NJC

RSC Publishing

PAPER

Applying vibrational spectroscopy to the study of nucleobases – adenine as a case-study†

Cite this: *New J. Chem.*, 2013, **37**, 2691

Rui P. Lopes,[a] Rosendo Valero,[a] John Tomkinson,[b] M. Paula M. Marques*[ac] and Luís A. E. Batista de Carvalho[a]

Received (in Montpellier, France) 29th April 2013,
Accepted 4th June 2013

DOI: 10.1039/c3nj00445g

www.rsc.org/njc

A full conformational study of solid-state anhydrous adenine is reported, using vibrational spectroscopy techniques coupled to DFT calculations, for the isolated molecule and the solid. In both cases, the N9H-amino tautomer was found to be the predominant species, followed by the N7H-amino form. An excellent agreement was achieved between experiment and theory, both for wavenumbers and intensities (without the need for scaling). A complete spectral assignment was performed, since all vibrational spectroscopic techniques were available to this study – FTIR, Raman and INS – allowing us to detect and interpret even the lowest frequency vibrational bands, not previously accessed. The quantum mechanical calculations presently carried out represent the highest theoretical level applied so far to the study of nucleobases.

1. Introduction

Nucleic acid bases are fundamental biological entities, as building blocks of the genetic code. Apart from their presence in DNA and RNA nucleosides, the purine bases adenine and guanine, in particular, play a major role as structural constituents of the second messengers cAMP and cGMP, and are often involved in mutations leading to carcinogenesis. In addition, they are the preferred biological targets of platinum-based drugs (such as cisplatin, carboplatin or oxaliplatin), thus mediating the cytotoxic effect of these anticancer agents.[1,2]

Therefore, the exact knowledge of their structure and conformational behaviour under physiological conditions is essential for understanding the biochemical processes in which they are involved, as well as the mechanisms of action of metallochemotherapeutics. Moreover, a successful rational design of new and more effective metal-based antineoplastic compounds relies on the accurate structural knowledge of these specific receptors.

In recent years, there has been growing interest in characterising DNA bases using spectroscopic experimental techniques coupled to dedicated theoretical methods, with a view to develop a model capable of assisting the study of larger systems comprising these building blocks (such as nucleotides and nucleic acids). Nucleobases have a very large range of protonation and tautomeric species, which justifies the difficulty in predicting their stability and relative population. Even using advanced spectroscopic methods, the subtle conformational changes that occur upon tautomeric equilibria are difficult to grasp, which renders the spectral assignment a complex task.

Regarding adenine (9*H*-purin-6-amine, $C_5H_5N_5$), work has been developed using Fourier transform infrared (FTIR) and Raman spectroscopic techniques aiming at elucidating its conformational preferences.[3–9] Fourteen different tautomeric species were found, both for the gas-phase and the aqueous solution,[10] comprising four amino and ten imino forms. In fact, adenine's tautomeric equilibrium has been studied in great detail and at different theoretical levels.[10–13] The most stable tautomer both for the isolated molecule and the solid was shown to be the N9H-amino (N9H), followed by the N7H-amino (N7H)[14,15] (Fig. 1), with an energy difference of *ca.* 36.0 kJ mol^{-1} (at the B3LYP/6-311G** level, for the gas phase).[14] X-ray data for polycrystalline adenine[16,17] and neutron diffraction results for the anhydrous compound[18] unequivocally evidenced that the N9H tautomer is its canonical form in the solid state. Studies on the determination of the exact tautomeric composition of adenine in the condensed phase are scarce to date, owing to the difficulty in interpreting spectroscopic results due to intermolecular interactions occurring in the solid lattice.[6]

In aqueous solution, while initial calculations supported that adenine should exist as a mixture of N7H and N9H tautomers,[19] later studies proposed N7H as the prevalent species due to its

[a] *Research Unit "Molecular Physical Chemistry", Faculty of Science and Technology, University of Coimbra, 3004-535 Coimbra, Portugal. E-mail: labc@ci.uc.pt; Fax: +351 239826541*
[b] *ISIS Facility, SFTC Rutherford Appleton Laboratory, Chilton, Didcot, OX 11 0QX, UK*
[c] *Department of Life Sciences, Faculty of Science and Technology, University of Coimbra, 3001-401 Coimbra, Portugal*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3nj00445g

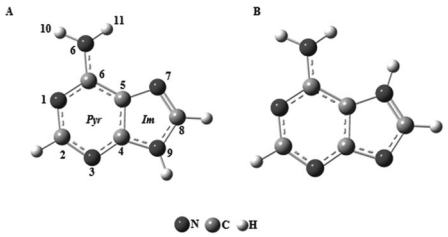

Fig. 1 Structural representation of the calculated (DFT/B3LYP**) N9H-amino (A) and N7H-amino (B) tautomeric forms of adenine, in the gas phase. (The atom numbering is included. Pyr and Im refer to the pyrimidine and imidazole rings, respectively.)

larger dipole moment.[20] However, a recent work by Raman and resonance Raman (RR) techniques, coupled to quantum mechanical calculations, unequivocally determined the tautomeric composition of adenine in solution (at pH 3):[9] N9H coexists with the N7H and N9HN1H$^+$ tautomeric species, the latter being the predominant one.

Vibrational spectroscopy is a reliable and accurate method for determining this type of conformational preferences, since it permits the analysis of samples in both the solid state and the solution (including aqueous solution), for distinct conditions (*e.g.* concentration, pH, ionic strength and temperature), hence allowing us to reproduce the physiologic medium.[21,22] Inelastic Neutron Scattering (INS) spectroscopy, in particular, is a well-suited technique to the study of hydrogenous compounds such as adenine. Actually, the neutron scattering cross-section of an atom (σ) is characteristic of that atom and independent of its chemical environment. Since the value for hydrogen (80 barns) far exceeds that of all other elements (typically *ca.* 5 barns), the modes of significant hydrogen displacement (u_i) dominate the INS spectra. For a mode at a given energy ν_i, the intensity from a powdered sample obeys the simplified relationship

$$S_i^{\bullet}(Q^2, \nu_i) = \frac{(Q^2 u_i^2)\sigma}{3} \exp\left(-\frac{Q^2 \alpha_i^2}{3}\right) \quad (1)$$

where Q (Å^{-1}) is the momentum transferred from the neutron to the sample and α_i (Å) is related to a weighted sum of all the displacements of the atom. INS is not limited by selection rules, and yields both the energies (the eigenvalues, ν_i) and the atomic displacements (the eigenvectors, u_i) of the vibrational transitions. This significantly enhances the information obtainable from the vibrational spectrum and adds to that from the complementary Raman and infrared vibrational spectroscopic methods, allowing us to detect some low frequency modes unavailable to these optical techniques. Since the spectral intensities can be quantitatively compared with those calculated using theoretical methods, by combining the INS results with quantum mechanical molecular orbital calculations it is possible to link molecular geometry with the experimental spectroscopic features, and produce a consistent conformation for the systems under investigation.

