COMMUNICATION

The importance of loop length on the stability of i-motif structures†

Sarah P. Gurung,[ab] Christine Schwarz,[a] James P. Hall,[ab] Christine J. Cardin[a] and John A. Brazier*[c]

Using UV and srCD spectroscopy it is found that loop length within the i-motif structure is important for both thermal and pH stability, but in contrast to previous statements, it is the shorter loops that exhibit the highest stability.

The i-motif is a quadruplex structure that is formed by cytosine-rich sequences of DNA (Fig. 1a) and has been found to form in several sequences complementary to G-quadruplexes.[1–3] The role of the i-motif in modulating gene expression has recently been elucidated, with the contribution of the loop regions within the structure key to the recognition and binding of transcriptional proteins.[4,5] The i-motif is not only of interest for its biological role in gene expression, especially due to its presence in the promoter regions of several genes associated with cancer, but also for its use as a switch within functional nanodevices.[6] The majority of studies using an i-motif based nanoswitch have used the human telomeric sequence (CCCTAA)$_3$CCC or a subtle variant. It is

known from research on biologically derived sequences, that i-motifs formed by different sequences exhibit different thermal and pH stability.[3] Recent research focussed on the design of a pH sensor based on i-motif structures has shown that changes in the sequence composition can change the pH sensitivity in a controlled way.[7] It has been reported in the literature that i-motif forming sequences containing longer loops (N = 5–8, Fig. 1b) are thermally more stable than those with short loops, with the length of the loop often suggested to be the contributing factor for this stability.[8] This is in contrast to research carried out on the loop length of G-quadruplex forming sequences, which showed that longer loops were less stable.[9] Recently we published results that showed unusually high stability for a sequence containing short loops (N = 1–4, Fig. 1b) which questioned the need for long loops to impart stability.[3] To probe this question further, we report here a systematic study of loop length *versus* both, thermal and pH stability. A series of i-motif forming sequences (Table 1), containing an increasing number of thymine bases within the loop region, with the general sequence CCCT(3-8)CCCT(3-8)CCCT(3-8)CCC,
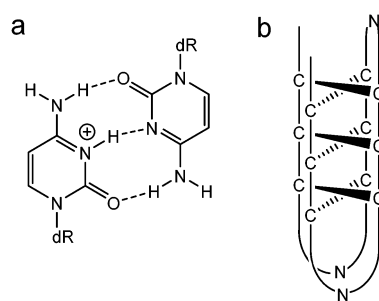

Fig. 1 (a) Protonated-cytosine: cytosine base pair. (b) Schematic representation of an i-motif structure, where N denotes any DNA heterocyclic base.

[a] *Department of Chemistry, University of Reading, Whiteknights, Reading, Berks, RG6 6AD, UK*

[b] *Diamond Light Source, Chilton, Didcot, Oxfordshire, UK*

[c] *School of Pharmacy, University of Reading, Whiteknights, Reading, Berks, RG6 6AD, UK. E-mail: j.a.brazier@reading.ac.uk*

† Electronic supplementary information (ESI) available: Experimental details, UV melting curves at 295 nm, and CD spectra at pH 5 and 8. See DOI: 10.1039/c4cc07279k
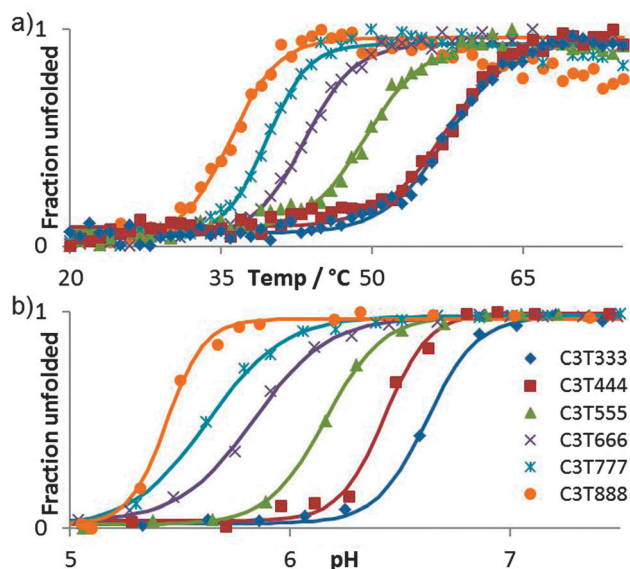
Table 1 Sequences used in this study, with associated T_m and T_{pH} values

