The EMBO Journal vol.10 no.7 pp.1891 - 1902, 1991

Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron-responsive element in the mRNA

Timothy C.Cox, Michael J.Bawden, Alison Martin and Brian K.May

Department of Biochemistry, University of Adelaide, GPO Box 498, Adelaide, South Australia 5001, Australia

Communicated by I.W.Mattaj

5-Aminolevulinate synthase (ALAS) catalyzes the first step of the heme biosynthetic pathway. cDNA clones for the human erythroid ALAS isozyme were isolated from a fetal liver library. It can be deduced that the erythroid ALAS precursor protein has a molecular weight of 64.6 kd, and is similar in size to the previously isolated human housekeeping ALAS precursor of molecular weight 70.6 kd. The mature mitochondrial forms of the erythroid and housekeeping ALAS isozymes are predicted to have molecular weights of 59.5 kd and 64.6 kd, respectively. The two isozymes show little amino acid identity in their N-terminal signal sequences but have considerable sequence identity in the C-terminal two-thirds of their proteins. An analysis of the immediate promoter of the human erythroid ALAS gene revealed several putative erythroid-specific *cis*-acting elements including both a GATA-1 and an NF-E2 binding site. An iron-responsive element (IRE) motif has been identified in the 5′-untranslated region of the human erythroid ALAS mRNA, but is not present in the housekeeping ALAS mRNA. Gel retardation experiments established that this IRE motif formed a protein–RNA complex with cytosolic extracts from human K562 cells and this binding was strongly competed with IRE transcripts from ferritin or transferrin receptor mRNAs. A transcript of the ALAS IRE, mutated in the conserved loop of the IRE, did not readily form this protein–RNA complex. These results suggest that the IRE motif in the ALAS mRNA is functional and imply that translation of the mRNA is controlled by cellular iron availability during erythropoiesis.

Key words: 5-aminolevulinate synthase/erythroid heme synthesis/iron-responsive element/post-transcriptional regulation

Introduction

During erythropoiesis, large amounts of hemoglobin are produced at the terminal stages of red cell differentiation and this requires a co-ordinated production of heme molecules and globin chains. The increased synthesis of globin results from increased globin gene transcription (Karlsson and Nienhuis, 1985) and translation (London *et al.*, 1987) while enhanced heme formation is a result of an increase in the levels of all enzymes of the heme pathway (Beaumont *et al.*, 1984), with at least three genes being transcriptionally activated (Beaumont *et al.*, 1984; Schoenhaut and Curtis, 1986; Elferink *et al.*, 1988; Raich *et al.*, 1989; Romeo *et al.*, 1986). Enzymes of the heme biosynthetic pathway must be expressed in all cell types to supply heme for respiratory cytochromes but there is a particular need for heme during erythropoiesis as greater than 80% of total body heme is synthesized in erythroid cells (Worwood, 1977).

We have focused on the regulation of expression of 5-aminolevulinate synthase (ALAS), the first enzyme of the heme biosynthetic pathway. The enzyme is synthesized in the cytosol as a precursor protein with an N-terminal signal sequence that is proteolytically cleaved upon transport of the protein into the mitochondria (May *et al.*, 1986). The resulting mature mitochondrial protein is catalytically active and catalyzes the formation of 5-aminolevulinate (ALA) from glycine and succinyl CoA in a reaction that is rate-limiting in the liver and probably other tissues (May *et al.*, 1986).

It has been recently established that there are two separate ALAS genes in chicken and mice encoding different ALAS isozymes (Riddle *et al.*, 1989; Schoenhaut and Curtis, 1989). One gene, the housekeeping gene, is expressed ubiquitously and expression of this gene in rat liver can be elevated by porphyrinogenic drugs and repressed by administration of the end-product hemin (Elliott *et al.*, 1989). The other gene, the erythroid-specific gene, is expressed in erythroid tissue to supply heme for hemoglobin; and transcription of this gene is not affected by heme (Elferink *et al.*, 1988; Elliott *et al.*, 1989).

We have previously isolated cDNA clones for the housekeeping ALAS of chick embryo (Borthwick *et al.*, 1985), rat (Srivastava *et al.*, 1988) and human (Bawden *et al.*, 1987) and determined the structure of the chicken housekeeping ALAS gene (Maguire *et al.*, 1986). As a first step towards understanding the molecular regulation of ALAS during erythropoiesis, we report here the characterization of cDNA clones and the gene promoter for human erythroid ALAS. The promoter contains both potential erythroid-specific and ubiquitous *cis*-acting DNA motifs which have been identified in other erythroid-specific promoters. In the 5′-untranslated region (UTR) of the erythroid ALAS mRNA there is an iron-responsive element (IRE), similar to those found in the mRNAs for ferritin (Aziz and Munro, 1987) and transferrin receptor (Casey *et al.*, 1988; Müllner *et al.*, 1989). Gel retardation analysis experiments have shown that the IRE in the ALAS mRNA is likely to be functional and our results imply that translation of this mRNA is regulated by iron availability in erythroid cells.

Results

Isolation of cDNA clones for human erythroid ALAS

In early mammalian development, the fetal liver is an erythroid tissue and is a major site of erythropoiesis (Karlsson and Nienhuis, 1985). To investigate the ALAS mRNA species expressed in human fetal liver, a 50

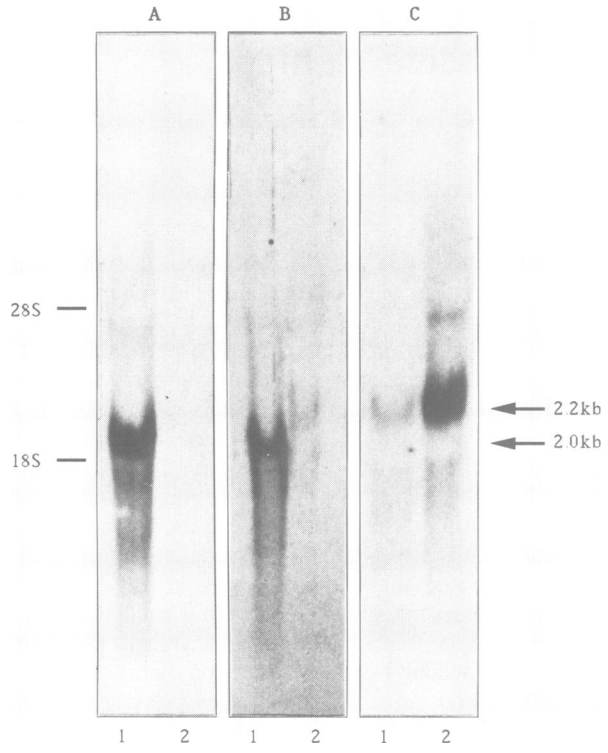


Fig. 1. Northern blot analysis of ALAS mRNAs in human fetal and adult liver. Total RNA (40 μg) for human fetal liver (lane 1) and human adult liver (lane 2) was electrophoresed on a 1.1 M formaldehyde−1% agarose gel and probed with oligolabelled human erythroid ALAS cDNA (pHEA-6) (panel A), chemically synthesized 50mer 5′-phosphorylated with [γ-^{32}P]ATP (panel B), or oligolabelled human housekeeping ALAS cDNA (panel C). The positions of the 2.2 kb housekeeping and 2.0 kb erythroid ALAS mRNAs are indicated by arrows. The relative rRNA positions are also indicated. Molecular size markers used were generated by *Acc*I digestion of pBR322.

nucleotide oligomeric probe (50mer) was synthesized, the sequence of which was based on a highly conserved region of 17 amino acids located in the putative catalytic domain of animal ALAS proteins (Elliott *et al.*, 1989). In initial experiments, total RNA isolated from both human fetal liver and human adult liver was examined by Northern blot analysis using as hybridization probes either the 50mer or an ALAS cDNA clone previously isolated from human adult liver (Bawden *et al.*, 1987). We have shown in separate studies that the latter cDNA clone detects a ubiquitous ALAS mRNA of size 2.2 kb in all human cell lines and human tissues examined and is therefore referred to here as the housekeeping ALAS. When Northern blots of human fetal and adult liver RNA were probed with this cDNA clone, a single predominant mRNA species of size 2.2 kb was seen in the adult liver whereas this species was barely detectable in fetal liver (Figure 1, panel C). By contrast, the 50mer, when used as probe, hybridized strongly to an mRNA species of size 2.0 kb present only in the fetal liver, while showing a weak hybridization signal for the 2.2 kb mRNA species in the adult liver (Figure 1, panel B). These results show that an ALAS mRNA of size 2.0 kb is expressed in fetal liver and that this species can be detected with the 50mer. A human fetal liver cDNA library in λgt11 was screened with the 50mer probe. From a total of 5×10^5 recombinant phage, fifty positive clones were identified. Of these, twelve strongly hybridizing clones were plaque purified through three subsequent rounds of screening and recloned in both orientations into the *Eco*RI site of the phagemid vector, pTZ19. The cDNA inserts ranged in size from ~800 bp to 1.9 kb. Restriction enzyme mapping showed that the cDNA clones were distinct from the cDNA clone encoding the human housekeeping ALAS enzyme (Bawden *et al.*, 1987) and that the clones contained common restriction enzyme fragments. One clone, pHEA-6, with an insert of ~1.9 kb was further examined. When pHEA-6 was expressed in *E.coli*, which does not contain endogenous ALAS (Avissar and Beale, 1989), large amounts of ALAS activity were detected in the cell lysate and this correlated with a foreign protein of molecular weight 68 kd (data not shown), the size expected for ALAS (May *et al.*, 1986). Northern blots of total RNA from human fetal liver and human adult liver were reprobed with pHEA-6 together with the housekeeping ALAS cDNA clone under the same hybridization conditions. As expected, pHEA-6 hybridized strongly to a major ALAS mRNA of size 2.0 kb in fetal liver but detected only traces of this mRNA in adult liver (Figure 1, panel A). The results obtained with the housekeeping ALAS cDNA clone were as described previously. In other Northern blot experiments it was established that the 2.0 kb ALAS mRNA is present in human bone marrow, human red blood cells and the erythroid/myeloid cell lines (K562 and HEL) but is absent in human platelets, B cells and the hemopoietic cell lines of monocytic (HL60 and U937) and lymphoid (JURKAT) origin and from the non-erythroid HepG2 cell line. These results confirmed that pHEA-6 encodes an ALAS isozyme that is most probably expressed only in erythroid cells with the trace amount of erythroid mRNA detected in adult liver resulting from circulating red cells.

