RESEARCH PAPER

PCCP
www.rsc.org/pccp

The mid-IR absorption spectrum of gas-phase clusters of the nucleobases guanine and cytosine

Joost M. Bakker,*†[a] Isabelle Compagnon,[a] Gerard Meijer,[b] Gert von Helden,[b] Martin Kabeláč,[c] Pavel Hobza[c] and Mattanjah S. de Vries[d]

[a] *FOM-Institute for Plasma Physics Rijnhuizen, Edisonbaan 14, NL-3439 MN, Nieuwegein, The Netherlands. E-mail: jmbakker@fhi-berlin.mpg.de*
[b] *Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195, Berlin, Germany*
[c] *Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Prague 6 and Center for Complex Molecular Systems and Biomolecules, 166 10, Prague 6, Czech Republic*
[d] *Department of Chemistry, University of California Santa Barbara, Santa Barbara, CA, USA*

Received 10th December 2003, Accepted 8th March 2004
First published as an Advance Article on the web 16th April 2004

The mid-infrared (IR) absorption spectrum of jet-cooled clusters of the nucleobases guanine and cytosine has been recorded in the 500–1800 cm^{-1} range by ion-dip spectroscopy. Some 30 clearly separated and sharp resonances are observed. The combination of the experimental data with new high-level *ab initio* calculations is consistent with a previous structural assignment and a tentative assignment is made to the K7-1E structure of guanine–cytosine. These data can serve as a test of the method in the mid-IR regime.

I. Introduction

One of the questions yet to be answered for the DNA system is what determines the ultimate stabilization of the classical Watson–Crick structure.[1] Several factors play a role and it is unclear what part in the total stabilization they play: (*i*) the multiple hydrogen bonds between the base pairs that cross-link the two strands, (*ii*) the interaction between the nucleobases on one strand, the so-called stacking mechanism, and (*iii*) the stabilization by the deoxyribosoic backbone and other environmental effects. This question has caused a considerable theoretical effort to evaluate the various factors.[2–4] To investigate this matter experimentally one can adopt the approach of trying to separate the factors that play a role, which can be achieved by studying isolated molecules in the gas phase. Spectroscopic studies of gas phase biological molecules have enabled the elucidation of fundamental interactions governing biological processes unobscured by such effects as the broadening of spectral lines by sample–solvent interactions.[5,6] The obvious start of gas phase studies involving the biological heredity system lies in interrogating the fundamental building blocks, the isolated purine and pyrimidine molecules. By extending the study to clusters of these building blocks, one can extract valuable information on the multiple hydrogen-bonded interactions, stabilizing the two strands.

Considerable spectroscopic knowledge on individual gas-phase nucleobases has been obtained in recent years.[7–12] One of the striking findings is the presence of many tautomers that are different from the form in which they are found in DNA in living matter. Using spectroscopic techniques, clusters of nucleobases have also been investigated.[13–15] In those studies the Watson–Crick geometries of the paired nucleobases have not been observed so far.

Fig. 1 shows the four lowest energy structures of the guanine–cytosine (G–C) cluster, out of more than 23 possible hydrogen bonded structures. Structure A represents the triply hydrogen bonded Watson–Crick structure. Structure C is very similar, but it has cytosine as an enol rather than as a keto tautomer. Structure D is again very similar: it is another Watson–Crick type bonded structure, now with an enol cytosine tautomer and a guanine tautomer where one hydrogen is attached at the N7 position (as opposed to N9 for structure C). Finally, in structure B the tautomers as found in DNA are hydrogen bonded in a way which would be impossible if the sugar backbone were present. To discern the structures, a nomenclature was introduced by Nir *et al.*[14] In this notation, structure C is called K9-1E, indicating that guanine is found in the keto form, with a hydrogen attached to N9; cytosine is in its enol form, and the OH group is oriented towards N1

