Excited state tautomerism of the DNA base guanine: a restricted open-shell Kohn-Sham study

Holger Langer and Nikos L. Doltsinis ⋆

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

⋆ corresponding author. email: nikos.doltsinis@theochem.ruhr-uni-bochum.de;

Fax: ++49 234 3214045

Preprint submitted to J. Chem. Phys. *12 December 2002*

The relative stabilities of the six lowest energy tautomers of the DNA base guanine have been investigated in the first excited singlet state, S_1, employing the restricted open-shell Kohn-Sham method. Comparison of the S_1 optimized geometries to the respective ground state structures reveals large distortions for the keto tautomers, whereas the enol tautomers remain essentially planar. Harmonic vibrational spectra in the S_1 states have been calculated using the ROKS potential energy surfaces. Adiabatic excitation energies together with characteristic vibrational features of the individual guanine tautomers enable us to unambiguously assign recent experimental IR-UV spectra. Velocity autocorrelation functions obtained from adiabatic excited state Car-Parrinello molecular dynamics simulations demonstrate that anharmonic effects only play a minor role.

1 Introduction

Nucleic acids are of fundamental biological importance due to the role they play in DNA. As first suggested by Chargaff [1–3] and later shown in detail by Watson and Crick [4], the sequence of the guanine – cytosine and adenine – thymine hydrogen bonded base pairs stores the genetic code.

All DNA bases can exist in a variety of tautomeric forms giving rise to a large number of possible base pair combinations. In the case of guanine, for example, ground state energies of the four most stable tautomers have been

calculated to lie within a range of 7 kJ/mol [5,6]. However, only a single guanine tautomer, namely the so–called N9H-*keto* form (see Fig.1 for a schematic representation of this and other guanine tautomers discussed in the following), is usually present in DNA, whereas other tautomeric forms may be responsible for genetic damage [7–9].

Experimental study of DNA base pairs has been complicated by interference of both the backbone and the solvent. Theory, on the other hand, is not yet capable of explicitly taking into account the biological environment. Therefore, comparison of experimental and theoretical data has been problematic. Recent advances in experimental techniques, however, have made possible the study of isolated DNA bases and base pairs through supersonic jet expansion techniques [6, 10–12]. These developments have paved the way to meaningful comparisons of *ab initio* electronic structure calculations to experimental measurements. In particular, the vibrational spectrum of guanine in the S_1 excited electronic state has been recorded independently by two different research groups [6, 11, 13]. Both groups agree that the spectrum is in fact a superposition of vibrational spectra of at least three different guanine tautomers. However, there is discrepancy regarding the assignment of spectral lines to specific tautomers.

This is where first principles electronic structure theory can make a significant contribution. The first issue that needs to be addressed are the relative ground state energies of the various tautomers. The energies of 36 stable tautomers

have been calculated at the Hartree-Fock and MP2 levels of theory [5,14]. The energetic ordering has been found to be very sensitive to the method used and the basis set chosen [5,14,15]. Accurate *ab initio* calculations on electronically *excited* states still present a major challenge to theoreticians even for relatively small molecules. For single point calculations of large molecules only the CASSCF, CASPT2, CIS, and TDDFT methods are applicable in view of current computational resources, in particular if, as in the present paper, many different isomers have to be considered. Unfortunately, no analytic gradients are publicly available yet for CASPT2 and TDDFT, although progress along these lines has been made [16–18]. CASSCF is computationally rather expensive; it may therefore be used only for a limited number of calculations. Excited state geometries and vibrational spectra can be computed cheaply using the CIS method. However, the accuracy of CIS for the type of system under consideration is highly questionable [19]. For these reasons only few theoretical excited state vibrational spectra of guanine have been reported in the literature [6,15]. Nir *et al.* [6] computed the normal mode frequencies of the *trans*-N9H-*enol* tautomer using CIS. Vibrational frequencies for additional tautomers were calculated in the electronic ground state using the MP2 method [6] and compared to the measured S_1 spectra in the hope that ground and excited state spectra are very similar. This would be the case if the S_1 geometry does not differ significantly from the ground state geometry.

As an alternative to the above-mentioned traditional *ab initio* quantum chem-

istry approaches, we employ in this paper the recently developed restricted open-shell Kohn-Sham (ROKS) method [20]. This density functional method has been used previously to optimize geometries and perform molecular dynamics simulations in the S_1 first excited singlet state [20–22]; see Refs. [23,24] for reviews. For the six most stable guanine tautomers (see Fig. 1) we have first optimized the ground state structures and compared their energies. In a second step, the vertical S_1 ($\pi \rightarrow \pi^*$) excitation energies have been computed before optimizing the excited state geometries. Changes in configuration and stability in relation to the ground state have been monitored. Finally, harmonic vibrational spectra have been computed for all tautomers and compared to the experimental spectra. We are thus able to suggest an assignment of measured peaks to individual tautomeric forms of guanine. Furthermore, we point out distinctive features of their vibrational spectra which have not yet been considered by experimentalists and may help in distinguishing between different tautomers.

