

Thermodynamic Stability of RNA Structures Formed by CNG Trinucleotide Repeats. Implication for Prediction of RNA Structure[†]

Magdalena Broda, Elżbieta Kierzek, Zofia Gdaniec, Tadeusz Kulinski, and Ryszard Kierzek*

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland

Received February 9, 2005; Revised Manuscript Received April 27, 2005

ABSTRACT: Trinucleotide repeat expansion diseases (TREDs) are correlated with elongation of CNG DNA and RNA repeats to pathological level. This paper shows, for the first time, complete data concerning thermodynamic stabilities of RNA with CNG trinucleotide repeats. Our studies include the stability of oligoribonucleotides composed of two to seven of CAG, CCG, CGG, and CUG repeats. The thermodynamic parameters of helix propagation correlated with the presence of multiple N-N mismatches within CNG RNA duplexes were also determined. Moreover, the total stability of CNG RNA hairpins, as well as the contribution of trinucleotide repeats placed only in the stem or loop regions, was evaluated. The improved thermodynamic parameters allow to predict much more accurately the thermodynamic stabilities and structures of CNG RNAs.

At present, almost 20 neurological diseases result from dynamic expansions of CNG (where N is A, C, U, and G) trinucleotide repeats (1). The pathological effect of those expansions can be related to biological processes involving CNG DNA and RNA, as well as proteins obtained during elongation of expanded CNG RNA tracks. The source of trinucleotide neurological diseases could also be associated with the structural changes of the CNG DNA and RNA tracks. So far, the investigations have mostly been focused on proteins, whereas CNG DNA and CNG RNA molecules, in particular, were much less intensively studied (2–5). Recently, it has been shown that for expanded CUG RNA tracks the pathogenesis process can be related to abnormal pre-mRNA splicing. Processing of pre-mRNA can be altered by the regulation and localization of CUG-binding proteins (2, 3).

A better understanding of the structures of DNA and RNA containing CNG repeats and, particularly, structural changes resulting from the expansion of trinucleotide repeats to the pathological level may be crucial for designing an effective therapy (4).

It was reported that CNG DNA molecules can adapt to various secondary structures such as hairpins, intra- and intermolecular triplexes, tetraplexes, and parallel duplexes (5–10). A very common phenomenon occurring in these molecules is the possibility of slipping of opposite strands and, as a consequence, formation of the slippery DNA (s-DNA) (11–13). The most important is the formation by CNG DNA of a structured hairpin, often accompanied with the slipping of the opposite strand of a hairpin stem (11–13). The analysis of thermodynamic stability of CNG DNA

hairpins formed by 10 and 30 repeats of CAG and CTG indicates that increasing the length of the oligonucleotides three times results in the enhancement of free energy by 40% only (14). In the cells of the patients with fragile X syndrome disease (FRAXA), the concentration of 0.18 mM aluminum salt was found. This is very interesting that CCG tracks, expansion of which is correlated to FRAXA, can form a very strong complex with aluminum salt. The formation of the complex dramatically changed the initial structure of CCG DNA, presumably to Z-form, which was confirmed by circular dichroism spectra (15).

At present, the studies concerning the structure of RNA with CNG trinucleotide repeats are very limited. Several known CNG RNAs containing more than 10 CNG RNA trinucleotide repeats formed hairpins; however, the slippery of the stem was observed as well. The formation of hairpins by CGG RNA trinucleotide repeats was also suggested, based on the cleavage pattern by T1 and S1 nucleases (16–21). On the other hand, it was found that the (CUG)₅ molecule adopts A-type RNA structure. Moreover, the formation of the structures similar to pseudoknots or triplexes was suggested, based on a spectroscopic analysis and mobility on nondenatured gel (6). The thermodynamic analyses of RNA fragments formed by 5, 10, 15, 20, 35, and 69 CUG repeats have shown that these molecules can adopt linear or linear and/or branched hairpin structures. The analysis of the melting process of (CAG)₂₉ and (CAG)₅₆ provided similar conclusions (22).

In this paper, the thermodynamic parameters of helix propagation correlated with the presence of multiple N-N mismatches within CNG RNA duplexes were determined. Moreover, the total stability of CNG RNA hairpins, as well as the contribution of trinucleotide repeats in the stem and/or loop regions, was evaluated. The improved thermodynamic parameters allow to predict much more accurately the thermodynamic stabilities and structures of CNG RNA.

[†] This work was supported by Polish State Committee for Scientific Research (KBN) Grant PZB-KBN 059/T09/14 to R.K. and NIH Grant 1R03 TW1068 to R.K. and D. H. Turner. E.K. was a recipient of a fellowship for young scientists from the Polish Science Foundation.

* To whom correspondence should be addressed. Phone: (061) 852-85-03. Fax: (061) 852-05-32. E-mail: rkierzek@ibch.poznan.pl.

EXPERIMENTAL PROCEDURES

Synthesis and Purification of Oligoribonucleotides. Oligoribonucleotides were synthesized on an Applied Biosystems DNA/RNA synthesizer, using β -cyanoethyl phosphoramidite chemistry (23). For the synthesis, standard and modified [inosine (I), purine riboside (P), 8-bromoadenosine (^{Br}A)] oligonucleotides, commercially available phosphoramidites containing the 2'-*O*-*tert*-butyldimethylsilyl groups, were used. The phosphoramidite of 8-bromoguanosine (^{Br}G) was synthesized according to the previously published procedure (24). The details of deprotection and purification of oligoribonucleotides were described previously (25).

UV Melting of Oligoribonucleotides. The oligoribonucleotides were melted in a standard buffer containing 1 M sodium chloride, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7.0. The oligoribonucleotide single strand concentrations were calculated from high-temperature (>80 °C) absorbencies and single strand extinction coefficients approximated by a nearest-neighbor model (25–27). The absorbency vs temperature melting curves were measured at 260 nm with a heating rate of 1 °C/min from 0 to 90 °C on a Beckman DU 640 spectrometer with a thermoprogrammer. There were nine melts for the duplexes and hairpins in a standard buffer at concentrations ranging from 10^{−3} to 10^{−6} M RNA. The melt curves were analyzed, and thermodynamic parameters were calculated using MeltWin 3.5 and Microcal Origin 6.0 programs.

Deconvolution of Experimental UV Melting Profiles Using Microcal Origin 6.0. In the melting data presented here, an unfolding of duplexes may take place in one or two two-state unfolding transitions due to the sequence and the presence of the mismatches. When multiple unfolding transitions are present, each transition is characterized with its melting temperature (T_m),¹ enthalpy (ΔH), entropy (ΔS), and a percent of hyperchromicity (ΔA). A precise determination of all these parameters may be difficult if transitions are closely spaced. The UV melting profiles were subjected to a nonlinear least-squares parameter estimation of T_m , ΔH_i , ΔS_i , and ΔA_i for each *i*th two-state unfolding transition within the multitransition-independent unfolding model, using the program created in the laboratory. The program is an extended version of the earlier program used for the analysis of RNA microcalorimetry data (28, 29).

The numerical fitting of the model function which describes a sum of independent two-state transitions to experimental data was performed on the basis of a nonlinear least-squares procedure using the Marquardt algorithm. The baseline position was also used as an independent fitting parameter. The routine for fitting takes the baselines into account by linear extrapolations of low- and high-temperature data into the transition region of the melting curve. The plots of the distribution of residuals between the experimental and fitted curves and the value of the reduced χ^2 were used to describe the appropriateness of fit of the data during the evaluation.

NMR Analysis. Oligonucleotides were dissolved in 90% H₂O/10% D₂O with 150 mM sodium chloride, 0.1 mM Na₂-

EDTA, and 10 mM phosphate buffered at pH 6.6. One-dimensional exchangeable proton NMR spectra were recorded on Bruker AVANCE 600 MHz spectrometer using a pulsed-field gradient watgate to suppress the water signal (30). The data were multiplied by a 1.0 Hz line broadening exponential function and Fourier transformed with Bruker XwinNMR software. In total, 64–128 scans were collected for each FID. The chemical shifts were indirectly referenced relative to DSS (30).

