Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 7729–7733, September 1991
Biochemistry

Cloning of the cDNA of the heme-regulated eukaryotic initiation factor 2α (eIF-2α) kinase of rabbit reticulocytes: Homology to yeast GCN2 protein kinase and human double-stranded-RNA-dependent eIF-2α kinase

(translational inhibitor/protein synthesis/cell cycle kinases)

JANE-JANE CHEN*†, MARK S. THROOP*, LEE GEHRKE*‡, IRENE KUO*, JAYANTA K. PAL*, MICHAEL BRODSKY*, AND IRVING M. LONDON*§

*Harvard-Massachusetts Institute of Technology Division of Health Sciences and Technology and §Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and ‡Department of Anatomy and Cellular Biology, Harvard Medical School, Boston, MA 02115

Contributed by Irving M. London, May 20, 1991

ABSTRACT **We have cloned the cDNA of the heme-regulated eIF-2α kinase (HRI) of rabbit reticulocytes. *In vitro* translation of mRNA transcribed from the HRI cDNA yields a 90-kDa polypeptide that exhibits eIF-2α kinase activity and is recognized by a monoclonal antibody directed against authentic HRI. The open reading frame sequence of the HRI cDNA contains all 11 catalytic domains of protein kinases with consensus sequences of protein-serine/threonine kinases in conserved catalytic domains VI and VIII. The HRI cDNA also contains an insert of ≈140 amino acids between catalytic domains V and VI. The HRI cDNA coding sequence has extensive homology to GCN2 protein kinase of *Saccharomyces cerevisiae* and to human double-stranded-RNA-dependent eIF-2α kinase. This observation suggests that GCN2 protein kinase may be an eIF-2α kinase in yeast. In addition, HRI has an unusually high degree of homology to three protein kinases (NimA, Wee1, and CDC2) that are involved in the regulation of the cell cycle.**

Protein synthesis in intact reticulocytes and their lysates is dependent on the availability of heme (1–4). In heme deficiency, protein synthesis is inhibited at the level of initiation due to the activation of a heme-regulated inhibitor (HRI), also called the heme-controlled repressor (refs. 4–6; for review, see refs. 7 and 8). HRI is a cAMP-independent protein kinase that specifically phosphorylates eIF-2α (9–12). Phosphorylation of eIF-2α [eIF-2(αP)] in reticulocyte lysates results in the binding and sequestration of reversing factor (RF), also designated as guanine nucleotide exchange factor or eIF-2B, in an RF·eIF-2(αP) complex. Since RF is required for the exchange of GTP for GDP in the recycling of eIF-2 and in the formation of the eIF-2·Met-tRNA^{Met}·GTP ternary complex, its unavailability results in the cessation of the initiation of protein synthesis (ref. 13; for review, see refs. 7 and 8).

Although the mechanism of regulation of protein synthesis by HRI has been extensively studied, little is known about the structure and regulation of HRI itself. We have reported (14) the amino acid sequences of three tryptic peptides of heme-reversible HRI. HRI peptide P-52 contains the sequence Asp-Phe-Gly, which is the most highly conserved short stretch in the conserved domain VII of protein kinases as presented by Hanks *et al.* (15). The N-terminal 14 amino acids of HRI peptide P-74 show 50–60% identity to the conserved domain IX of kinase-related transforming proteins (14). These findings are consistent with the autokinase and eIF-2α kinase activities of HRI. To study the structure–function relations of HRI, we undertook the molecular cloning of rabbit HRI cDNA,¶ which we report here.

METHODS

PCR Amplification of HRI cDNA Between P-52 and P-74. Poly(A)$^+$ mRNA (1 μg) was reverse-transcribed to obtain single-stranded cDNAs (16). The sense-strand oligodeoxynucleotide of P-52 and the antisense-strand oligodeoxynucleotide of P-74, deduced with preferred codon usage (17), were used as primers. PCR was carried out in the presence of single-stranded cDNA template and each primer at 1 μM for 40 cycles (94°C for 1 min, 47°C for 2 min, and 72°C for 3 min).

Preparation of the Rabbit Reticulocyte cDNA Library in λZap II and the Isolation of HRI cDNA Clones. Rabbit reticulocyte cDNAs were prepared using a Pharmacia cDNA synthesis kit. cDNAs >500 base pairs (bp) were pooled and were ligated to λZap II vector (Stratagene). The cDNA library obtained has 95% recombinant efficiency. The cDNA library on a nitrocellulose sheet was hybridized at 42°C overnight in a solution containing 5× Denhardt's solution/6× SSPE/salmon sperm DNA (500 μg/ml)/tRNA (1.7 mg/ml)/0.4% SDS/heat-denatured nick-translated ^{32}P-labeled HRI cDNA probe (10^6 cpm/ml). (1× SSPE = 0.18 M NaCl/10 mM sodium phosphate/1 mM EDTA, pH 7.4; 1× Denhardt's solution = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin). The nitrocellulose was then washed for three 5-min periods with 6× SSPE/0.1% SDS at room temperature, followed by two 10-min washes at 50°C under the same salt conditions. HRI cDNA was subcloned into pBlueScript plasmid by *in vivo* excision from the recombinant λZap II as described by Stratagene. The DNA sequence of HRI cDNA was determined by dideoxynucleotide chain-termination (18) as modified by Fawcett and Barlett (19).

