J Mol Model (2011) 17:3265–3274
DOI 10.1007/s00894-011-1002-y

ORIGINAL PAPER

Localization and anharmonicity of the vibrational modes for GC Watson–Crick and Hoogsteen base pairs

Attila Bende · Diana Bogdan · Cristina M. Muntean · Cristian Morari

Received: 5 November 2010 / Accepted: 28 January 2011 / Published online: 4 March 2011

Abstract We present an ab initio study of the vibrational properties of cytosine and guanine in the Watson–Crick and Hoogsteen base pair configurations. The results are obtained by using two different implementations of the DFT method. We assign the vibrational frequencies to cytosine or to guanine using the vibrational density of states. Next, we investigate the importance of anharmonic corrections for the vibrational modes. In particular, the unusual anharmonic effect of the H^+ vibration in the case of the Hoogsteen base pair configuration is discussed.

Keywords DFT · DNA · Watson–Crick · Hoogsteen · Vibrational density of states · Anharmonic frequencies

Introduction

The formation of DNA base pairs plays a crucial role in the realization of the main role of DNA: the storage and replication of genetic information [1]. Vibrational properties are among the most important observables used to investigate this process [2]. For instance, it is well known that the low-frequency vibrational modes of large molecules are determined by the global shape of the multidimensional potential energy. Therefore, the vibrational frequencies of low-frequency vibrational modes of DNA can be highly sensitive to the geometric structure of the molecule [3, 4]. On the other hand, Raman spectroscopy has proven its efficiency as a tool for the detection of Watson–Crick (W-C) to Hoogsteen GC structural transitions [5]. The structure of double-helical poly(dG-dC)·poly(dG-dC) was investigated at various pH values by Raman spectroscopy. These investigations emphasized the presence of Hoogsteen base pairing in DNA structure. Hoogsteen pairing can be accommodated in the double helix when the cytosine group is protonated and the sugar-guanine conformer has adopted a C2′-*endo/syn* conformation. It has been shown that this antiparallel-stranded Hoogsteen base-paired structure can be maintained under varying conditions, balancing the decrease in pH with an increased salt concentration [5]. This conformational change enables hydrogen bonding of N7 of guanine and N3 of cytosine in a Hoogsteen base pair [6].

Raman data on chromosomes and calf-thymus DNA also provide evidence for the reversible formation of protonated Hoogsteen GC base pairs under acidic conditions, in full agreement with NMR, optical rotatory dispersion and spectrophotometric pH-titration studies of DNA (see [6] and references therein). Also, a perfectly matched parallel Hoogsteen paired duplex (at pH 4) and a triple-helical structure with third strand bases that are Hoogsteen paired to the purine bases of the duplex were observed by IR spectroscopy [7].

From a theoretical perspective, a detailed description of the vibrational modes for GC base pairs can be obtained from ab initio calculations. We note here the assignment of vibrational properties [8, 9] and the simulation of Raman spectra [10]. The interaction energy between base pairs has been investigated in detail by

Electronic supplementary material The online version of this article (doi:10.1007/s00894-011-1002-y) contains supplementary material, which is available to authorized users.

A. Bende · D. Bogdan · C. M. Muntean · C. Morari (✉)
Molecular and Biomolecular Physics Department,
National Institute for Research and Development
of Isotopic and Molecular Technologies,
Donath Street, No. 65-103,
400293 Cluj-Napoca, Romania
e-mail: cristim@itim-cj.ro

Jurečka et al. [11], considering second-order Møller–Plesset perturbation theory level at the resolution of the identity approximation (RI-MP2) with a complete basis set extrapolation technique. They obtained a value of −28.8 kcal/mol for the interaction energy. Using symmetry-adapted perturbation theory (SAPT) [12, 13], Hesselmann [14] obtained a value of −30.5 kcal/mol for the GC base pair interaction energy. The planar structure of the GC base pair in the Hoogsteen configuration is stabilized by two H-bonds. The strength of the intermolecular interaction was described by Han et al. [15] at the B3LYP/6-31+G(d,p) DFT level of theory. They obtained −35.5 kcal/mol for the $G^{+}C$ configuration. Comparing the strengths of the intermolecular interactions for the Watson–Crick and Hoogsteen configurations, even though they were not obtained at the same level of theory, one can conclude that the Hoogsteen configuration presents greater stability due to the larger electrostatic attraction induced by the presence of the proton between the base pairs.

The localization of the vibrational modes and the anharmonic corrections yields a rich source of information revealing the properties of DNA, such as vibrational energy transfer [16]. The effect of anharmonic corrections in normal-mode vibrational analysis is very important for obtaining good agreement between theoretical and experimental results [17–19]. At the same time, the vibrational shifts derived from the anharmonic calculations show good agreement with the trends found in the experiments [20].

Our goal in the work reported here was to investigate, by ab initio methods, the localization properties of the vibrational modes in GC base pairs as well as their anharmonicities. To this end, we used two independent implementations of DFT and two different exchange-correlation functionals. The localization of the vibrational modes was described in terms of the vibrational density of states. A detailed analysis of the anharmonic corrections to the vibrational frequencies was performed using the VSCF method. The consequences of our results from an experimental point of view are also discussed below.

Computational models

DFT calculations

We carried out two independent calculations of the vibrational properties of GC base pairs of DNA in both the Watson–Crick and the Hoogsteen geometry configurations (see Figs. 1 and 2) using the *Siesta* [21, 22] and *Gaussian* codes [23]. *Siesta* employs pseudopotentials and expands the wavefunctions of valence electrons using linear combinations of numerical atomic orbitals. Periodic boundary

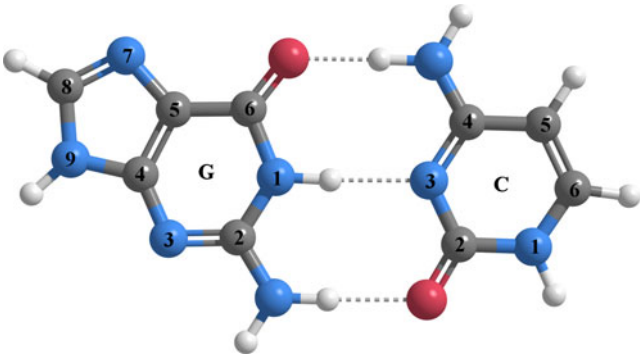

Fig. 1 The equilibrium geometry structure of the guanine–cytosine Watson–Crick base pair of DNA obtained at the B3PW91 level of theory using the TZVP basis set

conditions were imposed on the system; therefore, we used a large supercell (i.e., 30×30×30 $Å^3$) in order to get rid of the interactions between periodically repeated images of the molecule. We employed a double-zeta polarized (DZP) basis set with an energy cutoff of 150 meV for all atoms. For the exchange and correlation part, we used the newly developed functional of Dion et al. [24, 25], which includes van der Waals (vdW) effects. The second set of calculations were carried out using the B3PW91 [26, 27] exchange-correlation (XC) functional together with TZVP [28] basis set implemented in the *Gaussian* 03 software suite [23]. The B3PW91 XC functional was chosen based on its good accuracy in describing the intermolecular interaction energies of H-bonds between the nucleic acids presented by Guerra et al. [29]. All results were corrected for basis set superposition error (BSSE). To represent the normal modes graphically, we used the Gabedit program [30].

