Nucleic Acids Research, Vol. 20, No. 4 819–824

Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops

Vincent P.Antao+ and Ignacio Tinoco, Jr*
Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, CA 94720, USA

Received November 4, 1991; Revised and Accepted January 21, 1992

ABSTRACT

We determined the melting temperatures (T_m) and thermodynamic parameters of 15 RNA and 19 DNA hairpins at 1 M NaCl, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7. All these hairpins have loops of four bases, the most common loop size in 16S and 23S ribosomal RNAs. The RNA hairpins varied in loop sequence, loop-closing base pair (A·U, C·G, or G·C), base sequence of the stem, and stem size (four or five base pairs). The DNA hairpins varied in loop sequence, loop-closing base pair (C·G, or G·C), and base sequence of the four base-pair stem. Thermodynamic properties of a hairpin may be represented by nearest-neighbor interactions of the stem plus contributions from the loop. Thus, we obtained thermodynamic parameters for the formation of RNA and DNA tetraloops. For the tetraloops we studied, a free energy of loop formation (at 37°C) of about +3 kcal/mol is most common for either RNA or DNA. There are extra stable loops with $\Delta G°_{37}$ near +1 kcal/mol, but the sequences are not necessarily the same for RNA and DNA. The closing base pair is also important; changing from C·G to G·C lowered the stability of several tetraloops in both RNA and DNA. These values will be useful in predicting RNA and DNA secondary structures.

INTRODUCTION

Accurate prediction of secondary structure in RNA and DNA from thermodynamic data (1, 2) requires parameters for the favorable contributions from double-strand formation (3, 4) and the unfavorable contributions from loops and bulges (5–7). Phylogenetic comparisons of 16S and 23S ribosomal RNAs from many different species provide some interesting insights as to the size, sequence, and loop-closing base pair characteristics of hairpin loops in these RNAs (8, 9). Over 50% of the hairpins predicted have four-base loops. About 70% of these tetraloops have the consensus loop sequences (GNRA) or (UNCG) (where N = A, C, G, or U; and R = A or G). RNA hairpins with these sequences form unusually stable loop conformations (10–13). In spite of the importance of tetraloops in RNA structure, not much is known about their contribution to the stability of hairpins, and hence to the overall stability of RNA secondary structure. There is an equal, if not greater, paucity of information about the factors contributing to the stability of DNA hairpins.

We have determined the contributions of tetraloops to the thermodynamic stability of RNA and DNA hairpins. Included in this study are those loops which have been found to be unusually stable (10–13).

MATERIALS AND METHODS

The DNA template molecules for synthesis of the RNA hairpins and the DNA hairpin molecules were synthesized by the phosphoramidite method on an automated DNA synthesizer (Applied Biosystems, Inc.). After deprotection, the DNA oligomers were purified by polyacrylamide gel electrophoresis. The RNA hairpins were synthesized using T7 RNA polymerase (14) and were purified by polyacrylamide gel electrophoresis. The sequences of the RNA molecules were determined enzymatically.

For the melting profiles, the DNA and RNA stock solutions were extensively dialyzed, first against 0.01 M EDTA, 0.5 M NaCl, and 0.01 M sodium phosphate, second against 0.5 M NaCl, and 0.01 M sodium phosphate, third against 0.01 M sodium phosphate, all at pH 7, and finally against double-distilled water. For each melting profile, an appropriate volume of the oligonucleotide stock was dried in a speed-vac and the sample then dissolved in the buffer (at pH 7 ± 0.1) which contained 1.0 M NaCl, 0.01 M sodium phosphate, and 0.1 mM EDTA. UV absorbance melting profiles at 260 nm were obtained using a Gilford 250 Spectrophotometer. Samples were initially heated rapidly to above 90°C for a few minutes and then cooled to 1°C to begin the experiment. The heating rate was 0.5°C/min and was controlled by a Gilford 2527 Thermo-programmer. The data shown represent the average of at least 5 independent melting profiles for each hairpin; the data were processed according to Cheong (15). Thermodynamic parameters were determined from

* To whom correspondence should be addressed

+ Present address: Geraldine Brush Cancer Research Institute, Medical Research Institute of San Francisco, 2330 Clay Street, San Francisco, CA 94115, USA

plots of fraction single strand vs. temperature. The variation in the parameters from replicate experiments is within ± 1°C for the T_m, ± 0.2 kcal/mol for $\Delta G°_{37}$, and ± 5% for $\Delta H°$ and $\Delta S°$. The variation in the values for $\Delta H°$ and $\Delta S°$ for two of the DNA hairpins GGAG(GCTT)CTCC and GATC(AAAA)GATC are slightly greater, being within ± 7.5%. Additional significant figures are given for $\Delta H°$ and $\Delta S°$ to allow more accurate calculation of $\Delta G°$ at various temperatures.

