Vol. 210, No. 2, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 16, 1995 Pages 556-566

A NOVEL GENE ORIENTED IN A HEAD-TO-HEAD CONFIGURATION WITH THE HUMAN HISTIDYL-tRNA SYNTHETASE (HRS) GENE ENCODES AN mRNA THAT PREDICTS A POLYPEPTIDE HOMOLOGOUS TO HRS

Terrance P. O'Hanlon[1], Nina Raben*, and Frederick W. Miller

Molecular Immunology Laboratory, Center for Biologics Evaluation and Research, Food and Drug Administration and *National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892

Received April 10, 1995

The human histidyl-tRNA synthetase (HRS) gene encodes an enzyme that catalyzes the esterification of histidine to its cognate tRNA as an early step in protein biosynthesis. Previous reports have described a bidirectional promoter element which coordinates the transcription of both HRS and an unknown mRNA whose gene is oriented in a head-to-head configuration with HRS. We have isolated and characterized a human genomic DNA clone that encodes portions of these oppositely transcribed mRNAs and a putatively full-length cDNA clone (HO3) corresponding to the gene mapping immediately 5' of HRS. The largest open reading frame within HO3 (1518 bp) shares approximately 75% nucleotide sequence identity with human HRS (1527 bp) and predicts a polypeptide with extensive amino acid sequence homology with the HRS protein (72%). Moreover, amino acid sequence motifs characteristic of class II aminoacyl-tRNA synthetases are conserved within HO3. Despite their similarity, HRS and HO3 have divergent amino-terminal domains which correspond to the first two exons of each gene. RNA blot analysis revealed that HRS (2.0 kb) and HO3 (2.5 kb) exhibit distinct patterns of steady-state mRNA expression among multiple human tissues.

Aminoacyl-tRNA synthetases (aaRS) play a fundamental role in the translation of the genetic code by coordinating the appropriate attachment of individual amino acids (aa) to their cognate tRNAs. At least 20 different aaRSs are required to synthesize the full complement of aa-tRNA complexes required for protein biosynthesis. A combination of genetic and crystallographic studies suggest that two structurally distinct classes of aaRS genes may have evolved independently (1,2). Class I aaRSs share two consensus peptide sequences (referred to as HIGH and KMSKS) which border a Rossmann fold structure characteristic of these and other

[1] Corresponding author. FAX: (301) 827-0449.

0006-291X/95 $5.00

nucleotide binding proteins (3). In contrast, class II aaRSs lack the Rossmann fold domain but alternatively possess three distinct polypeptide motifs (designated motifs 1-3) which are positioned within the catalytic core of these enzymes (4). In addition, aaRSs specific for each amino acid possess unique and evolutionarily conserved sequences (signature regions) required for cognate amino acid and tRNA recognition functions (3,5).

The family of aaRSs are commonly referred to as "housekeeping" enzymes in keeping with their constitutive expression in all tissues. Despite this classification, several studies indicate that mammalian aaRSs have many structural and functional features not typically associated with housekeeping genes (6-9). Such features have recently been ascribed to human histidyl-tRNA synthetase (HRS), one of the more intensively studied human aaRSs owing to its recognition as an autoantigen in inflammatory muscle disease (10,11).

Genomic and cDNA sequences encoding HRS have been isolated from a variety of sources, including human cells, and have been shown by comparative evolutionary analysis to have well-conserved class II aaRS structural motifs (5). The human HRS gene directs the synthesis of a 2.0 kb mRNA which encodes the cytosolic form of the enzyme (12,13). Molecular analyses have identified a distinct, opposite-strand transcription unit oriented in a head-to-head configuration with the human HRS gene (13). Curiously, both transcripts originate from a small bidirectional promoter element mapping upstream of the HRS gene (13). In this report, we describe a cDNA clone which corresponds to the human HRS bidirectional transcript and predicts a polypeptide with strong homology to HRS.

MATERIALS AND METHODS

Genomic and cDNA clones: A 630 bp genomic DNA fragment spanning sequences upstream of the human HRS gene [Fig. 1A, map positions -554 to +76 relative to the HRS translational start codon (+1)ATG] was generated by inverse PCR chromosome walking as described previously (5). The HRS 630 bp fragment was labeled with α-[^{32}P]dATP by random priming (14) and used to screen both a human chromosome 5-specific genomic DNA library (ATCC #57720) and a human skeletal muscle cDNA library (Stratagene) by established methods (15). In each instance, positively-hybridizing clones were plaque-purified and subsequently cloned into the pBluescript KS+ vector (Stratagene) for DNA sequence analysis (Sequenase, Amersham Corp.). Protein and nucleic acid sequence analyses were performed using the PC/Gene software package (Intelligenetics, Inc.).

