Structure and Expression of the Chicken Ferritin H-Subunit Gene

PRISCILLA WILKINS STEVENS,¹ JERRY B. DODGSON,² and JAMES DOUGLAS ENGEL^{1*}

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60201,¹ and Departments of Microbiology and Biochemistry, Michigan State University, East Lansing, Michigan 48824²

Received 28 October 1986/Accepted 26 January 1987

Although the genomes of many species contain multiple copies of ferritin heavy (H)- and light (L)-chain sequences, the chicken genome contains only a single copy of the H-subunit gene. The primary transcription unit of this gene is 4.6 kilobase pairs and contains four exons which are posttranscriptionally spliced to generate a mature transcript of 869 nucleotides. Chicken and human ferritin H-subunit genomic loci are organized with similar exon-intron boundaries. They exhibit approximately 85% nucleotide identity in coding regions, which yield proteins 93% identical in amino acid sequence. We have identified a sequence of 22 highly conserved nucleotides in the 5' untranslated sequences of chicken, human, and tadpole ferritin H-subunit genes and propose that this conserved sequence may regulate iron-modulated translation of ferritin H-subunit mRNAs.

Ferritin is an ubiquitous iron storage protein which maintains intracellular iron in a form which is biologically accessible and readily mobilized for cellular metabolic processes. Each ferritin molecule is an essentially spherical arrangement of 24 protein subunits encompassing a core of up to 4,500 iron atoms (see references 28 and 38 for reviews). Ferritin subunits are of two types, heavy (H) and light (L), of molecular weight 21,000 and 19,000, respectively (2). The subunits are similar in three-dimensional structure; each is a bundle of four long helices (A through D) with a fifth short helix (E) forming an acute angle to the bundle (31). The H and L subunits appear to be interchangeable within the protein shell (29), generating a large number of isoferritins, certain populations of which are characteristic for particular cell types (2).

Tissue-specific differences in the ratios of H and L subunits appear to be determined by mRNA levels (11), but synthesis of protein subunits is regulated posttranscriptionally by the intracellular iron concentration. When iron is limiting, ferritin mRNAs are stored within the cell but are not translated. In the presence of iron, these preexisting RNAs are recruited by polysomes and actively translated (4, 36, 44).

H and L subunits are derived from different genes, but the number of genes encoding each subunit is not known (22). Multiple copies of H- and L-subunit sequences exist in human, rat, and tadpole genomes (8, 11, 13, 14), but it is not clear how many loci are transcribed and processed into mature mRNAs. Many of the human genomic loci contain intronless pseudogenes flanked by direct repeats. It is therefore possible that only a single human ferritin H-subunit locus is transcriptionally active, whereas 10 to 15 other loci harbor nontranscribed pseudogenes (12). We have isolated genomic recombinants and cDNA clones for the chicken ferritin H subunit and find no evidence for H-subunit pseudogenes in the chicken genome.

MATERIALS AND METHODS

Isolation of chicken ferritin H-subunit genomic and cDNA clones. In a difference screen (34) to obtain genes expressed at high levels in erythroid cells but not in other cell types, a group of 22 nonoverlapping, non-globin-encoding λ recom-

binants were selected from a λ Charon 4A library of chicken chromosomal DNA (15) when the library was probed with cDNA synthesized from anemic hen reticulocyte (RBC) poly(A)⁺ RNA. These 22 nonglobin, RBC-positive clones were further screened by DNA dot blots and RNA blot hybridization to identify recombinants whose sequences are more prevalent in RBC mRNA than in mRNA from other cell types. The genomic clone λ Ch4A-56D selected in this screen was subsequently shown by sequence analysis to contain the 3' portion of the gene for the chicken ferritin H subunit.

Ferritin H-subunit cDNA clones were isolated from a λ gt10 cDNA library prepared from anemic hen RBC mRNA (43) by using the exon-containing, 2.8-kilobase-pair (kbp) *Eco*RI-*Hin*dIII fragment of λ Ch4A-56D as probe. Of five cDNA clones plaque purified and sequenced (clones 5C, 20A, 21D, 22, and 29), two contain the entire coding sequence and 3' untranslated region (clones 20A and 22), but none contains the entire 5' untranslated region.

 λ EMBL3-56D, a genomic recombinant containing the entire ferritin H-subunit gene, was then isolated by using 5' and 3' subfragments of cDNA clone 21D to probe a λ EMBL3 library containing 18-kbp *Mbo*I partial digestion fragments of HD6 cell chromosomal DNA (M. Yamamoto, unpublished data) constructed as previously described (10). A 16.3-kbp *Bam*HI fragment in λ EMBL3-56D includes approximately 8.2 kbp 5' and 3.5 kbp 3' to the 4.6-kbp primary transcription unit of the ferritin H-subunit locus.

RNase and S1 nuclease protection analyses. $[\alpha^{-32}P]$ GTPlabeled antisense RNAs of different lengths were generated by Sp6 polymerase from a 3.9-kbp genomic *Eco*RI fragment of λ Ch4A-56D subcloned into pSp65 (26). For the shortest probe (0.7 kb), 10⁶ cpm was used for hybridization; for the other probes, the number of counts hybridized was directly proportional to probe length. RNA probes were hybridized to 0.25 µg of anemic hen RBC poly(A)⁺ RNA for 3 h at 37°C and then treated with RNase A and RNase T₁ at 37°C for 30 min as described previously (9, 45). Double-stranded RNA protection fragments were separated on a nondenaturing 3.5% polyacrylamide gel and visualized by autoradiography.

For the S1 nuclease protection, a 510-nucleotide *NcoI*-StuI fragment from λ EMBL3-56D was kinase labeled with $[\gamma^{-32}P]$ ATP at the *NcoI* site and hybridized to 5 µg of anemic hen RBC total RNA at 42°C for 12 h. Hybridization reactions were quenched on ice and then digested with 1,000 U of S1

^{*} Corresponding author.



FIG. 1. Structural analysis of genomic recombinants and the transcribed region of the chicken ferritin H-subunit gene. (A) Selected restriction endonuclease maps of the chicken ferritin H-subunit genomic locus shown in B. Vertical lines indicate cleavage positions of the enzyme abbreviated at the left. (B) A map of the chicken ferritin H-subunit gene locus from λ EMBL3-56D is shown, beginning at the *KpnI* site 5' to the first exon and terminating at the *EcoRI* site 3' to the final exon. Boxes depict the relative positions of the four exons encoding the ferritin H-subunit within a 7.1-kbp segment of the genome. Symbols: \Box , coding sequence; \blacksquare , 5' (left) and 3' (right) nontranslated regions. The 3.9-kbp fragment common to λ Ch4a-56D and λ EMBL3-56D is indicated above the map, together with the 2.8-kbp subfragment mentioned in the text. (C) Structure of the transcribed region of the ferritin H-subunit gene as deduced from data shown below (Fig. 2 and 3). Coding and noncoding sequences are represented as in B. (D) Selected restriction endonuclease maps of the transcribed region of the chicken ferritin H-subunit cDNA. Symbols on the first line mark the cleavage sites of enzymes which cut only once in this region. For the other enzymes listed at the left, vertical lines indicate cleavage positions. Abbreviations: A, AvaI; D, DdeI; H, HindIII; HF, HinfI; K, KpnI; N, NcoI; PS, PstI; PV, PvuII; RI, EcoRI; RS, RsaI; S3, Sau3AI; SP, SphI; ST, StuI; XB, XbaI; XH, XhoI.

nuclease per ml at 30° C for 2 h. Hybridization conditions and S1 nuclease buffers were as previously described (41).

DNA sequence determination. DNA sequencing was accomplished by the dideoxy chain termination method (33) by using M13 subcloning (27) and also by using synthetic primers complementary to Sp6 and T7 promoters to prime dideoxynucleotide-truncated Klenow synthesis from alkalidenatured pGEM plasmid subclones (Promega Biotec, Madison, Wis.). Some regions of the genomic clones were sequenced by a modified chemical cleavage method (25, 37). For all sequence determinations, both DNA strands were independently sequenced.

Cells and cell lines. Chicken embryo fibroblasts (CEF) were prepared from 11-day-old embryos by the procedure of Vogt (40). MSB-1 is a lymphocytic cell line derived from a splenic lymphoma of a chick with Marek's disease (1). HD6 is a continuous definitive erythroid cell line transformed with the ts167-AEV allele of avian erythroblastosis virus (5, 30). Transformation of erythroid precursor cells with avian erythroblastosis virus appears to block erythroid differentiation and arrest infected precursor cells at the CFU-E stage (32).

Isolation of RNA samples. Nonerythroid cells and tissues were homogenized in guanidine thiocyanate solution, and RNA was banded in CsCl as described previously (19). For isolation of total cellular RNA from erythroid cells, a modified procedure was developed. About 2 ml of wet-packed RBC isolated from chronically anemic hens (16) were lysed

in 16 ml of guanidine thiocyanate-imidazole solution (50%) guanidine thiocyanate, 0.5% Sarkosyl, 25 mM sodium citrate, 0.1 M \beta-mercaptoethanol, 0.1% antifoam A [Sigma Chemical Co.], 0.16 M imidazole; final solution adjusted to pH 5.5 with glacial acetic acid) at maximum speed for 2 min in an Omnitron tissumizer. Cell debris was pelleted and discarded. To the supernatant was added EDTA (pH 8.0) to 25 mM, sodium acetate (pH 5) to 50 mM, and solid CsCl to 1 g/ml. After centrifugation at 12,000 \times g for 20 min at 20°C (SS34 rotor), the protein rising to the top of the supernatant solution was discarded, as was the residual pellet. The cell lysate-CsCl solution was transferred into 38.5-ml Beckman tubes, underlaid with 3 to 10 ml of $\rho = 1.8$ CsCl solution, sealed, and then centrifuged at 178,000 \times g at 20°C for 40 to 50 h (60 Ti rotor). The opalescent RNA band was removed from the side of the tube with a syringe and processed as previously described (19). $Poly(A)^+$ RNA was selected by oligo(dT)-cellulose chromatography (3).

