## Isolation of a cDNA encoding human holocarboxylase synthetase by functional complementation of a biotin auxotroph of *Escherichia coli*

ALFONSO LEÓN-DEL-RIO, DANIEL LECLERC, BEVERLY AKERMAN, NOBUAKI WAKAMATSU, AND ROY A. GRAVEL\*

McGill University-Montreal Children's Hospital Research Institute and Departments of Biology, Human Genetics and Pediatrics, McGill University, Montreal, QC, Canada H3H 1P3

Communicated by Allan M. Campbell, Stanford University, Stanford, CA, December 27, 1994 (received for review July 11, 1994)

ABSTRACT Holocarboxylase synthetase (HCS) catalyzes the biotinylation of the four biotin-dependent carboxylases in human cells. Patients with HCS deficiency lack activity of all four carboxylases, indicating that a single HCS is targeted to the mitochondria and cytoplasm. We isolated 21 human HCS cDNA clones, in four size classes of 2.0-4.0 kb, by complementation of an Escherichia coli birA mutant defective in biotin ligase. Expression of the cDNA clones promoted biotinylation of the bacterial biotinyl carboxyl carrier protein as well as a carboxyl-terminal fragment of the  $\alpha$  subunit of human propionyl-CoA carboxylase expressed from a plasmid. The open reading frame encodes a predicted protein of 726 aa and  $M_r$ 80,759. Northern blot analysis revealed the presence of a 5.8-kb major species and 4.0-, 4.5-, and 8.5-kb minor species of poly(A)<sup>+</sup> RNA in human tissues. Human HCS shows specific regions of homology with the BirA protein of E. coli and the presumptive biotin ligase of Paracoccus denitrificans. Several forms of HCS mRNA are generated by alternative splicing, and as a result, two mRNA molecules bear different putative translation initiation sites. A sequence upstream of the first translation initiation site encodes a peptide structurally similar to mitochondrial presequences, but it lacks an in-frame ATG codon to direct its translation. We anticipate that alternative splicing most likely mediates the mitochondrial versus cytoplasmic expression, although the elements required for directing the enzyme to the mitochondria remain to be confirmed.

Holocarboxylase synthetase (HCS; EC 6.3.4.10) catalyzes the biotinylation of three mitochondrial and one cytoplasmic biotin-dependent carboxylases in mammalian cells (1). In contrast, Escherichia coli has a single biotinylated protein, the biotinyl carboxyl carrier protein (BCCP), a subunit of the bacterial acetyl-CoA carboxylase (2). The biotinylation of apocarboxylases by HCS occurs by the addition of a biotin molecule to a specific lysine residue located within a highly conserved sequence, Met-Lys-Met, present in all biotindependent enzymes (3). A deficiency of HCS activity in humans is responsible for the neonatal form of multiple carboxylase deficiency (4-6). Patients have life-threatening metabolic acidosis which, in almost all cases, can be successfully treated with pharmacologic doses of oral biotin (7). Biochemically, affected patients show a decrease in the activity of all four biotin-dependent carboxylases (8, 9).

Unlike the mammalian HCS, the biotin ligase of *E. coli*, BirA, has been purified and its gene has been cloned (10, 11). It is a versatile enzyme because, in addition to its role in biotinylation of BCCP, it also acts as the repressor of the biotin operon (12). Mutations are known in the *birA* gene that affect either the repressor or the biotin-ligase function (13–15). Indeed, the latter mutants are biotin auxotrophs, by analogy with the human disorder (16, 17). Based on these data and evidence that mammalian HCS is a monomeric protein (18, 19), we set out to clone the human HCS cDNA by expression in bacteria with a temperature-sensitive *birA* mutation. We report here the isolation of human HCS cDNA clones capable of replacing the biotin ligase function of *E. coli.*<sup>†</sup>

## **MATERIALS AND METHODS**

Materials. Radiolabeled compounds were from DuPont. A human multiple tissue Northern blot and  $\beta$ -actin cDNA probe kit were from Clontech. cDNA probes were labeled with a random priming kit from Boehringer Mannheim. Oligonucleotides were synthesized by R. Clarizio (Montreal Children's Hospital). Sequenase was from United States Biochemical. Temperature-sensitive E. coli birA mutants were provided by A. M. Campbell (Stanford University) (16, 17) and A. J. Otsuka (Illinois State University). A 67-aa carboxyl-terminal fragment (p-67) of the  $\alpha$  subunit of human propionyl-CoA carboxylase (PCC- $\alpha$ ), encoded in the vector pFLAG67, was used as substrate for HCS (20). A human cDNA library,  $\lambda$ YES, from Epstein–Barr virus-transformed B lymphocytes and E. coli strain BNN132 containing  $\lambda KC$  phage were provided by S. J. Elledge (Baylor College of Medicine, Houston) (21) and E. McIntosh (York University, Toronto). A cosmid genomic DNA library in pWE15 and made from human lymphocyte DNA was from Stratagene.

