Sequence Dependence for the Energetics of Terminal Mismatches in Ribooligonucleotides†

Naoki Sugimoto,‡ Ryszard Kierzek,§ and Douglas H. Turner*,‡

Department of Chemistry, University of Rochester, Rochester, New York 14627, and Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-704 Poznan, Noskowskiego 12/14, Poland

Received December 3, 1986; Revised Manuscript Received March 2, 1987

ABSTRACT: Stability increments of terminal mismatches on the core helixes AUGCAU and UGCGCA are reported. Enthalpy, entropy, and free energy changes of helix formation were measured spectrophotometrically for 15 oligoribonucleotides containing the core sequences and various mismatches. Free energy increments for mismatches in this series range from −0.5 to −1.1 kcal/mol. These increments for mismatches on AU base pairs are smaller than those measured previously on GC base pairs [Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Biochemistry 25*, 3209–3213]. The terminal GU mismatches in the sequences GAUGCAUUp and UAUGCAUGp add approximately the same stability increment as the corresponding terminal AU mismatch. The stability increments for pyrimidine–pyrimidine and pyrimidine–purine mismatches can be approximated within 0.3 kcal/mol by adding the stability increments for the corresponding 3′ and 5′ unpaired nucleotides (dangling ends). Stability increments for purine–purine mismatches are approximated well by the stability increment for the corresponding 3′ dangling end made more favorable by 0.2 kcal/mol. These approximations are used to provide a table of stability increments for all 48 possible sequences of mismatches.

Helixes in RNA secondary structures and associations often end with terminal mismatches. In most cases, helixes are short. For example, the average length of accepted helixes in the secondary structures of both 16S and 23S ribosomal RNAs from *Escherichia coli* (Noller, 1984) is 7 base pairs. Previous studies have shown the stability increments associated with terminal mismatches are similar to those for base pairs (Hickey & Turner, 1985; Freier et al., 1986a). Thus the energetics for terminal mismatches have a significant impact on the stability of many RNA helixes and presumably help determine both the secondary and tertiary structures of an RNA. On the basis of the frequency of their occurrence, GA and GU mismatches are thought to be particularly important (Traub & Sussman, 1982; Mizuno & Sundaralingam, 1978; Noller, 1984; Woese et al., 1983). Despite the importance, little is known about the sequence dependence of the energetics of terminal mismatches. This paper reports experimental determinations of free energy increments associated with 15 terminal mismatches added to the core helixes AUGCAU and UGCGCA. Combined with previous measurements (Hickey & Turner, 1985; Freier et al., 1986) and empirical rules, the results provide reasonable approximations for the free energy increments associated with all 48 possible sequences of A, C, G, and U containing terminal mismatches.

Materials and Methods

Oligonucleotide synthesis and measurement of thermodynamic parameters were carried out as described in the preceding paper (Sugimoto et al., 1987). Extinction coefficients ($\times 10^{-4}$ cm^{-1} M^{-1}) calculated at 260 nm with the nearest-neighbor approximation (Richards, 1975) are for AAUGCAUAp, 7.21; for AAUGCAUCp, 6.76; for AAUGCAUGp, 7.52; for CAUGCAUAp, 6.67; for CAUGCAUCp, 6.58; for CAUGCAUUp, 6.70; for GAUGCAUAp, 7.03; for GAUGCAUGp, 7.40; for GAUGCAUUp, 6.92; for UAUGCAUCp, 6.24; for UAUGCAUGp, 7.00; for UAUGCAUUp, 6.84; for AUGCGCAAp, 6.39; for AUGCGCACp, 6.04; and for AUGCGCAGp, 6.79.

Results and Discussion

Experimental results for stabilities of the 15 oligonucleotide duplexes containing mismatches are shown in the supplementary material (see paragraph at end of paper regarding supplementary material) and listed in Tables I and II in the manner described in the preceding paper (Sugimoto et al., 1987). All the oligomers melt in a two-state manner, except for CAUGCAUCp and UAUGCAUGp. While $\Delta H°$'s determined for CAUGCAUCp and UAUGCAUGp from plots of T_M^{-1} vs. log C_T and shapes of melting curves differ by 17% and 15%, respectively, $\Delta G°$'s at 37 °C differ by less than 1%. Thus it is reasonable to compare $\Delta G°_{37}$'s for all the oligomers.