RNA and DNA blot analysis. RNA samples were denatured in formaldehyde solution, electrophoresed on an agaroseformaldehyde gel, transferred to nitrocellulose (20, 23), and probed with radiolabeled antisense RNA generated by Sp6 polymerase transcription of the 5C ferritin H-subunit cDNA fragment cloned into pSp65 (26). The filter was washed and then treated with 8 μ g of RNase A per ml in fresh RNase buffer as described previously (43). The chromosomal DNA blot was prepared by standard procedures (18) and probed with a nick-translated insert from cDNA clone 20A. The filter was washed at 42°C four times for 30 min each in 50



FIG. 2. RNase and S1 nuclease mapping experiments. (A) Four radiolabeled antisense RNA probes of different lengths were transcribed from a pSp65 subclone of the 3.9-kbp EcoRI fragment of λ Ch4A-56D (Fig. 1B) truncated at restriction enzyme sites within the insert fragment (X, A, T) or at the BamHI site of the vector polylinker sequence (B). (B) Each probe was independently hybridized to 0.25 µg of anemic hen RBC poly(A)⁺ RNA and then treated with RNase A and RNase T₁ to digest all nonhybridized RNA; protected fragments were then separated on a neutral 3.5% polyacrylamide gel (9). The 0.7-kb probe (X) does not hybridize to RBC mRNA. The 1.5 kb probe (A) protects a fragment of 195 nucleotides, which is not seen with the next larger probes. Thus the Aval site interrupts an exon of 334 nucleotides. The 2.7-kb probe (T) protects two exons 334 and 126 nucleotides in length. No additional exons are protected by the full-length 3.9-kb probe (B). (C) Probe for the S1 nuclease mapping experiment was a 510-nucleotide NcoI-StuI genomic fragment (nucleotides 918 to 1428 of Fig. 3) radiolabeled at the NcoI site, which contains the translational start codon. The cDNA map indicates coding sequences () and noncoding regions) of the largest isolated cDNA recombinant, clone 22, which terminates 25 nucleotides 5' to the AUG start codon (nucleotide 127 of the transcribed sequence shown in Fig. 3). (D) Anemic hen RBC total RNA (5 µg) was hybridized to the probe shown in C, digested with S1 nuclease, and electrophoresed on a 5% polyacrylamide-50% urea gel in parallel with radiolabeled DNA restriction enzyme fragment markers (41). The protected fragment and DNA markers were visualized by autoradiography. Abbreviations: A, AvaI; B, BamHI; K, KpnI; N, NcoI; RI, EcoRI; ST, StuI; T, TaqI; X, XmnI; BP, base pairs; NT, nucleotides.

mM Tris (pH 7.5), 1 mM EDTA, 0.1% sodium dodecyl sulfate, 1 mM Na₂PO₄, and 1 mM sodium pyrophosphate.

RESULTS

Isolation and structural analysis of genomic and cDNA clones. Genomic clone λ Ch4A-56D was isolated from a chicken RBC chromosomal DNA library (15) by using dif-

ference methods (34; see Materials and Methods) to screen for genes transcribed at high levels in erythroid cells but at much lower levels in nonerythroid tissues and cells. A 2.8-kbp *Eco*RI-*Hin*dIII subclone of λ Ch4A-56D (Fig. 1B) shown to contain two exons of the chicken ferritin H-subunit gene (Fig. 2A and B) was used to probe a λ gt10 cDNA library (43). Five independent cDNA clones were plaque purified and sequenced; all five clones are of identical sequence in the overlapping regions. Fragments from the 5' and 3' ends of cDNA clone 21D were then used to screen a λ EMBL3 chicken chromosomal library to obtain a genomic clone spanning the chicken ferritin H-subunit gene (see Materials and Methods). The clone selected, λ EMBL3-56D, contains four exons which encode the entire chicken ferritin H subunit (Fig. 1B).

Nucleotide sequence analysis of the chicken ferritin Hsubunit gene. We have sequenced a 7.1-kbp segment of the chicken genome which contains the entire ferritin H-subunit locus (Fig. 3), including 1.2-kbp 5' and 1.3-kbp 3' flanking sequences (Fig. 1B). The ferritin H-subunit transcript is generated from four exons, which together with three introns span 4.6 kbp in the chicken genome and are processed to an mRNA of 869 nucleotides (Fig. 1C).

A single transcriptional initiation site 154 nucleotides 5' to the translational start codon was identified by S1 nuclease mapping (Fig. 2C and D). A standard TATA box is located 31 nucleotides 5' to the transcriptional initiation site, but we do not find an adjacent CCAAT box (7). Immediately abutting the first exon are sequences extremely rich in guanine (G) and cytidine (C) nucleotides. Whereas the average G+C content of the chicken ferritin H-subunit genomic locus is 40 to 45%, the 100 bases 5' to exon 1 are greater than 75% G+C, and the 150 bases 3' to exon 1 are nearly 90% G+C.

Locations of introns within the chicken ferritin H-subunit sequence correspond precisely to their positions within the intron-containing human ferritin H-subunit gene, whose genomic sequence has also been determined (12, 21). The first intron is roughly 1.5 kbp in both chicken and human sequences, but the latter two introns are each approximately 7 times longer in the chicken gene: 1,491 versus 256 nucleotides for intron 2 and 725 versus 95 nucleotides for intron 3 (12, 21). Nucleotide sequences of the chicken and human coding regions are quite homologous, with 85, 84, 79, and 82% identity of coding regions for exons 1 through 4, respectively. Untranslated regions of the chicken and human H-subunit genes are more divergent, with only about 60% identity of the 5' ends and about 40% identity of the 3' ends with the algorithm of Wilbur and Lipman (42).

Although overall homology in untranslated regions is poor, a segment of 22 nucleotides toward the 5' end is identical within the chicken and human ferritin H-subunit genes (nucleotides 41 to 62 of the transcribed region shown in Fig. 3). These 22 nucleotides occur within the context of 39 nucleotides, of which only 4 differ between the chicken and human sequences. The 22-nucleotide homology is maintained with only two substitutions (lowercase letters) within the 5' untranslated region of the tadpole RBC ferritin sequence (14):

chicken	³⁰ GGGTTCCTGCGTCAACAGTGCTTGGAC
	GGAACCGGCCGC
human	²⁸ GGtTTCCTGCtTCAACAGTGCTTGGACG
	GAACCcGgCGC
tadpole	²⁴ GaGTTCtTGCtTCAACAGTGtTTGaAC
	GGAACCctCtct

