Effect of Loop Sequence and Size on DNA Aptamer Stability†

Ivan Smirnov and Richard H. Shafer*

Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, San Francisco, California 94143

Received August 16, 1999; Revised Manuscript Received November 24, 1999

ABSTRACT: The thrombin aptamer is a 15-mer oligodeoxyribonucleotide that folds into a unimolecular quadruplex consisting of a stack of two guanine quartets connected by two external loops and one central loop and possesses a high affinity for thrombin. We have undertaken a systematic examination, in KCl, of the thermodynamic stability of thrombin aptamer analogues containing sequence modifications in one or more of the loops, as well as in the number of quartets. UV melting studies have been carried out to obtain the relevant thermodynamic parameters for these aptamers. van't Hoff analysis of these data, with a two-state model for unimolecular denaturation, gave excellent fits to the experimental observations. Thermodynamic analysis indicates that the central loop sequence in the parent aptamer is optimal for stability. Modifications in this or other loops can effect either $\Delta H°$, $\Delta S°$, or both. Addition of a single G at the 5′-end decreases stability while addition of a G at the 3′-end increases stability. Differential scanning calorimetry experiments on the thrombin aptamer reveal that a heat capacity change, not detected by UV measurements, accompanies the unfolding of the aptamer.

DNA and RNA aptamers are oligonucleotides usually identified by some selection process to specifically bind to, or fit, a desired target (*1*, *2*). Their target affinity usually derives from their particular three-dimensional shape, which, in the case of DNA aptamers, is often based on the guanine quartet as the fundamental secondary structure unit. In general, guanine quadruplex structures, stabilized by guanine quartets, can be constructed from one, two, or four oligonucleotide strands. When four separate strands are involved, a linear quadruplex results, with all deoxyguanosine nucleosides in the anti conformation about the glycosidic bond. In the other two cases, the quadruplexes involve folded strands and there is an alternation between syn and anti conformations for these residues (*3*).

The thrombin-binding aptamer, or TBA,[1] was identified from a randomized oligonucleotide library as a high-affinity ligand for the protein thrombin and is one of the first DNA aptamers to be studied (*4*). The structure of TBA, a 15-mer with the sequence d(GGTTGGTGTGGTTGG), has been determined alone in solution by NMR (*5*−*8*) and complexed with thrombin by X-ray diffraction (*9*, *10*). The oligonucleotide folds such that it forms two guanine quartets connected by a central three-base TGT loop and two smaller TT loops, as illustrated in Figure 1.

Several studies have appeared that involve selection of aptamers for recognition of a mutant thrombin as well as wild-type, resulting in similar or slightly different consensus sequences or modular aptamer structures combining a folded quadruplex and a duplex (*11*−*14*). Additionally, a three-quartet aptamer has been reported to exhibit similar thrombin binding activity as TBA, yet was not part of the consensus sequences derived from in vitro selection experiments (*15*).

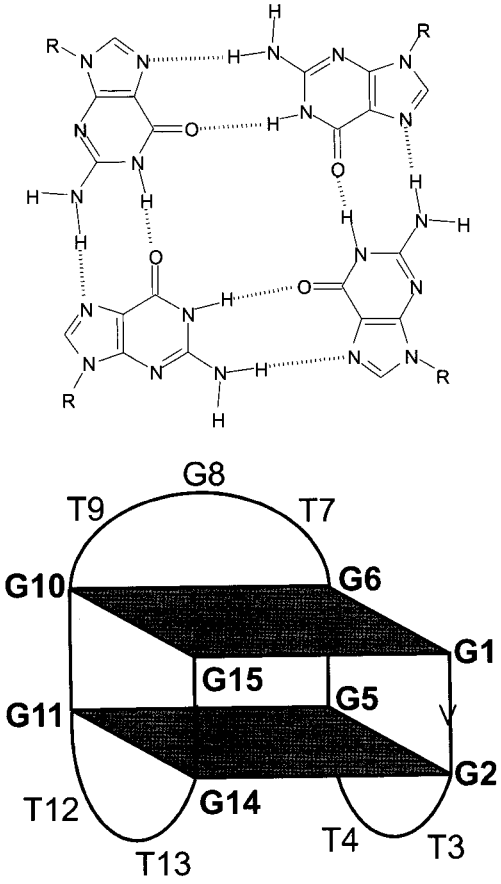

FIGURE 1: Gunaine quartet (top) and schematic drawing of the thrombin aptamer structure (bottom). Shaded squares represent guanine quartets.

† This work is supported by NIH Grant 1 P01 AI39152.

* To whom correspondence should be addressed. Phone: (415) 476-2761. Fax: (415) 476-0688. E-mail: shafer@socrates.ucsf.edu.

[1] Abbreviations: PAGE, polyacrylamide gel electrophoresis; DSC, differential scanning calorimetry; TBA, thrombin-binding aptamer.

