Human Mitochondrial C1-Tetrahydrofolate Synthase

GENE STRUCTURE, TISSUE DISTRIBUTION OF THE mRNA, AND IMMUNOLOCALIZATION IN CHINESE HAMSTER OVARY CELLS*

Priya Prasannan‡§§, Schuyler Pike‡§§, Kun Peng¶, Barry Shane¶, and Dean R. Appling‡¶

From the ‡Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, University of Texas, Austin, Texas 78712-0165 and the ¶Department of Nutritional Sciences, University of California, Berkeley, California 94720

Received for publication, April 24, 2003, and in revised form, August 20, 2003
Published, JBC Papers in Press, August 22, 2003, DOI 10.1074/jbc.M304319200

C1-tetrahydrofolate (THF) synthase is a trifunctional enzyme found in eukaryotes that contains the activities 10-formyl-THF synthetase, 5,10-methyl-THF cyclohydrolase, and 5,10-methylene-THF dehydrogenase. The cytoplasmic isozyme of C1-THF synthase is well characterized in a number of mammals, including humans; but a mitochondrial isozyme has been previously identified only in the yeast Saccharomyces. Here, we report the identification and characterization of the human gene encoding a functional mitochondrial C1-THF synthase. The gene spans 236 kilobase pairs on chromosome 6 and consists of 28 exons plus one alternative exon. The gene encodes a protein of 978 amino acids, including an N-terminal mitochondrial targeting sequence. The mitochondrial isoform is 61% identical to the human cytoplasmic isoform. Expression of the gene was detected in most human tissues, but transcripts were highest in placenta, thymus, and brain. Two mRNAs were detected, a 3.6-kb transcript and a 1.1-kb transcript, and both transcripts were observed in varying ratios in each tissue. The shorter transcript results from an alternative splicing event, where exon 7 is spliced to exon 8a instead of exon 8. Exon 8a is derived from an exonized Alu sequence, sharing no homology with exon 8 of the long transcript, and encodes just 15 amino acids followed by a stop codon and a polyadenylation signal. This short transcript potentially encodes a bifunctional enzyme lacking 10-formyl-THF synthetase activity. Both transcripts initiate at the same 5′-site, 107 nucleotides upstream of the ATG start codon. The full-length (2934 bp) cDNA fused to a C-terminal V5 epitope tag was expressed in Chinese hamster ovary cells. Immunoblots of subfractionated cells revealed a 107-kDa protein only in the mitochondrial fractions of these cells, confirming the mitochondrial localization of the protein. Yeast cells expressing the full-length human cDNA exhibited elevated 10-formyl-THF synthetase activity, confirming its identification as the human mitochondrial C1-THF synthase.

C1-tetrahydrofolate (THF) synthase is a trifunctional enzyme found in eukaryotes that contains the activities 10-formyl-THF synthetase (EC 6.3.4.3), 5,10-methyl-THF cyclohydrolase (EC 3.5.4.9), and 5,10-methylene-THF dehydrogenase (EC 1.5.1.5) (Fig. 1, reactions 1–3). These activities, along with serine hydroxymethyltransferase (Fig. 1, reaction 4), are central to the interconversion of the one-carbon units carried by the biologically active form of folic acid, THF. The activated one-carbon units are used in a variety of cellular processes, including de novo purine and thymidylate synthesis, serine and glycine interconversion, methionine biosynthesis, and protein synthesis in mitochondria and chloroplasts.

In eukaryotic cells, the mitochondrial and cytosolic compartments each contain a parallel set of one-carbon unit-interconverting enzymes (1). For example, in the yeast Saccharomyces cerevisiae, mitochondrial and cytoplasmic isozymes of C1-THF synthase (encoded by the nuclear genes MIS1 and ADE3, respectively) have been purified and characterized (2, 3). Both isozymes exist as homodimers of 100-kDa subunits. Each subunit consists of a C-terminal 10-formyl-THF synthetase domain of ~70 kDa and an N-terminal bifunctional dehydrogenase/cyclohydrolase domain of ~30 kDa linked via a proteolytically sensitive connector region. This subunit size and domain structure are shared by cytoplasmic isozymes from mammalian and avian sources (4–9).

All three activities of C1-THF synthase are found in mammalian mitochondria as well (10, 11). Our studies with intact rat liver mitochondria and mitochondrial extracts demonstrated the ability of these organelles to oxidize carbon 3 of serine to formate by a folate-dependent pathway (Fig. 1, reactions 1–4) (11). However, the existence, structure, and function of the folate-interconverting activities of C1-THF synthase in mammalian mitochondria have been controversial. MacKenzie and co-workers (12, 13) characterized a bifunctional NAD-dependent 5,10-methylene-THF dehydrogenase/5,10-methenyl-THF cyclohydrolase, originally isolated from ascites tumor cells. This bifunctional enzyme lacks the large C-terminal domain catalyzing the 10-formyl-THF synthetase activity and thus is unable to produce formate. This enzyme was shown to be a nuclear encoded mitochondrial protein (14, 15), detectable only in transformed mammalian cells and embryonic or non-differentiated tissues (12). Among adult differentiated tissues, NAD-dependent 5,10-methylene-THF dehydrogenase activity is detectable only in rat adrenal tissue (16), although the mRNA encoding this enzyme is present at low levels in all tissues examined (17). MacKenzie and co-workers (18, 19) have

* This work was supported by National Institutes of Health Grants DK61428 (to D. R. A.) and DK42033 (to B. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY274130 and AY774131.

§ Both authors contributed equally to this work.

¶ To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry, University of Texas, 1 University Station A5300, Austin, TX 78712-0165. Tel.: 512-471-5842; Fax: 512-471-5849; E-mail: dappling@mail.utexas.edu.

† The abbreviations used are: THF, tetrahydrofolate; CHO, Chinese hamster ovary; nt, nucleotide(s); SOE, splice overlap extension; HMS, homogenization solution; TBS, Tris-buffered saline; RACE, rapid amplification of cDNA ends; EST, expressed sequence tag; GCS, glycine cleavage system.
argued that mammalian mitochondria lack a C1-THF synthase and that the bifunctional NAD-dependent dehydrogenase/cyclohydrolase is the mammalian homolog of the trifunctional mitochondrial enzyme.

