

# *Long-lived excited-state dynamics of i-motif structures probed by time-resolved infrared spectroscopy*

Article

Accepted Version

Keane, P. M., Baptista, F. R., Gurung, S. P., Devereux, S. J., Sazanovich, I. V., Towrie, M., Brazier, J. A., Cardin, C. J., Kelly, J. M. and Quinn, S. J. (2016) Long-lived excited-state dynamics of i-motif structures probed by time-resolved infrared spectroscopy. *ChemPhysChem*, 17 (9). pp. 1281-1287. ISSN 1439-7641 doi: <https://doi.org/10.1002/cphc.201501183>  
Available at <https://centaur.reading.ac.uk/58037/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://onlinelibrary.wiley.com/doi/10.1002/cphc.201501183/full>

To link to this article DOI: <http://dx.doi.org/10.1002/cphc.201501183>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Long-lived excited state dynamics of i-motif structures probed by time-resolved infrared spectroscopy

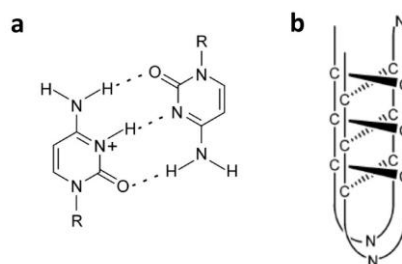
Páraic M. Keane,<sup>\*[a]</sup> Frederico R. Baptista,<sup>[b]</sup> Sarah P. Gurung,<sup>[a,c]</sup> Stephen J. Devereux,<sup>[b]</sup> Igor V. Sazanovich,<sup>[d]</sup> Michael Towrie,<sup>[d]</sup> John A. Brazier,<sup>[e]</sup> Christine J. Cardin,<sup>[a]</sup> John M. Kelly<sup>[f]</sup> and Susan J. Quinn<sup>\*[b]</sup>

**Abstract:** UV-generated excited states of cytosine (C) nucleobases are precursors to mutagenic photoproduct formation. The i-motif formed from C-rich structures is known to exhibit high yields of long-lived excited states following UV absorption. Here the excited states of several i-motif structures have been characterized following 267 nm laser excitation using time-resolved infrared spectroscopy (TRIR). All structures possess a long-lived excited state of ~300 ps and notably in some cases decays greater than 1 ns are observed. These unusually long-lived lifetimes are attributed to the interdigitated DNA structure which prevents direct base stacking overlap.

## Introduction

UV excitation of DNA is the first step in sunlight-induced skin cancer.<sup>[1]</sup> There is therefore significant interest in studying how the structure of DNA influences the nature and yield of disease causing DNA photoproducts<sup>[2]</sup> To this end one of the main concerns in the study of DNA photophysics is the lifetime of the excited states, where long-lived states are often suspected as being precursors to deleterious photoproducts. The availability of ultrafast deactivation processes are seen to act as a defense mechanism against DNA damage. In this context researchers have worked to resolve the deactivation channels accessible to double-stranded DNA through base-pairing and base-stacking interactions.<sup>[3,4]</sup> These studies point to the influence of base sequence. It may be noted that it has been proposed that rapid proton transfer in GC base pair mitigates the deactivation of longer lived excited states formed between vertically stacked nucleobases within the strand.<sup>[5-7]</sup>

Individual pyrimidine nucleotides possess longer-lived excited states not observed for the purine bases.<sup>[8,9]</sup> For example while the lifetime of the  $^1\pi\pi^*$  excited state of a cytosine nucleobase is less than 1 ps,<sup>[2]</sup> approx. 15% of excitations in dCMP decay via a non-emissive  $1n\pi^*$  state that is longer-lived (39 ps). This state has a distinctive absorption at  $1574\text{ cm}^{-1}$  (in  $\text{D}_2\text{O}$ ) that can be readily observed by picosecond time-resolved IR (ps-TRIR) spectroscopy.<sup>[9]</sup> UV irradiation of adjacent pyrimidine bases leads to the DNA photoproducts such as cyclobutane pyrimidine dimers (CPDs), and pyrimidine(6-4)pyrimidine photoproducts (64PPs). The predominant DNA photoproduct formed in cells and human skin upon UV absorption are CPDs.<sup>[10]</sup> While TRIR studies of thymine derivatives by the Zinth group revealed CPD formation from the singlet state,<sup>[11]</sup> the presence of long-lived excited states in pyrimidine rich sequences may be expected to be an important source of these photoproducts. A striking example of a DNA structure that is found to stabilize pyrimidine excited states is the tetraplex i-motif system.<sup>[12,13]</sup> This structure is formed by interdigitating base-pairs of hemi-protonated dC-rich sequences at weakly acid pH.<sup>[14]</sup> In the case of intramolecular i-motifs the core C-tracts are connected via a looped structure which is key to binding to transcriptional proteins, see Figure 1.<sup>[15]</sup> The i-motif structure has attracted much interest due to its possible role in oncogene transcription and in the onset of insulin-dependent diabetes mellitus<sup>[16]</sup> as well as its potential in nanoswitches,<sup>[17]</sup>



**Figure 1.** (a) Protonated cytosine base pair. (b) Schematic representation of an intramolecular i-motif structure where N denotes the base composition of the loop (adapted from Gurung et al.<sup>[19]</sup>).