Replacement of the 5′ Untranslated Leader Sequence of HRI mRNA. The 5′ untranslated leader sequence of the HRI cDNA was replaced by using PCR to introduce a unique *Nco* I site (CCATGG) at the initiating methionine [nucleotide (nt) 113], followed by ligation of the coding sequence to a vector containing the tobacco mosaic virus (TMV) untranslated leader sequence that was designed to provide both the initiating methionine and 3′-terminal *Nco* I site (L.G., unpublished data). The introduction of the *Nco* I site changes

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: eIF-2α, eukaryotic initiation factor 2α; HRI, heme-regulated eIF-2α kinase; dsI, double-stranded-RNA-dependent eIF-2α kinase; TMV, tobacco mosaic virus; nt, nucleotide(s).

†To whom reprint requests should be addressed at: Massachusetts Institute of Technology, E25-537, 77 Massachusetts Avenue, Cambridge, MA 02139.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. M69035).

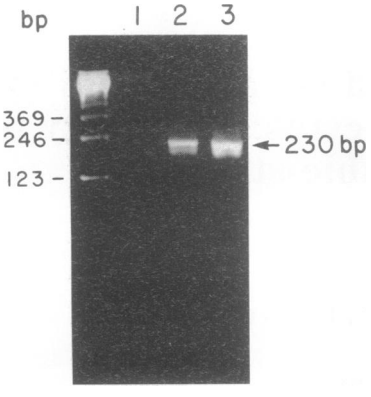

FIG. 1. Amplification of HRI cDNA between peptides P-52 and P-74 by PCR. Lanes: 1, P-52 primer only; 2 and 3, P-52 and P-74 primers; 1 and 2, 10 ng of template; 3, 7.5 ng of template. Amplified cDNA was analyzed by electrophoresis in a 1.5% agarose gel.

the second amino acid of HRI from leucine to valine, constituting a conservative substitution.

***In Vitro* Transcription and Translation and eIF-2α Kinase Assay of Translated Products.** Linearized HRI cDNAs were transcribed using T7 polymerase. *In vitro* translation of HRI mRNA (40 μg/ml) was carried out in the presence of [^{35}S]methionine as described by Promega using nuclease-treated reticulocyte lysates or wheat-germ extracts. Protein kinase assays were carried out in 40-μl reaction mixtures containing 10 μCi of [γ-^{32}P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq), 1.5 μl of translational mixture, and purified rabbit eIF-2 (1 μg) as indicated, at 30°C (reticulocyte lysate) or 25°C (wheat-germ extract) for 10 min as described (20).

RESULTS

PCR Amplification of HRI cDNA Between Peptides P-52 and P-74. The sequence homology of HRI tryptic peptides P-52 and P-74 to the conserved protein kinase domains VII and IX (14) makes it possible to predict that P-52 is positioned to the N-terminal side of P-74. This information permitted us to design primers for PCR amplification of a partial HRI cDNA. Using these two primers, we obtained two amplified cDNA fragments that were ≈230 bp long (Fig. 1, lanes 2 and 3). The upper band of the apparent doublet was much more prominent than the lower band. The lower band disappeared when extra time was provided at the end of 40 amplification cycles to complete the elongation of all the amplified cDNA. Omitting the P-74 primer in the PCR abolished the amplification of the 230-bp doublet (Fig. 1, lane 1), an indication that the 230-bp doublet is an amplification product of the two primers.

We have subcloned and sequenced this cDNA fragment. Excluding the 15-bp *Eco*RI restriction sites present on both primers, the remaining 219-bp sequence encodes an open reading frame for 73 amino acids. The newly obtained 38-amino acid sequence of HRI deduced from this cDNA sequence contains the consensus sequence Gly-Thr/Ser-Xaa-Xaa-Tyr/Phe-Xaa-Ala/Ser-Pro-Glu of protein-serine/threonine kinases located in the conserved domain VIII. This observation is consistent with the finding that HRI phosphorylates eIF-2α at Ser-51 (21). Furthermore, the amino acid sequences of HRI between conserved domains VII, VIII, and IX are unique to HRI.

