Jiří Šponer[1–3]
Jerzy Leszczynski[2]
Pavel Hobza[1]
[1] *J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic and Center for Complex Molecular Systems and Biomolecules, Dolejškova 3, 182 23 Prague, Czech Republic*

[2] *Department of Chemistry and Computational Center for Molecular Structure and Interactions, Jackson State University, Jackson, 39217 MS, USA*

Electronic Properties, Hydrogen Bonding, Stacking, and Cation Binding of DNA and RNA Bases

[3] *Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic*

Received 14 May 2001; accepted 24 August 2001

Abstract: *This review summarizes results concerning molecular interactions of nucleic acid bases as revealed by advanced ab initio quantum chemical (QM) calculations published in last few years. We first explain advantages and limitations of modern QM calculations of nucleobases and provide a brief history of this still rather new field. Then we provide an overview of key electronic properties of standard and selected modified nucleobases, such as their charge distributions, dipole moments, polarizabilities, proton affinities, tautomeric equilibria, and amino group hybridization. Then we continue with hydrogen bonding of nucleobases, by analyzing energetics of standard base pairs, mismatched base pairs, thio-base pairs, and others. After this, the nature of aromatic stacking interactions is explained. Also, nonclassical interactions in nucleic acids such as interstrand bifurcated hydrogen bonds, interstrand close amino group contacts, C—H · · · O interbase contacts, sugar–base stacking, intrinsically nonplanar base pairs, out-of-plane hydrogen bonds, and amino–acceptor interactions are commented on. Finally, we overview recent calculations on interactions between nucleic acid bases and metal cations. These studies deal with effects of cation binding on the strength of base pairs, analysis of specific differences among cations, such as the difference between zinc and magnesium, the influence of metalation on protonation and tautomeric equlibria of bases, and cation–π interactions involving nucleobases. In this review, we do not provide methodological details, as these can be found in our preceding reviews. The interrelation between*

Correspondence to: Jiří Šponer, J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague, Czech Republic; email: sponer@indy.jh–inst.cas.cz

Contract grant sponsor: Center for Complex Molecular Systems and Biomolecules, Ministry of Education of the Czech Republic (CCMSB); National Science Foundation (NSF); and National Institute of Health (NIH)

Contract grant number: LN00A032 (CCMSB), NSF CREST 9805465 and 9706268, and NIH RCMI G1 2RR13459–21

Biopolymers (Nucleic Acid Sciences), Vol. 61, 3–31 (2002)

advanced QM approaches and classical molecular dynamics simulations is briefly discussed. Biopoly (Nucleic Acid Sci) 61: 3–31, 2002; DOI 10.1002/bip.10048

Keywords: *nucleic acid bases; ab initio quantum chemical calculations; RNA; DNA*

INTRODUCTION

The structure and dynamics of nucleic acid molecules are influenced by a variety of contributions. Among them, interactions involving nucleic acid bases are of particular importance. Nucleobases are involved in two qualitatively different mutual interaction types: hydrogen bonding and aromatic base stacking. Further, nucleic acid bases interact with solvent molecules, metal cations, drugs, and other species.

In this review, we summarize recent results obtained by quantum-chemical (QM) studies of base stacking, base pairing, and cation binding of nucleic acid bases published in the period of 1994–2001. We mainly introduce selected results relevant to structural biology while detailed information about the methodologies and other rather technical aspects can be found in recent reviews.[1–4]

First, let us briefly comment on the history, purpose, advantages, and limitations of QM approaches used in studies of interactions of nucleobases.

Advance of High-Level Ab Initio Calculations of Nucleic Acid Bases and Base Pairs

QM studies of DNA bases have been attempted for more than 30 years. However, *before advance of powerful supercomputers in the beginning of the 1990s, no reliable calculations on medium-sized molecular clusters (such as base pairs) were possible.* Thus the old results were necessarily highly inaccurate, mutually contradicting, and method dependent. There has been a dramatic revival of QM studies in the 1990s and the presently available data are very reliable, often being close to ultimate predictions. The tremendous increase in computer power gives the researchers in this field a great advantage over all their predecessors. *Modern high-level ab initio calculations provide data of great accuracy and reliability, which for nucleobase interactions cannot be presently obtained by any other experimental or computational technique*.

Advantages of the Ab Initio Studies of Nucleobase Interactions

The major advantage of ab initio (first principle) QM calculations is absence of any empirical parameters. In contrast, all empirical potential and semiempirical QM methods heavily rely on parameterizations. Semiempirical QM approaches were often used in older studies of base–base interactions. However, even the most recent methods of semiempirical nature are not suitable to study molecular interactions.[4,5]

Accuracy of ab initio calculations is determined primarily by two factors: the size of the basis set of atomic orbitals, which is used to construct the molecular orbitals, and the inclusion of electron correlation effects. *Improving both factors in a balanced way leads to a systematic improvement of the results while solid evaluations of the error margins of the treatment are possible even before reaching convergence. This is a unique feature of high-level ab initio calculations, sharply distinguishing them from all parameterization-based techniques. Obviously, even the best QM calculations are influenced by approximations. Nevertheless, quality of high-level ab initio QM methods is comparable to accurate physicochemical experiments. The advanced QM calculations can be applied for systems where no relevant experiments exist.*

Modern electronic structure calculations of base pairing and stacking rely *on highly verified common methodology, which has provided reliable data for thousands of other systems.* There obviously are some basic rules to be followed in order to obtain meaningful results. These rules nevertheless are rather simple and essentially identical for all studies of closed-shell molecular clusters. *Thus, in fact, ab initio QM calculations are quite transparent and simple to understand, even for nonspecialists.* Ab initio calculations are much more transparent than any force-field methods where one always has to consider how the parametrization is conducted. We wish to underline that this comment concerns only studies utilizing standard high-level ab initio methods. We do not provide any guarantee regarding results of QM studies based on low quality QM methods, semiempirical studies, simplified and sometimes obscure approaches. Such calculations can, unfortunately, sometimes be found in the literature and cause confusions. QM methods are represented by a variety of approaches, ranging from exceptionally accurate treatments up to methods of literally no value.

The basic method used in QM calculations is the Hartree–Fock (HF) approximation, which solves the time-independent Schrödinger equation by assuming

that an electron moves in an averaged field of the other electrons. (Equivalent abbreviation is SCF, self-consistent field.) More sophisticated electron correlation "post-HF" methods consider explicitly that electron motions are correlated. The electron correlation effects are very important. For example, the London dispersion attraction is exclusively a consequence of intermolecular correlation of electrons (induced dipole–induced dipole interactions in the classical sense). Inclusion of electron correlation is also important to obtain accurate values of charge distributions and molecular dipole moments, as the HF method overestimates the dipole moments of DNA bases. Therefore, accurate studies of nucleobase interactions require post-HF methods.

The cheapest electron correlation method is the second-order Møller–Plesset perturbational method (MP2) considering single and double electron excitations to the second order of the perturbational approach. This method mostly provides (with a reasonable basis set of atomic orbitals) very reliable results. Basically all results presented in this review are based on a combination of HF and MP2 treatments. The third method to be noted is the coupled-cluster method with noniterative inclusion of triple electron excitations, abbreviated as CCSD(T). This method is often considered to be of a spectroscopic accuracy, but is prohibitively costly. It has been in few cases used for base pairing studies to verify the MP2 data. Further methodological details can be found elsewhere.[4]

Base pairs are closed shell systems and can be safely studied with a high accuracy using standard single determinant QM techniques. First pioneering ab initio studies on base pairing were published in the 1980s; however, these were still done with small basis sets and mostly without any consideration of electron correlation effects.[6–12] The second generation of ab initio calculations on base pairs published since 1994 is based on utilization of polarized basis sets of atomic orbitals and the MP2 method, verified by higher quality reference calculations for a representative set of relevant systems. These calculations provide correct picture of nucleobase interactions which, despite some refinements, is not going to be substantially changed by any future advances of quantum chemistry and computer equipment.[1–4]

In recent years, we have noticed an expansion of an entirely new class of QM methods, so-called density functional theory, DFT. DFT methods are considerably more economical compared with the MP2 method while they also include electron correlation effects. The best DFT approaches, such as Becke3LYP, can for some systems provide results superior to the MP2 approach. However, in contrast to conventional ab initio techniques, DFT methods utilize adjustable parameters. Unfortunately, DFT methods are unable to capture the London dispersion attraction. It means that for all molecular complexes with important dispersion contribution (such as base stacking) DFT techniques fail.[13,14] Thus, for base pairing, conventional QM ab initio methods are to be preferred. Thorough discussion of the applicability of DFT approaches for molecular clusters can be found in the literature.[4,13]

Energetics of Molecular Interactions—The Main Task of QM Calculations of Base Pairing and Stacking

The ab initio technique can be used to determine optimal structures of molecular clusters and to calculate energies for any single geometry of the cluster. QM calculations provide molecular wave functions, which can be used to derive physicochemical properties, such as vibrational spectra, dipole and higher multipole moments, polarizabilities, and proton affinities. *Nevertheless, the main achievement of QM calculations was description of the nature and energetics of nucleobase interactions.* This is because evaluation of the energetics of molecular interactions is quite difficult for experimental approaches while energetics is crucial for molecular structures.

Interaction Energy

Interaction energy of two bases A and B in a given geometrical arrangement, ΔE^{AB}, is the energy difference between the total electronic energy of the dimer E^{AB} and the electronic energies, E^{A} and E^{B}, of isolated bases.

$$\Delta E^{AB} = E^{AB} - E^{A} - E^{B} \tag{1}$$

In order to understand the physical origin of the interactions and to verify/parametrize force fields, we need to know the interaction energies over a substantial portion of the conformational space of the studied assembly with sufficient accuracy. Quantum chemistry can provide this data. It is rather easy to calculate interaction enthalpy ΔH by adding the change of zero-point energy to ΔE.[4] Formation of a base pair is, however, not driven by the interaction energy but by the change of free energy, ΔG, determining association constants. Thus, ΔG and not ΔE can be determined experimentally. ΔG is calculated as sum of enthalpy and entropy terms, that is, $\Delta G = \Delta H - T\Delta S$ (see below). Having a correct interaction energy func-

tion, it is possible to include free energies via computer experiments such as molecular dynamics.

The interaction energy ΔE can be decomposed into two terms: The Hartree–Fock interaction energy and the electron correlation interaction energy:

$$\Delta E = \Delta E^{\mathrm{HF}} + \Delta E^{\mathrm{cor}} \quad (2)$$

The HF component ΔE^{HF} contains the short-range exchange repulsion, electrostatic, induction, and part of the charge-transfer contributions. However, as explained in detail elsewhere, no unambiguous decomposition of the ΔE^{HF} values to obtain the individual terms is possible.[4] The electron correlation component ΔE^{cor} brings the dispersion attraction and corrections to the remaining contributions. The sum of the short-range exchange repulsion and dispersion attraction roughly corresponds to the van der Waals term of empirical force fields (neglecting contributions such as anisotropy of atoms, higher terms in the dispersion expansion, induction, and others).

The Gas Phase Nature of QM Calculations: Systems in a Complete Isolation

The calculations presented in this review have been carried out for systems in a complete isolation, i.e., assuming the gas phase condition. Gas phase calculations reveal the *intrinsic* interactions in the studied systems with no perturbation by external effects such as solvent, crystal packing and others. At first glance, the gas phase calculations may appear to be rather far fetched from physiological systems. However, one needs to correctly understand the *intrinsic* molecular interactions, as this knowledge constitutes an important prerequisite towards understanding the role of nucleobases in DNA structure.

QM Calculations and Experiments

The natural counterpart of QM calculations of base pairing is a gas phase experiment carried out at very low temperature. Many years ago mass field spectroscopy data provided gas phase enthalpies of selected base pairs,[15] being in a very good agreement with modern QM data.[4] The experiments did not provide any evidence about the structures of the studied species but it was believed that H-bonded base pairs were studied. Recent gas phase molecular dynamics (MD) simulation studies suggest that the gas phase experiments in fact deal with a mixture of simultaneously populated structures, involving not only the expected H-bonded dimers, but also T-shaped and stacked structures.[16–20] The calculations clearly demonstrate that potential energy and free energy surfaces of isolated base pairs differ considerably. On the potential energy surface, H-bonded structures are in general more stable compared to stacking arrangements. However, entropy contribution favors stacked arrangements. The balance between energy and free energy is quite delicate, and depends on composition of the base pairs. It can be substantially modulated by methylations blocking some often very stable H-bonded base pairs and enhancing population of dispersion-stabilized stacked structures. Thus, the experiments can often deal with stacked dimers, which in some cases could even be dominating.[16–20]

Another attempt to measure the intrinsic energetics of base pairs in the gas phase[21] failed because of the lack of thermodynamic equilibrium in the molecular beam expansion.[22] Several groups are currently advancing in experimental gas phase studies of nucleobase pairing.[23–25] These experiments will provide excellent data regarding mainly the spectroscopy of base pairs (including detection of rare tautomers) but the available methods do not appear to be designed to reveal the energetics of base pairing. Gas phase experimental studies have also been reported on interactions between nucleobases and selected monovalent metal cations.[26,27] There have been studies in argon matrices regarding the balance of stacking and H-bonding for model systems.[28,29]

Evidently, we do not want to neglect very valuable experimental studies in a condensed phase or in crystals showing nucleobases in relevant environments.[30–39] However, these experiments do not reveal the *intrinsic* interactions. Different condensed phase experimental studies could even capture different aspects of molecular interactions as the actual expression of the intrinsic molecular interactions depends on the environment and the way how the experiment is conducted.[40] For example, stacking of two protonated cytosines is a highly repulsive interaction in the gas phase due to a charge–charge repulsion.[41,42] However, in the condensed phase, stacking of protonated cytosines can be associated with both structure stabilization and destabilization. In the i-DNA quadruplex,[42–44] stacking of a number of consecutive closely spaced protonated cytosines occurs. Thus i-DNA has, in contrast to all other DNA forms, *repulsive intrinsic stacking energy terms*.[41,42] Clearly, one cannot predict i-DNA stability based solely on the gas phase data. At the same time, one cannot evaluate the *intrinsic* stacking energetics based on studies of i-DNA. Initially, it has been postulated that the intrinsic i-DNA stacking is stabilized by some unusual dipole–dipole molecular interactions involving base exocyclic groups.[43,44] Then QM calculations clearly re-

vealed a substantial stacking repulsion in the i-DNA stacking geometry and absence of the major hypothetical exocyclic group attraction.[41] We wish to underline this explicitly as the incorrect idea about exocyclic group attraction in i-DNA is still alive, leading to confusing interpretations of the sources of i-DNA stability. Solution of the i-DNA stacking paradox has been provided by large-scale explicit solvent MD simulations of $d(CCCC)_4$ i-motif. The simulations proved that the vertical repulsion between consecutive protonated base pairs in i-DNA is counterbalanced by solvent and backbone screening effects.[42] This screening, however, is specific for i-DNA. In a striking contrast to i-DNA consecutive protonated cytosines sharply destabilize DNA triplex, because in triplex the screening of the electrostatic repulsion by the backbone and environment is different than in i-DNA.[42,45] Thus in this particular case one cannot extrapolate numbers and experience concerning nucleobase interactions between two DNA forms and each case should be studied separately. Very instructive computational analysis has been recently published, dealing with difficulties in unambiguous interpretation of condensed phase experimental data on nucleobase stacking and explaining the origin of some conflicting opinions about base-stacking interactions.[46] The evident complexity of condensed phase results underlines the importance of having an accurate gas phase picture of the intrinsic molecular interactions.[40,46]

Inclusion of solvent effects into QM calculations is very difficult. One option is to extend the studied system by explicit water molecules. Such calculations can be done quite accurately but they still deal with a gas phase molecular cluster and not with a bulk solvent. In addition, a substantial problem of this approach is that the potential energy surface contains a large number of minima, and without an efficient sampling technique it is virtually impossible to verify the true global minimum. Interestingly, recent gas phase MD studies have demonstrated that inclusion of very few water molecules (two to six) into the cluster is sufficient to entirely convert the H-bonded base pairs into stacked structures.[20] This underlines the importance of microsolvation. The other option is to include the solvent as a polarizable continuum. There exist a number of such QM methods, and these include effects of the continuum on the electronic structure of the solute molecules, in contrast to classical continuum approaches.[46–48] Unfortunately, these techniques are, in contrast to QM gas phase methods, based on substantial approximations, making them considerably less accurate.[49–51] The outcomes are quite sensitive with respect to the choice of parameters such as the atomic radii used to define the "solute" cavity, etc. Nevertheless, several noteworthy QM studies of base stacking and H-bonding of bases in water emerged recently while solvent studies remain a challenge for future.[52–55]

Brief History of High-Level Ab Initio Studies of Base Stacking, H-Bonding, and Cation Binding

First geometry optimization of isolated cytosine with inclusion of electron correlation effects has been published in 1992. The optimization assumed C_s symmetry, while the harmonic vibrational analysis revealed that the base is nonplanar in its amino group.[56] Unconstrained optimizations of all standard nucleobases in 1994 have shown surprisingly nonplanar geometries of amino groups of DNA bases, adopting a partial sp^3 pyramidal hybridization.[57,58]

First base-pairing energies calculated with inclusion of electron correlation effects become available in the period of 1994–1996,[1,41,59–62] while corresponding studies of base stacking were published in the period of 1995–1997.[1,14,41,60,63–65] These studies ruled out several older concepts of aromatic base stacking, including so-called induction and sandwich "π–π" models. *It is one of the most important results of QM studies to convincingly clarify that base stacking does not show any unusual properties that would set it apart from other molecular interactions and can be well described by utilizing the simplest form of all-atom empirical force fields.* The ab initio data for base stacking and H-bonding of bases cumulated in the period 1995–1997 provide current reference values, with respect to which other cheaper computational tools (including major force fields) are to be compared or parameterized. Further studies analyzed base pairs containing thiobases,[65] protonated bases,[41] hydrophobic (nonpolar) bases pairs[66,67] H-bonded trimers,[68] quartets,[69–71] stacking in base-pair steps,[55,64] hydration shells around base pairs,[72–76] water-mediated base pairs,[77,78] nucleobase-intercalator stacking,[79] stacking and H-bonding of base pairs in a polar solvent,[52,55] and others.