RESULTS

Understanding the structure and thermodynamic stability of CNG tracks is important to apply CNG RNA as therapeutic target. To achieve this goal, studies of thermodynamic stability of RNAs formed by up to seven repeats of CAG, CUG, CCG, and CGG were performed. It is relatively easy to predict that the RNAs formed by many CNG trinucleotide repeats will adopt the hairpin structures, whereas the shorter CNG oligoribonucleotides will exist mainly as RNA duplexes or as a mixture of both structures. The most interesting were the studies on (i) thermodynamic stability of RNA duplexes formed by CAG, CUG, CCG, and CGG repeats and the influence of the nature of multiple N-N mismatches on their stability, (ii) thermodynamic stability of RNA hairpins with CAG, CUG, CCG, and CGG repeats, (iii) equilibrium between the CNG RNA duplex and hairpin structures as a function of number and type of trinucleotide repeats, and (iv) contribution of CNG stem and loop regions, and the size of RNA loops, as well as the presence of unpaired ends on the overall thermodynamic stability of CNG RNA hairpins.

In this study, oligoribonucleotides of the sequences G(C-NG)_{*n*}C (*n* = 2–7) were used. Each oligoribonucleotide, with internal CNG repeats, starts with a guanosine residue at the 5'-side and terminates with cytidine at the 3'-side because guanosine is the last nucleotide of previous repeat, whereas cytidine is the first nucleotide of next CNG trinucleotide repeat. That allowed to collect thermodynamic data which better reflect thermodynamic properties of long internal CNG RNAs.

The chemical synthesis and purification of the oligoribonucleotides containing many guanosine residues were complicated; however, the quantitative analysis of experimental results was even more difficult. Meltwin 3.5 software, which is usually used to analyze the experimental melting curves, allows to calculate thermodynamic parameters such as enthalpy (ΔH°), entropy (ΔS°), and free energy (ΔG°_{37}) by two methods: fitting of the experimental and theoretical melting curves and linear correlation of melting temperatures ($1/T_m$) and concentration of oligoribonucleotides ($\log C_T$) (31). In the case of CNG tracks, the analysis of the experimental data demonstrated that the melts were often performed in a non-two-state manner; moreover, simultaneous melting of two CNG RNA structures was observed. To solve the problem of simultaneous melting of two RNA structures, Microcal Origin 6.0 (Originlab) was applied to perform a deconvolution of the experimental melting curves and to calculate the thermodynamic parameters on the basis of the analysis of experimental melting curves only.

Thermodynamic Stability of CNG RNA Duplexes. The UV melting analysis demonstrates that short oligoribonucleotides

¹ Abbreviations: C_T , total concentration of all strands of oligonucleotides in solution; N, any nucleotide including A, C, G, or U; T_m , melting temperature in degrees Celsius; FID, free induction decay; DSS, 3-(trimethylsilyl)propionic acid sodium salt.

Table 1: Thermodynamic Parameters for Helix Formation^a

no. of CNG repeats	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ_{37}$ (kcal/mol)	T_m (°C) ^b	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ_{37}$ (kcal/mol)	T_m (°C) ^b
CAG					CCG			
2	40.1 ± 4.2	110.1 ± 13.4	6.02 ± 0.07	39.8	29.3 ± 6.2	74.9 ± 19.2	6.09 ± 0.48	41.6
3	30.1 ± 12.0	75.4 ± 38.5	6.69 ± 0.22	48.1	42.2 ± 12.0	114.4 ± 37.9	6.71 ± 0.29	44.9
4	32.0 ± 10.2	80.2 ± 32.5	7.07 ± 0.22	51.4	45.8 ± 9.9	66.7 ± 16.1	6.05 ± 0.35	46.0
CGG					CUG			
2	50.1 ± 3.9	139.5 ± 12.4	6.78 ± 0.17	44.2	49.6 ± 3.7	141.9 ± 12.2	5.08 ± 0.18	33.7
3	50.7 ± 3.2	134.3 ± 9.7	9.09 ± 0.45	59.5	54.9 ± 6.4	147.9 ± 19.5	7.01 ± 0.31	44.3
4	58.0 ± 7.6	153.0 ± 23.8	10.56 ± 0.20	65.6	43.9 ± 11.1	116.7 ± 35.7	7.70 ± 0.13	52.2

^a Solutions are 1 M sodium chloride, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7. ^b Melting temperatures are calculated for 10⁻⁴ M oligomer concentration.

Table 2: Thermodynamic Parameters for Helix Formation with Modified Nucleotides^a

RNA sequence	1/T _M vs log C _T parameters				curve fit parameters					
	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ_{37}$ (kcal/mol)	T_m^a (°C)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ_{37}$ (kcal/mol)	T_m (°C) ^b	$\Delta\Delta G^\circ_{37}$ (kcal/mol)	ΔT_m (°C)
GCGGCGGC	56.7 ± 2.6	161.1 ± 8.4	6.76 ± 0.04	43.0	48.9 ± 3.0	135.9 ± 9.6	6.70 ± 0.16	43.6	0	0
GCI ⁺ GCGGC	65.5 ± 7.8	192.6 ± 25.4	5.78 ± 0.19	37.5	53.9 ± 3.6	154.8 ± 11.8	5.93 ± 0.20	38.4	0.77	-5.2
GCGGCI ⁺ GC	74.8 ± 14.1	220.8 ± 45.4	6.29 ± 0.48	39.6	63.2 ± 9.1	183.4 ± 29.5	6.33 ± 0.27	40.3	0.37	-3.3
GCI ⁺ GCI ⁺ GC	46.4 ± 3.1	132.9 ± 10.1	5.15 ± 0.11	33.5	48.5 ± 7.4	139.7 ± 24.0	5.14 ± 0.12	33.6	1.56	-10.0
GCPGCGGC	57.1 ± 2.6	167.0 ± 8.4	5.29 ± 0.06	34.9	53.7 ± 5.6	155.8 ± 18.7	5.41 ± 0.18	35.5	1.29	-8.1
GCGGCPGC	64.3 ± 6.9	186.6 ± 22.1	6.39 ± 0.18	40.5	50.9 ± 7.4	143.3 ± 24.5	6.46 ± 0.29	41.8	0.24	-1.8
GCPGCPGC	36.5 ± 3.6	101.2 ± 11.8	5.09 ± 0.17	32.1	47.4 ± 15.2	136.8 ± 49.6	4.96 ± 0.26	32.4	1.74	-11.2
GC ^{Br} GGCGGC	98.7 ± 5.0	280.9 ± 15.4	11.53 ± 0.24	56.6	76.4 ± 3.2	212.7 ± 9.6	10.45 ± 0.33	57.7	-3.70	14.1
GCGGC ^{Br} GGC	74.8 ± 2.4	208.2 ± 7.5	10.20 ± 0.12	57.0	86.3 ± 8.3	243.3 ± 25.2	10.80 ± 0.53	56.6	-4.10	13.0
GC ^{Br} GGC ^{Br} GGC	46.5 ± 3.6	129.8 ± 11.7	6.24 ± 0.09	40.8	52.7 ± 4.6	150.0 ± 15.3	6.19 ± 0.23	40.0	0.51	-3.6
GCAGCAGC	35.0 ± 1.3	94.6 ± 4.3	5.69 ± 0.03	37.1	38.7 ± 3.7	106.4 ± 11.8	5.69 ± 0.06	37.1	0	0
GC ^{Br} AGCAGC	49.1 ± 2.7	142.2 ± 9.0	5.03 ± 0.10	33.0	51.4 ± 4.8	149.5 ± 15.7	5.01 ± 0.11	33.1	0.68	-4.0
GCAGC ^{Br} AGC	44.9 ± 1.7	128.2 ± 5.7	5.15 ± 0.31	33.4	52.5 ± 5.1	153.1 ± 16.6	4.96 ± 0.12	32.8	0.73	-4.3
GC ^{Br} AGC ^{Br} AGC	33.8 ± 5.8	97.4 ± 20.0	3.63 ± 0.48	19.3	42.9 ± 30.7	127.4 ± 103.1	3.45 ± 1.33	21.7	2.24	-15.4

^a Solutions are 1 M sodium chloride, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7. ^b Melting temperatures are calculated for 10⁻⁴ M oligomer concentration.

composed of two and three CNG trinucleotide repeats form duplex structures, whereas the oligonucleotides containing four and five CNG repeats exist as a simultaneous composition of duplexes and hairpins. The CNG RNAs carrying six and more trinucleotide repeats predominantly form a hairpin. Table 1 presents the data containing the thermodynamic parameters for two to four CNG trinucleotide repeats which were obtained from the deconvolution of experimental melting curves and from quantitative analysis of these parts of the melting curves that are related to the melting of CNG duplexes.

