Cleavage of Functionally Relevant Sites in Ferritin mRNA by Oxidizing Metal Complexes

H. Holden Thorp,*,† R. Ann McKenzie,‡ Peng-Nian Lin,§ William E. Walden,*,|| and Elizabeth C. Theil*,‡

Department of Chemistry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3290, Department of Biochemistry, North Carolina State University, Raleigh, North Carolina 27695-7622, and Department of Microbiology and Immunology, University of Illinois, Chicago, Illinois 60612

Received August 23, 1995[⊗]

IREs, a family of mRNA regulatory structures, control mRNA function and the synthesis of ferritin and the transferrin receptor. IRE secondary structure is a hairpin; in ferritin mRNA, the IRE also interacts with a base-paired flanking region (FL). Using $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$ as cleavage reagents (phen = 1,10-phenanthroline, phi = 9,10-phenanthrenequinone diimine, tpy = 2,2′,2″-terpyridine, bpy = 2,2′-bipyridine), two different substructures in the IRE or FL were detected that are related to regulation. $Rh(phen)_2phi^{3+}$, which intercalates in RNA or DNA, exhibited a single major cleavage site in the FL. Mutation FL2 altered negative IRE regulation and eliminated the specific $Rh(phen)_2phi^{3+}$ cleavage site; the FL2 derivative, FL2R, restored regulation and the $Rh(phen)_2phi^{3+}$ site. $Ru(tpy)(bpy)O^{2+}$, which interacts with nucleic acids electrostatically and cleaves on the basis of solvent accessibility and chemical reactivity, cleaved a single site, G14, in the IRE hairpin loop. G14 appeared to have a simple structure with classical probes but, on the basis of $Ru(tpy)(bpy)O^{2+}$ cleavage, appears to be in a region of high accessibility, possibly caused by distortion in the phosphate backbone. Mutation IL-2 created a heptaloop and destroyed the $Ru(tpy)(bpy)O^{2+}$ site; positive and negative regulations were also affected in IL-2. IRE mutation at other sites did not affect the $Ru(tpy)(bpy)O^{2+}$/RNA interaction. In addition to identifying substructure in the IRE+FL, $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$ should also be useful in determining substructure of potential functional significance in other mRNAs as well as ribozymes, rRNAs, viral RNAs, and snRNAs.

Probing nucleic acid structures with transition-metal complexes has been pursued with two major goals.[1−6] The first goal is understanding the relationship between the chemical reactivity of nucleotides as monomers and the structure of complex oligomers, which is relevant to mutagenesis,[7,8] chemotherapy,[9,10] and nucleotide metabolism[11] and is readily studied using the electrochemical and spectroscopic signatures of metal complexes that cleave DNA.[2,12] The second goal is the analysis of structure in nucleic acids using rigid transition-metal complexes with known recognition and reactivity properties.[1,4] The complex $Rh(phen)_2phi^{3+}$ was recently applied to studying the structure of the tat-1 binding site in the stem-loop TAR RNA of HIV-1 (phen = 1,10-phenanthroline; phi = 9,10-phenanthrenequinone diimine).[13] Binding of the tat-1 protein to a three-nucleotide bulge in the TAR RNA allows for complete transcription of the HIV-1 gene. $Rh(phen)_2phi^{3+}$ specifically recognizes a nearby site, and mutations in the bulge region change the cleavage site. Here we report studies on the stem-loop structure of the iron regulatory element (IRE) and base-paired flanking region (IRE+FL, Figure 1)[14−16] using the photocleavage agent $Rh(phen)_2phi^{3+}$ and the thermal agent $Ru(tpy)(bpy)O^{2+}$ (tpy = 2,2′,2″-terpyridine; bpy = 2,2′-bipyridine). The results show long-range effects of mutation on RNA structure, extending the results with the IRE and Fe-bleomycin.[17] In addition to describing the structural and functional relationships of RNA recognition by $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$, the results indicate differences in chemical reactivity between RNA and DNA. Continued use of primer extension analysis to visualize cleavage sites emphasizes the advantages when end-labeling is inappropriate for long RNA molecules.

The IRE (iron regulatory element) regions of the mRNAs encoding ferritin, the red cell heme biosynthetic enzyme aminolevulinate synthase (eALAS), and the transferrin receptor (TfR) are among the longest (n = 28, 30) conserved nucleotide sequences known in RNA.[14−16] These elements appear to be specific to contemporary animals, although the coding sequences for ferritin are highly conserved among plants, animals, and bacteria. Phylogenetic conservation among individual IREs indicates that the IREs in each type of mRNA comprise members of a family of isoIREs.[14] Several types of proteins

† University of North Carolina.
‡ North Carolina State University.
§ Present address: Department of Microbiology and Immunology, Duke University Medical Center, Durham, NC 27710.
|| University of Illinois.
⊗ Abstract published in *Advance ACS Abstracts,* April 15, 1996.

(1) Theil, E. C. *New J. Chem.* **1994**, *18*, 435−441.
(2) Thorp, H. H. *Adv. Inorg. Chem.* **1995**, *43*, 127−177.
(3) Barton, J. K. In *Bioinorganic Chemistry*; Bertini, I., Gray, H. B., Lippard, S. J., Valentine, J. S., Eds.; University Science Books: Mill Valley, CA, 1994; pp 505−584.
(4) Burrows, C. J.; Rokita, S. E. *Acc. Chem. Res.* **1994**, *27*, 295−301.
(5) Sigman, D. S.; Bruice, T. W.; Mazumder, A.; Sutton, C. L. *Acc. Chem. Res.* **1993**, *26*, 98.
(6) Hecht, S. M. *Bioconjugate Chem.* **1994**, *5*, 513−526.
(7) Sohal, R. S.; Agarwal, A.; Agarwal, S.; Orr, W. C. *J. Biol. Chem.* **1995**, *270*, 15671−15674.
(8) Pavlov, Y. I.; Minnick, D. T.; Izuta, S.; Kunkel, T. A. *Biochemistry* **1994**, *33*, 4695−4701.
(9) Stubbe, J.; Kozarich, J. W. *Chem. Rev.* **1987**, *87*, 1107.
(10) Hecht, S. M. *Acc. Chem. Res.* **1986**, *19*, 83.
(11) Lippard, S. J. *Science* **1993**, *261*, 699−700.
(12) Johnston, D. H.; Welch, T. W.; Thorp, H. H. *Met. Ions Biol. Syst.* **1995**, *33*, 297−324.
(13) Neenhold, H. R.; Rana, T. M. *Biochemistry* **1995**, *34*, 6303−6309.
(14) Theil, E. C. *Biochem. J.* **1994**, *304*, 1−11.
(15) Theil, E. C. *Biofactors* **1993**, *4*, 87−93.
(16) Klausner, R. D.; Rouault, T. A.; Harford, J. B. *Cell* **1993**, *72*, 19−28.
(17) Dix, D. J.; Lin, P.-N.; McKenzie, A. R.; Walden, W. E.; Theil, E. C. *J. Mol. Biol.* **1993**, *231*, 230−240.