Cloning of HCS cDNA and Genomic DNAs. The birA104 mutant was lysogenized with a helper phage ( $\lambda$ KC) harboring the kanamycin-resistance (kan) and cre genes. Kanamycinresistant colonies showed wild-type growth at 30°C in LB plates with 5  $\mu$ M biotin but no growth at 42°C. The resulting E. coli birA104( $\lambda$ KC) was infected with a human cDNA library in  $\lambda$ YES as described (21) and incubated for 2 hr at 30°C in nonselective LB medium [containing 5  $\mu$ M biotin and 2 mM isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG)]. The cells were collected by centrifugation at  $1500 \times g$ , washed twice with LB medium to eliminate biotin, and plated in selective LB agar medium (containing ampicillin at 50  $\mu$ g/ml and no added biotin) at 30°C for 24 hr. Survivor colonies were recovered and cDNA clones were isolated as pSE936 plasmid derivatives. The clones were grouped in classes by restriction analysis and subcloned into the EcoRI site of pBluescript (Stratagene) for sequencing. The cosmid genomic DNA library was used for the isolation of clones containing the 5' end of the HCS gene; a 5'-end, 366-bp Kpn I fragment of cDNA clone BL-11 was the probe.

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

Abbreviations: HCS, holocarboxylase synthetase; PCC, propionyl-CoA carboxylase; BCCP, biotinyl carboxyl carrier protein; IPTG, isopropyl  $\beta$ -D-thiogalactopyranoside; RT, reverse transcription.

<sup>\*</sup>To whom reprint requests should be addressed at: McGill University-Montreal Children's Hospital Research Institute, 2300 Tupper Street, Montreal, QC, Canada, H3H 1P3.

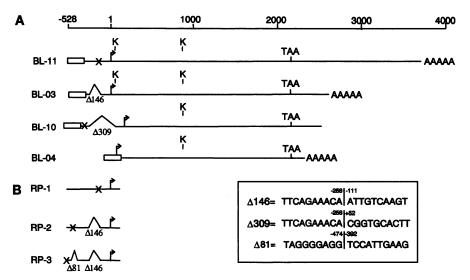
<sup>&</sup>lt;sup>†</sup>The sequence reported in this paper has been deposited in the GenBank data base [accession no. X80160 (HSHCSR)].

Functional Characterization of HCS cDNA. The candidate HCS cDNA clones in pSE936 were used to retransform birA104 cells. The cells were plated in biotin-restricted LB agar medium supplemented with 2 mM IPTG and evaluated for growth after overnight incubation. To assess the HCS activity, colonies were grown to an OD<sub>600</sub> of 0.4-0.5 in LB liquid medium and transferred to biotin-free minimal medium [Amedium (22)] with 2 mM IPTG at 30°C or 42°C for 10 min. D-[carbonyl-<sup>14</sup>C]Biotin (1  $\mu$ Ci/ml; 1  $\mu$ Ci = 37 kBq) was added to the cultures, and the incubation was continued for 2 hr. Crude cell extracts were prepared and their proteins were separated by electrophoresis in 0.1% SDS/12% polyacramide gels. The proteins were transferred to nitrocellulose membranes and the radioactivity was quantitated with a phosphor image analyzer (Fujix BAS 2000; Fuji). In some experiments, birA104 cells were transformed with pFLAG67, with or without prior transformation with candidate cDNAs, to test the HCS activity of cDNA clones on peptide p-67. Detection of the radiolabeled peptide was determined after electrophoresis as above.

Analysis of RNA. The multiple tissue Northern blot, prepared from poly(A)<sup>+</sup> RNA (2  $\mu$ g per lane) of the indicated human tissues, was probed with a *Kpn* I fragment spanning positions 37–873 of the HCS cDNA. Hybridization with human  $\beta$ -actin cDNA served as a control for the quantity and integrity of the RNA in the blot. HCS mRNA was also analyzed by agarose or polyacrylamide gel electrophoresis of PCR products derived from reverse-transcribed RNA. Total RNA was isolated from fibroblasts and reverse-transcribed with (dT)<sub>15</sub> as primer (23). The sequence and position of the sense (S) and antisense (A) oligonucleotide primers used in the PCR steps are shown in Fig. 2 [S1 (-528 to -504), S2 (-437 to -414), S3 (-293 to -268), and A1 (+83 to +62)].

