



# The Open Biochemistry Journal

Content list available at: <https://openbiochemistryjournal.com>



## RESEARCH ARTICLE

### Characteristics of the Iron-responsive Element (IRE) Stems in the Untranslated Regions of Animal mRNAs

Bin Wang<sup>1,\*</sup>, Michael S. Thompson<sup>1</sup> and Kevin M. Adkins<sup>1</sup>

<sup>1</sup>Department of Chemistry, Marshall University, Huntington, WV 25755, United States

#### Abstract:

#### Background:

Iron-responsive Elements (IREs) are hairpin structures located in the 5' or 3' untranslated region of some animal mRNAs. IREs have a highly conserved terminal loop and a UGC/C or C bulge five bases upstream of the terminal loop, which divides the hairpin stem into an upper stem and a lower stem.

#### Objective:

The objective of this study was to investigate the base-pair composition of the upper and lower stems of IREs to determine whether they are highly conserved among mRNAs from different genes.

#### Methods:

The mRNA sequences of six 5'IREs and five 3'IREs from several animal species were retrieved from the National Center for Biotechnology Information. The folding free energy of each IRE mRNA sequence was predicted using the RNAfold WebServer.

#### Results:

We found that the upper and lower stems of IREs are not highly conserved among the mRNAs of different genes. There are no statistically significant differences in the IRE structures or folding free energies between mammalian and non-mammalian species relative to either the ferritin heavy chain 5'IRE or ferroportin 5'IRE. There are no overall significant differences in the folding free energies between UGC/C-containing 5'IREs and C-bulge-containing 5'IREs, or between 5'IREs and 3'IREs.

#### Conclusion:

Further studies are needed to investigate whether the variations in IRE stem composition are responsible for fine-tuning the IRE/Iron-Regulatory Protein interactions among different mRNAs to maintain the balance of cellular iron metabolism, and to identify whether evolutionary processes drive the base-pair composition of the upper and lower stems of IREs toward any particular configuration.

**Keywords:** Iron-regulatory protein (IRP), Iron-responsive element (IRE), Untranslated region (UTR) of mRNA, IRE upper stem, IRE lower stem, Animal mRNA.

#### Article History

Received: December 2, 2020

Revised: February 21, 2021

Accepted: March 1, 2021

## 1. INTRODUCTION

The interactions between Iron-Responsive Elements (IREs) and Iron-Regulatory Proteins (IRPs) are among the most extensively investigated post-transcriptional regulatory mechanisms, as cellular iron metabolism is regulated through the

IRE/IRP system [1 - 8]. When intracellular free iron content is low, IRPs bind to IREs in two locations: 1) the 5' Untranslated Regions (5'UTRs) of certain messenger RNAs (mRNAs), where they inhibit ribosome binding and thus mRNA translation, and 2) the 3' Untranslated Regions (3'UTRs) of several other mRNAs, stabilizing those mRNAs against endonucleolytic degradation. When intracellular iron concentration is high, IRP-IRE binding is inhibited, which allows the 5'IRE-containing mRNAs to be translated to produce proteins involved in iron storage (ferritin heavy chain,

\* Address correspondence to this author at the Department of Chemistry, Marshall University, One John Marshall Drive, BBSC 241L, Huntington, WV 25755, United States; Tel: 001-304-696-3456; E-mail: wangb@marshall.edu

FTH; and ferritin light chain, FTL), iron export (Ferroportin, FPN), and heme biosynthesis catalysis (erythroid aminolevulinic acid synthase 2, ALAS2), as well as enzymes of the tricarboxylic acid cycle (mitochondrial aconitase 2, ACO2), and transcription factors produced in response to hypoxia (hypoxia-inducible transcription factor 2 $\alpha$ , HIF2 $\alpha$ ). The inhibition of IRP-IRE binding at high cellular iron concentrations also destabilizes the 3'IRE-containing mRNAs, including the Transferrin Receptor (TFRC), Divalent Metal Transporter 1 (DMT1), and cell cycle regulator (cell division cycle 14A, CDC14A), which facilitates the degradation of these mRNAs and inhibits further iron uptake [1, 2, 6, 9 - 13].

IREs are highly conserved hairpin stem-loop structures containing 26-30 nucleotides. The terminal loop sequence is 5'-CAGUGN-3' (N can be C, U, or A, but never G). There is a conserved UGC/C or C bulge five bases upstream of the terminal loop, which divides the hairpin stem into an upper stem and a lower stem [9, 14]. In recent years, IRE-containing mRNAs, such as the mRNAs of amyloid precursor protein and  $\alpha$ -synuclein, have been discovered by biochemical and computational approaches [15, 16]. Some newly discovered IREs do not contain the conserved 5'-CAGUG-3' terminal loop sequence, and the C bulge is not located exactly five bases upstream of the terminal loop [15 - 18].

The IRP mainly binds to IREs at two separate sites: the AGU apical loop and the C bulge [19, 20]. The upper stem of the IRE, between the terminal loop and the bulge, and the lower stem below the bulge, both play important roles in maintaining the orientation of the loop and the bulge for proper IRE-IRP binding. Few studies have investigated the stems of IREs, nor have the separate roles of the upper and lower stems been examined. We hypothesize that the base-pair composition of the upper and lower stems of IREs are not highly conserved among mRNAs from different genes (except those with the irregular terminal loop and bulge) and that the IREs with a UGC/C internal loop have a somewhat different stem composition than those with a single C bulge. To test this hypothesis, we investigated six 5'IREs and five 3'IREs from animal mRNAs and analyzed the number of Watson-Crick base-pairs, wobble base-pairs, and mismatched pairs in the upper and lower stems of each type of IRE.

## 2. METHODS

Among the IREs in the mRNAs of nine genes investigated in this study, six are located in the 5'UTRs of mRNAs (FTH, FTL, ACO2, FPN, ALAS2, and HIF2 $\alpha$ ), and the other three are located in the 3'UTRs of mRNAs (CDC14A, DMT1, and TFRC). The mRNA sequences were retrieved from the National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/nuccore/>). Each mRNA investigated was found under the "Nucleotide" category, with the following filters activated: Species-Animals, Molecule types-mRNA, and Source databases-RefSeq. Of the various mRNAs that were retrieved, only those that contain experimentally supported base sequences were analyzed; sequences labeled as "Predicted mRNA" were not included.

From the mRNAs that were retrieved using these parameters, only the IRE sequences containing a conserved 5'-

CAGUG-3' terminal loop and a conserved UGC/C or C bulge five bases upstream of the terminal loop were analyzed. For example, the 3'UTR of TRFC mRNA contains five IREs. However, only three of them have the exact 5'-CAGUG-3' loop sequence and a C bulge five bases upstream of the terminal loop; therefore, only these three 3'IRE sequences were analyzed.

For a complete comparison, we investigated IRE sequences in both mammals and non-mammalian animals for the mRNAs of nine genes. Some species have incomplete UTR information; therefore, either no IRE sequence or only partial sequences were retrieved. We did not include transcripts with absent or partial sequences in our analysis but included them in Tables 1-9 with a brief note.

The minimum folding free energy of each mRNA sequence in Tables 1-9 was predicted using the RNAfold WebServer (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>), which is a component of the ViennaRNA package developed by the Institute for Theoretical Chemistry at the University of Vienna.