Characterization of ALAS cDNA clones

The nucleotide sequence of pHEA-6 was determined and shown to contain an insert of 1890 bp, while sequence from two additional clones extended the sequence a further 12 bp in the 5′ direction and 35 bp in the 3′ direction, respectively (Figure 2). To determine the site of initiation of transcription of the erythroid ALAS gene and thereby the full length of the mRNA, primer extension analysis was carried out. A 5′-γ-^{32}P-labelled 19mer complementary to the sequence of nucleotides 19−37 in Figure 2 was used in a primer extension reaction on total RNA isolated from human fetal liver. Two major products of 37 and 38 nucleotides were seen following electrophoresis of the extension products and detection by autoradiography (Figure 3). The primer extension results indicate that the 5′-UTR of the erythroid ALAS mRNA is 52 or 53 nucleotides in length with pHEA-6 being 18 or 19 nucleotides short of full-length. The transcription start site was confirmed using a 20mer complementary to the sequence of nucleotides 79−98 in Figure 2 (data not shown). For convenience, the nucleotide at the 5′ end of the 37 nucleotide extension product has been designated as +1 (see Figure 5). The remaining sequence of the 5′-UTR was determined by sequencing an appropriately subcloned genomic fragment (see Figure 2).

From the full-length cDNA sequence it can be seen that there are two in-frame ATG codons at nucleotide positions 53 and 68; the first of these is considered to be the translation initiation codon since this codon is preceded by a sequence which more closely resembles the consensus start sequence

proposed by Kozak (1984). There is a polyadenylation signal sequence, AATAAA, at position 1923 and an in-frame TGA stop codon at position 1814. The open reading frame of 1761 nucleotides from the ATG codon at nucleotide 53 to the termination codon encodes a predicted ALAS precursor of 587 amino acids with a molecular weight of 64.6 kd, in agreement with the sizes estimated for other animal ALAS proteins (May *et al.*, 1986) and for the protein expressed from pHEA-6 in *E.coli*.

Comparison of ALAS nucleotide and protein sequences

At the nucleotide level, sequences for the human erythroid and housekeeping ALAS precursor proteins show about 60% identity, with the longest stretch of identical sequence being 21 nucleotides. This finding accounts for the fact that pHEA-6 and the human housekeeping ALAS cDNA clone, when used separately as probes in Northern blots of human fetal and adult liver RNA, only detected their specific mRNA

```
                                                  M  V  T  A  A  M  L  L  Q  C  C  P  V
CACCTGTCATTCGTTCGTCCTCAGTGCAGGGCAACAGGACTTTAGGTTCAAGATGGTGACTGCAGCCATGCTGCTACAGTGCTGCCCAGT
        10         20         30         40         50         60         70         80         90

  L  A  R  G  P  T  S  L  L  G  K  V  V  K  T  H  Q  F  L  F  G  I  G  R  C  P  I  L  A  T
GCTTGCCCGGGGCCCCACAAGCCTCCTAGGCAAGGTGGTTAAGACTCACCAGTTCCTGTTTGGTATTGGACGCTGTCCCATCCTGGCTAC
       100        110        120        130        140        150        160        170        180

  Q  G  P  N  C  S  Q  I  H  L  K  A  T  K  A  G  G  D  S  P  S  W  A  K  G  H  C  P  F  M
CCAAGGACCAAACTGTTCTCAAATCCACCTTAAGGCAACAAAGGCTGGAGGAGATTCTCCATCTTGGGCGAAGGGCCACTGTCCCTTCAT
       190        200        210        220        230        240        250        260        270

  L  S  E  L  Q  D  G  K  S  K  I  V  Q  K  A  A  P  E  V  Q  E  D  V  K  A  F  K  T  D  L
GCTGTCGGAACTCCAGGATGGGAAGAGCAAGATTGTGCAGAAGGCAGCCCCAGAAGTCCAGGAAGATGTGAAGGCTTTCAAGACAGATCT
       280        290        300        310        320        330        340        350        360

  P  S  S  L  V  S  V  S  L  R  K  P  F  S  G  P  Q  E  Q  E  Q  I  S  G  K  V  T  H  L  I
GCCTAGCTCCCTGGTCTCAGTCAGCCTAAGGAAGCCATTTTCCGGTCCCCAGGAGCAGGAGCAGATCTCTGGGAAGGTCACACACCTGAT
       370        380        390        400        410        420        430        440        450

  Q  N  M  M  P  G  M  Y  V  F  S  Y  D  Q  F  F  R  D  K  I  M  E  K  K  Q  D  H  T  Y  R
TCAGAACAATATGCCTGGAAACTATGTCTTCAGTTATGACCAGTTTTTCAGGGACAAGATCATGGAGAAGAAACAGGATCACACCTACCG
       460        470        480        490        500        510        520        530        540

  V  F  K  T  V  M  R  W  A  D  A  Y  P  F  A  Q  H  F  S  E  A  S  V  A  S  K  D  V  S  V
TGTGTTCAAGACTGTGAACCGCTGGGCTGATGCATATCCCTTTGCCCAACATTTCTCTGAGGCATCTGTGGCCTCAAAGGATGTGTCCGT
       550        560        570        580        590        600        610        620        630

  W  C  S  M  D  Y  L  G  M  S  R  H  P  Q  V  L  Q  A  T  Q  E  T  L  Q  R  H  G  A  G  A
CTGGTGTAGTAATGATTACCTGGGCATGAGCCGACACCCTCAGGTCTTGCAAGCCACACAGGAGACCCTGCAGCGTCATGGTGCTGGAGC
       640        650        660        670        680        690        700        710        720

  G  G  T  R  N  I  S  G  T  S  K  F  H  V  E  L  E  Q  E  L  A  E  L  H  Q  K  D  S  A  L
TGGTGGCACCCGCAACATCTCAGGCACCAGTAAGTTTCATGTGGAGCTTGAGCAGGAGCTGGCTGAGCTGCACCAGAAGGACTCAGCCCT
       730        740        750        760        770        780        790        800        810

  L  F  S  S  C  F  V  A  N  D  S  T  L  F  T  L  A  K  I  L  P  G  C  E  I  Y  S  D  A  G
GCTCTTCTCCTCCTGCTTTGTTGCCAATGACTCTACTCTCTTCACCTTGGCCAAGATCCTGCCAGGGTGCGAGATTTACTCAGACGCAGG
       820        830        840        850        860        870        880        890        900

  N  H  A  S  M  I  Q  G  I  R  N  S  G  A  A  K  F  V  F  R  H  N  D  P  D  H  L  K  K  L
CAACCATGCTTCCATGATCCAAGGTATCCGTAACAGTGGAGCAGCCAAGTTTGTCTTCAGGCACAATGACCCTGACCACCTAAAGAAACT
       910        920        930        940        950        960        970        980        990

  L  E  K  S  N  P  K  I  P  K  I  V  A  F  E  T  V  H  S  M  D  G  A  I  C  P  L  E  E  L
TCTAGAGAAGTCTAACCCTAAGATACCCAAAATTGTGGCCTTTGAGACTGTCCACTCCATGGATGGTGCCATCTGTCCCCTCGAGGAGTT
      1000       1010       1020       1030       1040       1050       1060       1070       1080

  C  D  V  S  H  Q  Y  G  A  L  T  F  V  D  E  V  H  A  V  G  L  Y  G  S  R  G  A  G  I  G
GTGTGATGTGTCCCACCAGTATGGGGCCCTGACCTTCGTGGATGAGGTCCATGCTGTAGGACTGTATGGGTCCCGGGGCGCTGGGATTGG
      1090       1100       1110       1120       1130       1140       1150       1160       1170

  E  R  D  G  I  M  H  K  I  D  I  I  S  G  T  L  G  K  A  F  G  C  V  G  G  Y  I  A  S  T
GGAGCGTGATGGAATTATGCATAAGATTGACATCATCTCTGGAACTCTTGGCAAGGCCTTTGGCTGTGTGGGCGGCTACATTGCCAGCAC
      1180       1190       1200       1210       1220       1230       1240       1250       1260

  R  D  L  V  D  M  V  R  S  Y  A  A  G  F  I  F  T  T  S  L  P  P  M  V  L  S  G  A  L  E
CCGTGACTTGGTGGACATGGTGCGCTCCTATGCTGCAGGCTTCATCTTTACCACTTCTCTGCCCCCCATGGTGCTCTCTGGAGCTCTAGA
      1270       1280       1290       1300       1310       1320       1330       1340       1350

  S  V  R  L  L  K  G  E  E  G  Q  A  L  R  R  A  H  Q  R  N  V  K  H  M  R  Q  L  L  M  D
ATCTGTGCGGCTGCTCAAGGGAGAGGAGGGCCAAGCCCTGAGGCGAGCCCACCAGCGCAATGTCAAGCACATGCGCCAGCTACTCATGGA
      1360       1370       1380       1390       1400       1410       1420       1430       1440
```

```
  R  G  L  P  V  I  P  C  P  S  H  I  I  P  I  R  V  G  N  A  A  L  N  S  K  L  C  D  L  L
CAGGGGCCTTCCTGTCATCCCCTGCCCCAGCCACATCATCCCCATCCGGGTGGGCAATGCAGCACTCAACAGCAAGCTCTGTGATCTCCT
      1450       1460       1470       1480       1490       1500       1510       1520       1530

  L  S  K  H  G  I  Y  V  Q  A  I  N  Y  P  T  V  P  R  G  E  E  L  L  R  L  A  P  S  P  H
GCTCTCCAAGCATGGCATCTATGTGCAGGCCATCAACTACCCAACTGTCCCCCGGGGTGAAGAGCTCCTGCGCTTGGCACCCTCCCCCCA
      1540       1550       1560       1570       1580       1590       1600       1610       1620

  H  S  P  Q  M  M  E  D  F  V  E  K  L  L  L  A  W  T  A  V  G  L  P  L  Q  D  V  S  V  A
CCACAGCCCTCAGATGATGGAAGATTTTGTGGAGAAGCTGCTGCTGGCTTGGACTGCGGTGGGGCTGCCCCTCCAGGATGTGTCTGTGGC
      1630       1640       1650       1660       1670       1680       1690       1700       1710

  A  C  N  F  C  R  R  P  V  H  F  E  L  M  S  E  W  E  R  S  Y  F  G  N  M  G  P  Q  Y  V
TGCCTGCAATTTCTGTCGCCGTCCTGTACACTTTGAGCTCATGAGTGAGTGGGAACGTTCCTACTTCGGGAACATGGGGCCCCAGTATGT
      1720       1730       1740       1750       1760       1770       1780       1790       1800

  T  T  Y  A  *
CACCACCTATGCCTGAGAAGCCAGCTGCCTAGGATTCACACCCCACCTGCGCTTCACTTGGGTCCAGGCCTACTCCTGTCTTCTGCTTTG
      1810       1820       1830       1840       1850       1860       1870       1880       1890

                 ←
TTGTGTGCCTCTAGCTGAATTGAGCCTAAAAATAAAGCACAAACCACAGCA(A)n
      1900       1910       1920
```

Fig. 2. Nucleotide and predicted amino acid sequence of human erythroid ALAS. The full-length nucleotide sequence (1955 nucleotides) and predicted amino acid sequence of human erythroid ALAS precursor protein (molecular weight 64.6 kd) are presented. The 5′-and 3′-UTRs are of 52 and 127 nucleotides, respectively. A polyadenylation signal (AATAAA) is present at position 1923. The termination codon at position 1814 is asterisked. The sequence derived from the cDNA clone, pHEA-6, extends from position 31 to 1921 (arrows). The sequence from positions 18 to 30 and 1921 to 1955 were obtained from two additional clones. The remaining sequence to the transcription start site was obtained by sequencing an appropriately subcloned genomic fragment.

species and showed no cross-hybridization as seen in Figure 1.