10.1021/bi9919044 CCC: $19.00
Published on Web 01/20/2000

Table 1: Thermodynamics of Aptamer Formation Determined by van't Hoff Analysis of UV Melting Curves

oligonucleotide		T_m (°C) (±1)	$\Delta H°$ (kcal/mol) (±1.4)	$\Delta S°$ (eu) (±4)	$\Delta G°_{37}$ (kcal/mol) (±0.1)
5′-d(GGTTGG**TGT**GGTTGG)	**I**	46.4	−40.3	−126	−1.19
5′-d(GGTTGG**GTT**GGTTGG)	**II**	39.6	−37.6	−120	−0.31
5′-d(GGTTGG**TTG**GGTTGG)	**III**	40.2	−42.5	−136	−0.44
5′-d(GGTTGG**TAT**GGTTGG)	**IV**	40.2	−37.6	−120	−0.385
5′-d(GGTTGG**TTT**GGTTGG)	**V**	42.2	−40.9	−130	−0.67
5′-d(GGTTGG**TT**GGTTGG)	**VI**	21.0	−30.5	−103.6	+1.67
5′-d(GGTTGG**TTTT**GGTTGG)	**VII**	34.0	−40.7	−132.5	+0.40
5′-d(GGTTGG**CTTG**GGTTGG)	**VIII**	45.2	−37	−116	−0.95
5′-d(GGTTGG**GTAA**GGTTGG)	**IX**	45.4	−35.3	−111	−0.93
5′-d(GG**T**GG**TGT**GG**T**GG)	**X**	nd[a]	nd	nd	nd
5′-d(GG**TTT**GG**TGT**GG**TTT**GG)	**XI**	46.4	−47.3	−148	−1.4
5′-d(GG**C**TGG**TGT**GGTTGG)	**XII**	46.8	−39.4	−123	−1.2
5′-d(GGT**C**GG**TGT**GGTTGG)	**XIII**	40.9	−38.1	−121	−0.47
5′-d(GGTTGG**TGT**GG**C**TGG)	**XIV**	47.1	−40.6	−127	−1.28
5′-d(GGTTGG**TGT**GGT**C**GG)	**XV**	39.5	−37	−118.5	−0.3
5′-d(GG**TTTT**GG**TTTT**GG**TTTT**GG)	**XVI**	31.4	−41.0	−134	+0.75
5′-d(**G**GGTTGG**TGT**GGTTGG)	**XVII**	43.1	−39.1	−123	−0.75
5′-d(GGTTGG**TGT**GGTTGG**G**)	**XVIII**	50.5	−41.6	−128.5	−1.75
5′-d(GGTTGG**TGT**GGTTG)	**XIX**	<10	nd	nd	nd
5′-d(**GGG**TT**GGGTGTGGG**TT**GGG**)	**XX**	64.5	−55	−165	−4.55

[a] nd, not determined.

DNA aptamers other than TBA that possess potential therapeutic or biological activity have recently been described. For example, a quadruplex-based aptamer has been shown to have inhibitory activity against HIV integrase (*16*, *17*). Other aptamers have been shown to be inhibitors of neutrophil elastase (*18*), to possess anti-HIV activity based on its binding to the Rev protein (*19*) or to inhibit DNA polymerases (*20*). Catalytic DNA aptamers have also been described (*21*) and a DNA aptamer that binds arginine has been identified by in vitro selection methods and likely involves a quadruplex-based structure (*22*). Thus, there is an increasing pool of biologically active aptamers about which relatively little is known concerning their structure and stability and their relationship to function.

The ability of a given sequence to survive the in vitro selection process depends on two requirements, thermal stability and target affinity, both of which need to be fulfilled. In the work presented here, we examine the thermodynamic stability of a series of aptamer sequences as a function of modifications in the various loop domains of the aptamer, as well as the effect of adding an additional quartet. Our results indicate that changes in loop sequences can have a significant impact on the aptamer stability, consistent with our recent report on single base changes in the loop region of a hairpin dimer quadruplex (*23*).

MATERIALS AND METHODS

Oligonucleotides **I**−**XX** (see Table 1) were purchased from either Genset Corp. (La Jolla., CA) or Genosys Biotechnologies, Inc. (The Woodlands, TX). Purity was estimated to be in the range of 90% or greater by denaturing PAGE (polyacrylamide gel electrophoresis). Samples were suspended in 10 mM Tris, pH 7.4 and dialyzed extensively against this buffer, and then KCl was added as indicated. Following dialysis and each subsequent addition of salt, samples were annealed by heating to 80 °C and slow cooling to room temperature. Extinction coefficients were estimated by nearest neighbor methods (*24*). In the case of **I**, we also measured the extinction coefficient by phosphate determination (*25*), and obtained a value for ϵ_{260} in water of 1.46×10^5, which agrees within 2% of the value found by the nearest neighbor calculation.

UV Melting. Temperature-dependent melting curves were measured at 295 nm on a Cary 3A spectrophotometer, using a heating/cooling rate of 0.5 or 0.2 °C/min. Experiments included both heating and cooling curves. Analysis of these curves was carried out by fitting the melting curves (absorbance as a function of temperature) to the following equation:

$$A(T) = (1 - f)A_H(T) + fA_C(T)$$

where $A(T)$ is the absorbance of the solution at temperature T, $A_H(T)$, and $A_C(T)$ are the absorbances of the helical (structured) and coil (unstructured) forms, respectively, and f is the fraction of oligonucleotide strands in the melted or unstructured state. Following the development of Jovin and co-workers (*26*), we represent $A_H(T)$ and $A_C(T)$ as linear functions of the absolute temperature, T, i.e., $A_H(T) = b_H[1 + m_H(T - 273.15)]$ and $A_C(T) = b_C[1 + m_C(T - 273.15)]$. The fraction of melted strands, f, satisfies the following equation:

$$f^n + \frac{f}{2^{n-1}E} - \frac{1}{2^{n-1}E} = 0$$

for self-complementary n-mers, where

$$E = \frac{K(T_m)}{K(T)} = e^{(\Delta H°/R)(1/T - 1/T_m)}$$

and T_m is the temperature at which $f = 1/2$. The best fitting values of $\Delta H°$ and T_m were determined by nonlinear curve fitting. In this model, both $\Delta H°$ and $\Delta S°$ are assumed to be constant, i.e., $\Delta C_P = 0$ for the melting transition. Entropy changes were then calculated from

$$\Delta S° = \frac{\Delta H°}{T_m} + R \ln K(T_m) = \frac{\Delta H°}{T_m} + R \ln\left[n\left(\frac{C_0}{2}\right)^{n-1}\right]$$

where C_0 is the total strand concentration. Note that, for unimolecular systems, $n = 1$ and determination of $\Delta H°$, $\Delta S°$, and $\Delta G°$ does not require knowledge of the oligonucleotide concentration, C_0.