Here we report the identification and characterization of the human gene encoding a functional mitochondrial C1-THF synthase. We show that it is expressed widely in adult human tissues and that the full-length cDNA encodes a protein that localizes to mitochondria when expressed in Chinese hamster ovary (CHO) cells. These data confirm the existence of C1-THF synthase in mammalian mitochondria, completing the folate interconverting pathway shown in Fig. 1.

EXPERIMENTAL PROCEDURES

Materials—All chemicals were of the highest available commercial quality. Difco media components were obtained from VWR (West Chester, PA). Restriction enzymes, shrimp alkaline phosphatase, calf intestinal alkaline phosphatase, and T4 DNA ligase were purchased from New England Biolabs. Tryptone, and 0.5% NaCl) plates containing 50 μg/ml ampicillin. The colonies were screened by PCR with a vector primer and a gene-specific primer, and positive plasmids were prepared using a QIAGEN miniprep kit with the DNA.pure kit (Stratagene). The repaired full-length cDNA clone, pcDNA3.1-humito, was sequenced completely, and the correct sequence was confirmed (GenBank accession number AY374130).

CHO Cell Transfection—CHO cells (1.5 × 10^6) were plated on 35-mm diameter dishes and cultured in α-minimal Eagle’s medium supplemented with 10% fetal bovine serum. Duplicate plates were then transfected with 2 μg of pcDNA3.1-humito plate using the LipofectAMINE 2000 reagent method (Invitrogen). After transfection, cells were cultured for an additional 48 h in regular medium before a 2418-containing selective medium (0.8 mg/ml) was applied. The selective medium was applied for ~1 week until antibiotic-resistant colonies developed. Resistant colonies were picked, replated, cultured, and collected.

Preparation of Cell Homogenates and Subcellular Fractions—Transfected cells were cultured in two 150-cm² T-flasks to yield 1–2 × 10^10 cells. The monolayer was rinsed with phosphate-buffered saline (4 × 5 ml) at 4 °C and then incubated with phosphate-buffered saline containing 10 mM EDTA (10 ml) at room temperature until the cells detached (5–10 min). The flasks were tapped gently to dislodge the cells, and the cells were transferred to a 50-ml plastic conical centrifuge tube. Cells were pelleted by centrifugation at 300 × g for 6 min. The supernatant was removed carefully, transferred to another centrifuge tube, and stored on ice. The pellet was resuspended in 1 ml, and the final viscous pellet (nuclear fraction) was resuspended in H3S (1 ml). The suspension of disrupted cells was collected into a 3-ml conical ground-glass Duall tissue grinder and further disrupted with four strokes of the homogenizer (23).

Nuclei and unbroken cells were sedimented by centrifugation at 900 × g for 6 min. The supernatant was removed carefully, transferred to another centrifuge tube, and stored on ice. The pellet was resuspended in H3S (1 ml) and further dispersed by four strokes in the grinder. After centrifugation at 900 × g for 6 min, the supernatant was combined with the first supernatant and stored on ice. The pellet was washed with H3S (3 × 1 ml), and the final viscous pellet (nuclear fraction) was resuspended in H3S (1 ml). The combined supernatants were centrifuged at 900 × g for 5 min, and any pellet was discarded.

2719-bp PCR products served as templates in the SOE-PCR using primers TOP05 and TOP03 and Tfl polymerase. The full-length cDNA product (2934 bp) was gel-purified and cloned into the mammalian expression vector pcDNA3.1/DV5-His-TOPO (Invitrogen) using directional TOPO cloning according to the manufacturer’s instructions. The TOPO cloning reaction was transformed into One-Shot chemically competent Escherichia coli (Invitrogen) by chemical transformation, and positive colonies were selected on YT (0.5% yeast extract, 0.8% tryptone, and 0.5% NaCl) plates containing 50 μg/ml ampicillin. The colonies were screened by PCR with a vector primer and a gene-specific primer, and positive plasmids were prepared using a QIAGEN miniprep kit. Sequence analysis revealed a base substitution in the full-length clone compared with the original cDNA and genomic sequences, presumably incorporated during the PCRs. The 7fl polymerase, which was chosen due to the high GC content of exon 1, lacks a 3′ → 5′ proofreading activity. This substitution was repaired using the QuikChange site-directed mutagenesis kit (Stratagene). The repaired full-length cDNA clone, pcDNA3.1-humito, was sequenced completely, and the correct sequence was confirmed (GenBank access number AY374130).

CHO Cell Transfection—CHO cells (1.5 × 10^6) were plated on 35-mm diameter dishes and cultured in α-minimal Eagle’s medium supplemented with 10% (v/v) fetal bovine serum. Duplicate plates were then transfected with 2 μg of pcDNA3.1-humito plate using the LipofectAMINE 2000 reagent method (Invitrogen). After transfection, cells were cultured for an additional 48 h in regular medium before a 2418-containing selective medium (0.8 mg/ml) was applied. The selective medium was applied for ~1 week until antibiotic-resistant colonies developed. Resistant colonies were picked, replated, cultured, and collected.

Preparation of Cell Homogenates and Subcellular Fractions—Transfected cells were cultured in two 150-cm² T-flasks to yield 1–2 × 10^10 cells. The monolayer was rinsed with phosphate-buffered saline (4 × 5 ml) at 4 °C and then incubated with phosphate-buffered saline containing 10 mM EDTA (10 ml) at room temperature until the cells detached (5–10 min). The flasks were tapped gently to dislodge the cells, and the cells were transferred to a 50-ml plastic conical centrifuge tube. Cells were pelleted by centrifugation at 300 × g for 5 min at room temperature, and the cell pellet was washed with 15 ml of homogenization solution (HMS; 250 mM sucrose and 1 mM EDTA (pH 6.9)) at 4 °C. The cell pellet was resuspended in H3S (2 ml) at 4 °C, transferred to a Kontes nitrogen cavitation device, and exposed to a pressure of 36 p.s.i. for 30 min at 4 °C. The suspension of disrupted cells was collected into a 3-ml conical ground-glass Duall tissue grinder and further disrupted with four strokes of the homogenizer (23).

