Consecutive GC base pairs determine the energy barrier of DNA duplex formation under molecularly crowded conditions†

Xiao-Bo Gu,[a] Shu-ichi Nakano[a] and Naoki Sugimoto*[ab]

Received (in Cambridge) 26th February 2007, Accepted 10th April 2007
First published as an Advance Article on the web 23rd April 2007
DOI: 10.1039/b702865b

The kinetics of DNA duplex formation was affected by the addition of PEGs with different masses (MW = 200–8000) to an aqueous solution; for each condition, two duplexes (5′-TAGGTTATAA-3′/5′-TTATAACCTA-3′ and 5′-CAGGTCACAG-3′/5′-CTGTGACCTG-3′) with different stabilities were formed after overcoming the same association activation energy barrier, suggesting that the formation of consecutive GC base pairs in the helices rather than the helix terminus is the initiation nucleus for DNA duplex formation not only in the absence, but also in the presence of PEGs.

Thermodynamics and kinetics are the fundamental aspects of DNA duplex formation. The former demonstrates the possibility of a conformational transition and the thermal stability of the duplex, which depends on the whole sequence, and can be predicted by the nearest-neighbor model.[1] On the other hand, the latter reflects the realizability and the mechanism of the transition. From a kinetic standpoint, it is known that DNA duplex formation does not belong to an elementary reaction, but follows the zip-up model in which the rate-determining step of duplex association is the formation of a nucleus of several base pairs, and the double helix can immediately "zip-up" after the nucleus formation.[2–4] The nucleation site is often regarded to be the nucleotide terminus, but it may also depend on the nucleotide sequence. Both the thermodynamic and kinetic aspects are essential for understanding the base pair formations, which are critical for many biological processes and technological applications, such as RNAi, antisense methods and the development of DNA chips.[5] Since these technologies are often aimed at in-cell use, quantitative information regarding both the thermodynamic and kinetic aspects for duplex formation under cell-like conditions are important. To mimic the molecular environment in living cells containing a number of biomolecules from giant proteins to small metabolites, poly(ethylene glycol) (PEG) is one of the most commonly used cosolutes added to aqueous solution,[6,7] because different molecular weight PEGs are available and PEGs are inert to nucleotides. We previously reported that the thermal stability of a short DNA duplex decreased with the addition of high concentrations of PEGs.[8,9] In this study, we investigated the kinetics of DNA duplex formation with different thermodynamic stabilities in the presence of various masses of PEGs (average molecular weight was 200–8000) to elucidate the kinetic aspects of DNA duplex formation under cell-like conditions.

The 10-mer DNA duplex strands[10] (**1** and **2** in Fig. 1) were designed so as not to form intramolecular secondary structures,[11] and there was a smaller number of the GC base pairs in duplex **1** than that in duplex **2**, leading to a significant difference in the thermodynamic stabilities of these duplexes. It is noted that both duplexes have one consecutive GC base pair near the middle of the strands, but have different base pairs at both ends. In all the solutions examined in this study including 1 M NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA at pH 7.0 in the absence and presence of 20 wt% of PEG 200,[12] PEG 600 or PEG 8000, the thermal stability of duplex **1** was found to be much lower than that of duplex **2**, as expected (Fig. S1, ESI†): At 50 μM DNA in the solution without the PEG, the melting temperature (T_m) of duplex **1** (37.1 °C) was 22.3 °C lower than that of duplex **2** (59.4 °C). In the solutions containing 20 wt% PEG 200, PEG 600 and PEG 8000, the T_m values of duplex **1** were 21.7, 20.6 and 22.6 °C lower than those of duplex **2**, respectively.