- 1				
1	ggtaccatccaccttaactt	aacgaccttaattactcctg	acggcttaattggcaggagc	cacggaaagccgtgagacca
81	aggaaggcagtggctgcccg	ttgctttcagacacccagct	acaaaagcaaacacacccct	gtgcctcacccgcgggcaat
161	cacaggttcaaagaccaagt	aggactttgcagaaggggct	gcttagaggaaaaattaagt	caattggagctcgtgggctc
241	agtgattcatgctgccggtg	cttggggtactccgacaaag	gtcagcattctctgttgtcc	tagcacaagtcgaggtctgg
321	ggagctcagtgttatggggt	gctctcagttcagggaggag	atttctggccctcgaggttg	ctcgtacagcctgcctgaaa
401	tgacagaaacctgaaagtca	aataaaacgacctttattgg	aaacaaatgaccgcgcaaat	gggactgagcagctgaaatc
481	accatcctggtatcagaagt	cccactcggacctttgattt	cccaacgacctctttatcat	tcggcgtgacgttaaactga
561	tgtcctaaagcccccaccag	gtaaatggaacagcagcagg	caggccgcgtttggtgtcac	tgtttctgcgtgcctctgtt
641	cattgccgtgctaacgtgag	gggagtttccttcccaaggt	cattaccacccagcagccat	ggcgccttacggtgtggaaa
721	gcactgctaaatggggcttc	gtatgcttttccagtgtaaa	ttaccacacgcagagtaggc	agtaagggtgaggagcagcc
801	agcacagtgcagctctcttc	ttagtgttacagtgagcgac	togagcacttoccacccaag	trtgagagrtgrargetget
881	ttagggaacaaacgcaacg	ctccctcgtcctcaggccta	raggrarggtgtrogtragg	
961	atcogtggctaggcggggag	cacacacacagetgeggeae	ggrotggcoagaactgctot	gragargtetetegateact
1041	ccgcacctctcactccgctt	tetgegeacegergeteteg		
1121	CECCECCEEEEEEEEEE	agtrarggrggtrrggrag	BCagcaggegecacagegeeet	
1201	cctccccgtgaggggggggg	catacgrggtgargrtgrgg		
		Cutucgegg tguegetgegg	50	BECECERCIECCOCHONOC
1281	BCGTCGGCGAGGCTGAGCGG	AGCGGGTTCCTGCGTCAACA		
	100	ACCOUNTION	OT OUT TO OACOGAACCOULC	150
1361	CCCCGACCGCCACCACCA	GCCGTCGTATCCACCGCATC	TOTOTOTOTOTOCOCCA	CCCCATCOCTACCCCTCCT
1001	COUCCACCOCACCOCACCA	deconcorrate accounte	ICICICICICICICCUCCA	GCGCCATGGCTACGCCTCCT
	1	200		MAIPP
1441	TCCCAGGTGCGCCAGAACTA			
	S O V B O N V	W O D C F A		
		нцрсва	AINKŲIN	LELYASY
1521	CGTGTACCTCAGCATGTTCA			I
		ECEECEE a EEEEECCCCECE	RERCARCERCERCELLAACE	REFECTEREFCCCERCERCE
1601				
1681	rggtgrtrrrggrgrage		Cgageeeegggeaaggeeeg	
1761				
1841		claccicgggcgagccitg	geegegegelaacigegegi	cgilggcigilaicagcgca
		****	<u> </u>	
1921	gcicagagcigcigggcggc	tctgcccttggggcagggct	ggggagcggttgcccaacgg	agcgcttggcagctctccca
1921	ctgcgttctctggaggcag cttgccttcttggattggc	tctgcccttggggcagggct gcttctcggggcctggtttt	ggggagcggttgcccaacgg gtgatgaattaggaggaatc	agcgcttggcagctctccca cgtggcatccccggttctgt
1921 2001 2081	ctgcgttctctggggggg gctgccttcttggctggc	tctgcccttggggcagggct gcttctcggggcctggtttt acgcaactcttgtttcagct	ggggagcggttgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg
1921 2001 2081 2161	ctgcgttctctggagcag gctgccttcttggctggc acctcgggcaacttcct	tctgcccttggggcagggct gcttctcggggcctggttt acgcaactcttgtttcagct agctgtgctgt	ggggagcggttgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttctggtgtagaattgaact	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc
1921 2001 2081 2161 2241	ctggcgttctctggagcag gcttgccttcttggctgc acctctgcaggcaacttcct aagatgtgaacctccactg	totgcccttggggcagggct gcttctcggggcctggttt acgcaactctgttcagct agctgtgctgt	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtgccagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgggctc
1921 2001 2081 2161 2241 2321	ctgcgttctctggagcag gctgcctcttggctgc acctctgcaggcaacttcct aagatggaacctccactg caggctgctgctgctgcagagc	totgcccttggggcagggct gcttctcggggcctggttt acgcaactcttgttccagct agctgtgctgt	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttotggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaagctttctgttg
1921 2001 2081 2161 2241 2321 2401	ctggcgttctctggagcag gctgcctcttggagcag acctctgcagcaacttcct aagatggaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtctttc	totgcccttggggcagggct gcttctcggggcctggttt acgcaactctgtttcagct agctgtgctgt	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagctg	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgcctgtgagcctc agatacaaagctttctgttg tcacagatctggctttcaag
1921 2001 2081 2161 2241 2321 2401 2481	ctgcgttctctggagcag gctgcctcttggagcag acctctgcaggcaacttcct aagatgtgaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtctttc agctggttttgtgttgctg	totgcccttggggcagggct gcttctcggggcctggttt acgcaactcttgttcagct agctgtgctgt	ggggagcggttgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac ggatggcttctgacagctg cagttaagcttttcgctttt	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggtcttg gcttgctgctgtcttactgc accggtgcctgtgggcctc agatacaaagctttctgttg tcacagatctggcttcaag aaatgtaatactagagttt
1921 2001 2081 2161 2241 2321 2401 2481 2561	ctggcgttctctggagcag gctgccttcttggctggc acctctgcagcaacttcct aagatgtgaacctccactg cagctgctgctgtcagagc aattctggtctgtgtcttc agctggttttgtgtttgctg ctagaagtaggaagcatgc	totgcccttggggcagggct gcttctcggggcctggttt acgcaactctgtttcagct agctgtgctgt	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagctg cagttaagcttttcgctttt aaatctggtggagtgtgcat	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaaagctttctgttg tcacagatctggcttcaag aaatgtaatactagagttt gaagcattggtgcagaagt
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641	ctggcgttctctggagcag gctgcctcttggctgcc acctctgcaggcaacttcct aaggatggaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtctttc agctggtttgtgtttgctg ctcagaagtaggaagcatgc acctaacgttctggccac	totgcccttggggcagggct gcttctcggggcctggttt agcaactcttgttcagct agctgtgctgt	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtgcaggagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagct cagttaagcttttcgctttt aaatctggtggagtgtgcat gtgtatgatcgaactttg	agcgcttggcagctctccca cgtggcatccccggttctgt gctgctgctgctgtcttactgc accggtgtcctgtgagcctc gatacaaagctttctgttg tcacagatctggctttcaag aaatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgatg
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641 2721	ctggcgttctctggagcag gctgcctcttggctgcc acctctgcaggcaacttcct aaagatgtgaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtcttc agotggtttgtgtttgctg ctcagaagtaggaagcatgc acctaaacgttctggccac accttggtccacaaaagga	totgcccttggggcaggct gcttctcggggcctggttt acgcaactctgttccagct gcagctggacgcagctgtg tgagagctctgcggttctt tctgaagaaaaggagctgt ttgataggtgttgtgcttc ttgctagtccatacctgct tcttaatttactgttagcag gtggaaagcactccttgaa	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttotggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac ggatggcttctgacagctg cagttaagctttcgattt aaatctggtggagtgtgcat gtgtatgattcgacttg	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaaagctttctgttg tcacagatctggctttcaag anatgtaatactagagttt gaaagcattggtgcagaagt acttgtgaccactgtgatg
1921 2001 2081 2161 2241 2321 2401 2401 2481 2561 2641 2721 2801	ctggcgttctctggagcag gctgcctcttggctgcc acctctgcagcaacttcct aagatgtgaacctccactg caggctgctgctgctgcagagc aattctggtttgtgttgctg ctagaagtaggaagcatgc acctaacgttctggcac acctaggtctgtgcacaaagga gtggtttccatgtattcag traacttgattcatgtattcag	totgcccttggggcaggct gcttctcggggcctggttt acgcaactctgttcagct agctgtgtgacgcagctgtg tgagagctctgcgggttctt ttgaagagaaaaggagctg ttgataggtgtgtgtctt ttgataggtctatacctgct ttcttaatttactgttagcag gtggaaaagcactccttgaa aagtgaccacaaataccag	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttotggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagctg cagttaagcttttcgctttt aaatctggtggagtgtgcat gtgttatgattcgaactttg agcaagttactcgaggaggc ttagttgtgcttactgatca	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaaagctttctgttg tcacagatctggctttcaag aaatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgatg ggcatggattaaaaacta
1921 2001 2081 2161 2241 2321 2401 2481 2561 2561 2561 2721 2881	ctggcgttctctggagcag gctgccttcttggctggc acctctgcagcaacttcct aagatgtgaacctccactg cagctgctgctgtcagagc aattctggtctgtgtcttc agctggttttgtgttgctg ctcagaagtaggaagcatgc acctaacgttctggcac acctagagtccacaaagg gtggtttccatgtattcag tgaacttaattctagctg	totgcccttggggcaggct gcttctcggggcctggttt acgcaactctgttcagct agctgtgtgacgcagctgtg tgagagctctgcgggtctt ttgaaggtgtgtgtgttc ttgctagtctaacctgct tcttaattactgtagcag gtggaaagcactccttga aagtgaaccaaataccag gaaactgaacttcccgg	ggggagcggttgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttotggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagctg cagttaagctttcgcttt aaatctggtggagtgcat gtgtatgattcgaactttg agcaagttactcgaggaggc ttagttgtgtcttactgatca agggaggactttgccttgt	agcgcttggcagcttccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgcctgtgagcctc agatacaaagctttctgttg tcacagatctggcttcaag aaatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgatg gggcatggattaaaaactaa tacttgctccctgttgtac
1921 2001 2081 2161 2241 2401 2481 2561 2641 2721 2801 2881	ctggcgttctctggagcag gctgccttcttggctggc acctctgcaggcaacttcct aagatgtgaacctccactg caggctgctgctgtcagagc aattctggtctgg	totgcocttggggcagggct gottotcggggcotggttt acgcaactottgttcagct agotgtgotgtattgcagcg caagotggacgcagotgttg totgaagaaaaggagotgt ttgaataggtgttgtgctto ttgctagtotcatacotgot tottaattactgttagcag gtggaaaagcactocttgaa agtgaaccacaaataccag ggaaactgaactttccacgg ggctotagcottottgaagg	ggggagcggttgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac ggatggcttctgacagct tagttaggttgtgagtgtgcat gtgttagatcgagtgtgcat gtgttagatcgaggggggg ttagtgtgcttactgatca agggaggactttggcttgt gctgagatgtttgaagttac	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaagctttctgttg tcacagatctggctttcaag aaatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgatg ggaacagaacttagtggtg ggcatggttaaaaactaa taccttgctcccgttgtac agcatagacttgctttgcg
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641 2721 2801 2881	ctggcgttctctggagcag gctgcctcttggctgcg acctctgcaggcaacttcct aaggatggaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtcttc agctggtttgtgtttgctg ctcagaagtaggaagcatgc acctaacgttctggccac accttggtccacaaagga gtggtttccatgtatcag tgaacttaattctagctg	totgcocttgggcaggct gottotogggcotggttt acgcaactottgttotagct gagotgtgacgcagctgtg taggaggctotgoggttott totgaagaaaaggagotgt ttgataggtgttgtgtotto ttgctagtotcatacotgct tottaattactgttagcag gtggaaagcactocttgaa aagtgaaccacaataccag gsottagcottottgaagg	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagct tagttagcttttcgattt aaatctggtggagtgtgcat gtgttagattcgaactttg agcagttactcgaggagc ttagtgtgctactgatca aggsaggacttggccttgt ctgatagtttgattcgacttg	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaagctttctgtg tcacagatctggtttcaag anatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgaatg agaacagaac
1921 2001 2081 2161 2241 2321 2401 2481 2561 2561 2721 2801 2881 2961	ctggcgttctctggagcag gctgcctcttggctgcg acctctgcaggcaacttcct aagatgtgaacctccactg caggctgctgctgctgtcagagc aattctggtctgtgtttgctg ctcagaagtaggaagcatgc acctaaacgttctggccac acctggtccagaagtaggaagcatgc acctaaacgttctggccac acctggttcacataaagga gtgggttccatgtatcag tgaagctagtaccacagctg tgctggttacccacagttg tggatgctagagtgagactt	totgcocttggggcagggot gottotogggootggttt acgcaactottgttocagot gagotgtgatgacgcagotgtg tgagagototgoggttott totgaagaaaaggagotg ttgataggtgttgtgotto ttgotagtotcatacotgot tottaattactgttagcag gtggaaagcactocttgaa aagtgaaccacaataccagg gaaactgaacttocagg ggottagottottagag ggotttotgagaatgaact	ggggagcggtgccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt tictggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac ggatggcttctgacagctg cagttaagctttccgactg gtgtatgatcgagtgtgcat gtgttatgatcgaactttg agcaagttactcgaggagc tiagttgtgcttactgatca agggggatttggccttgt gctgagatgttcgaattac gctagtagtgcttcgcctgt	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaaagcttctgttg tacagatctggctgtcagaagt aaatgtaatactagagttt gaaagcattggtgcagaagt agaacagaac
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641 2721 2801 2881 2961	ctggcgttctctggagcag gctgcctcttggctgcc acctctgcaggcaacttcct aaagatgtgaacctccactg caggctgctgctgctgctagagc aattctggtctgtgtttgctg ctcagaagtaggaagcatgc acctaaacgttctggccac accttgagtccacaaaagga gtggtttccatgtattcag tgaaacttaattctagctgg tgctggttacccacaacttg tggatgctagagtggagactt	totgcccttggggcaggct gcttctcggggcctggttt acgcaactctgttcagct gagcgtgtgatgacgcgctgtg tgagagctctgcgggttctt ttgaagagaaaaggagctg ttgataggtgtgtgtctc ttgcagtaccacactgct tcttaatttactgttagcag gagaaccacaaataccag gaaactgaacttccacgg ggctctagcctcttgaagg gagatttctgagcatatgaat	ggggagcggtgccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt tictggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac ggatggcttctgacagctg cagttaagctttcgacttt aaatctggtggagtgtgcat gtgttatgatccgaggaggc ttagtggctactgagaggc ttagtggcttactgatca aggaggactttggccttg gctgaatgtttgacttga tcttacttcagTCCTACTA S Y Y	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc agatacaaagcttctgttg tcacagatctggtgtcagagtt gaagcatggtgccagaagt actgtgaccactgtgaagt agaacagaac