2 Computational Details

For the calculation of S_1 excited state properties, i.e. vertical and adiabatic excitation energies, optimized molecular structures, and vibrational spectra, we have used the restricted open-shell Kohn-Sham method (ROKS) [20–24] implemented in the `CPMD` program package [23, 25]. In the ROKS framework,

the S_1 energy is calculated by optimizing the set of $n+1$ (n is half the number of electrons) orthonormal Kohn-Sham orbitals $\{\phi_i^{(1)}\}$ so as to minimize the expression

$$E(s_1) = 2E(m) - E(t) \quad , \tag{1}$$

where $E(m)$ is the energy expectation value of the mixed state

$$|m\rangle = |\phi_1^{(1)}\bar{\phi}_1^{(1)} \cdots \phi_n^{(1)}\bar{\phi}_{n+1}^{(1)}\rangle \tag{2}$$

and $E(t)$ is the energy of the triplet state

$$|t\rangle = |\phi_1^{(1)}\bar{\phi}_1^{(1)} \cdots \phi_n^{(1)}\phi_{n+1}^{(1)}\rangle \quad . \tag{3}$$

The corresponding singlet excited state wavefunction is given by

$$|s_1\rangle = \frac{1}{\sqrt{2}}\left\{|\phi_1^{(1)}\bar{\phi}_1^{(1)} \cdots \phi_n^{(1)}\bar{\phi}_{n+1}^{(1)}\rangle + |\phi_1^{(1)}\bar{\phi}_1^{(1)} \cdots \bar{\phi}_n^{(1)}\phi_{n+1}^{(1)}\rangle\right\} \tag{4}$$

using a common set of orbitals $\{\phi_i^{(1)}\}$. Based on this symmetry-adapted representation of the S_1 state the *orbital-dependent* ROKS density functional is minimized with respect to orbital variations while imposing their orthonormality [20]. A similar technique was pioneered by Ziegler et al. [26]; further variants and generalizations have been proposed recently [27–31].

The ROKS calculations presented here were carried out using a plane wave expansion of the one-electron orbitals in conjunction with Troullier-Martins pseudopotentials [32] for the core electrons. A plane wave cutoff of 100 Ry turned out to be necessary to obtain converged harmonic frequencies in the S_1 state, whereas excitation energies and geometry optimization only required

a 70 Ry cutoff. The dimensions $15\times12\times7.5\times$ Å were chosen for the simulation cell for all static calculations. Spurious interactions between periodic images were eliminated by employing Hockney's poisson solver [33,34]. The molecular dynamics simulations were performed in a periodically repeated cubic box of length 15 Å with a time step of 2 a.u. and a fictitious electron mass of 400 a.u. at the constant temperature of 100 K. All plane wave calculations were carried out using the BLYP exchange correlation functional [35,36] and the corresponding results are denoted "BLYP/p.w.".

Additional, more conventional quantum chemical methods, i.e. CIS, CASSCF, and TDDFT, were applied as implemented in the Gaussian 98 program package [37] using the 6-31G**, cc-pVDZ, 6-31++G**, and aug-cc-pVDZ basis sets. TDDFT calculations were performed with the BLYP and B3LYP [38,39] exchange-correlation functionals. For the CASSCF calculations, an active space of six π electrons in three π and three π^* orbitals was chosen.

3 Results and Discussion

3.1 Ground state structures and energies

The molecular structures of the six most stable guanine tautomers depicted schematically in Fig. 1 have been optimized in the electronic ground state using DFT with various basis sets and exchange-correlation functionals. All

tautomers have essentially planar aromatic rings; merely the amino group deviates from planarity exhibiting a considerable degree of pyramidalization. The corresponding absolute dihedral angles $\angle H_{21}N_{2'}C_2N_1$ (dh1) and $\angle H_{22}N_{2'}C_2N_3$ (dh2) (see Fig. 1 for nomenclature) are listed in Tab. 1 for three different calculations. First we observe that the BLYP plane wave/pseudopotential results with a 70 Ry cutoff are in qualitative agreement with those obtained using BLYP with the Gaussian-type basis set 6-31++G**. We have performed an additional geometry optimization of N9H-*keto* guanine using the aug-cc-pVTZ basis set yielding the dihedral angles 31.5° (dh1) and 13.4 ° (dh2), which are in even better agreement with the plane wave results. This suggests that our plane wave basis set is superior to the 6-31++G** basis, although the latter yields, for our purposes, sufficiently converged structural data. Similar results, i.e. 33.1° (dh1) and 13.2° (dh2), have been obtained by Guerra et al. [40] using the BP86 functional and the Slater-type TZ2P basis. Furthermore, as can be seen by Tab. 1, there is little difference between the geometries obtained with the GGA functional BLYP and those calculated using the hybrid functional B3LYP. We are therefore confident that the BLYP functional in combination with a plane wave basis set truncated at 70 Ry and pseudopotentials yields reliable structural information in the present case.

As the data in Tab. 1 demonstrate, there is a clear distinction between the structures of the *keto* and *enol* tautomers. The dihedral angle dh1 is significantly larger for the *keto* tautomers compared to the *enol* tautomers, whereas

the angle dh2 is somewhat larger for the *enol* tautomers. In contrast to the *enol* tautomers the dihedral angles of the *keto* tautomers are highly asymmetric. This can be easily understood; the presence of the hydrogen atom bonded to nitrogen N_1 in the *keto* case forces hydrogen H_{21} further out of plane. Moreover, we notice this asymmetry being most pronounced for the N7H-*keto* tautomer. Interestingly, our calculations also predict the N7H-*keto* tautomer to have the lowest energy of all six tautomers investigated, closely followed by N9H-*keto* guanine (see Tab. 2) which is roughly 3 kJ/mol higher in energy according to our BLYP calculations. All *enol* tautomers are seen to be significantly less stable than their *keto* counterparts. Moreover, for both N9H-*enol* and N7H-*enol* guanine the *cis* isomers are less stable than their *trans* counterparts. In particular, there is a large gap (29.2 kJ/mol in the case of BLYP/p.w.) between the two N7H-*enol* tautomers.