Nuclei and unbroken cells were sedimented by centrifugation at 900 × g for 6 min. The supernatant was removed carefully, transferred to another centrifuge tube, and stored on ice. The pellet was resuspended in H3S (1 ml) and further dispersed by four strokes in the grinder. After centrifugation at 900 × g for 6 min, the supernatant was combined with the first supernatant and stored on ice. The pellet was washed with H3S (3 × 1 ml), and the final viscous pellet (nuclear fraction) was resuspended in H3S (1 ml). The combined supernatants were centrifuged at 900 × g for 5 min, and any pellet was discarded.
Human Mitochondrial C1-THF Synthase

volume of the supernatant (total postnuclear supernatant fraction) was increased to 5 ml by the addition of HEPES.

The post-nuclear supernatant was centrifuged at 10,000 × g for 15 min, and the pellet was stored on ice. The supernatant was recentrifuged at 10,000 × g for 15 min to give a final supernatant (cytosolic fraction). The second pellet was combined with the first, washed with HEMS (2 ml), and resuspended in HEMS (1 ml) to give the mitochondrial fraction. Glutamate dehydrogenase activity (24) was used as a mitochondrial marker, and lactate dehydrogenase activity (25) was used as a cytoplasmic marker.

Immunoblotting—The protein concentration of the cytosolic and mitochondrial fractions was determined using the Bradford assay (26) with bovine serum albumin as a standard. Eighty μg of cytosolic and mitochondrial protein from transfected and untransfected cells were fractionated on a 7.5% SDS-polyacrylamide gel for 50 min at 180 V. One-half of the gel was stained, and the proteins on the other half were transferred onto a nitrocellulose membrane (Midwest Scientific, Valley Park, MO) by electroblotting for 90 min at 250 mA. The membrane was then washed with distilled water (3 × 5 min each) and blocked in 2% dry milk in Tris-buffered saline (TBS; 10 mM Tris base, 0.0025% Tween 20 (TBST; 3 × 5 min each) and incubated with goat anti-mouse primary antibody (1:1000 dilution; Zymed Laboratories Inc., San Francisco, CA) for 1 h at room temperature. The membrane was then washed with TBS and TBS (2 × 5 min each) and rinsed with water before visualizing the bands. Reactions were visualized by enhanced chemiluminescence detection (ECL, Amersham Bioscience).

Expression in Yeast and Enzyme Assays—The full-length human cDNA was subcloned from pcDNA3.1-humito into the BamHI and Xho sites of the yeast expression vector pVT103U (27). In the resulting construct, pVT-humito, the entire human mitochondrial C1-THF synthase open reading frame, including the mitochondrial presequence, is expressed at 42°C (following the kit manufacturer instructions) was used to synthesize the cDNA-specific primer (GSI, 5'-CGCTCTCAGGAAGCGCTGTTCTCAGGGGACAC-3', with the Xho site underlined) complementary to nt -9 to -30 in the 5'-untranslated region. The cDNA-specific outer primer (GSO, 5'-AGCGGCAACAGGCACAGGACAC-3') was complementary to nt +93 to +73. The 5'-RACE inner primer and the cDNA-specific inner primer had BamHI and Xho sites, respectively, at their 5'-ends to facilitate cloning.

For mapping the 3'-end of the 1.1-kb transcript, first-strand cDNA was synthesized from human placental total RNA using the supplied 3'-RACE adapter. Nested sense primers specific to the cDNA were designed for use with the two nested 3'-RACE primers provided in the kit. The cDNA-specific inner primer (3'-RACE GSI, 5'-CAGTGCTCGAG- GAACCTGTTCAGAAACAAAGTGCTC-3', with the Xho site underlined) was equivalent to nt +485 to +508. The cDNA-specific outer primer (3'-RACE GSO, 5'-CGCTCTCAGGCGCTCAGTATAGCAGTGAA-3') was equivalent to nt +590 to +410. The 3'-RACE inner primer and the cDNA-specific inner primer had BamHI and Xho sites, respectively, at their 5'-ends to facilitate cloning.

PCR fragments generated in the “inner” PCRs of both 5' and 3'-RACE were gel-purified, digested with BamHI and XhoI, and ligated separately into BamHI/XhoI-digested pBluescript II KS(+)-vector (Stratagene, La Jolla, CA). The ligation reactions were transformed into chemically competent XL1-Blue cells (Stratagene). The colonies were selected on YT/ampicillin plates. Colonies were screened by PCR using T7 reverse (5'-GAATACGCTACTATAGGGCC-3') and T3 forward (5'-ATAACACCTCATAAAGGCGG-3') vector primers, and plasmids were prepared for sequence analysis. The DNA fragment has been submitted to the GenBank(TM)/EBI Data Bank under accession number AY374131.

RESULTS

cDNA Identification and Cloning—A cDNA encoding an open reading frame with high similarity to human cytoplasmic C1-THF synthase was cloned from human uterine RNA by the German Genome Project (RZPD; GenBank(TM)/EBI accession number AL117452). The homology extends the length of the proteins, suggesting that the cDNA encodes another trifunctional C1-THF synthase (Fig. 2). This cDNA encodes 917 amino acids plus 390 nt of 3'-noncoding sequence and a poly(A) tail, but lacks a start codon, suggesting that it is truncated at the 5'-end. Blasting this sequence against the Human Genome Database (NCBI Protein Database) revealed the corresponding gene on chromosome 6 at 6q25.2. This gene spans 236 kilobase pairs and encodes the entire cDNA sequence in 27 exons plus an additional 5'-exon that encodes 60 additional N-terminal amino acids. The predicted initiator codon sits within a near-perfect expanded Kozak consensus sequence (32). The first half of this N-terminal extension has the characteristics of a mitochondrial leader sequence, including the potential to form a positively charged amphipathic α-helix. Truncation of the original cDNA clone was due to the presence of a NotI site near the 3'-end of the first exon; NotI was used in the cDNA cloning procedure (20). Subsequently, the RIKEN Mouse Gene Encyclopedia Project (33) identified a full-length mouse cDNA (ID22289) that predicts a protein with 88% identity to the human protein, including the N-terminal extension (Fig. 2). The mouse cDNA lacks the NotI site that caused truncation of the human cDNA. These data suggest that the gene on human chromosome 6 encodes a mitochondrial C1-THF synthase.

