The β Subunit of Human Chorionic Gonadotropin Is Encoded by Multiple Genes*

Paul Policastro, Catherine E. Ovitt, Makoto Hoshina, Hideoki Fukuoka, Mark R. Boothby, and Irving Boime

From the Departments of Pharmacology and Obstetrics and Gynecology, Washington University School of Medicine, St. Louis, Missouri 63110

The abbreviations used are: hCG, hormone chorionic gonadotropin; LH, lutropin; FSH, follitropin; TSH, thyrotropin; EtBr, ethidium bromide; kb, kilobase.

The human placental hormone hCG1 is one of a family of glycoproteins that includes the pituitary hormones LH, FSH, and TSH. Each of these hormones consists of a noncovalent dimer of α and β subunits. Within a mammalian species, the α polypeptide chain sequence is the same for all four hormones, and in humans a single α subunit gene has been identified (1, 2). The β subunits, however, have diverged considerably in developing their separate endocrine functions, although the homology of the amino acid sequence between the hCGβ and LHβ subunits is about 80% (3). The 12 cysteine residues that form intrachain disulfide bonds are highly conserved positions in the β chain family, presumably reflecting their importance to the secondary structure required for association with the α subunit (3). The hCGβ subunit is unique in that it contains an extension of 29 amino acids at its COOH end (4). The nucleotide sequence of a cDNA clone to hCGβ mRNA suggests this extension derives from loss of an upstream stop codon, since the triplet encoding glutamine, the terminal amino acid of the extension, forms part of the polyadenylation signal (5). This signal is normally present in the 3′ nontranslated region of eukaryotic mRNAs (6).

Since endocrine function of the various cell types in which the α subunit is expressed depends on coordinate expression with a specific β subunit, these glycoprotein hormones offer an opportunity to study a group of closely related genes whose expression may require the presence of a previously activated α gene. To examine the expression, structure, and organization of the β subunit genes, we isolated and sequenced clones bearing hCGβ genes from a human genomic library and probed the hCGβ-coding regions in human cellular DNA. Our nucleotide sequence data indicate that the hCGβ subunit is encoded by at least three nonallelic genes which may be linked in a single cluster with four other hCGβ-like genes. These findings are supported by the results of Boorstein et al. (7) who identified eight hCGβ-like genes by restriction enzyme analysis of several genomic DNA isolates. The three genes we report on here differ in their 5′ flanking sequence and none of them is completely homologous in sequence to either of two hCGβ cDNA clones we have isolated. One of the sequenced genes has 3 base pair differences in the translated region resulting in changes of 2 amino acids, and the two complete gene copies show that the COOH terminal extension of hCGβ does not result from a splice event during transcription.

METHODS

Preparation and Isolation of cDNA Clones—The construction of PUNY, a cDNA clone derived from hCGβ mRNA, has been previously described (8). In a second preparation of cDNAs from the same sucrose gradient-purified mRNA, primary strand synthesis was stopped with the addition of 20 mM EDTA, 0.2% sodium dodecyl sulfate, and 100 μg/ml of Escherichia coli tRNA. The products were desalted on a Sephadex G-50–80 column; concentrated with 1-butanol; hydrolyzed at 68 °C for 30 min in 0.1 M NaOH, 10 mM EDTA; neutralized with acetic acid; and precipitated with ethanol. Reverse transcriptase was used in this case to synthesize the second strand (9), and double-stranded DNA complementary to mRNA was inserted into PstI-digested pBR322 (8). Tetracycline-resistant colonies obtained from HB101 transformations (10, 11) were screened (12) with the PUNY plasmid (and Charon 4A clone CGβα, see below) labeled with α-[32P]dATP or dCTP by nick translation (13). Plasmid DNA isolated from cleared lysates (14, 15) was centrifuged in 4.7 M CsCl
Multiple Human Chorionic Gonadotropin β Subunit Genes

