NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reprcduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Molecular Genetics of Holocarboxylase Synthetase Deficiency

by Alfonso León Del Rio
Department of Biology

McGill University, Montreal

June, 1995

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Doctor of Philosophy

ISBN 0-612-08127-3

TABLE OF CONTENTS

CHAPTER 5

GENERAL DISCUSSION

ABSTRACT

The objective of this thesis was to determine the molecular basis of neonatal multiple carboxylase deficiency (MCD) produced by an impairment in holocarboxylase synthetase (HCS) activity and the origin of the biotin-responsiveness that characterizes this disease. To determine HCS activity, I developed a peptide substrate and used the biotinylation system of *E. coli* to determine its properties. C-terminal fragments of the α subunit of human propionyl-CoA carboxylase (PCC-α) were expressed in *E. coli* and site-directed mutagenesis was used to define the residues required for biotinylation by the bacterial biotin ligase, BirA. These experiments showed that the biotin region of PCC-α can act as an autonomous domain for biotinylation and suggested its use as substrate for human HCS. For the molecular characterization of MCD, I isolated several cDNA clones encoding human HCS by functional complementation of an *E. coli* mutant with a temperature-sensitive BirA. Comparison of the predicted amino acid sequence of HCS with bacterial biotin ligases allowed the identification of the putative biotin-binding domain of this protein. Mutation analysis of DNA from HCS deficient patients showed that most of the changes in the HCS sequence are clustered in the biotin-binding domain. All the patients tested in this study showed deficiency of HCS activity as determined using the PCC-α peptide as substrate for biotinylation. The biotin-responsiveness was demonstrated by obtaining a stimulation of HCS activity of MCD cells at high biotin concentrations while remaining unstimulated in extracts of normal cells. Together with the mutation studies, these results showed that neonatal MCD is caused by mutations in the biotin binding domain of HCS which reduce the affinity of the enzyme towards biotin. This change in the kinetic properties of HCS results in the inefficient biotinylation of carboxylases at physiological concentrations of biotin.

The defect can be overcome by increasing the amount of the vitamin in such a way as to reach the concentration required by the mutant HCS to restore the pool of active carboxylases.

RÉSUMÉ

L'objectif du travail décrit dans cette thèse était de comprendre la nature moléculaire de la déficience multiple néonatale des carboxylases (DMC), causée par une déficience en activité holocarboxylase synthétase (HCS) et l'origine de la dépendance en biotine qui caractérise cette maladie. Pour déterminer l'activité de HCS, j'ai développé un substrat peptidique et en ai testé l'efficacité en utilisant le système de biotinylation de *E. coli*. Pour ce faire, des fragments c-terminaux de la sous-unité α de la propionyl-CoA carboxylase humaine (PCC-α) ont été exprimés dans *E. coli* et une analyse par mutagénèse dirigée a permis de préciser quels résidus sont requis pour en permettre la biotinylation par BirA, l'analogue bactérien de HCS. Ces expériences ont montré que la région de PCC-α qui porte la biotine se comporte comme un domaine autonome pour la biotinylation et en ont suggéré l'utilité comme substrat pour l'enzyme humaine. Afin de permettre la caractérisation moléculaire de la DMC, j'ai isolé plusieurs clones d'ADNc codant pour l'HCS humaine par complémentation fonctionnelle d'une mutation thermosensible de BirA d'*E. coli*. La comparaison de la séquence prédite pour HCS avec celle des analogues bactériens a permis d'identifier un site possible de liaison de la biotine sur HCS. L'analyse des mutations présentes chez les patients atteints d'une déficience en HCS a révélé que la plupart des mutations peuvent être regroupées près de ce domaine de liaison de la biotine. Tous les patients testés lors de cette étude ont montré une déficience en HCS lorsque le peptide PCC-α était utilisé comme substrat. La dépendance en biotine a été mise en évidence par une stimulation de l'activité HCS d'extraits de cellules de patients en présence de concentrations élevées de biotine, alors que l'activité n'était pas stimulée dans des extraits de cellules normales. Ces résultats, associés à la caractérisation des sites mutés dans le DNA des patients, montrent

que la DMC est causée par des mutations dans HCS au site de liaison de la biotine, réduisant l'affinité de l'enzyme envers la biotine. Ce changement des propriétés cinétiques de HCS résulte en une biotinylation inefficace des carboxylases aux concentrations physiologiques en biotine. Ce défaut peut être compensé en augmentant la concentration de la vitamine de façon à permettre une activité HCS suffisante pour activer les carboxylases.

