© 1992 Oxford University Press

Nucleic Acids Research, Vol. 20, No. 1 33–39

Expression of active iron regulatory factor from a full-length human cDNA by *in vitro* transcription/translation

Harald Hirling, Alice Emery-Goodman, Nancy Thompson, Barbara Neupert, Christian Seiser and Lukas C.Kühn*
Swiss Institute for Experimental Cancer Research, Genetics Unit, CH-1066 Epalinges, Switzerland

Received November 11, 1991; Accepted December 6, 1991

ABSTRACT

Iron regulatory factor (IRF), also called iron responsive element-binding protein (IRE-BP), is a cytoplasmic RNA-binding protein which regulates post-transcriptionally transferrin receptor mRNA stability and ferritin mRNA translation. By using the polymerase chain reaction (PCR) and the sequence published by Rouault et al. (1990) a probe was derived which permitted the isolation of three human IRF cDNA clones. Hybridization to genomic DNA and mRNA, as well as sequencing data indicated a single copy gene of about 40 kb specifying a 4.0 kb mRNA that translates into a protein of 98,400 dalton. By *in vitro* transcription of a assembled IRF cDNA coupled to *in vitro* translation in a wheat germ extract, we obtained full sized IRF that bound specifically to a human ferritin IRE. *In vitro* translated IRF retained sensitivity to sulfhydryl oxidation by diamide and could be reactivated by β-mercaptoethanol in the same way as native placental IRF. An IRF deletion mutant shortened by 132 amino acids at the COOH-terminus was no longer able to bind to an IRE, indicating that this region of the protein plays a role in RNA recognition. Placental IRF has previously been shown to migrate as a doublet on SDS-polyacrylamide gels. After V8 protease digestion the heterogeneity was located in a 65/70 kDa NH_2-terminal doublet. The liberated 31 kDa COOH-terminal polypeptide was found to be homogeneous by amino acid sequencing supporting the conclusion of a single IRF gene.

INTRODUCTION

Coordinate regulation of key proteins in iron metabolism is one of the best studied cases for post-transcriptional control mechanisms of gene expression. In response to iron deprivation, the cytoplasmic stability of transferrin receptor mRNA is increased and ferritin mRNA translation inhibited (1–7). As a consequence, enhanced iron uptake and diminished iron storage tend to compensate for the lack of iron. This feedback regulation can be viewed as a protective mechanism that prevents nutritional starvation and permits the biosynthesis of essential iron- or heme-containing proteins. Under conditions of high iron supply, when cells may need to store excess iron in order to prevent adverse effects of iron overload, the regulatory balance is inversed: transferrin receptor mRNA decays more rapidly and ferritin translation is no longer inhibited.

The mechanisms affecting transferrin receptor mRNA stability and inhibiting ferritin mRNA translation are clearly distinct, but both require a highly conserved RNA structure, the iron responsive element (IRE) (3,5,6,8–10). Five IREs are located in the 3′ untranslated region of transferrin receptor mRNA and one IRE is close to the 5′ end of ferritin mRNA. Recently, additional IREs have been discovered near the 5′ end of aconitase and erythroid δ-aminolevulinic acid synthase mRNA (11–13). All these RNA structures represent specific binding sites for a cytoplasmic protein of about 95 kDa (3,9,14,15) called the iron regulatory factor (IRF) (3), also known as the IRE-binding protein (IRE-BP) (16) or ferritin repressor protein (FRP) (7). The total cellular IRF pool consists of RNA-bound molecules, an unbound active fraction, as well as the protein in an inactive state (17). Its IRE-binding affinity is modulated as a function of intracellular iron levels, and the fraction of active IRF molecules increases when iron concentration decreases (3,16,18). This post-translational change parallels precisely the *in vivo* accumulation of transferrin receptor mRNA (3), and the binding of IRF to the IRE directly prevents ribosomes from attaching to ferritin mRNA (7,19–21).

While there is no doubt that IRF activity is influenced by iron levels *in vivo*, some uncertainty prevails as to the mechanism by which this protein is modulated. Experiments with cultured cells strongly suggest a central role of chelatable iron that may directly bind to IRF. This hypothesis seems plausible since the recent isolation of an IRF cDNA clone (22) that has revealed sequence homology to pig heart aconitase (23,24). This enzyme of the citric acid cycle contains a [4Fe-4S] cluster capable of iron exchange (25). Experiments carried out *in vitro* have indicated, in addition, the involvement of redox reactions in IRF activation (16,26). Reducing agents, such as β-mercaptoethanol at unphysiologically high concentrations, activate IRF *in vitro* to the same extent as low iron levels would do *in vivo*. In contrast,

* To whom correspondence should be addressed

binding of IRF to an IRE is prevented by the sulfhydryl-oxidizing agent azodicarboxylic acid bis[dimethylamide] (diamide). Hemin, itself a compound capable of inactivating IRF *in vivo*, is a strong inhibitor of IRE-IRF interactions *in vitro* (27–29) and seems to exert its effect through a reversible oxidation of IRF (our unpublished observation). It remains unclear whether the *in vitro* activation or inactivation mimics a redox reaction that is allosterically facilitated *in vivo* by iron exchange; or whether, on the contrary, iron is directly involved in an intra- or intermolecular redox step which may represent a prerequisit for IRE-binding (16).

Further progress on the biochemistry of IRF will require the isolation of large amounts of IRF and the testing of IRF mutants. Thus far, IRF has been purified in relatively small quantities from liver and placenta (20,30,31), and its analysis has not yet revealed an iron-sulfur cluster. However, since isolation procedures were carried out in the presence of oxygen, where iron-sulfur clusters become unstable, it is possible that the cluster exists, but has remained undetected for technical reasons. It obviously would be of importance to express the protein *in vivo* or *in vitro* from its cDNA. In the present study, we isolated human IRF cDNA clones with a PCR probe based on a previous sequence determined by Rouault et al. (22). A reconstructed full-length cDNA has allowed us to obtain active IRF by *in vitro* transcription coupled to *in vitro* translation in a wheat germ extract.