RNA blot analysis: A 1.5 kb probe encompassing the entire HRS open reading frame (ORF) was generated by PCR using a previously described HRS cDNA clone as template (EMBL accession number Z11518)(5). A 1.4 kb probe which spans approximately 90% of the largest HO3 ORF was generated by restriction enzyme digestion of the HO3 cDNA clone (see Fig. 1B, KpnI/EcoRI fragment). RNA blot analysis was performed using a commercially-prepared human multiple tissue mRNA blot (Clontech Laboratories, Inc.). Random-primed probes specific for HRS (1.5 kb), and HO3 (1.4 kb), and a human genomic DNA fragment which spans the first exon of each gene (630 bp, see Fig. 1A) were utilized for successive hybridizations of the mRNA blot according to the manufacturer's instructions. Following each round of hybridization and autoradiography, the mRNA blot was stripped and verified to be free

of bound probe by autoradiography. HRS and HO3 cDNA probes of comparable length and specific activity were utilized for sequential hybridization of the identical mRNA blot in order to approximate relative differences in steady-state mRNA expression among multiple human tissues. Hybridization for human ß-actin mRNA expression was performed by both the manufacturer (Clontech Laboratories Inc, Lot Release #3X832) and our laboratory to confirm that relatively equivalent amounts of mRNA (2 ug per lane) were loaded for each human tissue (data not shown).

In vitro translation: Full-length cDNAs encoding HRS and the putative HO3 polypeptide were cloned downstream of either the T7 (HRS) or T3 (HO3) bacteriophage promoter in the pBluescript vector (Stratagene) and subsequently utilized as template for in vitro translation using reticulocyte lysate extracts (Promega Co.). [^{35}S]methionine-labeled translation products were size-fractionated in a 10% SDS-PAGE gel and visualized by standard autoradiography (15).

RESULTS

A 630 bp DNA fragment containing genomic sequences located 5' of the human HRS translational start codon [Fig. 1A; -554 to +76 relative to HRS (+1)ATG] was used to isolate complementary clones from both a human chromosome 5-specific genomic DNA library and a human skeletal muscle cDNA library as described in Materials and Methods. Two genomic clones (designated C5-1 and C5-2) were isolated from a lambda Charon21A library and shown to have identical 6.1 kb EcoRI insert fragments by restriction enzyme and DNA sequence analysis. DNA sequence analysis verified that C5-1/C5-2 shared 100% sequence identity with an independently derived human HRS genomic DNA fragment (630 bp, see Materials and Methods). Further analysis of genomic clone C5-1 confirmed an overlap with the promoter region, exon 1, and exon 2 of the human HRS gene as illustrated in Fig. 1A (GenBank accession number U18936). In addition, C5-1 contained 3.9 kb of genomic information mapping upstream of the HRS translational start codon, a region predicted to span portions of a previously uncharacterized gene located 5' of HRS (13). Consistent with this idea, we have simultaneously isolated a putatively full-length cDNA clone (designated HO3) which corresponds to the gene positioned 5' of the human HRS gene (Fig. 1B). The 5' end of the HO3 cDNA maps 195 bp upstream of the HRS translational start codon (ATG) and is transcribed in the direction opposite of HRS. The head-to-head configuration of these genes is consistent with the functional definition of a bidirectional promoter element mapping between HRS and HO3 (13). The boundries of HO3 exons I-IV were delineated by comparative sequence analysis of genomic clone C5-1 and the HO3 cDNA (Fig. 1A). The HO3 cDNA clone (2435 bp, GenBank accession number U18937) is illustrated in Fig. 1B depicting the putative boundries of the 5'and 3'-untranslated regions and the intervening open reading frame (ORF).

To investigate further the potential for bidirectional gene expression at the human HRS locus, we surveyed a variety of human tissues for both HRS and HO3 mRNA expression. A human multiple tissue mRNA blot (Fig. 2; lanes 1-8, heart, brain, placenta, lung, liver, skeletal