Attempts to construct a full-length cDNA by RACE using human uterine RNA were unsuccessful, probably due to the extremely high GC content (>80%) of the first exon. Instead, a

2 Details are available upon request from the corresponding author.
genomic P1 artificial chromosome clone (dJ44A20, Sanger Centre) was used to PCR-amplify the 5′/H11032-exon. This was then spliced to the remaining cDNA by SOE-PCR to construct a full-length cDNA encoding the human protein (GenBankTM/EBI accession number AY374130).

**CHO Cell Expression and Subcellular Localization**—To determine whether the protein encoded by this cDNA is, in fact, mitochondrial, we expressed the cDNA in CHO cells. The full-length cDNA was cloned into the mammalian expression vector pcDNA3.1D/V5-His-TOPO. This construct fused the 14-amino acid V5 epitope and a His6 tag to the C terminus of the 2934-bp coding region. Expression of the insert in mammalian cells is driven by the cytomegalovirus promoter. The resulting plasmid, pcDNA3.1-humito, was transfected into CHO cells, and G418-resistant colonies were selected and grown. The cytosolic and mitochondrial fractions from transfected and untransfected (control) CHO cells were isolated as described under "Experimental Procedures." Each fraction was assayed for the mitochondrial marker enzyme glutamate dehydrogenase and the cytoplasmic marker enzyme lactate dehydrogenase. Glutamate dehydrogenase activity ranged from 68 to 95 μmol/min/mg of protein in the mitochondrial fractions, compared with 2.4–4 μmol/min/mg of protein in the cytoplasmic fractions. The lactate dehydrogenase activity of the mitochondrial fraction was only one-seventh that of the cytoplasmic fraction. These results confirm that this cDNA encodes a protein that localizes exclusively to mitochondria in a mammalian cell line.

**Expression in Yeast**—The full-length human mitochondrial C_{1-THF} synthase cDNA, including the 62-codon N-terminal extension, was subcloned into a yeast expression vector (pVT103U) and transformed into an ade3 deletion strain (DAY3). Disruption of the ADE3 gene, which encodes the cytoplasmic C1-THF synthase, results in yeast cells with very low 10-formyl-THF synthetase and 5,10-methylene-THF dehydrogenase activities; the residual activity is due to the mitochondrial isozyme (34). DAY3 cells transformed with pVT-humito overexpressed 10-formyl-THF synthetase activity 9-fold compared with cells transformed with empty vector (64.6 versus 7.1 milliunits/mg of protein). However, we did not detect any increase in 5,10-methylene-THF dehydrogenase activity in cells carrying the pVT-humito plasmid using either NADP or NAD as cofactor. The third activity of C1-THF synthase, 5,10-methenyl-THF cyclohydrolase, was not assayed because it is difficult to accurately measure this activity in crude extracts. These results, together with the mitochondrial localization data above, confirm that this cDNA encodes a protein with 10-formyl-THF synthetase activity, further supporting its identification as the human mitochondrial C_{1-THF} synthase.

**Gene Structure**—The human gene encoding C_{1-THF} synthase spans 236 kilobase pairs on chromosome 6 (Fig. 4). The coding sequence consists of 28 exons and is interrupted by 27 introns ranging from 89 to 55,350 bp in length. The start codon is present in the first exon, and the 5′-end of exon 1 extends 107 bp upstream of the ATG start codon (see "Transcript Mapping" below). The stop codon is present in exon 27, and exon 28 encodes 360 nt of 3′-untranslated region, including a polyadenylation signal (AATAAA). Exon 1 is very GC-rich (>80% GC), containing a CpG island and a NotI restriction enzyme.
site (GGGGCCGC). The existence of this NotI site prevented the cloning of a full-length cDNA because NotI linkers were used in the cloning procedure (20). All of the intron/exon splice sites follow the GT/AG rule (35), except after the terminal exons, 8a and 28 (Table 1). A scan of the 5'-flanking sequences by the TESS web server3 using the TRANSFAC Version 4.0 Database predicts numerous potential transcription factor-binding sites, including Sp1, retinoic acid receptor-α, and CAAT/enhancer-binding protein-α. The 5'-flanking sequence contains a TATAAA sequence at position −985.

Northern Analysis—A Northern blot membrane prebound with human poly(A) RNA from several tissues was obtained from Ambion Inc. A 304-bp 5'-end probe spanning nt 215−518 of the human mitochondrial C₅-THF synthase cDNA revealed two bands: one at −3.6 kb and the other at −1.1 kb (Fig. 5A). The upper band corresponded to the expected size of the full-length transcript. To ensure that the 1.1-kb band was not an artifact, we washed the membrane for an additional 30 min with high stringency wash buffer at 50 °C. The additional wash did not eliminate either band. The upper and lower band distributions were very similar, with the highest transcript levels being in placenta, thymus, and brain. Expression was low in liver and skeletal muscle and barely detectable in heart.

To determine the relationship of the 3.6- and 1.1-kb transcripts, a 465-bp probe was synthesized that ended just before the stop codon. This 3'-probe detected only the 3.6-kb transcript (Fig. 5B), suggesting that the 1.1-kb transcript represents just the 5'-end of the cDNA.

We also compared the tissue distribution of the mitochondrial C₅-THF synthase transcript with that of the cytoplasmic isozyme. Using a 230-bp probe from the 3'-end of the cytoplasmic C₅-THF synthase cDNA (21), a 3.3-kb transcript was observed (Fig. 5C). The tissue distribution of this transcript differed from that of the mitochondrial isozyme, being highest in liver, kidney, and skeletal muscle. Thus, the human mitochondrial and cytoplasmic C₅-THF synthase isozymes are encoded by distinct transcripts that do not cross-hybridize under these probe and wash conditions.

Transcript Mapping—A 5'-RACE experiment was done to determine the transcriptional start site(s). 5'-RACE was performed as described under “Experimental Procedures” using 10 μg of human placential total RNA for first-strand cDNA synthesis by reverse transcription. This was followed by a first round of PCR (outer PCR), which gave no detectable specific product. Two μl of the outer PCR product were used in a second round of PCR with nested primers (inner PCR), yielding a specific product of <300 bp. The final PCR product was gel-purified and cloned. Nine colonies were screened by PCR, and all of them gave a product of between 220 and 298 bp. Three of the nine clones were sequenced, and all of them exhibited the same 5'-end 107 bp upstream of the ATG start codon (Fig. 6). These results suggest that the majority of the transcripts from this gene initiate at or near position −107, and it appears that both the 3.6- and 1.1-kb transcripts initiate from this site.

