Thermodynamics of DNA Triplex Formation in Oligomers with and without Cytosine Bases: Influence of Buffer Species, pH, and Sequence

W. David Wilson,*,† Harry P. Hopkins,*,† Shaikh Mizan,† David D. Hamilton,† and Gerald Zon‡

Department of Chemistry and Laboratory for Microbial and Biochemical Sciences Georgia State University, Atlanta, Georgia 30303 Lynx Therapeutics, 3832 Bay Center Place Hayward, California 94545

Received December 22, 1993

Triple-helical nucleic acids have provided interesting models for alternative DNA structures and base-pair hydrogen bonding.[1-3] The H-DNA triplex observed in natural DNA sequences has raised the possibility that triplexes may serve a role in gene expression.[4-6] Investigations of nucleic acid triplexes have also been stimulated by potential applications as highly specific nucleases,[7] and as agents for control of gene expression, through binding of either a single strand to duplex DNA (the antigene concept)[2,8,9] or two strands to RNA (an extension of antisense applications).[10-14] Rational application of triplexes in recognition and therapeutic strategies requires thermodynamic parameters for triplex formation;[2,14] however, the effects of sequence and T·A·T or C·G·C content on triplex stability and thermodynamics remain to be defined.[15,16]

Experimental enthalpies for formation of DNA triplexes from duplexes and single strands range from approximately −2 to −7 kcal/(mol of base).[17-22] This variation in enthalpy parameters may have arisen because partial protonation of C·G·C triplets makes the observed thermodynamic parameters dependent on the buffer species as well as the pH. The ΔH_i (ionization) of the buffer and the pH used in the experiments influence the observed triplex thermodynamic parameters. Although the effects of pH have been recognized, the influence of buffer species has not been systematically investigated. In this communication, we present results for triplex formation over an extended range of salt concentration, pH values, and buffer types. We compare isothermal titration calorimetric measurements and UV analysis of melting for dT_{19}·dA_{19}·dT_{19} with those for a corresponding 19mer with three separated C·G·C triplets: d[$(T_4C)_3T_4$]·d[$(A_4G)_3A_4$]·d-[$(T_4C)_3T_4$]. These sequences are long enough to form stable triplexes and to minimize end effects, and each C or G base has the same flanking sequence in both the 5′ and 3′ directions.

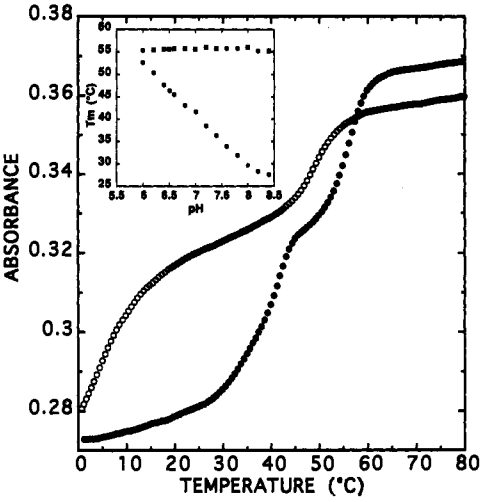

Figure 1. Absorbance (280 nm) vs temperature for 1:2 mixtures of purine (2×10^{-5} M) and pyrimidine oligomers (prepared as described[23]) in a 0.01 M PIPES buffer (pH 7.0) containing 0.001 M EDTA and 0.2 M NaCl: (○) dA_{19} and dT_{19}; (●) d(AAAAGAAAAGAAAAGAAAA) {d[$(A_4G)_3A_4$]} and d(TTTTCTTTTCTTTTCTTTT) {d[$(T_4C)_3T_4$]}. Inset: T_m vs pH for 1:2 mixtures of d[$(A_4G)_3A_3$] with d[$(T_4C)_3T_4$] in phosphate buffer with 0.2 M NaCl: (■) duplex (second transition) and (●) triplex (first transition).

In melting experiments, both sets of oligomers have two clearly resolved transitions,[23] one for the duplex (higher T_m) and one for the triplex (lower T_m) (Figure 1). A striking observation from Figure 1 is the marked increase in stability of the triplex, even at pH 7, caused by exchange of only three T·A·T for C·G·C triplets. A possible explanation for this result is that dA·dT sequences have an unusual conformation and hyration,[24-27] and disrupting the dA_{19}·dT_{19} structure may require significant Gibbs energy in agreement with the inverse correlation between triplex and duplex stability observed by Roberts and Crothers.[28]

Melting experiments were also conducted at 0.01 M < [Na^+] < 1.0 M, and T_m vs log [Na^+] plot are linear (not shown). Similar slopes are found for the duplexes, but the slope for the pure T·A·T triplex is significantly larger than for the triplex with three C·G·C.[29] In addition, the stability of the d[$(T_4C)_3T_4$]·d-[$(A_4G)_3A$]·d[$(T_4C)_3T_4$] triplex increases markedly as pH decreases from 8 to 6, whereas the T_m of the duplex is constant over this pH range (Figure 1, inset). These observations lead to several conclusions: C's in the third strand of d[$(T_4C)_3T_4$]·d[$(A_4G)_3A_4$]·d-[$(T_4C)_3T_4$] are partially protonated at pH 7; protonation stabilizes

† Georgia State University.
‡ Lynx Therapeutics.