As for the molecular structures, we observe good agreement among the BLYP data obtained with different basis sets and the B3LYP results. Comparison of the latter with the MP2/6-31G** energies by Ha et al. [5] reveals great similarity. Nir et al. [6] have performed MP2 calculations using the bigger basis set 6-311G** and obtained only slightly different energies. However, they find N9H-*keto* guanine to be energetically most favourable being 1 kJ/mol more stable than the N9H-*keto* tautomer. Shukla et al. [15] have calculated relative energies using MP2 with a mixed basis consisting of a 6-311+G* basis set for the nitrogen atom of the amino group and a 4-31G basis for all other atoms.

Although their energetic ordering is in accordance with most other sets of data, their relative energies indicate that the results are far from converged with respect to the size of the basis set.

3.2 Vertical electronic excitations

For all tautomers we have computed the excitation energies for the lowest $\pi \rightarrow \pi^*$ transition using the ROKS method in conjunction with the BLYP functional and plane waves at their ground state equilibrium structures described in the previous subsection. For comparison we have again carried out alternative (in this case TDDFT) calculations using conventional Gaussian-type basis sets. Our results are presented in Tab. 3 together with pertinent data from the literature.

Let us first compare the TDDFT BLYP/6-31++G** numbers to the ROKS BLYP/p.w. values. Both methods predict N9H-*keto* guanine to have the highest vertical excitation energy followed by the N9H-*enol* tautomers 0.2 – 0.3 eV below. There further is agreement on the N7H tautomers having lower vertical excitation energies than their N9H counterparts. ROKS finds N7H-*keto* guanine 0.35 eV below N9H-*keto* guanine in accord with the TDDFT value of 0.28 eV. The fact that the *trans* isomers always have higher excitation energies than the corresponding *cis* isomers also emerges as a common pattern. The most significant point at which the two approaches give different answers is

the energy gap between the N9H-*enol* and the N7H-*enol* tautomers. According to ROKS this gap is less than 0.05 eV whereas TDDFT finds differences as large as 0.45 eV. Of course we should point out that there is a systematic shift of all ROKS excitation energies to smaller values which has been discussed previously in the literature [20, 21]. In the present case, the differences between the ROKS and the TDDFT numbers range from 0.27 to 0.76 eV. The TDDFT results, on the other hand, are likely to be much closer to the true values. This becomes apparent when we compare the TDDFT vertical excitation energy for N9H-*keto* guanine of 4.39 eV to the experimental value for crystalline 9-ethyl guanine of 4.46 eV. We have noticed a significant reduction of TDDFT excitation energies upon inclusion of diffuse functions in the basis set by comparison of TDDFT BLYP 6-31++G** and 6-31G** results.

Use of the B3LYP hydrid functional gives the same energetic ordering; however the excitation energies are shifted to larger values by roughly half an electron Volt. In the basis set limit, the B3LYP value for N9H-*keto* guanine should be comparable to the CASPT2 result of 4.76 eV by Fülscher et al. [41]. Although the semiempirical CIPSI calculations by Menucci et al. [42] yield the same value for N9H-*keto* guanine, they predict the N7H-*keto* energy to be higher by 0.18 eV conflicting with all other methods. The CIS calculations of Shukla et al. [15] overestimate all excitation energies by approximately 2 eV. In particular, they find *trans*-N9H-*enol* guanine to have the highest excitation energy of all tautomers.

Although the focus of this work is on the lowest $\pi \rightarrow \pi^*$ transition, since this state has clearly been associated with the experimental vibronic spectra [6, 11, 13], we should also discuss briefly other close-lying electronic states. First of all, our TDDFT calculations confirm that the intensity of the $\pi \rightarrow \pi^*$ transition is typically one order of magnitude larger than that of the $n \rightarrow \pi^*$ transition. The latter is therefore difficult to observe experimentally [11]. Furthermore, we observe that for basis sets without diffuse functions, the $\pi \rightarrow \pi^*$ excitation is lowest in energy for all tautomers. Upon inclusion of diffuse basis functions we notice a small degree of mixing of the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ states for both *keto* tautomers as well as for *cis*-N7H-*enol* guanine. In the case of N9H-*keto* and *cis*-N7H-*enol* guanine the $n \rightarrow \pi^*$ transition becomes the lowest excitation. These observations are in agreement with recent theoretical studies on the radiationless decay of the $\pi \rightarrow \pi^*$ state of cytosine [43], where $n \rightarrow \pi^*$ states present an efficient route for relaxation to the ground state.

3.3 S_1 structures, relative energies and adiabatic electronic excitations

Geometry optimization of the six most stable guanine tautomers in the first excited singlet state S_1 using the ROKS BLYP/p.w. method leads to significant structural changes relative to the ground state. The S_1 structures of *keto* and *enol* tautomers exhibit distinctive features. Both *keto* tautomers are found to have out of plane distorted six-membered aromatic rings as well as

pyramidalized amino groups. In the case of the *enol* tautomers, on the other hand, both the aromatic rings and the amino group are planar, whereas the hydrogen atom bonded to carbon C_8 shows out of plane distortion.

The N9H-*keto* tautomer clearly exhibits the largest differences between excited state and ground state structures. The C_2N_3 bond length, for instance, is elongated by 0.12 Å in the S_1 state causing a decrease in the $C_2N_3N_{2'}$ bond angle by 7°. A measure for the strong out of plane distortion of the six-membered aromatic ring is the change of the dihedral angle $C_6N_1C_2N_3$ from $-0.7°$ in the ground state to $+32.2°$ in the S_1 state. In the case of N7H-*keto* guanine this distortion is much less pronounced, the above dihedral angle being just under 10°. Here, the hydrogen atoms H_1 and H_7 show the largest deviations from planarity with the corresponding dihedral angles ranging from 15° to 21°.