Calculation of $\Delta G°$ (310) was made via the following relationship, valid for constant $\Delta H°$:

$$\Delta G°(310) = \frac{310\Delta H°}{R}\left(\frac{1}{310} - \frac{1}{T_m}\right)$$

All systems were tested for reversibility by running heating and cooling curves at the same heating rate (either 0.5 or 0.2 °C/min). The heating and cooling curves superimposed very closely, matching the behavior observed for DNA duplexes run under the same conditions. The thermodynamic parameters in Table 1 represent averages from heating and cooling curves from 2 to 3 experiments. Experimental uncertainties in these parameters were determined from an analysis of a large number of melting experiments carried out on TBA, **I**, and are listed at the top of Table 1.

Microcalorimetry. Differential scanning calorimetry (DSC) was carried out on a MicroCal MC-2 differential scanning calorimeter. The thrombin aptamer sample was resuspended in 100 mM potassium phosphate buffer, pH 6.1, and dialyzed against the same buffer. The dialysate was used as the reference buffer in the calorimetry scans. The excess heat capacity curve was constructed by subtraction of the buffer versus buffer scan from the sample versus buffer. The resulting curve was fitted by nonlinear regression, using the manufacturer-supplied Origin software, with a variety of models (either supplied by the software or developed by us) as described in the text below.

RESULTS

CD of Oligonucleotides. The set of oligonucleotides we examined is described in Table 1. The modifications can be classified according to changes in the central loop alone (**II**−**IX**), outer loops only (**X**−**XV**), all loops (**XVI**), termini (**XVII**−**XIX**), and in the number of quartets (**XX**). The conformation of each oligonucleotide was examined both in the absence and presence of 25 mM KCl. For all sequences, the CD spectra, measured at 15 °C, showed characteristically low amplitude peaks, indicating little or no structure in the absence of added salt. In the presence of salt, however, most sequences exhibited a spectrum characterized by positive maxima near 293 and 247 nm and a negative minimum near 265 nm, as illustrated for three sequences in Figure 2. These features are typical of folded quadruplexes involving deoxyguanosines alternating between syn and anti conformations about the glycosidic bond (*3*).

The principal exception to this behavior was **VI**, which showed very low amplitude bands in the presence of KCl, indicating a partially structured oligonucleotide, **X**, where the CD spectrum in the presence of 25 mM KCl showed a broad maximum at 275 nm, and **XX**, which showed different features depending on how the sample was prepared. For example, if KCl was added after annealing **XX** in Tris alone, its CD spectrum consisted of a maximum at 293 nm, characteristic of folded quadruplexes. If the sample then was annealed again, a small peak appeared at 265 nm, typical of linear, four-stranded quadruplexes, in addition to the one at 293 nm. Over the period of several days, the peak at 265 nm increased in amplitude while that at 290 nm decreased. The equilibrium solution is probably a mixture of both folded and linear quadruplexes. A similar mixture of linear and folded species was observed for the sequence d(GGGTTTTGGG) in KCl (*27*). This behavior was not observed in the other sequences studied here containing runs of at most two guanines.

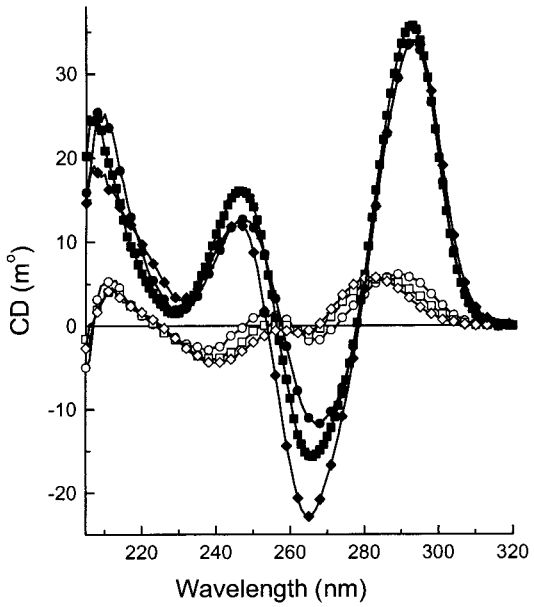

FIGURE 2: CD spectra at 15 °C of **I**, **V**, and **VII** (concentration ≈ 8 μM strand) in 10 mM Tris, pH, 6.8 alone (○, □, and ◇, respectively) and with 25 mM KCl (●, ■, and ◆, respectively.

UV Melting. Due to the extremely slow kinetics of forming linear, four-stranded quadruplexes, it is very difficult to carry out thermodynamic measurements, such as melting, which typically require the system to be constantly in equilibrium (*28, 29*). Dimeric hairpin quadruplexes, such as $[d(G_3T_4G_3)]_2$, associate at a considerably faster rate, leading to a slight degree of hysteresis at a heating/cooling rate of 0.5 °C/min (data not shown). Unimolecular quadruplexes, however, often exhibit much faster kinetics such that these measurements can be readily performed (*30*). UV heating and cooling curves for each sequence were superimposable, indicating rapid kinetics, and were fitted by models corresponding to denaturation of one-, two-, three-, or four-stranded structures, as indicated above. In all cases except **X** and **XIX**, which were not analyzed for thermodynamic parameters, the best fit was obtained for the model corresponding to melting of a unimolecular structure. Furthermore, several sequences were melted at varying strand concentrations and no changes in melting temperature were observed. Thus, for the most part, the sequence modifications imposed on the thrombin aptamer did not alter the overall molecularity of the resulting quadruplexes.

