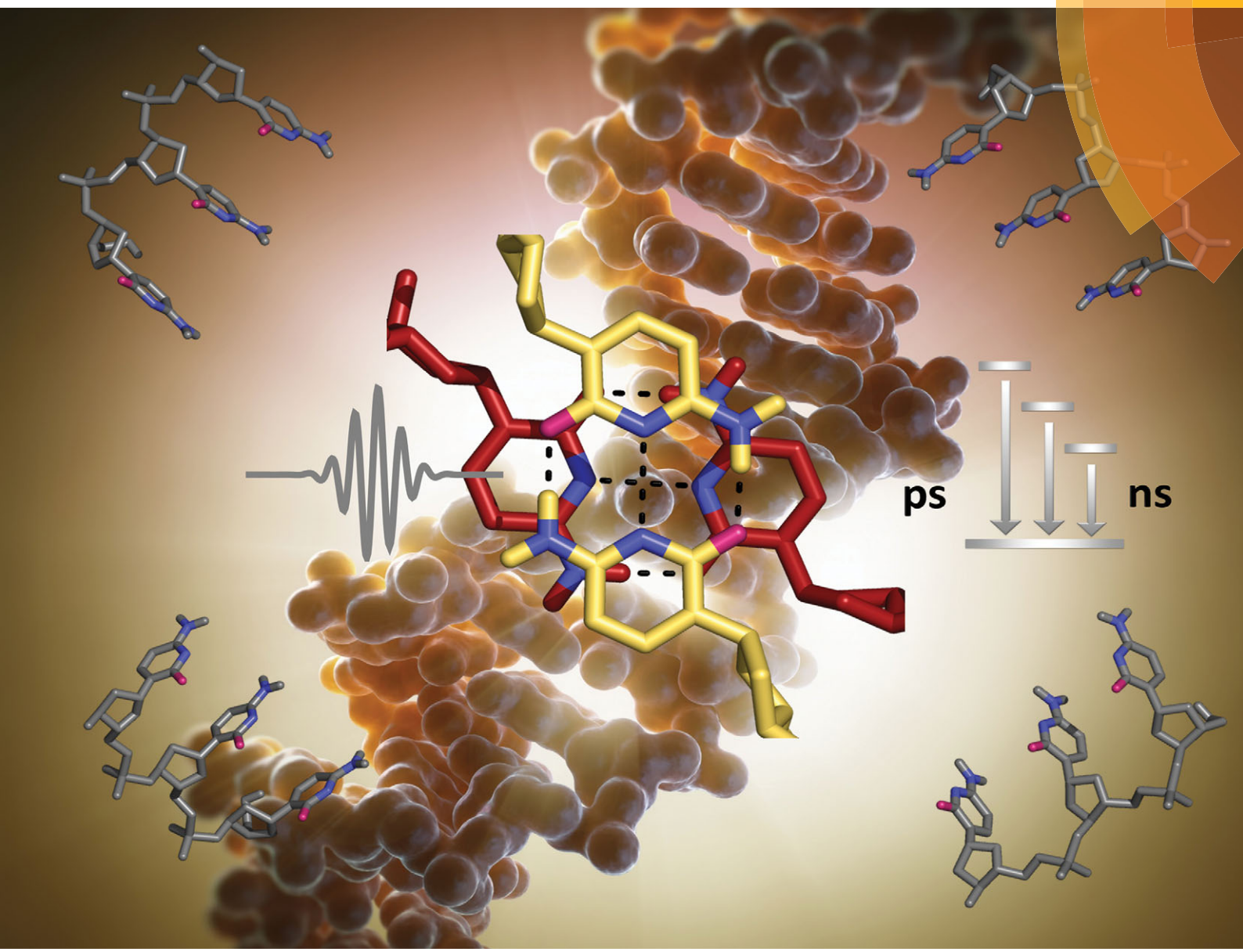
Volume 50 | Number 23 | 21 March 2014 | Pages 2971–3110

ChemComm

Chemical Communications

www.rsc.org/chemcomm



ISSN 1359-7345

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

Susan J. Quinn *et al.*

Long-lived excited states in i-motif DNA studied by picosecond time-resolved IR spectroscopy

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2014, **50**, 2990

Received 30th August 2013,
Accepted 18th September 2013

DOI: 10.1039/c3cc46594b

www.rsc.org/chemcomm

Long-lived excited states in i-motif DNA studied by picosecond time-resolved IR spectroscopy†

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

The transient IR absorption spectrum for UV-excited i-motif DNA is reported for the first time and found to possess complex dynamics pointing to multiple decay processes, including possible charge transfer between packed hemi-protonated C bases.