(1) Felsenfeld, G.; Davies, D.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024.
(2) Sun, J. S.; Helene, C. *Curr. Opin. Struct. Biol.* **1993**, *3*, 345–356.
(3) Cheng, Y.-K.; Pettit, M. *Prog. Biophys. Mol. Biol.* **1992**, *58*, 225–257.
(4) Frank-Kamenetskii, M. D. *Methods Enzymol.* **1992**, *211*, 180–191.
(5) Mirkin, S. M.; Lyamichev, V. I.; Drushlyak, K. N.; Dobrynin, V. N.; Filippov, S. A.; Frank-Kamenetskii, M. D. *Nature* **1987**, *330*, 495–497.
(6) Wells, R. D.; Collier, D. A.; Hanvey, J. C.; Shimizu, M.; Wohlrab, F. *FASEB J.* **1988**, *2*, 2939–2949.
(7) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645–650.
(8) Helene, C.; Toulme, J. J. *Biochem. Biophys. Acta* **1990**, *1049*, 99–125.
(9) Wilson, W. D.; Tanious, F. A.; Mizan, S.; Yao, S.; Kiselyov, A. S.; Zon, G.; Strekowski, L. *Biochemistry* **1993**, *32*, 10614–10621.
(10) Giovannangeli, C.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10013–10017.
(11) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543–584.
(12) Zon, G. *Pharm. Res.* **1988**, *5*, 539–548.
(13) Ts'o, P. O. P.; Aurelian, L.; Chang, E.; Miller, P. S. *Ann. N.Y. Acad. Sci.* **1992**, *660*, 159–177.
(14) *Antisene Research and Applications*; Crooke, S. T., Lebleu, B., Eds.; CRC Press: Boca Raton, 1993.
(15) Fossella, J. A.; Kim, Y. J.; Shih, H.; Richards, E. G.; Fresco, J. R. *Nucleic Acids Res.* **1993**, *21*, 4511–4515.
(16) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957–6965.
(17) Plum, G. E.; Park, Y. W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436–9440.
(18) Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625–5631.
(19) Rougee, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Helene, C. *Biochemistry* **1992**, *31*, 9269–9278.
(20) Manzini, G.; Xodo, L. E.; Gasparotto, D.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *J. Mol. Biol.* **1990**, *213*, 833–843.
(21) Pilch, D.; Brousseau, R.; Shafer, R. *Nucleic Acids Res.* **1990**, *18*, 5743–50.
(22) Shindo, H.; Torigoe, H.; Sarai, A. *Biochemistry* **1993**, *32*, 8963–8969.
(23) Kibler-Herzog, L.; Kell, B.; Zon, G.; Shinozuka, K.; Mizan, S.; Wilson, W. D *Nucleic Acids Res.* **1990**, *18*, 3545–3555.
(24) Hopkins, H. P.; Hamilton, D. D.; Wilson, W. D.; Zon, G. *J. Phys. Chem.* **1993**, *97*, 6555–6563.
(25) Aymami, J.; Coll, M.; Frederick, C. A.; Wang, A., H.-J.; Rich, A. *Nucleic Acids Res.* **1989**, *17*, 3229–3245.
(26) Nadeau, J. G.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2622–2626.
(27) Nelson, H. C. M.; Finch, J. T.; Luisi, B.; Klug, A. *Nature* **1987**, *330*, 221–226.
(28) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463–1466.
(29) T_m vs log [Na^+] slopes: (dA_{19}·dT_{19}) → strands, 17.6; (dT_{19}·dA_{19}·dT_{19}) → duplex and strand, 52.9; d[$(A_4G)_3A_4$]·[$(T_4C)_3T_4$] → strands, 17.4; d[$(T_4C)_3T_4$]·[$(A_4G)_3A_4$]·[$(T_4C)_3T_4$] → duplex and strand, 19.0.

0002-7863/94/1516-3607$04.50/0

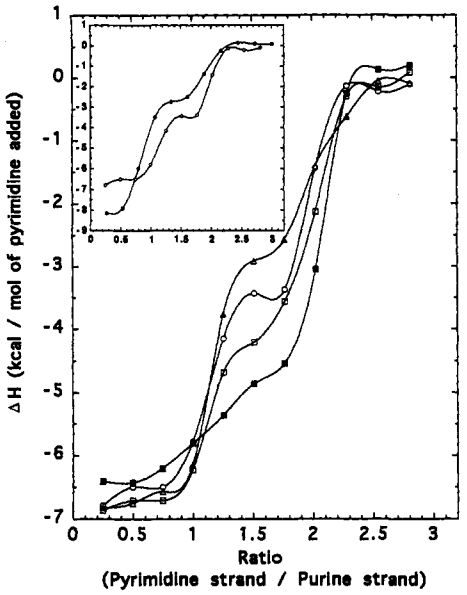

Figure 2. Isothermal calorimetric titration curves at 21 °C for d[$(T_4C)_3T_4$] into d[$(A_4G)_3A_3$] in various buffers and pH's with 1.0 M added NaCl: (△) TES, (○) PIPES, (□) phosphate (pH 7.0), and (■) phosphate (pH 6.0). Each data point represents the ΔH/(mol of pyrimidine) (10-μL injections, [base] = 0.625 mM) added into the purine strand (volume = 1.385 mL, [base] = 0.0183 mM). Inset: Calorimetric titrations of (○) d[$(T_4C)_3T_4$] into d[$(A_4G)_3A_3$] and (●) dT_{19} into dA_{19} into dA_{19} in PIPES.

