PCCP

Pairing of the nucleobases guanine and cytosine in the gas phase studied by IR–UV double-resonance spectroscopy and *ab initio* calculations

E. Nir,[a] Ch. Janzen,[b] P. Imhof,[b] K. Kleinermanns[b] and M. S. de Vries[c]

[a] *Department of Chemistry, The Hebrew University, Jerusalem 91904, Israel*
[b] *Institut für Physikalische Chemie und Elektrochemie I, Heinrich Heine Universität Duesseldorf, 40225, Düsseldorf, Germany*
[c] *Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106, USA*

Received 16th August 2001, Accepted 8th November 2001
First published as an Advance Article on the web 29th January 2002

We present R2PI, IR–UV and UV–UV double resonance measurements of the guanine–cytosine (G–C) dimer formed in a supersonic jet. We show that there is only one isomer of G–C in the investigated wavelength range from 33200 to 34100 cm^{-1}. We assigned the observed G–C isomer to a specific structure, based on comparisons of the IR spectra of the G and C monomers with the G–C dimer in the range of the OH and NH stretching vibrations and *ab initio*-calculated vibrational frequencies and dimer stabilities. The cluster exhibits an HNH···O/NH···N/C=O···HNH bonding similar to the Watson–Crick G–C base pair bonding but with C as the enol tautomer. We did not observe any keto–keto or enol–enol G–C dimers in the investigated wavelength region.

Introduction

It is well known that the two strands of DNA are held together by guanine–cytosine and adenine–thymine base pairs.[1] However, it is difficult to study the details of that interaction directly because they are masked by effects induced by the solvent and the DNA backbone.[2,3] In order to distinguish the inherent properties of the base pair interaction from those external effects, we study isolated gas-phase complexes composed of paired bases. Obtaining such data is also important for testing high-level computations, that are recently becoming available.[4–9] We synthesized the isolated dimers by using laser desorption of monomer mixtures followed by cooling in a supersonic jet. Subsequently, we applied different techniques of laser spectroscopy for detailed analysis of the clusters. From our measurements of the guanine monomer we know that three tautomers (9H-enol *trans*, 9H-keto, 7H-keto) are abundant in our jet experiment,[10] resulting in a large number of possible dimer configurations. Recently we have shown that only one isomer of G–C exists in the investigated wavelength range from 33200 to 34100 cm^{-1}.[11] Here we present a comparison of the infrared spectra obtained from the IR–UV experiments with *ab initio*-calculated frequencies for the different cluster structures which enables us to assign the observed isomer to a definite cluster structure.

Experimental and theoretical methods

The measurements were performed with an apparatus described in detail elsewhere.[12] Briefly, material is laser desorbed from a solid graphite sample in front of a pulsed nozzle. To obtain guanine–cytosine clusters the sample consisted of an approximately 1 : 1 mixture of guanine and cytosine on the graphite surface. Typical fluences of the Nd : YAG desorption laser operated at 1064 nm (where graphite absorbs but guanine does not) are about 1 mJ cm^{-2} or less, which is considerably lower than the fluences normally used for ablation. The laser is focused to a spot of about 0.5 mm diameter within 2 mm in front of the nozzle. We used a pulsed valve (General Valve; Iota One) with a nozzle diameter of 1 mm at a backing pressure of about 5 atm argon drive gas.

The skimmed molecular beam crossed the ionization laser at right-angles inside the source region of a reflectron time-of-flight (TOF) mass spectrometer. By monitoring the parent mass peak of the respective cluster while varying the two photon, one color ionization wavelength (resonant two photon ionization: R2PI), we obtained mass selected excitation spectra. We performed spectral hole burning (SHB) by using two counter-propagating dye laser pulses with a delay of about 150 ns. This resulted in two peaks in the TOF spectrum, the first from the "burn" laser and the second from the "probe" laser. When both lasers are tuned to the resonance of the same tautomer, the burn laser causes a decrease in the signal of the probe laser. We scan the burn laser while the probe laser frequency is fixed to an intense band of one of the cluster isomers. If a significant band of the R2PI spectrum is missing in the burn spectrum it belongs to another isomer. In the next step we probed at this frequency while scanning the pump laser to reveal the spectrum of the next isomer. Hot bands cannot be excluded by SHB but we did not observe any small bands at longer wavelengths than the intense electronic origin band of G–C which would be typical for hot bands.

We performed IR–UV SHB with the same method but taking a difference frequency IR laser as burn laser.[13,14] The radiation from an infrared dye (a mixture of Styryl 8 and Styryl 9) was aligned collinearly with the perpendicularly polarized Nd:YAG fundamental (1064 nm) and directed through an MgO-doped $LiNbO_3$ crystal to generate 3300–4000 cm^{-1} tunable IR light. Suitable dielectric mirrors separated the Nd:YAG fundamental and the dye laser beam behind the crystal. We typically used 50 mJ of the YAG fundamental and

 DOI: 10.1039/b107429f

10 mJ of the dye laser to obtain around 1 mJ per pulse IR radiation between 3300 and 4000 cm^{-1} with a bandwidth of <0.1 cm^{-1}. The IR laser was calibrated by recording a water vapor spectrum. Color centers in the $LiNbO_3$ crystal led to a decrease of the IR intensity from 3515 to 3550 cm^{-1}. In that spectral range we used another $LiNbO_3$ crystal with a gap in another region.

