# Stability of XGCGCp, GCGCYp, and XGCGCYp Helixes: An Empirical Estimate of the Energetics of Hydrogen Bonds in Nucleic Acids[†]

Susan M. Freier,[‡] Naoki Sugimoto,[‡] Alison Sinclair,[§] Dirk Alkema,[§] Thomas Neilson,[§] Ryszard Kierzek,[‖] Marvin H. Caruthers,[⊥] and Douglas H. Turner*[,‡]

*Department of Chemistry, University of Rochester, Rochester, New York 14627, Department of Biochemistry, McMaster University, Hamilton, Ontario L8N 3Z5, Canada, Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-704 Poznan, Noskowskiego 12/14, Poland, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309*

*Received October 14, 1985; Revised Manuscript Received February 3, 1986*

ABSTRACT: The stabilizing effects of dangling ends and terminal base pairs on the core helix GCGC are reported. Enthalpy and entropy changes of helix formation were measured spectrophotometrically for AGCGCU, UGCGCA, GGCGCCp, CGCGCGp, and the corresponding pentamers XGCGCp and GCGCYp containing the GCGC core plus a dangling end. Each 5′ dangling end increases helix stability at 37 °C roughly 0.2 kcal/mol and each 3′ end from 0.8 to 1.7 kcal/mol. The free energy increments for dangling ends on GCGC are similar to the corresponding increments reported for the GGCC core [Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., & Turner, D. H. (1985) *Biochemistry 24*, 4533–4539], indicating a nearest-neighbor model is adequate for prediction of stabilization due to dangling ends. Nearest-neighbor parameters for prediction of the free energy effects of adding dangling ends and terminal base pairs next to G·C pairs are presented. Comparison of these free energy changes is used to partition the free energy of base pair formation into contributions of "stacking" and "pairing". If pairing contributions are due to hydrogen bonding, the results suggest stacking and hydrogen bonding make roughly comparable favorable contributions to the stability of a terminal base pair. The free energy increment associated with forming a hydrogen bond is estimated to be −1 kcal/mol of hydrogen bond.

There is no general consensus on the relative importance of the various forces determining stabilities of double-helical nucleic acids. Theoretical calculations nearly 20 years ago predicted that stacking interactions are important (DeVoe & Tinoco, 1962; Pullman & Pullman, 1968, 1969). Studies of the effects of unpaired terminal nucleotides (dangling ends) on the stabilities of double helixes provide experimental evidence for stacking contributions (Martin et al., 1971; Romaniuk et al., 1978; Neilson et al., 1980; Alkema et al., 1981a,b; Petersheim & Turner, 1983; Freier et al., 1983, 1984, 1985a). The contributions of hydrogen bonding to stability are more controversial (Fersht et al., 1985). Studies on the association of small molecules by hydrogen bonding in solution (Schellman, 1955; Klotz & Franzen, 1962) suggest contributions of hydrogen bonds to stability are negligible because they only replace hydrogen bonds to water. On the basis of the difference in melting temperatures of polymers with three and two hydrogen bonds per base pair, Crothers and Zimm (1964) suggested hydrogen bonds contribute about −1 kcal/mol of hydrogen bond to helix stability. Additional experimental evidence for choosing between these possibilities is essentially nonexistent. There must also be an unfavorable configurational entropy that opposes helix formation (Longuet-Higgins & Zimm, 1960; DeVoe & Tinoco, 1962; Applequist & Damle, 1966). This arises from the limited number

of conformations available to the double helix relative to separated single strands.

In this paper, we report thermodynamic parameters of helix formation for the ribooligonucleotide GCGC, several pentanucleotides containing the GCGC core plus a dangling end, and four hexamers containing the GCGC core plus a terminal Watson–Crick pair. The results are combined with those from previous studies (Petersheim & Turner, 1983; Freier et al., 1983, 1985a) to suggest a partitioning of the free energy increment for a terminal base pair into contributions from stacking, hydrogen bonding, and configurational effects. In addition, the data provide tests of the nearest-neighbor model for helix stability (Gralla & Crothers, 1973; Borer et al., 1974) and thermodynamic parameters for prediction of RNA stability.

## MATERIALS AND METHODS

*Oligonucleotide Synthesis.* GGC, GCGC, and CGCG were synthesized chemically by solution-phase phosphotriester procedures and were characterized by ^{1}H NMR (Alkema et al., 1981a,b; Sinclair et al., 1984; Freier et al., 1983). AGCGCU, UGCGCA, and UGCGC were synthesized chemically on solid support by phosphoramidite procedures.

Pentanucleotides and hexanucleotides were synthesized stepwise from GGC, GCGC, and CGCG. The following procedure was followed: (1) A 5′–3′ nucleoside bisphosphate (pNp) was added by T4 RNA ligase; (2) the product was purified on a C-18 Sep-Pak cartridge (Waters); (3) the 3′-terminal phosphate was removed by calf alkaline phosphatase; (4) the phosphatase was deactivated by boiling, and the product was again purified on a C-18 Sep-Pak. Additional nucleotides were added by repetition of this four-step cycle. After the final addition, the procedure was stopped after step

