146–153 Nucleic Acids Research, 1995, Vol. 23, No. 1

The importance of a single G in the hairpin loop of the iron responsive element (IRE) in ferritin mRNA for structure: an NMR spectroscopy study

H. Sierzputowska-Gracz[1,2,*], R. Ann McKenzie[1] and Elizabeth C. Theil[1,3]

Departments of [1]Biochemistry, [2]Chemistry and [3]Physics, North Carolina State University, Raleigh, NC 27695-7622, USA

Received September 14, 1994; Revised and Accepted November 22, 1994

ABSTRACT

Noncoding sequences regulate the function of mRNA and DNA. In animal mRNAs, iron responsive elements (IREs) regulate the synthesis of proteins for iron storage, uptake and red cell heme formation. Folding of the IRE was indicated previously by reactivity with chemical and enzymatic probes. ^{1}H- and ^{31}P-NMR spectra now confirm the IRE folding; an atypical ^{31}P-spectrum, differential accessibility of imino protons to solvents, multiple long-range NOEs and heat stable subdomains were observed. Biphasic hyperchromic transitions occurred (52 and 73°C). A G–C base pair occurs in the hairpin loop (HL) (based on dimethylsulfate, RNAase T_1 previously used, and changes in NMR imino proton resonances typical of G–C base pairs after G/A substitution). Mutation of the hairpin loop also decreased temperature stability and changed the ^{31}P-NMR spectrum; regulation and protein (IRP) binding were previously shown to change. Alteration of IRE structure shown by NMR spectroscopy, occurred at temperatures used in studies of IRE function, explaining loss of IRP binding. The effect of the HL mutation on the IRE emphasizes the importance of HL structure in other mRNAs, viral RNAs (e.g. HIV-TAR), and ribozymes.

INTRODUCTION

Present understanding of RNA folding motifs is incomplete because knowledge of three-dimensional structures of RNAs is limited (1). However, structural analysis of different motifs of RNA by heteronuclear multidimensional NMR spectroscopy (^{13}C and/or ^{15}N-labeled RNA) is increasing rapidly (2–4). Folding of RNA secondary structure, known in group I introns /ribozymes (5,6), HIV (TAR) (7–9), t-RNA (10,11) and rRNA (sarcin/ricin loop) (12), appears to yield highly specific, three-dimensional structures stabilized by backbone and side chain interactions, much like proteins. The specific shape produced by RNA folding will contribute to the recognition between RNA and proteins, e.g. HIV-TAR/Tat (9), rRNA and elongation factors (12) as well as between RNA and RNA, exemplified by self-splicing introns (5,6). However, even before three-dimensional structures can be generated from a complete set of proton–proton internuclear distance restraints, it is possible to use other NMR parameters, such as chemical shifts, temperature studies, steady state NOE and relaxation parameters to generate useful insights about the structure and dynamics of nucleic acid fragments of interest (see e.g. references 13–15).

The iron responsive element (IRE) family is the best characterized group of eukaryotic mRNA regulatory elements and is one of the longest conserved RNA sequences known (n = 28–30) (16,17). The secondary structure is a hairpin loop with an internal loop or bulge. A folded IRE structure has been indicated by comparing wild-type ferritin mRNA to RNA with altered IRE sequences, using classical protein nucleases, an alkylating reagent and shape selective transition metal complexes as probes of structure (18–20); interactions as far away as 20 nucleotides upstream from the IRE have been observed (20). The results have been related to mRNA function measured as ferritin synthesis *in vitro* (20–21).

Among members of the IRE family, the ferritin IRE is a particularly attractive sequence to examine because it is distinct among IREs in encoding both positive and negative translational control (enhancement or diminution of protein synthesis, depending on the proteins bound) (21,23–25), because the ferritin IRE sequence allows quantitatively greater negative regulatory effects than other IREs (22), and because extensive studies of ferritin mRNA, both *in vitro* and *in vivo*, provide a useful base of information about function (16,17,22–29). We now report basic ^{1}H- and ^{31}P-NMR experiments with wild-type ferritin IRE and, by comparison to an IRE with a hairpin loop mutation, show the importance of G18 in hairpin hexaloop for stabilizing the IRE structure. Substitution of G18 with A18, previously shown to prevent recognition by the IRP protein (18,21), greatly decreased the temperature stability, and changed the backbone conformation, suggesting a structural basis for the loss of function in the mutant.

* To whom correspondence should be addressed

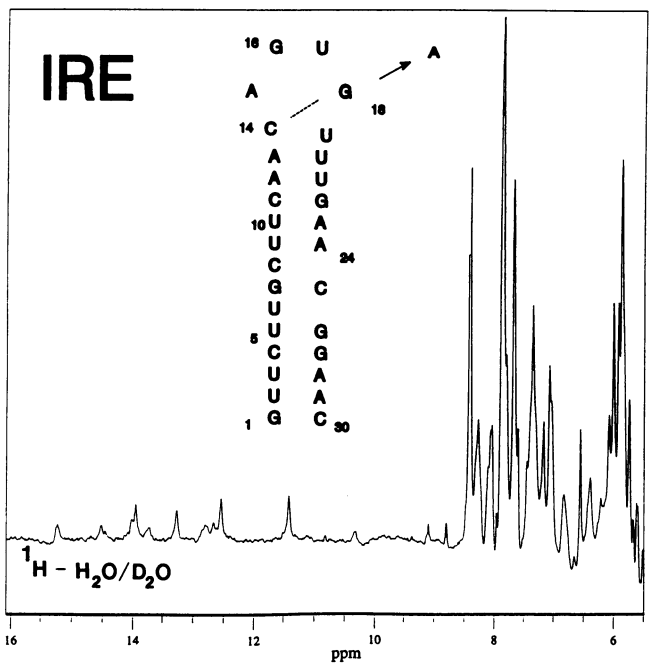

Figure 1. The ferritin IRE predicted secondary structure and ^{1}H NMR spectrum. The IRE predicted secondary structure of the wild-type and G18A mutant IREs superimposed on the exchangeable and part of nonexchangeable proton NMR spectrum of 30mer IRE wild-type sequence; only the H6/H2/H5/H1′ region of the nonexchangeable proton spectrum is shown.

MATERIALS AND METHODS

RNA design

NMR spectra were collected for two oligoribonucleotides (n = 30) one of which had the wild type ferritin IRE sequence and the other of which had an identical sequence except for a change in the terminal hexaloop from CAGUGU to CAGUAU; sequences of similar size have been shown to form an IRP-protein binding site (16–18,29). ^{1}H spectra were also collected for a 16mer with the wild-type sequence/hairpin loop and upper stem, but the peaks were broad. G18A was selected for the base substitution for five reasons: (1) the site had aberrant chemical/enzymatic reactivity even in full-length mRNA (18,19); (2) the mutation eliminated negative but not positive translational control (21); (3) the mutation occurred spontaneously in nature in a pseudogene (a genomic, full-length, ferritin cDNA) (27); (4) the site is crosslinked to the IRP by UV^{30}; and (5) the predicted secondary structure of the IRE (Figure 1) was not changed by the G18A substitution. Note that G18 in the 30mer, corresponds to G16 in the generic/consensus IRE numbering scheme, since some members of the IRE family have only 28 conserved nucleotides (16,17).

