Theor Chem Acc (2012) 131:1157
DOI 10.1007/s00214-012-1157-3

REGULAR ARTICLE

Modelling vibrational coupling in DNA oligomers: a computational strategy combining QM and continuum solvation models

Alessandro Biancardi · Roberto Cammi · Chiara Cappelli · Benedetta Mennucci · Jacopo Tomasi

Received: 20 July 2011 / Accepted: 30 September 2011 / Published online: 18 February 2012

Abstract In this paper, a computational strategy, based on DFT calculations at the M06-2X level, combined with the polarizable continuum model, the Hessian matrix reconstruction method and the Partial hessian vibrational approach is applied to evaluate inter- and intra-layer vibrational couplings between hydrogen bonded and stacked DNA base pairs. The present work demonstrates that this computational scheme can effectively predict and interpret the vibrational couplings of nucleic acids in solution. The effect of the environment described in a cluster or in a continuum manner is necessary in order to improve the agreement with the experimental values.

Keywords Vibrational coupling · Solvent effects · DNA pairs · DFT calculations · Hessian matrix reconstruction · Partial hessian vibrational approach · Guanine · Cytosine

1 Introduction

Two-dimensional infrared spectroscopy (2D-IR) [1–5] has been recently reported to provide valuable information on vibrational properties and dynamics of molecular systems, for which it is also a valuable tool to proof the local environment. 2D-IR somewhat overcomes limitations of conventional spectroscopic techniques, such as FTIR spectroscopy and circular dichroism, which need empirical qualitative relations in order to extract structural information from mono-dimensional spectra.

In addition to the usual diagonal peaks, the 2D-IR spectrum displays off-diagonal peaks that, similarly to the 2D-NMR spectroscopy, are produced by the coupling between oscillators and provide direct information about structure and dynamics. For these reasons, 2D-IR spectroscopy has been used to monitor the vibrational coupling in several molecular systems, such as hydrogen-bonded systems, nucleic acids [6–10], proteins and peptides.

Despite such potentialities, the extraction of structural information from 2D-IR spectra requires the coupling of experimental data with theoretical models, generally formulated in terms of quantum mechanical (QM) calculations, which permit to directly compute the vibrational properties of the molecular sub-units (frequencies and vibrational transition moments) and the vibrational coupling among them. The QM calculation of harmonic vibrational frequencies and transition dipole moments can be nowadays routinely performed, mainly due to the availability of analytical algorithms for the evaluation of energy second derivatives, for both isolated and solvated systems treated by means of continuum solvation models [11]. The theoretical modelling of the vibrational coupling is instead far from being a standard procedure and has been based on different approaches.

The simplest approach is the transition dipole coupling (TDC) model [12–14], based on the Coulomb interaction between the vibrational transition point dipoles associated with the interacting modes. However, due to its definition, the TDC is limited to describe vibrational couplings of electrostatic origin. A more accurate approach is the

Dedicated to Professor Vincenzo Barone and published as part of the special collection of articles celebrating his 60th birthday.

A. Biancardi · C. Cappelli · B. Mennucci · J. Tomasi (✉)
Dipartimento di Chimica e Chimica Industriale,
Università di Pisa, via Risorgimento, 35, 56126 Pisa, Italy
e-mail: tomasi@dcci.unipi.it

R. Cammi
Dipartimento di Chimica GIAF, Università di Parma,
viale delle Scienze, 43100 Parma, Italy

Hessian matrix reconstruction (HMR) method developed originally by Cho and co-workers [15–18], which is able to describe also the non-electrostatic components of the vibrational coupling. The basic formulation of the HMR requires an harmonic QM calculation of the vibrational modes of an appropriately chosen model system, followed by the computational vibrational analysis of the target system. The normal modes of the target system are then resolved in terms of vibrational modes localized on the molecular subunits, so that local modes frequencies and the inter-mode vibrational coupling can be computed. Both the TDC and the HMR methods have been originally used to describe the vibrational coupling for isolated systems, and only recently extended to solvated systems [19], by some of the present authors within the framework of the polarizable continuum model (PCM) [20], which describes the target solute at the QM level, while the solvent is described as a dielectric medium characterized by its macroscopic dielectric properties. The most recent implementations of the PCM manage to compute the molecular geometry and a large variety of properties of solvated systems, including vibrational normal modes and frequencies. In the formulation of the HMR within the PCM, the solvent effects are taken into account in terms of changes in the geometry of the system, of its normal modes and vibrational frequencies, due to the presence of the solvent. The extension of the TDC to PCM solutes relies instead on the evaluation of the vibrational transition dipoles of the local modes in the presence of the solvent and then explicitly includes medium screening effects on the Coulombic interaction between the transition dipoles. On the grounds of its formulation, the PCM–TDC model is suitable to model the vibrational coupling of weak interacting systems only.

In this paper, we will focus on the application of the PCM–HMR method to study the vibrational coupling between the nucleic acid bases in DNA double-stranded oligomers. It is in fact well known that infrared techniques are very sensitive to the structure of the polynucleotide helix [21, 22] and that DNA bases change in a unique way their vibrational properties after structural modifications, due to the different vibrational couplings and due to environment effects.

A good understanding of the structure and dynamics of a molecule is a key factor in order to correlate the change of couplings, to particular structural modifications [23]; in the case of DNA, this is fundamental, for instance, in order to develop more efficient DNA-binding drugs [24, 25].