The i-motif is a unique DNA secondary structure formed by cytosine-rich sequences which assemble through a network of interdigitated hemi-protonated base-pairs, see Chart 1.[1] The structure is a source of wide-ranging interest and has been exploited in areas such as switching and nanostructures.[2] However, the chief focus of interest is its role in the physiological properties of biologically relevant C-rich sequences, where the i-motif is speculated to participate in the onset of insulin dependent diabetes mellitus[3] and in oncogene transcription.[4] *In vivo*, C-rich sequences occur as complementary sequences to quadruplex-forming G-rich DNA. G-quadruplex formation has been observed for the gene promoter region of several genes associated with the development of cancer. Significantly, for several of these promoter genes the complementary C-rich DNA has been observed to form the i-motif.[5]

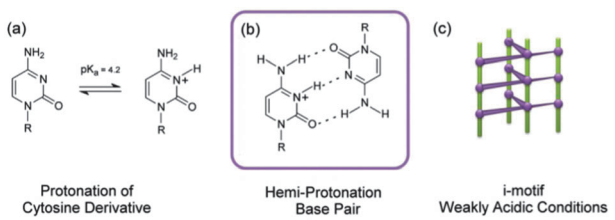


Chart 1 Structures of (a) protonation of cytosine (b) hemi-protonated C base-pairs (c) intercalated i-motif structure.

[a] *School of Chemistry and Chemical Biology, University College Dublin, Dublin 4, Ireland. E-mail: susan.quinn@ucd.ie; Tel: +353 17162407*
[b] *School of Chemistry, Trinity College, Dublin 2, Ireland*
[c] *Central Laser Facility, Research Complex at Harwell, Science & Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Oxford, Didcot, Oxfordshire, UK OX11 0QX*

† Electronic supplementary information (ESI) available: Experimental details and additional spectra are supplied. See DOI: 10.1039/c3cc46594b

C-rich DNA is susceptible to photodamage which may lead to the formation of mutagenic cyclobutane pyrimidine dimers (CPDs).[6] Consequently, there is interest in the study of the photophysics of cytosine systems in order to understand the mechanisms of damage.[7–10] The lifetime of the $^1\pi\pi^*$ excited state of the isolated cytosine nucleobase is less than 1 ps.[11] However approx. 15% of excitations in mononucleotide dCMP can decay *via* a relatively long-lived (39 ps) and non-emissive $^1n\pi^*$ state. This state has a distinctive absorption at 1574 cm^{-1} (in D_2O) that can be observed using picosecond time-resolved IR (ps-TRIR) spectroscopy. We have recently shown that this state is the major intermediate present on the picosecond timescale in single-stranded polymeric cytosine systems, where in the case of single-stranded (ss)-dC_{30} the lifetime is lengthened to 80 ps due to nearest neighbour interactions.[12] The formation of hemi-protonated C-tracts has been shown to dramatically affect the excited state properties, with reports of long-lived excited states and fluorescence on the nanosecond timescale.[7] However, the origin of long-lived excited states, and the role of the $^1n\pi^*$ localised 'dark' state in cytosine i-motif photochemistry remains unclear. To address these questions we have used ps-TRIR to probe the dynamics of two i-motif forming sequences, dC_{30} and 5′-d$(CCCTAA)_4$, the complementary sequence to the human telomeric sequence 5′-d$(TTAGGG)_4$ which was the subject of a previous study.[13]