ACKNOWLEDGEMENTS

This thesis was written as a result of material provided and invaluable suggestions made by many scientists, particularly Dr. Allan M. Campbell (Stanford University), Dr. Stephen Elledge (Baylor College of Medicine, Houston), Dr. Timothy Garrow (University of California, Berkeley), Dr. Daniel Leclerc (McGill University, Montreal), Dr. Evan McIntosh (York University, Toronto), Dr. Anthony Otsuka (Illinois State University) and Dr. Lawrence Sweetman (University of California, Los Angeles).

I have also benefited from the extraordinary work of the people of the Department of Human Genetics of the Montreal Children's Hospital-Research Institute, Dr. Daniel Leclerc, Lucie Dupuis, Dr. Nobuaki Wakamatsu, Beverly Akerman, Eric Campeau and Rino Clarizio. I would like to thank all of them and the following persons for surrounding me with a creative and stimulating environment and favouring me with their friendship. Dr. Barbara Triggs-Raine, Magali Loyer, Ija Italiano and of course Marlene Aardse.

I would like to thank particularly to Dr. Roy A. Gravel for his support and encouragement to do all that I wanted (experimentally speaking, of course) during the last four and a half years, for many hours of exciting conversations about molecular biology and astronomy and for skilfully applying the accelerator and the brake to my imagination when each was needed.

I would like to express my gratitude to Dr. Antonio Velazquez for nine years of friendship and invaluable advice. To Universidad Nacional Autonoma de Mexico my thanks for supporting my

studies through a scholarship from the Programa Universitario de Investigacion en Salud.

Personal, too, are some of my sentiments of gratitude. To my parents Alfonso and Maria de la Luz Leon for being there whenever I needed them and for their unconditional love and support. I think this is the best moment to thank you both for not saying anything and even helping in the hunt for all the bugs and animals that were unwittingly released in my room during my undergraduate years.

I would like to dedicate this thesis to my wife Gabriela as a token of gratitude for all these delightful years of love and companionship. In particular, for her commitment to molecular biology which rendered the cloning of the entire human genome: Bruno. I co-dedicate this thesis to my son Bruno who spent countless hours improving my knowledge in palaeontology by introducing me to Barney and friends and for being always an endless source of love, enthusiasm and strength.

CHAPTER 1

GENERAL INTRODUCTION AND BACKGROUND

I. BIOTIN AND BIOTIN-DEPENDENT CARBOXYLASES

A. Biotin as Prosthetic Group of Carboxylases

Biotin is a water soluble enzyme cofactor that belongs to the vitamin B complex (1,2). It was first isolated and characterized by Kogl in 1936 (3) as an essential growth factor for yeast. Its importance for higher organisms was recognized when biotin was shown to protect rats from the toxic effects of a diet rich in raw eggs (4,5). The toxicity was later found to be due to avidin, a glycoprotein present in the egg white with extraordinary affinity for biotin (4,5,10).

Although the importance of dietary biotin was known for many years, its role in metabolism remained obscure until Lardy and Peanasky (78) reported that biotin deficiency in rats reduced the ability of liver mitochondrial extracts to catalyze the ATP dependent carboxylation of propionate to succinate. In 1958, highly purified acetyl-CoA carboxylase (ACC), isolated from avian liver extracts, proved to be enriched with respect to biotin and to be inhibited by avidin (79). Finally, Lane and Lynen (8) demonstrated that biotin was covalently bound to the enzyme propionyl-CoA carboxylase (PCC). As a result of this observation, it was soon found that biotin is used in all living organisms as the prosthetic group of enzymes involved in carboxylation reactions (9,10,11,12). These enzymes are synthesized as inactive precursors or apocarboxylases, that are activated by the covalent addition of biotin (Fig.1). The role of biotin in carboxylases is to act as a vector for transfer of a carboxyl group between donor and acceptor molecules during the carboxylation reaction.