QM calculations of the vibrational coupling in DNA oligomers have been reported by Zanni and co-workers [7] and by Cho and co-workers [26–29] adopting different computational strategies. In particular, Zanni and co-workers have studied the vibrational coupling between normal modes involving carbonyl stretching in Watson–Crick (WC) base pairs guanine (G) and cytosine (C) and between nearest-neighbour bases G and C, in two different DNA oligomers. QM calculations were performed on isolated G and C bases, by neglecting the deoxyribose and phosphate groups to reduce the computational costs, and the vibrational coupling was obtained by using the electrostatic TDC [12–14] and transition dipole density distribution [30] models, and a finite difference evaluation of the PES curvature. In particular, for the vibrational coupling between the WC pairs, a single pair was used as model, while for the evaluation of the vibrational coupling between nearest-neighbour bases G and C, two stacked WC base pairs were considered. The vibrational coupling for the same DNA oligomers has also been studied by Cho and co-workers [26–29], by performing QM calculations on G and C and computing the vibrational coupling through a modified HMR method. Large vibrational coupling models consisting of GC stacked base pairs up to quadruple GC stacked base pairs have been considered, but solvent effects have been only marginally studied in the particular case of a GC pair surrounded by the explicitly included adjacent base pairs.

The consideration of the local environment has been reported as crucial in order to realistically describe vibrational couplings in solution [10], principally because the interaction with the environment modifies both the single local vibrations and the coupling between them. Similarly to a previous work of some of the present authors on amide vibrational couplings in small peptides [19], in this paper, we will apply a computational strategy still based on the PCM combined with the HMR method and Partial hessian vibrational approach (PHVA) [31–37] to evaluate inter- and intra-layer couplings between hydrogen bonded and stacked DNA base pairs. In particular, the coupling between carbonyl vibrations located on G and C bases will be considered in the case of a A-form DNA of well-known structure, resolved with X-ray crystallography [38].

The paper is organized as follows. In the next section, the description of the systems investigated and of the methods used for the analyses is reported. Then, numerical results will be shown, by also putting into evidence how the results are influenced by the choice of the pair geometry and orientation, on the extension of the molecular model and on the accounting/discarding of solvent effects. Some conclusions end the presentation.

2 Methods and systems

2.1 Modelling DNA oligomers

In this study, we have investigated a sequence of guanine and cytosine in the double-stranded helix DNA $d(G_nC_n)$. In

particular, we have used the structure of the DNA oligomer $d(G_4C_4)$ determined by McCall et al. [38] by single-crystal X-ray methods (PDB ID: 2ANA). In the structure, there are two strands in the asymmetric unit, and these coil about each other to form a right-handed *A*-type double helix with Watson–Crick hydrogen bonds between base pairs. The two d(G–G–G–G)·d(C–C–C–C) segments in the oligomer exhibit similar and uniform structures (see Fig. 1).

From this experimental structure, five different clusters were extracted, containing the coupled bases that from now on will be indicated as fragments 1–5. The borders of the helix were not considered.

To introduce the effect of the rest of DNA, two different models were used. In the first model (referred to as PCM), the environment was represented as a continuum dielectric, while in the second model (cluster), the adjacent base pairs were explicitly included in the QM calculation. More in details, for the H-bonded (HB) pairs, two additional bases were considered, while for the stacked pairs, four additional base pairs were included (see Fig. 2).

2.2 Methodologies

As pointed out by Lee and Cho [26], the application of the HMR approach to the study of coupled vibrations localized on base pairs presents critical points, due to the high delocalization of the vibrational modes. To overcome these critical points, we have combined the HMR approach with the PHVA [31–37].

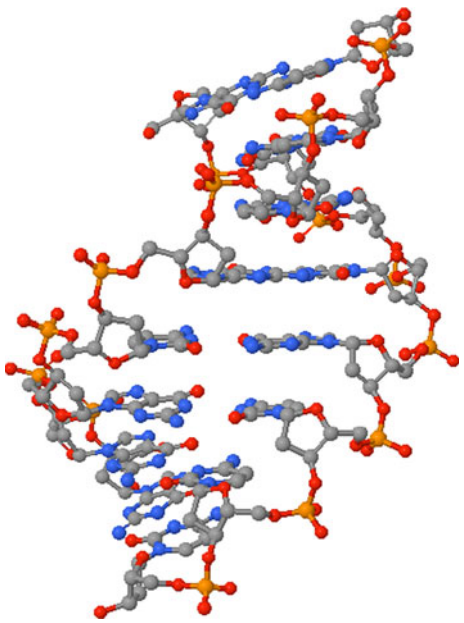

Fig. 1 Structure of the DNA oligomer investigated in this study. From this experimental structure, five different coupled bases were extracted, which from now on will be indicated as fragments 1–5

The PHVA approach is used both to perform the vibrational calculation on the coupled HB and stacked pairs and to localize the vibrational modes on each single basis. In this approach, the system is divided in two parts: a freely vibrating and a frozen part. The latter is taken frozen during the vibrational calculation and, therefore, does not directly contribute to the normal mode of the former. In agreement with the one-bond-distance rule [39], more than one bonds between the freely vibrating (the bases) and the fixed part (the furanoses and the adjacent base pairs in the cluster model) have been left in order not to bring the two parts too close (see Fig. 3).

The calculation of the vibrational coupling between two local modes located on different base pairs consists of the following steps:

- Normal modes calculation of the single uncoupled basis in the presence of the others kept frozen, using the PHVA approach. These are the local modes of the uncoupled basis, which are used as a basis for the determination of the vibrational coupling.
- Normal modes calculation of the coupled pair possibly in the presence of the other bases kept frozen, using the PHVA approach. These are the normal modes of the coupled pair.
- Application of the HMR using the results of the coupled and the uncoupled calculations to determine the vibrational coupling between the local vibrations.