Besides the harmonic vibrational frequencies, anharmonic corrections based on the VSCF theory with second-order perturbation expansion [31, 32] were also computed for the second case. VSCF calculations were performed starting from the analytical second derivatives of the energy at the given theoretical level. The third and semidiagonal fourth derivatives needed for second-order perturbation [33]

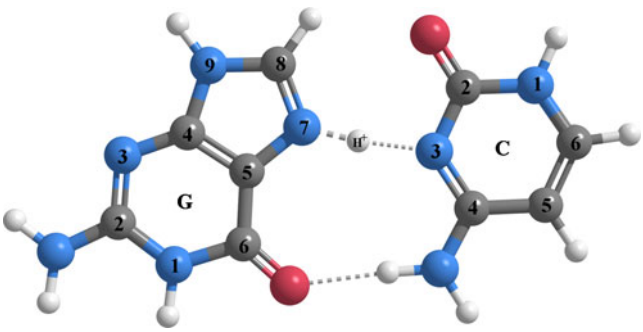

Fig. 2 The equilibrium geometry structure of the guanine–cytosine Hoogsteen base pair of DNA obtained at the B3PW91 level of theory using the TZVP basis set

can be effectively computed by a finite difference approach which scales linearly with the number of normal modes. To evaluate third derivatives of the energy with respect to normal coordinates, we performed numerical differentiations of analytical Hessian matrices at geometries displaced by small increments δq from the reference geometry [32]:

$$f_{ijk} = \frac{1}{3}\left(\frac{f_{jk}(+\delta q_i) - f_{jk}(-\delta q_i)}{2\delta q_i} + \frac{f_{ki}(+\delta q_j) - f_{ki}(-\delta q_j)}{2\delta q_j} + \frac{f_{ij}(+\delta q_k) - f_{ij}(-\delta q_k)}{2\delta q_k}\right). \quad (1)$$

In our calculations, δq was taken to be 0.01 Å (implicit value from [32]). The semidiagonal fourth-order derivatives of energy required to compute rovibrational energies via second-order perturbation theory are given as:

$$f_{ijkk} = \frac{f_{ij}(+\delta q_k) - 2f_{ij}(0) + f_{ij}(-\delta q_k)}{\delta q_k^2} \quad (2)$$

$$f_{iikk} = \frac{1}{2}\left(\frac{f_{ii}(+\delta q_k) - 2f_{ii}(0) + f_{ii}(-\delta q_k)}{\delta q_k^2} + \frac{f_{kk}(+\delta q_i) - 2f_{kk}(0) + f_{kk}(-\delta q_i)}{\delta q_i^2}\right). \quad (3)$$

For nonlinear molecules, these computations require at most the Hessian matrices at $6N-11$ different points, where N is the number of atoms in the molecule. Based on these cubic (f_{ijk}) and quartic (f_{ijkl}) force constants, diagonal (x_{ii}) and off-diagonal (x_{ij}) vibrational anharmonic constants could be obtained for the asymmetric top molecules as [34]:

$$x_{ii} = \frac{1}{16} f_{iiii} - \frac{1}{32}\sum_j f_{iij}^2\left[\frac{1}{2\omega_i + \omega_j} - \frac{1}{2\omega_i - \omega_j} + \frac{4}{\omega_j}\right] \quad (4)$$

while

$$x_{ij} = \frac{1}{4} f_{iijj} - \frac{1}{4}\sum_k \frac{f_{iik} f_{kjj}}{\omega_k} - \frac{1}{8}\sum_k f_{ijk}^2\left[\frac{1}{\omega_i+\omega_j+\omega_k} - \frac{1}{\omega_i+\omega_j-\omega_k} + \frac{1}{\omega_i-\omega_j+\omega_k} - \frac{1}{\omega_i-\omega_j-\omega_k}\right] + \left(A_e\left(\zeta_{i,j}^{(a)}\right)^2 + B_e\left(\zeta_{i,j}^{(b)}\right)^2 + C_e\left(\zeta_{i,j}^{(c)}\right)^2\right)\left(\frac{\omega_i+\omega_j}{\omega_j+\omega_i}\right) \quad (5)$$

where ω_i are the harmonic frequencies, A_e, B_e, and C_e are the equilibrium rotational constants, and $\zeta_{i,j}$ are the Coriolis zeta constants. Considering these vibrational anharmonic constants, one can easily obtain anharmonic corrections Δ_i and fundamental frequencies v_i:

$$v_i = \omega_i + \Delta_i = \omega_i + 2x_{ii} + \frac{1}{2}\sum_{j,j\neq j} x_{ij}. \quad (6)$$

Data analysis

Traditionally, the analysis of the vibrational properties of a molecule is rationalized in terms of the normal modes of vibration [35, 36]. This is a powerful approach, allowing for an accurate description of the vibrational eigenvectors in terms of geometrical representations. The disadvantage of it is that the analysis of large amount of data (i.e., all vibrational modes of relatively large molecules) involves the comparison of $3N-3$ geometrical representations of the vibrational eigenmodes, where N represents the number of atoms in the molecule.

The vibrational density of states (VbDOS) provides an alternative way to analyze the vibrational properties of a molecule, and is more suitable for a global comparison. For the present study, we followed the straightforward definition of the VbDOS given in [37]:

$$I_\Gamma(\omega) = \sum_{\chi\in\Gamma}\sum_i \left|A_i^\chi(\omega)\right|^2 \quad (7)$$

where $|A_i^\chi(\omega)|$ is the ith component of the vibrational eigenvector (i=1, 2, 3) of the modes with vibrational frequency ω. Here, Γ refers to a group of atoms that are indexed by χ.