RESULTS

Characterization of cDNA Probes and hCGβ mRNA—To identify genomic sequences homologous to hCGβ mRNA, we have used three different hCGβ cDNA clones in the plasmid vector pBR322 (Fig. 1). The first of these, pUNY, contains an insert of 110 nucleotides (8) encoding amino acids 13 to 48 of mature hCGβ. Two larger cDNAs were later isolated from the same pool of first trimester placental mRNA and sequenced. The first, pCG507, contains 192 nucleotides of the untranslated region and 241 nucleotides of translated sequence. The second, pCG474, encodes 151 nucleotides of the 5’ untranslated region and 429 nucleotides of translated sequence. These clones therefore contain information on the 5’ untranslated region not present in a previously analyzed cDNA, which extended 25 nucleotides upstream from the initiator methionine codon and contained the complete 498 nucleotide translated region (5). This cDNA showed that the final Gln codon CAA and the terminator codon UAA form part of the polyadenylation signal in the mRNA, which has a 3’ untranslated region of only 16 nucleotides before the polyadenylation site. Since synthesis of cDNA can terminate prematurely on the mRNA template, there is no assurance that pCG507 or pCG474 represent the complete 5’ untranslated sequence of hCGβ mRNA. In a primer extension experiment in which a 194-nucleotide DdeI fragment of the pCG507 insert was annealed to first trimester mRNA, reverse transcriptase gave a product which extended −165 nucleotides beyond the 5’ end of the cDNA insert. The 5’ untranslated portion of this hCGβ mRNA species therefore appears to be −360 nucleotides long (Fig. 1).

We expected identical 5’ untranslated sequences in the pCG474 and pCG507 clones if they were copies of a single mRNA species. In fact, the first 75 nucleotides upstream from the initiator methionine codon show one difference between the two cDNAs, at nucleotide −9 (see Fig. 5). At nucleotide −76, the sequences diverge. In the case of pCG474, nucleotides −73 to −110 seem to be an inverted repeat of the hCGβ coding sequence 17–54 (34 match out of 38) and nucleotides −111 to −151 share no homology with pCG507 or any portion of the hCGβ genes sequenced below. The inverted repeat in pCG474 contains the repeated leucine codons of the hCGβ signal sequence (21 bases) on one of its strands. While we have not excluded the possibility of nonlinear transcription during cloning, pCG507 and pCG474 cDNAs may represent copies of two distinct mRNA species.

Isolation and Southern Blot Analyses of hCGβ Genes—To identify and amplify genomic sequences encoding hCGβ subunit, a Charon 4A human genomic recombinant phage library (kindly provided by Dr. Maniatis) was probed with nick-translated pUNY cDNA. Two isolates, CGβa and CGβe, were identified and purified from 3 × 10^8 phage that were screened.

To determine the regions of CGβa and CGβe that hybridized with pUNY, the phase DNAs were digested with several restriction enzymes in single or double digests, electrophoresed in agarose gels, and blotted for hybridization with the same probe (Fig. 2). For CGβa, pUNY hybridized to a single fragment in each digest except for the EcoRI, EcoRI-BamHI, and XhoI digests. A prominent 13.6-kb fragment and a minor band of 11.3 kb are observed in EcoRI digests of CGβa. A faint XhoI fragment of 6.8 kb (below the 9.6-kb band) is also seen in the stained gel, that cannot be explained by incomplete digestion. We suspect that phase which banded at a lower density in CsCl gradients of CGβa are responsible for these fragments, since such phage are likely to contain a shorter...
DNA insert. The presence of a minor HindIII fragment of 4.5 kb is consistent with a deletion of \( \sim 2.9 \) kb in this variant, within the region marked in Fig. 3. In the EcoRI-BamHI and XbaI digests, the extra pUNY-hybridizing bands are the result of incomplete digest. The central XbaI site (brackets) is particularly resistant to cleavage for unknown reasons. The pUNY probe identifies a region of 0.6 kb between the rightward HindIII site and a PvuII site that contains all the sequence homologous to the 110-nucleotide hCGB cDNA (Fig. 3).

In contrast to this single region of homology, the CGPe insert contains two regions separated by 8 kb that hybridize to the pUNY probe. This is shown most directly by digests with EcoRI, HindIII, and PvuII (Fig. 3). The insert has no internal EcoRI sites and contains two HindIII sites that yield an 8-kb fragment not recognized by pUNY. The flanking HindIII fragments of 20.5 and 12 kb do hybridize, and each is cut by EcoRI to give hybridizing fragments of 0.75 to the left and 7.1 kb to the right of the nonhybridizing HindIII fragment. When digested with PvuII, CGPe yields pUNY-hybridizing fragments of 2.5 and 2.1 kb. Further digest with HindIII shows that the two regions of homology to the probe are of the same size (0.6 kb) as that found in HindIII-PvuII-digested CGPa. Restriction digests of both CGPa and CGPe with BglII also give common pUNY-hybridizing fragment sizes of 0.5 kb. These homologies between the three sites suggested each was a copy of the hCGB gene.

