

- and it should probably be included explicitly.
35. L. F. Kuyper, R. N. Hunter, D. Ashton, K. M. Merz, Jr., P. A. Kollman, *J. Phys. Chem.* **95**, 6661 (1991).
36. P. Cieplak, P. Bash, U. C. Singh, P. A. Kollman, *J. Am. Chem. Soc.* **109**, 6283 (1987).
37. P. Beak, F. S. Fry, Jr., J. Lee, F. Steele, *ibid.* **98**,

- 171 (1976).
38. We are grateful to C. Lim for helpful discussions. This work was supported in part by the National Science Foundation.

11 December 1991; accepted 19 February 1992

## Context Dependence of Hydrogen Bond Free Energy Revealed by Substitutions in an RNA Hairpin

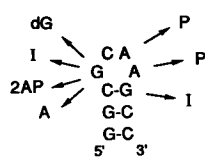
John SantaLucia, Jr., Ryszard Kierzek, Douglas H. Turner\*

Prediction and modeling of RNA structure requires knowledge of the free energy contributions of various interactions. Many unusual hydrogen bonds were recently proposed in the structure of a GCAA hairpin determined from nuclear magnetic resonance. The contributions of these hydrogen bonds to the folding stability of the hairpin formed by rG-GCGCAAGCC have now been investigated through the use of functional group substitutions. These and previous results suggest a strong context dependence for the free energy of hydrogen bond formation. The results also suggest that the phylogenetic preference for GNRA (where N = A, C, G, or U and R = A or G) tetraloops may have a functional rather than thermodynamic basis.

The functional diversity of RNA reflects diversity in three-dimensional (3-D) structure. The 3-D folding of RNA is apparently controlled by the RNA sequence, because proteins are usually not required for proper folding (1-4). Prediction of RNA structure from sequence has focused largely on secondary structure (5), and predictions have improved as relations between structure and free energy have been determined (5-7). In principle, prediction of 3-D folding is amenable to a similar approach and requires a knowledge of the standard free energy ( $\Delta G^\circ$ ) contributions of individual interactions, such as hydrogen bonding, stacking, and hydration in diverse structural contexts. We report  $\Delta G^\circ$  increments for hydrogen bonds in an RNA hairpin loop, GCAA, for which the structure has been determined with nuclear magnetic resonance (NMR) (8). The GCAA sequence is one of a group of sequences, designated "tetraloops," that occur commonly in ribosomal and other RNAs (9, 10).

Substitution of hydrogen-bonding groups provides a method for determining the free energy contributions of hydrogen bonds to macromolecular properties (11-15). This "atomic mutation" approach is used here. For example, hydrogen bonds involving the amino group of guanosine are studied by substituting guanosine with inosine (I),

thus replacing the G amino group with a hydrogen. The  $\Delta G^\circ$  increment for the hydrogen bonds is then defined as the difference in the  $\Delta G^\circ$  of folding for the hairpin with and without the amino group. The substitutions made are shown in Fig. 1.

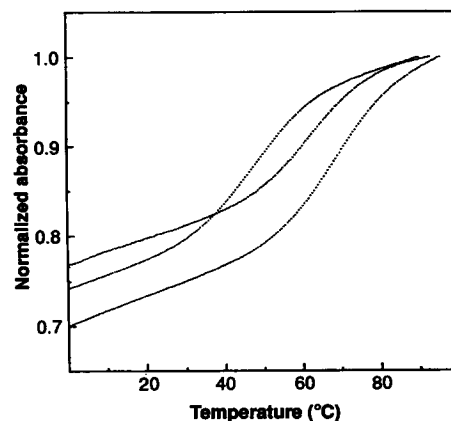


**Fig. 1.** Summary of individual substitutions made in the GGC-GCAAAGCC hairpin. P, purine; 2AP, 2-aminopurine; and dG, deoxyguanosine.