The DNA duplex formation can be represented as A + B $\rightleftharpoons$ AB, where A and B are the non-self complementary DNA single-strands and AB is the fully-matched DNA duplex. The rapid kinetics for nucleotide duplex formation in solution can be measured by stopped-flow and temperature-jump methods. For the stopped-flow experiment, two or more solutions are rapidly mixed, which often causes a problem when using the solutions containing a high concentration of cosolutes, because of a higher viscosity. Therefore, the temperature-jump method was used in this study to investigate the kinetics of DNA duplex formation. Fig. 2 shows the typical temperature-jump relaxation curve and the plot of τ^{-2} (τ, the relaxation time) *vs.* C_t (the total DNA concentration) of duplex **1** measured at 37 °C, thus providing the association (k_{+1}) and dissociation (k_{-1}) rate constants[4,11,13] (see ESI†). The k_{+1} and k_{-1} values of duplexes **1** and **2** at 37 °C in 1 M NaCl–phosphate buffer with and without 20 wt% PEGs are listed in Table 1. The values of k_{+1} (k_{-1}) for duplexes **1** and **2** in the solution without the PEG are 1.15×10^7 M^{-1} s^{-1} (128 s^{-1}) and 5.43×10^7 M^{-1} s^{-1} (0.170 s^{-1}), respectively, which are similar to

Duplex **1**:
5'-TAGGTTATAA-3'
3'-ATCCAATATT-5'

Duplex **2**:
5'-CAGGTCACAG-3'
3'-GTCCAGTGTC-5'

Fig. 1 The DNA duplex sequences.

[a]*Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe, 658-8501, Japan*
[b]*Department of Chemistry, Faculty of Science and Engineering, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe, 658-8501, Japan. E-mail: sugimoto@konan-u.ac.jp; Fax: (+81)78-435-2539; Tel: (+81)78-435-2497*

† Electronic supplementary information (ESI) available: Materials and methods. The melting curves. Plots of rate constants *vs.* viscosity. See DOI: 10.1039/b702865b

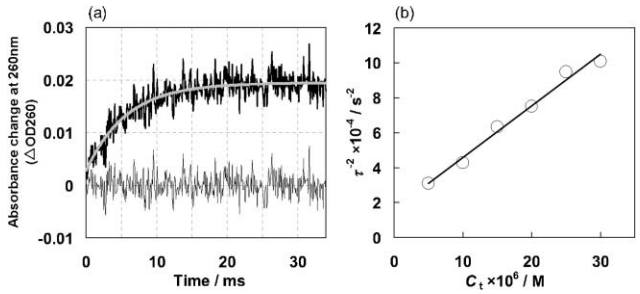

Fig. 2 (a) Temperature-jump kinetic trace (thick line) of a 10 μM 5′-TAGGTTATAA-3′/5′-TTATAACCTA-3′ (duplex **1**) at 37 °C in 1 M NaCl–phosphate buffer. The curve of a single exponential fitting (light grey smooth line) and the residuals of the fitting (thin line) are also indicated. (b) The τ^{-2} *vs.* C_t plot of duplex **1** at 37 °C in 1 M NaCl–phosphate buffer.

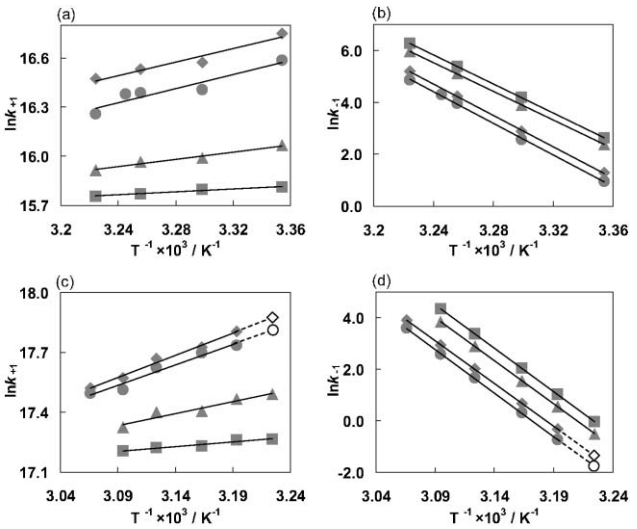

Fig. 3 Arrhenius plots of duplex **1** (5′-TAGGTTATAA-3′/5′-TTATAA CCTA-3′) (a and b) and duplex **2** (5′-CAGGTCACAG-3′/ 5′-CTGTGA CCTG-3′) (c and d) in the absence (circles) and in the presence of 20 wt% PEG 200 (squares), PEG 600 (triangles) or PEG 8000 (diamonds). The open symbols indicate the data at 37 °C extrapolated from the linear plots.