For both isomers of N9H-*enol* guanine the aromatic rings remain planar in the S_1 state. Nevertheless, there are significant geometrical changes. Besides an elongation of the N_7C_8 bond by 0.13 Å the biggest deviation from C_s symmetry is given by the hydrogen atom H_8 with corresponding dihedral angles of up to 40°. It should be noted that *cis* and *trans* isomers behave in a similar fashion.

The latter statement is not entirely true for the N7H-*enol* tautomers. In this case, the *trans* isomer clearly shows larger distortions relative to the ground state. As for N9H-*enol* guanine, the N_7C_8 bond stretches by 0.13 Å in the case of *trans*-N7H-*enol*, 0.01 Å more than for *cis*-N7H-*enol* guanine. The *cis*-

N7H-*enol* tautomer is found to be the only *enol* tautomer with an in plane hydrogen H_8, whereas *trans*-N7H-*enol* guanine shows dihedral angles of about 25° at this position.

The relative energies of the six most stable guanine tautomers in the S_1 excited state are compiled in Tab. 4. First of all, we observe that the energetic ordering is largely unchanged compared to the ground state N7H-*keto* guanine still being lowest in energy. However, the energy difference to the remaining tautomers is now larger than in the ground state by roughly 10 kJ/mol. On the other hand, the energy gap between N9H-*keto* guanine and the *cis*-N9H-*enol* tautomer has shrunk from 11 kJ/mol to 4 kJ/mol. Partially, the reason for this is the relative stabilization of *cis*-N9H-*enol* guanine in the S_1 state, which also leads to the *trans*-N9H-*enol* tautomer being the less stable N9H-*enol* tautomer, in contrast to the ground state.

We have computed adiabatic excitation energies by subtracting the minimum ground state energy from the minimum S_1 energy. Despite having neglected zero point vibration, the theoretical adiabatic excitation energies should be comparable to the experimentally measured 0–0 transitions [6,13]. Our ROKS BLYP/p.w. values are listed in Tab. 5 together with experimental 0–0 transition energies measured and assigned to the different tautomers by Nir et al. [6] and Mons et al. [13], respectively. Similar to the vertical excitation energies, the ROKS results appear to be too low by approximately 1 eV. However, the relative ordering already constitutes important information that may be used

in the assignment of experimental IR-UV spectra. We would like to point out the striking similarity between the two sets of experimental data. Basically, it seems that both groups have indeed recorded more or less identical spectra, merely the assignment is different. For example, if we interchanged the values for the N9H-*keto* and N7H-*keto* tautomers in the data of Nir et al. [6], there would be good agreement between all three sets of data. Both theory and experiment would then see the 0–0 transition of N9H-*keto* roughly 0.1 eV above that of N7H-*keto* guanine. Furthermore, if we swap the *trans*-N7H-*enol* and *trans*-N7H-*enol* numbers in the set of Mons et al. [13], both experiments match. According to our ROKS calculations the lower adiabatic excitation energy should be assigned to the *trans*-N7H-*enol* tautomer.

3.4 Excited state vibrational spectra

Harmonic vibrational spectra of all six guanine tautomers considered in this work have been calculated in the S_1 state by means of the ROKS BLYP/p.w. method. We would like to emphasize that a larger plane wave cutoff of 100 Ry has proven crucial in order to obtain converged frequencies. Moreover, extremely careful geometry optimization is required if imaginary frequencies are to be avoided. Again, we have performed additional quantum chemical calculations using the CIS and CASSCF methods for the sake of comparison. All the results are summarized in Tab. 6. In the following, we shall discuss characteristic features of the individual tautomers that may be used to identify

their experimental IR-UV spectra.

Distinguishing *keto* from *enol* vibrational spectra is straightforward; the former exhibit a double peak at the high energy end of the spectrum between 3500 and 3600 cm^{-1}, whereas the latter show a double peak shifted to higher energies by roughly 100 cm^{-1}, i.e. between 3600 and 3700 cm^{-1} (see ROKS data in Tab. 6). This is in good agreement with experimental observations in the ground state [6, 13], where the *enol* highest energy lines are measured at roughly 3590 cm^{-1}, approximately 80 cm^{-1}above the highest *keto* lines. Our ROKS values for the *enol* tautomers all lie between 3628 and 3637 cm^{-1}, with the exception of *cis*-N7H-*enol* guanine whose highest peak is at 3678 cm^{-1}, because of steric effects as we shall discuss below. Nir et al. [44] have previously stated that vibrational frequencies in the S_1 are on average 10 % larger compared to the ground state. Therefore, our ROKS results appear to be remarkably close to the true values. Thus, in contrast to the CIS and CASSCF numbers there is no need to rescale the ROKS frequencies by an empirical fudge factor.

In the case of the *enol* tautomers, the highest frequency vibration (mode 42) can be identified as the OH stretching mode, whereas mode 41 is the asymmetric stretching mode of the amino group. For all *enol* tautomers, the energy gap between these two vibrations is at most 4 cm^{-1} with the exception of *cis*-N7H-*enol* guanine, which has a gap of 56 cm^{-1}. This shift in the OH vibrational energy is due to the presence of the hydrogen atom H_7 making this

tautomer altogether less stable than the other five tautomers. It is therefore also less likely to be observed experimentally.

The highest energy normal modes of the *keto* tautomers are the N_9H (N9H-*keto*) and N_7H (N7H-*keto*) vibrations, respectively. Mode 41 is the asymmetric stretching mode of the amino group. Our ROKS calculation for N9H-*keto* guanine reproduces the experimental energy separation between these two lines of 16 cm^{-1} measured by Nir et al. [6] in the ground state rather well. However, ROKS finds a separation for N7H-*keto* guanine of 62 cm^{-1}, significantly larger than the experimental ground state value of 8 cm^{-1}. There is, at present, no conclusive evidence to decide whether this discrepancy can be attributed to inaccuracies of the ROKS method or if we are dealing with a real effect induced by geometrical changes in the S_1 state. As we have discussed in Sect. 3.3, the position of the hydrogen atom involved in the highest energy vibration, namely H_7, changes significantly in the excited state. Unfortunately, it has so far been impossible to probe the high energy spectral region experimentally in the S_1 state [6, 13].