Coupling the suitability of INS spectroscopy to detect low-wavenumber modes (below 1000 cm^{-1}, normally due to out-of-plane molecular vibrations) to Raman and FTIR (Fourier Transform Infrared) techniques, which enable the higher frequency vibrational modes to be accessed, leads to a total vibrational assignment, in the whole spectral range of interest. Regarding adenine, no such complete vibrational assignment has been reported to date. Also, no theoretical simulations were performed for this nucleobase in the solid state, in spite of the wealth of information that can be retrieved for the condensed phase from periodic density functional calculations – both geometrical and spectroscopical.

The present work aims to fill this gap, by carrying out a full conformational analysis of the two main tautomeric forms of adenine (N7H-amino and N9H-amino, Fig. 1), using all available vibrational spectroscopic techniques – FTIR, Raman and INS – coupled to quantum mechanical calculations – at the Density Functional Theory and Plane Wave levels. This paper is thus a significant contribution to the existing spectroscopic characterization of adenine, by unambiguously interpreting its complete vibrational profile. Additionally, the INS data presently reported were obtained using the TOSCA spectrometer of the ISIS pulsed neutron and muon source (United Kingdom), which represent a substantial improvement relative to the previously reported results that were acquired in the former TFXA configuration of this spectrometer.[3]

2. Experimental

2.1. Quantum mechanical calculations

The quantum mechanical calculations for the isolated molecule were performed using the GAUSSIAN 03W program[23] within the Density Functional Theory (DFT) approach, in order to account for the electron correlation effects. The widely employed hybrid method denoted by B3LYP, which includes a mixture of HF and DFT exchange terms and the gradient-corrected correlation functional of Lee, Yang and Parr[24] as proposed and parameterised by Becke,[25,26] was used, along with the double-zeta split valence basis set 6-31G**.[27] Molecular geometries were fully optimised by the Berny algorithm, using redundant internal coordinates:[28] the bond lengths to within *ca.* 0.1 pm and the bond angles to within *ca.* 0.1°. The final root-mean-square (rms) gradients were always less than 3×10^{-4} Hartree per Bohr or Hartree per radian. No geometrical constraints (*e.g.* planarity) were imposed on the molecule.

For the solid state, plane-wave expansions, as implemented in the PWSCF code from the Quantum Espresso package,[29] were used, based on DFT methods within the Perdew–Zunger local density approximation (LDA).[30] The atomic coordinates were fully optimised using the published crystal structure of anhydrous adenine as a starting point.[31] Anhydrous adenine crystallizes in a centrosymmetric monoclinic space group ($P2_1/c$) with 8 molecules in the unit cell (z = 8). The unit cell dimension vectors were conserved during the optimisation process. The pseudopotentials employed were of the norm-conserving type – a Von Barth–Car approach[32] was applied to the H and C atoms, and a Martins–Troullier[33] type was used for the N atoms. This choice of methods was guided by the fact that Raman activities can only be calculated with PWSCF methods,

using a DFT-LDA approach and norm-conserving pseudopotentials. A cutoff energy of 70 Ry and a Monkhorst–Pack grid[34] of 3 × 3 × 3 were found sufficient to attain convergence. The dynamical matrix was calculated for the optimised geometries within the Density Functional Perturbation theory,[35] and was diagonalised to obtain the vibrational normal mode wavenumbers, as well as the Raman activities, S_i. No scaling factors were applied to these calculated frequencies.

The harmonic vibrational wavenumbers, as well as the Raman activities and infrared intensities, were obtained at the same theory level as the geometry optimisation, and were scaled according to Merrick *et al.*[36] Raman activities, S_i, in particular, are straightforwardly derived from the program output and cannot be compared directly with the experiment. The theoretical Raman intensity was calculated according to the equation

$$I = C(\nu_0 - \nu_i)^4 \frac{S_i}{\nu_i} \quad (2)$$

C being a constant and ν representing frequency values. In order to simulate the linewidth of the experimental lines, an artificial Lorentzian broadening was introduced using the SWizard program (revision 4.6).[37,38] The Raman band half-widths were taken as 15, 25 and 30 cm^{-1}, respectively, below 1250 cm^{-1} between 1250 and 2000 cm^{-1} and above 2000 cm^{-1}. The theoretical INS transition intensities were obtained from the calculated normal mode eigenvectors and the spectra were simulated using the dedicated aCLIMAX program.[39]

2.2. Vibrational spectroscopy

The Fourier transform infrared (FTIR) spectrum of adenine (Sigma, anhydrous, ≥99%) was recorded using a Bruker Optics Vertex 70 FTIR spectrometer, in the range of 400–4000 cm^{-1}, using KBr disks (*ca.* 1% (w/w)), a KBr beamsplitter, and a liquid nitrogen cooled Mercury Cadmium Telluride (MCT) detector. The spectra were collected for 2 minutes (*ca.* 140 scans), with a 2 cm^{-1} resolution. Under these conditions, the error in wavenumbers was estimated to be less than 1 cm^{-1}.

The Raman spectra of adenine were obtained using a triple monochromator Jobin-Yvon T64000 Raman system (focal distance 0.640 m, aperture f/7.5), equipped with holographic gratings of 1800 grooves per mm. The spectral region above 200 cm^{-1} was acquired with the two initial monochromators (the premonochromator stage) in the subtractive mode. The detection system was a liquid nitrogen cooled non-intensified 1024 × 256 pixel (1″) Charge Coupled Device (CCD) chip. The entrance slit was set to 100 μm, while the slit between the premonochromator and the spectrograph was set to 400 μm. The lowest energy spectral region (20 to 200 cm^{-1}) was acquired with the triple monochromator configuration (additive mode), using a thermoelectrically cooled Burle C31034 photomultiplier tube (PMT) detector. In this configuration the entrance slit was set to 100 μm. In all cases, a 90° geometry between the incident radiation and the collecting system was employed. Under the above-mentioned conditions, the error in wavenumbers was estimated to be less than 1 cm^{-1}. The 514.5 nm line of an Ar^+ laser (Coherent, model Innova 300-05) was used as the excitation radiation, providing *ca.* 15 mW at the sample position (in order to avoid strong fluorescence emission). The spectra were recorded at room temperature, and the sample was sealed in a Kimax glass capillary tube of 0.8 mm inner diameter.