To determine if a hairpin was the only species present at 1 M NaCl, 0.01 M sodium phosphate and 0.1 mM EDTA, absorbance melting profiles for the hairpin were measured over at least a hundred-fold range in nucleic acid concentration (10 μM to 1 mM). The similarity of the melting profiles for most of these molecules indicated that the species involved were unimolecular. For some of the molecules, the melting profiles at ten-fold and one hundred-fold the usual nucleic acid concentration indicated the presence of a small amount of a second species with a much lower T_m; most likely the dimer (a duplex with an internal loop). However, at the nucleic acid concentrations used to determine the thermodynamic parameters, there was no indication of the presence of any species other than the hairpin and the single strand. (Hairpin and single strand refer to the two secondary structures of the unimolecular species.)

At 1 M NaCl, the hairpins with the G(UUUG)C and the G(UUUU)C loops showed biphasic melting profiles, indicating the presence of a small but significant amount of another specie, most likely the dimer. Our previous studies at 0.01 M sodium phosphate showed these two hairpins to be significantly less stable than the others in the group (13). This would explain why a small amount of dimer might be present with these two hairpins at the higher salt concentration. Melting curves for the RNA hairpin with a G(CUUG)C loop showed a large concentration dependence indicating significant duplex formation at 1 M NaCl; this molecule was not studied further. SantaLucia et al. (16) have found that internal loops with four U's stabilize a duplex at 37°C; $\Delta G°$ is negative for this internal loop.

RESULTS AND DISCUSSION

RNA hairpins

Certain RNA hairpins are unusually stable such as those containing the loop sequence C(UUCG)G (10, 13). UV absorbance melting studies carried out at 0.01 M sodium phosphate, showed that changing the cytosine of the (UUCG) loop to a uracil decreases the stability significantly as also does changing the loop-closing base pair from C·G to G·C (10, 13). However, since the nearest-neighbor parameters used to estimate the contribution of the stem are for 1 M NaCl, while the previous studies of the (UNCG) hairpins were carried out at 0.01 M sodium phosphate, thermodynamic contributions of the C(UUCG)G and related loops to the stability of their respective hairpins could not be determined. Our measurements in this study were therefore all made with samples at 1 M NaCl, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7.

The thermodynamic parameters for loop formation given in Table 2 were determined by subtracting the nearest-neighbor interactions estimated for the stem from those measured for the hairpin (Table 1). Thermodynamic parameters for RNA nearest-neighbor interactions were those obtained by Turner and coworkers (3) from a set of RNA duplexes at 1 M NaCl, 0.01 M sodium phosphate, 0.1 mM EDTA. The $\Delta H°$, $\Delta S°$ and $\Delta G°_{37}$ values in Table 2 characterize the thermodynamics of initiating a hairpin by forming the loop. The C(UUCG)G loop is clearly the most stable; it has the smallest free energy (+1.0 kcal/mol at 37°C) and the largest negative $\Delta H°$. At 37°C initiating a hairpin with this loop is slightly unfavorable, but because of the negative $\Delta S°$ value the calculated free energy becomes less positive as the temperature decreases. At temperatures below 22°C the free energy becomes negative hairpin initiation becomes spontaneous. The thermodynamic parameters thus predict that the hexanucleotide C(UUCG)G will form a stable loop closed by a single base pair below room temperature. The octanucleotide GC(UUCG)GC does form a loop with a stem of two base pairs melting at 54°C, but absorbance melting experiments on the hexanucleotide C(UUCG)G were inconclusive (17). The next most stable loop is also highly conserved in nature and extra stable: C(UACG)G with $\Delta G°_{37}$ = +1.6 kcal/mol. It is capable of initiating a hairpin spontaneously below 7°C.