The predicted amino acid sequence of the human erythroid ALAS precursor was aligned with that of the human housekeeping ALAS precursor and with sequences of other ALAS proteins from both mammalian and prokaryote sources (Figure 4). During transport into mitochondria, N-terminal signal sequences are removed to generate the mature mitochondrial enzymes (May *et al.*, 1986). As shown in Figure 4, chicken and rat housekeeping ALAS precursor proteins have N-terminal signal sequences of 56 amino acids (designated region 1 in Figure 4) with the proteolytic cleavage site in both cases located between two glutamines (Maguire *et al.*, 1986; Srivastava *et al.*, 1988). On this basis, the signal sequence of the human housekeeping ALAS precursor is also predicted to be 56 amino acids, with the proteolytic cleavage site occurring between two glutamine residues (Figure 4). An examination of the signal sequences of all of the housekeeping ALAS enzymes shown in Figure 4 reveals marked sequence identity as previously noted (Elliott *et al.*, 1989). The length of the signal sequence of the human erythroid ALAS precursor is not known, but from the alignment in Figure 4, it is predicted to be 49 amino acid residues with the cleavage site between serine and glutamine residues; this would result in a mature mitochondrial protein of molecular weight 59.5 kd, somewhat smaller than its housekeeping counterpart of molecular weight 64.4 kd. This putative signal sequence would have limited amino acid sequence identity with that of the human housekeeping isozyme but would closely resemble that proposed for the mouse erythroid enzyme (Schoenhaut and Curtis, 1986). Interestingly, the signal sequence of the chicken erythroid ALAS enzyme (Riddle *et al.*, 1989) resembles more closely the signal sequence of the chicken housekeeping isozyme (Borthwick *et al.*, 1985) than those of the other erythroid isozymes (Figure 4).

We have previously concluded (Elliott *et al.*, 1989) that the catalytically active core of all animal ALAS proteins lies within a C-terminal region of about 440 amino acids representing 75% of the mature protein and is designated

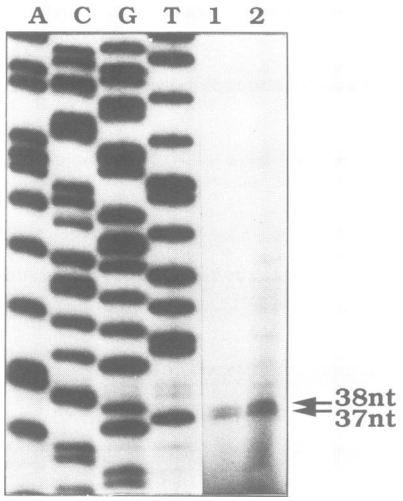


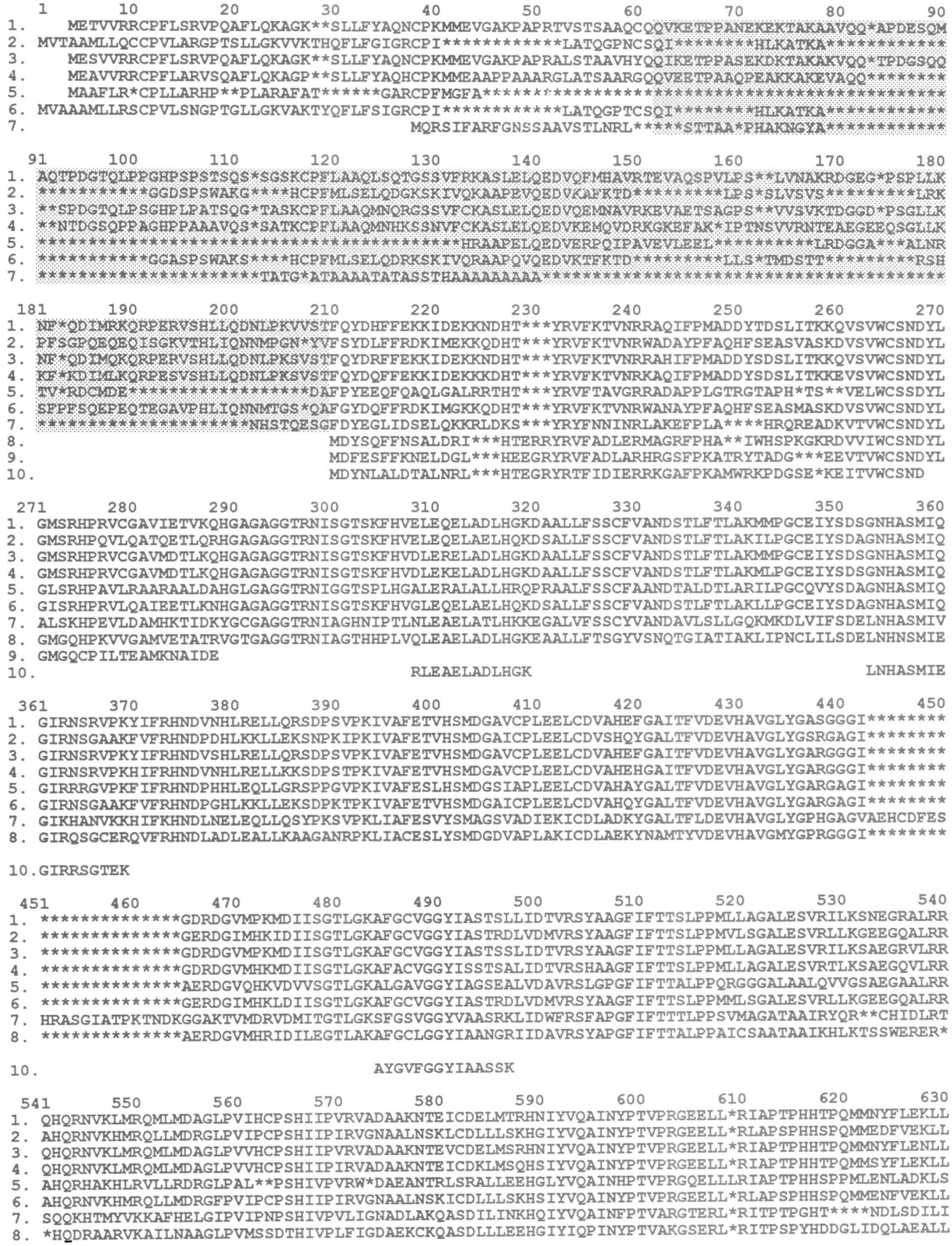
Fig. 3. Determination of the transcription start site of human erythroid ALAS gene by primer extension analysis on human fetal liver total RNA. A chemically synthesized 19mer 5′-phosphorylated with [γ-^{32}P]ATP was used to primer extend on 10 μg (lane 1) and 20 μg (lane 2) of total RNA from human fetal liver. Products were analyzed on a 7 M urea, 6% polyacrylamide gel with an M13 sequencing ladder as size markers. The two major extension products of 37 and 38 nucleotides are indicated by arrows.

region 3 in Figure 4. The N-terminus of this highly conserved domain corresponds approximately to the N-termini of the bacterial enzymes. In keeping with this proposal, region 3 of the human erythroid ALAS protein (positions 210−680 in Figure 4) shows strong sequence identity with the corresponding regions of the human housekeeping (73%) and mouse erythroid (94%) enzymes. The remaining N-terminal region of the human erythroid ALAS mature protein (designated region 2 in Figure 4) is 47 amino acids shorter than that of its housekeeping counterpart with limited stretches of sequence identity. However, it can be seen that this region of the human erythroid protein clearly resembles the corresponding region in the mouse erythroid isozyme. In contrast, region 2 of the

chicken erythroid protein exhibits limited identity with the corresponding region from human and mouse erythroid proteins and is also considerably shorter.

Analysis of the human erythroid ALAS gene promoter

We have recently established that the human erythroid and housekeeping isozymes of ALAS are encoded by distinct genes (Sutherland *et al.*, 1988; Cox *et al.*, 1990). Characterization of the entire gene for human erythroid ALAS has demonstrated that the gene is 22 kb in length and is organized into 11 exons (manuscript in preparation). The exons range in size from 37 bp to 270 bp and the introns from 561 bp to 6.0 kb, with the longest intron (intron 1) being located in the 5′-UTR of the gene. The gene structure bears remarkable similarity to that previously reported from mouse erythroid ALAS (Schoenhaut and Curtis, 1989).



```
 631      640       650       660       670       680       690
1. LTWKRVGLELKPHSSAECNFCRRPLHFEVMSEREKAYFSGMSKM*VSAQA
2. LAWTAVGLPLQDVSVAACNFCRRPVHFELMSEWERSYFGNMGPQYVTTYA
3. VTWKQVGLELKPHSSAECNFCRRPLHFEVMSEREKSYFSGLSKL*VSAQA
4. ATWKDVGLELKPHSSAECNFCRRPLHFEVMSERERSYFSGMSKL*LSVSA
5. ECWGAVGLPREDPPGPSCSSCHRPLHLSLLSPLERDQFG******VRGAAAG
6. LAWTGVGLPLQDVSVAACNFCHRPVHFELMSEWERSYFGNMGPQYVTTYA
7. NAVDDVFNELQLPRVRDWESQGGLLGVGESGFVEESNLWTSSQLSLTNDDLNPNVRDPIVKQLEVSSGIKQ
8. QVWDRLGLPLKQKSLAA
```

1. RAT HOUSEKEEPING
2. HUMAN ERYTHROID
3. HUMAN HOUSEKEEPING
4. CHICK HOUSEKEEPING
5. CHICK ERYTHROID
6. MOUSE ERYTHROID
7. *Saccharomyces cerevisiae*
8. *Bradyrhizobium japonicum*
9. *Rhizobium meliloti*
10. *Rhodopseudomonas spheroides*

Fig. 4. Comparison of amino acid sequences of eukaryotic and prokaryotic ALAS proteins as described from cDNA clones or direct protein sequencing. The eukaryote ALAS sequences have previously been assigned into three distinct regions (Elliott *et al.*, 1989): Region 1 corresponds to the N-terminal signal sequence that is cleaved upon transport into the mitochondrion. The known cleavage site for the chicken (Borthwick *et al.*, 1985) and rat (Srivastava *et al.*, 1988) housekeeping ALAS proteins occurs between two glutamine residues at positions 61–62. The predicted cleavage site for the human erythroid ALAS signal sequence is also between positions 61 and 62. Region 3 beginning at position 210 comprises the C-terminal two-thirds of the proteins and corresponds to the complete bacterial ALAS proteins. The remaining sequence from positions 62 to 209 (shaded) is designated Region 2. The amino acid sequences used for this comparison were obtained from: 1) Srivastava *et al.* (1988); 3) Bawden *et al.* (1987); 4) Borthwick *et al.* (1985); 5) Riddle *et al.* (1989); 6) Schoenhaut and Curtis (1986); 7) Urban-Grimal *et al.* (1986); 8) McClung *et al.* (1987); 9) Leong *et al.* (1985); 10) Elliott *et al.* (1989). The sequences of the human and rat housekeeping ALAS proteins have been corrected from those published previously (Bawden *et al.*, 1987; Srivastava *et al.*, 1988) with alterations occurring chiefly in region 2. The eighth amino acid of the human erythroid ALAS protein reported (May *et al.*, 1990) has also been corrected (Y to L).