The results in Table I can be used to derive free energy increments for all possible mismatches on an AUGCAU core and for AA, AC, and AG mismatches on a UGCGCA core. These are listed in Table III, along with values measured previously on CCGG, GGCC, and GCGC cores (Hickey & Turner, 1985; Freier et al., 1986). Hickey and Turner (1985) suggested that mismatches on AU and GC pairs might have the same stability increment. Inspection of Table III indicates this is not true. For all cases where a mismatch has been studied on both AU and GC pairs, the stability increment is larger on the GC pair. This parallels observations on dangling ends (Sugimoto et al., 1987). To improve predictions of RNA secondary structure, Papanicolaou et al. (1984) found it necessary to make all terminal GC pairs 0.8 kcal/mol more stable than internal pairs. No stability was assigned to terminal mismatches. Studies on oligonucleotides indicate terminal and internal base pairs usually have the same stability (Sugimoto et al., 1987; Freier et al., 1986a,b). Thus it is

† This work was supported by National Institutes of Health Grant GM22939.

* Author to whom correspondence should be addressed.
‡ University of Rochester.
§ Polish Academy of Sciences.

0006-2960/87/0426-4559$01.50/0

Table I: Thermodynamic Parameters of Helix Formation[a,b]

oligomer	log C_T parameters			temperature-independent parameters		
	$-\Delta H°$[c] (kcal/mol)	$-\Delta S°$[c] (eu)	T_M[c,d] (°C)	$-\Delta H°$[e] (kcal/mol)	$-\Delta S°$[e] (eu)	T_M[d,e] (°C)
AAUGCAUAp	49.6	138.6	43.3	49.2	137.3	43.4
AAUGCAUCp	50.3	142.3	40.0	48.0	134.8	40.3
AAUGCAUGp	49.2	136.2	45.3	51.0	142.0	45.2
CAUGCAUAp	54.2	154.6	40.5	52.1	147.6	40.8
CAUGCAUCp	51.9	148.3	38.6	47.7	134.5	39.1
CAUGCAUUp	44.4	124.1	39.0	43.7	121.5	39.4
GAUGCAUAp	59.5	169.5	43.6	58.1	165.1	43.8
GAUGCAUGp	59.5	169.2	44.2	59.7	169.7	44.3
GAUGCAUUp	62.9	180.8	42.6	60.8	174.0	42.9
UAUGCAUCp	45.2	127.5	36.8	43.1	120.3	37.6
UAUGCAUGp	62.3	180.1	41.0	57.7	165.1	41.4
UAUGCAUUp	39.0	107.2	38.0	40.6	112.1	38.4
AUGCGCAAp	59.3	160.0	59.3	57.5	154.4	59.4
AUGCGCACp	56.0	150.2	59.4	55.9	149.7	59.5
AUGCGCAGp	57.6	154.2	60.9	59.0	158.3	60.8

[a] Measurements were in 1 M NaCl, 10 mM Na_2HPO_4, and 0.1 mM Na_2EDTA, 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 strand concentration. [e] Temperature-independent thermodynamic parameters are the average of those from plots of T_M^{-1} vs. log C_T and those from averaging fits of individual melting curves to a two-state model with sloping base lines.

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

oligomer	$-\Delta H°_{37}$[c] (kcal/mol)	$-\Delta S°_{37}$[d] (eu)	$-\Delta C_p°$ (cal mol^{-1} K^{-1})
AAUGCAUAp	48.2	133.5	294
AAUGCAUCp	46.0	127.4	249
AAUGCAUGp	51.7	144.4	333
CAUGCAUAp	50.2	142.5	459
CAUGCAUCp	44.4	124.2	376
CAUGCAUUp	45.2	125.6	817
GAUGCAUAp	56.3	161.3	425
GAUGCAUGp	59.0	166.8	592
GAUGCAUUp	57.3	162.6	716
UAUGCAUCp	46.8	132.9	1282
UAUGCAUGp	52.4	149.0	652
UAUGCAUUp	46.2	129.2	941
AUGCGCAAp	54.1	144.0	91
AUGCGCACp	54.0	143.4	94
AUGCGCAGp	57.4	154.3	147