MATERIALS AND METHODS

Cloning and Sequencing of IRF cDNA clones

Screening of libraries was carried out with a PCR-probe that was amplified from a human placenta cDNA library. The opposite oligonucleotide primers corresponded to position 210 to 235 and 446 to 472 in the sequence published by Rouault et al. (22). The reaction mixture contained 20 mM Tris-HCl pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, 100 μg/ml BSA, 200 μM dNTP and 1 U Vent™-polymerase (New England Biolabs, Beverly, MA). The amplification was carried out at 67°C for annealing and 72°C for the polymerase reaction. The amplified product was cloned into pGEM-3zf(−) (Promega, Madison, WI) and confirmed by sequencing. This probe was used to screen two different cDNA libraries, one from human placenta in λgt11 (Clontech, Palo Alto, CA) and the other one from the fibroblast cell line EK4 (32) in λgt10 (gift of Dr. R. Tyrrell, ISREC, Epalinges, Switzerland). In both cases about 4×10^5 phages were plated. Duplicate filters (GeneScreen Plus, Du Pont, Boston, MA) were hybridized at 42°C with a random-primed probe (3×10^5 cpm/ml) in 5×SSPE, 0.3% SDS and 1% milkpowder and washed after 15 h with 5×SSPE, 0.1% SDS at 52°C twice for 15 min. Positive plaques were picked and rescreened 4 times. Finally, among five positive clones with overlapping restriction sites, we analysed in further detail two from the placenta library and one from the fibroblast library. These clones were digested with EcoRI and the fragments from inserts subcloned into pGEM-3zf(−)

Sequencing was carried out according to the dideoxynucleotide chain termination method of Sanger et al. (33), using Sequenase version 2.0 (United States Biochemical, Cleveland, OH). Single-stranded DNA from subclones of EcoRI-fragments in pGEM was prepared by using the helper phage VCS13 (Stratagene, Heidelberg, Germany). The inserts of λIRF2 and λIRF4 were entirely sequenced, where necessary with the help of internal sequencing primers. The part of the sequence matching exactly the one of Rouault et al. (22) was verified only on one strand, but any mismatches and new sequences were ascertained on both strands.

Reconstruction of a full-length IRF cDNA clone

The overlapping inserts of λIRF2 and λIRF4 (Fig.1) were used to reassemble a larger cDNA clone. The 5′ EcoRI-DraIII fragment from λIRF2 (previously subcloned as an entire EcoRI-fragment into pGEM-3zf(−)) and the central DraIII-KpnI fragment from λIRF4 were ligated together into the EcoRI/KpnI sites of pGEM-7zf(−). This new clone was linearized with KpnI in order to insert a KpnI-KpnI fragment containing the 3′ KpnI-EcoRI region of λIRF4 and polylinker sequences between the EcoRI and KpnI sites of pGEM-3zf(−). This plasmid will be referred to as pGEM-hIRF.

Northern and Southern blot analysis

Total cytoplasmic RNA from the fibroblast cell line EK4 was isolated according to the procedure of Chomczynski and Sacchi (34). 25 μg of RNA was loaded on a 1.2% agarose gel containing 6.6% formaldehyde in 20 mM morpholinoethane sulfonic acid pH 7.0, 1 mM sodium acetate, 1 mM EDTA. After migration, the RNA was blotted onto a GeneScreen membrane (Du Pont, Boston, MA). A random-primed IRF cDNA probe (10^6 cpm/ml) from bases 143 to 3040 (Fig.2) was hybridized to the RNA in 5×SSPE, 100 mM sodium phosphate pH 6.5, 5×Denhardt's solution, 50% formamide, 10 mM EDTA, 1% sarcosyl and 125 μg/ml tRNA for 16 h at 42°C. Random priming was done with a kit from Boehringer (Mannheim, Germany) which yielded a specific activity of 0.5 to 1×10^9 cpm/μg DNA. The blot was washed in 6×SSC, 0.1% SDS for 30 min at room temperature, followed by one wash in 2×SSC, 0.1% SDS for 15 min at 55°C.

Genomic DNA from human placenta was prepared as previously described (35). 20 μg of DNA was digested with 50 U EcoRI or HindIII and loaded on a 0.8% agarose gel in 1×TBE. After migration, the gel was treated for 15 min first in 250 mM HCl and then for 15 min in 0.4 M NaOH. The DNA was transferred overnight to a nylon membrane (ZETA-PROBE, BioRad, Richmond, CA) according to the manufacturer's instructions. The membrane was sequencially hybridized with several random-primed probes. In each case hybridization was carried out in 0.5 M sodium phosphate pH 7.2, 7% SDS, 1 mM EDTA and 1% BSA at 68°C for 16 h with a probe concentration of 10^6 cpm/ml. The membrane was washed twice for 30 min at 65°C in 40 mM sodiumphosphate pH 7.2, 1 mM EDTA, 0.5% BSA and 5% SDS. Prior to re-hybridization the previous probe was stripped with 0.1×SSC and 0.5% SDS.

In vitro transcription and translation of IRF

The plasmid pGEM-hIRF, which contains the IRF full-length cDNA, was linearized by various restriction enzymes and transcribed *in vitro* with T7 polymerase (Promega) in the presence of $m^7G(5')ppp5'G$ (Pharmacia, Uppsala, Sweden) (36) according to the supplier's instructions. 100 ng of capped RNA was translated *in vitro* in 25 μl reaction volume using a wheat germ extract (Promega). The reaction was carried out according to the manufacturer's protocol in the presence of ^{35}S-methionine for SDS-polyacrylamide gel-analysis or unlabelled methionine for gel retardation assays. The length of translated polypeptides in amino acids (aa) depended on the restriction site at which pGEM-hIRF was linearized: 889aa (HindIII or ClaI), 757aa

(KpnI), 682aa (PvuII), 559aa (NdeI), 546aa (SmaI), 392aa (StyI). Radiolabelled translation products were analysed on a 8% SDS-polyacrylamide gel. The gel was fixed with 10% TCA and treated with enhancer (EN^{3}HANCE, Du Pont), prior to drying and exposure for autoradiography.

RNA-protein bandshift assays

To analyse the biological activity of *in vitro* translated IRF, a 5 μl aliquot of the *in vitro* translation reaction was incubated at room temperature with a molar excess of [^{32}P]CTP-labelled RNA transcript from plasmid pSPT-fer as previously decribed (3,14). This plasmid contains the sequence corresponding to the IRE from the 5′ untranslated region of human ferritin H chain mRNA. After 20 minutes, RNase T_1 (1 U) was added for 10 min, followed by an incubation with heparin (5 mg/ml final) for another 10 min. RNA-protein complexes were analysed on a 6% non-denaturing polyacrylamide gel as described (3). Where indicated, *in vitro* translated IRF was preincubated for 10 min with 1.5 mM diamide on ice or with 2% β-mercaptoethanol. For measurements of specificity, RNA-protein bandshifts were performed in the presence of an excess of unlabelled pSPT-fer RNA or 4F2 antigen heavy chain RNA (37).

IRF isolation, partial protease digestion and peptide sequencing

Affinity-purified IRF from human placenta (31) was digested with V8 protease or trypsin sequencing grade protease (Boehringer Mannheim). V8 protease was added to IRF at a ratio of 1:20 (w/w), and digestion was carried out in 25 mM NH_4HCO_3 pH 7.8 at room temperature. For trypsin, the protease to IRF ratio was 1:100 (w/w), and digestion was done in 100 mM Tris-HCl pH 8.5 at 37°C. Partially digested IRF was analysed on a 7.5% SDS-polyacrylamide gel. For peptide sequencing V8 protease digestion products were transferred electrophoretically onto Immobilon P transfer membrane (Millipore, Bedford, MA) in 10 mM 3-(cyclohexylamino)-1-propanesulfonic acid pH 11 and 10% methanol. After transfer, the membrane was stained with 0.5% Ponceau S (Sigma) in 1% acetic acid and washed several times with water. The bands corresponding to the digestion products were exised, rinsed with water, air dried and submitted to automated amino acid sequencing.