We used Eq. 7 to project the VbDOS over the cytosine and guanine. To this end, we searched for the atom with the largest vibrational amplitude (in each vibrational mode) $A_{\max}(\omega)$. We set up a list including all atoms with vibrational amplitudes $A^\chi(\omega) > 0.25A(\omega)$. By inspecting this list, it was possible to establish the vibrational modes located on cytosine or guanine.

Note that we used an arbitrary factor of 0.25 as a cutoff to assign the vibrational amplitude. Note that this is equivalent to a limitation on the VbDOS contribution to about 6% per atom. For plotting purposes, we use a smearing function with a Lorentzian shape and a width of 5 cm^{-1}, as well as a discretization step of 1 cm^{-1}.

Results

The guanine–cytosine binary system in the Watson–Crick configuration (Fig. 1) has 29 atoms and presents 81 normal mode vibrations, of which 6 have intermolecular character. Guanine has a C=O group at C6 that acts as the hydrogen acceptor, while the group at N1 and the NH_2 group at C2 act as the hydrogen donors. In cytosine, the NH_2 group acts at C4 as the hydrogen donor, and the C2 carbonyl and the N3 amine act as the hydrogen-bond acceptors, which binds the cytosine to the guanine through three H-bonds.

The guanine–cytosine binary system in the Hoogsteen configuration has 30 atoms and 84 normal mode vibrations,

of which 6 have intermolecular character. In this geometric assemblage, guanine has a C=O group at C6 that acts as the hydrogen acceptor, while the N1 is the special hydrogen (proton) donor. In cytosine, the NH_2 group acts at C4 as the hydrogen donor, and the N3 amine acts as the hydrogen-bond acceptor, binding the cytosine to the guanine through two H-bonds. The guanine–cytosine Hoogsteen binary system is explicitly stabilized by the proton bridge (Fig. 2) between N7 of guanine and N3 of cytosine. As can be seen from Fig. 4, the proton has two equilibrium positions with similar total energies. This explains the fact that we found different equilibrium positions in the two simulations, each belonging to one of these minima. To be precise, in our *Gaussian* calculations we found the following distances: N7−H=1.13 Å and N3···H=1.50 Å, while in *Siesta* we found: N7···H=1.58 Å and N3−H=1.09 Å.

Localization of the vibrational modes

The results of our mode localization analysis, as described in "Data analysis," are given in Fig. 3. We plot the interval 0−1750 cm^{-1}. By inspecting it, we can immediately see that there is qualitative agreement between the two DFT calculations. The complete list of the vibrational modes localized in cytosine or guanine is given in the "Electronic supplementary material" (ESM).

As can be seen from Table 1, both calculations agree in the fact that for the Hoogsteen base pair, a slightly smaller fraction of the total number of modes are localized on the bases compared to the Watson–Crick case. This result can be explained by the fact that in the Hoogsteen base pairs the anharmonic corrections are larger, which may indicate, for instance, that the thermal transport mediated by vibrational coupling can be very different from that in the Watson–Crick case. Nevertheless, the present results are not sufficient to support this ansatz; further investigations are needed to clarify this issue.

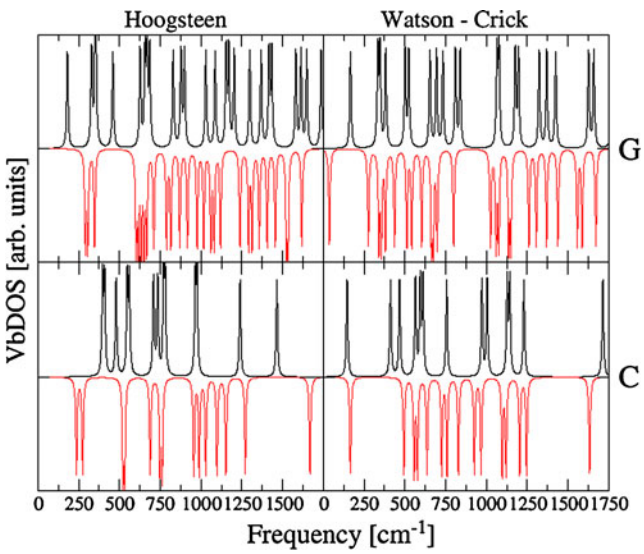

Fig. 3 Localization of the VbDOS for cytosine and guanine in the Hoogsteen (*left column*) and Watson–Crick base pairs (*right column*). *Black lines* represent the *Gaussian* calculation; *red lines* represent the *Siesta* calculation

Table 1 Statistical analysis of the localization of vibrational modes on cytosine and guanine. For each of the investigated models, the number of localized modes and their percentages of the total number of modes are shown

Base	Hoogsteen		W-C		
C	17	(19%)	18	(22%)	*Gaussian*
G	30	(34%)	26	(35%)	
C	18	(20%)	19	(23%)	*Siesta*
G	32	(36%)	29	(36%)	

By inspecting Fig. 3, we can see that both calculations agree with the fact that there are relatively wide frequency ranges (i.e., more than 100 cm^{-1}) where the vibrations are located only on the guanine for both configurations. For Hoogsteen, this is the case for the intervals 550−650 cm^{-1} and 770−900 cm^{-1}, where only guanine is vibrating. For Watson–Crick, the localization on guanine occurs in the 250−350 cm^{-1} range. Finally, at 1250−1600 cm^{-1}, vibrations are localized on guanine for both configurations (i.e., Hoogsteen and Watson–Crick). There is only one exception: at ~1400 cm^{-1} in the Hoogsteen configuration with the *Gaussian* calculation. All normal-mode frequencies of the studied systems are presented in tabular format in the ESM.