The quality of the base-pairing energies was verified in 2000.[80] The new reference calculations established the basis set limit and suggested that the base-pairing energies published in 1996[62] are underestimated by approximately 2–3 kcal/mol, due to a neglect of some fraction of the dispersion energy. Similar higher-level reevaluation of stacking energies will be completed within the next couple of years while reference results available so far show that there will be no qualitative changes of the present picture.[81,82]

Ab initio calculations were applied for analyses of bifurcated H-bonds in A-tracts and ApA B-DNA steps,[83] close amino group contacts in ApT and CpG B-DNA steps,[83] amino–acceptor interactions in biomolecules,[84,85] intrinsically nonplanar G · A mismatches,[86] cation–π interactions involving nucleobases,[87] sugar–base stacking,[64] metal-assisted tautomers of nucleobases,[88] C—H · · · O contacts in base pairs,[89] local conformational variability and sequence dependence of stacking in DNA crystals,[55,64] cooperativity of base stacking,[55,64] and others.

Ab initio calculations were further used to characterize proton-transfer processes in base pairs[90] and base-pair radical cations.[91–93] QM calculations are now routinely used in studies of various aspects of nucleobase H-bonding, and other valuable papers can be found in the literature.[93–110]

Advanced ab initio studies on metal–cation nucleobase interactions have been emerging since 1996.[26,27,68,70,71,87,88,104,111–119] The application of QM methods for metal–cation containing clusters is of primary importance, as these systems always show major nonelectrostatic contributions (induction and charge-transfer effects) completely neglected by conventional force fields. The calculations were initiated by evaluations of interaction between an isolated base and bare cations,[111] followed by studies of metal effects on base pairing,[68,112] inclusion of cation hydration shell,[113–115] and the sugar–phosphate backbone.[104,116] The calculations characterized effects of cation binding on the base-pairing strength and captured specific differences among cations, such as the difference between Zn^{2+} and Mg^{2+}.[111–115] The calculations advocate for considering nonelectrostatic effects as being one of the major factors causing diverse biological and biochemical roles of various cations.[111–116] Further attention has been paid to effects of metalation of bases on the tautomeric and protonation equilibria of nucleobases.[88,118]

ELECTRONIC STRUCTURE OF NUCLEOBASES

Molecular interactions of nucleobases are predetermined by their electronic structures. Therefore, we first discuss selected properties of isolated nucleobases.

Atomic Charges, Dipole Moments, and Polarizabilities

Table I provides atom-centered point charges, dipole moments, and polarizabilities of selected nucleobases. Note that in the N3-protonated cytosine the +1 charge is equally distributed over the whole molecule. When discussing charges we would like to stress that *due to basic principles of quantum mechanics, any atomic charge distribution is arbitrary. It is because there exists no quantum mechanical operator for atomic charges and thus there is no experiment (even a hypothetical one) to determine atomic charges. Partial atomic charges, even those derived from experimental electron densities, do not correspond to any real physicochemical quantity.* Thus charge distributions can be greatly manipulated by changing the way they are derived and there exist vast possibilities how to do define them. This should be considered when arguing about molecular interactions based on individual atomic charges, as popular in biological literature, especially in cases of weak molecular interactions such as C—H · · · X contacts. It makes no sense to compare individual atomic charges in different force fields. Here, the important issue is how the whole set of charges performs in interaction energy calculations.

The charges in Table I were derived by fitting to molecular electrostatic potential (MEP, ESP charges) around the monomers.[1,4,63–65,79,120] *This is one of the most reasonable ways to derive charge distributions, because the molecular electrostatic potential, dipole, and higher moments are uniquely defined, measurable quantities.* Thus the individual charges still remain arbitrary; nevertheless, the electrostatic field created by all of them together closely approximates the actual electrostatic field around the nucleobases.[120] The charges effectively include higher multipole moments, and one does not need to add point dipoles, quadrupoles, etc. ESP charge distributions are very useful for calculations of electrostatic interaction energies, and the popular AMBER force field uses this way to derive charges.[120] Disadvantage of the ESP approach is that for flexible molecules such as amino acids, the charges derived in one conformer do not reproduce the electrostatic potential in other conformers. This is one of the major problems of current molecular modeling. The most promising way to reduce this deficiency appears to be inclusion of an induction term into future generations of force fields. Nevertheless, for rigid nucleobases, ESP charges do an excellent job in calculations of molecular interactions.[5,63–65] Another noticeable way to derive charges is, for example, the Natural Bond Orbital method, NBO,[121] and there exist also other approaches, which under certain circumstances may be used to assess interactions and charge transfer. Probably the most rigorous way to derive net atomic populations is the Bader's "atoms in molecules" topological approach,[122] which provides very valuable insight into molecular interactions. *We strongly advocate against*

Table I Atomic Charges Evaluated by Fitting to Molecular Electrostatic Potential Obtained with the MP2 Ab Initio Method and Extended aug-cc-pVDZ Basis Set for Planar Nucleobases[a]

			Cytosine			
C2 +0.967	N1 −0.634	C6 +0.238	C5 −0.740	C4 +1.055	N3 −0.817	O2 −0.619
N4 −1.100	H1 +0.359	H6 +0.138	H5 +0.244	H41 +0.451	H42 +0.458	μ 6.39
$\alpha(\alpha_{zz})$ 103 (45)						
			Adenine			
N3 −0.715	C2 +0.465	N6 −0.694	C6 +0.603	C5 +0.128	C4 +0.591	N9 −0.568
C8 +0.243	N7 −0.578	A6 −0.897	H2 +0.071	H8 +0.129	H9 +0.402	H61 +0.421
H62 +0.399	μ 2.56	$\alpha(\alpha_{zz})$ 122 (53)				
			Guanine			
O6 −0.500	C6 +0.524	N1 −0.710	C2 +0.856	N3 −0.713	C4 +0.480	C5 +0.194
N7 −0.593	C8 +0.302	N9 −0.602	N2 −1.053	H1 +0.393	H8 +0.105	H9 +0.415
H21 +0.458	H22 +0.444	μ 6.55	$\alpha(\alpha_{zz})$ 119 (55)			
			Thymine			
O4 −0.531	C4 +0.683	N3 −0.618	C2 +0.744	N1 −0.520	C6 −0.043	C5 −0.041
O2 −0.559	C5M −0.474	H3 +0.361	H1 +0.364	H6 +0.186	HM1 +0.1575	HM2 +0.133
HM3 +0.1575	μ 4.31 D	$\alpha(\alpha_{zz})$ 112 (50)				
			Uracil			
O4 −0.555	C4 +0.817	N3 −0.609	C2 +0.749	N1 −0.505	C6 +0.133	C5 −0.554
O2 −0.554	H5 +0.222	H3 +0.349	H1 +0.351	H6 +0.156	μ 4.37	$\alpha(\alpha_{zz})$ 94 (41)
			6-Oxopurine			
O6 −0.514	C6 +0.601	N1 −0.644	C2 +0.334	N3 −0.639	C4 +0.550	C5 +0.126
N7 −0.548	C8 +0.251	N9 −0.549	H2 +0.123	H1 +0.383	H8 +0.126	H9 +0.400
μ 5.16	$\alpha(\alpha_{zz})$ 119 (50)					
			2-Amino adenine			
N6 −0.877	C6 +0.455	N1 −0.663	C2 +0.851	N3 −0.731	C4 +0.487	C5 +0.232
N7 −0.619	C8 +0.279	N9 −0.601	N2 −1.011	H8 +0.110	H9 +0.412	H61 +0.423
H62 +0.399	H21 +0.433	H22 +0.421	μ 0.91	$\alpha(\alpha_{zz})$ 131 (60)		
			6-Thioguanine			
S6 −0.272	C6 −0.129	N1 −0.295	C2 0.766	N3 −0.675	C4 0.480	C5 0.411
N7 −0.585	C8 0.241	N9 −0.587	N2 −1.081	H1 0.251	H8 0.130	H9 0.410
H21 0.465	H22 0.470	μ 7.28	$\alpha(\alpha_{zz})$ 169 (69)			
			Purine			
N3 −0.731	C2 0.518	N1 −0.658	C6 0.231	C5 0.106	C4 0.770	N9 −0.709
C8 0.361	N7 −0.605	H6 0.114	H2 0.061	H8 0.112	H9 0.430	
μ 3.75	$\alpha(\alpha_{zz})$ 115 (48)					
			2-Thiouracil			
O4 −0.517	C4 0.543	N3 0.016	C2 −0.145	N1 0.033	C6 −0.125	C5 −0.313
S2 −0.276	H5 0.184	H3 0.194	H1 0.218	H6 0.188	μ 4.58	$\alpha(\alpha_{zz})$ 95 (54)
			N3-protonated cytosine			
C2 0.699	N1 −0.455	C6 0.213	C5 −0.551	C4 0.816	N3 −0.532	O2 −0.430
N4 −0.933	H1 0.385	H6 0.188	H5 0.260	H41 0.482	H42 0.478	H3 0.380

[a] The table further contains dipole moments (μ, in D, MP2/aug-cc-pVDZ method) and polarizabilities (α, in a.u., values in parentheses show the vertical component of polarizability, α_{zz}, Becke3LYP/aug-cc-pVDZ method). The amino hydrogens are as follows: cytosine H41 is *cis* to C5, adenine H61 is *cis* to N1, guanine and thioguanine H21 is *cis* to N3, 2-amino adenine H21 is *cis* to N3, and H61 *cis* to N1. Thymine HM2 hydrogen is in the plane of the base and points to C6.

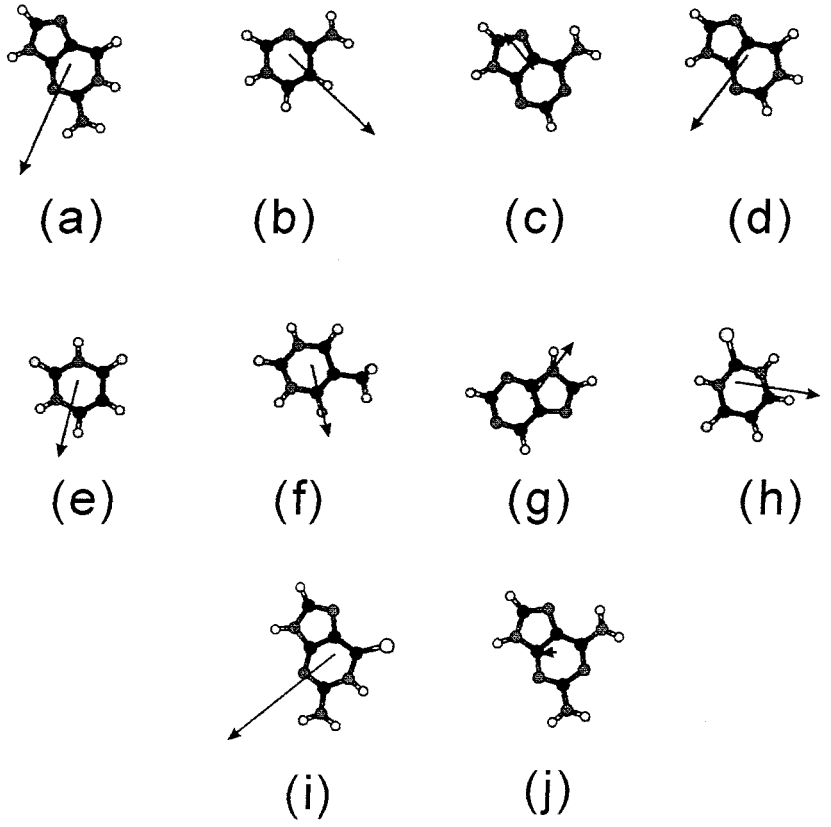

FIGURE 1 Molecular dipole moments of selected nucleobases. See Table I for the magnitude of dipoles and the charge distribution. (a) Guanine, (b) cytosine, (c) adenine, (d) 6-oxopurine (inosine), (e) uracil, (d) thymine, (g) purine, (e) 2-thiouracil, (i) 6-thioguanine, (j) 2-amino-adenine.

the popular Mulliken populational analysis. This atomic charge distribution is extremely method and basis set dependent[65] *and should be avoided, especially for metal–cation containing clusters. When investigating binding of hydrated metal cations to adenine, we observed that a minor adjustment of an exponent of a single inner s- atomic orbital on* Mg^{2+} *changed the calculated Mulliken charge on the cation by as much as 0.4 e.*

The dipole moments (μ) in the Table I have been calculated at the MP2/aug-cc-pVDZ level and can be considered as converged. Figure 1 shows orientation of the dipole moments. The most polar base is thioguanine with a dipole moment of 7.3D, closely followed by guanine and cytosine. 6-Oxopurine also has a large dipole moment with the same orientation as guanine. These polar bases thus can be involved in strong H-bond interactions. On the other hand, 2-amino adenine is a very apolar base and thus forms much weaker H-bonded base pairs than guanine. The polarity is determined by distribution of donor and acceptor functional groups around the base. The electrostatic part of the molecular interactions between nucleobases can be in the first approximation considered as a molecular dipole–molecular dipole interaction. This does not mean one can calculate the base pair strength as a product of the calculated *point* dipoles because the point dipole approximation is valid only when viewing bases from infinity. When referring to *molecular dipole* we mean the overall electrostatic potential distribution that mostly has a clear dipole component. Note we do not provide dipole moment of CH^+ because the dipole moment of non-neutral systems depends on the origin of the coordinate frame used for the calculations.

The last number in the Table I shows polarizabilities calculated using DFT Becke3LYP method with the aug-cc-pVDZ basis set of atomic orbitals. Note that the polarizability cannot be predicted based on additive atomic polarizabilities. Thus, thioguanine has a much larger polarizability compared to guanine, while the difference between thiouracil and uracil is much smaller.

Proton Affinities

Table II shows the most complete set of protonation energies of nucleobases published so far; never-

Table II Protonation Energies of Selected Nucleic Acid Bases Evaluated at the MP2/6-31G Level (kcal/mol)**

Cytosine	**P_{N3} −241.4**	**P_{O2} −241.9**[a]	P_{O2} −232.8[b]	P_{N4} −212.7	
Adenine	P_{N1} −234.8	**P_{N3} −237.1**	P_{N7} −228.6	P_{N6} −218.6	
Guanine	P_{N3} −223.5	**P_{N7} −239.8**	P_{O6} −233.5[c]	P_{O6} −224.3[d]	P_{N2} −205.3
Thymine	**P_{O4} −217.5**[b]	P_{O4} −214.9[a]	P_{O2} −211.1[b]	P_{O2} −209.8[a]	
Uracil	**P_{O4} −216.4**[b]	P_{O4} −213.4[a]	P_{O2} −208.0[b]	P_{O2} −206.6[a]	
6-Oxopurine	P_{N3} −217.7	**P_{N7} −232.8**	P_{O6} −228.3[c]	P_{O6} −220.1[b]	
2-Aminoadenine	P_{N1} −238.8	**P_{N3}** −239.2	P_{N7} −235.2	P_{N6} −223.3	P_{N2} −230.8
6-Thioguanine	P_{N3} −221.9	**P_{N7} −240.1**	P_{S6} −235.1[a]	P_{S6} 231.3[b]	P_{N2} −203.8
2-Thiouracil	**P_{O4} −215.5**[b]	P_{O4} −212.5[a]	P_{S2} 210.9[b]	P_{S2} 210.3[a]	
Purine	**P_{N1} −230.9**	P_{N3} −219.4	P_{N7} −225.2		

[a] *cis* with respect to N3.
[b] *trans* with respect the N3.
[c] *trans* with respect to N1.
[d] *cis* with respect to N1.

theless, we suggest also other related theoretical and experimental papers for reading.[123–132] Protonation energy is the electronic energy difference between protonated and unprotonated base, i.e., $P = E(BH^+) - E(B) - E(H^+)$. A long time ago Del Bene determined quite accurately the following order of protonation energies of standard bases: guanine > cytosine > adenine ⩾ uracil = thymine.[123] Table II shows the following order: cytosine > thioguanine = guanine = 2amino adenine > adenine ⩾ inosine > purine ⩾ uracils/thymine. The proton affinities of individual nucleobase sites are interrelated with the electrostatic potential of nucleobases and play a major role in determining the basicity of the sites, as defined by pKa values in condensed phase experiments. Obviously, protonation energies tell also a lot about preferable cation binding sites, though many cations show strong nonelectrostatic preferences for certain atoms. That is, zinc strongly prefers nitrogen sites over oxygen; other cations selectively bind to sulphur, etc.[113,115,116]

Table II presents the gas phase protonation energies. The gas phase protonation enthalpies could be obtained by adding the zero point energy (ZPE) correction calculated via harmonic vibrational analysis. This would change all numbers quite uniformly by ca. +8.5 kcal/mol at 0 K. The only exception is protonation of heavy sulfur atoms with the ZPE correction reduced to ca. +5.5 kcal/mol. Thermal corrections of protonation enthalpies and free energies do not change the results.