The region between the two hCGB-related sequences in CGPe contains sites for PvuII, XbaI, and XhoI, at 2.3, 2.6, and 3.2 kb from the leftward HindIII site and 1.9, 2.2, and 2.7 kb from the rightward HindIII site, respectively (Fig. 3). The XhoI positions were determined from Rf-XhoI double digests (Fig. 2), while those of the PvuII and XbaI sites relied on other restriction digests (data not shown) and on inferences from sequence analysis of the CGPe insert subcloned in pBR322 (see below). The array of these restriction sites indicate that sequences in each half are similar, inversely oriented to each other, and differentiated by a length of \( \sim 400 \) base pairs present in the leftward PvuII-HindIII region absent in the rightward region. Despite identical restriction sites to CGPa within two regions of at least 0.6 kb (HindIII, PvuII, and BglII data), the remaining restriction map of CGPe is distinct from CGPa. By probing a fraction of the human genomic library, we isolated two chromosomal segments bearing three distinct copies of hCGB-related sequences. Two of these putative genes are linked in an inverse orientation in the CGPa insert. Boorstein et al. constructed three composite maps containing eight hCGB-like genes from a series of genomic isolates (7); the first of these maps confirms the arrangement of hCGB-like sequences found in CGPe. Comparison of CGPa to their second and third maps, however, suggests the latter may overlap and share an identical gene. If so, this reduces the number of hCGB-like genes in their analysis to seven.  

**Orientation of Genes in CGPa and CGPe**—The putative hCGB gene structures in the CGPe and CGPa phage clones were characterized by sequence analysis. To amplify the genomic inserts, we subcloned their EcoRI fragments in pBR322 and subclones of the phage inserts were identified using nick-translated CGPa and pUNY. The large 13.6-kb EcoRI fragment of CGPa and the single 16-kb EcoRI insert fragment of CGPe were subjected to sequence analysis from the HindIII sites adjacent to the regions hybridizing to pUNY (Figs. 3 and 4). Although not present in the sequence of hCGB cDNA, it seemed likely that this restriction site was within either the long 5′ untranslated region, or an intervening sequence conserved between the three genes. EcoRI sites at the left and right ends of the CGPe subclone were also used as sequencing start sites. Extensive sequence homology was observed when reading from the HindIII sites toward the pUNY-hybridizing regions. None of these sequence runs were pursued into the translated regions identified by pUNY. Sequencing into the CGPe insert from the EcoRI site at its left end revealed a portion of the hCGB coding sequence beginning at nucleotide 320 and continuing through the 5′ end of an exon that begins at nucleotide 184 (Figs. 4 and 5). This established the leftward pUNY hybridization region of CGPe as an hCGB-like gene and allowed us to infer the 5′-3′ orientation of the three genes. Sequencing into the CGPe insert from the EcoRI site at its right end for 140 nucleotides showed that this region contains a small exon partially homologous to the first five codons of hCGB mRNA and continues in a 5′ direction with some homology to the untranslated regions of hCGB cDNAs (Fig. 4 and see below). This exon lies 7 nucleotides to the left of a PstI site and 85 nucleotides to the left of the insert end. Besides providing further information on the probable structure of the three genes identified by the hCGB cDNA probe, the presence of this exon increased the number of known hCGB-like gene structures to at least four, three linked in the CGPe and one in the CGPa inserts. Boorstein et al. (7) have reported that this fourth gene copy, which corresponds to hCGG4 in group a of their maps, encodes the LHβ subunit.