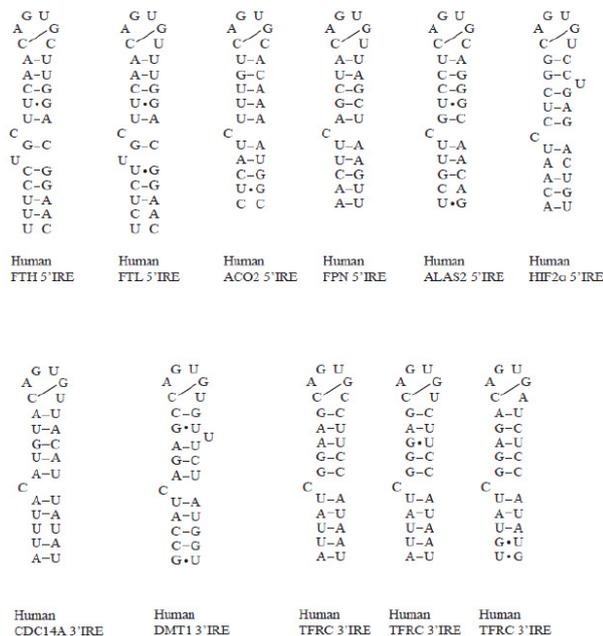
Statistical analyses were performed using Excel to compare the differences in IRE structure and folding free energy between mammalian and non-mammalian species for the transcripts of each of the nine genes studied. Seven analyses were run to compare the mammalian and non-mammalian species relative to 1) the number of AU pairs in the upper stems of their IREs, 2) the number of GU pairs in the upper stems of their IREs, 3) the number of GC pairs in the upper stems of their IREs, 4) the number of AU pairs in the lower stems of their IREs, 5) the number of GU pairs in the lower stems of their IREs, 6) the number of GC pairs in the lower stems of their IREs, and 7) the predicted folding free energies of the IRE transcripts of each gene in this study. An *F*-test was performed to determine whether the variance ( $s^2$ ) was significant at the 95% confidence level, followed by a *t*-test ("two-sample equal variance" or "two-sample unequal variance," according to the results from the *F*-test) to determine whether the difference between mammalian and non-mammalian species was significant at the 95% confidence level for each of the seven comparisons listed above. If Excel returned a calculated *p*-value of less than 0.05, a significant difference was indicated.

Statistical analyses were also performed to compare the folding free energies between 5'IREs and 3'IREs, and between UGC/C-containing IREs and C-bulge-containing IREs. Non-mammalian species could not be analyzed due to a lack of available IRE mRNA sequence data. The results are summarized in Table 10.

## 3. RESULTS AND DISCUSSION

Fig. (1) illustrates the secondary structures of eleven of the human IREs contained in the transcripts of the nine genes investigated in this study. Data presented in Tables 1-9 provide the base-pair compositions of the upper stem (*i.e.*, the five base-pair helix between the terminal loop and the bulge) and the lower stem (*i.e.*, the five base-pair helix below the bulge) in the 5'IREs of FTH, FTL, ACO2, FPN, ALAS2, and HIF2 $\alpha$  mRNAs, and in the 3'IREs of CDC14A, DMT1, and TFRC

mRNAs, respectively. The five base-pair upper stems of IREs are shown in bold; the five base-pair lower stems are underlined. The predicted folding free energy of each mRNA sequence is also presented in Tables 1-9.



**Fig. (1).** An illustration of the secondary structures of IREs in the 5'UTRs of human FTH, FTL, ACO2, FPN, ALAS2, and HIF2α mRNAs (top row), and in the 3'UTRs of human CDC14A, DMT1, and the three human TFRC mRNAs used in this study (second row). Note: A Watson-Crick base-pair is depicted by a dash, and a wobble base-pair is depicted by a dot. The line connecting the 1<sup>st</sup> nucleotide (C) and the 5<sup>th</sup> nucleotide (G) in the terminal loop of an IRE indicates the existence of a Watson-Crick base-pairing between these two nucleotides.

Regarding the upper stems, in the FTH 5'IRE (Table 1), five of the mammalian species studied (*Homo sapiens*, *Pongo abelli*, *Macaca mulatta*, *Sus scrofa*, and *Canis lupus familiaris*) contain three AU Watson-Crick pairs, one GU wobble pair, and one GC Watson-Crick pair in their upper stems. Three other mammalian species (*Bos taurus*, *Mus musculus*, and *Ovis aries*) contain four AU pairs and one GC pair in their upper stems. Four non-mammalian species (*Danio rerio*, *Salmo salar*, *Xenopus laevis*, and *Xenopus tropicalis*) also contain four AU pairs and one GC pair, whereas the other two non-mammalian species (*Gallus gallus* and *Anas platyrhynchos*) contain two AU pairs, one GU pair, one GC pair, and one mismatched GA pair in their upper stems. Statistical analyses indicate that there are no significant differences ( $p > 0.05$ ) between mammalian and non-mammalian species regarding the number of AU, GU, or GC pairs in the upper stem of the FTH 5'IRE.

In the FTL 5'IRE (Table 2), nine out of ten mammalian species studied contain three AU pairs, one GU pair, and one GC pair in their upper stems, whereas *Mus musculus* has four AU pairs and one GC pair. For the FTL mRNAs, no non-mammalian species contains a complete IRE sequence, so a statistical comparison to mammalian species could not be conducted. In the ACO2 5'IRE (Table 3), all of the mammalian

species studied contain four AU pairs and one GC pair in their upper stems, whereas the non-mammalian *Gallus gallus* contains three AU pairs, one GC pair, and one mismatched UC pair. Statistical analysis was not performed because data are available from only one of the non-mammalian species. In the FPN 5'IRE (Table 4), all of the species studied contain three AU and two GC pairs in their upper stems. There is no difference between mammalian and non-mammalian species. In the ALAS2 5'IRE (Table 5), all of the four mammalian species studied contain one AU, one GU, and three GC pairs in their upper stems, whereas the non-mammalian *Danio rerio* contains one AU, three GC, and one mismatched UU pair in its upper stem. Statistical analysis was not performed because data are available from only one of the non-mammalian species. In the HIF2α 5'IREs (Table 6), the 5'IRE structure was only identified in humans among the mammalian species, and contains one AU and four GC pairs in its upper stem. The non-mammalian *Ictalurus punctatus* contains one AU, one GU, and three GC pairs in its upper stem; *Danio rerio* contains two AU, one GU, and two GC pairs; the other two non-mammals, *Xenopus laevis* and *Xenopus tropicalis*, both contain two AU and three GC pairs in their upper stems. Statistical analysis was not performed because data are available from only one of the mammalian species.

In the CDC14A 3'IRE (Table 7), two of the mammalian species studied (*Homo sapiens* and *Rattus norvegicus*) contain four AU pairs and one GC pair in their upper stems, whereas *Mus musculus* contains three AU, one GC, and one mismatched AC pair. No non-mammalian species contain a complete IRE sequence in its CDC14A mRNA, so a statistical analysis was not conducted. In the DMT1 3'IRE (Table 8), three of the mammalian species studied (*Homo sapiens*, *Mus musculus*, and *Macaca fascicularis*) contain two AU, one GU, and two GC pairs in their upper stems. The *Rattus norvegicus* mRNA has two 3'IREs: one contains two AU, one GU, and two GC pairs in its upper stem; the other has three AU, one GU, and one GC pair in its upper stem. No non-mammalian species contain a complete IRE sequence in their DMT1 mRNA, so a statistical analysis was not conducted. Of the three 3'IREs in mammalian TFRC mRNAs studied (Table 9), most fall into one of two categories of base-pair composition in their upper stems: either two AU and three GC pairs, or one AU, one GU, and three GC pairs. Two exceptions are *Rattus norvegicus*, which has one of the three 3'IREs containing one AU pair, three GC pairs, and one mismatched AC pair, and *Cavia porcellus*, which has one of the three 3'IREs containing one AU, one GU, two GC, and one mismatched AC pair. No non-mammalian species contain a complete IRE sequence in their TFRC mRNA, so a statistical analysis was not conducted.