Alternative Splicing—A 3'-RACE experiment was performed to determine the 3'-end of the short 1.1-kb transcript observed on Northern blots (Fig. 5A). One μg of human placential total RNA was used for first-strand cDNA synthesis. This was followed by a first round of PCR (outer PCR), which gave no detectable specific RACE product. One μl of the outer PCR product was used in a second round of PCR with nested primers (inner PCR). Four distinct PCR products of 500, 350, 200, and 100 bp were detectable on a 2% agarose gel. (A smear at the top of the gel was also observed, produced from the full-length transcript.) Based on the 1.1-kb length of the short transcript and the position of the inner primer, the 500- and 350-bp RACE products were gel-purified and cloned separately. Six of the clones were sequenced to determine the 3'-extent of the clones. All of these clones represented the short transcript, in which exon 7 is spliced to a previously unrecognized exon, termed exon 8a, which sits in the intron between exons 7 and 8 (Fig. 7A). Exon 8a appears to be 139 bp long, although in one clone, the 3'-end extended 162 bp. It contains a stop codon after 45 nucleotides and a polyadenylation signal near its 3'-end. Thus, the 3.6- and 1.1-kb transcripts share the first seven exons and then diverge at exon 8/8a. The 1.1-kb transcript would be translated into a 275-amino acid protein in which the first 260 amino acids are identical to the full-length protein, followed by 15 unrelated amino acids (GenBank™/EBI accession number A574131) (Fig. 7, B and C).

Additional variation was observed upon sequencing the 3'-RACE clones. Several of the clones contained an extra exon at position +643, at the junction between exons 6 and 7 (Fig. 8). This extra valine codon appears to arise from variation in the 3'-splice acceptor site during the splicing of exon 6 to exon 7. The 5'-splice site has the GT consensus sequence as the first 2 nt of the intron. The 3'-splice site has two AG consensus dinucleotides at the 3'-end of the intron. If the first AG dinucleotide is used, exon 7 contains 3 additional nt; if the second is used, these 3 nt are not present in exon 7.

DISCUSSION

The experiments described here confirm that humans express a mitochondrial C₅-THF synthase, with properties very similar to those of the cytoplasmic homologs previously characterized. The full-length human cDNA encodes a protein of 978 amino acids, including an N-terminal mitochondrial targeting sequence. When the full-length cDNA was expressed in CHO cells, the targeting sequence directed the protein exclusively to mitochondria (Fig. 3). Alignment of the deduced amino acid sequence with the human cytoplasmic C₅-THF synthase (935 residues) reveals a 62-residue N-terminal extension in the putative mitochondrial protein (Fig. 2). PSORT II analysis predicts a mitochondrial targeting sequence with a cleavage site between residues 31 and 32. The next 31 residues, before

---

3 Available at www.cbil.upenn.edu/tess.

4 Available at psort.nibb.ac.jp/form2.html.
alignment with the cytoplasmic protein begins, include an unusual run of 9 consecutive glycines and several basic residues. A very similar N-terminal extension is predicted for the mouse protein (Fig. 2). Excluding this N-terminal extension, homology to the human cytoplasmic C1-THF synthase is quite high (61% identity), and the putative mitochondrial protein appears to possess the same domain structure. In the cytoplasmic protein, the N-terminal dehydrogenase/cyclohydrolase domain is 300 residues, and the C-terminal synthetase domain is 700 residues (9). The two human proteins share 31% identity in the dehydrogenase/cyclohydrolase domains and 73% identity in the synthetase domains, including conserved active-site residues and the 10-formyl-THF-binding site in the synthetase domain (31). However, the putative mitochondrial proteins from human and mouse lack 12 amino acids near the junction between the two domains (position 318) (Fig. 2).

**Table 1**

| Exon no. | Exon size | Intron sequence | Exon sequence 5'-junction 3'-junction | Introne size | Intron no. |
|----------|-----------|-----------------|----------------------------------------|--------------|------------|
| 1        | 356  b    | tcaatag          | AGAAGT-ATCCAG gtaagctgc               | 9860         | 1          |
| 2        | 85        | cttctctag        | GCCAGG-GAGGAG gtaagctgc               | 1460         | 2          |
| 3        | 51        | cttttttctag      | GTCTGT-GCCGAG gtaataaag               | 4936         | 3          |
| 4        | 54        | cttttttctag      | ATTAAAT-GAGGAG gtaagaaaaa             | 2748         | 4          |
| 5        | 70        | cttttttctag      | AATACCA-ATTACG gtaagctgc              | 2108         | 5          |
| 6        | 101       | cttttttctag      | GTTGCA-AGCAAG gtaaatcccc              | 10,879       | 7          |
| 7        | 137       | cttttttctag      | AGGAGG-GAGGAG gtaagcccc              | 6652         | 8a         |
| 8        | 139       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 13,286       | 9          |
| 9        | 92        | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 3537         | 9          |
| 10       | 98        | cttttttctag      | AAACAG-CCCAAG gtaagcccc              | 3820         | 10         |
| 11       | 174       | cttttttctag      | TGAACAT-GTGCTG gtaagcccc              | 10,509       | 11         |
| 12       | 137       | cttttttctag      | GATCAG-GTGAGA gtaagcccc              | 1759         | 12         |
| 13       | 47        | cttttttctag      | GCCGAG-GAGGAG gtaagcccc              | 5741         | 13         |
| 14       | 108       | tttttttctag      | TACACG-GATAAG gtaagcccc              | 874          | 14         |
| 15       | 75        | tttttttctag      | GTCTCG-CTAAAA gtaagcccc              | 3490         | 15         |
| 16       | 103       | tttttttctag      | AAAATG-AGAGAG gtaagcccc              | 6862         | 16         |
| 17       | 77        | cttttttctag      | TATTTG-GGGAGA gtaagcccc              | 4204         | 17         |
| 18       | 141       | tttttttctag      | GGGAGG-GGAGG gtaagcccc              | 4544         | 18         |
| 19       | 69        | cttttttctag      | GCAGAG-GGAGG gtaagcccc              | 6910         | 19         |
| 20       | 112       | tttttttctag      | GGGACA-TTGTAG gtaagcccc              | 37,900       | 20         |
| 21       | 174       | cttttttctag      | TACACG-GGAGA gtaagcccc              | 3536         | 21         |
| 22       | 47        | cttttttctag      | GCCGAG-GAGGAG gtaagcccc              | 536          | 22         |
| 23       | 139       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 536          | 23         |
| 24       | 108       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 536          | 24         |
| 25       | 108       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 536          | 25         |
| 26       | 153       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 536          | 26         |
| 27       | 121       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 536          | 27         |
| 28       | 165       | tttttttctag      | GTGAGG-GAGGAG gtaagcccc              | 536          | 28         |