The examination of the results obtained for CNG RNA duplexes formed by CAG and CCG repeats demonstrates their similar thermodynamic stabilities and shows that the free energies (ΔG°_{37}) oscillate around -6.5 ± 0.5 kcal/mol. The thermodynamic stabilities of the duplexes formed by CUG repeats (except those containing two CUG repeats) were ca. -7.5 ± 0.5 kcal/mol. The duplexes formed by three and four CGG repeats were up to 4 kcal/mol more stable than those formed by CAG and CCG repeats.

The noncanonical base pairs stabilized by hydrogen bonds were found in many RNA loops (32–35). It was expected that similar interactions within the mismatches such as A-A, U-U, C-C, and G-G might contribute to the enhancement of thermodynamic stability of CNG RNA duplexes. To probe the hydrogen bond interactions within N-N mismatches, the modified oligonucleotides were used. The thermodynamic stabilities of the following duplexes were measured: (GCI-GCGGC)₂, (GCGGCI⁺GC)₂, (GCI⁺GCI⁺GC)₂, (GCPGCG-

GC)₂, (GCGGCPGC)₂, and (GCPGCPGC)₂, where I and P correspond to inosine and purine riboside, respectively (35).

The analysis of thermodynamic stability of the modified CNG RNA duplexes collected in Table 2 provided evidence for the presence of hydrogen bond interactions within N-N mismatches, particularly strong within a G-G mismatch. In general, the replacement of a guanosine residue with inosine or purine riboside decreased the thermodynamic stability (ΔG°_{37}), and the effect depended on the position of the modified nucleotide within the duplex (36, 37). The stability (ΔG°_{37}) of (GCGGCGGC)₂, (GCI⁺GCGGC)₂, and (GCPGCGGC)₂ was -6.70 , -5.93 , and -5.41 kcal/mol, respectively. The deletion of the 2-amino group from guanosine resulted in a decrease of free energy by 0.77 kcal/mol, whereas the removal of both functional groups reduced the stability of (GCGGCGGC)₂ by 1.29 kcal/mol. The analyzed RNA duplexes were self-complementary, and the measured thermodynamic effect resulted from the presence of two G-G mismatches within the duplex. Owing to this, a hydrogen bond within a single G-G mismatch via the 6-oxygen contributed to the stability of ca. 0.25 kcal/mol whereas via the 2-amino group to ca. 0.38 kcal/mol. However, while making conclusions, one shall consider the fact that the contribution of the same functional groups in a hydrogen bonding can be different for guanosine and inosine residues. For unknown reasons, the substitution of guanosine at position 6 of (GCGGCGGC)₂ resulted in lower destabilization of the duplex. The thermodynamic stabilities (ΔG°_{37}) of (GCGGCI⁺GC)₂ and (GCGGCPGC)₂ are -6.33 and -6.46

kcal/mol, respectively. The substitution of guanosine-6 in (GCGGCGGC)₂ by inosine and purine riboside destabilized the reference duplex by 0.37 and 0.24 kcal/mol, respectively. Different thermodynamic effects of the substitution of guanosine by inosine and purine riboside simultaneously at positions 3 and 6 of (GCGGCGGC)₂ suggested that the structure of the duplex might not be symmetric. The substitution of guanosines with inosines and purine ribosides at positions 3 and 6 in the (GCGGCGGC)₂ duplex resulted in similar thermodynamic stability of (GCIGCIGC)₂ and (GCPGCPGC)₂, which came to -5.14 and -4.96 kcal/mol, respectively. The destabilization ($\Delta\Delta G^\circ_{37}$) was 1.56 and 1.74 kcal/mol for (GCIGCIGC)₂ and (GCPGCPGC)₂, respectively. Similar destabilization effect could suggest that within I-I and P-P mismatches hydrogen bond interactions do not exist or are of similar type.

The other group of modified nucleotides used to probe the interactions within G-G and A-A mismatches was 8-bromoguanosine and 8-bromoadenosine, respectively. It is well documented that the presence of a bulky substituent such as bromine at position 8 of guanosine and adenosine results in changing the conformation of the glycosidic bond from *anti* to *syn* (7, 38). 8-Bromo-modified analogues of guanosine and adenosine were placed at position 3 or 6 and, simultaneously, at positions 3 and 6 of (GCGGCGGC)₂ and (GCAGCAGC)₂, respectively. The data in Table 2 demonstrate that the replacement of one of the guanosines within a G-G mismatch by 8-bromoguanosine significantly increases the thermodynamic stability of the (GCGGCGGC)₂ duplex, and the stabilization depends on the position of 8-bromoguanosine within the duplex. The placement of 8-bromoguanosine at position 3 of the (GCGGCGGC)₂ duplex enhanced the stability ($\Delta\Delta G^\circ_{37}$) by 3.70 kcal/mol whereas at position 6 by 4.10 kcal/mol. When 8-bromoguanosine was located simultaneously at positions 3 and 6 of the duplex, the stability was quite similar to the reference duplex ($\Delta\Delta G^\circ_{37}$ was 0.51 kcal/mol). When 8-bromoadenosine was placed within (GCAGCAGC)₂ at position 3 or 6, the thermodynamic effect was less pronounced, and a destabilization ($\Delta\Delta G^\circ_{37}$) by 0.68 and 0.73 kcal/mol was observed, respectively (Table 2). However, the placement of two of the 8-bromoadenosines in an A-A mismatch resulted in the destabilization ($\Delta\Delta G^\circ_{37}$) of (GCAGCAGC)₂ by 2.06 kcal/mol. A minor destabilization effect introduced by a single 8-bromoadenosine suggests that the 6-amino group of adenosine does not contribute to the interactions within a A-A mismatch. This assumption is also supported by similar thermodynamic stability of the duplexes containing 8-bromoadenosine or purine riboside simultaneously at position 3 or 6 (Table 2). Significant destabilization of the duplex induced by the presence of 8-bromoadenosines simultaneously at positions 3 and 6 could result from the lack of interactions within an A-A mismatch or steric hindrance resulting from the presence of two 8-bromoadenosines in an opposite site of the strands.

The analysis of the thermodynamic stability of CNG-type duplexes containing G-G and A-A and duplexes modified with inosine, purine riboside, 8-bromoguanosine, and 8-bromoadenosine demonstrates that hydrogen bond interactions exist within the analyzed mismatches. Particularly, the interactions within G-G seem to be relatively strong.

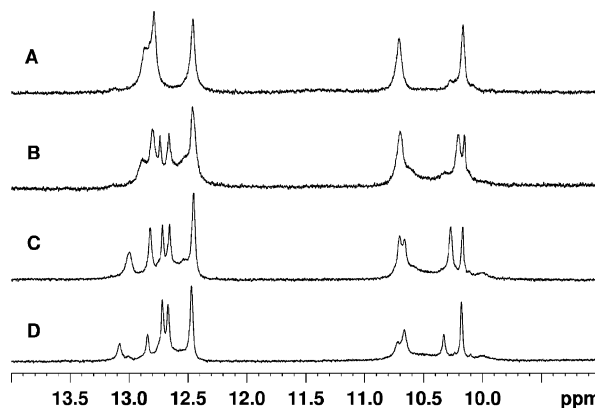


FIGURE 1: The exchangeable imino region of one-dimensional ¹H NMR spectra of (A) (GC^{Br}GGCGGC)₂ and (B–D) (GC^{Br}GGCGGCGGC)₂ recorded at 283 K in 90% H₂O and 10% D₂O with 0.1 mM EDTA and 10 mM phosphate buffered at pH 6.8. Spectra were taken at the following sodium chloride concentrations: 150 mM (A), 150 mM (B), 50 mM (C), and 0 mM (D).

NMR Spectra of 8-Bromoguanosine-Modified CGG RNA Duplexes. The NMR experiments were performed to better understand the unexpected stability of duplexes modified with 8-bromoguanosine. Figure 1 shows the exchangeable imino region of one-dimensional ¹H NMR spectra (9.0–14.0 ppm) for the sequences (GC^{Br}GGCGGC)₂ and (GC^{Br}GGCGGCGGC)₂. For both oligonucleotides studied, one observes two sets of imino protons (guanine imino protons in the range of 12.0–13.5 ppm typical for Watson–Crick GC base pairs and protons from the mismatched sites in the range of 9.7–11.0 ppm).