Isolation of HRI cDNA. Approximately 150,000 recombinant clones were screened with the 234-bp probe of HRI. Among the 12 positive clones of the primary screen, 5 are full-length and contain a cDNA insert of ≈2700 bp. The 2729-nt sequence of HRI cDNA is presented in Fig. 2. There are 112 nt preceding the first ATG. Starting from this first ATG (nt 113), the open reading frame continues to nt 1990 encoding 626 amino acids followed by multiple stop codons

```
1                                                    CGCACGGCGC
11   TCGCGACCCGGACGCGCGAGGAGGCGGTCCCGGAGTCGGGGAGCTGGCGGG
62   TGGGCTGTGGTCCCCGCATTTGCGCGCGCGGGCGCCCCGCGTGACCGGCG
113  ATGCTGGGGGGCAGCGCCGGGACCCGCGGGGGCGAAGCCGAGGGCGACGGG
     MetLeuGlyGlySerAlaGlyThrArgGlyGlyGluAlaGluGlyAspGly   17
164  GCGGGGGCGGTGGGGGCGGTGGCCCCGCCGCCCCGCCATCGACTTCCCCGCT
     AlaGlyAlaValGlyAlaValAlaProProProAlaIleAspPheProAla   34
215  GAGGTGTCGGATCCCAAGTATGATGAGTCGGATGTCCCGGCAGAGCTGCAG
     GluValSerAspProLysTyrAspGluSerAspValProAlaGluLeuGln   51
266  GTGCTGAAGGAGCCGCTGCAGCAGCCAGCCTTCCCCTTCGCCGTCGCCAAC
     ValLeuLysGluProLeuGlnGlnProAlaPheProPheAlaValAlaAsn   68
317  CAGCTGCTGCTCGTCTCCCTGCTGGAGCACCTGAGTCATGTGCACGAGCCA
     GlnLeuLeuLeuValSerLeuLeuGluHisLeuSerHisValHisGluPro   85
368  AACCCGCTTCGCTCCAGACAGGTGTTTAAACTGCTCTGTCAGACCTTCATC
     AsnProLeuArgSerArgGlnValPheLysLeuLeuCysGlnThrPheIle  102
419  AAAATGGGGCTGCTGTCTTCCTTCACCTGCAGCGACGAGTTTAGCTCATTG
     LysMetGlyLeuLeuSerSerPheThrCysSerAspGluPheSerSerLeu  119
470  AGGCTGCATCACAACAGAGCTATTACGCATCTGATGAGGTCCGCCAGAGAG
     ArgLeuHisHisAsnArgAlaIleThrHisLeuMetArgSerAlaArgGlu  136
521  AGAGTTCGGCAGGATCCCTGTGCTGATAATTCTCATATCCAGAAAATCAGG
     ArgValArgGlnAspProCysAlaAspAsnSerHisIleGlnLysIleArg  153
572  TCGCGAGAAGTTGCCTTGGAAGCACAGACCTCACGATACTTGAATGAGTTT
     SerArgGluValAlaLeuGluAlaGlnThrSerArgTyrLeuAsnGluPhe  170
                                                  P-56
623  GAAGAGCTCTCCATCCTGGGGAAAGGTGGCTATGGCCGAGTGTACAAGGTC
     GluGluLeuSerIleLeuGlyLysGlyGlyTyrGlyArgValTyrLysVal  187
674  AGGAATAAATTAGATGGCCAGTATTATGCAATTAAAAAAATTCTGATTAAA
     ArgAsnLysLeuAspGlyGlnTyrTyrAlaIleLysLysIleLeuIleLys  204
725  GGTGCAACTAAAACAGATTGCATGAAGGTATTACGAGAAGTGAAAGTGCTG
     GlyAlaThrLysThrAspCysMetLysValLeuArgGluValLysValLeu  221
776  GCGGGCCTCCAGCACCCTAATATCGTAGGCTATCACACCGCGTGGATAGAG
     AlaGlyLeuGlnHisProAsnIleValGlyTyrHisThrAlaTrpIleGlu  238
827  CATGTCCACGTTCACGTTCAAGCAGACAGAGTTCCGATTCAGTTGCCTTCT
     HisValHisValHisValGlnAlaAspArgValProIleGlnLeuProSer  255
878  CTGGAAGTGCTCTCTGACCAGGAAGAAGACAGAGATCAATATGGTGTTAAA
     LeuGluValLeuSerAspGlnGluGluAspArgAspGlnTyrGlyValLys  272
929  AATGATGCAAGCAGCAGCTCATCCATTATTTTCGCTGAGTTCTCCCCAGAA
     AsnAspAlaSerSerSerSerSerIleIlePheAlaGluPheSerProGlu  289
980  AAAGAAAAATCCTCTGACGAATGTGCCGTTGAGAGTCAGAATAACAAACTG
     LysGluLysSerSerAspGluCysAlaValGluSerGlnAsnAsnLysLeu  306
1031 GTGAACTACACCACCAACTTAGTGGTGAGGGACACCGGTGAGTTTGAATCG
     ValAsnTyrThrThrAsnLeuValValArgAspThrGlyGluPheGluSer  323
1082 TCCACGGAGCGCCAAGAGAACGGCTCGATCGTGGAGCGTCAGCTACTGTTC
     SerThrGluArgGlnGluAsnGlySerIleValGluArgGlnLeuLeuPhe  340
1133 GGGCATAACTCAGACGTAGAAGAGGATTTCACGTCCGCGGAGGAATCTTCT
     GlyHisAsnSerAspValGluGluAspPheThrSerAlaGluGluSerSer  357
1184 GAGGAAGACTTAAGCGCGTTGCGGCACACAGAGGTGCAGTACCACCTGATG
     GluGluAspLeuSerAlaLeuArgHisThrGluValGlnTyrHisLeuMet  374
1235 CTGCATATCCAGATGCAGCTGTGCGAGCTGTCCCTGTGGGACTGGATCGCC
     LeuHisIleGlnMetGlnLeuCysGluLeuSerLeuTrpAspTrpIleAla  391
1286 GAGAGGAACAGGCGGAGCCGAGAGTGCGTGGACGAATCTGCCTGTCCTTAT
     GluArgAsnArgArgSerArgGluCysValAspGluSerAlaCysProTyr  408
1337 GTTATGGTCAGTGTTGCAACAAAAATTTTTCAAGAACTGGTGGAAGGTGTG
     ValMetValSerValAlaThrLysIlePheGlnGluLeuValGluGlyVal  425
1388 TTTTACATACATAACATGGGCATCGTGCACAGAGACCTGAAGCCTAGAAAT
     PheTyrIleHisAsnMetGlyIleValHisArgAspLeuLysProArgAsn  442
1439 ATTTTTCTTCATGGTCCTGATCAACAAGTGAAAATAGGAGACTTTGGTCTG
     IlePheLeuHisGlyProAspGlnGlnValLysIleGlyAspPheGlyLeu  459
                                                  P-52
1490 GCCTGCGCCGACATCATCCAGAAGAATGCGGCCCGGACCAGCAGAAACGGG
     AlaCysAlaAspIleIleGlnLysAsnAlaAlaArgThrSerArgAsnGly  476
1541 GAGAGAGCACCCACACACACTTCCCGAGTGGGCACCTGTCTGTACGCCTCG
     GluArgAlaProThrHisThrSerArgValGlyThrCysLeuTyrAlaSer  493
1592 CCCGAGCAGTTGGAAGGATCGGAGTATGATGCCAAGTCAGACATGTACAGC
     ProGluGlnLeuGluGlySerGluTyrAspAlaLysSerAspMetTyrSer  510
                                                  P-74
1643 GTCGGCGTGATCCTGCTGGAGCTCTTCCAGCCCTTCGGGACAGAGATGGAG
     ValGlyValIleLeuLeuGluLeuPheGlnProPheGlyThrGluMetGlu  527
1694 CGGGCAGAGGTCCTGACGGGCGTGCGAGCTGGCCGCATACCCGACTCCCTC
     ArgAlaGluValLeuThrGlyValArgAlaGlyArgIleProAspSerLeu  544
1745 AGTAAGAGGTGCCCGGCGCAGGCCAAGTACGTCCAGCTGCTGACCAGGAGG
     SerLysArgCysProAlaGlnAlaLysTyrValGlnLeuLeuThrArgArg  561
1796 AACGCGTCCCAGCGGCCGTCCGCCCTTCAGCTGCTGCAGAGTGAGCTCTTC
     AsnAlaSerGlnArgProSerAlaLeuGlnLeuLeuGlnSerGluLeuPhe  578
1847 CAGAACTCCGCGCATGTTAACCTCACCCTACAGATGAAGATAATAGAGCAG
     GlnAsnSerAlaHisValAsnLeuThrLeuGlnMetLysIleIleGluGln  595
1898 GAAAGAGAAATCGAGGAACTCAAGAAGCAGCTGAGCCTCCTCTCCCAGGCC
     GluArgGluIleGluGluLeuLysLysGlnLeuSerLeuLeuSerGlnAla  612
                                                  ***
1949 CGAGGGGTGAGGAGTGACAGGCGAGACGGAGAGCTCCCTGCCTAGCCGTCA
     ArgGlyValArgSerAspArgArgAspGlyGluLeuProAla           626
2000 CTCGGCCACGTCACAGGGGAACGTGGACTTGCACTTGCAGCAGTCAACTGG
2051 AATGGACAATTTCAAGCCTCCTGAGGTTCAGGCGGCATAATCCTCACTTGG
2102 AATCACTCAGCCCGCATGACTCTCCCCTCATGCTGCTCTTCCCGGAGGTAC
2153 CTCCTGGTGACCTCCTGGTGACTGCTCCCAATTAAACTTACGCTTTTCCCT
2204 TTCCTATTCCGCAAGTCCCATTCCTGAGCCTCCTACCTAAGCATTAACTAA
2255 ATCTTAGGTATCGGTCTCCATTCTTTCTCCTTTGAATCCTGGCCACCTCGC
2306 TCCTTTAGAAGCACACTCACTGCCCCGCCACCACCCAAGGCCAGGCCTGCA
2357 CCCTGGCGCAACAGCTGCCAGTCTTAGTCCTTAGCTGCTGCTGCTGTTGCC
2408 AGAGACACCTGCTCCGTTCACTCCCTCCAGGGTGGAAGCTCAGCCTGTGAG
2459 CAGCGCCTCTGCTCTCCCCGGCTGCAGCCCAGCGCCACTCGGGCAGGCTTC
2510 ACACGCTCACCCCAGGTGGCCTCGGAACAGCTGCGACAGCATCTCCCCGCA
2561 CCCTTCTGCCTTCTCAGCACTTGGCTCTCCAGCCAGCCTCTCCACTCACTC
2612 GTTTTTGTTTCCCGGAGCTGTCTGCCACAATGTTGGCAGTCTTCATGGACT
2663 ACTGTACGTGATTCTGCTGAATTTTAAATAAATAAACCCTGCAAATCAAAA
2714 AAAAAAAAAAAAAAAA
```