From a photochemical viewpoint the i-motif is an interesting structure as cytosine bases are known to form mutagenic photoproducts that are not efficiently repaired by cellular machinery. In addition, a number of biologically relevant i-motif sequences also contain thymine rich loops. Thus the formation of appreciable yields of long-lived excited states in these systems raises questions about the potential for permanent DNA damage. Furthermore, the absorption spectra of cytosine rich hemi-protonated sequences are red-shifted to the UVA region, where skin penetration depths are greater and this radiation constitutes 95% of all UV radiation reaching the earth's surface.<sup>[1]</sup>

In our preliminary study<sup>[13]</sup> we probed the excited state dynamics of hemi-protonated dC<sub>30</sub> and the human telomeric sequence d(CCCTAA)<sub>4</sub> using ps-TRIR (i.e. < 2 ns) following UV laser excitation. That study revealed that the resulting excited states had a complex IR structure and that long-lived decay processes occurred beyond the time-limits of the apparatus. Notably the excited states appeared to differ from those in the corresponding single stranded sequences, where the dominant long-lived decay pathway (~15%) was assigned as a <sup>1</sup>nπ\* state, which has a lifetime of ca. 85 ps in single-stranded dC<sub>30</sub> and in poly(rC).<sup>[18]</sup>

In order to provide greater insight to the common photophysical characteristics of the i-motif we now expand our study to include additional sequences that have been shown to form stable intramolecular i-motifs in solution. To do this a C<sub>3</sub>T<sub>n</sub> family, namely, (C<sub>3</sub>T<sub>7</sub>)<sub>3</sub>C<sub>3</sub>, (C<sub>3</sub>T<sub>8</sub>)<sub>3</sub>C<sub>3</sub>, (C<sub>3</sub>T<sub>8</sub>C<sub>3</sub>T<sub>3</sub>C<sub>3</sub>T<sub>8</sub>C<sub>3</sub>).<sup>[19]</sup> together with a C<sub>5</sub> sequence, (C<sub>5</sub>T<sub>3</sub>)<sub>3</sub>C<sub>5</sub><sup>[20,21]</sup>, was chosen. These sequences also allow us to probe the influence of different T-rich loop sections on the deactivation dynamics and also allow us to establish whether the long-lived species previously observed for d(CCCTAA)<sub>4</sub> and dC<sub>30</sub> are present.<sup>[13]</sup> We then report the dynamics of a number of these systems on the extended ps to ns timescale, which allows for the first time the full decay process to be characterized in the same experiment by TRIR. Finally, we have also studied short sequences that have been shown to form intermolecular i-motifs in the crystal form (CCCT<sup>[22]</sup>, TAACCC<sup>[23]</sup>), see Table 1. Together these C-rich sequences (Table 1) allow insight into the deactivation processes of biologically relevant and synthetic i-motif structures in solution.

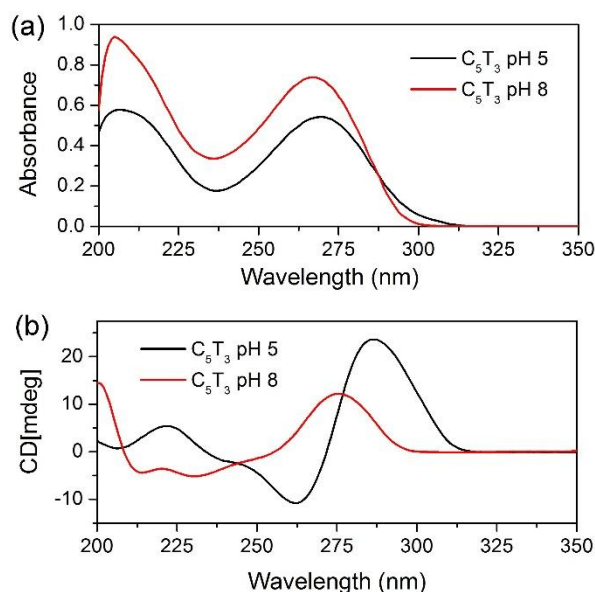
Table 1. Family of C-rich sequences used in this study

sequence	name
dC <sub>30</sub> <sup>[13]</sup>	i-dC <sub>30</sub>
d(CCCCCTTTCCCCCTTTCCCCCTTTCCCC) <sup>[20,21]</sup>	C <sub>5</sub> T <sub>3</sub>
d(CCCTTTTTTTT CCCTTTTTTTT CCCTTTTTTTTCCC) <sup>[19]</sup>	C <sub>3</sub> T <sub>7</sub>
d(CCCTTTTTTTTCCCTTTTTTTTCCCTTTTTTTTCCC) <sup>[19]</sup>	C <sub>3</sub> T <sub>8</sub>
d(CCCTTTTTTTT CCCTTT CCCTTTTTTTTCCC) <sup>[19]</sup>	C <sub>3</sub> T <sub>838</sub>
d(CCCTAACCCCTAACCCCTAACCCCTAA) <sup>[13]</sup>	HTS
d(TAACCC) <sup>[22]</sup>	TA <sub>2</sub> C <sub>3</sub>
d(CCCT) <sup>[23]</sup>	C <sub>3</sub> T

## Results and Discussion

### UV-visible Absorption and Circular Dichroism

The  $C_5T_3$  design favors the formation of one specific conformation of intramolecular i-motif in which the four core C tracts are linked by three all-thymine loops.<sup>[21]</sup> Hemi-protonation of the  $C_5T_3$  C-rich sequence results in a redshift in the cytosine absorption from 267 nm (pH 8) to 274 nm (pH 5), see Figure 2a. The presence of the i-motif structure in solution is most readily supported by circular dichroism (CD). The CD spectra recorded at pH 8 and pH 5 show significant differences. At pH 5 bands characteristic of the i-motif structure are observed with a strong positive Cotton effect at 286 nm and a negative Cotton effect at 265 nm, see Figure 2b. Similar behavior in the UV-visible and CD spectra was observed for  $dC_{30}$ , and  $C_3(T)_n$  family of sequences (Figure S1). However, in the case of the short sequences  $C_3T$  and  $TA_2C_3$  the CD did not reveal bands characteristic of i-motif formation at the same concentration of nucleotide (Figure S2).