There are four possible structures of G-G mismatches that involve two hydrogen bonds. Three of them require one of the bases to be inverted either by switching the base from *anti* to *syn* or by reversing the direction of the strand (39). It is expected that the presence of a bulky substituent such as a bromine atom at the C8 position of guanosine that influences the rotation about the glycosidic bond, disfavoring the normal *anti* orientation of the base, will promote the formation of the ^{Br}G(*syn*)-G(*anti*) base pair (38). In the case of the symmetric duplex (GC^{Br}GGCGGC)₂, two resonances (10.17 and 10.71 ppm) in a region typical for mismatched base pairs are observed (Figure 1A). The area under each high-field resonance corresponds to one proton indicating the formation of a stable duplex with ^{Br}G-G base pairs involving two protons.

For the duplex (GC^{Br}GGCGGCGGC)₂, five well-resolved resonances corresponding to six protons in the region of Watson–Crick guanine imino protons (resonance at 12.46 ppm arises from two protons) and four high-field-shifted, relatively sharp resonances arising from ^{Br}G-G base pairs were observed (Figure 1B,C). Except for those signals, the resonances exchange-broadened almost to the baseline in both the high-field and low-field regions can be distinguished. The duplex (GC^{Br}GGCGGCGGC)₂ cannot be symmetric if the central G-G mismatch base is involved in hydrogen bonding. The analysis of NMR spectra of the duplexes containing a single G-G mismatch indicates that the effect of the mismatch site can propagate out of the adjacent base pairs to a different extent (32, 36–41). Various studies concerning DNA and RNA duplexes with a single G-G mismatch have previously shown that the structure of the mismatch site is also context dependent (32, 41). The

Table 3: Thermodynamic Parameters for CNG Hairpin Formation^a

RNA CNG hairpins	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ_{37}$ (kcal/mol)	T_m (°C) ^b
C(CAG) ₅ G	29.3 ± 4.1	86.6 ± 12.4	2.45 ± 0.30	65.2
C(CAG) ₆ G	35.5 ± 5.9	104.0 ± 17.4	3.22 ± 0.53	68.0
C(CAG) ₇ G	30.6 ± 3.7	89.4 ± 10.9	2.84 ± 0.34	68.8
C(CGG) ₅ G	19.9 ± 4.5	58.2 ± 13.5	1.86 ± 0.29	69.0
C(CGG) ₆ G	23.9 ± 3.3	69.8 ± 9.6	2.24 ± 0.33	69.0
C(CGG) ₇ G	23.4 ± 4.0	67.5 ± 11.7	2.51 ± 0.39	74.2
C(CUG) ₅ G	42.8 ± 7.2	128.3 ± 21.5	3.04 ± 0.55	60.6
C(CUG) ₆ G	39.8 ± 5.8	118.6 ± 17.1	3.07 ± 0.48	62.9
C(CUG) ₇ G	48.6 ± 4.3	145.5 ± 12.7	4.06 ± 0.40	65.2
C(CCG) ₄ G	22.5 ± 5.3	123.4 ± 31.6	1.78 ± 0.41	63.7
C(CCG) ₅ G	32.8 ± 4.1	121.5 ± 27.3	2.71 ± 0.35	65.0
C(CCG) ₆ G	24.4 ± 2.4	72.0 ± 7.1	2.10 ± 0.19	65.6
C(CCG) ₇ G	27.3 ± 6.5	80.1 ± 9.3	2.40 ± 0.49	67.2

^a Solutions are 1 M sodium chloride, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7. ^b Melting temperatures are calculated for 10⁻⁴ M oligomer concentration.

data obtained for the duplex (GC^{Br}GGCGGCGGC)₂ can either reflect the loss of the duplex symmetry caused by the presence of a mismatched G-G pair or result from the coexistence of two different conformations in a slow exchange on the NMR time scale. To gain more insight into the structures involved, one-dimensional spectra at 10 °C as a function of salt concentration were recorded. The lowering of sodium chloride concentration from 150 to 0 mM results in the lowering of an amplitude for half of the resonances (Figure 1D). This observation suggests that, in the duplex (GC^{Br}GGCGGCGGC)₂, the central G-G base pair forms alternative base-paired structures that are in equilibrium with intermediate exchange rate and are not introduced by asymmetry of the duplex. However, further studies are necessary to explain the effect introduced by the central G-G mismatch.

Thermodynamic Stability of CNG RNA Hairpins. It has been shown that, depending on the type and number of CNG trinucleotide repeats, RNA molecules can exist as hairpin structures, mixtures of hairpins and duplexes, or multi-branched hairpins (16, 17, 22). Of particular importance were the following problems: (i) a relationship between oligonucleotide length, the nature of mismatched nucleotides, and preferred hairpin over duplex formation, (ii) the thermodynamic stability of CNG RNA hairpins containing four- and/or seven-nucleotide loops (17), (iii) the evaluation of the individual contribution of a stem and four-nucleotide and seven-nucleotide CNG loops on the overall thermodynamic stability of CNG hairpins, and (iv) the influence of trinucleotide 5'- and 3'-dangling ends as a consequence of a slippage of opposite strands on the thermodynamic stability of CNG hairpins.

The oligoribonucleotides consisting of five and more CNG trinucleotide repeats (in the case of CCG repeats, even four repeats) can form hairpins as well as duplexes. However, when the number of repeats increases, the contribution of CNG hairpins gradually becomes dominating. The ratio between the CNG hairpins and the duplexes depends on the concentrations of both CNG oligoribonucleotides and sodium chloride. Low concentrations of both oligonucleotides and sodium chloride favor the formation of CNG hairpins (42). The analysis of Table 3 reveals that the thermodynamic stabilities (ΔG°_{37}) of CNG hairpins marginally depend on a number and type of CNG trinucleotide repeats and that free




















energies (ΔG°_{37}) oscillate between -2.0 and -3.5 kcal/mol. Due to the simultaneous presence of two RNA forms (hairpin and duplex) and the process of deconvolution of melting curves, the experimental error of the collected data is sometimes considerable.

The stability of the hairpin loop and stem contributes to the overall thermodynamic stability of CNG hairpins (42, 43). To evaluate the contribution of the CNG stem within the hairpin, the thermodynamic stability of 5'-G-CNG-CNG-C-UUCG-G-CNG-CNG-C 3' was determined (italics mark the nucleotides in a hairpin loop). The hairpin stem is composed of two CNG repeats on both strands and is flanked by guanosine and cytidine residues to enhance the internal character of CNG repeats as was explained in the case of CNG duplexes. The results in Table 4A demonstrate that the thermodynamic stabilities (ΔG°_{37}) of hairpins with UUCG loops are, on average, 1 kcal/mol higher than those reported for the hairpins of the same length, fully composed of CNG repeats (Table 3). The hairpin loops (UNCG stable tertaloo) were the same for all set RNA molecules, so the difference in a stability reflects the interaction within CNG repeats in the stem region only. The obtained thermodynamic stability (ΔG°_{37}) of the hairpins was -3.85, -3.36, -5.21, and -3.90 kcal/mol for CAG, CCG, CGG, and CUG repeats, respectively. The stabilities of hairpins, except those containing the CGG repeat, were similar, and the order was coherent with the thermodynamic stability of CNG RNA duplexes (see Table 1).

Another problem analyzed concerns the contribution of CNG loops to the total thermodynamic stability of CNG RNA hairpins. The examination of the sequences of CNG oligoribonucleotides and some preliminary results suggested a possibility of formation of two different types of hairpins consisting of four-nucleotide and seven-nucleotide loops. The thermodynamic stability was measured for two model groups of CNG oligonucleotides: 5'-GAG C-NG CN-G CUC3' and 5'-GAG C-NG CNG CN-G CUC3' (italics mark nucleotides in the hairpin loop). The analysis of Table 4B shows that the thermodynamic stability (ΔG°_{37}) of hairpins containing four-nucleotide CNG hairpin loops oscillates between -1.88 and -3.30 kcal/mol, and it suggests that the stability of four-nucleotide CNG hairpins is strongly correlated with the effect of 3'-dangling ends.