## RESULTS

Isolation and Sequence of Human HCS cDNA. The *birA104* ( $\lambda$ KC) cells were infected with 25 × 10<sup>6</sup>  $\lambda$ YES bacteriophage expressing a human lymphocyte cDNA library and plated in selective LB plates without added biotin. After 24 hr of growth at 30°C, 31 bacterial colonies were isolated. Of these, 21 Cre-mediated pSE936 plasmids retained the ability to complement the biotin auxotrophy and temperature-sensitive growth of the *birA104* host. The cDNA clones were evaluated for relatedness by restriction enzyme analysis with *Eco*RI, which cleaved at sites flanking the inserts, and *Kpn* I. Four classes of cDNA inserts were identified, of length 4.0 kb (3 clones), 3.0 kb (8 clones), 2.2 kb (5 clones), and 2.0 kb (5 clones) (Fig. 1*A*). One member of each class (BL-11, BL-03,



BL-10, and BL-04, respectively) was subcloned and sequenced in both orientations.

All of the clones share the same internal sequence but are heterogeneous at their 5' or 3' ends (Fig. 1A). The longest, BL-11, contains an open reading frame of 2178 bp, encoding a predicted protein of 726 aa and  $M_r$  80,759 (Fig. 2). The first in-frame ATG is preceded by a nonsense codon at position -123. At the 3' end, there is an additional  $\approx$ 1.5 kb of sequence beyond the end of the open reading frame, terminating in a poly(A) tail. Two of the clones, BL-03 and BL-10, differ in the alternative deletion (seemingly spliced out) of an internal 146or 309-bp sequence (Fig. 1A,  $\Delta$ 146 and  $\Delta$ 309) beginning at -257 compared with BL-11 (Fig. 2). BL-04 is similar to the other clones but does not extend as far as the variable 5'-end region. At the 3' end, BL-03 and BL-04 contain poly(A) tails, each at a different location and far short of the poly(A) site in BL-11 (Fig. 1A). None of the cDNAs contains a consensus polyadenylylation signal.

The ATG codon used by the bacteria to express BL-11, BL-03, and BL-10 is located within the cDNA (+1 for the first two and +172 for BL-10; Fig. 2), defined either because a nonsense codon precedes the ATG (BL-11 and BL-10) or because the sequence is out of frame with respect to the single ATG in the pSE936 vector (BL-03). BL-04 is expressed as a fusion protein initiated by the ATG in the vector (9 aa). This cDNA begins at +165 relative to BL-11 and the first ATG encoded by the cDNA is located at +172 (Met<sup>58</sup>).

Functional Characterization of HCS cDNA. The clones were evaluated for their ability to restore biotinylation of BCCP in birA104 cells incubated at 30°C or 42°C. The untransformed birA104 cells showed incorporation of [14C]biotin into BCCP at 30°C but not at 42°C, reflective of the temperature-labile biotin ligase in these cells (Fig. 3, BirA104). In contrast, birA104 cells transformed with any of the cDNA clones could biotinylate BCCP at either temperature (Fig. 3). Further, when birA104 cells were transformed with pFLAG67, a very limited incorporation of [14C]biotin into p-67 was obtained at 30°C (Fig. 3, +p-67). Additionally, the presence of p-67 in these cells resulted in a 10-fold reduction in the biotinylation of BCCP by the mutant BirA protein (quantities in legend of Fig. 3). As found for BCCP, birA104 cells were also unable to carry out the biotinylation of the p-67 peptide at 42°C. On the other hand, when birA104 cells were cotransformed with pFLAG67 and pSE936 containing BL-11, BCCP and p-67 were efficiently biotinylated at both temperatures (Fig. 3; +p-67 + BL-11).

**RNA Analysis.** An 837-bp Kpn I fragment of the HCS cDNA (Fig. 1*A*) was used to probe a Northern blot prepared from several human tissues. The probe hybridized to a principal RNA species of 5.8 kb present in all tissue samples (Fig. 4),

FIG. 1. (A) HCS cDNA classes isolated by complementation of E. coli birA104. Alternatively spliced regions relative to BL-11 are indicated as  $\Delta 146$  and  $\Delta 309$ . In-frame ATG codons suitable for initiation of translation are indicated by arrows. Location of stop codons (TAA), polyadenylylation tails (AAAAA), and Kpn I restriction sites (K) used for determining cDNA classes are shown. Upstream nonsense codons are represented by X. Numbering at the top of the figure indicates cDNA length in bp and is based on the first in frame ATG codon on BL-11. (B) Diagram of reverse transcription (RT)-PCR products obtained by use of primers S1-A1. Alternatively spliced regions relative to RP-1 are shown as  $\Delta 81$ ,  $\Delta 146$ , and  $\Delta$ 309. Arrows indicate the position of the first ATG codon in relation to BL-11. (Inset) Sequences of the three splice junctions detected in HCS cDNA.

