Ultrafast IR Spectroscopy of Polymeric Cytosine Nucleic Acids Reveal the Long-lived Species is due to a Localised State

Páraic M. Keane, a Michal Wojdyla, a Gerard W. Doorley, a John M. Kelly, a Ian P. Clark, b Anthony W. Parker, b Gregory M. Greetham, b Michael Towrie, b Luís M. Magno c and Susan J. Quinn. c*

The decay pathways of UV-excited cytosine polymers are investigated using picosecond-time-resolved infrared spectroscopy. Similar yields of a non-emissive $^1n\pi^*$ state are found in the single-stranded dC$_{50}$ polymer as in the dCMP monomer, but with a longer lifetime in the polymer (80 ps vs. 39 ps). A longer lifetime is also found in the d(CpC) dinucleotide. No evidence of excimer states is observed, suggesting that localised $^1n\pi^*$excited states are the most significant intermediates present on the picosecond timescale.

Introduction

There is great interest in the photophysics of nucleic acids, both from a fundamental standpoint and in order to understand the role of various DNA excited states in carcinogenic photodamage. 1-3 The singlet excited states of the four DNA nucleobases and the purine nucleotides are all very short-lived on the order of hundreds of femtoseconds. A more complex picture exists for the pyrimidine nucleotides. 2 In the case of 5′-cytidine monophosphate, 5′-dCMP an additional and relatively long-lived species (39 ps, ~15% yield) has been observed and assigned as a $^1n\pi^*$ state. 4-6 This observation led us to probe the influence of ring substitution on the excited state dynamics of cytosine monomer species. Unexpectedly, replacing the hydrogen with a methyl group at the N1-position of cytosine yielded behaviour similar to that of the nucleobase. 7 The contrasting behaviour between 1-methylcytosine and 5′-dCMP demonstrates that the decay mechanisms in C-systems are strongly influenced by additional factors related to the nature of the N1 substituent.

While the precise role of the sugar group in the appearance of the long-lived dark state is yet to be resolved, a number of possibilities exist. It is possible that the mass of the sugar or some steric property may restrict certain ring motions or may result in a different nuclear configuration. It is also possible that hydrogen bonding between the sugar and the solvent may aid vibrational relaxation of the hot $^1n\pi^*$ state in pyrimidines, resulting in trapping on the $^1n\pi^*$ surface. The need to understand the influence of such factors on excited state behaviour fuels continued interest in these systems. Furthermore, the non-fluorescent (‘dark’) $^1n\pi^*$ state observed in 5′-dCMP is worthy of attention as it may be responsible for the formation of C-C photodimers, which are more mutagenic and less easily repaired than the better known T-T dimers. Therefore, it is important to develop our understanding of the nature of long-lived states of cytosine and what influences their behaviour in higher order structures where consecutive cytosine residues are present.

In this study we turn our attention to the role of 5′-dCMP excited states in C-rich polynucleotides. Long-lived species observed in polynucleotides are typically attributed to either an excimeric or charge-transfer (CT) state. 8,9 However, in the case of polymeric dC systems, ultrafast studies have yielded differing interpretations of the excited states. Early studies by Plessow et al. interpreted ps-fluorescence data as showing evidence for an excimer in dCpdC and dC$_{15}$. 8 Subsequent transient absorption (TA) measurements on single-stranded dC$_{18}$ (ss-dC$_{18}$) revealed a species with a lifetime of ca. 10 ps. This species was assigned to an excimer on the basis that it was unlikely that the transient band at 570 nm could be due to a $^1n\pi^*$ state. 9 The role of excimers was challenged by recent transient fluorescence experiments by Schwalb and Temps who found the lifetimes for ss-dC$_{20}$ were similar to those of the monomer with no emissive species observed at longer times. 10
and the excited state, together with its capability to provide structural information, makes it an excellent technique with which to interrogate the dynamics of DNA related systems. Our previous picosecond TRIR studies on UV excited 5′-dCMP identified a strong marker band at 1574 cm\(^{-1}\) assigned to the \(^1\)nπ* localised cytosine state. We now use this band to investigate further the dynamics of single-stranded dC\(_{30}\) (ss-dC\(_{30}\)) and compare it to the simple dinucleotide d(CpC). Figure 1 gives the structure of the various systems investigated.