According to our ROKS data, the spectral region between 1430 and 1571 cm^{-1} should be well suited to differentiate between the N9H-*keto* and N7H-*keto* tautomers. Here we find no vibrational modes for N9H-*keto*, whereas N7H-*keto* guanine has two spectral lines at 1479 and 1516 cm^{-1}, respectively. Our CIS calculations suggest that the former is sufficiently intense to be recognized in experiment.

Another interesting spectral region is the interval between 674 and 794 cm^{-1} where N7H-*keto* guanine is inactive, but the N9H-*keto* tautomer shows lines at 701, 736, 755, and 768 cm^{-1}. Again, by analyzing the corresponding intensities obtained from CIS calculations we are confident that at least some of those four lines are detectable experimentally.

Nir et al. [6] report S_1 vibrational modes at 333 and 448 cm^{-1}, which they assign to the N7H-*keto* tautomer, as well as lines at 336 and 467/470 cm^{-1} [6] assigned to the N9H-*keto* tautomer. ROKS predicts vibrations at 332 and 441 cm^{-1} for N9H-*keto* guanine and at 333, 459, and 469 cm^{-1} for N7H-*keto* guanine. Together with the positions of the respective 0–0 transitions (see Tab. 5), our results thus suggest that the assignment for the *keto* tautomers by Nir et al. [6] should be interchanged.

Do our calculations also provide enough spectral information to distinguish N7H-*enol* from N9H-*enol* tautomers? We have determined characteristic vibrational frequencies of N7H-*enol* guanine at 695 (*trans*) and 698 cm^{-1} (*cis*). The nearest lines of the N9H-*enol* tautomers are separated by at least 30 cm^{-1}. A further promising spectral region lies between 1073 and 1136 cm^{-1}, where the N7H-*enol* tautomer is inactive but N9H-*enol* guanine exhibits three bands at 1086, 1088, and 1090 cm^{-1}, the first of which should have relatively strong intensity.

The last issue that we would like to address is the distinguishability of *trans*

and *cis enol* tautomers. We have already discussed the pronounced blue shift of the OH vibrational frequency in the case of *cis*-N7H-*enol* guanine due to steric effects. Furthermore, there is a characteristic spectral line at 1624 cm^{-1} corresponding to mode 37 (predominantly stretching of the C_5C_6 bond) of *trans*-N7H-*enol* guanine. However, our CIS calculations predict the signal to be rather weak for this mode.

Since there is no steric hinderance in the case of N9H-*enol* guanine, *cis* and *trans* isomers are expected to have very similar spectra. On the whole, this assumption is confirmed by our theoretical vibrational frequencies. Moreover, preliminary molecular dynamics simulations in the S_1 state indicate that the rotational barrier between the two isomers may be sufficiently low to allow the OH group to rotate more or less freely [45]. The presence of *cis*-N9H-*enol* guanine in molecular beams may be verified experimentally by analyzing the spectral region around 1419 cm^{-1}, where our ROKS calculations predict a line for the *cis* tautomer separated by at least 36 cm^{-1} from the nearest lines of the *trans* tautomer. The vibrational mode at 1419 cm^{-1} has a large C_6N_1 stretch component and, according to our CIS data, has rather large intensity.

Finally, we have studied the significance of anharmonic effects by comparing the harmonic ROKS vibrational spectrum of *cis*-N9H-*enol* guanine to the Fourier transform of the velocity autocorrelation function obtained from an adiabatic Car-Parrinello molecular dynamics simulation in the first excited state also using the ROKS method. A graphical representation of the two

spectra is shown in Fig. 2. Considering the error bars (the anharmonic spectrum has a resolution of 13 cm^{-1} and has been corrected for unphysical spectral shifts due to the fictitious Car-Parrinello electronic degrees of freedom), our conclusion is that there are no dramatic anharmonic shifts and the harmonic approximation can be safely applied for most vibrational modes. A notable exception is the region around 1150 cm^{-1}, where the anharmonic peak is red-shifted by roughly 40 cm^{-1} compared to the harmonic frequency. Visualization of the respective harmonic eigenmode reveals a strongly mixed vibration with large contributions of the C_8H and OH stretching modes as well as out of plane distortions of the aromatic ring system.

4 Conclusions

We have employed the restricted open-shell Kohn-Sham approach to calculate structural, energetic and spectroscopic quantities of the six most stable guanine tautomers in the first excited singlet state S_1. In addition, we have carried out standard Kohn-Sham density functional calculations in the electronic ground state S_0.

Using the BLYP functional we have found the N7H-*keto* guanine to be the most stable tautomer both in the ground state and in the first excited $\pi \rightarrow \pi^*$ state, slightly lower in energy than the N9H-*keto* tautomer. We observe substantial geometrical distortions in the S_1 state compared to the ground

state, in particular for the N9H-*keto* tautomer whose six-membered aromatic ring is heavily nonplanar.

Our theoretical adiabatic S_1 excitation energies can be compared to experimental 0–0 transition energies providing hints as to the spectral positions of the individual guanine tautomers. In combination with our ROKS S_1 vibrational spectra, the present results facilitate the assignment of experimental IR-UV and REMPI spectra of jet-cooled guanine.

In addition, this work has demonstrated that excited state vibrational frequencies can be obtained fairly reliably using the ROKS method. In particular, unlike the more conventional CIS and CASSCF methods, ROKS does not require any rescaling of vibrational frequencies.