2.3 Computational details

All calculations were performed by using the Gaussian 09 package [40], by employing DFT and using the M06-2X functional [41]. This functional was chosen for its ability to describe non-covalent interactions [42, 43]. To describe the continuum environment, the integral equation formalism [44] version of PCM was used. PCM cavities were built as a series of interlocking spheres centred on the atoms. Two different set of radii were tested: the UFF radii (all multiplied by a cavity size factor of 1.1), which are the Gaussian 09 default settings (from now on indicated as StCav), and a set of enlarged radii (employing a cavity size factor of 1.7). Since the distance between staked bases is about 3.4 Å [45], these enlarged radii are able to mimic the void space around the base pairs in the DNA double helix (from now on indicated as LCav). In the particular case of the PCM calculations for the fragments, in order to consider the possibility that in a DNA double helix the volume between the base pairs cannot be accessible to solvent, a larger radius (equal to that of the carbon atom) was used for the hydrogen atoms involved in hydrogen bonds between bases. Two different values of the dielectric permittivity were tested, corresponding to a weakly polar ($\varepsilon = 6$) and a non-polar ($\varepsilon = 2$) environment.

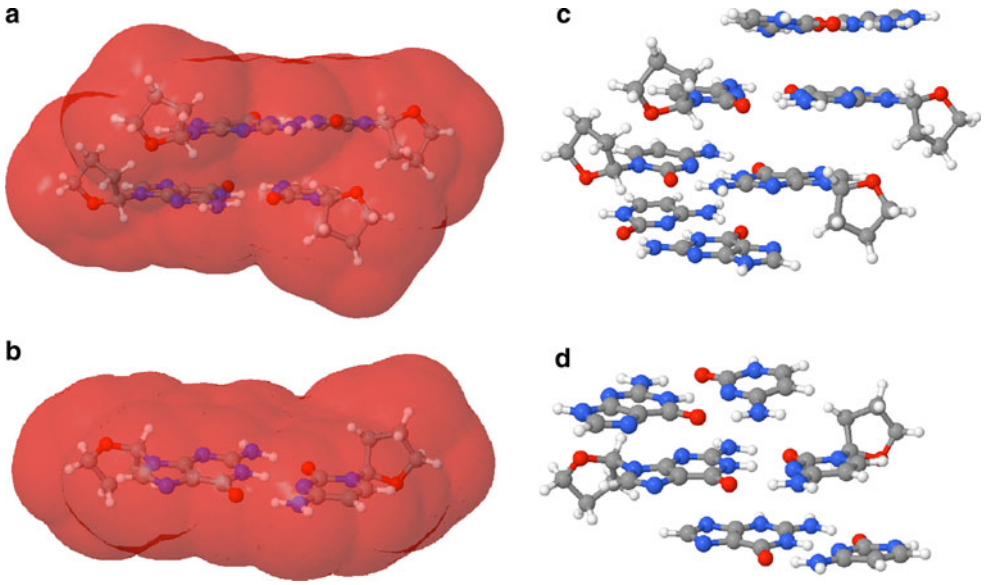

Fig. 2 Representations of the stacked and HB fragments within the PCM (**a** and **b**) and the explicitly included adjacent base pairs (**c** and **d**), respectively

The combination of three basis sets for the three different parts of the fragments was used, namely 6-311+G(d,p) for the freely vibrating bases, 6-31G(d) for the other vibrationally frozen bases and 3–21 g for the furanose rings. All the geometries were kept fixed to the experimental structure (taken from [38]), with the exception of the missing hydrogens in the experimental X-ray structure, that were optimized at the 6-311+G(d,p) level.

3 Results

3.1 Normal modes and frequencies of references bases in solution

Before discussing the couplings in the DNA oligomers, in this section, we will focus on the vibrational properties of the single bases composing the systems, that is, guanine (G) and cytosine (C), in their methylated form. The role of such results is twofold: in fact, besides being interesting per se, they can also give an idea of the accuracy of the results obtained in the following analysis on the DNA oligomers, since the single C and G are used as basis set for the PHVA and HMR analysis.

The vibrational properties of C and G have been previously studied in the literature, both from the experimental [7, 26, 46–50] and computational [7, 46, 51–54] points of view. In particular, values for the two selected modes, referred to in the following as Gs and Cs, have been reported in various environments. According to DFT calculations, these modes have predominately carbonyl stretch character, as shown in Fig. 4. Such modes are the ones on which the following discussion on vibrational couplings will be focused.

In Table 1, PCM calculated frequencies for G_s and C_s modes in different environments are reported, chosen according to the availability of experimental data. The outcomes of previous calculations reported in the literature are also shown for comparison's sake, as well as experimental values. In the case of PCM calculated data, two different choices of the definition of the PCM molecular cavity surrounding the solute in the dielectric are compared, that is, the cavity obtained by exploiting the default settings of Gaussian 03 [55] (G03,) and Gaussian 09 [40] (G09). In fact, the cavity size and shape are the only adjustable parameters in a PCM calculation, for a given solvent.