INS spectra were obtained in the Rutherford Appleton Laboratory (UK), at the ISIS pulsed neutron source, using the TOSCA spectrometer. This is an indirect geometry time-of-flight, high resolution (($\Delta E/E$) *ca.* 1.25%), broad range spectrometer.[40] The samples (Sigma-Aldrich, anhydrous, 99.9+%), weighing 2–3 grams, were wrapped in aluminium foil to make a 4 × 4 cm sachet and placed in thin walled aluminium cans, which filled the beam. To reduce the impact of the Debye–Waller factor (the exponential term in eqn (1)) on the observed spectral intensity, the samples were cooled to *ca.* 20 K. Data were recorded in the energy range of 16 to 4000 cm^{-1} and converted to the conventional scattering law, $S(Q, \nu)$ *vs.* energy transfer (in cm^{-1}) through standard programs.

3. Results and discussion

3.1. Conformational analysis

The optimised geometries of the N9H-amino and N7H-amino tautomers of adenine, in the gas phase, are represented in Fig. 1, while Fig. 2 shows the optimised crystal cell structure of anhydrous adenine. Table S1 (ESI†) comprises the geometrical parameters presently calculated both for the isolated molecule and the solid, as well as from its asymmetric unit previously obtained by X-ray diffraction.[31]

The lowest energy conformation calculated for isolated adenine, at the B3LYP/6-31G** level, is the N9H-amino tautomer (Fig. 1(A)). The asymmetric unit of anhydrous adenine comprises two geometrically non-equivalent molecules (Fig. 2(A)) which, by symmetry operations under a centrosymmetric monoclinic $P2_1/c$ space group, originate a unit cell fitting eight centroids (Fig. 2(B)). Two molecules in the unit cell are

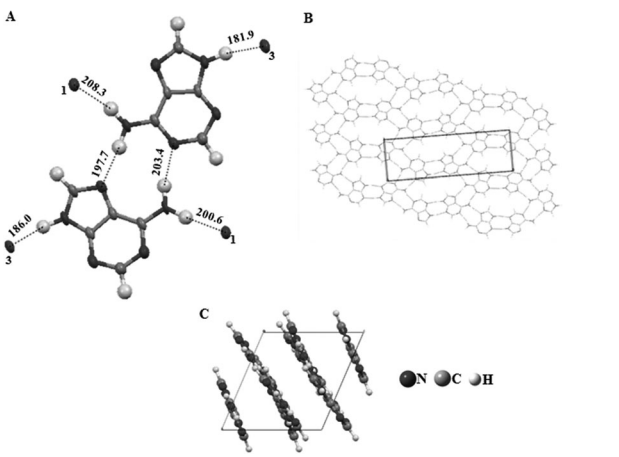

Fig. 2 (A) Optimised asymmetric unit of anhydrous adenine using the LDA functional and the PW methodology, showing the non-geometrically equivalent I and II arrangements. (The dashed lines represent intermolecular H-bonds (distances in pm).) (B) Longitudinal view, along the *c* axis, showing the crystalline lattice arrangement including the unit cell. (C) Lateral view, along the *b* axis, of the unit cell.

connected by two N–H···N hydrogen close contacts, and the crystal lattice is stabilised by two sets of N–H···N bonds – one across the centre of symmetry connecting the imidazole moiety (Im) to pyrimidine (Pyr) and the other linking the next asymmetric unit. Thus, in the condensed phase, adenine displays three intermolecular H-bonding donor sites (NH_2 and N9H) and three intermolecular acceptor sites (N1, N7 and N3 atoms).

The sum of the van der Waals contact radii for the nitrogen and hydrogen atoms (N···H) is found to be 275 pm,[41] and the calculated intramolecular N1···H10 distance for the two adenine molecules (I and II, Fig. 2(A)) in the asymmetric unit is 250.4 and 250.8 pm, respectively. The same distance is predicted for the isolated molecule to be 250.0 pm. These measurements suggest the presence of an intramolecular H-bond between H10 and N1. The N7···H11 intramolecular distance, in turn, is calculated to be 287.8 and 287.9 pm, respectively, in the condensed phase, and 279.8 pm in the isolated molecule, therefore excluding the presence of intramolecular H-bonds.

The calculated geometrical data presented in Table S1 (ESI†) evidence that isolated adenine is quasi-planar except for the amino group, which lies out of the plane defined by the Pyr–Im rings, similar to guanine.[42] In the case of adenine, however, the dihedral angles defining the position of the hydrogen atoms relative to the plane of the rings have almost the same value for both H10 and H11 (*ca.* 10.5° relative to the Pyr plane). This predicted value is considerably different from previously reported results obtained at the MP2/6-31G* level, yielding dihedrals of 16.46° and 17.96° out-of-Pyr plane, respectively.[43]

Interestingly, condensed phase calculations yield an almost planar adenine molecule and are in perfect agreement with the measurements on the asymmetric unit (*ca.* less than 1° out-of-Pyr plane for H10 and H11 atoms, Table S1, ESI†). It is also surprising that, despite having a less extensive H-bonding network than guanine (each adenine molecule contains three H-bond donor sites as compared to four in guanine), adenine is also quasi-planar in the anhydrous state. Thus, as previously found for guanine,[42] the influence of the H-bonding profile on the amine group planarity must be considered with caution: the higher NH_2 planarity in the solid as compared to the isolated molecule does not seem to be significantly affected by this type of close contacts, widely present in the crystal lattice.

The DFT calculated bond lengths involving H atoms in adenine are larger than the X-ray values.[31] In fact, since X-ray diffraction locates only electron density and not nuclear positions, the H atoms in the X-ray determined structure are systematically and considerably shifted along the X–H bond (X = N or C) towards the heavier atom. The amount of this displacement depends mainly on the nature of X (essentially on its electronegativity), and to a smaller extent on temperature due to vibration effects (atomic displacements). Typically, the shortening of the X–H bond will be about 10 to 20 pm: representative N–H refined distances from X-ray diffraction are around 84.0 pm *versus* calculated values of *ca.* 104 pm (Table S1, ESI†).