1921 2001 2081 2241 2321 2401 2481 2561 2561 2721 2881 2881 2961	ctggcgttctctggagcag gctggcgttctctggagcag gctggcttcttggctggc acctctgcaggcaacttcct aagatggaacctccactg agtggtggtgtgtgtctg aattctggtctgtgtcttc agctggttttgtgttgctg ctcagaagtaggaagcatgc acctaacgtttctggcag agtgggttccatgattcag tgaaacttaattctagctgg tgctggttacccacagcttg tggatgctagagtgagactt	totgcccttggggcaggct gcttctcgggcctggttt acgcaactcttgttcagct agctgtgctgt	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagctg cagttaagcttttcgctttt aaatctggtggagtgtgcat gtgtatgattaccgaacttg agcagttactcgagaggc ttagttggcttactgatca agggaggactttgccttgt gctgagatgtttgaagttac tcttactttcagTCCTACTA S Y Y 350	agggcttggcagctctccca cgtggcatcccggttctgt cagtccagacagcggctctg gctggctgtcttactgc accggtgtcctgtgagcctc agatacaagctttcgttg tcacagattggtgcagaagt actggaccactgtgaag agaacagaac
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641 2721 2801 2881 2961 3041	ctggcgttctctggagcag gctgcctcttggctgcc acctctgcagcaacttcct aagatgtgaacctccactg caggctgctgctgtcagac aattctggtctgg	totgcocttgggcaggcat gottotcgggcotggttt agcaactotgttotagct agctgtgctgtattgcagcg caagotggacgcagctgtg tgagggototgoggttott totgaagaaaaggagctgt tgtagtgtctcatacotgot tottaattactgtagcag gtggaaagcactocttgaa gggaaagcactottgaagg ggctotagcottottgaagg gggtttctgagcatatgaat	gggsagcggtgcccaacgg gtgatgaattaggaggaatc gtggtagagccgggtgcctt tictggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggctictgacagct gtgttaggtttcgattt aaatctggtggagtgtgcat gtgttagatccgaactttg agcagttactcgagagc tiagtggtcttaggagta gctgagatgtttgaagtac scagtagtttgacttggcttg tagtgtgcttaggagt cttacttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG	agcgcttggcagctctccca cgtggcatccccggttctgt gctgctgctgctgtcttactgc accggtgtcctgtgagcctc gatacaaagctttctgttg tcacagatcggcttcaag taatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgatg ggcatggattaaaactaa tacttgctccctgttgtac agcatagactttgctttg
1921 2001 2081 2161 2241 2321 2401 2561 2641 2641 2641 2641 2881 2961 3041	ctggcgttctctggagcag gctgcctcttggctgcg acctctgcaggcaacttcct aagatggaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtttgctg ctcagaagtaggaagcatgc acctaaacgttctggccac acctaaacgttctggccac acctagagtcagaatgagagt ggggtttccatgtatcag tgaagctagagtgagactt tggatgctagagtgagactt	totgcocttgggcaggcot gottotogggcotggttt acgcaactottgttotagot agotgtgotgtattgcagog caagotggacgcagotgttg totgaagaaaaggagotgt ttgaataggtgttgtgtotto ttgotagtotcatacotgot tottaattactgttagoag gtggaaagcactocttgaa aagtgaaccacaataccagg gaaaactgaacttoccacgg ggototagcottottgaagg gagtttotgagcatatgaat TACTTCCTGCACCAGTCCCA Y F L H Q S H	ggggagcggtgccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt tictggtgtagaattgaact gcgtggtggaaagtggcac gggatggcttctgacagctg cagttagctttcgacttg agcagttagtctacgagtgcat gtgttatgatcgagtgcat gtgttatgatcgagtgc tagttggtgtatcgacttg agcagttactcgaggggc ttagtgtgcttactgatca aggggggactttggcttat acttgatgtgttagattac cttactttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A	agcgcttggcagctctccca cgtggcatcccggttctct cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaagcttctgttg tacagatctggttctagagtt gaagcattggtgcagaagt acttgtgaccactgtgattg gggcatggattaaaactaa tacctgctcctgttgtac agcatagactttgctttg
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641 2721 2881 2961 3041	ctggcgttctctggagcag gctgcctcttggctgcg acctctgcaggcaacttcct aagatgtgaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtttgctg ctcagaagtaggaagcatgc acctaaacgttctggccac acctaacgttctggccac acctaggtccacaaagga gtgggttccatgtatcag tgaagctagatgaggagctt tggatgctagagtgagactt CTCTGAAAAACTTTGCCAAG A L K N F A K 	totgcccttgggcaggct gcttctcgggcctggttt acgcaactctgttcagct agctgtgctgt	ggggagcggtgccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt tictggtgtagaattgaact gcctggtggataagtggcac gggtggcttctgacagctg cagttaagctttcgacttg agcaagttactcgagggggc tagttagatcgagggg tagtatgatcgaacttg agcaagtactcgaggaggc tiagttgtgcttactgatca agggggatttggccttgt gctgagatgtttgaagttac tcttacttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaaagcttctgttg tcacagatctggtgtcagaagt aaatgtaatactagagttt gaagcattggtgcagaagt agaacagaac
1921 2001 2081 2161 2241 2321 2401 2561 2561 2561 2561 2881 2961 3041 3121	ctggcgttctctggagcag gctggcgttctctggagcag gctggcttcttggctggc acctctgcagcaacttcct aagatggaacctccactg agctggtgtgtgtctgt aattctggtctgtgtcttc agctggttttgtgttgctg ctcagaagtaggaagcatgc acctaacgttctggcag acctaacgttctggcag tgaggttaccacaaggag tggggttccatgtattcag tgaaacttaattctagctgg tgctggttacccacagcttg tggatgctagagtgagactt	totgecettgggeagget gettetegggeetggett agetaetettgtteaget agetgtgetgt	gggsagcggtgcccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt ttctggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac ggatggcttctgacagctg cagtaagcttttcgacttt aaatctggtggagtgtgcat gtgtatgattcgaacttg agcaagttactcgaggagc ttagtgtgcttctgctgt gctgagatgtttgaagttac tcttacttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A tgaatactagctctctgggt	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcct aaatgtaatactagagttt gaagcattggtgcagaagt actggtgccagagttggggcaggagt agaacagacttagtggtgt gggcatggattaaaaactaa tacttgctccctgttgtac agcatagacttgcttgcg .300 TTTTGACCGGGATGATGTGG F D R D D V AGAAGCTGATGAAGCTGCAA E K L M K L Q tgtcatgcctaatgcagaac
1921 2001 2081 2161 2241 2321 2401 2561 2641 2721 2801 2881 2961 3041 3121	ctggcgttctctggagcag gctggcgttctctggagcag gctggcttcttggctggc acctctgcaggcaacttcct aagatggaacctccactg cagctgctgtgtgtcgtg aattctggtctgtgtcttc agctggttttgtgttgctg ctcagaagtaggaagcatgc acctaacgttctggccac accttgagtccacaaagga tgggtttccatgattcag tgaacttaattctagctgg tgctggttacccacagcttg tggatgctagagtgagactt CTCTGAAAAACTTTGCCAAG A L K N F A K .400 AACCAGAGGGGTGGACGCAT N Q R G G R I	totgcocttgggcaggcot gottotcgggcotggttt agcaactotgttotagct agctgtgctgtattgcagcg caagotggacgcagctgtg tgagggototgcggttot totgaagaaaaggagctgt tgtgatggtotgtgtot ttgctagtotcatacotgct tottaattactgtagcag gtggaaaagcactocttgaa gggaaagcactottgaagg ggototagcottottgaagg gggtttctgagcatatgaat TACTTCCTGCACCAGTCCCA Y F L H Q S H CTTCTTGCAGGACATCAAG F L Q D I K	ggggagcggtgcccaacgg gtgatgaattaggaggaatc gtggcagagccgggtgcctt tictggtgtagaatgaac gcctggtggataagtggcac gggatggctictgacagct gtgataggtgtggatagc cagtaagctttcgacttt aaatcggtggagtgtgcat gtgtatgatcgaactttg agcagttactcaggggg tiagtgtgctactgatca agggaggactttggcttgt gctgagatgtttgaagttac tcttacttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A tgaatactagctctctggt	agcgcttggcagctctccca cgtggcatccccggttctgt gcttgctgctgtcttactgc accggtgtcctgtgagcctc gatacaaagcttctgtg tcacagatcggcttcaag aaatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgaatg agaacagaac
1921 2001 2081 2161 2241 2321 2401 2561 2641 2641 2641 2641 2641 2641 2641 3041 3121 3201	ctggcgttctctggagcag gctgcctcttggctgcg acctctgcaggcaacttcct aagagtggacctccactg caggctgctgctgctgcagagc acttctggttggtcttct agctggtttggtt	totgcoottgggcaggcot gottotogggcotggttt agcaactottgttoagot agctgtgotgtattgcagcg caagotggacgcagotgttg totgaagaaaaggagotgt ttgaataggtgttgtgtott ttgctagtotcatacotgot tottaattactgttagcag gtggaaaagcactoottgaa gaaactgaacttocaaataccag ggototagcottottgaagg gatttotgagcatatgaat TACTTCCTGCACCAGTCCCA Y F L H Q S H CTTCTTGCAGGACATCAAG F L Q D I K	gggsagcggtgccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt tictggtgtagaattgaact gtgataggtggtacacaaac gcctggtggataagtggcac gggatggctictgacagctg agcagttagcttitcgacttt aatctggtggagtgtgcat gtgtatgatcgagtgtgcat gtgtatgatcgagtgtgcat gtgtatgatcgagtgtgcat gtgtatgatcgagtgtgcat gtgtatgatcgagtgtgcat gtgtatgatcgagtgtgcat gtgagagtatcgagtgtgcat gctgagagtatcgagtg tcttacttcagTcCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A tgaatactagctctctggt	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagcagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaagctttctgttg tcacagatctggtttcaag anatgtaatactagagttt gaagcattggtgcagaagt agaacagaac
1921 2001 2081 2161 2241 2321 2401 2481 2561 2641 2721 2881 2961 3041 3121 3201 3281	ctggcgttctctggagcag gctgcctcttggctgcg acctctgcaggcaacttcct aaagatggaacctccactg caggctgctgctgctgcagagc aattctggtctgtgtttgctg ctcagaagtaggaagcatgc acctaaacgttctggccac acctaaacgttctggccac acctagagtcagaagtaggagcatg gtgggtttccatgtatcag tgaagctagagtgagactt tggatgctagagtgagactt CTCTGAAAAACTTTGCCAAG A L K N F A K AACCAGAGGGGGGGACGCAT N Q R G G R I tactcttggggcctgtctcc agtggttgagcaatgcgt	totgcccttgggcaggct gcttctcgggcctggttt acgcaactcttgttcagct agctgtgctgt	ggggagcggtgccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt tictggtgtagaattagac gcdggtggtacacaaac gcctggtggataagtggcac gggatggcttctgacagctg agcagttagctttcgacattg agcagtgatctgacacttg aggaggacttggctta tagttgtgttacgacttg ctagttgtgctactgagaggc ttagttgtgctactgatca aggaggacttggccttgt GCGAGGAGCGTGAACATGCTG E E R E H A tgaatactagctcttgggt	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagcagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaagcttctgttg tacagatctggttctagagttt gaagcattggtgcagaagt acttgtgaccactgtgatg gggcatggattaaaactaa tacctgctcctgttgtac agatagactttgctttg
1921 2001 2081 2161 2241 2321 2481 2561 2641 2721 2881 2961 3041 3121 3201 3281 3261 244	ctggcgttctctggagcag gctggcgttctctggagcag gctggcttcttggctggc acctctgcagcaacttcct aagagtggaacctccactg agctggtggtggtctgtgtcttc agctggtttggtgtttg ctcagaagtaggaagcatgc acctaacgttctggcac acctaacgttctggcac acctaacgttctggcac acctagagtcagaagtaggag tgggtttccatgattcag tgaaacttaattctagctgg tgctggttacccacagctg tggatgctagagtgagactt	totgcocttgggcaggcat gottotogggcotggttt acgcaactottgttoagot agotgtgotgtattgcagcg caagotggacgcagotgttg tgagagototgoggttott totgaagaaaaggagotgt tgtgataggtgtgtgtoto ttgotagtotcatacotgot tottaatttactgttagcag gagaacgaactocottgaa aagtgaaccacaataccagg gaaaactgaacttococgg ggototagcottottgaagg ggattotgggcatatgaat Y F L H Q S H CTTCTTGCAGGACATCAAG F L Q D I K aggotgtaggggaaggaatt gotgtgatotgagtggtgtgt	gggsagcggtgcccaacgg gtgatgaattaggaggaatc gtgcagagccgggtgcctt ttctggtgtagaattagact gtgataggtggtacacaaac gcctggtggatagtggcac ggatggcttctgacagctg cagtaagcttttcgacttt aaatctggtggagtgtgcat gtgtatgattcgacagtg agcagttactcgaggagc ttagtgtgctactgata aggaggactttggcttgt gctgagatgtttgaagttac tcttacttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A tgaatactagctctctgggt cccctgtgaggcctcctgca atcagtcggcaagtcta	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gcttgctgctgtcttactgc accggtgtcctgtgagcctc agatacaaagcttctgttg tcacagatctggtgtcagaagt aaatgtaatactagagttt gaagcattggtgcagaagt agaacagaac
1921 2001 2081 2161 2241 2321 2401 2561 2561 2641 2721 2801 2881 2961 3041 3121 3201 3281 3261 3261 3261	ctagagtictotgagcag gottgocttottggagcag gottgocttottggattgc acctotgoagcaacttoot aagatggaacotcoactg aggotgotgotgtogagc aattotggtotgtgtotto agotggttttgtgttgotg ctagaagtaggaagcatgo acctaaacgtttotggcag acctaacgtttotggcag acctagagtagaagcatgo acctagagtcacaaaggag tgaaacttaattotagotgg tgotggttaccacaagttg tgatgctagagtgagactt CTCTGAAAAACTTTGCAAG A L K N F A K .400 AACCAGAGGGGTGGACGCAT N Q R G G R I tactottggggcotgctoc agtggttgagcacatgogt gacgagaagcagcaaacta	totgecettgggeagget gettetegggeetggttt ageagetggaegeagetgtg taggaggetggaegeagetgtg taggaggetetggaegeagetgtt tetgaagaaaaggagetgt ttgaaggaegeagetgttgtgtette ttgaaggaaggaegetgttgtgtgtette ttgeaggaaggeetettetgageag gtggaaaageaeteetgaag gggaeageaetteetgaagg gggetetageettetgaagg gggetetggeetettetgaagg gggetetggeetettetgaagg gggetetggeetettettettgaagg ff L H Q S H CTTCTTGCAGGACATCAAG F L Q D I K aggetgtagggaaggaatt getggaetggggaggaatt getggaetggtgtgt tteetettaagaeagetg	ggggagcggtgccaacgg gtgatgaattaggaggaatc gtggcaggccgggtgcctt ttctggtgtagaatgaac gcctggtggataagtggcac ggatggcttctgacagcg gagtggcttctgacagct gtgatagctttcgacttt aatctggtggagtgtgcat gtgtatgattcgaacttg agcagttactcgagggg ttagtgtgctactgatca agggaggactttggcttgt gctgagatgtttgaagttac tcttactttcagTCCTACTA S Y Y 350 CGAGGAGCGTGAACATGCTG E E R E H A tgaatactagctctctgggt cccctgtgaggcctctgcg atagtcggcaagtcta	agcgcttggcagctctccca cgtggcatccccggttctgt cagtccagacagcggctctg gctggctgtcttactgc accggtgtcctgtgagcctc agatacaaagcttctgtg tcacagattggtgcagaagt aatgtaatactagagttt gaagcattggtgcagaagt acttgtgaccactgtgaatg agaacagaac