RESULTS AND DISCUSSION

Molecular cloning and structure of IRF cDNA clones

Previous purification of IRF from human placenta by affinity chromatography on an IRE-containing RNA-column has revealed two bands with molecular weights of about 95 and 100 kDa (31). Both proteins showed similar biochemical properties, and were capable of binding IREs from different sources *in vitro* (3,31). However, IRF purified from liver has been reported without such a heterogeneity and with a single molecular mass of 90 kDa (20, 30). Thus, there existed the possibility for tissue-specific variations among IRE-binding proteins. Particularly in view of the large number of mRNAs that contain IREs and are regulated by IRF, it was of interest to examine whether the observed IRF heterogeneity reflects distinct gene products or rather variants from a single gene. We, therefore, decided to isolate IRF cDNA clones from a human placenta library. For screening we amplified a PCR-probe encompassing bases 210 to 472 in the sequence published by Rouault et al. (22). This probe was chosen close to the 5′ end in order to select for rare full-length cDNAs. The

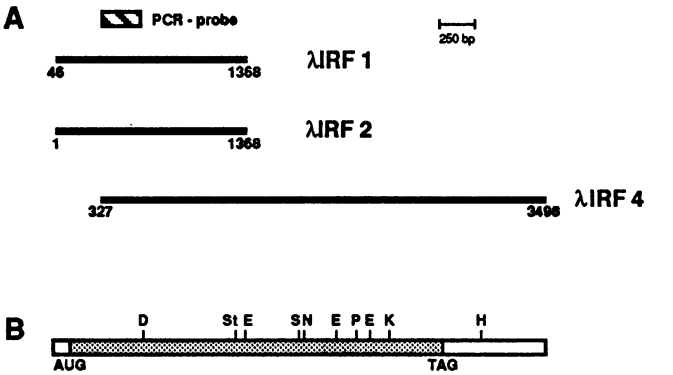

Figure 1. Isolation and assembly of human IRF cDNA clones. A PCR-probe was amplified according to the published sequence of Rouault et al. (17) (between positions 210–472). This probe was used to isolate phages with inserts λIRF1 and λIRF2 from a human placenta cDNA library and λIRF4 from a human EK4-fibroblast cDNA library (panel A). The inserts of clones λIRF2 and λIRF4 (Fig.1) were assembled to provide a full-length cDNA of IRF, which is 3498 bases in length (panel B). The start of translation is at position 108, the stop codon ends at nucleotide 2778. Restriction sites indicated are D, DraIII; E, EcoRI; H, HindIII; K, KpnI; N, NdeI; P, PvuII; S, SmaI; St, StyI.

PCR-probe was used to screen two independent cDNA libraries, a random primed one from human placenta and an oligo-dT primed library from the human fibroblast cell line EK4. In each case about 4×10^5 λ-phages were screened, and plaque hybridization yielded three clones in the placenta library (λIRF1, 2, and 3) and two clones in the fibroblast library (λIRF4 and 5) that were consistently positive after four rounds of re-screening. When the size and restriction sites of inserts were determined, we noticed in all three placenta clones a divergent pattern towards the 5′ end as compared to the reported clone of Rouault et al. (22). We did not continue to analyse λIRF3 and 5, because these two clones had similar inserts as λIRF1 and 4, respectively. EcoRI-fragments from λIRF1, 2 and 4 were subcloned into pGEM-vectors and entirely sequenced by the method of Sanger et al. (33).

Sequence data showed that the different clones are overlapping (Figs 1 and 2), and that the longest one from the placenta library contains a new sequence of 318 bases diverging from Rouault's clone (22) at their position 174. Due to this difference, our cDNA clones code for additional 99 amino acids upstream of the previously assumed initiation codon. We also found several other sequences to be different in λIRF4 compared to the originally published data (22). These changes concern sequences located at amino acids 532 to 543 and 712 to 727 (Fig.2) that have since been corrected by the authors (23). New sequences were verified on both strands.

