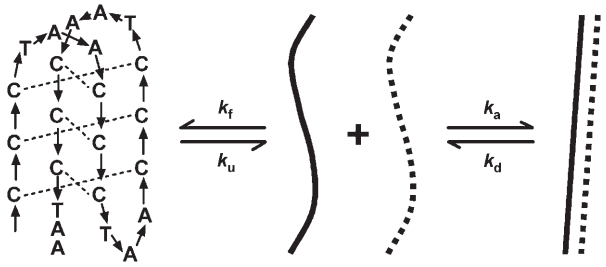
DOI: 10.1002/cbic.200500175

The Folding and Unfolding Kinetics of the i-Motif Structure Formed by the C-Rich Strand of Human Telomere DNA

Yong Zhao,[a] Zhi-xiong Zeng,[a] Zhong-yuan Kan,[a] Yu-hua Hao,[a] and Zheng Tan*[a, b]

Nucleic acids with tandem cytidine-rich repetitive stretches can fold into a four-stranded intramolecular structure, called an i-motif (Scheme 1), through intercalated C–C$^+$ base pairing at



Scheme 1. Formation of the i-motif structure through intercalated cytidine–cytidine base pairing in the C-rich strand of human telomere DNA. The unfolded C-rich strand also forms duplexes in the presence of a complementary strand.

acidic pH.[1–6] These sequences are widely dispersed in genomic DNA,[7] and proteins have been found that interact selectively with such sequences, possibly in their i-motif form;[8–12] this suggests that these sequences and the relevant i-motif structures might be involved in biological processes.[13] It has been reported, for example, that the i-motif upstream of the insulin gene might facilitate transcription.[14] Because of the biological implications, such nonduplex structures are emerging as a new class of therapeutic targets for cancers and other diseases.[15] The opening and closing of i-motifs, driven by protons, has also been proposed as a molecular nanomachinery.[16]

DNA in a cell is normally bound by proteins and exposed only under certain particular situations and for limited periods. As the most-studied i-motif-forming sequence, the C-rich strand of telomere DNA was recently reported to transiently exist as a 5′ single-stranded overhang at the chromosome ends in the S-phase of replicating human cells.[17] The rates of folding and unfolding are crucial in determining the opportunity or possibility of the formation of i-motif under in vivo situations within the narrow "time window", during which the DNA is transiently liberated. The folding- and unfolding-rate constants provide a quantitative description of the folding and unfolding processes and deliver information on 1) how quickly each structure forms, 2) for how long each structure is likely to exist once formed, and 3) the fraction of each structure at equilibrium. They also partly define the competition of i-motif formation with duplex formation in the presence of complementary strands. Therefore, the measurement of these kinetic parameters should provide quantitative information for the above processes for a better understanding of the biological functions of the i-motif structure. On the other hand, the folding and unfolding kinetics also determine how fast an i-motif-based DNA nanomachine[16] can operate.

While many previous studies on i-motifs have provided detailed information on their thermodynamics and structure,[2,18–25] few kinetic data have been reported for these structures.[26] In this work, we report the measurement of the folding- and unfolding-rate constants over a wide pH range for the i-motif structure formed by $(CCCTAA)_4$, which represent the C-rich strand of human and other vertebrate telomeres, using an optical biosensor based on surface plasmon resonance (SPR) by a method we recently developed.[27]

Formation of an i-motif is promoted by protons.[13] To explore how the folding and unfolding of $(CCCTAA)_4$ is affected by pH, measurements were carried out at pH 4.8–7.0. Figure 1 shows two representative sensorgrams obtained at two different pH values, and their fitting to the coupled-hybridization model (see Experimental Section). A summary of the kinetic parameters is provided in Table 1 and Figure 2. They show that both the folding- and unfolding-rate constants (k_f and k_u) are