**Figure 2.** (a) UV absorption and (b) CD spectra of  $C_5T_3$  (10 mM nucleotide) under weakly basic (pH 8) and acid (pH 5) conditions in aqueous phosphate buffered solution.

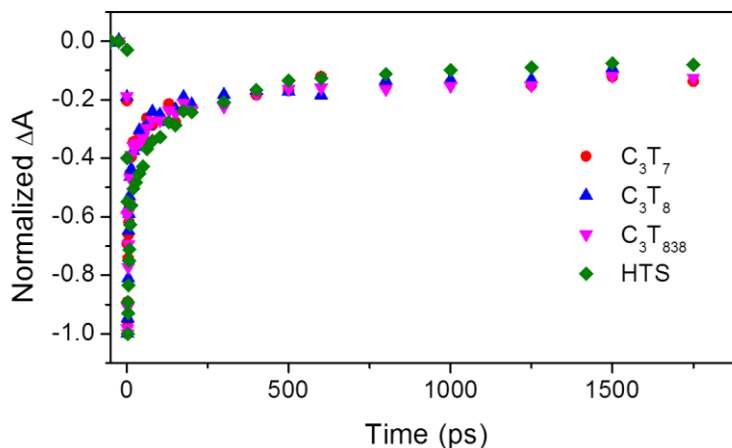
### IR and TRIR experiments

The mid infrared  $1700\text{--}1450\text{ cm}^{-1}$  reports on the carbonyl stretching vibrations and in-plane ring modes of the nucleic bases and is very sensitive to structural changes. Base pairing of the hemi-protonated cytosine sequences results in the cytosine carbonyl stretch ( $1650\text{ cm}^{-1}$ ) splitting into two bands and shifting to high wavenumbers and is accompanied by a slight shift in the ring stretch ( $1619\text{ cm}^{-1}$ ) to lower wavenumbers (see Figure S3).<sup>[24]</sup> In the TRIR spectra the formation of the excited state is also accompanied by bleaching (negative) of these ground-state bands as well as the corresponding appearance of transient features.<sup>[25]</sup>

**ps-TRIR of  $C_3(T)_n$  Variable thymine loop length sequences:** As indicated above we are interested in seeing the effect of  $dT_n$  loops on the behavior of the transient species, given their role in stabilizing i-motif structures. The IR spectra for the  $C_3(T)_7$ ,  $C_3(T)_8$  and  $C_3(T)_{838}$  at pH 5 show characteristic i-motif bands, see Table S1. The TRIR spectra, obtained following 267 nm excitation, of the three systems are quite similar, see Figure S4. The spectra show characteristic bleaching bands of the i-motif and distinctive transient species, namely a broad transient at  $1550\text{ cm}^{-1}$  and transients at  $1630\text{ cm}^{-1}$  and  $1680\text{ cm}^{-1}$ . The core cytosine ring and carbonyl bands are found to recover with similar lifetimes, with a fast component ( $\sim 80\%$ ) and a longer component ( $\sim 20\%$ ). Similar recovery kinetics are determined for the three  $C_3(T)_n$  samples, Table 2.

The fast component has a lifetime of approx.  $4.5 \pm 1\text{ ps}$  and we assign this as vibrational cooling of 'hot' ground states. While the second component has a lifetime in the region of 180 ps which is assigned to the base-paired cytosine core of the i-motif. The recovery dynamics for the main bleach at  $1666\text{ cm}^{-1}$  are compared in Figure 3 and show a small amount of non-recovery at 2 ns. The

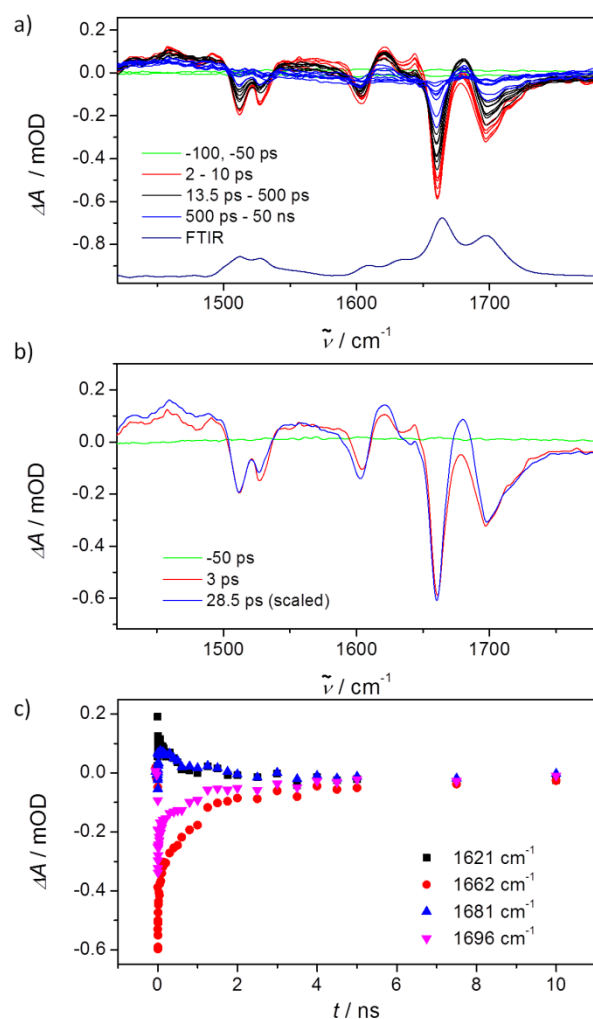
measured lifetimes compare well with the behavior observed for the biologically relevant human telomeric sequence (HTS) d(CCCTAA)<sub>4</sub>. This sequence also contains a C<sub>3</sub> based i-motif core and has a lifetime of 195 ± 30 ps. However, the contribution of the longer-lived component to the HTS deactivation is greater than that observed for the C<sub>3</sub>(T)<sub>n</sub> family.