The different 5′ end in our cDNA might have been due to alternative splicing or the existence of a second gene, but both assumptions seem unlikely. To test for the first possibility, we tried to amplify the 5′ end sequence of Rouault et al. by PCR in different cDNA libraries, but failed to obtain an amplification product. The second possibility is improbable due to the fact that our clones are identical over most of the remaining sequence with the one published by Rouault et al. (22). Therefore, we consider it likely that our IRF cDNAs and the IRE-BP clone from Rouault et al. are representative of the same gene. It is interesting, that our new NH_2-terminal amino acid sequences (Fig.2) as well as internal corrections improve the alignment to a related protein,

```
   1            GGGCTCGAACGCGCAGCGCACGGGAACCGGTCCCGCTGCTTGGGTCAGGTTCGCCGGTCGCGGGAGCCCCGCCGTGCAGTCGGAGGAACACGTGGCCATCAGTAATC

 108 ATG AGC AAC CCA TTC GCA CAC CTT GCT GAG CCA TTG GAT CCT GTA CAA CCA GGA AAG AAA TTC TTC AAT TTG AAT AAA TTG GAG GAT TCA
     Met Ser Asn Pro Phe Ala His Leu Ala Glu Pro Leu Asp Pro Val Gln Pro Gly Lys Lys Phe Phe Asn Leu Asn Lys Leu Glu Asp Ser  30

 198 AGA TAT GGG CGC TTA CCA TTT TCG ATC AGA GTT CTT CTG GAA GCA GCC ATT CGG AAT TGT GAT GAG TTT TTG GTG AAG AAA CAG GAT ATT
     Arg Tyr Gly Arg Leu Pro Phe Ser Ile Arg Val Leu Leu Glu Ala Ala Ile Arg Asn Cys Asp Glu Phe Leu Val Lys Lys Gln Asp Ile  60

 288 GAA AAT ATT CTA CAT TGG AAT GTC ACT CAG CAC AAG AAC ATA GAA GTG CCA TTT AAG CCT GCT CGT GTC ATC CTG CAG GAC TTT ACG GGT
     Glu Asn Ile Leu His Trp Asn Val Thr Gln His Lys Asn Ile Glu Val Pro Phe Lys Pro Ala Arg Val Ile Leu Gln Asp Phe Thr Gly  90

 378 GTG CCC GCT GTG GTT GAC TTT GCT GCA ATG CGT GAT GCT GTG AAA AAG TTA GGA GGA GAT CCA GAG AAA ATA AAC CCT GTC TGC CCT GCT
     Val Pro Ala Val Val Asp Phe Ala Ala Met Arg Asp Ala Val Lys Lys Leu Gly Gly Asp Pro Glu Lys Ile Asn Pro Val Cys Pro Ala 120

 468 GAT CTT GTA ATA GAT CAT TCC ATC CAG GTT GAT TTC AAC AGA AGG GCA GAC AGT TTA CAG AAG AAT CAA GAC CTG GAA TTT GAA AGA AAT
     Asp Leu Val Ile Asp His Ser Ile Gln Val Asp Phe Asn Arg Arg Ala Asp Ser Leu Gln Lys Asn Gln Asp Leu Glu Phe Glu Arg Asn 150

 558 AGA GAG CGA TTT GAA TTT TTA AAG TGG GGT TCC CAG GCT TTT CAC AAC ATG CGG ATT ATT CCC CCT GGC TCA GGA ATC ATC CAC CAG GTG
     Arg Glu Arg Phe Glu Phe Leu Lys Trp Gly Ser Gln Ala Phe His Asn Met Arg Ile Ile Pro Pro Gly Ser Gly Ile Ile His Gln Val 180

 648 AAT TTG GAA TAT TTG GCA AGA GTG GTA TTT GAT CAG GAT GGA TAT TAT TAC CCA GAC AGC CTC GTG GGC ACA GAC TCG CAC ACT ACC ATG
     Asn Leu Glu Tyr Leu Ala Arg Val Val Phe Asp Gln Asp Gly Tyr Tyr Pro Asp Ser Leu Val Gly Thr Asp Ser His Thr Thr Met 210

 738 ATT GAT GGC TTG GGC ATT CTT GGT TGG GGT GTC GGT GGT ATT GAA GCA GAA GCT GTC ATG CTG GGT CAG CCA ATC AGT ATG GTG CTT CCT
     Ile Asp Gly Leu Gly Ile Leu Gly Trp Gly Val Gly Gly Ile Glu Ala Glu Ala Val Met Leu Gly Gln Pro Ile Ser Met Val Leu Pro 240

 828 CAG GTG ATT GGC TAC AGG CTG ATG GGG AAG CCC CAC CCT CTG GTA ACA TCC ACT GAC ATC GTG CTC ACC ATT ACC AAG CAC CTC CGC CAG
     Gln Val Ile Gly Tyr Arg Leu Met Gly Lys Pro His Pro Leu Val Thr Ser Thr Asp Ile Val Leu Thr Ile Thr Lys His Leu Arg Gln 270

 918 GTT GGG GTA GTG GGC AAA TTT GTC GAG TTC TTC GGG CCT GGA GTA GCC CAG TTG TCC ATT GCT GAC CGA GCT ACG ATT GCT AAC ATG TGT
     Val Gly Val Val Gly Lys Phe Val Glu Phe Phe Gly Pro Gly Val Ala Gln Leu Ser Ile Ala Asp Arg Ala Thr Ile Ala Asn Met Cys 300

1008 CCA GAG TAC GGA GCA ACT GCT GCC TTT TTC CCA GTT GAT GAA GTT AGT ATC ACG TAC CTG GTG CAA ACA GGT CGT GAT GAA AAA TTA
     Pro Glu Tyr Gly Ala Thr Ala Ala Phe Phe Pro Val Asp Glu Val Ser Ile Thr Tyr Leu Val Gln Thr Gly Arg Asp Glu Glu Lys Leu 330

1098 AAG TAT ATT AAA AAA TAT CTT CAG GCT GTA GGA ATG TTT CGA GAT TTC AAT GAC CCT TCT CAA GAC CCA GAC TTC ACC CAG GTT GTG GAA
     Lys Tyr Ile Lys Lys Tyr Leu Gln Ala Val Gly Met Phe Arg Asp Phe Asn Asp Pro Ser Gln Asp Pro Asp Phe Thr Gln Val Val Glu 360

1188 TTA GAT TTG AAA ACA GTA GTG CCT TGC TGT AGT GGA CCC AAA AGG CCT CAG GAC AAA GTT GCT GTG TCC GAC ATG AAA AAG GAC TTT GAG
     Leu Asp Leu Lys Thr Val Val Pro Cys Cys Ser Gly Pro Lys Arg Pro Gln Asp Lys Val Ala Val Ser Asp Met Lys Lys Asp Phe Glu 390

1278 AGC TGC CTT GGA GCC AAG CAA GGA TTT AAA GGA TTC CAA GTT GCT CCT GAA CAT CAT AAT GAC CAT AAG ACC TTT ATC TAT GAT AAC ACT
     Ser Cys Leu Gly Ala Lys Gln Gly Phe Lys Gly Phe Gln Val Ala Pro Glu His His Asn Asp His Lys Thr Phe Ile Tyr Asp Asn Thr 420

1368 GAA TTC ACC CTT GCT CAT GGT TCT GTG GTC ATT GCT GCC ATT ACT AGC TGC ACA AAC ACC AGT AAT CCG TCT GTG ATG TTA GGG GCA GGA