RNA synthesis

RNA was synthesized, as previously described (18), using a DNA template which contained the complement of the desired sequence next to the T_7 RNA polymerase promoter (n = 17); 150 ml reaction mixtures were used with T_7 RNA polymerase and 4 mM nucleotide triphosphates as substrates. T_7 RNA polymerase was purified from *E.coli* strains, kindly provided by W. Studier, containing the cloned gene under the control of the lac operator; expression and purification of the RNA polymerase followed the procedure described in reference 31. RNA transcribed *in vitro* was purified, after alcohol precipitation by electrophoresis in urea–acrylamide gels, followed by gel filtration (Sephadex G-25) and extensive dialysis against water. Sequence of the RNA was confirmed by analyzing digests of ^{32}P-end labelled oligomer with RNAase U_2, A, Phy M and T_1 under denaturing conditions. The reactivity of the wild-type 30mer, probed with V_1 and S_1 nucleases and 1,10-phenanthroline-Cu under native conditions, was comparable to the reactivity of the IRE in full length ferritin mRNAs.

Two types of IRE 30mers were prepared. These sequences were a wild-type and a hairpin loop mutant (A substituted for G), which corresponded to the hairpin loop, internal loop and upper and lower stem of the IRE (Figure 1). NMR (1.5–2 mg) quantities of the pure sample of the 30mer IRE sequence of ferritin mRNA, were dialyzed against water, lyophilized and dissolved in 10 mM phosphate buffer pH 6.9, 0.1 mM EDTA 90% H_2O/10% D_2O, the final concentration of mRNA was 0.3 mM. For further NMR studies, samples were concentrated by vacuum and exchanged twice with D_2O. NMR sample homogeneity is illustrated in Figure 1. The IRE is 97% conserved among ferritin mRNAs; the bullfrog red cell sequence was selected because previous studies had shown that the translational regulation observed *in vivo* (26) could be reproduced *in vitro* with both natural polyA$^+$ RNA (24,25,28) and capped *in vitro* transcripts (20,21,28). In addition, a synthetic oligoribonucleotide with the IRE sequence of the bullfrog red cell H-subunit ferritin has been shown to bind the regulator protein (IRP) similarly to natural ferritin mRNA (18) and to have a similar structure as natural ferritin mRNA as probed with protein nucleases and transition metal complexes (18,19).

Absorbance (thermal denaturation) analysis and circular dichroism measurements (CD)

Thermal denaturation profiles were obtained using a Gilford Response System spectrometer controlled by the associated Response II Thermal Programming software. Profiles were measured in duplicate at 260 nm on three RNA samples for the wild-type IRE and on two samples for the mutant G18A; the concentration of mRNA was 10 μM. Melting curves of IRE 30mers with the wild-type and G18A mutant sequences were obtained over a 100× concentration range in 0.1 M potassium phosphate, pH = 6.9, and were corrected for background absorbance due to the buffer components. All transitions were reversible as indicated by hysteresis values of 1% obtained at the end of the cooling cycle to that prior to heating.

CD spectra of 10 μM solutions of IRE 30mers with the wild-type and G18A mutant sequences were obtained in 0.1 M potassium phosphate, pH = 6.9, with no Mg^{+2}. Spectra were collected on a JASCO J600 spectropolarimeter interfaced with an IBM PC computer. Samples were contained in a 1-cm path length cylindrical cell surrounded by an external jacket of water for temperature control. The CD spectra for two samples of wild-type and G18A mutant were monitored over a range of 200 nm to 320 nm and over temperature range of 10–80°C.

NMR methods

NMR spectra were obtained using a GE 500 MHz Omega Spectrometer interfaced with SUN 3/160 computer and GE 6.0

operating software. The temperature was controlled by a SUN 3/160 computer interfaced to the spectrometer and the variable-temperature control unit. Constant-flow cooling was provided by a Microsoft thermostat unit, and the temperature was monitored via a thermocouple implanted in the probe. Imino proton spectra were obtained either by presaturation of the HDO signal for 2s prior to applying the 6 μs observation pulse or by using the '1-1 hard pulse' solvent suppression method (32). One-dimensional ^1H spectra were collected in 16K data sets consisting of 1024 scans each. One-dimensional nuclear Overhauser effects (NOEs) (1D NOE) were measured using the decoupler for preirradiation and '1-1 hard pulse' sequence for HDO suppression. All of the NOE data were acquired at 10°C. The data were obtained in the interleaved on and off resonance preirradiation mode. Preirradiation times of 100, 200 and 600ms were used to monitor for possible spin-diffusion effects on the NOEs (and to achieve a steady state). The extent of decoupler power spillover was determined by preirradiating at variable offsets from the peak. Typically, 2048 scans were collected in 10%/90% D_2O/H_2O, and the difference free induction decays were processed using 4-Hz exponential line broadening. 2D NOESY data were also acquired at 10°C, using a jump-return echo (JRE) pulse sequence for water suppression with excitation maximum set to the imino resonances (33). Data were acquired on a Bruker 500 MHz spectrometer kindly made available by S. Brown at Glaxo, Inc. The spectrum contained 2048×256 complex data points with a sweep width of 12000 Hz, a mixing time of 250 ms, a recycle delay of 1.7 s, and 256 scans per slice. Spectra were processed with FELIX software (Biosym, Inc.) using exponential weighting function or shifted sine-bell function to resolve overlapped imino protons.

NOESY (34), DQF-COSY (35) and TOCSY (36) experiments were recorded in 99.996% D_2O on a 500 MHz GE Omega spectrometer. Data sets with 1024 complex points in t_2 and 512 complex points in t_1 were acquired with 5000 Hz sweep widths in both dimensions and 128 scans per slice. Two NOESY spectra were acquired with mixing times of 120 ms and 200 ms and a recycle delay of 1.2 s. The TOCSY spectrum was recorded with a 75-ms MILEV spin lock pulse and a recycle delay of 1.5 s. The DQF-COSY was recorded with WALTZ decoupling of ^{31}P during acquisition and recycle of 1.6 s. The diagonal and cross peaks of DQFCOSY spectra were phased with antiphase absorption line shape in both dimensions. All spectra were processed with combination of exponential and sine-skewed function and zero-filled to 2K×2K data points using GE Omega 6.0 software. ^{31}P spectra were acquired at 10°C with 5-mm inverse detection probe and standard parameters at 202.45 MHz frequency. 85% H_3PO_4 was used as the reference sample.