Sample melting curves are displayed in Figure 3, along with the best fit, and the resulting thermodynamic parameters are included in Table 1. The difference between the high- and low-temperature UV spectra of **I** is also included, showing that the largest change upon melting is the decrease in absorbance at 295 nm. With several notable exceptions, most modified sequences possess lower melting temperatures, T_m, than the original aptamer, and a correspondingly less favorable $\Delta G°$ at 37 °C. Further examination of Table 1

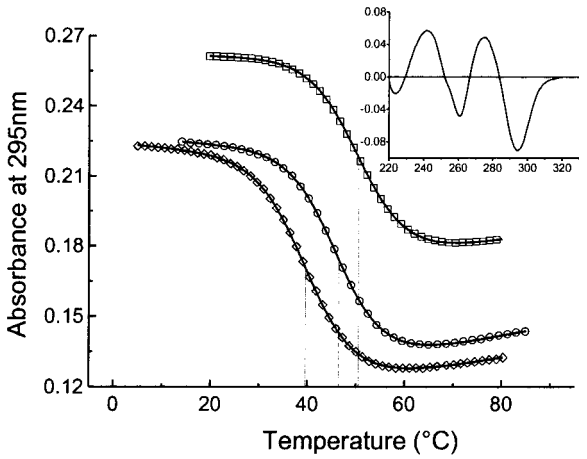

FIGURE 3: UV melting curves in 10 mM Tris, pH 6.8, 25 mM KCl for **I** (○), **XVIII** (□), and **II** (◇) at ≈ 8 μM strand concentration. Smooth curve represents best fit for a unimolecular all-or-none melting process. Only every 9th to 14th experimental point is plotted for clarity. Inset: UV difference spectrum (spectrum at 80 °C minus spectrum at 15 °C).

reveals that some modifications affect mainly the enthalpy change whereas others affect mainly the entropy change. For example, both **II** and **III** possess lower T_m values, by 6−7 °C, than **I**. However, $\Delta H°$ is about 3 kcal/mol less favorable for forming **II** compared to **I**, whereas $\Delta H°$ for forming **III** is about 2 kcal/mol more favorable than **I**. Compensating effects can be seen for $\Delta S°$ of formation, which is more favorable for **II**, and less favorable for **III**, than **I**.

It is difficult to interpret these results unambiguously in terms of differences in the structure of the folded quadruplexes, as sequence modifications may also affect the unfolded state, which still may harbor a small amount of structure via base stacking; we assume here that this effect is small. We note that $\Delta S°$ for folding **III** is the most negative of all 15-mer sequences. Oligonucleotides **II** and **III** are also unique in that they are the only ones, other than **XX**, that possess an internal run of three guanines. Only **III**, however, has a particularly unfavorable $\Delta S°$. One interpretation is that this particular sequence leads to a highly stabilized central loop, perhaps with additional stacking between the last G in that loop and the adjacent quartet. This enthalpic stabilization may be compensated by the entropic cost of additional ordering in that region. Similar interactions involving the first G in the loop apparently do not occur.

The all-T central loop in **V** leads to a very similar value for $\Delta H°$ as that of **I**, accompanied by a slightly less favorable $\Delta S°$, resulting in a T_m that is ∼4 °C lower than that of **I**. If that loop is increased to four Ts (**VII**), again the $\Delta H°$ term is very close to that of **I**, but the $\Delta S°$ term is more destabilizing, leading to a considerably lower T_m of 34 °C. If, in addition to four T's in the central loop, the external loops also consist of 4 Ts (**XVI**), then this trend continues with a $\Delta H°$ term that is barely more favorable than that of **VII** and a still more unfavorable $\Delta S°$. Thus, a trend is apparent with increasing the loop length, in both the central and the external loops, involving slightly more favorable enthalpic effects accompanied by substantially less favorable entropic effects such that the T_m gradually decreases, as does the stability at 37 °C. The unusually low stability of **VI**, possessing a −TT− central loop, arises from a significantly less favorable enthalpy change which may reflect the steric effects involved in such a short loop. We note that results on **VI** are less reliable than on the other sequences due to our inability to obtain a low-temperature baseline.

We have examined several oligonucleotides with a modified central loop consisting of a sequence identified as forming an unusually stable tetraloop (*31*) (**VIII** and **IX**). Here, it can be seen that these oligonucleotides are substantially more stable than the 4-T central loop sequence, **VII**, although they are not quite as stable as **I**. Relative to **VII**, the increase in stability for both sequences arises from a more favorable entropy of formation that is partially compensated by a less favorable enthalpy of formation. Thus, it appears that these tetraloop sequences are more stable than a 4-T loop in the context of a folded quadruplex as well as a hairpin duplex.

Oligonucleotides **X** and **XI** possess modified external loops, the former having external loops composed of a single T, while the latter has three Ts in its external loops. Since **X** exhibits an unusual CD spectrum and multiple bands on native PAGE (data not shown), we did not carry out thermodynamic analysis of its UV melting profile. Clearly, reducing the size of these loops to one nucleotide has a major impact on the folding. Changing the two external loops from 2 Ts to 3 Ts results in a structure (**XI**) with the same T_m as **I**, but with increased enthalpic stability compensated by entropic destabilization. This continues the trend observed above with increasing the size of the central, all-T loop.