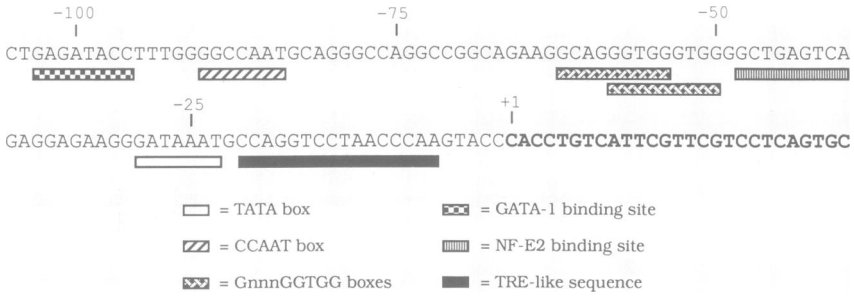


Fig. 5. Sequence of the immediate promoter region of the human erythroid ALAS gene. Promoter sequence is represented in plain text whereas exon 1 sequence is in bold text. The site of transcription initiation is designated as +1. Numerous putative *cis*-regulatory elements are indicated. The sequence shown was determined from sequencing both strands.

An examination of the immediate promoter of the human erythroid ALAS gene revealed several interesting putative *cis*-acting motifs (Figure 5). Unlike that of the mouse erythroid ALAS gene (Schoenhaut and Curtis, 1989), the human promoter contains a consensus TATA box at position −27 and a consensus CCAAT box at position −87. These elements are nearly identical in sequence and position to those found in the human α- and β-globin gene promoters where they are known to be functional (Leibhaber *et al.*, 1980; Myers *et al.*, 1986). A possible GATA-1 binding site is located at position −100. GATA-1 is an erythroid-specific *trans*-acting factor which has been shown to bind to multiple sites in the promoter of the human β-globin gene (deBoer *et al.*, 1988; Wall *et al.*, 1988; Orkin, 1990) and to the upstream enhancer of the β-globin locus (Wall *et al.*, 1988; Martin *et al.*, 1989; Philipsen *et al.*, 1990; Talbot *et al.*, 1990) as well as to the erythroid porphobilinogen deaminase (PBGD) promoter (Mignotte *et al.*, 1989a and b). At position −40 there is a perfect consensus for NF-E2, another erythroid-specific factor. The NF-E2 motif was first recognized in the erythroid PBGD promoter (Mignotte *et al.*, 1989a and b) and more recently, multiple sites have been found in the upstream enhancer of the β-globin locus (Philipsen *et al.*, 1990; Talbot *et al.*, 1990; Ney *et al.*, 1990). The NF-E2 sequence in the ALAS gene promoter is identical to that observed in the erythroid PBGD promoter (Mignotte *et al.*, 1989a). Centred around position −56, there are two matches to the GnnnGGTGG motif recently proposed to be a key determinant in erythroid gene expression (Philipsen *et al.*, 1990). Indeed, this motif is positioned ~35 bp from a GATA-1 binding site, therefore being consistent with its proposed functional role (Philipsen *et al.*, 1990). Alternatively, located in this position on the minus strand there is the sequence 5′-CCCCACCCACCCT-3′, a partially duplicated sequence having two overlapping matches with the consensus CAC box binding site (Dierks *et al.*, 1983). Of particular interest is the presence in the ALAS promoter of a potential thyroid hormone responsive element (Glass *et al.*, 1988; Beato, 1989) from position −21 to −6, located immediately downstream of the TATA box.

Presence of an iron-responsive element sequence in the human erythroid ALAS mRNA

As mentioned previously, a notable feature of the human erythroid ALAS gene is the presence of a 6.0 kb intron located between nucleotides +37 and +38 of the 52 bp 5′-UTR. An examination of this 5′-UTR revealed a palindromic RNA sequence with striking similarity to an iron-responsive element (IRE) (Aziz and Munro, 1987; Hentze *et al.*, 1987). IREs have been found in the 5′-UTR of mRNAs for the iron-storage protein, ferritin (Aziz and

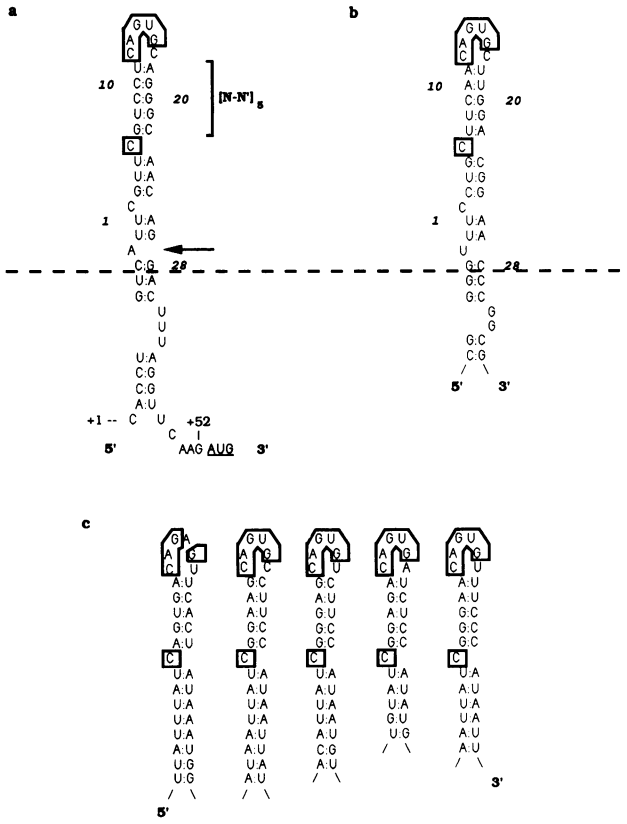


Fig. 6. Comparison of the postulated IRE found in the 5′-UTR of human erythroid ALAS mRNA (**a**) with the IREs found in the 5′- and 3′-UTRs of human ferritin H-chain (**b**) and human transferrin receptor (**c**) mRNAs, respectively. The complete 5′-UTR of ALAS is depicted and numbered +1 to +52. The IRE structure consists of an unpaired cytosine residue 5 bp 5′ of the 5′-CAGUGX-3′ motif of the loop. The conserved bases are boxed. The position of the 6.0 kb intron that interrupts the 5′-UTR of ALAS is indicated by the arrow. The putative AUG initiation codon is underlined and in bold type. The extended secondary structure of the ALAS 5′-UTR referred to in the text is that sequence below the dotted line. Above the dotted line are the highly conserved 28 nucleotides separately numbered 1 to 28 of the ferritin H-chain IRE and the analogous 28 nucleotides found in the ALAS IRE. The ferritin H-chain structure indicated has been modified slightly from that proposed by Casey *et al.* (1988) to maximize secondary structure similarity with the ALAS IRE.

Munro, 1987; Hentze *et al.*, 1987; Leibold and Munro, 1988), and the 3′-UTR of mRNAs for transferrin receptor (Casey *et al.*, 1988; Müllner *et al.*, 1989), a membrane receptor that mediates iron uptake. Interestingly, the proposed secondary structure for human erythroid ALAS (Figure 6a) involves essentially the entire 5′-UTR (nucleotide +2 to +46), despite the lower half of the stem being derived from both exon 1 and exon 2 sequence. Also in Figure 6, the proposed secondary structure for the IREs for human ferritin H-chain (Aziz and Munro, 1987; Hentze *et al.*, 1987) and transferrin receptor (Müllner *et al.*, 1989; Casey *et al.*, 1989) are compared with that for the human erythroid ALAS. The structure of the 5′-UTR of the ferritin mRNA depicted in Figure 6 contains a highly conserved IRE regulatory region of 28 nucleotides and an extended stem provided by the sequence flanking the IRE (Aziz and Munro, 1987; Casey *et al.*, 1988). Characteristic features of IREs for ferritin and transferrin receptor include an upper loop of six nucleotides of consensus 5′-CAGUGX-3′, an unpaired cytosine residue 5 nucleotide pairs 5′ of this loop, and an additional base paired stem structure (Casey *et al.*, 1988).

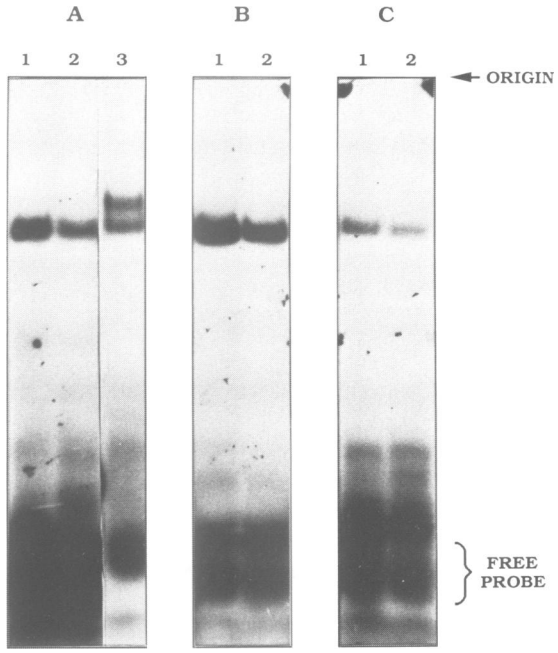


Fig. 7. Formation of specific RNA−protein complexes between the 5′-UTR of human erythroid ALAS and cytosolic extracts from human K562, human HL60 and mouse J2E-1 cell lines. [^{32}P]RNA transcripts (2.0 ng) of the ALAS 5′-UTR (panel **A**), ferritin H-chain IRE (panel **B**) and the human transferrin receptor IRE (panel **C**) were incubated at room temperature with 30 μg of cytosolic protein extract from either human K562 (lane 1), human HL60 (lane 2) or mouse J2E-1 (lane 3) cell lines. Following the sequential addition of RNase T_1 and heparin, the complexes were resolved in a 6% non-denaturing acrylamide gel.

