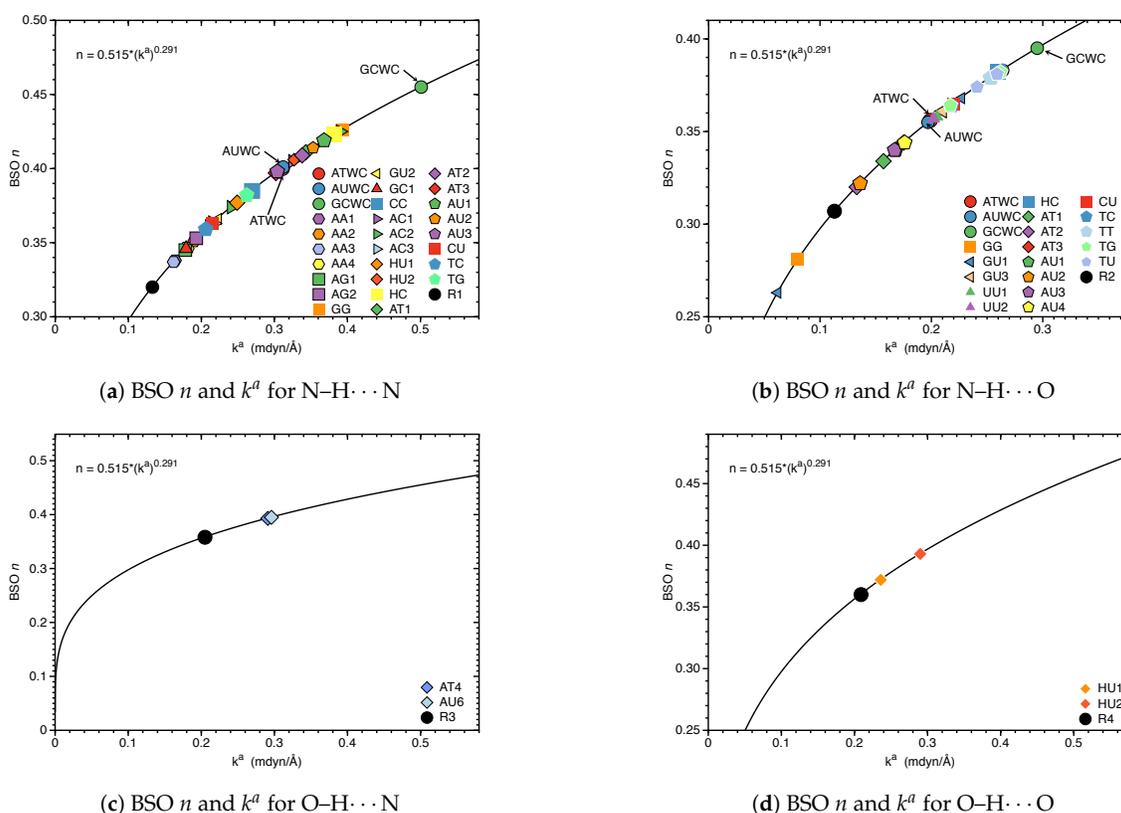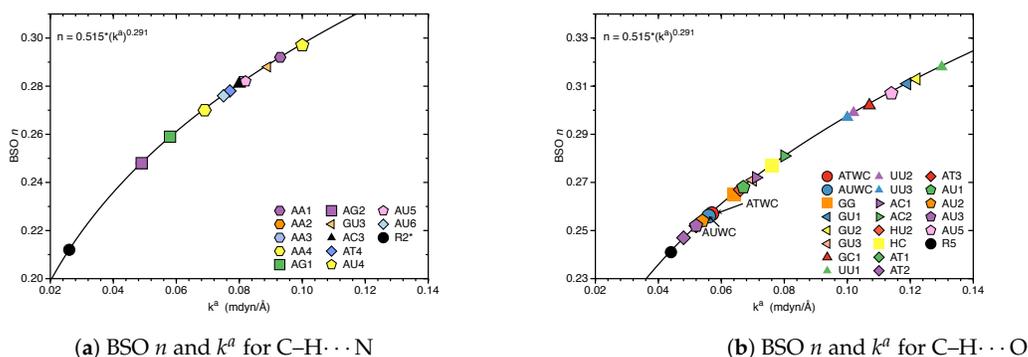

(b) BSO n and k^a for C–H$\cdots$O

Figure 5. Bond strength order BSO n of the HBs for the **NBP**$_s$ and **UBP**$_s$ as a function of the corresponding local stretching force constant k^a as determined via Equation (6). Calculated at the ωB97X-D/6-31++G(d,p) level of theory. For comparison the HB in the reference complexes are included.

3.3. Covalent Character of HBs

In the following we assess the covalent character of the **NBP** and **UBP** HBs via the normalized energy density H_c/ρ_c for all of HBs investigated in this work. The electron density analysis is complemented with NBO charges of all atoms X–H$\cdots$Y involved in hydrogen bonding, see Table S2 in the supporting information. In Figure 6, BSO n values are correlated with the corresponding H_c/ρ_c values. In case of N–H$\cdots$N bonds, H_c/ρ_c values range from -0.058 Hartree/electron to 0.002 Hartree/electron. All of the base pairs have N–H$\cdots$N bonds in the covalent region except the **GC1** with on the border value of $H_c/\rho_c = 0.002$ Hartree/electron for N–H$\cdots$N bond. Base pair **AU3** shows more covalent character (-0.058 Hartree/electron) of this HB compared to the same HB in other base pairs. According to our results H_c/ρ_c for strongest N–H$\cdots$N bond (belong to **GCWC**) is -0.021 Hartree/electron which is slightly less negative than the most covalent N–H$\cdots$N (belonging to **AU3**) with $H_c/\rho_c = -0.058$ Hartree/electron. However, most of the base pairs show that the increased strength of the N$\cdots$H bond in N–H$\cdots$N bond is correlated with a more covalent character of this bond (see Tables S1 and S2 in the supporting information).
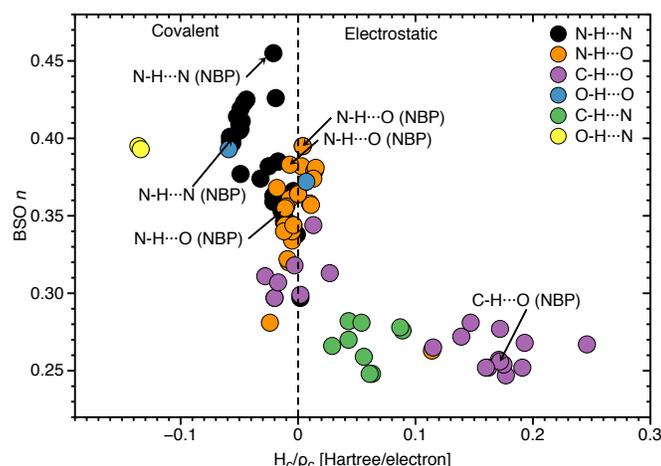
Figure 6. Correlation of BSO n and normalized energy density H_c/ρ_c. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

To be more specific, the N$\cdots$H bond of N–H$\cdots$N belonging to **GC1** has the less covalent character ($k_c = 0.002$ Hartree/electron) which shows the weakest BSO $n = 0.297$ for the same HB among all other base pairs. In Figure 6, the N$\cdots$H bond of O–H$\cdots$N shows the most covalent character ($H_c/\rho_c = -0.136$ Hartree/electron in **AU6** and -0.134 Hartree/electron in **AT4**) between all types of HB that was investigated in this work. According the HB strength analysis, these are stronger HBs compared other HBs since, the strength of the HB also depends on the nature of donor (N, O, C), in addition of the electron density distribution of the lone pair of the HB acceptor atom (N, O). However, as we see in Figure 6, there are several central N$\cdots$H bonds in the N–H$\cdots$N bond with less covalent characters. This leads to the conclusion that if a base pair is formed with three HBs, a strong N–H$\cdots$N bond is found in the middle. But the same HB in the base pair with two possible HBs is less covalent and weaker than the N$\cdots$H bond in O–H$\cdots$N. It should be noted that we didn't found O–H$\cdots$N bond in **WCBP**$_s$, it means that the electrostatic interactions are more strongly felt in the interior non-polar environment of DNA where the bases form a pair. According Figure 6, C–H$\cdots$N bonds are in the electrostatic region, but C–H$\cdots$O bonds are speared in both covalent region with H_c/ρ_c from -0.028 to -0.003 Hartree/electron and electrostatic regions with H_c/ρ_c from 0.002 to 0.246 Hartree/electron. The C–H$\cdots$O bonds belong to **ATWC** and **AUWC** are in the electrostatic region.

3.4. Intrinsic HB Strength and BEs

In Figure 7, the correlation between the average of BSO *n* of the HBs in each base pairs with BE_s is shown. The reason for using average BSO *n* is to account for the fact that the number of HBs differs in the base pairs, i.e., three or two HBs. There is some overall trend, i.e., stronger HBs are connected with larger BEs. However, the scattering of data points shows that there is no direct relationship between the two quantities, which is not surprising. The BE is a cumulative measure of the overall energy required to break a bond/weak interaction including the reorganization of the electron density and geometry relaxation of the dissociation product while the BSO *n* reflects intrinsic strength of the HB, as discussed above [101,102,105]. **GCWC** has the strongest BE (−32.88 kcal/mol) and the largest BSO *n* of 0.411 of all HBs investigated in this work. In contrast, the other two **NBP**$_s$ are found in the middle range. **AU5** has the weakest BE with −6.74 kcal/mol and the smallest average BSO *n* of 0.295.
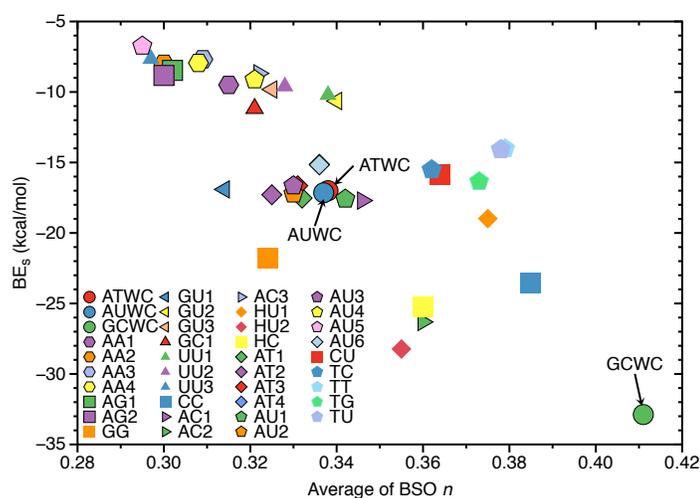
Figure 7. Correlation of the average of BSO *n* and binding energies (BE_s). Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

3.5. HBs in the DNA Environment

In order to evaluate the influence of the DNA environment on hydrogen bonding we compared the **AT** and **GC** base pairs in the gas phase (Figure 8a) and in the DNA (Figure 8b). The results are summarized in Tables 2 and 3.

According to our calculations, the strongest HB of the **AT** base pairs in DNA, is observed for the N–H···N bonds, which are even stronger than in the **AT** base pair in the gas phase (0.383 and 0.420 mdyn/Å in **AT1** and **AT2** in DNA, and 0.318 mdyn/Å in **AT3** in the gas phase). However, the force constant of the O···H bond in the N–H···O bonds of the **AT** base pairs in DNA are smaller than the corresponding HBs in the gas phase (0.128 and 0.102 mdyn/Å in **AT1** and **AT2** in DNA, and 0.201 mdyn/Å in **AT3** in the gas phase), and the opposite trend is observed for the C–H···O non-classical HBs (0.156 and 0.141 mdyn/Å in **AT1** and **AT2** in DNA, and 0.060 mdyn/Å in **AT3** the gas phase). These results indicate that the DNA environment increases the strength of the central N–H···N HB and the C–H···O non-classical HB, and at the same time it decreases the strength of the N–H···O bond in the **AT** base pairs. As it is seen in Table 2 the increased strength of the N···H bond of the central N–H···N bond in DNA, is also correlated with the decreased strength of the N–H bond in this hydrogen bond (4.496 and 3.338 mdyn/Å in **AT1** and **AT2** in DNA, and 4.673 mdyn/Å in **AT3** the gas phase). A similar effect of the DNA environment is observed in our calculations of the **GC** base pairs. According to Table 2, the strongest hydrogen bond of the **GC** base pairs in DNA, is observed for the central N–H···N bond, similarly as in the **AT** base pairs, where the DNA environment increases

the strength of the N$\cdots$H bond in this hydrogen bond of the **GC** base pairs, and this increase is also correlated with a decrease of the N–H bond strength in this HB. Therefore, based on our QM/MM calculations of the two **AT** and two **GC** base pairs in DNA, we conclude that the DNA environment changes the electronic structure the central N–H$\cdots$N bond of these base pairs, which makes the proton transfer between nitrogen atoms of the purine and pyrimidine bases easier. We can generally conclude that the gas phase calculations show the general features of HBs for the majority of the base pairs presented in this study. It has been confirmed in other studies that the Watson-Crick **AT** and **GC** base pairs are electronically complementary through proton transfer [147,148]. These results can be expanded to tautomeric base pairs where photoexcitation studies show a link between UV-excited DNA states and efficient charge production and transmission in DNA [147]. Base pair radical ions behave similarly to those created when ionizing radiation interacts with DNA [148–151]. Intermolecular hydrogen-bond distances in both tautomeric Watson-Crick base pairs are shorter than those in canonic base pairs. This means that after double-proton transfer in the canonic base pairs, the HBs become stronger [152–154].

Table 2. Comparison of the **AT** and **GC** base pair HBs in gas phase and DNA [a].

Base Pair	X–H Bond		Y$\cdots$H Bond		X$\cdots$Y Distance	
	d (Å)	k [a] (mdyn/Å)	d (Å)	k [a] (mdyn/Å)	d_{calc} (Å)	d_{exp} [b] (Å)
AT1$_{DNA}$						
N–H$\cdots$O	1.020	6.747	2.108	0.128	3.122	3.050
N$\cdots$H–N	1.050	4.496	1.764	0.383	2.812	2.776
C–H$\cdots$O	1.084	5.772	2.504	0.156	3.334	3.468
AT2$_{DNA}$						
N–H$\cdots$O	1.015	7.014	2.144	0.102	3.143	2.981
N$\cdots$H–N	1.065	3.338	1.638	0.420	2.700	2.761
C–H$\cdots$O	1.085	5.701	2.385	0.141	3.234	3.475
AT3$_{gas}$						
N–H$\cdots$O	1.022	6.554	1.903	0.201	2.921	-
N$\cdots$H–N	1.047	4.673	1.777	0.318	2.824	-
C–H$\cdots$O	1.087	5.635	2.738	0.060	3.570	-
GC1$_{DNA}$						
N–H$\cdots$O	1.020	6.508	1.855	0.267	2.866	2.880
N–H$\cdots$N	1.035	5.508	1.831	0.511	2.851	2.912
O$\cdots$H–N	1.030	5.663	1.748	0.282	2.763	2.852
GC2$_{DNA}$						
N–H$\cdots$O	1.027	6.096	1.758	0.387	2.784	2.789
N–H$\cdots$N	1.038	5.379	1.855	0.517	2.892	2.875
O$\cdots$H–N	1.025	6.241	1.868	0.230	2.888	2.839
GC3$_{gas}$						
N–H$\cdots$O	1.022	6.418	1.866	0.264	2.888	-
N–H$\cdots$N	1.033	5.695	1.891	0.486	2.924	-
O$\cdots$H–N	1.034	5.496	1.750	0.283	2.785	-

[a] QM/MM calculations in DNA: the base pair **AT1**$_{DNA}$, **AT2**$_{DNA}$, **GC1**$_{DNA}$, **GC2**$_{DNA}$, the ωB97X-D/6-31++G(d,p)/AMBER level of theory; QM calculations in the gas phase: the base pair **AT3**$_{gas}$ and **GC3**$_{gas}$, the ωB97X-D/6-31++G(d,p) level of theory. The left atomic symbol of the base pair label corresponds to the purine basis (A and G), and the right atomic symbol corresponds to the pyrimidine basis (T and C); the X and Y symbols correspond to the hydrogen donor and acceptor atoms, respectively. [b] Taken from the experimental X-ray structure [135].
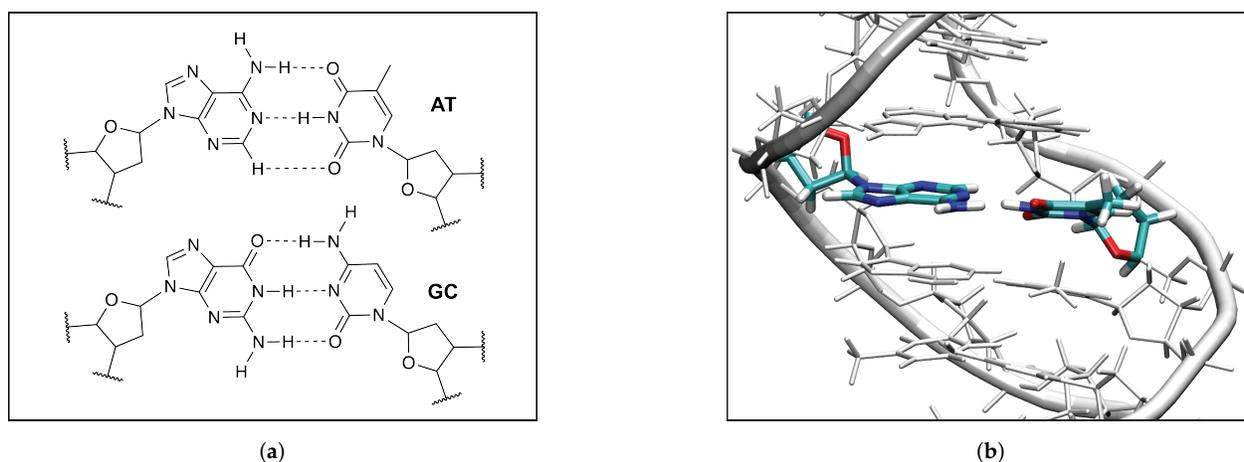
(a)

(b)

Figure 8. (**a**) The QM models of the **AT** and **GC** base pairs used in the QM/MM calculation; (**b**) The QM/MM optimized geometry of the **AT2**$_{DNA}$ base pair in DNA, the ωB97X-D/6-31++G(d,p)/AMBER level of theory, the water molecules and the Na$^+$ ions are not shown for clarity of the picture.

NBO charges of atoms X–H$\cdots$Y and energy densities of the X–H and the Y$\cdots$H HBs of the **AT** and **GC** base pairs in gas phase and DNA are compared in Table 3. Because the central N–H$\cdots$N bond is the strongest HB in the investigated base pairs, we focus in the following on a discussion of the NBO charges and H_c for this particular HB. According to Table 3, the NBO atomic charges of the atoms involved in the N–H$\cdots$N bond for the **AT** base pairs in DNA are similar to the NBO charges of these atoms based on the gas phase calculations (-0.605, 0.477, and -0.687 e in **AT1**; -0.606, 0.477, -0.685 e in **AT2**; -0.615, 0.477, -0.688 e in **AT3** for the N, H and N atoms, respectively; the first N atom belongs to the A base and the second N atom belongs to the T base). However, we observe changes in the H_c for the N$\cdots$H bond, when comparing the results from the DNA calculations and from the gas phase calculations (-0.0189 and -0.0661 Hartree/Å^3 in **AT1** and **AT2** in DNA, and -0.0169 Hartree/Å^3 in **AT3** in the gas phase). Similarly, there are also changes in the H_c for the N–H bond (-3.0368 and -2.8613 Hartree/Å^3 in **AT1** and **AT2** in DNA, and -3.0631 Hartree/Å^3 in **AT3** in the gas phase). Therefore, the changes of the energy density at the bond critical point shown in Table 3 are consistent with the changes of the local mode force constants presented in Table 2, showing that the increased strength of the N$\cdots$H bond in the DNA surrounding, is correlated with a more covalent character of this bond, and the decreased strength of the N–H bond in DNA is correlated with a less covalent character of this bond.

According to Table 3, the NBO atomic charges of the atoms involved in the N–H$\cdots$N bond for the **GC** base pairs in DNA are similar to the NBO charges of these atoms based on the gas phase calculations (-0.658, 0.472, -0.655 e in **GC1**; -0.668, 0.472, -0.671 e in **GC2**; -0.677, 0.467, -0.661 e in **GC3** for the N, H and N atoms, respectively; the first N atom belongs to the G and the second N atom belongs to the C base). Similar to the **AT** base pairs, there are changes in the energy density at the bond critical point between the values obtained from the DNA calculations and from the gas phase calculations for the N$\cdots$H bond (-0.0074 and -0.0067 Hartree/Å^3 in **GC1** and **GC2** in DNA, and -0.0047 Hartree/Å^3 in **GC3** in the gas phase), and for the N–H bond (-3.2082 and -3.1751 Hartree/Å^3 in **GC1** and **GC2** in DNA, and -3.2291 Hartree/Å^3 in **GC3** in the gas phase). Therefore, similar to the **AT** base pairs, the increased strength of the N$\cdots$H bond in the **GC** base pairs in the DNA surrounding, is correlated with a more covalent character of this bond, and the decreased strength of the N–H bond in DNA is correlated with a less covalent character of this bond. The similar NBO atomic charges of the atoms involved in the N–H$\cdots$N bond of the **AT** and **GC** base pairs in DNA and in the gas phase, confirm that the electrostatic interaction between these atoms is less important for the change of the strength in these bonds. Table 2 shows also a comparison of the QM/MM calculated in

our study and the experimentally measured [135] distance between the hydrogen donor atom (X) and the hydrogen acceptor atom (Y) of the X–H$\cdots$Y bond in the **AT** and **GC** base pairs in DNA. As a reference we also present in Table 2 this distance based on the QM gas phase calculations of the **AT** and **GC** base pairs. According to Table 2 the values of the calculated distance between the hydrogen donor and acceptor atoms are generally in the range the experimented values. Although this agreement is not perfect, a much better agreement with experiment is observed in our calculations of the **GC** rather than the **AT** base pairs, which can be explained by the smaller flexibility of the **GC** base pairs containing three classical HBs, in contrast to the **AT** base pairs having two classical and one non-classical HB.

Table 3. Comparison of NBO charges of atoms X–H$\cdots$Y and energy densities of the X–H and the Y$\cdots$H HBs of the **AT** and **GC** base pairs in gas phase and DNA [a].

Base Pair	NBO Atomic Charge (e)			H_c (Hartree/Å^3)	
	q_X	q_H	q_Y	X–H	Y$\cdots$H
AT1$_{DNA}$					
N–H$\cdots$O	-0.836	0.449	-0.665	-3.3479	-0.0047
N$\cdots$H–N	-0.605	0.477	-0.687	-3.0368	-0.0189
C–H$\cdots$O	0.275	0.243	-0.671	-2.1986	0.0054
AT2$_{DNA}$					
N–H$\cdots$O	-0.834	0.454	-0.708	-3.3701	-0.0027
N$\cdots$H–N	-0.606	0.477	-0.685	-2.8613	-0.0661
C–H$\cdots$O	0.289	0.239	-0.692	-2.1885	0.0040
AT3$_{gas}$					
N–H$\cdots$O	-0.829	0.458	-0.672	-3.2986	-0.0020
N$\cdots$H–N	-0.615	0.477	-0.688	-3.0631	-0.0169
C–H$\cdots$O	0.271	0.237	-0.668	-2.1493	0.0061
GC1$_{DNA}$					
N–H$\cdots$O	-0.835	0.456	-0.702	-3.3229	0.0007
N–H$\cdots$N	-0.658	0.472	-0.655	-3.2082	-0.0074
O$\cdots$H–N	-0.754	0.471	-0.791	-3.2149	0.0027
GC2$_{DNA}$					
N–H$\cdots$O	-0.851	0.454	-0.696	-3.2736	0.0027
N–H$\cdots$N	-0.668	0.472	-0.671	-3.1751	-0.0067
O$\cdots$H–N	-0.722	0.466	-0.800	-3.2716	0.0000
GC3$_{gas}$					
N–H$\cdots$O	-0.860	0.458	-0.711	-3.2831	-0.0007
N–H$\cdots$N	-0.677	0.467	-0.661	-3.2291	-0.0007
O$\cdots$H–N	-0.686	0.468	-0.810	-3.1677	-0.0047

[a] QM/MM calculations in DNA: the base pair **AT1**$_{DNA}$, **AT2**$_{DNA}$, **GC1**$_{DNA}$, **GC2**$_{DNA}$, the ωB97X-D/6-31++G(d,p)/AMBER level of theory; QM calculations in the gas phase: the base pair **AT3**$_{gas}$ and **GC3**$_{gas}$, the ωB97X-D/6-31++G(d,p) level of theory. X and Y correspond to the hydrogen donor and acceptor atoms, respectively. The left atomic symbol of the base pair label and the q_X charge in e correspond to the purine basis (A and G); the right atomic symbol of the base pair label and the q_Y charge in e correspond to the pyrimidine basis (T and C).

4. Conclusions and Outlook

We investigated in this work intermolecular hydrogen bonding in a diverse set of 36 unnatural and the three natural Watson Crick base pairs adenine (A)–thymine (T), adenine (A)–uracil (U) and guanine (G)–cytosine (C). The hydrogen bond strength was assessed utilizing local vibrational force constants derived from the local mode analysis, originally introduced by Konkoli and Cremer as a unique bond strength measure based on vibrational spectroscopy. The local mode analysis was complemented by the topological analysis of the electronic density and the natural bond orbital analysis. Our study led to the following interesting insights:

- Hydrogen bonding in Watson Crick base pairs is not exceptionally strong and the N–H$\cdots$N bond is the most favorable hydrogen bond in both unnatural and natural

base pairs while O–H··· N/O bonds are the less favorable in unnatural base pairs and not found at all in natural base pairs.

- In addition, non-classical C–H··· N/O bonds play an important role for the stabilization of base pairs, especially C–H··· O bonds in Watson Crick base pairs. This suggests that Nature's choice to combine classical and non-classical hydrogen bonding should also be copied in the design of new unnatural base pair combinations.

- Hydrogen bonding in Watson Crick base pairs modeled in the DNA via a QM/MM approach showed that the DNA environment increases the strength of the central N–H··· N bond and the C–H··· O bonds, and at the same time decreases the strength of the N–H··· O bond. However, the general trends observed in the gas phase calculations remain unchanged reflecting that electrostatic interactions with the environment are a less important factor determining the intermolecular hydrogen bond strength; an important validation of the gas phase model applied in this work.

- Natural base pairs do not possess larger binding energies than their unnatural counterparts. We also did not find a significant correlation between hydrogen bond strengths and binding energies, i.e., BSO n and BE values, as expected because these two quantities cannot directly be compared.

- We expect that the presence of base pairs with more nonclassical, i.e., weaker HBs in DNA will make the environment less covalent. During electron transfer these bonds will couple with specific vibrational modes of the DNA strand changing the electronic properties of the DNA. It has been documented [155–158] that these changes can stretch over 10 to 80 nucleobases accompanied by a decrease of the corresponding normal frequencies. When the DNA body lengthens, it becomes more mobile and less rigid. The experiment can only acquire normal vibrational frequencies of the backbone and the bases of DNA molecules characterized by coupled vibrational modes, while we can capture via LMA individual local frequencies from low to high and decode specific atomic motions, leading to more comprehensive and deeper insights into the stability of the DNA strand, which we will further explore in future work.

- The stability of the DNA double helix is mainly determined by (i) non-covalent interactions involving hydrogen bonds between A-T and G-C base pairs, (ii) stacking interactions between adjacent bases along the helix, and (iii) cross-interactions between base pairs [159]. Interactions outside the DNA double helix generally play a less important role [160]. The interplay between hydrogen bonding and stacking interactions in DNA has been the subject of several experimental [161–165] and theoretical investigations [159,166–173]. Based on DNA melting and energetics of the double helix [174], it has been recently suggested that in accordance with previous experiments [165,175] the stability of DNA double strands depends mainly on G-C base pair rich sequences. This is completely in line with our results identifying the hydrogen bonds of the G-C base pairs as one of the strongest. The local mode analysis can also quantitatively assess the strength of the stacking interactions between adjacent DNA bases along the helix, which is currently under investigation.

In summary, our study clearly reveals that not only the intermolecular hydrogen bond strength but also the combination of classical and non-classical hydrogen bonds play a significant role in natural base pairs, which should be copied in the design of new promising unnatural base pair candidates. Our local mode analysis, presented and tested in this work provides the bioengineering community with an efficient design tool to assess and predict the type and strength of hydrogen bonding in artificial base pairs.