The IR absorption spectroscopy of the cytosine monomer was performed on a Hewlett-Packard system consisting of a GC 5890 Series II gas chromatograph, a Fourier transform infrared (FTIR) spectrometer (IRD 5965 B), equipped with a wide-band detector with a frequency range between 550 and 4000 cm^{-1}, and a mass detector (MSD 5971).[15] Cytosine has to be heated to around 300 °C to obtain a sufficient vapor pressure for IR spectroscopy but decomposes to a great extent at this temperature. Hence an infrared spectrum obtained in a simple heated cell consists of IR bands of cytosine and decomposition products.[16] GC-FTIR-MS has the advantage that the intact cytosine and decomposition or reaction products of cytosine are separated gas chromatographically and that for each GC peak an IR spectrum and a directly correlated (15–20 s delay time) mass spectrum can be taken. With this method we were able to identify unambiguously the IR spectrum of the cytosine monomer.[5]

The calculations were carried out using the Gaussian 98 program package.[17] We performed Hartree–Fock (HF) calculations utilizing a 6-31G(d,p) basis set. All structures were fully optimized at this level with 10^{-8} E_h as the SCF convergence criterion and 1.5×10^{-5} E_h a_0^{-1} and E_h degree^{-1} as convergence criteria for the gradient optimization of the structures. The vibrational frequencies were obtained by performing a normal mode analysis on the optimized geometries using analytical gradients of the energy. The binding energies of the cluster were corrected for zero point energy (ZPE) at the HF level. We also performed basis set superposition error (BSSE) corrections for some of the clusters but noticed that the order of cluster stability was generally not changed by including this correction. Minimum energy structure calculations at the HF level cannot pretend to give reliable absolute binding energies and merely indicate relative stabilities. Therefore, the dissociation energies indicated in Table 1 do not include BSSE corrections.

Table 1 Dissociation energies D_0 of the most stable guanine–cytosine dimer structures obtained by correcting the electronic dissociation energy D_e at the HF/6–31G(d,p) level for the zero point vibrational energy (ZPE). All values in cm^{-1}

Structure	D_e	ZPE	D_0
K9K-1	8991	729	8262
K9K-2	7959	417	7542
K9K-3	4732	471	4261
K9K-4	4533	401	4132
K9K-5	4540	472	4068
K9E-1	6571	485	6086
K7E-1	6067	470	5597
K7E-2	5406	502	4904
K9E-3	5015	192	4823
K7E-3	4687	197	4490
K7E-4	4757	397	4360
E9K-1	5623	551	5072
E9K-2	5539	633	4906
E9K-3	5360	534	4826
E9K-4	5257	487	4770
E9K-5	5229	611	4618
E9K-6	4589	549	4040
E9E-1	4397	284	4113
E7E-2	4631	577	4054
E7E-3	4502	497	4005
E9E-4	3978	356	3622
E9E-5	3617	338	3279

Results and discussion

Fig. 1 shows the structures of the most stable guanine–cytosine (G–C) isomers obtained from calculations at the HF/6–31G(d,p) level. We obtained the initial structures for geometry optimization by circling the guanine molecule around the cytosine molecule and searching for possible hydrogen bond interactions. We repeated this procedure with the stationary guanine molecule inverted. Attempts to obtain G–C "sandwich" structures of comparable stability as the most stable hydrogen bonded structures failed, which is not surprising as this level of calculation does not take the dispersion energy into account. From our IR–UV experiments we know that we have three guanine tautomers in the jet, namely 9H-enol and 9H-/7H-keto guanine. From microwave spectroscopy and the IR spectrum shown in Fig. 3 we know that cytosine exists in the gas phase as keto and enol tautomers in comparable quantities.[18] Based on these experimental findings, we calculated more than 60 different G–C dimer structures with these monomers as moieties.

We adopt a nomenclature in which K and E denote a keto or enol tautomer of guanine and cytosine, respectively; 9 and 7 denote the positions of H substitution in guanine and the suffixes -1, -2, *etc.* indicate cluster families with the same H bond arrangement, ordered according to their stability. The 9H-keto guanine–keto cytosine dimer K9K-1 is the Watson–Crick G–C base pair with HNH···O=C/NH···N/C=O···HNH bonding. This structure turns out to be the most stable one. K9K-2 is the second most stable cluster with two symmetrically arranged NH···O=C hydrogen bonds. K9E-1 (K7E-1) designate the 9H(7H-)-keto guanine–enol cytosine dimers with the largest stability. This cluster exhibits an HNH···O/NH···N/C=O···HNH bonding similar to the Watson–Crick G–C base pair bonding, but with C in the enol form. K7E-2 exhibits C=O···HNH and N7H···N hydrogen bonds. K9E-3 (K7E-3) form HNH···OH and N1H···N hydrogen bonds. K7E-4 (K9E-4) are stabilized by HNH···N/N···HNH interactions. The enol–keto and enol–enol isomers, shown in Fig. 1(b), are generally less stable. The calculated binding energies are listed in Fig. 1. All calculated structures and binding energies are available as supplementary material.

Fig. 2(a) and (b) display the intermolecular vibrations and the OH and NH stretching vibrations of the K9E-1 dimer in the electronic ground state. The intermolecular modes consist of the out-of-plane torsion τ, the out-of-plane bending (butterfly) motion γ_s with the guanine and cytosine moieties as wings, the antisymmetric out-of-plane bending (alternating stairs) motion γ_{as}, the in-plane bending (gearing) δ of the two moieties, and the antisymmetric and symmetric stretching vibrations σ_{as} and σ_s. The intermolecular modes of the other G–C isomers basically have the same characteristics, but their frequencies and sometimes even their sequence differ considerably (*cf.*, Table 2). OH(G,C), NH_2^a(G,C), NH_2^s(G,C) label the OH and the antisymmetric and symmetric NH_2 stretching vibrations of the guanine and cytosine moiety, respectively. The calculated vibrational frequencies of the most stable guanine-cytosine isomers are listed in Tables 2 and 3. In what follows we will be able to assign these vibrations in the experimental spectra and use their frequency pattern to distinguish between different cluster structures.