**Figure 3.** Comparison of carbonyl bleach (1662 cm<sup>-1</sup>) recovery kinetics for C<sub>3</sub>T<sub>7</sub>, C<sub>3</sub>T<sub>8</sub>, C<sub>3</sub>T<sub>838</sub>, HTS. Signals have been normalized to equal bleach intensity.

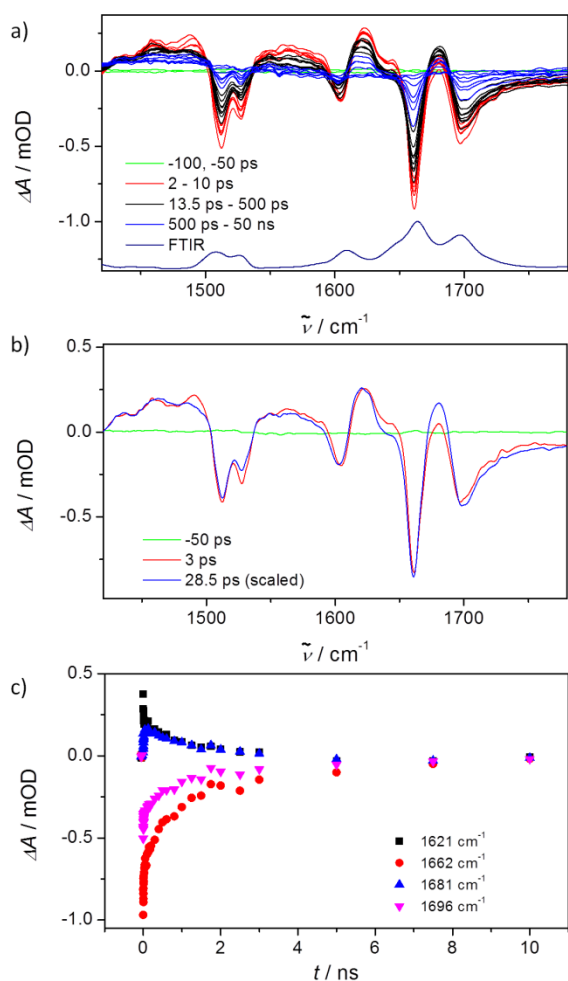
It is possible that the significant contribution of the short-lived component to the deactivation of the C<sub>3</sub>(T)<sub>n</sub> family is due to the relatively high population of thymine bases present in these sequences. The thymine bases reside in the loops where stacking interactions are expected to be minimal and as such the bases would be expected to recover predominantly by fast mechanisms.<sup>[25]</sup> It is noteworthy that the presence of long chains of T in these sequences means they may also be more susceptible to photodamage by TT dimer formation. In light of these factors we therefore decided to consider a more detailed TRIR study of the C<sub>5</sub>T<sub>3</sub> i-motif, which has both a larger number of core cytosine bases and a shorter, less flexible loop with reduced thymine contribution.

**ps-ns-TRIR C<sub>5</sub>T<sub>3</sub>:** The FTIR of i-C<sub>5</sub>T<sub>3</sub> shows two cytosine carbonyl bands at 1664 cm<sup>-1</sup> and 1697 cm<sup>-1</sup> and a ring mode at 1609 cm<sup>-1</sup> (note the carbonyl and ring modes of thymine also contribute to this region of the spectrum), see Figure 4a. The TRIR spectra following 267 nm excitation shows similar features to that of the C<sub>3</sub>(T)<sub>n</sub> family with transients at 1550 cm<sup>-1</sup>, 1630 cm<sup>-1</sup> and 1680 cm<sup>-1</sup> (Figure 4a). By comparing the early spectra with the later spectra (after the initial fast decay process has finished) it can be seen that the transient at 1680 cm<sup>-1</sup> grows in over the first ~30 ps (Figure 4b and ESI Fig S7). The excited state fully decays in approx. 10 ns. The core cytosine ring and carbonyl bands are found to recover with similar lifetimes, with a fast component and a longer component, see Figure 4c. The first component has a lifetime of 5 ± 1 ps consistent with vibrational cooling of 'hot' ground states. The second component has a lifetime of 360 ± 50 ps and we assign this to the base-paired cytosine core of the i-motif. A notable difference dynamics in C<sub>5</sub>T<sub>3</sub> to those observed for the C<sub>3</sub>(T)<sub>n</sub> family, see Table 2, is the relative contribution of the longer component is greater. This may be due to the more extended core region or possibly to the lower proportion of T bases in the sequence. Interestingly, a short-lived transient is also observed at 1647 cm<sup>-1</sup>. This may be attributed to vibrational cooling of an unstacked thymine base in the loop strand.<sup>[25]</sup>



**Figure 4.** (a) TRIR spectra of  $C_5T_3$  i-motif after excitation at 267 nm (150 fs, 1 kHz) in buffered  $D_2O$  (pH 5) at 10 mM nucleotide, (b) comparison of spectrum at 3 ps with that at 28.5 ps, scaled to equal intensity at maximum of the strongest bleach (1662  $\text{cm}^{-1}$ ), (c) comparison of the bleach and transient kinetics

**i-dC<sub>30</sub>.** Next we considered the dynamics of the dC<sub>30</sub> sequence whose excited state we previously observed did not recover by 2 ns.<sup>[13]</sup> This sequence presents a slightly more complicated situation as it may potentially form a number of structures in solution. The FTIR spectrum of the dC<sub>30</sub> sequence again shows characteristic bands of the i-motif, see Figure 5a,b and Table S1. The TRIR spectra of i-dC<sub>30</sub>, following 267 nm excitation contains similar transient bands to those observed for the previous systems (Figure 5a). As was the case with the  $C_5T_3$  i-motif system the sample is found to fully recover within approximately 10 ns with similar recovery kinetics for all bleach bands. However, in contrast to  $C_5T_3$  up to three exponentials were required to adequately fit the kinetics. The first component has a lifetime of  $5 \pm 1$  ps and we assign this as vibrational cooling of ‘hot’ ground states. Some slight changes in the spectra are observed for the vibrationally cooled system (Figure 5b). The two longer components fit as  $380 \pm 90$  ps and  $3 \pm 1$  ns.