Tautomeric Equilibria of Nucleic Acid Bases

Besides protonation and deprotonation, nucleic acid bases can also undergo proton shifts while keeping their neutrality and form rare tautomers.[133–150] Rare tautomers may be involved in proton transfer processes, stabilize mispairs, promote point mutations, and play other roles. Nevertheless, direct and unambiguous evidence of the presence of tautomers in biomolecules is rare. Note that in structural studies reporting tautomers sometimes protonated species rather than tautomers are observed. Strictly speaking, tautomers should be neutral.[88]

The computational literature devoted to gas phase tautomerism of nucleobases is an order of magnitude more extended compared with studies devoted to base pairing and stacking. Here we summarize basic tautomeric properties of bases relevant to nucleic acids. The reader can found additional information in a recent review[133] and other papers.[134–150]

Cytosine

Figure 2 shows three low-energy gas phase tautomers of cytosine including the canonical amino-oxo form (C1). In matrix isolation infrared studies, both (C1) and (C2) have been detected, with likely a small amount of (C3).[141] Tautomer (C2) has no biological relevance as it involves deprotonation of the N1 position where sugar is attached. Structure (C3) in its N3-*anti* orientation (C3b) could be, on the other hand, very important, as it has identical donor and acceptor distribution as the Watson–Crick edge of guanine and protonated cytosine.[41,145] QM and MD calculations suggested that this imino tautomer of cytosine could (perhaps temporarily) occur for example in C.GC triplexes,[45,68,145] i-DNA,[41,42] and parallel-stranded DNA[74]; however, no experimental evidence is available.

Table III summarizes the relative gas phase energies (with respect to the canonical form) of all low-energy cytosine tautomers obtained by reference CCSD(T) calculations with a quite extended basis set

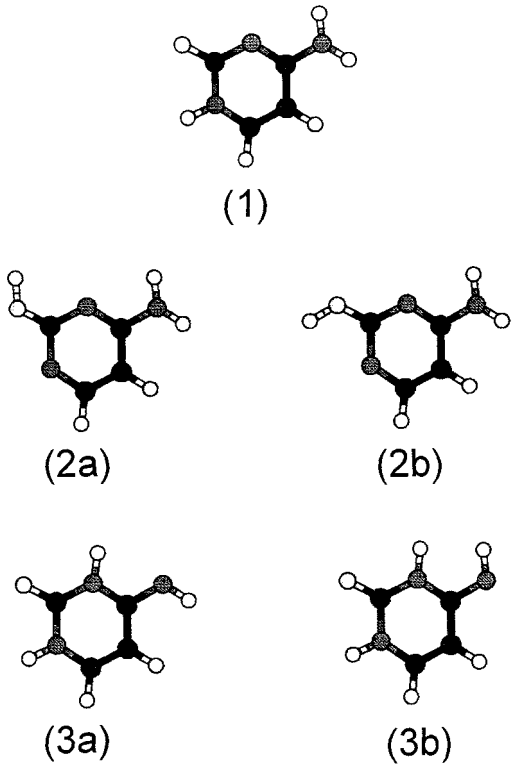

FIGURE 2 Low-energy tautomers of cytosine.

of atomic orbitals.[144] In agreement with experiments,[141] the calculations show that all species are going to coexist in the gas phase.

Solvation has a dramatic influence on the tautomerism of cytosine.[145,146] Polar solvent sharply stabilizes the canonical keto-amino form (C1) over the imino tautomer by as much as 7 kcal/mol, in agreement with experimental estimates of free energy of cytosine tautomerization.[145] This greatly reduces the chance to see this tautomer in biomolecules, which is not surprising, as living systems need to knock off rare tautomers in most situations.

Table III Relative Gas Phase Tautomeric Energies of Low-Energy Tautomers of Cytosine with Respect to the Canonical Amino-Oxo Form[a]

Tautomer (See Figure 3)	Relative Energy	Dipole Moment (D)
(C1)	0	6.2
(C2a)	−0.6	4.4
(C2b)	−1.2	3.2
(C3a)	+0.8	4.5
(C3b)	+2.1	2.3

[a] In (kcal/mol, negative values give tautomers more stable than the canonical form).[144]

Table IV Relative Gas Phase Energies of Low-Energy Tautomers of Guanine with Respect to the Canonical Amino-Oxo Form[a]

Tautomer (See Figure 3)	Relative Energy	Dipole Moment (D)
Gua H9 = oxo	0	6.3
Gua H7 = oxo	−0.7	1.8
Gua H7 = OH	+1.7	4.3
Gua H9 = OH[b]	+0.1	3.1
Gua H9 = OH[c]	+0.4	3.9

[a] MP2/6-311++G(2d,2p)//MP2/6-31G(d,p) method (kcal/mol, negative values give tautomers more stable than the canonical form).[140]

[b] OH *cis* with respect to N1.

[c] OH *trans* with respect to N1.

Condensed phase and X-ray bioinorganic experiments show that a structure formally identical to the cytosine imino tautomer (C3) is often induced by metalation of the cytosine amino group.[151–154] These species are known as metal-assisted imino tautomers. Quantum-chemical calculations revealed that the actual electronic structure and molecular interactions of these species are similar to N3-protonated cytosine rather than to the neutral imino tautomer.[88]

Guanine

Guanine can form several low-energy tautomeric forms in gas phase and to determine their equilibria accurately is difficult (Table IV and Figure 3).[133,138–140,145] The calculations usually show that the most stable gas phase tautomer of guanine is the amino-oxo form with hydrogen at N7 rather than at N9 (designated as Gua-H7-oxo).[133,140] Gua-H7 has no relevance to DNA but is going to be detected in the gas phase, unless the N9 position is methylated. The other low energy form is the canonical amino-oxo tautomer (Gua-H9-oxo). There are two other gas phase low energy tautomers where the H1 ring hydrogen moves to O6, i.e., amino-hydroxo tautomers (Gau-H9-OH and Gua-H7-OH). Some calculations even show these tautomers to be slightly more stable than the canonical form as the best available calculations are still subjected to uncertainty of the order of 1–2 kcal/mol.[140] Of course, the calculations can very safely discriminate between low- and high-energy tautomers. Upon solvation, the enol forms of guanine are largely destabilised leading to an exclusive occurrence of keto-amino species, as for cytosine.[145]

Interestingly, the gas phase tautomerism of 6-thioguanine is slightly different because, when considering enthalpies, the ZPE correction (see above) stabi-

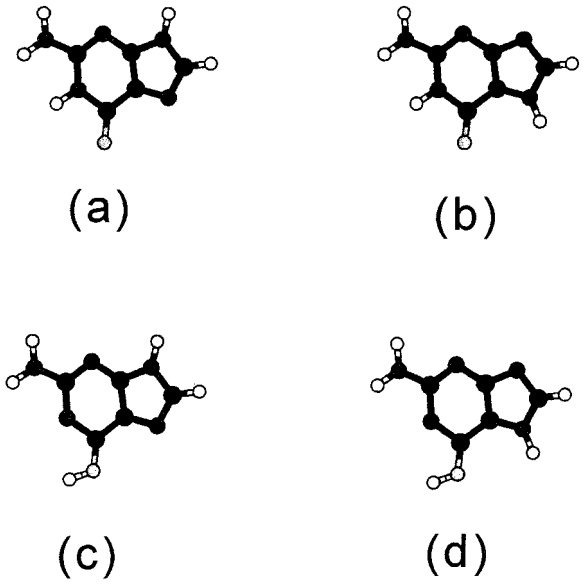

FIGURE 3 Low-energy tautomers of guanine. (a) Gua H9–oxo, (b) Gua H7–oxo, (c) Gua H9-OH, (d) Gua H7-OH.

lizes the thiol (S-H) tautomers compared to the thione ones by about 2.5 kcal/mol.[147–149] This still is not sufficient to stabilize the thiol forms in a polar solvent.[147]

Uracil and Thymine

Uracil and thymine do not posse any tautomer close in energy to the canonical base in the gas phase.[133–135] The second lowest energy tautomer of uracil, 2-hydroxo-4oxo form, is separated by ca. 11–12 kcal/mol from the canonical form, making this tautomer undetectable in gas phase experiments.[133–135] Tautomeric equlibria of thymine are close to identical. Inclusion of solvent effects does not change this picture.[150] The intrinsic tautomeric equilibria of uracil and thymine are not substantially changed by heavy atom substitution, i.e., when considering thio- and seleno-derivatives.[147–149] Advanced computational gas phase and polar solvent calculations suggest that 5-subtitutions of uracil such as bromination do not change the tautomerism.[150] The tautomers of 5-substituted uracils have been suggested to play interesting biological roles, but the electronic structure calculations provide a solid basis for ruling out the involvement of noncanonical enol tautomers as the origin of the mutagenic properties of 5-bromouridine.[150]

Adenine

For several years two forms of adenine were believed to coexist in the gas phase in comparable amounts. This belief was supported by the existence of two amino tautomers, N9(H) and N7(H) in polar solvents.[133] Note again that the N7(H) form has no biological relevance. Nevertheless, ab initio calculations revealed that the energy difference between these two forms is of the order of 7–8 kcal/mol, which has finally been confirmed by IR experiments.[136] This is not surprising considering the poor proton affinity of the N7 site of adenine (Table II). Adenine could also form an imino tautomer similar to cytosine, by shifting one amino hydrogen to the adenine N1 position. However, due to a different electronic structure of adenine, its imino tautomer is by 10–12 kcal/mol less stable compared to the canonical tautomer.[88,133,151–158] Thus, formation of this tautomer is unlikely. Nevertheless, similar to cytosine, such imino H-pattern is induced by direct metalation of the adenine amino group[88,155–158] leading to a protonation of N1 due to changes of the electronic structures of the adenine ring.[88] This could explain mutagenecity of some metals.

Pyramidalization of Amino Groups of Bases

Table V shows geometries of amino groups of isolated bases: dihedral angles between the amino group hydrogens and nucleobase rings and the energy difference between a fully optimised nonplanar and planar nucleobase. The MP2/6-311G(2df,p) data are considered to be largely converged.[1] All amino groups are pyramidal with a partial sp^3 hybridisation of their nitrogens (Figure 4).[1,4,57,58] This means both hydrogens deviate from the nucleobase plane in one direction, the nitrogen is slightly shifted in the opposite way and there exist a lone electron pair above the nitrogen. It resembles pyramidalization of aniline,[159] although the effect of sp^3 pyramidalization for amino groups of bases is smaller. The nonplanar hydrogens can be involved in out of plane H-bonds while the nitrogens can serve as (albeit rather weak) H-bond acceptors.[83–86] Our recent calculations suggest that a further pyramidalization may be associated with a charge transfer between the nucleobase ring and the negatively charged sugar phosphate segment.[160] Similarly, pyramidalization could be boosted by increasing the conjugation of systems which can be achieved by some intermolecular interactions when amino groups are not involved in H-bonding.[161] For the sake of completeness let us add that other parts of nucleobases are planar and the amino groups are planarized upon ring protonation.[41]

Table V Nonplanar Geometries of Isolated DNA Bases[a]

Base	Dihedral Angles		Sum of Amino Group Valence Angles	Inversion Barrier
Cytosine	C5C4N4H41 −21.4	N3C4N4H42 +12.6	351.9	−0.15
Adenine	C5C6N6H41 −15.3	N1C6N6H62 +16.5	352.9	−0.13
Guanine	N3C2N2H21 −13.3	N1C2N2N22 +39.2	339.6	−1.12
6-thioG	N3C2N2H21 −13.5	N1C2N2H22 +38.0	340.6	−0.98
2-aminoA	C5C6N6H61 +18.1	N1C6N6H62 −17.0	351.0	—
2-aminoA	N3C2N2H21 +22.2	N1C2N2H22 −22.1	345.0	−0.79[b]

[a] Current reference MP2/6-311G(2df,p) data. Inversion barrier (kcal/mol) is the energy difference between nonplanar and planar optimized structures. Sum of amino group valence angles is 360° for a planar amino group.

[b] Calculated for both amino groups.

Pyramidal-Rotated Geometries of Amino Groups

Intrinsic nonplanarity of the guanine amino group enveloped by two ring nitrogens is substantially larger compared to adenine and cytosine, where there is one carbon atom adjacent to the C-NH_2 group. Further, pyramidalization of guanine amino group is nonsymmetrical, i.e., one of the hydrogens is substantially more nonplanar (dihedral angle of ca 40°, Table V) than the other (ca. 10°).[58] It is because its repulsion with the adjacent polar (N1)-H1 ring hydrogen. Even the H1 ring hydrogen is bent by about 6° with respect to the guanine ring.[58] We call this amino group geometry as pyramidal-rotated.[58] Smaller rotation is caused by (C5)-H5 group of cytosine.[58] Pyramidal-rotated geometries of amino groups can be also induced by intermolecular interactions.[85] Note that rotation of the amino group weakens the double bond and facilitates a further pyramidalization.

Activation of Amino Groups in DNA and RNA Molecules

The amino groups usually adopt sp^2 planar geometry when involved in Watson–Crick base pairs though their flexibility can facilitate extreme values of propeller twisting and buckling.[83,86] However, the amino groups are often activated to adopt partial sp^3 geometry (even when participating in standard base pairing) when contacted with a donor or acceptor positioned above (below) the molecular plane of bases.[58,83–86] Flexible amino hydrogens smoothly follow the distribution of donors and acceptors in their proximity in an extent not anticipated in a vast majority of current x-ray, NMR, and MD studies.[83–86]

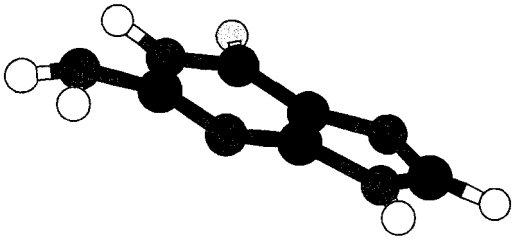

FIGURE 4 Pyramidal geometry of guanine amino group in a complete isolation.