The effect of dangling bases was explored in **XVII** and **XVIII**, which contain an additional G nucleotide at the 5′- or 3′-end, respectively. The presence of an extra G at the 3′-end led to increased stability for **XVIII** compared to the parent aptamer, **I**, with a higher T_m and an increased enthalpic stabilization, accompanied by a slightly less favorable entropy change. In contrast, an additional G at the 5′-end of **XVII** yielded an oligonucleotide with lower stability, with both a lower T_m and a less favorable $\Delta H°$, accompanied by a slightly more favorable $\Delta S°$. This is just the opposite effect observed in **XVIII**. We also investigated the effect of deleting the 3′-terminal G in **I**, to yield oligonucleotide **XIX**, which had a very low T_m (<13 °C) and which was not analyzed further. Such a dramatic loss of stability is consistent with an incomplete stack of guanine quartets. Furthermore, this result suggests that any failed sequences present in our samples would melt at a very low temperature, and thus have a negligible influence on our results.

Interloop Hydrogen Bonding. The high-resolution NMR structures revealed evidence for hydrogen bonding between the two external loops (*8*). In particular, hydrogen bonds between T4 and T13 were indicated, although it was difficult to assess whether they involved thymine O2 or O4 acceptor atoms. We probed these interactions by examining the stability of a series of oligonucleotides, **XII**−**XV**, containing one C in each of the four possible positions in the external loops. We expected that substituting C for T3 (**XII**) or T12 (**XIV**) to have little effect on the stability of the resulting sequence, while substituting C for T4 (**XIII**) or T13 (**XV**), would lead to decreased stability due to loss of hydrogen bonding.

The results summarized in Table 1 are consistent with these expectations. Both **XIII** and **XV** show reduced thermodynamic stability, while **XII** and **XIV** have stabilities very close to that of **I**. When examined at pH 5, TBA as

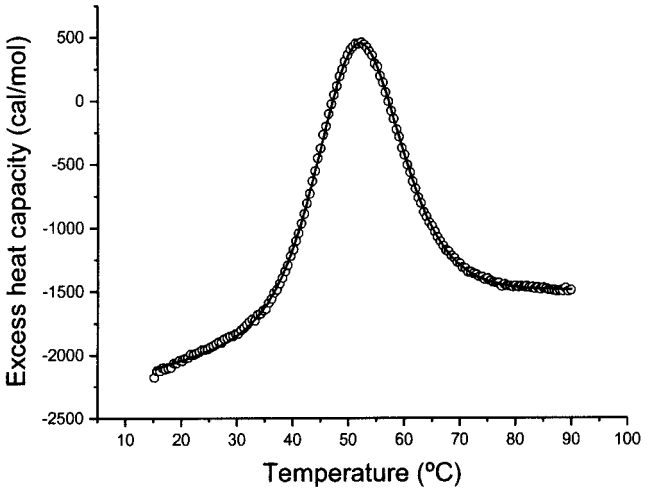

FIGURE 4: Excess heat capacity curve for **I** (300 μM strand concentration) in 100 mM potassium phosphate buffer, pH 6.1. Smooth curve is best fit to an all-or-none melting process accompanied by a change in heat capacity between initial and final states. Only every other experimental point is plotted for clarity.

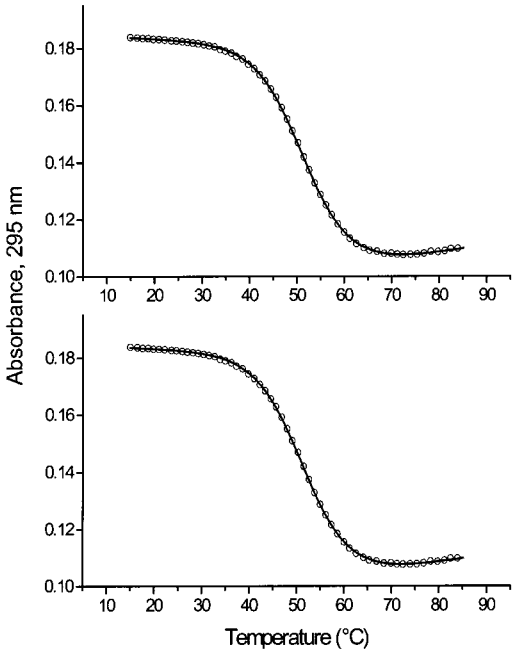

FIGURE 5: UV melting curve for **I** (6.5 μM strand concentration) in 100 mM potassium phosphate buffer, pH 6.1. Smooth curve is best fit to all-or-none unimolecular denaturation process with no heat capacity change, as described in the text for UV melting analysis (bottom) or with inclusion of a heat capacity change and using the thermodynamic parameters from the fit in Figure 4, optimizing only the absorbance baseline terms (top). Only every 10th point is displayed for clarity.

well as all four modified oligonucleotides (**XII**−**XV**) had similar thermodynamic stabilities and were more stable than **XIII** and **XV** at neutral pH (data not shown). This latter result suggests that **XIII** and **XV** are capable of restoring the hydrogen bonding at low pH, indicating that these hydrogen bonds involve the pyrimidine O2 atom. This conclusion was also suggested, but not proved, in NMR studies (*8*).

Effect of Adding a Third Quartet. The longest sequence examined in this study is **XX**, which allows for formation of a unimolecular quadruplex composed of three quartets, rather than two. As can be seen in Table 1, this results in a substantial increase in T_m and stability, arising from a large increase in enthalpic stability, again partially compensated by a corresponding increase in the entropy of folding. The enthalpy per quartet, however, is somewhat lower for this three-quartet quadruplex compared to the two-quartet structures. This difference may reflect the influence of the loops on the overall folding energetics.