The results presented above suggest that heterocyclic bases within the four-nucleotide hairpin loops are not involved in hydrogen bond interactions. To support this assumption, the thermodynamic stabilities of three oligoribonucleotides, GAGC-5'XGCX3'-GCUC (X = 8-bromoguanosine, italics mark nucleotides in the hairpin loop), were analyzed. 8-Bromoguanosine was placed on the 5'-side or 3'-side as well as on both the 5'- and 3'-sides of the hairpin loop. The analysis of thermodynamic stability and melting character (relation 1/ T_m and log C_T) of the oligoribonucleotides demonstrated that the monosubstituted oligoribonucleotides formed hairpins, whereas in the case of bisubstituted oligonucleotides the duplex structure was favored. The monosubstituted hairpins containing 8-bromoguanosine at the 5'-side or 3'-side of the loop were thermodynamically more stable ($\Delta \Delta G^\circ_{37}$) than the reference hairpin by -3.25 and -4.27 kcal/mol, respectively (Table 4D). The stabilization effect ($\Delta \Delta G^\circ_{37}$) of monobromoguanosine-substituted hairpins was twice as high as that measured for CNG duplexes

Table 4: Thermodynamic Parameters for Hairpin Formation with CNG Trinucleotide Repeats^a

Sequence and structure of CNG hairpins		$-\Delta H^0$ (kcal/mol)	$-\Delta S^0$ (eu)	$-\Delta G_{37}^0$ (kcal/mol)	T_m ($^{\circ}\text{C}$) ^b	
A		N = A	38.2 ± 2.7	110.9 ± 7.8	3.85 ± 0.28	71.7
		N = C	34.8 ± 4.3	101.2 ± 12.7	3.36 ± 0.38	70.2
		N = G	44.7 ± 3.5	127.4 ± 10.3	5.21 ± 0.37	77.9
		N = U	39.7 ± 2.7	115.5 ± 8.2	3.90 ± 0.22	70.7
B		N = A	38.4 ± 12.4	113.0 ± 37.0	3.30 ± 0.99	66.2
		N = C	24.5 ± 8.6	72.9 ± 26.5	1.88 ± 0.39	62.8
		N = G	34.4 ± 4.4	101.2 ± 13.2	3.03 ± 0.48	67.0
		N = U	31.6 ± 5.8	93.3 ± 17.2	2.72 ± 0.47	66.3
C		N = A	25.0 ± 7.7	75.0 ± 23.0	1.76 ± 0.57	60.5
		N = C	48.8 ± 8.4	143.4 ± 25.1	4.34 ± 0.66	67.2
		N = G	39.5 ± 4.1	117.4 ± 11.9	3.10 ± 0.47	63.4
		N = U	49.1 ± 2.3	147.2 ± 6.9	3.46 ± 0.16	60.5
			36.4 ± 3.9	109.7 ± 11.6	2.36 ± 0.26	58.5
D		N1, N2 = G	34.4 ± 4.4	101.2 ± 13.2	3.03 ± 0.48	67.0
		N1 = G ^{Br} N2 = G	61.3 ± 4.0	177.6 ± 11.3	6.28 ± 0.52	72.4
		N1 = G N2 = G ^{Br}	70.1 ± 5.3	202.4 ± 15.1	7.30 ± 0.61	73.0
E		N1, N2 = 0	24.5 ± 8.6	72.9 ± 26.5	1.88 ± 0.39	62.8
		N1 = 0 N2 = CCG	44.8 ± 3.5	130.0 ± 10.1	4.45 ± 0.33	71.2
		N1 = CCG N2 = 0	18.7 ± 4.0	55.9 ± 12.1	1.42 ± 0.28	62.4

^a Solutions are 1 M sodium chloride, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7. ^b Melting temperatures are calculated for 10⁻⁴ M oligomer concentration.

containing the same modification (Table 1). The larger stability of the monosubstituted CNG hairpin cannot be simply explained by stacking of guanosine derivatives on the top of the hairpin stem. If it was the case, the larger stability of the hairpin (by ca. 1.7 kcal/mol) with 8-bromoguanosine at the 5'-side of loop would be expected. Presumably, several other factors might contribute to the stabilization effect of 8-bromoguanosine (44).

In the second group of CNG hairpins, the loop was designed to consist of seven nucleotides. The measurements of thermodynamic stability of several hairpin structures formed by oligoribonucleotides GAG C-5'NG CNG CN3'-G CUC (italics mark the nucleotides in the hairpin loop) were performed. The analysis of the data obtained for those oligoribonucleotides suggests the appearance of more complex structures than was observed for the corresponding

hairpins containing a four-nucleotide loop (Table 4C). The duplex and hairpin were found to coexist in a solution when N was uridine, whereas two different hairpin structures were observed in the case of cytidine. Furthermore, the order of stabilization was different from that obtained for CNG tetraloops. The CCG hairpin ($\Delta G_{37}^0 = -4.34$ kcal/mol) was more stable, whereas the CAG hairpin ($\Delta G_{37}^0 = -1.76$ kcal/mol) turned out to be less stable.

As mentioned above, two different hairpin structures were formed in the case of GAGC-5'CGCCG CC3'-GCUC, and the corresponding free energies were -3.46 and -2.36 kcal/mol. Although the structures of both forms are not known, one of the hairpins could have the expected seven-nucleotide loop, whereas the second hairpin could contain the four-nucleotide loop and CUG as the 3'-dangling end. The problem of slippery of long CNG trinucleotide tracks of RNA

and DNA was already reported before (11–13). Due to this phenomenon, the hairpin with a seven-nucleotide CNG loop can be converted into a hairpin containing a four-nucleotide loop and a CNG unpaired fragment on the 3'- or 5'-side. However, there are no data available concerning the influence of multinucleotide 5'- and 3'-dangling ends on thermodynamic stability of RNA hairpins. The effect of mononucleotide unpaired ends on the stability of RNA duplexes is well documented, although the data are limited (45–47).

The results of the thermodynamic stability of model hairpins formed by 5'GAGCCGCCGCUC**CCG**3' and 5'**CCG**-GAGCCGCCGCUC3' (italics denote the loop fragment, whereas bolds refer to the unpaired fragment) are collected in Table 4E. The examination of the data indicates that the 3'-CCG dangling end reveals stabilizing properties, whereas when placed at the 5'-end it does not essentially change the thermodynamic stability of the hairpin. The difference in the stability ($\Delta\Delta G^\circ_{37}$) of both model hairpins was 3.03 kcal/mol.

It was also interesting to evaluate the contribution of the loop and stem to the overall thermodynamic stability of CNG RNA hairpins. It is known that the total free energy of the hairpin ($\Delta G^\circ_{37 \text{ hairpin}}$) is a sum of the free energy of the stem ($\Delta G^\circ_{37 \text{ stem}}$) and loop ($\Delta G^\circ_{37 \text{ loop}}$) fragments (42, 43). The average penalties of loop formation were calculated on the basis of the total stability of CNG hairpins containing four- and seven-nucleotide loops. The free energies of four-nucleotide CNG hairpin loop formation ($\Delta G^\circ_{37 \text{ loop}}$) were estimated to be ca. 5.6 kcal/mol. The average free energy of four-nucleotide hairpin loop formation was previously reported to be 5.6 kcal/mol (43, 46). For seven-nucleotide CNG hairpin loops, free energies of the loop formation were calculated to be ca. 5.7 kcal/mol, whereas the literature data reported 5.9 kcal/mol for the same size of a loop (43, 46). It means that the free energies of CNG hairpin loop formation were basically the same as for random (non-CNG-type) RNA hairpin loops of the same size.

Implication for Prediction of a Structure of RNA Composed of CNG Trinucleotide Repeat Tracks. According to the nearest-neighbor model, the total free energy of RNA duplexes is a sum of free energies of initiation and propagation of the RNA duplex, as well as the free energy related to the symmetry of a duplex and the presence of an AU base pair at the terminal position (25–27). On the basis of that relation, thermodynamic parameters of RNA helix propagation related to the presence of N-N mismatches can be calculated. The values of 5'CN/3'GN were evaluated on the basis of the thermodynamic stability of CNG duplexes as well as on the stability of a stem of CNG hairpins (Table 5).

The measured thermodynamic stabilities of CNG duplexes are much lower than that predicted by existing RNA folding programs such as Mfold or RNAstructure. At present, thermodynamic parameters included in algorithms were calculated on the basis of the previously reported, thermodynamic effects of single A-A, C-C, G-G, and U-U mismatches within RNA duplexes (32, 46). It is well documented that the thermodynamic stability of RNA duplexes with single mismatches depends on the nature of a mismatch as well as on a sequence of adjacent base pairs and a position within a duplex (32).