At pH 5.5 hemiprotonation of dC_{30} results in the formation of the i-motif (i-dC_{30}). This is accompanied by a red shift in the UV spectrum and identified by a characteristic CD spectrum with bands at 266 nm (negative) and 288 nm (positive),[14] see Fig. 1. The corresponding FTIR shows that the single carbonyl band of ss-dC_{30} at 1650 cm^{-1} is replaced by two carbonyl bands at 1665 and 1695 cm^{-1} in i-dC_{30}, while the loss of π-electron density in the ring results in the suppression and slight shifting of the ring stretches, see Fig. 2a and b.[15] This characteristic IR signature should be particularly helpful when we examine the transient IR spectra as it allows us to be certain about which ground state species has been excited – a particularly useful feature for oligocytidine systems where multiple species may be present.[7]

The TRIR spectra (1400–1650 cm^{-1}) following 267 nm excitation of dC_{30} at pH 5.5 and 8.5 are compared in Fig. 2. It may be noted

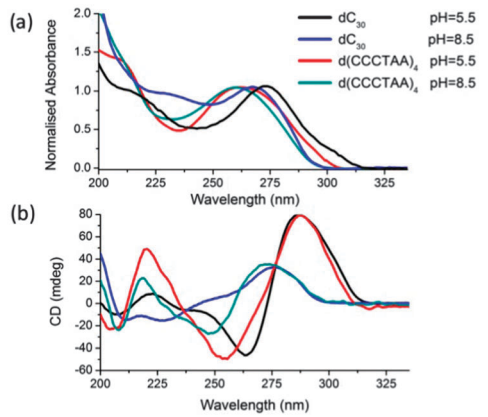


Fig. 1 (a) UV-visible absorption and (b) CD spectra of dC_{30} and $d(CCCTAA)_4$ under weakly basic and acidic conditions in phosphate buffered D_2O. (i-motif CD signals scaled for ease of comparison. For raw data see ESI,† Fig. S1).

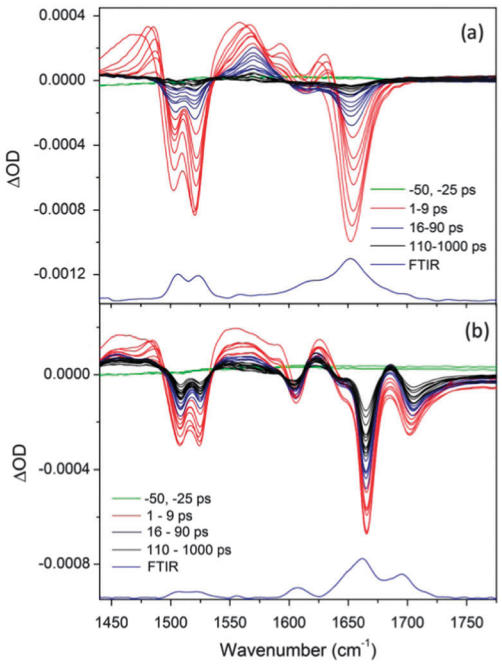


Fig. 2 ps-TRIR spectra of 10 mM (a) ss-dC_{30} (b) i-dC_{30}.

that at pH 5.5 the negative 'bleach' bands correspond to those caused by the removal of the i-motif ground state while those at pH 8.5 are consistent with the ss-DNA form.[15] The transient decay and bleach recovery over the initial 10 ps (shown in red in the figure) are assigned to cooling of the vibrationally 'hot' ground states.[16] This occurs at a similar rate for both ss-dC_{30} and i-dC_{30} (4–5 ps). It is notable that the amount of short-lived species, attributed to vibrational cooling of monomer (unstacked) bases, is significantly less for i-dC_{30} than ss-dC_{30}.

As previously reported the TRIR spectrum of ss-dC_{30} is dominated by a band at 1574 cm^{-1} (80 ± 15 ps, see ESI,† Fig. S2) assigned to the $^1n\pi^*$ excited state.[12] By contrast the TRIR spectra recorded for the sample at pH 5.5 display more complex behaviour. Thus after the initial vibrational cooling was complete, the shape of the transient band between 1540 and 1600 cm^{-1} changed with time (Fig. 2b). At earlier times (16–90 ps, navy in Fig. 2b) transient absorption peaked at 1574 cm^{-1}. At greater than 150 ps a much longer-lived broad absorption band with maximum at *ca.* 1545 cm^{-1} dominates the spectrum, with additional transient bands present at 1623 cm^{-1} and 1683 cm^{-1} (ESI,† Fig. S3 and S4). Again in contrast to the behaviour of ss-dC_{30} the ground state does not recover fully over the 1000 ps timescale. This indicates the presence of a very long-lived transient species.