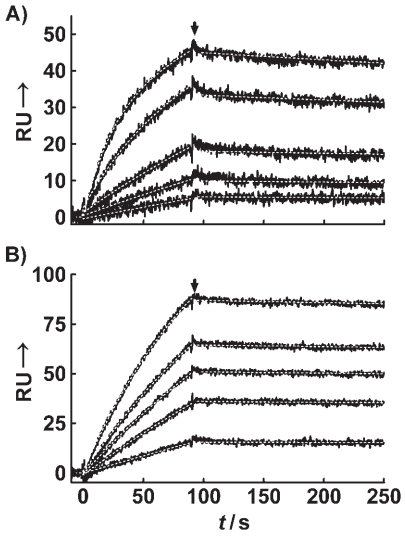


Figure 1. Representative sensorgrams (black lines) and their global fittings (dotted white lines) for the coupled-hybridization model. A) Measurement conducted at pH 5.3, $(TTAGGG)_4$ was injected at 1, 2, 4, 8, and 16 nM. B) Measurement conducted at pH 7.0, $(TTAGGG)_4$ was injected at 1, 2, 3, 4, and 6 nM. Arrows indicate the start of the dissociation phase when the sensing surface was washed with running buffer.

[a] *Dr. Y. Zhao, Z.-x. Zeng, Z.-y. Kan, Y.-h. Hao, Dr. Z. Tan*
Laboratory of Biochemistry and Biophysics, College of Life Sciences
Wuhan University, Wuhan 430072 (P. R. China)
Fax: (+86) 27-6875-6661
E-mail: tanclswu@public.wh.hb.cn

[b] *Dr. Z. Tan*
Institute of Zoology, Chinese Academy of Sciences
Beijing 100080 (P. R. China)
Fax: (+86) 10-6265-9958

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

Table 1. Kinetic rate, equilibrium constants, and half-lives of $(CCCTAA)_4$ obtained at different pH values.[a]

pH	k_f [s^{-1}]	k_u [s^{-1}]	K_F	$t_{f1/2}$ [s]	$t_{u1/2}$ [s]	k_a [$M^{-1}s^{-1}$]	k_d [s^{-1}]	K_A [M^{-1}]
4.8	1.6×10^{-2} (2.6×10^{-3})	2.0×10^{-3} (4.0×10^{-4})	8.2	347	43.3	2.5×10^{6} (8.2×10^{5})	2.7×10^{-3} (3.0×10^{-4})	9.2×10^{8}
5.0	9.3×10^{-3} (9.7×10^{-4})	1.7×10^{-3} (1.9×10^{-4})	5.4	408	74.5	2.5×10^{6} (3.8×10^{5})	1.6×10^{-3} (2.8×10^{-4})	1.5×10^{9}
5.3	7.6×10^{-3} (7.3×10^{-4})	2.3×10^{-3} (9.6×10^{-5})	3.4	301	91.2	2.6×10^{6} (1.7×10^{5})	1.1×10^{-3} (3.3×10^{-4})	2.4×10^{9}
5.6	6.0×10^{-3} (2.7×10^{-4})	2.9×10^{-3} (2.9×10^{-4})	2.1	239	116	3.5×10^{6} (4.7×10^{5})	7.2×10^{-4} (7.5×10^{-5})	4.8×10^{9}
6.0	3.5×10^{-3} (7.6×10^{-4})	8.4×10^{-3} (4.9×10^{-4})	4.2×10^{-1}	82.5	198	3.8×10^{6} (4.7×10^{5})	5.4×10^{-4} (1.5×10^{-4})	7.0×10^{9}
6.5	2.7×10^{-3} (2.0×10^{-4})	1.0×10^{-2} (2.1×10^{-3})	2.8×10^{-1}	69.3	257	3.6×10^{6} (1.2×10^{5})	4.4×10^{-4} (1.4×10^{-4})	8.2×10^{9}
7.0	1.9×10^{-5} (1.3×10^{-5})	1.3×10^{-2} (7.5×10^{-4})	1.4×10^{-3}	53.3	36500	3.0×10^{6} (5.5×10^{5})	1.5×10^{-4} (6.8×10^{-5})	2.0×10^{10}

[a] The rate constants at each pH are the mean of 3–4 measurements with standard error (in parenthesis). K_F is the folding equilibrium constant of the i-motif given by k_f/k_u, and K_A is the association equilibrium constant of the duplex given by k_a/k_d. The half-lives, $t_{f1/2}$ for the folded and $t_{u1/2}$ for the unfolded forms, were calculated by using $t_{f1/2}=\ln 2/k_u$ and $t_{u1/2}=\ln 2/k_f$, respectively.