<u>\$1</u>	
TCCTGCGGCCTGCAGAGGAGGAGCATCAGATAAGCTTATTTCCATGTCACTTGGAATTATAGGGGAGG gtatgttt 	-466
cagagggggttgaaatttatggcagttgtggtatgaagtgacaatctcattatatttctttc	-350
	-234
gatgtggtttctcatgaggttacttgctctaaaggactttatattttggaaccataagagcacccttgctggcccaggcatttatggatgatcccttttagtgctcccagtaacctt P	-117
ccaag   ATTGTCAAGTGGTCAGACTGTTGTTTGCCATTAGCTGCAGACCTGGGGATCCTTATCGGCTAATTGCTGAAGCAAGTGGACAACTTCAGCAAGCTGGGGGGGG	-1
SK <sup>°</sup> IVKWSDCCLPLACRPGDPYRLIA <b>B</b> ASVDNFSKLGVAF	
$\sim$ $h_1$	117
MEDRLHMDNGLVPQKIVSVHLQDSTLKBVKDQVSNKQAQ	39
•BL-04	
ATCCTAGAGCCGAAGCCTGAACCTTCTCTTGAGATTAAGCCTGAGCAGGCGGCATGGGGGCATGGCGGGAGAGACGCCCCAAAGGCTCTTGGTGAAGAACCCAAAGGAGAGGCC	234
I L E P K P E P S L E I K P E Q D G M E H V G R D D P K A L G E E P K Q R R G	78 351
AGTGCCTCTGGGAGTGAGCCTGCTGGGGACAGTGACAGGGAGGG	117
SA SO SE PA O SE LA CONCENCIA CONCEN	468
$\mathbf{G}$ A T A S A E N I P D L P Y D Y S S S L E S V A D E T S P E R E G R R V A	156
CTCACGGGAAAGGCACCAACCATCCTCCTCTATGTGGGCTCCGACTCCCAGGAAGCCCTCGGCCGGTTCCACGAGGTCCGGCTCTGGCCGACTGTGGCGACATTGACAGTTAT	585
L T G K A P N I L L V V G S D S O E A L G R F H E V R S V L A D C V D I D S Y	195
ATTCTCTACCACTGCTGGAGGACAGTGCTCTCAGAGACCCGTGGACGGAC	702
I L Y H L L E D S A L R D P W T D N C L L L V I A T R E S I P E D L Y O K F M	234
GCCTATCTTTCTCAGGGAAGGGAAGGGTGTTGGGCCTGGTCTTCATCCTTCACCTTTGGTGGCTTTCAGGTGACAAGGCAAGGAGGTGCACAAGACAAGCCAGAACTTGGGTTTCCTCC	819
A Y L S O G G K V L G L S S S F T F G G F O V T S K G A L H K T V O N L V F S	273
AAGGCTGACCAGGCGAGGTGAAGCTCAGCGTCTTGAGCAGTGCGCTGCAGGTACCAGGAAGGCCCCCGTCCGGCCAGCCCCGCCAGGCCACCTGGAGAATGAGGACAAG	936
K A D O S E V K L S V L S S G C R Y O E G P V R L S P G R L O G H L E N E D K	312
	1053
DRMIN VHVPFGCTTROMONOGOGOGOMANGIGIGIGIGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG	351
	1170
L L K S S N F R R Y E V L R E I L T T L G L S C D M K O V P A L T P L Y L L S	390
	1287
A A E E I R D F L M O W L G K H V D S E G E I K S G O L S L R F V S S Y V S E	429
	1404
VEITPSCIPVVTNMEAFSSEHFNLEIYRQNLQTKQLGKV	468
	1525
I L F A E V T P T T M R L L D G L M F Q T P Q E M G L I V I A A R Q T E G K G	507
	1638
R G G N V W L S P V G C A L S T L L I S I P L R S Q L G Q R I P F V Q H L M S	546
	1755
V A V V E A V R S I P E Y Q D I N L R V K W P N D I Y Y S D L M K I G G V L V	585
	1872
N S T L M G E T F Y I L I G C G F N V T N S N P T I C I N D L I T E Y N K O H	624
	1989
KAELKPLRADYLIARVVTVLEKLIKBFQDKGPNSVLPLY	663
	2106
Y R Y W Y H S G Q Q V H L G S A E G P K V S I V G L D D S G F L Q V H Q E G G	702
	2223
EVVTVHPDGNSFDMLRNLILPKRR***	726