**Fig. 2.** Typical melting curves. Sequences from highest to lowest melting temperature are GGC-GCAAAGCC (with 0.1 M NaCl), GGCICAAAGCC (with 0.1 M NaCl), and GGC-GCAAAGCC (without NaCl). Oligoribonucleotides were synthesized on solid support with the phosphoramidite method (12, 23). After deblocking with ammonia and with acid, the RNA was purified on a Si500F thin-layer chromatography plate (Baker) and eluted for 5 hours with *n*-propanol-ammonia-water (55:35:10 by volume) (24). Bands were visualized with an ultraviolet lamp, and the least mobile band was cut out and eluted three times with 1 ml of doubly distilled water. Oligomers were further purified with a Sep-pack C-18 cartridge (Waters). Oligomers for low-salt melts were desalted by continuous-flow dialysis (BRL). Purities were checked with analytical C-8 high-pressure liquid chromatography (Beckman) and were >95%. The buffer for most melting curves was 0.1 M NaCl, 10 mM sodium phosphate, and 0.5 mM Na<sub>2</sub>EDTA at pH 7. For low-salt experiments, 0.1 M NaCl was omitted. Absorbance versus temperature curves were measured at 280 nm with a heating rate of 1°C per minute on a Gilford 250 spectrophotometer. Experiments with a heating rate of 0.5°C per minute gave the same results as those at 1°C per minute, suggesting thermodynamic equilibrium was achieved. Oligomer concentration was varied roughly 100-fold. Absorbance versus temperature curves were fit to a two-state model with sloping baselines through the use of a nonlinear least-squares program.

The reference sequence is rGGCGCAA-GCC, nucleotides 377 to 386 in *Escherichia coli* 16S ribosomal RNA (rRNA). This is a shortened form of the pppGGCGCAA-GCCUUAU sequence used for structure determination with NMR (8). Thermodynamic parameters for folding of the hairpins were derived from fits of melting curves to a two-state model (Fig. 2 and Table 1). The  $\Delta G^\circ$  values at 70°C are more reliable because each requires a small extrapolation from the melting temperature,  $T_M$ . Except for rGGCGCAAAGCC, all of the sequences have concentration-independent  $T_M$ 's in 0.1 M NaCl, indicating hairpin formation. For rGGCGCAAAGCC in 0.1 M NaCl, concentration-dependent  $T_M$ 's were observed, consistent with duplex formation (see footnote to Table 1). In order to quantify the effect of the G to I substitution in the closing base pair, rGGCGCAAAGCC and rGGCGCAAAGCC were melted in 10 mM phosphate buffer without added NaCl. Under these conditions, the  $T_M$ 's are independent of concentration. The G to I substitution in the closing base pair results in a 16°C drop in  $T_M$ , which corresponds to a less favorable (by 1.3 kcal/mol)  $\Delta G^\circ$  at 70°C (see Table 1). This result is similar to  $\Delta G^\circ$  decrements observed upon removal of hydrogen-bonding amino groups in GC pairs and GA mismatches (11, 12).

Whereas a hydrogen bond in a stem base pair contributes more than 1 kcal/mol to hairpin stability, the results in Table 1 indicate that hydrogen bonds within the hairpin all contribute less than 1 kcal/mol (Fig. 3). The largest effect, 0.7 kcal/mol, is observed for the G to I substitution in the GCAA loop. The NMR structure of the loop (8) suggests that the amino group thus



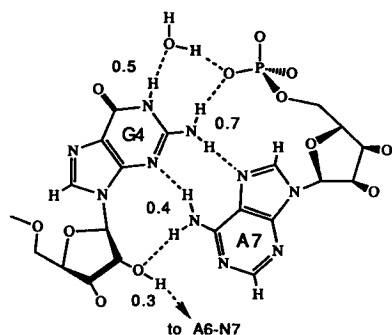
J. SantaLucia, Jr., and D. H. Turner, Department of Chemistry, University of Rochester, Rochester, NY 14627-0216.

R. Kierzek, Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-704 Poznan, Noskowskiego 12/14, Poland.

\*To whom correspondence should be addressed.

replaced is involved in hydrogen bonds to N7 of the last A in the loop and to a phosphate (Fig. 3). The thermodynamic results suggest that neither hydrogen bond contributes substantially to hairpin stability, in contrast with previous results for hydrogen bonds to charged groups. For example, a  $\Delta G^\circ$  decrement of 1.3 kcal/mol at 70°C was reported for a UUCG tetraloop when U was substituted for a C that had an amino group with a single hydrogen bond to a phosphate (16).