Calorimetry. We carried out differential scanning calorimetry (DSC) on **I** in an effort to confirm the two-state nature of the melting transition for this folded aptamer. In addition, we also wanted to clarify the apparent discrepancy between the previously reported DSC and UV melting results for this aptamer. In the former, a value of 22 kcal/mol was reported for $\Delta H°$ in 100 mM potassium buffer, pH 6.1 (*7*), while in the latter it is was found to be in the range of 39−46 kcal/mol in 20 mM Li_3PO_4 buffer, pH 7, containing 5−50 mM KCl (*32*) or 43.5 kcal/mol in 100 KCl (*30*).

Results from our DSC experiments, carried out in 100 mM potassium phosphate, pH 6.1, are shown in Figure 4. The simplest analysis assumes a two-state transition with $\Delta C_P = 0$, neglecting any difference in C_P between the two baseline regions. Application of this model involves smoothly connecting the low- and high-temperature segments of the excess heat capacity curve and subtracting the resulting baseline. The result of this analysis, using several different baseline constructions, led to an unacceptably poor fit of the resulting heat capacity curve. Extension to a non-two-state model yielded a very good fit, but with a van't Hoff enthalpy change, $\Delta H_{VH}°$, ranging from −41 to −46 kcal/mol along with a calorimetric enthalpy change, $\Delta H_{CAL}°$, ranging from −28 to −33.2 kcal/mol (the range of values is due to differences in baseline construction). When $\Delta H_{VH}°$ exceeds $\Delta H_{CAL}°$, however, the usual explanation involves dimerization or higher order aggregation (*33*, *34*). Several observations argue against this interpretation. First, the calorimetrically determined melting temperature, as measured by the maximum in the excess heat capacity curve, is very close to that observed in the UV melting studies carried out at much lower DNA concentrations. This is consistent with a unimolecular, not associated, melting system. Second, results from equilibrium ultracentrifugation experiments were also consistent with a single unimolecular species present at both high and low DNA concentrations (data not shown).

The next approach to analyzing the calorimetry data was to allow for a change in heat capacity to accompany the unfolding transition, i.e., $\Delta C_P \neq 0$. The most general case in this instance was to allow both folded and unfolded forms to have temperature-dependent heat capacities that vary linearly in T. This analysis gave excellent fits to the data, as shown by the solid line in Figure 4, with a resulting $\Delta H°$ of −41.6 kcal/mol and $\Delta C_P = -0.155$ kcal/mol K at $T_m = 51.6$ °C. Similarly, a very good fit was obtained with a constant ΔC_P (−0.192 kcal/mol K) in the transition region, but somewhat less good in the baseline regions, with a comparable enthalpy change, $\Delta H° = -42.9$ kcal/mol at at $T_m = 51.5$ °C.

UV melting studies were then carried out under the same solution conditions as the calorimetry experiments, except at lower DNA concentrations, and the melting profiles are shown in Figure 5. Analysis of these data was carried out three different ways. First, we fit these UV data as described above, for the other UV melting experiments, using the

simple, two-state model for melting of a unimolecular species. This led to a T_m of 52.0 °C and an enthalpy change, $\Delta H°$, of −42.6 kcal/mol (Figure 5, bottom), in good agreement with that obtained in the calorimetry experiments. Second, we fit these data to the same $\Delta C_P \neq 0$ model used to analyze the calorimetry data, imposing the nonzero ΔC_P parameters but fitting to the $\Delta H°$ and T_m parameters, as well as the UV baseline parameters. Here, we obtained a T_m value of 51.8 °C and a $\Delta H°$ of −42.6 kcal/mol at T_m, virtually identical to the results above from the simplest model.

Third, these UV melting curves were fit to the same $\Delta C_P \neq 0$ model used to analyze the calorimetry data, imposing all the thermodynamic parameters from the calorimetry fit, including the $\Delta H°$ and T_m parameters, and only adjusting the UV absorbance baseline terms. Here again, excellent results were also obtained in that these parameters produced an excellent fit to the data (see Figure 5, top). Thus, a two-state transition model with a change in heat capacity is able to fit both UV melting and differential scanning calorimetry data.

DISCUSSION

We have examined the effect of a variety of sequence changes in the loops of the thrombin aptamer on its thermodynamic stability. Most of the changes in the central loop lead to a decrease in thermodynamic stability, indicating that, at least among the sequences explored, the -TGT- sequence of **I** is optimal for stability. These effects may involve changes in both stacking interactions and cation binding. Interestingly, loop length appears to play an important role in overall stability as well. This is evident from the significant decrease in stability of **VII** compared to **I**. The substantially lower T_m for **VII** arises from a more negative entropy of folding, the enthalpy of folding being very similar to that of **I**. This implies that there is no additional stabilizing interactions within the 4-T loop, but there is an entropic cost to forming this loop, relative to a 3-T loop. Similarly, the -TT- central loop leads to a pronounced decrease in stability arising from a substantially lower enthalpy change, possibly reflecting the strain involved in decreasing the central loop from three to two nucleotides. Additionally, there may be a loss of stacking between the loop and neighboring quartet (*5*−*8*).

Altering the four-base sequence to correspond to one of the unusually stable tetraloop sequences, as exemplified by **VIII** and **IX**, can compensate for some of this destabilization. Interestingly, in these two cases, the stability is enhanced by virtue of a less unfavorable entropy of folding, as the enthalpy terms are slightly less favorable than in **VII**. This contrasts with the observations made in duplex hairpin tetraloops, indicating that the stabilization derives from a more favorable enthalpy change (*35*).