**Sequence Determination of the Cloned Genes**—To obtain further sequence information, we inserted PstI fragments from the pUNY-hybridizing regions of the above EcoRI subclones into pBR322. A 700-nucleotide fragment from CGPa
Multiple Human Chorionic Gonadotropin β Subunit Genes

Fig. 2. Hybridization of pUNY cDNA to restriction enzyme-digested human genomic clones. Restriction enzymes used in digests of E. coli phage λ and Charon 4A human genomic library clones CGßa (a) and CGße (e) are indicated above the lanes. The dark panels on the left are the 0.9% agarose gels containing marker and clone DNA visualized under UV light with ethidium bromide. The light panels on the right are autoradiograms of nitrocellulose filter transfers of the DNA which were hybridized to 32P-labeled nick-translated pUNY cDNA. Autoradiograms are matched to their respective gels. The left margin shows phage λ HindIII marker fragment sizes in kilobases.

Fig. 3. Restriction enzyme maps of CGßa and CGße inserts. Darkened lines refer to smallest defined regions hybridizing to pUNY cDNA. Restriction enzyme sites are shown above each insert: RI, EcoRI; Hd, HindIII; Xh, Xhol; BH, BamHI; Kp, KpnI; Xb, XbaI; Pv, PvuII. Asterisks above PvuII sites are shown to indicate analysis for this enzyme is incomplete. Only those sites defining pUNY-hybridizing regions are given. Bracketed regions above CGßa denote region of ~2.9-kb deletion (∆) in a portion of the CGßa phage; those below CGße denote size of 1 kilobase of DNA.

CGßa

CGße

Regions of the genomic insert was made, based on several base changes in the sequences reading 3'ward from the HindIII sites of the left and center gene copies present in CGße (see Fig. 5). Complete sequence determination of these three PstI inserts was pursued, and showed that they contained the second exon of the three gene copies, which extends from the sixth codon of the signal peptide to codon 41 of the mature protein, or nucleotide 16 through 183 of the mRNA. This exon was flanked in each case by appropriate splice recognition sequences (Ref. 25 and Fig. 5). At the extreme 3' end of the inserts, four codons from the next exon were detected, corresponding to the sequence next to the PstI site found in the coding region of each hCGβ cDNA. The 5' and 3' ends of hCGβ coding information therefore lie outside the PstI subclones (Fig. 4). These regions were sequenced directly from the larger EcoRI subclones, using restriction sites shown by arrows in Fig. 4. In the three complete or partially complete genes, the 15-nucleotide translated region of the first exon is separated from the 168-nucleotide second exon by an intron of 351–352 nucleotides (Fig. 5). A 234–236-nucleotide intron separates the second from the third exon. The translated
region of the latter is 315 nucleotides in length in the two complete gene copies. The coding sequences of these genes are in agreement with that of the previously reported hCGβ cDNA sequence (5), with the following exceptions (Fig. 5). Nucleotides 70 and 71 of the CGPa gene coding region are AT, whereas CC are expected from the cDNA sequence. The codon change CCG → ATG at nucleotides 70–72 indicates a Pro 4 → Met 4 substitution in the encoded protein. The CGPa gene also contains a base change at nucleotide 410 in the third exon, changing the codon GAC to GCC with the anticipated alanine change in the putative amino acid change Asp 117 in the putative gene product. The central gene of the CGPe genomic clone contains a single base change at nucleotide 438 in exon 3, changing the third base of the Pro 126 codon from C to G. Both of these full length gene copies also show a silent base change from C to T at nucleotide 448. Finally, the rightward gene in CGPe (an LHβ gene, see above) shows a T to C base change at nucleotide 10, indicating a Leu codon at this position in the signal peptide sequence. The last base in this exon was not determined.

A Possible Pseudogene—Several base changes and apparent deletions and insertions occur when the first and second introns of the sequenced genes are compared, but the only alteration that may affect transcript splicing is at the donor position in the first intron of the leftward CGPe gene. The GA, rather than GT, sequence at this site contradicts the donor consensus sequence present in 139 tabulated splice junctions (25). The 5' untranslated regions of the genes in the two inserts are compared in Fig. 5 to those of pCGβ4747 and pCGβ807. As mentioned above, these two cDNAs diverge in sequence 75 nucleotides upstream from the initiator methionine codon. Three of the four genes share this region of homology with the cDNAs, namely the CGβα gene, the central gene in CGPe, and the rightward gene in CGPe. The leftward CGβe gene shows homology to these sequences for only nine bases upstream from the translational start site, then diverges completely from the upstream cDNA sequences. The restriction map of the CGβe insert (Fig. 3) shows that the cluster of PvuII, XbaI, and XhoI sites upstream from this gene copy are ~400 nucleotides further from the pUNY-hybridizing region than those of the central CGβe gene 5' flanking region. Southern blot analysis of this region in CGβe, using pCGβ4747 and pCGβ807 as probes, will be necessary to determine if the leftward CGβe gene shows any further homology to these cDNAs in its 5' flanking region. The presence of an insertion at this site together with the splice junction mutation noted above would suggest that this copy is a pseudogene.