The calculated value of G_s in DMSO is sensitive to the choice of the PCM cavity. In particular, the use of a smaller cavity (the G03 default settings) decreases the calculated value of more than 20 cm^{-1}. As said, the Gs mode involves the C=O group, which, due to the geometry of the molecule, is exposed to the solvent and, therefore, very sensitive to the polarity of the solvent (compare also the previous calculations in vacuo reported in Table 1). However, with the default G09 settings, the cavity surrounding the C=O group is larger than in the case of the G03 default settings, so that a decrease in the solvent effect (which causes a decrease in the calculated value) is observed. As far as the comparison with the experiment is concerned, a better agreement is obtained with the use of the G03 cavity. However, both choices of the PCM cavity yield an exp/calc ratio in the range 0.96–0.97, which are in line with the reported scaling factors for the same DFT functional and similar basis sets [56]. The use of a low-polarity environment for the calculations ($\varepsilon = 2$) makes the calculated values approach the calculations for the isolated system, as it is expected. Also, a very low difference is observed with the two cavities, as the solvent effect being in both cases very low.

Also in the case of cytosine in water, the same trend as reported for DMSO as varying the cavity is observed. In

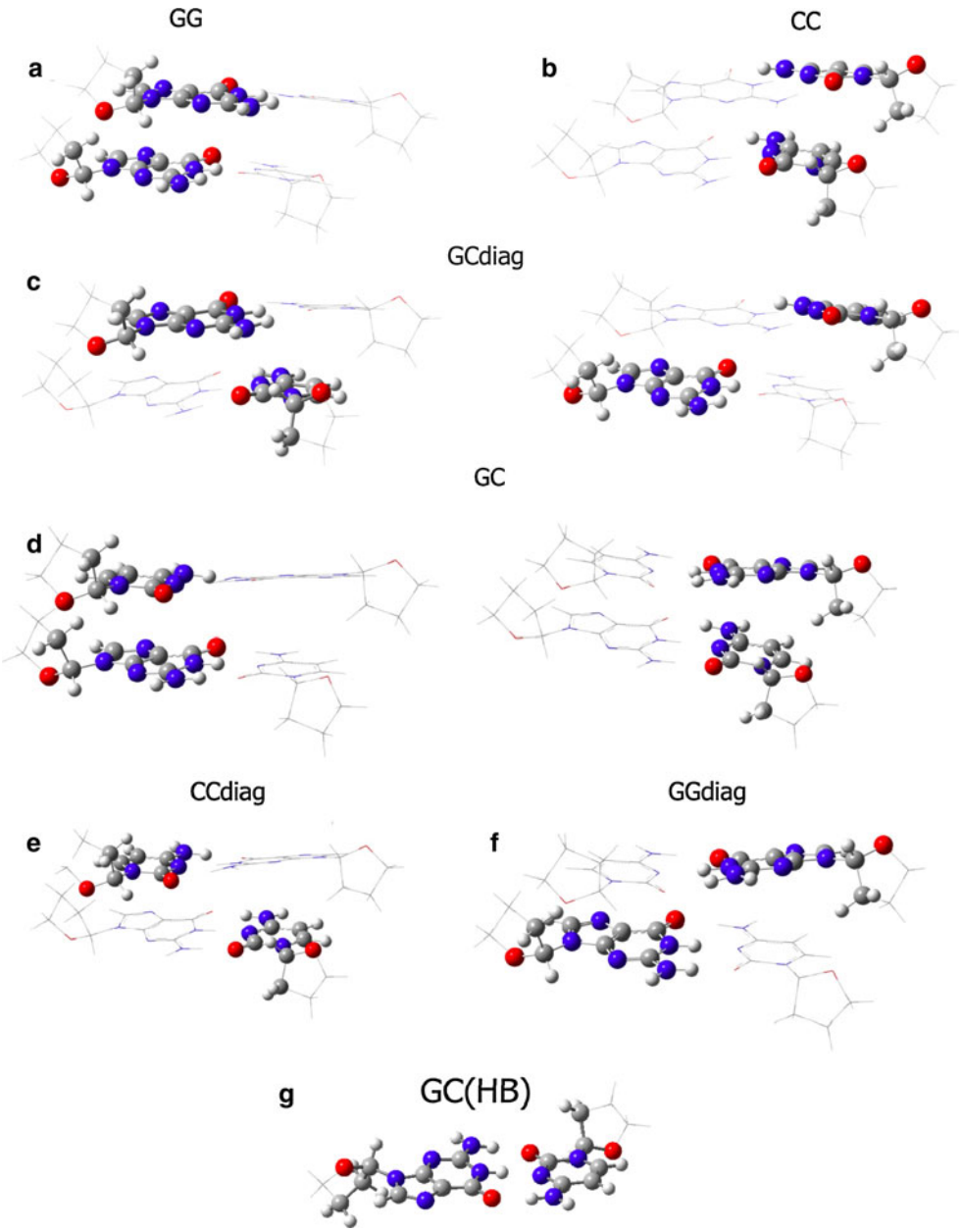

Fig. 3 Representations of the stacked (**a**–**f**) and HB fragments (**g**). The freely vibrating atoms are represented in "Ball and Stick", whereas the frozen atoms are shown in "wireframe". The systems **a**–**d** are obtained from the fragments 1, 2, 4, 5, whereas the systems represented in **d**–**f** are obtained from the fragment 3

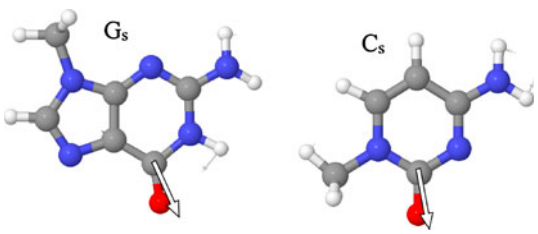

Fig. 4 Structure and pictorial view of the two selected normal modes of G (*left*) and C (*right*)

fact, a difference of about 20 cm^{-1} is reported for C_s, with the G03 once again approaching the experimental value, with an exp/calc ratio in the range 0.97–0.99, depending on the reference experimental value.