RESULTS

Verification of stem–loop structure

Information on higher order structure of the hairpin (HL) and internal loops (IL) of the IRE can only be obtained currently by NMR spectroscopy or chemical probing, since computer predictions of secondary structure show the HL and IL as single stranded regions. Reactivity of the HL and IL with nucleases and chemical probes indicated more complex structures (18–20).

In order to evaluate the predicted secondary structure of IRE and to explore the structure in solution, in this study absorbance thermal denaturation, circular dichroism and NMR spectroscopy

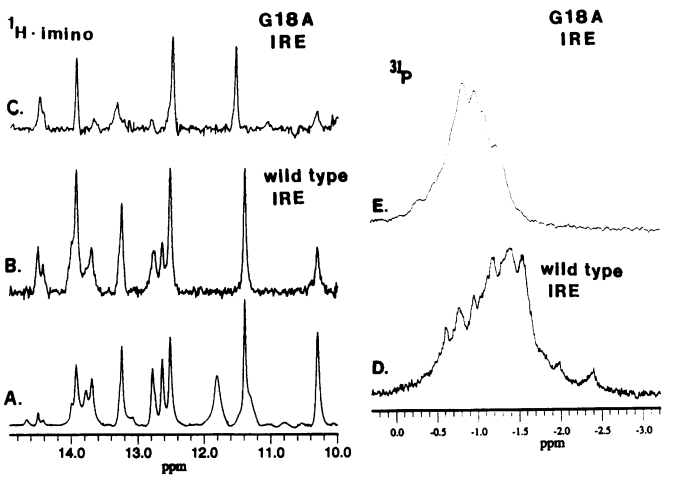

Figure 2. Stem/base pair structure in wild-type and mutant ferritin IREs analyzed by ^1H imino and ^{31}P NMR spectroscopy. RNA samples were dissolved in 10 mM sodium phosphate buffer, pH 6.9, 0.1 mM EDTA in H_2O/D_2O(90%/10%) and measured at 10°C (see Materials and methods). **(A)** Spectrum obtained with an 11-spin echo water suppression pulse sequence ('1-1 hard pulse'); **(B,C)** Spectrum obtained with presaturation of the water signal (pcp); (A,B) Wild-type sequence; **(C)** Mutant sequence with a G18A substitution; **(D)** ^{31}P spectrum of wild-type IRE; **(E)** ^{31}P spectrum of G18A mutant IRE.

have been used. Purity and conformational homogeneity of the IRE (n = 30) are indicated by the spectra in Figures 1 and 2.

It is necessary to show that the IRE samples are monomers under the conditions of the experiment, since all RNA stem–loops have the potential to form palindromic dimers, where the nucleotide corresponding to the loop in a stem–loop structure could form an internal loop in the dimer structure. A stem–loop of the IRE, e.g. is extremely difficult to distinguish from a symmetric dimer on basis of NMR data alone, since almost identical sets of interproton distances would be present in either structure. ^1H-NMR and T_m experiments at a variety of concentrations were examined to assess the presence of IRE dimers at high concentration. The similarity of the NMR spectra, over the concentration range 6 μM to 800 μM, and the linear relationship observed between $1/T_m$ vs log(RNA), over the concentration range 0.5 μM to 5 μM, indicate that the IRE was ≥90% monomeric over the range studied. The ^{31}P spectrum for the wild-type sequence also showed a pattern typical for a hairpin-loop (Figure 2). In addition, the ^{31}P spectrum of the G18A mutant IRE displayed features of a stem loop (Figure 2). However, the more narrow dispersion of the spectrum for the G18A mutant compared to the wild type indicates some changes in the backbone caused by the HL mutation. CD spectra of both the wild-type and G18A mutant IREs were typical of RNA stem–loops with maxima at 220 and 270 nm.

Studies of the IRE by Fe EDTA cleavage indicated differential accessibility of regions of the RNA to solvent (18). In order to evaluate relative access to solvent of protons along the IRE wild-type sequence by NMR spectroscopy, the 1-D ^1H spectra of exchangeable imino protons (Figure 2) were compared using either a 1-1 ('1-1 hard pulse') spin echo water suppression pulse sequence (33) (the excitation maximum was placed on the imino proton resonances) or presaturation of the water signal (pcp) for 3 s prior to applying the 7.4 μs observation pulse. Of the sharp peaks in the