The NMR structure (8) suggests the amino group of the last A in the GCAA loop hydrogen bonds to both the N3 and O2' of the loop G (Fig. 3). Replacing this amino group with a hydrogen results in a  $\Delta G^\circ$  decrement of only 0.4 kcal/mol, which



**Fig. 3.** Summary of  $\Delta G^\circ$  changes for folding of the GGCGCAAGCC hairpin upon substitution of individual functional groups. Except for the G carbonyl oxygen and imino proton, functional groups were individually replaced with H. The G carbonyl oxygen and imino proton are replaced by a H at C6. Dashed lines indicate hydrogen bonds suggested by the NMR structure (8). The water-mediated hydrogen bond between the G imino group and the phosphate between the As in the GCAA loop is proposed on the basis of the work reported here. The  $\Delta G^\circ$  decrements at 70°C are given in kilocalories per mole for substitution of the closest functional group (Table 1).

falls between the decrements of 1.2 and 0.0 kcal/mol mismatch at 70°C observed for the same substitution in the GA mismatches in the internal loop of (GGCGAGCC)<sub>2</sub> and at the helix termini of (AGCGCG)<sub>2</sub> (12). The NMR structure also suggests that hydrogen bonds are formed between the 2'-OH of G in the GCAA loop and the amino group of the second A and the N7 of the first A (Fig. 3). Replacing the 2'-OH with H gives a  $\Delta G^\circ$  decrement of only 0.3 kcal/mol, which is less than the decrements of ~1 kcal/mol observed in binding to a group I ribozyme upon substituting H for single 2'-OH groups in the substrate (14, 15). Evidently, pairs of hydrogen bonds in the GCAA loop are worth less than single hydrogen bonds in other contexts.

The NMR structure (8) of the GCAA hairpin suggests the N7's of the first and second loop A's are hydrogen-bonded to the 2'-OH and amino groups of the loop G, respectively. Simultaneous replacement of both loop adenosines with 7-deazaadenosine lowered  $T_M$  by 4°C (8), consistent with a weak hydrogen-bonding  $\Delta G^\circ$  of -0.5 kcal/mol at 70°C. The results for ICAA, GCAP, and (dG)CAA loops are also consistent with weak hydrogen bond free energies. It seems unlikely that the weak hydrogen bond from the first loop A N7 to the loop G 2'-OH is responsible for the strong phylogenetic preference for a purine in the third position in the loop. Alternative explanations are (i) steric requirements for a small five-membered ring, rather than a more bulky six-membered pyrimidine ring; (ii) greater stacking or space filling for a purine versus pyrimidine base; or (iii) both effects.

The NMR structure (8) has no hydrogen bonds to the amino protons of the first A in the GCAA loop. As expected, substituting H for the NH<sub>2</sub> of the first loop A has essentially no effect on hairpin stability at 70°C. In contrast, although the NMR struc-

ture has no hydrogen bonds to the carbonyl and imino protons of the G in the loop, replacing the carbonyl O with H and eliminating the imino proton by substituting 2-aminopurine for G results in a  $\Delta G^\circ$  decrement of 0.5 kcal/mol. Model building suggests that a water-mediated hydrogen bond could form between the G imino proton and the phosphate between the A's in the GCAA loop (Fig. 3). If this interpretation is correct, a water molecule that bridges two functional groups from different nucleotides can contribute significantly to RNA stability. Such a structure would also allow rationalization of the sharp NMR resonance that is observed for the G imino proton even though it is exposed to solvent (8). For internal loops with similar GA pairing, this imino resonance is broad (12, 17-19).

The above results indicate that  $\Delta G^\circ$  increments for hydrogen bonds are dependent on context and that hydrogen bonds within the GCAA hairpin contribute relatively little to thermodynamic stability. If the dielectric constant, the replacement of a deleted hydrogen-bonding group by water, or both, are factors, then the  $\Delta G^\circ$  increment for a hydrogen bond may depend on the position of a given motif in the 3-D structure of an RNA. For example, if a GCAA loop is buried inside an RNA or a protein, the hydrogen bonds may be more important in maintaining the structure of the loop. Chemical mapping of *E. coli* 16S rRNA indicates that 10 of 11 tetraloop sequences are either strongly or moderately modified by chemical probes (3), suggesting they are accessible to water. Thus the  $\Delta G^\circ$  increments measured here for an isolated loop are likely to be representative of those found in a large RNA.