Although highly conserved in ferritin H-subunit sequences, this sequence does not occur within the 5' untranslated regions of published human and rat ferritin L-subunit cDNA sequences (6, 17, 24). We therefore propose that this conserved 22-nucleotide sequence is the region of ferritin Hsubunit RNA which mediates iron induction of H-subunit protein synthesis (see Discussion).

Analysis of amino acid sequence of the chicken ferritin H subunit. The chicken ferritin H-subunit gene encodes a protein of 179 amino acids which, without posttranslational modification, has a calculated molecular weight of 20,941 and a pI of 6.1. Although the pI of 6.1 is substantially higher than the pI range of 4.2 to 5.6 determined for ferritins of other species (38), it is consistent with the pI of 6.14 determined experimentally for chicken spleen ferritin (35). The predicted chicken ferritin H-subunit sequence has 93% identity with the amino acid sequence of the human ferritin H subunit (6, 11, 12). Seventy-three consecutive amino acids (Glu-16 through Asp-88) are identical in the amino-terminal halves of the chicken and human H-subunit sequences. There are amino acid differences at 10 internal positions; Cys-15, Arg-89, Asn-94, Thr-97, Glu-122, Asp-138, Gln-146, Lys-161, Tyr-162, and Met-164 in the chicken sequence are replaced by Ser, Cys, Ser, Asn, Asp, Asn, Glu, Glu, Ser, and Leu at the corresponding positions of the human protein. H-subunit sequences vary most at amino-terminal (Ala-Thr-Pro-Pro in chicken, Thr-Thr-Ala-Ser-Thr in human) and carboxyl-terminal (Glu-Ser-Asp-Ser in chicken, Asp-Ser-Asp-Asn-Glu-Ser in human) ends.

When the chicken and human H-subunit sequences are aligned with the horse spleen L-subunit sequence and helical boundaries determined by X-ray crystallography of the L-subunit (31) are imposed upon H-subunit sequences, it is clear that the amino acid sequences of the four long helices