[‡] University of Rochester.
[§] McMaster University.
[‖] Polish Academy of Sciences.
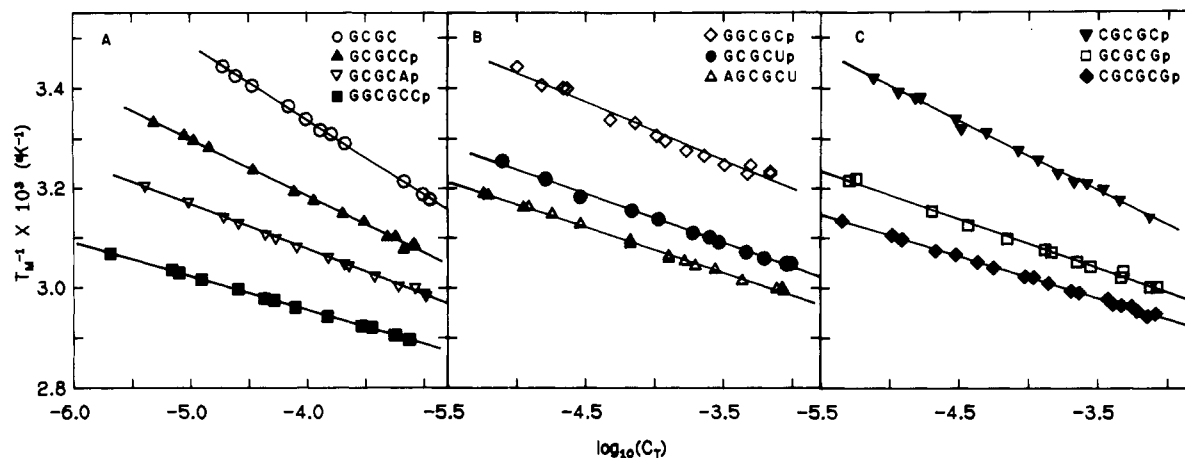[⊥] University of Colorado.

FIGURE 1: Reciprocal melting temperature vs. log concentration for (A) GCGC (O), GCGCCp (▲), GCGCAp (▽), and GGCGCCp (■), for (B) GGCGCp (◇), GCGCUp (●), and AGCGCU (△), and for (C) CGCGCp (▼), GCGCGp (□), and CGCGCGp (◆) in 1 M NaCl, 0.005 M Na₂HPO₄, and 0.5 mM Na₂EDTA, pH 7.

Table I: Thermodynamic Parameters of Helix Formation[a]

| oligomer | log $C_t$ parameters | | | "temperature-independent" parameters | | |
|---|---|---|---|---|---|---|
| | $-\Delta H°$ (kcal/mol)[b,c] | $-\Delta S°$ (eu)[b,c] | $T_M$ (°C)[c,d] | $-\Delta H°$ (kcal/mol)[b,e] | $-\Delta S°$ (eu)[b,e] | $T_M$ (°C)[d,e] |
| CGCG[f] | 31.8 | 91.6 | 16.6 | 28.3 | 79.6 | 16.0 |
| GCGC | 30.5 | 83.4 | 26.5 | 32.3 | 89.0 | 27.4 |
| CGCGCp | 32.9 | 89.1 | 33.1 | 33.3 | 90.4 | 33.4 |
| GGCGCp | 39.6 | 113.0 | 28.6 | 35.3 | 98.7 | 28.8 |
| UGCGC | 29.8 | 79.7 | 31.1 | 29.2 | 77.6 | 31.5 |
| GCGCAp | 50.6 | 137.6 | 51.8 | 50.0 | 135.7 | 51.9 |
| GCGCCp | 39.9 | 108.8 | 40.9 | 39.2 | 106.2 | 41.3 |
| GCGCGp | 46.9 | 126.4 | 50.7 | 47.9 | 129.5 | 50.7 |
| GCGCUp | 46.6 | 127.9 | 45.4 | 45.5 | 124.6 | 45.5 |
| UGCGCA | 51.5 | 139.7 | 53.1 | 51.2 | 138.5 | 53.2 |
| AGCGCU | 50.1 | 135.7 | 52.1 | 50.5 | 136.8 | 52.2 |
| CGCGCGp | 54.5 | 146.4 | 57.9 | 53.8 | 144.2 | 58.0 |
| GGCGCCp | 67.8 | 182.0 | 65.2 | 65.6 | 175.5 | 65.3 |

[a] Measurements were in 1 M NaCl, 0.005 M Na₂HPO₄, and 0.5 mM Na₂EDTA, pH 7. [b] Although estimated errors in $\Delta H°$ and $\Delta S°$ are ±5%, additional significant figures are given to allow accurate calculation of $T_M$. [c] From plots of reciprocal melting temperature vs. log $C_t$. [d] Calculated for $10^{-4}$ M oligomer concentration. [e] Temperature-independent thermodynamic paraters. These are the average of those from plots of $T_M^{-1}$ vs. log $C_t$ and those from fits of individual melting curves to a two-state model with sloping base lines. [f] Because the $T_M$ is low, error limits for CGCG are ±10% for $\Delta H°$ and $\Delta S°$ and ±4% for $\Delta G°$ at the $T_M$.

2. If necessary, the final product was purified by ion exchange chromatography on DEAE-Sephadex. Details of the enzymatic synthesis and purification procedures are given elsewhere (Freier et al., 1983, 1985a; Beckett & Uhlenbeck, 1984). Purity of all oligomers was confirmed by high-performance liquid chromatography.

*Oligonucleotide Solutions.* Oligonucleotide concentrations ($C_t$) are strand concentrations and were calculated from the high-temperature absorbance at 280 nm. Single-strand extinction coefficients were calculated from extinction coefficients for dinucleoside monophosphates and nucleotides (Freier et al., 1983). In units of $10^4$ M⁻¹ cm⁻¹, the calculated extinction coefficients are as follows: GCGC, 2.01; GGCGCp, 2.74; CGCGCp, 2.44; UGCGC, 2.20; GCGCAp, 2.17; GCGCCp, 2.52; GCGCGp, 2.57; GCGCUp, 2.30; AGCGCU, 2.56; GGCGCCp, 3.25; CGCGCGp, 2.99; UGCGCA, 2.36.