3.2 Normal modes and frequencies in DNA

The analysis reported in the previous section for the solvated G and C bases has shown the characteristics (including possible limits) of the QM model here adopted to describe vibrational motions in solvated systems. When moving to a more complex environment such as DNA, however, further difficulties appear due to the extremely inhomogeneous and anisotropic interactions between the bases and the embedding matrix. In particular, when focusing on vibrations, differences between solvated and embedded bases can appear not only in the frequencies of the investigated modes but also in the nature of the modes themselves. It is thus necessary to analyse the possible

Table 1 Calculated M06-2X/6-311+G(d,p) and experimental frequency values for a selected mode of C and G monomers in various environments

Mode	This work	Previous calculations in vacuo	Experimental
C_S	1747 G09 $\varepsilon = 2$	1750 (B3LYP/6-31G(d)) [26]	1712 (in N_2) [17, 46] 20 (in Ar) [52]
	1738 G03 $\varepsilon = 2$	1660 (BP86/6-311G) [46]	1730 (in Ne) [53]
	1702 G09 in H_2O	1722 (MP2/6-31G(d, p)) [49]	1645 in D_2O [7]
	1680 G03 in H_2O	1785 (B3LYP/6-31++G(d,p)) [7]	1660 in H_2O [54]
		1818 B3LYP/6-31G(d,p) [50]	
G_S	1799 G09 $\varepsilon = 2$	1686 (BP86/6-311G) [46]	1736, 1749 (in Ar) [51]
	1799 G03 $\varepsilon = 2$	1727 (B3LYP/6-31+G(d,p)) [47]	1742 (in N_2) [51]
	1760 G09 in DMSO	1762 (B3LYP/6-31G(d)) [26]	1693 in DMSO [7]
	1737 G03 in DMSO	1840 (B3LYP/6-31++G(d,p)) [7]	
		1833 (B3LYP/6-31G(d)) [48]	

All data are given in cm^{-1}

changes on the previously investigated modes before trying to simulate their couplings.

As explained in the methodological section, the computational strategy here adopted is based on the HMR approach, in which the couplings between vibrational modes are obtained by "projecting" the local modes of the uncoupled systems into those of the coupled pair. In present context, the local modes are those of the uncoupled G or C but in the presence of the embedding matrix, which will modify the local modes due to different effects, including the fact that now G and C are chemically bonded to furanose groups. The presence of a real chemical bond between the guanine or the cytosine to the DNA strand will in fact reduce the vibrational freedom of these molecules with important consequences in the displacements and the frequencies related to the modes. Here, in particular, we shall focus on the coupling between what we previously indicated as G_s and C_s modes. Notice that the effect of the side chain on the vibrational modes has been reported to be small for G_s, but more relevant for C_s (see [26]).

In Fig. 4, we previously reported the selected modes of G and C, whereas in Table 2, we report the corresponding frequencies obtained in vacuo and with the different models for mimicking the DNA environment.

With respect to the calculations in vacuo, the presence of the continuum dielectric, even if with a low dielectric constant, causes a decrease in the frequency absolute values, which is of the order of 6 cm^{-1} for C and 10 cm^{-1} for G. The use of a larger PCM cavity (i.e. LCav) causes the results to go towards the in vacuo limit, as expected. Overall, the differences among the results obtained with the two choices for the size of the PCM cavity are minor, in both cases of the order of 5–7 cm^{-1}, and independent on the choice of the fragment. The increase in the dielectric constant of the medium (moving from $\varepsilon = 2$ to $\varepsilon = 6$) also causes a shift of the same order. Noticeably, the calculations performed with the "cluster" model yield results very similar to the PCM $\varepsilon = 2$, thus showing that the actual environment experienced by the single bases in the DNA environment is nicely reproduced by the PCM.

Table 2 Vibrational frequencies of uncoupled linked cytosine and guanine moieties as obtained with different solvation models, in the fragment *1*

	C_{S1}	C_{S2}	Average
Cytosine			
Vacuum	1709	1700	1704
PCM $\varepsilon = 2$, StCav	1703	1694	1698
PCM $\varepsilon = 2$, LCav	1707	1699	1703
PCM $\varepsilon = 6$, StCav	1696	1687	1691
Cluster	1703	1695	1699
	G_{S1}	G_{S2}	Average
Guanine			
Vacuum	1764	1755	1760
PCM $\varepsilon = 2$, StCav	1754	1745	1750
PCM $\varepsilon = 2$, LCav	1761	1752	1756
PCM $\varepsilon = 6$, StCav	1743	1734	1738
Cluster	1757	1749	1753

C_{S1}, C_{S2} and G_{S1}, G_{S2} refer to the local vibrations of cytosine and guanine in fragment 1. All values are given in cm^{-1}

We first note that, even if in Table 2, we reported data for fragment 1, the analysis we can get is general as the differences obtained for the same system in different fragments are always small, of the order of 10 cm^{-1}.