Supplementary Materials: The following are available online. Cartesian coordinates of all **NBP**$_s$ and **UBP**$_s$ investigated in this work; Table S1: Bond distances R, local mode force constants k^a, local mode frequencies ω_a and bond strength orders BSO for all **UBP** and **NBP** BHs investigated in this work; Table S2: NBO charges of atoms X–H··· Y and energy density parameters for all **UBP** and **NBP** BHs investigated in this work; specific note Note about the QM/MM calculations.

Abbreviations

The following abbreviations are used in this manuscript:

HB	Hydrogen bond
NBP	natural base pair
UBP	Unnatural base pair

References

1. Minchin, S.; Lodge, J. Understanding biochemistry: Structure and function of nucleic acids. *Essays Biochem.* **2020**, *63*, 433–456. [CrossRef]
2. Chevizovich, D.; Michieletto, D.; Mvogo, A.; Zakiryanov, F.; Zdravkovic, S. A review on nonlinear DNA physics. *R. Soc. Open Sci.* **2020**, *7*, 200774. [CrossRef] [PubMed]
3. Varela, S.; Montañes, B.; López, F.; Berche, B.; Guillot, B.; Mujica, V.; Medina, E. Intrinsic Rashba Coupling Due to Hydrogen Bonding in DNA. *J. Chem. Phys.* **2019**, *151*, 125102. [CrossRef] [PubMed]
4. Hernandez, V. A Proposed Structure for the Nucleic Acids (1953) by Linus Pauling and Robert Brainard Corey. Embryo Project Encyclopedia, SSN: 1940-5030. Available online: http://embryo.asu.edu/handle/10776/13121 (accessed on 26 August 2019).
5. Watson, J. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*; Weidenfeld and Nicolson: London, UK, 2012.
6. Holmes, F.L. *Meselson, Stahl, and the Replication of DNA: A History of 'The Most BeautifulExperiment in Biology'*; Yale University Press: London, UK, 2001.
7. Cobb, M. A Speculative History of DNA: What If Oswald Avery Had Died in 1934? *PLoS Biol.* **2016**, *14*, e2001197. [CrossRef] [PubMed]
8. Pauling, L.; Corey, R.B. A Proposed Structure For The Nucleic Acids. *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 84–97. [CrossRef]
9. Watson, J.D.; Crick, F.H. Molecular Structure of Nucleic Acids. *Nature* **1953**, *171*, 737–738. [CrossRef] [PubMed]
10. Durmaz, A.A.; Karaca, E.; Demkow, U.; Toruner, G.; Schoumans, J.; Cogulu, O. Evolution of Genetic Techniques: Past, Present, and Beyond. *BioMed Res. Int.* **2015**, *2015*, 461524. [CrossRef] [PubMed]
11. Watson, J.D. *The DNA Story: A Documentary History of Gene Cloning*; WH Freeman and Co.: New York, NY, USA, 1981.
12. Watson, J.D.; Crick, F.H. Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature* **1953**, *171*, 964–967. [CrossRef]
13. Fonseca Guerra, C.; Bickelhaupt, F.M.; Snijders, J.G.; Baerends, E.J. The Nature of the Hydrogen Bond in DNA Base Pairs: The Role of Charge Transfer and Resonance Assistance. *Chem. Eur. J.* **1999**, *5*, 3581–3594. [CrossRef]
14. Heinemann, U.; Roske, Y. Symmetry in Nucleic-Acid Double Helices. *Symmetry* **2020**, *12*, 737. [CrossRef]
15. Raguseo, F.; Chowdhury, S.; Minard, A.; Di Antonio, M. Chemical-biology Approaches to Probe DNA and RNA G-quadruplex Structures in The Genome. *ChemComm* **2020**, *56*, 1317–1324. [CrossRef] [PubMed]
16. Subramanian, H.; Gatenby, R.A. Evolutionary advantage of anti-parallel strand orientation of duplex DNA. *Sci. Rep.* **2020**, *10*, 9883. [CrossRef] [PubMed]
17. Ramezani, H.; Dietz, H. Building Machines with DNA Molecules. *Nat. Rev. Genet.* **2020**, *21*, 5–26. [CrossRef]
18. Nie, P.; Bai, Y.; Mei, H. Synthetic Life with Alternative Nucleic Acids as Genetic Materials. *Molecules* **2020**, *25*, 3483. [CrossRef] [PubMed]
19. Ghosh, D.; Datta, L.P.; Govindaraju, T. Molecular Architectonics of DNA for Functional Nanoarchitectures. *Beilstein J. Nanotechnol.* **2020**, *11*, 124–140. [CrossRef] [PubMed]
20. Romesberg, F.E. Synthetic Biology: The Chemist's Approach. *Isr. J. Chem.* **2019**, *59*, 91–94. [CrossRef]
21. Hamashima, K.; Soong, Y.T.; Matsunaga, K.I.; Kimoto, M.; Hirao, I. DNA Sequencing Method Including Unnatural Bases for DNA Aptamer Generation by Genetic Alphabet Expansion. *ACS Synth. Biol.* **2019**, *8*, 1401–1410. [CrossRef]

22. Devine, K.G.; Jheeta, S. De Novo Nucleic Acids: A Review of Synthetic Alternatives to DNA and RNA That Could Act as Bio-Information Storage Molecules. *Life* **2020**, *10*, 346. [CrossRef]
23. Feldman, A.W.; Ledbetter, M.P.; Zhang, Y.; Romesberg, F.E. Reply to Hettinger: Hydrophobic Unnatural Base Pairs and the Expansion of the Genetic Alphabet. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6478–E6479. [CrossRef]
24. Feldman, A.W.; Dien, V.T.; Karadeema, R.J.; Fischer, E.C.; You, Y.; Anderson, B.A.; Krishnamurthy, R.; Chen, J.S.; Li, L.; Romesberg, F.E. Optimization of Replication, Transcription, and Translation in a Semi-Synthetic Organism. *J. Am. Chem. Soc.* **2019**, *141*, 10644–10653. [CrossRef]
25. Weaver, J. Expanding the genetic alphabet. *BioTechniques* **2017**, *62*, 252–253. [CrossRef]
26. Biondi, E.; Benner, S.A. Artificially Expanded Genetic Information Systems for New Aptamer Technologies. *Biomedicines* **2018**, *6*, 53. [CrossRef]
27. Eggert, F.; Kurscheidt, K.; Hoffmann, E.; Kath-Schorr, S. Towards Reverse Transcription with an Expanded Genetic Alphabet. *ChemBioChem* **2019**, *20*, 1642–1645. [CrossRef] [PubMed]
28. Dien, V.T.; Holcomb, M.; Romesberg, F.E. Eight-Letter DNA. *Biochemistry* **2019**, *58*, 2581–2583. [CrossRef] [PubMed]
29. Jahiruddin, S.; Datta, A. What Sustains the Unnatural Base Pairs (UBPs) with No Hydrogen Bonds. *J. Phys. Chem. B* **2015**, *119*, 5839–5845. [CrossRef] [PubMed]
30. Jahiruddin, S.; Mandal, N.; Datta, A. Structure and Electronic Properties of Unnatural Base Pairs: The Role of Dispersion Interactions. *ChemPhysChem* **2018**, *19*, 67–74. [CrossRef]
31. Hernandez, A.R.; Shao, Y.; Hoshika, S.; Yang, Z.; Shelke, S.A.; Herrou, J.; Kim, H.J.; Kim, M.J.; Piccirilli, J.A.; Benner, S.A. A Crystal Structure of a Functional RNA Molecule Containing an Artificial Nucleobase Pair. *Angew. Chem. Int. Ed.* **2015**, *127*, 9991–9994. [CrossRef]
32. Hoshika, S.; Leal, N.A.; Kim, M.J.; Kim, M.S.; Karalkar, N.B.; Kim, H.J.; Bates, A.M.; Watkins, N.E.; SantaLucia, H.A.; Meyer, A.J. Hachimoji DNA and RNA: A Genetic System with Eight Building Blocks. *Science* **2019**, *363*, 884–887. [CrossRef] [PubMed]
33. Beck, K.M.; Krogh, M.B.; Hornum, M.; Ludford, P.T., III; Tor, Y.; Nielsen, P. Double-headed nucleotides as xeno nucleic acids: Information storage and polymerase recognition. *Org. Biomol. Chem.* **2020**, *18*, 7213–7223. [CrossRef] [PubMed]
34. Hamashima, K.; Kimoto, M.; Hirao, I. Creation of Unnatural Base Pairs for Genetic Alphabet Expansion Toward Synthetic Xenobiology. *Curr. Opin. Chem. Biol.* **2018**, *46*, 108–114. [CrossRef]
35. Feldman, A.W.; Dien, V.T.; Romesberg, F.E. Chemical Stabilization of Unnatural Nucleotide Triphosphates for the in Vivo Expansion of the Genetic Alphabet. *J. Am. Chem. Soc.* **2017**, *139*, 2464–2467. [CrossRef]
36. Dien, V.T.; Holcomb, M.; Feldman, A.W.; Fischer, E.C.; Dwyer, T.J.; Romesberg, F.E. Progress Toward a Semi-Synthetic Organism with an Unrestricted Expanded Genetic Alphabet. *J. Am. Chem. Soc.* **2018**, *140*, 16115–16123. [CrossRef]
37. Taylor, A.I.; Houlihan, G.; Holliger, P. Beyond DNA and RNA: The Expanding Toolbox of Synthetic Genetics. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032490. [CrossRef] [PubMed]
38. Kubyshkin, V.; Budisa, N. Anticipating Alien Cells with Alternative Genetic Codes: Away from the Alanine World! *Curr. Opin. Chem. Biol.* **2019**, *60*, 242–249. [CrossRef]
39. Whitford, C.M.; Dymek, S.; Kerkhoff, D.; März, C.; Schmidt, O.; Edich, M.; Droste, J.; Pucker, B.; Rückert, C.; Kalinowski, J. Auxotrophy to Xeno-DNA: An exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications. *J. Biol. Eng.* **2018**, *12*, 13. [CrossRef]
40. Arunan, E. One Hundred Years After The Latimer and Rodebush Paper, Hydrogen Bonding Remains an Elephant! *Indian J. Sci.* **2020**, *100*, 249–255. [CrossRef]
41. Gibb, B.C. The Centenary (maybe) of The Hydrogen Bond. *Nat. Chem.* **2020**, *12*, 665–667. [CrossRef] [PubMed]
42. Scheiner, S. The Hydrogen Bond: A Hundred Years and Counting. *J. Indian Inst. Sci.* **2020**, *100*, 61–76. [CrossRef]
43. Scheiner, S. Comparison of Bifurcated Halogen with Hydrogen Bonds. *Molecules* **2021**, *26*, 350. [CrossRef] [PubMed]
44. Wain-Hobson, S. The Third Bond. *Nature* **2006**, *439*, 539. [CrossRef]
45. Jayaraman, A. 100th Anniversary of Macromolecular Science Viewpoint: Modeling and Simulation of Macromolecules with Hydrogen Bonds: Challenges, Successes, and Opportunities. *ACS Macro Lett.* **2020**, *2*, 656–665. [CrossRef]
46. Pauling, L.; Corey, R.B. Two hydrogen-bonded spiral configurations of the polypeptide chain. *J. Am. Chem. Soc.* **1950**, *72*, 5349. [CrossRef]
47. Pauling, L.; Corey, R.B. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205–211. [CrossRef]
48. Pauling, L.; Corey, R.B. Configurations of polypeptide chains with favored orientations around single bonds: Two new pleated sheets. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 729–740. [CrossRef] [PubMed]
49. Nobel Prize for Chemistry: Prof. Linus Pauling, For. Mem. R.S. *Nature* **1954**, *174*, 907–908. [CrossRef]
50. Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: Ithaca, NY, USA, 1960; Volume 260.
51. Oswald, S.; Suhm, M.A. Soft Experimental Constraints for Soft Interactions: A Spectroscopic Benchmark Data Set For Weak and Strong Hydrogen Bonds. *Phys. Chem. Chem. Phys.* **2019**, *21*, 18799–18810. [CrossRef] [PubMed]
52. Grabowski, S.J. *Understanding Hydrogen Bonds: Theoretical and Experimental Views*, 1st ed.; Royal Society of Chemistry: London, UK, 2020.
53. Karas, L.J.; Wu, C.H.; Das, R.; Wu, J.I.C. Hydrogen bond design principles. *WIREs Comput. Mol. Sci.* **2020**, *10*, e1477. [CrossRef]

54. Van der Lubbe, S.C.C.; Guerra, C.F. The Nature of Hydrogen Bonds: A Delineation of the Role of Different Energy Components on Hydrogen Bond Strengths and Lengths. *Chem. Asian J.* **2019**, *14*, 2760–2769. [PubMed]

55. Williams-Ashman, G. Review of Horizons in Biochemistry. *Perspect. Biol. Med.* **1963**, *6*, 264–267. [CrossRef]

56. Hettinger, T.P. Helix Instability and Self-Pairing Prevent Unnatural Base Pairs From Expanding the Genetic Alphabet. *PNAS* **2017**, *114*, E6476–E6477. [CrossRef]

57. Czyznikowska, Z.; Góra, R.; Zalesny, R.; Lipkowski, P.; Jarzembska, K.; Dominiak, P.; Leszczynski, J. Structural Variability and the Nature of Intermolecular Interactions in Watson- Crick B-DNA Base Pairs. *J. Phys. Chem. B* **2010**, *114*, 9629–9644. [CrossRef]

58. Brovarets', O.O.; Yurenko, Y.P.; Hovorun, D.M. Intermolecular CH$\cdots$O/N H-bonds in the Biologically Important Pairs of Natural Nucleobases: A Thorough Quantum-Chemical Study. *J. Biomol. Struct. Dyn.* **2014**, *32*, 993–1022. [CrossRef] [PubMed]

59. Stasyuk, O.A.; Solà, M.; Swart, M.; Fonseca Guerra, C.; Krygowski, T.M.; Szatylowicz, H. Effect of Alkali Metal Cations on Length and Strength of Hydrogen Bonds in DNA Base Pairs. *ChemPhysChem* **2020**, *21*, 2112–2126. [CrossRef]

60. Cerón-Carrasco, J.P.; Cerezo, J.; Jacquemin, D. How DNA Is Damaged by External Electric Fields: Selective Mutation vs. Random Degradation. *Phys. Chem. Chem. Phys.* **2014**, *16*, 8243–8246. [CrossRef] [PubMed]

61. Cerón-Carrasco, J.P.; Jacquemin, D. Electric-Field Induced Mutation of DNA: A Theoretical Investigation of the GC Base Pair. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4548–4553. [CrossRef] [PubMed]

62. Kounovsky-Shafer, K.L.; Hernandez-Ortiz, J.P.; Potamousis, K.; Tsvid, G.; Place, M.; Ravindran, P.; Jo, K.; Zhou, S.; Odijk, T.; De Pablo, J.J.; et al. Electrostatic Confinement and Manipulation of DNA Molecules for Genome Analysis. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13400–13405. [CrossRef]

63. Slocombe, L.; Al-Khalili, J.; Sacchi, M. Quantum and Classical Effects in DNA Point Mutations: Watson–Crick Tautomerism in AT and GC BBase Pairs. *Phys. Chem. Chem. Phys.* **2021**, *23*, 4141–4150. [CrossRef]

64. Florián, J.; Leszczyński, J. Spontaneous DNA Mutations Induced by Proton Transfer in the Guanine? Cytosine Base Pairs: An Energetic Perspective. *J. Am. Chem. Soc.* **1996**, *118*, 3010–3017. [CrossRef]

65. Konkoli, Z.; Cremer, D. A New Way of Analyzing Vibrational Spectra. I. Derivation of Adiabatic Internal Modes. *Int. J. Quantum Chem.* **1998**, *67*, 1–9. [CrossRef]

66. Konkoli, Z.; Larsson, J.A.; Cremer, D. A New Way of Analyzing Vibrational Spectra. II. Comparison of Internal Mode Frequencies. *Int. J. Quantum Chem.* **1998**, *67*, 11–27. [CrossRef]

67. Konkoli, Z.; Cremer, D. A New Way of Analyzing Vibrational Spectra. III. Characterization of Normal Vibrational Modes in terms of Internal Vibrational Modes. *Int. J. Quantum Chem.* **1998**, *67*, 29–40. [CrossRef]

68. Konkoli, Z.; Larsson, J.A.; Cremer, D. A New Way of Analyzing Vibrational Spectra. IV. Application and Testing of Adiabatic Modes within the Concept of the Characterization of Normal Modes. *Int. J. Quantum Chem.* **1998**, *67*, 41–55. [CrossRef]

69. Cremer, D.; Larsson, J.A.; Kraka, E. New Developments in the Analysis of Vibrational Spectra on the Use of Adiabatic Internal Vibrational Modes. In *Theoretical and Computational Chemistry*; Parkanyi, C., Ed.; Elsevier: Amsterdam, The Netherlands, 1998; pp. 259–327.

70. Mohajeri, A.; Nobandegani, F.F. Detection and Evaluation of Hydrogen Bond Strength in Nucleic Acid Base Pairs. *J. Phys. Chem. A* **2008**, *112*, 281–295. [CrossRef]

71. Bader, R. *Atoms in Molecules: A Quantum Theory*; Clarendron Press: Oxford, UK, 1995.

72. Popelier, P. *Atoms in Molecules: An Introduction*; Prentice-Hall: Harlow, UK, 2000.

73. Kraka, E.; Zou, W.; Tao, Y. Decoding Chemical Information from Vibrational Spectroscopy Data: Local Vibrational Mode Theory. *WIREs Comput. Mol. Sci.* **2020**, *10*, 1480. [CrossRef]

74. Wilson, E.B.; Decius, J.C.; Cross, P.C.M. *Molecular Vibrations. The Theory of Infrared and Raman Vibrational Spectra*; McGraw-Hill: New York, NY, USA, 1955; pp. 59–136.

75. Wilson, E.B.; Decius, J.C.; Cross, P.C.; Sundheim, B.R. Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra. *J. Electrochem. Soc.* **1955**, *102*, 235C. [CrossRef]

76. Zou, W.; Cremer, D. C$_2$ in a Box: Determining its Intrinsic Bond Strength for the X^1 Σ^+_g Ground State. *Chem. Eur. J.* **2016**, *22*, 4087–4097. [CrossRef]

77. Oomens, J.; Kraka, E.; Nguyen, M.K.; Morton, T.M. Structure, Vibrational Spectra, and Unimolecular Dissociation of Gaseous 1-Fluoro-1-phenethyl Cations. *J. Phys. Chem. A* **2008**, *112*, 10774–10783. [CrossRef]

78. Zou, W.; Kalescky, R.; Kraka, E.; Cremer, D. Relating Normal Vibrational Modes to Local Vibrational Modes: Benzene and Naphthalene. *J. Mol. Model.* **2012**, *19*, 2865–2877. [CrossRef] [PubMed]

79. Kalescky, R.; Kraka, E.; Cremer, D. Identification of the Strongest Bonds in Chemistry. *J. Phys. Chem. A* **2013**, *117*, 8981–8995. [CrossRef] [PubMed]

80. Kalescky, R.; Kraka, E.; Cremer, D. Description of Aromaticity with the Help of Vibrational Spectroscopy: Anthracene and Phenanthrene. *J. Phys. Chem. A* **2013**, *118*, 223–237. [CrossRef] [PubMed]

81. Kalescky, R.; Kraka, E.; Cremer, D. New Approach to Tolman's Electronic Parameter Based on Local Vibrational Modes. *Inorg. Chem.* **2013**, *53*, 478–495. [CrossRef]

82. Kalescky, R.; Kraka, E.; Cremer, D. Are Carbon-Halogen Double and Triple Bonds Possible? *Int. J. Quantum Chem.* **2014**, *114*, 1060–1072. [CrossRef]

83. Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Quantitative Assessment of the Multiplicity of Carbon-Halogen Bonds: Carbenium and Halonium Ions with F, Cl, Br, and I. *J. Phys. Chem. A* **2014**, *118*, 1948–1963. [CrossRef]

84. Humason, A.; Zou, W.; Cremer, D. 11,11-Dimethyl-1,6-methano[10]annulene—An Annulene with an Ultralong CC Bond or a Fluxional Molecule? *J. Phys. Chem. A.* **2014**, *119*, 1666–1682. [CrossRef]

85. Sethio, D.; Lawson Daku, L.M.; Hagemann, H.; Kraka, E. Quantitative Assessment of $B-B-B$, $B-H_b-B$, and $B-H_t$ Bonds: From BH_3 to $B_{12}H_{12}{}^{2-}$. *ChemPhysChem* **2019**, *20*, 1967–1977. [CrossRef] [PubMed]

86. Makoś, M.Z.; Freindorf, M.; Sethio, D.; Kraka, E. New Insights into $Fe-H_2$ and $Fe-H^-$ Bonding of a [NiFe] Hydrogenase Mimic—A Local Vibrational Mode Study. *Theor. Chem. Acc.* **2019**, *138*, 76. [CrossRef]

87. Makoś, M.Z.; Zou, W.; Freindorf, M.; Kraka, E. Metal-Ring Interactions in Actinide Sandwich Compounds: A Combined Normalized Elimination of the Small Component and Local Vibrational Mode Study. *Mol. Phys.* **2020**, *118*, e1768314. [CrossRef]

88. Verma, N.; Tao, Y.; Zou, W.; Chen, X.; Chen, X.; Freindorf, M.; Kraka, E. A Critical Evaluation of Vibrational Stark Effect (VSE) Probes with the Local Vibrational Mode Theory. *Sensors* **2020**, *20*, 2358. [CrossRef]

89. Freindorf, M.; Kraka, E. Critical Assessment of the FeC and CO Bond strength in Carboxymyoglobin—A QM/MM Local Vibrational Mode Study. *J. Mol. Model.* **2020**, *26*, 281. [CrossRef]

90. Kraka, E.; Freindorf, M. Characterizing the Metal Ligand Bond Strength via Vibrational Spectroscopy: The Metal Ligand Electronic Parameter (MLEP). In *Topics in Organometallic Chemistry—New Directions in the Modeling of Organometallic Reactions*; Lledós, A., Ujaque, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; Volume 67, pp. 1–43.

91. Freindorf, M.; Kraka, E.; Cremer, D. A Comprehensive Analysis of Hydrogen Bond Interactions Based on Local Vibrational Modes. *Int. J. Quantum Chem.* **2012**, *112*, 3174–3187. [CrossRef]

92. Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Local Vibrational Modes of the Water Dimer—Comparison of Theory and Experiment. *Chem. Phys. Lett.* **2012**, *554*, 243–247. [CrossRef]

93. Kalescky, R.; Kraka, E.; Cremer, D. Local Vibrational Modes of the Formic Acid Dimer—The Strength of the Double H-Bond. *Mol. Phys.* **2013**, *111*, 1497–1510. [CrossRef]

94. Kraka, E.; Freindorf, M.; Cremer, D. Chiral Discrimination by Vibrational Spectroscopy Utilizing Local Modes. *Chirality* **2013**, *25*, 185–196. [CrossRef] [PubMed]

95. Setiawan, D.; Kraka, E.; Cremer, D. Description of Pnicogen Bonding with the help of Vibrational Spectroscopy—The Missing Link Between Theory and Experiment. *Chem. Phys. Lett.* **2014**, *614*, 136–142. [CrossRef]

96. Setiawan, D.; Kraka, E.; Cremer, D. Strength of the Pnicogen Bond in Complexes Involving Group VA Elements N, P, and As. *J. Phys. Chem. A* **2014**, *119*, 1642–1656. [CrossRef] [PubMed]

97. Setiawan, D.; Kraka, E.; Cremer, D. Hidden Bond Anomalies: The Peculiar Case of the Fluorinated Amine Chalcogenides. *J. Phys. Chem. A* **2015**, *119*, 9541–9556. [CrossRef]

98. Kraka, E.; Setiawan, D.; Cremer, D. Re-Evaluation of the Bond Length-Bond Strength Rule: The Stronger Bond Is not Always the Shorter Bond. *J. Comput. Chem.* **2015**, *37*, 130–142. [CrossRef] [PubMed]

99. Zhang, X.; Dai, H.; Yan, H.; Zou, W.; Cremer, D. B-H π Interaction: A New Type of Nonclassical Hydrogen Bonding. *J. Am. Chem. Soc.* **2016**, *138*, 4334–4337. [CrossRef]

100. Setiawan, D.; Cremer, D. Super-Pnicogen Bonding in the Radical Anion of the Fluorophosphine Dimer. *Chem. Phys. Lett.* **2016**, *662*, 182–187. [CrossRef]

101. Oliveira, V.; Kraka, E.; Cremer, D. The Intrinsic Strength of the Halogen Bond: Electrostatic and Covalent Contributions Described by Coupled Cluster Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 33031–33046. [CrossRef]

102. Oliveira, V.; Kraka, E.; Cremer, D. Quantitative Assessment of Halogen Bonding Utilizing Vibrational Spectroscopy. *Inorg. Chem.* **2016**, *56*, 488–502. [CrossRef]

103. Tao, Y.; Zou, W.; Jia, J.; Li, W.; Cremer, D. Different Ways of Hydrogen Bonding in Water—Why Does Warm Water Freeze Faster than Cold Water? *J. Chem. Theory Comput.* **2017**, *13*, 55–76. [CrossRef] [PubMed]

104. Oliveira, V.; Cremer, D. Transition from Metal-Ligand Bonding to Halogen Bonding Involving a Metal as Halogen Acceptor: A Study of Cu, Ag, Au, Pt, and Hg Complexes. *Chem. Phys. Lett.* **2017**, *681*, 56–63. [CrossRef]

105. Oliveira, V.; Cremer, D.; Kraka, E. The Many Facets of Chalcogen Bonding: Described by Vibrational Spectroscopy. *J. Phys. Chem. A* **2017**, *121*, 6845–6862. [CrossRef] [PubMed]

106. Oliveira, V.; Kraka, E. Systematic Coupled Cluster Study of Noncovalent Interactions Involving Halogens, Chalcogens, and Pnicogens. *J. Phys. Chem. A* **2017**, *121*, 9544–9556. [CrossRef] [PubMed]

107. Zou, W.; Zhang, X.; Dai, H.; Yan, H.; Cremer, D.; Kraka, E. Description of an Unusual Hydrogen Bond Between Carborane and a Phenyl Group. *J. Organometall. Chem.* **2018**, *856*, 114–127. [CrossRef]

108. Yannacone, S.; Oliveira, V.; Verma, N.; Kraka, E. A Continuum from Halogen Bonds to Covalent Bonds: Where Do λ^3 Iodanes Fit? *Inorganics* **2019**, *7*, 47. [CrossRef]

109. Lyu, S.; Beiranvand, N.; Freindorf, M.; Kraka, E. Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleosides. *J. Phys. Chem. A* **2019**, *123*, 7087–7103. [CrossRef] [PubMed]

110. Oliveira, V.P.; Marcial, B.L.; Machado, F.B.C.; Kraka, E. Metal-Halogen Bonding Seen through the Eyes of Vibrational Spectroscopy. *Materials* **2020**, *13*, 55. [CrossRef]

111. Tao, Y.; Qiu, Y.; Zou, W.; Nanayakkara, S.; Yannacone, S.; Kraka, E. In Situ Assessment of Intrinsic Strength of X-I$\cdots$OA Type Halogen Bonds in Molecular Crystals with Periodic Local Vibrational Mode Theory. *Molecules* **2020**, *25*, 1589. [CrossRef]

112. Yannacone, S.; Sethio, D.; Kraka, E. Quantitative Assessment of Intramolecular Hydrogen Bonds in Neutral Histidine. *Theor. Chem. Acc.* **2020**, *139*, 125. [CrossRef]

113. Martins, J.; Quintino, R.P.; Politi, J.R.S.; Sethio, D.; Gargano, R.; Kraka, E. Computational Analysis of Vibrational Frequencies and Rovibrational Spectroscopic Constants of Hydrogen Sulfide Dimer using MP2 and CCSD(T). *Spectrochim. Acta A* **2020**, *239*, 118540-1–118540-9. [CrossRef] [PubMed]

114. Yannacone, S.; Freindorf, M.; Tao, Y.; Zou, W.; Kraka, E. Local Vibrational Mode Analysis of π-Hole Interactions between Aryl Donors and Small Molecule Acceptors. *Crystals* **2020**, *10*, 556. [CrossRef]

115. Bader, R.F.W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18*, 9–15. [CrossRef]

116. Bader, R. *Atoms in Molecules: A Quantum Theory*; International Series of Monographs on Chemistry; Clarendon Press: Oxford, UK, 1990.

117. Cremer, D.; Kraka, E. Chemical Bonds without Bonding Electron Density? Does the Difference Electron-Density Analysis Suffice for a Description of the Chemical Bond? *Angew. Chem. Int. Ed.* **1984**, *23*, 627–628. [CrossRef]

118. Cremer, D.; Kraka, E. A Description of the Chemical Bond in Terms of Local Properties of Electron Density and Energy. *Croat. Chem. Acta* **1984**, *57*, 1259–1281.