Regarding the conformation of the Pyr and Im rings within the nucleobase molecule, the calculated dihedrals (both for the gas and the condensed phase) show that these rings are almost but not perfectly planar (Table S1, ESI†): with a shift of *ca.* 0.3° for isolated adenine and *ca.* 0.5–1.1° for the solid, the latter being in a better agreement with the measured dihedrals on the asymmetric unit, as expected. Actually, this non-rigidity is a well recognised feature of aromatic heterocycles.[44] It is worth commenting on the relative stability of the N9H and N7H tautomers. Our B3LYP/6-31G** relative energies for the N9H and N7H tautomers (32.7 kJ mol^{-1} favouring N9H) and those reported in ref. 14 (36 kJ mol^{-1} at the B3LYP/6-311G** level and 33 kJ mol^{-1} at the MP2/6-311G** level) and ref. 43 (*ca.* 37 kJ mol^{-1} for B3LYP and 38–39 kJ mol^{-1} for MP2, depending on the basis set) are accurate enough for determining the accuracy of vibrational frequencies calculated for the isolated-molecule. Nevertheless, the main goal of the present study is the vibrational analysis of adenine in the solid state, which renders the results obtained at the plane-wave level (for the condensed phase) much more relevant.

3.2. Vibrational analysis

A complete spectral assignment of adenine was achieved, through combination of optical and acoustic vibrational data and assignment of these data in the light of the DFT predicted spectra, for both the solid and the isolated molecule. The experimental vibrational spectra presently obtained for adenine – FTIR, Raman and INS – are represented in Fig. 3–5, while Table 1 contains the corresponding wavenumbers and the corresponding calculated values and assignments.

Isolated adenine has 39 vibrational modes, 27 in-plane and 12 out-of-plane. The application of the PW methodology to the unit cell originates 360 harmonic vibrational frequencies that

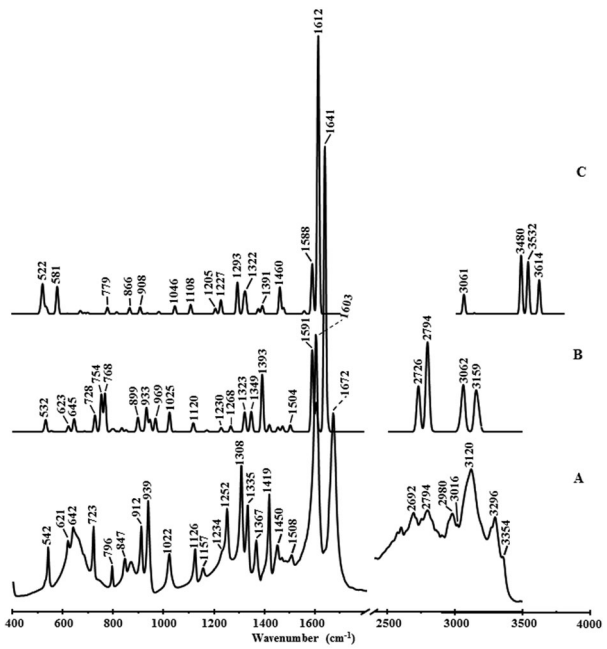

Fig. 3 Infrared spectra of anhydrous adenine (in the solid state). Experimental (A); calculated, for the condensed state (B) and for the isolated molecule (C). (The band at 1603 cm^{-1} in (A) was used as a reference for vertical scaling.)

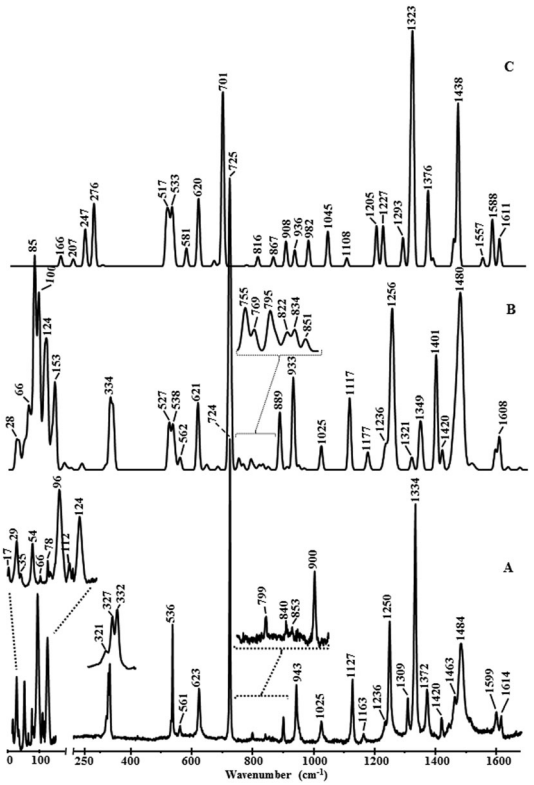

Fig. 4 Raman spectra of anhydrous adenine. Experimental (A); calculated, for the condensed state (B) and for the isolated molecule (C). (The band at 724 cm^{-1} in (A) was used as a reference for vertical scaling. The low frequency region (20–160 cm^{-1}) in (A) is expanded in order to allow the visualisation of spectral details, which are otherwise undetectable.)

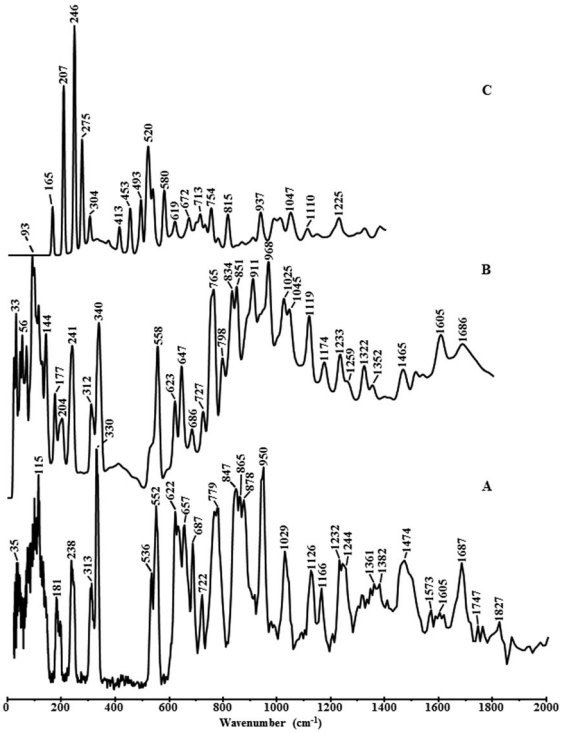

Fig. 5 Solid-state INS spectra of anhydrous adenine. Experimental (A); calculated, for the condensed state (B) and the isolated molecule (C).