It is clear from Figure 6 that the secondary structure proposed for the 5′-UTR of the ALAS mRNA is very similar to that for the ferritin mRNA IRE.

Analysis of IRE function by gel retardation analysis

Cellular ferritin and transferrin receptor levels are co-ordinately regulated by iron via post-transcriptional mechanisms. The IREs of the ferritin and transferrin receptor mRNAs bind a cytosolic protein, known as the iron-responsive element binding protein (IRE-BP), in response to a decrease in intracellular iron levels and this results in inhibition of ferritin mRNA translation and decreased degradation of the transferrin receptor mRNA, respectively (Theil, 1990). There is evidence that the same IRE-BP binds to the IREs of ferritin and transferrin receptor (Rothenberger *et al.*, 1990; Leibold *et al.*, 1990).

Gel retardation experiments were employed to determine if a cytosolic protein specifically binds to the putative IRE in the ALAS mRNA. For these experiments, transcripts were synthesized for the proposed IRE of ALAS (from position +1 to +51). Other transcripts containing the human ferritin H-chain IRE (Müllner *et al.*, 1989) and human transferrin receptor IRE were also employed (Müllner *et al.*, 1989). Cytosolic extracts from human monocyte HL60 cells, human erythroid/myeloid K562 cells and mouse erythroid J2E-1 cells (Klinken *et al.*, 1988) were incubated with radiolabelled RNA transcripts containing either the ALAS, ferritin or transferrin receptor IRE transcripts. Following digestion with ribonuclease T_1 (RNase T_1), complexes were separated in 6% non-denaturing acrylamide gels. In Figure 7, it can be seen that a prominent RNA−protein complex is obtained

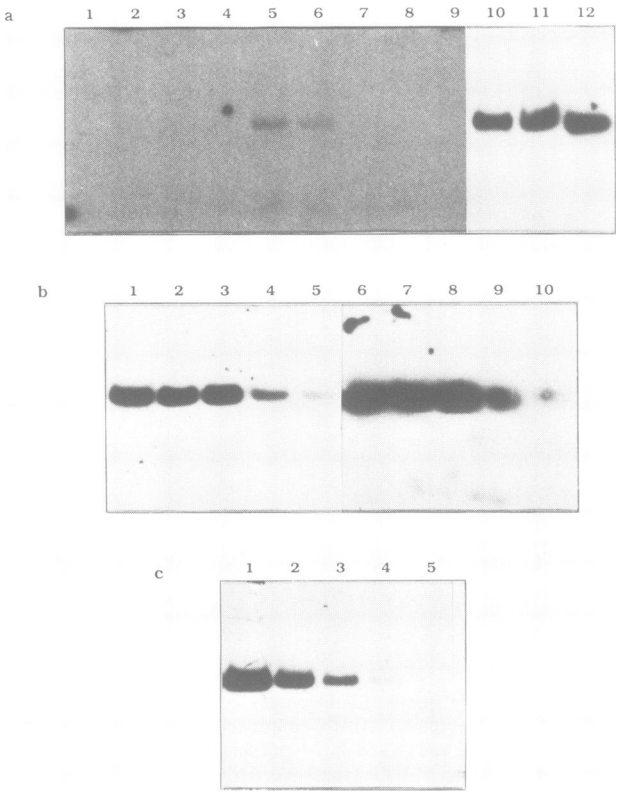


Fig. 8. The 5′-UTR of human erythroid ALAS specifically competes with the ferritin and transferrin receptor IREs for the binding of IRE-BP. (**a**) K562 cytosolic protein extract (10 μg) was incubated with 0.5 ng of radiolabelled ALAS 5′-UTR transcript in the presence of a 3-, 10-, 30- or 100-fold molar excess of unlabelled transcripts of either ferritin H-chain IRE (lanes 4−1, respectively) or transferrin receptor IRE (lanes 6−9, respectively). Lane 5 had no competitor added. To the 0.5 ng of ALAS 5′-UTR transcript in lanes 10−12 was added either 0-, 250- or 500-fold molar excess of *E.coli* tRNA. The complexes were resolved in a 6% non-denaturing acrylamide gel. (**b**) Reactions were carried out exactly as described in (a) except that unlabelled ALAS 5′-UTR transcript at 0-, 3-, 10- and 30- or 100-fold molar excess was used as the competitor with either radiolabelled ferritin H-chain IRE (lanes 1−5) or transferrin receptor IRE (lanes 6−10). (**c**) Control reactions showing that unlabelled transferrin receptor IRE at 0-, 3-, 10-, 30- or 100-fold molar excess (lanes 1−5) competes with ferritin H-chain IRE for binding of IRE-BP from K562 as reported (Casey *et al.*, 1988; Müllner *et al.*, 1989; Leibold *et al.*, 1990).

using cytosolic extracts from HL60 and K562 cells with each of the three radiolabelled transcripts and these complexes ran with similar mobilities. With J2E-1 extracts, two distinct RNA−protein complexes were observed with the ALAS transcript (Figure 7) while two complexes with similar mobilities were also seen with the ferritin and transferrin receptor IRE transcripts (results not shown). Such complexes with mouse extracts have been seen previously with ferritin and transferrin receptor IREs but the reason for this is not clear (Leibold and Munro, 1988; Rothenberger *et al.*, 1990).

In competition experiments, we next examined binding of the ^{32}P-labelled ALAS putative IRE to K562 cytosolic protein in the presence of 0- to 100-fold molar excess of unlabelled ferritin or transferrin receptor IRE transcripts. *E.coli* tRNA was used as a control competitor. A 100-fold molar excess of either the competitor ferritin or transferrin receptor IRE transcripts markedly inhibited the formation of RNA−protein complexes with the ALAS putative IRE transcript (Figure 8). By contrast, *E.coli* tRNA, even at a 500-fold molar excess, did not compete. In the reverse experiments, unlabelled ALAS transcript used as a competitor up to 100-fold molar excess, inhibited the formation of RNA−protein complexes with either radiolabelled ferritin or transferrin receptor IRE transcripts (Figure 8). These experiments suggest that the 5′-UTR of ALAS contains a functional IRE.

In other experiments, a shortened ALAS 5′-UTR transcript derived from *in vitro* transcription across exon 1 and lacking 5′-UTR sequence encoded by exon 2, was shown to give a similar RNA−protein complex with K562 extracts to that observed with the extended ALAS 5′-UTR stem (Figure 9). This result shows that for protein binding, the extended secondary structure of the ALAS IRE depicted in Figure 6 is not obligatory. This is in keeping with findings for the ferritin IREs, where nucleotides 1−28 (see Figure 6) are sufficient for IRE-BP binding (Casey *et al.*, 1988; Müllner *et al.*, 1989; Leibold *et al.*, 1990; Barton *et al.*, 1990) and also IRE function (Casey *et al.*, 1988; Caughman *et al.*, 1988).

Mutagenesis analysis of ferritin and transferrin receptor IREs has established the importance of the conserved nucleotide loop sequence 5′-CAGUGX-3′ in the binding of IRE-BP (Leibold *et al.*, 1990; Barton *et al.*, 1990). To investigate whether a similar requirement is needed for the binding of IRE-BP to the ALAS IRE, the consensus loop sequence was altered (see Figure 9c). Gel retardation analysis using this mutated transcript and K562 cytosolic extracts showed that formation of the specific RNA−protein complex was markedly reduced (Figure 9a). This result was supported by competition experiments using the mutated ALAS IRE as unlabelled competitor and the ferritin H-chain IRE as the radiolabelled probe (Figure 9b). In contrast to the wild type ALAS 5′-UTR which competed efficiently with the ferritin H-chain IRE at 30-fold molar excess (Figure 8b), the mutated ALAS 5′-UTR competed inefficiently with ferritin H-chain IRE even up to 180-fold molar excess (Figure 9b).

Overall, our results show that the ALAS IRE sequence binds an IRE-BP which is either identical or at least closely related to that which recognizes ferritin and transferrin receptor IREs and imply that the ALAS IRE sequence plays a functional role in the regulation of ALAS in erythroid cells.

Discussion

During erythropoiesis, large amounts of heme are required for hemoglobin and there is evidence that ALAS catalyzes the rate-controlling step in heme formation (Bottomley and Müller-Eberhard, 1988; Gardner and Cox, 1988; Ponka *et al.*, 1988). As a first step towards understanding the molecular mechanisms modulating the expression of ALAS during erythroid cell differentiation, we have isolated cDNA clones and the gene for the human erythroid ALAS isozyme. These ALAS clones encode a precursor protein of molecular weight 64.6 kd, similar in size to that of the human housekeeping precursor of molecular weight 70.6 kd (Bawden *et al.*, 1987). In region 3, which encompasses approximately the C-terminal two-thirds of both of these isozymes, there is considerable similarity and it seems that this region contains the catalytic site (Elliott *et al.*, 1989). The N-terminal signal sequences of the ALAS isozymes (region 1 in Figure 4), show limited similarity unlike those