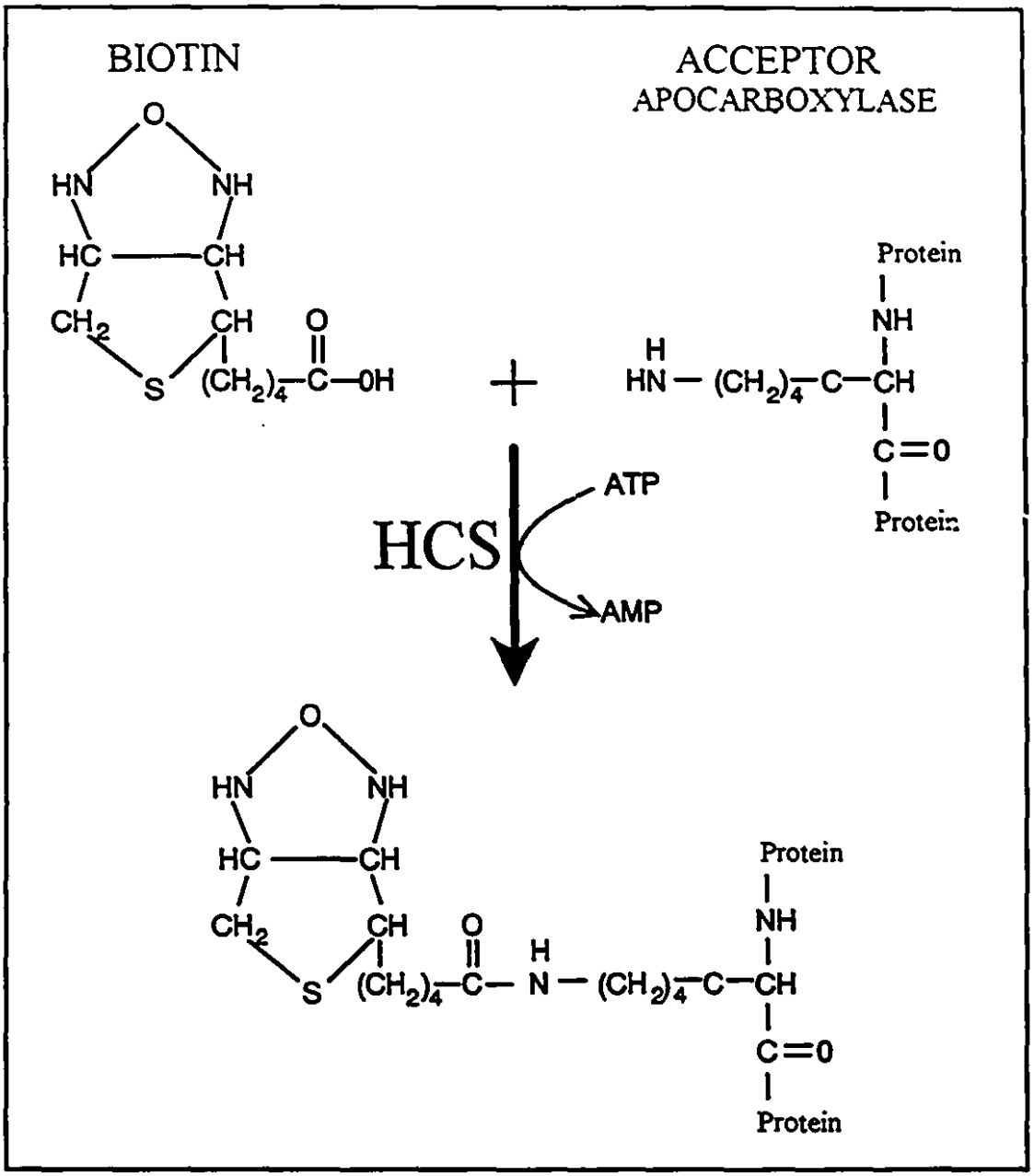


Fig. 1. Biotinylation of apocarboxylases. Activation of carboxylases occurs by the addition of a biotin molecule to a specific Lys residue in the polypeptide chain. This reaction is catalyzed by holocarboxylase synthetase (HCS).

The metabolic processes in which biotin-dependent enzymes are involved are quite divergent and can be divided into three classes (18,19):

Class 1. The proteins that catalyze this type of reaction transfer free bicarbonate and generate carboxybiotin as an intermediate product in an ATP-dependent fashion. All eukaryotic carboxylases belong to this class.

$$\text{E-biotin} + \text{ATP} + \text{HCO}_3 \rightarrow \text{E-biotin-CO}_2 + \text{ADP} + \text{Pi}$$
$$\text{E-biotin-CO}_2 + \text{R-H} \rightarrow \text{E-biotin} + \text{R-CO}_2$$

Class II. Proteins in this class are present in prokaryotes and are involved in energy production. These proteins are biotin-dependent, sodium transport decarboxylases which have as their main task to pump sodium against a concentration gradient without the involvement of ATP.

$$\text{R-CO}_2 + \text{E-biotin} \rightarrow \text{R-H} + \text{E-biotin-CO}_2$$
$$\text{E-biotin-CO}_2 + \text{H}^+ + 2\text{Na}^{+\text{in}} \rightarrow \text{E-biotin} + \text{HCO}_3 + 2\text{Na}^{+\text{out}}$$