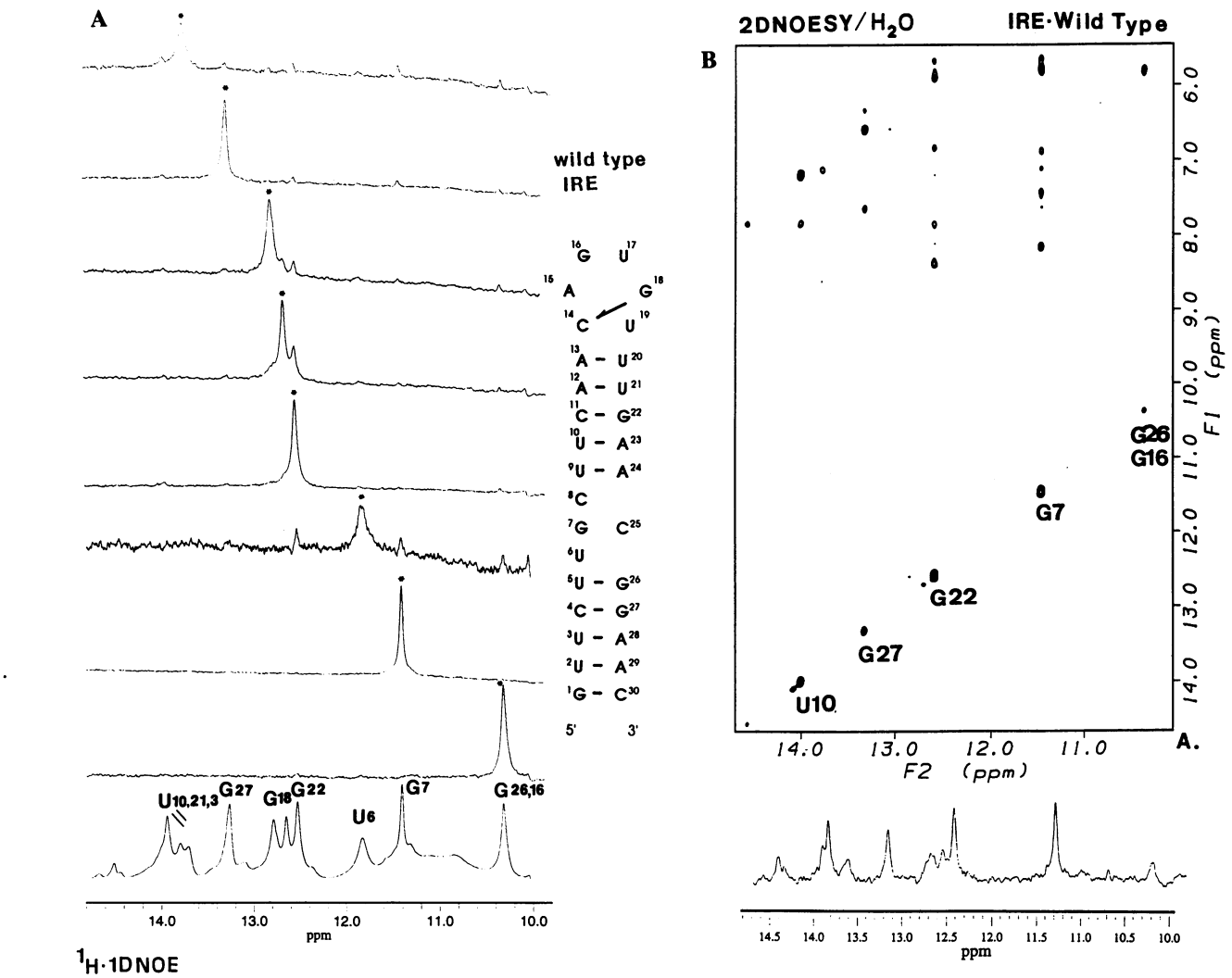

Figure 3. ^{1}H interactions in the ferritin IRE analyzed by 1D NOE and 2D NOESY. (**A**) 1D NOE spectra. The signal to noise ratio allowed detection in at least four independent measurements for each 1D NOE set. Spectra displayed correspond to peaks detected in 2D NOESY (Figure 3B) and isolated peaks; decoupler spillover assessment is described in Materials and Methods; (**B**) 2D NOESY of the wild-type IRE in 10 mM sodium phosphate buffer, pH 6.9, 0.1 mM EDTA, in H_2O/D_2O (90%/10%). Contour plot of the imino proton-aromatic/amino proton region obtained from 2D NOESY (jump and return) experiment.

exchangeable proton region, one was eliminated and several were broadened, when data were collected with presaturation of H_2O (Figure 2); peak broadening or disappearance of peaks is attributable to exchange with solvent. The differential changes in imino protons in the IRE spectra observed with presaturation of the water peak indicates that the hydrogens corresponding to the broadest or absent resonances were relatively more accessible to solvent than those which were minimally affected in presaturation experiment; the position of the imino resonances also confirms that the IRE sequences have the base paired stems predicted. Differences in solvent accessibility of imino IRE protons in the region of base pairs, 10–15 p.p.m., indicate a folded structure and confirm the inside/outside IRE structure detected as differential sensitivity to cleavage by Fe-EDTA (18).

Imino and amino proton assignment

The downfield exchangeable proton spectrum of the wild-type IRE, shown in Figure 2, (A) '1–1 hard pulse' and (B) (pcp) presaturation of HDO, indicates base pairs which appear to be typical for the predicted secondary structure (Figure 1). Signals of imino protons characteristic for G–C base pairs are typically in the region 12–13.5 p.p.m.. The signals of imino protons downfield of 13.5 p.p.m. are characteristic of A–U base pairs. Resonances upfield of 11.5 p.p.m. are most likely due to non-hydrogen bonded imino protons from uridine and guanosine residues; imino resonances below 11 p.p.m., corresponding to non-hydrogen bonded imino protons, have been observed in loop structures of DNA (37) and RNA (38) or exposed imino protons from guanosine residue in syn conformation (39). Some of the resonances present in the 30mer spectrum were also in the ^{1}H-NMR spectrum of a 16mer (hairpin loop and upper stem), but the peaks were broader, indicating the importance of the internal loop/lower stem in stabilizing the IRE structure.

Imino protons were assigned using data from 1D NOE and 2D NOESY in H_2O. Figure 3 shows the exchangeable proton, imino region of the 1D NMR spectrum, the results of 1D-NOE difference experiment, and part of the map of 2D NOESY (jump and return)

in 10%/90% D_2O/H_2O for the wild-type sequence of IRE. There are six A–U base pairs and four G–C base pairs predicted in the secondary structure of the IRE (Figure 1). The strategy to assign exchangeable imino protons in wild-type IRE by NOE connectivities was to start from a known G–C base pair. When the imino regions of the NMR spectrum of three IRE sequences were compared [wild-type 16mer (upper stem and hairpin loop only), wild-type 30mer (upper and lower stem hairpin loop and internal loop) and G18A mutant 30mer], a peak at 12.6 p.p.m. in the G–C base pair region, was present in all three sequences. The 16mer has two potential G–C base pairs, G22–C11 in the stem and G18–C14 in the hairpin loop. The presence of a resonance at 12.6 p.p.m. in the spectrum of both G18A (Figures 2 and 4) and the 16mer, as well as the 30mer indicates that it is likely G22. Irradiation of the resonance at 12.6 p.p.m. in the wild-type 30mer of the IRE produced a medium intensity NOE to neighboring U residues (Figure 3, Table 1) and a strong NOE to a guanine amino proton at 6.8 p.p.m.. The large number of NOEs (Figure 3, Table 1) also suggest that the imino proton with a resonance at 12.6 p.p.m. is on a G–C base pair with a central position in the structure. The only G–C base pair in 30mer which could have a 1D NOE to A–U base pairs and be present in the 16mer is G22, supporting the assignment of the resonance at 12.6 p.p.m. to G22. Detailed analysis of 1D NOE connectivities and 2D NOESY connectivities using the assignment of the 12.6 p.p.m. resonance to G22 allowed assignment of seven imino proton resonances to specific G residues and six predicted AU base pairs.

The signal at 13.35 p.p.m. could, based on previous results (38), belong to either a G–C or an AU base pair. However, inspection of the 2D NOESY (jump-return) in H_2O (Figure 3B) shows the resonance to belong to a G–C base pair since there were strong NOE cross-peaks from cytosine $4NH_2$ to guanine imino protons. In contrast, an AU base pair would have had a strong NOE from the U imino to the H2 proton of the paired adenine and a weaker NOE to the adenine $6NH_2$, which was observed for imino resonances downfield from 13.35 p.p.m.. The signal at 13.35 p.p.m. has been assigned to G27 based on NOE data used as follows. Upon irradiation, G27 had NOE connectivities to two Gs at 10.4 p.p.m. and 11.5 p.p.m. (Figure 3A, Table 1). The only G which could possibly have two G residues that close is G27 (Figure 1). The signal at 11.5 p.p.m. has been assigned to G7, based on 1D NOE, because this resonance had no strong NOE to neighboring base pairs (Figure 3A). The signal at 12.74 p.p.m. has been assigned to G18 from a G–C base pair in the HL because this signal is undetectable in the spectrum of the IRE G18A mutation (Figures 2 and 4); G18 is also inaccessible to RNAase T_1 (18) and C14 to methylation (19), supporting the conclusion of a G–C base pair in the HL. The signal at 10.4 p.p.m. has been assigned to G26 based on the pattern of 1D NOE upon irradiation of signal 13.35 p.p.m. (G27), which gives a strong NOE to G7 (11.5 p.p.m.) and G22 (12.6 p.p.m.) and to the signal at 10.4 p.p.m.. The only remaining G–C base pair is at 12.84 p.p.m., which is assigned to G1–C30 (see Table 1).