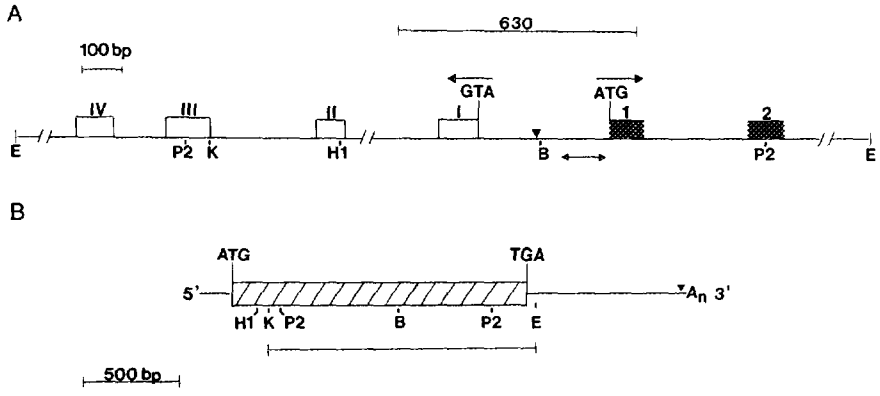

Figure 1. Structural overview of genomic and cDNA clones encoding HRS and HO3. A. Schematic illustration and partial restriction map of a 6.1 kb EcoRI DNA fragment isolated from a human chromosome 5-specific genomic library (designated C5-1). Genomic clone C5-1 was characterized by both restriction enzyme (E, EcoRI; K, KpnI; P2, PvuII; B, BamHI; H1, HpaI) and DNA sequence analyses as described in Materials and Methods. C5-1 was shown to include portions of the human HRS gene (exons 1 and 2, filled boxes) and a previously uncharacterized gene (exons I - IV, open boxes) mapping upstream and oriented opposite of human HRS (13). The direction of transcription of each gene (arrows) is illustrated above their respective translational start codons (ATG). Partial exon/intron boundries of each gene were defined by comparison of genomic sequence information with the human HRS cDNA sequence (EMBL accession number Z11518) and a novel cDNA sequence (HO3, described below) corresponding to the opposite-strand transcription unit. In each instance, exon/intron boundries conformed to expected splice site consensus sequences (data not shown). The position of the HRS bidirectional promoter described by Tsui *et al.* (13) is designated below the schematic (↔). A human genomic DNA fragment linking HRS exon 1 and HO3 exon I is illustrated above C5-1 (630 bp, described in Materials and Methods). The 5'-end of the HO3 cDNA (described below) is shown for reference (▾). B. A partial restriction map of a cDNA clone (designated HO3) corresponding to the gene mapping 5' of the human HRS gene is illustrated. 5'- and 3'-untranslated regions (149 bp and 760 bp, respectively) are shown flanking the putative open reading frame (1518 bp, hatched box). A consensus polyadenylation signal sequence (▾) was identified 17 bp upstream of a polyadenylate tract positioned at the 3'-terminus of HO3 (A_n). The position of the 1.4 kb KpnI/EcoRI HO3 probe is designated below the diagram.

muscle, kidney, and pancreas, respectively) was sequentially hybridized with cDNA probes specific for the HRS and HO3 ORFs (Fig. 2A and 2B, respectively), and a genomic DNA probe which spans the first exon of each gene (Fig. 2C) as described in Materials and Methods. Variable levels of steady-state HRS mRNA (2.0 kb) were detected among the multiple tissues surveyed; relatively higher levels of HRS mRNA were detected in heart, brain, liver, and kidney (Fig. 2A, lanes 1, 2, 5, and 7, respectively). As shown in Fig. 2B, HO3 mRNA levels also varied among different tissues, but in contrast to HRS, higher levels of HO3 mRNA (2.5 kb) were present in heart, skeletal muscle, and kidney (lanes 1, 6, and 7, respectively) while lower levels of HO3 mRNA were detected in brain and liver (lanes 2 and 5, respectively). As anticipated, the human genomic probe which includes the first exon of HRS and HO3 (630 bp, see Fig. 1A) hybridized to both the 2.0 kb (HRS) and 2.5 kb (HO3) transcripts confirming their close proximity within the HRS locus (Fig. 2C).

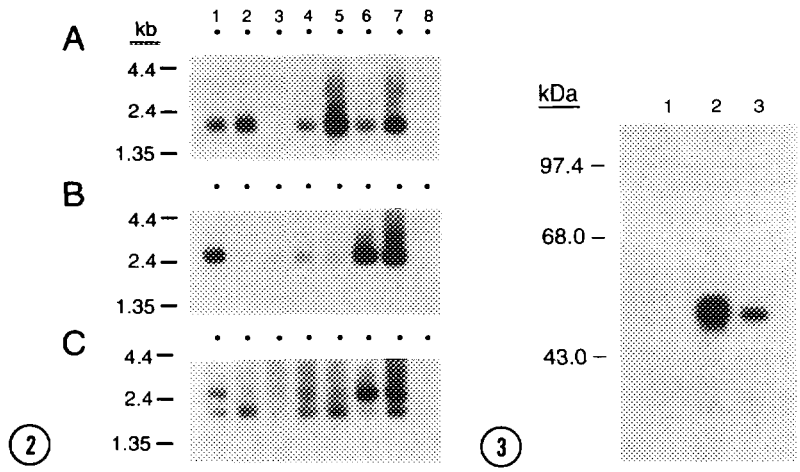

Figure 2. Human tissue survey of HRS and HO3 mRNA expression. A mRNA blot (2 μg polyadenylated mRNA per lane) representing multiple human tissues (lanes 1-8, heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas, respectively) was sequentially hybridized with cDNA probes specific for HRS (panel A, 48 hr exposure), and HO3 (panel B, 48 hr exposure), and a genomic DNA probe which spans the first exon of each gene (panel C, 72 hr exposure) as described in Materials and Methods. Although relative HRS and HO3 mRNA levels varied considerably between tissues (e.g., compare lanes 1 and 8 in each panel), each was clearly present in all tissues as demonstrated by longer film exposures. The positions of RNA size markers (diagrammed at the left of each panel) were used to estimate the sizes of the HRS (2.0 kb) and HO3 (2.5 kb) transcripts.