**Figure 5.** (a) TRIR spectra of dC<sub>30</sub> i-motif after excitation at 267 nm (150 fs, 1 kHz) in buffered D<sub>2</sub>O (pH 5) at 10 mM nucleotide. (b) Comparison of spectrum at 3 ps with that at 28.5 ps, scaled to equal intensity at the maximum of the strongest bleach (1662  $\text{cm}^{-1}$ ) (c) Comparison of the bleach and transient kinetics.

We had previously reported the bleach recovery as  $300 \pm 70 \text{ ps}^{[13]}$ , similar to the value obtained by the Kohler group for hemiprotonated poly(rC) ( $310 \pm 30 \text{ ps}$ ), measured by UV/visible pump/probe spectroscopy.<sup>[27]</sup> However, in a more detailed subsequent study they noted that the decay process in hemiprotonated poly(rC) and dC<sub>18</sub> could not be satisfactorily described by exponential kinetics.<sup>[12]</sup> Similarly, the requirement to use three exponentials in the current study implies that care must be taken in the interpretation of the data. For example, the data may represent a distribution of decaying species rather than a discrete number of components. This is a possibility considering that there are many ways that hemi-protonated dC<sub>30</sub> could associate in solution. A summary of the i-motif kinetics for the systems studied is given in Table 2.



**Table 2.** Comparison of the kinetics at 1666 cm<sup>-1</sup> for the i-motif systems

sequence	$\tau_1$ (ps)	A(%)	$\tau_2$ (ps)	A(%)
i-dC <sub>30</sub>	4.8 ± 0.9	30	380 ± 90 <sup>[a]</sup>	34
C <sub>5</sub> T <sub>3</sub>	7.4 ± 1.0	50	410 ± 60	50
C <sub>3</sub> T <sub>7</sub>	4.7 ± 0.5	82	220 ± 40	18
C <sub>3</sub> T <sub>8</sub>	4.5 ± 0.5	81	160 ± 40	19
C <sub>3</sub> T <sub>838</sub>	4.0 ± 0.3	84	180 ± 30	16
HTS	5.4 ± 0.7	58	195 ± 30	42

[a] an additional component of 3 ± 1 ns is included in the fit

### Short sequences and single stranded structures:

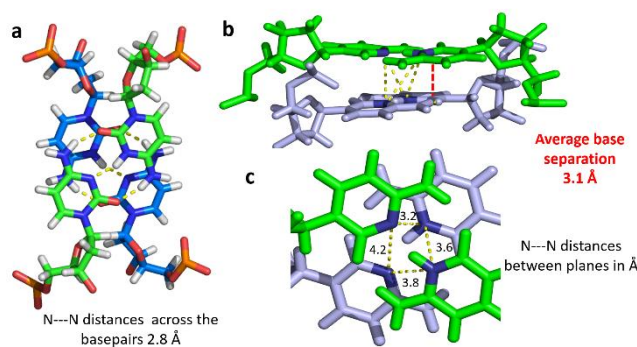
The TRIR spectra of all the sequences at pH 8 are characteristic of single-stranded C-rich DNA and are shown in Figure S4 (TRIR spectra of C<sub>3</sub>T and TA<sub>2</sub>C<sub>3</sub> both pH 5 & 8 are shown in Figure S6).<sup>[13]</sup> Common to all is the presence of a long-lived transient feature at 1574 cm<sup>-1</sup> and concomitant bleach recovery. We have previously assigned this transient band to a localized <sup>1</sup>nπ\* state, and we have shown that the lifetime observed in dCMP is lengthened in dCpdC and ss-dC<sub>30</sub> due to base-base interactions.<sup>[18]</sup> Furthermore we have shown this to be the dominant long-lived decay channel in dC polymers (~15%). In all the sequences examined in the current study, the lifetime of the C bleach recovery and the transient decay at 1574 cm<sup>-1</sup>, is 100 ps ± 20 ps, which is quite similar to the 80 ± 15 ps lifetime we previously reported in ss-dC<sub>30</sub> or poly(rC) (see Table 3; full kinetics for the 1574 cm<sup>-1</sup> transient are listed in Table S2). As we have shown earlier the lifetime of dCpdC (dC<sub>2</sub>) is 55 ± 6 ps, and this value reflects that the dinucleotide sample is a mixture of stacked and unstacked (open) dinucleotide. The results here now suggest that the long-lived component in single-stranded C tracts is generally invariant for dC<sub>n</sub> where n>2.

Under single-stranded conditions (pH 8, Figure S7) the contributions from thymine cannot be readily extracted from the data due to the significant overlap with the C bleaches. However, a small amount of the long-lived recovery may be attributed to thymine tracts. Kohler et al previously proposed a similar behavior with T-rich polymers, where a 100 ps lifetime was recorded in both TMP and dT<sub>18</sub>.<sup>[28]</sup> The similarity between the mono and the polynucleotide being attributed to the poor TT stacking. This suggests that dT<sub>n</sub> and dC<sub>n</sub> may behave quite differently, although it is not clear how this is connected to the differing photochemistry of either pyrimidine.