Table 5: Thermodynamic Parameters for CNG RNA Helix Propagation in 1 M Sodium Chloride

5'GN 3'CN	ΔG°_{37} (kcal/mol)			$\Delta\Delta G^\circ_{37}$ (kcal/mol) ^d
	duplex ^a	duplex ^b	single N-N mismatch ^c	
5'GA 3'CA	1.70	0.90	0.55	1.15
5'CC 3'GC	1.65	0.80	1.02 ^e	0.63
5'CG 3'GG	0.80	−0.10	−0.70	1.50
5'CU 3'GU	1.40	0.95	0.22	1.18

^a Free energy of propagation of 5'GN/3'CN calculated for CNG helix without contribution of first and last CNG trinucleotide repeats. ^b Free energy of propagation of 5'GN/3'CN calculated for CNG helix from contribution of every CNG trinucleotide repeat. ^c Original propagation parameters of 5'GN/3'CN. ^d Differences between original (column 4) and improved propagation parameters of 5'GN/3'CN (column 2). ^e Propagation parameters of 5'GC/3'CC were calculated by extrapolation.

Table 6: Comparison of Experimental and Calculated Free Energies for CUG and CAG Oligo- and Polynucleotides

CNG RNA sequences	ΔG°_{37} (kcal/mol)		
	calcd by RNAstructure 4.0 based on original 5'GN/3'CN parameters	calcd by RNAstructure 4.0 based on improved 5'GN/3'CN parameters ^a	exptl data (UV melting method)
(CUG) _n			
n = 10	−10.0	−5.8	−5.7
n = 15	−16.0	−7.9	−7.6
n = 20	−25.0	−11.0	−9.2
n = 35	−46.0	−20.4	−12.3
n = 69	−112.0	−47.9	−14.9
(CAG) _n			
n = 29	−28.8	−3.8	−3.4
n = 56	−61.0	−3.3	−3.2

^a Free energy of propagation of 5'GN/3'CN calculated for CNG helix without contribution of first and last CNG trinucleotide repeats.

The improved propagation parameters for 5'CN/3'GN were calculated on the basis of thermodynamic data of CNG RNA duplexes and hairpins, using two different methods. In the first method, the effects of every CNG repeat were included in a calculation. In the second, only the effects of internal CNG trinucleotide repeats were analyzed (the effects of terminal CNGs were omitted in the calculation). The parameters calculated by the second approach appear to be more reliable, since in long CNG RNA structures the effect of terminal CNG can be neglected because most of the trinucleotides form internal CNG repeats. The differences of parameters of free energy propagation ($\Delta\Delta G^\circ_{37}$) for 5'CN/3'GN used currently by RNA folding programs and calculated on the basis of the experiments described herein oscillate between 0.63 and 1.50 kcal/mol (Table 5).

To test the accuracy of the improved thermodynamic parameters, the stabilities of several RNA structures formed by CUG and CAG trinucleotide repeats were analyzed (22). The comparison of experimental free energies as well as those calculated by RNAstructure 4.0 software using original and improved thermodynamic parameters is presented in Table 6. For CUG polyribonucleotides, the experimental free energies and those calculated using the improved parameters

are consistent with up to at least 20 CUG repeats. The difference becomes significant for longer RNA fragments, but it is still a fewfold smaller than that calculated on the basis of original thermodynamic parameters. In case of molecules consisting of 29 and 56 CAG trinucleotide repeats, the free energies calculated with the improved parameters perfectly match the experimental data.

The differences observed between the experimental free energies of CUG-type RNA and those calculated on the basis of the improved thermodynamic parameters could suggest that longer RNA carrying more than 25–30 of the CUG trinucleotide repeats adopt different structure than that predicted for a smaller number of CUG repeats, for example, branched CNG hairpins.

DISCUSSION

CNG RNAs are undoubtedly a promising target for therapeutic treatment of trinucleotide repeat expansion diseases. The use of RNA as a target allows to move toward the application of antisense strategies, various types of ribozymes, or interference RNA methods as potentially practical approaches. Here, systematic thermodynamic studies have been performed in order to more accurately predict the structure and stability of CNG RNAs at pathological level. The G(CNG)_{2–7}C oligoribonucleotides were selected for the studies since they correspond to naturally occurring trinucleotide repeat tracks. The presence of guanosine and cytidine residues at the ends of oligoribonucleotides made the first and last CNG trinucleotide fragment more similar to the internal one.

The experimental data reported in this work have shown that the oligoribonucleotides containing up to four CNG trinucleotide repeats predominantly existed as duplexes. The only exception was found for the G(CCG)₄C oligonucleotide for which a significant amount of hairpin structure was detected. The shapes of melting curves of many G(CNG)_{2–4}C oligoribonucleotides often indicated the non-two-state character of those processes. This phenomenon is quite common for RNA duplexes containing the structure disturbing the elements such as mismatches, internal loops, or modified nucleotides (32, 34, 48, 49). An interesting feature of CNG RNA tracks is the fact that the stabilities of G(CNG)_{2–4}C are very similar and not very much dependent neither upon the number of CNG RNA trinucleotide repeats nor upon the nature of mismatches, which is contrary to the thermodynamic stability of perfectly matched duplexes. The measured thermodynamic stabilities (ΔG°_{37}) of G(CAG)₂C, G(CUG)₂C, G(CCG)₂C, and G(CGG)₂C were -6.02 , -5.08 , -6.09 , and -6.78 kcal/mol, respectively. The RNA duplexes formed by oligoribonucleotides containing Watson–Crick base pairs instead of N–N mismatches were at least twice more stable. For example, the calculated free energies of 5′G(CAG)₂C/3′C(GUC)₂G, 5′G(CUG)₂C/3′C(GAC)₂G, 5′G(CCG)₂C/3′C(GGC)₂G, and 5′G(CGG)₂C/3′C(GCC)₂G RNA duplexes were -13.8 , -13.8 , -16.6 , and -16.6 kcal/mol, respectively. Moreover, the differences in stability ($\Delta\Delta G^{\circ}_{37}$) between 5′G(CUG)₂G/3′C(GAC)₂C and 5′G(CUG)₃G/3′C(GAC)₃C, as well as 5′G(CGG)₂G/3′C(GCC)₂C and 5′G(CGG)₃G/3′C(GCC)₃C, were 6.9 and 8.3 kcal/mol for both pairs of duplexes, respectively. The free energies of CAG and CCG RNA duplexes oscillate around -6.5 ± 0.5 kcal/mol,

whereas those of CUG RNA duplexes oscillate around -7.5 ± 0.5 kcal/mol, except for the duplex formed by two CUG trinucleotide repeats for which the free energy (ΔG°_{37}) was only -5.08 kcal/mol. The thermodynamic stability of G(CGG)₂C (ΔG°_{37} was -6.78 kcal/mol) was similar to those observed for other CNG RNA duplexes, whereas the structures formed by three and four CCG repeats were more stable (in comparison to CAG and CCG RNA duplexes) by 2.3 and 3.8 kcal/mol, respectively.

The presence of hydrogen bonds within a G–G mismatch and the contribution of both the 2-amino group and 6-oxygen of guanosine to those interactions were demonstrated when the analogues of guanosine, inosine and purine riboside, were applied. Moreover, the replacement of guanosine and adenosine residues by 8-bromoguanosine and 8-bromoadenosine, respectively, visualizes the importance of the stereo orientation of guanine and adenine residues within G–G and A–A mismatches. That observation is consistent with the results reported by Burkard and Turner (33, 36, 37). The authors, on the basis of the NMR structure of (GCAGGCGUGC)₂ and functional group mutation of the G–G mismatch, proved that the mismatch simultaneously adopted G(4)*syn*–G(7)*anti* and G(4)*anti*–G(7)*syn* conformations and the population of both conformers depended on a position within a (GCAGGCGUGC)₂ duplex. The stabilizing effect of 8-bromoguanosine is presumably related to two phenomena. The first one concerns the presence of a bromine atom at position C-8 of guanosine promoting the change of glycosidic bond conformation from *anti* to *syn* (38). The second reason is most probably related to the observation that G–G mismatch might exist simultaneously in two dynamic conformations (37). This dynamic transition of the conformations is related to some energetic penalty. Since 8-bromoguanosine exists predominantly in the *syn* conformation, probably only one of those dynamic structures is dominant. As a consequence, the energy related to the conformational entropy effect within G–G now contributes to the overall stability of the RNA duplex.