Name	Sequence 5′-3′	$T_m/°C$	T_{pH}
C3T333	d(CCC**TTT**)$_3$CCCT	57	6.6
C3T444	d(CCC**TTTT**)3CCCT	57	6.4
C3T555	d(CCC**TTTTT**)$_3$CCCT	49	6.2
C3T666	d(CCC**TTTTTT**)$_3$CCCT	43	5.8
C3T777	d(CCC**TTTTTTT**)$_3$CCCT	39.5	5.6
C3T888	d(CCC**TTTTTTTT**)$_3$CCCT	36	5.4
C3T338	d(CCC**TTT**CCC**TTT**CCC**TTTTTTTT**CCC)	50	6.5
C3T383	d(CCC**TTT**CCC**TTTTTTTT**CCC**TTT**CCC)	55	6.6
C3T833	d(CCC**TTTTTTTT**CCC**TTT**CCC**TTT**CCC)	51	6.5
C3T883	d(CCC**TTTTTTTT**CCC**TTTTTTTT**CCC**TTT**CCC)	47	6.2
C3T838	d(CCC**TTTTTTTT**CCC**TTT**CCC**TTTTTTTT**CCC)	40	6.1
C3T388	d(CCC**TTT**CCC**TTTTTTTT**CCC**TTTTTTTT**CCC)	46	6.1

Fig. 2 Normalized absorbance at 260 nm *versus* (a) temperature, (b) pH.

Fig. 3 srCD spectra (molar ellipticity *vs.* wavelength (nm)) at (a) pH 5, (b) pH 6.5, (c) pH 7.5–8.