In order to get more insight into the trends just commented, it is relevant to recall here the actual molecular structures of the systems to which Table 2 refers. In particular, it is to note that all the calculations were performed on the same geometries, corresponding to what is extracted from the experimental X-ray structure taken from [38],

Table 3 Vibrational couplings calculated in vacuo and through different solvation models, in the fragment 1

	β_{GG}	β_{CC}	β_{GC}	β_{CG2}	β_{GC}(HB)
Vacuum	10.0	4.5	2.5	4.5	11.4
PCM $\varepsilon = 2$, StCav	9.1	3.7	0.9	4.0	11.8
PCM $\varepsilon = 2$, LCav	9.1	3.6	1.0	4.3	12.0
PCM $\varepsilon = 6$, StCav	8.1	3.5	0.5	2.9	13.7
Cluster	8.6	3.6	1.5	4.2	10.8

All values are in cm^{-1}

except for the positions of the hydrogen atoms, which were optimized. In addition, each basis is, in this case linked, to a partially frozen furanose group (see Fig. 3), and the corresponding frequencies have been obtained in the presence of the coupled bases, kept frozen. All these aspects make the comparison with what was previously shown in Table 1 for the "free" bases, quite difficult. In fact, in that case, the geometries were relaxed in each of the considered environments. As it can be seen, the differences are quite substantial: by referring to the PCM $\varepsilon = 2$ calculation, the calculated data for the "free" bases are about 45 cm^{-1} higher than the corresponding ones for the bonded bases in the frozen fragments.

3.3 Couplings

In Table 3, we report the vibrational coupling constants (β) obtained for the different pairs (stacked and HB) using different solvation models: the values reported in the different columns refer to G–G, C–C and G–C in the stacked pairs (see Fig. 3a–c) and G–C in the HB pair (see Fig. 3g), respectively. In the case of stacked G–C, two sets of values are reported corresponding to the two different diagonal arrangements (see Fig. 3c).

As far as the nomenclature in Table 3 is concerned, we note that the coupling terms β_{GG} and β_{CC} refer to stacked guanine and cytosine bases, respectively, and β_{GC}, β_{CG2} refer to the diagonal coupling between guanine and cytosine.

As it can be seen from the data reported in Table 3, the inclusion of the effects of the DNA environment always leads to a decrease in the coupling constants, with the only exception of the HB pair in a PCM environment.

Such a decrease in the β values can be ascribed to screening effects of the polarizable environment, and this is what found for all stacked pairs with all solvation models. As expected, the largest screening is found in the case of PCM with $\varepsilon = 6$ (i.e. the highest dielectric constant). For all stacked pairs, the cluster calculation is better reproduced by the PCM model with $\varepsilon = 2$, which can be defined a good continuum model to mimic the microscopic DNA-like environment here examined (which, however, does not fully consider the complexity of DNA, also comprising ribose, phosphate, surrounding waters and even counterions). In the following, the analysis of PCM environment will be limited to this description.

In the case of the HB pair, the analysis is less simple as the PCM description always leads to an increase of β also in the case of the smallest permittivity ($\varepsilon = 2$) and the largest cavity. This increase seems to be due to a large variation of the transition densities which is here opposite to the expected screening effect which is also diminished being the two coupled bases quite close [57].

Table 4 Calculated vibrational couplings for the investigated fragments in different environments

Fragment	β_{GG}			β_{CC}		
	Vac	$\varepsilon = 2$ LCav	$\varepsilon = 2$ StCav	Vac	$\varepsilon = 2$ LCav	$\varepsilon = 2$ StCav
1	10.0	9.1	9.1	4.5	3.7	3.6
2	11.7	10.7	9.9	3.5	3.4	3.4
3	6.7	6.7	6.5	2.9	2.6	2.5
4	13.9	12.9	12.0	3.4	3.3	3.3
5	10.6	9.8	9.0	2.9	2.7	2.7
$<\beta>$	10.6	9.8	9.3	3.4	3.1	3.1

Fragment	β_{GC}			β_{CG2}			β_{GC}(HB)		
	Vac	$\varepsilon = 2$ LCav	$\varepsilon = 2$ StCav	Vac	$\varepsilon = 2$ LCav	$\varepsilon = 2$ StCav	Vac	$\varepsilon = 2$ LCav	$\varepsilon = 2$ StCav
1	2.5	1.0	0.9	4.5	4.3	4.0	11.4	12.0	11.8
2	1.5	1.3	1.1	8.2	7.5	6.9	9.7	9.8	9.7
3	6.0	5.7	6.0	11.9	8.5	6.1			
4	1.5	1.3	1.3	3.4	3.6	5.0			
5	6.3	2.4	1.7	4.6	4.2	4.1			
$<\beta>$	3.6	2.3	2.2	6.5	5.6	5.2			

All values are given in cm^{-1}

Table 5 Internuclear C(O)–C(O) distances (in Å) and CO–CO dihedral angles (in degrees) in the coupled bases in the different fragments

Fragment	r_{GG}	r_{CC}	r_{GC}	r_{GC2}	r_{GC}(HB)	θ_{GG}	θ_{CC}	θ_{GC}	θ_{GC2}	θ_{GC}(HB)
1	4.2	4.4	7.3	4.2	4.5	24	24	174	162	174
2	3.7	4.7	7.2	4.5	4.4	32	40	168	154	170
3	4.6	5.5	4.7	4.2		47	47	91	96	
4	3.6	4.5	7.1	4.7		21	24	175	164	
5	3.8	4.4	7.1	4.5		27	31	178	156	
Avg	4.0	4.7	6.7	4.4	4.5	30	33	157	146	172

Moving now to the comparison among the different pairs, it is interesting to note that the GG stacked pair presents a coupling constant very similar to that of the GC HB pair, for which the largest interaction was expected.

In order to be properly compared with experiments, the results reported in Table 3 have to be further analysed so to estimate the effects of heterogeneity along the double strand. As explained in the methodological section, from the DNA oligomer, we have extracted five non-equivalent fragments (see Fig. 3). The corresponding β values obtained in vacuo and with PCM $\varepsilon = 2$ with the two cavities are reported in Table 4, whereas in Table 5, we collect the main geometrical parameters, namely the internuclear distances between the carbonyl carbons of the two coupled bases and the angles between the corresponding CO vectors.