*Melting Curves.* All melting curves were measured at 280 nm in 1 M NaCl, 0.005 M Na₂HPO₄, and 0.5 mM disodium ethylenediaminetetraacetic acid (Na₂EDTA), pH 7, as described previously (Freier et al., 1983). For each oligomer, at least 16 absorbance vs. temperature profiles (melting curves) were measured over a 100-foid range in strand concentration.

*Thermodynamic Parameters.* Thermodynamic parameters of helix formation were obtained from absorbance vs. temperature profiles by two methods: (1) enthalpies and entropies

from fits of individual melting curves to a two-state model with sloping base lines were averaged, and (2) linear plots of reciprocal melting temperature ($T_M^{-1}$) vs. log $C_t$ yielded enthalpies and entropies (Borer et al., 1974):

$$T_M^{-1} = (2.3R/\Delta H°) \log C_t + \Delta S°/\Delta H° \qquad (1)$$

We report thermodynamic parameters from plots of $T_M^{-1}$ vs. log $C_t$ and temperature-independent thermodynamic parameters, which are the average of methods 1 and 2. [See Petersheim and Turner (1983) and Freier et al. (1985a) for details.] Estimated errors are ±5% in the temperature-independent entropies and enthalpies and ±2% in $\Delta G°$ at the $T_M$. Errors in $\Delta G°$ at other temperatures are propagated from the errors in $\Delta H°$ and $\Delta G°_{T_M}$, assuming no correlation in $\Delta H°$ and $\Delta G°_{T_M}$.

Temperature-dependent thermodynamic parameters are also reported. They were obtained from plots of $\Delta H°$ vs. $T_M$ and $\Delta S°$ vs. ln $T_M$, where $\Delta H°$ and $\Delta S°$ are the parameters obtained from the fit of each curve to a two-state model with linear sloping base lines. Heat capacity changes are obtained from the slopes of these plots (Petersheim & Turner, 1983; Freier et al., 1983). Estimated errors in the temperature-dependent free energy, enthalpy, and entropy changes are ±5, ±10, and ±10%, respectively, at the $T_M$; errors in $\Delta C°_P$ are ±50%. Errors in $\Delta H°$ and $\Delta S°$ at other temperatures are

Table II: Temperature-Dependent Thermodynamic Parameters of Helix Formation[a]

| oligomer | $-\Delta H^{\circ}_{37}$ (kcal/mol)[b,d] | $-\Delta S^{\circ}_{37}$ (eu)[c,d] | $-\Delta C^{\circ}_P$ (cal mol$^{-}$/$^1$ K$^{-1}$) |
|---|---|---|---|
| CGCG | 26.9 | 76 | 160 |
| GCGC | 38.7 | 110 | 640 |
| CGCGCp | 35.0 | 95 | 260 |
| GGCGCp | 34.8 | 97 | 494 |
| UGCGC | 32.1 | 86 | 343 |
| GCGCAp | 47.6 | 127 | 120 |
| GCGCCp | 36.8 | 98 | 373 |
| GCGCGp | 47.4 | 129 | 94 |
| GCGCUp | 42.3 | 115 | 217 |
| AGCGCU | 44.9 | 119 | 406 |
| CGCGCGp | 49.0 | 129 | 206 |
| GGCGCCp | 58.4 | 154 | 185 |
| UGCGCA | 47.2 | 127 | 235 |

[a] Measurements were in 1 M NaCl, 0.005 M Na$_2$HPO$_4$, and 0.5 mM Na$_2$EDTA, pH 7. [b] Obtained from plots of $\Delta H^{\circ}$ vs. $T_M$ as described under Materials and Methods. [c] Obtained from plots of $\Delta S^{\circ}$ vs. ln $T_M$ as described under Materials and Methods. [d] Although estimated errors in $\Delta H^{\circ}$ and $\Delta S^{\circ}$ at the $T_M$ are ±10%, additional significant figures are given to allow accurate calculation of $T_M$.

propagated from the error in $\Delta C^{\circ}_P$ and the errors at the $T_M$.

## RESULTS

*Temperature-Independent Thermodynamic Parameters.* Plots of reciprocal melting temperature vs. log $C_t$ are shown in Figure 1. Enthalpy and entropy changes obtained from these plots are in Table I. These parameters were averaged with those derived from fits of individual melting curves to obtain the temperature-independent parameters, which are also listed in Table I. As shown in Table I, addition of two 5' ends to GCGC increases the $T_M$ only 1–6 °C. In contrast, addition of two nucleotides to the 3' end of GCGC increases the $T_M$ 14–24 °C, and two terminal pairs increase the stability even more, 25–38 °C.

*Temperature-Dependent Thermodynamic Parameters.* For most oligomers, enthalpy changes from plots of $T_M^{-1}$ vs. log $C_t$ agree within 10% with those from fits indicating the tran-

sition is two-state. The exception is GGCGCp, where enthalpies derived with the two analytical methods differ by 20%. Although the transitions appear to be two-state, when fitted enthalpy and entropy changes are plotted vs. $T_M$ and ln $T_M$, respectively, a small temperature dependence is observed. Slopes of these plots provide heat capacity changes, $\Delta C^{\circ}_P$, between helix and coil. These are listed in Table II, along with enthalpy and entropy changes for helix formation at 37 °C. The $\Delta C^{\circ}_P$ values in Table II are similar to those observed previously (Petersheim & Turner, 1983; Freier et al., 1983, 1985a; Hickey & Turner, 1985a,b) and probably are due to unstacking of the single-strand state.