were assessed for their thermal (Fig. 2a) and pH stability (Fig. 2b) using both UV and synchrotron CD spectroscopy. By confining the loop bases to thymine, the possibility of interaction between the bases was limited to thymine: thymine and were consistent throughout the systematic series of oligonucleotides. The results of the thermal stability experiments show that C3T333 has the highest stability and C3T888 has the lowest stability. The other sequences in the series follow a predominately linear trend with respect to thermal stability with a decreasing order of C3T555, C3T666 and C3T777. The only exception is C3T444 which has the same thermal stability as C3T333. The results clearly show that the stability of the i-motif structure decreases as the length of the loop increases, a finding which is in contrast with statements made in the literature. One rationale for this observation, is as the loop length increases, the structure gains more flexibility, and therefore stabilising interactions become weaker. This decrease in stability is also present in the results from the pH titration experiments, where the transitional pH (T_{pH}, the pH at which half of the i-motif is unfolded) decreases as loop length increases. Once again C3T333 shows the highest T_{pH}, with C3T888 demonstrating the lowest value. Other i-motif forming sequences show a systematic decrease between these two values, with an order of C3T444, C3T555, C3T666 and C3T777. The pH sensitivity was also confirmed by srCD,‡ the spectra in Fig. 3a measured at pH 5 show the characteristic maximum at 285 nm and minimum at 265 nm associated with i-motif formation for each sequence.[10] The data in Fig. 3b measured at pH 6.5 shows these characteristic peaks only for C3T333 and C3T444 consistent with the T_{pH} values measured using UV absorbance, as at pH 6.5 only the C3T333 and C3T444 sequences are expected to show i-motif formation. Data in Fig. 3c measured at pH 7.5–8 lack all the characteristic signals for an i-motif structure, confirming the absence of i-motif formation under basic conditions for these sequences. These results suggest that flexibility within loop regions is detrimental to the stability of the i-motif structure, and long loop regions are only stabilising if they
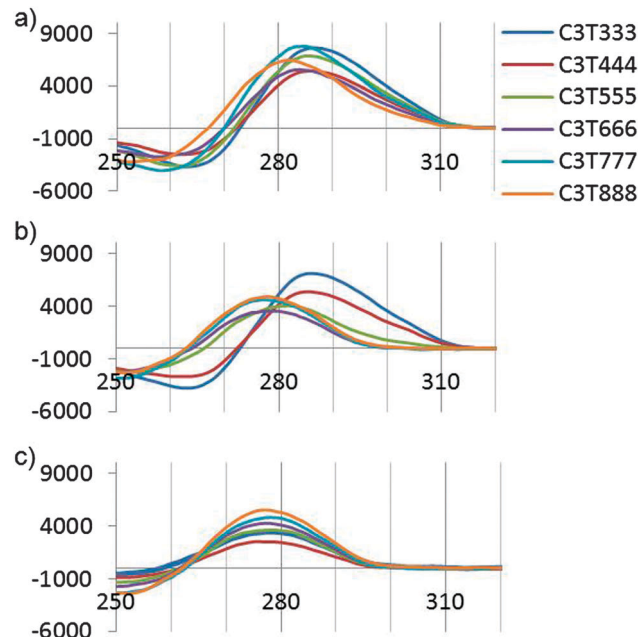
can form additional intramolecular interactions that limit flexibility, an explanation suggested for the high stability of the BCl2 promoter i-motif, which contains long loop with a loop pattern of 8-5-7.[8] Without the proposed stabilising interactions, this sequence would not be expected to have the high stability that it does. To further test this hypothesis we introduced an AA modification within a single T8 loop in the C3T838 i-motif to give a loop composition of TTTAATTT. It was expected that this would allow for a AT base pair, or other stabilising interaction to occur between the bottom two loops, and UV melting experiments show that this is the case, with an increase in T_m of 1 °C (see Fig. S4, ESI†) irrespective of which loop contains the AA bases. One further consideration is the ionic strength of the buffer solution, with high levels of salt known to destabilise i-motif structures.[11] Experiments with high levels of salt present showed a small decrease in stability with regard to the short loop C3T333 sequence (Fig. S6, ESI†), but did provide a stabilising effect with the C3T888 sequence (Fig. S7, ESI†). This again suggests that stabilising interaction within long loop regions, in this case, formation of salt bridges or ordering of water between the long loops by increased ionic components has led to a stabilisation of the loop structure, and ultimately stabilisation of the i-motif structure. These results confirm that the bases in the loop region are an important consideration when determining the stability of the structure of an i-motif forming sequence.[12] To elucidate the contribution from each loop region, a further series of sequences were investigated, in which the length of each loop was systematically altered between 3 and 8 thymine bases. It is clear from Fig. 4a that introducing a single long loop into a short loop i-motif sequence, has a minimal effect on the pH stability of the structure, with C3T833, C3T383 and C3T338 showing similar T_{pH} values to the parent sequence C3T333. In contrast, introducing a single short loop into a long loop sequence, results in a significant increase in the T_{pH} values, between C3T888 and the C3T833, C3T838
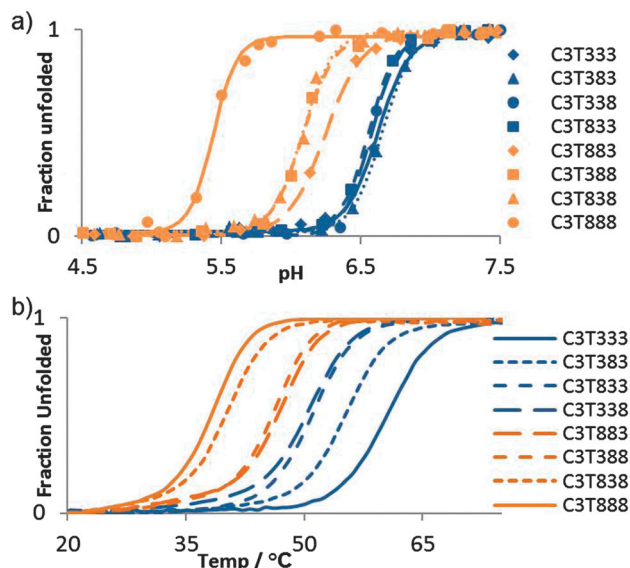
Fig. 4 Normalised absorbance at 260 nm *versus* (a) pH and (b) temperature.