     Glu Phe Thr Leu Ala His Gly Ser Val Val Ile Ala Ala Ile Thr Ser Cys Thr Asn Thr Ser Asn Pro Ser Val Met Leu Gly Ala Gly 450

1458 TTG TTA GCA AAG AAA GCT GTG GAT GCT GGC CTG AAC GTG ATG CCT TAC ATC AAA ACT AGC CTG TCT CCT GGG AGT GGC GTG GTC ACC TAC
     Leu Leu Ala Lys Lys Ala Val Asp Ala Gly Leu Asn Val Met Pro Tyr Ile Lys Thr Ser Leu Ser Pro Gly Ser Gly Val Val Thr Tyr 480

1548 TAC CTA CAA GAA AGC GGA GTC ATG CCT TAT CTG TCT CAG CTT GGG TTT GAC GTG GTG GGC TAT GGC TGC ATG ACC TGC ATT GGC AAC AGT
     Tyr Leu Gln Glu Ser Gly Val Met Pro Tyr Leu Ser Gln Leu Gly Phe Asp Val Val Gly Tyr Gly Cys Met Thr Cys Ile Gly Asn Ser 510

1638 GGG CCT TTA CCT GAA CCT GTG GTA GAA GCC ATC ACA CAG GGA GAC CTT GTA GCT GTT GGA GTA CTA TCT GGA AAC AGG AAT TTT GAA GGT
     Gly Pro Leu Pro Glu Pro Val Val Glu Ala Ile Thr Gln Gly Asp Leu Val Ala Val Gly Val Leu Ser Gly Asn Arg Asn Phe Glu Gly 540

1728 CGA GTT CAC CCC AAC ACC CGG GCC AAC TAT TTA GCC TCT CCC CCC TTA GTA ATA GCA TAT GCA ATT GCT GGA ACC ATC AGA ATC GAC TTT
     Arg Val His Pro Asn Thr Arg Ala Asn Tyr Leu Ala Ser Pro Pro Leu Val Ile Ala Tyr Ala Ile Ala Gly Thr Ile Arg Ile Asp Phe 570

1818 GAG AAA GAG CCA TTG GGA GTA AAT GCA AAG GGA CAG CAG GTA TTT CTG AAA GAT ATC TGG CCG ACT AGA GAC GAG ATC CAG GCA GTG GAG
     Glu Lys Glu Pro Leu Gly Val Asn Ala Lys Gly Gln Gln Val Phe Leu Lys Asp Ile Trp Pro Thr Arg Asp Glu Ile Gln Ala Val Glu 600

1908 CGT CAG TAT GTC ATC CCG GGG ATG TTT AAG GAA GTC TAT CAG AAA ATA GAG ACT GTG AAT GAA AGC TGG AAT GCC TTA GCA ACC CCA TCA
     Arg Gln Tyr Val Ile Pro Gly Met Phe Lys Glu Val Tyr Gln Lys Ile Glu Thr Val Asn Glu Ser Trp Asn Ala Leu Ala Thr Pro Ser 630

1998 GAT AAG CTG TTT TTC TGG AAT TCC AAA TCT ACG TAT ATC AAA TCA CCA CCA TTC TTT GAA AAC CTG ACT TTG GAT CTT CAG CCC CCT AAA
     Asp Lys Leu Phe Phe Trp Asn Ser Lys Ser Thr Tyr Ile Lys Ser Pro Pro Phe Phe Glu Asn Leu Thr Leu Asp Leu Gln Pro Pro Lys 660

2088 TCT ATA GTG GAT GCC TAT GTG CTG CTA AAT TTG GGA GAT TCG GTA ACA ACT GAC CAC ATC TCC CCA GCT GGA AAT ATT GCA AGA AAC AGT
     Ser Ile Val Asp Ala Tyr Val Leu Leu Asn Leu Gly Asp Ser Val Thr Thr Asp His Ile Ser Pro Ala Gly Asn Ile Ala Arg Asn Ser 690

2178 CCT GCT GCT CGC TAC TTA ACT AAC AGA GGC CTA ACT CCA CGA GAA TTC AAC TCC TAT GGC TCC CGC CGA GGT AAT GAC GCC GTC ATG GCA
     Pro Ala Ala Arg Tyr Leu Thr Asn Arg Gly Leu Thr Pro Arg Glu Phe Asn Ser Tyr Gly Ser Arg Arg Gly Asn Asp Ala Val Met Ala 720

2268 CGG GGA ACA TTT GCC AAC ATT CGC TTG TTA AAC AGA TTT TTG AAC AAG CAG GCA CCA CAG ACT ATC CAT CTG CCT TCT GGG GAA ATC CTT
     Arg Gly Thr Phe Ala Asn Ile Arg Leu Leu Asn Arg Phe Leu Asn Lys Gln Ala Pro Gln Thr Ile His Leu Pro Ser Gly Glu Ile Leu 750

2358 GAT GTG TTT GAT GCT GCT GAG CGG TAC CAG CAG GCA GGC CTT CCC CTG ATC GTT CTG GCT GGC AAA GAG TAC GGT GCA GGC AGC TCC CGA
     Asp Val Phe Asp Ala Ala Glu Arg Tyr Gln Gln Ala Gly Leu Pro Leu Ile Val Leu Ala Gly Lys Glu Tyr Gly Ala Gly Ser Ser Arg 780

2448 GAC TGG GCA GCT AAG GGC CCT TTC CTG CTG GGA ATC AAA GCC GTC CTG GCC GAG AGC TAC GAG CGC ATT CAC CGC AGT AAC CTG GTT GGG
     Asp Trp Ala Ala Lys Gly Pro Phe Leu Leu Gly Ile Lys Ala Val Leu Ala Glu Ser Tyr Glu Arg Ile His Arg Ser Asn Leu Val Gly 810

2538 ATG GGT GTG ATC CCA CTT GAA TAT CTC CCT GGT GAG AAT GCA GAT GCC CTG GGG CTC ACA GGG CAA GAA CGA TAC ACT ATC ATT ATT CCA
     Met Gly Val Ile Pro Leu Glu Tyr Leu Pro Gly Glu Asn Ala Asp Ala Leu Gly Leu Thr Gly Gln Glu Arg Tyr Thr Ile Ile Ile Pro 840

2628 GAA AAC CTC AAA CCA CAA ATG AAA GTC CAG GTC AAG CTG GAT ACT GGC AAG ACC TTC CAG GCT GTC ATG AGG TTT GAC ACT GAT GTG GAG
     Glu Asn Leu Lys Pro Gln Met Lys Val Gln Val Lys Leu Asp Thr Gly Lys Thr Phe Gln Ala Val Met Arg Phe Asp Thr Asp Val Glu 870

2718 CTC ACT TAT TTC CTC AAC GGG GGC ATC CTC AAC TAC ATG ATC CGC AAG ATG GCC AAG TAG GAGACGTGCACTTGGTCGTGCGCCCAGGGAGGAAGCCGC
     Leu Thr Tyr Phe Leu Asn Gly Gly Ile Leu Asn Tyr Met Ile Arg Lys Met Ala Lys                                          889

2817 ACCACCAGCCAGCGCAGGCCCTGGTGGAGAGGCCTCCCTGGCTGCCTCTGGGAGGGGTGCTGCCTTGTAGATGGAGCAAGTGAGCACTGAGGGTCTGGTGCCAATCCTGTAGGCACAAA
2936 ACCAGAAGTTTCTACATTCTCTATTTTTGTTAATCATCTTCTCTTTTTCCAGAATTTGGAAGCTAGAATGGTGGGAATGTCAGTAGTGCCAGAAAGAGAGAACCAAGCTTGTCTTTAAA
3055 GTTACTGATCACAGGACGTTGCTTTTTCACTGTTTCCTATTAATCTTCAGCTGAACACAAGCAAACCTTCTCAGGAGGTGTCTCCTACCCTCTTATTGTTCCTCTTACGCTCTGCTCAA
3174 TGAAACCTTCCTCTTGAGGGTCATTTTCCTTTCTGTATTAATTATACCAGTGTTAAGTGACATAGATAAGAACTTTGCACACTTCAAATCAGAGCAGTGATTCTCTCTTCTCTCCCCTT
3293 TTCCTTCAGAGTGAATCATCCAGACTCCTCATGGATAGGTCGGGTGTTAAAGTTGTTTTGATTATGTACCTTTTGATAGATCCACATAAAAAGAAATGTGAAGTTTTCTTTTACTATCT
3412 TTTCATTTATCAAGCAGAGACCTTTGTTGGGAGGCGGTTTGGGAGAACACATTTCTAATTTGAATGAAATGAAATCTATTTTCAGTG
```