FIG. 2. (*A*) Deoxynucleotide sequence and deduced amino acid sequence of HRI cDNA. Numbers to the left indicate the position of nucleotides and numbers to the right indicate the position of amino acids. Asterisk indicates the first stop codon. Portions of deduced amino acid sequences that exactly match the amino acid sequences of HRI tryptic peptides (P-52, P-56, and P-74) are underlined and identified. The overlapping and repeated polyadenylylation signal sequence in the 3′ untranslated region AATAAA is underlined.

in the 3′ untranslated region of 739 nt. It should be noted that the first 250 nt of HRI cDNA are very G+C-rich (80%). We

obtained the nucleotide sequence in this area of HRI cDNA by using terminal deoxynucleotidyltransferase and pyrophosphatase. The overlapping repeat of the AATAAA polyadenylylation signal is found at nt 2689–2698, 11 nt from the poly(A) tail. The deduced amino acid sequence of the HRI cDNA contains the exact amino acid sequences of the three tryptic peptides of HRI obtained by microsequencing (14). P-52 is located in domain VII, P-56 is in domain I, and P-74 is in domain IX (Fig. 2).

***In Vitro* Transcription and Translation of HRI cDNA.** To determine the apparent molecular size of the protein encoded by the HRI cDNA and also to test for protein kinase activity, *in vitro* transcription and translation were carried out. Translation of HRI clone 2B mRNA in a nuclease-treated rabbit reticulocyte lysate yielded a predominant 90-kDa product as observed by SDS/PAGE (Fig. 3, lane 2). Similarly, the translation of the mRNA transcribed from each of the other four HRI cDNA clones also yielded a 90-kDa polypeptide (data not shown).