S0020-1669(95)01094-9 CCC: $12.00

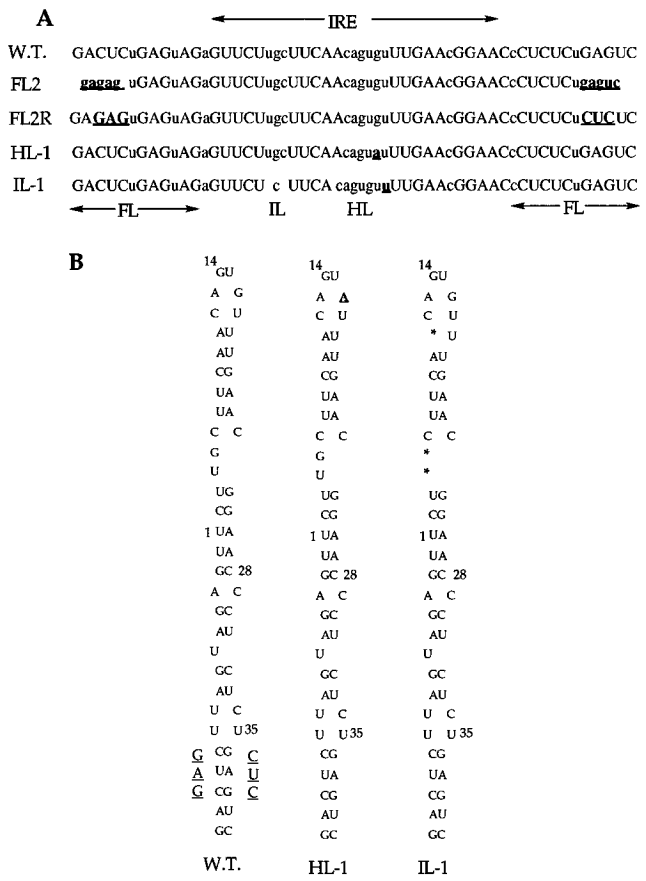

Figure 1. (A) Mutations in FL2, FL2R, HL-1, and IL-1. Paired bases are capitalized, and unpaired bases are shown in lower case letters. (B) Secondary structure predicted for wild-type and mutated ferritin IREs. HL-1 does not bind the IRP; see refs 24 and 25. Mutations are shown as underlined bold letters, and the deletions in IL-1 are shown by asterisks. Sites of the triple mutation in FL2 and the hextuple mutation in FL2R are indicated in the triplet of base pairs (TBP).

are recognized by the IRE regions: a specific regulatory system (the IRP),[18−20] initiation factors, and nucleases. Coordinate regulation of iron storage (ferritin synthesis) and iron uptake (transferrin receptor synthesis) are achieved though repression of initiation factor binding to ferritin mRNA and stabilization of transferrin receptor mRNA when the IRP binds to the IRE or enhancement of initiation factor binding to the ferritin mRNA and degradation of the transferrin receptor mRNA when the IRP is not bound.

IRE−protein interactions are studied by direct analysis of binding and by effects on rates of protein synthesis *in vitro* or on mRNA degradation in cells. Negative mRNA regulation is a decrease in protein synthesis. For ferritin, a decrease in synthesis occurs when the specific IRP protein binds to the IRE, interfering with binding of initiation factors and ribosomes to ferritin mRNA. For the transferrin receptor, a decrease in synthesis occurs when the IRP protein *does not* bind the IRE, allowing nucleases to bind and degrade the mRNA. Positive mRNA regulation of translation is enhanced protein synthesis, above the rate for average mRNAs. When initiation factors/ribosomes are rate-limiting for protein synthesis, such as in cells that grow slowly or *in vitro* where the ratio of total mRNA: initiation factors is selected to be high, mRNAs displaying positive control have a competitive advantage; i.e., more of the encoded protein is synthesized than for average or less competitive mRNAs. Positive control of ferritin mRNA translation was observed both *in vitro* and *in vivo*,[21−23] but only relatively recently was the IRE region identified as the positve control element.[24] Some viral mRNAs such as reovirus exhibit positive control, achieving preferential synthesis of viral proteins over host proteins.

A hairpin loop with an internal loop or bulge is the IRE secondary structure motif (Figure 1). However, the length of conserved primary sequence among members of the IRE family suggests higher order interactions, which were confirmed by analyzing the IRE structure using traditional approaches. For example, there were only two RNAase T_1 sites, one RNAse S_1 site, and two RNAse V_1 sites in the 55-nucleotide IRE region of ferritin mRNA ($n = 900$) that was analyzed.[25−27] In addition, alkylation with dimethyl sulfate modified several A residues in the IRE stem but modified only some of the A,C residues in the IRE loops.[26,27] Determining an IRP "footprint" was difficult with the more traditional probes, since so few sites were accessible to the classical probes even in the unbound mRNA.

Oxidizing transition metal complexes that cleave at or near the site of RNA binding or produce radicals that cleave solvent-accessible regions are an alternative to the large protein nucleases and small alkylating reagents.[1,28,29] Moreover, since the metal ligands in these metallonucleases can be designed to encompass a wide range of shapes, sizes, and electrostatic properties, the potential site specificity is large. The shape specificity of metallonucleases can be extremely specific; for example, an active-site residue in the Tetrahymena type I intron ribozyme can be identified from metallonuclease cleavage.[30] An important advantage of metallonuclease studies is that the quantity (nanograms) of RNA required for analysis by radiolabeling is much smaller than that required for structural studies by NMR (milligrams). The approach therefore provides a useful starting point for the design of detailed structural studies and also for the selection of sites for mutagenesis experiments, as discussed below. Finally, analysis of the cleavage products can be carried out either by direct radiolabeling of the RNA or by visualization using primer extension, which allows for the analysis of much larger RNAs and may ultimately be applicable for *in vivo* footprinting and structural analysis.