*Helixes with Two Possible Configurations.* The sequence CGCGCp can form two duplexes, a GCGC core helix with two 5' dangling cytidines or a CGCG core helix with two 3' dangling cytidines. Similarly, GCGCGp can have either two 5' dangling guanosines on CGCG or two 3' dangling guanosines on GCGC. Both oligomers, however, melt in a two-state manner, suggesting, for each, only one duplex species predominates. We assume both helixes form a GCGC core duplex. The basis of this assumption is comparison of the observed free energies of helix formation for CGCGCp and GCGCGp with that predicted for each possible structure. Previously, we reported free energy increments for dangling ends on GGCC and CCGG cores (Freier et al., 1985a). Comparison with results for UGCGC, GGCGCp, GCGCCp, GCGCAp, and GCGCUp demonstrate these increments depend only on the nature of the dangling base and the terminal pair. Thus, they can be used to predict free energies of helix formation for the four possible structures. Using free energy increments measured on CCGG or GGCC [see Freier et al. (1985a) or Table III] and the observed free energies of helix formation for CGCG and GCGC, we predict the stability at 37 °C of GCGCGp to be −8.1 kcal/mol for a GCGC core duplex with two 3' dangling guanosines. If the duplex contains a CGCG core with two 5' dangling guanosines, the predicted $\Delta G^{\circ}_{37}$ is −4.0 kcal/mol. The observed value is −7.7 kcal/mol, indicating the structure with 3' dangling ends is preferred.

Table III: Excess Stabilization ($-\Delta\Delta G^{\circ}_{37}$ in kcal/mol) by Dangling Ends and Terminal Base Pairs in 1 M NaCl[a]

| oligomer | pred terminal base pair[d] | meas terminal base pair | 3' dangling end | 5' dangling end | "empirical pairing"[f] | stabilization per H bond[g] |
|---|---|---|---|---|---|---|
| GCGC core | | | | | | |
| AGCGCU | 1.6 | 1.7 | 1.1 | (0.2)[e] | 0.4 | 1.2 |
| CGCGCGp | 2.0 | 2.2 | 1.5 | 0.3 | 0.4 | 0.8 [0.4][h] |
| GGCGCCp | 2.9 | 3.3 | 0.8 | 0.0 | 2.5 | 1.5 [1.6][h] |
| UGCGCA | 1.8 | 1.8 | 1.7 | 0.2 | −0.1 | 0.9 |
| GGCC core[b] | | | | | | |
| AGGCCUp | 1.6 | 1.6 | 1.2 | 0.2 | 0.2 | 1.1 |
| CGGCCCGp | 2.0 | 2.3 | 1.7 | 0.2 | 0.4 | 0.8 [0.7][h] |
| UGGCCAp | 1.8 | 1.6 | 1.8 | 0.0 | −0.2 | 0.9 |
| CCGG core[b] | | | | | | |
| ACCGGUp | 2.1 | 1.9 | 0.6 | 0.5 | 0.8 | 1.4 |
| GCCGGCp | 3.3 | 3.4 | 0.4 | 0.2 | 2.8 | 1.6 [1.5][h] |
| UCCGGAp | 2.2 | 1.6 | 1.1 | 0.1 | 0.4 | 1.2 |
| CGCG core[c] | | | | | | |
| GCGCGCp | 3.3 | 3.4 | (0.4)[e] | (0.2)[e] | 2.8 | 1.6 |

[a] $\Delta\Delta G^{\circ}_{37}$ is half the difference between the free energy of helix formation for the molecule containing the core helix plus the added termini and the free energy of helix formation for the tetramer core. No correction has been applied to oligomers without 3'-terminal phosphates since this correction is small (Freier et al., 1985a). Temperature-independent thermodynamic parameters were used to calculate $\Delta\Delta G^{\circ}_{37}$. Propagation of the errors in $\Delta H^{\circ}$ and $\Delta G^{\circ}_{T_M}$ lead to errors in $\Delta\Delta G^{\circ}_{37}$ of about 0.1 kcal/mol. Note that $-\Delta\Delta G^{\circ}_{37}$ is reported. [b] From Freier et al. (1983, 1985a) and Petersheim and Turner (1983). [c] From Freier et al. (1985b). [d] Helix propagation free energies from a linear least-squares fit to the nearest-neighbor model for thermodynamic data determined from log $C_T$ plots for 41 oligoribonucleotides (Freier et al., 1986a). [e] These values were measured on another core and are predicted to be valid from the nearest-neighbor model. For example, 0.4 kcal/mol for a 3' dangling C on CGCG is predicted from the measured $\Delta\Delta G^{\circ}_{37}$ for a 3' dangling C on CCGG. [f] "Empirical pairing" is defined as $\Delta\Delta G^{\circ}_{emp}$(pairing) = $\Delta\Delta G^{\circ}_{emp}$(bp) − $\Delta\Delta G^{\circ}_{emp}$(3' dangle) − $\Delta\Delta G^{\circ}_{emp}$(5' dangle). [g] The stabilization per H bond was calculated with the assumption that the configurational $\Delta\Delta G^{\circ}_{37}$ for fixing the 5'-nucleotide in a base pair is 1.9 kcal/mol. For a base pair with $n$ H bonds, the stabilization per H bond is defined as $\Delta\Delta G^{\circ}$(H bond) = $(1/n)[\Delta\Delta G^{\circ}_{emp}$(pairing) − 1.9]. [h] These values are calculated by subtracting the free energy increment for a terminal AU pair from the free energy increment for the corresponding terminal GC pair.

