PCCP

Accepted Manuscript

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Ultrafast excited-state dynamics of isocytosine[†]

Rafał Szabla,[*a] Robert W. Góra,[*b] and Jiří Šponer[a,c]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Alternative nucleobase isocytosine has for long been considered as a plausible component of hypothetical primordial informational polymers. To examine this hypothesis we investigated the excited-state dynamics of the two most abundant forms of isocytosine in the gas phase (*keto* and *enol*). Our surface-hopping nonadiabatic molecular dynamics simulations employing the algebraic diagrammatic construction to the second order [ADC(2)] method for the electronic structure calculations suggest that both tautomers undergo efficient radiationless deactivation to the electronic ground state with time constants which amount to τ_{keto}=182 fs and τ_{enol}=533 fs. The dominant photorelaxation pathways correspond to ring-puckering ($\pi\pi^*$ surface) and C=O stretching/N–H tilting ($n\pi^*$ surface) for the *enol* and *keto* forms respectively. Based on these findings, we infer that isocytosine is a relatively photostable compound in the gas phase and in these terms resembles the biologically relevant nucleobases. The estimated $S_1 \rightsquigarrow T_1$ intersystem crossing rate constant of $8.02 \cdot 10^{10}$ s^{-1} suggests that triplet states might also play an important role in the overall excited-state dynamics of the *keto* tautomer. The reliability of ADC(2)-based surface-hopping molecular dynamics simulations was tested against multireference quantum-chemical calculations and the potential limitations of the employed ADC(2) approach are briefly discussed.

1 Introduction

The origin of informational polymers on Earth have been a subject of intense studies over the past 50 years.[1,2] Initial difficulties in obtaining RNA nucleosides[1] and oligomers[1,3,4] from small organic substrates under prebiotically plausible conditions resulted eventually in the appearance of several competitive scenarios. Perhaps the most compelling hypotheses are related to the emergence of RNA from formamide[5,6] or the indirect but efficient synthesis of pyrimidine nucleosides from one- and two-carbon feedstock molecules.[7–9] Other ideas suggest that RNA might have evolved from a prebiotic ancestor,[10–13] sometimes referred to as XNA.[14] In this latter case, the possible participation of alternative nucleobases in molecular evolution attracted a lot of attention.[13,14] The most noticeable of the alternative nucleobases are hypoxantine and diaminopurines found in carbonaceous meteorites,[15] isocytosine which can be formed in the presence of iron-sulphur minerals,[16] and 2,4,6-triaminopyrimidine which readily undergoes glycosylation with ribose as opposed to many other nucleobases.[12]

Isocytosine (iC) was found as an important product of thermal condensation of formamide catalyzed by several types of meteoritic materials.[16–18] Interestingly, iC can form Watson–Crick base pairs with cytosine and isoguanine or a reversed Watson–Crick base pair with guanine.[16,19,20] Saladino and co-workers suggested that iC could serve as a bioisostere of guanine in primordial nucleic acids, since the latter one was obtained in a very few prebiotically plausible reactions and with rather low yields.[16,18] Final exclusion of iC from DNA and RNA was, thus, ascribed to molecular evolution which selected guanine as a substitute having only one stable tautomer in aqueous environment and being less prone to mispairings.[21,22] Even though this scenario is intriguing and plausible, it lacks strong evidence that could eventually confirm or deny its validity. Therefore, the knowledge of the possible prebiotic environmental factors that could affect the selection of nucleobases and their presumed prebiotic analogs is indispensable.

UV-irradiation was suggested as one of the most prominent selection factors that controlled early molecular evolution.[23–25] Owing to the implied higher activity of the young sun in the ultraviolet range[26] and the absence of the ozone layer in the Archean age,[27] only the most photostable molecules endured to

---

[a] Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 61265, Brno, Czech Republic; E-mail: rafal.szabla@gmail.com

[b] Department of Physical and Quantum Chemistry, Faculty of Chemistry, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland; E-mail: robert.gora@pwr.edu.pl

[c] CEITEC–Central European Institute of Technology, Masaryk University, Campus Bohunice, Kamenice 5, CZ-62500 Brno, Czech Republic

† Electronic Supplementary Information (ESI) available: [including geometrical parameters of selected structures, their relative energies, relevant spin-orbit coupling matrix elements, supplementary potential energy profiles, plots of cummulative populations and population fittings and results of preliminary MR-CISD calculations]. See DOI: 10.1039/b000000x/

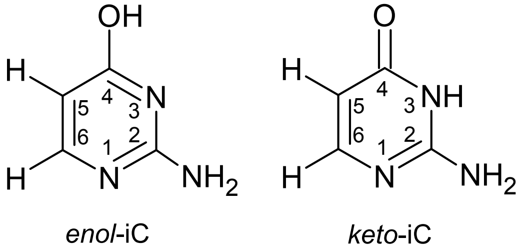


Fig. 1 Two dominant tautomers of iC in the gas phase.

contribute to the ensemble of biomolecular building blocks.[24] In this context, the photochemical reactivity of the postulated predecessors of nucleobases might improve our understanding why only five nucleobases (adenine, cytosine, guanine, uracil and thymine) were selected. It could also indicate which of the alternative nucleobases were photostable enough to constitute primordial informational polymers. For instance, the recent studies of 2-aminopurine, revealed that this analog of adenine has an increased excited-state lifetime by up to ~100 times upon microhydration, and is essentially much less photostable than its biologically relevant counterpart.[28,29]

Bakalska and Delchev investigated the photochemistry of iC by means of excited-state potential energy surface calculations and an UV-irradiation experiment in acetonitrile solution.[30] They suggested that the irradiated sample slowly undergoes phototautomerization reaction (within ~2h) driven by the dissociative $\pi\sigma^*$ state.[30] Here, we present the results of surface-hopping nonadiabatic molecular dynamics simulations of the two major tautomers of iC in the gas phase (see Fig. 1), employing the algebraic diagrammatic construction to the second order [ADC(2)][31] method for the electronic structure computations. These results enabled the observation of major photorelaxation pathways, which were not indicated by the previous static calculations, and provide theoretical estimates of the excited-state lifetimes of the title molecule. We also discuss the possible role of triplet manifold, and estimate the $S_1 \rightsquigarrow T_1$ intersystem crossing rate using the time-dependent approach based on the short-time approximation of correlation function proposed by Marian and co-workers.[32] We also compare our findings to the corresponding studies describing the excited-state dynamics of cytosine.

2 Computational Methods

The ground-state equilibrium geometries and vibrational frequencies of the *keto* and *enol* forms of iC were obtained using the Kohn–Sham density functional theory with the def2-TZVP basis set and the B3LYP hybrid functional. These geometries were further used for the computations of vertical excitation energies. Vertical excitation energy calculations, UV-vis spectra simulations, geometry optimizations on the S_1 hypersurface, computations of harmonic vibrational frequencies in the S_1 minima, and nonadiabatic molecular dynamics simulations were all performed with the algebraic diagrammatic construction method to the second-order [ADC(2)][31,33] and the aug-cc-pVDZ correlation consistent basis set.

The semi-classical nonadiabatic molecular dynamics simulations of both iC tautomers were performed using the Tully's fewest-switches surface hopping algorithm with decoherence correction of Granucci and Persico (with the decoherence parameter of 0.1 Hartree).[34] Initially, the UV-vis absorption spectra were simulated by employing the nuclear ensemble method.[35] For this purpose, 500 points per tautomer were sampled from a Wigner distribution for all vibrational normal modes of the ground-state geometries. The six low-lying excited electronic states were considered in the spectra simulations and in subsequent nonadiabatic molecular dynamics simulations. The initial conditions for the dynamics were selected from the 5.50±0.2 spectral domain (see Fig. 3). One hundred trajectories were started for each tautomer and were propagated for up to 1 ps. In the case of the *keto* tautomer, one trajectory was started from the S_2 state, 21 in the S_3, 23 in the S_4, 33 in the S_5, and 22 in the S_6. For the *enol* form, two trajectories were started in the S_2 state, 18 in the S_3, 46 in the S_4, 29 in the S_5, and 5 in the S_6. The time step for propagation of the classical equations for nuclear motion was 0.5 fs, while 0.025 fs time step was applied to the semi-classical approximation of the electronic time-dependent Schrödinger equation. The nonadiabatic transitions were enabled only between the excited states and the trajectories were terminated when the energy gap between the ground state and the first excited state dropped below 0.15 eV (since the ADC(2) method becomes unreliable in the vicinity of conical intersections with the ground state). Such an approach enables to identify the plausible photorelaxation channels and to estimate the excited state lifetime of the studied molecule, even though the lack of nonadiabatic couplings and surface hoppings between the S_1 and S_0 states imposes some limitations. Similar computational protocol was applied in several previous studies,[36–38] and compared well with both the experimental results and the data obtained from nonadiabatic dynamics simulations performed at the MRCIS level.[36,37]