Probing the IRE regulatory region of ferritin mRNA with metallonucleases has already yielded a set of useful observations. For example, cleavage of RNA by Fe-EDTA-generated radicals showed that the IRP-binding site was within the IRE, excluding the base-paired flanking regions (FL).[25] In addition, self-protection (folding) of the IRE+FL was observed and solvent-accessible regions were defined, as previously observed in a ribozyme.[31,32] The differential solvent accessibility of regions

(18) IRP was formerly called P-90, FRP, IRE-BP, or IRF. There are at least two IRPs, known as IRP-1 and IRP-2, of slightly different RNA-binding specificities, sharing 60% amino acid sequence and differing in genetic regulation.[19,20]

(19) Samaniego, F.; Chin, J.; Iwai, K.; Rouault, T. A.; Klausner, R. D. *J. Biol. Chem.* **1994**, *269*, 30904−30910.
(20) Guo, B.; Yu, Y.; Leibold, E. A. *J. Biol. Chem.* **1994**, *269*, 24542−24260.
(21) Schaefer, F. V.; Theil, E. C. *J. Biol. Chem.* **1981**, *256*, 1711−1715.
(22) Shull, G. E.; Theil, E. C. *J. Biol. Chem.* **1982**, *257*, 14187−14191.
(23) Shull, G. E.; Theil, E. C. *J. Biol. Chem.* **1983**, *258*, 7921−7923.
(24) Dix, D. J.; Lin, P.-N.; Kimata, Y.; Theil, E. C. *Biochemistry* **1992**, *31*, 2818−2822.
(25) Harrell, C. M.; McKenzie, A. R.; Patino, M. M.; Walden, W. E.; Theil, E. C. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4166−4170.
(26) Wang, Y.-H.; Sczeken, S. R.; Theil, E. C. *Nucleic Acids Res.* **1990**, *18*, 4463−4468.
(27) Wang, Y.-H.; Lin, P.-N.; Sczekan, S. R.; McKenzie, R. A.; Theil, E. C. *Biol. Met.* **1991**, *4*, 56−61.
(28) Sigman, D. S. *Biochemistry* **1990**, *29*, 9097−9105.
(29) Tullius, T. D. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 213−37.
(30) Chen, X.; Woodson, S. A.; Burrows, C. J.; Rokita, S. E. *Biochemistry* **1993**, *32*, 7610−7616.

of the IRE was recently confirmed for the imino protons by NMR spectroscopy.[33] Cleavage with (1,10-phenanthroline)Cu showed that multiple conformations of the IRE+FL region exist that can be accessed by changing the magnesium concentration during annealing;[26,27] alternate structures were not observed in secondary structure predictions for the wild-type sequence, although they were predicted for mutated sequences.[17] Fe-bleomycin detected an effect of mutation in the FL, which modulated translational regulation, on the reactivity of a site in the upper IRE stem 20 nucleotides away; the observation indicates long-range interactions between the FL and the IRE.[17]

The results reported here show that the two transition metal complexes, $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$, displayed remarkable specificity in recognizing unique sites within the IRE+FL that are important for function. $Rh(phen)_2phi^{3+}$, which intercalates in regions of double-stranded RNA that are distorted by higher order interactions,[34,35] cleaved U35 in the FL, which is flush with the phylogenetically conserved triplet of base pairs that modulate negative regulation by the IRP;[17] such specificity was not achieved with reagents such as Fe-EDTA, Cu-phen, Fe-bleomycin, or protein ribonucleases. Moreover, recognition by $Rh(phen)_2phi^{3+}$ faithfully reflected disruption or restoration of function by mutations in the FL; cleavage by $Rh(phen)_2phi^{3+}$ was unaffected by mutations elsewhere. The $Ru(tpy)(bpy)O^{2+}$ complex, which exhibits no specificity of cleavage in calf thymus DNA and single-stranded DNA,[36] cleaved the ferritin IRE at a single site in the hairpin hexaloop, G14. Apparently, a substructure in the IRE hairpin loop produced a region of special solvent accessibility at G14, although studies with ribonucleases S_1 and T_1 had indicated no unusual structure at the site.[25] (Cleavage with $Fe(EDTA)^{2-}$ at the site was at the control level, in part because of a reverse transcriptase pause site at G14.[25]) A mutation that increased the size of the IRE hairpin loop and eliminated the three-nucleotide bulge led to changes in both positive and negative regulations of translation and destroyed the $Ru(tpy)(bpy)O^{2+}$ site, but cleavage was unaffected by another mutation that preserved the hexaloop. The specificity of cleavage of sites in the ferritin IRE+FL displayed by $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$ should also be useful for detecting substructure of potential functional importance in other RNA molecules. A final point of interest is that the cleavage sites were visualized using primer extension in full-length mRNA; increased exploitation of this approach in other metallonuclease studies will open avenues for more detailed biological experiments.

Experimental Section

Mutagenesis, Transcription, Translation, and Protein Binding. Capped mRNA was prepared from wild-type and mutant DNA templates transcribed *in vitro* and translated in rabbit reticulocyte lysates as previously described;[24] negative control was assessed at low concentration of RNA, 3 nM, where the IRP was in excess, and positive control was assessed at high concentrations of mRNA, 27 nM, where [RNA] ≫ [IRP] but initiation factors, ribosomes, tRNA, and nucleotides were still in excess.[24] Globin mRNA was used as a control, since in embryonic red blood cells, ferritin mRNA is either negatively or positively regulated compared to globin mRNA depending on the iron concentration.[23] Site-directed mutagenesis followed the procedure of Kunkel,[37] using primers of 30 nucleotides with the mutations in the center of the primer; the sequence of all mutants was analyzed by sequencing and confirmed in the RNA transcripts by the primer extension analysis used in the past.[38] Binding of IRP-1 after incubation with the RNA at 30 °C for 30 min was analyzed in 4% acrylamide gels to accommodate the length (900 nucleotides) of the mRNAs as previously described[25] using RNA labeled during synthesis with [^{35}S]-uridine.[24] Note two differences in "gel shifts" between full-length functional ferritin mRNA used in Figure 3 and fragments of the IRE:[24] (1) all of the full-length ferritin mRNA is complexed with a relatively small excess of IRP, and (2) the change in mobility on protein binding is relatively small since the RNA mass is 315 kDa and the protein mass is 95 kDA; IRE-containing RNA fragments usually used range in size from 14 to 35 kDa.