can be numerically arranged in sets of 8 (since the unit cell fits eight centroids). Within each of these sets (or ''packs'') of grouped frequencies, however, the assignment is not uniform – there are small differences in the individual assignments, resulting from the presence of two non-equivalent entities in the asymmetric unit. Such an arrangement – each set of 8 frequencies becomes subdivided in a four plus four pattern – leads to a complex description of the harmonic vibrations in the solid and hinders an individual assignment for each of the 360 normal modes. Hence, for the sake of clarity, the PW calculated frequencies included in Table 1 are grouped into several intervals defined by the first and last of each set of 8 wavenumbers. A quite good agreement was obtained between the calculated and experimental frequencies (without any scaling). Scaling factors were applied solely to the isolated-molecule calculated frequency values since in this case standard factors are available for many density functionals and basis sets, while they do not exist for the LDA functional employed in the present plane-wave calculations. More importantly, for isolated molecules the difference between theoretical and experimental vibrational frequencies is highly systematic, unlike for solid-state systems (due to the numerous interactions experienced by each given molecule). Furthermore, all the experimental peaks could be unambiguously assigned to normal vibrational modes without the need for scaling. It is noteworthy that the present vibrational analysis differs from earlier reported studies[5,14,45] (especially in the low wavenumber region), since these considered a planar (C_s) calculated conformation for adenine which is a saddle point.

The vibrational features of adenine comprise three main groups: (i) the low frequency modes, up to 950 cm^{-1}, where mostly out-of-plane deformations are present, frequently coupled to movements that involve all the atoms in the molecule. The lowest wavenumber bands within this interval (below 350 cm^{-1}) are ascribed to skeletal deformations (*e.g.* ring torsion) and lattice vibrations, mainly detected by INS. (ii) The intermediate region (950 to 1700 cm^{-1}), characterised by the in-plane bending modes of ring atoms, which become extensively mixed with C–N and C–C stretching modes rendering their accurate assignment a quite difficult task. (iii) The high wavenumber range (2700 cm^{-1} to 3400 cm^{-1}) comprising the stretching vibrations involving the hydrogen atoms.

A very good quality experimental Raman spectrum was obtained for adenine (Fig. 4), as compared to the data found in the literature.[46,47] Particularly remarkable is the lower frequency spectral region (between 20 and 180 cm^{-1}, Fig. 4(A)), never assessed before, comprising lattice modes (librations). The experimental features between 200 and 1800 cm^{-1} are very well predicted by the PW approach, both for the frequencies and the intensities (Fig. 4(B) and Table 1), with the exception of the band at 1334 cm^{-1} (considerably underestimated in intensity). The calculated Raman spectra for the isolated molecule, as expected, do not show such a good accordance with the measured values, although the overall intensity distribution compares well with the experimental one.

Table 1 Experimental and calculated wavenumbers for anhydrous adenine

Experimental			Calculated		
INS	Raman	FTIR	[a]G03W	[b]PW	[c]Approximate description
	17				
	29			26–49	Lattice translation
35	35				
	54			46–69	Dimer translation
	66				
75	78				
	81				
100	96				
115	112			74–153	Dimer translation
118	117				
126	128				
133					
181					
197			166	177–209	γ(C6-N6)
238			207	229–249	[d]"Butterfly" mode
313	321		305	309–323	Γ(N1-C2-N3); Γ(C8-N7-C5)
330	327		276	332–349	Δ(C6-C5-N7) − Δ(N1-C6-N6) + Δ(N3-C4-N9)
	332		532	524–533	Δ(N1-C6-C5) + Δ(C2-N3-C4)
536	537	542	621	538–555	Δ(N6-C6-C5) + Δ(C6-C5-C4) + Δ(N3-C4-N9) − Δ(C6-C5-N7) − Δ(N1-C6-N6)
552	561		581	557–564	Γ(N1-C2-N3) + γ(C2-H) − γ(C8-H)
622	623	621		620–626	Δ(C4-C5-N7) − Δ(C4-N9-C8) + Δ(N1-C6-N6)
657			672	643–651	Γ(N7-C8-N9); Γ(C8-N7-C5); Γ(C8-N9-C4); γ(C8-H)
687		642	688	685–689	C4-C6 [d]"umbrella"; γ(N6-H11)
722	724	723	701	724–730	Pyr + Im ring breathing
779	799	796	247	754–770	ω(NH_2)
847					
865	840	847	779 816	794–823	γ(C8-H); C4 + C6-C5 [d]"umbrella"; τ(NH_2)
878		872	515 539	833–852	τ(NH_2)
918	900	912	867	888–900	Δ(N1-C2-N3) + Δ(C8-N9-C5)
			937	905–916	γ(C2-H)
950	943	939	908	932–936	Δ(N7-C8-N9); γ(N9-H)
			523	947–969	γ(N9-H)
1029	1025	1022	982	1024–1048	t(NH_2); ν(C6-N1)
1126	1127	1126	1108	1117–1123	δ(C8-H); ν(C3-N3) + ν(C4-N9) − ν(C5-C6) − ν(C5-N7)
1166	1163	1157	1046	1170–1184	δ(C8-H) − δ(N9-H); ν(C8-N9) + (C5-N7) − ν(C4-N9) − ν(C4-N3) − ν(C2-N3)
1232	1236	1234	1206	1230–1237	ω(NH_2); ν(C5-N7) + ν(C8-N7) − ν(C6-N1); δ(C8-H); δ(C2-H); δ(N9-H)
1244	1250	1252	1227	1253–1274	δ(C8-H) + δ(N9-H); ν(C2-N3) + ν(C8-N7) + ν(C2-N1) − ν(C5-N7) − ν(C4-N9)
1314	1309	1308	1293	1320–1355	δ(C2-H); ν(C2-N3) + ν(C5-N7) − ν(C8-N7) − ν(C4-N9)
1361	1334	1335	1319 1325		
1382	1372	1367	1375	1392–1404	δ(N9-H); δ(C8-H); ν(C8-N9) + ν(C4-C5) + ν(C2-N3) + ν(C6-N1) − ν(C5-C6) − ν(C2-N1) − ν(C4-N3) − ν(C8-N7)
1410	1420	1419	1391	1419–1422	ν(C4-N9) + ν(C8-N7) + ν(C6-N1) − ν(C4-C5) − ν(C4-N3) − ν(C8-N9)
1474	1463	1450	1460	1459–1473	δ(N9-H) − δ(C8-H) + δ(C2-H); ν(C8-N9) + ν(C2-N1) + ν(C6-N1) − ν(C2-N3) − ν(C8-N7) − ν(C6-N6)
1496	1484	1508	1475	1480–1519	α(NH_2); δ(N9-H) − δ(C8-H) + δ(C2-H); ν(C4-C5) + ν(C2-N1) + ν(C6-N1) + ν(C8-N9) − ν(C2-N3) − ν(C4-N9) − ν(C8-N9)
1605	1599	1603	1588 1556	1588–1608	α(NH_2); δ(N9-H); ν(C4-C5) + ν(C2-N1) − ν(C5-C6) − ν(C6-N1) − ν(C4-N3) − ν(C2-N3)
1687	(1614)	1672	1612	1608–1676	α(NH_2) − ν(C6-N6) + ν(C4-C5) + ν(C2-N1) − ν(C5-C6) + ν(C6-N1)
		2692			Combination mode
		2794			Combination mode
		2980			Combination mode
	3039	3016	3532	2726–2835	ν(N9-H)
	3129	3120	3060	3043–3048	ν(C2-H)
	3273	3267	3480	3034–3092	ν_s(NH_2)
	3300	3296	3137	3142–3159	ν(C8-H)
	3356	3354	3614	3154–3179	ν_a(NH_2)