Regarding the lower stems, for the FTH 5'IRE (Table 1), seven mammalian and two non-mammalian species studied (*Homo sapiens*, *Pongo abelli*, *Macaca mulatta*, *Sus scrofa*, *Bos taurus*, *Mus musculus*, *Ovis aries*, *Danio rerio*, and *Salmo salar*) contain two AU, two GC, and one mismatched UC or AC pair in their lower stems. One mammalian (*Canis lupus familiaris*) and two non-mammalian species (*Gallus gallus* and *Anas platyrhynchos*) contain two AU and three GC pairs. The non-mammalian *Xenopus laevis* contains two AU, one GU, and two GC pairs in its lower stem. Statistical analyses indicate that

there are no significant differences ( $p > 0.05$ ) between mammalian and non-mammalian species regarding the number of AU, GU, or GC pairs in the lower stem of the FTH 5'IRE.

In the FTL 5'IRE (Table 2), all ten mammalian species studied contain one AU, one GU, one GC, and two mismatched (UC and CA/AA/GA) pairs in their lower stems. No non-mammalian species contains a complete IRE sequence in its FTL mRNA, so a statistical analysis was not conducted. In the ACO2 5'IRE (Table 3), all of the mammalian species studied contain two AU, one GU, one GC, and one mismatched CC pair in their lower stems, whereas the non-mammalian *Gallus gallus* contains one AU and four mismatched pairs in its lower stem. Statistical analysis was not performed because data are available from only one of the non-mammalian species. In the FPN 5'IRE (Table 4), all of the mammalian species studied contain four AU pairs and one GC pair in their lower stems. The non-mammalian *Danio rerio* contains three AU, one GU, and one GC; *Gallus gallus* contains three AU and two GC pairs

in its lower stem. Statistical analyses indicate that there are no significant differences ( $p > 0.05$ ) between mammalian and non-mammalian species regarding the number of AU, GU, or GC pairs in the lower stem of the FPN 5'IRE. In the ALAS2 5'IRE (Table 5), all of the mammalian species studied contain two AU, one GU, one GC, and one mismatched (CA or GA) pair in their lower stems, while the non-mammalian *Danio rerio* contains two AU, one GC, and two mismatched (AG and AA) pairs. Statistical analysis was not performed because data are available from only one of the non-mammalian species. In the HIF2 $\alpha$  5'IREs (Table 6), the 5'IRE structure was only identified in humans among the mammalian species and contains three AU, one GC, and one mismatched AC pair in its lower stem, matching that of three non-mammalian species (*Danio rerio*, *Xenopus laevis*, and *Xenopus tropicalis*). The non-mammalian *Ictalurus punctatus* contains two AU, one GU, one GC, and one mismatched AC pair in its lower stem. Statistical analysis was not performed because data are available from only one of the mammalian species.

**Table 1. Characteristics of the 5'IRE stems in ferritin heavy chain (FTH) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the UGC/C internal loops are capitalized.**

Species	GenBank Accession No.	5'IRE Sequence (5'→3')	# of AU Pairs in the 5'IRE Stem	# of GU Pairs in the 5'IRE Stem	# of GC Pairs in the 5'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_002032	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	3 <u>2</u>	1	1 <u>2</u>	-6.70
<i>Pongo abelli</i>	NM_001132636	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	3 <u>2</u>	1	1 <u>2</u>	-6.70
<i>Macaca mulatta</i>	NM_001195380	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	3 <u>2</u>	1	1 <u>2</u>	-6.70
<i>Sus scrofa</i>	NM_213975	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	3 <u>2</u>	1	1 <u>2</u>	-6.70
<i>Canis lupus familiaris</i>	NM_001003080	<u>gu</u> u <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	3 <u>2</u>	1	1 <u>3</u>	-8.70
<i>Bos taurus</i>	NM_174062	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	4 <u>2</u>	0	1 <u>2</u>	-7.20
<i>Mus musculus</i>	NM_010239	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	4 <u>2</u>	0	1 <u>2</u>	-7.20
<i>Ovis aries</i>	NM_001009786	uu <u>cc</u> UGC <b>cuuca</b> acagugcu <u>ugga</u> Cggaac	4 <u>2</u>	0	1 <u>2</u>	-7.20
<i>Rattus norvegicus</i>	NM_012848	(partial IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Equus caballus</i>	NM_001252054	(no IRE sequence in the very short 5'UTR)	–	–	–	N/A
<i>Danio rerio</i>	NM_131585	<u>uu</u> accUGC <b>cuuca</b> acagugcu <u>ugga</u> Cggcaa	4 <u>2</u>	0	1 <u>2</u>	-6.80
<i>Salmo salar</i>	NM_001139722	<u>uu</u> accUGC <b>cuuca</b> acagugcu <u>ugga</u> Cggcaa	4 <u>2</u>	0	1 <u>2</u>	-6.80
<i>Xenopus laevis</i>	NM_001096738 NM_001086111 NM_001090588	<u>gu</u> u <u>cc</u> UGC <b>cuuca</b> acagugu <u>ugga</u> Cggaac	4 <u>2</u>	0 <u>1</u>	1 <u>2</u>	-8.40
<i>Xenopus tropicalis</i>	NM_203677	<b>Cu</b> ucaacagugu <u>ugga</u> Cggaac (partial IRE sequence in the 5'UTR)	4	0	1	N/A
<i>Gallus gallus</i>	NM_205086	<u>gu</u> u <u>cc</u> UGC <b>gu</b> ucaacagugcu <u>ugga</u> Cggaac	2 <u>2</u>	1	1 <u>3</u>	-7.30
<i>Anas platyrhynchos</i>	NM_001310377	<u>gu</u> u <u>cc</u> UGC <b>gu</b> ucaacagugcu <u>ugga</u> Cggaac	2 <u>2</u>	1	1 <u>3</u>	-7.30

**Table 2. Characteristics of the 5'IRE stems in ferritin light chain (FTL) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the UGC/C internal loops are capitalized.**

Species	GenBank Accession No.	5'IRE Sequence (5'→3')	# of AU Pairs in the 5'IRE Stem	# of GU Pairs in the 5'IRE Stem	# of GC Pairs in the 5'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_000146	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Pongo abelli</i>	NM_001133378	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Macaca mulatta</i>	NM_001261207	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Macaca fascicularis</i>	NM_001283240	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Sus scrofa</i>	NM_001244131	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Bos taurus</i>	NM_174792	uc <u>cu</u> UGC <b>cu</b> caacagug <u>cu</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Canis lupus familiaris</i>	NM_001024636	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Heterocephalus glaber</i>	NM_001279866	uc <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.30
<i>Rattus norvegicus</i>	NM_022500	ua <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> ggaCggaac	<b>3</b> <u>1</u>	<b>1</b> <u>1</u>	<b>1</b> <u>1</u>	-5.10
<i>Mus musculus</i>	NM_010240	ug <u>cu</u> UGC <b>cu</b> caacagugu <u>u</u> gaaCggaac	<b>4</b> <u>1</u>	<b>0</b> <u>1</u>	<b>1</b> <u>1</u>	-5.60
<i>Cavia porcellus</i>	NM_001172858	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Oryctolagus cuniculus</i>	NM_001101688	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Equus caballus</i>	NM_001114540	(the 5'UTR sequence is not available)	–	–	–	N/A
<i>Felis catus</i>	NM_001048150	(the 5'UTR sequence is not available)	–	–	–	N/A
<i>Ailuropoda melanoleuca</i>	NM_001304921	(no IRE sequence in the very short 5'UTR)	–	–	–	N/A
<i>Tursiops truncatus</i>	NM_001280630	(no IRE sequence in the very short 5'UTR)	–	–	–	N/A
<i>Xenopus laevis</i>	NM_001086458	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Gallus gallus</i>	NM_204383	(the 5'UTR sequence is not available)	–	–	–	N/A