* Introns are numbered starting with intron 1 between exons 1 and 2.
* The 5'-end of exon 1, and therefore its length, is based on the longest 5'-RACE clone isolated (see “Results”).
* In one 3'-RACE clone, exon 8a extended an additional 23 bp, for a total length of 162 bp.

**Fig. 4.** Intron/exon structure of the human mitochondrial C1-THF synthase gene. Exons are shown as numbered black bars, and introns as thin horizontal lines. Exon and intron sizes and positions are drawn roughly to scale, with the exception of intron 26, which is 55,350 bp. The entire gene spans 236 kilobase pairs. The actual sizes of each exon and intron are listed in Table I.

**Fig. 5.** Northern blot analysis of mitochondrial C1-THF synthase transcripts in adult human tissues. A human multiple-tissue RNA blot was hybridized with 32P-labeled probes to the 5'-end (A) or 3'-end (B) of human mitochondrial C1-THF synthase cDNA. In C, the membrane was hybridized with a probe from the 3'-end of the human cytoplasmic C1-THF synthase cDNA. The lanes in each panel contain RNA from (left to right) brain (b), placenta (p), skeletal muscle (s), heart (h), kidney (k), pancreas (p), liver (l), lung (lu), spleen (s), and thymus (t). The schematic diagram below shows the relative locations of the probes used for A and B on the 3.6- and 1.1-kb transcripts.

residues, and the C-terminal synthetase domain is – 700 residues (9). The two human proteins share 31% identity in the dehydrogenase/cyclohydrolase domains and 73% identity in the synthetase domains, including conserved active-site residues and the 10-formyl-THF-binding site in the synthetase domain (31). However, the putative mitochondrial proteins from human and mouse lack 12 amino acids near the junction between the two domains (position 318) (Fig. 2).
Expression of the full-length cDNA in yeast revealed elevated 10-formyl-THF synthetase activity, further supporting its identification as the human mitochondrial C1-THF synthase. We were unable to detect increased 5,10-methylene-THF dehydrogenase activity in these cells using either NADP⁺ or NAD⁺ as cofactor. Is the human mitochondrial enzyme multifunctional like its yeast counterpart? Given the low identity between the human cytoplasmic and mitochondrial isozymes in the dehydrogenase/cyclohydrolase domain (31%), it is conceivable that the mitochondrial protein has lost these activities. However, other members of this family have diverged as much as 30% between the human cytoplasmic and mitochondrial isozymes in the NADP-binding site and a critical element that produce a functional 3'-splice site (41). The exonization and alternative splicing of this Alu sequence in the human mitochondrial C1-THF synthase gene are apparently due to accumulated mutations in the Alu element that produce a functional 3'-splice site (41).

Assuming the short transcript is translated in vivo, it is unlikely that the resulting protein would retain 5,10-methylene-THF dehydrogenase or 5,10-methyl-THF cyclohydrolase activity. Modeling the human mitochondrial protein sequence onto the x-ray structure of the dehydrogenase/cyclohydrolase domain of the human cytoplasmic C1-THF synthase (42) reveals that exons 8 and 9, which are missing in the short transcript, encode the major portion of the Rossman fold of the NADP-binding site and a critical α-helix that forms one wall of the folate-binding site. It is likely that a truncated protein lacking these structural elements would not fold into a stable structure and would be rapidly degraded. However, without knowing how the 15 novel amino acids affect the structure, it remains possible that a stable protein with altered function could be produced. Experiments are underway to determine whether a truncated form of the protein is expressed in vivo.

Using RNA from human placenta, a single 5'-transcriptional start site at position -107 was identified by 5'-RACE (Fig. 6). It appears that both the 3.6- and 1.1-kb transcripts initiate from this site because only a single 5'-end was identified. A BLAST search of the Human EST Database with the 5'-end of the human cDNA revealed >100 entries. Four ESTs extended beyond position -107 (Fig. 6). BG481636 (position -276) and BE735249 (position -268) were isolated from a lymphoma cell line. Thus, it appears there may be some heterogeneity in the 5'-transcrip-
A. Gene structure:

B. Potential Protein products:

C. acggagtcgctctgtcaccaggccggagtgcagcgcgtgatctag

T E S R S V T R L E C R R V I *

Fig. 7. Alternative splicing of the human mitochondrial C1-THF synthase transcript. A, gene structure and splicing pattern. The alternative exon 8a is in the intron between exons 7 and 8. In the 4.6-kb transcript, exon 7 is spliced to exon 8a. Exon 8a contains a stop codon (black dot) after 15 sense codons and contains a polyadenylation signal (AATAAA) near its 3'-end. There is no homology between exons 8a and 8. B, potential protein products. The long transcript is translated into a 978-amino acid protein, whereas the short transcript is translated into a 579-amino acid protein. The amino acid sequences of the two proteins are identical through the first seven exons, encoding 260 amino acids. The short protein has 15 amino acids from exon 8a in place of the amino acids encoded by exon 8. The junction between the dehydrogenase/cyclohydrolase (D/C) and synthetase (SYN) domains is predicted to lie within amino acids 330–350 (Fig. 2). The asterisks indicate variable 3'-splice site selection at the exon 6/7 junction (see Fig. 8). C, nucleotide and amino acid sequences of the coding sequence of exon 8a. Sequences 3' to the stop codon are not shown.