3' Flanking Regions and Sequence Summary—The CGβα gene and the central CGβe gene show considerable homology in their 3' flanking regions: notably, each contains an 8-fold repeat of the doublet CA, located 51–52 nucleotides from the polyadenylation signal AATAAA and 36 nucleotides from the presumptive site of polyadenosine addition to the transcribed message. Although the 16-nucleotide 3' nontranslated region of the CGβe gene agrees with the previously reported cDNA sequence (5), the CGβα gene shows three base changes and one deletion in this comparison. We suggest that transcripts of this gene add polyadenosine at an equivalent distance from the termination codon. These three genes are clearly copies of the hCGβ-encoding sequence, though the CGβα gene displays two amino acid coding substitutions, Pro 4 → Met 4 and Asp 117 → Ala 117, and the leftward CGβe gene displays a 5' translated region that bears little homology to either cDNA clone and a splicing mutation. Both complete copies of coding sequence demonstrate continuity of the third exon and the COOH-terminal-encoding region. The Asp 117 → Ala 117 substitution in the CGβα gene copy occurs at the start of the sequence that encodes this extension.

Are the hCGβ Genes Clustered?—The multiplicity of genes present in the above inserts prompted us to ask what the total number and arrangement of hCGβ genes is in the human genome. We digested high molecular weight human DNA with HindIII, EcoRI, BamHI, XbaI, and KpnI and probed these fragments with a HhaI fragment of pCGβ4747 which contains the 151-nucleotide 5' untranslated sequence contiguous with 246 nucleotides of the translated region (Fig. 1). In comparison to known hCGβ gene sequences (Fig. 5), one might expect probe recognition of primarily nontranslated sequence to the left of the HindIII site in the first intron and of coding sequence to the right of this site (second exon and part of third exon). Except HindIII, none of the enzymes used cleave within the coding regions present in CGβα or CGβe nor in the immediate flanking or intervening sequences (Fig. 3). In the genomic analysis (Fig. 6), EcoRI appears to give one fragment that hybridizes to the hCGβ cDNA. Although this band seems only slightly larger than the 23.1-kb marker fragment, the EtBr-stained gel indicates that the largest fragments of genomic DNA in these preparations migrate at this position. Even if the hybridizing EcoRI fragment is larger than ~25 kb, the method of electrophoretic separation prevents an
**Fig. 5. Nucleotide sequence of pCGB507, pCGB474, and hCGβ-related sequences in genomic clones.**

The first line is the complete untranslated region of the pCGB507 insert. The second line is the equivalent portion of the pCGB474 insert. Lines 3–6 indicate the DNA sequence in the Charon 4A human genomic library clones CGβa and CGβe. Exons are defined by the encoded hCGβ amino acid sequence underneath. Introns are defined by the shaded boxes enclosing the beginning and end of their sequences. Nucleotide numbering originates at the initiator methionine codon and is continuous in the 5' upstream direction, but is interrupted in the 3' direction by introns. Only those bases differing from the top line sequence are indicated on their respective lines. Undetermined nucleotides are denoted by hyphens; spaces (s) have been inserted to maximize sequence homologies. CGβa left sequence ends at the leftward EcoRI site (nucleotide 320) and CGβe right sequence ends at the rightward EcoRI site of the genomic insert (at nucleotide 80 of Intron I). CGβa and CGβe center are intact copies of the coding region of hCGβ. Suggested sites of polyadenosine addition to gene transcripts are shown by arrows.