### Results and Discussion

The ps-TRIR spectra of ss-dC\(_{30}\) generated by pumping at 267 nm, are given in Figure 2b. The measurements were carried out at pH 8.5 to ensure the presence of single-stranded dC\(_{30}\) (ss-dC\(_{30}\)) which was confirmed by circular dichroism (CD) experiments, see Figure 3. The spectrum shows a positive band at 277 nm and two negative bands at 230 and 215 nm. This is similar to the spectra previously obtained for dC\(_n\) which is characteristic of the single strand form.

![Fig. 2](image-url)

**Fig. 2.** (a) The ps-TRIR spectra obtained following UV excitation (300 fs, 267 nm) of (a) 10 mM 5′-dCMP, at neutral pH. (b) ss-dC\(_{30}\) (pH 8.5) and (c) 10 mM d(CpC), all in D\(_2\)O in 50 mM phosphate buffer FTIR spectra are shown in blue below the TRIR.

However, transient fluorescence measurements convey a fraction of the story as they are blind to long-lived dark states such as \(^1\)nπ* states. The ability of time-resolved infrared spectroscopy (TRIR) to simultaneously monitor the ground state and the excited state, together with its capability to provide structural information, makes it an excellent technique with which to interrogate the dynamics of DNA related systems. Our previous picosecond TRIR studies on UV excited 5′-dCMP identified a strong marker band at 1574 cm\(^{-1}\) assigned to the \(^1\)nπ* localised cytosine state. We now use this band to investigate further the dynamics of single-stranded dC\(_{30}\) (ss-dC\(_{30}\)) and compare it to the simple dinucleotide d(CpC). Figure 1 gives the structure of the various systems investigated.

### Results and Discussion

The ps-TRIR spectra of ss-dC\(_{30}\) generated by pumping at 267 nm, are given in Figure 2b. The measurements were carried out at pH 8.5 to ensure the presence of single-stranded dC\(_{30}\) (ss-dC\(_{30}\)) which was confirmed by circular dichroism (CD) experiments, see Figure 3. The spectrum shows a positive band at 277 nm and two negative bands at 230 and 215 nm. This is similar to the spectra previously obtained for dC\(_n\) which is characteristic of the single strand form.

![Fig. 3](image-url)

**Fig. 3.** CD spectra of cytosine systems (10 mM per nucleotide).

Fitting the spectra reveals biexponential decay kinetics, see Table 1 and Fig. 4. The TRIR spectral changes over the first 10 ps are assigned to processes including vibrational cooling of hot ground states formed following subpicosecond internal conversion from the bright \(^1\)ππ* state. Interestingly, both the extent (82%) and the kinetics of this decay phase in the polymer (4.3 ± 0.3 ps) are very similar to that for the mononucleotide (84% and 5.0 ± 0.5 ps) indicating that the primary relaxation

<table>
<thead>
<tr>
<th>DNA/RNA</th>
<th>Bleach recovery</th>
<th>Transient decay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\tau_1) (ps)</td>
<td>(\tau_2) (ps)</td>
</tr>
<tr>
<td>5′-dCMP</td>
<td>5.0 ± 0.4 (84)</td>
<td>39 ± 5 (16)</td>
</tr>
<tr>
<td>5′-rCMP</td>
<td>4.0 ± 0.5 (86)</td>
<td>48 ± 8 (14)</td>
</tr>
<tr>
<td>ss-dC(_{30}) pH 8.5</td>
<td>4.3 ± 0.3 (82)</td>
<td>80 ± 15 (18)</td>
</tr>
<tr>
<td>poly(rC)</td>
<td>4.0 ± 0.3 (85)</td>
<td>85 ± 15 (15)</td>
</tr>
<tr>
<td>d(CpC)</td>
<td>4.5 ± 0.5 (84)</td>
<td>55 ± 6 (16)</td>
</tr>
<tr>
<td>d(CpC)*</td>
<td>4.7 ± 0.1 (80)</td>
<td>39 (70), 80 (30)</td>
</tr>
</tbody>
</table>

Fitted to biexponential function at carbonyl bleach maximum (ca. 1650 cm\(^{-1}\)) and transient (1574 cm\(^{-1}\)). Only the long component is shown for transient decay (see ESI Table S1). *Bleach data fitted to a constrained triexponential function.
process is not significantly altered in the polymer system. The longer-lived decay component, seen in the bleaching recovery (downward bands) of the carbonyl (ca. 1650 cm\(^{-1}\)) and ring stretches (ca. 1500-1525 cm\(^{-1}\)) was found to have a lifetime of 80 ± 15 ps. This is in good agreement with the value found for the loss of the transient absorption at 1574 cm\(^{-1}\), Table 1, Fig. 4 and ESI Fig. S1. This indicates the absence of a role of an additional species. This is longer than that observed in the monomer species yet the relative amplitude is similar to that observed for 5'-dCMP (18% vs. 16%, see Table 1).