Bifurcated Hydrogen Bonds and Close Amino Group Contacts

Partial pyramidalization of the amino groups likely helps to stabilize interstrand bifurcated H-bonds in ApA steps in oligo-adenine tracts (Figure 5).[83] Another interaction which is facilitated by amino group nonplanarity is the interstrand amino group contact in B-DNA (Figure 6).[83] Abundant occurrence of interstrand amino group contacts in CpG and mainly ApT B-DNA steps contradicts the original view that these interactions are purely repulsive steric clashes. The average N6-N6 distance in ApT steps in B-DNA crystal structures is 3.15 Å.[83] Quantum chemical calculations predict that close amino group contacts are inherently nonsymmetrical interactions and indeed they are systematically absent in those base-pair steps where a twofold symmetry is imposed by the crystal packing.[83] The major groove interstrand N6-N6 amino group contact in ApT B-DNA steps is the most frequent interbase-pair contact in DNA crystal structures.[83]

Unpaired Amino Groups and G · A Mismatches

Unpaired amino groups can easy form out-of-plane H-bonds and amino acceptor interactions (Figure 7). The out-of-plane bond between nonplanar guanine amino group in the highly propeller twisted G(*anti*) ·

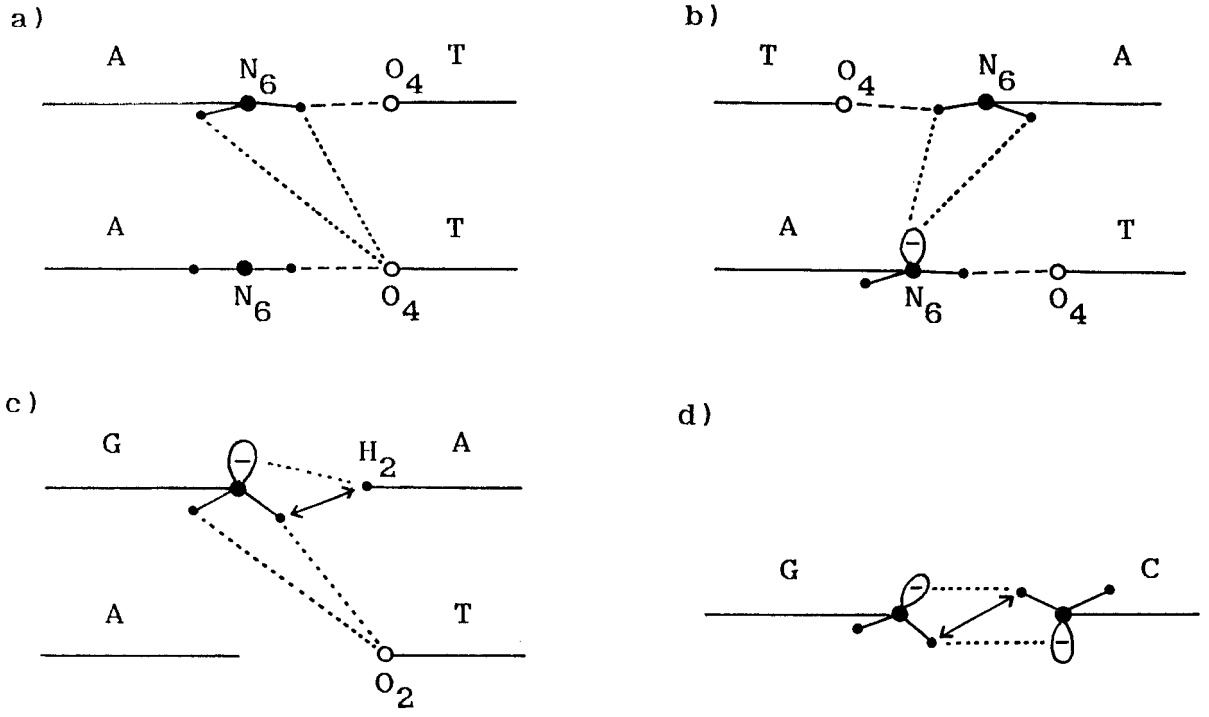

FIGURE 5 Activated nonplanar amino groups are assumed to be involved in stabilization of several interactions. (a) Interstrand major groove bifurcated H-bond in A-tract B-DNA structures. (b) Close major groove interstrand amino group contact seen in ApT steps of B-DNA crystal structures. (c) Strong out-of-plane H-bond characteristic for G(*anti*) · A(*anti*) mismatches. (d) The way how two amino groups forced in a close in-plane contact would eliminate a major part of the amino–amino repulsion.

A(*anti*) WC-like mismatch and the adjacent thymine in the d(CCAAGATTGG)$_2$ crystal structure[162] is one example.[86] QM calculations show that optimization of positions of the two guanine amino group hydrogens, invisible in x-ray experiments, improves the energy of the system by ca. 3 kcal/mol compared with planar amino group. The guanine amino group forms a *regular* H-bond with thymine.[86] The amino group hydrogens of bases can easy adopt dihedral angles about 50° with respect to the aromatic rings.[85] Note that the *crystal* geometry of the highly propeller-twisted G(*anti*) · A(*anti*) base pair[162] is almost identical to its gas phase global minimum (Figure 7).[86] Thus the observed nonplanarity is an intrinsic feature of the G ·

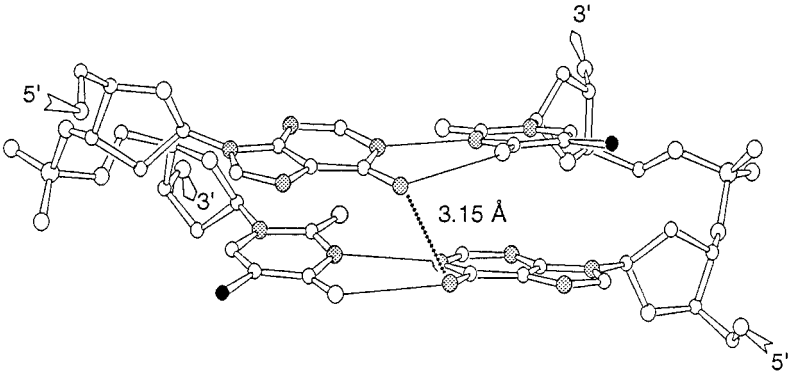

FIGURE 6 Close amino group contact in ApT steps as recurrently seen in B-DNA crystals. The average N6 · · · N6 distance is 3.15 Å. Because this interatomic contact contradicts the chemical intuition it has been mostly ignored, however, it is an important marker of B-DNA ApT step geometries.[40]

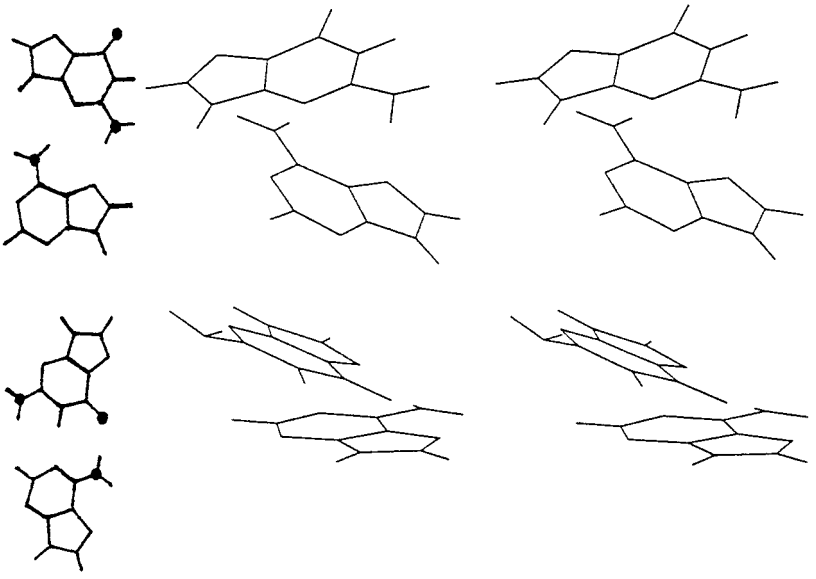

FIGURE 7 Optimal gas phase geometries of sheared (top) and *anti–anti* WC-like G · A mismatches are nonplanar. Their nonplanarity seen in DNA and RNA crystal structures is thus their inherent feature and is not related to stacking with adjacent bases. The substantially nonplanar amino group of guanine in G(*anti*) · A(*anti*) base pair forms strong out-of-plane H-bonds with adjacent base.

A base pair and is not a consequence of its environment. Analogous nonplanar G · A base pairs forming out-of-plane H-bonds with O2 of adjacent cytosines are seen in crystals of an RNA duplex[163] and a Holliday junction.[164] It is quite possible that the ability of the amino group in G(*anti*) · A(*anti*) base pair to form the out-of-plane H-bonds may contribute to the sequence dependence of the G · A base pairing. That is, when reversing the adjacent base pair, lack of the out-of-plane H-bond could destabilize the G(*anti*) · A(*anti*) with respect to the more common sheared arrangement. Note that also other G · A mismatches are intrinsically nonplanar,[86] and their available crystal geometries closely resemble their gas phase structures.[165,166]

DAPI Binding to Guanine

Recent crystal structure shows novel binding of DAPI (4′,6-diamidino-2-phenylindole) to an ATTG segment of a B-DNA duplex at a resolution of 1.9 Å, with a 3.04 Å contact between the amidinium nitrogen of DAPI and guanine amino group nitrogen atom.[85] Calculations predict that the binding is facilitated by a pyramidal-rotated geometry of the guanine amino group, which is simultaneously involved in out-of-plane H-bond and amino–acceptor interactions with solvent molecules, while still keeping its base pairing with cytosine essentially intact (Figure 8).

Amino Groups in the Force Fields

All major presently used molecular mechanical force fields assume that the amino groups are purely planar, greatly underestimating their flexibility. The force fields support neither out-of-plane H-bonds nor amino–acceptor interactions. Not surprisingly, neither bifurcated H-bonds nor amino-group contacts are properly reproduced by contemporary MD simulations.[167–170] MD simulations also show that the force

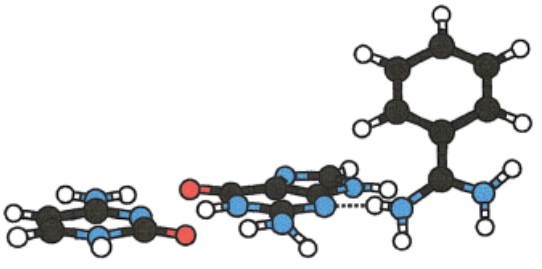

FIGURE 8 Crystal structure of $d(GGCCAATTGG)_2$ complexed with DAPI has revealed an unexpected 3.0 Å contact between amidinium nitrogen of DAPI and guanine amino group nitrogen N2.[85] QM calculations demonstrated that the DAPI amidinium group binds strongly to guanine N3 while its repulsion with the amino group is eliminated by bending the outer hydrogen of the amino group. On the picture, amidinium DAPI segment is on the right side. Crystal and QM data jointly suggest that two proximal water molecules interact with nonplanar guanine amino group, one as donor and the other as acceptor (not shown).[85]

field is incapable of reproducing the ATTG binding motif of DAPI mentioned above.[170] The standard assignment of hydrogen positions in crystallographic studies prevents the identification of out-of-plane interactions while the refinement protocols are assumed to penalize such geometries.

Amino Group sp^3 Pyramidalization and Basicity

It is widely assumed that the amino groups of bases are planar because they, in contrast to aniline amino groups, do not undergo protonation under solution experiments. This judgment based on basicity (pKa) studies is fundamentally wrong because the protonation of the amino group nitrogen (even considering the sp^3 hybridization) is considerably less favorable compared to bare ring nitrogen sites of the nucleobases (Table II).[171] Thus, the ring nitrogens are protonated first, and once any of them is protonated, the amino group becomes planar, due to electronic structure changes. Obviously, the lack of ring nitrogens permits a protonation of the aniline amino group. There is no reason to relate pyramidalization and basicity. Note (Table II) that the guanine amino group showing the largest degree of pyramidalization has in fact the worst amino nitrogen proton affinity among all bases studied. Proton affinity of adenine is substantially better and this explains why guanine appears to be more efficient in establishing out-of-plane H-bonds while adenine is more often seen in amino–acceptor contacts.[83–86] Note also that protonation energy of the N2 of 2-amino adenine (−230 kcal/mol) is better compared with aniline amino group (ca. −222 kcal/mol).[171]

Activated Amino Groups in Small Molecule Crystals

Interactions involving nonplanar amino groups typically occur when the amino groups are not fully satisfied by primary in-plane H-bonds. Thus, one finds these interactions only rarely in ultra high-resolution crystals of small compounds as these crystals have a maximized network of primary H-bonds. There is no reason to expect nonplanar amino groups to be statistically significant under such conditions. *A small-molecule dataset is not equivalent to situation in large systems.* There exists nevertheless solid x-ray evidence about an involvement of amino groups of bases in amino acceptor interactions.[84] Such interactions are likely to be more often utilized in unique three-dimensional shapes of large biomolecules and their complexes.

HYDROGEN BONDING OF DNA BASES

Hydrogen bonding of DNA bases is a fundamental interaction often challenged by QM calculations.

Nature of Base Pairing

Stability of H-bonding originates in the HF component of interaction energy, which means that the interaction is of electrostatic origin. A non-negligible contribution to base-pair stability comes also from dispersion attraction (intermolecular electron correlation). In view of our latest reference calculations, this contribution might be of the order of 4–5 kcal/mol for a base pair.[80] Several recent papers stressed the importance of a charge-transfer contribution to base pairing.[102–104] This agrees with the well established role of charge transfer in H-bonded complexes. It is, however, not trivial to extract the actual magnitude of the charge-transfer term as any energy decomposition is by definition ambiguous.[3,4] In fact, H-bonding of bases can be excellently reproduced by simple force fields neglecting charge transfer. It is, however, to be noted that the good performance of standard force fields for H-bonding of bases may be due to a compensation of errors. The concept of charge transfer is important for explanation of vibrational properties of H-bonding.[172]

Formation of a H-bonded base pair leads to a prolongation of the N-H bonds involved in the binding, which is of the order of 0.01 Å and accompanied with a significant red shift of the N-H stretch in the ir spectra.[89,172] Further, intensity of the N-H stretch increases upon formation of H-bond. Potential energy surfaces of base pairs are very complicated and contain large number of energy minima. To explore the gas phase potential energy surface completely, one needs an efficient searching technique and unexpected structures can be found as important minima.[16–20] Note also that especially RNA molecules show many additional H-bonding patterns, and there have been attempts to propose their classification based on geometry criteria.[173] Surprisingly, no systematic analysis of interaction energies of the RNA H-bonding patterns have been made although it is self-evident that many of the observed RNA base pair patterns do not correspond to minima on the intrinsic potential energy surfaces of base pairs.

Table VI and Figure 9 show interaction energies (see above for definition) of all H-bonded DNA base pairs formed by standard bases, with exclusion of base pairs involving the ring nitrogen atoms where sugar is attached. Table VII shows energetics of se-

Table VI The Interaction Energies of Planar DNA Base Pairs[a]

Pair	ΔE^{HF}	ΔE^{MP2}	ΔE^{T}	$\Delta E^{T,est}$	Pair	ΔE^{HF}	ΔE^{MP2}	ΔE^{T}	$\Delta E^{T,est}$
G · C WC	−24.6	−25.8	−23.8	−26.3	A · T WC	−9.7	−12.4	−11.8	−14.3
G · G1	−25.1	−24.7	−22.2	−24.7	G · G3	−16.0	−17.8	−17.0	−19.5
G · G4	−6.5	−10.0	−9.3	−11.8	G · A1	−12.2	−15.2	−14.1	−16.6
G · A2	−6.8	−10.3	−9.6	−12.1	G · A3	−10.8	−13.8	−13.1	−15.6
G · A4	−7.9	−11.4	−10.7	−13.2	A · T RH	−10.3	−13.2	−12.6	−15.1
A · T RWC	−9.6	−12.4	−11.7	−14.2	A · T H	−10.4	−13.3	−12.7	−15.2
C · C	−16.1	−18.8	−17.5	−20.0	G · C1	−11.6	−14.3	−13.4	−15.9
A · A1	−7.8	−11.5	−11.0	−13.5	A · A2	−7.2	−11.0	−10.3	−12.8
A · A3	−6.2	−9.8	−9.2	−11.7	G · T1	−14.2	−15.1	−13.9	−16.4
G · T2	−13.8	−14.7	−13.5	−16.0	T · C1	−8.7	−11.4	−10.7	−13.2
T · C2	−8.9	−11.6	−10.7	−13.2	T · T1	−9.3	−10.6	−10.0	−12.5
T · T2	−9.3	−10.6	−10.0	−12.5	T · T3	−9.3	−10.6	−9.9	−12.4
A · C1	−10.8	−14.3	−13.5	−16.0	A · C2	−10.4	−14.1	−13.2	−15.7

[a] ΔE^{HF} is the HF interaction energy, ΔE^{MP2} is the interaction energy after adding the electron correlation contribution, and ΔE^{T} is the total interaction energy after adding the monomer deformation energies.[62] The last column $\Delta E^{T,est}$ shows the estimated interaction energy based on very recent reference calculations on a small sample of base pairs and model systems.[86] These numbers differ from the ΔE^{T} values by −2.5 kcal/mol, and we suggest that $\Delta E^{T,est}$ is presently the most accurate estimate of base-pairing energetics to be used in force field verifications. Abbreviations: WC, Watson–Crick; H, Hoogsteen; R, reverse. Numbering of noncanonical base pairs according to Ref. 7 (see Figure 9). (All energies in kcal/mol.)

lected base pairs with non-natural bases. The last column of the Tables VI and VII shows the estimated interaction energy based on the preceding calculations with medium-sized bases sets of atomic orbitals[62,65] combined with an extrapolation based on latest reference calculations.[80]

Energies of Watson–Crick Base Pairs

The G · C WC base pair is almost twice more stable than the A · T WC base pair (Table VI). This reflects the number of H-bonds as well as the strong polarity of G and C compared with T and mainly A