2D NOESY correlations in H_2O (Figure 3B) showed that for predicted base pair G26–U5 NOEs was not detected among three sets of imino resonances. The signature of a GU base pair is an imino proton crosspeak at least as strong as crosspeaks in the region of non- exchangeable protons (38). In Figure 3B, crosspeaks can be clearly seen for protons in the non-exchangeable region but none for a G26–U5 base pair. The reactivity of G26 to transition metal complexes supports the conclusion of the absence of a typical G–U base pair at the junction of the stem/internal loop: 1,10-phenanthrolene-Cu readily cleaved G26 (19) and Fe-EDTA cleaved G26 more extensively than neighboring stem residues (18), indicating that around G26, the backbone is at the surface of the folded structure. In addition, the imino resonance at 10.4 p.p.m. currently assigned to G26 is broadened in the pcp experiment in 10% D_2O at 10°C (Figure 2) and is significantly broadened between 30 and 40°C in the '1-1 hard pulse' experiment (Figure 4). In contrast, G22 (12.6 p.p.m.) paired to C11 in the center of the upper stem is relatively inaccessible to cleavage by Fe-EDTA (18) and 1,10-phenanthrolene-Cu (19), showed little perturbation after presaturation of H_2O peak in the pcp experiment (Figure 2) and is present in the '1-1 hard pulse' imino spectrum at temperatures above 60°C (Figure 4). Sharpness of the peaks suggest that the internal loop region is highly structured.

Table 1. 1D NOEs and current assignments of imino protons for the wild-type ferritin IRE

Assignment[a]	Chemical shift (p.p.m.)	NOE observed (p.p.m.)[b]
G26,G16	10.4	12.61, 8.06, 11.9, 5.8
G7	11.5	13.35, 8.1, 7.66, 9.1, 8.12, 7.42, 6.87, 5.8, 6.65
U6	11.9	12.6, 11.5, 10.4, 7.86
G22–C11[c]	12.61	14.02, 10.4, 8.3, 7.84, 6.85, 5.87, 5.69
G18–C14	12.74	10.4, 7.81
G1–C30	12.84	11.5, 10.4, 12.6, 8.08, 7.84
G27–C4	13.35	12.6, 11.5, 10.4, 8.3, 7.66, 6.59
AU	13.81	11.5, 7.05
A–U	13.88	7.62
2AU	14.02	12.6, 7.14, 7.62, 7.07
		13.3, 12.8, 11.9, 11.5, 10.4
A–U	14.60	12.5, 7.76, 7.84
A–U	14.68	12.6

[a]Assignment based on current information. 1D NOE data for exchangeable proton region are shown in Figure 3.

[b]The signal to noise ratio allowed detection in at least four independent measurements for each 1D NOE set. When irradiated peaks were overlapping, the extent of decoupler power spillover was determined by preirradiating at variable offsets from peak.

[c]Since the resonance at 12.6 p.p.m. is present in the wild-type 16mer (upper stem), 30mer wild type, and the G18A mutant, it is assigned to G22; only G22 is present in all three sequences and is the only G–C base pair in the 30mer with 1D NOEs to 2 AU base pairs.

Nonexchangeable protons

Partial analysis of the region of nonexchangeable protons in the ^1H NMR spectrum supported the assignments in Table 1. For example, sequential base to H1′ NOEs typical of A-form RNA were observed for the lower and upper stems, with an H8–H8