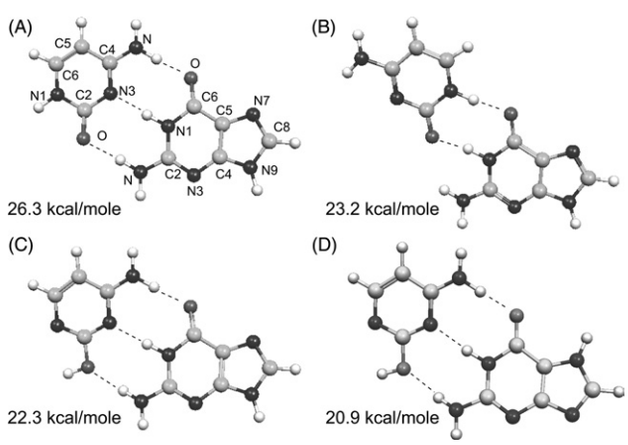

Fig. 1 Structures of the four lowest-energy G–C pairs. The numbers in the figures are the stabilization energies calculated at RI-MP2/TZVPP level. Also included is the atom labeling convention which is used throughout this paper.

† Present address: Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6, D-14195, Berlin, Germany.

DOI: 10.1039/b316158g

(*cf.* Fig. 1). Using REMPI spectroscopy and UV–UV holeburning so far two structures of G–C clusters have been distinguished. With the help of IR–UV holeburning in the near IR between 3300 cm^{-1} and 3700 cm^{-1} one of these could tentatively be identified as structure C[14] and the other one as structure B. The identification of structure C is not unambiguous: the X–H (X = N, C) stretching vibrations of structures C and D are very similar, and assignment was done based on the relative energies of the structures. Both tautomers of guanine have been observed in a jet experiment, and assignment of the spectrum to structure D can not be ruled out.[12] No spectrum is observed that can be assigned to structure A, however. This could be caused by a considerable difference in the UV spectral properties or by a dramatically reduced lifetime of the G–C cluster in structure A. Broad REMPI spectra have been observed for clusters of 1-methylcytosine with 9-methylguanine and of cytidine with guanine that can tentatively be assigned to structure A.[29]

The general strategy to elucidate the structure of a molecular system in the gas phase is to combine UV spectroscopy and UV–UV/IR–UV holeburning spectroscopy with high-level theoretical calculations. UV–UV holeburning spectroscopy ascertains that a specific UV transition results from one single structure, whereas IR–UV holeburning spectroscopy is used to identify this structure. Recently we presented results on two systems in the molecular fingerprint region of the IR, where we made use of a Free-Electron Laser (FEL) as a source of tunable IR radiation.[16,17] These experiments demonstrated that the mid-IR fingerprint region contains a wealth of resonances with considerable structure-sensitive information. Here, we present the IR absorption spectrum of the gas-phase cluster of guanine and cytosine, recorded in the 500–1800 cm^{-1} region. At the same time, we present a set of new high-level calculations to interpret the presented data. These data serve the dual purpose of testing the original assignment of this cluster and testing the comparison of the high-level computations with the experiment in this region of the IR spectrum.

Our calculations represent a new step in theoretical studies of DNA bases. All calculations on DNA bases and base pairs should be performed at a single, common level that accurately describes all interactions in the system. Hartree–Fock (HF) and Density Functional Theory (DFT) calculations, for example, only correctly describe hydrogen bonded interactions but are not able to describe the equally important stacking interactions. Surprisingly, up to now no DFT functional exists that describes the stacking of DNA bases. The first acceptable level for a correct description of both hydrogen bonding and stacking interactions is the MP2 level using at least a DZ + P basis set. This level provides accurate characteristics (structure, stabilization energies, and vibrational frequencies) of hydrogen bonded pairs but underestimates stabilization energies for stacking. Reliable values for all characteristics of any type of DNA base pair are generated, however, if the MP2 calculations are performed with the TZ + P basis set,[18] *i.e.* a triple-zeta basis set with two sets of first polarization functions and one set of second polarization functions on all the atoms. Such a treatment is feasible for obtaining the structure as well as stabilization energy of a pair but computationally too costly for generating vibrational frequencies. Therefore, in the present paper, we adopt the following computational strategy: structure and stabilization energy are calculated at the MP2/TZ + P level, allowing for a comparison between the stabilization energies of the studied hydrogen bonded complexes with those of stacked complexes. The vibrational frequencies are then determined at a cheaper MP2/DZ + P level.