Cleavage by Transition Metal Complexes. Reaction mixtures contained 95 nM RNA in 25 mM Hepes·Na, 40 mM KCl, pH = 7.2. $Ru(tpy)(bpy)O^{2+}$, prepared as described previously,[36] was added to a final concentration of 0.36 mM, and the solutions were incubated for 40 min at 30 °C. $Rh(phen)_2phi^{3+}$, kindly provided by Professor J. K. Barton, was added to a concentration of 0.10 mM, and the RNA solution was irradiated with 365 nm light (xenon lamp) for 10 min at 11 °C. The conditions were selected to cleave <1 site per molecule to avoid creation of secondary cleavage sites. Since the length/structure of the RNA precluded efficient labeling of either the 5′ or 3′ ends, cleavage sites were detected by synthesizing cDNAs corresponding to the cleavage fragments and separating the cDNA fragments by electrophoresis in calibrated acrylamide gels, as previously described.[17,25−27]

Results

Functional Effects of Mutation IL-1. Mutation IL-1 has a smaller internal loop (IL) and a larger hairpin loop (HL) than the wild-type ferritin IRE (Figure 1). The design was suggested by the magnesium-dependent difference in IRE reactivity with Cu-phen[26,27] and phylogenetic comparisons of members of the IRE family, which had either the four-nucleotide internal loop (ferritin IREs) or a single C bulge on the 5′ side of the stem (TfR and eALAS IREs).[14,15] In IL-1, U4 and G5 on the 5′ side are deleted,[39] leaving only a single unpaired C6 on the 5′ side and unpaired C23 on the 3′ side, adjacent to the major cut site for (1,10-phenathroline)Cu at G-24, G-25 (Figure 1). An enlarged HL ($n = 7$) in IL-1 is created by deletion of A11 at the 5′ junction of the HL and stem. The importance of structure at A11 was emphasized by the change of conformation, based on reactivity with Fe-bleomycin at A11, in mutation FL2 which also modulated IRE function.[17] In IL-1, the consequence of deleting A11 is creation of a heptaloop (CAGUGUU), replacing the conserved hexaloop (CAGUGU) (Figure 1).

Functional effects of the IL-1 mutation were analyzed in full-length, capped transcripts of ferritin mRNA.[24] Two types of effects were monitored. First, negative regulation (or translation repression) was measured as ferritin synthesis in rabbit reticulocyte lysates at [mRNA] < [IRP][24] and as IRP binding using mobility in acrylamide gels (Figure 2). Second, positive control or translation enhancement was analyzed as ferritin synthesis at [mRNA] > [IRP]; previous studies had shown that ferritin mRNA was translated more efficiently than other mRNAs when initiation factors and ribosomes were limiting and that the IRE was required for the effect.[14,21−24]

(31) Latham, J. A.; Cech, T. R. *Science* **1989**, *245*, 276−282.
(32) Celander, D. W.; Cech, T. R. *Science* **1991**, *251*, 401.
(33) Sierzputowska-Gracz, H.; McKenzie, R. A.; Theil, E. C. *Nucleic Acids Res.* **1995**, *23*, 146−153.
(34) Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972−982.
(35) Chow, C. S.; Barton, J. K. *Methods Enzymol.* **1992**, *212*, 219−241.
(36) Cheng, C.-C.; Goll, J. G.; Neyhart, G. A.; Welch, T. W.; Singh, P.; Thorp, H. H. *J. Am. Chem. Soc.* **1995**, *117*, 2970−2980.
(37) Kunkel, T. A.; Roberts, J. D.; Zakour, R. A. *Methods Enzymol.* **1987**, *154*, 367−382.
(38) Didsbury, J. R.; Theil, E. C.; Kaufman, R. E.; Dickey, L. F. *J. Biol. Chem.* **1986**, *261*, 949−955.
(39) Note that the traditional numbering system for the IRE family ($n = 28$) starts two nucleotides from the 5′ end of ferritin IREs that have 30 conserved nucleotides; the NMR experiments[33] used a 30-mer corresponding to the ferritin IRE where position 1 in the oligomer corresponds to −2 in the generic number scheme and the mutation G16A in the generic numbering system is G18A in the NMR experiments.[33]

Table 1. mRNA Mutations That Effect IRE-Mediated Positive or Negative Translational Control Change Structure at Rare Sites Targeted by Transition-Metal Complexes

mRNA[a]	IRP binding[b]	protein synthesis		$Rh(phen)_2phi^{3+}$	$Ru(tpy)(bpy)O^{2+}$
		positive control[c]	negative control[c]		
wild type	+	1.0	1.0	U35	G14
IL-1: negative control decreased, positive control enhanced	−	**2.7 ± 0.3**[c,d]	**0.030 ± 0.002**[c,d]	U35	**G14, G16**
HL-1: negative control decreased	−	1.1 ± 0.2	**0.089** ± 0.03[c]		G14
FL2: negative control decreased	+	1.2 ± 0.1	**0.41 ± 0.04**[c]	**nonspecific**	G14
FL2R: negative control restored	+	1.2 ± 0.1	0.78 ± 0.06	U35	G14
unregulated [globin mRNA-reticulocyte PolyA+RNA]	N/A	1.3 ± 0.1	**0.11 ± 0.02**[c]		

[a] H-type subunit ferritin, bullfrog red cell model. IL-1: decreases the internal loop in the secondary structure (deletes U and G) and also converts the hairpin loop from a hexa- to a heptaloop—CAGUGU → CAGUGUU (see Figure 1). No IRP binding was observed (Figure 2). HL-1: G/A mutation in the HL converting CAGUGU → CAGUAU (Figure 1). FL2: a triple mutation (CUC → GAG) eliminating pairing between the conserved triplet of base pairs in the IRE flanking region (FL). FL2R: a triple mutation of FL2 (hextuple mutation of wild type) that restores the triplet of base pairs in the IRE flanking region—CUC(5′)/GAG(3′) → GAG(5′)/CUC(3′) (see Figure 1). Bold type is used to emphasize difference from the wild type. [b] IRP binding is described for IL-1 in Figure 2, this work; for HL-1 in ref 26 (pseudogene) and in Figure 2, this work; and for FL2 in ref 17. [c] Protein synthesis was measured at high [RNA] and IRP deficiency to assess positive control or at low [RNA] and IRP excess to assess negative control by the IRE sequence (translation inhibition). [d] Significantly different from the wild-type ferritin mRNA at >99% confidence level. The error is presented as the standard deviation. The results are the average of three or four experiments with two or three independently prepared RNA samples. Data for HL-1 are from ref 24 and for FL2 and FL2R from ref 17. [e] Negative control of IL-1 is significantly different from that of other deregulated mRNAs HL-1, FL2, and globin at the 99% confidence level. Positive control of IL-1-directed synthesis is significantly higher than that of wild type, globin, HL-1, or FL2 at the 99% confidence level and explains the apparent difference in negative control between IL-1 and HL-1. The error is presented as the standard deviation.