Fig. 3(a) and (b) show the IR spectrum of the cytosine monomer, recorded at the cytosine parent mass of $m/z = 111$ and with the GC-FTIR-MS system operated at 300 °C. The spectrum shows the spectral region of the OH, NH and CH stretching vibrations and the region of the C=O stretching and skeleton vibrations below 2000 cm^{-1}. The band with the highest frequency is located at 3610 cm^{-1}. All NH vibrations of the other nucleobases, guanine,[10] uracil,[19] thymine[19] and adenine,[19] occur below 3600 cm^{-1}. The only vibration

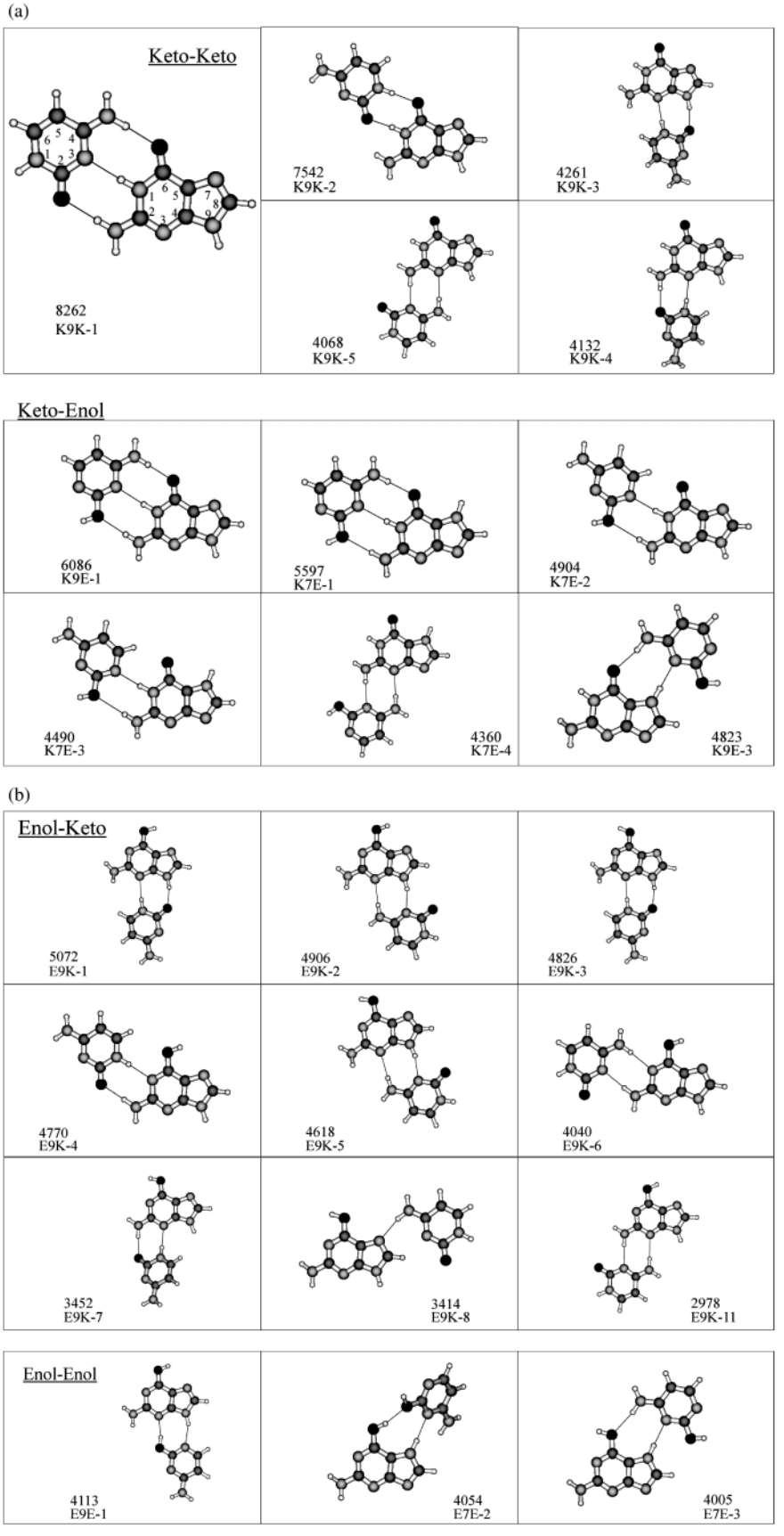
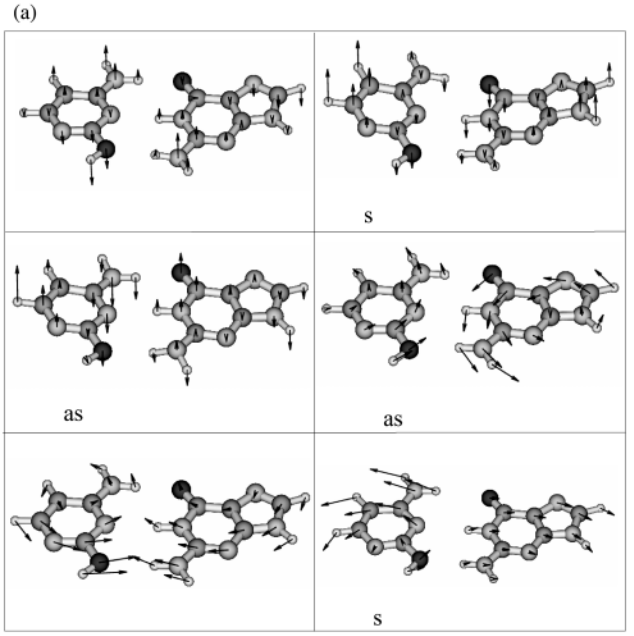


Fig. 1 Structures and stabilities of the most stable guanine–cytosine isomers at the HF/6–31G(d,p) level including ZPE correction. K and E denote the keto or enol tautomer of guanine and cytosine; 9 and 7 denote the positions of H substitution in guanine and -1, -2, *etc*., label clusters with the same H bond arrangement and orders these cluster families according to their stability. For example, K9K-1 labels the 9H-keto guanine–keto cytosine dimer with greatest stability. This is the Watson–Crick G–C base pair with HNH···O=C/NH···N/C=O···HNH bonding, which is the most stable cluster in this family of H bond arrangements (and also the overall the most stable structure). K9E-1 designates the 9H-keto guanine–enol cytosine dimer with greatest stability. The numbers in the figures are the cluster dissociation energies (cm^{-1}) with ZPE included. (a) Keto–keto and keto–enol dimers; (b) enol–keto and enol–enol dimers.