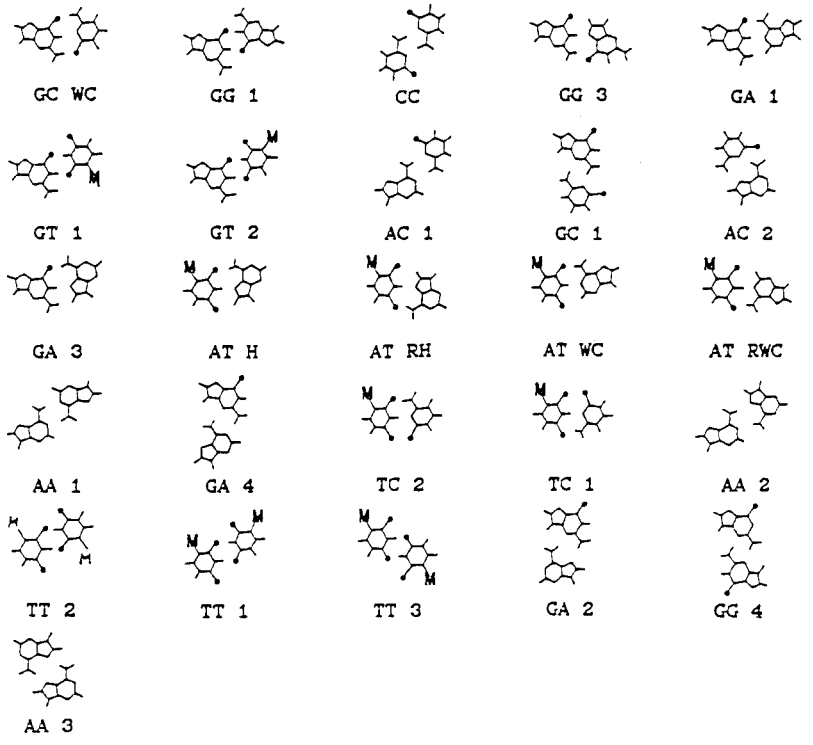

FIGURE 9 Structure and numbering of H-bonded base pairs.[7,62]

Table VII The Interaction Energies of Planar-Modified DNA Base Pairs[63,65,a]

Pair	ΔE^{HF}	ΔE^{MP2}	ΔE^{T}	$\Delta E^{T,est}$	Pair	ΔE^{HF}	ΔE^{MP2}	ΔE^{T}	$\Delta E^{T,est}$
^{6S}G · C WC	−23.1	−25.0	−22.5	−25.0	2amA · T WC	−11.9	−15.1	−14.1	−16.6
I · C WC	−18.5	−19.4	−18.0	−20.5	A · ^{4S}U WC	−8.4	−11.8	−11.1	−13.6
A · ^{2S}U WC	−9.6	−12.8	−12.1	−14.6	^{2S}U · ^{2S}U2	−8.7	−10.2	−9.6	−12.1
^{2S}U · ^{2S}U1	−6.9	−9.3	−8.8	−11.3	^{6S}G · ^{6S}G	−19.3	−22.3	−19.9	−22.4
C · CH+	−43.2	−44.8	−41.7	−44.2	A · F WC	−2.7	−3.9	−3.8	−6.4

[a] ΔE^{HF} is the HF interaction energy, ΔE^{MP2} is the interaction energy after adding the electron correlation contribution, and ΔE^{T} is the total interaction energy after adding the monomer deformation energies. The last column $\Delta E^{T,est}$ shows the estimated total interaction energy.[80] Abbreviations: WC, Watson–Crick; H, Hoogsteen; R, reverse; 2amA, 2-aminoadenine; I, 6-oxopurine (inosine); ^{6S}G, 6-thioguanine; ^{2S}U and ^{4S}U, 2- and 4-thiouracils; F, difluorotoluene; C · CH$^+$, triply bonded base pair (as seen in i-DNA). (All energies in kcal/mol.)

(see above). The 2-amino adenine–thymine base pair (Table VII) is, despite having three H-bonds, only slightly stronger compared to A · T WC base pair. This is due to the very low dipole moment of 2-amino adenine (Table I) caused by alternating pattern of donors and acceptors on its Watson–Crick edge. The inosine–cytosine base pair, despite having only two H-bonds, is considerably stronger compared to 2-amino adenine–thymine base pair and this again can be rationalized by the electronic distribution of inosine closely resembling that of guanine.

Energetics of Mismatches

Table VI further shows energetics of all possible mismatches consisting of standard bases. Note, that the mismatches range from very strong base pairs such as the symmetrical G · G 1 with two N1—H1 · · · O6 H-bonds up to very weak A · A base pairs.[62] Polar solvent greatly reduces the differences in base pairing stabilities.[52]

H-Bonding of Thioguanine and Thiouracils[65]

Table VII shows that, in contrast to a widespread opinion, substitution of a carbonyl group by a thiogroup does not alter the base-pairing energetics significantly. Note that thioguanine has a larger dipole moment than guanine (see above) and although the electrostatic component of H-bonding is slightly reduced by the bulky sulfur atom, the dispersion attraction is enhanced. Note also, that thiobases stack equally well as standard bases.[65] *Thus, destabilization of some nucleic acid forms by thiobases should not be explained as a result of a weak intrinsic H-bonding capability, and alternative explanations should be searched for.* Note also, that thioguanine does not destabilize triplexes.[174] QM calculations demonstrated that the thiogroup is, in contrast to the carbonyl group, very poorly hydrated.[65] Thus, 6-thioguanine could substantially change the hydration pattern in the duplex major groove.[65] The know ability of thioguanine to inhibit G-DNA formation[175–178] is a purely steric effect.[178] As revealed by MD simulations, the bulky thiogroup expels monovalent cations from the central ion channel of the G-DNA stem, leading to its collapse.[178] Thioguanine in isolation interacts with monovalent cations equally well as guanine.[178]

Protonated Base Pairs

Table VII also shows interaction energy of protonated triply-bonded cytosine dimer involving N3-protonated cytosine. This base pair is exceptionally strong due to a significant molecular ion–molecular dipole and induction contributions.[41] In protonated (A · C)H$^+$ base pair, the excessive proton occurs on adenine despite higher proton affinity of cytosine. This is a consequence of molecular interactions.[41] For more details about H-bonding of protonated base pairs, see Ref. 41. The role of protonated base pairs in i-DNA and triplexes has also been investigated using MD simulations.[42,45]

Hydrophobic Base Pairs

Electron correlation calculations clearly confirm that so-called hydrophobic bases such as diflurotoluene–adenine are not involved in true H-bonding and their gas-phase complexes are rather weak.[66,67] The gas phase studies were nicely complemented by large-scale MD simulations of oligonucleotides involving nonpolar analogues of bases, giving an excellent insight into the structural dynamics as well as thermodynamics of these base pairs when inserted into nucleic acids.[179–181]

Water-Mediated Base Pairs

QM calculations were also used to characterize gas phase structures and energies of water-mediated base

pairs occurring in RNA.[77,78] The water-mediated base pairs are rather strong H-bonded complexes and the inserted water molecule appears to be an integral part of the base pairing. The gas phase studies were nicely completed by explicit solvent simulation of an RNA molecule with water-mediated base pairs.[182] Note also, that in an isolated gas phase cluster 2–6 explicit water molecules are capable to convert H-bonded base pairs into stacked ones, thus microsolvation is capable to exert enormous effect on the nucleobase interactions.[20] Other recent QM studies investigated the structure of primary hydration shell around base pairs and possible effect of water insertion on base pair opening,[72–76] the later, however, at a semiempirical level.[76]

C—H · · · O Contacts in Base Pairs

Several base pairs contain C—H · · · O contacts. QM calculations revealed no C—H · · · O H-bonding contribution in the A · T and A · U WC base pairs.[89] The C2-H2(A) and O2(T) groups are separated and do not interact with each other. This conclusion follows from in-depth evaluation of electron topology and ir spectra of the A · U WC base pair.[89] On the other hand, C—H · · · O contact occurring in RNA U · U base pairs represents true, albeit weak H-bond, as demonstrated by red shift of the C—H stretch in the calculated ir spectra of this base pair as well as by electron topology analysis.[89] C—H · · · O contacts are delicate interactions, which require very high-accuracy methods to analyse their physicochemical nature. Low-quality methods as well as superficial calculations might be misleading. We advice a caution to judge about C—H · · · O interactions based on MD simulations utilising classical force fields.

Trimers and Tetramers

Several QM papers investigated H-bonded triads and quartets of DNA bases.[68–71,183] The calculations were compared with empirical force fields[68,71,178,183] and interactions of triads and quartets with cations have been studied.[68,70,71] Due to inherent neglect of polarization effects, the standard force fields are deficient in description of the interaction between guanine carbonyl groups and monovalent cations in guanine quartets.[71,183–186] Structure and dynamics of nucleobase quartets as well as their interaction with monovalent metal cations have been investigated in series of explicit-solvent MD studies of G-DNA molecules.[178,183–186]

The Effect of Solvent

Tables VI and VII show the intrinsic gas phase H-bonding energies. Water environment obviously compensates for the electrostatic contribution to H-bonding. Inclusion of solvent into QM calculations is a difficult task. Nevertheless, a thorough QM study has been published, based on the ab initio/Langevin-dipole approach, evaluating the thermodynamics parameters of 28 H-bonded base pairs and 10 stacked base dimers in a water environment.[52] Considerable attention has been recently paid to microsolvation of base pairs.[20,54,72–76]

G · A Mismatches and Other Intrinsically Nonplanar Base Pairs

Crystal structures often show base pairs substantially nonplanar, propeller twisted, and buckled. These nonplanarities are mostly explained as a result of base-stacking forces. It likely is correct for A · T and G · C WC Crick base pairs, as these base pairs are planar in isolation.[62] Note, however, that QM vibrational studies indicate that the base pairs are exceptionally flexible with respect to buckle and propeller motions, with out-of-plane vibrational frequencies around 20 cm^{-1}.[187] Isolated A · T WC base pair would adopt an effectively nonplanar geometry at room temperature.[187] On the other hand, many noncanonical base pairs are nonplanar intrinsically.[86] The nonplanarity is often induced/facilitated by nonplanarity of the amino groups as well as by secondary electrostatic interactions. Note that especially all G · A mismatches are nonplanar, including sheared G · A and WC-like G(*anti*) · A(*anti*) binding patterns (see above).[86] Structure and dynamics of sheared G · A mismatch base pairs inside a DNA molecule have been characterized by MD simulations.[166]

AROMATIC BASE STACKING

One of the most important tasks for electronic structure calculations was to reveal the physicochemical nature of aromatic base stacking and to clarify which empirical potential form is sufficiently accurate do describe base stacking. The QM calculations of base stacking were thoroughly reviewed elsewhere[40]; thus, here we present only a brief overview.

Physicochemical Nature of Stacking

Ab initio calculations with inclusion of electron correlation revealed the following picture of base stacking interactions between two bases.[40,63–65] *Base stacking is basically determined by three contributions: dispersion attraction, short-range exchange repulsion, and electrostatic interaction. No specific π–π interactions have been revealed. Dispersion attraction originates in the intermolecular correlation of*

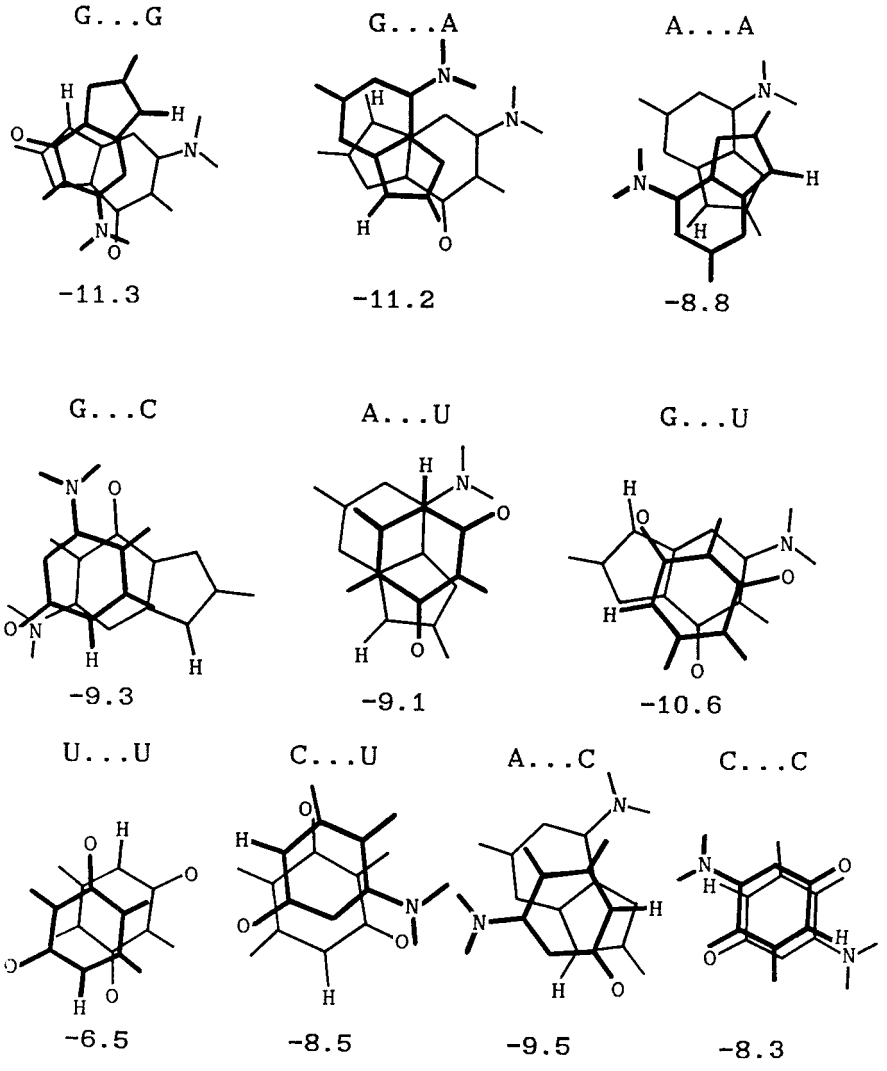

FIGURE 10 Optimal gas phase stacking arrangements in ten nucleobase dimers and the corresponding stacking energies in kcal/mol.[63]

electron motions classically called "induced dipole–induced dipole interactions." Stabilization of base stacking is dominated by the dispersion attraction, which is rather isotropic and proportional to the geometrical overlap of the bases. The vertical distance between stacked bases is determined by balance between dispersion attraction and short-range exchange repulsion. Finally, the mutual orientation of bases and their displacement are primarily determined by the electrostatic interactions (Figure 10). Note that response of polar solvents to the electric field of the bases would lead to a destabilization of the most stable gas phase arrangements.[52]

Empirical Potential Description of Base Stacking

Simple empirical potential consisting of a Lennard–Jones van der Waals term and a Coulombic term with atom-centered point charges [Eq. (3)] is able to reproduce the ab initio stacking energy within the range of ± 1.5 kcal/mol per stacked dimer over the major portion of the conformational space[63–65]:

$$\Delta E(\text{kcal/mol}) = \sum 332 q_i q_j / r_{ij} + \sum - A_{ij}/r_{ij}^n + \sum B_{ij}/r_{ij}^m \quad (3)$$

where $n = 6$, $m = 9$–12, q are atom-centered point charges, A and B are constants of the van der Waals term, and r is the interatomic distance between atoms i and j. The summation is performed over all atoms of both interacting monomers. The surprisingly good agreement of the simple empirical potential with the ab initio calculations bolster the current molecular modeling, since this analytical potential function is employed in almost all parametrizations used in studies of DNA. We wish to point out here that already the

comprehensive empirical potential characterization of stacked base dimers reported by Poltev and Shulyupina[188] using the Zhurkin et al. force field[189] was qualitatively correct. However, this does not mean that each empirical potential is accurate. Differences of force fields used in the past in studies of nucleic acids are rather large.[5,190] This uncertainty can be substantially reduced these days using the ab initio reference data, and a thorough comparison of widely used empirical potentials has been published.[5,190]

Base Stacking in Duplex DNA

High-level QM calculations have been used to characterize base stacking in standard and high-resolution x-ray B-DNA and Z-DNA geometries. *Surprisingly, the local conformational variations seen in the highest-resolution crystals of duplexes do not improve the intrinsic stacking energy terms*.[55,64] All geometries observed in high-resolution crystals are rather isoenergetical. It is especially apparent for the CpA(TpG) step, which shows unprecedented variability of stacking patterns and yet the different stacking geometries fall within 1 kcal/mol.[55] Another notable step is the GpG one, characterized by intrastrand electrostatic repulsion and a repulsive four-body term.[55,64] It should be noted that medium and lower resolution crystal structures often show deterioration of stacking compared to standard geometries caused by data and refinement inaccuracies.[190] We have argued against relying of visual inspection of base overlaps in evaluation of stacking as the overlap-dependent part of stacking shows only minor variability.[55] In contrast, the electrostatic stacking terms show a substantial sequence dependence leading to large but mutually compensating variability of intrastrand and interstrand stacking.[55,64]

Base Stacking in i-DNA

In contrast to all other DNA forms, i-motif is characterized by *repulsive* intrinsic base stacking, due to a molecular ion–molecular ion repulsion.[41] The vertical repulsion is of the order of +20 to +30 kcal/mol (Figure 11).[41] Nevertheless, large-scale explicit solvent MD simulations demonstrated that in the water environment and considering the backbone arrangement, the intrinsic base–base repulsion is compensated for, leading to very stable and rigid i-DNA assemblies in water (see above).[42]

Sugar–Base Stacking

Nucleic acid bases can also be involved in sugar–base stacking. QM calculations demonstrated that this interaction should be classified as a dispersion-controlled contact, so its physical nature is close to regular base stacking. No unusual molecular orbital contribution to sugar–base stacking was found.[64]

Intercalator–Base Stacking

Recently, the first electron correlation study of the interaction between monocationic intercalator amiloride and DNA bases has been published.[79] The extended calculations confirm that intercalator–base interactions are very similar to base stacking, and can be very satisfactorily reproduced by the AMBER force field. This conclusion could still change when considering other intercalators and work is in progress in our laboratory to evaluate more systems.