It has been shown that the total stability of CNG RNA hairpins was mostly determined by the stability of the stems. The thermodynamic stability of the CNG RNA hairpin stem was found to be similar to the stability of CNG RNA duplexes. That correlation was particularly strongly supported by the analysis of the thermodynamic stability of the hairpins constructed in such a way that stem regions were formed by two consecutive CNG trinucleotide repeats, whereas the sequences of the loop were identical (UUCG) (Table 4A).

In this paper, it has been shown that stabilization effects of four- and seven-nucleotide CNG loops in CNG RNA hairpins were compatible to those reported in “random” (non-CNG-type) hairpin loops (43, 46). Some differences in the stability between CNG hairpins reflect the effect of 5′-side nucleotides of the loop. The nucleotide at this position can also be considered as a 3′-unpaired nucleotide (3′-dangling end) to the helical hairpin stem. That can be the source of the stabilization effect of the nucleotide at that position. It was reported previously that the stabilization effect of 3′-unpaired ends of 5′CX/G terminated RNA duplexes ($X = A, U, C$, and G) was -1.7 , -1.2 , -0.8 , and -1.7 kcal/mol, respectively (45–47). The measured thermodynamic stabilities of four-nucleotide CNG hairpins were -3.30 , -2.72 , -1.88 , and -3.03 kcal/mol for $N = A, U, C$, and G in the

loops, respectively (Table 4B). Since the 5'-nucleotide of the four-nucleotide hairpin loop can be considered as a 3'-dangling end of the helical stem, the correlation of hairpin stability and stacking ability of the 5'-nucleotide in the loop is understandable.

The factors important for the stability of the hairpins containing seven-nucleotide CNG loops are less clear. It seems that CNG RNA hairpins containing seven-nucleotide hairpin loops are formed less favorably than those composed of four nucleotides. So far, there are no data concerning the thermodynamic stability of hairpins containing seven-nucleotide loops. However, there is relatively a lot of information concerning six-nucleotide loops (50–52). It was shown that the loops closed by C-C are more stable than those closed by U-U and G-G (there are no data for the A-A closing loop) (50). Moreover, as published previously, the hairpins containing the U₉ loop are more stable than the A₉ loop (53). The thermodynamic data presented in Table 4C are consistent with those two observations.

For some of the CNG hairpins (especially CCG), due to the slippery of the stem, the formation of tetraloop hairpins with a trinucleotide 3'-dangling end was postulated (11–13). The presence of a trinucleotide 3'-unpaired end in the CNG RNA hairpin stem strongly stabilized that structure ($\Delta\Delta G^\circ_{37}$ ca. 2.5 kcal/mol), whereas the 5'-side effect was negligible. Previously, Bibillo et al. observed a similar stabilization effect of the RNA duplex ($\Delta\Delta G^\circ_{37} = 2.63$ kcal/mol) when UAA was placed as the 3'-dangling end (54). Moreover, Sugimoto et al. demonstrated that the 3'-AAA dangling end stabilized ($\Delta\Delta G^\circ_{37}$) the RNA duplex by 2.4 kcal/mol, whereas when placed on the 5'-side, it enhanced its thermodynamic stability by 0.8 kcal/mol only (55).

The analysis of an influence of CNG RNA trinucleotide repeats on thermodynamic stability of double-stranded (duplexes and hairpin stems) and single-stranded (hairpin loops) regions of CNG RNAs demonstrates that the free energies of CNG double-stranded fragments are significantly lower than those corresponding to matched RNA, whereas the thermodynamic stability of single-stranded (hairpin loop) CNG and “random” (non-CNG) RNA is compatible. This observation allows to conclude that, for longer tracks of CNG RNA repeats, formation of “branched” CNG hairpins (with at least two loops) should be thermodynamically more favorable than “linear” CNG hairpin structures (with only one loop).

RNAstructure 4.0 software was used for prediction of CNG RNA structures with improved 5'CN/3'GN thermodynamic parameters. The comparison of the structures and thermodynamic stabilities of CNG RNA predicted by software with original and improved 5'CN/3'GN propagation thermodynamic parameters indicated large discrepancies among them.

Interesting results were obtained from the comparison of the thermodynamic stability of several CNG RNAs calculated on the basis of (i) original parameters, (ii) two types of improved propagation thermodynamic parameters 5'CN/3'GN, and (iii) experimentally measured thermodynamic stabilities. For CUG RNAs containing 10 and 15 repeats, the calculated (using the improved thermodynamic parameter for internal CUG) thermodynamic stabilities (ΔG°_{37}) matched exactly the experimental free energy values. For 20 CUG repeats, the free energy was different by 20%, whereas for

35 and 69 CUG repeats, the differences were 66% and 320%, respectively. The comparison of experimental free energies and those calculated using original parameters demonstrates the difference of 2 to almost 8 times. For CAG RNAs (29 and 56 repeats), the free energy calculated on the basis of the improved parameters matched very well ($\Delta\Delta G^\circ_{37}$ up to 12%) the experimental data, whereas the free energy calculated on the basis of original parameters differs up to 19 times (22).

The presented results demonstrate that double-stranded CNG RNAs are thermodynamically very unstable and their stabilities are not much dependent upon the nature of N-N mismatches and the length of trinucleotide tracks. That observation may be important for understanding the many biological processes involving CNG RNAs. Any therapeutic strategy targeted toward CNG RNA will benefit from the knowledge of their thermodynamic stability.

ACKNOWLEDGMENT

We thank Tianbing Xia and Douglas H. Turner (University of Rochester) for the data concerning thermodynamic stability of CUG and CAG RNA.

REFERENCES

1. Everett, C. M., and Wood, N. M. (2004) Trinucleotide repeats and neurodegenerative disease, *Brain* 127, 2385–2405.
2. Ranum, L. P. W., and Day, J. W. (2004) Myotonic dystrophy: RNA pathogenesis comes into focus, *Am. J. Hum. Genet.* 74, 793–804.
3. Ranum, L. P. W., and Day, J. W. (2004) Pathogenic RNA repeats: an expanding role in genetic disease, *Trends Genet.* 20, 506–512.
4. Masino, L., and Pastore, A. (2001) A structural approach to trinucleotide expansion diseases, *Brain Res. Bull.* 56, 183–189.
5. Cummings, Ch. J., and Zoghbi, H. Y. (2000) Fourteen and counting: unraveling trinucleotide repeat diseases, *Hum. Mol. Genet.* 9, 909–916.
6. Pinheiro, P., Scarlett, G., Rodger, A., Rodger, P. M., Murray, A., Brown, T., Newbury, S. F., and McClellan, J. A. (2002) Structures of CUG repeats in RNA. Potential implications for human genetic diseases, *J. Biol. Chem.* 277, 35183–35190.
7. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) The trinucleotide repeat sequence d(CGG)₁₅ forms a heat-stable hairpin containing Gsyn-Ganti base pairs, *Biochemistry* 34, 12803–12811.
8. Mitas, M. (1997) Trinucleotide repeats associated with human disease, *Nucleic Acids Res.* 25, 2245–2254.
9. Sinden, R. R. (1999) Biological implications of the DNA structures associated with disease-causing triplet repeats, *Am. J. Hum. Genet.* 64, 346–353.
10. LeProust, E. M., Pearson, Ch. E., Sinden, R. R., and Gao, X. (2000) Unexpected formation of parallel duplex in GAA and TTC trinucleotide repeats of Friedreich's ataxia, *J. Mol. Biol.* 302, 1063–1080.
11. Pearson, C. E., Tam, M., Wang, Y. H., Montgomery, S. E., Dar, A. C., Cleary, J. D., and Nichol, K. (2002) Slipped-strand DNAs formed by long (CAG)_n(CTG) repeats: slipped-out repeats and slip-out junctions, *Nucleic Acids Res.* 30, 4534–4547.
12. Pearson, Ch. E., Wang, Y.-H., Griffith, J. D., and Sinden, R. R. (1998) Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)_n(CAG)_n repeats from the myotonic dystrophy locus, *Nucleic Acids Res.* 23, 816–823.
13. Gacy, A. M., and McMurray, C. T. (1998) Influence of hairpins on template reannealing at trinucleotide repeat duplexes: a model for slipped DNA, *Biochemistry* 37, 9426–9434.
14. Petruska, J., Arnheim, N., and Goodman, M. F. (1996) Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases, *Nucleic Acids Res.* 24, 1992–1998.
15. Latha, K. S., Anitha, S., Rao, K. S., and Viswamitra, M. A. (2002) Molecular understanding of aluminum-induced topological changes