[a] Calculated at the DFT B3LYP/6-31G** level of theory, scaled according to ref. 17. [b] Calculated using the LDA functional and the PW methodology, unscaled. These frequencies are grouped into intervals defined by the first and last of each set of 8 frequencies. [c] According to the PW results. ω: wagging; δ: in-plane deformation; Δ: in-plane ring deformation of skeletal atoms; γ: out-of-plane deformation; Γ: out-of-plane ring deformation of skeletal atoms; α: scissoring; τ: torsion; t: twisting; ν_s: symmetric stretching; ν_a: antisymmetric stretching. Atoms are labeled according to Fig. 1. [d] Fig. 6.

Regarding the FTIR spectrum (Fig. 3), while experimental to PW calculated agreement is quite good, the intensities are not so well reproduced by theory as for Raman. This may be due to the hydrogen bonding network in the solid, which affects the infrared pattern more significantly and leads to a broadening of the observed bands. In fact, the presently calculated IR data

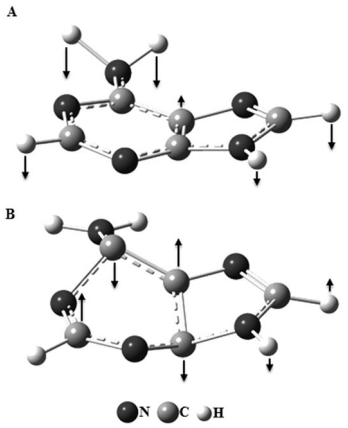

Fig. 6 Schematic representation of the low-frequency butterfly mode (A) and the out-of-plane symmetric deformation – umbrella mode (B) – for isolated adenine.

show a better agreement with the spectrum previously obtained by Nowak and coworkers[5] in an argon matrix, within which the intermolecular H-bonding is greatly reduced owing to a larger spatial separation between adenine molecules (yielding sharper lines).

The INS spectrum recorded for adenine is represented in Fig. 5, and constitutes a significant improvement relative to the previously reported one,[3] which was acquired in 1992 at the previous version of the TOSCA spectrometer (TFXA) at ISIS (UK)[48] with a far lower resolution and signal-to-noise ratio. PW simulations of the INS profile evidence an excellent accordance with the measured data: the calculated INS eigenvectors closely align with the experimental intensities, even for the higher frequencies (above 1000 cm^{-1}). Interestingly, the calculated to experimental agreement in the lowest frequency region (below 250 cm^{-1}) is better for adenine than for guanine,[42] which may result from a more complex H-bonding network in the latter.

The low frequency spectral range comprises characteristic vibrations, such as the ''butterfly''-like motion of the Pyr and Im rings (Fig. 6(A)). This mode (not detected by Raman) is ascribed to the 238 cm^{-1} INS band (Fig. 5(A)), and is predicted between 229 and 249 cm^{-1} for the solid (Fig. 5(B)) and at 207 cm^{-1} for the isolated molecule (Fig. 5(C)). This agrees with earlier assignments[33] that reported this feature at 249 cm^{-1} in Raman. Other studies, however, calculated this mode at 200 cm^{-1} (MP2/6-31G*) and assigned it to the experimental 193 cm^{-1} Raman band.[2]

The most intense INS band, at 330 cm^{-1} (Fig. 5), and the feature at 313 cm^{-1} are unequivocally related to the 321, 327 and 332 cm^{-1} Raman signals (Table 1 and Fig. 4). Normal coordinate analysis based upon PW calculations shows a contribution from the out-of-plane deformation modes (N1-C2-N3) and (C8-N7-C5) to the 313 cm^{-1} INS band, and from several in-plane skeletal deformations to the most intense 330 cm^{-1} signal (Table 1). Regarding the latter, it is interesting to point out that whereas Δ(C6-C5-N7) and Δ(N1-C6-N6) occur simultaneously, but each of them in one adenine molecule of the asymmetric unit (either I or II, Fig. 2(A)), the (C6-C5-N7) angle decreases in adenine I while (N1-C6-N6) increases in adenine II. This leads to a strong in-plane displacement of the amino group, which originates a change in the N6-H10$\cdots$N7 and N6-H11$\cdots$N1 hydrogen bond lengths – ''H-bond effect'', also observed for guanine.[42] Such an effect, along with the strong in-plane NH_2 displacement, may account for the very strong intensity of this INS band. This vibration was previously reported in the Raman spectrum at 329 cm^{-1} and inaccurately assigned to the Pyr puckering mode.[43] Studies in aqueous solution, in turn, feature this mode at 300 cm^{-1} in Raman.[49]

The two very strong INS bands at 536 and 552 cm^{-1} are straightforwardly correlated with the infrared feature at 542 cm^{-1} (Fig. 3(A)) and the Raman signals at 536 and 561 cm^{-1}, respectively (Fig. 4 and Table 1), the former being PW-calculated as a splitting at 527/538 cm^{-1} (Fig. 4(B)). These modes are assigned to in-plane and out-of-plane skeletal deformations mainly involving the Pyr ring.

Other remarkable features of adenine's vibrational pattern are the ''umbrella''-type Pyr deformation modes (687 and 865 cm^{-1}, Fig. 6(B)), and the 724 cm^{-1} Raman line (predicted at *ca.* 725 and 701 cm^{-1} for the solid and the isolated molecule, respectively, Fig. 4) resulting from symmetric stretching modes of all Pyr/Im atoms and usually defined as the ''breathing'' mode (Table 1).