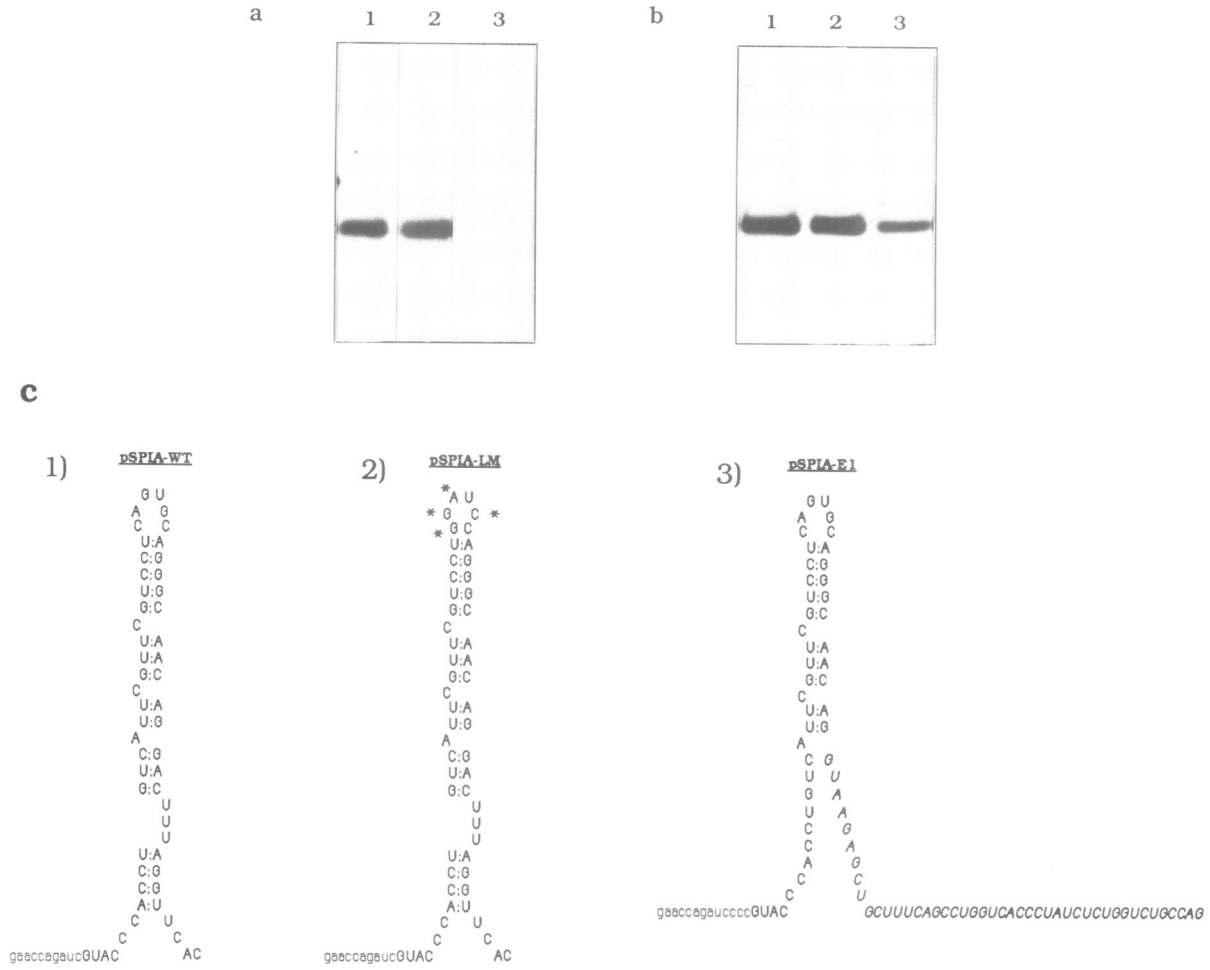


Fig. 9. Structural requirements of the human erythroid ALAS 5′-UTR for binding of IRE-BP. (**a**) K562 cytosolic protein extract (30 μg) was incubated with 2.0 ng of radiolabelled ALAS transcripts. These transcripts [see (c)] comprised either the entire wild-type 5′-UTR (lane 1), 5′-UTR encoded by exon 1 (lane 2), or the mutated 5′-UTR (lane 3). The complexes were resolved in a 6% non-denaturing acrylamide gel. (**b**) Reactions were carried out as described in Figure 8, except that ferritin H-chain IRE was used as the radiolabelled probe and unlabelled mutated ALAS 5′-UTR as the competitor transcript in 0-, 90- or 180-fold molar excess (lanes 1−3, respectively). (**c**) For use in (a) and (b) above, transcripts were generated *in vitro* for the entire human erythroid ALAS 5′-UTR containing wild type ALAS IRE (pSPIA-WT) (lane 1) the mutated ALAS IRE (pSPIA-LM (lane 2), or the truncated 5′-UTR encoded by exon 1 (pSPIA-E1) (lane 3). pSPIA-WT and pSIA-LM were linearized with *Nco*I, whereas pSPIA-E1 was internally linearized with *Pvu*II prior to their addition to transcription reactions. Asterisks indicate the mutated nucleotides in pSPIA-LM. Lower case characters represent pSP73 polylinker sequence. Upper case characters represent ALAS sequence. Uppercase italicized characters in pSPIA-E1 represent intron 1 derived sequence.

of the housekeeping ALAS proteins from different sources which display remarkable identity and have been proposed to be involved in the heme-inhibition of liver ALAS transport into mitochondria (May *et al.*, 1990). Whether there is in fact a functional basis for the dissimilarity in the presequences of the ALAS isozymes remains to be determined. Region 2 of the mature mitochondrial erythroid ALAS is considerably shorter than that of its housekeeping counterpart and these regions show limited sequence identity. The role, if any, of region 2 is not known at present.

We have recently shown that the gene for human erythroid ALAS is located on the X-chromosome (Cox *et al.*, 1990; Bishop *et al.*, 1990) and evidence from the present work strongly suggests that this gene is expressed exclusively in erythroid cells. Characterization of the gene for this ALAS isozyme (manuscript in preparation) has revealed it to consist of 11 exons spanning 22 kb, with a feature being the presence of a 6.0 kb intron in the 5′-UTR. A second gene, that for human housekeeping ALAS, has been localized on chromosome 3 (Sutherland *et al.*, 1988; Bishop *et al.*, 1990) and is apparently expressed ubiquitously to produce heme for respiratory cytochromes and other hemoproteins. This gene is also probably expressed in erythroid cells since the housekeeping mRNA can be detected in K562 and mouse erythroleukemic (MEL) cell lines (our unpublished data). However, the results of our Northern blots on human fetal liver show that while expression of erythroid mRNA is high in this erythroid tissue, that of the housekeeping mRNA is very low, indicating that the housekeeping ALAS isozyme does not contribute significantly to heme formation during erythropoiesis.

From their structural organization (data not shown), it seems probable that the human erythroid and housekeeping ALAS genes may well have evolved by duplication of an ancestral gene encoding a primitive catalytic protein with subsequent additions of DNA sequences encoding the different regions 1 and 2.

It is of interest that while two distinct genes exist for ALAS in the human haploid genome, there is only a single copy gene for PBGD, the third enzyme of the heme pathway. This

gene has two overlapping transcription units each with its own promoter (Chretien *et al.*, 1988) such that there is an upstream ubiquitous promoter and another downstream promoter active only in erythroid cells. Two mRNAs are generated by alternative splicing and these encode the PBGD isozymes found in non-erythroid and erythroid tissue (Mignotte *et al.*, 1989a; Raich *et al.*, 1989).

During erythroid cell differentiation it has been established that enhanced synthesis of the β-like globin chains results from increased transcriptional gene activation through the interaction of *trans*-acting factors with regulatory regions in the globin gene promoters (deBoer *et al.*, 1988; Wall *et al.*, 1988) and also in the far upstream erythroid-specific enhancer (Wall *et al.*, 1988; Martin *et al.*, 1989; Philipsen *et al.*, 1990; Talbot *et al.*, 1990). Studies on the regulation of the erythroid promoter for PBGD have identified *trans*-acting factors which are common to the β-globin gene promoter and upstream enhancer (Philipsen *et al.*, 1990; Talbot *et al.*, 1990; Mignotte *et al.*, 1989a and b). A structural analysis of the immediate promoter for the human erythroid ALAS gene has revealed a number of *cis*-acting control motifs. In addition to consensus TATA and CCAAT boxes at positions −27 and −87, respectively, there is a possible GATA-1 motif at −100 and a perfect NF-E2 consensus sequence at −40. GATA-1 is an erythroid-specific factor that has been shown to bind to multiple sites in the β-globin gene promoter and upstream enhancer (Wall *et al.*, 1988; Martin *et al.*, 1989; Philipsen *et al.*, 1990; Talbot *et al.*, 1990) and to sites in the erythroid promoter of PBGD (Mignotte *et al.*, 1989b). NF-E2 is an erythroid-specific factor that recognizes an AP-I-like sequence (Mignotte *et al.*, 1989a) and it was first identified in the erythroid PBGD promoter (Mignotte *et al.*, 1989a and b) and more recently in the upstream enhancer of the β-globin gene cluster (Philipsen *et al.*, 1990; Talbot *et al.*, 1990; Ney *et al.*, 1990). Also present in the promoter of the human erythroid ALAS gene are sequences around positions −50 and −55 which match the GnnnGGTGG sequence proposed by Philipsen *et al.* (1990) to play a key role in erythroid-specific expression through co-operation with GATA-1. Alternatively, on the minus strand at the same position, the sequence resembles overlapping consensus CAC boxes (Dierks *et al.*, 1983). Although the CAC box motif is recognized by ubiquitous factors (Dierks *et al.*, 1983) it is present in other erythroid promoters such as β-globin (Myers *et al.*, 1986) and PBGD (Mignotte *et al.*, 1989a). The presence in the human erythroid ALAS gene promoter of *cis*-acting elements identified in other erythroid-specific promoters, indicates that globin genes and genes for enzymes of the heme pathway will be subject to control by common *trans*-acting factors during erythropoiesis. Another interesting feature of the promoter for the human erythroid ALAS gene, is the presence of a possible thyroid hormone responsive element between nucleotides −21 and −6. Recent studies in avian erythroid cells have demonstrated that the thyroid hormone plays a role in the induction of erythroid differentiation (Zenke *et al.*, 1990) and evidence indicated that the thyroid hormone receptor, in the absence of thyroid hormone, repressed transcription of a subset of erythroid-specific genes including ALAS. The presence of a possible thyroid hormone responsive element near the TATA box of the human erythroid ALAS gene suggests that this gene may be regulated by thyroid hormone.

An interesting finding in the present work is the identification of an IRE in the 5′-UTR of the human erythroid ALAS gene. IREs are found in the 5′-UTR of ferritin mRNAs (Aziz and Munro, 1987; Hentze *et al.*, 1987; Leibold and Munro, 1988) and the 3′-UTRs of transferrin receptor mRNAs (Casey *et al.*, 1988; Müllner *et al.*, 1989) from different species and are important for controlling iron metabolism by the regulation of ferritin and transferrin receptor synthesis through iron-dependent post-transcriptional mechanisms. As stated earlier, a cytosolic protein (IRE-BP) binds to the IREs and inhibits ferritin mRNA translation and increases stability of the transferrin receptor mRNA; cellular iron levels modulate IRE-BP activity with decreased activity observed when levels are elevated (Theil, 1990). The ALAS IRE sequence displays features characteristic of the ferritin and transferrin receptor IRE motifs. A comparison of different ferritin IRE motifs has shown that nucleotides 1−28 (see Figure 6b) are highly conserved and *in vivo* experiments have established that these nucleotides are obligatorily required for the functional response of the IRE to altered iron levels (Theil, 1990). However, extended secondary structures for the 5′-UTR of ferritin mRNAs, consisting of sequences flanking the IRE, have also been proposed (Aziz and Munro, 1987; Casey *et al.*, 1988, and see Figure 6) and may assist in the inhibition of translational initiation by bound IRE-BP (Wang *et al.*, 1990). From Figure 6 it can be seen that the ALAS IRE and its flanking sequences can form a secondary structure similar to the 5′-UTR of ferritin mRNAs although the nucleotide sequences differ markedly. It has recently been shown that iron-dependent repression is not observed when the ferritin IRE is positioned in the 5′-UTR at a site 67 nucleotides or more from the 5′ end of the mRNA (Goossen *et al.*, 1990). Although the reason for this remains unclear, it is noteworthy that the human ALAS IRE (nucleotides 1−28) is located 10 nucleotides from the 5′- end of its mRNA and thus might be expected to be functionally active. It is also of interest that the IRE motif reported recently in the 5′-UTR of the mouse erythroid ALAS mRNA (Dierks, 1990) is almost identical to that of the human ALAS IRE and that an IRE motif is not present in the mRNAs for human (Bawden *et al.*, 1987), rat (Srivastava *et al.*, 1988) or chicken (Borthwick *et al.*, 1985) housekeeping ALAS isozymes.