The remaining loops shown in Table 2 (with two exceptions) are characterized by $\Delta G°_{37}$ = +3.1 ± 0.2 kcal/mol and negative values of $\Delta H°$ and $\Delta S°$. The loop sequence A(GAAA)U is in this group, although it is more stable than other loops closed by an A·U base pair, as discussed in the Appendix. The negative values of $\Delta H°$ for all these loops mean heat is released on forming the loops and implies an increase in stacking and hydrogen bonding in the loop relative to the single strand. The entropy decrease is consistent with the loss of conformational freedom

Table 1. RNA and DNA Hairpins of the (UNCG)/(TNCG) and the (CUUG)/(CTTG) Families in 1 M sodium chloride, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7.

RNA GGAX(NNNN)X'UCC Hairpin Parameters					DNA GGAX(NNNN)X'TCC Hairpin Parameters				
RNA Loop Sequence	Tm (°C)	ΔH° (kcal/mol)	ΔS° (e.u.)	ΔG°(37) (kcal/mol)	DNA Loop Sequence	Tm (°C)	ΔH° (kcal/mol)	ΔS° (e.u.)	ΔG°(37) (kcal/mol)
C(UUCG)G	76.2	−55.9	−159.9	−6.3	C(TTCG)G	60.4	−31.3	−93.9	−2.2
C(UUUG)G	70.3	−44.0	−128 0	−4.2	C(TTTG)G	59.8	−31.0	−93.0	−2.1
C(UUUU)G	69.6	−44.3	−129.3	−4.2	C(TTTT)G	61.4	−33.5	−100.1	−2.4
G(UUCG)C	67.7	−44.8	−131.4	−4.0	G(TTCG)C	53.8	−30.2	−92.5	−1.5
G(UUUG)C	*	*	*	*	G(TTTG)C	54.6	−31.7	−96.8	−1.7
G(UUUU)C	*	*	*	*	G(GCTT)C	51.2	−27.2	−83.7	−1.2
G(CUUG)C	*	*	*	*	G(CTTG)C	70.7	−42.0	−122.2	−4.1
C(GCUU)G	70.9	−45.0	−130.8	−4.4	C(GCTT)G	58.5	−29.3	−88.3	−1.9
C(UACG)G	73.8	−53.6	−154.5	−5.7	C(dUdUCG)G	60.8	−31.8	95.3	−2.3

*Evidence for duplex formation was seen in the melting curves

in loop formation. The extra stable loops have the largest decrease in enthalpy and entropy, as expected.

The exceptionally *unstable* loops, C(CCCC)G (6) and A(CCCC)U (18) have unfavorable loop free energies of greater than +6.3 kcal/mol at 37°C and positive values of $\Delta H°$ (> +25 kcal/mol) and $\Delta S°$ (> +58 eu). The stability of these loops *increases* with increasing temperature. The positive value of $\Delta H°$ means that energy must be added to form the loop. Steric repulsions or other repulsive interactions may be caused by formation of the loop, or stem base pairs may be disrupted by loop formation. Alternatively, favorable stacking and/or hydrogen bonding interactions in the single strand may be lost on loop formation. The increase in entropy on forming the loop requires that water molecules or other bound species are released when the tetra-C loop forms. Groebe and Uhlenbeck (6) suggested that the apparent thermodynamic instability of their cytosine-containing loops was caused by formation of C·G base pairs in the partially melted hairpins. However, this cannot explain the similar thermodynamic instability of $A_6(CCCC)U_6$ hairpins (18). The similar results seen in Table 2 for the different tetra-C loops indicate that the thermodynamics of loops containing only cytosines are unique among all the RNA loops studied. It is unfortunate that this was the first RNA hairpin loop studied and served as the model RNA hairpin for many years (18, 19).