INTERACTIONS OF METAL CATIONS WITH BASES, BASE PAIRS, AND NUCLEOTIDES

With increasing power of computers we are in a position to study the interactions of metals with bases and base pairs using reliable quantum chemical methods. Complexes with metal cations are characterized by very large nonadditivities of interactions. The strength of these effects increases with the charge of the cation and is substantially larger for transition metal elements. For such systems one cannot use simple force fields and methods considering the electronic structure of the complexes must be applied. On the other hand, metal cation containing complexes are mostly *non-neutral.* For non-neutral ionic systems the gas-phase nature of the ab initio technique creates problems concerning a comparison with the situation present in condensed phase. While the ionic electrostatic effects dominate in the gas phase, they are almost eliminated in polar solvents and in crystals. It should be noted that the same problem exists also for gas phase experiments, and gas phase studies of complexes between bases and bare cations obviously show a different picture compared to the binding modes relevant to nucleic acids. Thus, quantum chemical studies of metal cation containing systems are not trivial and it is essential to study extended complexes. Nevertheless, our recent studies show that when gas phase data are properly analyzed they provide surprisingly good correlation with condensed phase bioinorganic data.[171]

Nonadditivity of Interactions in the Cation Coordination Sphere

Nonadditivity of interactions is exceptionally important to properly account for the metal cation coor-

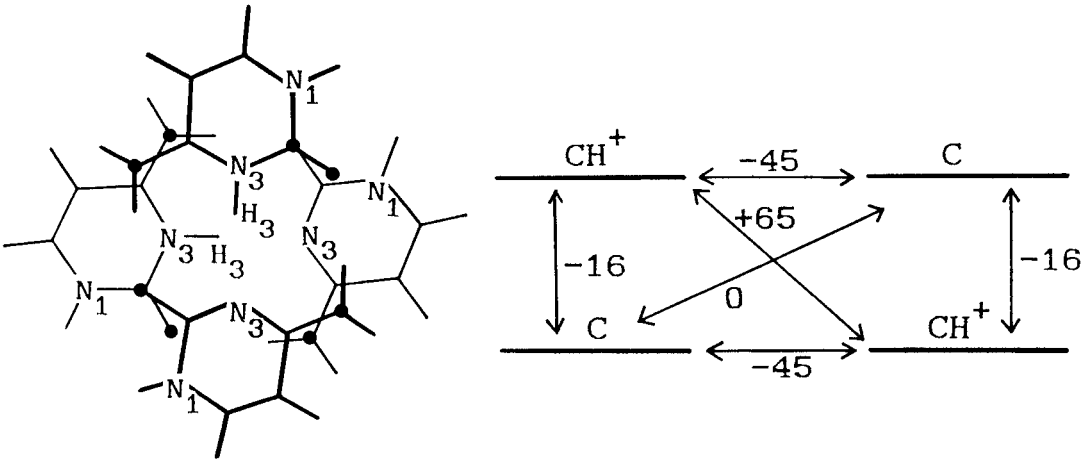

FIGURE 11 Gas phase molecular interactions of bases in I-DNA geometries (kcal/mol). The protonated H-bonded base pairs are very stable due to electrostatic and induction interactions. However, the vertical stacking is highly repulsive due to the molecular ion–molecular ion repulsion.[41] Explicit solvent MD simulations have demonstrated that solvent screening effects and negatively charged backbone overcome this repulsion.[42]

dination and hydration, and to describe the balance of interactions of a given cation with various chemical groups in biopolymers. Nonadditive interactions cannot be treated as a sum of the dimer (pairwise) interactions, but there also are substantial many body terms. For example, interaction of a trimer may be expressed as a sum of three pairwise dimer terms and a three-body term. Many-body terms are neglected by standard force fields and the same is true for induction and charge-transfer contributions to the pairwise terms. Thus the force fields neglect all nonadditivities, specific electronic differences among cations (see below) and strongly underestimate dimer cation–base binding energies.[68,113] The importance of many body terms may be well documented on a simple hydration of a cation. Hydration of a divalent cation by six water molecules is associated with a hydration energy of ca. −300 to −350 kcal/mol. This contribution consists of six attractive cation– water pairwise energies and a repulsive many-body term reflecting the interligand repulsion caused by polarization of the water molecules by the cation. This interligand repulsion is typically on a scale of +70 kcal/mol.[113] The polarized water molecules (or other ligands) around the cation have a greatly enhanced ability to act as strong H-bond donors outside the primary hydration shell.[116] Quantum chemical calculations are able to accurately capture all these nonadditivities and represent these days the basic tool to study interactions of metal cations with fragments of biopolymers (Figure 12). In some instances specialized nonadditive potentials can be used.[191,192]

The Effect of Cation Binding on the Base Pair Strength

Ab initio calculations predicted that direct binding of a cation to the N7 site of guanine significantly influences the strength of the guanine-containing base

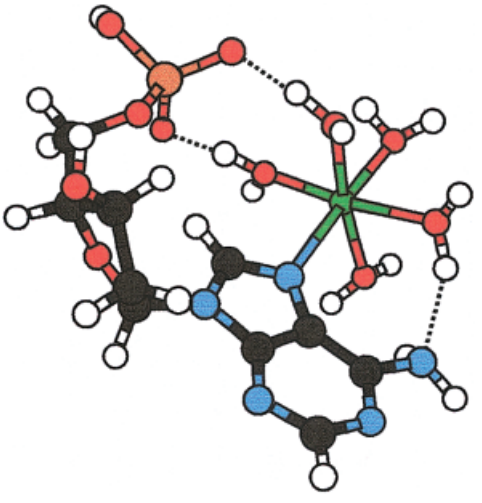

FIGURE 12 Inner-shell binding of hydrated divalent cation to $dAMP^{-1}$. The polarized water molecules of the cation hydration shell form very strong ionic water bridges with the anionic oxygens of the backbone segment. These water bridges are characterized by large red shifts and lengthening of the respective O–H bonds of the order of 0.04 Å. There is a strong amino–acceptor interaction between another polarized water molecule and the highly pyramidal-rotated adenine amino group. The pyramidalization has essentially no effect on Watson–Crick base pairing of adenine. For more details, see Refs. 116 and 172.

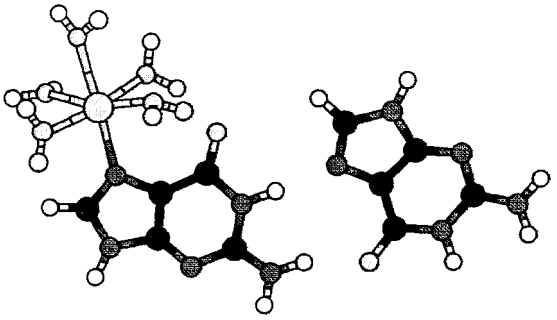

FIGURE 13 G · G base pair interacting with pentahydrated Zn^{2+}. The cation binding strengthens this particular H-bonded base pair by ca 100% due to a combination of induction and electrostatic effects. For more details, see Ref. 113.

pairs and provides about 5–10 kcal/mol of additional polarization stabilization.[112–114] The strength (the energy necessary to disrupt the base pair) of isolated reverse Hoogsteen G · G base pair is around 18 kcal/mol. When a hydrated divalent metal cation binds to the N7 position of guanine (Figure 13) the G · G base-pairing strength is doubled to 36 kcal/mol.[112] Approximately half of this metal-ion induced stabilization (cooperativity) is due to polarization effects, the rest stems from electrostatic attraction between the hydrated cation and the nonmetalated guanine. The polarization effect is reduced but not eliminated when including the negatively charged backbone into the calculations.[116] Interestingly, the cation binding to N7 has no influence on the stability of the adenine–adenine reverse Hoogsteen base pair.[114] The effect of stabilization of base pairs by N7-metal binding has been recently confirmed for platinated base pairs in condensed phase experiments.[194,194]

Zinc vs Magnesium Difference

Ab initio techniques can be efficiently used to study specific differences among cations. A textbook example is the difference between Zn^{2+} and Mg^{2+}. Both cations have the same charge and approximately the same ionic radius. Thus, they would have very similar properties when treated by pair-additive empirical force fields. However, quantum chemical calculations clearly show that, compared to magnesium, zinc has a much larger affinity to bind to nitrogen sites, while both metal cations have rather similar affinities to interact with oxygen atoms.[111–113,195,196] The partial covalent bonding involving the 3d electrons of zinc and nitrogen lone electron pairs is the most important contribution. This explains why zinc and other transition metals often interact with nucleobases in DNA, while magnesium tends to bind almost exclusively to the anionic oxygens of the phosphate groups. When hydrated Mg^{2+} binds directly (inner-shell binding) to the N7 position of purine bases, the interaction can be viewed as interaction between nucleobase and a hydrated cation.[113] It is rather easy to release the Mg^{2+} back to the solvent. Zn^{2+} in the same position establishes a different interaction. Although binding geometries of hydrated Zn^{2+} and Mg^{2+} look very similar, zinc forms a considerably stronger interaction with N7 while weakening its binding with the hydrating water molecules. Thus, compared to the Mg^{2+} binding, the hydrated Zn^{2+}–nucleobase complex is shifted considerably to what could be described as hydration of a metalated base. Once bound to N7, it is more difficult to separate Zn^{2+} away this position, compared to Mg^{2+}. The quantitative picture is illustrated in Table VIII. On the other hand, the sulfur atom is capable of very efficiently discriminating, for example, between Zn^{2+} and Hg^{2+}.[115] Modern QM methods allow us to analyze such differences over a wide range of cations and binding sites, and provide a direct quantitative energetic perspective of theories such as HSAB (hard and soft acids and bases).

Cation–π Interactions Involving Nucleobases?

Recent crystallographic study suggested a presence of cation–π interaction between hydrated magnesium and a cytosine in a B-DNA duplex.[197] Thus, we have characterized such interactions using ab initio methods, assuming idealized and crystal geometries of the interacting species.[87] In principle, hydrated metal cations can interact with nucleobases in a cation–π manner. The stabilization energy of such complexes would be large and comparable to a cation–π complex with benzene. However, in contrast to the benzene–cation complexes, the cation–π configurations are highly unstable for nucleobases since the conventional in plane binding of hydrated cations to the acceptor sites on the nucleobase is strongly preferred.[87] Thus, a cation–π interaction with a nucleobase can occur only if the position of the cation is locked above the nucleobase plane by another strong interaction. In the particular crystal structure the cytosine – cation geometry is determined by the primary interactions between the hydrated cation and N7/O6 guanine sites. The cation is rather far from the cytosine and their mutual interaction is weak. Thus it is not a cation–π interaction, though it does not rule out that the electric field of the cation influences the cytosine position.[197] Eventually, the cytosine position may be influenced by polarization of the adjacent guanines interacting with the hydrated cation. We do

Table VIII Interaction Energies in Selected Complexes Involving Zinc and Magnesium Divalent Cations (M; metal)[a]

	Zn^{2+}	Mg^{2+}	Difference (Zn^{2+}–Mg^{2+})
$G \cdots M^{2+}$	−185.1	−148.9	−36.2
$H_2O \cdots M^{2+}$	−94.9	−82.9	−12.0
$M^{2+} \cdots 6H_2O$[b]	−339.8	−329.6	−10.2
$G\text{-}M^{2+} \cdots 5H_2O$[c]	−203.2	−220.3	+17.1
$G \cdots (M^{2+} + 5H_2O)$[d]	−94.4	−89.9	−4.5
$G \cdots M^{2+} \cdots 5H_2O$[e]	−388.4	−373.7	−14.7

[a]Direct (inner-shell) binding of pentahydrated M to guanine N7 is considered. The last three rows show different ways how the complex between guanine and pentahydrated cation can be divided into subsystems to highlight the difference between zinc and magnesium. All energies in kcal/mol. For more details see Ref. 113.

[b]Hexa-hydration of the cation (seven subsystems).

[c]Hydration of the "metalated base" $G\text{-}M^{2+}$ (subsystems: five water molecules and the metalated base). Note the substantial reversal of the zinc–magnesium difference compared with the preceding line. Replacement of one water molecule by guanine N7 in the primary hydration shell substantially reduces the hydration energy of zinc compared with Mg^{2+}.

[d]Interaction between the hydrated cation and base (two subsystems). Although this term appears to be similar for both cations, when considering the preceding two lines we clearly see the different balance of water–cation and base–cation interactions for zinc and magnesium.

[e]Interaction energy of the whole complex (seven subsystems).

not assume that cation–π interactions involving DNA bases are common and a convincing example has yet to be shown.[87]

Metal-Cross-Linked Base Pairs

Ab initio calculations have been recently used to investigate base-pair mismatches stabilized by a metal cross-link.[119,198] The calculations complement the experimental studies, evaluate the intrinsic energetics of the interactions, and analyze the potential energy surface of the base pairs in isolation. Studies of base pairs stabilized by Ag(I) metal cation (Figure 14) revealed a surprising flexibility of this metalated base pair that is larger compared with nonmetalated base pairs.

The Effect of Metalation on the Electronic Structure of Bases, Protonation, Deprotonation, and Tautomerism

It is well established by experiments that metalation of nucleobases can cause shifts of nucleobase hydrogens, such as protonation, deprotonation, and formation of rare tautomers of bases.[151–158] Proton shifts could induce formation of mismatched base pairs and this could contribute to mutagenecity of certain metals. It has been suggested that many of the proton shifts could actually be caused by nonelectrostatic effects—that is, by substantial changes of the electronic structure of nucleobases by metalation. Due to the number of possible cation binding modes as well as variability of environmental effects, it is quite tempting to complement the bioinorganic experiments by reliable gas phase calculations. QM calculations can rather unambiguously identify the nature of the interactions and the forces causing the proton shifts. QM calculations can also provide valuable data for parametrization of empirical force fields for molecular modeling of metalated bases.[191,192,199]

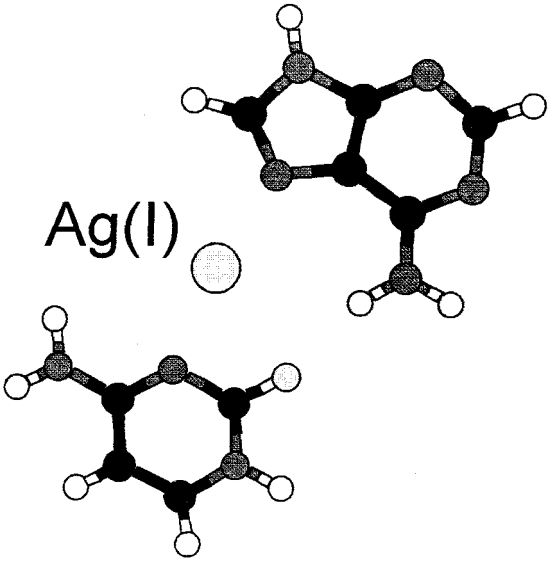

FIGURE 14 Cytosine–Ag(I)–adenine mispair. This base pair is dominated by two rather strong Ag(I)–N(base) contributions. The metalated base pair shows a pronounced flexibility around the N(A)–Ag(I)–N(C) bond. The Ag(I) cation is capable to switch from nucleobase ring nitrogen sites to exocyclic oxygens atoms. For more details, see Ref. 119.

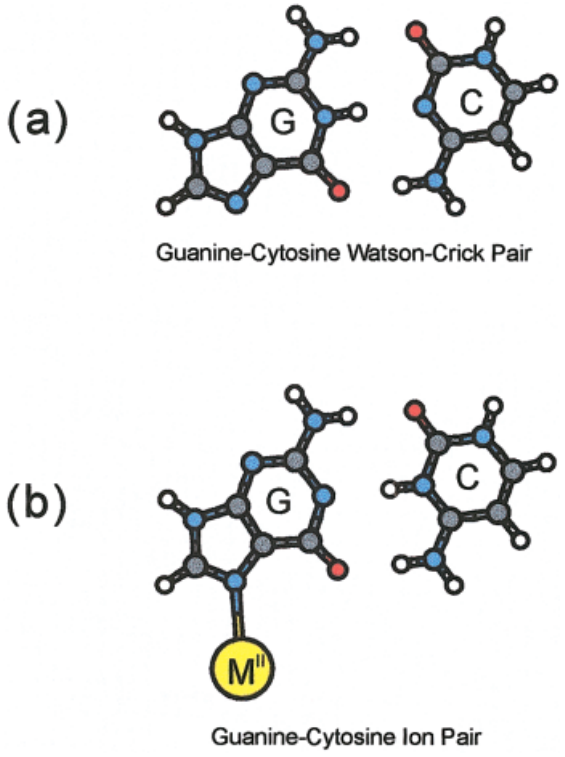

FIGURE 15 Metalation of the guanine N7 position can acidify the H1 hydrogen of guanine and increase probability of N1 deprotonation or hydrogen transfer to N3 position of cytosine. For more details, see Ref. 116.