The nucleotide sequence data (Fig. 2) demonstrate that the 5′ untranslated leader sequence is extremely G+C-rich with the potential to form significant secondary structure. Secondary structure at the 5′ terminus of mRNAs is known to diminish mRNA translational efficiency (22). Indeed, we found that the HRI mRNA was not translatable in a wheat germ extract (Fig. 3, lane 5). Unlike the reticulocyte lysate, the wheat germ extract does not contain an endogenous HRI enzyme; therefore, expression of the HRI protein in the wheat germ system should facilitate analysis of kinase activity in the HRI translation products. The translational efficiency of mRNA transcripts can be increased by the use of untranslated leader sequences of some plant viral RNAs such as TMV that have been shown to act in cis (23, 24). Accordingly, we replaced the G+C-rich HRI untranslated leader sequence with that of TMV. The chimeric TMV–HRI mRNA was translated with ≈10-fold greater efficiency than HRI mRNA in the reticulocyte lysate (Fig. 3, lane 3), and translation in the wheat germ extract is also clearly evident (Fig. 3, lane 6). In all cases, the translated product of HRI mRNA migrated slightly faster than authentic purified phosphorylated HRI (Fig. 3, lane 4). This slight difference in mobility is most likely due to a lower level of phosphorylation in the translation products.

The Translational Product Encoded by HRI cDNA Is an eIF-2α Kinase. To determine whether the translational product derived from the mRNA of HRI cDNA is an eIF-2α kinase, a small portion of the total translation mixture was incubated with purified rabbit reticulocyte eIF-2 and [γ-^{32}P]ATP in the absence of added hemin under protein kinase assay conditions. The results presented in Fig. 4*A* show that translational products of HRI 2A and HRI 2B mRNAs have enhanced eIF-2α kinase activity (lanes 2 and 3) as compared to the control in the absence of added mRNA (lane 1). It should be emphasized that, under our kinase assay conditions (final hemin concentration, 0.75 μM), the activity of newly synthesized HRI exceeds the low activity of endogenous preformed HRI in the nuclease-treated lysate and makes it possible to detect enhanced phosphorylation of eIF-2α. In the absence of added purified rabbit eIF-2, only slight phosphorylation in the region of eIF-2α is observed (Fig. 4*B*, lane 4). Furthermore, the HRI polypeptide synthesized in the wheat-germ extracts exhibits eIF-2α kinase activity (Fig. 4*B*, lane 2) as does purified HRI (Fig. 4*B*, lane 3). It should be noted that there is no mammalian eIF-2α kinase activity in the wheat-germ extracts (Fig. 4*B*, lane 1), and our purified reticulocyte HRI phosphorylates purified wheat-germ eIF-2α very inefficiently (unpublished observation). In addition, the 90-kDa polypeptide expressed from HRI cDNA is immunoprecipitated by monoclonal antibodies to HRI (25) (data not shown). We conclude, therefore, that HRI cDNA is expressed as a 90-kDa protein with eIF-2α kinase activity.

Conserved Catalytic Domains of Protein Kinases in HRI cDNA. Hanks *et al.* (15) have compared and aligned the protein sequences of 65 protein kinases. They have identified 11 domains of protein kinases with invariant amino acid residues in each domain. The alignment of the HRI sequence with the sequences of a protein-serine/threonine kinase (Ca^{2+}/calmodulin protein kinase) and of a protein-tyrosine kinase (Src) is shown in Fig. 5. HRI cDNA contains all 11 catalytic domains with invariant amino acid residues (Fig. 5). The consensus ATP-binding sequence Gly-Xaa-Gly-Xaa-Xaa-Gly and the invariant valine residue located two positions downstream of the Gly-Xaa-Gly-Xaa-Xaa-Gly are conserved in HRI. In domain II, the invariant lysine residue has been shown to be indispensable and to be involved in the phosphotransferase activity of protein kinases (for review, see ref. 15). In HRI this invariant residue is Lys-199. Domain VI contains the consensus sequence that specifies either protein-serine/threonine kinases or protein-tyrosine kinases. HRI possesses Asp-Leu-Lys-Pro-Arg-Asn in domain VI, which is characteristic of protein-serine/threonine kinases (15). Asp-Phe-Gly located in domain VII is the most conserved short stretch in the catalytic domains of protein kinases and is probably involved in ATP binding. It is found in HRI as Asp-Phe-Gly (positions 456–458). In domain VIII the Ala/Ser-Pro-Glu consensus sequence essential for catalytic activity of protein kinases is also found in HRI. Domain

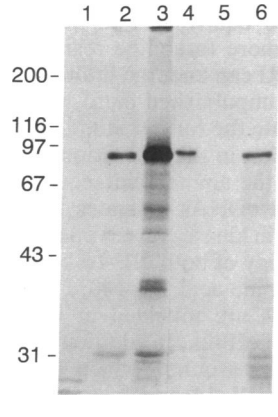

FIG. 3. *In vitro* translation of HRI mRNA transcribed from HRI cDNA. Lanes: 1–3, translation in nuclease-treated reticulocyte lysates; 1, minus mRNA; 2, HRI 2B mRNA; 3, TMV–HRI mRNA; 5 and 6, translation in wheat-germ extracts; 5, HRI 2B mRNA; 6, TMV–HRI mRNA; 4, ^{32}P-labeled HRI as a marker. Molecular sizes in kDa are shown.