Kinetics of the recovery of the strongest bleach at 1664 cm^{-1} were analysed using biexponential fitting yielded, in addition to the rapid vibrational cooling, yielded a lifetime of 300 ± 70 ps (ESI,† Fig. S5) and a very long-lived species which contributes 18% to the overall signal (Fig. 3 and Table 1). Similar analysis of the two principal transient bands gave lifetimes of 162 ± 53 ps at 1574 cm^{-1} and 241 ± 81 ps at 1545 cm^{-1} (ESI,† Fig. S6). It should be noted that these are average lifetimes and that there is some overlap in the absorption of the transient bands between 1540 cm^{-1} and 1570 cm^{-1}.

In considering the nature of these transient species it is necessary to acknowledge a number of key features of the i-motif structure. Firstly, it is a rigid structure which is held tightly through backbone interactions. The increased rigidity compared to single-stranded and B-DNA may inhibit the cytosine excited state adopting a non-planar structure, which is instrumental in the ultrafast decay of the nucleotide.[9] Secondly, there is overlap between the exocyclic keto groups and amino groups of the interdigitated cytidine bases (held at a reduced inter-base spacing of 3.1 Å), while stacking interactions are reduced in the i-motif structure as adjacent bases are positioned perpendicular to each other.[17] Thirdly, the structure may be considered to contain protonated cytosine and unprotonated cytosine moieties.

The transient species peaking at 1574 cm^{-1} may be assigned to a $^1n\pi^*$ localised excited state, possibly due to the presence of some single-stranded dangling C tracts. [A similar profile but with a greater proportion of the 1574 cm^{-1} transient is seen for dC_{30} at neutral pH where a mixture of ss-dC_{30} and i-dC_{30} is present (ESI,† Fig. S7).]

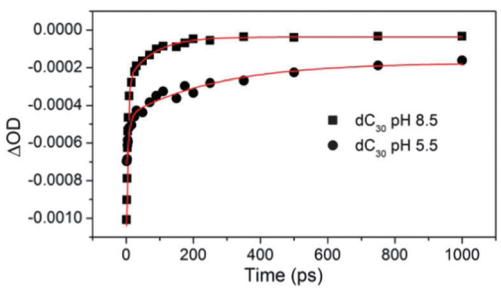


Fig. 3 Comparison of bleach recovery kinetics of 10 mM ss-dC_{30} (1650 cm^{-1}) and 10 mM i-dC_{30} (1664 cm^{-1}) fitting by a biexponential function.

Table 1 Summary of recovery kinetics for C-rich systems

DNA	τ_1 (ps)	A_1 (%)	τ_2 (ps)	A_2 (%)	A_3 (%)
5′-dCMP	5.0 ± 0.4	84	39 ± 5	16	0
5′-$HdCMP^+$	4.3 ± 0.3	100	—	—	0
ss-dC_{30}	4.3 ± 0.3	82	80 ± 15	18	0
i-dC_{30}	5.5 ± 1.1	46	300 ± 70	36	18
ss-$d(CCCTAA)_4$	4.0 ± 0.5	63	95 ± 15	27	0
i-$d(CCCTAA)_4$	5.8 ± 0.6	48	175 ± 25	38	14

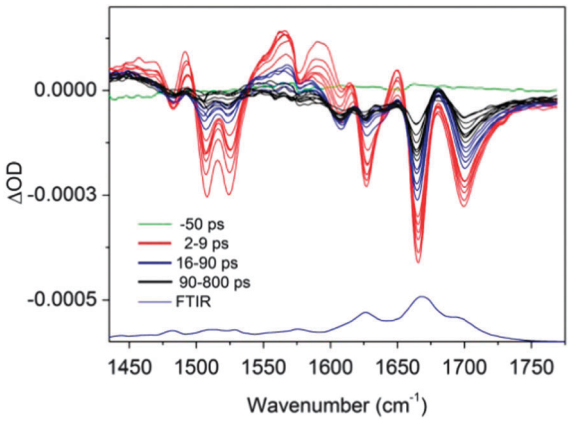


Fig. 4 ps-TRIR spectrum of 10 mM i-d(CCCTAA)$_4$ at pH 5.5.