II. Experimental

The experiment is performed in a previously described molecular beam setup[16,17,19] combined with a newly constructed laser-desorption source. A mixture of solid guanine and cytosine crystals and graphite powder is deposited on the surface of a solid graphite bar ($50 \times 15 \times 2$ mm) that is placed directly under the orifice of a pulsed valve. The 10 Hz pulsed valve releases gas pulses of argon (typically 50 μs long; backing pressure 3.5 bar) through a 0.5 mm diameter nozzle into a vacuum. Directly after opening the nozzle, a pulsed Nd:YAG laser (Thales Diva-2, 1064 nm, 5 ns, <1 mJ per pulse) desorbs sample molecules from the graphite matrix. The desorbed molecules are entrained in the supersonically expanding carrier gas. In the adiabatic expansion the internal degrees of freedom in the nucleobase molecules are cooled to about 10 K. The expansion conditions enable the formation of clusters of guanine and cytosine, which are internally cooled as well. About 4 cm downstream from the source the molecular beam is skimmed and enters a differentially pumped Wiley-McLaren type linear Time-Of-Flight (TOF) mass spectrometer.

The molecules in the beam interact with UV laser beams as well as with an IR laser beam at the crossing point of the mutually perpendicular molecular beam axis, laser beam axis and TOF axis. Ions are produced in this region and accelerated towards a microchannel plate (MCP) detector, generating mass spectra with a resolution of $M/\Delta M \approx 200$. The signal from the MCP detector is amplified and fed into a 10 bit, 100 MS s^{-1} digital oscilloscope that is read out by a PC. Two digital delay/pulse generators are used to synchronize the molecular beam to the various laser sources.

For the measurement of the IR absorption spectra of the jet-cooled guanine–cytosine (G–C) clusters we employ IR ion-dip spectroscopy. Ions are produced from ground state molecules using a two-photon ionization scheme. For this, the molecules are first excited to the lowest vibrational level in the first electronically excited singlet state (S_1) using a frequency-doubled Nd:YAG pumped pulsed dye laser (Rhodamine 640, 5 ns pulses, spectral width around 0.4 cm^{-1}). A second photon from the excitation laser ionizes the molecules.

A few μs before the excitation laser is fired, the IR laser interacts with the molecules in the beam. If a vibrational transition is induced by the IR light, molecular population is transferred from the ground state into an excited vibrational state, resulting in a depletion of ground state population. This results in a reduction in the number of produced G–C ions. By measuring the ion yield of G–C ions, while varying the wavelength of the IR laser, the ion-dip spectrum is obtained.

The IR radiation is produced at the Free Electron Laser for Infrared eXperiments (FELIX) user facility at the FOM Institute for Plasma Physics in Nieuwegein, the Netherlands.[20] The temporal output of this 10 Hz pulsed laser system consists of a few-μs long burst (macropulse) of micropulses. The micropulse spacing within the burst is set to 1 ns. The spectral bandwidth is adjusted to approximately 0.5% (fwhm) of the central frequency which corresponds to a micropulse duration of about 100 optical cycles. The frequency range that can be covered extends from 40 cm^{-1} to 2000 cm^{-1}, although only the range from 500–1800 cm^{-1} is used in the present study. Typically, energies of up to 100 mJ can be reached in the macropulse. In the present experiment, the UV detection laser is running at a 10 Hz repetition rate, while FELIX is set to 5 Hz. By independently recording both the IR-*on* signal and the IR-*off* signal a normalized ion-dip spectrum can be obtained, that is insensitive to long-term drifts in UV laser power or source conditions.

III. Theoretical methods

The use of DZ + P or even TZ + P basis sets in a combination with the exact MP2 treatment for DNA base pairs is difficult if not impractical. Recently we explored the applicability of the approximative resolution of identity (RI) MP2 methods[21,22]

for nucleic acid base pairs and larger DNA fragments[23] and we have shown that this method is capable of an accurate description of hydrogen bonded and stacked DNA base interactions. The results obtained with the RI-MP2 method differ only marginally from those evaluated with the exact MP2 method, while the efficiency is larger by up to one order of magnitude.

The potential energy surface (PES) of the G–C pair is investigated using the molecular dynamics/quenching (MD/Q) method with the Cornell *et al.* empirical force field.[24] We perform MD/Q simulations in the *NVE* microcanonical ensemble (*N*, *V* and *E* refer to the number of particles, volume and energy, respectively) within the quaternion formalism. The MD simulations are carried out at a constant total energy corresponding to an average temperature of 298 K which is high enough to allow crossing over relatively high energy barriers and thus to sample the whole potential energy surface. Every 1 ps the MD run is interrupted, the kinetic energy is removed and the structure of the complex of guanine and cytosine is optimized using the conjugate gradient method. The geometry and the corresponding energy are then stored, while the MD run is restarted from the point where it has been interrupted. The total simulation time is 250 ns.