119. Kraka, E.; Cremer, D. Chemical Implication of Local Features of the Electron Density Distribution. In *Theoretical Models of Chemical Bonding. The Concept of the Chemical Bond*; Maksic, Z.B., Ed.; Springer: Berlin/Heidelberg, Germany, 1990; Volume 2, pp. 453–542.

120. Reed, A.; Curtiss, L.; Weinhold, F. Intermolecular Interactions from A Natural Bond Orbital, Donor-Acceptor Viewpoint. *Chem. Rev.* **1988**, *88*, 899–926. [CrossRef]

121. Weinhold, F.; Landis, C.R. *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*; Cambridge University Press: Cambridge, UK, 2003.

122. Glendening, E.D.; Badenhoop, J.K.; Reed, A.E.; Carpenter, J.E.; Bohmann, J.A.; Morales, C.M.; Landis, C.R.; Weinhold, F. *NBO6*; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, USA, 2013.

123. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16*; Gaussian Inc.: Wallingford, CT, USA, 2016.

124. Chai, J.D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620. [CrossRef]

125. Chai, J.D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, 084106. [CrossRef]

126. Hehre, W.J.; Ditchfield, R.; Pople, J.A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian? Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261. [CrossRef]

127. Ditchfield, R.; Hehre, W.J.; Pople, J.A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728. [CrossRef]

128. Clark, T.; Chandrasekhar, J.; Spitznagel, G.W.; Schleyer, P.V.R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+ G Basis Set for First-Row Elements, Li–F. *J. Comput. Chem.* **1983**, *4*, 294–301. [CrossRef]

129. Frisch, M.J.; Pople, J.A.; Binkley, J.S. Self-Consistent Molecular Orbital Methods 25. Supplementary Functions for Gaussian Basis Sets. *J. Chem. Phys.* **1984**, *80*, 3265–3269. [CrossRef]

130. Gräfenstein, J.; Cremer, D. Efficient Density-Functional Theory Integrations by Locally Augmented Radial Grids. *J. Chem. Phys.* **2007**, *127*, 164113. [CrossRef] [PubMed]

131. Zou, W.; Tao, Y.; Freindorf, M.; Makoś, M.Z.; Verma, N.; Kraka, E. *Local Vibrational Mode Analysis (LModeA)*; Computational and Theoretical Chemistry Group (CATCO), Southern Methodist University: Dallas, TX, USA, 2020.

132. Keith, T.A. *AIMAll (Version 17.01. 25)*; TK Gristmill Software: Overland Park, KS, USA, 2017.

133. Boys, S.F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553–566. [CrossRef]

134. Senn, H.M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 1198–1229. [CrossRef] [PubMed]

135. Delgado, J.L.; Vance, N.R.; Kerns, R.J. Crystal Structure of DNA Dodecamer D(CGCGAATTCGCG). National Institutes of Health/National Institute of Allergy and Infectious Diseases. 2018. Available online: http://www.rcsb.org/structure/6CQ3 (accessed on 10 April 2021).

136. Chung, L.W.; Sameera, W.M.C.; Ramozzi, R.; Page, A.J.; Hatanaka, M.; Petrova, G.P.; Harris, T.V.; Li, X.; Ke, Z.; Liu, F.; et al. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796. [CrossRef] [PubMed]

137. Huang, M.; Li, H.; He, H.; Zhang, X.; Wang, S. An Electrochemical Impedance Sensor for Simple and Specific Recognition of G-G Mismatches in DNA. *Anal. Methods* **2016**, *8*, 7413–7419. [CrossRef]

138. Sun, H.; Bennett, R.J.; Maizels, N. The Saccharomyces Cerevisiae Sgs1 Helicase Efficiently Unwinds G-G Paired DNAs. *Nucleic Acids Res.* **1999**, *27*, 1978–1984. [CrossRef]

139. Mondal, S.; Bhat, J.; Jana, J.; Mukherjee, M.; Chatterjee, S. Reverse Watson-Crick G-G base pair in G-quadruplex formation. *Mol. BioSyst.* **2016**, *12*, 18–22. [CrossRef]

140. Brown, T.; Woodward, P.; Murphy, C.J.; Bursten, B.E.; LeMay, J.H.E. *Chemistry: The Central Science*, 11th ed.; Pearson Prentice Hall: Boston, MA, USA, 2009.

141. Reece, J.B.; Urry, L.A.; Cain, M.L.; Wasserman, S.A.; Minorsky, P.V.; Jackson, R.B. *Campbell Biology*; Pearson: Boston, MA, USA, 2014; Volume 9.

142. Arunan, E.; Desiraju, G.R.; Klein, R.A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D.C.; Crabtree, R.H.; Dannenberg, J.J.; Hobza, P.; et al. Defining the hydrogen bond: An account (IUPAC Technical Report). *Pure Appl. Chem.* **2011**, *83*, 1619–1636. [CrossRef]

143. Desiraju, G.R. A Bond by Any Other Name. *Angew. Chem. Int. Ed.* **2011**, *50*, 52–59. [CrossRef]

144. Horowitz, S.; Trievel, R.C. Carbon-Oxygen Hydrogen Bonding in Biological Structure and Function. *J. Biol. Chem.* **2012**, *287*, 41576–41582. [CrossRef]

145. Nick Pace, C.; Scholtz, J.M.; Grimsley, G.R. Forces Stabilizing Proteins. *FEBS Lett.* **2014**, *588*, 2177–2184. [CrossRef] [PubMed]

146. Cheng, R.; Loire, E.; Fridgen, T.D. Hydrogen Bonding in Alkali Metal Cation-Bound i-Motif-Like Dimers of 1-Methyl Cytosine: An IRMPD Spectroscopic and Computational Study. *Phys. Chem. Chem. Phys.* **2019**, *21*, 11103–11110. [CrossRef] [PubMed]

147. Karas, L.J.; Wu, C.H.; Ottosson, H.; Wu, J.I. Electron-Driven Proton Transfer Relieves Excited-State Antiaromaticity in Photoexcited DNA Base Pairs. *Chem. Sci.* **2020**, *11*, 10071–10077. [CrossRef]

148. Zhang, Y.; de La Harpe, K.; Beckstead, A.A.; Improta, R.; Kohler, B. UV-Induced Proton Transfer Between DNA Strands. *J. Am. Chem. Soc.* **2015**, *137*, 7059–7062. [CrossRef] [PubMed]

149. Lewis, F.D.; Wasielewski, M.R. Dynamics and Efficiency of Photoinduced Charge Transport in DNA: Toward the Elusive Molecular Wire. *Pure Appl. Chem.* **2013**, *85*, 1379–1387. [CrossRef]

150. Kumar, A.; Sevilla, M.D. Proton-Coupled Electron Transfer in DNA on Formation of Radiation-Produced Ion Radicals. *Chem. Rev.* **2010**, *110*, 7002–7023. [CrossRef] [PubMed]

151. Black, P.J.; Bernhard, W.A. Excess Electron Trapping in Duplex DNA: Long Range Transfer via Stacked Adenines. *J. Phys. Chem. B* **2012**, *116*, 13211–13218. [CrossRef]

152. Gorb, L.; Podolyan, Y.; Dziekonski, P.; Sokalski, W.A.; Leszczynski, J. Double-Proton Transfer in Adenine- Thymine and Guanine-Cytosine Base Pairs. A Post-Hartree- Fock ab Initio Study. *J. Am. Chem. Soc.* **2004**, *126*, 10119–10129. [CrossRef]

153. Li, P.; Rangadurai, A.; Al-Hashimi, H.M.; Hammes-Schiffer, S. Environmental Effects on Guanine-Thymine Mispair Tautomerization Explored with Quantum Mechanical/Molecular Mechanical Free Energy Simulations. *J. Am. Chem. Soc.* **2020**, *142*, 11183–11191. [CrossRef]

154. Shekaari, A.; Jafari, M. Modeling the Action of Environment on Proton Tunneling in the Adenine–Thymine Base Pair. *Prog. Biophys. Mol. Biol.* **2020**, *150*, 98–103. [CrossRef] [PubMed]

155. Marvi, M.; Ghadiri, M. A Mathematical Model for Vibration Behavior Analysis of DNA and Using a Resonant Frequency of DNA for Genome Engineering. *Sci. Rep.* **2020**, *10*, 3439. [CrossRef] [PubMed]

156. Schmidt, B.B.; Hettler, M.H.; Schön, G. Influence of Vibrational Modes on the Electronic Properties of DNA. *Phys. Rev. B* **2007**, *75*, 115125. [CrossRef]

157. Jo, S.; Son, J.; Lee, B.H.; Dugasani, S.R.; Park, S.H.; Kim, M.K. Vibrational Characteristics of DNA Nanostructures Obtained Through a Mass-Weighted Chemical Elastic Network Model. *RSC Adv.* **2017**, *7*, 47190–47195. [CrossRef]

158. Guchhait, B.; Liu, Y.; Siebert, T.; Elsaesser, T. Ultrafast Vibrational Dynamics of the DNA Backbone at Different Hydration Levels Mapped by Two-Dimensional Infrared Spectroscopy. *Struct. Dyn.* **2016**, *3*, 043202. [CrossRef] [PubMed]

159. Poater, J.; Swart, M.; Bickelhaupt, F.M.; Fonseca Guerra, C. B-DNA structure and stability: The role of hydrogen bonding, π-π–stacking interactions, twist-angle, and solvation. *Org. Biomol. Chem.* **2014**, *12*, 4691–4700. [CrossRef]

160. Barone, G.; Fonseca Guerra, C.; Bickelhaupt, F.M. B-DNA Structure and Stability as Function of Nucleic Acid Composition: Dispersion-Corrected DFT Study of Dinucleoside Monophosphate Single and Double Strands. *ChemistryOpen* **2013**, *2*, 186–193. [CrossRef]

161. Kool, E.T. Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replication. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1–22. [CrossRef]

162. SantaLucia, J.; Hicks, D. The Thermodynamics of DNA Structural Motifs. *Ann. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440. [CrossRef]

163. Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M.D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **2006**, *34*, 564–574. [CrossRef]

164. Zhang, T.B.; Zhang, C.L.; Dong, Z.L.; Guan, Y.F. Determination of Base Binding Strength and Base Stacking Interaction of DNA Duplex Using Atomic Force Microscope. *Sci. Rep.* **2015**, *5*, 9143. [CrossRef] [PubMed]

165. Vologodskii, A.; Frank-Kamenetskii, M.D. DNA melting and energetics of the double helix. *Phys. Life Rev.* **2018**, *25*, 1–21. [CrossRef] [PubMed]

166. Florián, J.; Šponer, J.; Warshel, A. Thermodynamic Parameters for Stacking and Hydrogen Bonding of Nucleic Acid Bases in Aqueous Solution: Ab Initio/Langevin Dipoles Study. *J. Phys. Chem. B* **1999**, *103*, 884–892. [CrossRef]

167. Oliva, R.; Cavallo, L.; Tramontano, A. Accurate energies of hydrogen bonded nucleic acid base pairs and triplets in tRNA tertiary interactions. *Nucleic Acids Res.* **2006**, *34*, 865–879. [CrossRef]

168. van Mourik, T.; Hogan, S.W.L. DNA base stacking involving adenine and 2-aminopurine. *Struct. Chem.* **2016**, *27*, 145–158. [CrossRef]

169. Lee, C.; Park, K.H.; Cho, M. Vibrational dynamics of DNA. I. Vibrational basis modes and couplings. *J. Chem. Phys.* **2006**, *125*, 114508. [CrossRef] [PubMed]

170. Svozil, D.; Hobza, P.; Šponer, J. Comparison of Intrinsic Stacking Energies of Ten Unique Dinucleotide Steps in A-RNA and B-DNA Duplexes. Can We Determine Correct Order of Stability by Quantum-Chemical Calculations? *J. Phys. Chem. B* **2010**, *114*, 1191–1203. [CrossRef]

171. Chakraborty, K.; Mantha, S.; Bandyopadhyay, S. Molecular dynamics simulation of a single-stranded DNA with heterogeneous distribution of nucleobases in aqueous medium. *J. Chem. Phys.* **2013**, *139*, 075103. [CrossRef] [PubMed]

172. Brown, R.F.; Andrews, C.T.; Elcock, A.H. Stacking Free Energies of All DNA and RNA Nucleoside Pairs and Dinucleoside-Monophosphates Computed Using Recently Revised AMBER Parameters and Compared with Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 2315–2328. [CrossRef]

173. Karwowski, B.T. The AT Interstrand Cross-Link: Structure, Electronic Properties, and Influence on Charge Transfer in dsDNA. *Mol. Ther. Nucl. Acids* **2018**, *13*, 665–685. [CrossRef] [PubMed]

174. Zacharias, M. Base-Pairing and Base-Stacking Contributions to Double-Stranded DNA Formation. *J. Phys. Chem. B* **2020**, *124*, 10345–10352. [CrossRef] [PubMed]

175. Rice, S.A.; Doty, P. The Thermal Denaturation of Desoxyribose Nucleic Acid. *J. Am. Chem. Soc.* **1957**, *79*, 3937–3947. [CrossRef]

Appendix C

Paper on the formation of CN bonds in Titan's atmosphere

Marek Freindorf, Nassim Beiranvand, Alexis A.A. Delgado. Yunwen Tao and Elfi Kraka, On the formation of CN bonds in Titan's atmosphere - A Unified Reaction Valley Approach study.

has been submitted to the Journal.

On the formation of CN bonds in Titan's atmosphere - A Unified Reaction Valley Approach study

Marek Freindorf · Nassim Beiranvand ·
Alexis A. A. Delgado · Yunwen Tao ·
Elfi Kraka[*]

Abstract In this work we investigated the formation of protonated hydrogen cyanide $HCNH^+$ and methylene amine cation $CH_2NH_2^+$ (both identified in Titan's upper atmosphere) from three different pathways which stem from the interaction between CH_4 and $N^+(^3P)$. As a mechanistic tool we used the Unified Reaction Valley Approach (URVA) complemented with the Local Mode Analysis (LMA) assessing the strength of the CN bonds formed in

Marek Freindorf
Computational and Theoretical Chemistry Group (CATCO),
Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas
75275-0314, United States
Tel.: +1-214-768-1109
E-mail: mfreindorf@smu.edu

Nassim Beiranvand
Computational and Theoretical Chemistry Group (CATCO),
Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas
75275-0314, United States
E-mail: nbeiranvand@smu.edu

Alexis A. A. Delgado
Computational and Theoretical Chemistry Group (CATCO),
Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas
75275-0314, United States
E-mail: aadelgado@smu.edu

Yunwen Tao
Computational and Theoretical Chemistry Group (CATCO),
Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas
75275-0314, United States
E-mail: yunwent@smu.edu

Elfi Kraka
Computational and Theoretical Chemistry Group (CATCO),
Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas
75275-0314, United States
Tel.: +1-214-768-1609
E-mail: ekraka@smu.edu

these reactions. Our URVA studies could provide a comprehensive overview on bond formation/cleavage processes relevant to the specific mechanism of eight reactions **R1** - **R8** that occur across the three pathways. In addition we could explain the formation of $CH_2NH_2^+$ and the appearance of $HCNH^+$ and $CHNH_2^+$ along these paths. Although only smaller molecules are involved these in reactions including isomerzation, hydrogen atom abstraction and hydrogen molecule capture, we found a number of interesting features, such as *roaming* in reaction **R3** or the primary interaction of H_2 with the carbon atom in $HCNH^+$ in reaction **R8** followed by migrating of one of the H_2 hydrogen atoms to the nitrogen which is more cost effective than breaking the HH bond first; a feature often found in catalysis. In all cases, charge transfer between carbon and nitrogen could be identified as a driving force for the CN bond formation. As revealed by LMA the CN bonds formed in reactions **R1** - **R8** cover a broad bond strength range from very weak to very strong, with the CN bond in protonated hydrogen cyanide $HCNH^+$ identified as strongest of all molecules investigated in this work. Our study demonstrates the large potential of both, URVA and LMA to shed new light into these extraterrestrial reactions to help better understand prebiotic processes as well as develop guidelines for future investigations involving areas of complex interstellar chemistry. In particular the formation of CN bonds as a precursor to the extraterrestrial formation of amino acids [110, 111] will be the focus of future investigations.

1 Introduction

At least 200 specific molecular constituents including neutral molecules and charged species have been identified so far in the interstellar medium (ISM) [1–5]. The number is constantly increasing because of advanced detection techniques onboard space crafts such as the *Cassini* [6,7], space-based and terrestrial telescopes (ranging from the Spitzer Space Telescope [8], Hubble Space Telescope [9], the James Webb Telescopes [10] to the planned Giant Magellan Telescope [11] (just to name a few), being complemented with sophisticated terrestrial experimental investigations [12–14]. The interstellar chemistry involving these species, which is often triggered by intense UV light, has recently attracted a lot of attention [15–19]. One area of interest pertains to the formation of organic aerosols (tholins) [20], as such molecules have been considered possible prebiotic precursors of nucleic acids, proteins, amino sugars, and other compounds [17,21–23]. It is assumed that early earth was a tholin-rich place [24] with thiolin-like materials having been found on Saturn's largest moon, Titan [4,25–27], Neptune's largest moon, Triton [20], and on smaller icy bodies [28,12].
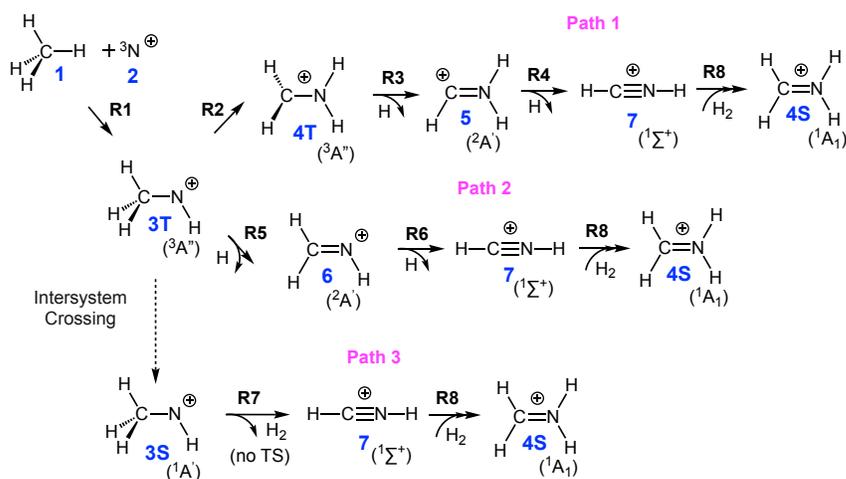
Fig. 1: Reaction pathways investigated in this work.

Titan's dense atmosphere consists primarily of N_2 and a small percentage of CH_4 [29] which have been presumed to have formed thiolins. In Titan's upper atmosphere, Cassini's Ion and Neutral Mass Spectrometer (INMS) detected neutral and positive ion signatures [30] primarily for protonated hydrogen cyanide (i.e., $HCNH^+$, iminomethyl cation) [31,32], the methylene amine cation (i.e, $CH_2NH_2^+$) [33–35], the methanimine cation (i.e., CH_2NH^+), and its isomer the aminomethylene cation (i.e., $CHNH_2^+$) [36]. A key component

for the formation of HCNH$^+$ has been attributed to the ionic dissociation of N$_2$ into N$^+$(^3P) via extreme ultraviolet radiation [15], ions, and energetic electrons [31,32]. High-pressure mass spectrometry experiments, in combination with computational studies, have revealed that N$^+$(^3P) reacts with CH$_4$ to form HCNH$^+$ [37]. However, the formation of CH$_2$NH$_2$$^+$ and the appearance of HCNH$^+$ and CHNH$_2^+$ along this path has yet to be comprehensively understood.

From the reaction between CH$_4$ (**1**) and N$^+$(^3P) (**2**) three different pathways for the synthesis of HCNH$^+$ (**7**) and CH$_2$NH$_2$ (**4S**) result (see Figure 1). In this work we investigate these three pathways. *Path 1* includes the reactions **R2**, **R3**, **R4**, and **R8**, *Path 2* includes reactions **R5**, **R6**, and **R8**. In addition to these two paths we also consider *Path 3* which starts from the singlet state of CH$_3$NH$^+$ (**3S**), resulting from intersystem crossing of the triplet state CH$_3$NH$^+$ (**3T**) , and includes reactions **R7** and **R8**. The main focus of this study is to explore the detailed mechanism of the formation of HCNH$^+$ (**7**) and CH$_2$NH$_2$$^+$ (**4S**) using the Unified Reaction Valley Approach (URVA) [38] and to quantitatively assess and compare the strength of the CN bonds formed in these reactions through local vibrational mode analysis [39]. Our results will give new insights into these extraterrestrial reactions to help better understand prebiotic processes as well as develop guidelines for future investigations involving areas of complex interstellar chemistry.

2 Computational Method

The main tool applied in this work is URVA [38,40,41]. A comprehensive review can be found in Ref. [38] therefore, in the following only the key essentials will be summarized. URVA analyzes the curvature of the reaction path traced out by the reaction complex (i.e. the union of reacting molecules) on the potential energy surface (PES) moving from the entrance channel up to the transition state (TS), from here, down into the product channel. Any electronic structure change including bond breaking/forming processes, re-hybriziation, charge polarization and transfer, etc. lead to distinct curvature peaks which are directly reflected in the scalar reaction path curvature calculated at each path point; whereas regions of minimal electronic change are reflected by curvature minima where the region from one curvature minimum to the next embedding a curvature peak is called a reaction phase. Each chemical reaction has a unique pattern of curvature maxima and minima with a different number of reaction phases, which can be used as its characteristic *fingerprint* [38]. We then decompose the reaction curvature along the reaction path into individual components such bond lengths, bond angles, and dihedral angles or puckering coordinates to get a deeper understanding of each chemical event [40].The sign of a component indicates if the parameter in question supports the chemical event (positive sign) or if it resists the chemical change (negative sign) [40]. For a detailed mathematical derivation and recent advances of URVA, interested readers are referred to Ref. [38].

Another important tool used for the assessment of the CN bond strength in molecules **3** - **7** and TSs **TS1** - **TS6** and **TS8** of the corresponding reactions **R1** - **R6** and **R8** investigated in this work is the Local Vibrational Mode Analysis (LMA). A comprehensive review of the underlying theory can be found in Ref. [39]. Therefore, in the following only some essentials are summarized. Normal vibrational modes are generally delocalized as a result of kinematic and electronic coupling [42–44] In other words, if one considers a particular normal stretching mode between two atoms of interest, it can be coupled to other normal modes such as bending or torsion, which inhibits the direct correlation between stretching frequency and bond strength as well as the comparison between stretching modes in related molecules. As a consequence, the normal stretching force constant cannot be used as a direct bond strength measure. One needs to derive local counterparts that are free form any mode-mode coupling. Konkoli, Cremer and co-workers solved this problem by solving the mass-decoupled analogue of Wilson's equation of vibrational spectroscopy [45–49] leading to local vibrational modes, associated local mode frequencies and local mode force constants. Zou and Cremer showed that the local stretching force constant k^a reflects the curvature of the PES in the direction of the bond stretching [50]. This important result qualifies the local stretching force constants k^a as unique quantitative measures of the intrinsic strength of chemical bonds and/or weak chemical interactions based on vibrational spectroscopy, which has been extensively applied in previous work as documented in Ref. [39]. For some more recent work see also Refs. [51–57].

Geometry optimizations and frequency calculations including a local mode analysis for all stationary points (i.e., reactants, products and TSs) of each reaction **R1** - **R8** were performed with second Møller-Plesset perturbation theory (MP2) [58,59] and complete active space perturbation theory of second order (CASPT2) [60–62] using Dunning's aug-cc-pVTZ basis set [63] as well as with Coupled Cluster theory [64–66] at the CCSD(T) level [67–69] using Dunning's cc-pVDZ basis set [63]. The active space in the CASPT2 calculations included 4-electrons with 4-orbitals, 5-electrons with 5-orbitals, and 6-electrons with 6-orbitals, depending on the molecular system. For the reaction path calculations and URVA analysis MP2/aug-cc-pVTZ was applied. For the reaction path the intrinsic reaction coordinate (IRC) of Fukui [70] was used with a step size of $s = 0.03$ amu$^{1/2}$ Bohr, applying the improved reaction path following the algorithm of Hratchian and Kraka, enables one to follow a chemical reaction far out into entrance and exit channel [71]. All IRC calculations were performed with the Gaussian16 program package [72] at the MP2/aug-cc-pTZ level of theory using a tight convergence criterion. The CASPT2 calculations were performed with MOLPRO [73–75]. For the CCSD(T) calculations the CFOUR program package [66] was utilized. The URVA analysis was carried out with the program pURVA [76] and the local mode analysis with the program LModeA [77]. The computation and analysis of atomic charges was done with the NBO program of Weinhold and co-workers[78–80].

3 Results and discussion

In the following, the energetics of reactions **R1** - **R8**, their specific mechanism as revealed by URVA, and the bond strength of all CN bonds formed during these reactions via local vibrational force constants will be discussed. In addition to tables and figures presented in the manuscript, further information on geometries, NBO charges, and reaction movies are collected in the Supporting Information.

3.1 Energetics Results

Table 1 presents activation energies E^a, activation enthalpies H^a, and activation free energies G^a, as well as reaction energies E_r, reaction enthalpies H_r, and reaction free energies G_r, calculated with CASPT2/aug-cc-pVTZ and CCSD(T)/cc-pVDZ model chemistries. The corresponding MP2/aug-cc-pVTZ results can be found in the Supporting Information. The energetics presented in Table 1 were calculated in each case relative to the reactant of the corresponding reaction. Although both model chemistries have a different focal point, i.e. CASPT2 on multireference character [81] and CCSD(T) on dynamic correlation [64,66,65], results are quite similar, confirming that both approaches lead to a reliable description of reactions **R1** - **R8**. The following discussion is limited to the CCSD(T) results.

Table 1: Activation energies E^a, enthalpies H^a, and free energies G^a; Reactions energies E_r, enthalpies H_r, and free energies G_r (kcal/mol) of the reaction **R1** - **R8** investigated in our study. Optimal geometries and frequencies calculated at the CASPT2/aug-cc-pVTZ and CCSD(T)/cc-pVDZ levels of theory.

Reaction	CASPT2/aug-cc-pVTZ						CCSD(T)/cc-pVDZ					
	E^a	E_r	H^a	H_r	G^a	G_r	E^a	E_r	H^a	H_r	G^a	G_r
R1[1]	1.62	-68.5	0.72	-66.83	0.62	-66.57	3.78	-66.14	1.97	-64.64	3.08	-63.95
R2	42.62	-7.82	39.42	-7.46	40.13	-7.49	46.12	-4.31	42.89	-3.91	43.53	-3.51
R3	42.50	40.09	38.58	36.10	39.07	29.46	41.20	37.39	37.19	31.56	37.35	26.26
R4	49.55	38.69	42.97	32.35	42.56	26.07	50.93	37.84	44.46	31.56	44.06	25.27
R5	39.50	39.24	34.16	33.84	34.26	27.00	40.86	38.61	35.6	33.29	35.73	26.50
R6	35.48	31.72	30.51	27.15	30.09	21.05	38.14	32.21	33.26	27.61	32.87	21.48
R7	-	-72.53	-	-74.00	-	-81.37	-	-67.94	-	-71.61	-	-73.24
R8	42.23	-54.86	45.64	-46.02	53.16	-38.35	39.60	-55.87	43.18	-47.14	50.70	-39.50

[1] Values relative to the van der Waals complex.

Reaction **R1** starts from a van der Waals complex (as shown in Figure 2a) which is formed directly from methane (**1**) and $N^+(^3P)$ (**2**) in a barrier less, strongly exothermic reaction with a reaction energy of -88.34 kcal/mol. From the van der Waals complex product **3T** is formed with a small activation energy of only 3.78 kcal/mol. The reaction energy with -66.14 kcal/mol indicates the

strong exothermic character of **R1**. In summary, the formation of **3T** from **1** and **2** leads to an overall energy gain of 155 kcal/mol; this large amount of excess energy facilitates all following reactions and counterbalances activation energies in the range of 40 - 50 kcal/mol as found for **R2** - **R6** and **R8**.

As shown in Figure 1, **3T** serves as a reactant for three different pathways, which start either on a triplet (*Path 1* and *Path 2*) or directly on a singlet potential energy surface (PES) after intersystem crossing (*Path 3*). Reaction **R2** is the first reaction of *Path 1* where **3T** undergoes isomerization to **4T** with an activation energy of 44.36 kcal/mol, and a reaction energy of -5.79 kcal/mol. By loosing a hydrogen atom **4T** transforms in reaction **R3** into **5** with an activation and reaction energy of 41.20 and 37.39 kcal/mol, respectively, i.e. this reaction is strongly endothermic. The next step of this pathway is reaction **R4**, in which **5** looses another H atom transforming into **7** with an activation and reaction energy of 50.93 and 37.84 kcal/mol, respectively, revealing again the strong endothermic character of this reaction.