[a] Measurements were in 1 M NaCl, 10 mM Na_2HPO_4, and 0.1 mM Na_2EDTA, pH 7. [b] Although estimated errors in $\Delta H°$ and $\Delta S°$ at the T_M are ±10%, additional significant figures are given to allow accurate calculation of T_M. [c] From plots of $\Delta H°$ vs. T_M. [d] From plots of $\Delta S°$ vs. ln T_M.

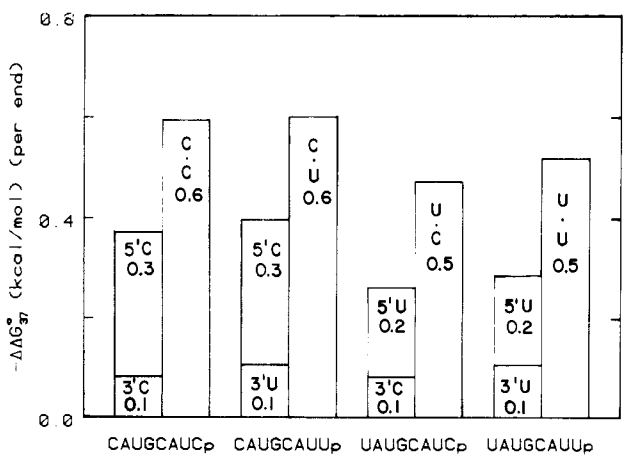

FIGURE 1: Free energy increments at 37 °C for adding a pyrimidine–pyrimidine terminal mismatch to an AUGCAU core. The left-hand columns represent the free energy increments for the dangling ends; the right-hand columns represent the free energy increments for the terminal mismatches. The data are from Table III and Sugimoto et al. (1987).

possible that some of the dissimilarity in treatment of terminal AU and GC pairs required by Papanicolaou et al. (1984)

Table III: Excess Stabilization (in Kilocalories per Mole) by Terminal Base Mismatches in 1 M NaCl[a]

added terminus	$-\Delta\Delta G°_{37}$ for core helix: AUGC-AU[b]	UGCG-CA[b]	GCGC or GGCC[b,c]	CCGG[b,c]
G·U mismatches				
5′Gp + 3′Up	1.0	(0.9)	1.9	2.3
5′Up + 3′Gp	0.8	(0.9)	1.5	1.4
purine–pyrimidine mismatches				
5′Ap + 3′Cp	0.7	0.6	(1.0)	1.1
5′Cp + 3′Ap	0.7	(0.9)	(2.1)	(1.3)
purine–purine mismatches				
5′Ap + 3′Ap	0.9	0.7	(2.0)	1.1
5′Ap + 3′Gp	1.1	0.9	(1.9)	1.6
5′Gp + 3′Ap	1.1	(0.9)	(2.0)	1.3
5′Gp + 3′Gp	1.1	(0.9)	(1.9)	1.5
pyrimidine–pyrimidine mismatches				
5′Cp + 3′Cp	0.6	(0.7)	(1.1)	(0.6)
5′Cp + 3′Up	0.6	(0.8)	(1.5)	(0.8)
5′Up + 3′Cp	0.5	(0.7)	(0.9)	(0.5)
5′Up + 3′Up	0.5	(0.8)	1.2	(0.7)

[a] $\Delta\Delta G°_{37}$ 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 core. Temperature-independent thermodynamic parameters were used to calculate $\Delta\Delta G°_{37}$. [b] Values in parentheses were not measured. They are estimated on the basis of the rules described in the text. [c] From Hickey and Turner (1985) and Freier et al. (1986). No correction has been applied to core with 3′-phosphate since this correction is small (Freier et al., 1985).

actually results from the differences in stability increments for terminal mismatches and dangling ends on AU and GC pairs.