Class III. This class is represented only by the prokaryotic enzyme, transcarboxylase (TC). It is composed of three different polypeptides and uses methylmalonyl-CoA as the donor of the carboxyl group that is to be transferred in an ATP-independent reaction.

$$\underset{\text{Methylmalonyl-CoA}}{\text{CH}_3\text{CH(COO}^-\text{)COSCoA}} + \underset{\text{Pyruvate}}{\text{CH}_3\text{COCOO}^-} \rightarrow \underset{\text{Propionyl-CoA}}{\text{CH}_3\text{CH}_2\text{COSCoA}} + \underset{\text{Oxaloacetate}}{^-\text{OOCCH}_2\text{COCOO}^-}$$

B. Architecture of Biotin-Dependent Carboxylases

The number of different biotin-dependent carboxylases present in the cell has not been constant during evolution, varying widely among different species. Mammalian cells have four biotin containing proteins: PCC, ACC, pyruvate carboxylase (PC), and methylmalonyl-CoA carboxylase (MCC). In other species, such as yeast, there are two biotin-dependent carboxylases, PC and ACC. There is only one in *E. coli*, ACC,which, as with its eukaryotic counterpart, catalyses the first committed step in fatty acid synthesis (24).

The polypeptide structure of biotin-dependent enzymes has also been extensively modified during evolution. Based on their molecular architecture, the members of the family of biotin enzymes can be divided in three different groups (80). The first group is represented by ACC of *E. coli* and transcarboxylase from *Propionibacterium shermanii* (Fig.2B). These enzymes are formed by three functionally distinct subunits which, represented by *E. coli* ACC, are: (1) the biotin carboxyl carrier protein (BCCP) which is the subunit that carries the biotin cofactor, (2) biotin carboxylase (BC), which carries out biotin carboxylation, and (3) carboxyl transferase (CT), responsible of the transfer of the carboxyl group from biotin to the acceptor molecule of carboxylation (20,21,22,23).

In the second group, represented by human PCC and MCC, *Achromobacter* MCC and *Pseudomonas citronellolis* PC, only two types of polypeptides are present. One of them, the α subunit, contains the biotin cofactor together with the biotin carboxylase activity.