The GCAA tetraloop is only marginally more stable than other hairpin loops of four nucleotides. The GCAA hairpin has a melting temperature 8°C higher than

**Table 1.** Thermodynamic parameters of hairpin formation; experimental conditions described in Fig. 2. Values are averages of parameters measured for four different oligomer concentrations. Errors are standard deviations. Errors in  $T_M$  are estimated as  $\pm 1^\circ\text{C}$ . P is purine and 2 is 2-aminopurine.

Sequence	$\Delta G^\circ_{70}$ (kcal/mol)	$\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (cal/mol · K)	$T_M$ (°C)	$\Delta\Delta G^\circ_{70}^*$ (kcal/mol)
No NaCl						
rGGCGCAAGCC	0.21 ± 0.04	-2.95 ± 0.08	32.7 ± 0.5	96.0 ± 1.5	67.8	
rGGCGCAAICCT†	1.49 ± 0.13	-1.19 ± 0.31	26.4 ± 4.6	81.3 ± 13.2	51.5	-1.28 ± 0.14
0.1 M NaCl						
rGGCGCAAGCC	-0.09 ± 0.05	-3.00 ± 0.23	30.3 ± 1.9	88.0 ± 5.6	71.0	
rGGCICAAGCC	0.56 ± 0.05	-2.46 ± 0.17	30.8 ± 1.8	91.5 ± 5.3	63.9	-0.65 ± 0.07
rGGCGCAPGCC	0.30 ± 0.04	-2.72 ± 0.15	31.1 ± 1.1	91.4 ± 3.0	66.8	-0.39 ± 0.06
rGGC(dG)CAAGCC	0.19 ± 0.03	-3.00 ± 0.06	33.0 ± 1.0	96.6 ± 3.0	68.1	-0.28 ± 0.06
rGGC2CAAGCC	0.39 ± 0.11	-2.33 ± 0.25	27.9 ± 1.1	82.3 ± 2.7	65.2	-0.48 ± 0.12
rGGCGCPAGCC‡	-0.14 ± 0.16	-3.60 ± 0.22	36.1 ± 1.4	104.7 ± 4.4	71.4	0.05 ± 0.17
rGGCAACAAGCC	0.64 ± 0.04	-2.23 ± 0.12	29.2 ± 1.4	86.9 ± 4.2	62.7	-0.73 ± 0.06

\* $\Delta\Delta G^\circ_{70} = \Delta G^\circ_{70}(\text{rGGCGCAAGCC}) - \Delta G^\circ_{70}(\text{sequence with substitution})$ . Errors are the square root of the sum of the squares of the errors of  $\Delta G^\circ_{70}$ . †Under conditions of 0.1 M NaCl, 10 mM sodium phosphate, 0.5 mM Na<sub>2</sub>EDTA, and pH 7, rGGCGCAAICCT melts like a duplex with the following thermodynamic parameters:  $\Delta G^\circ_{37} = -8.3$  kcal/mol,  $\Delta H^\circ = -46.7$  kcal/mol,  $\Delta S^\circ = -124.0$  cal/mol · K, and  $T_M$  (10<sup>-4</sup> M) = 55.4°C. ‡Results for rGGCGCPAGCC are less reliable than those for other sequences because the low-temperature baseline slopes for all concentrations are very steep, suggesting competition between alternate conformations, which thus compromises the two-state approximation.