DNA hairpins

Thermodynamic parameters for the loops in Table 3 were obtained by subtracting the nearest-neighbor contributions estimated for the stem from those measured for the hairpin. DNA nearest-neighbor parameters were those obtained by Breslauer and co-workers (4) from a set of DNA duplexes at 1 M NaCl, 0.01 M sodium phosphate, 0.1 mM EDTA. The patterns of DNA loop stabilities are different from those of the RNA loops, but the magnitudes of the thermodynamic parameters are similar. The most stable loops have $\Delta G°_{37} = +1.5 \pm 0.2$ kcal/mol, which include the sequences C(GNAA)G, C(GATA)G, and G(C-TTG)C. These loops are also extra stable in RNA (J. Haney and O. Uhlenbeck, personal communication, and Ref. 13). There is a wide range of loop sequences with $\Delta G°_{37} = +3.0 \pm 0.2$ kcal/mol; these vary from C(TTTT)G to C(AAAA)G and include C(TTCG)G, corresponding to the most stable RNA tetraloop. Neither of the two DNA C(TTCG)G hairpins, nor the C(dUdUCG)G hairpin, were more stable than the corresponding C(TTTT)G hairpins, indicating that these DNA hairpins did not form unusually stable loop conformations, unlike the RNA hairpins with C(UUCG)G loops. Switching the closing base pair from C·G to G·C for three DNA loops increased the free energy of loop formation by about 1 kcal/mol. All the loops have negative enthalpies and entropies of formation.

For the tetraloops we studied, a free energy of loop formation (at 37°C) of about +3 kcal/mol is most common for either RNA or DNA. There are extra stable loops with $\Delta G°_{37}$ near +1 kcal/mol, but the sequences are not necessarily the same for RNA and DNA. The closing base pair is also important; changing from C·G to G·C lowered the stability of several tetraloops in both RNA and DNA. We are not certain if this holds for RNA loops in general that close with a G·C base pair, since the only member of the RNA group for which we were able to obtain data (at 1 M NaCl) is unusually stable. However, it seems likely that RNA loops that close with G·C (and are not extra stable) would have a $\Delta G°_{37}$ close to +4 kcal/mol since the hairpin with the G(UUCG)C loop was shown to be more stable (at the lower salt concentration) by about −1.7 kcal/mol than hairpins with G(UUUG)C and G(UUUU)C loops (13). Thus the difference in stability between the DNA and RNA hairpins could almost entirely be attributed to the differences in nearest-neighbor interactions of the stem. Although more extensive studies are necessary, as an approximation one might be able to use similar stability values for DNA and RNA hairpin loops, provided one takes into account the loop-closing base pair and whether the loop is unusually stable or not.

RNA ...G(CUUG)C... and DNA ...G(CTTG)C... hairpins

G(CUUG)C is not a common loop sequence in ribosomal RNAs. However, it is very highly conserved at position 83 of 16S RNAs

Table 2. Thermodynamics of Tetraloop Formation in RNA in 1 M sodium chloride, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7.

RNA Loop Sequence	Hairpin Sequence	$\Delta H°$ (kcal/mol)	$\Delta S°$ (e.u.)	$\Delta G°(37)$ (kcal/mol)
C(UUCG)G	(a)	−20.2	−68.5	1.0
C(UACG)G	(a)	−17.9	−63.1	1.6
G(UUCG)C	(b)	−11.7	−47.0	2.9
C(GCUU)G	(a)	−9.3	−39.4	2.9
C(UUUU)G	(a)	−8.6	−37.9	3.1
C(UUUG)G	(a)	−8.3	−36.6	3.1
A(GAAA)U	(c)	−12.9±1.7	−51.7±6.1	3.2±0.1
C(UUUU)G	(d)	−4.3	−23.3	3.0
C(AAAA)G	(d)	−0.7	−13.5	3.6
C(CCCC)G	(d)	24.9	58.4	6.9
A(CCCC)U	(e)	28	70	6.3

(a) GGAC(NNNN)GUCC; data from Table 1.
(b) GGAG(NNNN)CUCC; data from Table 1.
(c) This represents the average of data from six hairpins with different stems given in Table A1.
(d) Data from Ref. 6; the hairpin sequence is GGGAUAC(NNNN)GUAUCCA.
(e) Data from Ref. 18; the hairpin sequence is AAAAAA(CCCC)UUUUUU.

Table 3. Thermodynamics of Tetraloop Formation in DNA in 1 M sodium chloride, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7.