FIG. 2. DNA sequence and predicted amino acid sequence of human HCS cDNA. The sequence shown was derived from BL-11 and RP-1 (Fig. 1). Alternatively spliced out sequences  $\Delta 81$  and  $\Delta 146$  are shown in small characters flanked by vertical lines. Arrowheads indicate the spliced-out sequence in BL-10 ( $\Delta 309$ ). The 5' end of each cDNA class is indicated by a dot. Oligonucleotide primers used in RT-PCR are overlined. Numbering at right gives positions with counting starting at the first in-frame ATG. Stop codon is indicated by three stars.

with the strongest signals apparent in RNA from skeletal muscle, kidney, and pancreas. Two fainter bands, possibly degradation products, of 4.6 and 4.0 kb were detected in most lanes. An 8.5-kb band was also detected in brain, skeletal muscle, kidney, and lung.

To test for alternative splicing, as suggested by the internal deletions in BL-03 and BL-10, a RT-PCR experiment was

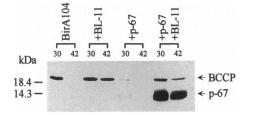


FIG. 3. In vivo [<sup>14</sup>C]biotin incorporation into BCCP and p-67 by BirA and human HCS in *E. coli birA104* cells. Lanes: BirA104, *birA104* cells not transformed with any expression vector; +BL-11, *birA104* cells transformed with pSE936 carrying HCS clone BL-11; +p-67, *birA104* cells transformed with pFLAG67 expressing the p-67 fragment of human PCC- $\alpha$ ; +p-67 + BL-11, *birA104* cells cotransformed with pFLAG67 and pSE936BL-11. Numbers below the brackets indicate the temperature used during biotin incorporation by the bacterial cultures, 30°C or 42°C. Radiolabel in BCCP or p-67 was detected by overnight exposure of nitrocellulose blots to a phosphor image analyzer. Data are for a representative experiment. In arbitrary units, the band densities for BCCP are (left to right) 344, 12; 465, 290; 32, 3; and 278, 101. For p-67, they are 23, 0; 36, 24; 93, 11; and 1908, 808. (S3 and A1) were positioned flanking  $\Delta 309$ , the larger of the two deletions (Fig. 2). Two PCR products were obtained and sequenced. The longer one matched BL-11, whereas the shorter one had a deletion of 146 bp corresponding exactly to the excised sequence of BL-03 (Fig. 1*A*,  $\Delta 146$ ). The  $\Delta 309$  deletion was not detected in the fibroblast RNA.

done on fibroblast RNA in which the oligonucleotide primers

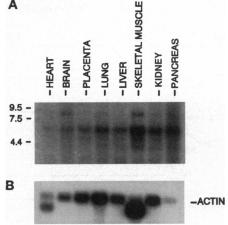


FIG. 4. Northern analysis of HCS mRNA. A Northern blot of poly(A)<sup>+</sup> RNA (2  $\mu$ g per lane) from the indicated human tissues was hybridized with a *Kpn* I-generated 837-bp fragment of clone BL-11 (A) or with a human  $\beta$ -actin probe (B). Positions of molecular size (kb) markers are indicated at left.

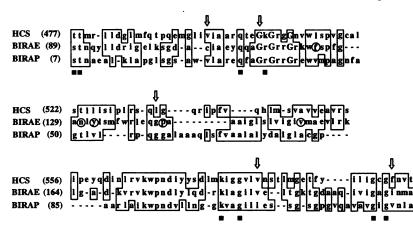
Identification of Additional Transcribed Sequences. The longest cDNA clones were far short of the 5.8-kb major species detected by Northern blotting. No additional sequence was detected in cDNA libraries. However, we obtained additional 5' genomic DNA sequence by using the 5' *Eco*RI-*Kpn* I cDNA fragment of BL-11 to probe a cosmid human genomic DNA library. A 9.5-kb genomic clone overlapping with the 5' end of the cDNA was isolated, and a 2.0-kb *Eco*RI fragment which hybridized to the HCS probe was subcloned and sequenced. This fragment matched the 5' end of the cDNA from -258 to -363. At -258, at the donor site of the  $\Delta146$  and  $\Delta309$  deletions, the genomic clone continued into an intron with the splice donor sequence TTCAGAAACA|gtgagtacca (| = splice site).