We limit the investigation of the G–C PES to 16 different combinations of base pairs, obtained by combining the 4 most stable tautomers of guanine and cytosine.[25] The four most stable structures of the complex are further reoptimized by gradient optimizations at the RI-MP2 level using DZ + P (cc-pVDZ [3s2p1d/2s1p]) and TZ + P (TZVPP [5s3p2d1f/3s2p1d]) basis sets with a standard (default) auxiliary basis set. Harmonic vibration frequencies are determined numerically using the smaller basis set while stabilization energies are evaluated with the larger one. The stabilization energy is corrected for the basis set superposition error using the function counterpoise procedure[26] and also the deformation energy is included. We are aware of the fact that the *a posteriori* inclusion of the BSSE can be a source of error, but this error is certainly much smaller than when the counterpoise-corrected procedure is not used. Performing a counterpoise-corrected gradient optimization would solve the problem, but that type of optimization for the complexes presented is much more time consuming and cannot be used for routine calculations. All the calculations are carried out using the TURBOMOLE, version 5.6, program suite.[27]

IV. Results and discussion

The IR ion-dip spectrum for G–C in the 500–1800 cm^{-1} region exhibits a wealth of sharp resonances. It is recorded with the UV probe laser at 33 314 cm^{-1}. This frequency corresponds to the $S_1 \leftarrow S_0$ (0–0) transition of the cluster structure that was earlier assigned to structure C, based on IR–UV holeburning spectroscopy in the near IR region between 3300 and 3700 cm^{-1}.[14] That earlier spectrum contained the frequency of the OH stretching vibration of the enol-cytosine as a distinctive feature. As the OH stretching vibration of structure D was calculated to be almost identical, no definitive assignment was made. The present data in the mid-IR may assist in accomplishing a final assignment as the current spectral range contains the CO stretching vibrational frequency as well as a large number of skeletal vibrational frequencies that can serve for further structural identification.

As demonstrated in a previous publication, it is possible to convert the IR ion-dip spectrum into an IR absorption spectrum if it is checked that the depletions due to each resonance are the result of a one-photon absorption process.[16] For a limited number of resonances this has been verified by performing measurements of the depletion as a function of IR laser intensity. We can therefore transform the normalized ion-dip signal $s(\nu)$ into relative IR absorption cross-sections $\sigma(\nu)$ using the conversion $\sigma(\nu) = -\log(s(\nu))$ and by correcting for the FELIX output power. Fig. 2 shows the IR absorption spectrum thus obtained for the G–C cluster in the 500–1800 cm^{-1} region. The spectrum shows a large number of clear resonances, the line widths of which are mainly determined by the FELIX spectral profile. Below the experimental spectrum four traces are shown that represent the theoretical spectra of the four lowest-energy configurations. The IR intensities obtained from the calculations are convoluted with a Gaussian function representing the FELIX spectral profile. The bandwidth used in the convolution is 0.5% of the FELIX frequency. Furthermore, the frequencies of the theoretical modes are scaled by a factor of 0.956 to obtain the best agreement with the observed spectrum and the vertical axis scaling for the three theoretical curves is the same.

From a comparison between the experimental spectrum and the theoretical curves one can observe that the spectra of structures C and D resemble the data the best. Particularly in the region of 1000 cm^{-1} to 1800 cm^{-1} the experimental spectrum is rather well predicted, as in this region for both structures all observed lines can directly be attributed to calculated modes and only minor differences in the spectra calculated for structures C and D are observed. In the lower energy region, below 1000 cm^{-1}, the agreement is less convincing but equal to or better than that for the calculated spectra of either structure A or B. The only serious disagreement between