Table 2. Structural Effects of Mutation in the IRE Region of Ferritin mRNA[a]

probe	wild type	FL2	HL-1	IL-1
IRP binding	+	+	−	−
		Hairpin Loop (HL)		
$Ru(tpy)(bpy)O^{2+}$	G14	G14	G14	G14, **G16**
RNAse T_1	G14	G14	G14	G14, **G16**
dimethyl sulfate[b]	A13	A13	A13, **A16**	not determined
		Junction Upper Stem/HL		
Fe-bleomycin	U17	**A10,11**[c]	U17	**A10**
		Internal Loop Region		
(1,10-phenanthroline)Cu[b]	G24,G25	G24,G25	G24,G25	**no sites**
		Flanking Region		
$Rh(phen)_2phi^{3+}$	U35	**no specific sites**	not determined	U35

[a] A site was counted as a cleavage site if the intensity on the autoradiogram was >3 times the intensity in the primer extension lane (to account for reverse transcriptase pausing). Two or three independently prepared capped, full-length mRNA transcripts and three or four cleavage reactions were used for each type of mRNA. Sites which differ from wild type are shown in bold. HL-1 has been shown by examination of the imino region of the proton NMR spectrum to have an overall structure similar to that of the wild type but to denature at much lower temperatures than the wild type,[33] which explains the loss of regulation and IRP binding measured at 30 °C. [b] Data from refs 26 and 27. [c] Data from ref 17.

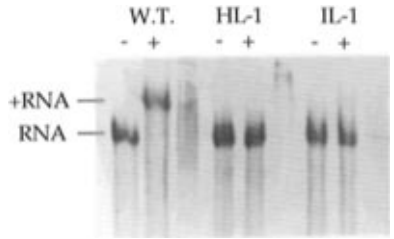

Figure 2. Autoradiogram of the polyacrylamide gel showing the results of incubation of full length ($n = 900$), capped ferritin mRNA (labeled during transcription with [^{3}H]uridine) with IRP at a ratio of 1:12. When the protein:RNA ratio was increased >25 times, large aggregates formed that apparently did not enter the gel and are represented in the third lane (L → R) for each type of RNA. Previous studies show that the RNA with mutation FL2 binds the IRP, but the structure of the protein/RNA complex is altered; see ref 17.

Negative translational control was essentially absent in IL-1; repression of ferritin synthesis directed by IL-1 is 0.03 times that of synthesis directed by wild-type mRNA (Table 1). IL-1 RNA did not bind the IRP (Figure 2), which explains the absence of negative control. Another ferritin mRNA that cannot bind the IRP is HL-1 (a single-base substitution, G/A, in the HL);[24] repression of ferritin synthesis directed by HL-1 mRNA is 0.089 times that of wild-type mRNA. Modulation of negative translational control was observed in a third mutation, FL2, which is mutated 20 nucleotides downstream from the IRE hairpin loop of the IRE. FL2 mRNA binds the IRE, but the RNA—protein complex has a different structure for the wild type.[17] The effect of IRP binding on repression in FL2 is diminished to 0.41 times that of the wild type (Table 1); repression is restored to wild-type levels by the FL2R mutation, which restores the wild type pairing of the FL that is absent in the FL2 mutant.

Positive translational control was increased 2.7 times compared to that of the wild type in the IL-1 mRNA. In the case of the HL-1 and FL2 mutations, it was possible to alter negative control without altering positive control whereas the IL-1 structure changed both positive and negative control. Whether it will be possible to change positive control without also changing negative control is not yet known. On the basis of the reactivity with RNAse T_1, dimethyl sulfate, Fe-bleomycin, and $Ru(tpy)(bpy)O^{2+}$ (Table 2), the hairpin loop structure is changed in the IL-1, HL-1, and FL2 mRNAs, which likely contributes to the altered translational control.

$Ru(tpy)(bpy)O^{2+}$ Cleavage. Complexes based on $Ru(tpy)(bpy)O^{2+}$ oxidize nucleotides via activation of the 1′-hydrogen and oxidation of the guanine base;[36,40−42] the latter pathway is 7 times more efficient on a per nucleotide basis in single-

(40) Grover, N.; Thorp, H. H. *J. Am. Chem. Soc.* **1991**, *113*, 7030.

stranded DNA.[43] The base oxidation proceeds through an inner-sphere Ru−O−G intermediate, which increases the steric demand of the guanine oxidation pathway. As a result, the ratio of cleavage at guanine compared to cleavage at sugar in duplexes is only half that in single strands. Because of the size (6 Å), shape of the organic ligand, and the double positive charge, RNA cleavage by $Ru(tpy)(bpy)O^{2+}$ is restricted to a small number of surface sites. The complex therefore targets solvent-accessible sites that are prone to cation binding. For example, loop sites in $tRNA^{Phe}$ that are near strong Mg^{2+}-binding sites are targeted by $Ru(tpy)(bpy)O^{2+}$.[44]

Cleavage by $Ru(tpy)(bpy)O^{2+}$ occurred at a single site in the IRE+FL. Such specificity most likely represents the higher-order folding of the IRE previously implicated by cleavage studies[25] and by differential accessibility of D_2O observed in the regions of the ^1H NMR representing duplex segments.[33] The results with RNA contrast with those obtained with duplex and single-stranded DNAs, which are cleaved at nearly every base. Further, a DNA analog of the hexaloop and first five stem base pairs is cleaved at nearly every site, although the site analogous to G14 is the strongest followed by the G16 site.[43] Single-stranded guanines are particularly reactive toward the reagent, so it is noteworthy that the internal loop G5 is not targeted by the reagent. The $Ru(tpy)(bpy)O^{2+}$ site (G14) at the top of the hairpin loop is a site for which no distinctive substructure had been previously detected. For example, G14 is predicted to have no secondary interactions and was cleaved by nucleases that cleave single-stranded G, RNAse T_1, or cleave nonspecifically in single-stranded regions, RNAse S_1.[1] *Thus, Ru(tpy)(bpy)O*$^{2+}$ *identified a specific substructure in the IRE hairpin loop that had been transparent to the other probes used.*

A mutation that converted the hexaloop CAGUGU to CAGUGUU (IL-1) and eliminated the internal loop decreased specificity of cleavage by $Ru(tpy)(bpy)O^{2+}$ so that both G14 and G16 were reactive (Figure 2). However, mutations in the FL (FL2 and FL2R) as well as base substitutions in the hairpin loop that maintained the hexaloop (CAGU*G*U → CAGU*A*U) had no effect on cleavage of G14 by $Ru(tpy)(bpy)O^{2+}$ (Table 1). Moreover, the IL-1 mutation had no effect on $Rh(phen)_2phi^{3+}$ cleavage in the FL (Table 2), emphasizing the specificity of recognition.