Fig. 8. Variable 3'-splice site selection at the exon 6/7 junction. Upper, intron/exon junctions for the 3'-end of exon 6 and the 5'-end of exon 7. The first 2 nt of the intron and the alternative AG splice acceptors sites are underlined. Lower, alternative splicing products if the first (left) or second (right) AG acceptor site is used. The amino acid sequence encoded by each product is shown below the nucleotide sequence, and the extra codon and amino acid are in boldface.

One additional splicing variation was discovered. Some transcripts contained an extra codon at position +643, at the junction between exons 6 and 7 (Fig. 8). This valine codon appears to arise from alternative usage of AG splice acceptor sites separated by 1 nt. This type of variation in splice site selection has been seen in several mammalian genes, including human prothymosin-α (43) and the rat transforming growth factor-β type I receptor (44). The extra codon was observed in most of the 3'-RACE clones we sequenced and can be found in numerous human ESTs that represent the full-length transcript. There is no evidence to suggest that this alternative splice site selection is a regulated process. It may simply be due to “sloppiness” in the splicing mechanism when two AG splice acceptor sites fall so closely together.

Based on the x-ray structure of the dehydrogenase/cyclohydrolase domain of the human cytoplasmic C1-THF synthase (42), the extra valine is predicted to reside on the exposed loop between α-helix D2 and β-strand e. This loop is not part of the dehydrogenase/cyclohydrolase active site or the dimerization interface for this domain. It is thus possible that an extra valine at this position could be tolerated without affecting stability or activity of the protein. On the other hand, we do not know how the dehydrogenase/cyclohydrolase and synthetase domains interact, so it will be necessary to express the protein containing the extra valine to determine its effect.

The tissue distribution of the mitochondrial C1-THF synthase is quite different from that of the cytoplasmic isozyme (Fig. 5). Whereas the cytoplasmic transcript is most abundant in liver and kidney, the transcripts for the mitochondrial isozyme are relatively low in those tissues, but highest in placenta, followed by thymus, spleen, brain, and lung. The low expression of the mitochondrial isozyme in liver probably contributed to our earlier difficulties in purifying the protein from liver mitochondria. Although the ratio of the two transcripts varies somewhat from tissue to tissue, both are present in every tissue assayed, even heart (Fig. 5). The short transcript is significantly reduced in brain. Future work will be directed toward understanding the metabolic role of the mitochondrial isozyme and how that role relates to the observed tissue distribution.

The discovery of the human gene for this mitochondrial C1-THF synthase confirms our model for the compartmentation of folate-mediated one-carbon metabolism in mammalian cells (Fig. 1). Based on the well documented existence of a mitochondrial C1-THF synthase in yeast (3, 45), we proposed that mammalian mitochondria also contain this bifunctional enzyme (10). All three activities of C1-THF synthase are found in mammalian mitochondria (10). More importantly, intact rat liver mitochondria and mitochondrial extracts were shown to oxidize carbon 3 of serine to formate by the folate-dependent pathway outlined in Fig. 1 (mitochondrial reactions 1–4) (11). However, all our attempts to purify these activities from rat liver mitochondria were unsuccessful.

During this same time period, MacKenzie and co-workers (12, 13) characterized a mammalian bifunctional NAD-depend-
cytoplasmic reactions to the 5-methyl-THF required for homocysteine remethylation. Consistent with this explanation is the observation that GCS activity is stimulated by glucagon (53), and glucagon lowers plasma homocysteine in rats (54), presumably by increasing the mitochondrial production of formate. An elegant stable isotope study in humans (55) provides further support for the role of mitochondrial one-carbon units in the remethylation of homocysteine. Gregory et al. (55) showed that both cytoplasmic and mitochondrial one-carbon units end up in the methyl group of methionine following infusion of deuterated serine. This result strongly supports mitochondrial for- 
mate production as a significant contributor to cytoplasmic one-carbon units in vivo in mammals and places the mitochondrial C\textsubscript{5}-THF synthase in the center of this pathway.