The fact that both the intensity and position of the transient IR bands (1574 cm\(^{-1}\)) in ss-dC\(_{30}\) and 5'-dCMP are strikingly similar suggests strongly that the same excited state is present in both species, as the presence of charge transfer (CT) or excimeric states in ss-dC\(_{30}\) would be expected to give rise to new transient absorption bands that would be significantly different from those observed in 5'-dCMP. Taken together, these facts do not support the existence of a modified excited state such as an excimer or a charge transfer species but rather indicate that the monomer and single-stranded dC\(_{30}\) have a common long-lived excited state. Based on our previous work,\(^5\)\(^6\) we assign this as a localised \(^1\)nπ* state.

For completeness, and to correlate our results with other workers in this field, we have measured the ribose analogues of the C systems. 5'-rCMP possesses a long-lived transient at 1574 cm\(^{-1}\), similar to 5'-dCMP (Table 1, Fig. 5 and ESI Fig. S2). However the lifetime of this transient is lengthened slightly in 5'-rCMP. This result indicates that the decay of the \(^1\)nπ* state is influenced by structural changes in the (deoxy) ribose substituent. This is further evidence of the sensitivity of the dark state to subtle structural effects at the N1 position.

Interestingly, we also find that the TRIR spectrum and kinetics of ss-dC\(_{30}\) agree closely with those recorded for poly(rC) (Table 1, Fig. 5 and ESI Fig. S3). This is somewhat surprising as the solution-phase secondary structure of poly(dC) and poly(rC) has been the subject of some debate.\(^{13,14}\) In particular, it has been suggested that stacking interactions are more prevalent in poly(rC) than poly(dC). Importantly, the similar nature of the excited state lifetimes indicates that any secondary structural difference in the single-stranded ribo- and deoxyribose cytosine polymers has little effect on the excited state dynamics. This would certainly be expected for a localised excited state and this is discussed later.

The TRIR for both ribose and deoxyribose cytosine systems provide strong evidence for the role of a common long-lived excited state in the monomer and polymeric species. To understand why this is the case it is important to consider the interactions that exist between cytosine bases in the polymer species. The ground state interaction of adjacent cytosine bases has recently been investigated by circular dichroism for a range of dC\(_n\) oligonucleotides (n=1 to 11). Importantly, the findings established that the cytosine bases participate in pair-wise coupling in the neutral single-strand form.\(^{12}\) We can now consider that if this “pair-wise” coupling were to be maintained in the electronic excited state of the polymer then the increased lifetime (ca. 80 ps) would be a manifestation of interactions between adjacent bases, affecting the non-radiative decay from a localised excited state.

To test this hypothesis, ps-TRIR measurements on the cytosine dinucleotide, d(CpC) were performed. The dinucleotide was chosen to allow us to consider the interaction with one adjacent base. This model system at the simplest level provides a ‘pure’ pair-wise interaction which, if our hypothesis is correct, would be expected to have a lengthened (modified) excited state with similar properties to ss-dC\(_{30}\). As expected the TRIR spectrum of d(CpC) is found to be very similar to those of 5'-dCMP and ss-dC\(_{30}\) and a transient band at 1574 cm\(^{-1}\) was again observed, see
Fig. 2c. Treatment of the data with a biexponential model (as was used for the monomer and polymer systems) again shows a fast process (4.5 ps) attributed to vibrational relaxation and a longer-lived species with a lifetime of 55 ± 6 ps. This latter value lies between the values for 5′dCMP (39 ± 5 ps) and ss-dC₃₀ (80 ± 15 ps), see Table 1, Fig. 3c, ESI Fig. S4. In solution dinucleotides exist in a dynamic equilibrium between ‘stacked’ (some base overlap more polymer like) and ‘unstacked’ (little or no overlap more monomer like) forms. Therefore the kinetics displayed by d(CpC) should reflect deactivation processes associated with the stacked and unstacked forms in addition to the fast initial cooling process, which is seen at early times. Hence the 55 ps component should reflect combined dynamics of the 1nπ* state stacked and unstacked systems. To demonstrate the feasibility of this hypothesis we assumed the 55 ps component of d(CpC) to be a weighted average of the lifetimes of the monomer (39 ps, 70%) and the polymer (80 ps, 30%). The bleach recovery was then fitted to a constrained triexponential, with lifetimes and amplitudes fixed for the vibrational cooling, monomer and polymer contributions (see Experimental). A good fit was obtained that was similar to that of the unconstrained biexponential fit (see Table 1 and ESI Fig. S5). A value of 30% stacked is reasonable, although the overlap of the bases in the ‘stacked’ form is expected to be weak.