QM studies of N7-metalated adenine and guanine have revealed that binding of Pt(II) can substantially affect tautomeric equilibria of bases. The calculations however suggested that the effects are caused primarily by electrostatic forces, as they vanished when the base interacted with a neutral Pt(II) adduct.[118] Thus, it is predicted that in a polar solvent the effect of N7-platination on base tautomerism would be attenuated. The QM calculations allowed for the first time to study in detail specific interactions between ligands attached to metal and nucleobase exocyclic groups.[118,171]

QM calculations also revealed that binding of hydrated divalent metal cations to N7 position of guanine substantially acidifies the guanine H1 hydrogen and facilitates its possible transfer to N3 site of cytosine in the G · C WC base pair.[116] This would mean a formation of ion-pair noncanonical form of the G · C WC base pair, one of the possible intermediates involved in formation of point mutations (Figure 15).[90] However, calculations carried out with more complete systems show, that at least for metals such as Zn^{2+}, the marked acidification of H1(G) is exclusively caused by electrostatic effects.[116] Thus proximity of the negatively charged backbone as well as solvent screening effects reduce the probability of the ion-pair formation. The H1-acidification due to N7 metal binding can nevertheless occur in DNA and its magnitude could depend on composition of the adjacent base pairs as well as on the DNA architecture and environment.[116]

QM calculations revealed that metalation of the exocyclic amino groups of adenine and cytosine substantially increases the proton affinity of the ring nitrogen sites of bases.[88] In contrast to the N7 binding, the proton shifts caused by amino-metalation are due to substantial nonelectrostatic effects and therefore are expected to remain fully expressed in polar solvents.[88] Further, the magnitude of the effect depends on the cation type substantially.[88] The computational results complemented series of experimental bioinorganic studies of metal-assisted rare tautomers of nucleobases.[151–158]

QM calculations were also used in a systematic study of the effect of platination of adenine on its gas phase proton affinity, investigating several binding sites and metal adducts.[171] A systematic comparison of the gas phase proton affinities with available condensed phase pKa values revealed a surprisingly good correlation.[171] The calculations help to understand why the effect of platination on adenine pKa values

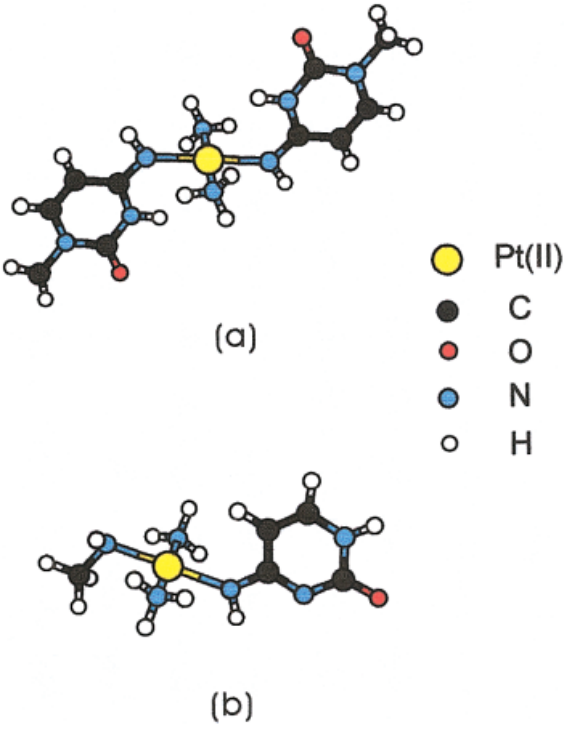

FIGURE 16 Platination of the cytosine amino group substantially changes electronic structure of the nucleobase and improves proton affinity of the N3 cytosine site by ca 30 kcal/mol.[88] (a) Crystal structure and (b) model complex utilized in the calculations.

shows essentially no dependence on the charge carried out by the metal entity. The study revealed number of specific ligand–base interactions including amino–acceptor interactions and intrasystem proton transfers. Their significance in the condensed phase situation remains to be clarified.[171]

CONCLUDING REMARKS

With the advance of high-level ab initio calculations with inclusion of electron correlation effects we for the first time have reliable nonempirical technique to study intrinsic interactions of nucleic acid bases. An important advantage of the QM approach is absence of any empirical parameters and physicochemical completeness of the data. This allows us to unambiguously determine nature of molecular interactions in the gas phase condition, i.e., in absence of any environmental perturbations. Within just a couple of years all basic interactions of nucleic bases have been characterized with accuracy and reliability not affordable by any other experimental or computer-based technique. The calculations provide an ultimate confirmation that stabilization of stacked nucleobase structures can be ascribed to the dispersion forces, that is, intermolecular electron correlation effects. The electrostatic contribution is essential to determine the mutual orientation of stacked bases. The advanced quantum-chemical theory ruled out some previously postulated models of base stacking (namely π–π and "induction" models)[200,201] and show that a simple empirical potential relying on the use of atom-centered point charges is *qualitatively* sufficient to describe base stacking. The ab initio calculations furnished accurate data for H-bonding of DNA bases. The ab initio calculations revealed intrinsic nonplanarity and flexibility of amino groups of DNA bases implying their ability to be involved in efficient out-of-plane H-bonds and amino–acceptor interactions. The ab initio calculations provide reference data for parametrization and verification of force fields utilized in molecular modeling and molecular dynamics simulations. The calculation also convincingly demonstrated drastic failures of all semiempirical and all standard DFT techniques for interactions of DNA bases. Ab initio calculations have been successfully applied also in studies of interactions between metal cations and nucleic acid fragments. Outcomes of recent gas phase QM calculations often substantially changed opinions built up on purely phenomenological evaluations of structural data and chemical intuition. We hope that the QM results will be more often considered in future x-ray, NMR, and MD studies of nucleic acids. This would certainly improve quality and relevance of discussions of molecular interactions and eliminate unnecessary speculations and misinterpretations.

Undoubtedly, various aspects of nucleic acid structure and interactions will be studied by QM methods in future. The standard techniques described in this review can be now used, with modest computer equipment, for any medium-sized system. In the near future, we expect the final refinement of QM data for energetics of H-bonding and stacking of nucleobase dimers, though it is assumed that the current values should be no more than 1–2 kcal/mol away the exact values. We also expect substantial efforts devoted to studies of a wide range of metal complexes of nucleic acids including open shell systems. Also, improvements in studies of excited states, electronic spectra, electron transfer, and other electronic properties of bases and base pairs is expected.[92,107,108,202–210] Finally, we hope that this decade will bring some substantial methodological advances allowing a proper evaluation of solvent effects and execution of QM-based MD simulations with some fast but yet reasonably accurate QM methods. This study was supported by grant LN00A032 (Center for Complex Molecular Systems and Biomolecules) by Ministry of Education of the Czech Republic, by the National Science Foundation (NSF CREST 9805465 and 9706268) (JL), and by the National Institute of Health (NIH RCMI G1 2RR13459-21). We would like to thank Supercomputer Center, Brno, and Mississippi Center for Supercomputer Research for a generous allotment of computer time.

REFERENCES

1. Šponer, J.; Leszczynski, J.; Hobza, P. J Biomol Struct Dynam 1996, 14, 117–136.
2. Šponer, J.; Hobza, P. In Encyclopedia of Computational Chemistry; Schleyer, P. v. R.; Allinger, N. L; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer, H. F., III,; Schreiner, P. R., Eds., John Wiley & Sons: Chichester, UK 1998; pp 777–789.
3. Šponer, J.; Leszczynski, J.; Hobza, P. In Theoretical Computational Chemistry, Vol. 8, Computational Molecular Biology; Leszczynski, J., Ed.; Elsevier, New York, 1999; pp 85–117.
4. Šponer, J.; Hobza, P. Chem Rev 1999, 99, 3247–3276.
5. Hobza, P.; Kabeláč, M.; Šponer, J. Mejzlik, P.; Vondrášek, J. J Comput Chem 1997, 18,1136–1150
6. Del Bene, J. E. Theochem—J Mol Struct 1985, 25, 201–212.
7. Hobza, P.; Sandorfy, C. J Am Chem Soc 1987, 109, 1302–1307.
8. Aida, M.; Nagata, C. Int J Quatum Chem 1986, 29, 1253–1261.

9. Aida, M. J Theor Biol 1988, 130, 327–335.
10. Aida, M. J Comput Chem 1988, 9, 362–368.
11. Sagarik, K. P.; Rode, B. M. Inorg Chim Acta 1983, 78, 177–180.
12. Anwander, E. H. S.; Probst, M. M.; Rode, B. M. Biopolymers 1990, 29, 757–760.
13. Hobza, P.; Šponer, J.; Reschel, T. J Comput Chem 1995, 16, 1315–1325.
14. Šponer, J.; Leszczynski, J.; Hobza, P. J Comput Chem 1996, 17, 841–850.
15. Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. Biopolymers 1979, 18, 1149–1160.
16. Kratochvil, M.; Engkvist, O.; Sponer, J.; Jungwirth, P., Hobza, P. J Phys Chem A 102, 1998, 6921–6926.
17. Kratochvil, M.; Šponer, J.; Hobza, P. J Am Chem Soc 2000, 122, 3495–3499.
18. Gervasio, F. L.; Procacci, P.; Cardini, G.; Guarna, A.; Giolitti, A., Schettino, V. J Phys Chem B 2000, 104, 1108–1104.
19. Kabelac, M.; Hobza, P. Chem Eur J 2001, 7, 2069–2074.
20. Kabelac, M.; Ryjacek, F.; Hobza, P. Phys Chem Chem Phys 2000, 2, 4906–4909.
21. Dey, M.; Grotemeyer, J.; Schlag, E. W. Z Naturforsch A 1994, 49, 776–784.
22. Šponer, J.; Hobza, P. Chem Phys Lett 1996, 261, 379–384.
23. Desfrancois, C.; Abdoul-Carime, H.; Schulz, C. P.; Schermann, J. P. Science 1995, 269, 1707–1709.
24. Desfrancois C.; Carles, C.; Schermann, J. P. Chem Rev 2000, 100, 3943–3962.
25. Nir, E.; Kleinernmanns, K.; de Vries, M. S. Nature 2001, 408, 949–951.
26. Hoayu, S.; Norman, K.; McMahon, T. B.; Ohanessian, G. J Am Chem Soc 1999, 121, 8864–8875.
27. Rodgers, M. T.; Armentrout, P. B. J Am Chem Soc 2000, 122, 8548–8558.
28. Plokhotnichenko, A. M.; Radchenko, E. D.; Stepanian, S. G. Adamowicz, L. J Phys Chem A 1999, 103, 11052–11059.
29. McCarthy, W.; Plokhotnichenko, A. M.; Radchenko, E. D., Smets, J.; Smith, D. M. A.; Stepanian, S. G.; Adamowicz, L. J Phys Chem A 1997, 101, 7208–7216.
30. Kuechler, E.; Derkosch, J. Z Naturforsf B Anorg Chem Org Chem 1966, 21b, 209–220.
31. Pitha, J.; Jones, R. N.; Pithova, P. Can J Chem 1966, 44, 1045–1051.
32. Ts'o, P. O. P.; Melvin, I. S.; Olson, A. C. J Am Chem Soc 1963, 85, 1289–1296.
33. Gray, D. M. Biopolymers 1997, 42, 783–793, and references therein.
34. Bommarito, J.; Peyret, S.; SantaLucia, J. Nucleic Acid Res 1998, 28, 1929–1934.
35. SantaLucia, J. Proc Natl Acad Sci USA 1998, 95, 1460–1465.
36. Nakano, S.; Fujimoto, M.; Hara, H.; Sugimoto, N. Nucleic Acids Res 1999, 27, 2957–2965.
37. Sartorius, J.; Schneider, H.-J. J Chem Soc Park Tr 2, 1997, 2319–2327 and references therein.
38. Mizutani, M.; Kubo, I.; Jitsukawa, K.; Masuda, H. Einaga, H. Inorg Chem 1999, 38, 420–421.
39. Newcomb, L. F.; Gellman, S. H. J Am Chem Soc 1994, 116, 4993–4994.
40. Šponer, J.; Berger, I.; Špackova, N.; Leszczynski, J.; Hobza, P. J Biomol Struct Dynam 2000, Conversation 11, Special Issue 2, 383–407.
41. Šponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. J Biomol Struct Dynam 1996, 13, 695–707.
42. Špackova, N.; Berger, I.; Egli, M.; Šponer, J. J Am Chem Soc 1998, 120, 6147–6151.
43. Gehring, K.; Leroy, J. L.; Gueron, M. Nature 1993, 363, 561–565.
44. Berger, I.; Egli, M.; Rich, A. Proc Natl Acad Sci USA 1996, 93, 12116–12126.
45. Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. J Am Chem Soc 1998, 120, 11226–11233.
46. Luo, R.; Gilson, H. S. R.; Potter, M. J.; Gilson, M. K. Biophys J 2001, 80, 140–148.
47. Arora, N.; Jayaram, B. J Phys Chem B 1998, 102, 6139, 6144.
48. Friedman, R. A.; Honig, B. Biopolymers 1992, 32, 145–159.
49. Cramer, C. J.; Truhlar, D. G. Chem Rev 1999, 99, 2161–2200.
50. Luque, F. J.; LopezBes, J. M.; Cemeli, J.; Aroztegui, M.; Orozco, M. Theor Chem Acc 1997, 96, 105–113.
51. Orozco, M.; Luque F. J. Chem Rev 2000, 100, 4187–4226.
52. Florian, J.; Šponer, J.; Warshel, A. J Phys Chem B 1999, 102, 884–892.
53. Sivanesan, D.; Subramanian, V.; Nair, B. U.; Ramasami, T. Ind J Chem A 2000, 39, 132–138.
54. Sivanesan, D.; Babu, K.; Gadre, S. R.; Subramanian, V. J Phys Chem A 2000, 104, 10887–10894.
55. Šponer, J.; Florian, J.; Ng, H.– L., Šponer, J. E., Špackova, N. Nucleic Acid Res 2000, 24, 4893–4902.
56. Leszczynski, J. Int J Quantum Chem Quantum Biol Symp 1992, 19, 43–52.
57. Šponer, J.; Hobza, P. Theochem—J Mol Struct 1994, 305, 35–40.
58. Šponer, J. Hobza, P. J Phys Chem 1994, 98, 3161–3164.
59. Gould, I. R.; Kollman, P. A. J Am Chem Soc 1994, 116, 2493–2499.
60. Hobza, P.; Šponer, J.; Polasek, M. J Am Chem Soc 1995, 117, 792–798.
61. Florian J., Leszczynski J. J Biomol Struct Dynam 1995, 12, 1055–1062.
62. Šponer, J.; Leszczynski, J.; Hobza, P. J Phys Chem 1996, 100, 1965–1971.
63. Šponer, J.; Leszczynski, J.; Hobza, P. J Phys Chem 1996, 100, 5590–5596.
64. Šponer, J.; Gabb, H. A.; Leszczynski, J.; Hobza, P. Biophys J 1997, 73, 76–87.
65. Šponer, J.; Leszczynski, J.; Hobza, P. J Phys Chem A 1997, 101, 9489–9495.
66. Meyer, M.; Suhnel, J. J Biomol Struct Dynam 1997, 15, 619–624.