$Rh(phen)_2phi^{3+}$ Cleavage. Complexes based on $Rh(phen)_2phi^{3+}$ cleave DNA upon UV irradiation by abstraction of the 3′-hydrogen from the sugar functionality.[45] The excited-state lifetime of the complex is extremely short (<1 ns), so cleavage only occurs in binding orientations that are favorable with regard to the cleavage reaction. On the basis of studies with tRNA, $Rh(phen)_2phi^{3+}$ cleaves RNA at sites in which the secondary structure is distorted by higher-order interactions.[34] The specificity with RNA[34,35,46−48] contrasts with that with DNA, where the reagent cleaves at sites of intercalation;[49,50] the $(phen)_2phi$ ligand set is apparently too large to fit into the major groove of the RNA A-helix unless the helix is distorted. Cleavage of the TAR RNA occurs adjacent to and opposite of a bulge,[13] which could be categorized as a higher-order site where a helix−bulge junction occurs.

The major cleavage site in the ferritin mRNA IRE regulatory region by $Rh(phen)_2phi^{3+}$ occurred at U35 in the base-paired flanking region (FL) (Figure 1). The structure of the FL at U35 therefore presents an "open" intercalation site that is distorted by some higher-order interactions, possibly because of proximity of the UU bulge to the helical segment. The structure in and around U35 in ferritin mRNA also has several other distinctive properties: (1) U35 is adjacent to the conserved CUC/GAG conserved triplet of base pairs, which modulate negative regulation;[17] (2) Formation of the IRE+IRP RNA/protein complex changed the conformation of U35 and the triplet of base pairs;[25] and (3) U35 and the conserved triplet GAG are hyperreactive with the hydroxyl radical generated by Fe-EDTA, indicating an outer surface location in the folded IRE+FL structure.[25]

The substructure at U35 recognized by $Rh(phen)_2phi^{3+}$ is destroyed when the triplet of base pairs is disrupted by mutation FL2 (5′FL-CUC → GAG, in the conserved triplet of base pairs, Figure 1); IRE mutations HL-1 and IL-1 showed no effect on the cleavage pattern (Table 1). In the FL2 mutant, many cleavage sites were observed throughout the IRE+FL, which may reflect the multiple secondary structures predicted.[17] The FL2 mutation also displayed diminished translational regulation (Table 1) and changed the structure in the IRE itself, 20 nucleotides away.[17]

A second set of mutations in FL2 that restored the triplet base pairs FL2R (3′FL-GAG → CUC) reversed in 3′−5′ positions also restored the substructure recognized by $Rh(phen)_2phi^{3+}$ (Figure 1) as well as normal translational regulation.[17] U35 is cleaved in the mutant derivative of FL2R that restores the triple base pairs, albeit with the base sequence reversed (Figure 1). Since $Rh(phen)_2phi^{3+}$ cleavage paralleled the loss and restoration of normal regulation, $Rh(phen)_2phi^{3+}$ recognized a substructure in the IRE+FL related to function.

Discussion

Chemical Reactivity. The chemical reactivity of nucleobases and nucleotides underlies many biological reactions that are of importance in metabolism and chemotherapy.[7−10] The chemistry of nucleic acid cleavage and mononucleotide oxidation by $Ru(tpy)(bpy)O^{2+}$ has been discussed in detail elsewhere.[2] In DNA, oxidation of the 1′-hydrogen by $Ru(tpy)(bpy)O^{2+}$ in the minor groove yields a characteristic furanone product and a 3′-modified terminus detectable on sequencing gels prior to piperidine treatment.[36] In mononucleotides, oxidation of the 1′-hydrogen is 5-fold slower for ribonucleotides containing the 2′-hydroxyl compared to the deoxy nucleotides.[43] The difference in oxidation rates for ribo- and deoxyribonucleotides can be ascribed to destabilization of the 1′ oxidation product by the polar effect of the 2′-hydroxyl, as observed in other nonbiological systems. $Ru(tpy)(bpy)O^{2+}$ also oxidizes guanine via an inner-sphere Ru(III)−O−G adduct, a reaction less susceptible to attenuation by substituents on the sugar ring.

Differences in the kinetics of oxidation of ribonucleotides and deoxyribonucleotides lead to the prediction that RNA should show considerably fewer cleavage sites attributable to sugar oxidation than DNA of the same length. In the IRE model, the $Ru(tpy)(bpy)O^{2+}$ cleavage pattern for the DNA analog of the IRE hexaloop and first five base pairs of the stem shows cleavage at nearly every site in the 16-mer.[43] In contrast,

(41) Welch, T. W.; Neyhart, G. A.; Goll, J. G.; Ciftan, S. A.; Thorp, H. H. *J. Am. Chem. Soc.* **1993**, *115*, 9311−9312.
(42) Neyhart, G. A.; Grover, N.; Smith, S. R.; Kalsbeck, W. A.; Fairley, T. A.; Cory, M.; Thorp, H. H. *J. Am. Chem. Soc.* **1993**, *115*, 4423.
(43) Neyhart, G. A.; Cheng, C.-C.; Thorp, H. H. *J. Am. Chem. Soc.* **1995**, *117*, 1463−1471.
(44) Cheng, C.-C.; Thorp, H. H. Unpublished results.
(45) Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303.
(46) Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534−3542.
(47) Chow, C. S.; Barton, J. K. *Biochemistry* **1992**, *31*, 5423−5429.
(48) Lim, A. C.; Barton, J. K. *Biochemistry* **1993**, *32*, 11029−11034.
(49) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432.
(50) Krotz, A. H.; Hudson, B. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12577−12578.

cleavage of the IRE and mutants is observed only at loop guanines, indicating that the 2′-hydroxyl dramatically attenuates the reactivity of the RNA sugars and explaining in part the lower levels of RNA cleavage observed for a variety of oxidiative cleavage agents.[28,32,48,51] Indeed, cleavage by $Rh(phen)_2phi^{3+}$, which occurs via activation of the 3′-hydrogen,[45] should also be subject to deactivation by the 2′-hydroxyl, and considerably higher specificities are observed with RNA compared to DNA. For example, cleavage of tRNA by $Rh(phen)_2phi^{3+}$ shows considerably fewer sites than cleavage of the analogous synthetic tDNA.[48] Other explanations for differences in the number of cleavage sites in RNA and DNA with the same primary sequence include RNA folding and physical accessibility of sites to cleavage reagents.[1]

Structural Information. Understanding the structure of the IRE family of mRNA isoregulatory sequences requires knowledge of the tertiary interactions, in addition to the primary and secondary structures now known. The presence of higher-order interactions is indicated by the high phylogenetic conservation of primary structure among the members of the IRE family.[14,15] Moreover, mutagenesis studies show that any change, including replacing a single base or base pair with another, alters IRP binding.[24,52−55] In addition, probing the IRE structure revealed few sites for RNAses either with specificity for single strands (S_1) or stacked bases (V_1),[25,54] and methylation patterns deviated from that expected for the predicted secondary structure.[26,27] Additional evidence for folding and higher-order interactions was provided by the demonstration of differential solvent accessibility along the IRE shown by self-protection to hydroxyl radical cleavage,[25] differential exchange with D_2O, long-range NOEs, and subdomains of different thermal stability.[33]