Given some of the uncertainty in the stacking interactions within a dC-rich sequence, we next examined alternative explanations for the excited state behaviour of d(CpC) and dC₃₀. Firstly, we must consider the fact that the dinucleotide d(CpC) differs from dC₃₀ in that each C-base is flanked on only one side by a neighbouring base. In our recent article we noted the sensitivity of the 1nπ* state dynamics to the N1 substituent. We argued that the presence of a deoxyribose group in dCMP allowed a favourable geometry for formation of the 1nπ* state, which was not possible in 1-methyl cytosine. Similarly, it may be noted that the N1 substituent is effectively altered as we progress from dCMP to d(CpC) and dC₃₀ as can be seen in Fig. 1. It is therefore also possible that the extended backbone imposes some conformational restriction on the cytosine base or on the deoxyribose substituents, which retards the dissipation of energy from the 1nπ* state. Alternatively, differing interactions with the solvent in each case may contribute to the different lifetimes observed. However it is interesting that the quantum yields of decay via the 1nπ* state are similar in each of dCMP, d(CpC) and ss-dC₃₀ and this fact implies that the ππ* / nπ* conical intersection is not significantly affected by the linking of the dC nucleotides.

In a recent TRIR study on the excited state behaviour of double-stranded [poly(dG-dC)]₂ we observed a transient at 1597 cm⁻¹ with a lifetime of 21 ± 4 ps, which suggested that it could be a 1nπ* state. This lifetime is considerably different to that observed here for the single-stranded dC₃₀. There are two main key factors that could be the origin of this difference. Firstly, the difference in lifetimes may reflect the influence of Watson-Crick base-pairing and improved base-stacking. Secondly, the presence of a different (guanine) base with different redox properties will greatly change the electronic interactions between the bases in [poly(dG-dC)]₂. It is important to acknowledge that the nature of the 1nπ* in cytosine systems is expected to be modified due to the interaction with the neighbouring base.

Conclusions

In conclusion, we have demonstrated that the infrared absorption band characteristic of cytosine’s 1nπ* state (at ca. 1574 cm⁻¹) is present in single-stranded dC₃₀ (and also in d(CpC)). In the case of d(CpC) the excited state kinetics can be adequately fitted to two components from 5′dCMP and ss-dC₃₀, representing the dinucleotide in unstacked and stacked forms in solution, respectively. While the yield and transient spectra are similar to those found for the nucleotide 5′-dCMP the lifetime is lengthened in both ss-dC₃₀ and d(CpC) due to nearest neighbour interactions. As no other IR bands are observed in ss-dC₃₀ and d(CpC) in the region studied (1450 – 1750 cm⁻¹), and because the transient decay and ground state recovery kinetics show similar trends, we propose that the 1nπ* state is the only major intermediate present on the picosecond timescale. Furthermore, the relative increase in excited state lifetime of linked C-rich sequences may have biological implications for photo-mutation and localisation of damage.