C·G·C-containing triplexes; and protonated C·G·C triplets lower the triplex charge density.

From titration calorimetric curves[30] for d[$(A_4G)_3A_4$] with d[$(T_4C)_3T_4$] in three different buffers (Figure 2) it can be seen that the model-independent calorimetric ΔH for duplex formation (plateau between 0 and 1 ratio) is independent of buffer, whereas that for triplex formation (plateau between 1 and 2 ratio) depends strongly on buffer species. The calorimetric ΔH's for triplex formation at pH 7 are approximately (per base triplet) −2.9 kcal/mol in TES, −3.4 kcal/mol in PIPES, and −4.2 kcal/mol in $H_2PO_4^-$. Also shown in Figure 2 is the titration curve at pH 6, in which the plateau for triplex formation is at a lower ΔH than at pH 7. These results demonstrate both the buffer and pH dependence of the observed ΔH for triplex formation. A complete analysis of the thermodynamics of these systems at a range of concentrations is in progress. The value for duplex formation for the G·C-containing strands is less negative than for the pure A·T strands (Figure 2, inset) as predicted by the data base of Breslauer et al.;[31] however, the value for triplex formation is less negative for the pure T·A·T triplex.[24] Unlike the results with the mixed sequence, the T_m values and calorimetric ΔH for triplex formation from the A·T duplex are independent of pH and buffer species (not shown).

The results presented here emphasize that ΔH values for triplex formation from strands with cytosine, derived from plots in Figure 2, as well as those generally reported in the literature, are *apparent* thermodynamic values. The observed reaction for triplex formation can be represented as $D + S + n BH^+ \rightarrow T(nH^+) + nB$; where D, S, T, and B are duplex, free strand, triplex, and buffer, respectively. The value of n is related to the difference in pK_a between the C bases in the single strand and the triplex, and to the pH.[4] Dissecting this reaction into proton transfer and binding steps yields

$$nBH^+ \Leftrightarrow nB + nH^+: \quad n\Delta H_i \tag{1}$$

$$S + nH^+ \Leftrightarrow S(nH^+): \quad n\Delta H_{pC} \tag{2}$$

$$S(nH^+) + D \Leftrightarrow T(nH^+): \quad \Delta H_T \tag{3}$$

The *apparent* ΔH for the overall reaction of any pH can be expressed as $\Delta H_{app} = n\Delta H_i + n\Delta H_{pC} + \Delta H_T$, which includes a direct contribution from the buffer. A plot of ΔH_{app} from triplex formation at constant pH vs ΔH_i[32] is linear with a slope equal to n. Our data yields n in the range 2–2.5 at pH 7 for d[$(T_4C)_3T_4$]·d[$(A_4G)_3A_4$]·d[$(T_4C)_3T_4$], which indicates that the pK_a for CH^+ is between 7 and 7.5 (relative to the pK_a in single strands of approximately 4.7[33]), in good agreement with the observed changes of T_m with pH.

The results presented here indicate that some of the controversy over thermodynamic parameters of triplex formation may be due to variations in *buffer species* as well as pH and base sequence used by different workers. Careful consideration must be given in the future to the influence of *buffer type* as well as pH on the ΔH and T_m for triplex formation with strands that have cytosine bases. The same requirement holds for any DNA or RNA complex that involves protonation of a base.

Acknowledgment. This research was supported by NIH Grants AI-27196 and SBIR-AI/CA-35441 (Lynx Therapeutics).

Supplementary Material Available: Experimental details on synthesis of oligomers, preparation of oligomer stock solutions, buffers, UV spectrophotometry, and determination of pK_a's and ΔH_i for buffers and of binding enthalpies for oligomer duplex and triplex (3 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

(30) Hopkins, H. P.; Fumero, J.; Wilson, W. D. *Biopolymers* **1990**, *29*, 449–444.

(31) Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3746–3750.

(32) Values for ΔH_i and pK_a in 1.0 M NaCl were determined as described in the supplementary material to be as follows: $H_2PO_4^-$ (ΔH_i = 1.34 kcal/mol and pK_{a2} = 6.2); PIPES (ΔH_i = 3.65 kcal/mol and pK_a = 6.6); and TES (ΔH_i = 8.85 kcal/mol and pK_a = 7.4).

(33) Manzini, G.; Xodo, L.; Gasparotto, D.; Quadrifoglio, F.; van der Marel, G.; van Boom, J. *J. Mol. Biol.* **1990**, *213*, 833–843.