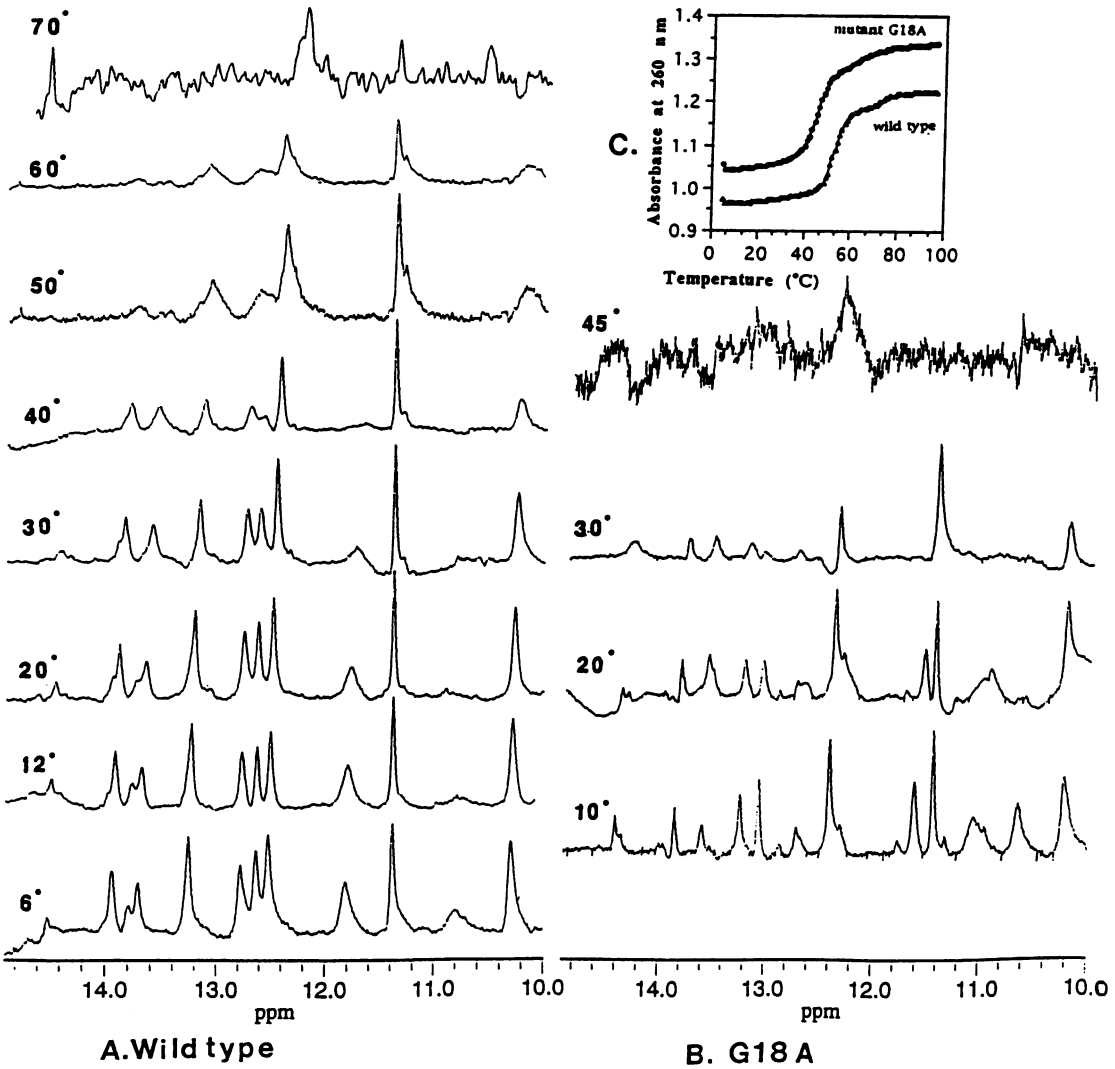

Figure 4. The influence of a mutation in the hairpin loop of the ferritin IRE on the temperature stability of the RNA (n = 30). Compared are the imino regions of the NMR spectra of the wild-type and G18A-IRE sequences at 10, 20, 30 and 45°C; spectra measured up to 70°C are included for the wild-type IRE. Melting curves were measured three to four times, using three independently prepared samples (wild type) and one sample of the mutant. Samples for measurement of melting behavior were in the same buffer used for the NMR spectra (10 mM sodium phosphate, pH 6.9, 0.1 mM EDTA). The melting curves were the same over a 100× concentration range. (**A**) Wild-type IRE; (**B**) G18A Hairpin loop mutant IRE; (**C**) Melting curves. Note for the G18A mutant a 10°C decrease in the T_m (from 52° to 42°C), and the absence of a second transition at high temperatures (ca. 73°C), when A is substituted for G in the hairpin loop at position 18.

NOE placing the position of an A stacked above G_{27} and an A stacked above G_{22}. However, analysis of the non-exchangeable proton of the ^{1}H-NMR spectrum provided little additional information about assignments in the internal loop and lower stem because of overlapping NOEs. Similar ambiguities and broad natural line widths occurred for resonances in the hairpin hexaloop. Complete assignment will require analysis, in the future, with samples enriched in ^{15}N and ^{13}C.

Effect of mutation on temperature stability of IRE and the role of hairpin loop

^{1}H-NMR spectra can be studied to assess conformational stability of the hairpin loop, even before the complete structure can be determined. The dramatic effect on the hairpin loop stability of the IRE was shown by measuring the ^{1}H spectrum in water at temperatures from 10 to 75°C (Figure 4). One heat stable substructure persisted up to 70°C in the wild-type IRE but was absent in the G18A mutant IRE (Figure 4). The presence of a heat stable subdomain in the wild-type IRE was also indicated by a secondary hyperchromic transition (T_m=73°; Figure 4C), although the major hyperchromic transition was at 52°C. The most heat stable signals are currently assigned to G22 in the upper stem (12.6 p.p.m.) and to G7 in the internal loop (11.5 p.p.m.).

In contrast to the wild-type IRE, the NMR spectrum of the G18A mutant IRE collapsed at 45°C, although the resonances at 12.61 p.p.m. and 11.5 p.p.m. were present at lower temperatures. The major hyperchromic transition of the G18A mutant IRE was monophasic, in contrast to the wild-type IRE, as well as being 10°C lower (t_m=42°C; Figure 4C). Note that at the temperatures typically used to measure protein synthesis and IRP binding (18,20,21), the structure of the G18A mutant IRE is altered based on the ^{1}H NMR spectrum and the T_m (Figure 4).

Changes in the IRE structure caused by the HL mutation (G18A) that were indicated by ^{1}H NMR and UV-vis spectrometry were extended by comparing the ^{31}P spectrum of the wild-type and mutant IRE (Figure 2). The ^{31}P spectrum of the G18A mutant is typical of an RNA hairpin loop. The larger dispersion of the ^{31}P spectrum of the wild-type IRE suggests folding of the stem backbone in the hairpin loop structure.

DISCUSSION

Comparison of the ^{1}H (imino region) NMR spectra, the ^{31}P NMR spectra, and the temperature stability of the two IRE oligoribonucleotides, wild-type and the G18A mutation, (Figures 1–4) showed the importance of G18 in the hairpin hexaloop for stabilizing the native IRE structure. The loss of IRP binding and negative regulation in ferritin mRNA with G18A mutation (18,20,21) emphasizes the functional importance of the structure of the hairpin loop.

Folding of the IRE secondary structure, suggested by the ^{1}H-imino spectrum (Figures 2– 4), ^{31}P-NMR spectrum (Figure 2), and previously by differential sensitivity along the structure to Fe-EDTA cleavage (18), was associated with high temperature stability (Figure 4A), and subdomains with different melting behavior (Figure 4B, C). For example, the ^{1}H imino resonance at 11.5 p.p.m. and 12.6 p.p.m., involving the internal loop G7 and the upper stem G22 were observable at 70°C (Figure 4) and appear to be related to the minor perturbation in the plot of A_{260nm} vs T at 73°C; the major hyperchromic transition (T_m) occurred at 52°C (Figure 4). Such data suggest that distinctive structure in the IL and upper stem of the wild-type IRE structure is stabilized by interactions which are relatively independent of the majority of stacked base pairs. However, the heat stable subdomain appears to depend in part on interactions involving the hairpin loop, since the heat stable environment was lost by mutation in the HL. Thus, the resonances, 12.6 p.p.m. (G22) and 11.5 p.p.m. (G7), which could be detected at 70°C in the wild-type IRE disappeared at ≤45°C in the G18A IRE (Figure 4B).