Table IV: Temperature-Dependent Enthalpy and Entropy Increments for 3' Dangling Ends and Terminal Pairs[a]

| core sequence | 3' dangling end | | | A·U terminal pair | | | 3' dangling end | | | G·C terminal pair | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | end | $-\Delta\Delta H^{\circ}_{37}$ (kcal/mol) | $-\Delta\Delta S^{\circ}_{37}$ (eu) | pair | $-\Delta\Delta H^{\circ}_{37}$ (kcal/mol) | $-\Delta\Delta S^{\circ}_{37}$ (eu) | end | $-\Delta\Delta H^{\circ}_{37}$ (kcal/mol) | $-\Delta\Delta S^{\circ}_{37}$ (eu) | pair | $-\Delta\Delta H^{\circ}_{37}$ (kcal/mol) | $-\Delta\Delta S^{\circ}_{37}$ (eu) |
| GCGC | 3'pUp | 1.8 | 2.7 | 5'Ap + 3'pU | 3.1 | 4.4 | 3'pGp | 4.3 | 9.5 | 5'Cp + 3'pGp | 5.1 | 9.3 |
| GCGC | 3'pAp | 4.4 | 8.5 | 5'Up + 3'pA | 4.3 | 8.5 | 3'pCp | -1.0 | -5.9 | 5'Gp + 3'pCp | 9.9 | 22.1 |
| GGCC | 3'pUp | 2.5 | 5.0 | 5'Ap + 3'pUp | 3.5 | 6.5 | 3'pGp | 2.5 | 3.0 | 5'Cp + 3'pGp | 7.6 | 18.0 |
| GGCC | 3'pAp | 3.5 | 5.6 | 5'Up + 3'pAp | 3.1 | 4.0 | 3'pCp | -0.7 | -4.5 | | | |
| CCGG | 3'pUp | 2.4 | 5.5 | 5'Ap + 3'pUp | 4.4 | 8.8 | 3'pGp | 4.6 | 10.9 | | | |
| CCGG | 3'pAp | 3.5 | 7.8 | 5'Up + 3'pAp | 3.2 | 5.1 | 3'pCp | 1.6 | 4.0 | 5'Gp + 3'pCp | 11.6 | 26.6 |
| CGCG | | | | | | | 3'pCp[b] | 1.6 | 4.0 | 5'Gp + 3'pCp | 15.6 | 37.3 |

[a] Data in Table II, Table I of Petersheim and Turner (1983), Table IV of Freier et al. (1983), and Table III of Freier et al. (1985a) were used to calculate $\Delta\Delta H^{\circ}$ and $\Delta\Delta S^{\circ}$. For example, the enthalpy increment for adding a 3' dangling pUp to GCGC is $\Delta\Delta H^{\circ}(3'U) = 0.5[\Delta H^{\circ}(GCGCUp) - \Delta H^{\circ}(GCGC)]$. [b] Increments for a 3' dangling C were measured on CCGGCp.

Similarly, for CGCGCp, the predictions are −5.1 or −4.4 kcal/mol for duplexes with 5' or 3' dangling cytidines, respectively. The observed value is −5.3 kcal/mol, again indicating the duplex with a GCGC core is formed.

## DISCUSSION

*Nearest-Neighbor Model and Free Energy Parameters for Adding Dangling Ends and Terminal Base Pairs to RNA Helixes.* The results in Table I can be used to derive free energy increments for adding a dangling end or terminal pair to a GCGC core. For example, the free energy increment for adding a 3' dangling pUp is $\Delta\Delta G^{\circ}(3'\text{-}U) = 0.5[\Delta G^{\circ}(GCGCUp) - \Delta G^{\circ}(GCGC)]$. These increments are listed in Table III and can be compared with those previously determined for GGCC and CCGG cores (Petersheim & Turner, 1983; Freier et al., 1983, 1985a). The nearest-neighbor model (Gray & Tinoco, 1970; Borer et al., 1974) predicts the free energy increment for adding an end to a core helix depends only on the end and the adjacent base pair. Thus, a 3' dangling U on GCGC is predicted to have the same $\Delta\Delta G^{\circ}$ as a 3' dangling U on GGCC. In all cases, free energy increments on the GCGC and GGCC cores agree within 0.2 kcal/mol, supporting the nearest-neighbor model. Except for UCCGGAp and GGCGCCp, the free energy increments for terminal base pairs listed in Table III are within 0.3 kcal/mol of the increments derived from fitting free energy changes of duplex formation for 41 oligoribonucleotides to the nearest-neighbor model (Freier et al., unpublished results). This agreement indicates the model is reasonable in most, but perhaps not all, cases.

Papanicolaou et al. (1984) find that making terminal AU pairs adjacent to GC pairs 0.7 kcal/mol less favorable than the corresponding internal pair improves predictions of tRNA and 5S RNA secondary structure. In contrast, the results in Table III show this is not valid for ribooligonucleotides.

*Stacking of Dangling Ends.* A dangling end will seek a conformation with the most favorable free energy. This conformation may be different in the single- and double-strand states. Thus, the free energy increments for dangling ends in Table III represent differences in free energies for the terminal nucleotides between single and double strands. This analysis assumes the $\Delta G^{\circ}$ associated with duplex formation of the core tetramer is independent of the nature of the ends (Freier et al., 1985a).

The roughly −0.2 kcal/mol increments for most 5' dangling ends are similar to those for 5'-phosphates (Freier et al., 1983) and suggest little or no additional stacking in the duplex. In contrast, the −0.4 to −1.8 kcal/mol increments for 3' dangling ends suggest considerably enhanced stacking due to interstrand interactions. Pure interstrand interactions may actually be more favorable than the free energy increments because some free energy may be required to disrupt favorable intrastrand interactions in order to reach the most favorable duplex conformation.