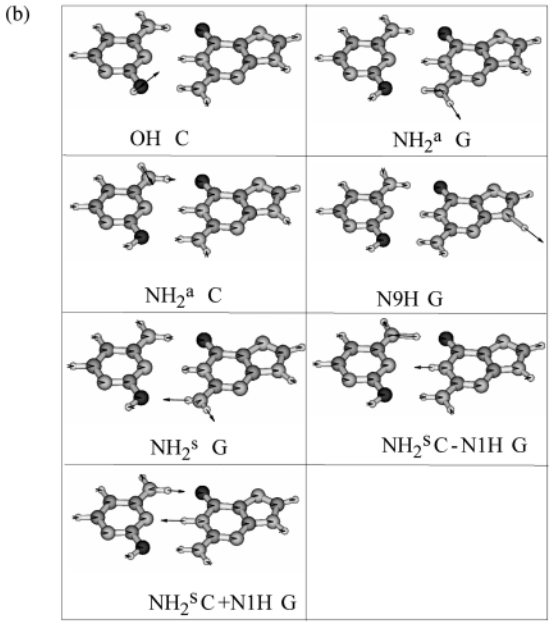


Fig. 2 Selected normal-mode vibrations of the K9E-1 dimer in the electronic ground state (a) Intermolecular vibrations. These are the out-of-plane bending (butterfly) motion γ_s with the guanine and cytosine moieties as wings, the out-of-plane torsion τ, the antisymmetric out-of-plane bending (alternating stairs) motion γ_{as}, the in-plane bending (gearing) δ of the two moieties and the antisymmetric and symmetric stretching vibrations σ_{as} and σ_s. (b) OH and NH stretching vibrations. OH (G,C), NH_2^a (G,C) and NH_2^s(G,C) label the OH and the antisymmetric and symmetric NH_2 stretching vibrations of the guanine and cytosine moiety, respectively.

expected at such a high frequency is the OH stretching mode. For comparison, the OH stretching vibration of the enol tautomer of guanine[1] absorbs at 3587 cm^{-1}. In a matrix isolation study of cytosine the highest frequency band at 3591 cm^{-1} was assigned to the OH stretching frequency of the enol tautomer of cytosine..[19] A prominent band at 1720 cm^{-1} was assigned to the C=O stretching vibration of the keto tautomer of cytosine.[20] The enol tautomer has the next vibration below the CH stretching vibrations at 1623 cm^{-1} (C5=C6 stretching vibration). We observe an intense band at 1740 cm^{-1} [*cf.*, Fig. 3(b)] and assign it to the C=O stretching vibration of the keto tautomer of cytosine in the gas phase.

Table 2 Vibrational assignment of the intermolecular vibrations of the guanine–cytosine dimer in the excited (S_1) state. All values are given in cm^{-1}. For comparison, the theoretical harmonic intermolecular frequencies are calculated for the S_0 state, for the most stable dimer structures. The calculated frequencies are not scaled

		HF/6–31G(d,p)					
Assignment	Experiment (GC)	K9E-1	K9K-1	K9K-2	K7E-2	E9K-1	E9K-2
τ		11	33	38	26	41	31
γ_s		23	25	14	24	19	23
$\gamma_s+\gamma_s$	56						
—	68						
γ_{as}		58	69	63[a]	64	68	66[a]
σ_{as}	82	79	84	86	76	81	68[a]
—	107						
δ	115	99[a]	115	62[a]	70	50	53
σ_s	120	116[a]	123	127	108	118	105
$\sigma_{as}+68$	151						
$\sigma_{as}+\sigma_{as}$	164						
—	168						
$\sigma_{as}+107$	189						
$\sigma_{as}+\delta$	196						
—	201						
—	224						
$\sigma_{as}+68$	232						

[a] Coupled motions.

The matrix and gas-phase results agree reasonably well considering possible shifts due to the matrix effect and the different temperatures of the two experiments. The spectrum displayed in Fig. 3 shows that both the enol tautomer of cytosine (OH stretching vibration at 3610 cm^{-1}) and its keto tautomer (C=O stretching vibration at ~1740 cm^{-1}) exist in comparable amounts at 300 °C. Microwave spectroscopy at 295 °C in the gas phase also showed that the keto and enol forms of cytosine have similar abundances in the gas phase, whereas the abundance of a third tautomer, the imino form, is considerably lower, approximately one quarter of that of the other tautomers.[18] The enol tautomer has the OH group *cis* to the N1–C2 bond.[7] The high frequency of the OH stretching vibration of enol cytosine at 3610 cm^{-1} turns out to be very important for structural assignment of the G–C cluster.

Fig. 4 shows a typical TOF mass spectrum obtained from non-resonant laser ionization with a mixture of guanine and cytosine desorbed in the jet. We observe the G, G–C, G–G, G–G–C and G–G–G parent peaks and a band at $m/z = 112$ which can be assigned to C + H. This species will be discussed below. By setting our mass gate on the parent peak of G–C at $m/z = 262$ and scanning the excitation laser in the region of the electronic absorption of guanine, we were able to obtain the resonant electronic spectrum of the G–C pair displayed in Fig. 5. The hole burning spectrum in Fig. 5 demonstrates that we only observed one G–C isomer in the investigated spectral range. The electronic origin band of this isomer is at 33 314 cm^{-1} and therefore blue-shifted by 446 cm^{-1} relative to the G origin at 32 868 cm^{-1}. The vibronic bands below 200 cm^{-1} can be assigned to hydrogen-bond vibrations of the G–C complex and will be discussed below.