To test the accuracy of the ADC(2) method in the vicinity of the S_1/S_0 intersection seam, we performed optimizations of all the relevant conical intersections using the ADC(2) and MP2 methods (for the S_1 and S_0 states, respectively) with aug-cc-pVDZ basis set, and compared the optimized minimum-energy crossing points (MECPs) to the corresponding conical intersections optimized at the multireference configuration interaction with single and double excitations (MR-CISD/cc-pVDZ) level.[39] The complete active space self consistent field (CASSCF) reference wave function was used in the MR-CISD calculations with six electrons distributed among five orbitals and with density averaging over the two lowest states (SA-2-CASSCF(6,5)/cc-pVDZ). The 8 core (1s) orbitals were kept frozen in the MR-CISD calculations. The MR-CISD MECPs optimizations were performed using the Columbus 7.0 package while for the purpose of ADC(2)/MP2 calculations we implemented the method of Levine, Coe and Martinez,[40] which allows the optimization of MECPs without evaluating the nonadiabatic couplings. This optimization scheme in conjunction with the ADC(2)/MP2 was recently used as a primary approach for locating minimum-energy crossing points in several instances.[41–45] We interfaced our implementation with the Turbomole 7.0.2 package,[46] and used the internal Broyden–Fletcher–Goldfarb–Shanno quasi-Newton scheme to perform the optimizations.

The intersystem crossing (ISC) rate was estimated using time-dependent approach based on the short-time approximation of correlation function proposed by Etinsky *et al.* which proved to provide good estimates of intersystem crossing rates for flavonoids, aromatic ketones and nucleobases.[32,47] The required vibrational normal modes and vibrational frequencies for the S_1 and T_1 states of *keto*-iC were obtained at the ADC(2)/aug-cc-pVDZ level. To take into account the mixing of vibrational normal modes upon electronic transition, we employed the Duschinsky transformation, which expresses the final-state normal modes as a linear combination of the initial-state normal modes. The spin-orbit coupling (SOC) between the S_1 and T_1 states was computed at the CASPT2/SA-CASSCF(10,8)/cc-pVDZ-DK level, while the scalar relativistic effects were accounted for by means of the 2nd order Douglas–Kroll–Hess Hamiltonian.

Computations of potential energy (PE) surface cuts were performed at the ADC(2) and RI-NEVPT2/SA-CASSCF(8,7) levels,[48–50] using aug-cc-pVDZ and cc-pVTZ basis sets, respectively. In the latter case the reference energies were obtained employing the CASSCF wavefunction averaged over four lowest-lying electronic states for *keto*-iC and 5 lowest-lying electronic states for *enol*-iC. The active space was composed of 8 electrons distributed among 7 orbitals (3π, one n_N, and $3\pi^*$). In the SA-CASSCF calculations the Split-RI-J (resolution of identity)[51] and the COSX (chain of spheres exchange) approximations were used.[52]

The choices of different active spaces were based on the rules proposed by Veryazov, Malmqvist and Roos.[53] In particular, we assigned the largest possible active spaces to keep the natural orbital occupations of active orbitals in the range 0.02-1.98. 10 electrons correlated in 8 orbitals were needed in the S_1 minimum of *keto*-iC and such active space was used in calculations of the SOC matrix elements at the CASPT2/SA-CASSCF level. Having regard to these rules, somewhat smaller active spaces were necessary in the vicinity of the MECPs. Therefore, the reference space in our MR-CISD optimizations was obtained by distributing 6 electrons among 5 orbitals. The active spaces at varied MECPs consisted of slightly different sets of orbitals, due to the changing character of the excited electronic states crossing with the ground state. These sets of orbitals contained predominantly occupied π and n_N orbitals, and virtual π^* orbitals along with the virtual σ^*_{CO} orbital showing non-negligible population at the $n\pi^*/S_0$ conical intersection. Since the natural orbital occupation numbers varied along the reaction paths corresponding to different deactivation channels, we applied the NEVPT2/SA-CASSCF(8,7) approach for the calculation of PE profiles in order to keep the active spaces consistent and to avoid unwanted orbital rotations along the paths (which led to discontinuous PE profiles).

All the ADC(2) and MP2 calculations were performed with the Turbomole 7.0.2 program.[46] The Newton-X 1.4 package[54] was utilized for calculation of the UV-vis spectra and nonadiabatic molecular dynamics simulations. The optimizations of conical intersections at the MR-CISD level were done using the Columbus 7.0 program.[55] The SOC matrix elements were computed with the Molcas 8.0 package.[56] The PE surface cuts at the RI-NEVPT2 level were performed using Orca 3.0.3 package.[57] Our in-house codes used for the optimization of MECPs at the ADC(2)/MP2 level and the evaluation of ISC rates are available upon request.

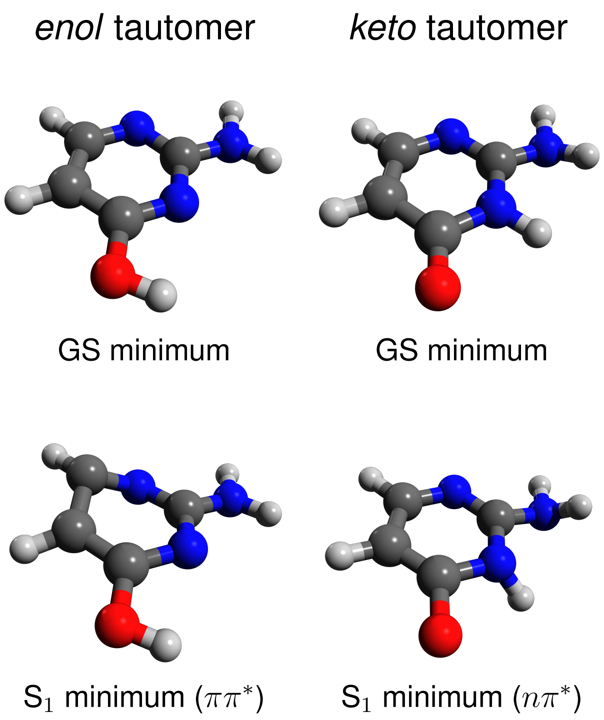


Fig. 2 Optimized geometries of the ground and lowest-lying excited states of the considered tautomers of iC.

3 Results and Discussion

S_0 and S_1 equilibrium geometries

Isocytosine exists in several relatively stable tautomeric forms both in the gas phase and in aqueous environment.[58–60] The *enol* and *keto* tautomers considered in this study were found as the main forms of iC existing in the gas phase, with the dominant proportion of the *enol* form.[58,59] The *keto* tautomer was also reported as one of the two main forms of iC present in aqueous environment.[58] The remaining major form of iC found in bulk water differs from the *keto* form investigated herein by the position of N(3)–H proton, which is located at the N(1) atom of the iC ring instead. In this work, however, we focused on the two dominant tautomers of iC observed in gas phase and optimized the ground-state equilibrium geometries at the B3LYP/def2-TZVP level.

The minimum-energy geometries of *enol*-iC and *keto*-iC in the S_1 electronic states obtained at the ADC(2) level, reveal several interesting discrepancies from the respective ground-state geometries. The optimized S_1 minimum of *enol*-iC is characterized by sizeable pyramidalization of the C(6) carbon atom entailing puckering of the pyrimidine ring. Moreover, the C(5)–C(6) and C(6)–N(1) distances are elongated with respect to the ground-state equilibrium geometry by 0.08 and 0.10 Å. The most characteristic structural features of the S_1 minimum of *keto*-iC are the out-of-plane tilting of the N(3)–H bond and significant elongation of the C=O carbonyl bond (from 1.22 Å in the S_0 minimum to 1.41 Å in the S_1 minimum), while the aromatic ring remains nearly planar. We also observe elongation of the C(5)–C(6) and C(2)–N(3) bonds by 0.06 Å, and shortening of the C(4)–C(5) distance by 0.09 Å. Interestingly, the S_1 minima of the two tautomers of iC have different characters, namely $\pi\pi^*$ in the case of *enol*-iC and $n\pi^*$ in the case of *keto*-iC. Furthermore, the S_1-S_0 energy gap in the S_1 minima amounts to 2.45 eV and 1.80 eV for the *enol* and *keto* tautomers, respectively. These observations indicate that rather significant differences should be expected with regard to the accessible photodeactivation pathways and excited-state lifetimes of the two tautomeric forms of iC.

Table 1 Vertical excitation energies (in eV) of the two considered tautomers of iC, computed using ADC(2)/aug-cc-pVDZ method assuming the ground-state minimum energy structures optimized at the B3LYP/def2-TZVP level.