The alternative *Path 2*, starting from **3T**, proceeds directly to loosing a hydrogen atom and transforming into **6** in reaction **R6**. The activation and the reaction energies are 40.86 and 38.14 kcal/mol, respectively, comparable with the energetics found for reaction **R3**. The next step along *Path 2* is reaction **R6**, where the loss of another hydrogen atom from **6** leads to **7**, with the activation and reaction energy values of 38.14 and 32.21 kcal/mol. Comparing the number of reactions in both path ways, the cumulative activation energies and endothermicities, *Path 2* seems to be more favorable.

In contrast to *Path 1* and *Path 2*, *Path 3* starts on a singlet PES. **3T** undergoes intersystem crossing forming **3S**. However, **3S** is unstable (ST splitting: 35.59 kcal/mol) and transforms directly into **7** via H_2 abstraction in barrierless reaction **R7** with a reaction energy –67.94 kcal/mol, definitely favoring *Path 3* as the most likely path.

Protonated hydrogen cyanide **7**, one of the key species found in Titan's atmosphere being formed in all three pathways, can then transform into the methylene amine cation **4S**, another key species in Titan's atmosphere via the addition of H_2 in reaction **R8**. The activation and the reaction energies are 39.60 and -55.87 kcal/mol, respectively, revealing the strong exothermic character of this chemical transformation. Whereas the discussion of the energetics provides valuable overall trends, it cannot lead to specific insights into the actual reaction mechanism, due to the nature of the energy being a cumulative property [82–87]. Therefore, in the following a detailed mechanistic study will be presented based on URVA.

3.2 Reaction Mechanisms

The mechanism of each of the reactions **R1** - **R8** will be described in detail, in particular in terms of the energy profiles, curvature profiles and their decomposition as well as changes of important geometry parameters and changes of the NBO charges monitoring charge transfer along the reaction parameter s.

The decomposition of the reaction path direction along the reaction parameter s for each reaction can be found in the Supporting Information. In addition, reaction movies following the changes of the reaction complex along the reaction parameter s for each of the reactions are included in the Supporting Information, facilitating the discussions.
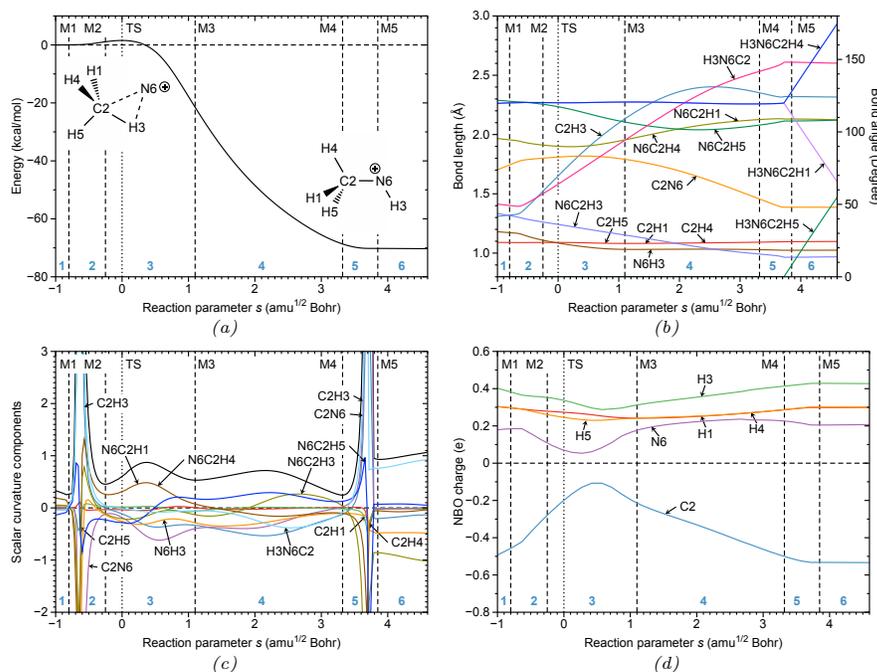

Fig. 2: Reaction **R1**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

Reaction **R1** describes the isomerization of the van der Waals complex originally formed from **1** and **2** to methylamino radical cation **3T**. As shown in Figure 2a in the van der Waals complex the migrating methyl hydrogen atom H3 is in a bridging position between carbon atom C2 and nitrogen atom N6 (R(C2N6) = 1.693 Å, R(C2H3 = 1.308 Å, R(N6H3) = 1.183 Å). During the reaction (see reaction movie **SR1**, Supporting Information) H3 swings around to the nitrogen atom, the C2H3 bond is cleaved and the new C2N6 and N6H3 bonds are formed. Furthermore, the molecules rotate from an eclipsed into a more stable staggered conformation. Figure 2c shows the changes of the most relevant geometry parameters along with the reaction parameter s. It

is interesting to note that the distance of the C2N6 bond to be formed first increases, reaching a maximum value of 1.818 Å at s= 0.4 $amu^{1/2}Bohr$, i.e., shortly after the TS before decreasing to its final value of 1.386 Å at s= 3.7 $amu^{1/2}Bohr$. Another interesting feature of the reaction is the rotation from an eclipsed into the staggered form starting at s= 3.8 $amu^{1/2}Bohr$ and being completed at the end of the reaction path at s= 4.7 $amu^{1/2}Bohr$. Figure 2c, shows the corresponding reaction profile. As expected from the small barrier, the entrance channel is relatively short, (stretching only over 1 path unit) compared to the exit channel of this strongly exothermic reaction (stretching over almost 5 path units). Figure 2b shows the scalar curvature along s, resulting in 5 curvature minima, which divide the reaction path into 6 reaction phases with phase 3 containing the TS. Only chemical events before the TS account for the reaction barrier, i.e., the large curvature peak in phase 2 is composed of the events determining the barrier which are further characterized by decomposition of the reaction path curvature into components. The cleavage of the already weakened C2H3 bond starts here, characterized by supportive contributions (positive values) of the C2H3 bond length and N6C2H3 angle, whereas the formation of the new C2N6 and N3H3 bonds is resisting (negative values). Interesting to note is that there is no pronounced peak at the TS, in accordance with our frequent findings that important chemical changes do not necessarily occur at the TS, they may occur in the entrance [88,89] or exit channel of a reaction [90,40]. The smaller curvature peaks in phases 3 and 4 are dominated by CNH angle contributions describing the migration of the H3 atom accompanied with the bond breaking/forming processes. The large curvature peak in phase 5 far out in the exit channel characterizes the simultaneous finalization of the C2H3 bond breaking and C2N6 and N6H3 bond formation events, accompanied by the rotation of the molecule into the final staggered conformation. Stretching over the whole reaction rather than occurring in an abrupt fashion is more energy conserving, a phenomenon which we also have frequently observed for other chemical reactions, in particular catalysis reactions [91,92]. Changes of the NBO charges as a function of s, reflecting charge-transfer, are shown in Figure 2d At the beginning and at the end of the reaction the N atoms have a positive charge of 0.2 e, i.e. the remaining 0.8 e are distributed over the methyl group, the carbon atom has a negative charge of -0.5 e (compared to -0.7 e in methane), each of the spectator hydrogen atoms have a positive charge of 0.3 e (compared to 0.18 e in methane) and the migrating H atom has the largest positive charge of 0.4 e. Whereas the charges on the H-atoms remain fairly constant during the course of the reaction, there is a transfer of negative charge from the C to the N atom with a maximum reached at s= 0.4 $amu^{1/2}Bohr$ (charge on C2: -0.1 e and on N6: 0.3 e with Δ(NBO) = 0.4 e compared to Δ(NBO) = 0.7 e at the start). This weakens the electrostatic interaction between the two atoms. s= 0.420 $amu^{1/2}Bohr$ is also the point where the maximum CN distance is reached. After that point, negative charge is transferred back to the C atom, Δ(NBO) and with that the bond polarity increases again supporting the formation of the CN bond. Figure 1 depicts the central role of the methylamino radical cation **3T** which

serves as a starting point for each of the following 3 pathways leading to the formation of protonated hydrogen cyanide HCNH$^+$ **7**; *Path 1* with reactions **R2** - **R4**, *Path 2* with reactions **R5** and **R6**, and *Path 3* with reaction **R7**.
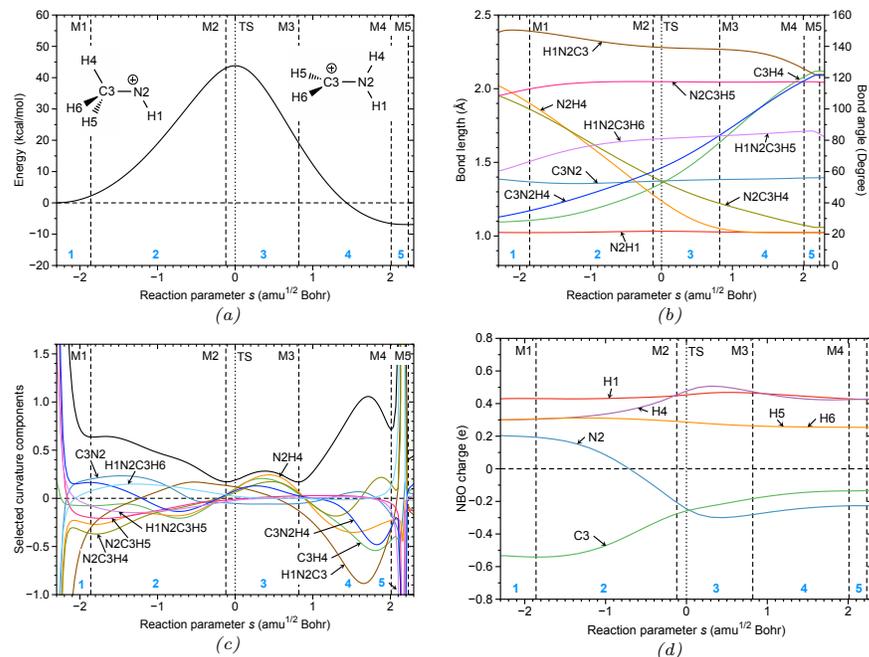

Fig. 3: Reaction **R2**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

Reaction **R2** describes the isomerization of the methylamino cation **3T** to the methylene amino cation **4T**. The reaction is only slightly exothermic by about 6 kcal/mol leading to entrance and exit channels of about the same lengths (2.3 $amu^{1/2}Bohr$). During this reaction, the H4 atom of the methyl group migrates from the C3 to the N2 atom, the C3H4 bond is cleaved and a new N2H4 bond is formed (see Figure 3a and reaction movie **SR2**, Supporting Information). The carbon atom changes from sp^3 to sp^2 hybridization. However, since the methyl and amino hydrogen atoms of the resulting **4T** are arranged almost perpendicular to each other, (H1N2C3H5=81°), the formation of a CN double bond is not possible. As shown in Figure 3b, the CN bond distance remains fairly constant during the whole reaction (1.386 Å in **3T** compared to 1.396 Å in **4T**) which corresponds to a typical CN single bond

distance. Most changes are observed for the N2H4 distance starting from 2.025 Å in **3T** and reaching its final value of 1.031 Å at s= 1.1 $amu^{1/2}Bohr$ and the C3H4 distance, starting from 1.094 Å in **3T** and reaching its final value of 2.118 Å at the end of the reaction at $s=$ 2.3 $amu^{1/2}Bohr$. At **TS2** the migrating hydrogen atom is in a bridging position, with a C3H4 distance of 1.358 Å, a N2H4 distance of 1.239 Å, and a C3N2H4 angle of 62°. Figure 3c presents the scalar reaction path curvature and its decomposition into components along s, which are characterized by six reaction phases. Interesting to note that also for this reaction, there is no significant curvature at the TS, major events occur in both entrance and exit channel. The large curvature peaks in phase 1 as well as in phase 6 are predominantly characterized by rotations of the methyl hydrogens H5 and H6 adjusting to the rehybrizidation, at the carbon atom. The cleavage of the C3H4 bond stretches over phases 3-5, starting with a small supporting contribution of the C3H4 component in phase 3, a small resisting contribution in phase 4, and a larger resisting contribution in phase 5 in which the C3H4 distance of 2.117 Å is reached as confirmed by Figure 3b. With the breakage of the C3H4 bond, the new N2H4 bond starts to form, reaching the final distance of 1.023 Å already in phase 4. As reflected by the large curvature contributions shown in Figure 3c, the H1N2C3 angle plays a central role for the reaction. Changing from 147° to 122° helps to push the carbon hydrogens H5 and H6 into an sp^2 position and make room at the nitrogen for the incoming H4. This is also visualized in the reaction movie **SR2** of the Supporting Information. Figure 3d shows the changes of the NBO charges as a function of s. Starting in phase 2 there is a large transfer of negative charge from the C to the N atom. The original difference Δ(NBO) decreases from 0.74 e to zero at the TS, where both C and N have a negative charge of -0.25 e. After that point the negative charge of N is slightly increasing and that of the C atom is slightly decreasing to the final values of -0.23 e for N and -0.13 e for C in **4T**, where the positive charge is equally distributed among the hydrogen atoms (charges on H5 and H6 0.25 e, respectively and on H1 and H4 0.43 e, respectively).

In reaction **R3** the methylene amino cation **4T** looses one hydrogen atom at the carbon atom and transforms into the planar aminomethylene cation **5** with a CN double bond. C2H6 bond breakage is initiated by a rotation of both the carbon and nitrogen hydrogens from an almost perpendicular to planar arrangement plane (see Figure 4a and reaction movie **SR3**, Supporting Information). As revealed by Figure 4b, the largest geometry changes in the entrance channel include the dihedral angles H1C2N3H4 approaching 180° and H1C2N3H5 approaching 0° during the course of the reaction. At s= -1.5 $amu^{1/2}Bohr$ the rotation has has reached a critical point (H1C2N3H4= 138° and H1C2N3H5=50°) and the C2H6 bond starts to lengthen almost linearly. The C2N3 bond distances decreases from 1.396 Å in **4T** to its final value of 1.250Å at the path endpoint s= 2.0 $amu^{1/2}Bohr$. Figure 4c shows the reaction path curvature which is composed of 6 reaction phases and its decomposition into components. The curvature peaks in phase 1 are characterized by methyl
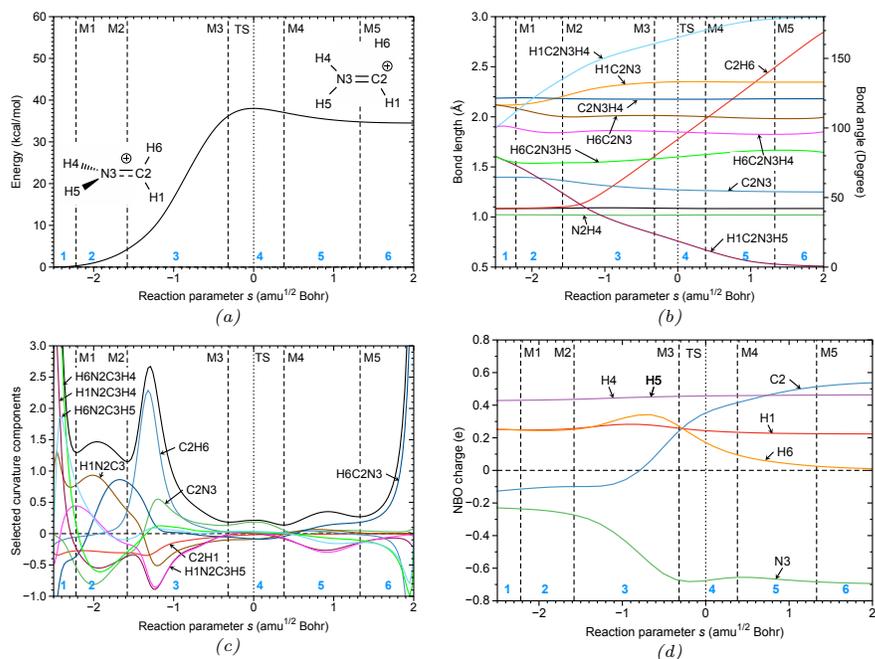
Fig. 4: Reaction **R3**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

and amino hydrogen rotations which are complemented in phase 2 by changes of the H1C2N3 and at the end of this phase by changes of the H6C2N3 bond angle. In phase 3 the cleavage of the C2H6 bond takes place as indicated by a large supporting contribution of the C2H6 component with a smaller contribution from C2N3. The curvature maximum occurs at s= -1.3 $amu^{1/2}Bohr$ where the C2H6 distance starts to increase (see Figure 4b). In summary, the energy demanding cleavage of a CH bond is assisted by methyl and amino hydrogen rotations guiding the CH$_2$ group in a position facilitating the bond breaking process. Phase 4 including **TS3** and phase 5 show only small curvature enhancements predominately made up from C2N3 bond adjustments and further hydrogen bond rotations. Phase 6 is dominated by final H6C2N3 changes and rotations about the CN bond needed to achieve the planar structure of resulting **5**. Figure 4d depicts changes of the NBO charges along s. Major changes occur for the carbon and nitrogen atoms. In phase 3, negative charge is transferred from C2 to N3 where at the end of phase 3 N3 almost reaches its final charge of -0.7 e and C2 adapts a positive charge of 0.25 e. In

the following phases 4-6, i.e. to the end of the reaction, additional negative charge is transferred from the C2 atom to the leaving H6 atom bringing its charge to zero, a small portion is also transferred to H1. At the end of the reaction with NBO charges of -0.7 e on N and 0.54 e on C2 a strongly polar CN double bond results in **5**.

Fig. 5: Reaction **R4**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

In reaction **R4** the aminomethylene cation **5** looses an additional amino hydrogen atom leading to the protonated hydrogen cyanide **7**. In resemblance with reaction **R3**, here the NH bond breakage is assisted by rotating the C2H1 bond and the N3H4 bonds into a linear arrangement which facilitates the cleavage of the N3H5 bond (see Figure 5a and reaction movie **SR4**, Supporting Information). The reaction movie also shows an interesting feature of this reaction. Hydrogen atom H5 orbits around the NH end instead of leaving on a more straight pathway (i.e. so-called *roaming*) which causes the long exit channel of almost eight s units. For a recent review on the phenomenon of roaming reactions, their experimental capture and potential importance of as-

trochemistry, see e.g. Refs. [93–95]. Figure 5b depicts changes of bond lengths and bond angles as a function of s. The H1C2N3 angles starts to widen from $133°$ in **5** with a steep incline until s=-1.1 $amu^{1/2}Bohr$, where a value of $162°$ is reached. After that path point the H1C2N3 angle slowly increases to its final value of $180°$. Starting from a value of $121°$ the C2N3H4 angle first moves through a shallow minimum, then steeply increases and reaches the value of $178°$ at s=2.0 $amu^{1/2}Bohr$. The N3H5 bond length starts to linearly increase at s=-1.2 $amu^{1/2}Bohr$ until it reaches a maximum value of 3.111 Å at s=3.3 $amu^{1/2}Bohr$, where the C2N3H5 angle adapts a value of $120°$. From there on the roaming of H5 starts, leading to a small decrease of the N3H5 distance to 2.896 Å and a steep incline of the C2N3H5 angle to its final value of $180°$. The C2N3 bond decreases slightly in the entrance channel and shortly after **TS4** at s=0.5 $amu^{1/2}Bohr$ attains it final value of 1.144 Å corresponding to a typical CN triple bond. In addition to roaming, reaction **R4** has another interesting feature, more frequency observed in transition metal catalysis [96], the involvement of formation of a dihydrogen interaction [97,98] in the reaction mechanism. The H4H5 distance first decreases to 1.687 Å at s=-1.1 $amu^{1/2}Bohr$, the point where N3H5 bond breakage starts, adapts a maximum value of 2.767 Å at s=3.0 $amu^{1/2}Bohr$ and then decreases again caused by the roaming motion of the H5 atom (see reaction movie **SR4**, Supporting Information). The reaction path curvature profile and its decomposition shown in Figure 5c quantifies these observations. The reaction is composed of seven distinct phases. Phase 1 is characterized by angle rotations, then in phase 2 the N3H5 bond starts to break, supported by H4H5 interactions leading to a large curvature peak. Another important contribution to this curvature peak is negative contribution of the H1C2N3 angle, resisting to become linear. The curvature peak in phase 3 shortly after the TS is dominated by the C2N3 contribution. In phase 6, the cleavage of the N3H5 bond is finalized with a large resisting contribution of the N3H5 and H4H5 components and supported by the C2N3H5 angle contribution. From there one H5 roaming is taking place, which does not lead to any pronounced curvature. Figure 5d presents the changes of the NBO charges along s. Most pronounced is the transfer of charge from the nitrogen atom (-0.7 e at the start of the reaction, -0.3 e at the end) to the leaving H5 atom, which starts with a positive charge of 0.5 e and becomes neutral at around s=1.0 $amu^{1/2}Bohr$.

Reaction **R5** is the first reaction of *Path 2* which leads to (7) in only two reaction steps. Starting from **3T** the planar methylene amine cation **6** is formed via the loss of one methyl hydrogen atom (see Figure 6a and reaction movie **SR5**, Supporting Information). As it is revealed in Figure 6b the C3H4 bond starts to elongate at s=-1.8 $amu^{1/2}Bohr$, then the C3H4 increases linearly over the remaining reaction until at s=2.0 $amu^{1/2}Bohr$ where the final value of 2.936 Å is reached. The N2C3 bond length decreases from 1.386 Å to 1.221 Å to transform from a single into a double bond. The most notable additional change is that of the H1N2C3H6 dihedral angle decreasing from $60°$ to $0°$ accompanied with the rehybridization of the carbon atom from sp^3 to
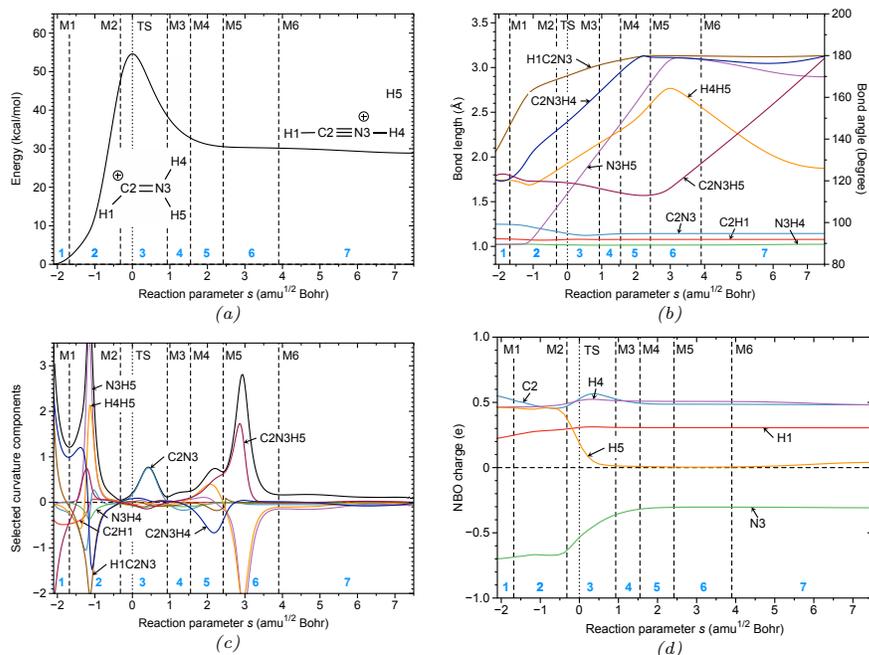
Fig. 6: Reaction **R5**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

sp^2 and the planarization of the molecule. Overall, most changes occur at the methyl group as shown in the reaction movie. Figure 6c shows the reaction path curvature profile and its decomposition. In contract to previous reactions, the curvature profile is relatively simple. There are 5 phases with two enhanced curvature peaks in phases 2 and 3, then a long phase 4 including **TS5** without noticeable curvature, i.e. without any major chemical events, and phase 5 with a pronounced curvature peak characterizing the finalization of this reaction. Phase 2 is characterized by H1N2C3 and N2C3H5 angle changes helping to get the C3H4 cleavage started, which occurs in phase 3, denoted by a supporting C3H4 components the C3N2 bond changes its character, denoted by a supporting C3N2 component. C3H4 cleavage is completed in phase 5 characterized by a large resisting C3H4 contribution and a supporting N2C3H4 contribution. In Figure 6d changes of the NBO charges along s are shown. As in reaction **R3** negative charge is transferred from the carbon atom the nitrogen and to H4, which has to become neutral. However, compared to reaction **R3** the resulting CN double bond in **6** is less polar with 0.09 e on C3 and -0.15 e on N.
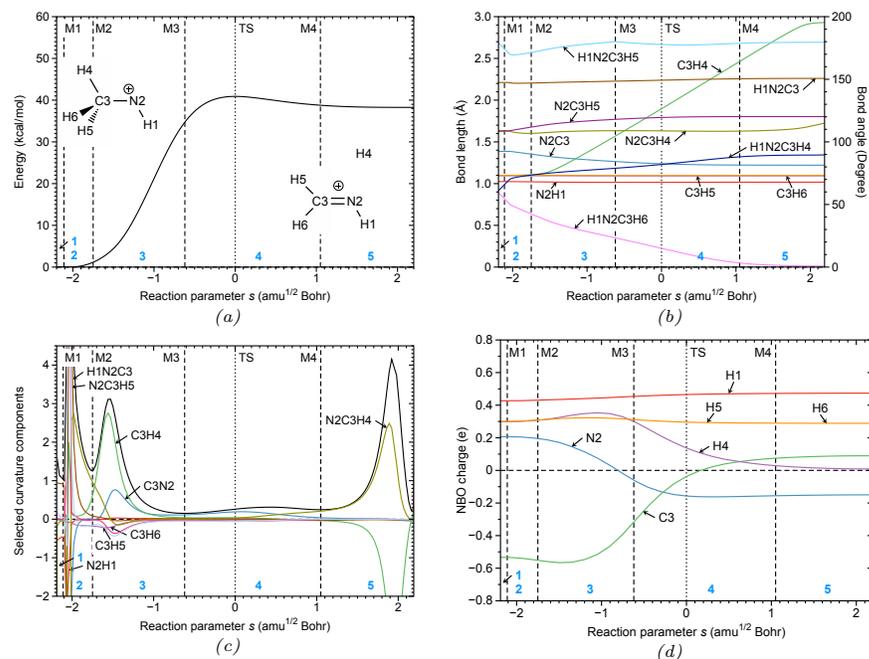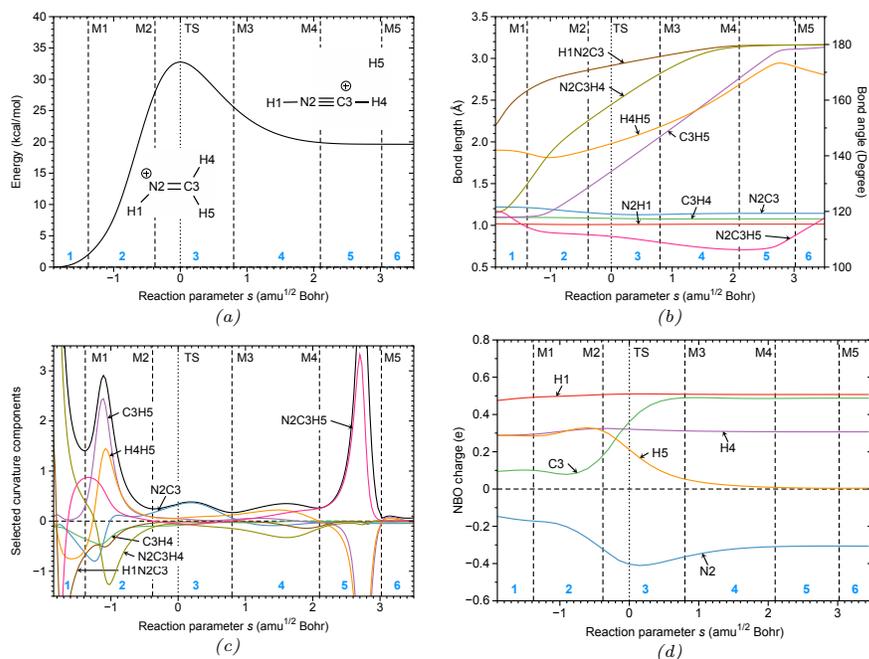
Fig. 7: Reaction **R6**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

Reaction **R6** can be considered as a counterpart of reaction **R4**, both aiming at the formation of protonated hydrogen cyanide **7**, the difference is that in this reaction the methylene loses a hydrogen atom. The CH bond breakage is assisted by rotating the H1N2 and the C3H4 bonds into a linear arrangement which facilitates the cleavage of the C3H5 bond (see Figure 11a and reaction movie **SR6**, Supporting Information). As shown in Figure 11b the H1N2C3 angle increases from 151° and the N2C3H4 angle from 120° in **6** to reach their final values of 180° at s=2.1 $amu^{1/2}Bohr$. The C3H5 bond starts to elongate at s=-1.2 $amu^{1/2}Bohr$, the C3H5 distance increases linearly until reaching the value of 3.1 Å at s=2.8 $amu^{1/2}Bohr$, which is close to the final bond length. Interesting to note again is the assistance of a dihydrogen interaction. The H4H5 distance of 1.901 Å in **6** first decreases to a minimum value of 1.812 Å at s=-1.0 $amu^{1/2}Bohr$, then increases to a maximum value of 2.944 Å at s=2.8 $amu^{1/2}Bohr$ and then decreasing again, showing the onset of a roaming motion which is however less pronounced than in in reaction **R4**, as shown in the reaction movie, which is also reflected by a much smaller change of the N2C3H5 angle compared to the more drastic C2N3H5 angle changes

in reaction **R4**. The reaction path curvature pattern of both reactions look quite similar, there are three distinct curvature peaks, two in the entrance and one in the exit channel. The big difference is that reaction **R6** has only six reaction phases and a shorter exit channel of 3.5 s units (see Figure 7c) than **R4** because of the less pronounced roaming. Phase 1 is characterized by angle rotations, then in phase 2 the C3H5 bond starts to break, supported by H4H5 interactions leading to a large curvature peak. Another important contribution to this curvature peak is negative contribution of the H1N2C3 angle, resisting to become linear. The curvature peak in phase 3 shortly after the TS is dominated by the N2C3 contribution. In phase 6, the cleavage of the C3H5 bond is finalized with a large resisting contribution of the C3H5 and H4H5 components and supported by the N2C3H5 angle contribution. The short phase 6 describes the onset of a roaming motion of the leaving H5 atom. Figure 6d presents the changes of the NBO charges along s. In contrast to reaction **R4** here the neutralization of the leaving hydrogen atom H5 occurs via transfer of negative charge from the carbon atom.