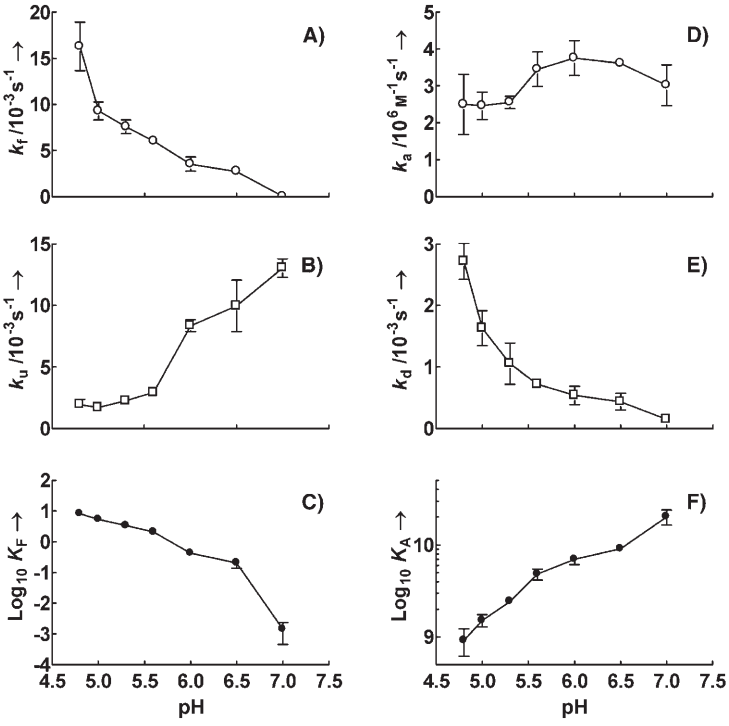


Figure 2. Kinetic rate and equilibrium constants of the immobilized $(CCCTAA)_4$ as a function of pH. The data are from Table 1.

strongly dependent on pH. The promotion of i-motif formation by protons is achieved by a combination of increased k_f (Figure 2A) and decreased k_u (Figure 2B), which result in a rapid increase in the folding equilibrium constant, K_F, with decreasing pH (Figure 2C). These results are, in general, in agreement with the fact that both the formation[25,28] and thermal stability[2,21] of the i-motif increase with decreasing pH. To access the stability of the folded and unfolded form of the $(CCCTAA)_4$, their half-lives were derived from k_u and k_f, respectively (Table 1). From these values, we can see that both the folding and unfolding of $(CCCTAA)_4$ occur on the timescale of subminutes or minutes at acidic pH. These data agree with the value in a previous report of about 5 s for a similar sequence in solution,[16] if taking into account the effect of immobilization on the kinetics.[27] In an earlier work, the lifetime of a similar sequence was measured by NMR at pH 5.0 within the temperature range 50–70 °C.[29] By extrapolation, the lifetime of this sequence at 25 °C would be roughly 10 times longer than the value we obtained (thousands vs. 408 s). Our previous work has shown that the quadruplex formed by an immobilized telomere sequence is less stable than the same structure in the solution phase.[27] Comparison of our data with published results[16,29] shows this is also true for the immobilized i-motif. In agreement with this, immobilized $(CCCTAA)_4$ was found to have a lower half-transition pH than $(CCCTAA)_4$ in solution. When the fraction of the unfolded form of the immobilized $(CCCTAA)_4$ at equilibrium was deduced from the k_f and k_u as a function of pH, it showed a half-transition pH at 5.8 (Figure 3); this agrees with the value in a previous study in which DNA-binding protein rather than $(TTAGGG)_4$ was used as probe in the mobile phase.[30] The conformational transition of the i-motif in solution can be followed by absorbance at 260 nm.[31] By this method, $(CCCTAA)_4$ in solution was found to have a half-transition pH at 6.3 (Figure 3), which is close to the reported value of 6.2 for the sequence $(CCCTAA)_3CCCT$.[25]