Figure 3. In vitro transcription/translation of HRS and HO3 cDNAs. cDNA clones spanning the full-length HRS and HO3 ORFs were cloned into a T3/T7 bacteriophage expression plasmid and subsequently used as templates in coupled in vitro transcription/translation reactions (see Materials and Methods). [^{35}S]methionine-labeled reaction products corresponding to a translation control (lane 1, plasmid vector minus a DNA insert), HO3 (lane 2), and HRS (lane 3) were size-fractionated in a 10% SDS polyacrylamide gel and developed by standard autoradiographic technique (24 hr exposure). The migrations of protein molecular weight standards are indicated to the left (kDa).

The largest ORF within the HO3 cDNA (1518 bp) predicts a polypeptide with a molecular weight of approximately 56.9 kDa. To examine the protein-coding potential of HO3, the full-length cDNA was cloned into a bacteriophage T3/T7 expression vector and subsequently utilized as template in a coupled in vitro transcription/translation reaction as described in Materials and Methods. Parallel reactions were performed using a full-length HRS ORF expression vector as a molecular weight reference (predicted MW = 57.4 kDa) and the plasmid vector devoid of insert DNA (negative control). As shown in Fig. 3, the primary HO3 and HRS translation products (lanes 2 and 3, respectively) comigrate in a size range consistent with that observed for HRS upon SDS-PAGE (16). No signal was detected from the plasmid vector lacking a DNA insert (lane 1).