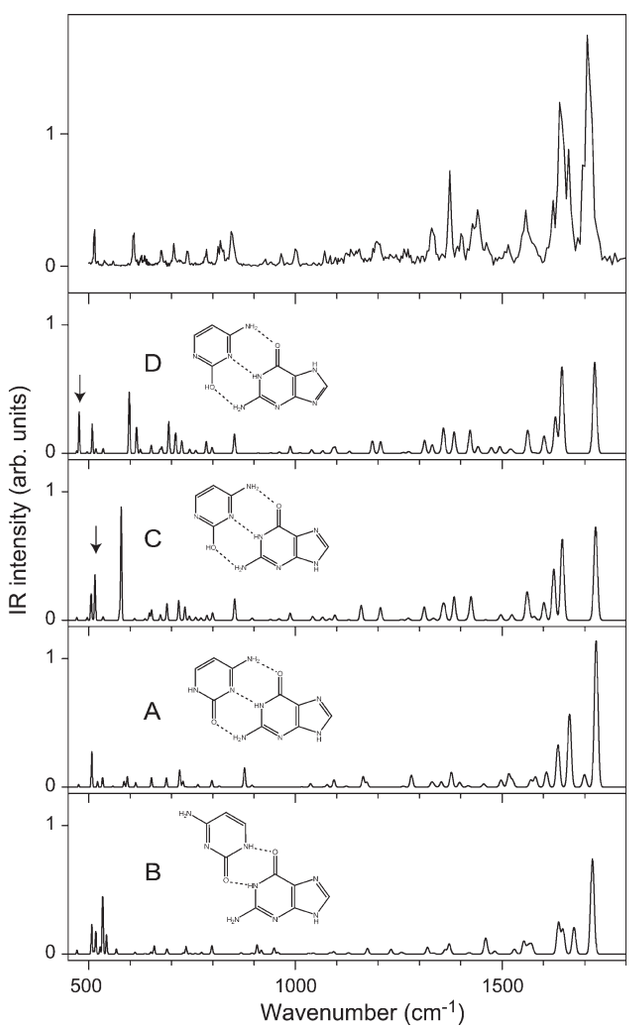

Fig. 2 IR absorption spectrum of the guanine–cytosine cluster. Below the figure are theoretical spectra of the four lowest-energy configurations of guanine–cytosine clusters. Theoretical frequencies for all structures are scaled by a factor of 0.956 and convoluted with a Gaussian function representing the FELIX spectral profile.

theory and experiment lies in the abnormally large intensity of the resonance calculated at 578 cm^{-1} (scaled value) for structure C. One could be tempted to assign the spectrum to structure D, based upon this discrepancy. It is necessary, however, to have information about the structures and the nature of the vibrational modes before such a decision can be made.

Structure C is a slightly non-planar (deviation of planes of the bases is about 20°) Watson–Crick-type structure of canonical guanine and a tautomer of cytosine in which a hydrogen atom is detached from the N1–H group and attached to the C2=O group, forming an enol side group. The presence of the OH group as well as the change of the cytosine electronic structure leads to a weaker interaction between G and C than in structure A. The distances between the heavy atoms in the hydrogen bonds (cyt N–H–gua O), (cyt N3–gua H–N1) and (cyt O–gua H–N) in structure C are 2.84, 2.97 and 3.07 Å, respectively (*cf.* Fig. 1). Here, all geometric parameters refer to the calculations done at the RI-MP2/cc-pVDZ level. Both amino groups in structure C are non-planar with respect to the base plane (deviation in both subsystems is approximately 20°).

Structure D is very similar to structure C but a hydrogen atom is detached from the N9–H group and attached to the N7. The base planes are tilted with respect to each other by about 23°, the hydrogen bond distances are 2.85, 2.95 and 3.10 Å and the amino hydrogens are oriented out-of-plane by about 25°.

Structures C and D correspond to the second and third local minima at the G–C potential energy surface (PES). Their (absolute) stabilization energies (including deformation energy and difference in tautomerization energies, a sum of relative energies of isolated tautomers with respect to the canonical form) are 22.3 kcal mol^{-1} and 20.9 kcal mol^{-1}, respectively. This is about 4.0 kcal mol^{-1} and 5.4 kcal mol^{-1} less stable than the global minimum structure, the canonical Watson–Crick structure A. Structure A is almost planar with only a small deviation (5°) of amino hydrogens from the plane of the base. The distances between the heavy atoms participating in the hydrogen bonds are 2.78, 2.92 and 2.92 Å, respectively and they are slightly shorter than in the cases of structures C and D.