A G–C base pair across the hairpin hexaloop of the native IRE between G18 and C14 is indicated by the absence of RNAse T_1 cutting at G18 (in contrast to G16) (18), the absence of methylation of C14 (in contrast to A15) (19), as well as absence of an imino resonance in the region typical of G paired to C, in the ^{1}H-NMR spectrum of the IRE with G18A (Figures 2 and 4). Altered structure in the hairpin loop of IRE-G18A is also indicated by methylation of A18 (19) and cleavage of G16, U17, and A18 by 1,10-phenanthroline-Cu (19). Base pairs or stacked bases in RNA hairpin loops are known in the hepta- and octa-hairpin loops of t-RNA (10), the tetra-hairpin loops in r-RNA (12–15), and model RNA hairpin loops (40,41). In fact the decrease of the IRE T_m from 52 to 42°C with G18A mutation is comparable to changes in T_m for model hairpin loops with a G–C pair across the loop when I is substituted for G. However, the sequence dependance of hairpin loop structure (40–44), the high phylogenetic conservation of IREs (16,17), and the dependence of the heat stable subdomain on the HL sequence (Figure 4), suggests that the ferritin IRE structure involves the interactions of the CAGUGU/C sequence in the HL with the stem and internal loop and may be distinctive among known hairpin loops. The importance of G at position 18 in the IRE is emphasized by a recently published study of randomly mutated RNA stem/hexaloops which bind the IRP (45). Among the 26 binding RNAs selected with single base substitutions in the HL, all had G at position 18 (45); only when two or more loop substitutions occurred was G18A observed in the binding pool of randomly mutated IREs. G18A substitutions which bound the IRP only occurred in double substitutions where C14U also occurred, which allows prediction of a potential A–U pair across the hairpin loop. However, only RNAs with the native C14–G18 HL base pair could bind both known IRPs, whereas the U14–A18 RNAs could only bind IRP-1 (45). Thus, a G–C base pair in the hairpin loop of the IRE seems to be associated with the widest range of interactions with IRPs and appears to have a specific effect on structure. Another deviation of the ferritin IRE from the predicted secondary structure (Figure 1), in addition to a G–C base pair in the HL, is the apparent absence of a base pair between G26–U5 at the junction of the lower stem internal loop. Absence of the characteristic G–U crosspeak in the 2D NOESY in H_2O (Figure 3) and strong cleavage of G26 by 1,10-phenanthrolene-Cu (18) and Fe-EDTA (19) all indicate an open or distorted secondary structure at the IL/stem junction.

Distinctive structural and functional properties of the folded IRE hairpin which depend on G18 in the hairpin loop, are manifested by the temperature stability of imino protons in IRE (Figure 4), long-range interactions indicated by the NOEs connectivities (Table 1, Figure 3), the large fraction of imino protons which are relatively inaccessible to solvent (Figure 2), binding of the IRP (18), and negative translational control (20,21). The decreased temperature stability of the mutant structure likely explains the absence of IRP binding and negative translational control (20,21), since much of the structure appears to be disrupted (Figure 4) in the mutant at temperatures where protein synthesis and IRP binding have been measured. A working model which fits the data from both NMR spectroscopy (Figures 1–4) and probing with transition metal complexes (18–20,46) is a folded hairpin stabilized by interactions between residues in the hairpin hexaloop and residues in the region of the four nucleotide internal loop and the stem (Figure 5). The TAR element in HIV RNA, which controls transcription, is also a hairpin hexaloop with an internal loop (7–9). However, the interaction of TAR with Tat, the transcription regulatory protein, excludes the hairpin hexaloop and depends more on the internal loop and stem (7–9). In contrast to TAT, the specificity and the stability of the folded ferritin IRE structure, required for the recognition between the mRNA and the IRP regulator protein, involves the entire IRE structure (hairpin hexaloop, internal loop, and stems) (18) and depends, in part at least, on the G in the fifth position of the hairpin hexaloop.

REFERENCES

1 Leclerc, F., Cedergren, R. and Ellington, A.D. (1994) *Nature Struc. Biol.* **1**, 293–300.
2 Sklenar, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1994) *J. Biomol. NMR* **4**, 117–122.
3 Nikonowicz, E.P. and Pardi, A. (1993) *J. Mol. Biol.* **232**, 1141–1156; Farmer II, B.T., Muller, L., Nikonowicz, E. and Pardi, A. (1994) *J. Biomol. NMR* **4**, 129–133.
4 Wijmenga, S.S., Heus, H.A., Werten, B., van Der Marel, G.A., Van Boom, J.H. and Hilbers, C.W. (1994) *J. Mag. Res. Series B* **103**, 134–141.
5 Murphy, F.L. and Cech, T.R. (1993) *Biochemistry* **32**, 5291.
6 Murphy, F.L. and Cech, T.R. (1994) *J. Mol. Biol.* **236**, 49–63.
7 Puglisi, J.D., Tan, R., Colman, B.J., Frankel, A.D. and Williamson, J.R. (1992) *Science* **257**, 76.
8 Jaeger, J.A. and Tinoco Jr., I. (1993) *Biochemistry* **32**, 12522; Michnicka, M.J., Harper, W.J. and King, G.C. (1993) *Biochemistry* **32**, 395–400.

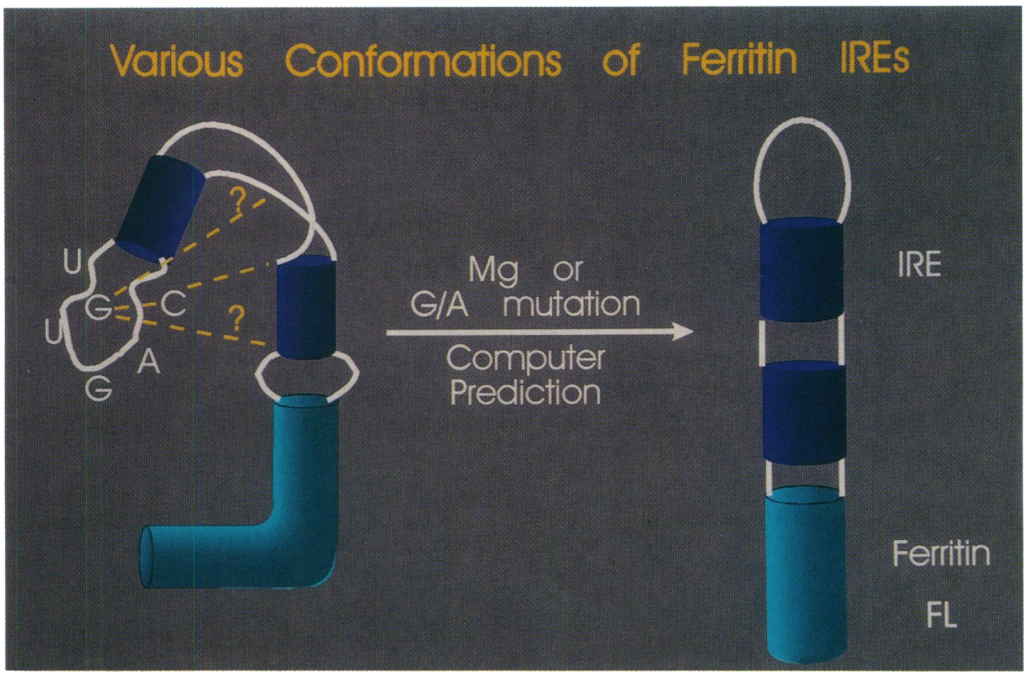

Figure 5. Cartoon of the folded ferritin IRE+FL (flanking region) structure. Base paired helical regions are shown as cylinders, blue for the IRE and turquoise for the ferritin-specific base-paired flanking region (FL). Loops are shown as white lines. The internal loop in the IRE is extended to reflect the accessibility of residues flanking the internal loop U9 to RNAase V_1 (18), A23, A24 to dimethyl sulfate (19), G26, G27 to 1,10-phenanthroline-Cu (19) and the hypersensitivity of U6, G7, C25, G26, and G27 to Fe-EDTA (18) (see Figure 1 for numbering scheme). The bend in the FL reflects the results of probing with RNAase V1 and Fe-EDTA (18) as well as the effect of mutations (20). Interactions of the hairpin hexaloop are drawn to reflect inaccessibility of C14 to methylation (19), G18 to RNAase T_1 (18), methylation of A18 in the G18A IRE (19), and the effect of the G18A substitution on the ^{1}H-NMR spectrum (Figures 1–4). The putative bend in the flanking region (turquoise) is based on hypersensitivity to cleavage by Fe-EDTA, 1,10-phenanthroline-Cu [and $Rh(phen)_2Phi^{3+}$], as well as increased accessibility to RNAase V_1 when the IRP covers the IRE. Computer prediction shows no bend, the hyperreactivity to 1,10-phenanthroline-Cu is absent if 0.5 or 5.0 mM Mg are present during folding or in the G/A mutation.