Acknowledgments

The authors are indebted to Prof. D. Marx for his interest in the subject and for lending his support to this study. We would like to thank Prof. K. Kleinermanns as well as Drs. R. Rousseau and H. Forbert for stimulating discussions. The John von Neumann Institute for Computing (FZ Jülich), the Rechnerverbund NRW, and RUB are gratefully acknowledged for providing computational resources. This work was partially supported by DFG and FCI.

References

[1] E. Chargaff, Experimentia **6**, 201 (1950).

[2] E. Chargaff, Experimentia **26**, 810 (1970).

[3] E. Chargaff, Science **172**, 637 (1971).

[4] J. D. Watson and F. H. C. Crick, Nature **171**, 737 (1953).

[5] T. K. Ha, H. J. Keller, R. Gunde, and H. H. Gunthard, J. Phys. Chem. A **103**, 6612 (1999).

[6] E. Nir, C. Janzen, P. Imhof, K. Kleinermanns, and M. S. de Vries, J. Chem. Phys. **115**, 4604 (2001).

[7] J. D. Watson and F. H. C. Crick, Nature **171**, 964 (1953).

[8] R. Knippers, *Molekulare Genetik* (Thieme, Stuttgart, 1997).

[9] E. S. Kryachko, Int. J. Quant. Chem. **90**, 910 (2002).

[10] E. Nir, K. Kleinermanns, and M. S. de Vries, Nature **408**, 949 (2000).

[11] F. Piuzzi, M. Mons, I. Dimicoli, B. Tardivel, and Q. Zhao, Chem. Phys. **270**, 205 (2001).

[12] E. Nir, C. Janzen, P. Imhof, K. Kleinermanns, and M. S. de Vries, Phys. Chem. Chem. Phys. **4**, 732 (2002).

[13] M. Mons, I. Dimicoli, F. Piuzzi, B. Tardivel, and M. Elhanine, J. Phys. Chem. **106**, 5088 (2002).

[14] M. Sabio, S. Topiol, and J. W. C. Lumma, J. Phys. Chem. **94**, 1366 (1990).

[15] M. K. Shukla, S. K. Mishra, A. Kumar, and P. C. Mishra, J. Comp. Chem. **21**, 826 (2000).

[16] C. van Caillie and R. D. Amos, Chem. Phys. Lett. **308**, 249 (1999).

[17] C. van Caillie and R. D. Amos, Chem. Phys. Lett. **317**, 159 (2000).

[18] F. Furche and R. Ahlrichs, J. Chem. Phys, **117**, 7433 (2002).

[19] A. L. Sobolewski and W. Domcke, Phys. Chem. Chem. Phys. **1**, 3065 (1999).

[20] I. Frank, J. Hutter, D. Marx, and M. Parrinello, J. Chem. Phys. **108**, 4060 (1998).

[21] M. Odelius, D. Laikov, and J. Hutter, preprint.

[22] N. L. Doltsinis and D. Marx, Phys. Rev. Lett. **88**, 166402 (2002).

[23] D. Marx and J. Hutter, in *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotendorst (NIC, Jülich, 2000), for downloads see `http://www.theochem.ruhr-uni-bochum.de/go/cprev.html`.

[24] N. L. Doltsinis and D. Marx, J. Theor. Comp. Chem. **1**, 319 (2002).

[25] `CPMD 3.4:` J. Hutter, P. Ballone, M. Bernasconi, P. Focher, E. Fois, S. Goedecker, D. Marx, M. Parrinello, and M. Tuckerman; MPI für Festkörperforschung, Stuttgart and IBM Zurich Research Laboratory.

[26] T. Ziegler, A. Rauk, and E. J. Baerends, Theor. Chim. Acta **43**, 261 (1977).

[27] M. Filatov and S. Shaik, Chem. Phys. Lett. **288**, 689 (1998).

[28] M. Filatov and S. Shaik, J. Chem. Phys. **110**, 116 (1999).

[29] M. Filatov and S. Shaik, Chem. Phys. Lett. **304**, 429 (1999).

[30] J. Gräfenstein, E. Kraka, and D. Cremer, Chem. Phys. Lett. **288**, 593 (1998).

[31] J. Gräfenstein and D. Cremer, Phys. Chem. Chem. Phys. **2**, 2091 (2000).

[32] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[33] R. W. Hockney, Comp. Phys. **9**, 136 (1970).

[34] R. N. Barnett and U. Landman, Phys. Rev. B **48**, 2081 (1993).

[35] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[36] C. Lee, W. Yang, and R. C. Parr, Phys. Rev. B **37**, 785 (1988).

[37] Gaussian 98, Revision A.9, M. J. Frisch et al., Gaussian, Inc., Pittsburgh PA, 1998.

[38] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[39] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[40] C. F. Guerra, F. M. Bickelhaupt, J. G. Snijders, and E. J. Baerends, J. Am. Chem. Soc. **122**, 4117 (2000).

[41] M. P. Fülscher, L. Serrano-Andrés, and B. O. Roos, J. Am. Chem. Soc. **119**, 6168 (1997).

[42] B. Mennucci, A. Toniolo, and J. Tomasi, J. Phys. Chem. A **105**, 7126 (2001).

[43] N. Ismail, L. Blancafort, M. Olivucci, B. Kohler, and M. A. Robb, J. Am. Chem. Soc. **124**, 6818 (2002).

[44] E. Nir, L. Grace, B. Brauer, and M. S. de Vries, J. Am. Chem. Soc. **121**, 4896 (1999).

[45] H. Langer and N. L. Doltsinis, to be published.

[46] L. B. Clark, J. Am. Chem. Soc. **99**, 3934 (1977).