State / Transition		E_{exc}/[eV]	f_{osc}	λ/[nm]
enol tautomer of iC				
S_1	$\pi\pi^*$	4.81	$9.18 \cdot 10^{-2}$	257.8
S_2	$n_N\pi^*$	5.04	$3.67 \cdot 10^{-3}$	246.0
S_3	$\pi\sigma^*_{NH}$	5.50	$1.57 \cdot 10^{-2}$	225.4
S_4	$n\pi^*$	5.78	$5.60 \cdot 10^{-3}$	214.5
S_5	$\pi\pi^*$	5.81	0.278	213.4
keto tautomer of iC				
S_1	$\pi\pi^*$	4.60	0.196	269.5
S_2	$n_O\pi^*$	4.85	$1.94 \cdot 10^{-4}$	255.6
S_3	$\pi\sigma^*_{NH}$	5.22	$8.15 \cdot 10^{-3}$	237.5
S_4	$n\pi^*$	5.62	$2.83 \cdot 10^{-4}$	220.6
S_5	$n\sigma^*$	5.89	$5.14 \cdot 10^{-2}$	210.5

Vertical excitation energies and UV-vis spectra

In contrast to the different characters of the S_1 minima of *enol*-iC and *keto*-iC, the vertical excitation energies presented in Table 1 reveal some qualitative similarities. In particular, the ordering of the first four excited states of *enol*-iC and *keto*-iC is the same, and in both cases the S_1 states are optically bright $\pi\pi^*$ transitions. The S_2 states are of $n\pi^*$ character and are separated from the lowest-lying $\pi\pi^*$ excitations by merely ~0.15 eV. This indicates the potential involvement of the $n\pi^*$ states in the photodeactivation mechanisms of iC. It is interesting to note, that the n orbital of *enol*-iC is associated with the nitrogen heteroatoms within the pyrimidine ring (n_N), whereas the corresponding n orbital of *keto*-iC belongs to the carbonyl group (n_O). The position of the dark $\pi\sigma^*$ states (S_3) indicates that the N–H or O–H bond fission photorelaxation channels should be accessible upon higher excitation energies. It is also worth noting, that the vertical excitations of the *enol* tautomer are slightly blue-shifted with respect to those of the *keto* form.

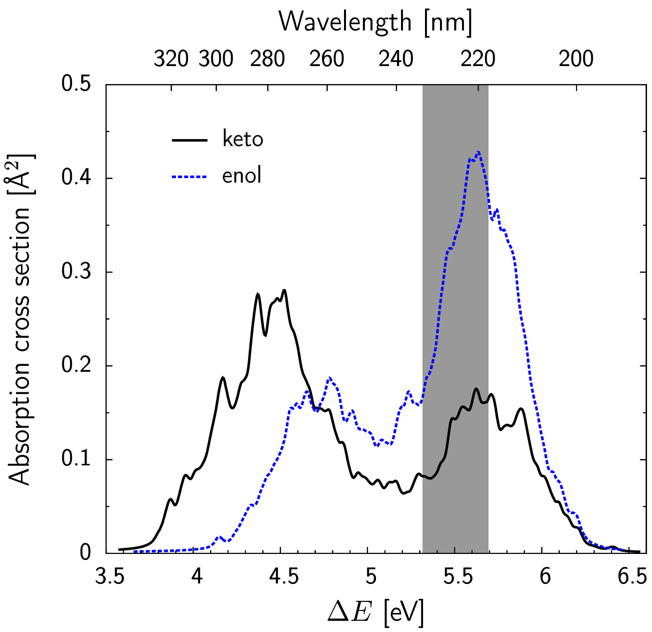


Fig. 3 UV absorption spectrum of the two considered tautomers of iC simulated at the ADC(2)/aug-cc-pVDZ level. The initial conditions were sampled from the shaded area of the spectrum.

The simulated UV-vis absorption spectra presented in Fig. 3 reflect the general features of electronic excitations manifold described above. First of all, two absorption maxima can be distinguished in each case, and the spectrum of the *enol* form is slightly blue-shifted when compared to that of *keto*-iC. The intensities and positions of the first absorption maxima can be compared to the energies and oscillator strengths of the corresponding S_1 excitations listed in Table 1. The second absorption maximum of *enol*-iC centered around 220 nm has also much higher intensity than the respective maximum of the *keto* form. This feature can be assigned to the presence of another optically bright $\pi\pi^*$ state (denoted as S_5) in *enol*-iC. The lack of a comparable electronic excitation in *keto*-iC results in much lower absorption of this tautomer in this spectral region. The two absorption maxima of *keto*-iC were also observed experimentally in acetonitrile solution (at 287 and 222 nm respectively).[30] The simulated spectrum presented in Fig. 3 is in good agreement with these measurements with rather insignificant discrepancies caused by the fact that we performed our simulations in the gas phase. To follow the excited-state dynamics of iC at shorter excitation wavelengths we sampled the initial conditions in the 5.50±0.2 eV spectral domain, around the absorption maxima of both tautomers (as marked in Fig. 3). This allows us to investigate the photodynamics of iC after irradiation with high-energy ultraviolet light which due to lack of ozone layer played and an important role in the Archean age.[26,27]

Nonadiabatic molecular dynamics simulations

As expected from the structural discrepancies between the S_1 minima, the *enol* and *keto* tautomers exhibit substantial differences in the excited-state dynamics. The evolution of the ground and excited state populations presented in Fig. 4 indicate much longer excited-state lifetime of *enol*-iC. The population decays were fitted between 0 and 1000 fs with the following sigmoid function (see also Fig. 4):

$$f(t) = f_0 + (1 - f_0)\exp\left[-\left(\frac{t}{\tau}\right)^2\right]. \tag{1}$$

According to the resulting time constants $\tau_{keto} \approx 185$ fs and $\tau_{enol} \approx 533$ fs, we infer an ultrafast excited-state photodeactivation for the *keto* tautomer, and clearly longer but still very short excited state lifetime of the *enol* tautomer. The electronic ground state of *keto*-iC is populated monoexponentially, and can be also fitted with the one-parameter exponential function:

$$g(t) = 1 - \exp\left[-\frac{t}{\tau_{exp}}\right]. \tag{2}$$

The resulting τ_{exp} amounts to 182 fs and is consistent with the value obtained by fitting of the sigmoid function. The time evolution of the ground-state population of *enol*-iC is more complex and the data acquired within 1 ps simulations are insufficient to perform a reliable fit with a multiexponential function.

Relatively effective population of the electronic ground state of *enol*-iC begins after 100 fs of the dynamics, and only 3 trajectories are terminated earlier. Interestingly, after the initial 60 fs the population of the S_1 state exceeds 80% and remains nearly constant for subsequent 180 fs. This implies that most trajectories are trapped around the S_1 minimum during this period of time (see Fig. 4). In contrast, nearly one third (32%) of trajectories obtained for *keto*-iC reaches the ground state within the first 100 fs after the photoexcitation. The fractional population of the S_1 state of *keto*-iC never exceeds 0.5, what suggests the presence of easily accessible S_1/S_0 state crossings in our simulations of this tautomer. The rather short relaxation time constant estimated for *keto*-iC indicates that it might undergo very efficient radiationless photodeactivation within the singlet manifold, just as the biologically relevant *keto* tautomer of cytosine. For instance, Barbatti and co-workers postulated a biexponential decay with the average lifetime of 580 fs based on nonadiabatic dynamics simulations of *keto*-cytosine involving singlet states only.[61,62] Biexponential decay was also proposed by Mai *et al.*, who included ISCs to the triplet manifold in their simulations.[63] Nonadiabatic dynamics simulations of cytosine also indicated a longer excited-state lifetime of the *enol* tautomer, similarly as it is observed for *enol*-iC.[63] It is worth noting here, that our benchmark calculations described in the following sections suggest that the ADC(2)/MP2 methods might not provide a reliable description of one of the photorelaxation channels operating in *keto*-iC. Therefore, the excited-state properties of *keto*-iC described in this paragraph should be treated with caution.

Contributions of the most important photorelaxation channels along with the corresponding MECP geometries of iC are presented in Fig. 5. We identified in total six deactivation pathways of *enol*-iC and three for *keto*-iC. In the case of *keto*-iC, the fractions presented below the optimized MECP geometries in Fig. 5 can be regarded as approximate quantum yields. The fractions shown for *enol*-iC do not include 19% of trajectories that did not relax to the electronic ground state within the simulation time, and 4% of trajectories that correspond to minor photorelaxation channels.