Experimental

ps-TRIR measurements were performed at the ULTRA and PIRATE ultrafast infrared absorption facilities at the Rutherford Appleton Laboratory, Chilton which have been described in detail elsewhere. ULTRA: The TRIR spectrometer comprises of a 10 kHz repetition rate titanium sapphire dual output amplifier (Thales), producing 0.8 mJ output with 40 fs pulse duration, at 800 nm. Optical parametric amplifiers (Light Conversion, TOPAS) and third harmonic generation of the 800 nm crystals created the mid infrared radiation and 266 nm femtosecond UV pump pulses used in these experiments. The pump pulses at the sample were at magic angle with energy of 1.5 μJ. The IR probe beam was split to form reference the probe beams which were passed through spectrographs onto MCT array detectors (IR Associates). The 5 kHz 266 nm pump pulses were focussed (~100 μm spot sizes) and overlapped with the probe beam (~50 μm spot size) in the sample cell. High speed data acquisition systems (Quantum Detectors) allowed 10 kHz acquisition and processing of the probe and reference pulses to generate a pump-on pump-off infrared absorption difference signal. The difference signal was calibrated using the characteristic cis-stilbene absorption lines. The sample was raster scanned in x- and y- directions at an approximate rate of >10 mm/s. PIRATE: The sample was excited with 267 nm, 200 fs pulses with 2 μJ of energy at 0.5 kHz repetition rate, generated from the third harmonic of part of the output from a 1 kHz, 800 nm, 200 fs, 1 mJ regenerative amplifier (Spectra Physics Tsunami/Spitfire) and probed with 150 cm⁻¹ FWHM broadband infrared pulses generated by difference frequency mixing the signal and idler outputs of a BBO (β-BaB₂O₄) optical parametric amplifier, pumped with some residual 800 nm, in AgGaSe₂ at 1 kHz. The spot size was 200 and 150 μm diameter for the pump and the probe beam, respectively. The difference signal pump-on minus pump-off was normalized on a shot-by-shot basis and typically accumulated for four successive rounds of 30 s data integration for a single time delay. The infrared beams were dispersed by 150 1/mm, 4000 nm blaze, gold grating monochromators and imaged onto 64 element MCT arrays. The data were collected in a number of 150 cm⁻¹ spectral
windows using the delay line for optical delays between 2 ps and 1.5 ns. The sample was raster scanned in x- and y- directions at an approximate rate of 100 mm/s. The difference signal was calibrated using water vapour lines present in the probe spectrum, and the spectral windows were interleaved after scaling using overlapping transients recorded at the same delay time.

Samples with an approximate volume of 40 μL were recorded between two 25 mm diameters CaF₂ plates (Crystalm Ltd), separated by a 56 micron Teflon Spacer and mounted in a demountable liquid cell (Harrick Corp.) FTIR spectra were recorded on a Perkin-Elmer FTIR2000. CD spectra were recorded on a JASCO J810 spectropolarimeter. d(CpC) and d(CpG) were synthesised by ATDBio (Southampton,UK). dCMP, rCMP and poly(rC) were purchased from Sigma-Aldrich. Data analysis:

Decay fitting was performed on OriginPro8.0/8.5. Biexponential fitting models were applied at single wavenumbers at band maxima. Fitting of d(CpC) kinetics: The decay parameters were fitted to a triexponential decay (Equation 1), where A₁, A₂ and A₃ are preexponential terms for the rapid initial cooling of the excited state, the monomer (and unstacked lifetime) and A₃ is the stacked/polymer respectively.

\[ y = A_1 \exp \left( -\frac{x}{\tau_1} \right) + A_2 \exp \left( -\frac{x}{\tau_2} \right) + A_3 \exp \left( -\frac{x}{\tau_3} \right) + y_0 \]  

(1)

\( \tau_1, \tau_2 \) and \( \tau_3 \) are the corresponding lifetimes, \( y_0 \) is an offset. The component values locked were \( A_1 = 0.8, A_2 = 0.14, A_3 = 0.06, \tau_2 = 39 \text{ ps}, \tau_3 = 80 \text{ ps} \) and \( \tau_1 \) was freely fitted. The fitting gave an equivalent \( \chi^2 \) (goodness of fit) as the unrestrained fitting. Note the A terms for the long component are 30% of the longer two lifetimes and this correlates with values commonly quoted for dinucleotides.

Acknowledgements

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant Numbers 06/RFP/CHP035, 07/RFP/CHEF437 and 10/RFP/CAP2915. We thank the EU FP7 for funded access to the Central Laser Facility, STFC (App81030).

Notes and References

1. School of Chemistry and Centre for Synthesis and Chemical Biology, Trinity College, Dublin 2, Ireland.
2. Central Laser Facility, Research Complex at Harwell, Science & Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, Oxfordshire, UK OX11 0QX.
3. School of Chemistry and Chemical Biology, Centre for Synthesis and Chemical Biology, University College Dublin, Dublin 4, Ireland.
4. Electronic Supplementary Information (ESI) available: [Biexponential kinetic fits for systems studied]. See DOI: 10.1039/b000000x/
5. Footnotes should appear here. These might include comments relevant to but not central to the matter under discussion, limited experimental and spectral data, and crystallographic data.