and C3T388 series, suggesting that the short loop could be acting to orient the intercalating strands together to enhance stability. The thermal stability of the mixed loop structures is far more striking, with clear changes in stability in a loop dependent manner. Alterations in length of the central loop leads to the most significant changes in stability, with C3T383 showing the lowest decrease in stability, and C3T838 showing the largest decrease. The structures in which changes are made to the first and last looping region show similar stability irrespective of which loop is changed, although structures with two long loops are always less stable than those with two short loops. These results demonstrate that the important loops for stability must be the first and last loop, with the middle loop able to accommodate a longer sequence with minimal disruption to the stability. It has recently been shown that a small molecule can bind to the middle loop and stabilise the i-motif structure in the BCl2 promoter.[5] This contains a loop pattern of 8, 5 and 7. The presence of a longer sequence in the middle loop will therefore be beneficial in terms of small molecule binding, as it may provide a unique site to distinguish between different loop sequences. It is also clear that if an i-motif forming sequence contains long sequences in the first or last loop region that it must have the ability to form additional interactions for it to form a more stable i-motif structure, such as base stacking, hydrogen bonding, or both. The difference in both thermal and pH stability as a function of loop length also provides the opportunity to tune the temperature and pH at which

the i-motif structure undergoes conformational change and may be useful in developing more sensitive i-motif based switches for nano devices. A recent report has shown that tuning the sequence using different length cytosine tract and different sequences of 3–4 bases in the loop region allows switching over a well-defined pH range between pH 7 and 6.5.[7] Using different length loops would allow this pH range to extend to pH 5.5, and introduce the additional parameter of heat responsive conformational change. Altering loop length would also have the added advantage of changing the distance the terminal ends of the switch can travel.

These results demonstrate that i-motif stability is clearly related to the sequence composition. The length of the loop region is a large contributing factor to the overall stability of the i-motif structure, but in contrast to the widely held belief, it is not long loops that give the most stable structures, but short loops. In addition, the central loop is less important in providing stability to the i-motif structure, with the first and last loop shown to influence stability more. These findings suggest that in biologically relevant sequences with long loops the sequence composition of the loops must be able to form interactions that reduce the influence of loop length and effectively mimic short loop sequences and the stability they impart. The thermal and pH responsiveness of different loop lengths may also be used to tune the responsive nature of molecular switches based on the i-motif structural transition.

Notes and references

1 K. Gehring, J. L. Leroy and M. Gueron, *Nature*, 1993, **363**, 561–565.
2 T. A. Brooks, S. Kendrick and L. Hurley, *FEBS J.*, 2010, **277**, 3459–3469.
3 J. A. Brazier, A. Shah and G. D. Brown, *Chem. Commun.*, 2012, **48**, 10739–10741.
4 H.-J. Kang, S. Kendrick, S. M. Hecht and L. H. Hurley, *J. Am. Chem. Soc.*, 2014, **136**, 4172–4185.
5 S. Kendrick, H.-J. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht and L. H. Hurley, *J. Am. Chem. Soc.*, 2014, **136**, 4161–4171.
6 Y. Dong, Z. Yang and D. Liu, *Acc. Chem. Res.*, 2014, **47**, 1853–1860.
7 I. V. Nesterova and E. E. Nesterov, *J. Am. Chem. Soc.*, 2014, **136**, 8843–8846.
8 S. Kendrick, Y. Akiyama, S. M. Hecht and L. H. Hurley, *J. Am. Chem. Soc.*, 2009, **131**, 17667–17676.
9 A. Guédin, P. Alberti and J.-L. Mergny, *Nucleic Acids Res.*, 2009, **37**, 5559–5567.
10 G. Manzini, N. Yathindra and L. E. Xodo, *Nucleic Acids Res.*, 1994, **22**, 4634–4640.
11 J. L. Mergny, L. Lacroix, X. G. Han, J. L. Leroy and C. Helene, *J. Am. Chem. Soc.*, 1995, **117**, 8887–8898.
12 Y. Yang, Y. Sun, Y. Yang, Y. Xing, T. Zhang, Z. Wang, Z. Yang and D. Liu, *Macromolecules*, 2012, **45**, 2643–2647.