Table 1

Dihedral angles [°] of the amino groups of the six guanine tautomers from Fig. 1 in the ground state. The plane-wave (p.w.)/pseudopotential calculations are compared to results obtained using Gaussian-type basis sets. Dh1 is the dihedral angle $H_{21}N_{2'}C_2N_1$ and dh2 stands for the dihedral angle $H_{22}N_{2'}C_2N_3$; see Fig. 1 for the labelling of the atoms.

method	BLYP		BLYP		B3LYP	
basis set	p.w.(70 Ry)		6-31++G**		6-31++G**	
	dh1	dh2	dh1	dh2	dh1	dh2
N7H keto	35.0	11.8	37.6	11.4	36.2	10.7
N9H keto	30.7	13.1	33.6	12.8	31.5	12.0
trans-N9H-*enol*	17.2	16.1	18.2	17.2	16.6	15.7
cis-N9H-*enol*	18.8	16.2	16.3	16.4	14.5	14.6
trans-N7H-*enol*	15.1	15.2	19.8	17.1	18.4	15.8
cis-N7H-*enol*	16.8	15.6	17.9	16.2	16.5	15.3

Table 2

Relative ground state energies [kJ/mol] of the six guanine tautomers from Fig. 1. The plane-wave (p.w.)/pseudopotential calculations are compared to results obtained using Gaussian-type basis sets. Shukla et al. [15] used the optimized structure at HF level for a single point calculation at MP2 level, both with a mixed basis set (m.b.) where 6-311+G* was used for the nitrogen atom of the amino group and 4-31G for all other atoms.

method	BLYP	BLYP	B3LYP	MP2 [5]	MP2 [6]	MP2 [15]
basis set	p.w.	6-31++G**	6-31++G**	6-31G**	6-311G**	m.b.
N7H-*keto*	0.0	0.0	0.0	0.0	+ 1	0.0
N9H-*keto*	+ 3.3	+ 2.9	+ 2.4	+ 0.2	0	+ 1.7
trans-N9H-*enol*	+12.8	+ 9.2	+ 5.1	+ 4.7	+ 4	+29.9
cis-N9H-*enol*	+14.4	+11.9	+ 7.9	+ 7.3	+ 7	
trans-N7H-*enol*	+23.0	+20.0	+17.5	+18.7	+18	+35.6
cis-N7H-*enol*	+52.2	+52.7	+51.8	+55.2		

Table 3. Lowest singlet vertical $\pi \rightarrow \pi^*$ excitation energies [eV] of six guanine tautomers obtained from TDDFT calculations using different exchange-correlation functionals. The ROKS/plane-wave (p.w.) calculations are compared to TDDFT results obtained using Gaussian-type basis sets as well as to different methods and the experimental excitation energy of 9-ethylguanine [46].

method	ROKS	TDDFT	TDDFT	CIS	CIPSI	CASPT2	Exp.
	BLYP	BLYP	B3LYP	[15]	[42]	[41]	[46]
basis set	p.w.	6-31++G**	6-31++G**	m.b.			
N9H-*keto*	3.70	4.39	4.87	6.32	4.76	4.76	4.55
trans-N9H-*enol*	3.52	4.19	4.72	6.51			
cis-N9H-*enol*	3.50	4.13	4.68				
trans-N7H-*enol*	3.48	3.80	4.36	6.15			
cis-N7H-*enol*	3.47	3.74	4.31				
N7H-*keto*	3.35	4.10	4.63	6.15	4.94		

Table 4

Relative energies [kJ/mol] in the S_1 excited state of the six most stable guanine tautomers obtained from ROKS calculations.

method	ROKS
functional	BLYP
basis set	p.w. (70 Ry)
N7H-*keto*	0.0
N9H-*keto*	+16.3
cis-N9H-*enol*	+20.7
trans-N9H-*enol*	+22.5
trans-N7H-*enol*	+35.0
cis-N7H-*enol*	+63.3

Table 5

Adiabatic excitation energies [eV] for the lowest $\pi \rightarrow \pi^*$ transition of six guanine tautomers obtained with ROKS BLYP/p.w. compared with experimental 0–0 transition energies

	ROKS	exp. [6]	exp. [13]
N9H-*keto*	3.11	4.13	4.20
trans-N7H-*enol*	3.10		4.07
cis-N7H-*enol*	3.09		
trans-N9H-*enol*	3.07	4.08	4.31
cis-N9H-*enol*	3.04		
N7H-*keto*	2.97	4.21	4.12

Table 6

Harmonic vibrational frequencies [cm^{-1}] of the different guanine tautomers calculated using the ROKS, CIS, and CASSCF methods. For the CIS frequencies, the corresponding intensities are given in parentheses [km/mol].