Specificity of cleavage among G residues in the IRE has been observed with nuclease T_1,[25] with Cu-phen,[26] and now with $Ru(tpy)(bpy)O^{2+}$. Among these three reagents, only $Ru(tpy)(bpy)O^{2+}$ is delivered to RNA in a form that is activated toward oxidation prior to binding. A series of events involving binding of the reagent to the nucleic acid, reaction with hydrogen peroxide, and possible formation of a diffusible intermediate must occur for oxidation by Cu-phen.[28] Likewise, enzyme nucleases must bind in a specific conformation to effect cleavage and could change the structure of the IRE. Kinetic studies have shown that $Ru(tpy)(bpy)O^{2+}$ is reduced each time it binds to DNA,[42] so the reagent is a direct probe of the number of collisions that lead to productive DNA oxidation; i.e., sites of greatest accessibility will be cleaved more frequently.

The most likely explanation for the $Ru(tpy)(bpy)O^{2+}$ specificity is that the guanine residues other than G14 are spatially inaccessible to the reagent.[30] Inaccessibility has been previously suggested to explain the small number of either RNAse T_1, S_1, or V_1 sites in the IRE+FL.[25,26] Of particular interest is the lack of cleavage at the single-stranded G16 and G5, which may be explained by a structural model that was advanced recently on the basis of NMR studies.[33] In this model, the IRE is folded at the internal loop, resulting in a tertiary interaction between G16 and sites lower in the stem, which would render G16 inaccessible. If the folding occurs because of flexibility in the internal loop (U4−C6), G5 may also therefore be inaccessible. Substitution of A for G at position 16 greatly decreased the temperature stability in the specific folded structure as measured by NMR.[33] This explanation also accounts for the observed

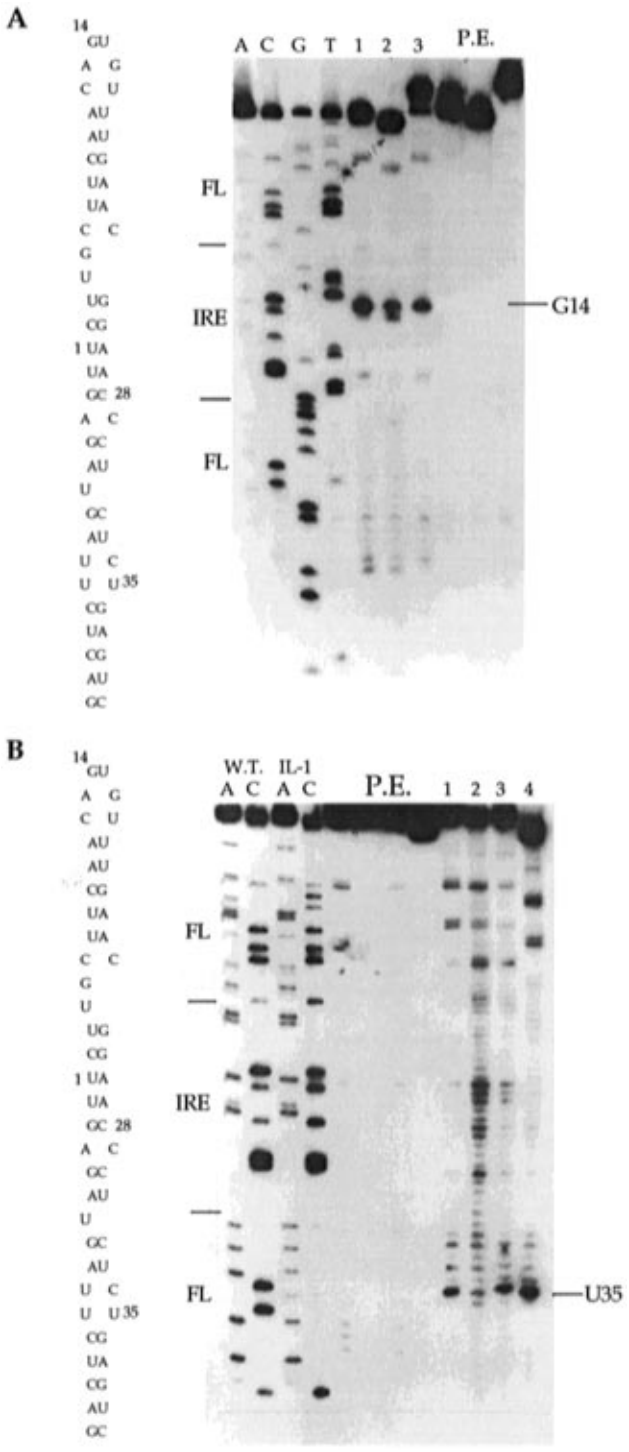

Figure 3. Autoradiogram showing the analysis of the ferritin IRE (iron regulatory element) (from three or four reactions; data represent at least three or four independent experiments). Results with IL-1 and FL2 are significantly different ($p > 99$) from wild type. (A) Cleavage by $Ru(tpy)(bpy)O^{2+}$ of ferritin mRNAs: wild type (lane 1), IL-1 (lane 2), and HL-1 (lane 3). (B) Cleavage by $Rh(phen)_2phi^{3+}$ of ferritin RNAs: wild type (lane 1), FL2 (lane 2), FL2R (lane 3), and IL-1 (lane 4). Mutations in IL-1, HL-1, FL2, and FL2R are shown in Figure 1. P.E. refers to the reverse transcription analysis on uncleaved RNA and indicates natural pause sites in the primer extension (P.E.) reaction.