Structure B, possessing a different hydrogen bond pattern between canonical forms of guanine and cytosine, is about 1 kcal mol^{-1} more stable than structure C (23.2 kcal mol^{-1}) and this structure thus corresponds to the first local minimum at the PES. Structure B is slightly nonplanar and it contains only 2 hydrogen bonds with the following distances between heavy atoms : O(gua)···N1H(cyt) = 2.77 Å, and N1H(gua)···O(cyt) = 2.71 Å.

The present RI-MP2/TZVPP stabilization energies are systematically larger than the respective RI-MP2/cc-pVDZ ones by about 3–4 kcal mol^{-1}. More detailed information on all 50 complexes obtained at the G–C PES by considering the four most stable tautomers of guanine and of cytosine will be published later.[25]

To discuss the nature of the resonances in the experimental spectrum the calculated vibrational modes are inspected using visualization software to identify the associated motions of the atoms. Line positions of all observed resonances and (scaled) frequencies of the assigned theoretical vibrations, including short descriptions of the modes, can be found in Table 1. Most of the assigned modes are identical for both structure C and D; only when they differ is an explicit mode description given for structure D. Throughout the rest of the discussion scaled values will be given for theoretical frequencies to facilitate comparison to Fig. 1. The dominating feature in the experimental spectrum, the highest frequency resonance found at 1706 cm^{-1}, is associated with the guanine C=O stretching mode, with a slight coupling to a bending motion of the guanine N1–H group. As this is the only C=O group in structures C and D, its assignment is unambiguous. For structure A the two C=O stretching modes are coupled, and result in one symmetric and one asymmetric stretching mode of which the latter (calculated at 1727 cm^{-1}) has a high IR intensity, and the former (1699 cm^{-1}) only little. The two C=O stretching vibrations in structure B are not coupled to each other, but to an N–H in-plane bending mode. This interaction is less favorable for the cytosine stretching mode (1674 cm^{-1}), which has considerably lower IR intensity than its guanine counterpart (1718 cm^{-1}). The modes lying at somewhat lower energy, between 1000 cm^{-1} and 1700 cm^{-1}, are all associated with skeletal motions in the plane of the molecules, coupled in varying strength with in-plane NH_2 or NH bending vibrations. Most of the modes here are rather well reproduced, except for a somewhat strong IR intensity for the resonance experimentally found at 1374 cm^{-1}, which might be due to overlapping resonances. The region below 1000 cm^{-1} contains a number of clear resonances that cannot so easily be assigned to a specific theoretical mode. Most vibrations here are out-of-plane bending motions of the various hydrogen atoms.

There is one mode with appreciable IR activity that could make an assignment to either structure C or D possible. The N9–H and N7–H out-of-plane bending vibrations are calculated to be very localized in structures C and D, respectively, and should be sensitive to their environment: in structure C the N9–H group is rather isolated, whereas in structure D the N7–H group could be influenced by the presence of the C6=O group. For both structures the respective N–H out-of-plane bending vibration is indicated with an arrow in the calculated spectrum. In structure C, around 500 cm^{-1} two vibrational modes are expected: the C–O–H in-plane bending mode at 506 cm^{-1} and the N9–H out-of-plane bending mode at 515 cm^{-1}. In structure D, only the C–O–H in-plane bending mode is expected here, at 508 cm^{-1} while the N7–H out-of-plane bending mode is calculated to lie at 477 cm^{-1}. From the fact that there is only one observed resonance, at 514 cm^{-1}, we are tempted to assign the experimental spectrum to structure D. Unfortunately, the data do not cover the range where the N7–H bending mode is expected in structure D, and a definite assignment can still not be made.

It must be pointed out that calculated vibration at 578 cm^{-1} in the spectrum of structure C, which is predicted with a larger IR intensity than what is observed cannot be invoked to finalize the assignment: the motions associated with this mode are those of the two hydrogens of the guanine amino group moving in-phase through the plane of the molecule. It has been observed earlier that such inversion-like modes that have a double minimum in the PES are poorly reproduced in the harmonic approximation.[17,28] It therefore is better not to use this resonance to assign the structure based on the comparison between calculated and experimental spectra. The cytosine NH_2 group does not exhibit a simultaneous bending vibration or tunneling mode in this structure nor in the three others; in structure C one localized N–H out-of-plane bending mode is predicted at 401 cm^{-1}.