An alignment of human HRS and HO3 cDNAs revealed a marked degree of primary nucleotide sequence identity (approximately 75%) within their protein-coding domains (1527 bp

```
HO3         MPLLGLLPRRAWASLLSQLLRPPCASCTGAVRCQSQVAEAV-LTSQLKAHQEKPN    54
HRS         MAERAAL--EELVKLQGERVRGLKQQKASAELIEEEVAKLLKLKAQLGPDESKQK    53
YEA         MS----------------------SATAAATSAPTANAANALKASKAPKKGKLQ    32
ECO         MAKN--------------------------------------------------     4
            *.

HO3         FIIKTPKGTRDLSPQHMVVREKILDLVISCFKRHGAKGMDTPAFELKETLTEKYG   109
HRS         FVLKTPKGTRDYSPRQMAVREKVFDVIIRCFKRHGAEVIDTPVFELKETLMGKYG   108
YEA         VSLKTPKGTKDWADSDMVIREAIFSTLSGLFKKHGGVTIDTPVFELREILAGKYG    87
ECO         --IQAIRGMNDYLPGETAIWQRIEGTLKNVLGSYGYSEIRLPIVEQTPLFKRAIG    57
              ... .* .*    . .... . . .  . ..*   .  * *    .    *

HO3         EDSGLM----YDLKDQGGELLSLRYDLTVPFARYLAMNKV---KKMKRYHVGKVW   157
HRS         EDSKLI----YDLKDQGGELLSLRYDLTVPFARYLAMNKL---TNIKRYHIAKVY   156
YEA         EDSKLI----YNLEDQGGELCSLRYDLTVPFARYVAMNNI---QSIKRYHIAKVY   135
ECO         EVTDVVEKEMYTFEDRNGDSLTLRPEGTAGCVRAGIEHGLLYNQEQRLWYIGPMF   112
            * . ..    *...*..*.  .** . *.  .*    . .    . . .... ..

HO3         RRESPTIVQGRYREFCQCDFDIAGQFDPMIPDAECLKIMCEILSGLQLGDFLIKV   212
HRS         RRDNPAMTRGRYREFYQCDFDIAGNFDPMIPDAECLKIMCEILSSLQIGDFLVKV   211
YEA         RRDQPAMTKGRMREFYQCDFDVAGTFESMVPDSECLSILVEGLTSLGIKDFKIKL   190
ECO         RHERPQ--KGRYRQFHQLGCEVFGLQGPDI-DAELIMLTARWWRALGISE-HVTL   163
            *.. *   .** *.* * . .. *  .. . *.* . .      .* . .  ...

HO3         NDRRIVDGMFAVCGVPESKFRAICSSIDKLDKMAWKDVRHEMVVKKGLAPEVADR   267
HRS         NDRRILDGMFAICGVSDSKFRTICSSVDKLDKVSWEEVKNEMVGEKGLAPEVADR   266
YEA         NHRKILDGIFQIAGVKDEDVRKISSAVDKLDKSPWEAVKKEMTEEKGQSEETADK   245
ECO         -ELNSIGSLEARANYRDALVAFLEQHKEKLD----EDCKRRMYTN-------PLR   206
             . . .... . . ..    .    .***    .. .. * .       . .

HO3         IGDYVQCHGGVSLVEQMF-QDPRLSQ--NKQALEGLGDL-KLLFEYLTLFGIADK   318
HRS         IGDYVQQHGGVSLVEQLL-QDPKLSQ--NKQALEGLGDL-KLLFEYLTLFGIDDK   317
YEA         IGEYVKLNGSLKEIHAVLSADANITS--NEKAKQGLDDI-ATLMKYTEAFDIDSF   297
ECO         V-----LDSKNPEVQALLNDAPALGDYLDEESREHFAGLCKLLESAGIAYTVNQR   256
            .      ..    ..... ... ..   .... . ....   * .    ....

HO3         ISFDLSLARGLDYYTGVIYEAVL----------LQTPTQAGEEP---LNVGSVA   359
HRS         ISFDLSLARGLDYYTGVIYEAVL----------LQTPAQAGEEP---LGVGSVA   358
YEA         ISFDLSLARGLDYYTGLIYEVVTSASAPPENASELKKKAKSAEDASEFVGVGSIA   352
ECO         ------LVRGLDYYNRTVFEWVTNSLGSQ----------------------GTVC   283
                  *.******.  ..* *                              *..

HO3         AGGRYDGLVGMF-DPKGHK---VPCVGLSIGVERIFYIVEQRMKTKGEKVRTTET   410
HRS         AGGRYDGLVGMF-DPKGRK---VPCVGLSIGVERIFSIVEQRLEALEEKIRTTET   409
YEA         AGGRYDNLVNMFSEASGKKSTQIPCVGISFGVERIFSLIKQRINS-STTIKPTAT   406
ECO         AGGRYDGLVEQLGGRAT------PAVGFAMGLERLVLLVQ----AVNPEFKADPV   328
            ******.**. . .  .      * **...*.**.  ...    . . . ... .

HO3         -QVFVAT--PQKN---F-LQERLKLIAELWDSGIKAEMLYKNNPKLLTQLHYCES   458
HRS         -QVLVAS--AQKK---L-LEERLKLVSELWDAGIKAELLYKKNPKLLNQLQYCEE   457
YEA         -QVFVMAFGGGKDWTGY-LPERMKVTKQLWDAGIEAEYVYKAKANPRKQFDTTKK   459
ECO         VDIYLVASGADTQSAAMALAERLR--DEL--PGVKL-MTNHGGGNFKKQFARADK   378
             .. . .    ..     * **..   .*  .*..     .   .  .*.   ..

HO3         TGIPLVVIIGEQELKEGVIKIRSV----ASREEVAIKRENFVAEIQKR----LSE   505
HRS         AGIPLVAIIGEQELKDGVIKLRSV----TSREEVDVRREDLVEEIKRRTGQPLCI   508
YEA         AGCHIAVILGKEEYLEGKLRVKRLGQEFADDDGELVSAADIVPIVQEKLSQIHED   514
ECO         WGARVAVVLGESEVANGTAVVK----DLRSGEQTAVAQDSVAAHLRTLLG-----   424
             * ......*. *  .*   ..       . .   .  .. .  ..

HO3         S-----------   506
HRS         C-----------   509
YEA         GLNEVTRLIKGL   526
ECO         ------------   424
```

Figure 4. Primary amino acid sequence alignment of HO3 and HRS gene sequences. A translation of the largest HO3 ORF (506 aa) was aligned with the primary amino acid sequence of human HRS (509 aa) using the PC/Gene CLUSTAL program. Sequences of yeast HRS (YEA) and E. coli HRS (ECO) genes were included for comparison. Sites of primary amino acid sequence identity (*) and conservative substitution (.) are designated below the alignment. Spaces were introduced by CLUSTAL to optimize the alignment (-).

and 1518 bp, respectively)[data not shown]. An alignment of the primary amino acid sequence predicted by the largest HO3 ORF (506 aa) and the amino acid sequence of human HRS (509 aa) is shown in Fig. 4. HRS gene sequences derived from yeast (YEA) and E. coli (ECO) were aligned for comparison. Sites of amino acid sequence identity (*) and conservative amino acid substitutions (.) are designated below the alignment. Despite the considerable degree of amino acid sequence identitiy between HO3 and human HRS (approximately 72%), a prominent divergence was observed within the corresponding amino-terminal domains (approximately 20% identity for aa 1-60). While no discernable sequence similarity was identified between the unique amino-terminus of HO3 and any other published aaRS gene sequences, both the human HRS and predicted HO3 polypeptides had a similar degree of primary amino acid sequence identity with the corresponding yeast and bacterial HRS sequences (approximately 50% and 25%, respectively).