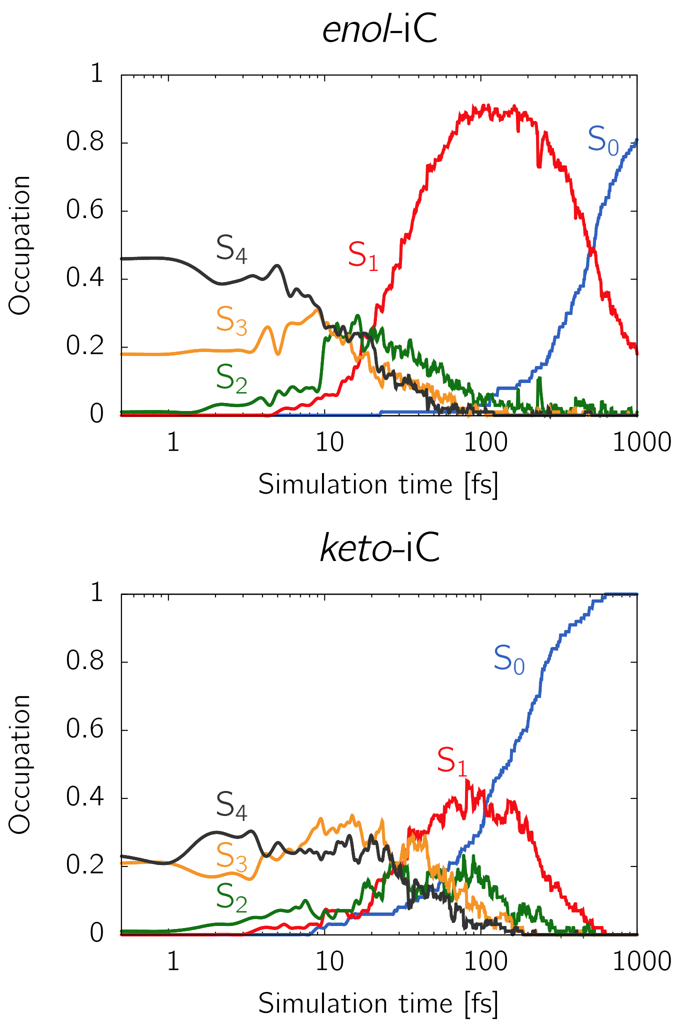
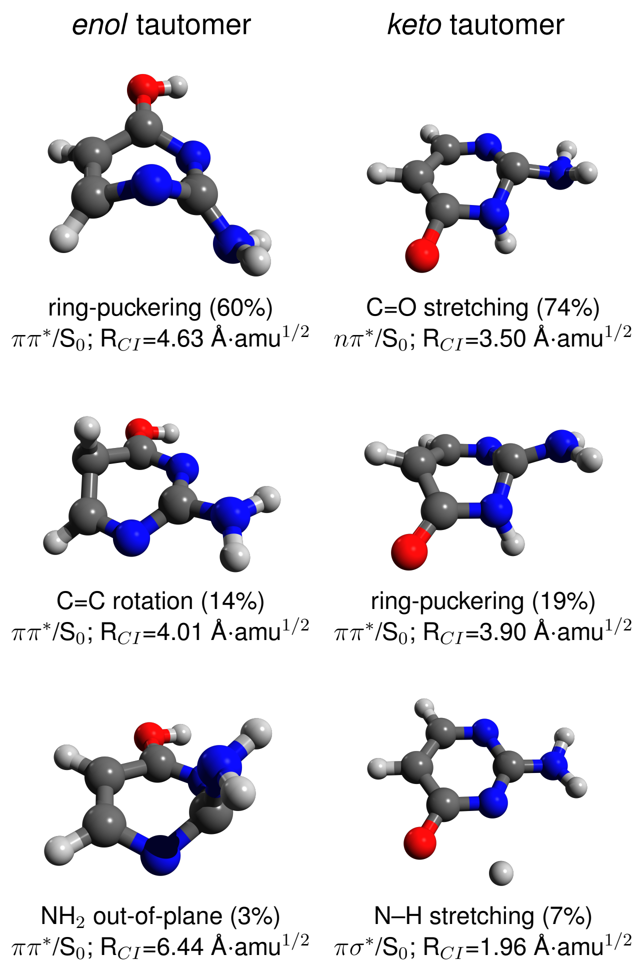


Fig. 4 Time-evolution of the ground and excited state populations of the *enol*- and *keto*- tautomers of iC. Only four excited electronic states are shown for clarity.

All of the major photodeactivation pathways observed for *enol*-iC occurred on the $\pi\pi^*$ hypersurface. The dominant mechanism is related to puckering of the pyrimidine ring at the C(4) and N(1) atoms and was followed by 60% of the trajectories. 14% of the observed S_1 to S_0 photorelaxation events were driven by twisting motion about the C(5)=C(6) bond, whereas all the remaining pathways played a minor role in the photodynamics of *enol*-iC. Among them are the deformation of the NH_2 group out of the ring plane (3%), and $\pi\sigma^*$-driven O–H, N–H or C–H bond fission mechanisms (2%, 1% and 1%, respectively).

The dominant photorelaxation mechanism of *keto*-iC occurs on the $n\pi^*$-state surface and is mainly related to C=O bond stretching, N(3)–H bond tilting, and slight pyramidalization of the C(2) carbon atom which is particularly evident in the case of the MECP



enol tautomer

keto tautomer

ring-puckering (60%)
$\pi\pi^*/S_0$; R_{CI}=4.63 Å·amu$^{1/2}$

C=O stretching (74%)
$n\pi^*/S_0$; R_{CI}=3.50 Å·amu$^{1/2}$

C=C rotation (14%)
$\pi\pi^*/S_0$; R_{CI}=4.01 Å·amu$^{1/2}$

ring-puckering (19%)
$\pi\pi^*/S_0$; R_{CI}=3.90 Å·amu$^{1/2}$

NH_2 out-of-plane (3%)
$\pi\pi^*/S_0$; R_{CI}=6.44 Å·amu$^{1/2}$

N–H stretching (7%)
$\pi\sigma^*/S_0$; R_{CI}=1.96 Å·amu$^{1/2}$

Fig. 5 Dominant photodeactivation channels of the considered tautomers of iC. The structures presented above are the minimum-energy crossing points (MECPs) optimized using the ADC(2)/MP2 methods. Estimated quantum yields, characters of the S_1 state at the MECPs and mass-weighted distances (R_{CI}) from the Franck-Condon region to the MECPs are given below each structure.

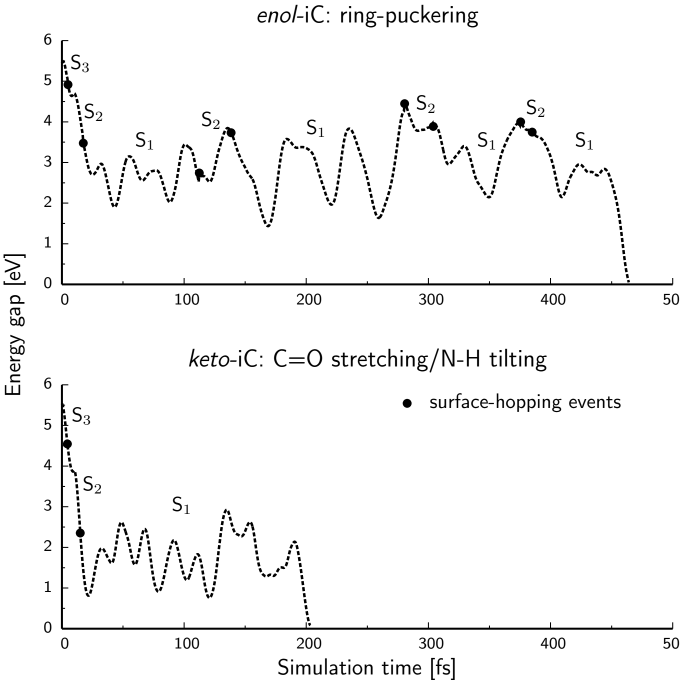


Fig. 6 Representative trajectories for *enol*-iC and *keto*-iC. The graphs show evolution of the energy gap between the populated excited state and the ground state. Both the trajectories were initiated in the S_3 state. Black dots indicate hopping events, while the 'current' state is indicated above each segment of the trajectories.

optimized at the ADC(2)/MP2 level (cf. Fig. 5). This reaction pathway was followed by almost three-quarters of all the trajectories. 19% of trajectories deactivated via an alternative relaxation channel driven by ring-puckering at the C(2) and C(5) atoms which leads to a $\pi\pi^*/S_0$ conical intersection. The remaining fraction of trajectories (7%) followed the N(3)–H bond fission pathway driven by the repulsive $\pi\sigma^*_{\mathrm{NH}}$ state. Thus, it becomes evident that $\pi\sigma^*_{\mathrm{NH}}$ play a rather minor role in radiationless deactivation of the investigated tautomers of iC.