The spectral band peaking at 1545 cm^{-1} is assumed to represent the vibrational spectrum of a transient species in the i-motif. One might consider that this is a protonated cytosine excited state. However it should be noted that protonation of dCMP prevents the formation of the ^{1}nπ^* state and the sole species observed after 2 ps is the vibrationally excited electronic ground state (ESI,† Fig. S8), so we regard this as unlikely.[10]

Another possibility we considered is that the 1545 cm^{-1} band is due to a hydrogen-bonded neutral cytosine held rigidly in the i-motif structure. Indeed it has been shown by Schwalb *et al.* that some C–C base pairing motifs result in a modest increase in fluorescence lifetime (~10–20 ps) in non-aqueous solvents.[18] However, in those cases the lifetimes are still relatively short and also the influence of the solvent prevents easy comparison. Further it may also be noted that base-pairing (for example in G–C systems) is often a path for ultrafast decay *via* proton transfer.[19]

We therefore consider whether the 1545 cm^{-1} species in d(C$_{30}$) could result from the formation of charge transfer states, as similar species have been proposed for other polynucleotide and dinucleotide systems.[20] For the i-motif the likely charge transfer process is between C and CH$^+$ moities yielding C$^{\bullet-}$ CH$^\bullet$. As pointed out by Cohen *et al.*[7] the structure of the i-motif positions the electron-rich amino group near the electron deficient pyrimidine ring of the nucleobase below and thus provide a favourable environment for the formation of CT states. It may also be noted that the poor overlap of the bases could reduce the rate of back electron transfer and lead to a relatively long-lived charge-separated state. We now turn to the long-lived ns species, the presence of which is consistent with previous observations for hemiprotonated dC$_{18}$ using visible transient absorption spectroscopy.[7] The lifetime suggests that this could be a triplet state possibly formed by the decay of the 1545 cm^{-1} species.

Finally we consider whether similar behaviour is observed for a biologically relevant i-motif sequence d(CCCTAA)$_4$. The lifetimes of the transient absorption and the bleaching bands for the TRIR spectrum of ss-d(CCCTAA)$_4$ are quite similar to those recorded for ss-dC$_{30}$ (*ca.* 100 ps, see ESI,† Fig. S9 and S10), see Table 1. However, upon formation of the i-motif, (*i.e.* at pH 5.5), we find evidence for both a longer-lived species (175 ± 25 ps) and a very long-lived transient that did not recover fully on the measurement timescale, see Fig. 4 and ESI,† Fig. S11. These features are similar to those found for i-dC$_{30}$. We therefore propose this transient behaviour is associated with the cytosine moieties in the i-motif structure.

In summary this TRIR study has provided identification of signature IR bands for the long-lived species found upon UV-excitation of the i-motif in C-rich DNA. The slow decay is not due to simple protonation, but is rationalised in terms of the specific structural features of the i-motif. The most likely origin is charge transfer between closely packed C bases. Finally, it is noteworthy that the substantially red-shifted absorption spectrum of i-motif DNA means that it may be selectively excited by UVB light above 300 nm, where protection from atmospheric ozone decreases (as has also been suggested for the G-quadruplex sequence[21]). This, combined with the very long-lived excited states present in hemiprotonated C tracts, suggests that these systems may have an important role in the photochemistry of cytosine *in vivo*.

This research was supported by Science Foundation Ireland (grants 06/RFP/CHP035, 07/RFP/CHEF437, 10/RFP/CAP2915). We thank the EU FP7 for funded access to the Central Laser Facility, STFC (App81030).