9 Puglisi, J.D., Chen, L., Frankel, A.D. and Williamson, J.R. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3680–3684; Colvin, R.A., White, S.W., Garcia-Blanco, M.A. and Hoffman, D.W. (1993) *Biochemistry* **32**, 1105–1112; Williamson, J.R. (1994) *Nature Struc. Biol.* **1**, 270–272.
10 Holbrook, S.R. Sussman, J.L., Warrant, R.W. and Kim, S.H. (1978) *J. Mol. Biol.* **123**, 631.
11 Westhof, E., Dumas, P. and Moras, D. (1985) *J. Mol. Biol.* **184**, 119.
12 Szewczak, A.A., Moore, P.B., Chan, Y.-L. and Wool, I.G. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 9581.
13 Legault, P. and Pardi, A. (1994) *J. Mag. Res. Series B* **103**, 82–86.
14 Moore, P.B. (1993) *Curr. Opin. in Struc. Biol.* **3**, 340–344.
15 Variani, G. and Tinoco Jr., I. (1991) *Quart. Rev. of Biophysics* **24**, 439.
16 Theil, E.C. (199) J. Biol. Chem. 265, 4771; E.C. Theil (1993) *Biofactors* **4**, 87.
17 Klausner, R.D., Roault, T.A. and Harford, J.B. (1993) *Cell* **72**, 19. Haile, D.J., Roault, T.A., Harford, J.B., Kennedy, M.C., Blondin, G.A., Beinert, H. and Klausner. R.D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11735.
18 Harrell, C.M., McKenzie, A.R., Patino, M.M., Walden, W.E. and Theil, E.C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4166.
19 Wang, Y.H., Sczekan, S.R. and Theil, E.C. (1990) *Nucl. Acids Res.* **18**, 4463; Wang, Y.H., Lin, P.N., Sczekan, S.R., McKenzie, R.A. and Theil, E.C. (1990) *Biol. of Metals* **4**, 56.
20 Dix, D.J., Lin, P.N., McKenzie, A.R., Walden, W.E. and Theil, E.C. (1993) *J. Mol. Biol.* **231**, 230.
21 Dix, D.J., Lin, P.N., Kimata, Y. and Theil, E.C. (1992) *Biochemistry* **31**, 2818.
22 Melefors, O., Goosen, B., Johanssen, H.E., Strepecke, R., Gray, N.K. and Hentze, M.W. (1993) *J. Biol. Chem.* **268**, 5974.
23 Theil, E.C. (1987) In Ilan, J. (ed.), Translational Regulation of Gene Expression. Plenum Press, New York, pp. 141–163.
24 Shull, G.E. and Theil, E.C. (1982) *J. Biol. Chem.* **257**, 14187.
25 Shull, G.E. and Theil, E.C. (1983) *J. Biol. Chem.* **258**, 7921.
26 Schaefer, F.V. and Theil, E.C. (1981) *J. Biol. Chem.* **256**, 1711.
27 Dickey, L.F., Shull, S., Didsbury, J.R., Wang, Y.H. and Theil, E.C. (1987) *J. Biol. Chem.* **262**, 7901.
28 Dickey, L.F., Wang, Y.-H., Shull, G.E., Wortman III, I.A. and Theil, E.C. (1988) *J. Biol. Chem.* **263**, 3071–3074.
29 Leibold, E.A., Landano, A. and Yu, Y. (1990) *Nucl. Acids Res.* **18**, 1819–1824.
30 Basilion, J.P., Roault, T.A., Massinople, C.M., Klausner, R.D. and Burgess, W.H. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 574.
31 Grodberg, J. and Dunn, J.T. (1988) *J. Bact.* **170**, 1245–1253.
32 Hore, P.J. (1983) *J. Mag. Res.* **54**, 539–542.
33 Sklenar, V. and Bax, A. (1987) *J. Mag. Res.* **74**, 469–479.
34 Nagayama, K., Kumar, A., Wuthrich, K. and Ernst, R.R. (1980) J. Magn. Reson. 40, 321; Macura, S. and Ernst, R.R. (1979) *J. Magn. Reson.* **46**, 269–282.
35 Kumar, A., Ernst, R.R. and Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun.* **103**, 3654.
36 Davis, D.G. and Bax, A. (1985) *J. Amer. Chem. Soc.* **107**, 2820–2821.
37 Haasnoot, C.A.G., den Hartog, J.H.J., de Rooij, J.F.M., van Boom, J.H. and Altona, C. (1980) *Nucl. Acids Res.* **8**, 169–181.
38 Hoffman, D.W., Colvin, R.A., Garcia-Blanco, M.A. and White, S.W. (1993) *Biochemistry* **32**, 1096–1104.
39 Carbonnaux, C. van der Marel, G.A., van Boom, J.H., Guschlbauer, W. and Fazakerley, G.V. (1991) *Biochemistry* **30**, 5449–5458.
40 Heus, H.A. and Pardi, A. (1991) *Science* **253**, 191.
41 Cheong, C., Variani, G. and Tinoco, I. (1990) *Nature* **346**, 680.
42 Santa-Lucia Jr., J., Kierzek, R. and Turner, D.H. (1992) *Science* **256**, 217.
43 Serra, M.J., Lyttle, M.H., Axenson, T.J., Schadt, C.A. and Turner, D.H. (1993) *Nucl. Acids Res.* **21**, 3845.
44 Tuerk, C., Gauss, P., Thermes, C., Groebe, D.R., Gayle, M., Guild, N., Stormo, G., D'Aubenton-Carafa, Y., Uhlenbeck, O.C., Tinoco Jr., I., Brody, E.N. and Gold, L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1364.
45 Henderson, B.R., Menotti, E., Bonnard, C. and Kuhn, L.C. (1994) *J. Biol. Chem.* **269**, 17481–17487.
46 Theil, E.C. (1994) *New Journal of Chem.* **18**, 435–411.