To determine whether any additional sequence 5' of -363 was represented in HCS mRNA, reverse transcription–PCR experiments were done on fibroblast RNA, using oligonucleotides as primers that bridged sequences in the genomic clone and the 3' side of  $\Delta 309$  (Fig. 1B). Three oligonucleotides (S1–S3 used with A1) which extended the transcribed sequence at least to -528 gave PCR products with longer 5' extensions but which were structurally similar to BL-11 or BL-03 (Fig. 1B, RP-1 and RP-2). A third PCR product, observed only with S1 as the sense primer, had a unique structure (Fig. 1B, RP-3). It contained two internal deletions,  $\Delta 146$  and another 81 bp in length ( $\Delta 81$ ) beginning at position -473. The sequence of  $\Delta 81$  has the structure of a conventional intron with consensus donor and acceptor splice motifs (Fig. 2).

## DISCUSSION

We have utilized the cross-species activity of biotin ligases to clone a cDNA encoding human HCS by functional complementation of a mutant *birA* strain of *E. coli*. The HCS activity encoded by the cDNA clones was specifically responsible for the biotinylation of *E. coli* BCCP and a fragment of human PCC- $\alpha$  coexpressed in the bacterial strain. During revision of this manuscript, a similar cDNA was reported by Suzuki *et al.* (24), who used oligonucleotides deduced from peptides determined from bovine HCS. Their cDNA is similar in structure to clone BL-11 reported here.

Previous studies have demonstrated HCS activity in both the mitochondria and the cytosol (19, 25). This implied either distinct HCS proteins encoded by different genes or different forms of the same HCS targeted to the two compartments. The absence of mitochondrial or cytosolic carboxylase activities in patients with HCS deficiency effectively excludes the two-enzyme model (7). Further, the cDNA isolated by Suzuki *et al.* (24) was cloned from sequences derived from the cytoplasmic form of HCS. Yet they identified mutations, one of them a frameshift, in a patient with multiple carboxylase deficiency. These data indicate that a single HCS is expressed in the mitochondria and cytosol.



Mechanistically, there are precedents for targeting proteins to both compartments. In general, this is achieved by initiating translation from different AUG codons, one to specify mitochondrial import and the other cytoplasmic targeting. This, in turn, can be achieved through the use of different transcription initiation sites (26, 27) or by taking advantage of secondary structure or sequence context to allow translation from multiple AUGs (28, 29). A third mechanism might be the use of alternative splicing to introduce different amino-terminal sequences into the protein. This has been documented for proteins translated on bound versus free polysomes (30, 31).

It is not obvious which of the above mechanisms might be utilized by HCS for its dual targeting. One possibility, based on the alternative splicing we have observed, is suggested by the identification of two different first in-frame ATG codons (Met<sup>1</sup> and Met<sup>58</sup>) among the four classes of cDNAs. These sites could represent the mitochondrial and cytosolic translation initiation sites, respectively. Inspection of the predicted aminoterminal sequence defined by Met<sup>1</sup> (or Met<sup>58</sup>) does not appear compatible with the amphiphilic nature of mitochondrial leader sequences (32, 33), although there are numerous exceptions to this empiric rule (32). Assessment of the potential of Met<sup>1</sup> to direct import into mitochondria will require its expression in mammalian cells and determination of the cellular fate of the encoded protein.

A second possibility is that alternative splicing of an unidentified upstream sequence could introduce the required mitochondrial import sequence 5' of the currently designated Met<sup>1</sup>. When the predicted protein sequence upstream of Met<sup>1</sup> was evaluated for a potential mitochondrial presequence with the subroutine TRANSPEP from the program PCGENE (Intelli-Genetics), the amino acid sequence PSKIVKWSDC-CLPLACRPG (encoded by nt -120 to -64) was identified. However, the nucleotide sequence lacks a start ATG and is preceded by several nonsense codons, the first at -123. The alternative splicing we have documented successfully intercepts these nonsense codons. Yet in every case, the new sequence is also interrupted by a nonsense codon. None introduce the required methionine residue. Therefore, validation of this model would require the introduction of a still-unidentified exon encoding the amino-terminal portion of a mitochondrial leader sequence.