Fig. 5 displays amino acid sequence alignments between HO3 and HRS sequences (HRS, human; YEA, yeast; ECO, E. coli) within structural motifs conserved among and defining the class II aaRS gene family (panel A, motifs 1-3) as well as within structural domains evolutionarily conserved among histidyl-tRNA synthetases (panel B, signature regions 1 and 2)(3,5). Amino acid sequence identity (*) and conservative amino acid substitutions (.) between HO3 and human HRS are designated above each alignment. Amino acid residues that are invariably conserved among all class II structural motifs are also indicated (*). While human HRS and HO3 sequences are generally well-conserved in the aaRS class II structural motif regions, an even higher degree of structural identity was observed within HRS signature regions 1 and 2 (92% and 96% identity, respectively). In addition, the position and spacing of structural motifs and signature regions within the complete primary sequences of human HRS and HO3 were also well-conserved. For comparison, an alignment of yeast and bacterial motif and signature regions is illustrated indicating amino acid residues that are invariably conserved (♦) or conservatively substituted (▲) among HRS gene sequences.

DISCUSSION

A growing body of literature suggests that aaRS gene expression is subject to more elaborate regulatory controls than those generally associated with housekeeping genes (6-9). Such observations extend to the human HRS gene which is reportedly linked in a head-to-head configuration with a previously uncharacterized gene (13). In this report, we describe a genomic clone (C5-1) linking these genes and a novel cDNA clone (HO3) corresponding to an opposite-strand transcript mapping 5' of human HRS.

The gene encoding HO3 resides in close proximity of the human HRS gene and is oriented in the opposite direction of HRS transcription (Fig. 1A). Genomic clone C5-1 was

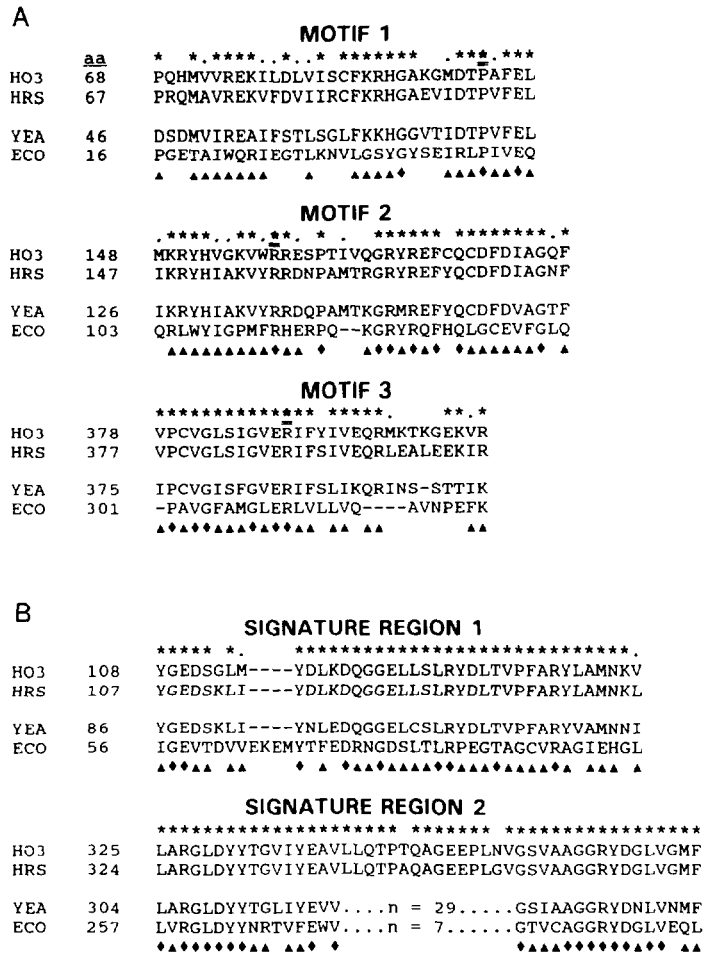

Figure 5. Alignment of class II aaRS consensus domains. Human HRS and HO3 polypeptide domains corresponding to sequences conserved among and defining class II aaRSs (panel A, motifs 1-3) and histidyl-tRNA synthetases (panel B, signature regions 1 and 2) were aligned using the PC/Gene PALIGN program. Regions of amino acid sequence identity (*) and conservative amino acid substitutions (.) are designated above the alignments. Amino acid residues invariably conserved among all class II aaRS sequence motifs are highlighted (motifs 1-3, $\underline{*}$). Alignments of corresponding HRS structural domains from yeast (YEA) and E. coli (ECO) are provided for comparison. Amino acids residues that are invariably (♦) or generally (▲) conserved among the aligned sequences are designated. The positions of the first amino acid of each structural domain within the complete sequences are designated to the left of each alignment. The boundries of HRS signature regions and class II aaRS structural motifs were based upon previous descriptions (3,5).