### Assignment of long-lived transients in the i-motif.

Our current understanding of DNA relaxation dynamics is informed by studies which have focused on resolving the contribution of base pairing and base stacking interactions to the deactivation processes. In the case of B-DNA the general picture that has emerged is one in which stacking interactions act to stabilize excited states<sup>[3,4,13]</sup> while base-pairing interactions act as a means of rapid deactivation.<sup>[5-7]</sup> Base-pair facilitated deactivation seem particularly important for Watson-Crick (WC) cytosine-guanine base pairs.<sup>[5]</sup> A number of recent publications have invoked the role of proton transfer in the deactivation of excited state energy in GC-rich double stranded DNA, noting that the process may quench the long lived state formed between stacked bases.<sup>[5-7]</sup> However, gas-phase experiments and calculations on GC base pairs have suggested that ultrafast decay may be specific to Watson-Crick base pairing.<sup>[29-32]</sup>

In this context the i-motif presents the intriguing case, not readily explained by the dynamics of B-DNA, as a structure that comprises proton rich base-pairs and minimal base stacking yet it is found to stabilize long-lived excited states. One of the difficulties in studying the i-motif is separating the influences of the CH<sup>+</sup>-C base-pairing from the base overlap. The base-pair environment of the i-motif is quite different to that of B-DNA. The binding base-pair energy (BPE) for the hydrogen bonds in the i-motif hemi-protonated bases has been calculated to be 169.7 kJ mol<sup>-1</sup>. This is significantly higher than the BPE calculated for the WC GC and neutral CC base pair which are 96.6 and 68.0 kJ mol<sup>-1</sup>, respectively.<sup>[33]</sup> NMR studies of the i-motif show the proton in the singly protonated base pair to exchange between both cytidine nucleobases with a hopping rate of at least 10<sup>8</sup> s<sup>-1</sup>.<sup>[34]</sup>



**Figure 6.** i-motif base pairing and base stacking interactions for the solution structure and internal motions of a fragment of the cytosine-rich strand of the human telomere.<sup>[37]</sup> Coordinates taken from pdb file 1ELN.

The question arises as to the influence of this base pair on the deactivation dynamics. Firstly, in contrast to dCMP the lifetime of the dCMPH<sup>+</sup> nucleotide is very short lived.<sup>9b</sup> It has also been observed that a variety of H-bonded C dimers result in only a modest increase in fluorescence lifetime (albeit in solvents such as n-hexane or CHCl<sub>3</sub>).<sup>[35]</sup> Calculations have indicated that the ability of the cytosine excited state to adopt a non-planar structure is instrumental in the ultrafast decay of the nucleotide.<sup>[9a, 36]</sup> However, access to this configuration is restricted by the i-motif's tightly held structure arising from significant backbone interactions.<sup>[37]</sup> In addition to contributing to the rigid structure the backbone interactions hold the adjacent bases perpendicular to each other. This introduces overlap between the exocyclic keto groups and amino groups of the interdigitated cytosine bases at the expense of ring overlap. In this structure bases held at a reduced inter-base spacing of 3.1 Å, see Figure 6.<sup>[38]</sup>

In our previous study, we suggested that the rich transient IR spectrum of photo-excited i-dC<sub>30</sub> may arise due to a charge transfer interaction along the sequence.<sup>[13]</sup> A possible result of such a charge transfer would be the generation of CH<sup>•</sup> and C<sup>•+</sup> species. Recently, Kohler and co-workers have calculated that the IR band of CH<sup>•</sup> appears around 1620 cm<sup>-1</sup>.<sup>[5]</sup> Although the cytosine containing systems described above were not performed in hemi-protonated systems we can use the TRIR bands as a guide to assigning the species in our systems. The TRIR spectra for the i-motif systems show prominent transient bands at 1550 cm<sup>-1</sup>, 1630 cm<sup>-1</sup> and 1680 cm<sup>-1</sup>. In the case of the 1680 cm<sup>-1</sup> the band appears to grow in with time. It is therefore intriguing to suggest that initial excitation generates C<sup>•-</sup> (1550 cm<sup>-1</sup>) which then undergoes proton transfer process producing the CH<sup>•</sup> (1630 cm<sup>-1</sup>) and corresponding C<sup>•+</sup> (1680 cm<sup>-1</sup>) species. The positions of the bands would also be expected to be influenced by how equally 'shared' the proton is.

## Conclusions

UV excitation of a variety of i-motif structures result in formation of significant populations of long-lived excited states with a lifetime greater than 300 ps. This is one of the longest-lived singlet excited states observed in DNA structures and significantly longer-lived than those observed for base-paired B-DNA. The results indicate that the formation of this excited state is influenced by the core C-tract length. The use of TRIR also allows probing of the loop bases which appear to recover as distinct single strands. Our results also suggest that the lifetime of the second stabilized nπ\* component in single-stranded C tracts is generally invariant for dC<sub>n</sub> where n>2. It is interesting that long-lived excited states are more prevalent in tetraplex structures (i-motif, G-quadruplex<sup>[39]</sup>) than B-DNA. This suggests that, from a molecular evolution point of view, these structures may be less well optimized for dealing with UV excitation via ultrafast deactivation.