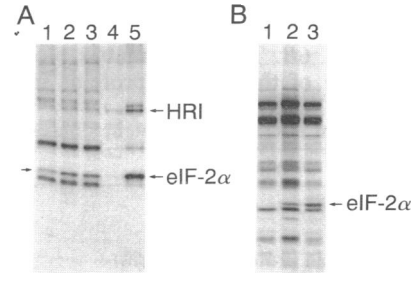

FIG. 4. *In vitro*-translated product of HRI cDNA is an eIF-2α kinase. (*A*) Protein kinase assays of translational products of HRI cDNAs from nuclease-treated reticulocyte lysates. Lanes: 1, minus mRNA control; 2, translational product of HRI 2A mRNA; 3, translational product of HRI 2B mRNA; 4, translational product of HRI 2B mRNA in the absence of added eIF-2; 5, purified HRI and eIF-2 as markers. (*B*) Protein kinase assays of translational products of HRI cDNA from wheat-germ extracts. Lanes: 1, minus mRNA control; 2, translational product of TMV–HRI mRNA; 3, minus mRNA control with added purified reticulocyte HRI.

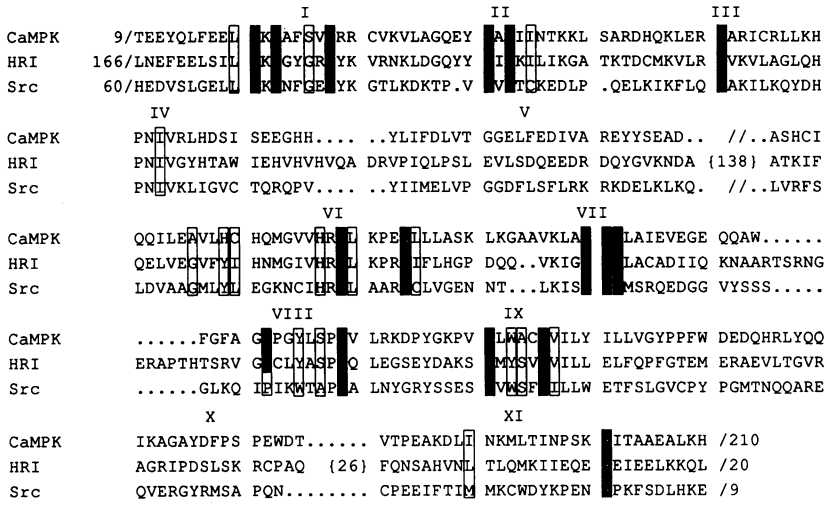

FIG. 5. Alignment of the conserved catalytic domains of HRI with other protein kinases. The conserved catalytic domains are indicated by the Roman numerals (I → XI). The conserved invariant amino acid residues are shown as solid boxes with white letters. The semiconserved amino acid residues of similar structure are shown as open boxes. Small gaps are shown by dots. There is an insertion of 138 amino acids in HRI between domains V and VI as indicated. The additional amino acids beyond the conserved domains are indicated by the numbers on both N and C termini. Single-letter amino acid code is used. CaMPK, Ca^{2+}/calmodulin protein kinase.

VIII of HRI contains the other consensus sequence for protein-serine/threonine kinases, Gly-Thr-Cys-Leu-Tyr, where underlined residues are conserved amino acids. The conserved amino acids in domain IX are also found in HRI. Thus, the homology of the deduced amino acid sequence of HRI cDNA to the conserved domains of other protein-serine/threonine kinases provides confirmatory evidence that HRI cDNA encodes a protein-serine/threonine kinase.

Kinase Insertion Sequence. As shown in Fig. 5, HRI cDNA contains an insertion of ≈140 amino acids between catalytic domains V and VI (amino acids 276–413). Similar large inserts have been reported for subclass III and IV receptor tyrosine kinases in which the kinase domains are divided into two halves by insertion of up to 100 mostly hydrophilic amino acid residues (for review, see ref. 26).

Extensive Homology of HRI to Yeast GCN2 Protein Kinase and Human Double-Stranded-RNA-Dependent eIF-2α Kinase (dsI). We have searched the GenBank for homology to other protein sequences of the amino acid sequence of HRI deduced from its cDNA. Of the 10 proteins with the highest scores (Table 1), 9 are protein-serine/threonine kinases, and of these, 3 are involved in regulation of the cell cycle (NimA, Wee1, and CDC2).

It is especially noteworthy that GCN2 protein kinase of yeast displays more homology to HRI than does dsI, the other known eIF-2α kinase (Table 1). The homology scores of HRI to GCN2 and dsI are significantly higher than scores to other protein kinases (Table 1). Human dsI cDNA has been cloned (28). A dot-matrix homology analysis of HRI and dsI coding sequences is shown in Fig. 6*A*, and a similar analysis of HRI and the kinase moiety of GCN2 coding sequences (30, 31) is shown in Fig. 6*B*. These dot-matrix plots reveal the extensive homology of these three proteins in the protein kinase catalytic domains I through X except for domain V where HRI has a large kinase insertion sequence. Homology in domains IX and X is observed with HRI, dsI, and GCN2, but not with the other eight protein kinases with the best scores. The significant homology in these regions suggests that these amino acids may be involved in the binding and phosphorylation of eIF-2 and also suggests that GCN2 protein kinase may be an eIF-2α kinase in yeast.

Table 1. Homology of HRI to other protein kinases

Kinase	Score
GCN2 protein kinase (yeast)	383
dsI (human)	331
Ca^{2+}/calmodulin protein kinase (rat)	252
Never-in-mitosis (*NimA*) gene product (yeast)	249
Wee1 gene product (yeast)	246
CDC2 gene product (human)	206
cAMP-dependent protein kinase (yeast)	205
Protein kinase C (rat)	192
Muscle light chain kinase (rat)	180
Src (rat)	180

Homology of the protein sequence of HRI to those of other proteins in GenBank was determined using the FASTA program (27).