ACAA, but this represents only a 0.7-kcal/mol decrease in  $\Delta G^\circ$  of folding at 70°C. An equal or larger gain in  $\Delta G^\circ$  could be achieved by changing any AU base pair in a stem to GC (6). With the  $\Delta G^\circ$ s of Freier *et al.* (6) for base pairs and the approximation of a terminal GA mismatch as AG (12), the hairpin loops of CAAAAG and CUUUUG in 1 M NaCl studied by Groebe and Uhlenbeck (20) have loop free energies of 4.2 and 3.9 kcal/mol at 70°C, slightly more stable than the 5.1 kcal/mol for the GCAA loop in 0.1 M NaCl studied here. Thermodynamic parameters measured by Antao *et al.* (21) show that a CGAAAG hairpin loop is only 0.3, 0.5, and 0.2 kcal/mol more stable at 70°C than CGC-UUG, CUUUUG, and CUUUGG loops, respectively. The phylogenetic preference for GNRA tetraloops (9) is probably due to nonthermodynamic factors, such as the ability to form tertiary contacts (22). Woese *et al.* (9) show that given tetraloops are replaced specifically rather than randomly by other tetraloops. For example, 93% of the hairpin loops at position 83 in 16S rRNA have the sequence CUUG, UUCG, or GCAA (9). NMR studies indicate both UUCG and GCAA loops have defined structures (8, 16). Thus, an alternative explanation for the context dependence of hydrogen bonding is that the  $\Delta G^\circ$  contribution may help to hold the loop in a single conformation.

#### REFERENCES AND NOTES

1. D. M. Crothers, P. E. Cole, C. W. Hilbers, R. G. Shulman, *J. Mol. Biol.* **87**, 63 (1974).
2. D. Riesner and R. Römer, in *Physico-Chemical Properties of Nucleic Acids*, J. Duchesne, Ed. (Academic Press, London, 1973), vol. 2, pp. 237–318.
3. D. Moazed, S. Stern, H. F. Noller, *J. Mol. Biol.* **187**, 399 (1986).
4. K. Kruger *et al.*, *Cell* **31**, 147 (1982).
5. D. H. Turner, N. Sugimoto, S. M. Freier, *Annu. Rev. Biophys. Chem.* **17**, 167 (1988).
6. S. M. Freier *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9373 (1986).
7. J. A. Jaeger, D. H. Turner, M. Zuker, *ibid.* **86**, 7706 (1989).
8. H. A. Heus and A. Pardi, *Science* **253**, 191 (1991).
9. C. R. Woese, S. Winker, R. R. Gutell, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8467 (1990).
10. C. Tuerk *et al.*, *ibid.* **85**, 1364 (1988).
11. D. H. Turner, N. Sugimoto, R. Kierzek, S. D. Dreiker, *J. Am. Chem. Soc.* **109**, 3783 (1987).
12. J. SantaLucia, Jr., R. Kierzek, D. H. Turner, *ibid.* **113**, 4313 (1991).
13. A. R. Fersht, *Trends Biochem. Sci.* **12**, 301 (1987); T. Alber *et al.*, *Nature* **330**, 41 (1987); T. Alber *et al.*, *Science* **239**, 631 (1988); B. A. Shirley, P. Stanssens, U. Hahn, C. N. Pace, *Biochemistry* **31**, 725 (1992); D. R. Lesser, M. R. Kurpiewski, L. Jen-Jacobson, *Science* **250**, 776 (1990).
14. A. M. Pyle and T. Cech, *Nature* **350**, 628 (1991).
15. P. C. Bevilacqua and D. H. Turner, *Biochemistry* **30**, 10632 (1991).
16. G. Varani, C. Cheong, I. Tinoco, Jr., *ibid.*, p. 3280.
17. J. SantaLucia, Jr., R. Kierzek, D. H. Turner, *ibid.*, p. 8242.
18. Two-dimensional NMR experiments (19) for the duplex (rGGCGAGCC)<sub>2</sub> are consistent with the G to A hydrogen bonding shown in Fig. 3. The

- duplex (rGGCGAGCC)<sub>2</sub>, however, is more stable than (rGGCGAGCC)<sub>2</sub> (12), which is surprising for the hydrogen bonding in Fig. 3. NMR spectra of (GGCAGCC)<sub>2</sub>, however, show a resonance at 14.3 ppm for the imino proton of I (19), which indicates the hydrogen bonding in the IA mismatch involves the imino proton. Thus the increase in stability for (rGGCAGCC)<sub>2</sub> as compared to (GGCGAGCC)<sub>2</sub> is associated with a change in structure.
19. J. SantaLucia, Jr., and D. H. Turner, unpublished observations.
  20. D. R. Groebe and O. C. Uhlenbeck, *Nucleic Acids Res.* **16**, 11725 (1988).