The measurements also gave the association and dissociation rate constants (k_a and k_d) for the duplex formed between $(CCCTAA)_4$ and $(TTAGGG)_4$ (Table 1

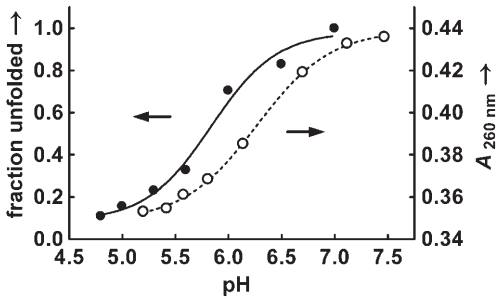


Figure 3. Structural transition of immobilized and free $(CCCTAA)_4$ in solution induced by changes in pH. The fraction of unfolded species for the immobilized oligonucleotide (●) was calculated as $k_u/(k_f+k_u)$. Changes in absorbance were recorded for the free nucleotide in solution (○) when titrated with acetic acid. Half-transition pH values (5.8 for the immobilized and 6.3 for the nucleotide in solution) were obtained by fitting the corresponding data to the Sigmoidal dose–response function.

and Figure 2D–F). Literature values for the k_a of complementary oligonucleotides range from 4×10^5 to $2\times10^7\,\text{M}^{-1}\text{s}^{-1}$.[32] The measured value of this constant for the $(CCCTAA)_4$ is around $10^6\,\text{M}^{-1}\text{s}^{-1}$, which is within the reported range. Our data also show that k_a is virtually independent of pH (Figure 2D), while k_d increases rapidly with decreasing pH (Figure 2E). These results are also in agreement with the reports that the DNA association rate is essentially independent of pH within the range 5–9,[33] while the stability of the DNA duplex decreases with decreasing pH;[34] this implies an increased dissociation with decreasing pH.

Many proteins have been found that specifically bind the C-rich telomere strand. Among them, several have an explicit requirement for at least four consecutive repeats,[8–12] which are necessary for the formation of an i-motif. The presence of such proteins suggests that the i-motif structure might be present in vivo. Animal cells maintain an intracellular pH ranging from 7.0–7.5 under normal physiological conditions. The extremely long half-life (>10 h, Table 1) of the unfolded $(CCCTAA)_4$ at pH 7.0 indicates that the formation of an i-motif structure would be very difficult under physiological pH. Relatively fast folding and unfolding rates are necessary for i-motifs to function as a regulatory element. Our data suggest that unassisted folding/unfolding of i-motifs is too slow to be a regulatory element at physiological pH.

In the genome, DNA is present with a complementary strand. In this case, the formation of an i-motif, if possible, will have to compete with the formation of a duplex. Our study provides insight into how pH influences these processes for $(CCCTAA)_4$. It reveals that acidic pH favors increased i-motif formation over that of the duplex (Figure 2). As is seen in Table 1, decreasing the pH from 7.0 to 4.8 increased the folding equilibrium constant, K_F, by over 5800-fold and decreased the association equilibrium constant, K_A, by over 20-fold.

Our data in this and a previous[27] work show that immobilized oligonucleotides have different kinetic properties from the free oligonucleotide in solution. This difference might be caused by the microenvironment created by the carboxymethylated dextran polymer matrix at the sensor-chip surface. It has been reported that the melting temperature of the quadruplex formed by the G-rich strand of human telomere DNA can be reduced by 9 °C when fluorescent dye is attached to the termini of the sequence.[35] Therefore, it is also possible that the physical restraint imposed by immobilization could contribute to the altered kinetics. If this is true, immobilized nucleotide might then better mimic the following situations. The C-rich telomere strand was reported to exist transiently as a 5′ single-stranded overhang at the chromosome ends in the S-phase of replicating human cells.[17] Such an overhang can be regarded as attached at one end to the hosting chromosome. Recently, the construction of a DNA nanomotor based on the folding and unfolding of oligonucleotides has been proposed, and the operation of these devices has been realized by using free molecules in solution under no-load conditions.[16,36] As driving device, a DNA nanomotor could need supporting and has to be interfaced or connected with other parts in order to function. Our data suggest that such engagement is likely to affect the operation kinetics of a DNA nanomotor.