Freier et al. (1986a) suggested stability increments for terminal mismatches could be approximated from stability increments for dangling ends and terminal AU base pairs. In particular, results for mismatches on GC pairs suggested stability increments for pyrimidine–pyrimidine and pyrimidine–purine mismatches were about equal to the sum of stability increments for the corresponding 5′ and 3′ dangling ends. Terminal GU mismatches were an exception. They are roughly equivalent to the corresponding terminal AU pair. Purine–purine mismatches were approximated by the corresponding 3′ dangling end made more favorable by 0.2 kcal/mol. The results in Table III provide tests of these rules for

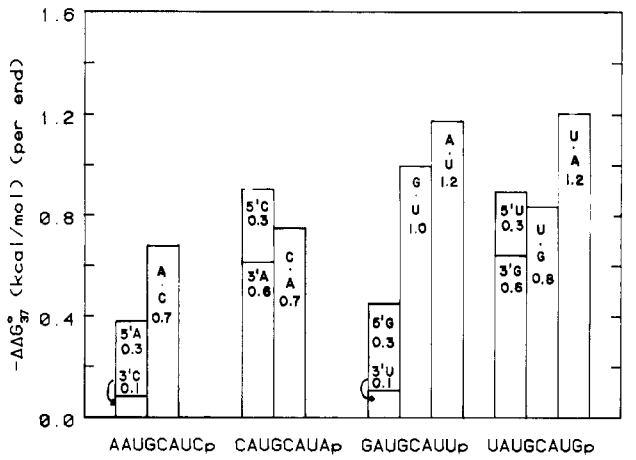

FIGURE 2: Free energy increments at 37 °C for adding a purine–pyrimidine terminal mismatch to an AUGCAU core. The left-hand columns represent the free energy increments for the dangling ends. For GU mismatches, the right-hand columns represent the free energy increments for the equivalent AU base pair.

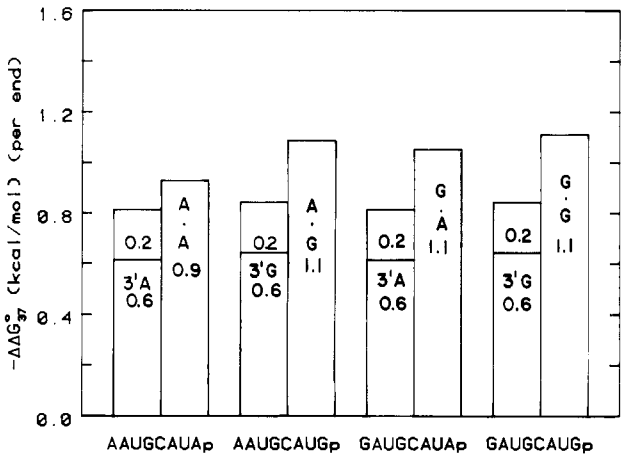

FIGURE 3: Free energy increments at 37 °C for adding a purine–purine terminal mismatch to an AUGCAU core. The left-hand columns represent the increments predicted by adding −0.2 kcal/mol to the increment for the appropriate 3′ dangling end (see text).

mismatches on AU pairs. Comparisons of the measured stability increments with those predicted are illustrated in Figures 1–4. In all cases, the predicted and measured values are within 0.3 kcal/mol. Thus the rules are reasonable for mismatches on both AU and GC pairs. These rules have therefore been used to predict stability increments for the 23 mismatches that have not been measured. The predictions are listed in parentheses in Table III to provide a complete catalog for stabilities of terminal mismatches.

The stability increments for terminal mismatches range from −0.5 to −2.3 kcal/mol, a large sequence dependence. This sequence dependence may be one factor affecting the prevalence of terminal mismatches in RNA. For example, in ribosomal RNA from *E. coli* (Noller, 1984), terminal mismatches on GC pairs are almost twice as frequent as on AU pairs. Moreover, there are almost 4 times as many purine–purine as pyrimidine–pyrimidine terminal mismatches. Both trends favor sequences with larger stability increments. Presumably, stability is important for much of ribosomal RNA structure, so natural selection may favor the most stable mismatches. If true, then opposite trends may be observed with messenger RNA where helixes must be disrupted for translation. Unfortunately, little is currently known about secondary structure of messenger RNAs, so this suggestion cannot presently be tested.