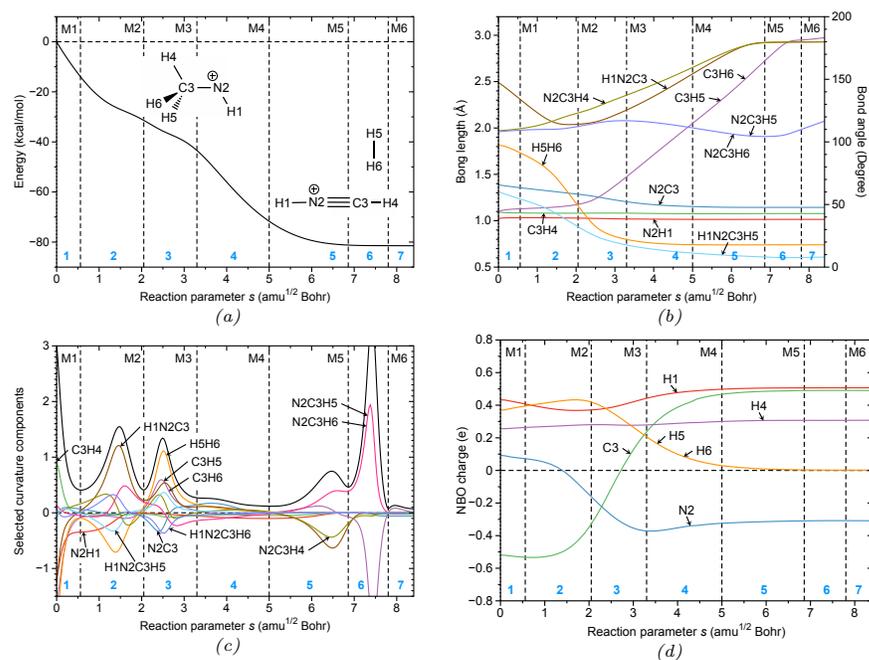Fig. 8: Reaction **R7**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. MP2/aug-cc-pVTZ level of theory.

Path 3 offers an attractive alternative for the formation of **R7** on the singlet PES via just one barrier-less reaction. As sketched in Figure 1, after intersystem crossing of **3T** into **3S**, **7** is formed via the loss of H_2. The out of plane H5 and H6 methyl atoms (see Figure 8a) leave simultaneously as depicted in reaction movie **SR7** (see Supporting Information) and combine to a hydrogen molecule. The process is initiated by the amino hydrogen atom, which moves towards the carbon atom helping to kick out H5 and H6 before swinging back into its final linear position in **R7**. As shown in Figure 8b both C3H5 and C3H6 distances change simultaneously, up to $s=2.0$ $amu^{1/2}Bohr$ more gradually, and thereafter in a steep linear fashion, until reaching their final value of 2.976 Å at $s=8.4$ $amu^{1/2}Bohr$. At $s=1.8$ $amu^{1/2}Bohr$ the H1N2C3 angle reaches a minimum of 114° before steadily increasing again reaching its final value of 180° at $s=6.9$ $amu^{1/2}Bohr$. The C2C3 single bond of **3S** decreases smoothly from 1.386 Å to 1.144 Å, the triple bond distance of **7**, which is reached after $s=5$ $amu^{1/2}Bohr$. The H1N2C3H5 and H1N2C3H6 dihedral angles of 60° approach a zero value upon leaving of the newly formed H_2 as documented by the decreasing H5H6 distance which reaches the HH distance of 0.742 Å already at $s=5.0$ $amu^{1/2}Bohr$ before the final departure from the carbon atom. The reaction path curvature profile and its decomposition is shown in Figure 8c. The reaction consists of seven distinct reaction phases with four distinct curvature peaks, in line with our previous findings that also reactions without a barrier possess detailed mechanistic features [99–101]. Phases 1 and 2 are dominated by a large negative H5H6 components resisting first the formation of H_2 which is balanced in phase 2 by a large supporting contribution of the H1N2C3 angle. The actual cleavage of the C3H5 and C3H6 bonds and the combination of both leaving hydrogen atoms to form an H_2 molecules start in a concerted fashion in phase 3 with supporting components. Interesting to note is that, as already predicted in Figure 8b, the H_2 formation is completed at the end of phase 4, whereas the final CH bond cleavage occurs in phase 6, as denoted by the large resisting C3H5 and C3H6 components supported by large N2C3H5 and N2C3H6 contributions. The shoulder in phase 5 is characterized by adjustments of the N2C3H4 and H1N2C3 angles to become linear, which again is achieved before final CH bond cleavage. As denoted by the small N2C3 curvature components, the transformation from CN single to triple bond occurs mainly in phases 2 and 3 and is completed at the end of phase 4. Figure 8d shows the change of the NBO charges along s. There is a large transfer of negative charge from the carbon to the nitrogen atom as well as to the leaving hydrogen atoms which both have a zero charge in the H_2 molecule. C3 starts with a negative charge of -0.5 e and ends with a positive charge of 0.5 e, whereas N2 starts with a positive charge of 0.1 e and negative charge of -0.31 e in **R7**.

Protonated hydrogen cyanide **7** formed in each of the three paths can recombine with molecular hydrogen and to the methylene amine cation **4S** which is planar in contrast to **4T**, and therefore can form a CN double bond, see Figure 1. As shown in reaction movie **SR8** (see Supporting Information) a
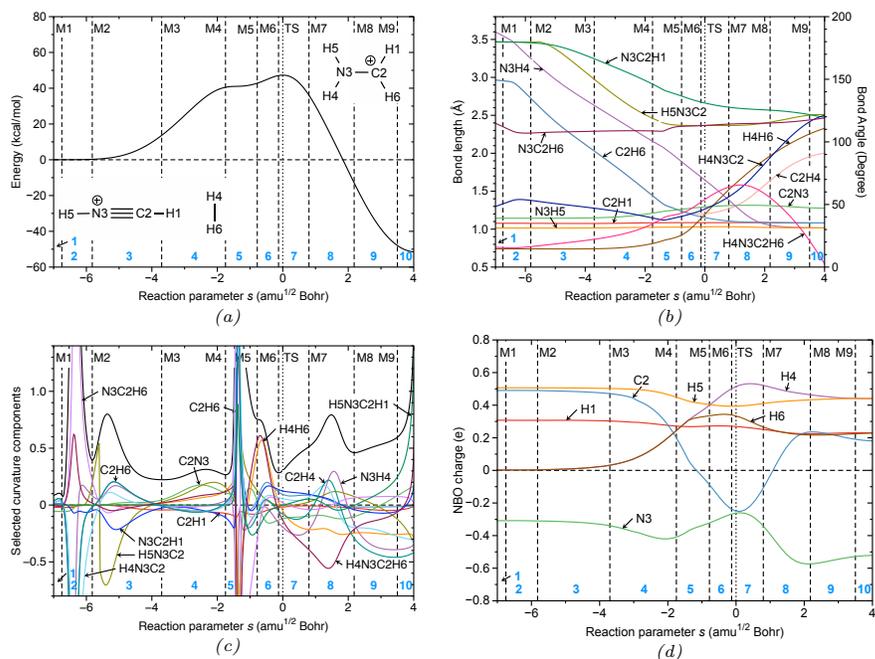
Fig. 9: Reaction **R8**; a) energy profile; b) changes in geometry parameters, c) reaction path curvature profile (black solid line) and its decomposition into components; d) changes in NBO charges as a function of the reaction path parameter s. Reaction phases are indicated by vertical dashed lines at curvature minima M1, M2, ..., Mn, and are labeled by blue numbers. The TS is indicated as a dotted line. MP2/aug-cc-pVTZ level of theory.

hydrogen molecule approaches **7** at the carbon end and causes the HCN and CNN bonds to bend into the opposite direction. In the first part of the reaction H_2 is attached to the carbon atom resembling a reversion of reaction **R7**. In the second part, one of the H_2 hydrogen atoms migrates to the nitrogen forming **4S**. Obviously, this mechanism, which been found e.g., in transition metal catalysis [96] is more energy effective than starting via breaking the hydrogen bond in H_2, which costs more than 100 kcal/mol. As denoted by the small shoulder in energy profile at about s=-2.0 $amu^{1/2}Bohr$ (see Figure 9a) reaction **R8** is at the onset for the two-step reaction mechanism. Figure 9b shows the geometric changes of reaction **R8** as a function of s. The H4H6 distance of 0.740 Å between the hydrogen atoms of the incoming H_2 molecule does not change for more than 3 s units, confirming that the H_2 molecule is captured by the carbon atom. At s=-2.5 $amu^{1/2}Bohr$ the H4H6 distance slowly starts to increase, after s=-0.8 $amu^{1/2}Bohr$ this distance is more pronounced as the split of H_2 and the migration of H4 to the nitrogen atom proceeds. At s=4.0 $amu^{1/2}Bohr$ the final H4H6 distance of 2.331 Å is reached. The C2H4/C2H6 distance linearly decrease from 2.977 Å. At s=-0.15 $amu^{1/2}Bohr$ the C2H6

distance further decreases to the value of 1.164 Å, whereas the C2H4 distance starts to increase from the value 1.201 Å in accordance with the starting H-migration. The N3H4 distance steadily decreases, until the final value of 1.016 Å is reached at s=3.8 $amu^{1/2}Bohr$. Another interesting parameter is the H4N3C2H6 dihedral angle. Before reaching its final value of 0° in **4S**, the H4N3C2H6 dihedral angle adapts a maximum of 66° at s=1.1 $amu^{1/2}Bohr$ denoting that the migration of H4 occurs out the molecular plane. Further details about this complex reaction mechanism, including a quantification of the geometry discussion, can be gained by inspection of the Figure 9c showing the reaction path curvature profile and its decomposition. The reaction proceeds in 10 distinct phases, with **TS8** included in phase 7, i.e., several important chemical events occur before the TS. The large curvature peak in phase 2 is characterized by the adjustment of the incoming H_2 as reflected by supporting N3C2H6/N3C2H4, H4N3C2H6, and H4N3C2 resisting contributions. Phase 3 is dominated by the bending of **7**, resisting H5N3C2 and N3C2H1 components and a supporting C2H6/C2H4 component. The smaller curvature enhancement in phase 4 predominately results from changes in the CN bond and a start of H4H6 elongation. In phase 5 more drastic changes take place as reflected by the large curvature peak. The cleavage of the H4H6 bond and the migration of H4 start with a strongly resisting H4H6 component and supportive C2H6 and C2H4 components. Phase 6 is dominated by HH bond cleavage (supportive H4H6 component) and hydrogen migration (supportive H4N3C2H6 component). After the TS, in phase 8 the final N3H4 distance is reached, denoted by large N3H4 and C2H4 components as well as the final C2H6 distance. There is still resistance to bring the molecule in its final planar form, i.e. large negative H4N3C2H6, which is then completed in phase 10 together with final bond angle adjustments. Figure 9d shows the NBO charges as a function of s. It is interesting to note changes of the NBO charges start only at the beginning of phase 4, where both hydrogens of the H_2 pick up positive charge predominately from C2 and to a lesser extend from N3. At s=-1.4 $amu^{1/2}Bohr$ H4 starts to adapt more positive charge than H6 in accordance with its migration to the nitrogen atom. At the end of the reaction the H4 and H5 hydrogens on N3 adapt a positive charge of 0.44 e and the H1 and H6 hydrogens on carbon a positive charge of 0.23 e. Interesting to note is that shortly after the TS, both carbon and nitrogen are equally charged with -0.26 e. After that point negative charge is transferred back from C2 to N3 resulting in a polar CN double bond in **4S** with 0.18 e on C2 and -0.52 on N3.

4 CN chemical bond analysis

The URVA analysis of chemical reactions **R1** – **R8** in this study was complemented by the assessment of the strength all CN bond formed during **R1** - **R8**. For this purpose we calculated local mode force constants k^a with the CASPT2/aug-cc-pVTZ and CCSD(T)/cc-pVDZ model chemistries utilizing the LmodeA program [77].
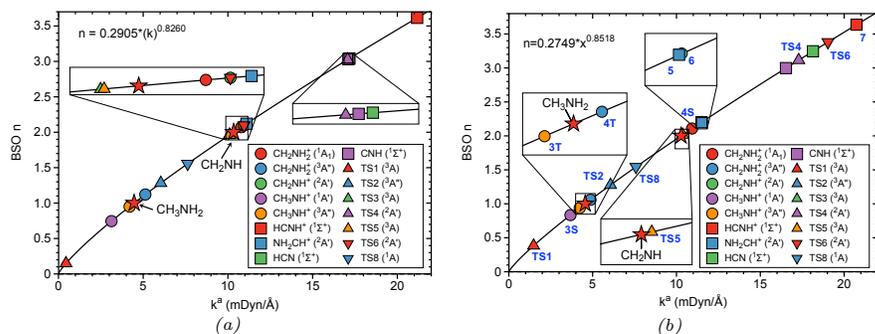
Fig. 10: Bond strength order BSO n as a function of the local mode force constant k^a for the CN bonds of all stationary points reactions **R1** – **R8** and some reference compounds (see legend). a) CASPT2/aug-cc-pVTZ level of theory; b) CCSD(T)/cc-pVDZ level of theory.

In order to facilitate the discussion local mode force constants k^a can be transformed into bond strength orders (BSO n) via a power relationship derived from the generalized Badger rule [102,82]: $BSO\ n = A * (k^a)^B$; where the constants A and B are determined by two reference molecules with known BSO n and k^a values and the requirement that for a zero force constant k^a the corresponding BSO n value is also zero. We used in this work as a reference CH_3NH_2 with a CN bond order of 1 and CH_2NH with a CN bond order of 2. This led to constants $A = 0.291$ and $B = 0.826$ for CASPT2/aug-cc-pVTZ with k^a(C–N) of $CH_3NH_2 = 4.467$ mDyn/Å and k^a(C=N) of $CH_2NH = 10.339$ mDyn/Å and $A = 0.275$ and $B = 0.852$ for CCSD(T)/cc-pVDZ with k^a(C–N) of $CH_3NH_2 = 4.555$ mDyn/Å and k^a(C=N) of $CH_2NH = 10.278$ mDyn/Å.

The (BSO n, k^a) power relationship for all stationary points of reactions **R1** – **R8** and some reference compounds is shown in Figure 10a for CASPT2/aug-cc-pVTZ and in Figure 10b for CCSD(T)/cc-pVDZ. Bond distances R(CN), local mode force constants k^a(CN) and corresponding bond strength orders BSO n(CN) of all CN bonds in molecules **3**–**7**, TSs, and some reference compounds, calculated at the CASPT2/aug-cc-pVTZ and CCSD(T)/cc-pVTZ levels of theory are collected in Table 2. Noteworthy is the broad CN bond order spectrum observed in our study, ranging from almost zero for the van der Waals complex and **TS1** to 3.7 found for the protonated hydrogen cyanide **7** in accordance with previous work comparing the strength of the CN bond in **7** with that of the triple bond of double-protonated carbon monoxide [103]. The following we will limit the discussion to the CCSD(T)/cc-pVDZ results.

According to Figure 1 we expect a very weak CN bond in both the van der Waals complex formed from methane (**1**) and $N^+(^3P)$ as well as in the **TS1** which is confirmed by the BSO n values of 0.059 and 0.384, respectively (see Table 2 and Figure 10b. Next follows a group of molecules with CN bonds in the single bond range , **3S** and **3T** and **4T** with BSO n values of 0.834, 0.939, and 1.059 respectively, followed by **TS2** (BSO n = 1.277) which has a

Table 2: Bond distances R(CN) in Å, local mode force constants k^a(CN) in mDyn/Å, and bond strength orders BSO n(CN) for molecules **3**–**7**, **TS1** - **TS6**, **TS8**, and some reference compounds, calculated at the CASPT2/aug-cc-pVTZ and CCSD(T)/cc-pVDZ levels of theory.

Molecule	CASPT2/aug-cc-pVTZ			CCSD(T)/cc-pVDZ		
	R(CN)	k^a(CN)	BSO n(CN)	R(CN)	k^a(CN)	BSO n(CN)
vdW CH_3NH^+ ($^3A''$)	1.825	0.257	0.095	2.027	0.163	0.059
3T CH_3NH^+ ($^3A''$)	1.388	4.200	0.950	1.409	4.231	0.939
3S CH_3NH^+ ($^1A_1$)	1.388	3.129	0.745	1.409	3.678	0.834
4T $CH_2NH_2^+$ ($^3A''$)	1.392	5.127	1.121	1.413	4.872	1.059
4S $CH_2NH_2^+$ ($^1A_1$)	1.280	10.789	2.072	1.290	10.930	2.108
5 $CHNH_2^+$ ($^2A'$)	1.239	10.954	2.098	1.272	11.496	2.201
6 CH_2NH^+ ($^2A'$)	1.261	11.096	2.120	1.252	11.520	2.205
7 $CHNH^+$ ($^1\Sigma^+$)	1.139	21.171	3.615	1.155	20.736	3.637
TS1 (^3A)	1.838	0.445	0.149	1.610	1.480	0.384
TS2 ($^3A''$)	1.379	6.041	1.283	1.394	6.068	1.277
TS4 ($^2A'$)	1.165	17.068	3.026	1.180	17.289	3.115
TS5 (^3A)	1.251	10.107	1.963	1.266	10.370	2.016
TS6 ($^2A'$)	1.239	10.954	2.098	1.167	19.032	3.381
TS8 (^1A)	1.313	7.623	1.555	1.332	7.585	1.544
CH_3NH_2 ($^1A_1$)	1.465	4.467	1.000	1.473	4.555	1.000
CH_2NH ($^1A_1$)	1.275	10.339	2.000	1.289	10.278	2.000
CH_2NH_2 ($^2A'$)	1.390	6.132	1.299	1.413	5.852	1.238
$CHNH$ ($^2A'$)	1.239	12.144	2.284	1.255	12.035	2.288
HCN ($^1\Sigma^+$)	1.165	17.175	3.042	1.175	18.140	3.246
CNH ($^1\Sigma^+$)	1.170	17.119	3.033	1.192	16.529	2.998

slightly shorter CN bond than found in **3T** and **4T** (see Table 2. **TS8** has with BSO n = 1.544 a somewhat weaker CN bond than expected from inspection of Figure 1, which is the result of a widening of the CN bond caused by the migrating hydrogen atom, as underlined in the reaction movie **SR7** (see Supporting Information). The third group of CN bonds are in the CN double bond range, **TS5** with BSO n = 2.016, followed by **4S** where in contrast to **4T** the CN double can be formed, as reflected by a BSO n value of 2.108. **5** and **6** with BSO n values of 2.201 and 2.205 respectively also belong to this group. The final group contains molecules in the CN triple bond range. The first member of this group is **TS4** with a BSO value of 3.115 which is somewhat smaller than the BSO n value of the CN bond in CNH, followed by **TS6** which has a somewhat stronger CN bond than the CN bond in the HCN molecule (BSO n values of 3.381 and 3.246, respectively). **7** has with BSO n = 3.637 the strongest CN bond found in this study. As discussed in a previous study [103] protonated hydrogen cyanide **7** belongs to the strongest bonds found in chemistry. In order to assess the effect of protonation on the bond strength we calculated the CN bond strength in CHNH radical and CH_2NH_2 radical, the counterpart of **4**. In both cases, we observed a drastic decrease in the CN bond strength (BSO n values of 2.288 compared to 3.637 in **7** and 1.238 compared to 2.108 in **4**, respectively, see Table 2) confirming the influence of protonation which strongly increasing the electronegativity of the heavy atom involved.
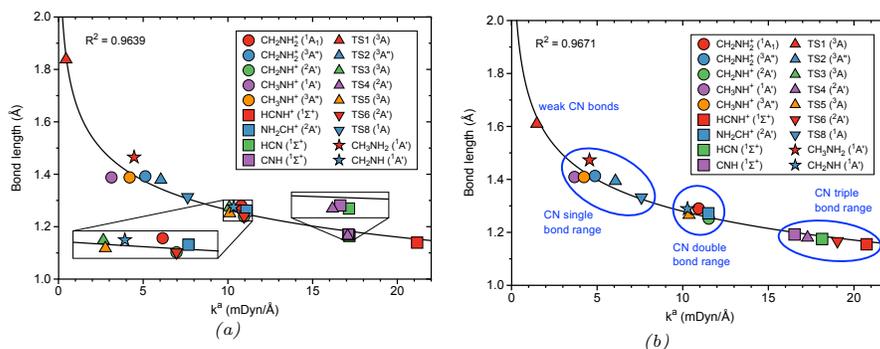
Fig. 11: Correlation between the CN bond length and local mode force constant $k^a(CN)$ for all stationary points of reactions **R1** – **R8** and some reference compounds (see legend). a) CASPT2/aug-cc-pVTZ level of theory; b) CCSD(T)/cc-pVDZ level of theory.

Popular strategies to assess the strength of a chemical bond or weak chemical interaction often refer to bond distances [104–106]. However, a caveat is appropriate. A number of cases have been reported illustrating that a shorter bond is not always a stronger bond [107–109]. In the following we discuss the situation for the CN bonds of this study using local mode force constants as a suitable assessment tool. In Figure 11a $R(CN)$ and $k^a(CN)$ values are correlated for CASPT2/aug-cc-pVTZ level of theory and in Figure 11b for CCSD(T)/cc-pVDZ level of theory. For both model chemistries similar trends are obtained and a good correlation between $R(CN)$ and $k^a(CN)$ with $R^2=0.964$ and $R^2=0.967$, respectively. We also see the separation into clusters as denoted in Figure 11b, very weak and long CN bonds, one cluster in the CN single bond, one in the CN double bond, and one in the CN triple bond range, confirming that for the CN bonds investigated in this study the bond length-bond strength relationship is valid.

5 Conclusions and Outlook

In this work we investigated the formation of protonated hydrogen cyanide $HCNH^+$ (**7**) and methylene amine cation $CH_2NH_2^+$ (**4S**) from three different pathways which stem from the interaction between CH_4 (**1**) and $N^+(^3P)$ (**2**). Our URVA studies could provide a comprehensive overview on bond formation/cleavage processes relevant to the specific mechanism of eight reactions **R1** - **R8** that occur across the three pathways. In addition we could explain the formation of $CH_2NH_2^+$ and the appearance of $HCNH^+$ and $CHNH_2^+$ along these paths. Although only smaller molecules are involved these in reactions including isomerzation, hydrogen atom abstraction and hydrogen molecule capture, we found a number of interesting features, such as *roaming* in reaction **R3** or the primary interaction of H_2 with the carbon atom in **7** in reaction

R8 followed by migrating of one of the H_2 hydrogen atoms to the nitrogen, which is more cost effective than breaking the HH bond first; a feature often found in catalysis. In all cases, charge transfer between carbon and nitrogen could be identified as a driving force for the CN bond formation. Among the three possible pathways, the path via the singlet potential energy surface is the shortest and therefore, most favorable path for the formation of **7** and **4**. Several new specific insights emerged from our study.

1. For reactions **R1** and **R2** no pronounced peaks were observed at the corresponding TS and significant chemical changes were seen to occur in both the entrance and exit channels revealing that major chemical events do not always take place at the TS.

2. From reaction **R4** we observed the hydrogen atom H5 to orbit around the NH molecule rather than exiting on a straight pathway which resulted in a long exit channel due to this so-called *roaming* mechanism. Alongside roaming, **R4** involves the formation of a dihydrogen interaction during the reaction mechanism which is more commonly observed in transition metal catalysis. We note that the during the reaction **R6** there is similar assistance of such a dihydrogen interaction, where **R6** has one less phase and a much shorter exit channel in contrast to **4** due to a lesser extent of *roaming*.

3. *Path 3* was shown to be an appealing alternative for the formation of **R7** through one barrier-less reaction which involves intersystem crossing of the triplet state CH_3NH^+ into the singlet state CH_3NH^+ and a loss of H_2. Furthermore, **R7** was shown to consist of seven reaction phases and four curvature peaks demonstrating that reactions with a barrier also possess elaborate mechanistic features, where notably, at the end of phase 4 the H_2 formation is finished and the final CH bond cleavage takes place in phase 6 and the change from a $C-N$ bond to a $C{\equiv}C$ bond happens during phases 2 and 3 and is completed by phase 4.

4. Reaction **R8** is a complicated reaction mechanism that takes place across 10 distinct phases where the TS is within phase 7 with various important chemical events taking place prior to the TS, this reaction also features the formation of a dihydrogen interaction. An interesting feature of this reaction pertains to the start of phase 4 where the H_2 atoms (H4 and H6) collect positive charge mainly from C2 where H4 become more positive than H6 as it migrates towards the N atom.

5. As revealed by LMA the CN bonds formed in reactions **R1** - **R8** cover a broad bond strength range from very weak (BSO n values < 0.06) to very strong (BSO n values > 3.5) with the CN bond in protonated hydrogen cyanide $HCNH^+$ (**7**) identified as strongest of all molecules investigated in this work.

In summary, our study shows the large potential of both, URVA and LMA to shed new light into these extraterrestrial reactions to help better understand prebiotic processes as well as develop guidelines for future investigations involving areas of complex interstellar chemistry. In particular the formation of

CN bonds as a precursor to the extraterrestrial formation of amino acids [110, 111] will be the focus of future investigations.

6 Data availability

All data generated or analyzed during this study are included in this published article. Additional information is provided in the Supporting Information which contains i) the decomposition of the reaction path direction along the reaction parameter s for each reaction; ii) reaction movies following the changes of the reaction complex along the reaction parameter s for each reaction; iii) Cartesian coordinates of molecules **1**, **3** - **7**, transition states **TS1** - **TS6** and **TS8**, and some reference molecules.

7 Code availability

The local mode analysis package LModeA can be obtained by the authors upon request.