The mass-weighted distances between the Franck-Condon region of *keto*-iC and *enol*-iC and the optimized MECPs (Fig. 5) provide preliminary explanation for the observed difference in excited-state lifetimes. The MECPs corresponding to major photorelaxation channels of *enol*-iC require greater rearrangement of the nuclei than the MECPs found for *keto*-iC. The most accessed conical intersection of *keto*-iC lies over 1.0 Å·amu$^{1/2}$ closer to the FC region than the dominant ring-puckered conical intersection of *enol*-iC. Even though the N–H and O–H bond fission reaction pathways proceed through the nearest conical intersections (with respect to the FC region in terms of mass-weighted internal coordinates), these mechanisms were scarcely triggered. This phenomenon is caused by the fact that the repulsive $\pi\sigma^*$ states responsible for these processes are accessible within the first 50 fs after photoexcitation, *i.e.* before the initial vibrational relaxation to the S_1 minima (being of $\pi\pi^*$ and $n\pi^*$ character in *enol*-iC *keto*-iC). When the vicinity of the respective S_1 minimum is reached the $\pi\sigma^*$ states are significantly destabilized. In consequence, N–H and O–H bond rupture from the S_1 minima is connected with energetic barriers large enough for other photorelaxation mechanisms to prevail.

More complete understanding of the photodynamics of iC emerges after the analysis of representative trajectories shown in Fig. 6. The two trajectories correspond to the dominant radiationless deactivation pathways of *enol*-iC and *keto*-iC, and were initiated in the S_3 state. In both cases, the molecules need less than 20 fs to reach the lowest lying excited singlet state, and soon afterwards they start oscillating around the respective S_1 minima. Even though the S_2 state can be temporarily populated at later stages of the dynamics (as shown for *enol*-iC), the S_1/S_0 state crossings are predominantly accessed from the vicinity of the S_1 minima. In the case of *enol*-iC, the S_1-S_0 energy gap does not drop below 1.0 eV until the intersection seam with the ground state is reached. For most of the simulation time (∼400 fs), *enol*-iC circulates around the S_1 minimum and the energy gap between the

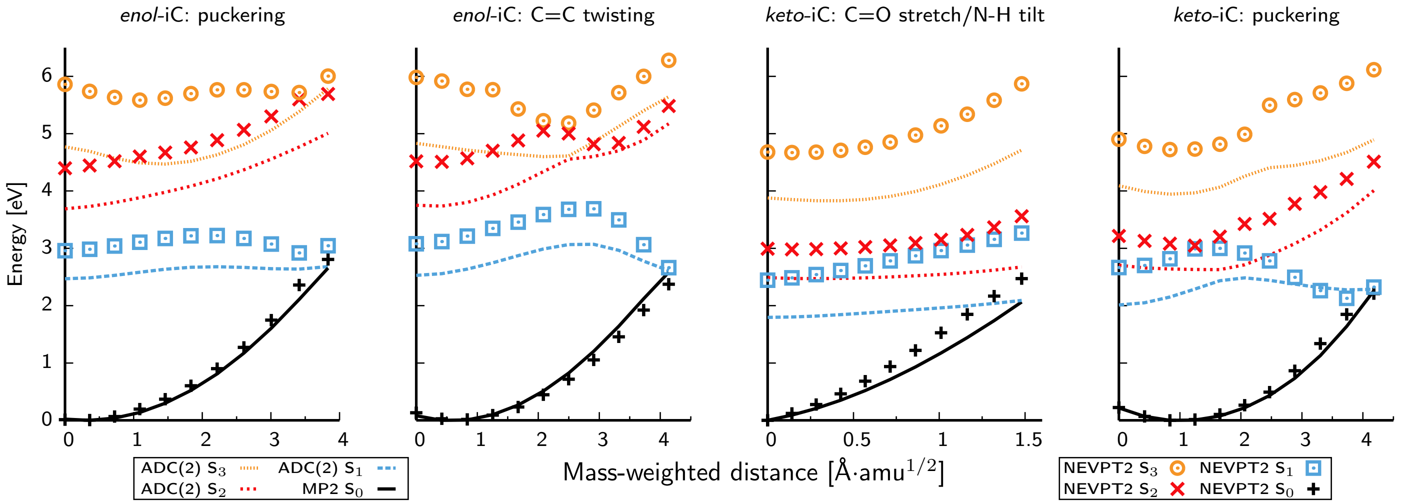


Fig. 7 Linear interpolation in internal coordinates between the S_1 minima of *enol*-iC and *keto*-iC and the most important MECPs optimized using the ADC(2)/MP2 methods. The lines present PE curves for the ADC(2)/aug-cc-pVDZ method, while the points show the reference values obtained at the NEVPT2/SA-CASSCF(8,7)/cc-pVTZ level.

populated state and the ground state varies between 1.5 eV and 3.5 eV. As shown for *keto*-iC, the S_1-S_0 energy gap drops several times below 1.0 eV within the 200 fs long trajectory, what indicates that the corresponding intersection seam can be accessed much more easily in our simulations of this tautomer. This observation is consistent with the lower S_1-S_0 energy difference found in the S_1 minimum of *keto*-iC in comparison to its *enol* counterpart.

Benchmark calculations of PE profiles for the dominant photorelaxation pathways

To assess the reliability of our dynamics simulations in describing the routes leading from the S_1 minima of *enol*-iC and *keto*-iC to the most significant conical intersections, we computed PE profiles using the ADC(2) and NEVPT2/SA-CASSCF(8,7) methods (see Fig. 7). The PE profiles were constructed by linear interpolation in internal coordinates between the stationary points. The corresponding S_1 minima were optimized at the ADC(2) level, while the MECPs were obtained using the ADC(2) and MP2 methods for the S_1 and S_0 states respectively.

The reaction paths computed for *enol*-iC reveal very good qualitative consistency of the ADC(2) and NEVPT2 results. In particular, the energies of the S_1 and S_0 states derived from NEVPT2 computations show near degeneracy at the MECPs, and suggest rather good performance of ADC(2) and MP2 methods in the vicinity of the S_1/S_0 crossing seam. The ADC(2) and NEVPT2 computations consistently predict a barrier of about 0.5 eV, which separates the S_1 minimum from the C=C-twisted conical intersection. The energetic barrier on the path leading to the ring-puckered MECP is rather insignificant. This latter finding explains the large fraction of trajectories that followed the ring puckering relaxation channel. NEVPT2 PE profiles for higher-lying excited states are shifted to higher energies, but in general correlate well with the profiles obtained using the ADC(2) method. Small discrepancies in the higher states visible in proximity of the ring-puckered MECP of *enol*-iC should not affect the final results of our nonadiabatic molecular dynamics simulations.

The NEVPT2-derived PE profiles for *keto*-iC are also qualitatively consistent with the corresponding ADC(2) results. The largest disagreement can be found along the C=O stretching/N–H tilting reaction pathway, for which the degeneracy of the S_1 and S_0 states is not reached at the $n\pi^*$/S_0 MECP located at the ADC(2)/MP2 level. The corresponding S_1-S_0 energy gap obtained with the NEVPT2 method amounts to ~0.8 eV, albeit it is clear that the energy gap is significantly reduced along this reaction coordinate and the S_1/S_0 intersection seam is nearly reached. Furthermore, the S_1 PE profile for this relaxation channel is more ascending in the case of NEVPT2 calculations. This indicates that the description of this photorelaxation channel might be less accurate at the ADC(2) level and further tests are necessary. We provide additional information about the $n\pi^*$/S_0 intersection seam of *keto*-iC in the next section and compare these results to the corresponding conical intersection optimized at the MR-CISD level. The ADC(2) and NEVPT2 calculations performed for the ring puckering photorelaxation channel of *keto*-iC are in a much better agreement. Both methods predict a moderate energetic barrier separating the S_1 minimum and the ring-puckered conical intersection, which is a consequence of the change in the S_1 character (from $n\pi^*$ to $\pi\pi^*$). Similarly as in the case of *enol*-iC, the NEVPT2 energies of S_1 and S_0 states exhibit near degeneracy at the ring-puckered MECP of *keto*-iC.

Comparative optimizations of conical intersections at the ADC(2)/MP2 and MR-CISD levels

The MECPs optimized at the ADC(2)/MP2 level and conical intersections obtained using the MR-CISD/SA-CASSCF(6,5) method are summarized in Fig. 8. In agreement with the PE profiles discussed above, the ADC(2)/MP2 approach provided reliable geometries of all the $\pi\pi^*$/S_0 MECPs. Only very subtle differences can be spotted between the MECPs optimized with the

ADC(2)/MP2 and MR-CISD methods. The corresponding RMSD values do not exceed 0.07 Å (see Fig. 8) and confirm the very good performance of the method applied for nonadiabatic dynamics simulations in the vicinity of $\pi\pi^*/S_0$ intersection seam.