Solvent perturbation studies indicate bulk solvent effects like solvophobic or classical hydrophobic bonding are not major contributors to nucleic acid stability (Turner et al., 1985; Hickey & Turner, 1985a; Dewey & Turner, 1980; Freier et al., 1981). Thus, the free energy increments for 3' dangling ends presumably contain primarily favorable enthalpic contributions from vertical electronic (van der Waals) interactions and unfavorable entropic contributions from configurational considerations. Except for GCGCCp, which involves terminal repeating C's [e.g., see Freier et al. (1983)], the thermodynamic data in Table II are consistent with this interpretation. The favorable $\Delta\Delta G^{\circ}$ of 3' dangling end stacking is due to a favorable enthalpy change that overcomes an unfavorable entropy change.

Although the enthalpy and entropy changes are difficult to measure precisely for each 3' dangling end, the averages for all that have been measured provide an estimate of these changes. The average $\Delta\Delta H^{\circ}$ and $\Delta\Delta S^{\circ}$ at 37 °C for 3' dangling ends without repeating C's are −3 kcal/mol and −6 eu, respectively (see Table IV). These values are probably relevant to the transition from a partially unstacked single strand to a fully stacked double helix (Brahms et al., 1967; Dewey & Turner, 1979; Romaniuk et al., 1979; Stone et al., 1981; Shum & Crothers, 1983; Doornbos et al., 1983). The measured entropy can be compared with that expected from restriction of backbone bonds in a single stacked configuration. An approximate counting of the number of conformational states available to the random coil (Applequist & Damle, 1966; Devoe & Tinoco, 1962; Sundaralingum, 1973, 1975; Dewey & Turner, 1979) gives for a helix propagation step $\Delta S \approx -R$ ln $(3 \times 3 \times 2 \times 6) = -9$ eu. Comparison of the theoretical and experimental values suggests configurational entropy is a major source of the entropy change associated with 3' dangling end stacking.

For 5' dangling ends, the roughly −0.2 kcal/mol free energy increment is associated with a small favorable entropy that overcomes an unfavorable enthalpy. This may be due to reduced flexibility in the coil state for the sugar of the 5'-terminal residue of the core (Sundaralingam, 1973, 1975) and would be expected to contribute to duplex stability whenever a phosphate is at the 5' terminus of the core sequence. The 5'-A in ACCGGp is an exception (Petersheim & Turner, 1983; Freier et al., 1985a). The $\Delta\Delta G^{\circ}$ of −0.5 kcal/mol is due to a favorable enthalpy and an unfavorable entropy, suggesting the 5'-adenosine may stack weakly in this duplex.

*Comparison of Free Energy Increments for Dangling Ends and Terminal Base Pairs.* Figure 2 presents free energy increments associated with adding dangling ends and terminal base pairs to GCGC helixes. The $\Delta\Delta G^{\circ}$'s for 3' dangling ends
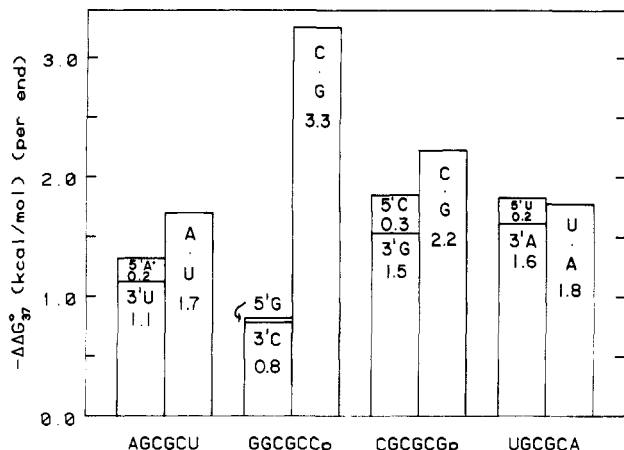
FIGURE 2: Free energy increments at 37 °C for adding a terminal base pair or dangling end to a GCGC tetramer core. The left-hand column represents free energy increments for 5' and 3' dangling ends; the right-hand column is the free energy of base pair formation. The free energy increments are from Table III. (*) This free energy increment for a 5' dangling A was measured on AGGCCp.

provide an empirical measure of the free energy associated with stacking. In three of four cases, free energy increments for 3' dangling ends are at least 65% of those for base pairs, indicating stacking is an important determinant of nucleic acid stability. It is conceivable, however, that the most favorable conformation for a base pair differs from that for a dangling end. Thus, the dangling end $\Delta\Delta G°$'s in Figure 2 provide the maximum favorable contribution of stacking to base pair stability.

Another contributor to base pair stability is the -0.2 kcal/mol favorable free energy associated with the phosphate bound to the 5' residue of the core. The free energy increments for 5' dangling ends shown in Figure 2 are largely attributable to this effect, which accounts for about 10% of the free energy increment for a terminal base pair.

*Role of Hydrogen Bonding in Duplex Stability.* When the free energy increments for 3' and 5' dangling ends are subtracted from the free energy increment for a terminal base pair, the remaining free energy can be assigned to "empirical pairing": $\Delta\Delta G°_{emp}(pairing) = \Delta\Delta G°_{emp}(bp) - \Delta\Delta G°_{emp}(3' dangle) - \Delta\Delta G°_{emp}(5' dangle)$. For example, for GGCGCCp, the empirical pairing contribution is $(-3.3) - [-0.8 + (-0.0)] = -2.5$ kcal/mol (see Table III). Calculations by Pullman and Pullman (1968, 1969) indicate hydrogen bonding is the dominant interaction in pairing. It is possible that specific solvation effects are also important. There is not enough information presently available, however, to consider them explicitly. For seven sequences (see Table III and Figure 2), this pairing free energy is small, between 0.2 and -0.4 kcal/mol. For GGCGCCp and GCCGGCp, it is large, -2.5 and -2.8 kcal/mol, respectively. Results for GCGCGCp coupled with the predicted $\Delta\Delta G°$ for a 3' dangling C and a 5' dangling G give a similar pairing free energy of -2.8 kcal/mol. This suggests hydrogen bonding can provide about -0.9 kcal/mol to $\Delta G°_{37}$ for each hydrogen bond in a G·C pair.