Importantly, we have shown using gel retardation experiments with IRE transcripts and cytosolic extracts that the ALAS IRE sequence binds a cytosolic protein that is most likely identical or closely related to that which binds to the IREs of ferritin and transferrin receptor mRNAs, although the possibility of an erythroid-specific IRE-BP that binds to the ALAS IRE cannot be discounted. The precise regions of the ALAS IRE required for binding the protein, remain to be investigated but an alteration of the conserved loop in the IRE markedly reduces binding. We have shown that a truncated stem of the ALAS IRE containing potential stem base pairing from nucleotide +8 to +38 (including nucleotides 1 to 28 in Figure 6) can still bind IRE-BP, as found for ferritin (Casey *et al.*, 1988; Müllner *et al.*, 1989; Barton *et al.*, 1990; Leibold *et al.*, 1990). However, it is possible that the remainder of the proposed stem secondary structure enhances binding of the IRE-BP to the ALAS IRE *in vivo*. It is perhaps noteworthy that this additional sequence required for the extended base pairing in the stem is encoded by exon 2 of the human erythroid ALAS gene with exon

1 essentially encoding nucleotides 1−28 of the IRE and the nucleotides 5′ of this element (see Figure 6a).

On the basis of the gel retardation results, we propose that the ALAS IRE is functional *in vivo* and that at low intracellular iron concentrations, binding of an IRE-BP prevents translation of the mRNA. It can be envisaged that the amount of ALAS activity in differentiating erythroid cells is subject to at least two controls—mRNA levels will be increased by transcriptional activation of the ALAS gene through the action of erythropoietin but subsequent translation of this mRNA will be dependent on iron availability. There is evidence that iron uptake by erythroid cells limits the overall rate of heme synthesis (Gardner and Cox, 1988; Ponka *et al.*, 1988). Hence, when iron enters the cell, translation of ALAS mRNA will be increased resulting in increased protoporphyrin production so that protoporphyrin formation is coupled to iron availability for subsequent heme formation. In turn, translation of globin chains are dependent upon the availability of heme (London *et al.*, 1987). Current experiments are in progress to confirm that in differentiating erythroid cells, translation of erythroid ALAS mRNA responds to altered iron levels in the fashion predicted. In addition, the roles of erythroid-specific and ubiquitous *trans*-acting factors in regulating expression of the human erythroid ALAS gene during erythroid differentiation are under investigation.

Materials and methods

Materials

Nitrocellulose (BA85) was purchased from Schleicher and Schuell; restriction endonucleases were from Integrated Sciences Inc. and New England Biolabs; [γ-^{32}P]ATP, [α-^{32}P]dATP and [α-^{32}P]rUTP, T_4 polynucleotide kinase and oligolabelling kits (OLK-A) were purchased from Bresatec, Thebarton, South Australia.

Northern blot analysis and cDNA library screening

A chemically synthesized 50mer (5′-CGAATGCCTTGGATCATGG AGGCATGATTGCCTGCATCTGAGTAGATCTC-3′) complementary to the mouse erythroid ALAS mRNA (Schoenhaut and Curtis, 1986; positions 347−363 in Figure 4, this manuscript) but derived from a highly conserved region of 17 amino acids found in all animal ALAS proteins (Elliott *et al.*, 1989) was 5′ phosphorylated with [γ-^{32}P]ATP and T_4 polynucleotide kinase.

Northern blot analysis was carried out using 40 μg total RNA and electrophoresis on 1.0% agarose gels containing 1.1 M formaldehyde as previously described (Elferink *et al.*, 1988). The RNA was transferred to nitrocellulose filters and hybridized to either the labelled 50mer or [α-^{32}P]dATP oligolabelled human housekeeping ALAS cDNA (Bawden *et al.*, 1987) in a solution containing 5 × SSPE (0.9 M NaCl, 50 mM sodium phosphate buffer, pH 7.0, 5 mM EDTA), 5 × Denhardt's (0.04% Ficoll, 0.04% polyvinylpyrrolidone, 0.04% bovine serum albumin), 0.1% SDS, 0.5% sodium pyrophosphate, 200 μg/ml denatured salmon sperm DNA and 50% formamide for the cDNA probe only at 42°C for 18 h. The filter probed with the cDNA clone was washed finally in 0.1 × SSPE containing 0.1% SDS at 65°C for 40 min. The filter probed with the 50mer was washed finally in 2 × SSPE containing 0.1% SDS at 48°C for 20 min. The filters were then exposed, using Kodak XAR-5 film. The 50mer filter was subsequently stripped (Maniatis *et al.*, 1982) and reprobed with [α-^{32}P]dATP oligolabelled human erythroid ALAS cDNA (pHEA-6) under the same conditions as described above. The filter was washed as for the housekeeping cDNA probed filter, followed by autoradiography. The ^{32}P-labelled 50mer was used to screen a human first trimester fetal liver library in λgt11, according to the standard procedure of Maniatis *et al.* (1982). Briefly, following plaque transfer to nitrocellulose filters and subsequent phage lysis, filters were baked for 1 h at 80°C *in vacuo* and hybridized to the 50mer under the conditions described above for the Northern blot analysis, in the absence of formamide. From a total of 5×10^5 recombinant phage, fifty clones with positive hybridization signals were picked. Of these, the twelve strongest were purified through three subsequent rounds of screening.

Nested deletion generation and sequencing of cDNA clones

Restriction enzyme analysis of the twelve purified clones showed that they shared common restriction fragments with the inserts ranging in size from ~800 bp to 1.9 kb. Following subcloning of the inserts into the *Eco*RI site of the phagemid vector, pTZ19 (Pharmacia), generating both orientations, the ends of each clone were sequenced using the Sequenase Version 2.0 sequencing kit (US Biochemicals). To determine the remaining internal sequence of human erythroid ALAS, the longest confirmed clone, pHEA-6, was chosen and following linearization with the restriction endonucleases *Sph*I and *Bam*HI, nested deletion generation was carried out according to the Erase-A-Base protocol (Promega) and the resultant clones sequenced. In addition, sequence of the ends of the *Sac*I and *Pst*I restriction fragments was carried out after subcloning into pTZ19.

Primer extension analysis

Primer extension analysis was carried out according to the method of McKnight *et al.* (1981) using AMV reverse transcriptase (Pharmacia), total RNA from human first trimester fetal liver and a 5′-γ-^{32}P-labelled synthetic 19mer. (5′-CTGTTGCCCTGCACTGAGG-3′) complementary to nucleotides 19−37 in Figure 2.

Cells and preparation of cytosolic extracts

Human K562 and HL60 cells were grown in RPMI with 10% heat-inactivated fetal calf serum. Mouse J2E-1 cells were grown in Dulbecco's modified Eagle's medium with 10% heat-inactivated fetal calf serum. All cells were incubated in 5% CO_2 at 37°C. Cytosolic extracts were prepared from each of these cell lines using the procedure of Müllner *et al.* (1989) as modified from that of Leibold and Munro (1988). Protein concentrations were determined using the Bio-Rad protein micro-assay (Bio-Rad, Richmond, CA).

Synthesis and mutagenesis of the 5′-UTR of human erythroid ALAS

A 70mer (5′-GTACCCACCTGTCATTCGTTCGTCCTCAGTGCAGG-GCAACAGGACTTTAGGTTCACCATGGTGACTGCAG-3′) and its complement were synthesized. These comprised the entire 5′-UTR of human erythroid ALAS and flanked by the native *Rsa*I site just 5′ of the transcription start site and the native *Pst*I site just 3′ of the putative ATG codon. For subsequent linearization in *in vitro* transcription reactions, the A and G residues at positions −2 and −1, respectively, relative to the ATG were replaced by C residues in the design of the 70mers to introduce an *Nco*I site. Hybridization of the two 70mers was carried out in an FTS-1 Thermal Sequencer (Corbett Research) at 95°C for 5 min, followed by 80°C for 15 min then snap cooled on ice. The hybridized oligomers were then blunt-end ligated into the *Sma*I site of the transcription vector, pSP73 (Promega), generating pSPIA-WT. An *Rsa*I genomic fragment spanning exon 1 was similarly cloned into pSP73 generating pSPIA-E1. Site-directed mutagenesis of the loop of the ALAS IRE was done by creation of a *Bam*HI site generating pSPIA-LM using the following 33mer: 5′-AGTCCTGTTGCCCTGG-ATCCAGGACGAACGAAT-3′. Mutant clones were detected by digestion of plasmid DNA from the resulting colonies with *Bam*HI. A second round of transformation was carried out in order to select pure mutant clones (Maniatis *et al.*, 1982). All clones were confirmed by sequencing.

In vitro transcription

For generation of ALAS IRE transcripts, pSPIA-WT and pSPIA-LM were linearized with *Nco*I and pSPIA-E1 was internally linearized with *Pvu*II. To ensure that pSP73 polylinker sequence was not affecting the subsequent protein−RNA interactions (described below), this sequence was minimized in some clones by purification of *Bgl*II and *Kpn*I digested vector, end-filling and religation. pSPT-fer (human ferritin heavy chain IRE) and pSPT-TR34 (human transferrin receptor, IRE palindrome E) were used as described (Müllner *et al.*, 1989) for synthesis of ferritin and transferrin receptor IRE transcripts. Transcription reactions were carried out using a Bresatec RTK-U RNA transcription kit, utilizing 100 μCi of [α-^{32}P]rUTP (3000 Ci/mmol), 15 μM rUTP and 0.5 mM rATP, rGTP and rCTP with 4−8 U or either SP6 or T7 polymerase where required. Unlabelled competitor RNAs were synthesized in the presence of all four rNTPs (0.5 mM). To aid in quantification of the competitor, a trace amount of [α-^{32}P]rUTP was also added to the synthesis reaction. Full-length transcripts were purified by elution into TE (10 mM Tris pH 7.5, 1 mM EDTA pH 8.0) buffer for at least 4 h at 37°C following fractionation on a 7 M urea, 12% (24:1) polyacrylamide gel. The specific activity of labelled transcripts was routinely $3-5 \times 10^8$ c.p.m./μg.