If the alternative splicing at the 5' end of the HCS mRNA is not for generating a mitochondrial presequence, then its role remains obscure. It cannot be specifically associated with the mRNA pattern observed by Northern blotting, since the alternatively spliced exons are small. Indeed, when the composite, unspliced cDNA and reverse transcription–PCR sequences are taken into account, there remains >1 kb to be accounted for when compared with the 5.8-kb major mRNA detected by Northern blotting. The missing sequence may be present at either end. While we suggest that additional 5' sequence may be required, the presence of several polyade-

> FIG. 5 Amino acid sequence comparison among biotin ligases. HCS, human HCS; BIRAE, *E. coli* BirA protein; BIRAP, *P. denitrificans* putative biotin-ligase. Conserved and similar amino acid residues (A, G; S, T; D, E; N, Q; V, L, I, M; K, R; and F, W, Y) are enclosed in boxes. Positions of amino acids that contact biotin in BirA ( $\blacksquare$ ), as determined by crystallographic analysis (35), are indicated below the protein sequence. Locations of mutations associated with an increase in the K<sub>m</sub> for biotin in *E. coli birA815, birA1, birA215*, and *birA104* (in order from the amino terminus) are indicated by arrows (13–15). Amino acid position for each protein is shown at left.

nylylation sites and the absence of consensus polyadenylylation signals also suggest that a polyadenylylation signal could remain to be identified farther downstream. Resolution of the complex relationship between the cDNA and RNA structures will require the analysis of a tissue with more abundant HCS expression than the skin fibroblasts examined here.

Comparison of the predicted protein sequence of HCS cDNA with the GenBank and EMBL data bases showed that it shares specific areas of homology with the BirA protein of *E. coli* (11) and a candidate BirA sequence in *Paracoccus denitrificans* (34) (Fig. 5). The similarities with HCS are restricted to a region of *E. coli* BirA shown by x-ray crystallography to contain the biotin binding site (35). The two proteins share 36% identity across a 129-aa region, with 6 of the 8 residues involved in direct contact with biotin in BirA identical in human HCS and the remaining 2 showing conservative changes (Thr<sup>477</sup> to Ser; Lys<sup>506</sup> to Arg). Several bacterial mutations associated with an increase in  $K_m$  for biotin fall within this region in BirA (Fig. 5). A similar localization might be anticipated for some mutations in patients with biotin-responsive multiple carboxylase deficiency.

The HCS cDNA has a consensus GXGXXG sequence at aa 505–510 which was also noted in BirA (35) (Fig. 5). This structure has been associated with ATP binding in several enzymes (36, 37), although it appears to be involved in contact with biotin in BirA (35). It is possible (35) that this reflects the requirement that ATP and biotin be spatially close to permit formation of biotinyl-5'-adenylate, an intermediate during the biotinylation reaction. The retention of this site in the evolutionarily distant human and bacterial biotin ligases underscores the potential for ATP binding at this site.

The isolation of a cDNA encoding human HCS will make it possible to determine the mechanism of biotinylation of apocarboxylases and the details of the events leading to mitochondrial versus cytoplasmic targeting. In particular, knowledge of the nucleotide sequence will allow determination of the mutations responsible for multiple carboxylase deficiency and the mechanism of the biotin responsiveness in these patients.

We thank Drs. A. M. Campbell (Stanford University) and A. J. Otsuka (Illinois State University) for supplying *birA* mutants and for helpful discussion, Dr. S. J. Elledge (Baylor College of Medicine, Houston) and Dr. Evan McIntosh (York University, Toronto) for providing us with a human cDNA library in  $\lambda$ YES, Dr. H.-S. Chen (Canadian Genetic Diseases Network) for invaluable sequencing assistance, Dr. T. Garrow (University of California, Berkeley) for helpful comments, and Dr. B. Triggs-Raine (University of Manitoba) for reading the manuscript. These studies were supported by a grant from the Medical Research Council of Canada. A.L.-D.-R. is recipient of a scholarship from the Programa Universitario de Investigación en Salud de la Universidad Nacional Autonoma de México.

- Wood, H. G. & Barden, R. E. (1977) Annu. Rev. Biochem. 46, 385-413.
- Li, S. J. & Cronan, J. E., Jr. (1992) J. Biol. Chem. 267, 855–863.
   Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase,
- F. C. & Wood, H. G. (1988) J. Biol. Chem. 263, 6461–6464.
  4. Burri, B. J., Sweetman, L. & Nyhan, W. L. (1981) J. Clin. Invest.
- 68, 1491-1495.
   Saunders, M. E., Sherwood, W. G., Duthe, M., Surh, L. & Gravel.
- Saunders, M. E., Sherwood, W. G., Duthe, M., Surh, L. & Gravel, R. A. (1982) Am. J. Hum. Genet. 34, 590-601.
- 6. Ghneim, H. K. & Bartlett, K. (1982) Lancet i, 1187-1188.