**Table 3. Characteristics of the 5'IRE stems in mitochondrial aconitase 2 (ACO2) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.**

Species	GenBank Accession No.	5'IRE Sequence (5'→3')	# of AU Pairs in the 5'IRE Stem	# of GU Pairs in the 5'IRE Stem	# of GC Pairs in the 5'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_001098	cucau <b>Cuu</b> ugucagugcaca <u>aaa</u> uggc	<b>4</b> <u>2</u>	<b>0</b> <u>1</u>	<b>1</b> <u>1</u>	-2.80
<i>Macaca mulatta</i>	NM_001261164	cucau <b>Cuu</b> ugucagugcaca <u>aaa</u> uggc	<b>4</b> <u>2</u>	<b>0</b> <u>1</u>	<b>1</b> <u>1</u>	-2.80
<i>Sus scrofa</i>	NM_213954	cucau <b>Cuu</b> ugucagugcaca <u>aaa</u> uggc	<b>4</b> <u>2</u>	<b>0</b> <u>1</u>	<b>1</b> <u>1</u>	-2.80
<i>Bos taurus</i>	NM_173977	cucau <b>Cuu</b> ugucagugcaca <u>aaa</u> uggc	<b>4</b> <u>2</u>	<b>0</b> <u>1</u>	<b>1</b> <u>1</u>	-2.80
<i>Mus musculus</i>	NM_080633	cucau <b>Cuu</b> ugucagugcaca <u>aaa</u> uggc	<b>4</b> <u>2</u>	<b>0</b> <u>1</u>	<b>1</b> <u>1</u>	-2.80
<i>Heterocephalus glaber</i>	NM_001308662	<b>uu</b> ugucagugcaca <u>aaa</u> uggcg (partial IRE sequence in the 5'UTR)	<b>4</b>	<b>0</b>	<b>1</b>	N/A
<i>Rattus norvegicus</i>	NM_024398	(no IRE sequence in the very short 5'UTR)	–	–	–	N/A
<i>Gallus gallus</i>	NM_204188	au <u>uu</u> C <b>uu</b> ucaguguc <u>aa</u> gauc <u>uc</u> g	<b>3</b> <u>1</u>	<b>0</b>	<b>1</b>	-0.30

(Table 3) contd....

<i>Danio rerio</i>	NM_198908	(no IRE sequence in the very short 5'UTR)	-	-	-	N/A
<i>Xenopus laevis</i>	NM_001092794	(no IRE sequence in the very short 5'UTR)	-	-	-	N/A

**Table 4. Characteristics of the 5'IRE stems in ferroportin (FPN, also called solute carrier family 40 member 1, SLC40A1) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.**

Species	GenBank Accession No.	5'IRE Sequence (5'→3')	# of AU Pairs in the 5'IRE Stem	# of GU Pairs in the 5'IRE Stem	# of GC Pairs in the 5'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_014585	<u>aacuu</u> <b>Cagcu</b> acagugu <b>uagcu</b> <u>aaguu</u>	<b>3</b> <u>4</u>	<b>0</b>	<u>2</u> <u>1</u>	-10.20
<i>Pongo abelii</i>	NM_001132161	<u>aacuu</u> <b>Cagct</b> acagugu <b>uagcu</b> <u>aaguu</u>	<b>3</b> <u>4</u>	<b>0</b>	<u>2</u> <u>1</u>	-10.20
<i>Mus musculus</i>	NM_016917	<u>aacuu</u> <b>Cagcu</b> acagugu <b>uagcu</b> <u>aaguu</u>	<b>3</b> <u>4</u>	<b>0</b>	<u>2</u> <u>1</u>	-10.20
<i>Rattus norvegicus</i>	NM_133315	<u>aacuu</u> <b>Cagcu</b> acagugu <b>uagcu</b> <u>aaguu</u>	<b>3</b> <u>4</u>	<b>0</b>	<u>2</u> <u>1</u>	-10.20
<i>Bos taurus</i>	NM_001077970	(no IRE sequence in the 5'UTR)	-	-	-	N/A
<i>Danio rerio</i>	NM_131629	<u>gacuu</u> <b>Cagcu</b> acagug <b>auagcu</b> <u>aaguu</u>	<b>3</b> <u>3</u>	<b>0</b> <u>1</u>	<u>2</u> <u>1</u>	-8.80
<i>Gallus gallus</i>	NM_001012913	<u>gacuu</u> <b>Cagcu</b> acagug <b>cuagcu</b> <u>aaguc</u>	<b>3</b> <u>3</u>	<b>0</b>	<u>2</u> <u>2</u>	-11.10
<i>Xenopus laevis</i>	NM_001093357	(no IRE sequence in the 5'UTR)	-	-	-	N/A
<i>Xenopus tropicalis</i>	NM_001097277	(no IRE sequence in the very short 5'UTR)	-	-	-	N/A

**Table 5. Characteristics of the 5'IRE stems in erythroid aminolevulinate synthase 2 (ALAS2) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.**

Species	GenBank Accession No.	5'IRE Sequence (5'→3')	# of AU Pairs in the 5'IRE Stem	# of GU Pairs in the 5'IRE Stem	# of GC Pairs in the 5'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_000032	<u>ucguu</u> <b>Cgucc</b> acagug <b>cagggca</b> <u>acag</u>	<b>1</b> <u>2</u>	<b>1</b> <u>1</u>	<b>3</b> <u>1</u>	-7.00
<i>Pongo abelii</i>	NM_001134158	<u>ucguu</u> <b>Cgucc</b> acagug <b>cagggca</b> <u>acag</u>	<b>1</b> <u>2</u>	<b>1</b> <u>1</u>	<b>3</b> <u>1</u>	-7.00
<i>Bos taurus</i>	NM_001035103	<u>ucguu</u> <b>Cgucc</b> acagug <b>cagggca</b> <u>acag</u>	<b>1</b> <u>2</u>	<b>1</b> <u>1</u>	<b>3</b> <u>1</u>	-7.00
<i>Mus musculus</i>	NM_009653	<u>ugguu</u> <b>Cgucc</b> acagug <b>cagggca</b> <u>acag</u>	<b>1</b> <u>2</u>	<b>1</b> <u>1</u>	<b>3</b> <u>1</u>	-6.90
<i>Rattus norvegicus</i>	NM_013197	(no IRE sequence in the very short 5'UTR)	-	-	-	N/A
<i>Danio rerio</i>	NM_131682	<u>aaguu</u> <b>Cgucc</b> acagug <b>cagguca</b> <u>acag</u>	<b>1</b> <u>2</u>	<b>0</b>	<b>3</b> <u>1</u>	-3.60
<i>Xenopus laevis</i>	NM_001094030	(no IRE sequence in the 5'UTR)	-	-	-	N/A
<i>Xenopus tropicalis</i>	NM_001006925	(partial IRE sequence in the 5'UTR)	-	-	-	N/A
<i>Gallus gallus</i>	NM_001018012	(no IRE sequence in the 5'UTR)	-	-	-	N/A

**Table 6. Characteristics of the 5'IRE stems in the hypoxia-inducible transcription factor 2α (HIF2α, also called endothelial PAS domain protein 1, EPAS1) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.**

Species	GenBank Accession No.	5'IRE Sequence (5'→3')	# of AU Pairs in the 5'IRE Stem	# of GU Pairs in the 5'IRE Stem	# of GC Pairs in the 5'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
---------	-----------------------	------------------------	---------------------------------	---------------------------------	---------------------------------	---

(Table 6) contd....