References

1. Öberg, K.I., Bergin, E.A.: Astrochemistry and compositions of planetary systems. Phys. Rep. (2020)
2. Okada, K., Sakimoto, K., Takada, Y., Schuessler, H.A.: A study of the translational temperature dependence of the reaction rate constant between ch$_3$cn and ne$^+$ at low temperatures. J. Chem. Phys. **153**(12), 124305 (2020)
3. de Barros, A., Bergantini, A., Domaracka, A., Rothard, H., Boduch, P., da Silveira, E.: Radiolysis of nh$_3$:co ice mixtures - implications for solar system and interstellar ices. MNRAS **499**(2), 2162–2172 (2020)
4. Dubois, D., Carrasco, N., Jovanovic, L., Vettier, L., Gautier, T., Westlake, J.: Positive ion chemistry in an n2-ch4 plasma discharge: Key precursors to the growth of titan tholins. Icarus **338**, 113437 (2020)
5. Tielens, A.: The molecular universe. RMP **85**(3), 1021 (2013)
6. Matson, D.L., Spilker, L.J., Lebreton, J.P.: The cassini/huygens mission to the saturnian system. The Cassini-Huygens Mission pp. 1–58 (2003)
7. NASA: Cassini/Huygens mission. See the website for a comprehensive overview
8. Werner, M.W., Roellig, T.L., Low, F.J., Rieke, G.H., Rieke, M., Hoffmann, W.F., Young, E., Houck, J.R., Brandl, B., Fazio, G.G., Hora, J.L., Gehrz, R.D., Helou, G., Soifer, B.T., Stauffer, J., Keene, J., Eisenhardt, P., Gallagher, D., Gautier, T.N., Irace, W., Lawrence, C.R., Simmons, L., Van Cleve, J.E., Jura, M., Wright, E.L., Cruikshank, D.P.: The spitzer space telescope mission. APIS **154**(1), 1–9 (2004)
9. France, K., Schindhelm, E., Herczeg, G.J., Brown, A., Abgrall, H., Alexander, R.D., Bergin, E.A., Brown, J.M., Linsky, J.L., Roueff, E., et al.: A hubble space telescope survey of h2 emission in the circumstellar environments of young stars. ApJl **756**(2), 171 (2012)
10. Gardner, J.P., Mather, J.C., Clampin, M., Doyon, R., Greenhouse, M.A., Hammel, H.B., Hutchings, J.B., Jakobsen, P., Lilly, S.J., Long, K.S., et al.: The james webb space telescope. Space Sci. Rev. **123**(4), 485–606 (2006)
11. Fanson, J., Bernstein, R., Angeli, G., Ashby, D., Bigelow, B., Brossus, G., Bouchez, A., Burgett, W., Contos, A., Demers, R., et al.: Overview and status of the giant magellan telescope project. In: Ground-based and Airborne Telescopes VIII, vol. 11445, p. 114451F. International Society for Optics and Photonics (2020)

12. Heazlewood, B.R., Softley, T.P.: Towards chemistry at absolute zero. Nat. Rev. Chem. pp. 1–16 (2021)
13. Cooke, I.R., Sims, I.R.: Experimental studies of gas-phase reactivity in relation to complex organic molecules in star-forming regions. ACS Earth Space Chem. pp. 1109–1134 (2019)
14. Smith, I.W.: Laboratory astrochemistry: Gas-phase processes. ARAA **49**(1), 29–66 (2011)
15. Bourgalais, J., Carrasco, N., Vettier, L., Gautier, T., Blanchet, V., Petit, S., Descamps, D., Fedorov, N., Delos, R., Gaudin, J.: On an euv atmospheric simulation chamber to study the photochemical processes of titan's atmosphere. Sci **10**(1), 1–14 (2020)
16. Bourgalais, J., Carrasco, N., Vettier, L., Comby, A., Descamps, D., Petit, S., Blanchet, V., Gaudin, J., marty Bernard: Aromatic formation promoted by ion-driven radical pathways in euv photochemical experiments simulating titan's atmospheric chemistry. J. Phys. Chem. A **125**, 3159–3168 (2021)
17. Sandford, S.A., Nuevo, M., Bera, P.P., Lee, T.J.: Prebiotic astrochemistry and the formation of molecules of astrobiological interest in interstellar clouds and protostellar disks. Chem. Rev. **120**(11), 4616–4659 (2020)
18. Haupa, K.A., Ong, W.S., Lee, Y.P.: Hydrogen abstraction in astrochemistry: Formation of ch_2conh_2 in the reaction of h atom with acetamide (ch_3conh_2) and photolysis of ch_2conh_2 to form ketene (ch_2co) in solid para-hydrogen. Phys. Chem. Chem. Phys. **22**(11), 6192–6201 (2020)
19. Noelle, A., Vandaele, A.C., Martin-Torres, J., Yuan, C., Rajasekhar, B.N., Fahr, A., Hartmann, G.K., Lary, D., Lee, Y.P., Limão-Vieira, P., et al.: Uv/vis+ photochemistry database: Structure, content and applications. J Quant Spectrosc Radiat Transf **253**, 107056 (2020)
20. Lavvas, P., Lellouch, E., Strobel, D., Gurwell, M., Cheng, A., Young, L., Gladstone, G.: A major ice component in pluto's haze. Nat. Astron. **5**(3), 289–297 (2021)
21. DIschia, M., Manini, P., Moracci, M., Saladino, R., Ball, V., Thissen, H., Evans, R.A., Puzzarini, C., Barone, V.: Astrochemistry and astrobiology: materials science in wonderland? Int. J. Mol. Sci. **20**(17), 4079 (2019)
22. Thripati, S., Ramabhadran, R.O.: Pathways for the formation of formamide, a prebiotic biomonomer: Metal—ons in interstellar gas-phase chemistry. J. Phys. Chem. A pp. 3457–3472 (2021)
23. Vastel, C., Loison, J.C., Wakelam, V., Lefloch, B.: Isocyanogen formation in the cold interstellar medium. Astronomy and Astrophysics p. A91 (2019)
24. Gavilan, L., Broch, L., Carrasco, N., Fleury, B., Vettier, L.: Organic aerosols in the presence of $co_2$2 in the early earth and exoplanets: Uv–vis refractive indices of oxidized tholins. Astrophys. J. Letters p. L5 (2017)
25. Waite, J., Young, D., Cravens, T., Coates, A., Crary, F., Magee, B., Westlake, J.: The process of tholin formation in titan's upper atmosphere. Science **316**(5826), 870–875 (2007)
26. Cable, M.L., Hörst, S.M., Hodyss, R., Beauchamp, P.M., Smith, M.A., Willis, P.A.: Titan tholins: Simulating titan organic chemistry in the cassini-huygens era. Chem. Rev. **112**(3), 1882–1909 (2012)
27. Neish, C.D., Somogyi, A., Smith, M.A.: Titan's primordial soup: Formation of amino acids via low-temperature hydrolysis of tholins. Astrobiology **10**(3), 337–347 (2010)
28. Dalton, J., Cruikshank, D., Stephan, K., McCord, T., Coustenis, A., Carlson, R., Coradini, A.: Chemical composition of icy satellite surfaces. Space Sci. Rev. **153**(1-4), 113–154 (2010)
29. Cravens, T.E., Robertson, I.P., Waite Jr., J.H., Yelle, R.V., Kasprzak, W.T., Keller, C.N., Ledvina, S.A., Niemann, H.B., Luhmann, J.G., McNutt, R.L., Ip, W.H., De La Haye, V., Mueller-Wodarg, I., Wahlund, J.E., Anicich, V.G., Vuitton, V.: Composition of titan's ionosphere. Geophys. Res. Lett **33**(7), E01003 (2006)
30. Waite, J., Lewis, W., Kasprzak, W., Anicich, V., Block, B., Cravens, T.E., Fletcher, G., Ip, W.H., Luhmann, J., McNutt, R., et al.: The cassini ion and neutral mass spectrometer (inms) investigation. Space Sci. Rev. **114**(1-4), 113–231 (2004)
31. Ennis, C., Cable, M.L., Hodyss, R., Maynard-Casely, H.E.: Mixed hydrocarbon and cyanide ice compositions for titan?s atmospheric aerosols: A ternary-phase co-crystal

predicted by density functional theory. ACS Earth and Space Chem **4**(7), 1195–1200 (2020)

32. Kronrod, V., Dunaeva, A., Gudkova, T., Kuskov, O.: Matching of models of the internal structure and thermal regime of partially differentiated titan with gravity field. Sol. Syst. Res. **54**(5), 405–419 (2020)
33. Markus, C.R., Thorwirth, S., Asvany, O., Schlemmer, S.: High-resolution double resonance action spectroscopy in ion traps: Vibrational and rotational fingerprints of $ch_2nh_2^+$. Phys. Chem. Chem. Phys. **21**(48), 26406–26412 (2019)
34. Yuen, C.H., Ayouz, M.A., Balucani, N., Ceccarelli, C., Schneider, F., Kokoouline, V.: Dissociative recombination of $ch_2nh_2^+$: a crucial link with interstellar methanimine and titan ammonia. MNRAS pp. 659–664 (2019)
35. Markus, C.R., Thorwirth, S., Asvany, O., Schlemmer, S.: High-resolution double resonance action spectroscopy in ion traps: vibrational and rotational fingerprints of $ch_2nh_2^+$. Phys. Chem. Chem. Phys. pp. 26406–26412 (2019)
36. Richardson, V., Alcaraz, C., Geppert, W., Polasek, M., Romanzin, C., Sundelin, D., Thissen, R., Tosia, P., Zabkae, J., Ascenzia, D.: The reactivity of methanimine radical cation (h_2cnh^+) and its isomer aminomethylene ($hcnh_2^+$) with methane. Chem. Phys. Lett. p. 138611 (2021)
37. Pei, L., Farrar, J.M.: Ion imaging study of reaction dynamics in the $n^+ + ch_4$ system. Int. J. Chem. Phys. **137**(15), 154312 (2012)
38. Kraka, E., Zou, W., Tao, Y., Freindorf, M.: Exploring the mechanism of catalysis with the unified reaction valley approach (urva) - a review. Catalysts **10**, 691 (2020)
39. Kraka, E., Zou, W., Tao, Y.: Decoding chemical information from vibrational spectroscopy data: Local vibrational mode theory. WIREs: Comput. Mol. Sci. **10**, 1480 (2020)
40. Zou, W., Sexton, T., Kraka, E., Freindorf, M., Cremer, D.: A New Method for Describing the Mechanism of a Chemical Reaction Based on the Unified Reaction Valley Approach. J. Chem. Theory Comput. **12**, 650–663 (2016)
41. Kraka, E.: Reaction Path Hamiltonian and the Unified Reaction Valley Approach. WIREs: Comput. Mol. Sci. **1**, 531–556 (2011)
42. Wilson, E.B., Decius, J.C., Cross, P.C.: Molecular Vibrations. McGraw-Hill, New York (1955)
43. Califano, S.: Vibrational States. Wiley, London (1976)
44. Herzberg, G.: Molecular Spectra and Molecular Structure, II. Infrared and Raman Spectra of Polyatomic Molecules. Van Nostrand, New York (1945)
45. Konkoli, Z., Cremer, D.: A New Way of Analyzing Vibrational Spectra. I. Derivation of Adiabatic Internal Modes. Int. J. Quant. Chem. **67**, 1–9 (1998)
46. Konkoli, Z., Larsson, J.A., Cremer, D.: A New Way of Analyzing Vibrational Spectra. II. Comparison of Internal Mode Frequencies. Int. J. Quant. Chem. **67**, 11–27 (1998)
47. Konkoli, Z., Cremer, D.: A New Way of Analyzing Vibrational Spectra. III. Characterization of Normal Vibrational Modes in terms of Internal Vibrational Modes. Int. J. Quant. Chem. **67**, 29–40 (1998)
48. Konkoli, Z., Larsson, J.A., Cremer, D.: A New Way of Analyzing Vibrational Spectra. IV. Application and Testing of Adiabatic Modes within the Concept of the Characterization of Normal Modes. Int. J. Quant. Chem. **67**, 41–55 (1998)
49. Cremer, D., Larsson, J.A., Kraka, E.: New Developments in the Analysis of Vibrational Spectra on the Use of Adiabatic Internal Vibrational Modes. In: C. Parkanyi (ed.) Theoretical and Computational Chemistry, pp. 259–327. Elsevier, Amsterdam (1998)
50. Zou, W., Cremer, D.: C_2 in a Box: Determining its Intrinsic Bond Strength for the $X^1\Sigma^+_g$ Ground State. Chem. Eur. J. **22**, 4087–4097 (2016)
51. Freindorf, M., Kraka, E.: Critical assessment of the fec and co bond strength in carboxymyoglobin - a qm/mm local vibrational mode study. J. Mol. Model. **26**, 281–1–281–15 (2020)
52. Yannacone, S., Freindorf, M., Tao, Y., Zou, W., Kraka, E.: Local vibrational mode analysis of π-hole interactions between aryl donors and small molecule acceptors. Crystals **10**, 556–1–556–25 (2020)
53. Tao, Y., Zhang, L., Zou, W., Kraka, E.: Equilibrium geometries, adiabatic excitation energies and intrinsic c=c/c-h bond strengths of ethylene in lowest singlet excited states described by tddft. Symmetry **12**, 1545–1–1545–13 (2020)

54. Pekar, K.B., Lefton, J.B., McConville, C., Burleson, J., Sethio, D., Kraka, E., Runčevski, T.: Peritectic phase transition of benzene and acetonitrile into a cocrystal relevant to titan, saturn's moon. Chem. Comm. (2020). DOI 10.1039/d0cc04999a

55. Tao, Y., Zou, W., Sethio, D., Verma, N., Qiu, Y., Tian, C., Cremer, D., Kraka, E.: In situ measure of intrinsic bond strength in crystalline structures: Local vibrational mode theory for periodic systems. J. Chem. Theory Comput. **15**, 1761–1776 (2019)

56. Delgado, A.A.A., Humason, A., Kalescky, R., Freindorf, M., Kraka, E.: Exceptionally Long Covalent CC Bonds - A Local Vibrational Mode Study. Molecules **26**, 950–1–950–25 (2021)

57. Beiranvand, N., Freindorf, M., Kraka, E.: Hydrogen Bonding in Natural and Unnatural Base Pairs - Explored with Vibrational Spectroscopy. Molecules **26**, 2268–12268–22 (2021)

58. Head-Gordon, M., Pople, J.A., Frisch, M.J.: Mp2 energy evaluation by direct methods. Chem. Phys. Lett. **153**(6), 503–506 (1988)

59. Cremer, D.: Møller-Plesset Perturbation Theory: From Small Molecule Methods to Methods for Thousand of Atoms. WIREs: Comput. Mol. Sci. **1**, 509–530 (2011)

60. Celani, P., Werner, H.J.: Multireference perturbation theory for large restricted and selected active space reference wave functions. J. Chem. Phys. **112**(13), 5546–5557 (2000)

61. Celani, P., Werner, H.J.: Analytical energy gradients for internally contracted second–order multireference perturbation theory. J. Chem. Phys. **119**(10), 5044–5057 (2003)

62. Győrffy, W., Shiozaki, T., Knizia, G., Werner, H.J.: Analytical energy gradients for second–order multireference perturbation theory using density fitting. J. Chem. Phys. **138**(10), 104104 (2013)

63. Dunning, T.H.: Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys. **90**(2), 1007–1023 (1989)

64. Bartlett, R.J., Musiał, M.: Coupled-cluster theory in quantum chemistry. Rev. Mod. Phys. **79**(1), 291–352 (2007)

65. Christiansen, O.: Coupled cluster theory with emphasis on selected new developments. Theor Chem. Acc. **116**, 106–123 (2006)

66. Matthews, D.A., Cheng, L., Harding, M.E., Lipparini, F., Stopkowicz, S., Jagau, T.C., Szalay, P.G., Gauss, J., Stanton, J.F.: Coupled-cluster techniques for computational chemistry: The CFOUR program package. J. Chem. Phys. **152**(21), 214108–1–214108–36 (2020)

67. Raghavachari, K., Trucks, G.W., Pople, J.A., Head-Gordon, M.: A fifth-order perturbation comparison of electron correlation theories. Chem. Phys, Lett. **157**, 479–483 (1989)

68. Bartlett, R.J., Watts, J., Kucharski, S., Noga, J.: Non-iterative fifth-order triple and quadruple excitation energy corrections in correlated methods. Chem. Phys, Lett. **165**, 513–522 (1990)

69. Stanton, J.: Why ccsd(t) works: a different perspective. Chem. Phys, Lett. **281**, 130–134 (1997)

70. Fukui, K.: The Path of Chemical Reactions - The IRC Approach. Acc. Chem. Res. **14**(12), 363–368 (1981)

71. Hratchian, H.P., Kraka, E.: Improved Predictor-Corrector Integrators For Evaluating Reaction Path Curvature. J. Chem. Theory Comput. **9**, 1481–1488 (2013)

72. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A.V., Bloino, J., Janesko, B.G., Gomperts, R., Mennucci, B., Hratchian, H.P., Ortiz, J.V., Izmaylov, A.F., Sonnenberg, J.L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V.G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery Jr, J.A., Peralta, J.E., Ogliaro, F., Bearpark, M.J., Heyd, J.J., Brothers, E.N., Kudin, K.N., Staroverov, V.N., Keith, T.A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A.P., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Millam, J.M., Klene, M., Adamo, C., Cammi, R., Ochterski, J.W., Martin, R.L., Morokuma, K., Farkas, O., Foresman, J.B., Fox, D.J.: Gaussian 16. Gaussian Inc. Wallingford CT (2016)

73. Werner, H.J., Knowles, P.J., Knizia, G., Manby, F.R., Schütz, M., Celani, P., Györffy, W., Kats, D., Korona, T., Lindh, R., Mitrushenkov, A., Rauhut, G., Shamasundar, K.R., Adler, T.B., Amos, R.D., Bennie, S.J., Bernhardsson, A., Berning, A., Cooper, D.L., Deegan, M.J.O., Dobbyn, A.J., Eckert, F., Goll, E., Hampel, C., Hesselmann, A., Hetzer, G., Hrenar, T., Jansen, G., Köppl, C., Lee, S.J.R., Liu, Y., Lloyd, A.W., Ma, Q., Mata, R.A., May, A.J., McNicholas, S.J., Meyer, W., Miller III, T.F., Mura, M.E., Nicklass, A., O'Neill, D.P., Palmieri, P., Peng, D., Pflüger, K., Pitzer, R., Reiher, M., Shiozaki, T., Stoll, H., Stone, A.J., Tarroni, R., Thorsteinsson, T., Wang, M., Welborn, M.: Molpro, version 2021.1, a package of ab initio programs. Https://www.molpro.net

74. Werner, H.J., Knowles, P.J., Manby, F.R., Black, J.A., Doll, K., Heßelmann, A., Kats, D., Köhn, A., Korona, T., Kreplin, D.A., Ma, Q., Miller, T.F., Mitrushchenkov, A., Peterson, K.A., Polyak, I., Rauhut, G., Sibaev, M.: The molpro quantum chemistry package. J. Chem. Phys. **152**(14), 144107 (2020)

75. Werner, H.J., Knowles, P.J., Knizia, G., Manby, F.R., Schütz, M.: Molpro: a general–purpose quantum chemistry program package. WIREs Comput. Mol. Sci. **2**(2), 242–253 (2012)

76. Tao, Y., Zou, W., Freindorf, M., Cremer, D., Kraka, E.: pURVA. Southern Methodist University: Dallas, TX, USA (2020)

77. Zou, W., Tao, Y., Freindorf, M., Makoś, M.Z., Verma, N., Kraka, E.: Local Vibrational Mode Analysis: LModeA. Computational and Theoretical Chemistry Group (CATCO), Southern Methodist University: Dallas, TX, USA (2020)

78. Reed, A., Curtiss, L., Weinhold, F.: Intermolecular Interactions from A Natural Bond Orbital, Donor-Acceptor Viewpoint. Chem. Rev. **88**, 899–926 (1988)

79. Weinhold, F., Landis, C.R.: Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective. Cambridge University Press (2003)

80. Glendening, E.D., Badenhoop, J.K., Reed, A.E., Carpenter, J.E., Bohmann, J.A., Morales, C.M., Landis, C.R., Weinhold, F.: NBO6 (2013). Theoretical Chemistry Institute, University of Wisconsin, Madison

81. Park, J.W., Al-Saadon, R., MacLeod, M.K., Shiozaki, T., Vlaisavljevich, B.: Multireference electron correlation methods: Journeys along potential energy surfaces. Chem. Rev. **120**, 5878–5909 (2020)

82. Cremer, D., Kraka, E.: From Molecular Vibrations to Bonding, Chemical Reactions, and Reaction Mechanism. Curr. Org. Chem. **14**, 1524–1560 (2010)

83. Kalescky, R., Kraka, E., Cremer, D.: Are Carbon-Halogen Double and Triple Bonds Possible? Int. J. Quant. Chem. **114**, 1060–1072 (2014)

84. Kalescky, R., Zou, W., Kraka, E., Cremer, D.: Quantitative Assessment of the Multiplicity of Carbon-Halogen Bonds: Carbenium and Halonium Ions with F, Cl, Br, and I. J. Phys. Chem. A **118**, 1948–1963 (2014)

85. Oliveira, V., Kraka, E., Cremer, D.: Quantitative Assessment of Halogen Bonding Utilizing Vibrational Spectroscopy. Inorg. Chem. **56**, 488–502 (2016)

86. Setiawan, D., Sethio, D., Cremer, D., Kraka, E.: From Strong to Weak NF Bonds: On the Design of a New Class of Fluorinating Agents. Phys. Chem. Chem. Phys. **20**, 23913–23927 (2018)

87. Sethio, D., Oliveira, V., Kraka, E.: Quantitative Assessment of Tetrel Bonding Utilizing Vibrational Spectroscopy. Molecules **23**, 2763 (2018)

88. Sexton, T.M., Freindorf, M., Kraka, E., Cremer, D.: A Reaction Valley Investigation of the Cycloaddition of 1,3-Dipoles with the Dipolarophiles Ethene and Acetylene: Solution of a Mechanistic Puzzle. J. Phys. Chem. A **120**, 8400–8418 (2016)

89. Reis, M.C., López, C.S., Kraka, E., Cremer, D., Faza, O.N.: Rational Design in Catalysis: A Mechanistic Study of β-Hydride Eliminations in Gold(I) and Gold(III) Complexes Based on Features of the Reaction Valley. Inorg. Chem. **55**, 8636–8645 (2016)

90. Sexton, T., Kraka, E., Cremer, D.: Extraordinary Mechanism of the Diels-Alder Reaction: Investigation of Stereochemistry, Charge Transfer, Charge Polarization, and Biradicaloid Formation. J. Phys. Chem. A **120**, 1097–1111 (2016)

91. Makoś, M.Z., Freindorf, M., Tao, Y., Kraka, E.: Theoretical Insights into [NHC]Au(I) Catalyzed Hydroalkoxylation of Allenes: A Unified Reaction Valley Approach Study. J. Org. Chem. **86**, 5714–5726 (2021)

92. Freindorf, M., Cremer, D., Kraka, E.: Gold(I)-Assisted Catalysis - A Comprehensive View on the [3,3]-Sigmatropic Rearrangement of Allyl Acetate. Mol. Phys. **116**, 611–630 (2017)
93. Bowman, J.M., Houston, P.L.: Theories and simulations of roaming. Chem. Soc. Rev. **46**, 7615–7624 (2017)
94. Quinn, M.S., Nauta, K., Jordan, M.J.T., Bowman, J.M., Houston, P.L., Kable, S.H.: Rotational resonances in the h_2co roaming reaction are revealed by detailed correlations. Science **369**, 1592–1596 (2020)
95. Endo, T., Neville, S.P., Wanie, V., Beaulieu, S., Qu, C., Deschamps, J., Lassonde, P., Schmidt, B.E., Fujise, H., Fushitani, M., Hishikawa, A., Houston, P.L., Bowman, J.M., Schuurman, M.S., Legare, F., Ibrahim, H.: Capturing roaming molecular fragments in real time. Science **370**, 1072–1077 (2020)
96. Makoś, M.Z., Freindorf, M., Sethio, D., Kraka, E.: New insights into Fe–H_2 and Fe–H^- bonding of a [nife] hydrogenase mimic – a local vibrational mode study. Theor. Chem. Acc. **138**, 76 (2019)
97. Bakhmutov, V.I.: Dihydrogen Bonds - Principles, Experiments, and Applications. Wiley Ingterscience, New York (2008)
98. Karas, L.J., Wu, C.H., Das, R., Wu, J.I.C.: Hydrogen bond design principles. WIREs Comput Mol Sci. **10**, e1477–1–e1477–15 (2020)
99. Kraka, E., Joo, H., Cremer, D.: A stunning example for a spontaneous reaction with a complex mechanism: The vinylidene-acetylene cycloaddition reaction. Mol. Phys. **108**, 2667–2685 (2010)
100. Joo, H., Kraka, E., Quapp, W., Cremer, D.: The Mechanism of a Barrierless Reaction: Hidden Transition State and Hidden Intermediates in the Reaction of Methylene with Ethene. Mol. Phys. **105**, 2697–2717 (2007)
101. Quapp, W., Kraka, E., Cremer, D.: Finding the Transition State of Quasi-Barrierless Reactions by a Growing String Method for Newton Trajectories: Application to the Dissociation of Methylenecyclopropene and Cyclopropane. J. Chem. Phys. A **111**, 11287–11293 (2007)
102. Kraka, E., Larsson, J.A., Cremer, D.: Generalization of the Badger Rule Based on the Use of Adiabatic Vibrational Modes. In: J. Grunenberg (ed.) Computational Spectroscopy, pp. 105–149. Wiley, New York (2010)
103. Kalescky, R., Kraka, E., Cremer, D.: Identification of the Strongest Bonds in Chemistry. J. Phys. Chem. A **117**, 8981–8995 (2013)
104. Zhao, L., Pan, S., Holzmann, N., Schwerdtfeger, P., Frenking, G.: Chemical Bonding and Bonding Models of Main-Group Compounds. Chem. Rev. **119**, 8781–8845 (2019)
105. Gavezzotti, A.: Comparing the strength of covalent hydrogen bonds, intermolecular hydrogen bonds and other intermolecular interactions for organic molecules: X-ray diffraction data and quantum chemical calculations. New J. Chem. **40**, 6848–6853 (2016)
106. Kraka, E., Cremer, D.: Characterization of CF Bonds with Multiple-Bond Character: Bond Lengths, Stretching Force Constants, and Bond Dissociation Energies. ChemPhysChem **10**, 686–698 (2009)
107. Kraka, E., Cremer, D.: Weaker Bonds with Shorter Bond Lengths. Rev. Proc. Quim. pp. 39–42 (2012)
108. Kraka, E., Setiawan, D., Cremer, D.: Re-Evaluation of the Bond Length-Bond Strength Rule: The Stronger Bond Is not Always the Shorter Bond. J. Comp. Chem. **37**, 130–142 (2015)
109. Kaupp, M., Danovich, D., Shaik, S.: Chemistry is about energy and its changes: A critique of bond-length/bond-strength correlations. Coord. Chem. Rev. **344**, 355–362 (2017)
110. Takeuchi, Y., Furukawa, Y., Kobayashi, T., Sekine, T., Terada, N., Kakegawa, T.: Impact-induced amino acid formation on Hadean Earth and Noachian Mars. Sci. Rep. **10**, 9220–1–9220–7 (2020)
111. Puzzarini, C., Salta, Z., Tasinato, N., Lupi, J., Cavallotti, C., Barone, V.: A twist on the reaction of the CN radical with methylamine in the interstellar medium: new hints from a state-of-the-art quantum-chemical study. MNRAS **496**, 4298–4310 (2020)

Appendix D

Paper on Ring Puckering and Hydrogen Bonding in Deoxyribonucleotides

Nassim Beiranvand and Marek Freindorf and Elfi Kraka, The Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleotides.
To be submitted to the Journal of .

The Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleotides

Nassim Beiranvand, Marek Freindorf, and Elfi Kraka*

Computational and Theoretical Chemistry Group (CATCO),

Department of Chemistry, Southern Methodist University,

3215 Daniel Ave, Dallas, Texas 75275-0314, USA

E-mail: ekraka@smu.edu

Abstract

Introduction

Deoxyribonucleotides play essential roles in the regulation and modulation of a wide variety of biological systems, and their biological importance has increased significantly in recent years.[1–4] It consist of three major components: a sugar ring, a phosphate group and a nitrogenous base and joined together by intermolecular hydrogen bonds (H-bonds). The study of the H-bonds between DNA unit has long been a topic of significant investigation.[5–9] However, only a few studies have been devoted to evaluate the intramolecular H-bonds within a deoxyribonucleotide. The investigation of these intramolecular H-bonds in the case of deoxyribonucleotides is very complicated, there are numerous intramolecular H-bonds like O–H$\cdots$O and C–H$\cdots$O bond between phosphate group and nitrogenous base as well as sugar ring . In this situation this is quite reasonable that the orientation of the phosphate

group and nitrogenous base with respect to the sugar ring can determine the the formation and existence of those H-bonds which is also concerns the equilibrium conformation of deoxyribonucleotides. Also, dependency of sugar ring puckering and the formation of internal H-bonding and vise versa should be consider as an important factor.[10–12] We used in this work as efficient tool a combination of the Cremer-Pople *Ring Puckering Analysis*[13–19] and the local mode analysis of Konkoli and Creme (LVM) of Konkoli and Cremer,[20–24] to determine the CES and pseudo-rotation path and the location of the global on the CES for the Deoxycytidine monophosphate (**dCMP** (Figure 1) which is a deoxyribonucleotides, and one of the four monomers that make up DNA[25]), in order to answering the following questions:

1. How the conformational energy, the puckering amplitude change along the pseudorotation path of sugar ring?

2. How the intramolecular H-bonds change along the pseudorotation path of sugar ring? Is there any connection?

3. What type of H-bonds contribution is possible between phosphate group and nitrogenous base and sugar ring?

4. Compare with our previous work,[26] Deoxyribonucleosides to understand the role of phosphate group?

The paper is structured in the following way. In the second section the methods used in this work are described as well as computational details. The third section presents the results and discussion part. Conclusions and an outlook are made in the final section.