- Wolf, B. & Heard, G. S. (1995) in *The Metabolic and Molecular Basis of Inherited Disease*, eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw-Hill, New York), pp. 3151-3177.
- Saunders, M., Sweetman, L., Robinson, B., Roth, R., Cohn, R. & Gravel, R. A. (1979) J. Clin. Invest. 64, 1695–1702.
- 9. Feldman, G. L. & Wolf, B. (1981) Clin. Chim. Acta 111, 147-151.
- Eisenberg, M. A., Prakash, O. & Hsiung, S. C. (1982) J. Biol. Chem. 257, 15167–15173.
- 11. Howard, P. K., Shaw, J. & Otsuka, A. J. (1985) Gene 35, 321-331.
- 12. Cronan, J. E. J. (1989) Cell 58, 427-429.
- 13. Barker, D. F. & Campbell, A. M. (1981) J. Mol. Biol. 146, 451-467.
- 14. Barker, D. F. & Campbell, A. M. (1981) J. Mol. Biol. 146, 469-492.
- Buoncristiani, M. R., Howard, P. K. & Otsuka, A. J. (1986) Gene 44, 255–261.
- Campbell, A., del Campillo-Campbell, A. & Chang, R. (1972) Proc. Natl. Acad. Sci. USA 69, 2219–2223.
- 17. Barker, D. F. & Campbell, A. M. (1980) J. Bacteriol. 143, 789-800.
- 18. Chiba, Y., Suzuki, Y. & Narisawa, K. (1991) Jpn. Soc. Inherit. Metab. Dis. 183.
- 19. Chiba, Y., Suzuki, Y., Aoki, Y., Ishida, Y. & Narisawa, K. (1994) Arch. Biochem. Biophys. 313, 8-14.
- 20. León-Del-Rio, A. & Gravel, R. A. (1994) J. Biol. Chem. 269, 22964-22968.
- Elledge, S. J., Mulligan, J. T., Ramer, S. W., Spottswood, M. & Davis, R. W. (1991) Proc. Natl. Acad. Sci. USA 88, 1731–1735.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1991) in *Current Protocols in Molecular Biology*, eds. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (Greene & Wiley, New York), pp. 1.1.1–1.1.2.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 7.12–7.15.
   Suzuki, Y., Aoki, Y., Ishida, Y., Chiba, Y., Iwamatsu, A., Kishino,
- Suzuki, Y., Aoki, Y., Ishida, Y., Chiba, Y., Iwamatsu, A., Kishino, T., Niikawa, N., Matsubara, Y. & Narisawa, K. (1994) Nat. Genet. 8, 122-128.
- Chang, H. I. & Cohen, N. D. (1983) Arch. Biochem. Biophys. 225, 237–247.
- 26. Purdue, P. E., Lumb, M. J. & Danpure, C. J. (1992) Eur. J. Biochem. 207, 757-766.
- Courchesne Smith, C., Jang, S. H., Shi, Q., DeWille, J., Sasaki, G. & Kolattukudy, P. E. (1992) Arch. Biochem. Biophys. 298, 576– 586.
- Slusher, L. B., Gillman, E. C., Martin, N. C. & Hopper, A. K. (1991) Proc. Natl. Acad. Sci. USA 88, 9789–9793.
- 29. Šuzuki, T., Yoshida, T. & Tuboi, S. (1992) Eur. J. Biochem. 207, 767-772.
- Kwiatkowski, D. J., Mehl, R. & Yin, H. L. (1988) J. Cell Biol. 106, 375–384.
- Stella, M. C., Schauerte, H., Straub, K. L. & Leptin, M. (1994) J. Cell Biol. 125, 607–616.
- 32. Hartl, F.-U., Pfanner, N., Nicholson, D. W. & Neupert, W. (1989) Biochim. Biophys. Acta 988, 1-45.
- 33. Gavel, Y. & von Heijne, G. (1990) Protein Eng. 4, 33-37.
- 34. Xu, X., Matsuno Yagi, A. & Yagi, T. (1993) Biochemistry 32, 968-981.
- Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. (1992) Proc. Natl. Acad. Sci. USA 89, 9257– 9261.
- Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) EMBO J. 1, 945–951.
- Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W. & Hennodson, M. A. (1986) *Nature (London)* 323, 448-450.