In order to check the accuracy of our calculations, we compare them with the experimental results [5]. Two Raman markers for the Watson–Crick configuration are located at 681 cm^{-1} (cytosine) and 783 cm^{-1} (cytosine backbone). In the Hoogsteen configuration, they are present at 675 cm^{-1} and 786 cm^{-1}. The band at 1259 cm^{-1} is strong for the Hoogsteen configuration but suppressed in the Watson–Crick configuration. Also, an important marker for both configurations occurs at 1488 cm^{-1}. From our analysis (see also the ESM), we found vibrational bands of guanine in the Watson–Crick configuration at 674.7 and 698.4 cm^{-1} (*Siesta* calculations) and at 694.1 cm^{-1} (*Gaussian* calculations). This last frequency was corrected to 686 cm^{-1} using anharmonic calculations (see below). We conclude that the values at 698.4 cm^{-1} (*Siesta*) and 694.1 cm^{-1} (*Gaussian*)—which are remarkably close—correspond to the experimental frequency of 681 cm^{-1}. A similar discussion occurs for the guanine in the Hoogsteen configuration: *Siesta* values close to this band are 657.0 and 710.2 cm^{-1}. By inspecting the *Gaussian* values, we found a value of 657 cm^{-1}, which is then corrected to 687.2 cm^{-1}. Again we note the excellent agreement between the frequencies computed with the harmonic approximation by the two codes on the one hand, and

between the value corrected for anharmonicity and the experimental value (i.e., 675 cm^{-1} from experiment; 687.2 cm^{-1} from the *Gaussian* calculation with anharmonic correction) on the other hand.

For the markers close to 785 cm^{-1}, the *Siesta* calculation gives 758.9 cm^{-1} (Watson–Crick) and 762.4 cm^{-1} (Hoogsteen). For the same band, the *Gaussian* results indicate 756.4 (835.2) cm^{-1} for Watson–Crick and 780.7 (815.6) cm^{-1} for Hoogsteen (the results in parentheses are corrected for anharmonicity). In this case we again note the good agreement between the two simulations; we also note the large correction for anharmonicity in the case of the Watson–Crick pair (about 80 cm^{-1}). We now move to the experimental band found at 1259 cm^{-1} (cytosine), which is an indicator of the presence of the Hoogsteen configuration [5]. Here the theoretical results are 1271.6 cm^{-1} (*Siesta* calculation) and 1239.8 (1224.9) cm^{-1} (*Gaussian* calculation). We note a relatively large difference (about 50 cm^{-1}) between the results of the two calculations and the relatively small correction for anharmonicity (15 cm^{-1}).

Finally, for the marker band located at 1488 cm^{-1}, we found it difficult to assign the computed frequencies to the experimental ones. This is due to the the differences that were found between the results of the two calculations (i.e., larger than 100 cm^{-1}) on the one hand and between the experimental results and numerical simulations on the other. We can explain this by noting that this mode is assigned to the H-bonding at N7 of guanine, so harmonic approaches are not suitable for this frequency. This point is discussed further below.

We note that the method presented here allows fast and accurate assignment of all vibrational modes to cytosine or to guanine. Therefore, we think it is well suited for the analysis of the vibrational properties of large systems, such as DNA chains.

Anharmonic correction

The importance of performing anharmonic corrections when predicting the vibrational frequencies of DNA base pairs has been pointed out by Špirko [38] and Brauer [39], although a systematic assessment of multiple anharmonic parameters on the basis of ab initio computations is missing from their works. More detailed studies of anharmonic corrections for A–T and G–C Watson–Crick base pairs were performed by Bende [40] and Wang [41]. They consist of the analysis of NH_2, N−H, and C=O stretching vibrational modes that are involved in the multiple H-bonds present in G–C base pairs, the sensitivities of their diagonal and off-diagonal anharmonicities, anharmonic vibrational couplings, as well as the intermolecular effects of the complementary base. Bende has also shown that frequency shifts in the normal-mode vibrations that are obtained from theoretical calculations are a common result of different physical (dimer and anharmonic) and unphysical effects (BSSE). In addition, in order to obtain the BSSE-corrected anharmonic frequencies, the anharmonic and BSSE corrections can be performed independently and then simply summed at the end.

Hereafter, in the Watson–Crick case, we will focus only on those normal modes for which vibrational motion is related to the NH_2, N−H, and C=O groups of guanine or cytosine, situated mostly in the guanine–cytosine molecular plane. The complete list of all 81 normal-mode frequencies and their anharmonic corrections can be found in Table 1 of the ESM. At the same time, it is possible to identify several other normal modes specific to purine and pyrimidine ring vibrational deformations that can significantly disturb the H-bond vibrations. In the frequency domain of 3000−3800 cm^{-1}, ten normal modes of C−H or N−H covalent bond stretching vibrations can be found. From an H-bond point of view, there are three normal modes (ν_5, ν_9 and ν_{10}) that can be associated with the N−H or NH_2 stretching vibrations found in the intermolecular interaction region. The harmonic and anharmonic frequencies, the diagonal and the sum of the off-diagonal anharmonic constants, as well as the assignments of these normal modes are presented in Table 2. Each of the normal mode vibrations could be assigned to one of the three guanine–cytosine binary system's H-bonds, but they are not explicitly located in some H-bonds; they just show larger vibrational amplitudes than the other two vibrations. Compared with the other seven bond-stretching vibrations, these three normal modes present larger anharmonic frequency shifts, as they have larger diagonal anharmonic constants and stronger couplings with the other normal modes. Similar findings were obtained by Del Bene et al. [42] for the N–H stretching vibration. Regarding these three normal modes, the ν_5 vibration is not very strong coupled with the ν_9 and ν_{10} modes, while in the latter two cases ν_9 and ν_{10} show very strong anharmonic coupling ($x_{9,10} = -158.3$ cm^{-1}). In the frequency domain of 1300−1800 cm^{-1}, we selected eight normal modes (ν_{11}, ν_{12}, ν_{13}, ν_{14}, ν_{15}, ν_{16}, ν_{22} and ν_{29}) after a detailed analysis using the Gabedit [30] molecular graphics program. These are mainly N–H or NH_2 group angle-bending vibrations and C=O bond-stretching vibrations, respectively, but small ring deformation motions are also present. Their characteristic feature is that their diagonal anharmonic constants are small (less than 5 cm^{-1}), but they show strong couplings (off-diagonal anharmonic constants) between them. This is true of both the frequency group selected in the present case and the group of previously selected stretching vibrations (for example: $\Sigma_{i=12-16,22,29} x_{11,i} = 14.4\, cm^{-1}$ and $\Sigma_{i=5,9,10} x_{11,i} = -39.3\, cm^{-1}$). In the far-infrared (90−150 cm^{-1}) frequency domain, there are three normal-

Table 2 Selected normal modes for the GC Watson–Crick base pair of DNA, including their harmonic and anharmonic frequencies, their diagonal and the sum of off-diagonal anharmonic constants, as well as their vibrational assignments, as obtained with the *Gaussian* program using the B3PW91/TZVP method