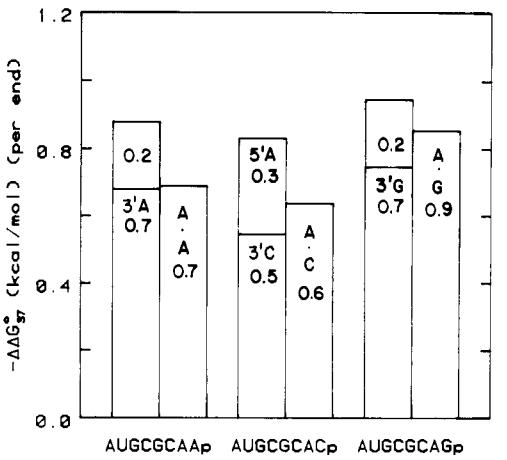

FIGURE 4: Free energy increments at 37 °C for adding a terminal mismatch to an UGCGCA core. The left-hand columns represent the increments predicted by the approximations described in the text.

While the stability increments listed in Table III are consistent with general trends of mismatch frequency in ribosomal RNA, they do not explain the overabundance observed for terminal GA mismatches (Traub & Sussman, 1982). Within experimental error, the measured GA mismatches have the same stability increments as the corresponding AA or GG mismatch with the same 3′ base. This suggests the specific structure of the GA mismatch is important.

It has been suggested that GA (Quigley & Rich, 1976) and GU (Crick, 1966) mismatches form hydrogen bonds in RNA. Studies of terminal Watson–Crick base pairs indicate formation of hydrogen bonds can make stability increments for pairs larger than the sum of the corresponding dangling ends (Freier et al., 1986a; Sugimoto et al., 1987). Inspection of Figure 2 indicates GAUGCAUUp is more stable than expected from the sum of the corresponding dangling ends. This suggests hydrogen bonds are formed in this terminal GU mismatch. Results for the other terminal mismatches are ambiguous. If hydrogen bonds are formed, the favorable free energy increment is only enough to counterbalance the unfavorable configurational increment from increased rigidity (Freier et al., 1986a). The results are consistent with the suggestion of Freier et al. (1985) that hydrogen bonds will be particularly strong when stacking interactions are weak.

In the preceding paper (Sugimoto et al., 1987), we suggested that stability increments for dangling ends may be useful for predicting stacking in the tertiary structure of RNAs. Nucleotides in terminal mismatches were treated as dangling ends. The results in Figures 1–4 suggest most mismatches are thermodynamically equivalent to their dangling ends, consistent with this treatment. The results for GAUGCAUUp, however, suggest a 3′ U mismatched with a G will form hydrogen bonds so that it is more likely to behave as a terminal AU pair (Crick, 1966) than as the equivalent dangling U. The parameters in Table III and in the preceding paper (Sugimoto et al., 1987) should be useful for further refining rules for prediction of tertiary structure when more three-dimensional structures are determined for RNA.

Supplementary Material Available

Four figures showing reciprocal melting temperature vs. log C_T (4 pages). Ordering information is given on any current masthead page.

References

Crick, F. H. C. (1966) *J. Mol. Biol. 19*, 548–555.

Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., & Turner D. H. (1985) *Biochemistry 24*, 4533-4539.
Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986a) *Biochemistry 25*, 3209-3213.
Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986b) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373-9377.
Hickey, D. R., & Turner, D. H. (1985) *Biochemistry 24*, 2086-2094.
Mizuno, H., & Sundaralingum, M. (1978) *Nucleic Acids Res. 5*, 4451-4461.
Noller, H. F. (1984) *Annu. Rev. Biochem. 53*, 119-162.
Papanicolaou, C., Gouy, M., & Nino, J. (1984) *Nucleic Acids Res. 12*, 31-44.
Quigley, G. J., & Rich, A. (1976) *Science (Washington, D.C.) 194*, 796-806.
Richards, E. G. (1975) *Handb. Biochem. Mol. Biol., 3rd Ed. 1*, 197.
Sugimoto, N., Kierzek, R., & Turner, D. H. (1987) *Biochemistry* (preceding paper in this issue).
Traub, W., & Sussman, J. L. (1982) *Nucleic Acids Res. 10*, 2701-2708.
Woese, C. R., Gutell, R., Gupta, R., & Noller, H. F. (1983) *Microbiol. Rev. 47*, 621-669.