Figure 2. Nucleotide and protein sequences of the human IRF cDNA. The cDNA inserts from clones λIRF2 and λIRF4 were fully sequenced by the method of Sanger et al. (33). The number of nucleotides (left side) and amino acids (right side) are indicated.

clone 10.1 published by Rouault et al. (22). No cross-hybridizing clones were found in our screening.

The overlapping inserts of clones λIRF2 and λIRF4 were assembled into a larger cDNA for IRF (Fig.1) according to the description in Materials and Methods. The complete sequence is 3498 nucleotides in length with an open reading frame coding for 889 amino acids (Fig.2). The first initiation codon is located at position 108 and matches the second most frequent sequence found for mammalian translation start sites (38), and the stop codon ends at position 2778. The deduced amino acid sequence specifies a protein of 98,398 Da (Fig.2), a value close to the size of our previous determinations for purified placental IRF (3,31). Computer calculation estimates the pI of this protein at 6.21, close to the experimental pI of 6.45 to 6.55 found for purified IRF from human placenta (31). As pointed out by Rouault et al. (23), the IRF protein sequence shows homology to pig aconitase. The new NH_2-terminal region, however, does not improve the sequence similarity between the two proteins.

Southern and Northern blot hybridizations were carried out in order to determine whether the IRF cDNA would hybridize

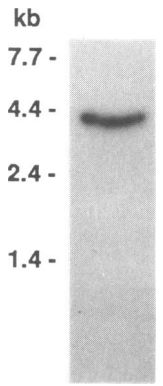

Figure 3. Northern blot analysis of IRF mRNA. Total RNA (25 μg) from human fibroblast cell line EK4 was separated on a 1.2% agarose gel containing 6.6% formaldehyde and transferred to nylon filters. The blot was hybridized to a probe corresponding to base positions 143 to 3040 of the human IRF cDNA.

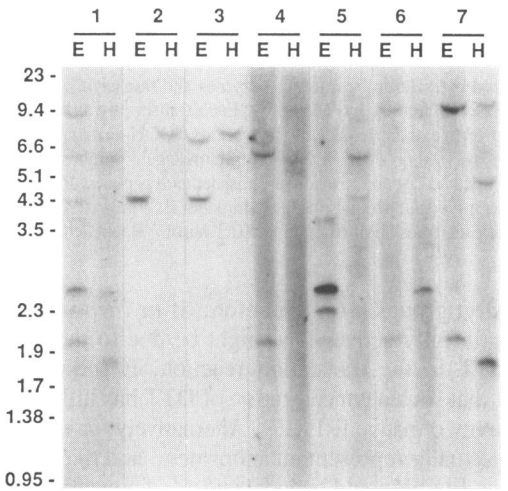

Figure 4. Southern blot hybridization of IRF cDNA fragments to human DNA. Genomic DNA from human placenta was digested with EcoRI (E) or with HindIII (H), separated on a 0.8% agarose gel and transferred to a nylon membrane. The blot was sequencially hybridized with various IRF cDNA probes at base positions 143–3040, (lane 1), 44–344 (lane 2), 210–472 (lane 3), 644–1368 (lane 4), 1368–2015 (lane 5), 2015–2220 (lane 6), and 2220–3040 (lane 7). Fragments of λ-DNA digested with HindIII/HindIII-EcoRI were used as molecular weight markers.

to a single gene or mRNA species. The radio-labelled cDNA probe recognized a single mRNA species of about 4.0 kb in a Northern blot with total RNA from the human fibroblast cell line EK4 (Fig.3). For Southern blot analysis, human genomic DNA from placenta was digested to completion with EcoRI or HindIII. The restriction fragments were resolved in an agarose gel, blotted and sequencially probed with a series of cDNA fragments (Fig.4). The results show that the IRF gene is at least 40 kb in size and represents a single copy gene. This is particularly evident in lanes 2 and 3 of Fig.4, where a short cDNA probe hybridizes with a single genomic restriction fragment, or two adjacent fragments, respectively.

From Southern hybridizations and sequence data it was not possible to decide whether the observed protein doublet of affinity-purified human placental IRF (31) is derived from a single

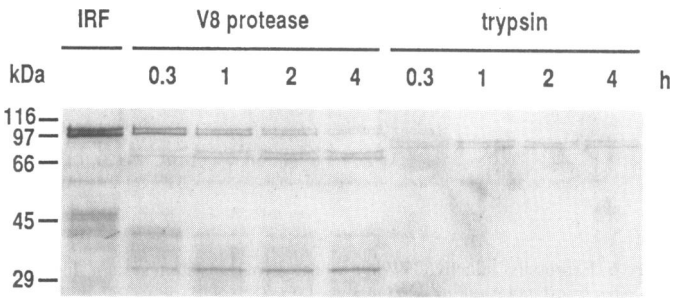

Figure 5. Partial digestion with V8 protease or trypsin of native IRF from human placenta. IRF was purified by affinity chromatography from human placenta (31), and 1 μg aliquots were partially digested for different lengths of time as indicated with V8 protease (1:20 w/w) or trypsin (1:100 w/w). The products were analysed on a 7.5% SDS-polyacrylamide gel and stained with Coomassie blue.

gene or two non-crosshybridizing genes. Therefore, we took a biochemical approach with proteases to analyse this heterogeneity. Highly enriched IRE-binding proteins from human placenta were partially digested with V8 protease or trypsin for different lengths of time (Fig.5). IRF was cleaved by both enzymes resulting in large sized protein doublet and, in the case of V8 protease, an apparently homogeneous fragment of 31 kDa. In order to determine the V8 protease cleavage site, both the large and the small fragments were sequenced by automated Edman degradation. The NH_2-terminus of the large doublet, as previously observed for the entire protein, was found to be blocked. The NH_2-terminus of the small 31 kDa-V8 protease fragment, however, gave a homogeneous sequence: XXNALATPSDKLFFXN. From this result it was concluded that partial digestion with V8 protease cuts native IRF at position 621. We deduced that the fragments of 65 and 70 kDa (Fig.5) correspond to the NH_2-terminal region of IRF and carry the observed heterogeneity of the placental protein. Since the COOH-terminal fragment did not show any heterogeneity in the sequenced amino acids , it is likely that the two proteins in the placental IRF doublet are encoded by a single gene. This is consistent with the above results from genomic Southern blots (Fig.4) and the previous identification of a single locus on chromosome 9 (22,39). The heterogeneity of the placental IRF doublet could be due to alternative splicing, a post-translational modification or allelic forms of IRF.

Expression of IRF by *in vitro* translation and its binding to an IRE

Several questions concerning the structure and function of IRF remain to be solved, in particular, the modulation of its activity by iron levels and the physical features of its interaction with IREs. Studying these questions would be facilitated by expression of IRF from cloned cDNA sequences. Since there are no previous reports of a successful expression, we decided to investigate IRF made in an *in vitro* transcription/translation system. The IRF cDNA was assembled (Fig.1) and subcloned behind the T7-promoter into pGEM-7zf(−). This plasmid, pGEM-hIRF, was linearized downstream of the cDNA insert and transcribed *in vitro* with T7-polymerase to obtain a capped RNA of 3.5 kb. The RNA was translated *in vitro* in a wheat germ translation reaction in the presence of ^{35}S-methionine. When the synthesized polypeptides were analysed by SDS-polyacrylamide gel electrophoresis, the major band migrated at about 97 kDa