Analysis of protein–RNA interactions

Binding reactions were carried out using the modified protocol of Leibold and Munro (1988) as described by Müllner *et al.* (1989). The subsequent RNA–protein complexes were treated with 1 U of ribonuclease T1 (Calbiochem) for 10 min then 5 mg/ml heparin (Sigma) and analyzed on 1.5 mm, 6% non-denaturing polyacrylamide gels (80:1 acrylamide:bis-acrylamide). Gels were pre-electrophoresed for 20–30 min at 10 V/cm prior to running for 3 h at the same voltage. Gels were then dried onto 3MM chromatography paper (Whatman) and autoradiographed as required. For competition reactions, unlabelled RNA was added together with the labelled RNA prior to addition of the cytosolic extract.

Acknowledgements

The authors wish to thank Dr Lukas Kühn (Swiss Institute for Experimental Cancer Research) for kindly supplying the plasmids, pSPT-fer and pSPT-TR34, and for information regarding protein–RNA interactions. The contribution of Dr Raman Bhasker in this work is greatly appreciated. We also wish to thank Professor Nader Abraham (New York Medical College) for generously providing the cDNA library and RNA samples and Dr Peter Klinken for supplying the mouse erythroid J2E-1 cell line. T.C.C. is the recipient of a National Health and Medical Research Council Biomedical Postgraduate Scholarship.

References

Avissar,Y.J. and Beale,S.I. (1989) *J. Bacteriol.*, **171**, 2919–2924.
Aziz,N. and Munro,H.N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8478–8482.
Barton,H.A., Eisenstein,R.S., Bomford,A. and Munro,H.N. (1990) *J. Biol Chem.*, **265**, 7000–7008.
Bawden,M.J., Borthwick,I.A., Healy,H.M., Morris,C.P., May,B.K. and Elliott,W.H. (1987) *Nucleic Acids Res.*,**15**, 8563.
Beato,M. (1989) *Cell*, **56**, 335–344.
Beaumont,C., Deybach,J.C., Grandchamp,B., Silva,V.D., de Verneuil,H. and Nordmann,Y. (1984) *Exp. Cell Res.*, **154**, 474–484.
Bishop,D.F., Henderson,A.S. and Astrin,K.H. (1990) *Genomics*, **7**, 207–214.
Borthwick,I.A., Srivastava,G., Day,A.R., Pirola,B.A., Snoswell,M.A., May,B.K. and Elliott,W.H. (1985) *Eur. J. Biochem.*, **150**, 481–484.
Bottomley,S. and Müller-Eberhard,U. (1988) *Semin. Hematol.*, **25**, 282–302.
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.
Casey,J.L., Koeller,D.M., Ramin,V.C., Klausner,R.D. and Harford,J.B. (1989) *EMBO J.*, **8**, 3693–3699.
Caughman,S.W., Hentze,M.W., Rouault,T.A., Harford,J.B. and Klausner,R.D. (1988) *J. Biol. Chem.*, **263**, 19048–19052.
Chretien,S., Dubart,A., Beaupain,D., Raich,N., Grandchamp,B., Rosa,J., Goossens,M. and Romeo, P.-H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6–10.
Cox,T.C., Bawden,M.J., Abraham,N.G., Bottomley,S.S., May,B.K., Baker,E., Chen,L.Z. and Sutherland,G.R. (1990) *Am. J. Hum. Genet.*, **46**, 107–111.
deBoer,E., Antoniou,M., Mignotte,V., Wall,L. and Grosveld,F. (1988) *EMBO J.*, **7**, 4203–4212.
Dierks,P. (1990) In Dailey,H.A. (ed.), *Biosynthesis of Heme and Chlorophylls*. McGraw-Hill, New York, pp. 201–233.
Dierks,P., van Ooyen,A., Cochran,M.D., Dobkin.C., Reiser,J. and Weissmann,C. (1983) *Cell*, **32**, 695–706.
Elferink,C.J., Sassa,S. and May, B.K. (1988) *J. Biol. Chem.*, **263**, 13012–13016.
Elliott,W.H., May,B.K., Bawden,M.J. and Hansen,A.J. (1989) In Kay,J., Ballard,F.J. and Mayer,R.J. (eds), *Gene Expression: Regulation at the RNA and Protein Levels*. Biochemical Society Symposium 55, The Biochemical Society, London, pp. 13–27.
Gardner,L.C. and Cox,T.M. (1988) *J. Biol. Chem.*, **263**, 6676–6682.
Glass,C.K., Holloway,J.M., Devary,O.V. and Rosenfeld,M.G. (1988) *Cell*, **542**, 313–323.
Gooseen,B., Caughman,S.W., Harford,J.B., Klausner,R.D. and Hentze,M.W. (1990) *EMBO J.*, **9**, 4127–4133.
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.
Karlsson,S. and Nienhuis,A.W. (1985) *Annu. Rev. Biochem.*, **54**, 1071–1108.
Klinken,S.P., Nicola,N.A. and Johnson,G.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8506–8510.
Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857–872.
Leibhaber,S.A., Goossens,M.J. and Wai Kan,Y. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7054–7058.
Leibold,E.A. and Munro,H.N. (1988) *J. Biol. Chem.*, **262**, 7335–7341.
Leibold,E.A., Laudano,A. and Yu,Y. (1990) *Nucleic Acids Res.*, **18**, 1819–1824.
Leong,S.A., Williams,P.H. and Ditta,G.S. (1985) *Nucleic Acids Res.*, **13**, 5965–5976.
London,I.M., Levin,D.H., Matts,R.L., Thomas,N.S.B., Petryshyn,R. and Chen,J.-J. (1987) In Boyer,P.D. and Krebs,E.G. (eds), *The Enzymes*. Academic Press, Orlando, FL, Vol. XVIII, Part B, pp. 360–380.
Maguire,D.J., Day,A.R., Borthwick,I.A., Srivastava,G., Wigley,P.L., May,B.K. and Elliott,W.H. (1986) *Nucleic Acids Res.*, **14**, 1379–1391.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Martin,D.I.K., Tsai,S. and Orkin,S.H. (1989) *Nature*, **38**, 435–438.
May,B.K., Borthwick,I.A., Srivastava,G., Pirola,B.A. and Elliott,W.H. (1986) In Horecker,B.L. and Stadtman,E.R. (eds), *Current Topics in Cellular Regulation*. Academic Press, New York, Vol. 28, pp. 233–262.
May,B.K., Bhasker,C.R., Bawden,M.J. and Cox,T.C. (1990) *Mol. Biol. Med.*, **7**, 405–421.
McClung,C.R., Sommerville,J.E., Guerinot,M.L. and Chelm,B.K. (1987) *Gene*, **54**, 133–139.
McKnight,S.L., Gavis,E.R. and Kingsbury,R. (1981) *Cell*, **25**, 385–398.
Mignotte,V., Eleouet,J.F., Raich,N. and Romeo,P.-H. (1989a) *Proc. Natl. Acad. Sci. USA*, **86**, 6548–6552.
Mignotte,V., Wall,L., deBoer,E., Grosveld,F. and Romeo,P.-H. (1989b) *Nucleic Acids Res.*, **17**, 37–54.
Müllner,E.W., Neupert,B. and Kühn,L.C. (1989) *Cell*, **58**, 373–382.
Myers,R.M., Tilly,K. and Maniatis,T. (1986) *Science*, **232**, 613–618.
Ney,P.A., Sorrentino,B.P., Lowrey,C.H. and Nienhuis,A.W. (1990) *Nucleic Acids Res.*, **18**, 6011–6017.
Orkin,S.H. (1990) *Cell*, **63**, 665–672.
Philipsen,S. Talbot,D., Fraser,P. and Grosveld,F. (1990) *EMBO J.*, **9**, 2159–2167.
Ponka,P., Schulman,H.M. and Martinez-Medellin,J. (1988) *Biochem. J.*, **251**, 105–109.
Raich,N., Mignotte,V., Dubart,A., Beaupain,D., Leboulch,Ph., Romana,M., Chabret,C., Charnay,P., Papayannopoulou,T., Goossens,M. and Romeo,P.-H. (1989) *J. Biol. Chem.*,**264**, 10186–10192.
Riddle,R.D., Yamamoto,M. and Engel,J.D. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 792–796.
Romeo,P.-H., Raich,N., Dubart,A., Beaupain,D., Pryor,M., Kusner,J.. Cohen-Solal,M. and Goossens,M. (1986) *J. Biol. Chem.*, **261**, 9825–9831.
Rothenberger,S., Müllner,E.W. and Kühn,L.C. (1990) *Nucleic Acids Res.*, **18**, 1175–1179.
Schoenhaut,D.S. and Curtis,P.J. (1986) *Gene*, **48**, 55–63.
Schoenhaut,D.S. and Curtis,P.J. (1989) *Nucleic Acids Res.*, **17**, 7013–7028.
Srivastava,G., Borthwick,I.A., Maguire,D.J., Elferink,C.J., Bawden,M.J., Mercer,J.F.B. and May,B.K. (1988) *J. Biol. Chem.*, **263**, 5202–5209.
Sutherland,G.R., Baker,E., Callen,D.F., Hyland,V.J., May,B.K., Bawden,M.J., Healy,H.M. and Borthwick,I.A. (1988) *Am. J. Hum. Genet.*, **43**, 331–335.
Talbot,D., Philipsen,S., Fraser,P. and Grosveld,F. (1990) *EMBO J.*, **9**,. 2169–2178.
Theil,E.C. (1990) *J. Biol. Chem.*, **265**, 4771–4774.
Urban-Grimal,D., Volland,C., Garnier,T., Dehoux,P. and Labbe-Bois,R. (1986) *Eur. J. Biochem.*, **156**, 511–519.
Wall,L., deBoer,E. and Grosveld,F. (1988) *Genes Dev.*, **2**, 1089–1100.
Wang,Y.-H., Sczekan,S.R. and Theil,E.C. (1990) *Nucleic Acids Res.*, **18**, 4463–4467.
Worwood,M. (1977) In Müller-Eberhard,U., Miescher,P.A. and Jaffé,E.R. (eds), *Iron Excess: Aberrations of Iron and Porphyrin Metabolism*. Grune and Stratton, Inc., New York, pp. 3–30.
Zenke,M., Munoz,A., Sap,J., Vennstrom,B. and Beug,H. (1990) *Cell*, **61**, 1035–1049.

Received on March 19, 1991