3601	gcttgaagtccctttactca	tttcctgctgttttgaatta	ctaatagtgtgatcagtttt	aatgtatagccattcttcat
3681	ggggagaacagtgacagcca	ctttagaaaaaatatttttc	aagttgcatatttggatatc	ttaaactgctttgttgtctt
3761	tgagattgtacctatgcagt	ggcttaacatggcaggcagc	tcagcaccacacagctgttc	actccatcccctccccacag
3841	gcatgggaggagtgttagca	aaaggactttaggggctgag	acaaagacagttcacaggac	agaaatgaaagggaatatta
3921	gggatatacaaaacaagtga	tgtgcaattgcttactgctt	gctggctggtgctcagccaa	tctcagggcagcagcccacc
4001	cctagcaaactcctccagtt	ctattattcagcgcgatacc	atatggcatggactatcttg	cagtttcagtcagctgcact
4081	gcttctgtgcagccccagct	ccttgctggcacaacaggat	gaggagctgaaaagtggttg	ccatctagtgcaagcactgc
4161	tcgacaactcaaacatcagt	gttatcacgctgtttttttt	ctcagtggcatcataccaac	cactatgaacaaagtttagc
4241	acgtttgagccaaaaacagg	acactactgagatgtttagg	ccatgagattacaaggaagg	aagaaagtagtcctaagttt
4321	caaccttccaaggtttcact	ggacgtaatgtgagaagaac	ttacagaacagatacgtgtc	tgttctctgtgacttttagt
4401	ctgcattcaattctgtccta	agtgcaacgtgagtgctcag	gtgttgttttatttttttt	ccaagacccaaatgcttgag
4481	cttggttgtgttgaaggagt	tttgagtarragagrottar	tegagaarteretaeterae	caagttgcttttttagatgg
4561	actttgagarrgarrtaaa	gantitgetetgttgattg	tttgetttstssasssate	tgggarttggreatgggttt
4001	accergagacegaceraaa	gattergettergatte	450	- CEEE COULE E E COULE
4641	tooottttadAACCGGATC		CTGACTGCAATGGAGTGTGC	CCTGCACCTGGAGAAGAATG
4041				I N I E K N
			LIANECA	
4721	TGAACCAGTCGCTGTTAGAG	CTGCACAAATTGGCAACTGA	AAAGAATGACCCACAC	at age t act as a got seaso
4/61				g igga igi igaggi iaagag
4001				
4001		geligialgiggelalggal		
4001		gcicilccaggaggaalgic		agaggaccaattgtattagg
4901	gigacaaaggicigicatag	agcatggggtgtgtagctgt	gettaaaccetttgtteet	guttututagaatagaaggta
5041	ctcatgaacagcactttgca	gagaagtatgtttatcagta	tgcaataggcttgattctct	ggggactictgctataggat
5121	tccaaggtcactttgggga	tgataaaattgctgacact	tcagaccgctgcaatccaga	cacatgetgetaacetette
5201	ctgatagctaaaaaagcctc	agtccagccttggcaatgat	gtctgtagtaaagccttcag	gtgtgaattaacttgtgttt
5281	tcagactactcttgaaaata	ccctgtgtgagctagaagga	tttattttaaaagcagcaac	atagetgacegttetacatg
5361	tataaatgctcgtgatcgtg	aagtgctgtgctaagtgctt	cttaaataccttcagagatg	tactigattcatctttcctt
5441	agtaggacttttcagcagga	tgagcgtgtttgggcctaaa	cttctttattctcccttcca	gTTGTGTGTGACTTCATTGAGA
5441	agtaggacttttcagcagga	tgagcgtgtttgggcctaaa	<u>cttctttattctcccttcca</u>	ETTGTGTGTGACTTCATTGAGA L C D F I E
5441			.600 .	EITGTGTGTGACTTCATTGAGA
5441 5521	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG	GTGAAAGCCATCAAGCAGCT	.600 . GGGTGACCAACC	ETGCGGAAGATGGGGGGCACCC
5441 5521	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L	CTTCTTTATTCTCCCTTCCA .800 . GGGTGACCATGTGACCAACC G D H V T N	ETTGTGTGACTTCATTGAGA L C D F I E
5441 5521	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L	CTTCTTTATTCTCCCTTCCA .600 GGGTGACCATGTGACCAACC G D H V T N	ETTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
5441 5521 5601	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA	GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC	CTTCTTTATTCTCCCTTCCA 600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA	ETTGTGGGAGGATGGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGCACCC L R K M G A P .700 AggCTTGCCAAGAACCACCC
5441 5521 5601	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGAGTA K Y G M A E Y	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T	CTTCTTTATTCTCCCTTCCA GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S *	ETTGTGTGACTTCATTGAGA L C D F I E TGCGGAAGATGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC
5441 5521 5601	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T 	CTTCTTTATTCTCCCTTCCA 600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S *	ETTGTGGGAGATGGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGCACCC L R K N G A P .700 AGGCTTGCCAAGAACCACCC
5441 5521 5601 5681	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGGTA K Y G M A E Y TGTGGGGTTTCAGGGGGACTT	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T 	CTTCTTTCAGCAACGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC	ETTGTGTGACTTCATTGAGA L C D F I E TGCGGAAGATGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGGACTT .800	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T .750 CTACTTCACTGGTCCAGCAA	CTTCTTTCAGCAAAGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC .850	ETTGTGGGAGGAGGATGGGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGGTTTCAGGGGGGACTT .800 TCTGTACCAAATATCAGCCC	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T 	CTTCTTTATCCCCTCCA 600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC 850 GTACTGCCCTCCATCCATA	ETTGTGGGAAGATGGGGGGGCACCC L R K M G A P .700 AggCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGGACTT .800 TCTGTACCAAATATCAGCCC gcttctttggcttttttctt	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T .750 CTACTTCACTGGTCCAGCAA <u>TCTTTTCTCTTGTGTTTTGT</u> ccccttgcgaagtgtttgtg	CTTCCTTCACCAACC G D H V T N TCGGGGAAAAGTGACAACCG G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC 850 GTACTGCCCTCCATCCAATAC Caggactgtgctctggtctt	EITGTGGGAAGATGGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGGCACCC L R K N G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGGAGTA K Y G N A E Y TGTGGGTTTCAGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttctttggctttttcctt gctgagcgtggatgaagaaa	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T 	CTTCTTCACCATCCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC .850 GTACTGCCCTCCATCCATATC caggactgtgctctggtctt aacttgagtattttagtgga	EITGTGGGAAGATGGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921 6001	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttcttggctgatgaagaaa tcatgcaaggatcgcagagg	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACACCC L F D K H T 	CTTCTTCACCATCTCACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC <u>850</u> GTACTGCCTCCATCCAATATC caggactgtgctctggtctt aacttgagtatttagtgga ttcctttgccaccaaagctt	ETTGTGGGAGATGGACTTCATTGAGA L C D F I E TGCGGGAAGATGGGGGGGCACCC L R K N G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC AGGTGACTTGGTTCAgatg aactgagaaagtgatcttga taccactggatggggagaaa ctgtcttaatgtacaagaag
5441 5521 5601 5681 5761 5841 5921 6001 6081	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGGACTT .800 TCTGTACCAAATATCAGCCC gcttcttggctttttctt gctgacgtggatgaagaaa tcatgcaaggatcgcagagg aggaaggcttgaggctctat	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T 	Cttctttattctcccttcca .600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC .850 <u>GTACTGCCCTCCATCC</u> AATA Caggactgtgctctggtctt aacttgggtatttagtgga ttcctttgccaccaaagctt tgaatttaggtgagatggtg	ETTGTGGTGACTTCATTGAGA L C D F I E TGCGGAAGATGGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC TAGATGACTTGGTTCA aactgagaaagtgatcttga taccactggagagagaaa taccactggagagagaaa taccactggagagagaaa taccactggagagagaaa taccactggagaagtgatcttga
5441 5521 5601 5681 5761 5841 5841 5921 6001 6081 6161	Agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGGACTT .800 TCTGTACCAAATATCAGCCC gcttctttgctttttctt gctgagcgtggatgaagaa tcatgcaaggatcgcagagg aggaaggcttgaggctctat ctaccctggcatacaaagtc	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCACACCC L F D K H T 	Cttctttattctcccttcca .600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC .850 GTACTGCCCTCCATCCATATA caggactgtgctctggtctt aacttgagtatttagtgga ttccttgccaccaagctt tgaatttaggtggatggtg cctccttcacaacgtaagct	EITGTGGGAAGATGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTAC
5441 5521 5601 5681 5761 5841 5921 6001 6081 6161 6241	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q . 650 AAGTATGGCATGGCAGAGTA K Y G N A E Y TGTGGGTTTCAGGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttctttggctttttcctt gctgagcgtggatgaagaaa tcatgcaaggatcgcagagg aggaaggcttgaggctctat ctaccctggcatacaaagtcacagtca	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACACCC L F D K H T 	CTTCTTTCAGCTGACCAAGCTGA G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC GTACTGCCCTCCATCCAATAC Caggactgtgctctggtctt aacttgagtattttagtgga ttcctttgccaccaaggtagtg cctccttcacaagtaggtg tgtagtgccatttgcccagc	EITGTGGGAAGATGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGCACCC L R K N G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921 6081 6081 6161 6241 6321	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGACTT .800 TCTGTACCAAATATCAGCCC gcttctttgctttttctt gctgagcgtggatgaagaa tcatgcaaggatcgcaggg aggaaggcttgaggctctat ctaccctgcatacaaagtc actgtctcaaaggtaggaatattg	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACACCC L F D K H T 	CTTCTTCACCATCTCCCTCCCATCC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC <u></u>	EITGTGGGACTTCATTGAGA L C D F I E TGCGGGAAGATGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921 6001 6081 6161 6241 6321 6401	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACCC L F D K H T 	CTTCTTTCACCATCCCATCC G D H V T N TCGGGGAAAGTGACAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC <u></u>	EITGTGGGAAGATGGGGGGGCACCC L R K M G A P .700 . AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC AGGCTGGACTGGGTCAgagatg aactgagaaagtgatcttga taccactggatggggagaaa ctgtcttaatgtacaagaag cctgtctaacaaaaatggca gagaaaagtggtgcaggt aagaacagtggctgggcaggt ggaaactaaaaacgaggaag
5441 5521 5601 5681 5761 5841 5921 6001 6081 6161 6321 6321 6401 6481	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCA V K A I K Q L CCTGTTTGACAAGCAAGCACCCC L F D K H T 	CTTCTTTCACCATCTCACCAACC G D H V T N TCGGGGAAAGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC 	EITGTGGGAAGATGGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921 6081 6161 6321 6481 6481 6481	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGGACTT .800 TCTGTACCAATATCAGCCC gcttctttggctttttctt gctgagggtggatgaagaa tcatgcaaggatgggctcat tcatccctggcatacaagtc actgtctaaaaggtaggatacgc ggcaaagggttggaatatg aacacacatttctgtgaa	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCA V K A I K Q L CCTGTTTGACAAGCAGCACCC L F D K H T 	CTTCTTTCAGCAACGCAACGCAACC G D H V T N TCGGGGAAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC .850 GTACTGCCCTCCATCCAATAC caggactgtgctctggtctt aacttgagtattttagtgga ttccttgccaccaaggtagtg cctccttcacaacgtaagct tgtagtgccatttgcccagc ggtgggtggactgaacttat ttcattaaactttattgtg ttttctttttttagggaag tttcattattagtgaattatttattgtga	EITGTGGGAAGATGGGGGGCACCC L C D F I E TGCGGAAGATGGGGGGCACCC L R K N G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTGGTCAAGTAGT TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921 6081 6081 6161 61241 6321 6481 6481 65641	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGGAGTA K Y G N A E Y TGTGGGTTTCAGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttctttggctttttctt gctgagcgtggatgaagaa tcatgcaaggatggatgaagaa tcatgcaaggatggatgaagaa tcatgccagggttgaatattg aacacacaactttcgtgaa caagtgggcctgacatcgc ttgttgcctgttaccatc tcaataccattgatggcagc	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAGCACCC L F D K H T 750 CTACTTCACTGGTCCAGCAA TCTTTTCTCTTGTGTTTTGT ccccttgcgaagtgtttgtg taatatggaggaattgttt aattgaagcacattcaaaa gcttcctaaactgctgcctg ataatacagagctcttgcat atagggtgttaggtagggt tgcaaggtgttggaaggga aagaggaagaactcatggt ctagttttaggtggtgg caattgtttgagtaggtag cacttgtttgagtagctt	CTTCTTTCACCATCTCACCAACC G D H V T N TCGGGGGAAAGTGACAACGCTGA L G E S D S * TGCATGCATCTTCAGCTATC GTACTGCCCTCCATCCAATAC Caggactgtgctctggtctt aacttgagtattttagtgga ttcctttgccaccaagct tgagtgccattgccagc ggtggtggactgaacttat ttcattaactttattgtg tttcaaatcagcagtatgg gtttcattttttttagggaag	EITGTGGGACTTCATTGAGA L C D F I E TGCGGAAGATGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC AGGTGACTTGGTTCA agatgaaagtgatcttga taccactggatggggagaaa ctgtcttaatgtacaagaag cctgtcttaatgtacaagaag cctgtcttaatgtacaagaag ctgtcttgtgattgaagga agataaacagtggtctgggcaggt agaaacagctgtgctgggcaggt tagaaactgagaacttgaagcaatg ggagaacttgaagcaag
5441 5521 5601 5681 5761 5841 5921 6001 6081 6081 6241 6321 6401 6481 6461 6541 6721	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGACTT .800 TCTGTACCAAATATCAGCCC gcttcttggctgttttctt gctgagcgtggatgaagaaa tcatgcaaggatcgcagagg aggaaggcttgaggctctat ctaccctggcatacaagta cactgtccaaaagtacgcagg ggcaagggttggaatattg aacacacaactttctgtgaa caagtgggcctgacatctgc ttgattgctgttaccatc tcataccatgatgccagc ggcattcttggtgcagc	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACCC L F D K H T 	Cttctttattctcccttcca .600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC 	EITGTGGGAAGATGGGGGGGCACCC L R K M G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC TAGATAAACAGTTCTTTTAC AGGCTGCCTGGGTCA aactgagaaagtgatcttga taccactggatggggagaa ctgtcttaatgtacaagaag cctgtctaacaaaatggca gataaacagtggtgcaggt aagtaaacagtggtgcaggt ggaaactaaaacgaggaag acttgttaggatcagctggca gtgctcttgtgattgaaggg tggaactaaaacgaggaag acttgttaggaccttagct atagagaacttgaccagcagg
5441 5521 5601 5681 5761 5841 5921 6001 6081 6081 6321 6401 6481 6481 6561 6641 6561 6801	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACCC L F D K H T 	CTTCTTTCACCATCCCATCC G D H V T N TCGGGGAAAGTGACCAACC G D H V T N TCGGGGAAAGTGACGACGAA L G E S D S * TGCATGCATCTTCAGCTATC 	EITGTGGGAAGATGGGGGGCACCC L C D F I E
5441 5521 5601 5681 5761 5841 5921 6081 6081 6481 6481 6881	agtaggacttittcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q .650 AAGTATGGCATGGCAGAGTA K Y G M A E Y TGTGGGTTTCAGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttctttggctttttctt gctgagggtggatgaagaa tcatgcaggatgggctgagg aggaaggcttgaggctcat ctatccctggcatacaagtc aacaacattctgtga acaacacattctgtgaa caagtgggcctgacactgc ttgtgttgcctgttaccatc ctaataccattgatggcagatg gctnggccttaggcaatattg acaacacattgatggcagtgg actacctggaggctgag gcttcttggttgcctgtaccatcgc ttgtgttgcctgtaccatcgc ttgtgttgcctgtagcagtgg acticctgaagcgctggg ctnggccttatggcaatattg ctnggccttatggcaatattgcagc	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAGCACCC L F D K H T 	Cttctttattctcccttcca .600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCCTTCCATCCTATC .850 GTACTGCCCTCCATCCATATA caggactgtgctctggtctt aacttgagtatttagtgga ttccttgccaccaaggtagtg cctccttcacaacgtaagct tgagtgccattgcccagc ggtgggtggactgaacttat ttcattaaactttattgtg tttcattatttttttagggaag tttcattattttttttagggaag tttcattatttttttttagggaag tttcattatttttttttttttttttttttttttttttt	EITGTGGGAAGATGGGGGGCACCC L R K N G A P .700 AGGCTTGCCAAGAACCACCC TAGATAAACAGTTCTTTTAC
5441 5521 5601 5681 5761 5841 5921 6081 6321 6481 6481 6481 6481 6561 6881 6861 68961	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGGAGTA K Y G N A E Y TGTGGGTTTCAGGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttctttggctttttctt gctgagcgtggatgaagaaa tcatgcaaggatggatgaagaaa tcatgccagggtggatgaagaaa tcatgccagggttggatacatcg ggcaagggttggatattg aacacacaacttctgtgaa caagtggcctgacatcgc ttgttgctctttagcgcagtgg ctcatccatgaggcctggg ctcatctgggcttacatctgc tcaataccattgatggcagtgg ctcatcttgggctatatct tcactctggctatacatctgc ttgttgctttggctatatct tcactctggcatacacttctggcaata tctgcttgggctaagcgctgtgg cctnggccttatggccaata ttctgcttgggctaacttct	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAGCACCC L F D K H T 	Cttctttattctcccttcca .600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC .850 GTACTGCCCTCCATCCATATA caggactgtgctctggtctt acttgagtattttagtgga ttcctttgccaccaagct tgagtgccattgcccagc ggtggtggactgaacttat ttcattaactttattgtg tttcatttatttttttaggaag ttcctttgccaccatgctgt ttcattagcagtatttagtga ttcctttgccaccatgtagct ttcattaactttattgg ccccatgcctcactgctgtctt ttcttggtaatttgttcc cagtagtctccttgttcc cattccgctaattagtt ttcttggtaatttgttcc cattccgctaattagtt ttccttgttattgtttcc cattccgctgtaattaggt	EITGTGGGAAGATGGGGGGCACCC L C D F I E
5441 5521 5601 5681 5761 5841 5921 6001 6161 6161 6241 6321 6401 6401 6561 6561 6561 6561 6561 7041	agtaggacttttcagcagga CTCACTACCTGGATGAGCAG T H Y L D E Q 650 AAGTATGGCATGGCAGAGTA K Y G N A E Y TGTGGGTTTCAGGGGGACTT .800 <u>TCTGTACCAAATATCAGCCC</u> gcttctttggctttttctt gctgagcgtggatgaagaa tcatgcaaggatcgcagagg aggaaggcttgaggctctat ctaccctggcatacaagtc accgtccaaaggtaggatacgc ggcaagggttggatattg aacacacactttctgtgaa caagtgggcctgacatctgc tcgatcctagagctgtacacagtc tcatgctcaaaggtagcagc gattcctgaacgtggatacagc gcttcttggcttatctt tcacctggcatacaagtc accgctacaacttctgtgaa caagtggccgcacattgc gattcctgaaccgctgtg ctnggcttagcagcatattg tcactctgaaccttagtgcaatat ttctgcttggcatatct tacctaatgctagaagctct agccaaatgcctgaagcctta	tgagcgtgtttgggcctaaa GTGAAAGCCATCAAGCAGCT V K A I K Q L CCTGTTTGACAAGCAAGCACCC L F D K H T 	Cttctttattctcccttcca .600 GGGTGACCATGTGACCAACC G D H V T N TCGGGGAAAGTGACAGCTGA L G E S D S * TGCATGCATCTTCAGCTATC 	EITGTGGGACTTCATTGAGA L C D F I E