A series of features are detected in the INS spectrum between 600 and 750 cm^{-1}, with decreasing intensity (Fig. 5). The one at 622 cm^{-1} correlates with the Raman line at 623 cm^{-1} (observed in FTIR at 621 cm^{-1}, Table 1) and is ascribed to the Pyr ring in-plane deformation. Those at 657 and 687 cm^{-1}, which relate to the very broad FTIR band centered at *ca.* 642 cm^{-1} and to the calculated modes for the isolated molecule at 672 and 688 cm^{-1} (Table 1), are assigned to the out-of-plane deformations of the Im ring. Another interesting feature in the INS spectrum is the one comprising three components at 847, 865 and 878 cm^{-1}, correlated with the PW calculated eigenvectors at 834/851 cm^{-1} (Fig. 5(B)) but not predicted within this region for the isolated molecule (Fig. 5(C)). Based on the PW approach and attending to its high intensity, these INS bands are ascribed to the out-of-plane motions involving H atoms such as γ(C8-H) and NH_2 wagging. These vibrations are calculated for the gas phase at 515 and 539 cm^{-1} (Table 1), underestimated by more than 300 cm^{-1} relative to the solid. The same occurs for the strong 779 cm^{-1} INS signal (Fig. 5(A)) due to NH_2 wagging, which is predicted for isolated adenine at 247 cm^{-1}.

The 900 cm^{-1} Raman band (912 cm^{-1} in FTIR) results from the (N1-C2-N3) + (C8-N9-C5) deformation modes. These in-plane, highly symmetric, vibrations do not involve significant H displacements, giving rise to the very weak INS signal at 918 cm^{-1} (Fig. 5(A)). In contrast, the (C2)H and (N9)H out-of-plane deformations yield a very intense 950 cm^{-1} INS feature (Fig. 5(A)). The (N9)H motion is particularly interesting, as this group is involved in H-bond interactions: this vibration is correctly predicted by condensed phase calculations (between 905 and 969 cm^{-1}, Table 1) but underestimated by more than

400 cm^{-1} when considering the isolated molecule (523 cm^{-1}, Table 1). Earlier studies incorrectly assigned this band to the NH_2 rocking vibration.[46]

Above 1000 cm^{-1}, spectral transitions are dominated by in-plane modes: the 1250 cm^{-1} Raman band (Fig. 4(A)) correlated to the INS signals at 1232 and 1244 cm^{-1}, and to the broad 1252 cm^{-1} FTIR feature, is ascribed to the (C8)H in-plane deformation combined with the C–N stretching and NH_2 twisting modes (Table 1), in agreement with some previous studies[5,43] but in opposition to another reported work that ascribed these features exclusively to C–N stretching.[46] The very strong Raman feature at 1334 cm^{-1}, as well as those at 1309 and 1372 cm^{-1} are due to a mixture of C–N stretching modes and the in-plane (C2)H deformation. The broad 1484 cm^{-1} Raman band, in turn, is mainly ascribed to the overlapped C–N and C–C stretching modes of the Pyr and Im rings.[43] Since this correlates well with the intense INS signals at 1474 cm^{-1} (Fig. 5(A)), an additional out-of-plane vibration involving H atom displacements is suggested to contribute to these features. Such motion is in fact predicted by PW calculations as the NH_2 scissoring mode (along with contributions from N–H and C–H in-plane deformations). The NH_2 scissoring mode gives rise to an intense infrared band at 1672 cm^{-1} and an INS strong feature centered at 1687 cm^{-1} (Fig. 3(A) and 5(A)), which also have contributions from C–N stretching modes (Table 1) in agreement with earlier studies.[3,46]

In the high frequency spectral range, five bands are expected to occur due to the N–H/C–H stretching vibrations (including the $\nu_s(NH_2)$ and $\nu_a(NH_2)$, Table 1). In the light of the PW calculations, the observed FTIR signals at 3016 and 3120 cm^{-1} (3039 and 3129 cm^{-1} in Raman) are assigned to the N9-H and C2-H stretching modes, respectively. Seuvre and collaborators[46] related these vibrations to the 2980 and 2800 cm^{-1} infrared bands, while Mohamed *et al.*[43] associated them to the 3347 and 2980 cm^{-1} signals, respectively. The fact that the N9-H stretching motion is very prone to be affected by intermolecular H-bonds can explain the red shift from the reported value of 3347 cm^{-1} (ref. 43) to the presently measured frequencies of 3039 cm^{-1} (Raman) and 3016 cm^{-1} (FTIR) and the large difference between calculated values for the isolated molecule (3532 cm^{-1}) and the solid (2726–2835 cm^{-1}) (Table 1). In turn, the (C2)H stretching, virtually unaffected by intermolecular close contacts, is assigned to the 3120 cm^{-1} FTIR band (2800 cm^{-1} and 2980 cm^{-1} by Seuvre and Mohamed, respectively). Finally, it should be emphasised that there is no accordance regarding the assignment of the $\nu_s(NH_2)$ and ν(N9-H) bands among the studies performed to this date: $\nu_a(NH_2)$ has been related to either the 3016 cm^{-1} (ref. 46) or 3426 cm^{-1} (ref. 43) infrared features of the solid 9*H*-keto-amino adenine tautomer, in contrast to the present assignment at 3354 cm^{-1} (Table 1).

4. Conclusions

In the present work a full vibrational spectroscopic study of the N9H-amino tautomeric form of adenine was performed, in the light of DFT calculations performed for both the isolated molecule and the condensed phase. A complete and accurate assignment of the experimental spectra was achieved, due to the combination of all available spectroscopic vibrational techniques (FTIR, Raman and INS) with state-of-the-art theoretical approaches. Within the latter, condensed-phase plane-wave DFT calculations were used, yielding an excellent agreement between experimental and theoretical data both for wavenumbers and intensities.

The results thus obtained clearly evidence the need for using periodic approaches such as plane-wave approaches for the representation of this kind of molecules in the solid state. In particular, the low energy region of the spectrum, comprising external (lattice) modes, can only be accurately predicted through such a methodology.

In summary, this study represents the most complete vibrational assignment of anhydrous adenine published to date, including the lowest frequency spectral region and based on calculations performed at the highest theoretical level applied so far to the study of nucleobases.

Acknowledgements

The authors acknowledge financial support from the Portuguese Foundation for Science and Technology – PEst-OE/QUI/UI0070/2011. The INS work was supported by the European Commission under the 7th Framework Programme through the Key Action: Strengthening the European Research Area, Research Infrastructures (Contract no: CP-CSA_INFRA-2008-1.1.1 Number 226507-NMI3).