<i>Homo sapiens</i>	NM_001430	<u>acaau</u> <b>Ccucgg</b> caguguccagagacugu	<u>1</u> 3	0	<u>4</u> 1	-4.80
<i>Sus scrofa</i>	NM_001097420	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Bos taurus</i>	NM_174725	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Mus musculus</i>	NM_010137	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Rattus norvegicus</i>	NM_023090	(no IRE sequence in the 5'UTR)	–	–	–	N/A
<i>Ictalurus punctatus</i>	NM_001350107	<u>acgau</u> <b>Ccucgg</b> caguguuucagagacugu	<u>1</u> 2	<u>1</u> 1	<u>3</u> 1	-5.30
<i>Danio rerio</i>	NM_001039806	<u>acaau</u> <b>Ccucag</b> caguguuucagagacugu	<u>2</u> 3	1	<u>2</u> 1	-4.70
<i>Xenopus laevis</i>	NM_001092249	<u>acaau</u> <b>Ccucag</b> cagugaccugagacugu	<u>2</u> 3	0	<u>3</u> 1	-4.90
<i>Xenopus tropicalis</i>	NM_001005647	<u>acaau</u> <b>Ccucag</b> cagugcccugagacugu	<u>2</u> 3	0	<u>3</u> 1	-4.90
<i>Gallus gallus</i>	NM_204807	(no IRE sequence in the very short 5'UTR)	–	–	–	N/A

In the CDC14A 3'IRE (Table 7), two of the mammalian species studied (*Homo sapiens* and *Rattus norvegicus*) contain four AU pairs and one mismatched UU pair in their lower stems, while *Mus musculus* has one AU, one GU, and three mismatched pairs. No non-mammalian species contains a complete IRE sequence in its CDC14A mRNA, so a statistical analysis was not conducted. In the DMT1 3'IRE (Table 8), three of the mammalian species studied (*Homo sapiens*, *Mus musculus*, and *Macaca fascicularis*) contain two AU, one GU, and two GC pairs in their lower stems. The *Rattus norvegicus* mRNA has two 3'IREs; one contains two AU, one GU, and two GC pairs in its lower stem; the other contains three GU and

two mismatched pairs. No non-mammalian species has a complete IRE sequence in its DMT1 mRNA, so a statistical analysis was not conducted. Of the three 3'IREs in mammalian TFRC mRNAs (Table 9), there are two major categories of base-pair arrangements in their lower stems: 1) five AU pairs, or 2) three AU and two GU pairs. The exceptions are *Mus musculus* and *Rattus norvegicus*, both of which have one of their three 3'IREs that contains four AU and one GU pair in the lower stem. No non-mammalian species has a complete IRE sequence in its TFRC mRNA, so a statistical analysis was not conducted.

**Table 7. Characteristics of the 3'IRE stems in cell division cycle 14A (CDC14A) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.**

Species	GenBank Accession No.	3'IRE Sequence (5'→3')	# of AU Pairs in the 3'IRE Stem	# of GU Pairs in the 3'IRE Stem	# of GC Pairs in the 3'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_001319210	<u>auuuu</u> <b>Caugu</b> acagugu <u>uacau</u> <u>uuau</u>	<u>4</u> 4	0	1	-5.00
<i>Rattus norvegicus</i>	NM_001134856	<u>auuuu</u> <b>Caugu</b> acagugu <u>uacau</u> <u>uuau</u>	<u>4</u> 4	0	1	-5.00
<i>Mus musculus</i>	NM_001080818 NM_001173553	<u>auuga</u> <b>Caugu</b> acagugu <u>uacacau</u> <u>ua</u>	<u>3</u> 1	<u>0</u> 1	1	-3.90
<i>Pan troglodytes</i>	NM_001280195	(no IRE sequence in the short 3'UTR)	–	–	–	N/A
<i>Macaca fascicularis</i>	NM_001319384	(no IRE sequence in the short 3'UTR)	–	–	–	N/A
<i>Gallus gallus</i>	NM_001177736	(the 3'UTR sequence is not available)	–	–	–	N/A

**Table 8. Characteristics of the 3'IRE stems in divalent metal transporter 1 (DMT1, also called solute carrier family 11 member 2, SLC11A2) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.**

Species	GenBank Accession No.	3'IRE Sequence (5'→3')	# of AU Pairs in the 3'IRE Stem	# of GU Pairs in the 3'IRE Stem	# of GC Pairs in the 3'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_001174128 NM_001174129 NM_001174130	<u>gccau</u> <b>Ca</b> gagcagugugu <u>uucua</u> <u>uggu</u>	<u>2</u> 2	<u>1</u> 1	<u>2</u> 2	-6.10

(Table 8) contd.....

<i>Mus musculus</i>	NM_001146161 NM_001356952	<u>gccauCagagccaguguguuuc<u>u</u>auggu</u>	<u>2</u> <u>2</u>	<u>1</u> <u>1</u>	<u>2</u> <u>2</u>	-6.10
<i>Rattus norvegicus</i>	NM_013173	<u>gccauCagagccaguguguuuc<u>u</u>auggu</u>	<u>2</u> <u>2</u>	<u>1</u> <u>1</u>	<u>2</u> <u>2</u>	-6.10
		<u>auguuCg<u>uuu</u>acagugauagacgguuc</u>	<u>3</u>	<u>1</u> <u>3</u>	<u>1</u>	-5.80
<i>Macaca fascicularis</i>	NM_001284816	<u>gccauCagagccaguguguuuc<u>u</u>auggu</u>	<u>2</u> <u>2</u>	<u>1</u> <u>1</u>	<u>2</u> <u>2</u>	-6.10
<i>Bubalus bubalis</i>	NM_001290899	(partial IRE sequence in the very short 3'UTR)	–	–	–	N/A
<i>Macaca mulatta</i>	NM_001266689	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Bos taurus</i>	NM_001101103	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Sus scrofa</i>	NM_001128440	(no IRE sequence in the very short 3'UTR)	–	–	–	N/A
<i>Danio rerio</i>	NM_001040370	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Xenopus tropicalis</i>	NM_001123466	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Gallus gallus</i>	NM_001128102	(no IRE sequence in the short 3'UTR)	–	–	–	N/A

Table 9. Characteristics of the 3'IRE stems in transferrin receptor (TFRC) mRNAs of selected animals. The base pairs in the upper stem are shown in bold; the base pairs in the lower stem are underlined; the C bulges are capitalized.