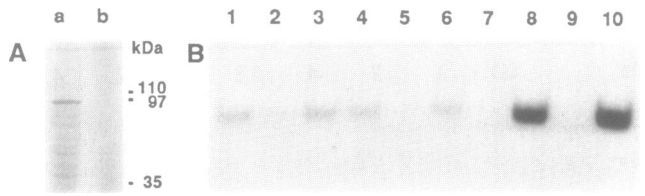

Figure 6. Expression of active IRF in a wheat germ translation system. In panel A, *in vitro* translated proteins were analysed on a 8% SDS-polyacrylamide gel. The translation was either with a full-length IRF RNA transcribed from the reassembled cDNA behind a T7 promoter (a) or without any RNA (b). A translation of Brome Mosaic Virus RNA was used for molecular weight markers. In panel B, a gel retardation-assay with the ^{32}P-labelled IRE of human ferritin H chain (transcript of pSPT-fer (3)) and aliquots from an *in vitro* transcription/translation reaction of IRF (lanes 1–6) or from a human placenta extract (lanes 8–10). Lane 1: incubation of *in vitro* translated IRF with 0.1 ng IRE RNA; lane 2: as in lane 1, but after pre-incubation of IRF with the sulfhydryl oxidizing reagent diamide at 1.5 mM; lane 3: as in lane 2, but followed by the reduction of oxidized IRF with 2% β-mercaptoethanol; lane 4: same as lane 1, after reduction with 2% β-mercaptoethanol only; lane 5: like lane 4, reduced IRF in the presence of 60 ng unlabelled competitor pSPT-fer RNA; lane 6: reduced IRF, as in lane 4, in the presence of 60 ng 4F2 antigen heavy chain RNA; lane 7: aliquot of a translation reaction without IRF RNA; lane 8: incubation of probe with human placenta extract (1 μg); lane 9: as in 8, but in the presence of 60 ng unlabelled pSPT-fer RNA; and lane 10: in presence of 60 ng 4F2 antigen heavy chain RNA.

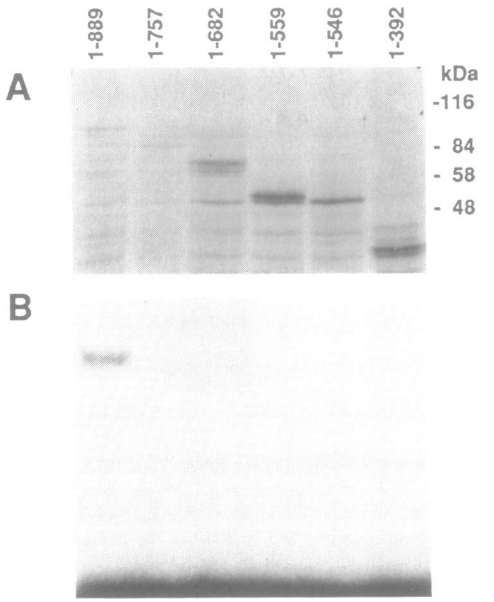

Figure 7. Requirement of the COOH-terminal part of *in vitro* translated IRF for the RNA-protein interaction. The assembled IRF cDNA plasmid pGEM-hIRF was linearized by different restriction enzymes and transcribed *in vitro* with T7 polymerase (see Materials and Methods). Each transcript was translated *in vitro* by the wheat germ extract. In panel A, the various COOH-terminal deletion mutants of IRF were analysed on a 8%-SDS-polyacrylamide gel. Numbers above the panel indicate the first and last amino acids of translated polypeptides. Prestained proteins were used as molecular weight markers. In panel B, a gel retardation assay with an aliquot of each translation reaction and 0.1 ng of ^{32}P-labelled pSPT-fer probe.

as expected for the open reading frame of the IRF cDNA (Fig.6A, lane a). No protein was translated in the absence of RNA (Fig.6A, lane b).

In order to test whether the translated protein corresponded to active IRF, a parallel unlabelled batch of the same protein was made and examined in gel retardation assays. An aliquot of the translation reaction was incubated with a ^{32}P-labelled RNA transcript of pSPT-fer containing the human ferritin H chain IRE (Fig.6B). The translated protein bound to this probe (lane 1), and the RNA-protein complex migrated at the same position as a complex between the IRE and IRF from a human placenta extract (lane 8). However, no complex could be detected with a translation reaction carried out in the absence of IRF RNA (lane 7). When the translated protein was pre-incubated with the sulfhydryl-oxidant diamide, binding to the IRE was inhibited (lane 2), but could be restored after reduction with β-mercaptoethanol (lane 3). Thus, *in vitro* translated IRF had the same properties as those described for affinity-purified IRF from human placenta (31). This was also true for the specificity of its interaction with the IRE. Binding could be entirely inhibited by an excess of unlabelled pSPT-fer RNA (lanes 5 and 9), but not with an unrelated RNA (lanes 6 and 10).

The addition of β-mercaptoethanol to the translated products without prior oxidation did not increase the IRE-binding activity (lane 4) compared to an untreated aliquot (lane 1). Therefore, *in vitro* translated IRF is synthesized immediatly with features of an active RNA-binding protein. In the case of IRF, this finding is not trivial. Recent evidence from sequence comparisons have revealed homology to iron-sulfur cluster containing isomerases (23,24), suggesting that IRF is most likely also an iron-sulfur cluster protein. IRF activity could be controlled as aconitase by iron exchange in such a cluster. This postulate is supported by the *in vivo* data indicating that chelatable iron has a central role in the final setting of IRF activity (3,14,15). Other lines of evidence mainly based on *in vitro* data suggest, however, that the modulation of IRF involves also the oxidation and reduction of sulfhydryl groups (16). Therefore, if *in vitro* translated IRF is inherently fully active, this might be due to the presence of 5 mM DTT in the translation reaction. It should be noted, however, that such a concentration of DTT has little or no effect on the activity of native IRF (17). Alternatively, *in vitro* translated IRF may actually represent an apoprotein, and the redox effects observed in Fig.6B could be independent of a putative iron-sulfur cluster. In support of this hypothesis, we and others found that changes in the redox state of the sulfhydryls on IRF are less permanent than changes induced *in vivo* by iron levels (17, our unpublished observation). In addition, it is known that [4Fe-4S] clusters are unstable in an aerobic environment (40). This makes it highly unlikely that such a cluster is incorporated into the nascent polypeptide during the *in vitro* translation. As a consequence, it would seem that apo-IRF can still bind to IREs.

A COOH-terminal region of IRF is required for interaction with IREs

To examine how IRF interacts with the IRE-element and which sites of the protein are necessary for this activity, deletions were made in the full-length cDNA. We shortened IRF at its COOH-terminus in order to analyse polypeptides of various length for their RNA-binding activity. pGEM-hIRF was linearized at particular restriction sites located in the open reading frame of the cDNA (see Fig.1). Each linearized plasmid was transcribed and translated *in vitro* creating deleted versions of IRF (Fig.7A). These polypeptides were incubated with the labelled probe of pSPT-fer in a gel retardation assay (Fig.7B). Interestingly, only the wild type form of IRF bound to the RNA (lane 1), whereas a deletion of just 132 amino acids abolished the binding entirely (lane 2). This means that the COOH-terminal region is required for the interaction of IRF with the IRE.