21. V. P. Antao, S. Y. Lai, I. Tinoco, Jr., *ibid.* **19**, 5901 (1991).
22. F. Michel and E. Westhof, *J. Mol. Biol.* **216**, 585 (1990).
23. R. Kierzek *et al.*, *Biochemistry* **25**, 7840 (1986).
24. S.-H. Chou, P. Flynn, B. Reid, *ibid.* **28**, 2422 (1989).
25. We thank A. Pardi for stimulating discussions and for supplying the structure of pppGGGCGCAA-GCCUUAU before publication. We also thank E. Kool, T. R. Krugh, and L. A. Lindahl for comments on the manuscript. Supported by NIH grant GM22939.

2 December 1991; accepted 21 February 1992

## Centriole Duplication in Lysates of *Spisula solidissima* Oocytes

Robert E. Palazzo,\* Eugeni Vaisberg, Richard W. Cole, Conly L. Rieder

A cell-free system has been developed that executes centriole duplication. Surf clam (*Spisula solidissima*) oocytes, arrested at late prophase of meiosis I, do not contain centrioles, centrosomes, or asters. Serial section high-voltage electron microscopy (HVEM) of asters and spindles isolated from potassium chloride-activated oocytes indicates that within 4 minutes oocytes assemble a single centriole that is duplicated by 15 minutes when assembly of the first meiotic spindle is complete. A mixture of lysates from unactivated oocytes and potassium chloride-activated oocytes induces centriole formation and duplication. Astral microtubule content in these lysate mixtures increases with time.

Centrioles are structurally complex cylindrical organelles associated with centrosomes, asters, cilia, and flagella of animal cells (1, 2). Within the centrosome, centrioles are commonly found as a mother and daughter pair intimately associated with ancillary pericentriolar structures. During interphase these pericentriolar structures serve as nucleation sites for cytoplasmic microtubules (MTs), whereas during mitosis they generate radial astral arrays of MTs that form the spindle (3, 4). The initiation of centrosome replication, which is a preparatory step for mitosis, is signaled by centriole replication; in echinoderm embryos the ability of a centrosome to replicate depends on the presence of a preexisting centriole (5, 6). The controlled replication of centrioles in sea urchin (7, 8) and frog zygotes (9) and their subsequent separation to form functionally independent centrosomes occurs in the absence of both mRNA and protein synthesis. Thus, in these developing embryos new centrioles and functional centrosomes are assembled within the zygote from preexisting pools of

subunits. Unfortunately, centrioles and their associated structures have yet to be isolated in purified form, and their composition and replication are poorly understood (10–12). Here we describe a cell-free system based on *Spisula solidissima* oocytes that supports centriole assembly and duplication.

When *Spisula* oocytes are arrested at late prophase of meiosis I, centrosomes cannot be detected in living oocytes by high-resolution light microscopy (13, 14) or in fixed oocytes by electron microscopy (EM) (16, 17) or tubulin immunofluorescence (IMF) (15). Moreover, lysates from unactivated oocytes do not contain or form asters (18). When oocytes are parthenogenetically activated with KCl they complete meiosis I and II (13, 14), and within 2 to 5 min of activation two astral arrays of MTs can be detected in the cytoplasm near the germinal vesicle (15). These asters are also found in lysates prepared from eggs 4 min after activation (Figs. 1 and 2) and appear to contain a single centriole when examined in random thin sections (18, 19). Indeed, high-voltage EM (HVEM) analyses (20) of serial thick sections through ten of these asters confirmed that each contains only a single centriole (Fig. 2).

Approximately 15 min after activation oocytes contain a mature meiosis I metaphase spindle (13–17). All ten centrosomes examined by serial-section HVEM that

R. E. Palazzo, The Marine Biological Laboratory, Woods Hole, MA 02543.  
E. Vaisberg, Protein Research Institute, Academy of Sciences, Moscow, Russia.  
R. W. Cole and C. L. Rieder, Wadsworth Center for Labs and Research, Albany, NY 12201, and Department of Biomedical Sciences, State University of New York, Albany, NY 12222.

\*To whom correspondence should be sent.