DNA Loop Sequence	Hairpin Sequence	$\Delta H°$ (kcal/mol)	$\Delta S°$ (e.u.)	$\Delta G°(37)$ (kcal/mol)
C(GTAA)G	(a)	−13.9	−49.0	1.3
C(GATA)G	(a)	−14.3	−50.6	1.4
C(GCAA)G	(a)	−13.6	−48.4	1.4
C(GAAA)G	(a)	−11.1	−41.9	1.8
G(CTTG)C	(b)	−17.6	−61.3	1.5
C(TTTT)G	(a)	−10.9	−43.5	2.6
C(TTTT)G	(c)	−10.4	−42.7	2.9
C(dUdUCG)G	(a)	−10.4	−42.1	2.7
C(dUdUCG)G	(c)	−8.7	−37.9	3.0
C(dUTCG)G	(a)	−9.8	−40.6	2.7
C(TdUCG)G	(a)	−8.6	−36.7	2.8
C(TTCG)G	(a)	−9.8	−40.8	2.8
C(TTCG)G	(c)	−5.9	−36.5	3.1
C(TTTG)G	(c)	−7.9	−35.6	3.2
C(GCTT)G	(c)	−6.2	−30.9	3.4
C(AAAA)G	(a)	−3.3	−21.6	3.4
G(TTTG)C	(b)	−7.3	−35.9	3.9
G(TTCG)C	(b)	−5.9	−31.6	4.1
G(GCTT)C	(b)	−2.8	−22.8	4.4

(a) GATC(NNNN)GATC; data from Ref. 13.
(b) GGAG(NNNN)CTCC; data from Table 1.
(c) GGAC(NNNN)GTCC; data from Table 1.

(9). G·C is almost exclusively the loop-closing base pair. At low salt concentrations (0.01 M sodium phosphate) we found the hairpin containing the G(CUUG)C loop to be extra stable (13), but we were not able to study it in 1 M NaCl because of dimer formation.

The DNA G(CTTG)C loop was much more stable than most of the DNA loops shown in Table 3. In fact, it was as stable as the DNA C(GNRA)G loops, showing yet again that a C·G loop-closing base pair is not a prerequisite for obtaining a very stable loop conformation.

DNA GATC(NNNN)GATC hairpins

The loop sequences (GAAA) and (GCAA) are the most frequently-occurring loop sequences in ribosomal RNA (8, 9). We have determined that RNA (GAAA) loops are unusually stable when the loop-closing base pair is A·U (see Appendix). RNA hairpins with the consensus loop sequence C(GNRA)G also formed unusually stable hairpins (Haney and Uhlenbeck, personal communication). Hirao *et al.* (20) found that the DNA oligomer GCGAAAGC formed a hairpin with a C(GAAA)G loop. Among the DNA loops we considered to be typical [$\Delta G°_{37}$ = +3.0 ± 0.2 kcal/mol], C(TTTT)G was the most stable with $\Delta G°_{37}$ = +2.6 kcal/mol. The top four loops listed in Table 3 contributed free energies of −0.8 to −1.3 kcal/mol more than the C(TTTT)G loop to the stability of their respective hairpins, and so we designated them as being unusually stable tetra-loops. It should be noted that C(GATA)G does not fall into the (GNRA) consensus loop sequence, nevertheless it is extra stable. The $\Delta G°_{37}$ values for these extra stable DNA loops were similar to those for the extra stable RNA (UUCG) and (UACG) loops (with C·G closing the loops), as well as the DNA G(CTTG)C loop (with G·C closing the loop).

SUMMARY

The thermodynamic data obtained for the hairpins at 1 M NaCl allow for the direct comparison of the stabilities of various loops; the contributions of the loop being the difference between the parameters estimated for the stem at 1 M NaCl and those we measured for the hairpin at the same ionic strength.

With the exception of the (UNCG) loops, our data show that the RNA tetraloops and their DNA sequence analogs have similar stabilities, provided they have the same loop-closing base pair. The DNA (TNCG) loops are not extra stable and thus have thermodynamic parameters similar to those of the other RNA and DNA loops that are not extra stable, but with the same loop-closing base pair. Thus, the fact that RNA hairpins are considerably more stable than the analogous DNA hairpins appears to be completely accounted for by the differences between the stabilities of the stems for RNA and DNA.