Notes and references

1 J. E. Sponer, J. Leszczynski, F. Glahé, B. Lippert and J. Sponer, *Inorg. Chem.*, 2001, **40**, 3269.
2 M. G. Añorbe, M. S. Lüth, M. Roitzsch, M. M. Cerdà, P. Lax, G. Kampf, H. Sigel and B. Lippert, *Chem.–Eur. J.*, 2004, **10**, 1046.
3 Z. Dhaouadi, M. Ghomi, J. Austin, R. Girling, R. Hester, P. Mojzes, L. Chinsky, P. Turpin and C. Coulombeau, *J. Phys. Chem.*, 1993, **97**, 1074.
4 A. Toyama, H. Takeuchi and I. Harada, *J. Mol. Struct.*, 1991, **242**, 87.
5 M. J. Nowak, L. Lapinski, J. S. Kwiatkowski and J. Leszczynski, *J. Phys. Chem.*, 1996, **100**, 3527.
6 G. N. Ten, V. V. Nechaev, N. B. Zotov and V. I. Baranov, *Opt. Spectrosc.*, 2009, **107**, 58.
7 T. Burova, G. Ten, V. Ermolenkov, R. Shcherbakov and I. Lednev, *Opt. Spectrosc.*, 2010, **109**, 853.
8 G. Ten, T. Burova, R. Shcherbakov and V. Baranov, *Opt. Spectrosc.*, 2010, **109**, 845.
9 T. G. Burova, V. V. Ermolenkov, G. N. Ten, R. S. Shcherbakov, V. I. Baranov and I. K. Lednev, *J. Phys. Chem. A*, 2011, **115**, 10600.
10 M. Hanus, M. Kabeláč, J. Rejnek, F. Ryjáček and P. Hobza, *J. Phys. Chem. B*, 2004, **108**, 2087.

11 C. F. Guerra, F. M. Bickelhaupt, S. Saha and F. Wang, *J. Phys. Chem. A*, 2006, **110**, 4012.
12 M. Y. Choi, F. Dong, S. W. Han and R. E. Miller, *J. Phys. Chem. A*, 2008, **112**, 7185.
13 Y. Xue, D. Xu, D. Xie and G. Yan, *Spectrochim. Acta, Part A*, 2000, **56**, 1929.
14 C. Plützer, E. Nir, M. Vries and K. Kleinermanns, *Phys. Chem. Chem. Phys.*, 2001, **3**, 5466.
15 K. Rajabi, K. Theel, E. A. L. Gillis, G. Beran and T. D. Fridgen, *J. Phys. Chem. A*, 2009, **113**, 8099.
16 T. J. Kistenmacher and T. Shigematsu, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1974, **30**, 1528.
17 T. C. Lewis and D. A. Tocher, *Acta Crystallogr., Sect. E: Struct. Rep. Online*, 2005, **61**, o1052.
18 R. J. Fram, *Curr. Opin. Oncol.*, 1992, **4**, 1073.
19 M. Majoube, M. Henry and G. Vergoten, *J. Raman Spectrosc.*, 1994, **25**, 233.
20 M. Majoube, P. Millie, L. Chinsky, P. Y. Turpin and G. Vergoten, *J. Mol. Struct.*, 1995, **355**, 147.
21 C. Lee, K.-H. Park and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114508.
22 C. Lee and M. Cho, *J. Chem. Phys.*, 2007, **126**, 145102.
23 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03*, Gaussian, Inc., Pittsburgh PA, 2003.
24 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.
25 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
26 A. D. Becke, *Chem. Phys.*, 1993, **98**, 5648.
27 G. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley and J. Mantzaris, *J. Chem. Phys.*, 1988, **89**, 2193.
28 C. Peng, P. Y. Ayala, H. B. Schlegel and M. J. Frisch, *J. Comput. Chem.*, 1996, **17**, 49.
29 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys.: Condens. Matter*, 2009, **21**, 395502.
30 J. P. Perdew and A. Zunger, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1981, **23**, 5048.
31 S. Mahapatra, S. K. Nayak, S. J. Prathapa and T. N. G. Row, *Cryst. Growth Des.*, 2008, **8**, 1223.
32 A. D. Corso, S. Baroni, R. Resta and S. de Gironcoli, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1993, **47**, 3588.
33 N. Troullier and J. L. Martins, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **43**, 1993.
34 H. J. Monkhorst and J. D. Pack, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1976, **13**, 5188.
35 P. Gianozzi and S. Baroni, *Handbook of Materials Modeling. Methods and Models*, Springer, 2005, p. 195.
36 J. P. Merrick, D. Moran and L. Radom, *J. Phys. Chem. A*, 2007, **111**, 11683.
37 S. I. Gorelsky, http://www.sg-chem.net/, University of Ottawa, Ottawa, Canada, 2010.
38 S. I. Gorelsky and A. B. P. Lever, *J. Organomet. Chem.*, 2001, **635**, 187.
39 A. J. Ramirez-Cuesta, *Comput. Phys. Commun.*, 2004, **157**, 226.
40 ISIS – A world centre for neutrons and muons, http://www.isis.stfc.ac.uk/, 31 October, 2012, UK.
41 A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441.
42 R. P. Lopes, M. P. M. Marques, R. Valero, J. Tomkinson and L. A. E. Batista de Carvalho, *Spectrosc. Int. J.*, 2012, **27**, 273.
43 T. A. Mohamed, I. A. Shabaan, W. M. Zoghaib, J. Husband, R. S. Farag and A. E. N. Alajhaz, *J. Mol. Struct.*, 2009, **938**, 263.
44 O. V. Shishkin, L. Gorb and J. Leszczynski, *Chem. Phys. Lett.*, 2000, **330**, 603.
45 R. Santamaria, E. Charro, A. Zacarías and M. Castro, *J. Comput. Chem.*, 1999, **20**, 511.
46 M. Mathlouthi and L. Seuvre Jack, *Carbohydr. Res.*, 1984, **131**, 1.
47 A. Hirakawa, H. Okada, S. Sasagawa and M. Tsuboi, *Spectrochim. Acta, Part A: Mol. Spectrosc.*, 1985, **41**, 209.
48 ISIS – TOSCA, http://www.isis.stfc.ac.uk/instruments/tosca/tosca4715.html, 31 October, 2012, UK.
49 T. G. Burova, V. V. Ermolenkov, G. N. Ten, R. S. Shcherbakov, V. I. Baranov and I. K. Lednev, *J. Phys. Chem. A*, 2011, **115**, 10600.