Methodologies

The important tools applied in this work are introduced in this section, the Cremer-Pople ring puckering analysis[13–19] , the local mode analysis of Konkoli and Cremer[20–24] and the topolog-

ical analysis of the electron density $\rho(\mathbf{r})$ via Bader's quantum theory of atoms in molecules (QTAIM) approach.[27,28] Then the computational details of this study are described.

Ring Puckering Analysis

Configurational space is defined by 3N-6 internal parameters that completely describe the geometry of the molecule. But it is difficult to describe the puckered N-membered ring motion via internal coordinate or Cartesian coordinate. In 1972, Altona and Sundaralingam[29] introduced a procedure describing both the puckering angle and puckering amplitude of a five-membered ring via torsional angles. However, the definition of the ring puckering amplitude expressed via an angle does not allow the appropriate description of the ring pseudo-rotation process. In this process, the maximum puckering amplitude is rotating around the edge of the ring without raising a substantial potential energy change. Such a movement can be best described with a puckering amplitude defined as displacement from a reference. Cremer and Pople[13] proposed to use N-3 puckering coordinates for extending a puckered N-membered ring conformational space. For odd-membered rings, it is possible to divide the N-3 puckering coordinates into pairs of pseudo-rotational coordinates $\{q_m, \phi_m\}$ ($m = 2, 3, \dots$ N-3) and for even-membered rings, an additional puckering amplitude $q_{N/2}$ that would describe ring inversion. The mth puckering mode's ring puckering degree is represented by the puckering amplitude q_m. The ring puckering mode is defined by the puckering angle ϕ_m.[30,31] Proposed by Cremer and Pople, the plane of the planar ring can apply to functioning as a mean plane as reference. The N-3 puckering coordinate pairs q_m, ϕ_m can help determine the out-of-plane coordinates z_j that describe the displacement perpendicular to the mean plane.[13,30,32–34] Despite the Infinite number of ring conformations placed on a pseudorotation cycle for deoxycytidine' sugar ring, it is sufficient to investigate a representative number of conformations. Referring to previous research on description of five-membered sugar ring in deoxyribose[26], a subset of 20 ring conformations were optimized using a mixed set of puckering and internal coordinates. Since the entire study of deoxycytidine in this research

have C_1 symmetry, it can be concluded that the conformational energy surface (CES) should also have C_1 symmetry. The application of ring puckering coordinates yields two main advantages.[35,36] At first, it is the optimization phase during which, for any given value of q_2 and ϕ_2, sugar ring's geometry is optimized, despite the possible lacking a stationary point on the CES. A Cartesian or internal coordinates based description of the ring would not make this possible.[16] A mixed set of puckering coordinates and internal/Cartesian coordinates for the substituents can apply to specifically identifying any conformer located on the CES for substituted ring systems such as Deoxycytidine. Besides, it is possible to gain physically meaningful pseudo-rotation path as a function of two. The expression of any property P of a puckered ring including geometry, magnetic properties, energy, dipole moment, charge distribution, and vibrational frequencies as a Fourier series of the puckering coordinates is the other advantage. For a detailed mathematical derivation, the interested readers are referred to ref.[26]

Local Vibrational Mode Theory

The normal vibrational modes (3N-L) of an N-atomic molecule has all important molecule's electronic structure and therefore, should be well suited as a measure of bond strength. However, in systems larger than diatomic molecules, the normal vibrational modes cannot assess bond strength because it is difficult to decode this information into individual atom-atom interactions. These modes are delocalized because of electronic and mass mode–mode coupling.[37–40] Dieter Cremer and Zoran Konkoli have developed an approach to localize vibrational motions to isolate the frequency of a given internal coordinate (q_n) by decoupling normal modes, including puckering coordinate, bond angle (θ), bond length (r), and dihedral (ϕ).[20–24,41] The fundamental aspect was to derive constants of local stretching force from local vibrational modes (LVMs), free of any mode-mode coupling, which would then serve as a direct measure of a chemical bond's intrinsic strength. In the Wilson equation's mass–decoupled analog, the mass–coupling was solved[20] that gives an LVM free of any mode–mode

coupling.[24] LVM was used soon after its initial formulation in 1997 to measure the chemical bond's intrinsic strength.[42–45] Each LVM is associated with a q_n corresponded to a local mode force constant (k^a) and a local mode frequency (ω^a) . There is a unique, direct, and vibrational spectroscopy-based relationship between the local mode force constant and a bond's intrinsic strength.[46] In 2012, Wenli Zou et al. demonstrated a one–to–one relationship between a non–redundant set of local vibrational modes and normal vibrational modes by an adiabatic connection scheme (ACS) known as the foundation for the decomposition of normal modes into local mode contributions (CNM), leading to a basic physical foundation for LVM.[47–49] As such, it resulted in a detailed analysis of a vibrational spectrum from a computational and measurement point of view.[22,23] The new theory was off and running at this point, which is applied to several chemical systems with both noncovalent and covalent interactions (NCI).Thus, LVM analysis has advanced as a powerful analytical tool and ever since has been extensively applied to a broad range of chemical systems covering both covalent bonds[46,50–61] and in particular, weak chemical interactions including hydrogen bond,[26,47,62–68] halogen bond,[69–74] pnicogen bond,[49,75,76] chalcogen bond,[60,77,78] tetrel bond,[79] and atom-π interactions.[80,81] LVM analysis was performed for the investigation of the hydrogen bonds strength in our projects. In my projects the bond strength assessment was simplified[55,57] by converting k^a to bond strength order (BSO n) with the extended Badger[55,57,82] and two reference bonds:

$$BSO\ n\quad =\quad a\,(k^a)^{\,b} \tag{1}$$

The constants a and b in Eq. (1) were determined via Eq. (2) where for hydrogen bonds $n = 1$ for the FH bond and $n = 0.5$ for the FH bond in $[\text{F}\cdots\text{H}\cdots\text{F}]^-$ anion were used,[65] leading to $a = 0.515$ and $b = 0.291$, and two reference values and the requirement that for a zero force constant the BSO n is zero.

$$\text{BSO}n(scaled) = 0.515(k^a)^{0.291} \tag{2}$$

According to Eq. (1) the OH bond in H_2O has a BSO n value of $= 0.966$. Therefore, we scaled the reference values, so that the BSO n of the OH bond in H_2O is 1.

QTAIM analysis

The Quantum Theory of Atoms-In-Molecules developed by Bader[83,84] provides a theoretical scheme for identifying, evaluating and characterizing chemical bonds and interactions through the total electron density topological features $\rho(\mathbf{r})$. In this work we used QTAIM as a complementary tool to the LVM analysis to evaluate the covalent/electrostatic character of intramolecular H-bonds via the Cremer-Kraka criterion[85–87] of covalent bonding.

The Cremer-Kraka criterion for covalent bonding is composed of two conditions: necessary condition: existence of a bond path and bond critical point $\mathbf{r_c} = c$ between two atoms; sufficient condition: energy density $H(\mathbf{r_c}) = H_c$ is smaller than zero. $H(\mathbf{r})$ is defined as:

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) \tag{3}$$

where $V(\mathbf{r})$ represents the potential energy density and $G(\mathbf{r})$ represents the kinetic energy density. A stabilizing accumulation of density leads to a negative $V(\mathbf{r})$ while depletion of electron density refers to the positive $G(\mathbf{r})$.[86] Consequently, the H_c sign indicates which term is the dominant expression.[87] The interaction is considered covalent in nature If $H_c < 0$, while $H_c > 0$ is representative of electrostatic interactions.

Computational methods

All calculations were performed with the ωB97X-D functional[88,89] and Pople's 6-31++G(d,p) basis set.[90–93] Tight convergence criteria were applied, (SCF iterations: 10^{-10} Hartree and geometry optimizations: 10^{-7} Hartree/Bohr) and an ultra-fine grid was used for the DFT numerical integration.[94] First, the most stable conformer for dC, e.g. the global minimum, was found by fully optimizing the structures derived from all possible H-bond combinations

using Cartesian coordinates. The pseudo-rotation path was then calculated in the following way. A puckering analysis was performed for the global minimum to obtain the corresponding phase angle ϕ_2 and the puckering amplitude q_2. Starting from this phase angle 20 representative path points were calculated by constrained optimizations with a fixed phase angle ϕ_2, modified in increments of 18° for e.g. $\phi_2 = 0, 18, 36, \ldots 342°$ to determine the pseudo-rotation path. For the constrained optimizations a mixed set of internal coordinates for the substituents and puckering coordinates for deoxyribose sugar ring was used.

Harmonic frequency calculations were carried out in each case to confirm the optimized conformers located on the pseudo-rotation path as minima (no imaginary frequency) or transition states (one imaginary frequency). The CES were also mapped via constrained optimizations. For each fixed phase angle ϕ_2, the puckering amplitudes q_2 were constrained to four values, namely 0.1, 0.2, 0.3, 0.4 and 0.5Å, and the remainder of the molecule was optimized, leading to 80 data points for CES. Then on the pseudo-rotation path, a local mode analysis was performed for the investigation of the H-bond strength.

The RING Puckering program[15] was used for analysis of the deoxyribose sugar ring conformations. For the geometry optimizations and frequency calculations, a combination of the RING Puckering program and the program package Gaussian09[95] was used. All local mode analyses calculations were carried out with the program package COLOGNE2019.[96]

Results and discussion

In Table 1, the properties of the dC calculated at different phase angle ϕ_2 are listed. These comprise the puckering amplitude q_2, relative energy ΔE with regard to the planar form and H-bond distance R, force constant k^a, vibrational frequency ω^a. Table 1 also contains the electron density $\rho(c)$ and energy density H_c at the bond critical points c.

Table 1: Energies, geometries, vibrational and electron density data for **dCMP** along their pseudo-rotation path as calculated at the ωB97X-D/6-31++G(d,p) level of theory.

ϕ_2	q_2	ΔE	Bond	R	k^a	ω^a	BSO n	ρ_c	H_c
180°	0.159	11.96	O—H$\cdots$O	1.952	0.062	332	0.263	0.171	-0.004
			C1–H$\cdots$O	2.283	0.247	666	0.377	0.120	0.009
			C2–H$\cdots$O	2.471	0.212	616	0.362	0.063	0.005
198°	0.237	10.42	O—H$\cdots$O	1.902	0.135	491	0.321	0.189	-0.003
			C1–H$\cdots$O	2.296	0.239	654	0.373	0.119	0.008
			C2–H$\cdots$O	2.355	0.035	251	0.228	0.083	0.003
216°	0.246	9.00	O—H$\cdots$O	1.862	0.175	559	0.344	0.205	-0.003
			C1–H$\cdots$O	2.302	0.256	678	0.380	0.119	0.008
			C2–H$\cdots$O	2.175	0.145	509	0.327	0.118	-0.002
234°	0.272	7.00	O—H$\cdots$O	1.871	0.164	542	0.338	0.199	-0.002
			C1–H$\cdots$O	2.375	0.245	663	0.376	0.105	0.009
			C2–H$\cdots$O	2.002	0.261	684	0.382	0.167	-0.006
252°	0.365	3.31	O—H$\cdots$O	1.855	0.162	539	0.337	0.205	-0.002
			C2–H$\cdots$O	1.961	0.288	718	0.392	0.182	-0.006
270°	0.408	0.28	O—H$\cdots$O	1.839	0.167	547	0.340	0.213	-0.002
			C2–H$\cdots$O	2.060	0.205	605	0.358	0.149	-0.006
288°	0.400	0.56	O—H$\cdots$O	1.835	0.162	538	0.337	0.216	-0.002
			C2–H$\cdots$O	2.245	0.094	411	0.292	0.102	-0.002
306°	0.370	3.04	O—H$\cdots$O	1.827	0.179	567	0.346	0.221	-0.002
			C2–H$\cdots$O	2.566	0.060	329	0.261	0.052	0.005
324°	0.355	5.62	O—H$\cdots$O	1.799	0.208	611	0.360	0.233	-0.001
342°	0.355	7.68	O—H$\cdots$O	1.802	0.192	586	0.352	0.231	-0.001

[*] Phase angle ϕ_2 in degree, puckering amplitude q_2 in Å, ΔE in kcal/mol with respect to the planar form as reference, bond length values R in Å, force constant k^a in mdyne/Å, vibrational frequency ω^a in cm^{-1}, electron density $\rho(c)$ at the bond critical points in e/Å^3, energy density H_c at the bond critical points in Hartree/Å^3. (Numbering in C1 and C2 is just for distinguishing between the C–H$\cdots$O bond of phosphate group with deoxyribose ring and nitrogenous base respectively.)
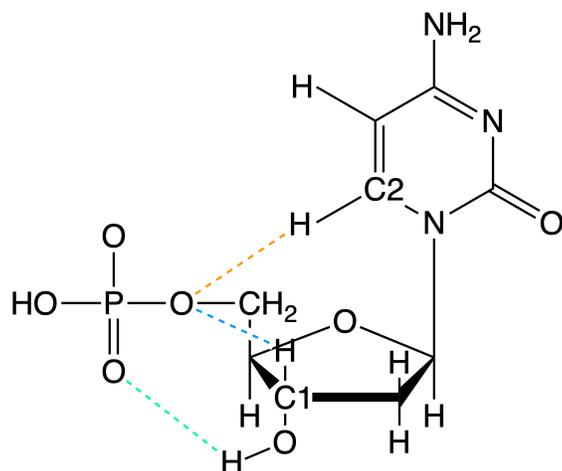
Figure 1: Deoxycytidine structure investigated in this work. Numbering in C1 and C2 is just for distinguishing between the C–H···O bond of phosphate group with ribose ring and nitrogenous base respectively.

CESs and pseudo-rotation (pseudo-libration) paths

The pseudo-rotation paths on the corresponding CES of the **dCMP** is shown in Figure 2. We found negative puckering amplitudes on the pseudo-rotation paths in the range of $\phi_2 = 0$ to $162°$ indicating a ring inversion. While in the case of deoxyribonucleosides we found negative puckering amplitudes in the range of $\phi_2 = 108$ to $270°$ which is longer path. But similar to the deoxyribonucleosides the pseudo-rotation path is an incomplete open curve, which we made as pseudo-libration path. Global minimum was identified in the $\phi_2 = 270$ °, $\Delta E = 0.28$ kcal/mol on the pseudo-rotation path (Figure 3). At the beginning of path where $\phi_2 = 180.0°$ with the highest energy ($\Delta E = 11.96$ kcal/mol) respect to the planar form, right after this point the conformer undergoes a complete rotation as shown in Figure 3.
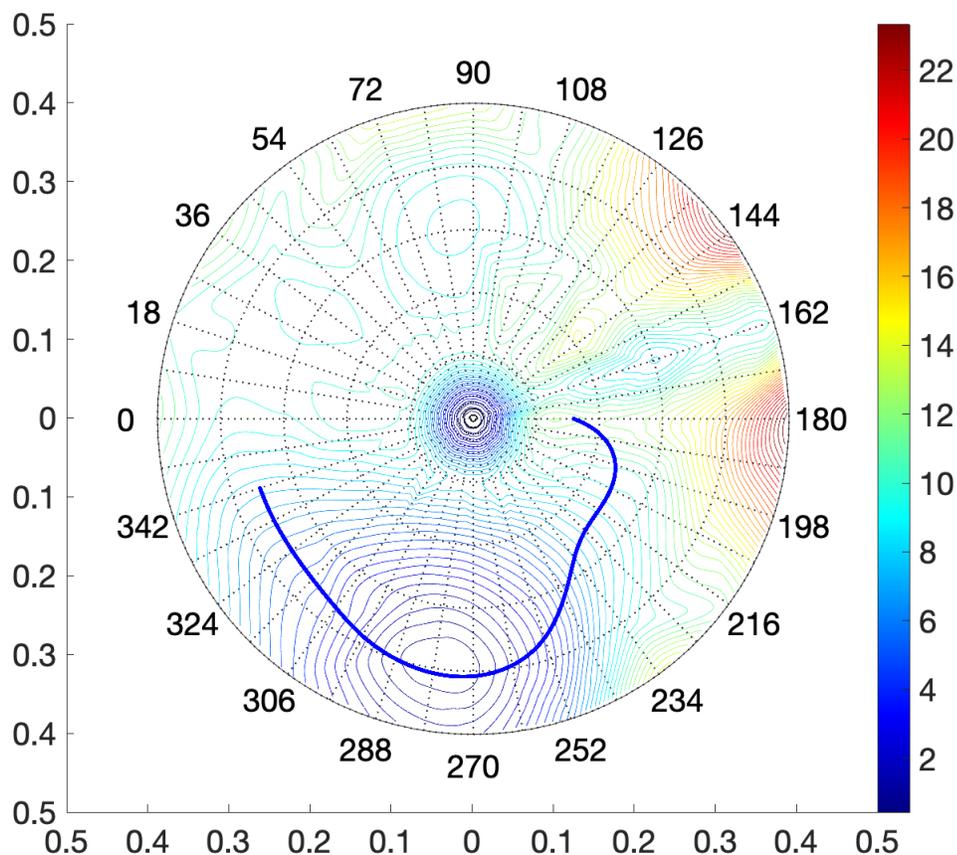
Figure 2: Conformational Energy Surfaces (CESs) and pseudo-rotation path of **dCMP**. At the center of CES, the planar deoxyribose ring form is located. The corresponding energy is used as reference. The X axe shows the magnitude of puckering amplitude q_2 in Å. The labels around CES circle denote the value of the phase angle ϕ_2 in Degree. The solid blue line indicates the pseudo-rotation path. The color bar represents the energy on the CES in kcal/mol relative to the planar form; *yellow to red regions:* location of conformers higher in energy than the planar form; *green to blue regions:* location of conformers lower in energy than the planar form. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.
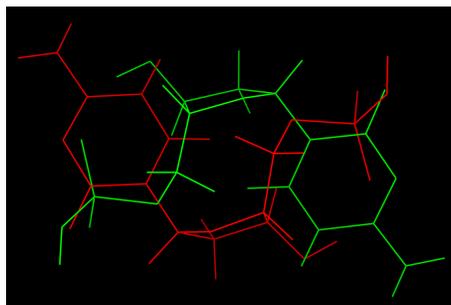

Figure 3: Rotation of **dCMP**; conformer 1 shown in green, $\phi_2 = 180°$ and $q_2 = 0.159$Å, conformer 2 shown in red, $\phi_2 = 198°$ and $q_2 = 0.237$ Å.

More detailed information is obtained by plotting changes of ΔE, q_2, R, and BSO n as a function of ϕ_2 as shown in Figure 4. A comparison of Figures 4a and 4b reveals how the puckering amplitude and the conformational energy are related. The largest puckering amplitude $q_2 = 0.408$ Åoccurs in the $\phi_2 = 270$ ° which is the global minimum region. Suggesting that the larger the puckering amplitude, the lower the conformational energy which is in agreement with our pervious work on Deoxyribonucleosides. Figures 4c and 4d show the change of the internal H-bond distance R and the corresponding H-bond strength BSO n along the pseudo-libration path. Our findings show that the possibility of forming three H-bonds, O–H···O bond and two C–H···O bond which occurs between phosphate group with ribose ring and nitrogenous base whereas just Deoxyribonucleosides had the possibility of forming one H-bond. Hence, the additional H-bonding constraints flexibility and it affects conformational energy. The number of H-bonds decreases along the pseudo-rotation path, as there are three H-bonds at the starting path while it reduces to two bonds in global minimum and at the end points of the path reaches to one bond. Our BSO n analysis show the global minimum doesn't necessarily poses the strongest H-bond, although the strongest H-bond (C–H···O bond, BSO $n = 0.392$) happens between phosphate group and nitrogenous base at the $\phi_2 = 252$ ° which is very close to the global minimum. The O–H···O bond between phosphate group and OH group of deoxyribose ring, constantly present with increasing trend of strength along the pseudo-rotation path.
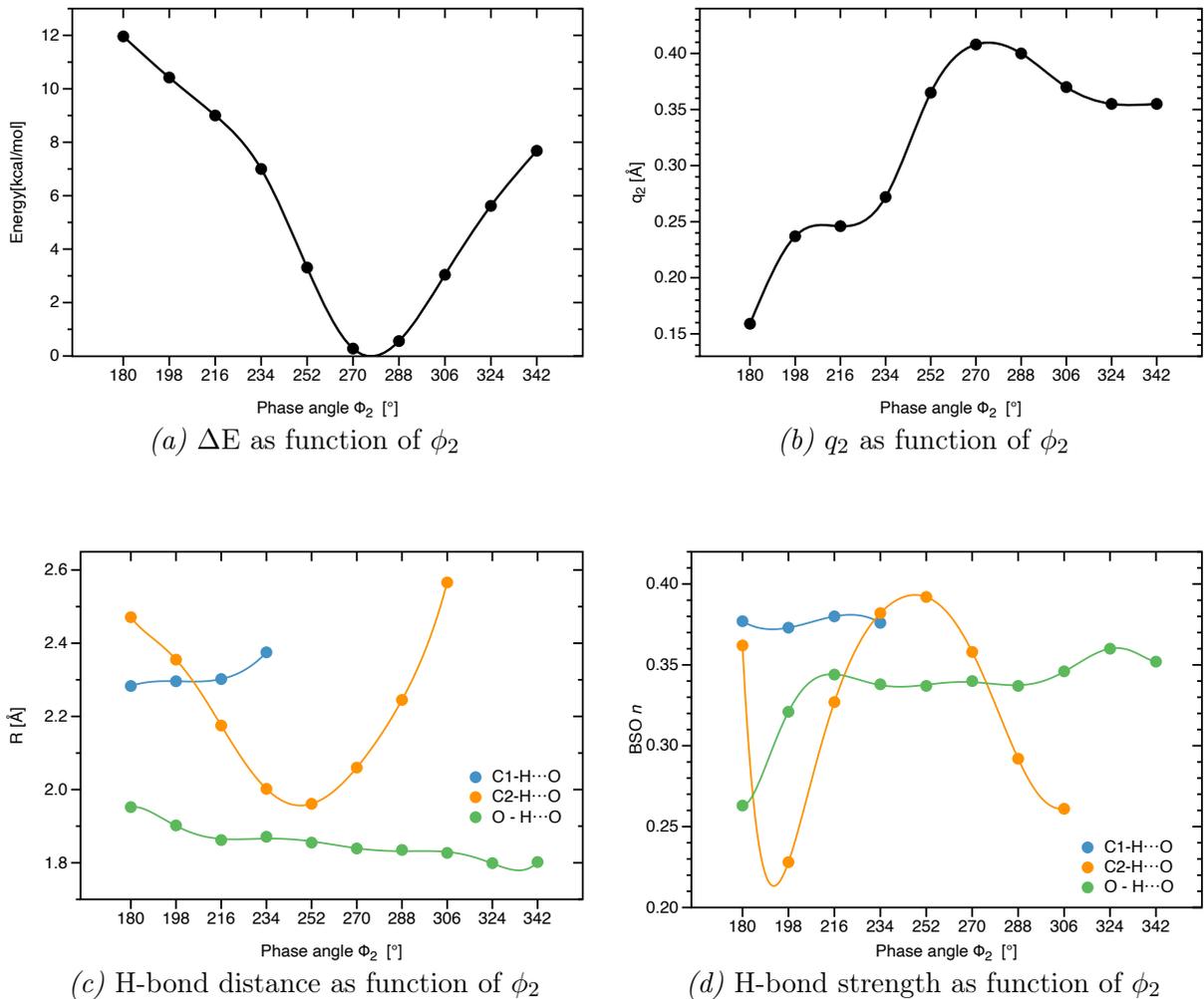
(a) ΔE as function of ϕ_2

(b) q_2 as function of ϕ_2

(c) H-bond distance as function of ϕ_2

(d) H-bond strength as function of ϕ_2

Figure 4: (a) Changes of the conformational energy ΔE with regard to the planar form; (b) the puckering amplitude q_2, (c) the H-bond distance R, and (d) H-bond strength order BSO n along the pseudo-libration path described as a function of the puckering angle ϕ_2. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

As reflected by Figure 5, there is direct relationship between H-bond length and H-bond strength for O–H$\cdots$O bonds, and to some extent for C2–H$\cdots$O bond but not in the case of C1–H$\cdots$O bonds. And the direct relationship is not always true as numerous examples have shown.[59,61]
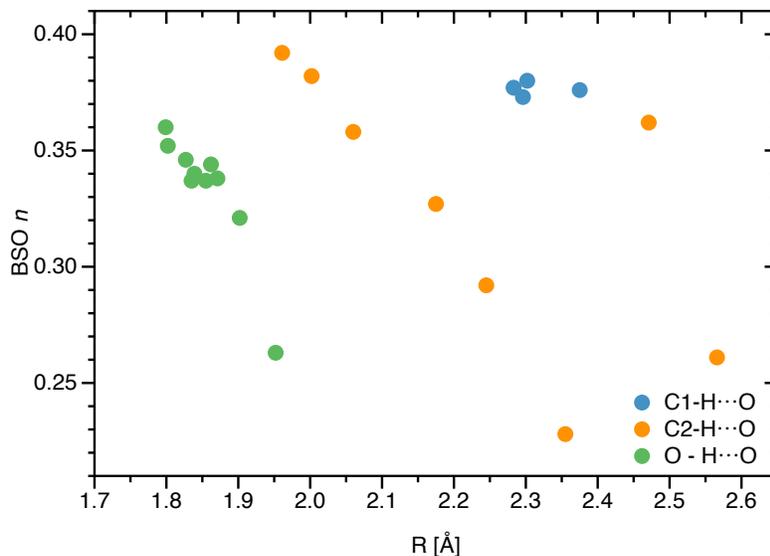
12

Figure 5: Correlation between H-bond strength and H-bond length. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

Internal H-bond strength and covalent/electrostatic character

The weakest H-bond (BSO $n = 0.228$) is found for the C2–H$\cdots$O bond at ϕ_2=198.0° which is weaker than the H-bond in the formaldehyde dimer **R3**, (BSO $n = 0.241$, see Table 2). The C2–H$\cdots$O bond at ϕ_2=252.0° has the strongest H-bond (BSO $n = 0.392$) see Figure 6, which is even stronger than the H-bond in the water dimer R1 with a BSO n value of 0.360, see Table 2. The percentage of conformers with BSO n values between 0.3 and 0.40 is 82.0%, in the range of the most common H-bond types in water clusters.[65] This clearly shows that the intermolecular H-bonds are of considerable strength and that they are an important feature of **dCMP**.
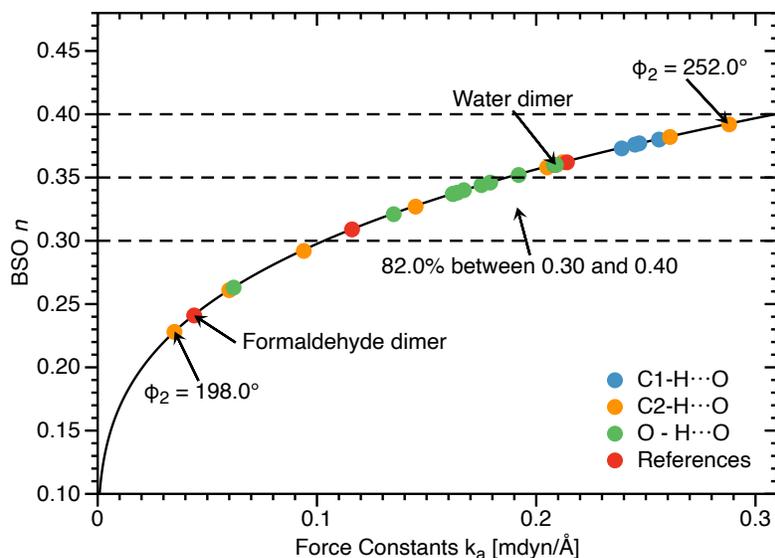
Figure 6: Bond strength BSO n of the H-bonds of **dCMP** as a function of the corresponding local stretching force constant k^a as determined via Equation 2. For comparison H-bonded complexes **R1** - **R4** defined in Figure 7 are included. Dashed horizontal lines mark BSO n values of 0.3, 0.35, and 0.4 respectively. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.
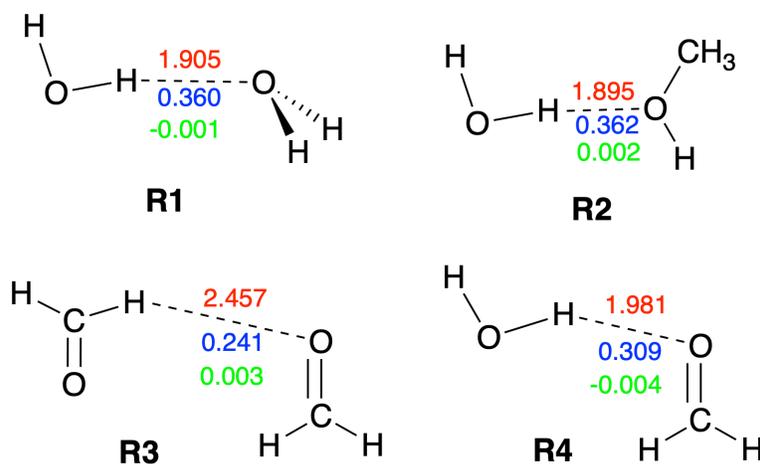
Figure 7: Reference complexes **R1** - **R4**. The intramolecular H-bond distance (Å) is given in red color, the corresponding BSO n value in blue color, and energy density $H(c)$ (Hartree/Å^3) in green color. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