Species	GenBank Accession No.	3'IRE Sequence (5'→3')	# of AU Pairs in the 3'IRE Stem	# of GU Pairs in the 3'IRE Stem	# of GC Pairs in the 3'IRE Stem	Predicted Folding Free Energy of the RNA (kcal/mol)
<i>Homo sapiens</i>	NM_003234 Transcript variant 1  NM_001313966 Transcript variant 4	<u>auuauC<b>gga</b>agcagugccu<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>5</u>	<u>0</u>	<u>3</u>	-6.20
		<u>auuauC<b>ggg</b>agcagugcu<u>u</u>cc<u>u</u>auau</u>	<u>1</u> <u>5</u>	<u>1</u>	<u>3</u>	-5.50
		<u>uguauC<b>gg</b>agacagugau<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>3</u>	<u>0</u> <u>2</u>	<u>3</u>	-9.00
<i>Homo sapiens</i>	NM_001128148 Transcript variant 2	<u>auuauC<b>gga</b>agcagugccu<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>5</u>	<u>0</u>	<u>3</u>	-6.20
		<u>auuauC<b>ggg</b>agcagugcu<u>u</u>cc<u>u</u>auau</u>	<u>1</u> <u>5</u>	<u>1</u>	<u>3</u>	-5.50
		<u>auuauC<b>ggg</b>aacagugu<u>u</u>cc<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>5</u>	<u>0</u>	<u>3</u>	-10.30
<i>Homo sapiens</i>	NM_001313965 Transcript variant 3	<u>auuauC<b>ggg</b>agcagugcu<u>u</u>cc<u>u</u>auau</u>	<u>1</u> <u>5</u>	<u>1</u>	<u>3</u>	-5.50
		<u>uguauC<b>gg</b>agacagugau<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>3</u>	<u>0</u> <u>2</u>	<u>3</u>	-9.00
		<u>auuauC<b>ggg</b>aacagugu<u>u</u>cc<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>5</u>	<u>0</u>	<u>3</u>	-10.30
<i>Pongo abelii</i>	NM_001131591	<u>auuauC<b>gga</b>agcagugccu<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>5</u>	<u>0</u>	<u>3</u>	-6.20
		<u>uguauC<b>gg</b>agacagugau<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>3</u>	<u>0</u> <u>2</u>	<u>3</u>	-9.00
		<u>auuauC<b>ggg</b>aacagugu<u>u</u>cc<u>u</u>cc<u>u</u>auau</u>	<u>2</u> <u>5</u>	<u>0</u>	<u>3</u>	-10.30

(Table 9) contd....

<i>Mus musculus</i>	NM_011638 Transcript variant 1	<u>uuu</u> auC <b>ggaag</b> cagugcc <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-6.20
		<u>uuu</u> auC <b>ggag</b> acagugau <u>cu</u> cc <u>uu</u> auau	2 4	0 1	3	-8.90
		<u>uuu</u> auC <b>gggaac</b> cagugu <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-10.30
<i>Mus musculus</i>	NM_001357298 Transcript variant 2	<u>uuu</u> auC <b>ggaag</b> cagugcc <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-6.20
		<u>uuu</u> auC <b>gggag</b> cagug <u>cu</u> cc <u>uu</u> auau	1 5	1	3	-5.50
		<u>uuu</u> auC <b>gggaac</b> cagugu <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-10.30
<i>Rattus norvegicus</i>	NM_022712	<u>uuu</u> auC <b>ggaag</b> cagugcc <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-6.20
		<u>uuu</u> auC <b>gggag</b> cagug <u>cu</u> cc <u>uu</u> auau	1 5	1	3	-5.50
		<u>uuu</u> auC <b>ggag</b> acagug <u>ac</u> cc <u>uu</u> auau	1 4	0 1	3	-6.10
<i>Heterocephalus glaber</i>	NM_001267852	<u>uuu</u> auC <b>ggaag</b> cagugcc <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-6.20
		<u>uuu</u> auC <b>gggag</b> cagug <u>cu</u> cc <u>uu</u> auau	1 5	1	3	-5.50
		<u>uuu</u> auC <b>gggaac</b> cagugu <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-10.30
<i>Cavia porcellus</i>	NM_001251822	<u>uuu</u> auC <b>ggaag</b> cagugcc <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-6.20
		<u>uuu</u> auC <b>aggag</b> cagug <u>cu</u> cc <u>uu</u> auau	1 5	1	2	-2.00
		<u>uuu</u> auC <b>gggaac</b> cagugu <u>uu</u> cc <u>uu</u> auau	2 5	0	3	-10.30
<i>Cricetulus griseus</i>	NM_001246819	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Sus scrofa</i>	NM_214001	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Bos taurus</i>	NM_001206577	(no IRE sequence in the 3'UTR)	–	–	–	N/A
<i>Macaca mulatta</i>	NM_001257303	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Callithrix jacchus</i>	NM_001301847	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Mustela putorius furo</i>	NM_001310181	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Felis catus</i>	NM_001009312	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Canis lupus familiaris</i>	NM_001003111	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Equus caballus</i>	NM_001081913	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Danio rerio</i>	NM_001009917	(the 3'UTR sequence is not available)	–	–	–	N/A
<i>Gallus gallus</i>	NM_205256	(no IRE sequence in the 3'UTR)	–	–	–	N/A

In canonical Watson-Crick RNA base pairing, A forms a base pair with U through two hydrogen bonds; G forms a base pair with C through three hydrogen bonds. Therefore, an RNA with high GC content is more thermodynamically stable than one with low GC content. The non-Watson-Crick wobble pair GU is a common element in the secondary structure of RNA, where G forms a pair with U through two hydrogen bonds. The thermodynamic stability of a GU pair is less than that of a GC pair but comparable to that of an AU pair [21, 22]. Since the glycosidic bond angles (*i.e.*, the angle between the base and C1' sugar atom) in a GU wobble pair are different from those in an AU or GC pair, an RNA stem containing a GU pair is

conformationally softer because the backbone is more easily distorted or altered at the site of a GU pair.

Among the IREs studied from the transcripts of nine different genes, three 5'IREs (FTH, FTL, and ACO2) and one 3'IRE (CDC14A) have lower GC content (only one GC pair) in their upper stems; the FPN 5'IRE and DMT1 3'IRE have medium GC content (two GC pairs); the ALAS2 5'IRE and TFRC 3'IRE have higher GC content (three GC pairs) in their upper stems. The 5'IRE in the human HIF2 $\alpha$  mRNA is a special case. It contains the highest GC content (four GC pairs in its upper stem) among the IREs investigated. The non-

mammalian species of HIF2 $\alpha$  mRNA, however, have only two or three GC pairs in their 5'IRE upper stems. Overall, we found that the base-pair compositions of the upper stems of IREs are not highly conserved among mRNAs from the genes we investigated.

Data analysis of the IREs indicates that, in general, the lower stems contain fewer GC pairs than the upper stems and often have mismatched pairs. The only exception is the FTH 5'IRE, which contains two or three GC pairs in its lower stem, and only one GC pair in its upper stem (Table 1). The more tightly bound (*i.e.*, stiffer) lower stem may be necessary for the existence of a UGC/C internal loop instead of the single C bulge. However, another UGC/C-containing 5'IRE, FTL, has only one GC pair but two mismatched (UC and CA/AA/GA) pairs in its lower stem. Further studies are needed to elucidate this seeming contradiction.