We also observed the electronic origin band of G–C as a spectral peak at the mass of $(C+H)^+$ ($m/z = 112$). This implies a fast G → C proton or hydrogen transfer after either excitation or ionization followed by dissociation. Proton transfer between bases is assumed to be a major source of damage of DNA by ionizing radiation.[21] This important effect will be investigated further in future work.

The R2PI spectra of guanine–5-methylcytosine (G–C5M) and guanine–3-methylcytosine (G–C3M) clusters also appear in Fig. 5. The G–C5M spectrum closely resembles that of G–C

Table 3 Harmonic frequencies (cm^{-1}) and approximate descriptions of the OH and NH stretching vibrations in the S_0 state of the most stable guanine–cytosine dimers. The calculated frequencies are scaled by a factor of 0.893 for all NH stretching frequencies and 0.867 (0.866) for the OH stretching frequency of enol guanine (cytosine) obtained from the best fit to the experimental guanine and cytosine monomer frequencies. Dissociation energies D_0 (cm^{-1}) obtained by correcting the electronic dissociation energy (D_e) at the HF/6–311G(d,p) level for the zero point vibrational energy (ZPE)

	OH		NH_2[a]		9H/7H[b]	1H		NH_2^s		
	G	C	G	C	G	G	C	G	C	D_0
K9K-1			3538	3526	3505	3264	3476	3343	3243	8262
K9K-2			3563	3575	3503	3201	3249	3429	3445	7542
K9K-3			3522	3573	3280	3449	3331	3417	3443	4261
K9K-4			3525	3574	3508	3450	3301	3263	3444	4132
K9K-5			3528	3537	3511	3451	3475	3321	3274	4068
Experiment[a]		3503	3352	3532	3510			3426		
K9E-1		3595	3558	3530	3503	3288		3422	3305	6086
K7E-1		3595	3547	3531	3511[b]	3288		3419	3312	5597
K7E-2		3602	3508	3535	3289[b]	3453		3402	3342	4904
K9E-3		3605	3557	3576	3505	3274		3427	3450	4823
K7E-3		3605	3546	3576	3512[b]	3268		3422	3450	4490
K7E-4		3605	3514	3534	3505[b]	3451		3297	3320	4360
E9K-1	3582		3541	3575	3259		3297	3426	3445	5072
E9K-2	3582		3551	3533	3289		3475	3434	3316	4906
E9K-3	3588		3539	3574	3270		3301	3426	3444	4826
E9K-4	3583		3540	3571	3501		3279	3329	3443	4770
E9K-5	3590		3548	3534	3295		3475	3432	3319	4618
E9K-6	3585		3542	3536	3501		3478	3338	3308	4040
E9E-1	3585	3330	3576	3570	3338			3449	3444	4113
E7E-2	3577	3591	3555	3558	3394[b]			3434	3437	4054
E7E-3	3577	3601	3555	3556	3338[b]			3434	3417	4005
E9E-4	3585	3300	3552	3565	3487			3370	3441	3622
E9E-5	3588	3379	3546	3563	3501			3345	3439	3279

[a] Experimental data from this work are tabulated according to their tentative assignment [b] N7–H

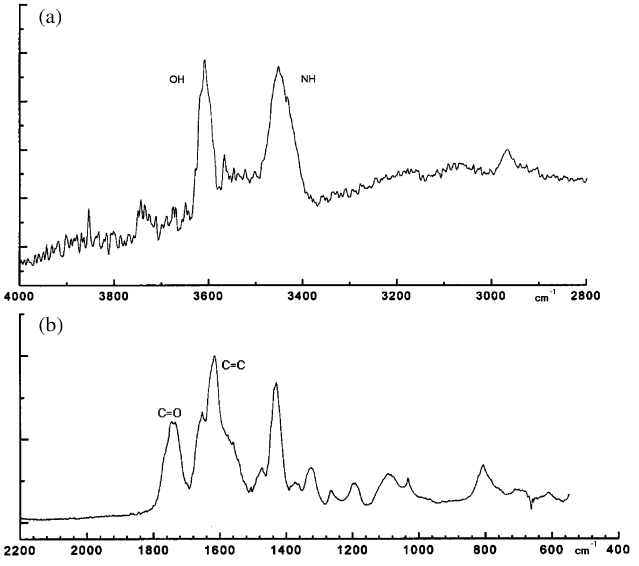


Fig. 3 The IR spectrum of the cytosine monomer at the cytosine parent mass $m/z = 111$ obtained by GC-FTIR-MS at 300 °C. (a) in the spectral region of the OH, NH and CH stretching vibrations; (b) in the region below 2000 cm^{-1}. The spectrum shows that both the enol tautomer of cytosine (OH stretching vibration at 3610 cm^{-1}) and its keto tautomer (C=O stretching vibration at ~1740 cm^{-1}) are present in comparable amounts at 300 °C (see text). Further prominent bands in the IR spectrum are C=C stretching vibrations at ~1610 $cm^{-1}$and in-plane ring deformation vibrations at ~1420 cm^{-1}.

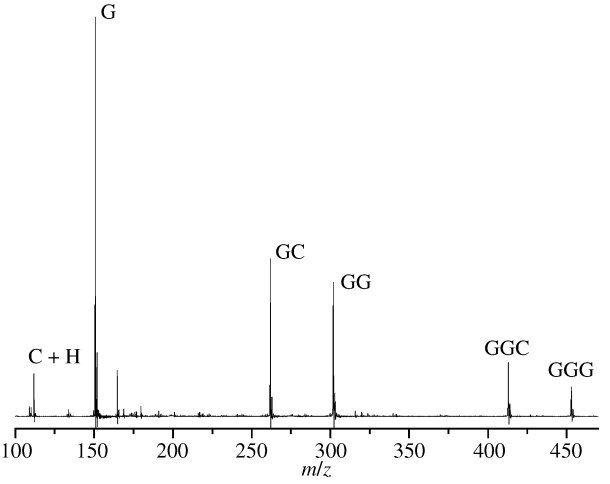


Fig. 4 TOF mass spectrum obtained from non-resonant laser ionization of guanine (G) and cytosine (C) clusters, obtained by laser desorbing a mixture of G and C crystals from a solid graphite sample into a supersonic jet.

and therefore we assign it to the same cluster structure. The similarity of the two spectra clearly suggests that methyl substitution in position 5 of cytosine does not hinder G–C dimer formation and thus the H atom in position 5 of C is not involved in the G–C hydrogen bonds. In contrast, the G–C3M spectrum does not resemble the G–C spectrum at all. It is only weakly but reproducibly structured. Ostensibly the bare nitrogen atom in position 3 is necessary to form the observed G–C dimer. This then implies that the G–C3M structure is very different and we might even speculate that stacking could be involved.