Structure D shows the same pattern: one tunneling-like mode is found at 598 cm^{-1}, (although with substantially less IR intensity than its equivalent in structure C) and a localized N–H out-of-plane bending mode at 392 cm^{-1}. The guanine NH_2 group in structure A does not exhibit the tunneling mode. Rather, its out-of-plane bending vibration is symmetric and yields almost no IR intensity. This different behavior is most likely caused by the fact that in this structure the cytosine is found as a keto tautomer, and the hydrogen bonding is substantially different.

Although the guanine tautomer in structure D is energetically not favored, it has been observed in a jet-experiment.[12] Moreover, the GC origin is found to be much closer to the 7H keto form of guanine than the 9H keto tautomer.[12] Differently phrased: GC complexes with G in the K9 form might

Table 1 Table of observed frequencies for the guanine–cytosine cluster, the assigned theoretical mode (from structures C and D) and a description of the associated vibration. When an observed resonance is assigned to multiple theoretical modes the line positions and descriptions of these modes are separated by a semicolon. The following abbreviations are used: ip: in-plane; oop: out-of-plane; cyt: cytosine; gua: guanine

	Calculated		Mode description	
Observed	Structure C	Structure D	Structure C	Structure D (if different)
514	515 or 506	508	gua N9H oop bending or cyt CO oop bending	cyt CO oop bending
538	534	535	cyt ip skeletal	
559	—	—		
628	578	598	gua NH_2 oop bending	
636	611	616	gua NH_2 oop bending, gua ip skeletal	
675	652	651	gua oop skeletal	gua ip skeletal
705	689	694	gua N1–C2–N3 oop bending	
740	717	710	cyt NH_2 oop bending	
785	—	—		
818	—	—		
845	853	852	gua N1H oop bending	
928	895	957	gua N7–C8–N9 ip bending	
965	940	941	cyt C5H, C6H oop bending	
1000	988	987	cyt C5H, C6H, NH_2 ip bending	
1071	1043	1039	gua C8H, N9H ip bending	gua C8H, N7H ip bending
1129	1095	1096	cyt C5H, C6H, NH_2 ip bending	
1152	1159	1186	gua ip skeletal	gua NH7, NH8 ip bend
1199	1205	1206	cyt OH ip bending	
1269	1273	1274	cyt C6H ip bending	
1330	1311	1312	cyt C5H, C6H, NH_2 ip bending	
1374	1355; 1360	1359	gua N1H, N2H, N9H ip bending; gua ip skeletal	gua ip skeletal
1401	1384	1384	cyt ip skeletal	
1429				
1441	1425	1423	cyt ip skeletal, gua N8H ip bending	cyt ip skeletal, gua N1H ip bending
1462				
1515	1523	1525	gua ip skeletal	
1558	1559	1562	cyt ip skeletal	
1639	1645;	1645;	cyt NH_2 scissor;	
	1625;	1629;	gua NH_2 scissor, N1H ip bending;	
	1601	1602	cyt NH_2 scissor, cyt skeletal	
1706	1726	1724	gua C=O stretching, gua N1H ip bending	

have resonances that are found at larger frequencies in the UV spectrum, and hence may be too difficult to observe, either because they could be hidden by spectral congestion due to resonances of the K7–1E complex. Alternatively, their higher-lying excited state could favor fast dynamic processes that affect the efficiency of the R2PI process.

V. Conclusions

We have measured the IR absorption spectrum for the cluster of the nucleobases guanine and cytosine isolated in the gas phase in the mid-IR fingerprint region, between 500 cm^{-1} and 1800 cm^{-1}. The IR spectrum contains a large number of sharp resonances that can be used to assign the spectrum. A comparison with theoretical spectra of a set of possible geometries is consistent with a previous assignment to a Watson–Crick type hydrogen bonded structure that contains an enol cytosine tautomer. A tentative assignment is made to a structure with a guanine tautomer in which a hydrogen atom is detached from the N9–H group and attached to N7, but a final assignment can only be made when further experimental work in the frequency regime below 500 cm^{-1} is done. The favorable comparison between theoretical and experimental spectra suggests that the hydrogen bonding properties of gas phase clusters can be treated in a satisfactory fashion with the state-of-the-art theoretical methods used.