We note that the significant differences found in the fragment n = 3 are due to the change in the sequence of bases. In fact, we note that, in fragments 1,2,4,5, the coupling terms β_{GG} and β_{CC} refer to stacked guanine and cytosine bases, respectively, and β_{GC}, β_{CG2} refer to the diagonal coupling between guanine and cytosine, whereas in fragment 3, the coupling terms β_{GC}, β_{CG2} refer to the coupling between stacked guanine and cytosine, respectively, and β_{GG} and β_{CC} refer to the coupling between diagonal guanine and cytosine (see Fig. 3d–f).

The comparison of the β values with the structural data reported in Table 5 reveals that the vibrational couplings are sensitive not only to the inter-layer distances, but also on the specific DNA base pair, that is, on the specific environment acting on the specific DNA base.

Finally, in Table 6, a comparison between our data and the outcomes of previous calculations are reported, as well as experimental data available in the literature.

We first note that previous calculations were performed by exploiting a different approach with respect to our. In fact, the data reported in [7] were obtained by using a finite difference (FD) approach, in which the coupling is calculated through the numerical evaluation of total energy second derivatives with respect to the local modes of the single nucleic acid bases. In this case, the local modes are normal modes of the isolated nucleic basis.

The calculations reported in [26] were performed with an extension of the HMR approach (Extended-HMR), which assumes the Hessian matrix in the atomic coordinate system to be divided into blocks, each associated with the local nucleic acid bases modes. Such an Hessian submatrix generates the corresponding local modes, which are used to transform the Hessian matrix from the global basis modes to the local modes. The same approach was also exploited to obtain the data reported in [29], where environmental effects are taken into account through molecular dynamics techniques.

The differences between the various approaches explain the variability of the results, being the set of values for the β_{GG} vibrational coupling the most significative under this comparative viewpoint. Also, due to the large differences between the computational approaches previously exploited and our method, it is not simple to clearly evidence the origins of the discrepancies. However, the trend reported for the extended-HMR/B3LYP approach by Lee and Cho [26] is in good agreement with our data, whereas the largest discrepancies are noticed for FD calculations [7], which is, however, the most different with respect to our model.

Table 6 also reports the comparison between the calculated and the fitted experimental vibrational couplings taken from [6, 7]. Overall, a general good agreement with experiments is found, especially if we consider that the experimental findings refer to a dG_5C_5 sequence that is not identical to the similar sequence dG_4C_4 used in the present study. The effect of the environment described in an explicit or in a continuum manner is not large, due to the low polarity of the environment to be modelled; however, a proper account of environmental effects improves the agreement with the experimental data.

4 Summary and conclusions

In this paper, the vibrational couplings of nucleic acids were investigated by carrying out QM calculations, combined with: a) the HMR approach in order to determine the coupling between vibrations, b) the PHVA in order to

Table 6 Comparison among experimental and calculated vibrational couplings

Method	Environment	β_{GG}	β_{CC}	β_{GC}	β_{CG2}	β_{GC}(HB)	References
HMR/M06-2X	Gas phase	10.0	4.5	2.5	4.5	11.4	This work
On fragment 1	$\varepsilon = 2$, LCav	9.1	3.7	1.0	4.3	12.0	This work
	$\varepsilon = 2$, StCav	9.1	3.6	0.9	4.0	11.8	This work
	Cluster	8.6	3.6	1.5	4.2	10.8	This work
HMR/M06-2X	$\varepsilon = 2$, LCav	9.8	3.1	2.3	5.6	10.9	This work
Averaged values on fragments 1–5	$\varepsilon = 2$, StCav	9.3	3.1	2.2	5.2	10.6	This work
FD/MP2	Gas phase	12.6	0.5	3.0	4.7	11.6	[7]
FD/B3LYP	Gas phase	21.8	1.0	1.6	5.0	10.4	[7]
Extended-HMR/M06-2X	Gas phase	8.4	1.3	0.4	2.5	13.9	[26]
	Cluster	8.3	1.0	0.2	2.3	13.7	[26]
	Classical MD simulations in D_2O solvent	17.4	3.2	1.8	4.3	12.8	[29]
Exp. (± 2 cm^{-1})		9.5	0.7	1.9	2.7	9.6	[7]
Exp.		9.7	2.3	>1.0	5.0		[6]

All values are given in cm^{-1}

accurately determine the vibrational local modes, c) an explicit description of the nearest parts of the DNA helix in order to account for a part of the DNA environment effect and d) a continuous description in order to account (in an average manner) for the effect of the rest of the DNA helix and of the solvent.

The present work demonstrates that this strategy can be effectively used in order to predict and interpret some features of the 2D-IR spectra of nucleic acids in solutions. The good correlation of our approach with both previous calculations and experimental findings shows that this method can indeed be used in combination with experimental measurements to gain an accurate simulation of the vibrational coupling between vibration localized on different DNA bases. However, many extensions of our work are possible. For example, the inclusion of water molecules potentially hydrogen bonded to the carbonyl groups could represent an important improvement towards a more realistic description: notice in fact that the authors of the X-ray structure [38] pointed out that "almost all exposed functional groups on the DNA are hydrated". Moreover, the notable change in the coupling term for fragment 3, which is attributed to the alteration in the base sequence, is an index on the important role played by the sequence which is known to affect the local structure and flexibility properties of DNA base pairs. Future studies could address this issue so to better understand the real impact of sequence on vibrational couplings.