The present results leave open how other parts of IRF are involved in the RNA binding. In preliminary experiments with ultra-violet crosslinking, we observed complexes between the NH_2-terminal 65/70 kDa V8 protease fragments (Fig.5) and an IRE probe (unpublished results). This suggests, that both NH_2- and COOH-terminal regions of IRF participate in the RNA-protein interaction. Thus, unlike certain transcription factors (41) or hnRNA-binding proteins (42) IRF does not seem to contain a restricted domain for its binding activity. It will be of interest to mutagenize IRF in order to precisely locate which regions are necessary for its interaction with IREs and its activation by iron levels. Expression *in vitro*, as described in this paper, will provide a fast method to analyse such mutants.

ACKNOWLEDGEMENTS

We wish to thank Dr. R. Tyrrell for the gift of the EK4 cDNA library and S. Modak for EK4 RNA. We also acknowledge Drs C. Dargemont and M. Nabholz for carefully reading the manuscript. The present study was supported by the Swiss National Foundation for Scientific Research. C.S. is a recipient of an Erwin-Schrödinger-fellowship.

REFERENCES

1. Owen, D. and Kühn, L.C. (1987) EMBO J., **6**, 1287–1293.
2. Müllner, E.W. and Kühn, L.C. (1988) Cell, **53**, 815–825.
3. Müllner, E.W., Neupert, B. and Kühn, L.C. (1989) Cell, **58**, 373–382.
4. Zähringer, J., Baliga, B.S. and Munro, H.N. (1976) Proc. Natl. Acad. Sci.USA, **73**, 857–861.
5. Hentze, M.W., Rouault, T.A., Caughman, S.W., Dancis, A., Harford, J.B. and Klausner, R.D. (1987) Proc. Natl. Acad. Sci. USA, **84**, 6730–6734.
6. Aziz, N. and Munro, H.N. (1987) Proc. Natl. Acad. Sci. USA, **84**, 8478–8482.
7. Brown, P.H., Daniels-McQueen, S., Walden, W.E., Patino, M.M., Gaffield, L., Bielser, D. and Thach, R.E. (1989) J. Biol. Chem., **264**, 13383–13386.
8. Casey, J.L., Hentze, M.W., Koeller, D.M., Caughman, S.W., Rouault, T.A., Klausner, R.D. and Harford, J.B. (1988) Science, **240**, 924–928.
9. Koeller, D.M., Casey, J.L., Hentze, M.W., Gerhardt, E.M., Chan, L.-N., Klausner, R.D. and Harford, J.B. (1989) Proc. Natl. Acad. Sci. USA, **86**, 3574–3578.
10. Casey, J.L., Koeller, D.M., Ramin, V.C., Klausner, R.D. and Harford, J.B. (1989) EMBO J., **8**, 3693–3699.
11. May, B.K., Bhasker, C.R., Bawden, M.J. and Cox, T.C. (1990) Mol.Biol.Med., **7**, 405–421.
12. Cox, T.C., Bawdon, M.J., Martin, A. and May, B.K. (1991) EMBO J., **10**, 1891–1902.
13. Dandekar, T., Stripecke, R., Gray, N.K., Goossen, B., Constable, A., Johansson, H.E. and Hentze, M.W. (1991) EMBO J., **10**, 1903–1909.
14. Leibold, E.A. and Munro, H.N. (1988) Proc. Natl. Acad. Sci. USA, **85**, 2171–2175.
15. Rouault, T.A., Hentze, M.W., Caughman, S.W., Harford, J.B. and Klausner, R.D. (1988) Science, **241**, 1207–1210.
16. Hentze, M.W., Rouault, T.A., Harford, J.B. and Klausner, R.D. (1989) Science, **244**, 357–359.
17. Barton, H.A., Eisenstein, R.S., Bomford, A. and Munro, H.N. (1990) J. Biol. Chem., **265**, 7000–7008.
18. Haile, D.J., Hentze, M.W., Rouault, T.A., Harford, J.B. and Klausner, R.D. (1989) Mol. Cell. Biol., **9**, 5055–5061.
19. Aziz, N. and Munro, H.N. (1986) Nucl. Acids Res., **14**, 915–927.
20. Walden, W.E., Patino, M.M. and Gaffield, L. (1989) J. Biol. Chem., **264**, 13765–13769.
21. Goossen, B., Caughman, S.W., Harford, J.B., Klausner, R.D. and Hentze, M.W. (1990) EMBO J., **9**, 4127–4133.
22. Rouault, T.A., Tang, C.K., Kaptain, S., Burgess, W.H., Haile, D.J., Samaniego, F., McBride, O.W., Harford, J.B. and Klausner, R.D. (1990) Proc. Natl. Acad. Sci. USA, **87**, 7958–7962.
23. Rouault, T.A., Stout, C.D., Kaptain, S., Harford, J.B. and Klausner, R.D. (1991) Cell, **64**, 881–883.
24. Hentze, M.W. and Argos, P. (1991) Nucl. Acids Res., **19**, 1739–1740.
25. Beinert, H. and Kennedy, M.C. (1989) Eur. J. Biochem., **186**, 5–15.
26. Rothenberger, S., Müllner, E.W. and Kühn, L.C. (1990) Nucl. Acids Res., **18**, 1175–1179.
27. Lin, J.-J., Daniels-McQueen, S., Patino, M.M., Gaffield, L., Walden, W.E. and Thach, R.E. (1990) Science, **247**, 74–77.
28. Lin, J.J., Daniels-McQueen, S., Gaffield, L., Patino, M.M., Walden, W.E. and Thach, R.E. (1990) Biochim. Biophys. Acta, **1050**, 146–150.
29. Haile, D.J., Rouault, T.A., Harford, J.B. and Klausner, R.D. (1990) J. Biol. Chem., **265**, 12786–12789.
30. Rouault, T.A., Hentze, M.W., Haile, D.J., Harford, J.B. and Klausner, R.D. (1989) Proc. Natl. Acad. Sci. USA, **86**, 5768–5772.
31. Neupert, B., Thompson, N.A., Meyer, C. and Kühn, L.C. (1990) Nucl. Acids Res., **18**, 51–55.
32. Tyrrell, R.M. and Pidoux, M. (1986) Cancer Res., **46**, 2665–2669.
33. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, **74**, 5463–5467.
34. Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem., **162**, 156–159.
35. Kühn, L.C., Barbosa, J.A., Kamarck, M.E. and Ruddle, F.H. (1983) Mol. Biol. Med., **1**, 335–352.
36. Krieg, P.A. and Melton, D.A. (1987) Meth. Enzymol., **155**, 397–415.
37. Teixeira, S., Di Grandi, S. and Kühn, L.C. (1987) J. Biol. Chem., **262**, 9574–9580.
38. Kozak, M. (1983) Microbiol. Rev., **47**, 1–45.
39. Hentze, M.W., Seuanez, H.N., O'Brien, S.J., Harford, J.B. and Klausner, R.D. (1989) Nucl. Acids Res., **17**, 6103–6108.
40. Kennedy, M.C., Emptage, M.H., Dreyer, J.-L. and Beinert, H. (1983) J. Biol. Chem., **258**, 11098–11105.
41. Ptashne, M. (1988) Nature, **335**, 683–689.
42. Bandziulis, R.,J., Swanson, M.,S. and Dreyfuss, G. (1991) Genes Dev. **3**, 431–437.