**accurate determination for the length of restriction fragments above this size. BamHI also yields a hybridizing band in this range, plus a fragment of 6.9 kb. Double digestion with EcoRI and BamHI gives the same pattern as BamHI alone. These data indicate that BamHI sites are limited within the EcoRI fragment(s) containing the hCGβ genes. The CGβa insert may represent one end of a hybridizing genomic EcoRI fragment, since it contains two EcoRI sites and three BamHI sites, all to the left side of its hCGβ gene (Fig. 3). The CGβe insert is internal to a genomic EcoRI fragment, because it**
lacks sites for this enzyme. CGβε contains two BamHI sites near the right end of the insert, however, separating this LHα-like gene from the two others in this insert. This rightward gene may be part of the 6.9-kb BamHI genomic fragment or part of another >23-kb fragment. If the region of the genome carrying these target sequences is not polymorphic for BamHI and EcoRI sites, the CGβα and CGβε inserts may represent regions near opposite ends of a single hCGβ gene cluster.

**Seven hCGβ-like Genes in Genomic Digests**—The question of how many genes lie in this putative cluster is further delineated by digests with KpnI. Six different KpnI fragments were identified in the genomic blot (12.5, 9.1, 7.7, 6.5, and 6.1 kb, Fig. 6). We also mapped these sites in the CGβα and CGβε inserts (data not shown; Fig. 3). A hybridizing KpnI band of 10.4 kb is expected from the CGβα insert, which probably corresponds to the 12.5-kb genomic fragment that is cleaved by BamHI in double digest. The 2.1-kb disparity in fragment size may be related to the deletion region in CGβα discussed above. The central gene in CGβε is located in a KpnI fragment of 7.7 kb, corresponding to a genomic fragment of the same size which hybridizes more strongly than any other of the six bands. The central gene in CGβε could have greater affinity for the 5′ untranslated sequence in the probe or, alternatively, the 7.7-kb genomic band represents more than one gene equivalent. Secondary restriction with HindIII results in loss of the signal at 7.7 kb, and appearance of the 3.6-kb HindIII-KpnI coding region fragment from CGβε, not the 4.1-kb KpnI-HindIII fragment containing the 5′ nontranslated region. The pCGβ474 nontranslated region is thus not responsible for the strong KpnI 7.7-kb signal and more than one gene equivalent seems to be present in this fragment. Presuming there are two 7.7-kb fragments that contain one gene each, linkage of the seven KpnI fragments would yield a continuous length of ~57 kb for the cluster.

The number of hybridizing KpnI fragments may reflect polymorphisms in the region containing the hCGβ genes. Such polymorphism has been previously observed in the hCGα gene with EcoRI and HindIII (1). To address this point, we used a trophoblast tissue that is primarily homozygous. The genome of hydatidiform mole is of androgenetic origin (26–28). It usually develops from duplications of one of the spermatids in an enucleate egg. By this criterion, we have recently shown that hydatidiform mole is consistently homozygous at the hCGα locus. KpnI digestion of molar DNA and hybridization of the blot transfer to pCGβ474 probe gives the same pattern of six bands as normal placental DNA (data not shown). Moreover, we have seen this KpnI pattern in digests of DNAs from several individual placentae and moles. These data indicate that the genome contains at least seven nonallelic hCGβ-like genes. It should be noted that none of these genes hybridized preferentially to the 5′ untranslated region of the pCGβ474 probe. The 38-nucleotide segment of this untranslated region which resembles part of the conserved translated portion of hCGβ may be hybridizing intramolecularly in the probe, or to this conserved region in each of the gene copies, rather than to the specific, 5′ flanking region of the gene from which it is transcribed.

**DISCUSSION**

There is no reason to expect that all of the structural genes encoding the β subunits of the glycoprotein hormone family should be clustered at one locus, since even the closely related hCGβ and hLHβ chains share only 80% homology. Nevertheless, it seemed possible the pUNY cDNA probe would recognize LHβ genes, since it encodes a portion of hCGβ which is 86% homologous to the equivalent amino acid sequence of hLHβ (31 of 37 residues). This was not the case in those Charon 4A recombinant phage clones which we isolated with this probe. Instead, we found that the pUNY-hybridizing regions represented nonallelic copies of the hCGβ gene. Though the gene copy in the CGβα insert contains three nucleotide substitutions that change two amino acids of the encoded protein, no frameshift or stop codon mutations were detected in any of the coding regions that would clearly indicate the existence of pseudogenes. However, examples exist of β-thalassemias in which changes of intron or exon sequence alter the normal transcript splicing patterns, with drastic effects on stability of transcribed sequences (29, 30). Thus, even if the leftward gene of the CGβε insert is transcribed, it may produce little or no stable mRNA, considering the splicing mutation in its first intron.