In addition to the H-bond strength we evaluated the covalent/electrostatic character of the H-bonds applying the Cremer-Kraka criterion of covalent bonding[85–87,97] described above. As reflected by data in Table 1 H_c values range from -0.006 Hartree/Å^3 (covalent

14

Table 2: Properties of hydrogen bonded reference complexes **R1** - **R4**

Molecule	k^a	BSO n	R(HB)	ρ_c	H_c
R1	0.209	0.360	1.905	0.183	-0.001
R2	0.214	0.362	1.895	0.176	0.002
R3	0.044	0.241	2.457	0.071	0.003
R4	0.116	0.309	1.981	0.162	-0.004

Force constant k^a in mdyne/Å, H-bond length R(HB) in Å, ρ_c in e/Å^3 and H_c in Hartree/Å^3. Reference complexes **R1** - **R4** are shown in Figure 7. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

character) for C2—H$\cdots$O at the $\phi_2 = 234°$, $252°$ (strongest H-bond, BSO $n = 0.392$) and $270°$ to 0.009 Hartree/Å^3 (electrostatic character) for C1—H$\cdots$O at the $\phi_2 = 180°$ starting point (BSO $n = 0.377$) and $\phi_2 = 1234°$ (BSO $n = 0.376$). Similar values were also found for the H-bonds of the reference compounds **R1** - **R4**, see Table 2. Figure 8a does not show
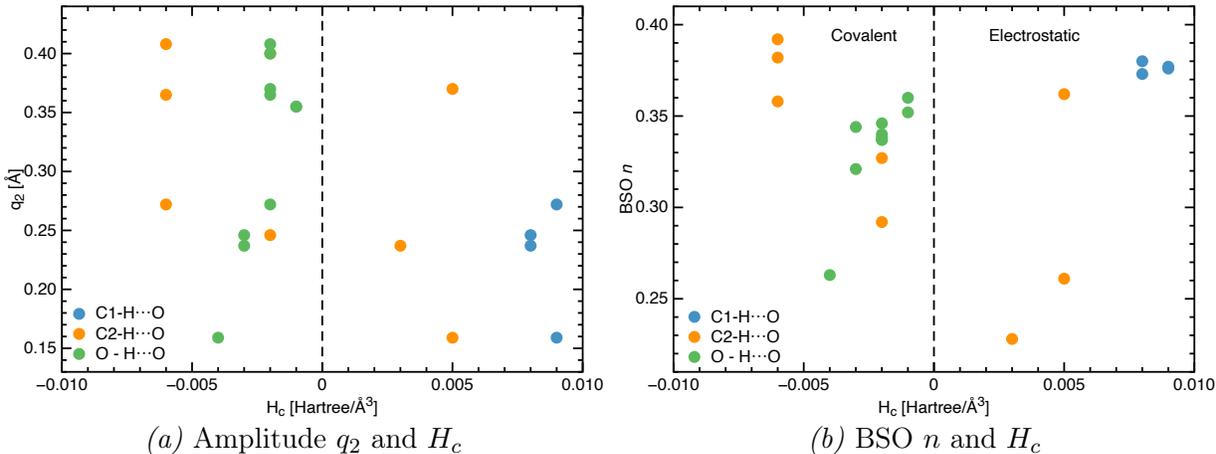

(a) Amplitude q_2 and H_c

(b) BSO n and H_c

Figure 8: (a) Correlation between amplitude q_2 and energy density H_c. (b) Correlation of BSO n and energy density H_c. Calculated at the ωB97X-D/6-31++G(d,p) level of theory.

general correlation between energy density H_c values and the amplitudes q_2 for H-bonds of **dCMP** and it was not far from our expectations. Although H_c is taken at a single point, the amplitude $q2$ is a more global property representing complex geometry changes caused by ring puckering. However, one finds a clear separation into different types of H-bond

type, which according to the correlation of BSO n values with H_c (the Figure 8b), leads to a different electronic environment. O–H$\cdots$O bonds are spread from -0.004 Hartree/Å^3 to -0.001 Hartree/Å^3 in covalent region and C1—H$\cdots$O bonds with $H_c = 0.008$ Hartree/Å^3 and 0.009 Hartree/Å^3 has the less covalent character. C2—H$\cdots$O bonds are scattered in both covalent and electrostatic regions, starting with the most covalent character -0.006 Hartree/Å^3 and ending with 0.005 Hartree/Å^3. As we described earlier the C2—H$\cdots$O bond is formed between phosphate group and nitrogenous base which at the beginning of pseudo-rotation path $\phi_2 = 180°$ has less covalent character and becomes more covalent as close as it gets to the global minimum.

Conclusions and outlook

For the first time, the Cremer-Pople ring puckering analysis and the Konkoli-Cremer local mode analysis, accompanied by topological analysis of the electron density, were used to systematically investigate the interplay between deoxyribose ring puckering and intramolecular H-bonding in Deoxycytidine (**dCMP**). Our work has led to the following conclusions:

- We found incomplete pseudo-rotation paths on the CESs caused by ring inversion for **dCMP**, similar to deoxyribonucleosides in our previous study[26] where we we coined an open curve pseudo-rotation path as pseudo-libration. But due to the phosphate group, there are more H-bonds, which shortens the pseudo-libration path. Also, on the pseudo-libration paths a global minimum in the range of $\phi_2 = 270$- $288°$ could be identified but not a local minimum or transition state.

- Our work showed that ring puckering plays the dominant role on determining the conformational flexibility of **dCMP** among two major factors (ring puckering and internal H-bonding). The largest puckering amplitude $q_2 = 0.408$ Åoccurs in the $\phi_2 = 270$ ° which is the global minimum region. This suggests that the lower the conformational energy, the larger the puckering amplitude, which is consistent with our previous

work.[26]

- We found three H-bonds (O–H$\cdots$O bond and two C–H$\cdots$O bond which occurs between phosphate group with ribose ring and nitrogenous base of **dCMP**), then quantitatively assessed the H-bonds strength along the pseudo-rotation paths via bond strength orders BSO n derived from local vibrational force constants. The number of H-bonds decreases along the pseudo-rotation path, from three at the start to two at the global minimum and one at the end of the path. The O–H$\cdots$O bond between phosphate group and OH group of deoxyribose ring, constantly present with increasing trend of strength along the pseudo-rotation path but C1–H$\cdots$O bond just found in $\phi_2 = 180$-234 °. We found a direct correlation between H-bonds strength and H-bonds length for O–H$\cdots$O bond and C2–H$\cdots$O bond but we did not find it in C1–H$\cdots$O. We found the shortest and strongest O–H$\cdots$O bond and C2–H$\cdots$O bond at $\phi_2 = 324$ ° outside the global minimum region close to the end point and $\phi_2 = 252°$ respectively. These results clearly reveal that H-bonding determines the shape and length of the pseudo-libration paths.

  This new analysis could be extended to different types of deoxyribonucleotides to understand the influence of puric and pyrimidic bases on formation internal H-bonding.

Acknowledgement

Supporting Information Available

References

(1) Sengupta, R.; Coppo, L.; Sircar, E.; Mishra, P.; Holmgren, A. S-Denitrosylation by the C-Terminal Swinging Arm of R1 Subunit: A Novel Mechanism to Restore Ribonucleotide Reductase Activity. *ChemistrySelect* **2021**, *6*, 1845–1851.

(2) Chingbiaknem, E.; Lyngdoh, R. D. Isomorphic Building Blocks For Information-bearing

Duplexes?part 2: Pyrimidine Base pairs With Sugar Phosphate Backbones. *Structural Chemistry* **2021**, 1–12.

(3) Brutus, L. C.; Jamieson, E.; Suarez, C.; Nunez, M. E. Investigation of the Spiroiminodi-hydantoin Lesion's Structural and Dynamic Effects on an 11-Mer Deoxyribonucleotide Duplex. *Biophys. J.* **2020**, *118*, 63a–64a.

(4) Chen, S.; Zhao, J.; Yang, X.; Zhao, S.; Liu, Y.-M. A Novel Intracellular Signal Amplification Strategy For the Quantification of ATP in Single Cells by Microchip Electrophoresis with Laser-induced Fluorescence Detection. *Chem. Commun.* **2020**, *56*, 6579–6582.

(5) Brovarets′, O. O.; Yurenko, Y. P.; Hovorun, D. M. Intermolecular CH···O/N H-bonds in the Biologically Important Pairs of Natural Nucleobases: A Thorough Quantum-Chemical Study. *J. Biomol. Struct. Dyn.* **2014**, *32*, 993–1022.

(6) Rimola, A.; Sodupe, M.; Ugliengo, P. Role of Mineral Surfaces in Prebiotic Chemical Evolution. In Silico Quantum Mechanical Studies. *Life* **2019**, *9*.

(7) Dąbkowska, I.; Jurečka, P.; Hobza, P. On geometries of stacked and H-bonded nucleic acid base pairs determined at various DFT, MP2, and CCSD(T) levels up to the CCSD(T)/complete basis set limit level. *J. Chem. Phys.* **2005**, *122*, 204322.

(8) Aray, Y.; Aguilera-García, R.; Izquierdo, D. R. Exploring the nature of the H-bonds between the human class II MHC protein, HLA-DR1 (DRB*0101) and the influenza virus hemagglutinin peptide, HA306-318, using the quantum theory of atoms in molecules. *J. Biomol. Struct. Dyn.* **2018**, *37*, 48–64.

(9) Hocquet, A.; Ghomi, M. The Peculiar Role of Cytosine in Nucleoside Conformational Behavior: Hydrogen Bond Donor Capacity of Nucleic Bases. *Phys. Chem. Chem. Phys.* **2000**, *2*, 5351–5353.

(10) Altona, C. t.; Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.

(11) Detecting, R. modifications in the epitranscriptome: predict and validate Helm, Mark; Motorin, Yuri. *Nature Reviews Genetics* **2017**, *18*, 275–291.

(12) Zhu, Y.; Hamlow, L.; He, C.; Roy, H.; Cunningham, N.; Munshi, M.; Berden, G.; Oomens, J.; Rodgers, M. Conformations and N-glycosidic bond stabilities of sodium cationized 2'-deoxycytidine and cytidine: solution conformation of [Cyd$^+$ Na]$^+$ is preserved upon ESI. *Int. J. Mass Spectrom.* **2018**, *429*, 18–27.

(13) Cremer, D.; Pople, J. A. General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.

(14) Cremer, D.; Pople, J. A. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. XXIII. Pseudorotation in Saturated Five-Membered Ring Compounds. *J. Am. Chem. Soc.* **1975**, *97*, 1358–1367.

(15) Cremer, D. RING - A Coordinate Transformation Program for Evaluating the Degree and Type of Puckering of a Ring Compound. *Quantum Chemical Program Exchange* **1975**, 1–8.

(16) Cremer, D. Calculation of Puckered Rings with Analytical Gradients. *J. Phys. Chem.* **1990**, *94*, 5502–5509.

(17) Cremer, D. Theoretical Determination of Molecular Structure and Conformation. XI. the Puckering of Oxolanes. *Isr. J. Chem.* **1983**, *23*, 72–84.

(18) Essén, H.; Cremer, D. On the Relationship Between the Mean Plane and the Least-Squares Plane of an N-membered Puckered Ring. *Acta Crystallogr. B* **1984**, *40*, 418–420.

(19) Wu, A.; Cremer, D. New Approach for Determining the Conformational Features of Pseudorotating Ring Molecules Utilizing Calculated and Measured NMR Spin-Spin Coupling Constants. *J. Phys. Chem. A* **2003**, *107*, 1797–1810.

(20) Konkoli, Z.; Cremer, D. A New Way of Analyzing Vibrational Spectra. I. Derivation of Adiabatic Internal Modes. *Int. J. Quant. Chem.* **1998**, *67*, 1–9.

(21) Konkoli, Z.; Larsson, J. A.; Cremer, D. A New Way of Analyzing Vibrational Spectra. II. Comparison of Internal Mode Frequencies. *Int. J. Quant. Chem.* **1998**, *67*, 11–27.

(22) Konkoli, Z.; Cremer, D. A New Way of Analyzing Vibrational Spectra. III. Characterization of Normal Vibrational Modes in terms of Internal Vibrational Modes. *Int. J. Quant. Chem.* **1998**, *67*, 29–40.

(23) Konkoli, Z.; Larsson, J. A.; Cremer, D. A New Way of Analyzing Vibrational Spectra. IV. Application and Testing of Adiabatic Modes within the Concept of the Characterization of Normal Modes. *Int. J. Quant. Chem.* **1998**, *67*, 41–55.

(24) Cremer, D.; Larsson, J. A.; Kraka, E. In *Theoretical and Computational Chemistry*; Parkanyi, C., Ed.; Elsevier: Amsterdam, 1998; pp 259–327.

(25) Tsuchiya, K.; Sano, T.; Tomioka, N.; Kohzu, A.; Komatsu, K.; Shinohara, R.; Shimode, S.; Toda, T.; Imai, A. Incorporation Characteristics of Exogenous 15N-labeled thymidine, Deoxyadenosine, Deoxyguanosine and Deoxycytidine into Bacterial DNA. *PloS one* **2020**, *15*, e0229740.

(26) Lyu, S.; Beiranvand, N.; Freindorf, M.; Kraka, E. Interplay of Ring Puckering and Hydrogen Bonding in Deoxyribonucleosides. *J. Phys. Chem. A* **2019**, *123*, 7087–7103.

(27) Bader, R. F. W.; Slee, T. S.; Cremer, D.; Kraka, E. Description of Conjugation and Hyperconjugation in terms of Electron Distributions. *J. Am. Chem. Soc.* **1983**, *105*, 5061–5068.

(28) Cremer, D.; Kraka, E.; Slee, T. S.; Bader, R. F. W.; Lau, C. D. H.; Dang, T. T. N.; MacDougall, P. J. Description of Homoaromaticity in Terms of Electron Distributions. *J. Am. Chem. Soc.* **1983**, *105*, 5069–5075.

(29) Altona, C.; Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.

(30) Geise, H. J.; Adams, W. J.; Bartell, L. S. Electron Diffraction Study of Gaseous Tetrahydrofuran. *Tetrahedron* **1969**, *25*, 3045–3052.

(31) McQuarrie, D. A. *Statistical Thermodynamics*; Harper & Row New York, 1973.

(32) Luger, P.; Buschmann, J. Twist Conformation of Tetrahydrofuran in the Crystal. *Angew. Chem. Int. Ed.* **1983**, *22*, 410–410.

(33) David, W. I. F.; Ibberson, R. M.; Dennis, T. J. S.; Hare, J. P.; Prassides, K. Structural Phase Transitions in the Fullerene C_{60}. *Europhys. Lett.* **1992**, *18*, 219–225.

(34) Han, S. J.; Kang, Y. K. Pseudorotation in Heterocyclic Five-membered Rings: Tetrahydrofuran and Pyrrolidine. *J. Mol. Struct.* **1996**, *369*, 157–165.

(35) Lambert, J. B.; Papay, J. J.; Khan, S. A.; Kappauf, K. A.; Magyar, E. S. Conformational Analysis of Five-membered Rings. *J. Am. Chem. Soc.* **1974**, *96*, 6112–6118.

(36) Kalinowski, H. O.; Berger, S.; Braun, S. ^{13}C-NMR-Spektroskopie [13 hoch C-NMR-Spektroskopie]: 200 Tabellen; Thieme, 1984.

(37) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations : The Theory of Infrared and Raman Vibrational Spectra*; McGraw-Hill: New York, NY, USA, 1955.

(38) Woodward, L. A. *Introduction to the Theory of Molecular Vibrations and Vibrational Spectroscopy*; Oxford University Press: Oxford, United Kingdom, 1972.

(39) Herzberg, G. *Molecular Spectra and Molecular Structure. Volume II: Infrared and Raman Spectra of Polyatomic Molecules*; Krieger Publishing Co.: New York, NY, USA, 1991.

(40) Herzberg, G.; Huber, K. P. *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*; Springer: New York, NY, USA, 1979.

(41) Zou, W.; Tao, Y.; Cremer, D.; Kraka, E. Systematic Description of Molecular Deformations with Cremer-Pople Puckering and Deformation Coordinates Utilizing Analytic Derivatives: Applied to Cycloheptane, Cyclooctane, and Cyclo[18]carbon. *J. Chem. Phys.* **2020**, *152*, 154107–1–54107–15.

(42) Quapp, W.; Kraka, E.; Cremer, D. Finding the Transition State of Quasi-Barrierless Reactions by a Growing String Method for Newton Trajectories: Application to the Dissociation of Methylenecyclopropene and Cyclopropane. *J. Chem. Phys. A* **2007**, *111*, 11287–11293.

(43) Joo, H.; Kraka, E.; Quapp, W.; Cremer, D. The Mechanism of a Barrierless Reaction: Hidden Transition State and Hidden Intermediates in the Reaction of Methylene with Ethene. *Mol. Phys.* **2007**, *105*, 2697–2717.

(44) Gräfenstein, J.; Cremer, D. Unusual Long-Range Spin-Spin Coupling in Fluorinated Polyenes: A Mechanistic Analysis. *J. Chem. Phys.* **2007**, *127*, 174704.

(45) Gräfenstein, J.; Cremer, D. Efficient Density–Functional Theory Integrations by Locally Augmented Radial Grids. *J. Chem. Phys.* **2007**, *127*, 164113.

(46) Zou, W.; Cremer, D. C_2 in a Box: Determining its Intrinsic Bond Strength for the $X^1\Sigma^+_g$ Ground State. *Chem. Eur. J.* **2016**, *22*, 4087–4097.

(47) Kalescky, R.; Kraka, E.; Cremer, D. Local Vibrational Modes of the Formic Acid Dimer - The Strength of the Double H-Bond. *Mol. Phys.* **2013**, *111*, 1497–1510.

(48) Kalescky, R.; Kraka, E.; Cremer, D. New Approach to Tolman's Electronic Parameter Based on Local Vibrational Modes. *Inorg. Chem.* **2013**, *53*, 478–495.

(49) Setiawan, D.; Kraka, E.; Cremer, D. Description of Pnicogen Bonding with the help of Vibrational Spectroscopy-The Missing Link Between Theory and Experiment. *Chem. Phys. Letters* **2014**, *614*, 136–142.

(50) Kalescky, R.; Kraka, E.; Cremer, D. Identification of the Strongest Bonds in Chemistry. *J. Phys. Chem. A* **2013**, *117*, 8981–8995.

(51) Setiawan, D.; Sethio, D.; Cremer, D.; Kraka, E. From Strong to Weak NF Bonds: On the Design of a New Class of Fluorinating Agents. *Phys. Chem. Chem. Phys.* **2018**, *20*, 23913–23927.

(52) Humason, A.; Zou, W.; Cremer, D. 11,11-Dimethyl-1,6-methano[10]annulene-An Annulene with an Ultralong CC Bond or a Fluxional Molecule? *J Phys Chem A.* **2014**, *119*, 1666–1682.

(53) Kalescky, R.; Kraka, E.; Cremer, D. Are Carbon-Halogen Double and Triple Bonds Possible? *Int. J. Quant. Chem.* **2014**, *114*, 1060–1072.

(54) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Quantitative Assessment of the Multiplicity of Carbon-Halogen Bonds: Carbenium and Halonium Ions with F, Cl, Br, and I. *J. Phys. Chem. A* **2014**, *118*, 1948–1963.

(55) Kraka, E.; Larsson, J. A.; Cremer, D. In *Computational Spectroscopy*; Grunenberg, J., Ed.; Wiley: New York, 2010; pp 105–149.

(56) Larsson, J. A.; Cremer, D. Theoretical Verification and Extension of the Mckean Relationship between Bond Lengths and Stretching Frequencies. *J. Mol. Struct.* **1999**, *485-486*, 385–407.

(57) Cremer, D.; Kraka, E. From Molecular Vibrations to Bonding, Chemical Reactions, and Reaction Mechanism. *Curr. Org. Chem.* **2010**, *14*, 1524–1560.

(58) Cremer, D.; Wu, A.; Larsson, J. A.; Kraka, E. Some Thoughts about Bond Energies, Bond Lengths, and Force Constants. *J. Mol. Model.* **2000**, *6*, 396–412.

(59) Kraka, E.; Cremer, D. Weaker Bonds with Shorter Bond Lengths. *Rev. Proc. Quim.* **2012**, 39–42.

(60) Setiawan, D.; Kraka, E.; Cremer, D. Hidden Bond Anomalies: The Peculiar Case of the Fluorinated Amine Chalcogenides. *J. Phys. Chem. A* **2015**, *119*, 9541–9556.

(61) Kraka, E.; Setiawan, D.; Cremer, D. Re-Evaluation of the Bond Length-Bond Strength Rule: The Stronger Bond Is not Always the Shorter Bond. *J. Comp. Chem.* **2015**, *37*, 130–142.

(62) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Local Vibrational Modes of the Water Dimer - Comparison of Theory and Experiment. *Chem. Phys. Letters* **2012**, *554*, 243–247.

(63) Freindorf, M.; Kraka, E.; Cremer, D. A Comprehensive Analysis of Hydrogen Bond Interactions Based on Local Vibrational Modes. *Int. J. Quant. Chem.* **2012**, *112*, 3174–3187.

(64) Kalescky, R.; Zou, W.; Kraka, E.; Cremer, D. Vibrational Properties of the Isotopomers of the Water Dimer Derived from Experiment and Computations. *Aust. J. Chem.* **2014**, *67*, 426.

(65) Tao, Y.; Zou, W.; Jia, J.; Li, W.; Cremer, D. Different Ways of Hydrogen Bonding in Water - Why Does Warm Water Freeze Faster than Cold Water? *J. Chem. Theory Comput.* **2017**, *13*, 55–76.

(66) Tao, Y.; Zou, W.; Kraka, E. Strengthening of Hydrogen Bonding With the Push-Pull Effect. *Chem. Phys. Lett.* **2017**, *685*, 251–258.

(67) Freindorf, M.; Tao, Y.; Sethio, D.; Cremer, D.; Kraka, E. New Mechanistic Insights into the Claisen Rearrangement of Chorismate - A Unified Reaction Valley Approach Study. *Mol. Phys.* **2018**, *117*, 1172–1192.

(68) Makoś, M. Z.; Freindorf, M.; Sethio, D.; Kraka, E. New Insights into Fe–H$_2$ and Fe–H$^-$ Bonding of a [NiFe] Hydrogenase Mimic – A Local Vibrational Mode Study. *Theor. Chem. Acc.* **2019**, *138*, 76.

(69) Oliveira, V.; Kraka, E.; Cremer, D. The Intrinsic Strength of the Halogen Bond: Electrostatic and Covalent Contributions Described by Coupled Cluster Theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 33031–33046.

(70) Oliveira, V.; Kraka, E.; Cremer, D. Quantitative Assessment of Halogen Bonding Utilizing Vibrational Spectroscopy. *Inorg. Chem.* **2016**, *56*, 488–502.

(71) Oliveira, V.; Cremer, D. Transition from Metal-Ligand Bonding to Halogen Bonding Involving a Metal as Halogen Acceptor: A Study of Cu, Ag, Au, Pt, and Hg Complexes. *Chem. Phys. Letters* **2017**, *681*, 56–63.

(72) Yannacone, S.; Oliveira, V.; Verma, N.; Kraka, E. A Continuum from Halogen Bonds to Covalent Bonds: Where Do λ^3 Iodanes Fit? *Inorganics* **2019**, *7*, 47.

(73) Oliveira, V. P.; Kraka, E.; Machado, F. B. C. Pushing 3c-4e Bonds to the Limit: A Coupled Cluster Study of Stepwise Fluorination of First-Row Atoms. *Inorg. Chem.* **2019**, *58*, 14777–14789.

(74) Oliveira, V. P.; Marcial, B. L.; Machado, F. B. C.; Kraka, E. Metal-Halogen Bonding Seen through the Eyes of Vibrational Spectroscopy. *Materials* **2020**, *13*, 55.

(75) Setiawan, D.; Kraka, E.; Cremer, D. Strength of the Pnicogen Bond in Complexes Involving Group VA Elements N, P, and As. *J. Phys. Chem. A* **2014**, *119*, 1642–1656.

(76) Setiawan, D.; Cremer, D. Super-Pnicogen Bonding in the Radical Anion of the Fluorophosphine Dimer. *Chem. Phys. Letters* **2016**, *662*, 182–187.

(77) Oliveira, V.; Cremer, D.; Kraka, E. The Many Facets of Chalcogen Bonding: Described by Vibrational Spectroscopy. *J. Phys. Chem. A* **2017**, *121*, 6845–6862.

(78) Oliveira, V.; Kraka, E. Systematic Coupled Cluster Study of Noncovalent Interactions Involving Halogens, Chalcogens, and Pnicogens. *J. Phys. Chem. A* **2017**, *121*, 9544–9556.

(79) Sethio, D.; Oliveira, V.; Kraka, E. Quantitative Assessment of Tetrel Bonding Utilizing Vibrational Spectroscopy. *Molecules* **2018**, *23*, 2763.

(80) Zhang, X.; Dai, H.; Yan, H.; Zou, W.; Cremer, D. B-H π Interaction: A New Type of Nonclassical Hydrogen Bonding. *J. Am. Chem. Soc.* **2016**, *138*, 4334–4337.

(81) Zou, W.; Zhang, X.; Dai, H.; Yan, H.; Cremer, D.; Kraka, E. Description of an Unusual Hydrogen Bond Between Carborane and a Phenyl Group. *J. Organometal. Chem.* **2018**, *856*, 114–127.

(82) Badger, R. M. A Relation Between Internuclear Distances and Bond Force Constants. *J. Chem. Phys.* **1934**, *2*, 128–131.

(83) Bader, R. F. W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18*, 9–15.

(84) Bader, R. *Atoms in Molecules: A Quantum Theory*; International series of monographs on chemistry; Oxford:Clarendon Press, 1990.

(85) Cremer, D.; Kraka, E. Chemical Bonds without Bonding Electron Density? Does the Difference Electron-Density Analysis Suffice for a Description of the Chemical Bond? *Angew. Chem. Int. Ed.* **1984**, *23*, 627–628.

(86) Cremer, D.; Kraka, E. A Description of the Chemical Bond in Terms of Local Properties of Electron Density and Energy. *Croatica Chem. Acta* **1984**, *57*, 1259–1281.

(87) Kraka, E.; Cremer, D. *Theoretical Models of Chemical Bonding. The Concept of the Chemical Bond*; Z.B. Maksic, ed., Springer Verlag, Heidelberg, 1990; Vol. 2; p 453.

(88) Chai, J. D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(89) Chai, J. D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

(90) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(91) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.

(92) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li–F. *J. Comput. Chem.* **1983**, *4*, 294–301.

(93) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-Consistent Molecular Orbital Methods 25. Supplementary Functions for Gaussian Basis Sets. *J. Chem. Phys.* **1984**, *80*, 3265–3269.

(94) Gräfenstein, J.; Cremer, D. Efficient Density-Functional Theory Integrations by Locally Augmented Radial Grids. *J. Chem. Phys.* **2007**, *127*, 164113.

(95) others,, et al. GAUSSIAN 09 (Revision B. 01), Gaussian, Inc., Wallingford, CT, Gaussian. 2010.

(96) Kraka, E.; Zou, W.; Filatov, M.; Gräfenstein, J.; Izotov, D.; Gauss, J.; He, Y.; Wu, A.; Konkoli, Z.; Polo, V.; Olsson, L.; He, Z.; Cremer, D. COLOGNE. 2019; see `http://www.smu.edu/catco`.

(97) Cremer, D. *Modelling of Structure and Properties of Molecules*; edt. by Z.B. Maksic, Ellis Horwood, Chichester, 1987; p 125.