## Experimental Section

Oligonucleotides were purchased from ATDBio (Southampton UK) in HPLC-purified form. Samples for CD and TRIR experiments were prepared by placing approx. 40 μl of 10 mM (per nucleotide) ODN in D<sub>2</sub>O between two CaF<sub>2</sub> plates (25 mm x 2 mm dia.) separated by a 50 micron Teflon spacer. Samples were recorded at pH 5 or pH 8 by use of phosphate (to directly compare to previous results)

or cacodylate buffers (to compare to systems recently studied<sup>[19]</sup> and those in the crystallization media). FTIR spectra were recorded on a Nicolet Avatar FTIR spectrometer. TRIR spectra were recorded at the Lasers for Science Facility using the time-resolved multiple probe spectroscopy<sup>[40]</sup> capability on the ULTRA apparatus, previously described in detail elsewhere. The sample was excited by a 267 nm pulse (150 fs, 1  $\mu$ J, beam diameter approx. 150  $\mu$ m). Experiments were performed at the magic angle with respect to the pump and probe beams. Samples were checked before and after the experiment using UV spectroscopy (Perkin-Elmer lambda 950) CD spectra were recorded on a JASCO J810 spectropolarimeter.

## Acknowledgements

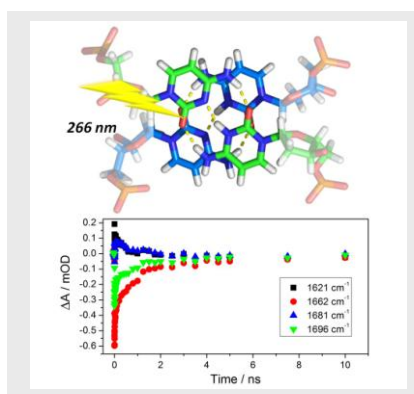
The work was supported by the BBSRC grants BB/K019279/1 and BB/M004635/1 and a Royal Irish Academy/Royal Society International Exchange Scheme award. These experiments were carried out thanks to the programmed access approved by the CLF (Application No. 14230014). S.J.Q. & F.R.B. acknowledge financial support from the University College Dublin College of Science.

**Keywords:** i-motif • DNA • transient infrared • excited-state • cytosine

- [1] J. Cadet, J. R. Wagner *Cold Spring Harb Perspect Biol.* **2013**, *5*, a012559.
- [2] C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, B. Kohler, *Annu. Rev. Phys. Chem.* **2009**, *60*, 217-239
- [3] G. W. Doorley, M. Wojdyla, G. W. Watson, M. Towrie, A. W. Parker, J. M. Kelly, S. J. Quinn, *J. Phys. Chem. Lett.* **2013**, *4*, 2739-2744.
- [4] D. B. Bucher, B. M. Pillés, T. Carell, W. Zinth, *Proc. Natl. Acad. Sci. U. S. A* **2014**, *11*, 4369-4374
- [5] Y. Zhang, K. de la Harpe, A. A. Beckstead, R. Improta, B. Kohler, *J. Am. Chem. Soc.* **2015**, *137*, 7059-6062
- [6] D. Bucher, A. Schlueter, T. Carell, W. Zinth, *Angew. Chem. Int. Ed.* **2014**, *53*, 11366-11369
- [7] K. Rottger, H. J. B. Marroux, M. P. Grubb, P. M. Coulter, H. Bohnke, A. S. Henderson, M. C. Gala, F. Temps, A. J. Orr-Ewing, G. M. Roberts, *Angew. Chem. Int. Ed.* **2015**, *54*, 14719-14722; *Angew. Chem.* **2015**, *127*, 14932-14935
- [8] P. M. Hare, C. E. Crespo-Hernández, B. Kohler, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 435-440
- [9] (a) P. M. Keane, M. Wojdyla, G. W. Doorley, G. W. Watson, I. P. Clark, G. M. Greetham, A. W. Parker, J. M. Kelly, S. J. Quinn, *J. Am. Chem. Soc.* **2011**, *133*, 4212-4215; (b) S. Quinn, G. W. Doorley, G. W. Watson, A. J. Cowan, M. W. George, A. W. Parker, K. L. Ronayne, M. Towrie, J. M. Kelly, *Chem. Commun.* **2007**, 2130-2132
- [10] (a) G. M. Halliday, J. Cadet *J. Invest. Dermatol.* **2012**, *132*, 265–267; (b) J. Cadet, S. Mouret, J. L. Ravanat, T. Douki, *Photochem Photobiol C* **2012**, *88*, 1048–1065.
- [11] W. J. Schreier, T. E. Schrader, F. O. Koller, P. Gilch, C. E. Crespo-Hernández, V. N. Swaminathan, T. Carell, W. Zinth, B. Kohler, *Science* **2007**, *315*, 625–629.
- [12] B. Cohen, M. H. Larson, B. Kohler, *Chem. Phys.* **2008**, *350*, 165-174.
- [13] P. M. Keane, M. Wojdyla, G. W. Doorley, J. M. Kelly, A. W. Parker, I. P. Clark, G. M. Greetham, M. Towrie, L. M. Magno, S. J. Quinn, *Chem. Commun.* **2014**, *50*, 2990-2992
- [14] S. Benabou, A. Avino, R. Eritja, C. Gonzalez, R. Gargallo *RSC Adv.* **2014**, *4*, 26956-26980
- [15] (a) H.-J. Kang, S. Kendrick, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4172–4185; (b) S. Kendrick, H.-J. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4161–4171.
- [16] P. Catasti, X. Chen, L. L. Deaven, R. K. Moyzis, E. Bradbury G. Gupta, *J. Mol. Biol.* **1997**, *272*, 369-382
- [17] (a) D.-L. Ma, H.-Z. De, D. S.-H. Chana, C.-H. Leung, *Chem. Sci.* **2013**, *4*, 3366-3380; (b) D. Zikich, K. Liu, L. Sagiv, D. Poirath, A. Kotlyar, *Small* **2011**, *7*, 1029-1034; (c) H. B. Ghodke, R. Krishnan, K. Vignesh, G. V. P. Kumar, C. Narayana, Y. Krishnan, *Angew. Chem. Int. Ed.*, **2007**, *46*, 2646–2649.
- [18] P. M. Keane, M. Wojdyla, G. W. Doorley, M. W. George, A. W. Parker, G. M. Greetham, L. M. Magno, S. J. Quinn, *Phys. Chem. Chem. Phys.* **2012**, *12*, 6307-6311
- [19] S. P. Gurung, C. Schwartz, J. P. Hall, C. J. Cardin, J. A. Brazier, *Chem. Commun.* **2015**, *51*, 5360-5362
- [20] J. L. Mergny, L. Lacroix, X. G. Han, J. L. Leroy, C. Hélène, *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898
- [21] J. Völker, H. H. Klump, K. J. Breslauer, *Biopolymers*, **2007**, *86*, 136-147
- [22] C. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11636-11640
- [23] C. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 3874-3878
- [24] (a) D. Tsankov, M. Krasteva, V. Andrushchenko, J.H. van de Sande, H. Wieser, *Biophys. Chem.* **2006**, *119*, 1-6; (b) E. Taillandier, J. Liquier, *Handb. Vib. Spectrosc.* **2002**, 3465-3480
- [25] M. Towrie, G. W. Doorley, M. W. George, A. W. Parker, S. J. Quinn, J. M. Kelly, *Analyst* **2009**, *134*, 1265-1273
- [26] B. M. Pillés, D. B. Bucher, L. Liu, P. Gilch, W. Zinth, W. J. Schreier, *Chem. Commun.* **2014**, *50*, 15623-15625
- [27] C. E. Crespo-Hernández, B. Cohen, P. M. Hare, B. Kohler, *Chem. Rev.* **2004**, *104*, 1977–2019.