Table 6

	N9H-*enol*				N7H-*enol*		N9H-*keto*	N7H-*keto*
	trans			*cis*	*trans*	*cis*		
	CIS	CAS	ROKS	ROKS	ROKS	ROKS	ROKS	ROKS
1	64 (29)	91	43	119	86	120	65	123
2	102 (17)	152	132	136	167	160	120	153
3	215 (5)	178	176	181	201	200	190	166
4	249 (82)	194	223	228	233	250	217	250
5	267 (145)	238	256	238	287	300	267	288
6	275 (8)	289	282	268	333	327	317	306
7	311 (14)	325	327	329	362	336	332	325
8	359 (155)	361	356	348	410	349	411	333
9	365 (6)	400	434	416	448	382	414	410
10	492 (11)	461	465	474	461	405	441	431
11	521 (88)	498	486	515	508	475	496	459
12	554 (15)	537	508	522	511	515	505	469
13	625 (43)	590	512	534	520	547	559	502
14	634 (39)	650	535	555	556	559	600	529
15	653 (11)	659	538	567	593	611	619	551
16	697 (0)	670	593	617	616	618	637	584
17	722 (5)	693	626	640	635	631	701	638
18	766 (27)	728	656	667	695	698	736	646
19	805 (6)	867	748	753	724	731	755	674
20	878 (1)	1006	784	782	768	764	768	794
21	985 (27)	1009	792	801	787	789	846	813
22	1036 (135)	1027	929	919	851	835	891	852
23	1101 (128)	1105	948	940	968	969	999	941
24	1142 (37)	1156	996	1007	999	995	1020	993
25	1204 (26)	1198	1015	1016	1073	1072	1068	1077
26	1250 (6)	1260	1088	1090	1138	1136	1141	1103
27	1302 (251)	1354	1150	1177	1173	1163	1184	1171
28	1398 (30)	1397	1245	1236	1252	1214	1223	1229
29	1451 (104)	1443	1293	1283	1258	1263	1247	1267
30	1471 (48)	1477	1316	1323	1281	1278	1261	1294
31	1528 (77)	1520	1328	1327	1332	1305	1311	1335
32	1562 (199)	1580	1358	1373	1352	1326	1355	1381
33	1595 (140)	1606	1381	1419	1384	1397	1398	1429
34	1635 (307)	1688	1455	1446	1421	1425	1430	1479
35	1715 (84)	1721	1485	1468	1493	1494	1571	1516
36	1799 (12)	1775	1556	1560	1575	1567	1588	1593
37	1808 (904)	1875	1581	1579	1624	1579	1641	1619
38	3445 (3)	3406	3112	3086	3126	3108	3167	3191
39	3856 (217)	3821	3504	3503	3490	3472	3421	3455
40	3913 (137)	3905	3536	3537	3517	3487	3477	3466
41	4006 (88)	3974	3636	3631	3625	3622	3516	3512
42	4153 (105)	4137	3637	3635	3628	3678	3532	3574

Fig. 1. Langer/Doltsinis, J. Chem. Phys.

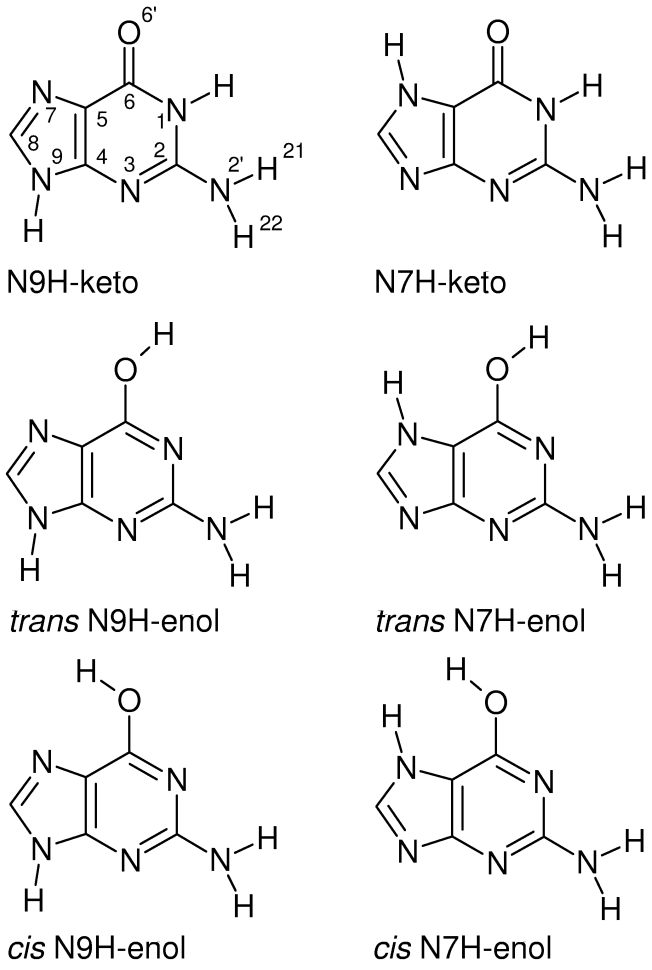

Fig. 2. Langer/Doltsinis, J. Chem. Phys.

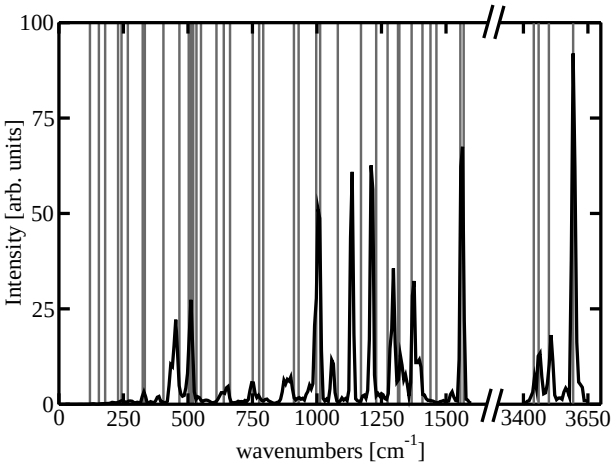

FIGURE CAPTIONS:

Fig. 1: Structures and nomenclature of the six lowest energy guanine tautomers. The atomic numbering scheme is illustrated for the N9H-*keto* tautomer.

Fig. 2: Comparison of the anharmonic (black line) and harmonic (grey lines) theoretical vibrational spectra calculated using the ROKS BLYP/p.w. method with a cutoff of 70 Ry. No intensities are available for the harmonic frequencies; the heights of the peaks of the anharmonic spectrum do not correspond to IR intensities.