Our results on the impact of replacing single Ts in the external loops with Cs provide indirect evidence for hydrogen bond formation between these loops, as observed in the NMR structures (*8*). Feigon and co-workers carried out similar substitutions, although they were unable to detect directly resonances corresponding to the C^+ imino proton at low pH (*8*). Nonetheless, the fact that the C-substituted aptamers exhibited the same stability as TBA at low pH suggests that protonation of the cytosines restored the hydrogen bonding present when all residues are T in the external loops. This pH-induced stabilization is also consistent with the T−T hydrogen bonding in TBA involving the O2 carbonyls as acceptor groups. Reduction of the external loops to a single T, as in **X**, led to a completely different CD spectrum and hence folding pattern, as might be expected due to the difficulty of forming a single base loop.

Finally, the calorimetry results warrant further discussion. From the data shown in Figure 4, it is apparent that the low- and high-temperature baselines possess different slopes, and hence, the folded and unfolded states of TBA possess different heat capacities. This behavior was observed in all of our calorimetry scans. If this difference in heat capacity is neglected, the only model that can fit the data is a non-two-state model that leads to a low estimate of $\Delta H_{CAL}°$, significantly smaller than $\Delta H_{VH}°$. The usual interpretation of such results invokes some type of aggregation, but several lines of evidence indicate that no such aggregation occurs. Accounting for the differences in heat capacity allows for a good fit of the data, and produces a $\Delta H°$ value at the corresponding T_m in good agreement with that determined by UV melting.

Most thermodynamic studies of DNA denaturation have been analyzed with the assumption that there is no difference in heat capacity between the initial and final states (*29*, *36*−*39*), although several studies have reported heat capacity changes associated with denaturation of duplexes and higher order structures (*40*, *41*). The question of heat capacity changes accompanying duplex melting has been examined quantitatively in several recent reports. Breslauer (*42*) and co-workers have observed an average ΔC_P of about 65 cal (mol base pair)$^{-1}$ for denaturation of a series of polynucleotide duplexes. Record and co-workers (*43*) observed a positive ΔC_P of about 57 cal (mol base pair)$^{-1}$ accompanying denaturation of a short duplex, in good agreement with the polynucleotide result. This was interpreted in terms of changes in solvent exposure of both polar and nonpolar groups. The overall process of duplex formation was described as involving several steps which influence the final result: folding of individual stands, partial stacking of strands to produce a stacked helical structure and docking of two such stacked helical structures to form a duplex. For the unimolecular quadruplexes studied here, only the first process is involved. However, the heat capacity differences between the folded and unfolded state are still expected to depend on the net changes in solvent exposure of polar and nonpolar groups. Formation and stacking of two guanine quartets likely involves sequestration from solvent of a significant nonpolar surface area and thus melting of such structures may be expected to entail an increase in heat capacity, as observed in our DSC experiments.

The similarity in the enthalpy changes for **I** determined by the simple analysis of UV melting curves and by the more complicated analysis of DSC data suggests that the simpler analysis provides a reasonable estimate of thermodynamic effects of the indicated sequence differences. To explore this question in greater detail, we reanalyzed the UV melting data for eight of the sequences included in Table 1, covering the range of T_m and $\Delta H°$ values. We applied the $\Delta C_P \neq 0$ model, imposing the heat capacity parameters determined from analysis of **I**, but fitting for $\Delta H°$, T_m, and the UV baseline parameters. The resulting $\Delta H°$ values at their respective T_m's

agreed to within 1 kcal/mol or less with those in Table 1. When these enthalpy changes were calculated at a common temperature of 37 °C, they differed by 3 kcal/mol or less from those in Table 1. Finally, when the $\Delta G°$ (37 °C) was recomputed for these eight sequences, their ranking with regard to stability was unchanged from that inferred from Table 1. Therefore, our conclusions concerning the effects of loop changes on thermodynamic stability are valid even though they are based on the simpler analysis.

With regard to previously reported data on other quadruplexes (*44*), our values for $\Delta H°$ per quartet range from 15 to 23.5 kcal/mol, with most values falling between 18 and 21 kcal/mol, and are in good agreement. While our value for $\Delta H_{CAL}°$ for **I**, determined from an analysis that assumes $\Delta C_P = 0$, is larger than that reported previously (*7*), they are both significantly smaller than $\Delta H°$ determined by van't Hoff analysis of UV melting curves (our results and ref *32*) or determined by analysis of calorimetric melting curves with $\Delta C_P \neq 0$. These differences in $\Delta H_{CAL}°$ may arise from differences in baseline construction, which can be significant when initial and final baselines have different slopes. Finally, our results indicate a very strong enthalpy−entropy compensation effect throughout the entire set of oligonucleotides, as has been reported in other DNA melting studies (*45*).

In summary, the relatively rapid kinetics characteristic of unimolecular quadruplex formation/dissociation observed in our studies facilitates the thermodynamic analysis of these systems. UV melting studies can provide good estimates of the thermodynamic parameters, although DSC studies may be necessary to assess the importance of heat capacity changes accompanying the unfolding of these small quadruplexes. Changes in loop sequence and composition have a significant effect on the stability of the resulting structure, although the largest increase in stability arises from addition of a third quartet. Although such a three-quartet sequence was not identified by the in vitro selection process as a tight binding ligand for thrombin, it may well possess this property along with increased thermodynamic stability.