However, the general features of the $n\pi^*/S_0$ conical intersection of *keto*-iC are only partly retained and some discrepancies between the ADC(2)/MP2 and MR-CISD are apparent. First of all, the geometry optimized at the MR-CISD level exhibits moderate pyramidalization of the C(6) atom and no pyramidalization at the C(2) atom. The opposite was found for the respective MECP optimized at the ADC(2)/MP2 level. Other features present in the geometry optimized at the MR-CISD level that are not found for the MECP obtained with the ADC(2)/MP2 method are: tilting of the C(6)-H hydrogen atom and lower value of the N(3)-C(4)-O valence angle (94.3°). This is reflected by much larger RMSD between the MECPs optimized with these two approaches, which amounts to 1.05 Å. In Fig. 9 we present a linearly interpolated path between the S_1 minimum to the $n\pi^*/S_0$ conical intersection optimized at the MR-CISD/SA-CASSCF(6,5) level. The corresponding single-point energies were computed using the ADC(2) and NEVPT2/SA-CASSCF(8,7) methods (similarly as was done for the MECPs obtained with the ADC(2)/MP2 approach in Fig. 7). The shapes of the PE surface cuts shown in Fig. 9, confirm our assumptions from the previous section regarding the inaccuracy of the ADC(2)/MP2 methods in the vicinity of $n\pi^*/S_0$ crossing seam in *keto*-iC. In particular, this crossing seam is reached much earlier when the ADC(2)/MP2 methods are used. Furthermore, both the NEVPT2 and MR-CISD calculations indicate a considerably more sloped topology of the $n\pi^*/S_0$ conical intersection. The corresponding energy difference between the S_1 minimum and this MECP amounts to $\sim$1.5 eV and is sizeably larger than the one visible in the third panel of Fig. 7.

According to Fig. 7, the $n\pi^*/S_0$ MECP is the nearest state crossing to the S_1 minimum of *keto*-iC. The mass-weighted distance between these stationary points optimized using the ADC(2)/MP2 and ADC(2) methods, respectively, amounts to 1.48 Å·amu$^{1/2}$. In contrast, the mass-weighted distance between these stationary points optimized at the MR-CISD level amounts to 3.86 Å·amu$^{1/2}$. Thus, even though the $n\pi^*/S_0$ MECP is still closer to the S_1 minimum than any of the remaining MECPs considered in this section, the mass-weighted distance obtained from MR-CISD optimizations is noticeably larger. It is interesting to note that clear indications of discrepancies between the ADC(2)/MP2 and MR-CISD geometries could be observed already at the MR-CISD/SA-2-CASSCF(2,2) level (cf. ESI of this article).

Having regard to the foregoing, it is instructive to discuss the results of the D_1 diagnostic, which is often used as a measure indicating the multireference character of the electronic ground state.[64] Jansen and Nielsen suggested that the values of D_1 diagnostic exceeding 0.04 might indicate that the MP2 method becomes inadequate.[64] Later, Plasser *et al.*[36] observed that this threshold is often surpassed during nonadiabatic molecular dynamics simulations of adenine with the ADC(2) method (MP2 was used for the ground state), but the description of lowest-lying electronic states remains reliable. Recently, Tuna *et al.*[41] suggested establishing a new trust region for the D_1 diagnostic for medium-sized organic molecules, also to assess the applicability of the approximate MECP optimization scheme using the ADC(2)/MP2 approach. Here, the D_1 diagnostic for the MP2 method amounts to about 0.06 in the case of the ring-puckered and C=C-twisted (both $\pi\pi^*/S_0$) MECPs of *enol*-iC and 0.07 for the ring-puckered $\pi\pi^*/S_0$ MECP of *keto*-iC, all optimized at the ADC(2)/MP2 level. On the other hand, in the case of the problematic $n\pi^*/S_0$ MECP of *keto*-iC optimized with the ADC(2)/MP2 method, the D_1 diagnostic amounts to 0.09. Therefore, the postulated "safety threshold" of the D_1 diagnostic for the approximate MECP optimization with the ADC(2)/MP2 method might lie somewhere inbetween 0.07 and 0.09. However, these values should be treated with caution, since it is unclear how transferable they are to systems of different composition and size.

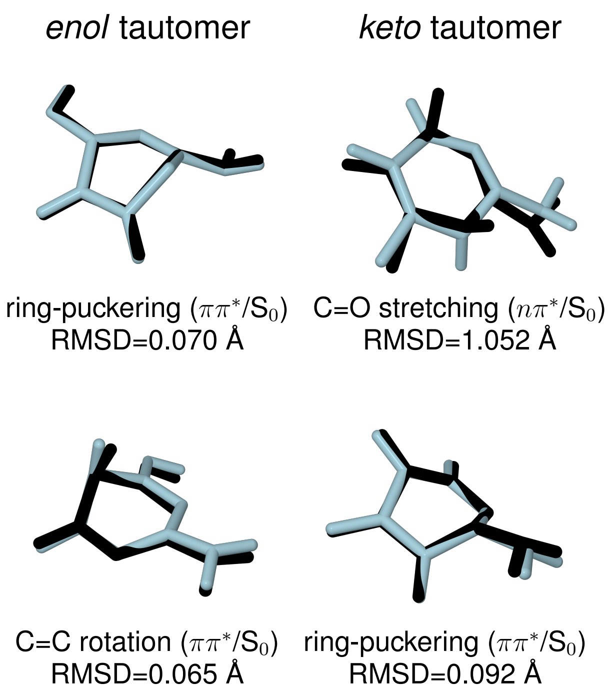


Fig. 8 Comparison of the geometries of conical-intersections optimized at the MR-CISD level (black) and MECPs obtained at the ADC(2)/MP2 level (light blue).

Interestingly, the semi-quantitative agreement of the MECP geometries obtained with the ADC(2)/MP2 and MR-CISD methods was found for all the considered $\pi\pi^*/S_0$ conical intersections, whereas the inconsistencies were observed in the case of the $n\pi^*/S_0$ MECP. Therefore, it is an interesting question for future studies whether the ADC(2)/MP2 approach is generally less applicable to optimizations of $n\pi^*/S_0$ MECPs and nonadiabatic dynamics simulations involving such deactivation channels, or this behavior is system-specific.

Estimate of the $S_1 \rightsquigarrow T_1$ intersystem crossing rate for *keto*-iC

Despite many advantages of nonadiabatic molecular dynamics simulations utilizing the ADC(2) method, currently these are lim-

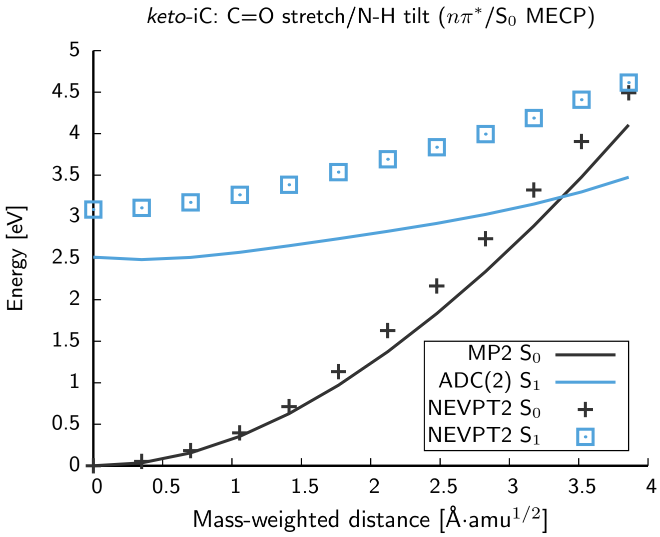


Fig. 9 Linear interpolation in internal coordinates between the S_1 minimum of *keto*-iC and the $n\pi^*$/S_0 conical intersection optimized using the MR-CISD/SA-CASSCF(6,5) method. The lines present PE curves for the ADC(2)/aug-cc-pVDZ method, while the points show the reference values obtained at the NEVPT2/SA-CASSCF(8,7)/cc-pVTZ level.

ited to the manifold of singlet states. Thus, we attempt to estimate the $S_1 \rightsquigarrow T_1$ ISC rate in *keto*-iC using the time-dependent approach based on the short-time approximation of correlation function.[32,47] The sum of squares of electronic spin-orbit coupling (SOC) matrix elements computed between S_1 ($^1n\pi^*$) state and all fine-structure components of T_1 ($^3\pi\pi^*$) state at the optimized S_1-minimum geometry amounts to 3538 cm^{-2}, which is a relatively large value for a nucleobase. For instance, Etinski *et al.* reported 2391 cm^{-2} and 2319 cm^{-2} as the corresponding values for uracil and thymine, respectively.[65] The corresponding ISC rate constant of *keto*-iC (k_{ISC}) amounts to $8.02{\cdot}10^{10}$ s^{-1} and is larger than the ISC rate constants estimated for thymine ($1.10{\cdot}10^{10}$ s^{-1}) and uracil ($2.73{\cdot}10^{10}$ s^{-1}) using the same approximation.[32,47] This indicates that triplet manifold might play a non-negligible role in the photodynamics of *keto*-iC.