The RNA tetraloops that fall into the (UNCG) and (GNRA) families are not the only ones that are unusually stable. Also to be included in this list are the G(CUUG)C (13) and the C(G-AUA)G (Haney and Uhlenbeck, personal communication) loops. Except for the (TNCG) loops, all the other DNA analogs of these unusually stable RNA loops are also extra stable. Caution must therefore be exercised when assigning stability bonuses to loops when predicting RNA and DNA secondary structures.

Our data clearly show that for unusually stable loops, the sequence is very important. This is consistent with the structural studies of the (UUCG)- and (GAAA)-containing hairpins, where it is evident that specific interactions between nucleotides result in the increased stability of these hairpins (11, 12). However, for hairpins that are not extra stable, the sequence of the loop (at least in the cases we studied) did not appear to make a significant difference.

Our study contains three pairs of DNA hairpins with the same loop and loop-closing base pair, but different (50% vs. 75% G·C) stems (Table 2 and Ref. 13), and six RNA hairpins with the same loop and loop-closing base pair, but different sequence (all A·U/U·A) and lengths (4 or 5 base pairs) of the stem (Table A1). The loops with the same sequence have similar thermodynamic parameters, suggesting that, besides the identity of the loop-closing base pair, the rest of the stem sequence, or its length, may not greatly influence the stability of the loop.

Two major factors appear to determine the contribution of different loops to the thermodynamic stability of hairpins; the identity of the loop-closing base pair and whether or not an extra stable loop structure can form. Given these criteria, small variations within the loop sequence, or the stem sequence or size, or even if the hairpin is DNA or RNA, do not seem to significantly change the thermodynamic parameters of the loop. Thus, extra stable loops with C·G as the closing base pair (2 RNA and 4 DNA) had $\Delta G°_{37}$ values ranging from +1.0 to +1.8 kcal/mol, while the DNA loop with G·C closing the loop was not too different, with a $\Delta G°_{37}$ value of +1.5 kcal/mol. For the six extra stable RNA loops closed by A·U [sequence A(GAAA)U], the mean $\Delta G°_{37}$ was +3.2 ± 0.1 kcal/mol, significantly different from those described above for RNA and DNA loops closed with a C·G base pair ($\Delta G°_{37}$ = +1.4 ± 0.3 kcal/mol) or a G·C base pair ($\Delta G°_{37}$ = +1.5 kcal/mol). The mean $\Delta G°_{37}$ value for loops with C·G as the closing base pair and not extra stable (3 RNA and 11 DNA) was +3.0 ± 0.3 kcal/mol. For loops with G·C as the closing base pair and not extra stable (3 DNA), the $\Delta G°_{37}$ ranged from +3.9 to +4.4 kcal/mol. Our study had no RNA loops with G·C as the closing base pair which were not extra stable. However, from our previous study of these hairpins at 0.01 M sodium phosphate (13), we estimate that the free energy for these RNA loops would lie in the same range as for their DNA counterparts. The thermodynamic parameters determined for the 15 RNA and 19 DNA hairpins add significantly to the thermodynamic data that is presently available for predicting RNA and DNA secondary structures, and should make such predictions more accurate.

ACKNOWLEDGEMENTS

We thank Mr. David Koh for synthesizing the DNA molecules used in this study, and Ms. Barbara Dengler for managing the laboratory. We are grateful to Ms. Kristina Downing, Ms. Arlene Lee, Ms. Sandy Lai, and Mr. Marc Canabou, all undergraduate students in the laboratory, for their help with some of the experiments, and Dr. John Jaeger, Dr. John SantaLucia, and Dr. Gabriele Varani for their helpful suggestions. Professor Olke Uhlenbeck, U. of Colorado, Boulder kindly made his unpublished results available to us. This research was supported in part by the National Institutes of Health grant GM 10840 and the Department of Energy, Office of Energy Research, Office of Health and Environmental Research grant DE-FG03-86ER60406.