Comparison of the Conformation and Orientation of Alamethicin and Melittin in Lipid Membranes

Horst Vogel
Department of Biophysical Chemistry, Biocenter, University of Basel, CH-4056 Basel, Switzerland
Received December 15, 1986; Revised Manuscript Received March 5, 1987

ABSTRACT: The secondary structure of alamethicin in lipid membranes below and above the lipid phase transition temperature T_t is determined by Raman spectroscopy and circular dichroism (CD) measurements. In both cases structural data are obtained by fitting the experimental spectra by a superposition of the spectra of 15 reference proteins of known three-dimensional structure. According to the Raman experiments, in a lipid bilayer above T_t alamethicin is helical from residue 1 to 12, whereas below T_t the helix extends from residue 1 to 16. The remaining C-terminal part is nonhelical up to the end residue 20 both above and below T_t. A considerable lower helix content is derived from CD, namely, 38% and 46% above and below T_t, respectively, in agreement with several reported values for CD in the literature. It is shown that the commonly used set of CD spectra of water-soluble reference proteins is unsuitable to describe the CD spectra of alamethicin correctly. Therefore the secondary structure of alamethicin as derived from CD measurements is at the present state of analysis unreliable. In contrast to the case of alamethicin, the CD spectra of melittin in lipid membranes are correctly described by the reference protein spectra. The helix content of melittin is determined thereby to be 72% in lipid membranes above T_t and 75% below T_t. The data are in accord with a structure where the hydrophobic part of melittin adopts a bent helix as determined recently by Raman spectroscopy [Vogel, H., & Jähnig, F. (1986) *Biophys. J. 50*, 573]. The orientational order parameters of the helical parts of alamethicin and of melittin in a lipid membrane are deduced from the difference between a corresponding CD spectrum of a polypeptide in planar multibilayers and that in lipid vesicles. The presented method for determining helix order parameters is new and may be generally applicable to other membrane proteins. The orientation of the helical part of both polypeptides depends on the physical state of the lipid bilayer at maximal membrane hydration and in the ordered lipid state furthermore on the degree of membrane hydration. Under conditions where alamethicin and melittin are incorporated in an aggregated form in a fluid lipid membrane at maximal water content the helical segments are oriented preferentially parallel to the membrane normal. Cooling such lipid membranes to a temperature below T_t changes the orientation of the helical part of alamethicin as well as melittin toward the membrane plane. On the contrary in dried planar membranes at 2% (w/w) water content both polypeptide molecules are oriented with their helical parts preferentially parallel to the membrane normal, similar to the case of fluid lipid membranes.

Several polypeptides exhibit similar properties to certain naturally channel-forming membrane proteins in the sense that after incorporation into lipid membranes they increase the membrane permeability depending on the transbilayer electrical potential. One of the most thoroughly investigated model compound in this context is the polypeptide alamethicin. Its channel properties have been characterized by a variety of electrical conductance measurements on planar bilayer membranes [for a review see Hall et al. (1984)]. From such experiments it was concluded that the alamethicin ion channel is composed of several polypeptide molecules. However, the structure of this channel is unknown.

The structure of alamethicin in crystals grown in a solvent mixture of methanol-acetonitrile has been determined. The molecule adopts a largely α-helical conformation with a bend in the helix axis near an internal proline residue (Fox & Richards, 1982). According to CD[1] measurements alamethicin is only about 40% helical when dissolved in apolar organic

[1] Abbreviations: CD, circular dichroism; NMR, nuclear magnetic resonance; IR, infrared; FA, fluorescence anisotropy; UV, ultraviolet; DMPC, dimyristoylphosphatidylcholine; DPPC, dipalmitoylphosphatidylcholine; DTPC, ditetradecylphosphatidylcholine; NRMSD, normalized standard deviation.

0006-2960/87/0426-4562$01.50/0