Acknowledgements

We gratefully acknowledge the support by the *Stichting voor Fundamenteel Onderzoek der Materie* (FOM) in providing the required beam time on FELIX and highly appreciate the skilful assistance by the FELIX staff. This material is based upon work supported by the National Science Foundation under Grant No. CHE-0244341. This work is also supported by the Ministry of Education of the Czech Republic, project LN00A032 (Center for Complex Molecular Systems and Biomolecules). Finally, we thank the referee for bringing structure D to our attention.

References

1 J. D. Watson and F. H. C. Crick, *Nature (London)*, 1953, **171**, 737.
2 C. F. Guerra, F. M. Bickelhaupt, J. G. Snijders and E. J. Baerends, *J. Am. Chem. Soc*, 2000, **122**, 4117.
3 P. Hobza and J. Sponer, *J. Am. Chem. Soc*, 2002, **124**, 11 802.
4 D. Sivanesan, I. Sumathi and W. J. Welsh, *Chem. Phys. Lett.*, 2003, **367**, 351.
5 T. S. Zwier, *J. Phys. Chem. A*, 2001, **105**, 8827.
6 E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2001, **3**, 1.
7 E. Nir, L. I. Grace, B. Brauer and M. S. de Vries, *J. Am. Chem. Soc*, 1999, **121**, 4896.
8 E. Nir, M. Muller, L. I. Grace and M. S. de Vries, *Chem. Phys. Lett.*, 2002, **355**, 59.
9 N. J. Kim, G. Jeong, Y. S. Kim, J. Sung, S. K. Kim and Y. D. Park, *J. Chem. Phys.*, 2000, **113**, 10 051.
10 C. Plützer, E. Nir, M. S. de Vries and K. Kleinermanns, *Phys. Chem. Chem. Phys.*, 2001, **3**, 5466.
11 D. C. Luhrs, J. Viallon and I. Fischer, *Phys. Chem. Chem. Phys.*, 2001, **3**, 1827.

12 W. Chin, M. Mons, I. Dimicoli, F. Piuzzi, B. Tardivel and M. Elhanine, *Eur. Phys. J. D*, 2002, **20**, 347.
13 E. Nir, K. Kleinermanns and M. S. de Vries, *Nature (London)*, 2000, **408**, 949.
14 E. Nir, C. Janzen, P. Imhof, K. Kleinermanns and M. S. de Vries, *Phys. Chem. Chem. Phys.*, 2002, **4**, 732.
15 C. Plützer and K. Kleinermanns, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4877.
16 J. M. Bakker, L. Mac Aleese, G. von Helden and G. Meijer, *J. Chem. Phys.*, 2003, **119**, 11 180.
17 J. M. Bakker, L. Mac Aleese, G. Meijer and G. von Helden, *Phys. Rev. Lett.*, 2003, **21**, 2032 003.
18 P. Jurecka and P. Hobza, *J. Am. Chem. Soc.*, 2003, **125**.
19 J. M. Bakker, R. G. Satink, G. von Helden and G. Meijer, *Phys. Chem. Chem. Phys.*, 2002, **4**, 24.
20 D. Oepts, A. F. G. van der Meer and P. W. van Amersfoort, *Infrared Phys. Technol.*, 1995, **36**, 297.
21 M. Feyereisen, G. Fitzgerald and A. Komornicki, *Chem. Phys. Lett.*, 1993, **208**, 359.
22 F. Weigend and M. Häser, *Theor. Chem. Acc.*, 1997, **97**, 331.
23 P. Jurecka, P. Nachtigall and P. Hobza, *Phys. Chem. Chem. Phys.*, 2001, **3**, 4578.
24 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc*, 1995, **117**, 5179.
25 M. Kabelác, K. Kleinermanns, M. S. de Vries and P. Hobza, in preparation.
26 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
27 R. Ahlrichs, M. Bar, M. Haser, H. Horn and C. Kolmel, *Chem. Phys. Lett.*, 1989, **162**, 165.
28 H. Piest, G. von Helden and G. Meijer, *J. Chem. Phys.*, 1999, **110**, 2010.
29 A. Abo-Riziq, L. Grace and M. S. de Vries, to be published.