those of previous observations that the k_{+1} value for oligo-DNA duplex formation is typically around 10^5 to 10^7 M^{-1} s^{-1} and that the k_{-1} value largely depends on the DNA sequence.[14]

Table 1 indicates that the rate constants of the two duplexes are affected by the PEGs in solution. For both duplexes, k_{+1} decreased and k_{-1} increased in the presence of low mass cosolutes of PEG 200 and PEG 600, and the lower mass one (PEG200) was more effective. On the other hand, both k_{+1} and k_{-1} of duplexes **1** and **2** in the presence of PEG 8000 were not very different from those obtained in the solution without PEG. The kinetic properties of the protein associations and refoldings in the presence of cosolutes have been studied. Kozer and Schreiber reported that the association rate constant between β-lactamase and its inhibitor protein was influenced by the viscosity alteration when adding low mass PEGs, while the association rate constant in the presence of high mass PEGs only slightly changed even under a 12-fold higher viscosity than water, supposedly because of the reaction in the porous medium structure formed by large-sized PEGs.[15] In contrast to the protein results, the plots of k_{+1} (or k_{-1}) *vs.* the solution viscosity in this study showed that the viscosity alteration by adding low or high mass PEGs did not primarily determine the rate constants of the DNA duplexes formation (Fig. S2, ESI†).

To discern the reaction mechanism of the DNA duplexes formation in the absence and presence of the PEGs, the k_{+1} and k_{-1} of duplexes **1** and **2** were determined at different temperatures to develop the Arrhenius plots. The association (E_{a+1}) and dissociation (E_{a-1}) activation energies were derived from the slope of the Arrhenius plots (Fig. 3), and are listed in Table 1. The E_{a+1} obtained for each solution was negative, suggesting that even in the presence of the PEGs, the duplex formation follows the zip-up model, in which the rate-determining step is the formation of a nucleus consisting of several base pairs.[2–4,14] Similar to the tendency of the rate constant data, the E_{a+1} values increased more in the presence of the lower mass PEGs (Fig. S3, ESI† and the slopes of the Arrhenius plot in Fig. 3), and the E_{a+1} obtained in the presence of PEG 8000 was almost the same as that of the solution without the PEG. The results could be attributed to the difference in the solution properties. Because the solutions containing low mass molecules (PEG 200 and PEG 600) are supposed to be homogeneous,[16] the solution properties are changed throughout the buffer solution by the addition of PEG 200 or PEG 600. The increase in the E_{a+1} value due to the addition of PEG 200 or PEG 600 is consistent with previous reports that the low mass PEGs increase the ΔH° (the enthalpy change) for DNA duplex formation,[8] because E_{a+1} reflects the enthalpy energy before the initiation nucleus for DNA duplex formation.[14] In contrast, PEG 8000 can form flexible networks and produce porous media in a

Table 1 The kinetic parameters of the duplexes of 5′-TAGGTTATAA-3′/5′-TTATAACCTA-3′ (duplex **1**) and 5′-CAGGTCACAG-3′/ 5′-CTGTGACCTG-3′ (duplex **2**) in different solutions