Experimental Section

Measurements of kinetic parameters: Measurements were carried out on a BIAcore X optical biosensor (BIAcore, Switzerland) as previously described.[27] Synthetic $(CCCTAA)_4$ was labeled with biotin at the 5′ end though a C_{12} spacer and immobilized on a CM5 sensor chip by a biotin–streptavidin interaction. Each sensorgram recording was started with a brief pulse of NaOH (5 μL, 20 mM) followed by equilibration for 1200 s with running buffer containing Tris-acetate (10 mM), EDTA (1 mM), and LiCl (150 mM) at the indicated pH. Then $(TTAGGG)_4$ in running buffer was injected across the sensor-chip surface for 90 s, followed by a flow of running buffer for 250–300 s. All measurements were carried out at 25 °C with a flow rate of $30\,\mu\text{L}\,\text{min}^{-1}$ with a blank cell as reference. At each specified pH, sensorgrams from five injections of $(TTAGGG)_4$ at different concentrations were recorded.

In the presence of $(TTAGGG)_4$, the following reactions are assumed to occur on the sensing surface.

$$Q \underset{}{\overset{k_u}{\rightleftharpoons}} R \qquad (1)$$

$$R + C \underset{k_d}{\overset{k_f \quad k_a}{\rightleftharpoons}} D \qquad (2)$$

here Q and R denote the immobilized $(CCCTAA)_4$ in the folded and unfolded forms, C the $(TTAGGG)_4$ injected, and D the duplex formed between R and C (Scheme 1). This coupled hybridization model assumes that the $(TTAGGG)_4$ in the mobile phase only interacts with the unfolded, but not the folded form of $(CCCTAA)_4$. Li^+-containing buffer was used to prevent $(TTAGGG)_4$ from forming a quadruplex (see Supporting Information). Global fittings of sensorgrams to the above model were carried out to obtain the kinetic constants by using the BIAevaluation 3.0 software supplied by the manufacturer of BIAcore. The k_d value was extracted from the dissociation phase, and then k_f, k_u, and k_a were extracted from the association phase, as previously described.[27]

i-Motif formation in solution: The pH dependence of i-motif formation in solution was determined by titrating $(CCCTAA)_4$ in Tris-HCl buffer (10 mM, pH 7.5, EDTA (1 mM), LiCl (150 mM)) with acetic acid while the absorbance was recorded on a Beckman DU-640 UV-VIS spectrophotometer at 25 °C. The increase in volume by addition of acetic acid was less than 1%.

Acknowledgements

This work was supported by grant 30270314 from the NSFC and G2000057001 from the MSTC.

Keywords: folding · i-motif · kinetics · nucleic acids · secondary structure · telomeres

[1] K. Gehring, J. L. Leroy, M. Guéron, *Nature* **1993**, *363*, 561–565.
[2] J. L. Leroy, M. Guéron, J. L. Mergny, C. Hélène, *Nucleic Acids Res.* **1994**, *22*, 1600–1606.
[3] C. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, A. Rich, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3874–3878.
[4] K. Snoussi, S. Nonin-Lecomte, J. L. Leroy, *J. Mol. Biol.* **2001**, *309*, 139–153.
[5] I. Berger, C. Kang, A. Fredian, R. Ratliff, R. Moyzis, A. Rich, *Nat. Struct. Biol.* **1995**, *2*, 416–425.