Fig. 6 shows the IR–UV double resonance spectrum of the one observed G–C isomer. The vibrational frequencies of the most stable G–C dimers calculated at the HF/6–31G(d,p) level are shown for comparison. The highest experimental frequency is at 3603 cm^{-1}, which is near the OH stretching frequency of the cytosine enol tautomer at 3610 cm^{-1} (see above). This frequency is significantly higher than the OH stretching frequency of the guanine enol tautomer at 3587 cm^{-1} and higher

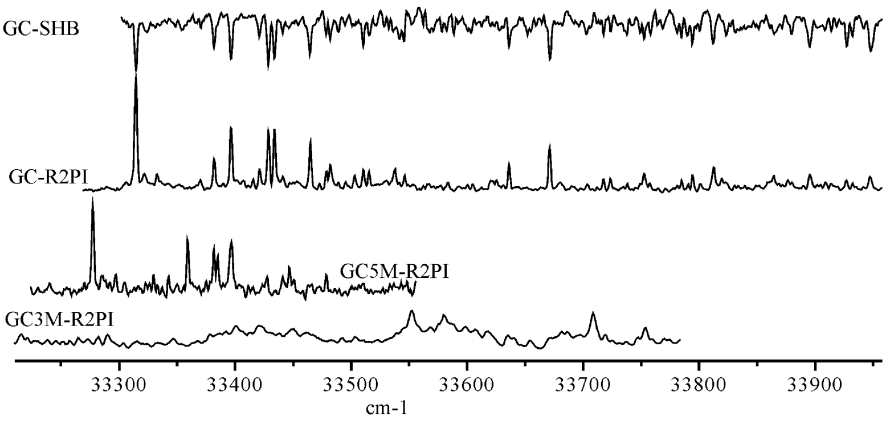


Fig. 5 R2PI and UV–UV hole burning spectrum of the G–C pair ($m/z = 262$) in the range 33300–33700 cm^{-1}. The experiment shows that only one G–C isomer could be observed in this spectral range. The electronic origin band of this isomer is at 33314 cm^{-1} and therefore blue shifted by 446 cm^{-1} relative to the G origin at 32868 cm^{-1}. The vibronic bands below 200 cm^{-1} can be assigned to hydrogen-bond vibrations of the G–C complex. The R2PI spectra of guanine with 5-methylcytosine (G–C5M) and 3-methylcytosine (G–C3M) are shown for comparison.