The discussion above is not meant to suggest that pairing makes no contribution to stability in cases where the empirical pairing free energy is small. This empirical definition could underestimate the attractive forces due to pairing because the empirical pairing free energy contains an unfavorable configurational free energy. We consider this likely because base pair formation and 3' dangling end stacking require fixing the positions of two and one nucleotide(s), respectively, but a 5' dangling end probably remains flexible (Petersheim & Turner,

1983a,b; Freier et al., 1985a). Thus, the empirical free energy change for pairing includes the unfavorable configurational free energy, $\Delta\Delta G°_{config}(5'-nt)$, associated with fixing the 5'-terminal nucleotide in the base pair. Consequently, the free energy increment from the additional bonding associated with pairing, $\Delta\Delta G°_b(bp)$, can be estimated as $\Delta\Delta G°_b(bp) = [\Delta\Delta G°_{emp}(pairing) - \Delta\Delta G°_{config}(5'-nt)]$. It was suggested above that the average $\Delta\Delta S°$ (37 °C) of -6 eu measured for 3' dangling end stacking is largely a configurational entropy (see Table IV). This provides an approximation for $\Delta\Delta G°_{config}(5'-nt)$ of $-(-6)(310) = 1.9$ kcal/mol. If the bonding associated with pairing is due to hydrogen bonding, then dividing $\Delta\Delta G°_b(bp)$ by the number of hydrogen bonds in the terminal pair gives an estimate of the free energy increment for a hydrogen bond. These values are listed in Table III.

The free energy increments per hydrogen bond range from -0.8 to -1.6 kcal/mol and average -1.2 kcal/mol. These increments become more favorable as free energy increments for stacking of 3' dangling ends become less favorable. This suggests strong interstrand stacking may favor conformations that are not optimal for hydrogen bonding. This type of competition has been suggested by Levitt (1978), on the basis of energy minimization studies of DNA.

Another empirical estimate of $\Delta\Delta G°$ for a hydrogen bond can be derived from the change in $\Delta G°$ of helix formation when a terminal GC pair is replaced with an AU pair. This assumes stacking interactions for GC and AU pairs are equal, so the extra stability of a terminal GC pair is due to the additional hydrogen bond. Inspection of free energy increments for dangling ends in Table III suggests this assumption is reasonable, though not exact. This method for estimating the free energy change for a hydrogen bond gives values ranging from -0.4 to -1.6 kcal/mol, consistent with those derived above (see Table III). Similar values are also derived by comparing terminal GC pairs with GU mismatches (Freier et al., 1986).

It is somewhat surprising that hydrogen bonding appears to be important for base pair stability. In aqueous solutions, the free energy of association of model compounds via hydrogen bonding is about equal to the free energy of hydrogen bonding between the model compound and water (Schellman, 1955; Klotz & Franzen, 1962). If these results are appropriate for hydrogen bonding in nucleic acids, one would predict a very small net $\Delta G°$ for the contribution of hydrogen bonds to the free energy of base pair formation. Values similar to those reported above have been derived, however, from studies of protein–substrate interactions (Fersht et al., 1985; Rupley et al., 1966). For nucleic acids, a $\Delta\Delta G°$ (30 °C) of -1.4 kcal/mol of hydrogen bond has been deduced for guanosine binding to the self splicing intervening sequence from *Tetrahymena thermophila* (Bass & Cech, 1984), although the nature of the binding site is not known. A value of -1 kcal/mol of hydrogen bond was estimated by Crothers and Zimm (1964) on the basis of the difference in melting temperatures of polynucleotides with three and two hydrogen bonds with the assumption that the enthalpy changes for melting are equal. Furthermore, it has been suggested that either the low dielectric environment of the helix or cooperative effects could lead to a favorable $\Delta G°$ and $\Delta H°$ for hydrogen bond formation (Nemethy et al., 1963; Jencks, 1969; Page & Jencks, 1971; Creighton, 1983). The dielectric constant inside a helix has been estimated to be 2 (Cech et al., 1976). The results in Table IV indicate the favorable $\Delta G°$ from pairing in G·C base pairs is associated with a favorable $\Delta H°$ and unfavorable $\Delta S°$, suggesting the low dielectric effect rather than a chelation

effect may account for the stability.

The net free energy change of base pair formation includes favorable bonding and unfavorable configurational terms, typically summing to −1 to −3 kcal/mol of base pair. That stacking contributes to this free energy change is indicated by the −0.4 to −1.8 kcal/mol stability increments measured for 3′ dangling ends. The pure bonding contribution from stacking is even larger because the measured free energy increments for 3′ dangling ends also include an unfavorable configurational free energy. Hydrogen bonding appears to contribute about −1 kcal/mol of hydrogen bonds. Thus, stacking and hydrogen bonding can make roughly equal contributions to the stability of a terminal base pair.

**Registry No.** CGCG, 89435-89-2; GCGC, 73942-16-2; CGCGCp, 101696-87-1; GGCGCp, 101696-88-2; UGCGC, 101696-89-3; GCGCAp, 101696-90-6; GCGCCp, 101696-91-7; GCGCGp, 101696-92-8; GCGCUp, 101696-93-9; UGCGCA, 101696-94-0; AGCGCU, 101696-95-1; CGCGCGp, 99508-76-6; GGCGCCp, 101696-96-2.

## REFERENCES

Alkema, D., Bell, R. A., Hader, P. A., & Neilson, T. (1981a) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) Vol. I, pp 417–428, Adenine, Guilderland, NY.