(51) Holmes, C. E.; Hecht, S. M. *J. Biol. Chem.* **1993**, *268*, 25909−25913.
(52) Leibold, E. A.; Landano, A.; Yu, Y. *Nucleic Acids Res.* **1990**, *18*, 1819−1824.
(53) Barton, H. A.; Eisenstein, R. S.; Bomford, A.; Munro, H. N. *J. Biol. Chem.* **1990**, *265*, 7000−7008.
(54) Bettany, A. J. E.; Eisenstein, R. S.; Munro, H. N. *J. Biol. Chem.* **1992**, *267*, 16531−16537.
(55) Hentze, M. W.; Rouault, T. A.; Wright, S. W.; Dancis, A.; Harford, J. B.; Klausner, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6730−6734.
(56) Kalsbeck, W. A.; Thorp, H. H. *Inorg. Chem.* **1994**, *33*, 3427−3429.
(57) Sharp, K. A.; Honig, B.; Harvey, S. C. *Biochemistry* **1990**, *29*, 340−346.
(58) Jayaram, B.; Sharp, K. A.; Honig, B. *Biopolymers* **1989**, *28*, 975−993.

cleavage of both G14 and G16 in the IL-1 mutant, because folding of the IRE (and the associated tertiary interaction at G16) may no longer be possible once the internal loop is changed. Changes due to expansion of the hairpin loop by one base are also likely to contribute to stabilizing folding as indicated by changes in stability of the HL-1 mutation where a G/A substitution occurred at G16 in the hairpin loop.[33] IRP SELEX analysis of IREs with alterted hexaloops further emphasizes the special role of G16 in stabilizing the IRE structure.[59]

Cleavage of nucleic acids by $Rh(phen)_2phi^{3+}$ requires binding and intercalation. In the case of DNA, binding is in the major groove; however, the major groove of RNA is too small for $Rh(phen)_2phi^{3+}$ to bind.[34,35,46] Thus, cleavage in tRNA occurred only in helical regions that had two properties: (1) stacked bases to permit intercalation and (2) distortion of the base stack by higher-order interactions, which facilitated binding of the complex.[34,35,46] The photoactivated cleavage reaction with $Rh(phen)_2phi^{3+}$ involves a short-lived (<100 ps) UV-excited state.[45,60] Only those bound RNA sites in the correct orientation will be cleaved, since the Rh excited state will decay long before conformational change or equilibration with other sites can occur. In the ferritin IRE+FL, four sets of stacked bases occur, on the basis of the secondary-structure prediction (Figure 1). $Rh(phen)_2phi^{3+}$ cleaved the IRE+FL only at an FL site adjacent to the conserved triplet of base pairs (Figure 1), indicating a specific structure at the site that was absent elsewhere. This result is reminiscent of studies of intercalator binding to tRNA, where only intercalation into the accessible acceptor stem is realized for the folded structure.[61−65] Intercalation leads to lengthening and unwinding of the helix, so intercalation can destroy important tertiary interactions between helical segments in folded structures. Since folding of the IRE in the current structural models involves tertiary interactions between the hairpin loop and sites lower in the stem, lengthening and unwinding of the helical segments of the IRE may not be tolerable. The observation of an intercalation site in the FL region therefore suggests that this region, like the acceptor stem of tRNA, has sufficient tertiary structure to distort the base stack and facilitate binding of $Rh(phen)_2phi^{3+}$.

Functional Relevance. IRE+FL cleavage by $Ru(tpy)(bpy)O^{2+}$ and $Rh(phen)_2phi^{3+}$ was distinct for each reagent and identified sites, shown by mutation, to have different functional effects (Table 1). For example, $Rh(phen)_2phi^{3+}$ recognized a site in the FL that modulated negative translational control (Table 1) but not IRP binding[17] while $Ru(tpy)(bpy)O^{2+}$ identified a site in the hairpin loop at G14 that affected IRP binding/negative control and positive control (Table 1, Figures 2 and 3). In earlier studies with more classical reagents, G14 exhibited no special reactivity. The cleavage of G14 by RNAse T_1 and S_1,[26,27] for example, was predicted from the secondary structure for the IRE. Specificity of the substructure at G14 is emphasized by the fact that mutation IL-2, which increases the size of the hairpin loop, changed the structure recognized by $Ru(tpy)(bpy)O^{2+}$. In contrast, the mutations in the FL (FL2, FL2R) or base substitutions in the hairpin loop (HL-1) had no effect on cleavage of G14 by $Ru(tpy)(bpy)O^{2+}$.

Design of a mutation in the FL site of the IRE depended solely on phylogenetic conservation of a specific triplet of base pairs that did indeed influence IRE function.[17,25,27] In spite of the phylogenetic conservation of the triplet of base pairs, no substructure had been detected to distinguish the base pairs from others in the FL. The specific cleavage by $Rh(phen)_2phi^{3+}$ now observed (Figure 1, Table 1) indicates the presence of a distinctive structure around the triplet of base pairs that is associated with wild-type regulation, since the only region in the IRE+FL that was recognized by $Rh(phen)_2phi^{3+}$ was proximal to the triplet of base pairs and was changed by a mutation that altered regulation. Concomitant restoration of the $Rh(phen)_2phi^{3+}$ site and normal regulation by derivative mutation emphasize the specificity of structure recognition by $Rh(phen)_2phi^{3+}$ also recently observed with the TAR element of HIV.[13]

The two substructures detected by probing with the transition metal complexes $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$ have different functional effects, indicating that the ferritin IRE+FL can be thought of as a composite of at least two structure/function subdomains; subdomains with specific structural features have previously been observed in the group I intron of Tetrahymena.[66] The subdomain with sites recognized by $Rh(phen)_2phi^{3+}$ has regulatory activity for negative control as well as a conserved sequence[25] although it is in the FL, outside the binding site for the IRP. In contrast, the IRE domain with a site recognized by $Ru(tpy)(bpy)O^{2+}$ has regulatory activity for both negative and positive controls, has conserved sequence common to other members of the IRE family,[15] and is within the binding site for the IRP.[25] In addition to new information on IRE+FL substructure and function, the use of $Rh(phen)_2phi^{3+}$ and $Ru(tpy)(bpy)O^{2+}$ should help identify substructures with potential functional significance in ribozymes, viral RNAs, rRNAs, and other mRNAs.

Acknowledgment. This work was supported in part by NIH Grant DK-20251 (E.C.T.), the National Science Foundation (H.H.T.), and NSF Grant DMB-8818203 (W.E.W.).

IC951094U

(59) Henderson, B. R.; Menotti, E.; Bonnard, L.; Kühn, L. C. *J. Biol. Chem.* **1994**, *269*, 17481−17489.
(60) Sitlani, A.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12589−12560.
(61) Jones, C. R.; Bolton, P. H.; Kearns, D. R. *Biochemistry* **1978**, *17*, 601−607.
(62) Kean, J. M.; White, S. A.; Draper, D. E. *Biochemistry* **1985**, *24*, 5062−5070.
(63) Tanner, N. K.; Cech, T. R. *Nucleic Acids Res.* **1985**, *13*, 7759−7779.
(64) Urbanke, C.; Römer, R.; Maass, G. *Eur. J. Biochem.* **1973**, *33*, 511−516.
(65) Wells, B. D.; Cantor, C. R. *Nucleic Acids Res.* **1977**, *4*, 1667−1680.
(66) Murphy, F. L.; Cech, T. R. *J. Mol. Biol.* **1994**, *236*, 49−63.