DISCUSSION

The molecular cloning of the cDNA of HRI reveals an open reading frame that contains conserved catalytic domains of protein kinases and the consensus sequences of protein-serine/threonine kinases (Figs. 2 and 5). In the HRI cDNA, we find the exact amino acid sequences of the HRI tryptic peptides P-52, P-56, and P-74 that we obtained (14) by amino acid microsequencing. *In vitro* translation of HRI mRNA transcribed from HRI cDNA yields a predominant 90-kDa polypeptide that exhibits eIF-2α kinase activity (Figs. 3 and 4) and is recognized by an HRI monoclonal antibody (25). These properties are characteristic of authentic HRI.

Activation of HRI in reticulocyte lysates is accompanied by its phosphorylation (7, 25). Our purified nonrecombinant HRI undergoes heme-regulated autophosphorylation and eIF-2α phosphorylation (14). The sites of autophosphorylation of many protein kinases are located within 20 amino acids of the conserved Ala/Ser-Pro-Glu sequence in catalytic domain VIII (e.g., Thr-197 of cAMP-dependent protein kinase). Thr-483 of HRI and Thr-197 of cAMP-dependent protein kinase are equivalent. In addition, there are two serine and three more threonine residues in the vicinity of Thr-483. Since HRI can undergo multiple phosphorylation *in vitro* (ref. 32 and unpublished data), the availability of HRI cDNA will facilitate the further study of the sites and role of autophosphorylation in the activation of HRI.

Comparison of the amino acid sequences of HRI and dsI deduced from the cDNAs indicates that in addition to the general homology in kinase conserved domains, there is a very significant homology of both eIF-2α kinases around domains IX and X (HRI amino acid 511–540; Fig. 6*A*). It is likely that these amino acids are involved in eIF-2 binding and the phosphorylation of eIF-2α. In addition, HRI synthetic peptide P-74, which resides around domain IX, inhibits the eIF-2α kinase activity of HRI (14) and dsI (unpublished observation).

GCN2 protein kinase of yeast displays very significant homology to HRI (Table 1 and Fig. 6*B*) especially in domains IX and X in which considerable homology is observed only in eIF-2α kinases. GCN2 protein kinase stimulates the expression of amino acid biosynthetic genes under conditions of amino acid starvation by derepressing GCN4, a transcrip-

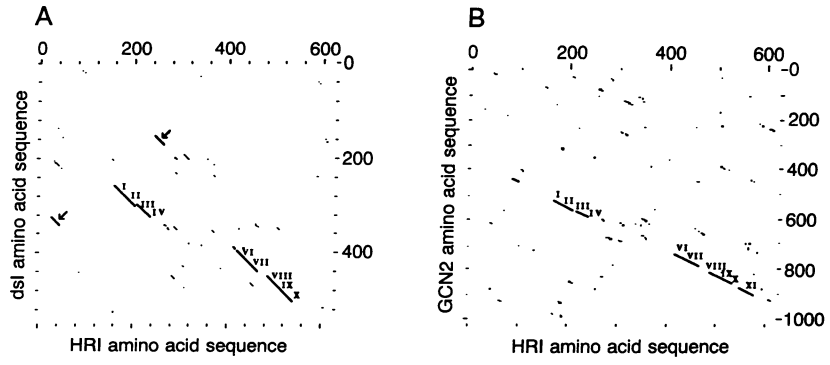

FIG. 6. (*A*) Dot-matrix analysis of the amino acid sequences of HRI and GCN2. (*B*) Dot-matrix analysis of the amino acid sequences of HRI and dsI. The dot matrix was performed using the Compare program (29) with a window of 30 and a stringency of 15. Locations of the conserved catalytic domains of protein kinases are indicated.

tional activator of these genes (for review, see ref. 33). The derepression of GCN4 by GCN2 protein kinase occurs at the level of translation of GCN4 mRNA (34, 35). The activation of the translation of GCN4 mRNA coincides with a decrease in the rate of general polypeptide chain initiation at the level of the eIF-2-dependent 43S preinitiation complex formation (36). Furthermore, a yeast strain that overexpresses GCN2 protein kinase has been reported to have a lower rate of protein synthesis (36). Thus, the effect of GCN2 protein kinase on protein synthesis is very similar to that of HRI. The molecular cloning of yeast eIF-2α (37) reveals 58% homology of its amino acid sequence to human eIF-2α (38). In addition, consensus phosphorylation site Ser-51 is conserved in yeast eIF-2α, and the phosphorylation of yeast eIF-2α has been demonstrated (37). The possibility that GCN2 protein kinase may phosphorylate eIF-2 has been raised by Cigan *et al.* (37) and Tzamarias *et al.* (36). The alignment of the amino acid sequences of HRI and GCN2 indicates 52% similarity and 28% identity in the kinase moiety of GCN2. This extensive homology of HRI and GCN2 further supports the view that GCN2 may be an eIF-2α kinase in yeast.

HRI cDNA contains a unique insertion sequence of $\approx$140 amino acids located between domains V and VI (Fig. 5). A similar insert in this location has been reported for subclasses III and IV of receptor tyrosine kinases (for review, see ref. 26) that include the platelet-derived growth factor receptor, the colony-stimulating factor 1 receptor, and the c-*kit* protooncogene product. Since kinase insertion sequences are highly conserved among species for each specific receptor, the kinase insert may play an important role in the action of receptor kinases. Indeed, the platelet-derived growth factor receptor kinase insert contains an autophosphorylation site (Tyr-751), and mutation of Tyr-751 to Phe or Gly blocks association of the platelet-derived growth factor receptor with phosphatidylinositol kinase and three other cellular proteins (39). For HRI, heme binds to HRI and regulates its kinase activities (7, 32). It will be of interest to determine whether the kinase insertion sequence of HRI is involved in the binding of heme and the regulation of the autokinase and eIF-2α kinase activities.