REFERENCES

1. Ellington, A. D., and Szostak, J. W. (1990) *Nature 346*, 818−822.
2. Tuerk, C., and Gold, L. (1990) *Science 249*, 505−510.
3. Williamson, J. R. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 703−730.
4. Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., and Toole, J. J. (1992) *Nature 355*, 564−566.
5. Wang, K. Y., Krawczyk, S. H., Bischofberger, N., Swaminathan, S., and Bolton, P. H. (1993) *Biochemistry 32*, 11285−11292.
6. Wang, K. Y., McCurdy, S., Shea, R. G., Swaminathan, S., and Bolton, P. H. (1993) *Biochemistry 32*, 1899−1904.
7. Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A., and Feigon, J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3745−3749.
8. Schultze, P., Macaya, R. F., and Feigon, J. (1994) *J. Mol. Biol. 235*, 1532−1547.
9. Padmanabhan, K., Padmanabhan, K. P., Ferrara, J. D., Sadler, J. E., and Tulinsky, A. (1993) *J. Biol. Chem. 268*, 17651−17654.
10. Padmanabhan, K., and Tulinsky, A. (1996) *Acta Crystallogr., Sect. D 52*, 272−282.
11. Tasset, D., Kubik, M., and Steiner, W. (1997) *J. Mol. Biol. 272*, 688−698.
12. Wu, Q., Tsiang, M., and Sadler, J. E. (1992) *J. Biol. Chem. 267*, 24408−24412.
13. Macaya, R., Waldron, J., Beutel, B., Gao, H., Joesten, M., Yang, M., Patel, R., Bertelsen, A., and Cook, A. (1995) *Biochemistry 34*, 4478−4492.
14. Tsiang, M., Gibbs, C. S., Griffin, L. C., Dunn, K. E., and Leung, L. L. (1995) *J. Biol. Chem. 270*, 19370−19376.
15. Sumikura, K., Yano, K., Ikebukuro, K., and Karube, I. (1997) *Nucleic Acids Symp. Ser. 9*, 257−258.
16. Cherepanov, P., Este, J. A., Rando, R. F., Ojwang, J. O., Reekmans, G., Steinfeld, R., David, G., De Clercq, E., and Debyser, Z. (1997) *Mol. Pharmacol. 52*, 771−780.
17. Mazumder, A., Neamati, N., Ojwang, J. O., Sunder, S., Rando, R. F., and Pommier, Y. (1996) *Biochemistry 35*, 13762−13771.
18. Bless, N. M., Smith, D., Charlton, J., Czermak, B. J., Schmal, H., Friedl, H. P., and Ward, P. A. (1997) *Curr. Biol. 7*, 877−880.
19. Konopka, K., Düzgünes, N., Rossi, J., and Lee, N. S. (1998) *J. Drug Target. 5*, 247−259.
20. Lin, Y., and Jayasena, S. D. (1997) *J. Mol. Biol. 271*, 100−11.
21. Breaker, R. R., and Joyce, G. F. (1994) *Chem. Biol. 1*, 223−229.
22. Harada, K., and Frankel, A. D. (1995) *EMBO J. 14*, 5798−5811.
23. Keniry, M. A., Owen, E. A., and Shafer, R. H. (1997) *Nucleic Acids Res. 25*, 4389−4392.
24. Gray, D. M., Hung, S. H., and Johnson, K. H. (1995) *Methods Enzymol 246*, 19−34.
25. Griswold, B. L., Humoller, F. L., and MacIntyre, A. R. (1951) *Anal. Chem. 23*, 192−194.
26. Evertsz, E. M., Rippe, K., and Jovin, T. M. (1994) *Nucleic Acids Res 22*, 3293−3303.
27. Strahan, G. D., Keniry, M. A., and Shafer, R. H. (1998) *Biophys. J. 75*, 968−981.
28. Wyatt, J. R., Davis, P. W., and Freier, S. M. (1996) *Biochemistry 35*, 8002−8008.
29. Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8832−8836.
30. Mergny, J. L., Phan, A. T., and Lacroix, L. (1998) *FEBS Lett. 435*, 74−78.
31. Antao, V. P., and Tinoco, I., Jr. (1992) *Nucleic Acids Res 20*, 819−824.
32. Jing, N., Rando, R. F., Pommier, Y., and Hogan, M. E. (1997) *Biochemistry 36*, 12498−12505.
33. Privalov, P. L., and Potekhin, S. A. (1986) *Methods Enzymol. 131*, 4−51.
34. Plum, G. E., Pilch, D. S., Singleton, S. F., and Breslauer, K. J. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 319−350.
35. Varani, G. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 379−404.
36. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746−3750.
37. Marky, L. A., and Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
38. SantaLucia, J., Jr. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 1460−1465.
39. Sugimoto, N., Nakano, S., Yoneyama, M., and Honda, K. (1996) *Nucleic Acids Res. 24*, 4501−4505.
40. Lu, M., Guo, Q., Marky, L. A., Seeman, N. C., and Kallenbach, N. R. (1992) *J. Mol. Biol. 223*, 781−789.
41. Ladbury, J. E., Sturtevant, J. M., and Leontis, N. B. (1994) *Biochemistry 33*, 6828−6833.
42. Chalikian, T. V., Völker, J., Plum, G. E., and Breslauer, K. J. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7853−7858.
43. Holbrook, J. A., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (1999) *Biochemistry 38*, 8409−8422.
44. Pilch, D. S., Plum, G. E., and Breslauer, K. R. (1995) *Curr. Opin. Struct. Biol. 5*, 334−342.
45. Petruska, J., and Goodman, M. F. (1995) *J. Biol. Chem. 270*, 746−750.

BI9919044