REFERENCES

1. Jaeger, J., Turner, D. H., and Zuker, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7706–7710.
2. Zuker, M., Jaeger, J. A., & Turner, D. H. (1991) *Nucleic Acids Res.* **19**, 2707–2714.
3. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
4. Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
5. Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17**, 167–192.
6. Groebe, D. R. & Uhlenbeck, O. C. (1988) *Nucleic Acids Res.* **16**, 11725–11735.
7. Longfellow, C. E., Kierzek, R., & Turner, D. H. (1990) *Biochemistry* **29**, 278–285.
8. Gutell, R. R., Weiser, B., Woese, C. R., & Noller, H. F. (1985) *Prog. Nucleic Acids Res. and Mol. Biol.* **32**, 155–216.
9. Woese, C. R. , Winker, S., & Gutell, R. R. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8467–8471.
10. Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., d'Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I. Jr., Brody, E. N., & Gold, L. (1988) *Proc. Natl . Acad. Sci. USA* **85**, 1364–1368.
11. Varani, G., Cheong, C., & Tinoco, I. Jr. (1991) *Biochemistry*, **30**, 3280–3289.
12. Heus, H. H., and Pardi, A. (1991) *Science*, **253**, 191–194.
13. Antao, V. P., Lai, S. Y., & Tinoco, I. Jr. (1991) *Nucleic Acids Res.* **19**, 5901–5905.
14. Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
15. Cheong, C. (1990) Ph.D. Thesis. University of California, Berkeley, CA.
16. SantaLucia, J. Jr., Kierzek, R., & Turner, D. H. (1991) *Biochemistry* **30**, 8242–8251.
17. Molinaro, M., & Tinoco, I. Jr. (1991) *Biophys. J.* **59**, 492a.
18. Uhlenbeck, O. C., Borer, P. N., Dengler B., & Tinoco, I. Jr. (1973) *J. Mol. Biol.* **73**, 483–496.
19. Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol.* **73**, 497–506.
20. Hirao, I., Nishimura, Y., Naraoka, T., Watanabe, K., Arata, Y., & Miura, K. (1989) *Nucleic Acids Res.* **17**, 2223–2231.
21. Sugimoto, N., Kierzek, R., & Turner, D. H. (1987) *Biochemistry* **26**, 4554–4558.

APPENDIX

RNA hairpins and the nearest-neighbor approximation

Although RNA secondary conformations consist largely of stem and loop structures, their stabilities are often predicted using a nearest-neighbor model and free-energy parameters derived from short RNA duplexes. The hairpins in this section were designed to test this model and determine if the thermodynamic parameters for the three nearest-neighbors (AA/UU), (AU/AU), and (UA/UA) obtained using hairpins were in agreement with those derived using small RNA duplexes (5). All these hairpins have A(GAAA)U loops, with A·U being the loop-closing base pair. The stems consist of A·U and U·A base pairs. We also synthesized RNA molecules with the same stem sequences as hairpins A1(a), A1(d), and A1(h) of Table A1, but with A(UUUU)U loops. Although the melting profiles for these three molecules were biphasic, it was clear that these A(UUUU)U hairpins were much less stable than the corresponding hairpins with A(GAAA)U loops (ΔT_m were about 10 to 15°C). Thus, the hairpins with A(GAAA)U are extra stable; Uhlenbeck and Haney (personal communication) have previously shown that RNA hairpins with C(GAAA)G are extra stable.

In addition to having the same loop, loop-closing base pair, and 5′ dangling nucleotides, hairpins A1(a) through A1(d) also have the same nearest-neighbors in the stem. Since the thermodynamic parameters of a hairpin can be expressed as the sum of stem and loop contributions, all four hairpins should have similar melting temperatures and thermodynamic parameters. The fact that hairpins A1(a), A1(b), and A1(c) do have similar thermodynamics (Table A1), indicates that the nearest-neighbor model for the stacking and base pairing of the stem of these hairpins is valid. However, there is a definite difference in the T_m (of about 2°C) and $\Delta G°_{37}$ (of about 0.3 kcal/mol) between hairpin A1(d) and hairpins A1(a), A1(b), and A1(c). Hairpin A1(d) differs from the other three in that the (AA/UU) nearest-neighbor is adjacent to the loop, resulting in a run of six purines, five of them adenines. The tendency of adenines to stack in the single-stranded form could shift the equilibrium more towards the single-stranded form for A1(d). Hairpins A1(e) and A1(f) have the same nearest neighbors and have similar thermodynamic properties as expected. There is a slight destabilization by adding another adenine to the run of purines, but the effect is small (ΔT_m = 0.5°C and $\Delta G°_{37}$ = +0.1 kcal/mol). However, about 20% higher (unfavorable) $\Delta G°_{37}$ values would be obtained for (AA/UU) nearest neighbors if hairpins A1(d), A1(e), and A1(f) were used in the data set. As they are relatively unstable compared with the other hairpins with the same nearest-neighbors, but without a run of purines extending from the stem into the loop, we omitted them in the calculation of stem nearest-neighbor parameters (Table A2). We also omitted them in the data for the A(GAAA)U loop parameters given in Table 2.