References

1. Cho M (2008) Chem Rev 108:1331
2. Kim YS, Hochstrasser RM (2009) J Phys Chem B 113:8231
3. Jeon J, Yang S, Choi JH, Cho M (2009) Acc Chem Res 42:1280
4. Ganim Z, Chung HS, Smith AW, Deflores LP, Jones KC, Tokmakoff A (2008) Acc Chem Res 41:432
5. Jansen TIC, Knoester J (2009) Acc Chem Res 42:1405
6. Krummel AT, Mukherjee P, Zanni MT (2003) J Phys Chem B 107:9165
7. Krummel AT, Zanni MT (2006) J Phys Chem B 110:13991
8. Yang M, Szyc Y, Rottger K, Fidder H, Nibbering ETJ, Elsaesser T, Temps F (2011) J Phys Chem B 115:5484
9. Yang M, Szyc Y, Elsaesser T (2011) J Phys Chem B 115:1262
10. Szyc L, Yang M, Nibbering ETJ, Elsaesser T (2010) Angew Chem Int Ed 49:3598
11. Mennucci B, Cammi R, Tomasi J (1999) J Chem Phys 110:6858
12. Krimm S, Abe Y (1972) Proc Natl Acad Sci USA 69:2788
13. Torii H, Tasumi M (1992) Chem Phys 96:3379
14. Torii H, Tasumi M (1996) In: Mantsch HH, Chapman D (eds) Infrared spectroscopy of biomolecules. Wiley, New York
15. Gorbunov RD, Kosov DS, Stock G (2005) J Chem Phys 122:224904
16. Hamm P, Wouterson S (2002) Bull Chem Soc Jpn 75:985
17. Ham S, Cha S, Choi JH, Cho M (2003) J Chem Phys 119:1451
18. Hahn S, Ham S, Cho M (2005) J Phys Chem B 109:11789
19. Biancardi A, Cappelli C, Mennucci B, Cammi R (2010) J Phys Chem B 114:4924
20. Tomasi J, Mennucci B, Cammi R (2005) Chem Rev 105:2999
21. Taillandier E, Liquier J (1992) In: Lilley HJ, Dahlberg JJ (eds) Methods in enzymology. Academic Press, New York
22. Banyay M, Sarkar M, Gräslund A (2003) Biophys Chem 104:477
23. Hunt NeilT (2009) Chem Soc Rev 38:1837
24. Richards AD, Rodger A (2007) Chem Soc Rev 36:471
25. Biancardi A, Biver T, Marini A, Mennucci B, Secco F (2011) Phys Chem Chem Phys 13:12595
26. Lee C, Park K, Cho M (2006) J Chem Phys 125:114508
27. Lee C, Park K, Cho M (2006) J Chem Phys 125:114509
28. Lee C, Park K, Cho M (2006) J Chem Phys 125:114510
29. Lee C, Park K, Cho M (2007) J Chem Phys 126:145102
30. Moran A, Mukamel S (2004) Proc Natl Acad Sci USA 101:506
31. Besley NA, Metcalf KA (2007) J Chem Phys 126:035101
32. Besley NA, Bryan JA (2008) J Phys Chem C 112:4308
33. Jin S, Head JD (1994) Surf Sci 318:204
34. Calvin MD, Head JD, Jin S (1996) Surf Sci 345:161
35. Head JD (1997) Int J Quantum Chem 65:827

36. Head JD, Shi Y (1999) Int J Quantum Chem 75:815
37. Head JD (2000) Int J Quantum Chem 77:350
38. McCall M, Brown T, Kennard O (1985) J Mol Biol 183:385
39. Ghysels A, Van Speybroeck V, Verstraelen T, Van Neck D, Waroquier M (2008) J Chem Theory Comput 4:614
40. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JJA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, revision B.01, Gaussian, Inc., Wallingford
41. Zhao Y, Truhlar DG (2008) Theor Chem Acc 120:215
42. Hohenstein EG, Chill ST, Sherrill CD (2008) J Chem Theory Comput 4:1996
43. Rutledge LR, Wetmore SD (2010) Can J Chem 88:815
44. Cances E, Mennucci B, Tomasi J (1997) J Chem Phys 107:3032
45. Watson JD, Crick FH (1953) Nature 171:737
46. Santamaria R, Charro E, Zacarías A, Castro M (1999) J Comput Chem 20:511
47. Chandra AK, Nguyen MT, Uchimaru T, Zeegers-Huyskens T (1999) J Phys Chem A 103:8853
48. Pelmenschikov A, Hovorun DM, Shishkin OV, Leszczynski J (2000) J Chem Phys 113:5986
49. Ha TK, Gunthard HH (2001) Spectrochim Acta Part A 57:55
50. Kwiatkowski JS, Leszczynski J (1996) J Phys Chem 100:941
51. Szczepaniak K, Szczesniak M (1987) J Mol Struct 156:29
52. Szczesniak M, Szczepaniak K, Kwiatkowski JS, KuBulat K, Person WB (1988) J Am Chem Soc 110:8319
53. Nowak MJ, Lapinski L, Fulara J (1989) Spectrochim Acta Part A 45:229
54. Tena GN, Baranov VI (2005) J Appl Spectr 72:155
55. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2004) Gaussian 03. Gaussian, Inc., Wallingford
56. Alecu IM, Zheng J, Zhao Y, Truhlar DG (2010) J Chem Theory Comput 6:2872
57. Scholes GD, Curutchet C, Mennucci B, Cammi R, Tomasi J (2007) J Phys Chem B 111:6978