[6] L. Cai, L. Chen, S. Raghavan, R. Ratliff, R. Moyzis, A. Rich, *Nucleic Acids Res.* **1998**, *26*, 4696 – 4705.
[7] P. Catasti, X. Chen, S. V. Mariappan, E. M. Bradbury, G. Gupta, *Genetica* **1999**, *106*, 15 – 36.
[8] E. Marsich, A. Piccini, L. E. Xodo, G. Manzini, *Nucleic Acids Res.* **1996**, *24*, 4029 – 4033.
[9] E. Marsich, L. E. Xodo, G. Manzini, *Eur. J. Biochem.* **1998**, *258*, 93 – 99.
[10] L. Lacroix, H. Lienard, E. Labourier, M. Djavaheri-Mergny, J. Lacoste, H. Leffers, J. Tazi, C. Hélène, J. L. Mergny, *Nucleic Acids Res.* **2000**, *28*, 1564 – 1575.
[11] J. F. Cornuel, A. Moraillon, M. Guéron, *Biochimie* **2002**, *84*, 279 – 289.
[12] E. Marsich, A. Bandiera, G. Tell, A. Scaloni, G. Manzini, *Eur. J. Biochem.* **2001**, *268*, 139 – 148.
[13] For a review, see: M. Guéron, J. L. Leroy, *Curr. Opin. Struct. Biol.* **2000**, *10*, 326 – 331.
[14] P. Catasti, X. Chen, L. L. Deaven, R. K. Moyzis, E. M. Bradbury, G. Gupta, *J. Mol. Biol.* **1997**, *272*, 369 – 382.
[15] L. H. Hurley, *Biochem. Soc. Trans.* **2001**, *29*, 692 – 696.
[16] D. Liu, S. Balasubramanian, *Angew. Chem.* **2003**, *115*, 5912 – 5914; *Angew. Chem. Int. Ed.* **2003**, *42*, 5734 – 5736.
[17] G. Cimino-Reale, E. Pascale, E. Alvino, G. Starace, E. D'Ambrosio, *J. Biol. Chem.* **2003**, *278*, 2136 – 2140.
[18] I. Berger, M. Egli, A. Rich, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12 116 – 12 121.
[19] K. Kanaori, N. Shibayama, K. Gohda, K. Tajima, K. Makino, *Nucleic Acids Res.* **2001**, *29*, 831 – 840.
[20] S. Ahmed, A. Kintanar, E. Henderson, *Nat. Struct. Biol.* **1994**, *1*, 83 – 88.
[21] J. L. Mergny, L. Lacroix, *Nucleic Acids Res.* **1998**, *26*, 4797 – 4803.
[22] J. L. Mergny, *Biochemistry* **1999**, *38*, 1573 – 1581.
[23] X. Han, J. L. Leroy, M. Guéron, *J. Mol. Biol.* **1998**, *278*, 949 – 965.
[24] J. L. Leroy, M. Guéron, *Structure* **1995**, *3*, 101 – 120.
[25] G. Manzini, N. Yathindra, L. E. Xodo, *Nucleic Acids Res.* **1994**, *22*, 4634 – 4640.
[26] M. Mills, L. Lacroix, P. B. Arimondo, J.-L. Leroy, J.-C. François, H. Klump, J.-L. Mergny, *Curr. Med. Chem. Anti-Cancer Agents* **2002**, *2*, 627 – 644.
[27] Y. Zhao, Z. Y. Kan, Z. X. Zeng, Y. H. Hao, H. Chen, Z. Tan, *J. Am. Chem. Soc.* **2004**, *126*, 13 255 – 13 264.
[28] S. Nonin-Lecomte, J. L. Leroy, *J. Mol. Biol.* **2001**, *309*, 491 – 506.
[29] A. T. Phan, M. Guéron, J.-L. Leroy, *J. Mol. Biol.* **2000**, *299*, 123 – 144.
[30] Z. X. Zeng, Y. Zhao, Y. H. Hao, Z. Tan, *J. Mol. Recognit.* **2005**, *18*, 267 – 271.
[31] S. Ahmed, E. Henderson, *Nucleic Acids Res.* **1992**, *20*, 507 – 511.
[32] S. Howorka, L. Movileanu, O. Braha, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12 996 – 13 001.
[33] J. G. Wetmur, N. Davidson, *J. Mol. Biol.* **1968**, *31*, 349 – 370.
[34] J. R. Hutton, *Nucleic Acids Res.* **1977**, *4*, 3537 – 3555.
[35] J. J. Green, L. Ying, D. Klenerman, S. Balasubramanian, *J. Am. Chem. Soc.* **2003**, *125*, 3763 – 3767.
[36] P. Alberti, J. L. Mergny, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1569 – 1573.

Received: April 24, 2005
Published online on October 5, 2005