Mode	Harmonic frequency (cm^{-1})	Anharmonic frequency (cm^{-1})	^{x}ii (cm^{-1})	Σx_{ij} (cm^{-1})	Assignment
ν_5	3385.1	3173.1	−98.3	−113.7	$N2^{(G)}$–H stretching
ν_9	3170.7	2903.5	−112.9	−154.3	$N1^{(G)}$–H stretching
ν_{10}	3087.8	2786.2	−152.5	−149.1	$N4^{(C)}$–H stretching
ν_{11}	1770.8	1739.9	−2.2	−28.7	$N1^{(G)}$–H bending
ν_{12}	1739.7	1694.9	−2.7	−42.1	$C4^{(C)}$=O stretching
ν_{13}	1714.5	1678.2	−0.6	−35.7	$N4^{(C)}$–H_2 bending
ν_{14}	1690.4	1650.0	−2.0	−38.4	$N2^{(G)}$–H_2 bending
ν_{15}	1676.4	1634.6	−1.1	−40.7	$N4^{(C)}$–H_2 bending
ν_{16}	1659.0	1629.3	0.0	−29.7	$N1^{(G)}$–H, $N2^{(G)}$–H_2 bending
ν_{21}	1536.8	1498.9	−3.6	−34.3	$N1^{(G)}$–H bending
ν_{22}	1458.2	1429.4	−2.6	−28.8	$N1^{(G)}$–H bending
ν_{29}	1315.7	1273.5	−4.6	−37.6	$C4^{(C)}$–$N3^{(C)}$ stretching
ν_{75}	131.2	126.3	−0.5	−4.4	G···C stretching
ν_{77}	127.7	127.0	−0.3	−0.4	G···C opening
ν_{78}	95.8	93.2	−0.2	−2.4	G···C shearing

mode frequencies (ν_{76}, ν_{76} and ν_{78}) that are characteristic of the intermolecular mode and have in-plane molecular vibrations. In this case, the guanine and the cytosine molecules show rigid motion compared to each other. These normal modes do not show large anharmonic frequency shifts, since these anharmonic effects are the results of different anharmonic couplings with opposite signs. Positive shifts are usually given by coupling with the ν_5, ν_9 and ν_{10} stretching vibrations ($\Sigma_{i=5,9,10} x_{76,i} = +10.7\,cm^{-1}$, $\Sigma_{i=5,9,10} x_{77,i} = +12.5\,cm^{-1}$, and $\sum_{i=5,9,10} x_{78,i} = +13.7 cm^{-1}$), while the negative effect is the contribution of couplings with normal modes that have lower frequency values ($\nu \leq 1800$ cm^{-1}). All of these anharmonic couplings suggest to us that there is an important coupling between the intramolecular N–H covalent bond-stretching vibrations with high frequency values and the intermolecular in-plane rigid vibrations of guanine and cytosine components with low frequencies. Similar findings were obtained for adenine–thymine Watson–Crick base pairs [40]. A triangular matrix with homogeneous and off-diagonal elements for the normal modes discussed above are presented in Table 2 of the ESM.

For the Hoogsteen configuration, the proton bridge shows significant intermolecular interaction energy (−44.8 kcal/mol), much stronger than in the Watson–Crick case (28.2 kcal/mol). Without this proton bridge, the Hoogsteen configuration shows repulsive behavior (ΔE_{int} = +5.9 kcal/mol). Just as we did for the Watson–Crick configuration, we will describe in detail here only those normal-mode vibrations for which the vibrational motion is related to the NH_2, N–H, and C=O groups of guanine or cytosine, and their vibrational motion is situated mostly in the guanine–cytosine molecular plane. The complete list of all 84 normal mode frequencies and their anharmonic corrections can be found in Table 3 of the ESM. The frequency domain of 3000−3800 cm^{-1} again contains ten normal modes characteristic of C–H, N–H and $N–H^+$ covalent bond-stretching vibrations. In the intermolecular interaction region, in this frequency range, only one stretching vibration (ν_7) can typically be found for N–H stretching of the C=O···H–N H-bond. The harmonic and anharmonic frequencies, the diagonal and the sum of off-diagonal anharmonic constants, as well as the assignments of these normal modes are presented in Table 3. Upon comparing them with the other nine normal modes, ν_7 presents the largest homogeneous (diagonal) anharmonic effect (x_{77} = −101.8 cm^{-1}) and the largest total anharmonic frequency shift ($a_7 = \nu_{anharm} - \nu_{harm}$ = −302.1 cm^{-1}). On the other hand, it also shows the strongest vibrational coupling, having 15 normal modes with off-diagonal anharmonic constants larger than 10 cm^{-1}. From an anharmonicity point of view, the frequency range of 1400−1900 cm^{-1} is the most interesting spectral domain. Here, we found 13 frequency lines (ν_{11}–ν_{16}, ν_{18}, ν_{20}, ν_{21}, ν_{23}–ν_{26}), most of them with strong homogeneous anharmonic effects and vibronic couplings. Among these 13 frequencies, the most interesting is the ν_{11} normal mode, which has a huge anharmonic shift (1231 cm^{-1}). To understand this strange behavior, we drew the potential energy curve along the N7 (G)–H bond for both the Watson–Crick and the Hoogsteen configurations (Fig. 4). As seen in Fig. 4, the energy curves show totally different profiles. In the Watson–Crick case, there is only one, well-defined minimum for the N–H

Table 3 Selected normal modes for the GC Hoogsteen base pair of DNA, including their harmonic and anharmonic frequencies, their diagonal and the sum of off-diagonal anharmonic constants, as well as their vibrational assignments, as obtained with the *Gaussian* program using the B3PW91/TZVP method