It is intriguing that HRI has significant homology to three yeast protein kinases (NimA, Wee1, and CDC2) that are involved in regulation of the cell cycle. This finding raises the possibility that HRI may also play a role in erythroid proliferation and differentiation.

We are most grateful to Dr. Mason Freeman for his advice in amplification of HRI cDNA by PCR. This investigation was supported by grants from the National Institutes of Health (DK 16272 and GM 42504) and from the National Science Foundation (DMB-890538). I.K. and M.B. were supported by the Undergraduate Research Opportunity Program of the Massachusetts Institute of Technology.

1. Bruns, G. P. & London, I. M. (1965) *Biochem. Biophys. Res. Commun.* **18**, 236–242.
2. Zucker, W. V. & Schulman, H. M. (1968) *Proc. Natl. Acad. Sci. USA* **59**, 582–589.
3. Adamson, S. D., Herbert, E. & Kemp, S. F. (1969) *J. Mol. Biol.* **42**, 247–258.
4. Hunt, T., Vanderhoff, G. A. & London, I. M. (1972) *J. Mol. Biol.* **66**, 471–481.
5. Howard, G. A., Adamson, S. D. & Herbert, E. (1970) *Biochim. Biophys. Acta* **213**, 237–240.
6. Maxwell, C. R., Kamper, C. S. & Robinovitz, M. J. (1971) *J. Mol. Biol.* **58**, 317–327.
7. London, I. M., Levin, D. H., Matts, R. L., Thomas, N. S. B., Petryshyn, R. & Chen, J.-J. (1987) in *The Enzymes*, eds. Boyer, P. D. & Krebs, E. G. (Academic, New York), 3rd Ed., Vol. 17, pp. 359–380.
8. Pain, V. M. (1986) *Biochem. J.* **235**, 625–637.
9. Levin, D. H., Ranu, R. S., Ernst, V. & London, I. M. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3112–3116.
10. Ranu, R. S. & London, I. M. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4349–4353.
11. Kramer, G., Cimadevilla, M. & Hardesty, B. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3078–3082.
12. Farrell, P., Balkow, K., Hunt, T., Jackson, R. J. & Trachsel, H. (1977) *Cell* **11**, 187–200.
13. Amesz, H., Goumans, H., Haubrich-Morree, T., Voorma, H. O. & Benne, R. (1979) *Eur. J. Biochem.* **98**, 513–520.
14. Chen, J.-J., Pal, J. K., Petryshyn, R., Kuo, I., Yang, J. M., Throop, M. S., Gehrke, L. & London, I. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 315–319.
15. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) *Science* **241**, 42–52.
16. Frohman, M. A., Dush, M. V. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
17. Lathe, R. (1985) *J. Mol. Biol.* **183**, 1–12.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
19. Fawcett, T. W. & Bartlett, S. G. (1990) *BioTechniques* **9**, 46–48.
20. Chen, J.-J., Yang, J. M., Petryshyn, R., Kosower, N. & London, I. M. (1989) *J. Biol. Chem.* **264**, 9559–9564.
21. Pathak, V. K., Schindler, D. & Hershey, J. W. B. (1988) *Mol. Cell. Biol.* **8**, 993–995.
22. Pelletier, J. & Sonenberg, N. (1985) *Cell* **40**, 515–526.
23. Gallie, D. R., Sleat, D. E., Watts, J. W., Turner, P. C. & Wilson, T. M. A. (1987) *Nucleic Acids Res.* **15**, 8693–8711.
24. Gehrke, L. & Jobling, S. A. (1990) in *Post-Transcriptional Control of Gene Expression*, eds. McCarthy, J. E. G. & Tuite, M. (Springer, Berlin), Vol. 49, pp. 389–398.
25. Pal, J. K., Chen, J.-J. & London, I. M. (1991) *Biochemistry* **30**, 2555–2562.
26. Ullrich, A. & Schlessing, J. (1990) *Cell* **61**, 203–212.
27. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
28. Meurs, E., Chong, K., Galabru, J., Thomas, N. S. B., Kerr, I. M., Williams, B. R. G. & Hovanessian, A. G. (1990) *Cell* **62**, 379–390.
29. Maizel, J. V., Jr., & Lenk, R. P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7665–7669.
30. Roussou, I., Thireos, G. & Hauge, B. M. (1988) *Mol. Cell. Biol.* **8**, 2132–2139.
31. Wek, R. C., Jackson, B. M. & Hinnebusch, A. G. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4579–4583.
32. Fagard, R. & London, I. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 866–870.
33. Hinnebusch, A. G. (1988) *Microbiol. Rev.* **52**, 248–273.
34. Thireos, G., Driscoll-Penn, M. & Greer, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5096–5100.
35. Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6442–6446.
36. Tzamarias, D., Roussou, I. & Thireos, G. (1989) *Cell* **57**, 947–954.
37. Cigan, A. M., Pabich, E. K., Feng, L. & Donahue, T. F. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2784–2788.
38. Ernst, H., Duncan, R. F. & Hershey, J. W. B. (1987) *J. Biol. Chem.* **262**, 1206–1212.
39. Kazlauskas, A. & Cooper, J. A. (1990) *Cell* **58**, 1121–1133.