Alkema, D., Bell, R. A., Hader, P. A., & Neilson, T. (1981b) *J. Am. Chem. Soc. 103*, 2866–2868.

Applequist, J., & Damle, V. (1966) *J. Am. Chem. Soc. 88*, 3895–3900.

Bass, B. L., & Cech, T. R. (1984) *Nature (London) 308*, 820–826.

Beckett, D., & Uhlenbeck, O. C. (1984) in *Oligonucleotide Synthesis: A Practical Approach* (Gait, M. J., Ed.) pp 185–197, IRL Press, Oxford.

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843–853.

Brahms, J., Maurizot, J. C., & Michelson, A. M. (1967) *J. Mol. Biol. 25*, 465–480.

Cech, C. L., Hug, W., & Tinoco, I., Jr. (1976) *Biopolymers 15*, 131–152.

Creighton, T. E. (1983) *Biopolymers 22*, 49–58.

Crothers, D. M., & Zimm, B. H. (1964) *J. Mol. Biol. 9*, 1–9.

DeVoe, H., & Tinoco, I., Jr. (1962) *J. Mol. Biol. 4*, 500–517.

Dewey, T. G., & Turner, D. H. (1979) *Biochemistry 18*, 5757–5762.

Dewey, T. G., & Turner, D. H. (1980) *Biochemistry 19*, 1681–1685.

Doornbos, J., Wreesmann, C. T. J., van Boom, J. H., & Altona, C. (1983) *Eur. J. Biochem. 131*, 571–579.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature (London) 314*, 235–238.

Freier, S. M., Hill, K. O., Dewey, T. G., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1419–1426.

Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983) *Biochemistry 22*, 6198–6206.

Freier, S. M., Petersheim, M., Hickey, D. R., & Turner, D. H. (1984) *J. Biomol. Struct. Dyn. 1*, 1229–1242.

Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., & Turner, D. H. (1985a) *Biochemistry 24*, 4533–4539.

Freier, S. M., Sinclair, A., Neilson, T., & Turner, D. H. (1985b) *J. Mol. Biol. 185*, 645–647.

Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Biochemistry* (preceding paper in this issue).

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 73*, 497–511.

Gray, D. M., & Tinoco, I., Jr. (1970) *Biopolymers 9*, 223–244.

Hickey, D. R., & Turner, D. H. (1985a) *Biochemistry 24*, 2086–2094.

Hickey, D. R., & Turner, D. H. (1985b) *Biochemistry 24*, 3987–3991.

Jencks, W. P. (1969) *Catalysis in Chemistry and Enzymology*, Chapter 6, McGraw-Hill, New York.

Klotz, I. M., & Franzen, J. S. (1962) *J. Am. Chem. Soc. 84*, 3461–3466.

Levitt, M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 640–644.

Longuet-Higgins, H. C., & Zimm, B. H. (1960) *J. Mol. Biol. 2*, 1–4.

Martin, F. H., Uhlenbeck, O. C., & Doty, P. (1971) *J. Mol. Biol. 57*, 201–215.

Neilson, T., Romaniuk, P. J., Alkema, D., Hughes, D. W., Everett, J. R., & Bell, R. A (1980) *Nucleic Acids Symp. Ser. No. 7*, 293–311.

Nemethy, G., Steinberg, I. Z., & Scheraga, H. A. (1963) *Biopolymers 1*, 43–69.

Page, M. I., & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1678–1683.

Papanicolaou, C., Gouy, M., & Ninio, J. (1984) *Nucleic Acids Res. 12*, 31–44.

Petersheim, M., & Turner, D. H. (1983a) *Biochemistry 22*, 256–263.

Petersheim, M., & Turner, D. H. (1983b) *Biochemistry 22*, 269–277.

Pullman, A., & Pullman, B. (1968) *Adv. Quantum Chem. 4*, 267–325.

Pullman, B., & Pullman, A. (1969) *Prog. Nucleic Acid Res. Mol. Biol. 9*, 327–402.

Romaniuk, P. J., Hughes, D. W., Gregoire, R. J., Neilson, T., & Bell, R. A. (1978) *J. Am. Chem. Soc. 100*, 3971–3972.

Romaniuk, P. J., Hughes, D. W., Gregoire, R. J., Bell, R. A., & Neilson, T. (1979) *Biochemistry 18*, 5109–5116.

Rupley, J. A., Butler, L., Gerring, M., Hartdegen, F. J., & Pecoraro, R. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1088–1095.

Schellman, J. A. (1955) *C. R. Trav. Lab. Carlsberg, Ser. Chim. 29*, 223–229.

Shum, B. W.-K., & Crothers, D. M. (1983) *Biopolymers 22*, 919–933.

Sinclair, A., Alkema, D., Bell, R. A., Coddington, J. M., Hughes, D. W., Neilson, T., & Romaniuk, P. J. (1984) *Biochemistry 23*, 2656–2662.

Stone, M. P., Johnson, D. L., & Borer, P. N. (1981) *Biochemistry 20*, 3604–3610.

Sundaralingam, M. (1973) *Jerusalem Symp. Quantum Chem. Biochem. 5*, 417–455.

Sundaralingam, M. (1975) in *Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions* (Sundaralingum, M., & Rao, S. T., Eds.) pp 487–524, University Park Press, Baltimore, MD.

Turner, D. H., Freier, S. M., Sugimoto, N., Hickey, D. R., Jaeger, J. A., Sinclair, A., Alkema, D., Neilson, T., Caruthers, M. H., & Kierzek, R. (1985) in *3 D Structure and Dynamics of RNA* (van Knippenberg, P. H., Hilbers, C. W., & Bosch, L., Eds.) Plenum, New York (in press).