Mode	Harmonic frequency (cm^{-1})	Anharmonic frequency (cm^{-1})	^{x}ii (cm^{-1})	Σx_{ij} (cm^{-1})	Assignment
ν_7	3422.1	3119.4	−101.8	200.3	$N4^{(C)}$–H stretching
ν_{11}	1858.2	627.2	−265.7	−965.3	$N7^{(G)}$–H^+ stretching
ν_{12}	1784.7	1707.9	−6.6	−70.2	$C6^{(G)}$=O stretching+$N7^{(G)}$–H^+ stretching
ν_{13}	1736.8	1284.1	−31.5	−421.2	$N7^{(G)}$–H^+ bending+$C2^{(C)}$=O stretching
ν_{14}	1692.5	1435.5	−12.6	−244.4	$N4^{(C)}$–H_2 bending
ν_{15}	1687.9	1570.6	−6.2	−111.2	$N7^{(G)}$–H^+ bending
ν_{16}	1680.1	1519.7	−7.5	−152.8	$N2^{(G)}$–H_2 bending+$N4^{(C)}$–H_2 bending ++ $N7^{(G)}$–H^+ stretching
ν_{18}	1648.4	1463.7	−6.1	−178.5	$C4^{(G)}$–$N3^{(G)}$ stretching+$N7^{(G)}$–H^+ stretching
ν_{20}	1581.8	1529.4	−2.5	−49.9	$N7^{(G)}$–$C8^{(G)}$ stretching+$N7^{(G)}$–$^+$ stretching
ν_{21}	1571.4	1517.7	−3.3	−50.3	$C4^{(C)}$–$C5^{(C)}$ stretching
ν_{23}	1519.0	1156.6	−20.8	−341.6	$C4^{(C)}$–$N4^{(C)}$ stretching+$N7^{(G)}$–H^+ stretching
ν_{24}	1465.1	1387.4	−2.2	−75.5	$N1^{(C)}$–$C6^{(C)}$ stretching
ν_{25}	1434.0	1290.9	−6.0	−137.1	$C8^{(G)}$–$N9^{(G)}$ stretching+$N7^{(G)}$–H^+ stretching
ν_{26}	1417.4	1365.0	−1.6	−50.8	$N1^{(G)}$–$C2^{(G)}$ stretching+$N7^{(G)}$–H^+ stretching
ν_{45}	877.6	492.3	−20.6	−364.6	G five-member ring deformation ++ $N7^{(G)}$–H^+ stretching
ν_{46}	829.6	609.6	−15.7	−204.3	G six-member ring deformation ++ $N7^{(G)}$–H^+ stretching
ν_{64}	531.3	475.7	−0.9	−54.7	G five- and six-member ring deformations
ν_{79}	137.3	103.0	−1.3	−33.0	G···C stretching
ν_{80}	102.0	87.3	−0.1	−14.6	G···C shearing
ν_{81}	87.1	14.8	−2.7	−69.6	G···C opening

distance. In the Hoogsteen geometry case, this profile shows two local minima: the energy difference between them is 0.12 kcal/mol and the potential energy barrier is 0.81 kcal/mol. Considering the amount of thermal energy at 300 K (~0.6 kcal/mol), the H^+ can easily overcome this energy barrier and jump from one minimum to another. This peculiar motion of the H^+ can substantially distort the harmonic approximation of the N−H bond stretching vibration. However, we consider that the anharmonic shift is quite large compared with the harmonic frequency, and we believe that further vibrational analysis based on a higher theoretical level (e.g., the Morse potential model) is needed. Similar conclusions were drawn by Brisker et al. [43] upon studying vibrational anharmonicity effects associated with electronic tunneling through molecular bridges. Since the proton can easily roam between the N7

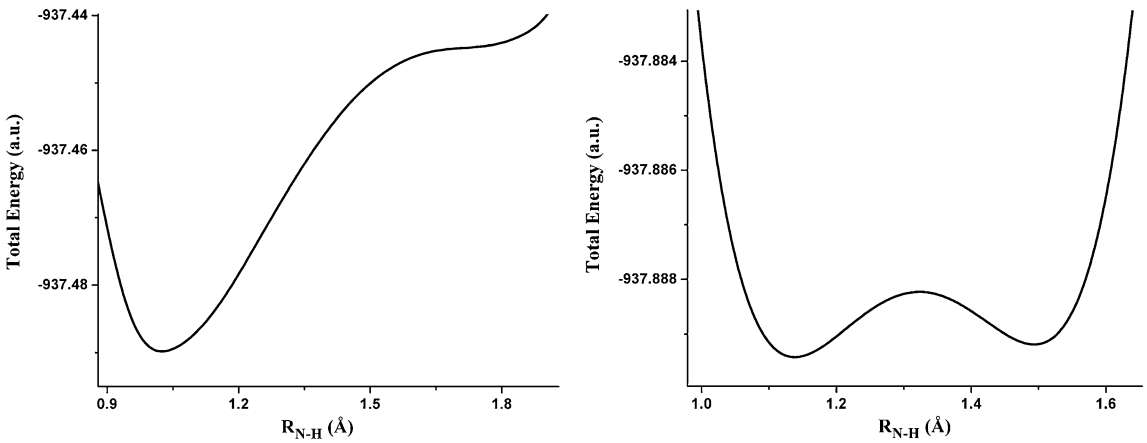

Fig. 4 **a**–**b** The potential energy curve along the $N7^{(G)}$–H bond for Watson–Crick (**a**) and Hoogsteen (**b**) configurations

atom of guanine and the N3 atom of cytosine, it perturbs other normal-mode vibrations located in the intermolecular region between the bases. Accordingly, its anharmonic coupling with other vibrations from the frequency set specified above is also quite strong. Apart from the influence of the ν_{11} normal mode on the vibrational anharmonicity, some other frequencies (ν_{13}, ν_{16}, ν_{18} and ν_{23}) from this frequency set also have significant diagonal and off-diagonal anharmonic constants. In the lower frequency domain (500−1000 cm^{-1}), three characteristic vibrations (ν_{45}, ν_{46} and ν_{64}) should be noted. These are ring-deformation vibrations of the guanine. These normal modes are also very strong when coupled with the ν_{11} vibration, but their couplings with other vibrations from the frequency range 1400−1900 cm^{-1} are also significant. If we consider the far-infrared (90−150 cm^{-1}) frequency domain, we can see three normal-mode frequencies (ν_{79}, ν_{80} and ν_{81}) that are characteristic of the intermolecular mode and have in-plane molecular vibrations. In this case, the guanine and the cytosine molecules show rigid motions compared to each other. They are strongly coupled with the ν_7 and ν_{11} stretching modes, with the ν_{16}, ν_{23} and ν_{25} bending vibrations, as well as with the ν_{45}, ν_{46} and ν_{64} ring deformations of the guanine, while their diagonal anharmonic constants are small. All of these anharmonic couplings suggest to us that there is significant coupling between the intramolecular stretching, bending and ring-deformation vibrations with high frequency values and the intermolecular in-plane rigid vibrations of guanine and cytosine components with low frequencies. A triangular matrix with homogeneous and off-diagonal elements for the normal modes discussed above is presented in Table 4 of the ESM.