REFERENCE

1. Appling, D. R. (1991) FASEB J. 5, 2645–2651
2. Paukert, J. L., Williams, G. R., and Rabinowitz, J. C. (1977) Biochem. Biophys. Res. Commun. 77, 147–154
3. Shannon, K. W., and Rabinowitz, J. C. (1986) J. Biol. Chem. 261, 12296–12271
4. Paukert, J. L., Straus, L. D. A., and Rabinowitz, J. C. (1976) J. Biol. Chem. 251, 5104–5111
5. Tan, L. U. L., and MacKenzie, R. E. (1979) Can. J. Biochem. 57, 806–812
6. Smith, G. K., Mueller, W. T., Wasserman, G. F., Taylor, W. D., and Benkovic, S. J. (1980) Biochemistry 19, 4313–4321
7. Villar, E., Schuster, B., Peterson, D., and Schirch, V. (1985) J. Biol. Chem. 260, 2245–2252
8. Cheek, W. D., and Appling, D. R. (1989) Arch. Biochem. Biophys. 276, 504–512
9. Hum, D. W., and MacKenzie, R. E. (1991) Proc. Natl. Acad. Sci. U.S.A. 88, 493–500
10. Barlowe, C. K., and Appling, D. R. (1988) Bioessays 1, 171–176
11. Garcia-Martinez, L. F., and Appling, D. R. (1993) Biochemistry 32, 4671–4676
12. Mejia, N. R., and MacKenzie, R. E. (1985) J. Biol. Chem. 260, 14616–14620
13. Mejia, N. R., Ross-Orlando, E. M., and MacKenzie, R. E. (1986) J. Biol. Chem. 261, 9509–9513
14. Mejia, N. R., and MacKenzie, R. E. (1988) Biochem. Biophys. Res. Commun. 155, 1–6
15. Belanger, C., and MacKenzie, R. E. (1989) J. Biol. Chem. 264, 4837–4843
16. Smith, G. K., Banks, S. D., Monaco, T. J., Rigual, R., Dach, D. S., Mullin, R. J., and Huber, B. E. (1989) Arch. Biochem. Biophys. 283, 367–371
17. Peri, K. G., and MacKenzie, R. E. (1993) Biochim. Biophys. Acta 1171, 281–287
18. Yang, X.-M., and MacKenzie, R. E. (1993) Biochemistry 32, 11118–11123
19. Di Pietro, E., Sirio, J., Trebouy, M. L., and MacKenzie, R. E. (2002) Mol. Cell. Biol. 22, 4158–4166
20. Wiemann, S., Weil, S., Beilnreuter, R., Gassenhuber, J., Gläsle, S., Ansorge, W., Bocher, M., Blocher, H., Barorsche, S., Blum, H., Laufer, J., Duester- 
haft, A., Beyer, A., Körner, K., Straack, N., Miewes, H. W., Ottenwalder, B., Obermaier, B., Tampe, J., Heubner, D., Wambutt, R., Korn, B., Klein, M., and Fouska, A. (2001) Genome Res. 11, 422–435
21. Hum, D. W., Bell, A. W., Rosen, R., and MacKenzie, R. E. (1988) J. Biol. Chem. 263, 15946–15950
22. Horton, R. M., Ho, S. N., Pullen, J. K., Hunt, H. D., Cai, Z., and Pease, L. R. (1993) Methods Enzymol. 219, 470–479
23. Lin, B.-F., Huang, R.-S. F., and Shane, B. (1993) J. Biol. Chem. 268, 21674–21679
24. Schmidt, E. (1974) in Methods of Enzymatic Analysis (Bergmeyer, H. U., ed) 2nd Ed., Vol. 2, pp. 650–656, Academic Press, Inc., New York
25. Kernberg, A. (1955) Methods Enzymol. 1, 441–443
26. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
27. Verret, T., Dignard, D., and Thomas, D. Y. (1987) Gene (Amst.) 52, 225–233
28. West, M. G., Horne, D. W., and Appling, D. R. (1996) Biochemistry 35, 3122–3132
29. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) J. Bacteriol. 153, 163–168
30. West, M. G., Barlowe, C. K., and Appling, D. R. (1993) J. Biol. Chem. 268, 153–169
31. Kricksey, J. T., and Appling, D. R. (1996) Arch. Biochem. Biophys. 333, 251–259
32. Rozak, M. (1987) J. Mol. Biol. 196, 947–950
33. Koyano, J., Shimazawa, K., Shibata, K., Yoshino, M., Ishii, M., Ishii, Y., Arai, T., Hara, A., Fukunishi, Y., Kanno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Nishi, K., et al. (2001) Nature 409, 685–690
34. Barlowe, C. K., and Appling, D. R. (1990) Mol. Cell. Biol. 10, 5679–5687
35. Bohnsach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383
36. Patel, H., Christensen, K. E., Mejia, N. J., and MacKenzie, R. E. (2002) Arch. Biochem. Biophys. 403, 145–148
37. Howard, K. M., Muga, S., Zheng, X., Thigpen, A. E., and Appling, D. R. (2003) Gene (Amst.), in press
38. Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agar- 
wala, P., Agarwala, R., Ainscough, R. L., Akanni, W., Am, P., Antonar, 
sis, K. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., et al. (2002) Nature 420, 520–562
39. Chang, D.-J., Chang, Y., K. Yamashita, S. S., Salazar, F. H., Kosaka, A. H., Khare, R., Bhakti, S. J., Shieh, I. S., Nesnick, J. D., Ford, A. P., Daniels, D. V., Egle, R. M., Clarke, D. E., Bach, C., and Chan, H. W. (1998) PERS LETT 423, 273–278
40. Soroker, A., Ast, G., and Grazer, D. (2002) Genome Res. 12, 1060–1067
41. Lev-Maor, G., Soroker, R., Shomron, N., and Ast, G. (2003) Science 300, 1288–1291
42. Schmidt, A., Wu, H., MacKenzie, R. E., Chen, V. J., Blevy, J. R., Ray, J. E., et al. (1995) J. Biol. Chem. 270, 1671–1678
Human Mitochondrial C$_7$THF Synthase

43187

Toth, J. E., and Cygler, M. (2000) *Biochemistry* **39**, 6325–6335

43. Manrow, R. E., and Berger, S. L. (1993) *J. Mol. Biol.* **234**, 281–288

44. Agrotis, A., Condron, M., and Bobik, A. (2000) *FEBS Lett.* **467**, 128–132

45. Shannon, K. W., and Rabindran, J. C. (1988) *J. Biol. Chem.* **263**, 7717–7725

46. Galper, J. B., and Darnell, J. E. (1969) *Biochem. Biophys. Res. Commun.* **34**, 205–214

47. Bianchetti, R., Lucchini, G., Crosti, P., and Tortora, P. (1977) *J. Biol. Chem.* **232**, 2519–2523

48. Lewis, K. F., Randolph, V. M., Nemeth, E., and Frisell, W. R. (1978) *Arch. Biochem. Biophys.* **185**, 443–449

49. Hamosh, A., Johnston, M. V., and Valle, D. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scrivener, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) 7th Ed., Vol. I, pp. 1337–1348, McGraw-Hill Book Co., New York

50. Van Hove, J. L., Lazeyras, F., Zeisel, S. H., Bottiglieri, T., Hylland, K., Charles, H. C., Gray, L., Jaeken, J., and Kahler, S. G. (1998) *J. Inherit. Metab. Dis.* **21**, 799–811

51. Randak, C., Rosschinger, W., Rolinski, B., Hadorn, H. B., Applegarth, D. A., and Roscher, A. A. (2000) *J. Inherit. Metab. Dis.* **23**, 520–522

52. Pasternack, L. B., Laude, D. A., Jr., and Applinging, D. R. (1994) *Biochemistry* **33**, 74–82

53. Jois, M., Hall, B., Fewer, K., and Brosnan, J. T. (1989) *J. Biol. Chem.* **264**, 3347–3351

54. House, J. D., Jacobs, R. L., Stead, L. M., Brosnan, M. E., and Brosnan, J. T. (1999) *Adv. Enzyme Regul.* **39**, 69–91

55. Gregory, J. F., III, Cuskelly, G. J., Shane, B., Toth, J. P., Baumgartner, T. G., and Starkpoole, P. W. (2000) *Am. J. Clin. Nutr.* **72**, 1535–1541

56. Corpet, F. (1988) *Nucleic Acids Res.* **16**, 10881–10890