Table A1. RNA ...A(GAAA)U... Hairpins in 1 M sodium chloride, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7.

RNA Hairpin	RNA Stem Sequence	Tm (°C)	ΔH° (kcal/mol)	ΔS° (e.u.)	ΔG°(37) (kcal/mol)
A1(a)	GGUUAUA AAUAU	44.1	−41.4	−130.4	−0.9
A1(b)	GGUAAUA AUUAU	44.7	−40.8	−128.4	−1.0
A1(c)	GGUAUUA AUAAU	43.9	−40.5	−127.8	−0.9
A1(d)	GGUAUAA AUAUU	42.3	−38.3	−121.4	−0.6
A1(e)	GGAAUAA UUAUU	41.9	−39.0	−124.0	−0.6
A1(f)	GGAUAAA UAUUU	41.4	−35.4	−112.6	−0.5
A1(g)	GGAUAUA UAUAU	46.7	−42.8	−133.8	−1.3
A1(h)	GGUAUA AUAU	38.1	−33.7	−108.3	−0.1
A1(i)	GGAAUA UUAU	37.9	−34.4	−110.7	−0.1

Table A2. Comparison of Hairpin-derived RNA Nearest-Neighbors with Duplex-derived values (Ref. 5) in 1 M sodium chloride, 0.01 M sodium phosphate, 0.1 mM EDTA, at pH 7.

Nearest Neighbor	ΔH° (kcal/mol)		ΔS° (e.u)		ΔG°(37) (kcal/mol)	
	This work	Ref. 5	This work	Ref. 5	This work	Ref. 5
5′-AA-3′ 3′-UU-5′	−7.2	−6.6	−20.6	−18.4	−0.8	−0.9
5′-AU-3′ 3′-UA-5′	−9.8	−5.7	−28.3	−15.5	−1.0	−0.9
5′-UA-3′ 3′-AU-5′	−5.7	−8.1	−15.4	−22.6	−1.0	−1.1

Thermodynamic values for (AA/UU), (AU/AU), and (UA/UA) nearest-neighbors (Table A2) were obtained by combining the parameters for a four base pair hairpin, either A1(h) or A1(i), with the parameters from the appropriate five base pair hairpin (Table A1). The (AU/AU) and (UA/UA) nearest-neighbor values also require knowledge of the difference in thermodynamic values for a 5′ dangling G stacked on an A·U base pair vs. a U·A base pair. Values are available for GA·U ($\Delta G°_{37}$ = −0.4 kcal/mol, $\Delta H°$ = +0.7 kcal/mol, $\Delta S°$ = +3.4 eu) (5, 21), but only $\Delta G°_{37}$ = −0.2 kcal/mol) is given for GU·A (5). We arbitrarily set $\Delta H°$ equal to zero for GU·A; the average value for all 5′ dangling ends measured is +0.6 kcal/mol.

The $\Delta G°_{37}$ nearest-neighbor values agree with the (much broader based) values derived from small RNA duplexes (within 0.1 kcal/mol). This agreement is encouraging and does not require any changes in the duplex nearest-neighbor free energy parameters used in secondary structure prediction programs. The $\Delta H°$ and $\Delta S°$ values differ significantly from the duplex-derived values (5), especially for the (AU/UA) and (UA/UA) nearest-neighbors. The greater uncertainty (compared to $\Delta G°$) in the values of $\Delta H°$ and $\Delta S°$ derived from either hairpins or duplexes does not allow a conclusion about which values are best for calculating secondary structures at temperatures other than 37°C.