Sequence	5′-TAGGTTATAA-3′/5′-TTATAACCTA-3′				5′-CAGGTCACAG-3′/5′-CTGTGACCTG-3′			
Condition	k_{+1} (37 °C)/ 10^7 M^{-1} s^{-1}	k_{-1} (37 °C)/s^{-1}	E_{a+1}/ kcal mol^{-1}	E_{a-1}/ kcal mol^{-1}	k_{+1} (37 °C)/ 10^7 M^{-1} s^{-1}	k_{-1} (37 °C)/ s^{-1}	E_{a+1}/ kcal mol^{-1}	E_{a-1}/ kcal mol^{-1}
1 M NaCl	1.15 ± 0.10	128 ± 8	−4.3 ± 0.4	60.7 ± 0.4	5.43[a]	0.170[a]	−4.1 ± 0.4	67.3 ± 0.4
1 M NaCl with 20 wt% PEG 8000	1.34 ± 0.09	168 ± 9	−4.1 ± 0.3	60.0 ± 0.3	5.79[a]	0.260[a]	−4.4 ± 0.2	66.1 ± 0.2
1 M NaCl with 20 wt% PEG 600	0.816 ± 0.053	391 ± 17	−2.2 ± 0.3	55.1 ± 0.3	3.96 ± 0.27	0.607 ± 0.043	−2.4 ± 0.3	66.7 ± 0.3
1 M NaCl with 20 wt% PEG 200	0.697 ± 0.041	532 ± 21	−0.89 ± 0.10	55.7 ± 0.1	3.16 ± 0.21	0.976 ± 0.058	−0.96 ± 0.13	67.1 ± 0.1

[a] The k_{+1} and k_{-1} values of duplex **2** in the absence of the PEGs and in the presence of PEG 8000 were calculated from the Arrhenius plots shown in Fig. 3(c) and (d) due to the very low ΔOD_{260} values at 37 °C in the solutions.

solution,[15,17] and the free space within the porous medium, in which the properties might be similar to that in the solution without the PEG, is large enough for the 10-mer DNA duplex,[15] explaining the kinetic parameters for the 20 wt% PEG 8000 being similar to those of the solution without PEG.

Next, we focused on the E_{a+1} and E_{a-1} values of duplexes **1** and **2** in the same solutions (Table 1). Significant differences in the E_{a-1} decrease due to the addition of the PEG between duplexes **1** and **2** were observed, while the E_{a+1} values of duplexes **1** and **2** were almost the same in the solutions with or without PEG. Because E_{a+1} reflects the enthalpy energy before the initiation nucleus for the DNA duplex formation,[14] the same E_{a+1} values of duplexes **1** and **2** suggest that the initiation events of these two duplexes are identical (Fig. S3, ESI†), although duplex **2** has terminal GC base pairs and duplex **1** has terminal AT base pairs. Therefore, the initiation nucleus for duplex **1** and **2** formation may exist in identical parts of the consecutive GC base pairs near the middle of the strands. Since the nearest-neighbor GC base pairs are much more stable than the other nearest-neighbor base pairs, the stable consecutive GC base pairs can be the initiation nucleus rather than the helix termini. Overall, duplexes **1** and **2** show different thermodynamic properties, but the consecutive GC base pairs near the middle of the strand rather than the strand termini determine the energy barrier of the duplex formation.

Formation of base pairs by short nucleotide sequences are fundamental to antisense, RNAi, DNA chip and nano-material technologies. Since these technologies are aimed at cells, quantitative information regarding the duplex formation under cell-like conditions is required. Our results demonstrate that the addition of a high concentration of the PEGs making a molecularly crowded media substantially changes the kinetics for a DNA duplex formation depending on the mass of the PEG. Moreover, the two DNA duplexes with different thermal stabilities can be formed after overcoming the same association activation energy barrier, suggesting the formation of the initiation nucleus formed in the middle of the helices. These findings are useful not only for understanding the relationships between the thermodynamics and kinetics of nucleic acids, but also in designing a DNA chip and constructing nucleic acid nano-materials.

This work was supported in part by Grants-in-Aid for Scientific Research and for the Academic Frontier Project (2004–2009) from MEXT (Ministry of Education, Culture, Sports, Science and Technology), Japan.

Notes and references

1 S. M. Freier, R. Kierzek, M. J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson and D. H. Turner, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 9373; N. Sugimoto, S. Nakano, M. Katoh, A. Matsumura, H. Nakamuta, T. Ohmichi, M. Yoneyama and M. Sasaki, *Biochemistry*, 1995, **34**, 11211; N. Sugimoto, S. Nakano, M. Yoneyama and K. Honda, *Nucleic Acids Res.*, 1996, **24**, 4501.