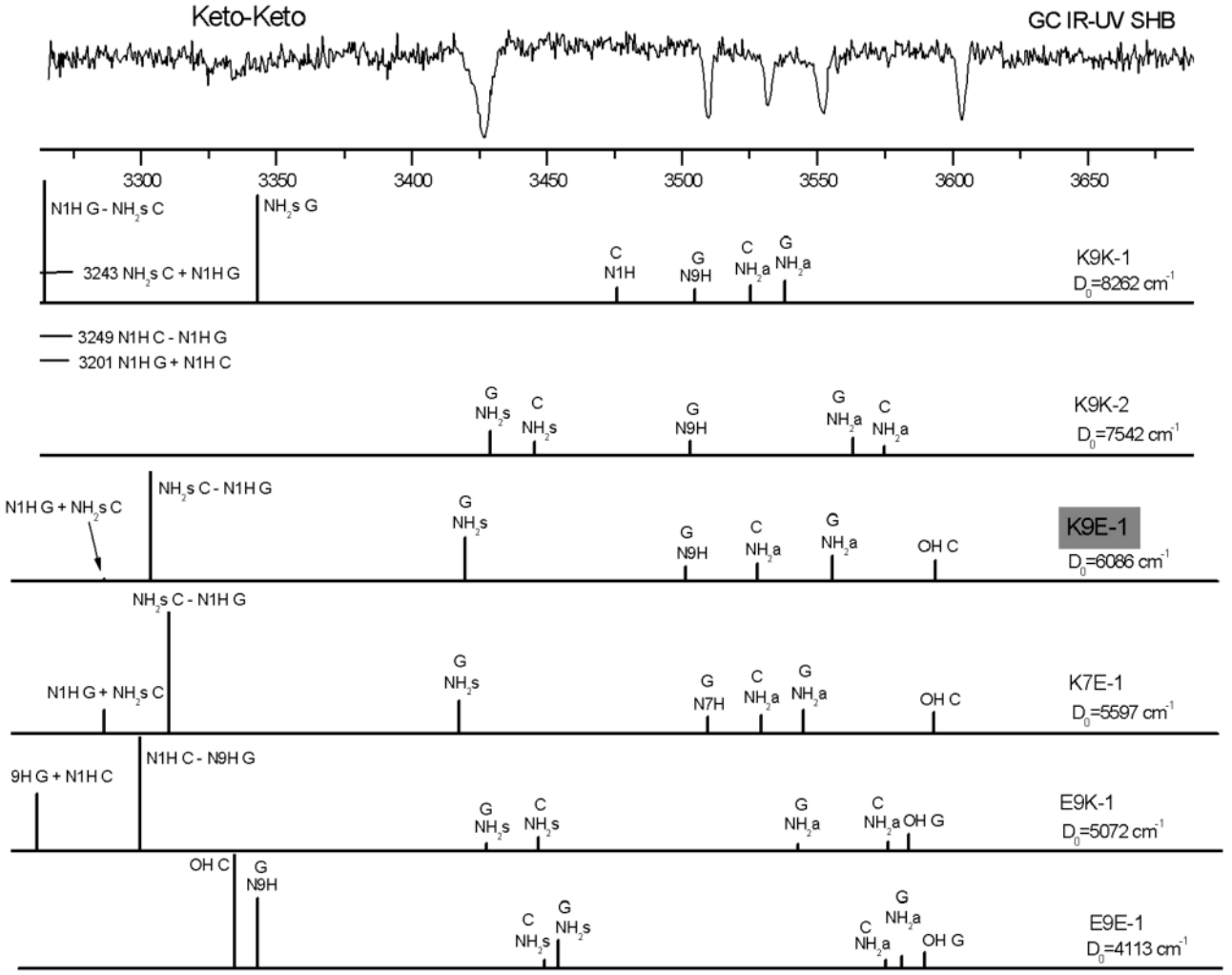


Fig. 6 IR–UV double resonance spectrum of the observed G–C isomer. The vibrational frequencies of the most stable dimers calculated at the HF/6–31G(d,p) level are shown for comparison. The calculated frequencies are scaled by a factor of 0.893 for all NH stretching frequencies and 0.867 (0.866) for the OH stretching frequency of enol guanine (enol cytosine) obtained from the best fit to the experimental guanine and cytosine monomer frequencies. All values are in cm^{-1}.

than any IR frequency of the isomers of the guanine dimer. Hence it is reasonable to assign the band at 3603 cm^{-1} to a free OH stretch of cytosine slightly perturbed by weak interactions with other parts of the dimer. Fig. 6 indeed shows that none of the keto–keto spectra resembles the experimental spectrum. The calculated frequencies for all dimer structures are available as supplementary material.

The IR spectra of K9E-1 and K7E-1 agree well with the experimental spectrum. The calculations show a red shift of 12 cm^{-1} if the OH group of cytosine is proton acceptor in

the HNH···OH hydrogen bond of K9E-1 in good agreement with the small red shift observed (3603 cm^{-1} compared with 3610 cm^{-1}). Isomer K9E-1 exhibits an HNH···OH/NH···N/C=O···HNH bonding similar to the Watson–Crick G–C base pair bonding but with C as enol tautomer. According to the calculations the band at 3552 cm^{-1} is the antisymmetric NH_2 stretching vibration (NH_2^a) of guanine involved in a hydrogen bond with the oxygen atom of enol cytosine. The band at 3532 cm^{-1} can be assigned to the NH_2^a vibration of cytosine involved in an HNH···O=C interaction. According to the calculations the band at 3510 cm^{-1} is the N9H (or N7H) stretching vibration of guanine. The intense and rather broad band at 3426 $cm^{-1}$is the symmetric NH_2 stretching vibration of guanine involved in a hydrogen bond with the oxygen atom of enol cytosine. The NH_2^s vibration of cytosine and the N1H vibration of guanine are coupled owing to the strong hydrogen bonds at these positions. We expect that the scaling factors obtained by fitting the monomer spectra are not appropriate here and the two hydrogen bonds are strong enough to shift the NH_2^s (C)–N1H (G) and NH_2^s (C) + N1H (G) vibrations outside the experimental range. Table 4 summarizes the experimental and calculated vibrational frequencies of the observed guanine–cytosine dimer in the range of the OH and NH stretching vibrations.

In all other respects there exists nearly perfect agreement between the experimental IR spectrum and the calculated spectrum of K9E-1, and also with the less stable K7E-1 pair. We cannot differentiate between K9E-1 and K7E-1 on the basis of the available experimental data.

Comparisons between the experimental spectrum and each of the other calculated spectra further substantiate our assignment. From the nearly 60 calculated cluster arrangements we can exclude those with an H bond donor OH group of guanine and cytosine because their OH stretching vibrations shift to much lower frequencies and fail to explain the experimental band at 3603 cm^{-1}. For the same reason we can exclude dimers with both guanine and cytosine in the keto form. We can also exclude dimers that exhibit free –OH or H acceptor –OH emanating from enol guanine because the calculations show that those OH stretching frequencies remain unchanged relative to the enol monomer or are at most only slightly red shifted. However, compared with the highest experimental G–C frequency of 3603 cm^{-1} the experimental monomer frequencies are 3587 and 3590 cm^{-1}.[22] The only cluster containg enol guanine and exhibiting a blue shift of the –OH stretching frequency is E9K-8 with a repulsive interaction between the –OH group of guanine in the *trans* position and the NH_2 group of cytosine (*cf.*, Fig. 1). Here the calculated –OH stretching vibration of guanine shifts from 3591 cm^{-1} (monomer) to 3600 cm^{-1}. However, this cluster exhibits comparatively little binding energy and its vibrational pattern does not match the experimental pattern at all. Table 3 lists the calculated NH and OH stretching frequencies of the enol–keto and enol–enol dimers. The comparison of experimental and calculated vibrational patterns shows no good agreement for any of the calculated enol–keto and enol–enol dimers.