The large anharmonic shift can explain the presence of H-bond stretching of the N–H bond at 1489 cm^{-1} in the Raman spectra of both the Watson–Crick and the Hoogsteen configurations (see Fig. 5 in Ref. [5]). In our normal-mode analysis, this bond-stretching vibration combined with other bending vibrations can be identified in the case of Hoogsteen pair as ν_{18}, which has an anharmonic frequency of 1463.7 cm^{-1}. In this case, the anharmonic correction is 178.5 cm^{-1}. For the Watson–Crick pair, we found the mode ν_{21} with an anharmonic frequency of 1498 cm^{-1} and a correction of 47 cm^{-1}. From these values, it is clear that anharmonic corrections are crucial to the correct assignment of this mode.

In the same experimental Raman spectra, two bands can be found for the Hoogsteen configuration in the region 600–700 cm^{-1}. The band with ν=616 cm^{-1} was not explicitly assigned to any normal modes by the authors of [5]. Considering our anharmonic frequency calculations, we suggest that this experimental line can be associated with the very large frequency shift of ν_{11} (see Table 3). Due to the unusual anharmonic effect, this mode is impossible to assign in the absence of anharmonic corrections. We think that this ansatz needs further, more detailed experimental and theoretical investigation.

Conclusions

We have presented a detailed investigation of the vibrational properties of Watson–Crick and Hoogsteen GC base pairs using two different DFT approaches. The vibrational density of states was used to describe the localization of the vibrational modes in the two configurations. We found that a relatively large fraction of the vibrational spectrum can be assigned to guanine alone. This is the case for the region 250−350 cm^{-1} in the Watson–Crick base pair. In the Hoogsteen configuration, the same localization in guanine occurs at 550−650 cm^{-1} and 770−900 cm^{-1}. Finally, in both configurations, only guanine's atoms exhibit significant vibrational amplitude in the region 1250−1600 cm^{-1}. We note that these results were obtained using the two different computational approaches.

On the other hand, a detailed investigation of the anharmonic corrections was performed. We have demonstrated that there are strong vibrational couplings between different normal modes in both the Watson–Crick and the Hoogsteen configurations. The presence of H^+ can significantly perturb the other normal-mode vibrations through its N−H^+ stretching mode (ν_{11}). In particular, we have shown that theoretical results cannot be correlated with experimental RAMAN spectra in the absence of anharmonic corrections. This is especially true for the Hoogsteen configuration. We noted the vibrational bands at 616 cm^{-1} and 1488 cm^{-1}. Large anharmonic corrections (frequency shifts of about 200–400 cm^{-1}) were found for all normal modes where the proton (H^+) is involved in the normal-mode vibration.

Acknowledgements We acknowledge financial support from CNCSIS-UEFISCDI, project PNII-IDEI ID 875/2008, contract no. 519/2009, as well as Core Program, project PN 09-440101, contract no. 44 N/2009. Thanks are due to NIRDIMT, Cluj-Napoca Data Center for providing computer facilities.

References

1. Jeffrey GA, Saenger W (1991) Hydrogen bonding in biological structure. Springer, Berlin
2. Marcus MA, Corelli JC (1974) Infrared spectroscopy of the photo- and radiobiology of DNA bases and their derivatives. Radiat Res 57(1):20–37
3. Fischer BM, Walther M, Jepsen PU (2002) Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy. Phys Med Biol 47(21):3807–3814. doi:10.1088/0031-9155/47/21/320