2 D. Porschke and M. Eigen, *J. Mol. Biol.*, 1971, **62**, 361.

3 M. E. Craig, D. M. Crothers and P. Doty, *J. Mol. Biol.*, 1971, **62**, 383.

4 A. P. Williams, C. E. Longfellow, S. M. Freier, R. Kierzek and D. H. Turner, *Biochemistry*, 1989, **28**, 4283.

5 A. Fire, S. Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello, *Nature*, 1998, **391**, 806; M. Chicurel, *Nature*, 2001, **412**, 580.

6 C. H. Spink and J. B. Chaires, *J. Am. Chem. Soc.*, 1995, **117**, 12887; C. H. Spink and J. B. Chaires, *Biochemistry*, 1999, **38**, 496.

7 R. Goobes, N. Kahana, O. Cohen and A. Minsky, *Biochemistry*, 2003, **42**, 2431; R. Goobes and A. Minsky, *J. Am. Chem. Soc.*, 2001, **123**, 12692.

8 S. Nakano, H. Karimata, T. Ohmichi, J. Kawakami and N. Sugimoto, *J. Am. Chem. Soc.*, 2004, **126**, 14330.

9 D. Miyoshi, H. Karimata and N. Sugimoto, *J. Am. Chem. Soc.*, 2006, **128**, 7957.

10 In fact, another 10-mer DNA duplex **3** (5′-CGGGCCGCGG-3′/5′-CCGCGGCCCG-3′) was studied, which has the same length, but different GC base pair content. However, duplex **3** showed a high melting temperature, and the ΔOD_{260} of duplex **3** was too low to determine using the temperature-jump method at 37 °C. Therefore, we only show the results of duplexes **1** and **2** in this study.

11 T. Ohmichi, H. Nakamuta, K. Yasuda and N. Sugimoto, *J. Am. Chem. Soc.*, 2000, **122**, 11286.

12 PEG 200 indicates poly(ethylene glycol) with an average molecular weight of 200. All the cosolutes were purchased from Wako Pure Chemical Co., Ltd., (Japan), and used without further purification.

13 The temperature-jump instrument was manufactured by Unisoku (Japan). The range of the temperature jump discharge at 15 kV in different buffers was determined by calibrating a test sample using phenolphthalein. Each relaxation curve was measured eight times and the relaxation time (τ) was evaluated from the average value obtained from the fits of the eight relaxation curves. For a DNA duplex formation, the plot of τ^{-2} *vs.* C_t (Fig. 2(b)) follows the equation $\tau^{-2} = 2k_{+1}k_{-1}C_t + k_{-1}^2$. Because the intercepts of the plots were very low and produced a large error, we calculated k_{+1} and k_{-1} from the slope (S) of the plot and the equilibrium association constant (K), which was obtained from the melting experiments: $k_{+1} = (S \cdot K/2)^{1/2}$ and $k_{-1} = k_{+1}/K$ (details described in ESI†).

14 D. H. Turner, in *Nucleic Acids: Structure, Properties and Functions*, ed. V. A. Bloomfield, D. M. Crothers and I. Tinoco, University Science Book, Sausalito, CA, 2000, ch. 8, pp. 259–334.

15 N. Kozer and G. Schreiber, *J. Mol. Biol.*, 2004, **336**, 763; Y. Y. Kuttner, N. Kozer, E. Segal, G. Schreiber and G. Haran, *J. Am. Chem. Soc.*, 2005, **127**, 15138; F. Du, Z. Zhou, Z.-Y. Mo, J.-Z. Shi, J. Chen and Y. Liang, *J. Mol. Biol.*, 2006, **364**, 469.

16 M. Burg, *Cell. Physiol. Biochem.*, 2000, **10**, 251.

17 K. Devanand and J. C. Selser, *Nature*, 1990, **343**, 739.