FIG. 3. Sequence of the chicken ferritin H-subunit locus. Sequence is of the 7.1-kbp ferritin H-subunit locus mapped in Fig. 1B. The four exons are individually outlined in boxes within the locus. Lowercase letters are used for nontranscribed regions; the transcribed sequence is shown in uppercase letters. The genomic sequence is numbered to the left of each line; arbitrarily, nucleotide 1 is assigned as the first base of the KpnI site. Above the DNA sequence, dots mark every 10th base of the mRNA sequence, with a number at every 50th nucleotide. Listed below the transcribed sequence is the amino acid sequence predicted from the single open reading frame. Amino acids are indicated by the standard one-letter code; * marks the translational stop codon. The TATA box 5' to the transcriptional start site and the AATAAA polyadenylation signal in the fourth exon are highlighted by shading.

which form the core of ferritin subunit structure are strictly conserved between the two H-subunit sequences. Amino acids of helix B are identical at every position. Helices A, C, and E each have one or two amino acid substitutions, all within the first or last helical turns, whereas helix D sustains conservative substitutions for two internal helical residues: Asn for Asp-138 and Glu for Gln-146. It is noteworthy that the chicken ferritin H subunit maintains the sequence Asp-Pro-His-Leu-Cys-Asp-Phe at residues 125 through 131 as well as the tetrapeptide Glu-Tyr-Leu-Phe at positions 166 through 169. No amino acid variability has been observed in either of these regions in any published ferritin H- and L-subunit sequences (14).

Expression of ferritin H-subunit RNA. RNA blot analysis (Fig. 4) shows that in total cellular RNA, the chicken ferritin H-subunit transcript is expressed at essentially the same level in all erythroid cells examined: normal cells of the primitive (4.5-day embryonic RBC) and definitive (11-day embryonic and anemic hen RBC) erythroid lineages as well as transformed erythroid precursor cells (HD6). Liver RNA samples from three developmental stages (15-day embryo, 1-month-posthatching chicken, and adult hen) have approx-



FIG. 4. Expression of ferritin H-subunit mRNA. Total cellular RNA (10 μ g/lane) was electrophoresed on a 1.4% agarose-2.2 M formaldehyde gel, transferred to nitrocellulose (20, 23), and probed with an antisense RNA transcript of ferritin H-subunit cDNA clone 5C (nucleotides 202 to 869 of the transcribed region shown in Fig. 3). Lanes A through D contain erythroid RNA samples: A, 4.5-day embryonic RBC; B, 11-day embryonic RBC; C, anemic hen RBC; and D, HD6 RNA. Lanes E through H contain RNA samples from nonerythroid, nonliver cells and tissues: E, MSB-1 cells; F, 4.5-day total chicken embryo (from which embryonic RBC were perfused); G, 11-day CEF; and H, 11-day embryonic brain RNA. Lanes I through K contain liver RNA samples: I, 15-day embryonic liver; J, 1-month-posthatching chicken liver; and K, adult hen liver RNA.

imately equal levels of ferritin H-subunit RNA, although the amount of the transcript per microgram of total RNA is lower (approximately threefold) for liver RNA than for erythroid RNA. Transcripts from the ferritin H-subunit gene are also evident in each of the nonerythroid, nonliver RNA samples (MSB-1 cells, 4.5-day total embryo, 11-day CEF, 11-day embryonic brain), although at much lower levels.

Analysis of ferritin H-subunit sequences in chicken genome. When chicken genomic DNA was cleaved with various restriction enzymes and probed with an H-subunit cDNA clone at low stringency (Fig. 5), the hybridization pattern was precisely that predicted from the map of λ EMBL3-56D (Fig. 1B). A single *Bam*HI fragment encompasses the entire ferritin H-subunit gene, and the 16-kbp *XbaI* fragment includes all but 95 nucleotides of the fourth exon. The 3.9-kbp



FIG. 5. Southern blot analysis of chicken genomic DNA. Chicken RBC DNA (5 μ g per lane) was digested with *Bam*HI (B), *Eco*RI (R), *Hin*dIII (H), or *Xba*I (X); fragments were separated on a 0.8% agarose gel and blotted to nitrocellulose (18). The probe was nick-translated cDNA clone 20A (nucleotides 137 to 869 of the transcribed region shown in Fig. 3). DNA size markers (λc I857S7 digested with *Eco*RI) were electrophoresed in parallel with the genomic DNA.

*Eco*RI fragment and the 3.6-kbp *Hin*dIII fragment contain 3' exons, whereas minor bands in these lanes at 5.8 and 5.6 kbp contain the first exon, of which only 125 nucleotides are included in the cDNA clone used as probe. The 0.24-kbp *Eco*RI fragment which contains the second exon is too small to be seen in this blot. In the chicken genome, therefore, ferritin H-subunit sequences exist in a single locus from which mRNA encoding the protein is transcribed. We find no evidence for multiple copies of the ferritin H-subunit gene in chicken, although chicken ferritin H-subunit cDNA probes detect multiple sequences in *Eco*RI-digested human genomic DNA (data not shown).

DISCUSSION

Our interest in the regulation of genes specifically induced during erythroid differentiation led us to isolate and characterize the chicken ferritin H-subunit gene described in this report. This gene is of special interest to us not only because its mRNA is abundant in erythroid cells but also because its expression is controlled at multiple levels.

First, RNA blot analysis of ferritin H-subunit transcripts in various chicken tissues and cell types indicates that the level of ferritin H-subunit mRNA is controlled in a tissuespecific manner. We therefore predict that the varied amounts of RNA result from differential transcription in the various tissues and cell types and that tissue-specific regulatory sequences modulate transcription from the single ferritin H-subunit gene in individual chicken tissues.

Second, other investigators have demonstrated that the ratio of ferritin H and L subunits in particular cell types appears to be determined by RNA levels for the two subunits and that the ratio of H mRNA to L mRNA may vary greatly during the course of cellular differentiation in any given tissue type (11). Although we find essentially equivalent amounts of ferritin H-subunit transcript in the total cellular RNA of erythroid cells from three different ages of organism development and from early and late stages of erythroid maturation, we have not assessed the ratio of H to L subunits in these samples. It will be of interest to determine whether transcription from the ferritin H- and L-subunit loci is regulated independently or coordinately during erythroid development and differentiation.