- [28] C. E. Crespo-Hernández, B. Cohen, B. Kohler, *Nature* **2005**, *436*, 1141-1144
- [29] A. L. Sobolewski, W. Domcke, C. Hattig *Proc. Natl. Acad. Sci. U. S. A.*, **2005**, *102*, 17903-17906
- [30] A. Abo-Riziq, L. Grace, E. Nir, M. Kabelac, P. Hobza, M. S. de Vries, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 20-23
- [31] A. Kumar, M. D. Sevilla, *Chem. Rev.* **2010**, *110*, 7002-7023
- [32] S. Steenken, *Chem. Rev.* **1989**, *89*, 503-520
- [33] B. Yang, M. T. Rodgers, *J. Am. Chem. Soc.* **2014**, *136*, 282–290.
- [34] A. L. Lieblein, M. Kramer, A. Dreuw, B. Furtig, H. Schwalbe, *Angew. Chem. Int. Ed.* **2012**, *51*, 4067–4070.
- [35] (a) N. K. Schwalb, T. Michalak, F. Temps, *J. Phys. Chem. B* **2009**, *113*, 16365-16376; (b) N. K. Schwalb, F. Temps, *J. Photochem. Photobiol. A* **2009**, *208*, 164-170
- [36] (a) L. Serrano-Andrés, M. Merchán, *J. Photochem. Photobiol. C* **2009**, *10*, 21–32; (b) T. Gustavsson, A. Banyasz, E. Lazzarotto, D. Markovitsi, G. Scalmani, M. J. Frisch, V. Barone, R. Improta, *J. Am. Chem. Soc.* **2006**, *128*, 607-619
- [37] I. Berger, M. Egli, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 12116.
- [38] A. T. Phan, M. Guéron, J. L. Leroy, *J. Mol. Biol.* **2000**, *299*, 123-144
- [39] D. A. McGovern, S. Quinn, G. W. Doorley, A. M. Whelan, K. L. Ronayne, M. Towrie, A. W. Parker, J. M. Kelly, *Chem. Commun.* **2007**, 5158-5160
- [40] G. M. Greetham, D. Sole, I. P. Clark, A. W. Parker, M. R. Pollard, M. Towrie, *Rev. Sci. Instrum.* **2012**, *83*, 103107.

**Exciting longevity:** UV excitation of the four-stranded cytosine i-motif produces excited states lasting into the nanosecond region, suggesting that they are vulnerable to photodamage. A variety of i-motif forming structures are probed using time-resolved infrared spectroscopy. Distinctive transient IR features are observed and their possible origins discussed with reference to the unique structure of the i-motif.



- [a] Dr. P. M. Keane (s), Ms. S. P. Gurung, Prof. C. J. Cardin  
Department of Chemistry, University of Reading  
Whiteknights, Reading, RG6 6AD (UK)  
E-mail: [p.keane@reading.ac.uk](mailto:p.keane@reading.ac.uk)
- [b] Mr. F. R. Baptista, Mr. S. J. Devereux, Dr. S. J. Quinn  
School of Chemistry, University College Dublin, Belfield, Dublin 4,  
(Ireland)  
E-mail: [susan.quinn@ucd.ie](mailto:susan.quinn@ucd.ie)
- [c] Ms. S. P. Gurung, Diamond Light Source, Harwell Science and  
Innovation campus, Didcot, Oxfordshire, OX11 0QX, (UK)
- [d] Dr. I. V. Sazanovich, Prof. M. Towrie  
Central Laser Facility, Research Complex at Harwell, STFC  
Rutherford Appleton Laboratory, Harwell Science and Innovation  
campus, Didcot, Oxfordshire, OX11 0QX, (UK)
- [e] Dr. J. A. Brazier,  
Department of Pharmacy, University of Reading  
Whiteknights, Reading, RG6 6AD (UK)
- [f] Prof J. M. Kelly,  
School of Chemistry, Trinity College Dublin, Dublin 2, (Ireland)

Supporting information for this article is given via a link at the end of the document.