Isolation of two hCGβ cDNAs with differing 5′ nontranslated sequences (pCGβ507 and pCGβ474) from a pool of first trimester cDNAs suggests that at least two hCGβ genes are transcribed to produce stable message. Previously, we observed a form of hCGβ translated from purified fractions of first trimester mRNA that differs in methionine content from the predominant precursor form of hCGβ (31). However, the inverted repeat in the 5′ region of pCGβ474 could arise from readback of the primary cDNA strand on itself. Readback does not explain base mismatch at four sites in the 3′-5′ readback sequence on itself. Readback does not explain base mismatch at four sites in the 3′-5′ readback sequence on itself. This was not the case in those Charon 4A recombinant phage clones which we isolated with this probe. Instead, we found that the pUNY-hybridizing regions represented nonallelic copies of the hCGβ gene. Though the gene copy in the CGβα insert contains three nucleotide substitutions that change two amino acids of the encoded protein, no frameshift or stop codon mutations were detected in any of the coding regions that would clearly indicate the existence of pseudogenes. However, examples exist of β-thalassemias in which changes of intron or exon sequence alter the normal transcript splicing patterns, with drastic effects on stability of transcribed sequences (29, 30). Thus, even if the leftward gene of the CGβε insert is transcribed, it may produce little or no stable mRNA, considering the splicing mutation in its first intron.

Isolation of two hCGβ cDNAs with differing 5′ nontranslated sequences (pCGβ507 and pCGβ474) from a pool of first trimester cDNAs suggests that at least two hCGβ genes are transcribed to produce stable message. Previously, we observed a form of hCGβ translated from purified fractions of first trimester mRNA that differs in methionine content from the predominant precursor form of hCGβ (31). However, the inverted repeat in the 5′ region of pCGβ474 could arise from readback of the primary cDNA strand on itself. Readback does not explain base mismatch at four sites in the 3′-5′ readback sequence on itself. Instead, we found that the pUNY-hybridizing regions represented nonallelic copies of the hCGβ gene. Though the gene copy in the CGβα insert contains three nucleotide substitutions that change two amino acids of the encoded protein, no frameshift or stop codon mutations were detected in any of the coding regions that would clearly indicate the existence of pseudogenes. However, examples exist of β-thalassemias in which changes of intron or exon sequence alter the normal transcript splicing patterns, with drastic effects on stability of transcribed sequences (29, 30). Thus, even if the leftward gene of the CGβε insert is transcribed, it may produce little or no stable mRNA, considering the splicing mutation in its first intron.
pCGβ474. We are sequencing the 5' structures of the cloned hCGβ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNAs arise from these particular hCGβ gene copies. We can already rule out the copy in CGβa as the source of either clone, since neither contains the Pro 4 → Met 4 mutation of this gene.

Another mutation found in the CGβa gene occurs near the start of the COOH-terminal extension of hCGδ. A residue 117 in the hCGβ chain is seven amino acids beyond the last conserved cysteine residue of the β chain family. The hTSHβ chain terminates at an equivalent position to this Asp residue, while hLHβ terminates one amino acid before this position and hFSHβ continues for 7 more residues. The origin of the relatively long peptide extension in hCGδ, continuing 29 amino acids past the terminus of hLHβ, remains obscure. The evolution of this COOH-terminal peptide extension probably predates the divergence of horses and primates since the hCGδ genes to determine if one or both of the cDNA
The beta subunit of human chorionic gonadotropin is encoded by multiple genes.
P Policastro, C E Ovitt, M Hoshina, H Fukuoka, M R Boothby and I Boime

*J. Biol. Chem.* 1983, 258:11492-11499.

Access the most updated version of this article at [http://www.jbc.org/content/258/19/11492](http://www.jbc.org/content/258/19/11492)

Alerts:
- When this article is cited
- When a correction for this article is posted

[Click here](http://www.jbc.org/content/258/19/11492.full.html#ref-list-1) to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/258/19/11492.full.html#ref-list-1](http://www.jbc.org/content/258/19/11492.full.html#ref-list-1)