4. Nir E, Kleinermans K, de Vries MS (2000) Pairing of isolated nucleic-acid bases in the absence of the DNA backbone. Nature 408(6815):949–951. doi:10.1038/35050053
5. Segers-Nolten GMJ, Sijtsema NM, Otto C (1997) Evidence for Hoogsteen GC base pairs in the proton-induced transition from right-handed to left-handed poly(dG-dC)·poly(dG-dC). Biochemistry 36 (43):13241–13247. doi:10.1021/bi971326w
6. Puppels GJ, Otto C, Greve J, Robert-Nicoud M, Arndt-Jovin DJ, Jovin TM (1994) Raman microspectroscopic study of low-pH-induced changes in DNA structure of polytene chromosomes. Biochemistry 33(11):3386–3395. doi:10.1021/bi00177a032
7. Taillandier E, Liquier J (2002) Vibrational spectroscopy of nucleic acids. In: Chalmers JM, Griffiths PR (ed) Handbook of vibrational spectroscopy, vol. 5. Wiley, New York, pp 3465–3480
8. Ten GN, Burova TG, Baranov VI (2009) Calculation and analysis of vibrational spectra of adenine–thymine, guanine–cytosine, and adenine–uracil complementary pairs in the condensed state. J Appl Spectrosc 76(1):73–81. doi:10.1007/s10812-009-9149-3
9. Santamaria R, Charro E, Zacarías A, Castro M (1998) Vibrational spectra of nucleic acid bases and their Watson–Crick pair complexes. J Comput Chem 20(5):511–530. doi:10.1002/(SICI)1096-987X(19990415)20:5<511::AID-JCC4>3.0.CO;2-8
10. Morari CI, Muntean CM (2003) Numerical simulations of Raman spectra of guanine–cytosine Watson–Crick and protonated Hoogsteen base pair. Biopolymers 72(5):339–344. doi:10.1002/bip.10418
11. Jurečka P, Hobza P (2003) True stabilization energies for the optimal planar hydrogen-bonded and stacked structures of guanine···cytosine, adenine···thymine, and their 9- and 1-methyl derivatives: complete basis set calculations at the MP2 and CCSD (T) levels and comparison with experiment. J Chem Am Soc 125:15608–15613. doi:10.1021/ja036611j
12. Jeziorski B, Moszynski R, Szalewicz K (1994) Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. Chem Rev 94:1887. doi:10.1021/cr00031a008
13. Misquitta AJ, Podeszwa R, Jeziorski B, Szalewicz K (2005) Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. J Chem Phys 123:214103. doi:10.1063/1.2135288, and references therein
14. Hesselmann A, Jansen G, Schütz M (2006) Interaction energy contributions of H-bonded and stacked structures of the AT and GC DNA base pairs from the combined density functional theory and intermolecular perturbation theory approach. J Chem Am Soc 128:11730–11731. doi:10.1021/ja0633363
15. Han SY, Lee SH, Chung J, Oh HB (2007) Base-pair interactions in the gas-phase proton-bonded complexes of C^+G andC^+ GC. J Chem Phys 127:245102. doi:10.1063/1.2817604
16. Bikwood R, Gruebele M, Leitner DM, Wolynes PG (1998) The vibrational energy flow transition in organic molecules: theory meets experiment. Proc Natl Acad Sci USA 95(11):5960–5964
17. Rekik N, Oujia B, Wócik MJ (2008) Theoretical infrared spectral density of H-bonds in liquid and gas phases: anharmonicities and dampings effects. Chem Phys 352:65–76. doi:10.1016/j.chemphys.2008.05.009
18. Pelel L, Gerber RB (2008) On the number of significant mode-mode anharmonic couplings in vibrational calculations: correlation-corrected vibrational self-consistent field treatment of di-, tri-, and tetrapeptides. J Chem Phys 128:165105. doi:10.1063/1.2909558
19. Watanabe Y, Maeda S, Ohno K (2008) Intramolecular vibrational frequencies of water clusters $(H_2O)_n$ (n=2–5): Anharmonic analyses using potential functions based on the scaled hypersphere search method. J Chem Phys 129:074315. doi:10.1063/1.2973605
20. Lundell J, Latajka Z (2008) Vibrational calculations for the H_2O···CO complex. J Mol Struct 887:172–179. doi:10.1016/j.molstruc.2007.12.013
21. Ordejón P, Artacho E, Soler JM (1996) Self-consistent order-N density-functional calculations for very large systems. Phys Rev B53(16):R10441. doi:10.1103/PhysRevB.53.R10441
22. Soler JM, Artacho E, Gale JD, García A, Junquera P, Ordejón P, Sánchez-Portal D (2002) The SIESTA method for ab initio order-N materials simulation. J Phys Cond Mat 14(11):2745–2779. doi:10.1088/0953-8984/14/11/302
23. Gaussian 03, Revision C.02, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr., Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels A. D, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian, Inc., Wallingford
24. Dion M, Rydberg H, Schröder E, Langreth DC, Lundqvist BI (2004) van der Waals density functional for general geometries. Phys Rev Lett 92(2):246401. doi:10.1103/PhysRevLett.92.246401
25. Román-Pérez G, Soler JM (2009) Efficient implementation of a van der Waals density functional: application to double-wall carbon nanotubes. Phys Rev Lett 103(9):096102. doi:10.1103/PhysRevLett.103.096102
26. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A38(6):3098–3100. doi:10.1103/PhysRevA.38.3098
27. Perdew JP, Burke K, Wang Y (1996) Generalized gradient approximation for the exchange-correlation hole of a many-electron system. Phys Rev B54(16):16533–16539. doi:10.1103/PhysRevB.54.16533
28. Schaefer A, Horn H, Ahlrichs R (1992) Fully optimized contracted Gaussian-basis sets for atoms Li to Kr. J Chem Phys 97(4):2571–2577. doi:10.1063/1.463096
29. Guerra CF, Bickelhaupt FM, Snijders JG, Baerends EJ (2000) Hydrogen bonding in DNA base pairs: reconciliation of theory and experiment. J Am Chem Soc 122:4117–4128. doi:10.1021/ja993262d
30. Allouche AR (2011) Gabedit—a graphical user interface for computational chemistry software. J Comp Chem 32:174–182. doi:10.1002/jcc.21600
31. Clabo DA Jr, Allen WD, Remington RB, Yamaguchi Y, Schaefer HF III (1988) A systematic study of molecular vibrational anharmonicity and vibration-rotation interaction by self-consistent-field higher-derivative methods. Asymmetric top molecules. Chem Phys 123:187–239. doi:10.1016/0301-0104(88)87271-9
32. Barone V (2005) Anharmonic vibrational properties by a fully automated second-order perturbative approach. J Chem Phys 122 (01):014108. doi:10.1063/1.1824881
33. Christiansen O (2003) Møller–Plesset perturbation theory for vibrational wavefunctions. J Chem Phys 119(12):5773–5781. doi:10.1063/1.1601593
34. Neugebauer J, Hess BA (2003) Fundamental vibrational frequencies of small polyatomic molecules from density-functional calculations and vibrational perturbation theory. J Chem Phys 118:7215–7225. doi:10.1063/1.1561045

35. Herzberg G (1945) Molecular spectra and molecular structure. II. Infrared and Raman spectra of polyatomic molecules. D. Van Nostrand Co. Inc., New York
36. Wilson EB, Decius JC, Cross PC (1955) Molecular vibrations. McGraw–Hill, New York (reprinted by Dover in 1980)
37. Postnikov AV, Pagés O, Hugel J (2005) Lattice dynamics of the mixed semiconductors (Be, Zn)Se from first-principles calculations. Phys Rev B 71:115206. doi:10.1103/PhysRevB.71.115206
38. Špirko V, Šponer J, Hobza P (1997) Anharmonic and harmonic intermolecular vibrational modes of the DNA base pairs. J Chem Phys 106:1472–1479. doi:10.1063/1.473296
39. Brauer B, Gerber RB, Kabeláč M, Hobza P, Bakker JM, Riziq AGA, de Vries MS (2005) Vibrational spectroscopy of the G-C base pair: experiment, harmonic and anharmonic calculations, and the nature of the anharmonic couplings. J Phys Chem A 109:6974–6984. doi:10.1021/jp051767m
40. Bende A (2010) Hydrogen bonding in the urea dimers and adenine–thymine DNA base pair: anharmonic effects in the intermolecular H-bond and intramolecular H-stretching vibrations. Theor Chem Acc 125(3–6):253–268. doi:10.1007/s00214-009-0645-6
41. Wang G-X, Ma X-Y, Wang J-P (2009) Anharmonic vibrational signatures of DNA bases and Watson–Crick base pairs. Chin J Chem Phys 22(6):563–570. doi:10.1088/1674-0068/22/06/563-570
42. Del Bene JE, Jordan MJT (1998) A comparative study of anharmonicity and matrix effects on the complexes XH:NH_3, X=F, Cl, and Br. J Chem Phys 108(8):3205–3212. doi:10.1063/1.476370
43. Brisker D, Peskin U (2006) Vibrational anharmonicity effects in electronic tunneling through molecular bridges. J Chem Phys 125 (11):111103. doi:10.1063/1.2353148