67. Barsky D.; Kool, E. T.; Colvin, M. E. J Biomol Struct Dynam 1999, 16, 1119–1134.
68. Šponer, J.; Burda, J. V.; Mejzlik, P.; Leszczynski, J.; Hobza, P. J Biomol Struct Dynam 1997, 14, 613–628.
69. Gu, J.; Leszczynski, J. J Phys Chem A 2000, 104, 1889–1904.
70. Gu, J.; Leszczynski, J. J Phys Chem A 2000, 104, 6308–6314.
71. Meyer, M.; Steinke, T.; Brandl, M.; Suhnel, J. J Comput Chem 2001, 22, 109–124.
72. Zhanpeisov, N. U.; Leszczynski, J. J Phys Chem A 1998, 102, 6167–6172.
73. Zhanpeisov, N. U.; Leszczynski, J. J Phys Chem B 1998, 102, 9109–9118.
74. Zhanpeisov, N. U.; Šponer, J.; Leszczynski, J. J Phys Chem A 1998, 102, 10374–10379.
75. Moroni, F.; Famulari, A.; Raimondi, M. J Phys Chem A 2001, 105, 1169–1174.
76. Kryachko, E. S.; Volkov, S. N. Int J Quantum Chem 2001, 82, 193–204.
77. Brandl, M.; Meyer, M.; Suhnel, J. J Am Chem Soc 1999, 121, 2605–2606.
78. Brandl, M.; Meyer, M.; Suhnel, J. J Phys Chem A 2000, 104, 11177–11187.
79. Bondarev, D. A.; Skawinski, W. J.; Venanzi, C. A. J Phys Chem 2000, 104, 815–822.
80. Šponer, J.; Hobza, P. J Phys Chem A 2000, 104, 4592–4597.
81. Šponer, J.; Hobza, P. Chem Phys Lett 1997, 267, 263–270.
82. Hobza, P.; Šponer, J.; Leszczynski, J. J Phys Chem B 1997, 101, 8038–8039.
83. Šponer, J.; Hobza, P. J Am Chem Soc 1996, 114, 709–714.
84. Luisi, B.; Orozco, M.; Šponer, J.; Luque, F. J.; Shakked, Z. J Mol Biol 1988, 279, 1123–1136.
85. Vlieghe, D.; Šponer, J.; van Meervelt, L. Biochemistry 1999, 38, 16443–16451.
86. Šponer, J.; Florian, J.; Leszczynski, J.; Hobza, P. J Biomol Struct Dynam 1996, 13, 827–833.
87. Šponer, J.; Šponer, J. E.; Leszczynski, J. J Biomol Struct Dynam 2000, 17, 1087–1096.
88. Šponer, J.; Šponer, J. E.; Gorb, L.; Leszczynski, J.; Lippert, B. J Phys Chem A 1999, 103, 11406–11413.
89. Hobza, P.; Šponer, J.; Cubero, E.; Orozco, M.; Luque, F. J. J Phys Chem B 2000, 104, 6286–6292.
90. Florian J., Leszczynski J. J Am Chem Soc 1996, 118, 3010–3017.
91. Colson, A.-O.; Sevilla, M. D. In Theoretical Computational Chemistry, Vol. 8, Computational Molecular Biology; Leszczynski, J., Ed.; Elsevier: Amsterdam, 1999; pp 245–275.
92. Colson, A.-O.; Sevilla, M. D. Int J Radiat Biol 1995, 67, 627–645, and references therein.
93. Bertran, J.; Oliva, A.; Rodriguez-Santiago, L.; Sodupe, M. J Am Chem Soc 1998, 120, 8159–8167.
94. Brameld, K.; Dasgupta, S.; Goddard, W. A., III. J Phys Chem B 1997, 101, 4851–4859.
95. Pundlik, S. S.; Gadre, S. R. J Phys Chem B 1997, 101, 9657–9662.
96. Gadre, S. R.; Pundlik, S. S.; Limaye, A. C.; Rendell, A. P. Chem Commun 1998, 573–574.
97. Raimondi, M.; Famulari, A.; Gianinetti, E. Int J Quantum Chem 1999, 74, 259–269.
98. Famulari, A.; M., Moroni, F.; Sironi, M.; Raimondi, M, Comput Chem 2000, 24, 351–357.
99. Kawahara S.; Wada, T.; Kawauchi, S.; Uchimaru, T.; Sekine, M. J Phys Chem A 1999, 103, 8516–8523.
100. Kawahara, S.-I.; Uchimaru, T.; Sekine, M. Theochem—J Mol Struct 2000, 530, 109–117.
101. Kawahara, S.;Uchimaru, T. Chem Phys Phys Chem 2000, 2, 869–2872.
102. Guerra, C. F.; Bickelhaupt, F. M. Ang Chem Int Ed 1999, 38, 2942–2945.
103. Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. Chem Eur J 1999, 5, 3581–3594.
104. Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. J Am Chem Soc 2000, 122, 4117–4128
105. Fellers R. S.; Barsky, D.; Gygi, F.; Colvin, M. Chem Phys Lett 1999, 312, 548–555.
106. Barsky, D.; Colvin, M. E. J Phys Chem A 2000, 104, 8570–8576.
107. Czernek J. J Phys Chem A 2001, 105, 1357–1365.
108. Saito, I.; Nakamura, T.; Nakatani, K.; Yoshioka, Y.; Yamaguchi, K.; Sugiyama, H. J Am Chem Soc 1998, 120, 12686–12687.
109. Leulliot, N.; Ghomi, M.; Jobic, H.; Bouloussa, O.; Baumruk, V.; Couloumbeau, C. J Phys Chem A 1999, 103, 8716–8724.
110. Hocquet, A.; Ghomi, M. Phys Chem Chem Phys 2000, 2, 5351–5353.
111. Burda, J. V.; Šponer, J.; Hobza, P. J Phys Chem 1996, 100, 7250–7255.
112. Burda, J. V.; Šponer, J.; Leszczynski, J.; Hobza, P. J Phys Chem B 1997, 101, 9670–9677.
113. Šponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. J Phys Chem A 1998, 102 5951–5957.
114. Šponer, J.; Sabat, M; Burda, J. V.; Leszczynski, J.; Hobza, P. J Biomol Struct Dynam 1998, 16, 139–144.
115. Šponer, J.; Burda, J. V.; Leszczynski, J.; Hobza, P. J Biomol Struct Dynam 1999, 17, 61–79.
116. Šponer, J.; Sabat, M.; Gorb, L.; Leszczynski, J.; Hobza, P. J Phys Chem B 2000, 104, 2224–2232.
117. Pelmenchikov, A.; Zilberberg, I.; Leszczynski, J.; Famulari, A.; Sironi, M.; Raimondi, M. Chem Phys Lett 1999, 314, 496–500.
118. Burda, J. V.; Šponer, J.; Leszczynski, J. J Biol Inorg Chem 2000, 5, 178–188.
119. Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P.; Lippert, B. J Biol Inorg Chem 1999, 4, 537–545.
120. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, D. S., Jr.; Weiner, P. J Am Chem Soc 1984, 106, 765–784.
121. Reed, A. E.; Weinstock, R. B.; Weinhold, F. J Chem Phys 1985, 73, 735–746.
122. Bader, R. F. W. Chem Rev 1991, 91, 893–928.
123. Del Bene, J. E. J Phys Chem 1983, 87, 367–371.

124. Chandra, A. K.; Nguyen, M. T.; Uchimaru, T.; Zeegers-Huyskens, T. J Mol Struct 2000, 555, 61–66.
125. Florian, J.; Baumruk, V.; Leszczynski, J. J Phys Chem 1996, 100, 5578–5589.
126. Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. J Phys Chem 1998, 102, 6010–6016.
127. Chandra, A. K.; Nguyen, M. T.; Uchimaru, T.; Zeegers-Huyskens, T. J Phys Chem A 1999, 103, 8853–8860.
128. Podolyan, Y.; Gorb, L.; Leszczynski, J. J Phys Chem A 2000, 104, 7346–7352.
129. Greco, F.; Liguori, A.; Sindona, G.; Ucella, N. J Am Chem Soc 1990, 112, 9092–9096.
130. Lias, S. G.; Liebman, J. F.; Levin, R. D. J Phys Ref. Data 1984, 13, 695–808.
131. Meot-Ner (Mautner) M. J Am Chem Soc 1979, 101, 2396–2401.
132. Liguori, A.; Napoli, A.; Sindona, G. J Mass Spectr 2000, 35, 139–144.
133. Leszczynski, J. In Encyclopedia of Computational Chemistry; Schleyer, P. v. R., Ed.; John Wiley: Chichester, 1998; pp 2951–2960.
134. Leszczynski, J. J Phys Chem 1992, 96, 1649–1653.
135. Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. J Mol Struct (Theochem) 1995, 331, 147–154.
136. Nowak, M. J.; Lapinski, L.; Kwiatkowski, J. S.; Leszczynski, J. J Phys Chem 1996, 100, 3234–3257.
137. Ha, T. K., Keller, M. J., Gunde, R.; Gunthard, H. H. J Mol Struct (Theochem) 1996, 364, 161–181.
138. Szczepaniak, K.; Szczesniak, M. J Mol Struct 1987, 156, 29–49.
139. Dolgounitcheva, Zakrzewski, V. G.; Ortiz, J. V. J Am Chem Soc 2000, 122, 12304–12309.
140. Leszczynski, J. J Phy Chem A 1998, 102, 2357–2362.
141. Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; KuBulat, K.; Person, W. B. J Am Chem Soc 1988, 110, 8319–8330.
142. Gorb, L.; Leszczynski, J. J Am Chem Soc 1998, 120, 5024–5032.
143. Aleman, C. Chem Phys 2000, 253, 13–29.
144. Kobaysashi, R. J Phys Chem A 1998, 102, 10831–10817.
145. Colominas, C.; Luque, F. J.; Orozco, M. J Am Chem Soc 1996, 118, 6811–6821.
146. Young, P.; Green, V. S.; Hillier, I. H.; Burton, N. A. Mol Phys 1993, 80, 503–513.
147. Alhambra, C.; Luque, F. J.; Estelrich, J.; Orozco, M. J Org Chem 1995, 60, 969–976.
148. Leszczynski, J. J Mol Struct (Theochem) 1994, 311, 37–43.
149. Leszczynski, J.; Sponer, J. J Mol Struct (Theochem) 1996, 388, 237–243.
150. Orozco, M.; Hernandez, B.; Luque, F. J J Phys Chem B 1998, 102, 5228–5223.
151. Clarke, M. J. J Am Chem Soc 1978, 100, 5068–5077.
152. Lippert, B.; Schöllhorn, H.; Thewalt, U. J Am Chem Soc 1986, 108, 6616–6621.
153. Pichierri, F.; Holthenrich, D.; Zangrando, E.; Lippert, B.; Randaccio, L. J Biol Inorg Chem 1996, 1, 439–445.
154. Müller, J.; Zangrando, E.; Pahlke, N.; Freisinger, E.; Randaccio, L.; Lippert, B. Chem Eur J 1998, 4, 397–405.
155. Zamora, F.; Kunsman, M.; Sabat, M.; Lippert, B. Inorg Chem 1997, 36, 1583–1587.
156. Arpalahti, J.; Klika, K. D. Eur J Inorg Chem 1999, 8, 1199–1121.
157. Day, E. F.; Crawford, C. A.; Folting, K.; Dunbar, K. R.; Christon, G. J Am Chem Soc 1994, 116, 9449–9340.
158. Velders, A. H.; van der Geest, B.; Kooijman, H.; Spek, A. L.; Haasnoot, J. G.; Reedijk, J. Eur J Inorg Chem 2001, 2, 369–372.
159. Bludsky, O.; Šponer, J.; Leszczynski, J.; Spirko, V.; Hobza, P. J Chem Phys 1996, 105, 11042–11050.
160. Šponer, J.; Leszczynski, J.; Hobza, P. J Mol Struct (Theochem), 2001, 573, 43–53.
161. Trygubenko, S. A.; Bogdan, T. V.; Samijlenko, S. P.; Hovorun, D. M.; Kabelac, M.; Hobza, P. J Phys Chem, submitted.
162. Prive, G. G.; Heinemann, U.; Chandrasegaran, S.; Kan, L.-S.; Kopka, M. L.; Dickerson, R. E. Science 1987, 38, 498–504.
163. Ennifar, E.; Yusupov, N.; Walter, P.; Marquet, R.; Ehresmann, B.; Ehresmann, C.; Dumas, P. Struct Fold Des 1999, 7, 1439–1449.
164. Ortiz-;Lombardia, M.; Gonzalez, A.; Eritja, R.; Aymami, J.; Azorin, F.; Coll, M. Nature Struct Biol 1999, 6, 913–917.
165. Shepard, W.; Cruse, W. B. T.; Fourme, R.; de la Fortelle, E.; Prange, T. Structure Fold Des 1998, 6, 849–861.
166. Spackova, N.; Berger, I.; Sponer, J. J Am Chem Soc 2000, 122, 7564–7572.
167. Štefl, R.; Koca, J. J Am Chem Soc 2000, 122, 5025–5033.
168. Strash, D.; Schlick, T. J Mol Biol 2000, 301, 643–663.
169. Lankas, F.; Cheatham, T. E., III; Hobza, P.; Langowski, J.; Špackova, N.; Šponer, J., submitted.
170. Špackova, N.; Ryjacek, F.; Lankas, F.; Hobza, P.; Šponer, J., submitted.
171. Šponer, J. E.; Glahe, F.; Leszczynski, J.; Lippert, B.; Šponer, J. Inorg Chem, 40, 2001, 3269–3278.
172. Manor, J.; Orozco, M.; Hobza, P.; Šponer, J.; Luque, F. J. J Phys Chem B 2001, 105, 6051–6060.
173. Leontis, N. B.; Westhof, E. Quart Rev Biophys 1998, 31, 399–455.
174. Gee, J. E.; Revankar, G.; Rao, T.; Hogan, M. E. Biochemistry 1995, 34, 2042–2048.
175. Olivas, W. M.; Maher, L. J. Nucleic Acid Res 1995, 23, 1936–1941.
176. Rao, T. S.; Durland, R. H.; Seth, D. M.; Myrick, M. A.; Bodepudi, V.; Revankar, G. R. 1995, 34, 765–772.
177. Marathias, V. M.; Sawicki, M. J.; Bolton, P. H. Nucleic Acid Res 1999, 27, 2860–2867.
178. Štefl, R.; Špackova, N.; Berger, I.; Koca, J.; Šponer, J. Biophys J 2001, 80, 455–468.
179. Cubero, E.; Laughton, C. A.; Luque, F. J.; Orozco, M. J Am Chem Soc 2000, 122, 6891–6899.

180. Cubero, E.; Sherer, E. C.; Luque F. J.; Orozco, M.; Laughton, C. A. J Am Chem Soc 1999, 121, 8653–8654.
181. Florian, J.; Goodman, M. F.; Warshel, A. J Phys Chem B 2000, 104, 10092–10099.
182. Schneider, C.; Brandl, M.; Suhnel, J. J Mol Biol 2001, 305, 659–667.
183. Spackova, N.; Berger, I.; Šponer, J. J Am Chem Soc 1999, 121, 5519–5534.
184. Špackova, N.; Berger, I.; Šponer, J. J Am Chem Soc 2001, 123, 3295–3307.
185. Chowdhury, S.; Bansal, M. J Biomol Struct Dynam 2000, 18, 11–28.
186. Read, M. A.; Neidle, S. Biochemistry 2000, 39, 13422–13432.
187. Špirko, V.; Šponer, J.; Hobza, P. J Chem Phys 1997, 106, 1472–1479.
188. Poltev, V. I.; Shulyupina, N. V. J Biomol Struct Dynam 1986, 3, 739–765.
189. Zhurkin, V. B.; Poltev, V. I.; Florentev, V. L. Mol Biol (USSR) 1980, 14, 1116–1140.
190. Šponer, J.; Kypr, J. J Biomol Struct. Dynam 1993, 11, 277–292.
191. Gresh, N.; Garmer, D. R. J Comput Chem 1996, 17, 1481–1495.
192. Gresh, N.; Šponer, J. J Phys Chem B 1999, 103, 11415–11427.
193. Sigel, R. K. O.; Lippert, B. Chem Commun 1999, 2167–2168.
194. Sigel, R. K. O.; Freisinger, E.; Lippert, B. J Biol Inorg Chem 2000, 5, 287–299.
195. Garmer, D. R.; Gresh, N. J Am Chem Soc 1994, 116, 3556–3567.
196. Bock, C. W.; Katz, A. K.; Glusker, J. P. J Am Chem Soc 1995, 117, 3754–3763.
197. McFail-Isom, L.; Shui, X. Q.; Williams, L. D. Biochemistry 1998, 37, 17105–1711.
198. Šponer, J. E.; Glahe, F.; Leszczynski, J.; Lippert, B.; Šponer, J., J Phys Chem B 2001, 105, 12171–12179.
199. Kozelka, J.; Savinelli, R.; Berthier, G.; Flament, J. P.; Lavery, R. J Comput Chem 1993, 14, 45–53.
200. Bugg, C. E.; Thomas, J. M.; Sundaralingam, M.; Rao, S. T. Biopolymers 1971, 10, 175–210.
201. Hunter, C. A. J Mol Biol 1993, 230, 1025–1054.
202. Fulscher, M. P.; Serrano-Andres, L.; Roos B. O. J Am Chem Soc 1997, 119, 6168–6176.
203. Shukla, M. K.; Mishra S. K.; Kumar, A.; Mishra, P. C. J Comput Chem 2000, 21, 826–846.
204. Shukla, M. K.; Leszczynski, J. Int J Quantum Chem 2000, 77, 240–254.
205. Broo, A.; Holmen, A. J Phys Chem A 1997, 101, 3589–3600.
206. Yoshioka, Y.; Kitagawa, Y.; Takano, Y.; Yamaguchi, K.; Nakamura, T.; Saito, I. J Am Chem Soc 1999, 121, 8712–8719.
207. Prat, F.; Houk, K. N.; Foote, C. S. J Am Chem Soc 1998, 120, 845–846.
208. Voityuk, A. A.; Jortner, J.; Bixon, M.; Rosch, N. J Chem Phys 2001, 114, 5614–5620.
209. Dannenberg, J. J.; Tomasz, M. J Am Chem Soc 2000, 122, 2062–2068.
210. Barfield, M.; Dingley, A. J.; Feigon, J.; Grzesiek, S. J Am Chem Soc 2001, 123, 4014–4022.