Table 4 Vibrational assignment of the ground state fundamentals of the observed guanine–cytosine dimer in the range of the OH and NH stretching vibrations. The calculated frequencies are scaled by a factor 0.893 for all NH stretching frequencies and 0.867(0.866) for the OH stretching frequency of enol guanine (cytosine) obtained from the best fit to the experimental guanine and cytosine monomer frequencies. All values are in cm^{-1}

Assignment	Experiment	HF/6-31G(d,p) (K9E-1)
OH (C)	3603	3595
NH_2^a (G)	3552	3558
NH_2^a (C)	3532	3558
N9H (G)	3510	3503
NH_2^s (G)	3426	3422
NH_2^s (C) – N1H (G)	—	3305
N1H G + NH_2^s (C)	—	3288

Table 5 Vibrational assignment of the intermolecular vibrations of the guanine–cytosine dimer in the excited (S_1) state. All values are given in cm^{-1}. The displayed theoretical intermolecular frequencies are calculated for the S_0 state

Assignment	Experiment	HF/6-31(d,p) (K9E-1)
$\gamma_s + \gamma_s$	56	23
—	68	
σ_{as}	82	79
—	107	
δ	115	99
σ_s	120	116
$\sigma_{as} + 68$	151	
$\sigma_{as} + \sigma_{as}$	164	
—	168	
$\sigma_{as} + 107$	189	
$\sigma_{as} + \delta$	196	
—	201	
—	224	
$\sigma_{as} + \sigma_{as} + 68$	232	

Fig. 5 and Table 5 display the intermolecular vibrations of the observed G–C base pair in the S_1 state. The two low-frequency peaks in the spectrum, marked with a circle, are absent when we use krypton instead of argon as the drive gas. This suggests that they may be due to hot bands or result from dissociating clusters with argon,[2] rather than G–C vibrations. The bands at 82, 115 and 120 cm^{-1} may be assigned to modes σ_{as}, δ and σ_s, respectively. The intermolecular vibrations are displayed schematically in Fig. 2(a). The agreement between the experimental (S_1 state) and calculated (S_0 state) vibrational patterns is not very good and other G–C pairs match the experimental pattern just as well or better. Unfortunately, quality *ab initio* calculations of the S_1 state vibrations of a cluster the size of G–C are out of reach at the moment. For this reason, the intermolecular vibrations in the S_1 state cannot yet be called on for structural assignment.

Conclusions

From comparison of the IR spectra of the G and C monomers and the G–C dimer in the range of the OH and NH stretching vibrations with *ab initio*-calculated vibrational frequencies and dimer stabilities, we were able to assign the observed G–C isomer to a cluster with C as enol tautomer and HNH···OH/NH···N/C=O···HNH G···C hydrogen bonding. The observed cluster is apparently not the Watson–Crick G–C base pair, which according to calculations should be the most stable cluster when isolated in the gas phase. Why did we not appear to detect the Watson–Crick G–C base pair despite its greater stability? One reason could be that it absorbs outside the investigated spectral range. Further work will be aimed at investigating a broader spectral range for other G–C isomers. Another reason may be that laser desorption leads to a tautomer distribution of cytosine which differs from that following thermal heating. Laser desorption leading to cytosine predominantly in the enol form could explain our results. We will also investigate this aspect further, for example by combining thermal and laser vaporization to produce G–C clusters in different ways. However, we note that when desorbing from a mixture of 1-methylcytosine and guanine, we did not observe any G–C1M clusters. Since in this derivative the enol tautomer is blocked, this finding could be consistent with the explanation that the Watson–Crick cluster in the keto form does not absorb in the given wavelength range.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft and the Israel Science Foundation, founded by the Israel Academy of Sciences and Humanities.

References

1 J. D. Watson and F. H. C. Crick, *Nature (London)*, 1953, **171**, 737.
2 H. Urabe, H. Hayashi, Y. Tominaga, Y. Nishimura, K. Kubota and M. Tsuboi, *J. Chem. Phys.*, 1985, **82**, 531.
3 S. P. A. Fodor and T. G. Spiro, *J. Am. Chem. Soc.*, 1986, **108**, 3198.
4 M. Elstner, P. Hobza, T. Frauenheim, S. Suhai and E. Kaxiras, *J. Chem. Phys.*, 2001, **114**, 5149.
5 M. Kratochvil, J. Sponer and P. Hobza, *J. Am. Chem. Soc.*, 2000, **122**, 3495.
6 O. Shishkin , J. Sponer and P. Hobza, *J. Mol. Struct.*, 1999, **477**, 15.
7 P. Hobza and J. Sponer, *Chem. Rev.*, 1999, **99**, 3247.
8 V. Spirko, J. Sponer and P. Hobza, *J. Chem. Phys.*, 1997, **106**, 1472.
9 M. Kabelá and P. Hobza, *J. Phys. Chem., B*, 2001, **105**, 5804.
10 E. Nir, C. Janzen, P. Imhof, K. Kleinermanns and M. S. de Vries, *J. Chem. Phys.*, 2001, **115**, 4604.
11 E. Nir, K. Kleinermanns and M. S. de Vries, *Nature (London)*, 2000, **408**, 949.
12 G. Meijer, M. S. de Vries, H. E. Hunziker and H. R. Wendt, *Appl. Phys., B*, 1990, **51**, 395.
13 R. N. Pribble, C. Gruenloh and T. S. Zwier, *Abstr. Pap. Am. Chem. Soc.*, 1996, **211**, 176.
14 S. Ishikawa, T. Ebata, H. Ishikawa, T. Inoue and N. Mikami, *J. Phys. Chem.*, 1996, **100**, 10531.
15 S. Sommer, R. Kamps, S. Schumm and K. Kleinermanns, *Anal. Chem.*, 1997, **69**, 1113.
16 C. Janzen, Thesis, Düsseldorf, 1996, .
17 M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez and J. A. Pople, *Gaussian 98, Revision A.4*, Gaussian, Inc., Pittsburgh, PA, 1998.
18 R.D. Brown, P.D. Godfrey, D. McNaughton and A. Pierlot, *J. Am. Chem. Soc.*, 1989, **111**, 2308.
19 P. Colarusso, Z. KeQing, G. Buji and P.F. Bernath, *Chem. Phys. Lett.*, 1997, **269**, 39.
20 K. Szczepaniak, M. Szczesniak and W.B. Person, *Chem. Phys. Lett.*, 1988, **153**, 39.
21 M. Hutter and T. Clark, *J. Am. Chem. Soc.*, 1996, **118**, 7574.
22 M. Mons, personal communication, 2001.