- in (CCG)₁₂ triplet repeats: relevance to neurological disorders, *Biochim. Biophys. Acta* 1588, 56–64.
16. Jasinska, A., Michlewski, G., de Mezer, M., Sobczak, K., Kozłowski, P., Napierala, M., and Krzyzosiak, W. J. (2003) Structures of trinucleotide repeats in human transcripts and their functional implications, *Nucleic Acids Res.* 31, 5463–5468.
 17. Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W. J. (2003) RNA structure of trinucleotide repeats associated with human neurological diseases, *Nucleic Acids Res.* 31, 5469–5482.
 18. Michlewski, G., and Krzyzosiak, W. J. (2004) Molecular architecture of CAG repeats in human disease related transcripts, *J. Mol. Biol.* 340, 665–679.
 19. Sobczak, K., and Krzyzosiak, W. J. (2004) Imperfect CAG repeats form diverse structures in SCA1 transcripts, *J. Biol. Chem.* 279, 41563–41572.
 20. Sobczak, K., and Krzyzosiak, W. J. (2004) Patterns of CAG repeat interruptions in SCA1 and SCA2 genes in relation to repeat instability, *Hum. Mutat.* 24, 236–247.
 21. Napierala, M., and Krzyzosiak, W. J. (1997) CUG repeats present in myotonin kinase RNA form metastable “slippery” hairpins, *J. Biol. Chem.* 272, 31079–31085.
 22. Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., and Thornton, Ch. A. (2000) Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR, *RNA* 6, 79–87.
 23. McBride, L. J., and Caruthers, M. H. (1983) An investigation of several deoxynucleosides phosphoramidites useful for synthesizing oligodeoxynucleotides, *Tetrahedron Lett.* 24, 245–249.
 24. Proctor, D. J., Kierzek, E., Kierzek, R., and Bevilacqua, P. C. (2003) Restricting the conformational heterogeneity of RNA by specific incorporation of 8-bromoguanosine, *J. Am. Chem. Soc.* 125, 2390–2391.
 25. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, Ch., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplex with Watson–Crick base pairs, *Biochemistry* 37, 14719–14735.
 26. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, Th., and Turner, D. H. (1986) Improved free-energy parameters for predictions of RNA duplex stability, *Proc. Natl. Acad. Sci. U.S.A.* 83, 9373–9377.
 27. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices, *J. Mol. Biol.* 86, 843–853.
 28. Kulinski, T., Bratek-Wiewiórska, M. D., Wiewiórowski, M., Zielenkiewicz, A., Zółkiewski, M., and Zielenkiewicz, W. (1991) Comparative calorimetric studies on the dynamic conformation of plant 5S rRNA: Structural interpretation of the thermal unfolding patterns for lupin seeds and wheat germ, *Nucleic Acids Res.* 19, 2449–2455.
 29. Kulinski, T., Bratek-Wiewiórska, M. D., Zielenkiewicz, A., and Zielenkiewicz, W. (1997) Mg²⁺ Dependence of the structure and thermodynamics of wheat germ and lupin seeds 5S rRNA, *J. Biomol. Struct. Dyn.* 14, 495–507.
 30. Pitto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR* 2, 661–662.
 31. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp, *Biochemistry* 22, 256–263.
 32. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of single mismatches in RNA duplexes, *Biochemistry* 38, 14214–14223.
 33. Burkard, M. E., Turner, D. H., and Tinoco, I., Jr. (1999) Structures of base pairs involving at least two hydrogen bonds, in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) 2nd ed., pp 675–680, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
 34. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (1999) The energetics of small internal loops in RNA, *Biopolymers* 52, 157–167.
 35. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A-C, C-C, G-G, U-C and U-U mismatches in RNA internal loops: evidence of stable hydrogen-bonded U-U and C-C+ pairs, *Biochemistry* 30, 8242–8251.
 36. Burkard, M. E., and Turner, D. H. (2000) NMR structures of r(GCAGGCGUGC)₂ and determinants of stability for single guanosine-guanosine base pairs, *Biochemistry* 39, 11748–11762.
 37. Burkard, M. E., Xia, T., and Turner, D. H. (2001) Thermodynamics of RNA internal loops with a guanosine-guanosine pair adjacent to another noncanonical pair, *Biochemistry* 40, 2478–2483.
 38. Tavale, S. S., and Sobell, H. M. (1970) Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, two purine nucleosides in the *syn* conformation, *J. Mol. Biol.* 48, 109–123.
 39. Cognet, J. A. H., Gabarro-Arpa, J., Le Bert, M., van der Marel, G. A., van Boom, J. H., and Fazakerley, G. V. (1991) Solution conformation of an oligonucleotide containing a G-G mismatch determined by nuclear magnetic resonance and molecular mechanics, *Nucleic Acids Res.* 19, 6771–6779.
 40. Bhattacharya, P. K., Cha, J., and Barton, J. K. (2002) H-1 NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches, *Nucleic Acids Res.* 30, 4740–4750.
 41. Peyret, N., Seneviratne, A., Allawi, H. T., and SantaLucia, J., Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A-A, C-C, G-G, and T-T mismatches, *Biochemistry* 38, 3468–3477.
 42. Turner, D. H. (2000) Conformational changes, in *Nucleic Acids: Structures, Properties and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 259–334, University Science Books, Sausalito, CA.
 43. Serra, M. J., and Turner, D. H. (1995) Predicting thermodynamic properties of RNA, *Methods Enzymol.* 259, 242–261.
 44. Proctor, D. J., Ma, H., Kierzek, E., Kierzek, R., Gruebele, M., and Bevilacqua, Ph. C. (2004) Folding thermodynamics and kinetics of YNMG RNA hairpins: specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate, *Biochemistry* 43, 14004–14014.
 45. Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., and Turner, D. H. (1983) Effects of 3-dangling and stacking on the stability of GGCC and CCGG double helices, *Biochemistry* 22, 6198–6206.
 46. Mathews, D. H., Sabina, J., Zucker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol.* 288, 911–940.
 47. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) Stability of XGCGCp, GCGCYp, and XGCGCYp helices: an empirical estimate of the energetics of hydrogen bonds in nucleic acids, *Biochemistry* 25, 3214–3219.
 48. Kierzek, E., and Kierzek, R. (2001) Influence of N6-isopentenyladenosine on thermal stability of RNA duplexes, *Biophys. Chem.* 91, 135–140.
 49. Ziomek, K., Kierzek, E., and Kierzek, R. (2002) The thermal stability of RNA duplexes containing modified base pairs placed at internal and terminal positions of the oligoribonucleotides, *Biophys. Chem.* 97, 233–241.
 50. Fountain, M. A., Serra, M. J., Krugh, T. R., and Turner, D. H. (1996) Structural features of a six-nucleotide RNA hairpin loop found in ribosomal RNA, *Biochemistry* 35, 6539–6548.
 51. Zhang, H., Fountain, M. A., and Krugh, T. R. (2001) Structural characterization of a six-nucleotide RNA hairpin loop found in *Escherichia coli*, r(UUAAGU), *Biochemistry* 40, 9879–9886.
 52. Huang, S., Wang, Y. X., and Draper, D. E. (1996) Structure of a hexanucleotide RNA hairpin loop conserved in ribosomal RNAs, *J. Mol. Biol.* 258, 308–321.
 53. Krol, A., and Carbon, Ph. (1989) A guide for probing native small nuclear RNA and ribonucleoprotein structure, *Methods Enzymol.* 180, 212–226.
 54. Bibillo, A., Figlerowicz, M., Ziomek, K., and Kierzek, R. (2000) The nonenzymatic hydrolysis of oligoribonucleotides VII. Structural elements affecting hydrolysis, *Nucleosides, Nucleotides Nucleic Acids* 19, 977–994.
 55. Ohmichi, T., Nakano, S., Miyoshi, D., and Sugimoto, N. (2002) Long RNA dangling end has large energetic contribution to duplex stability, *J. Am. Chem. Soc.* 124, 10367; positions of the oligoribonucleotides, *Biophys. Chem.* 97, 233–241.