Third, the synthesis and accumulation of ferritin protein subunits is controlled posttranscriptionally by the level of intracellular iron (4, 36, 44). We presume that just as ferritin subunit amino acid sequences are maintained with few substitutions in species as divergent as chickens and humans, likewise the cellular mechanism for iron-modulated translational control of ferritin subunits must be highly conserved among various species. Thus it is conceivable that the 22-nucleotide identity located in the otherwise marginally homologous 5' untranslated regions of the chicken and human ferritin H-subunit genes may be functionally conserved to effect iron-modulated sequestration and translation of ferritin H-subunit messages.

Fourth, there may exist yet another level in the control of ferritin synthesis during chicken embryonic development, specifically during the switch from the primitive to the definitive erythroid series. Ferritin protein levels have been measured to be 10-fold higher in the primitive RBC of 3-day chicken embryos than in the circulating definitive RBC of 10-day chicken embryos or adult roosters. This decline in ferritin content of the circulating RBC correlates with the replacement of the primitive embryonic erythroid lineage by the definitive adult erythroid lineage (39). Interestingly, however, at the RNA level we do not observe the difference in ferritin content of the two different erythroid lineages. The amount of ferritin H-subunit mRNA is essentially equivalent in 4.5-day (primitive) and 11-day (definitive) chicken embryo RBC as well as in circulating definitive RBC from anemic hens (Fig. 4). To adequately dissect the various levels of control of ferritin gene expression, it will be necessary to determine whether the differing amounts of ferritin protein in cells of the primitive and definitive erythroid series are a consequence of variation in iron concentration within cells of the two lineages or whether there is an additional cellular regulatory mechanism which differentially modulates the ferritin content of primitive and definitive erythroid cells.

Other investigations of the regulation of ferritin subunit gene expression have been complicated by the fact that the rat, tadpole, and human genomes contain multiple loci which harbor sequences complementary to ferritin H- and Lsubunit cDNAs. The chicken genome, however, contains only a single copy of the gene encoding the ferritin H subunit. Since our studies provide the DNA sequence of the entire chicken ferritin H-subunit locus (including flanking sequences and introns) as well as a complete transcriptional map, it should now be possible to probe the multiple levels of regulation which govern expression from this ferritin Hsubunit gene.

ACKNOWLEDGMENTS

We gratefully acknowledge expert technical assistance from Susan M. Selbach and Susan J. Stamler, and we thank our laboratory colleagues for many fruitful discussions.

This work was supported by a National Research Service Award individual postdoctoral fellowship from the National Institutes of Health to P.W.S. and by Public Health Service grant HL24415 from the National Institutes of Health to J.D.E. J.D.E. also acknowledges the support of an American Cancer Society Faculty Research Award.

LITERATURE CITED

- 1. Akiyama, Y., and S. Kato. 1974. Two cell lines from lymphomas of Marek's disease. Biken J. 17:105–116.
- 2. Arosio, P., T. G. Adelman, and J. W. Drysdale. 1978. On ferritin heterogeneity: further evidence for heteropolymers. J. Biol. Chem. 253:4451-4458.
- 3. Aviv, H., and P. Leder. 1972. Purification of the biologically active globin messenger RNA by chromatography on oligo-thymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA 69: 1408–1412.
- 4. Aziz, N., and H. N. Munro. 1986. Both subunits of rat liver ferritin are regulated at a translational level by iron induction. Nucleic Acids Res. 14:915–927.
- Beug, H., S. Palmieri, C. Freudenstein, H. Zentgraf, and T. Graf. 1982. Hormone-dependent terminal differentiation *in vitro* of chicken erythroleukemia cells transformed by *ts* mutants of avian erythroblastosis virus. Cell 28:907–919.
- Boyd, D., C. Vecoli, D. M. Belcher, S. K. Jain, and J. W. Drysdale. 1985. Structural and functional relationships of human ferritin H and L chains deduced from cDNA clones. J. Biol. Chem. 260:11755-11761.
- Breathnach, R., and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. 50:349–383.
- Brown, A. J. P., E. A. Leibold, and H. N. Munro. 1983. Isolation of cDNA clones for the light subunit of rat liver ferritin: evidence that the light subunit is encoded by a multigene family. Proc. Natl. Acad. Sci. USA 80:1265–1269.
- Brush, D., J. B. Dodgson, O.-R. Choi, P. Wilkins Stevens, and J. D. Engel. 1985. Replacement variant histone genes contain intervening sequences. Mol. Cell. Biol. 5:1307-1317.
- 10. Choi, O.-R., C. Trainor, T. Graf, H. Beug, and J. D. Engel.

1986. A single amino acid substitution in v-erbB confers a thermolabile phenotype to ts167 avian erythroblastosis virus-transformed erythroid cells. Mol. Cell. Biol. **6:**1751–1759.

- Chou, C.-C., R. A. Gatti, M. L. Fuller, P. Concannon, A. Wong, S. Chada, R. C. Davis, and W. A. Salser. 1986. Structure and expression of ferritin genes in a human promyelocytic cell line that differentiates in vitro. Mol. Cell. Biol. 6:566–573.
- Costanzo, F., M. Colombo, S. Staempfli, C. Santoro, M. Marone, R. Frank, H. Delius, and R. Cortese. 1986. Structure of gene and pseudogenes of human apoferritin H. Nucleic Acids Res. 14: 721-736.
- Costanzo, F., C. Santoro, V. Colantuoni, G. Bensi, G. Raugei, V. Romano, and R. Cortese. 1984. Cloning and sequencing of a full length cDNA coding for a human apoferritin H chain: evidence for a multigene family. EMBO J. 3:23-27.
- 14. Didsbury, J. R., E. C. Theil, R. E. Kaufman, and L. F. Dickey. 1986. Multiple red cell ferritin mRNAs, which code for an abundant protein in the embryonic cell type, analyzed by cDNA sequence and by primer extension of the 5'-untranslated regions. J. Biol. Chem. 261:949–955.
- Dodgson, J. B., J. Strommer, and J. D. Engel. 1979. Isolation of the chicken β-globin gene and a linked embryonic β-like globin gene from a chicken DNA recombinant library. Cell 17:879–887.
- Dolan, M., J. B. Dodgson, and J. D. Engel. 1983. Analysis of the adult chicken β-globin gene. J. Biol. Chem. 258:3983–3990.
- Dorner, M. H., J. Salfeld, H. Will, E. A. Leibold, J. K. Vass, and H. N. Munro. 1985. Structure of human ferritin light subunit messenger RNA. Proc. Natl. Acad. Sci. USA 82:3139–3143.
- Engel, J. D., and J. B. Dodgson. 1978. Analysis of the adult and embryonic chicken globin genes in chromosomal DNA. J. Biol. Chem. 253:8239–8246.
- Fyrberg, E. A., K. L. Kindle, N. Davidson, and A. Sodja. 1980. The actin genes of *Drosophila*: a dispersed multigene family. Cell 19:365–378.
- Goldberg, D. A. 1980. Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. Proc. Natl. Acad. Sci. USA 77:5794-5798.
- Hentze, M. W., S. Keim, P. Papadopoulos, S. O'Brien, W. Modi, J. Drysdale, W. J. Leonard, J. B. Harford, and R. D. Klausner. 1986. Cloning, characterization, expression, and chromosomal localization of a human ferritin heavy-chain gene. Proc. Natl. Acad. Sci. USA 83:7226-7230.
- Jain, S. K., K. J. Barrett, D. Boyd, M. F. Favreau, J. Crampton, and J. W. Drysdale. 1985. Ferritin H and L chains are derived from different multigene families. J. Biol. Chem. 260:11762– 11768.
- Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry 16:4743-4751.
- Leibold, E. A., N. Aziz, A. J. P. Brown, and H. N. Munro. 1984. Conservation in rat liver of light and heavy subunit sequences of mammalian ferritin. J. Biol. Chem. 259:4327-4334.
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499-560.
- Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12:7035–7056.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19:269-276.
- Munro, H. N., and M. C. Linder. 1978. Ferritin: structure, biosynthesis, and role in iron metabolism. Physiological Rev. 58:317-396.
- Otsuka, S., I. Listowsky, Y. Niitsu, and I. Urushizaki. 1980. Assembly of intra- and interspecies hybrid apoferritins. J. Biol. Chem. 255:6234–6237.
- Palmieri, S., H. Beug, and T. Graf. 1982. Isolation and characterization of four new temperature-sensitive mutants of avian erythroblastosis virus (AEV). Virology 123:296–311.

- Rice, D. W., G. C. Ford, J. L. White, J. M. A. Smith, and P. M. Harrison. 1983. The spatial structure of horse spleen apoferritin. Adv. Inorg. Biochem. 5:39–50.
- 32. Samarut, J., and L. Gazzolo. 1982. Target cells infected by avian erythroblastosis virus differentiate and become transformed. Cell 28:921–929.
- Sanger, F., S. Nicklen, and A. R. Coulsen. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- Sargent, T. D., and I. B. Dawid. 1983. Differential gene expression in the gastrula of *Xenopus laevis*. Science 222:135–139.
- 35. Shinjyo, S. 1973. Isolation and properties of ferritin from chicken (Gallus domesticus, broiler) spleen. Seikagaku 45: 289–295.
- Shull, G. E., and E. C. Theil. 1982. Translational control of ferritin synthesis by iron in embryonic reticulocytes of the bullfrog. J. Biol. Chem. 257:14187-14191.
- Smith, D. R., and J. M. Calvo. 1980. Nucleotide sequence of the E. coli gene coding for dihydrofolate reductase. Nucleic Acids Res. 8:2255-2274.
- Theil, E. C. 1983. Ferritin: structure, function, and regulation. Adv. Inorg. Biochem. 5:1-38.

- Theil, E. C., and G. M. Tosky. 1979. Red cell ferritin and iron storage during chick embryonic development. Dev. Biol. 69: 666-672.
- 40. Vogt, P. K. 1969. Focus assay of Rous sarcoma virus, p. 198-211. In K. Habel and N. P. Salzman (ed.), Fundamental techniques in virology. Academic Press, Inc., New York.
- 41. Weaver, R. F., and C. Weissmann. 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. 7:1175–1193.
- Wilbur, W. J., and D. J. Lipman. 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA 80:726-730.
- 43. Yamamoto, M., N. S. Yew, M. Federspiel, J. B. Dodgson, N. Hayashi, and J. D. Engel. 1985. Isolation of recombinant cDNAs encoding chicken erythroid δ-aminolevulinate synthase. Proc. Natl. Acad. Sci. USA 82:3702-3706.
- 44. Zahringer, J., B. S. Baliga, and H. N. Munro. 1976. Novel mechanism for translational control in regulation of ferritin synthesis by iron. Proc. Natl. Acad. Sci. USA 73:857–861.
- Zinn, K., D. DiMaio, and T. Maniatis. 1983. Identification of two distinct regulatory regions adjacent to the human β-interferon gene. Cell 34:865–879.