The nonadiabatic dynamics simulations of isolated uracil including both singlet and triplet manifold at the CASSCF level indicated ∼20% population of the T_1 state after 1 ps.[66] Similar simulations performed for the *keto* tautomer of cytosine, also suggest the existence of ultrafast ISC channels in this molecule.[63] In contrast, experimentally measured triplet yields of canonical pyrimidine nucleobases and nucleosides in aqueous solution are much lower (<2%).[67] Theoretical and experimental investigation of ISC rates in cytosine in the gas phase also reported on relatively low triplet quantum yields, which should not exceed 5%.[68] Even though the role ISCs in canonical pyrimidine nucleobases is still under debate, our calculations suggest that triplet states should be of at least comparable (or even higher) importance in the case of *keto*-iC.

Previous dynamical studies of the *keto* tautomer of cytosine reported on the significant role of other ISC channels (*e.g.* $S_1 \rightsquigarrow T_2$),[63,69] than the one considered in this section. Our calculations indicate, that the $S_1 \rightsquigarrow T_2$ transition would be very inefficient due to significantly lower value of the sum of squares of the SOC matrix elements (23 cm^{-2}, in contrast to 3538 cm^{-2} obtained for the $S_1 \rightsquigarrow T_1$ ISC). However, if the $^1n\pi^*$ and $^3\pi\pi^*$ states associated with sizeable values of SOC matrix elements exchange their order during the excited-state dynamics, $S_1 \rightsquigarrow T_2$ transitions might occur. We emphasize, that such effects can be studied only by means of nonadiabatic molecular dynamics simulations including singlet and triplet states, and transitions between them simultaneously.

The S_1-T_1 and S_1-T_2 SOCs computed in the in the S_1 minimum of *enol*-iC are all below 5.0 cm^{-1}. This indicates that consistently with the nonadiabatic dynamics simulations of *enol* cytosine including singlet and triplet manifolds at the CASSCF level,[63] triplet states should play a minor role in the radiationless deactivation of *enol*-iC.

4 Conclusions

In summary, we performed nonadiabatic molecular dynamics simulations of two most abundant tautomers of isocytosine (iC) in the gas phase, using the ADC(2) method for electronic structure calculations. Our results indicate that both tautomers undergo efficient radiationless deactivation to the electronic ground state, with the time constants τ_{keto}=182 fs and τ_{enol}=533 fs. Similarly as was found for *enol* cytosine,[63] the investigated *enol*-iC is longer lived than the *keto* counterpart. It is evident that iC is a photostable compound in the gas phase and its behavior upon UV excitation is in many terms similar to biologically relevant nucleobases.[62] Even though additional studies are necessary to elucidate the photochemistry of iC in water solution, some of the photorelaxation channels described above might be also transferable to aqueous environment. At this stage our results indicate that iC is relatively resistant to UV irradiation, sufficiently enough to survive under space conditions, and might have been discarded during early molecular evolution owing to other factors than UV-irradiation. Therefore, the most probable reason for the absence of iC in the ensemble of biological nucleobases is most likely the co-existence of more than one tautomeric forms in bulk water and the resulting susceptibility to mispairings.[21,22] At this stage it is also impossible to refute the theory of iC acting as a bioisostere of guanine in the hypothetical early informational polymer.

The observation of longer excited-state lifetime for *enol*-iC is in agreement with the longer lifetime of *enol* cytosine.[63] We suggest that another longer time constant might be observed in *enol*-iC, which we cannot quantify owing to the limited time of dynamics simulations performed herein. It is worth noting that all the benchmark calculations performed with multireference and multiconfigurational approaches confirmed the reliability of the ADC(2) method in nonadiabatic dynamics simulations of *enol*-iC.

The MECP optimizations and PE profile calculations performed at the MR-CISD and NEVPT2 levels revealed that *keto*-iC is a challenging case for the ADC(2)-based nonadiabatic dynamics and for approximate optimization of S_1/S_0 state crossings with the ADC(2)/MP2 method. In particular, we show that the $n\pi^*$/S_0 MECP geometry optimized at the ADC(2) level is partly inconsistent with the conical intersection optimized using the reference MR-CISD method. In fact, the latter one is located over two times further away from the S_1 minimum than it is indicated by the ADC(2)/MP2 approach. According to the NEVPT2 computations, the topography of this MECP is more sloped and more energy is needed to access this critical point than indicated by the ADC(2) calculations. Since the deformation related to C=O stretching and N–H tilting leading to the $n\pi^*/S_0$ MECP was found as the dominant photorelaxation pathway of *keto*-iC, we conclude that our τ_{keto} time constant is probably underestimated. In other words, *keto*-iC should have longer excited-state lifetime than predicted by our simulations. Nevertheless, we think that *keto*-iC should still deactivate efficiently due to presence of other barrierless photorelaxation channels (e.g. ring-puckering). This example shows that verification of ADC(2) simulations at critical points (e.g. S_1/S_0 MECPs) against results of multireference methods is often crucial for complete understanding of the photochemistry of biomolecules.

In addition to all the foregoing, the estimated $S_1 \rightsquigarrow T_1$ ISC rate constant of $8.02{\cdot}10^{10}$ s^{-1} suggests that triplet states might play an important role in the overall excited-state dynamics of *keto*-iC. In contrast, the relatively small SOC values obtained for *enol*-iC indicate that the triplet manifold might be of minor importance in this tautomer. However, these findings would require further validation with spectroscopic techniques or involving simulations with full treatment of surface hoppings between singlet and triplet states.

5 Acknowledgments

This work was funded by the Grant 14-12010S from the Grant Agency of the Czech Republic and by the project CEITEC 2020 (LQ1601) with financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the National Sustainability Programme II. Support from a statutory activity subsidy from the Polish Ministry of Science and Higher Education for the Faculty of Chemistry of Wrocław University of Technology is gratefully acknowledged. Part of the calculations was performed at the Wrocław Center for Networking and Supercomputing (WCSS) and Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw (ICM).

References

1 L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.*, 2004, **39**, 99–123.

2 J. D. Sutherland, *Angew. Chem. Int. Ed.*, 2016, **55**, 104–121.

3 J. E. Šponer, J. Šponer, A. Giorgi, E. Di Mauro, S. Pino and G. Costanzo, *J. Phys. Chem. B*, 2015, **119**, 2979–2989.

4 C. B. Mast, S. Schink, U. Gerland and D. Braun, *Proc. Natl. Acad. Sci.*, 2013, **110**, 8030–8035.

5 R. Saladino, G. Botta, S. Pino, G. Costanzo and E. Di Mauro, *Chem. Soc. Rev.*, 2012, **41**, 5526–5565.

6 S. Pino, J. E. Sponer, G. Costanzo, R. Saladino and E. Di Mauro, *Life*, 2015, **5**, 372–384.

7 M. W. Powner, B. Gerland and J. D. Sutherland, *Nature*, 2009, **459**, 239–242.

8 M. Powner, J. Sutherland and J. Szostak, *Synlett*, 2011, **2011**, 1956–1964.

9 M. Powner and J. Sutherland, *ChemBioChem*, 2008, **9**, 2386–2387.

10 G. Mittapalli, Y. Osornio, M. Guerrero, K. Reddy, R. Krishnamurthy and A. Eschenmoser, *Angew. Chem. Int. Ed.*, 2007, **46**, 2478–2484.

11 K. U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347–1351.

12 M. C. Chen, B. J. Cafferty, I. Mamajanov, I. Gállego, J. Khanam, R. Krishnamurthy and N. V. Hud, *J. Am. Chem. Soc.*, 2014, **136**, 5640–5646.

13 B. J. Cafferty, I. Gállego, M. C. Chen, K. I. Farley, R. Eritja and N. V. Hud, *J. Am. Chem. Soc.*, 2013, **135**, 2447–2450.

14 B. J. Cafferty and N. V. Hud, *Isr. J. Chem.*, 2015, **55**, 891–905.

15 M. P. Callahan, K. E. Smith, H. J. Cleaves, J. Ruzicka, J. C. Stern, D. P. Glavin, C. H. House and J. P. Dworkin, *Proc. Natl. Acad. Sci.*, 2011, **108**, 13995–13998.