shown by comparative sequence analysis to span the first four exons of the HO3 cDNA (Fig. 1B, nts 1-548) in addition to HRS exons 1 and 2, thus confirming a physical link between HO3 and HRS bidirectional gene expression. The head-to-head configuration of HRS and HO3 genes within C5-1 predicts that the synthesis of each gene is initiated within a short expanse of intervening sequence information (<200 bp). Consistent with this finding, Tsui *et al.* (13) have

defined a bidirectional promoter element which maps between human HRS and the HO3 gene described here (see Fig. 1A)(GenBank accession number M96646). Moreover, these same authors have described a similar arrangement of HRS bidirectional gene transcription in rodents (17).

Bidirectional promoters, although common within the compact genomes of prokaryotes and viruses, are seemingly rare among higher eukaryotes. In mammals, other examples of genes sharing a bidirectional promoter include WIT-1/WIT-2 of the Wilms tumor locus (18), dihydrofolate reductase/mismatch repair protein 1 (19), and the genes encoding the $\alpha 1$ and $\alpha 2$ chains of type IV collagen (20,21). The human HRS promoter, aside from its bidirectional character, is typical of most housekeeping promoters in lacking canonical RNA polymerase II transcription factor binding sites. Recently, a *cis*-acting element has been identified within the HRS promoter (-92 to -120 relative to the HRS translational start codon) which binds a putatively novel class of DNA binding proteins that may potentially influence HRS bidirectional gene transcription (13).

Our RNA blot analysis indicated that the gene corresponding to the HO3 cDNA encodes a 2.5 kb mRNA (Fig. 2B). Using a genomic probe which spans HO3 exon 1, Tsui *et al.* (13) estimated the size of the human HRS opposite-strand transcript at 3.5 kb. The reliability of this estimate is questionable as it was based upon comparative migration with 28S and 18S rRNA markers, a generally less precise means of estimating mRNA size compared to the use of predefined molecular size standards. By this latter approach, we found that the estimated size of the HO3 mRNA (2.5 kb) agreed with that predicted by the HO3 cDNA described above (2.44 kb). Together, the data indicate that the HO3 cDNA corresponds to an approximately full-length mRNA. In addition, we have shown that while both HRS and HO3 mRNAs (2.0 and 2.5 kb, respectively) are constitutively expressed in multiple human tissues, their steady-state mRNA levels vary considerably among different tissues (Fig. 2). More interestingly, HRS and HO3 exhibited distinctive profiles of mRNA expression. Most notably, inverse patterns of HRS and HO3 mRNA expression were observed in brain, liver, and skeletal muscle (Fig. 2A and 2B, compare lanes 2, 5, and 6, respectively). While it is possible that both transcriptional and/or post-transcriptional mechanisms may account for these differences, the data suggest that HRS bidirectional transcription may be regulated in a tissue-specific manner.

Bidirectional transcription units often encode polypeptides sharing structural and functional characteristics (21,22). A comparison of the amino acid sequences predicted by the human HRS and HO3 cDNAs revealed a striking degree of primary sequence identity (72%) within the corresponding ORFs (Fig. 4). Regions of amino acid sequence identity/similarity included structural motifs (Fig. 5A, motifs 1-3) comprising conserved elements of the class II aaRS catalytic domain (3). Most notably, amino acid residues that are invariably conserved

among class II aaRSs are likewise conserved within HO3. Of additional interest is the near perfect conservation of human HRS and HO3 amino acid identity within HRS signature regions 1 and 2 (Fig. 5B; 92% and 96%, respectively). HRS signature regions are evolutionarily conserved among HRS genes isolated from a variety of species and are likely to contribute to substrate recognition (i.e. His and $tRNA^{His}$ binding)(3,5). Together, the comparative structural data suggest that HO3 represents an evolutionary descendent of the class II aaRS gene family, and more specifically an HRS homologue.

Currently, the function of the predicted HO3 polypeptide is unknown. The high degree of homology detected between HO3 and other HRS sequences suggests that HO3 may possess HRS enzymatic activity. Yet, despite their overall similarity, human HRS and HO3 possess considerably divergent amino-terminal (Nt) domains (Fig. 4). The Nt-domain of human HRS (aa 1-60) contains a high density of charged amino acids (23/60) located within an α-helical coiled-coil structure predicted for this region (11). In contrast, the Nt-domain of HO3 contains a preponderance of hydrophobic, hydroxylated, and interspersed basic amino acids; a sequence similar to that associated with targeting of nuclear-encoded polypeptides to mitochondria (TRANSPEP algorithm, PC/Gene)(23).