Statistical analyses indicate that there are no significant differences ( $p > 0.05$ ) in the IRE structures between mammalian and non-mammalian species for the FTH 5'IRE or FPN 5'IRE. In addition, there are no significant differences ( $p > 0.05$ ) in the folding free energies between mammalian and non-mammalian species for the FTH 5'IRE or FPN 5'IRE. Statistical analyses were not performed for the other IREs due to the lack of sequence data for the mRNAs of either the non-mammalian or mammalian species.

Table 10 lists the results of comparing the folding free energies of mammalian IREs in the following groups: 1) the FTH 5'IRE and FTL 5'IRE both contain a UGC/C internal loop, but their folding free energies are significantly different; 2) the ACO2 5'IRE, FPN 5'IRE, ALAS2 5'IRE, and HIF2 $\alpha$  5'IRE all contain a C bulge, but their folding free energies are significantly different. Note that the HIF2 $\alpha$  5'IRE was not compared with other C-bulge-containing 5'IREs because data are only available for the HIF2 $\alpha$  mRNA of one species; 3) between UGC/C-containing 5'IREs and C-bulge-containing 5'IREs, only the FTH 5'IRE and ALAS2 5'IRE comparison was not statistically significant. Nonetheless, the pooled analysis of the UGC/C-containing 5'IREs *versus* C-bulge-containing 5'IREs did not indicate a significant difference in their folding free energies.

As for the 3'IREs, 1) TFRC contains three 3'IREs, and a comparison of the folding free energies of each of its IREs to the others resulted in statistically significant results for all analyses except between the first and second IREs; 2) there was no significant difference between the folding free energies of the CDC12A 3'IRE and DMT1 3'IRE; 3) the folding free energies of the CDC12A 3'IRE and DMT1 3'IRE were separately compared with those of each TFRC 3'IRE, and with the pooled TFRC 3'IREs, with varying results (Table 10). Overall, the folding free energies are not significantly different between the pooled 5'IREs and the pooled 3'IREs.

**Table 10. Significance results of the minimum folding free energies of selected mammalian IREs.**

Group	Name of IREs whose Folding Free Energies were Compared	Is the Energy Difference Significant ( $p < 0.05$ )?
5'IREs containing a UGC/C internal loop	FTH 5'IRE vs FTL 5'IRE	Yes
5'IREs containing a C bulge	ACO2 5'IRE vs. FPN 5'IRE	Yes
	ACO2 5'IRE vs. ALAS2 5'IRE	Yes
	ACO2 5'IRE vs. HIF2 $\alpha$ 5'IRE	N/A
	FPN 5'IRE vs. ALAS2 5'IRE	Yes
	FPN 5'IRE vs. HIF2 $\alpha$ 5'IRE	N/A
	ALAS2 5'IRE vs. HIF2 $\alpha$ 5'IRE	N/A
5'IRE containing a UGC/C internal loop vs. 5'IRE containing a C bulge	FTH 5'IRE vs. ACO2 5'IRE	Yes
	FTH 5'IRE vs. FPN 5'IRE	Yes
	FTH 5'IRE vs. ALAS2 5'IRE	No
	FTH 5'IRE vs. HIF2 $\alpha$ 5'IRE	N/A
	FTL 5'IRE vs. ACO2 5'IRE	Yes
	FTL 5'IRE vs. FPN 5'IRE	Yes
	FTL 5'IRE vs. ALAS2 5'IRE	Yes
	FTL 5'IRE vs. HIF2 $\alpha$ 5'IRE	N/A
Pooled 5'IREs containing a UGC/C internal loop vs. Pooled 5'IREs containing a C bulge		No
TFRC 3'IREs	1 <sup>st</sup> TFRC 3'IRE vs. 2 <sup>nd</sup> TFRC 3'IRE	No
	1 <sup>st</sup> TFRC 3'IRE vs. 3 <sup>rd</sup> TFRC 3'IRE	Yes
	2 <sup>nd</sup> TFRC 3'IRE vs. 3 <sup>rd</sup> TFRC 3'IRE	Yes

(Table 10 ) contd.....

3'IREs	CDC14A vs. DMT1	No
	CDC14A vs. 1 <sup>st</sup> TFRC 3'IRE	No
	CDC14A vs. 2 <sup>nd</sup> TFRC 3'IRE	No
	CDC14A vs. 3 <sup>rd</sup> TFRC 3'IRE	Yes
	CDC14A vs. all of the TFRC 3'IREs	Yes
	DMT1 vs. 1 <sup>st</sup> TFRC 3'IRE	No
	DMT1 vs. 2 <sup>nd</sup> TFRC 3'IRE	No
	DMT1 vs. 3 <sup>rd</sup> TFRC 3'IRE	Yes
	DMT1 vs. all of the TFRC 3'IREs	Yes
Pooled 5'IREs vs. Pooled 3'IREs		No

## CONCLUSION

In summary, the base pairs within the upper and lower stems of IREs are not highly conserved among the mRNAs investigated for this study. Both AU-rich and GC-rich upper stems exist. The lower stems, in general, contain fewer GC pairs than the upper stems. One exception is the UGC/C-containing FTH 5'IRE, whose lower stem includes more GC content than its upper stem. No statistically significant differences were found in the IRE structures or the folding free energies when comparing either the FTH 5'IRE or FPN 5'IRE of mammalian *versus* non-mammalian species. In addition, there were no overall significant differences between the folding free energies of UGC/C-containing 5'IREs and C-bulge-containing 5'IREs, or between 5'IREs and 3'IREs. Future studies may focus on investigating whether the evolutionary characteristics of the IRE stems in animal mRNAs differentially fine-tune the IRE/IRP interactions among different mRNAs to maintain the balance of cellular iron metabolism and whether evolutionary processes drive the base-pair composition of the upper and lower stems of IREs toward any particular outcome (*e.g.*, AU-rich, GC-rich, or a balanced composition).

## LIST OF ABBREVIATIONS

<b>IRE</b>	= Iron-Responsive Element
<b>IRP</b>	= Iron-Regulatory Protein
<b>mRNA</b>	= Messenger RNA
<b>5'UTR</b>	= 5' Untranslated Region
<b>3'UTR</b>	= 3' Untranslated Region
<b>FTH</b>	= Ferritin Heavy Chain
<b>FTL</b>	= Ferritin Light Chain
<b>FPN</b>	= Ferroportin
<b>ALAS2</b>	= Erythroid Aminolevulinate Synthase 2
<b>ACO2</b>	= Mitochondrial Aconitase 2
<b>HIF2<math>\alpha</math></b>	= Hypoxia-Inducible Transcription Factor 2 $\alpha$
<b>TFRC</b>	= Transferrin Receptor
<b>DMT1</b>	= Divalent Metal Transporter 1
<b>CDC14A</b>	= Cell Division Cycle 14A
<b>SLC40A1</b>	= Solute Carrier Family 40 Member 1
<b>EPAS1</b>	= Endothelial PAS Domain Protein 1
<b>SLC11A2</b>	= Solute Carrier Family 11 Member 2

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

## HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

## CONSENT FOR PUBLICATION

Not applicable.

## AVAILABILITY OF DATA AND MATERIALS

All data generated and analyzed during this study are included in this published article.

## FUNDING

This work is supported by the National Science Foundation under Awards No. EPS-1003907 and OIA-1458952.

## CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

## ACKNOWLEDGEMENTS

The authors are grateful to the financial support received from the National Science Foundation (United States). Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- [1] Muckenthaler, M.U.; Galy, B.; Hentze, M.W. Systemic iron homeostasis and the iron-responsive element/iron-regulatory protein (IRE/IRP) regulatory network. *Annu. Rev. Nutr.*, **2008**, *28*, 197-213. [<http://dx.doi.org/10.1146/annurev.nutr.28.061807.155521>] [PMID: 18489257]
- [2] Wilkinson, N.; Pantopoulos, K. The IRP/IRE system in vivo: Insights from mouse models. *Front. Pharmacol.*, **2014**, *5*, 176. [<http://dx.doi.org/10.3389/fphar.2014.00176>] [PMID: 25120486]
- [3] Anderson, C.P.; Shen, M.; Eisenstein, R.S.; Leibold, E.A. Mammalian iron metabolism and its control by iron regulatory proteins. *Biochim. Biophys. Acta*, **2012**, *1823*(9), 1468-1483. [<http://dx.doi.org/10.1016/j.bbamcr.2012.05.010>] [PMID: 22610083]
- [4] Khan, M.A.; Walden, W.E.; Goss, D.J.; Theil, E.C. Direct Fe<sup>2+</sup> sensing by iron-responsive messenger RNA: Repressor complexes weakens binding. *J. Biol. Chem.*, **2009**, *284*(44), 30122-30128. [<http://dx.doi.org/10.1074/jbc.M109.041061>] [PMID: 19720833]
- [5] Goforth, J.B.; Anderson, S.A.; Nizzi, C.P.; Eisenstein, R.S. Multiple determinants within iron-responsive elements dictate iron regulatory

- protein binding and regulatory hierarchy. *RNA*, **2010**, *16*(1), 154-169. [http://dx.doi.org/10.1261/rna.1857210] [PMID: 19939970]
- [6] Ma, J.; Haldar, S.; Khan, M.A.; Sharma, S.D.; Merrick, W.C.; Theil, E.C.; Goss, D.J. Fe<sup>2+</sup> binds iron responsive element-RNA, selectively changing protein-binding affinities and regulating mRNA repression and activation. *Proc. Natl. Acad. Sci. USA*, **2012**, *109*(22), 8417-8422. [http://dx.doi.org/10.1073/pnas.1120045109] [PMID: 22586079]
- [7] Khan, M.A.; Ma, J.; Walden, W.E.; Merrick, W.C.; Theil, E.C.; Goss, D.J. Rapid kinetics of iron responsive element (IRE) RNA/iron regulatory protein 1 and IRE-RNA/eIF4F complexes respond differently to metal ions. *Nucleic Acids Res.*, **2014**, *42*(10), 6567-6577. [http://dx.doi.org/10.1093/nar/gku248] [PMID: 24728987]
- [8] Khan, M.A.; Walden, W.E.; Theil, E.C.; Goss, D.J. Thermodynamic and kinetic analyses of iron response element (IRE)-mRNA binding to iron regulatory protein, IRP1. *Sci. Rep.*, **2017**, *7*(1), 8532. [http://dx.doi.org/10.1038/s41598-017-09093-5] [PMID: 28819260]
- [9] Piccinelli, P.; Samuelsson, T. Evolution of the iron-responsive element. *RNA*, **2007**, *13*(7), 952-966. [http://dx.doi.org/10.1261/rna.464807] [PMID: 17513696]
- [10] Theil, E.C. The IRE (iron regulatory element) family: Structures which regulate mRNA translation or stability. *Biofactors*, **1993**, *4*(2), 87-93. [PMID: 8347279]
- [11] Goss, D.J.; Theil, E.C. Iron responsive mRNAs: A family of Fe<sup>2+</sup>-sensitive riboregulators. *Acc. Chem. Res.*, **2011**, *44*(12), 1320-1328. [http://dx.doi.org/10.1021/ar2001149] [PMID: 22026512]
- [12] Theil, E.C. Ferritin: The protein nanocage and iron biomineral in health and in disease. *Inorg. Chem.*, **2013**, *52*(21), 12223-12233. [http://dx.doi.org/10.1021/ic400484n] [PMID: 24102308]
- [13] Khan, M.A.; Malik, A.; Domashevskiy, A.V.; San, A.; Khan, J.M. Interaction of ferritin iron responsive element (IRE) mRNA with translation initiation factor eIF4F. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **2020**, *243*, 118776. [http://dx.doi.org/10.1016/j.saa.2020.118776] [PMID: 32829157]
- [14] Ke, Y.; Wu, J.; Leibold, E.A.; Walden, W.E.; Theil, E.C. Loops and bulge/loops in iron-responsive element isoforms influence iron regulatory protein binding. Fine-tuning of mRNA regulation? *J. Biol. Chem.*, **1998**, *273*(37), 23637-23640. [http://dx.doi.org/10.1074/jbc.273.37.23637] [PMID: 9726965]
- [15] Rogers, J.T.; Randall, J.D.; Cahill, C.M.; Eder, P.S.; Huang, X.; Gunshin, H.; Leiter, L.; McPhee, J.; Sarang, S.S.; Utsuki, T.; Greig, N.H.; Lahiri, D.K.; Tanzi, R.E.; Bush, A.I.; Giordano, T.; Gullans, S.R. An iron-responsive element type II in the 5'-untranslated region of the Alzheimer's amyloid precursor protein transcript. *J. Biol. Chem.*, **2002**, *277*(47), 45518-45528. [http://dx.doi.org/10.1074/jbc.M207435200] [PMID: 12198135]
- [16] Friedlich, A.L.; Tanzi, R.E.; Rogers, J.T. The 5'-untranslated region of Parkinson's disease alpha-synuclein messengerRNA contains a predicted iron responsive element. *Mol. Psychiatry*, **2007**, *12*(3), 222-223. [http://dx.doi.org/10.1038/sj.mp.4001937] [PMID: 17325711]
- [17] Cho, H.H.; Cahill, C.M.; Vanderburg, C.R.; Scherzer, C.R.; Wang, B.; Huang, X.; Rogers, J.T. Selective translational control of the Alzheimer amyloid precursor protein transcript by iron regulatory protein-1. *J. Biol. Chem.*, **2010**, *285*(41), 31217-31232. [http://dx.doi.org/10.1074/jbc.M110.149161] [PMID: 20558735]
- [18] Campillos, M.; Cases, I.; Hentze, M.W.; Sanchez, M. SIREs: Searching for iron-responsive elements. *Nucleic Acids Res.*, **2010**, *38*, W360-7. [PMID: 20460462]
- [19] Walden, W.E.; Selezneva, A.I.; Dupuy, J.; Volbeda, A.; Fontecilla-Camps, J.C.; Theil, E.C.; Volz, K. Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science*, **2006**, *314*(5807), 1903-1908. [http://dx.doi.org/10.1126/science.1133116] [PMID: 17185597]
- [20] Walden, W.E.; Selezneva, A.; Volz, K. Accommodating variety in iron-responsive elements: Crystal structure of transferrin receptor 1 B IRE bound to iron regulatory protein 1. *FEBS Lett.*, **2012**, *586*(1), 32-35. [http://dx.doi.org/10.1016/j.febslet.2011.11.018] [PMID: 22119729]
- [21] Varani, G.; McClain, W.H. The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep.*, **2000**, *1*(1), 18-23. [http://dx.doi.org/10.1093/embo-reports/kvd001] [PMID: 11256617]
- [22] Xu, D.; Landon, T.; Greenbaum, N.L.; Fenley, M.O. The electrostatic characteristics of G.U wobble base pairs. *Nucleic Acids Res.*, **2007**, *35*(11), 3836-3847. [http://dx.doi.org/10.1093/nar/gkm274] [PMID: 17526525]