16 R. Saladino, V. Neri, C. Crestini, G. Costanzo, M. Graciotti and E. Di Mauro, *J. Am. Chem. Soc.*, 2008, **130**, 15512–15518.

17 R. Saladino, C. Crestini, C. Cossetti, E. Di Mauro and D. Deamer, *Origins Life Evol. Biospheres*, 2011, **41**, 437–451.

18 R. Saladino, G. Botta, M. Delfino and E. Di Mauro, *Chem. – Eur. J.*, 2013, **19**, 16916–16922.

19 N. U. Zhanpeisov, J. Šponer and J. Leszczynski, *J. Phys. Chem. A*, 1998, **102**, 10374–10379.

20 N. U. Zhanpeisov and J. Leszczynski, *J. Mol. Struct.: THEOCHEM*, 1999, **487**, 107–115.

21 A. Jaworski, J. S. Kwiatkowski and B. Lesyng, *Int. J. Quantum Chem.*, 1985, **28**, 209–216.

22 C. Roberts, R. Bandaru and C. Switzer, *J. Am. Chem. Soc.*, 1997, **119**, 4640–4649.

23 K. Kleinermanns, D. Nachtigallová and M. S. de Vries, *Int. Rev. Phys. Chem.*, 2013, **32**, 308–342.

24 C. E. Crespo-Hernández, B. Cohen, P. M. Hare and B. Kohler, *Chem. Rev.*, 2004, **104**, 1977–2020.

25 A. L. Sobolewski and W. Domcke, *Europhys. News*, 2006, **37**, 20–23.

26 K. J. Zahnle and J. C. G. Walker, *Rev. Geophys.*, 1982, **20**, 280–292.

27 C. S. Cockell and G. Horneck, *Photochem. Photobiol.*, 2001, **73**, 447–451.

28 S. Lobsiger, S. Blaser, R. K. Sinha, H.-M. Frey and S. Leutwyler, *Nat. Chem.*, 2014, **6**, 989–993.

29 M. Barbatti and H. Lischka, *Phys. Chem. Chem. Phys.*, 2015, **17**, 15452–15459.

30 R. I. Bakalska and V. B. Delchev, *J. Mol. Model.*, 2012, **18**, 5133–5146.

31 A. B. Trofimov and J. Schirmer, *J. Phys. B: At. Mol. Opt. Phys.*, 1995, **28**, 2299.

32 M. Etinski, J. Tatchen and C. M. Marian, *J. Chem. Phys.*, 2011, **134**, 154105.

33 C. Hättig, *Adv. Quantum Chem.*, Academic Press, 2005, vol. 50, pp. 37–60.

34 G. Granucci and M. Persico, *J. Chem. Phys.*, 2007, **126**, 134114.

35 R. Crespo-Otero and M. Barbatti, *Theor. Chem. Acc.*, 2012, **131**, 1–14.

36 F. Plasser, R. Crespo-Otero, M. Pederzoli, J. Pittner, H. Lischka and M. Barbatti, *J. Chem. Theory Comput.*, 2014, **10**, 1395–1405.

37 M. Barbatti, *J. Am. Chem. Soc.*, 2014, **136**, 10246–10249.

38 R. Szabla, J. Šponer and R. W. Góra, *J. Phys. Chem. Lett.*, 2015, **6**, 1467–1471.

39 H. Lischka, M. Dallos and R. Shepard, *Mol. Phys.*, 2002, **100**, 1647–1658.

40 B. G. Levine, J. D. Coe and T. J. Martínez, *J. Phys. Chem. B*, 2008, **112**, 405–413.

41 D. Tuna and W. Domcke, *Phys. Chem. Chem. Phys.*, 2015, **18**, 947–955.

42 D. Tuna, D. Lefrancois, Ł. Wolański, S. Gozem, I. Schapiro, T. Andruniów, A. Dreuw and M. Olivucci, *J. Chem. Theory Comput.*, 2015, **11**, 5758–5781.

43 D. Tuna, A. L. Sobolewski and W. Domcke, *J. Phys. Chem. A*, 2014, **118**, 122–127.

44 S. Chaiwongwattana, M. Sapunar, A. Ponzi, P. Decleva and N. Došlić, *J. Phys. Chem. A*, 2015, **119**, 10637–10644.

45 R. Crespo-Otero, N. Kungwan and M. Barbatti, *Chem. Sci.*, 2015, **6**, 5762–5767.

46 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165–169.

47 M. Etinski, *J. Serb. Chem. Soc.*, 2011, **76**, 1649–1660.

48 C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger and J.-P. Malrieu, *J. Chem. Phys.*, 2001, **114**, 10252–10264.

49 C. Angeli, R. Cimiraglia and J.-P. Malrieu, *J. Chem. Phys.*, 2002, **117**, 9138–9153.

50 C. Angeli, R. Cimiraglia and J.-P. Malrieu, *Chem. Phys. Lett.*, 2001, **350**, 297–305.

51 F. Weigend, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4285–4291.

52 F. Neese, F. Wennmohs, A. Hansen and U. Becker, *Chem. Phys.*, 2009, **356**, 98–109.

53 V. Veryazov, P. Å. Malmqvist and B. O. Roos, *Int. J. Quantum Chem.*, 2011, **111**, 3329–3338.

54 M. Barbatti, M. Ruckenbauer, F. Plasser, J. Pittner, G. Granucci, M. Persico and H. Lischka, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 26–33.

55 H. Lischka, R. Shepard, I. Shavitt, R. M. Pitzer, M. Dallos, T. Müller, P. G. Szalay, F. B. Brown, R. Ahlrichs, H. J. Böhm, A. Chang, D. C. Comeau, R. Gdanitz, H. Dachsel, C. Ehrhardt, M. Ernzerhof, P. Höchtl, S. Irle, G. Kedziora, T. Kovar, V. Parasuk, M. J. M. Pepper, P. Scharf, H. Schiffer, M. Schindler, M. Schüler, M. Seth, E. A. Stahlberg, J.-G. Zhao, S. Yabushita, Z. Zhang, M. Barbatti, S. Matsika, M. Schuurmann, D. R. Yarkony, S. R. Brozell, E. V. Beck, J.-P. Blaudeau, M. Ruckenbauer, B. Sellner, F. Plasser and J. J. Szymczak, *Columbus, Relase 7.0 2012, an Ab Initio Electronic Structure Program*, 2012.

56 F. Aquilante, L. De Vico, N. Ferre, G. Ghigo, P.-A. Malmqvist, P. Neogrady, T. B. Pedersen, M. Pitonak, M. Reiher, B. O. Roos, L. Serrano-Andres, M. Urban, V. Veryazov and R. Lindh, *J. Comput. Chem.*, 2010, **31**, 224–247.

57 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73–78.

58 M. Dračínský, P. Jansa, K. Ahonen and M. Buděšínský, *Eur. J. Org. Chem.*, 2011, **2011**, 1544–1551.

59 L. Gorb, Y. Podolyan and J. Leszczynski, *J. Mol. Struct.: THEOCHEM*, 1999, **487**, 47–55.

60 A. Y. Ivanov, S. G. Stepanian and L. Adamowicz, *J. Mol. Struct.*, 2012, **1025**, 92–104.

61 M. Barbatti, A. J. A. Aquino, J. J. Szymczak, D. Nachtigallová and H. Lischka, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6145–6155.

62 M. Barbatti, A. J. A. Aquino, J. J. Szymczak, D. Nachtigallová, P. Hobza and H. Lischka, *Proc. Natl. Acad. Sci.*, 2010, **107**, 21453–21458.

63 S. Mai, P. Marquetand, M. Richter, J. González-Vázquez and L. González, *ChemPhysChem*, 2013, **14**, 2920–2931.

64 C. L. Janssen and I. M. B. Nielsen, *Chem. Phys. Lett.*, 1998, **290**, 423–430.

65 M. Etinski, T. Fleig and C. M. Marian, *J. Phys. Chem. A*, 2009, **113**, 11809–11816.

66 M. Richter, S. Mai, P. Marquetand and L. González, *Phys. Chem. Chem. Phys.*, 2014, **16**, 24423–24436.

67 P. M. Hare, C. E. Crespo-Hernández and B. Kohler, *Proc. Natl. Acad. Sci.*, 2007, **104**, 435–440.

68 S. Lobsiger, M. Etinski, S. Blaser, H.-M. Frey, C. Marian and S. Leutwyler, *J. Chem. Phys.*, 2015, **143**, 234301.

69 M. Richter, P. Marquetand, J. González-Vázquez, I. Sola and L. González, *J. Phys. Chem. Lett.*, 2012, **3**, 3090–3095.