A recent report implicated the Nt-domain of human HRS (aa 1-60) as playing an essential role in enzymatic activity (11). In fact, a consensus motif identified within the Nt-domain of HRS (aa 14-45) is conserved among multiple class I and II aaRSs and is presumed to participate in tRNA binding (11). Nevertheless, it is unclear what structural elements a mitochondrial isoform of HRS might require for activity and whether the Nt-domain of HO3 may function in a similar capacity. Curiously, in yeast, both cytosolic and mitochondrial HRS are encoded by a single gene which maintains two in-frame translational start codons (24). The upstream translational start of the yeast HRS gene initiates a 20 amino acid leader sequence responsible for mitochondrial targeting. Although no recognizable sequence similarity exists between the yeast mitochondrial signal peptide and HO3, we are investigating the possibility that HO3 encodes a mitochondrial isoform of human HRS.

In the event HO3 encodes an enzymatically inactive homologue of HRS, it may conceivably retain one or more attributes associated with HRS function (e.g. ATP, His and/or $tRNA^{His}$ binding). In this capacity, HO3 might conceivably play an accessory role in the regulation of protein biosynthesis as proposed for other aaRSs and aaRS-like proteins (6-8,25).

ACKNOWLEDGMENTS

The authors thank Drs. Mary Whiteley and Steven Bauer for their critical review of the manuscript.

REFERENCES

1. Nagel, G.M. and Doolittle, R.F. (1991) Proc. Natl. Acad. Sci. U.S.A. 88, 8121-8125.
2. Carter, C.W.,Jr. (1993) Annu. Rev. Biochem. 62, 715-748.
3. Delarue, M. and Moras, D. (1993) Bioessays 15, 675-687.
4. Cusack, S., Hartlein, M., and Leberman, R. (1991) Nucleic Acids Res. 19, 3489-3498.
5. Raben, N., Borriello, F., Amiyn, J., Horwitz, R., Fraser, D., and Plotz, P. (1992) Nucleic Acids Res. 20, 1075-1081.
6. Miseta, A., Woodley, C.L., Greenberg, J.R., and Slobin, L.I. (1991) J. Biol. Chem. 266, 19158-19161.
7. Schray, B. and Knippers, R. (1991) Nucleic Acids Res. 19, 5307-5312.
8. Kisselev, L., Frolova, L., and Haenni, A.L. (1993) Trends. Biochem. Sci. 18, 263-267.
9. Kaiser, E., Hu, B., Becher, S., Eberhard, D., Schray, B., Baack, M., Hameister, H., and Knippers, R. (1994) Genomics 19, 280-290.
10. Miller, F.W., Waite, K.A., Biswas, T., and Plotz, P.H. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 9933-9937.
11. Raben, N., Nichols, R., Dohlman, J., McPhie, P., Sridhar, V., Hyde, C., Leff, R., and Plotz, P. (1994) J. Biol. Chem. 269, 24277-24283.
12. Shi, M.H., Tsui, F.W., and Rubin, L.A. (1991) J. Rheumatol. 18, 252-258.
13. Tsui, H.W., Mok, S., de Souza, L., Martin, A., and Tsui, F.W. (1993) Gene 131, 201-208.
14. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
15. Current Protocols in Molecular Biology, (1989) Green Publishing Associates and Wiley-Interscience, New York.
16. Miller, F.W., Twitty, S.A., Biswas, T., and Plotz, P.H. (1990) J. Clin. Invest. 85, 468-475.
17. Tsui, F.W. and Siminovitch, L. (1987) Gene 61, 349-361.
18. Huang, A., Campbell, C.E., Bonetta, L., McAndrews-Hill, M.S., Chilton-MacNeill, S., Coppes, M.J., Law, D.J., Feinberg, A.P., Yeger, H., and Williams, B.R. (1990) Science 250, 991-994.
19. Fujii, H., Shinya, E., and Shimada, T. (1992) FEBS Lett. 314, 33-36.
20. Burbelo, P.D., Martin, G.R., and Yamada, Y. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 9679-9682.
21. Poschl, E., Pollner, R., and Kuhn, K. (1988) EMBO J. 7, 2687-2695.
22. Lennard, A.C. and Fried, M. (1991) Mol. Cell. Biol. 11, 1281-1294.
23. Gavel, Y. and von Heijne, G. (1990) Protein Eng. 4, 33-37.
24. Natsoulis, G., Hilger, F., and Fink, G.R. (1986) Cell 46, 235-243.
25. Ramirez, M., Wek, R.C., Vazquez de Aldana, C.R., Jackson, B.M., Freeman, B., and Hinnebusch, A.G. (1992) Mol. Cell. Biol. 12, 5801-5815.