Notes and references

1 (*a*) K. Gehring, J. L. Leroy and M. Guéron, *Nature*, 1993, **363**, 561; (*b*) J. L. Leroy, M. Guéron, J. L. Mergny and C. Hélène, *Nucleic Acids Res.*, 1994, **22**, 1600.
2 (*a*) D.-L. Ma, H.-Z. He, D. S.-H. Chana and C.-H. Leung, *Chem. Sci.*, 2013, **4**, 336; (*b*) D. Zikich, K. Liu, L. Sagiv, D. Porath and A. Kotlyar, *Small*, 2011, **7**, 1029.
3 P. Catasti, X. Chen, L. L. Deaven, R. K. Moyzis, E. Bradbury and G. Gupta, *J. Mol. Biol.*, 1997, **272**, 369.
4 (*a*) K. Guo, A. Pourpak, K. Beetz-Rogers, V. Gokhale, D. Sun and L. H. Hurley, *J. Am. Chem. Soc.*, 2007, **129**, 10220; (*b*) T. A. Brooks, S. Kendrick and L. Hurley, *FEBS J.*, 2010, **277**, 3459.
5 J. A. Brazier, A. Shah and G. D. Brown, *Chem. Commun.*, 2012, **48**, 10739.
6 T. Douki and J. Cadet, *Biochemistry*, 2001, **40**, 2495.
7 B. Cohen, M. H. Larson and B. Kohler, *Chem. Phys.*, 2008, **350**, 165.
8 P. M. Hare, C. E. Crespo-Hernández and B. Kohler, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 435.
9 P. M. Keane, M. Wojdyla, G. W. Doorley, G. W. Watson, I. P. Clark, G. M. Greetham, A. W. Parker, J. M. Kelly and S. J. Quinn, *J. Am. Chem. Soc.*, 2011, **133**, 4212.
10 S. Quinn, G. W. Doorley, G. W. Watson, A. J. Cowan, M. W. George, A. W. Parker, K. L. Ronayne, M. Towrie and J. M. Kelly, *Chem. Commun.*, 2007, 2130.
11 J.-M. L. Pecourt, J. Peon and B. Kohler, *J. Am. Chem. Soc.*, 2000, **122**, 9348.
12 P. M. Keane, M. Wojdyla, G. W. Doorley, J. M. Kelly, I. P. Clark, A. W. Parker, G. M. Greetham, L. M. Magno and S. J. Quinn, *Phys. Chem. Chem. Phys.*, 2012, **14**, 6307.
13 D. A. McGovern, S. Quinn, G. W. Doorley, A. M. Whelan, K. L. Ronayne, M. Towrie, A. W. Parker and J. M. Kelly, *Chem. Commun.*, 2007, 5158.
14 A. I. S. Holm, L. M. Nielsen, B. Kohler, S. V. Hoffmann and S. B. Nielsen, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3426.
15 (*a*) E. Taillandier and J. Liquier, *Handb. Vib. Spectrosc.*, 2002, 3465–3480; (*b*) D. Tsankov, M. Krasteva, V. Andrushchenko, J. H. van de Sande and H. Wieser, *Biophys. Chem.*, 2006, **119**, 1.
16 M. Towrie, G. W. Doorley, M. W. George, A. W. Parker, S. J. Quinn and J. M. Kelly, *Analyst*, 2009, **134**, 1265.
17 I. Berger, M. Egli and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 12116.
18 (*a*) N. K. Schwalb and F. Temps, *J. Photochem. Photobiol., A*, 2009, **208**, 164; (*b*) N. K. Schwalb, T. Michalak and F. Temps, *J. Phys. Chem. B*, 2009, **113**, 16365.
19 A. L. Sobolewski and W. Domcke, *Phys. Chem. Chem. Phys.*, 2004, **6**, 2763.
20 (*a*) C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández and B. Kohler, *Annu. Rev. Phys. Chem.*, 2009, **60**, 217; (*b*) W. Doorley, M. Wojdyla, G. W. Watson, M. Towrie, A. W. Parker, J. M. Kelly and S. J. Quinn, *J. Phys. Chem. Lett.*, 2013, **4**, 2739.
21 K. Kawai, M. Fujitsuka and T. Majima, *Chem. Commun.*, 2005, 1476.