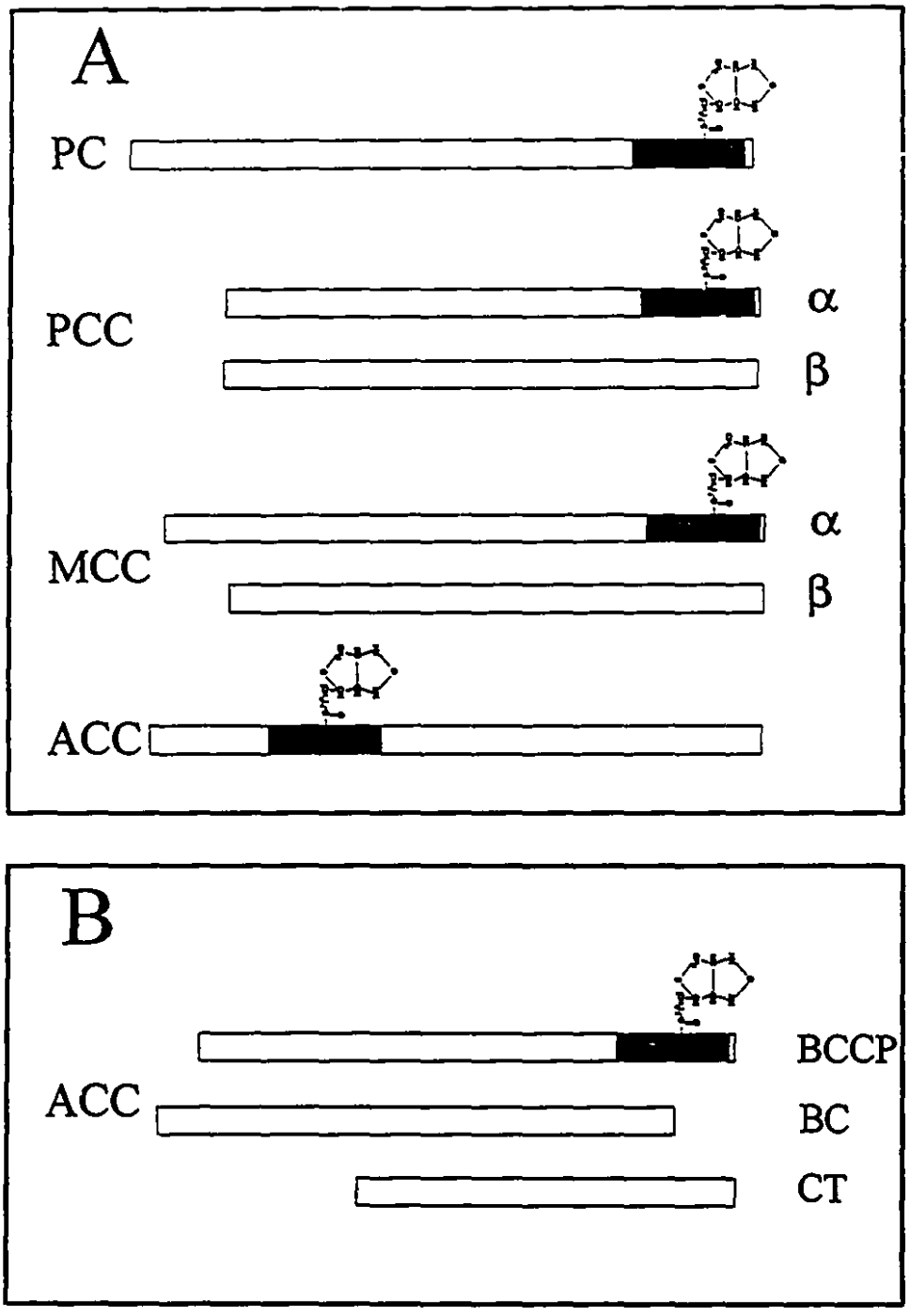


Fig. 2. Architecture of biotin-dependent carboxylases

The other, the ß subunit, carries the carboxyl transferase activity (Fig. 2A). In the third group, all three functions have been incorporated in one multifunctional polypeptide chain, which occurs several times in the native enzyme. ACC and PC from human and yeast cells are examples of enzymes composed of multifunctional polypeptides (Fig.2A).

It is interesting that structural variation occurs not only between different types of carboxylases, but also among equivalent enzymes from different species. For example, human ACC is composed of a single type of polypeptide while the bacterial ACC is, as described above, a heteropolymeric protein.

Given that biotin-dependent enzymes catalyze similar reactions, it is reasonable to consider that the genes that encode biotin-carrier polypeptides are derived from a common ancestor. It is possible that, through a process of one or more duplications, the primordial gene gave rise to two or more biotin-polypeptide genes that then evolved independently to code for biotin enzymes with different but still related functions (80). Under this scenario, the structurally different carboxylases would represent various stages in the evolution of the biotinylation system. If this is true, the original carboxylation reaction may have been catalyzed by an enzyme complex composed of an easily dissociable biotin carboxyl carrier protein which was able to recruit other polypeptides (81). However, part of the carboxylase evolution must also have been convergent, ie., through a process of gene fusion. This could have led to the formation of enzymes with multifunctional polypeptide chains. These last transformations of the biotin enzymes may have occurred because they facilitated the subunit association process and enzymatic

action upon different substrates. The biosynthesis of multifunctional polypeptides is also, from the kinetic and regulatory points of view, more favourable to the cell than the formation of an enzyme complex. Multifunctional peptides overcome the problem of stoichiometry with respect to the regulation of the synchronous synthesis of multiple component proteins.

C. Role of Carboxylases in Human Metabolism

The four biotin-dependent carboxylases of mammalian cells play a key role in different pathways of intermediary metabolism (Fig.3). PCC is involved in the catabolism of branched-chain amino acids and fatty acids of odd-numbered chain length. MCC is involved in the catabolism of the amino acid leucine. PC converts pyruvate to oxaloacetate, an intermediate product in the synthesis of phosphoenolpyruvate and, hence, is essential for glucose synthesis. ACC is responsible for the formation of malonyl-CoA in the biosynthesis of fatty acids. In eukaryotic cells, PC, PCC and MCC are targeted to the mitochondrial matrix by the presence of a mitochondrial leader sequence at the amino terminus of their constituent polypeptides. ACC is the only biotin-dependent carboxylase present in the cytosol.

D. Biotinylation of Apocarboxylases

Biotinylation of carboxylases is catalyzed in all organisms by an enzyme generically named biotin ligase in prokaryotes or holocarboxylase synthetase (HCS) in eukaryotes. The first evidence for a synthetase was obtained by Kosow and Lane (82,83,84) using a soluble apo-PCC holocarboxylase synthesizing system obtained from biotin-deficient

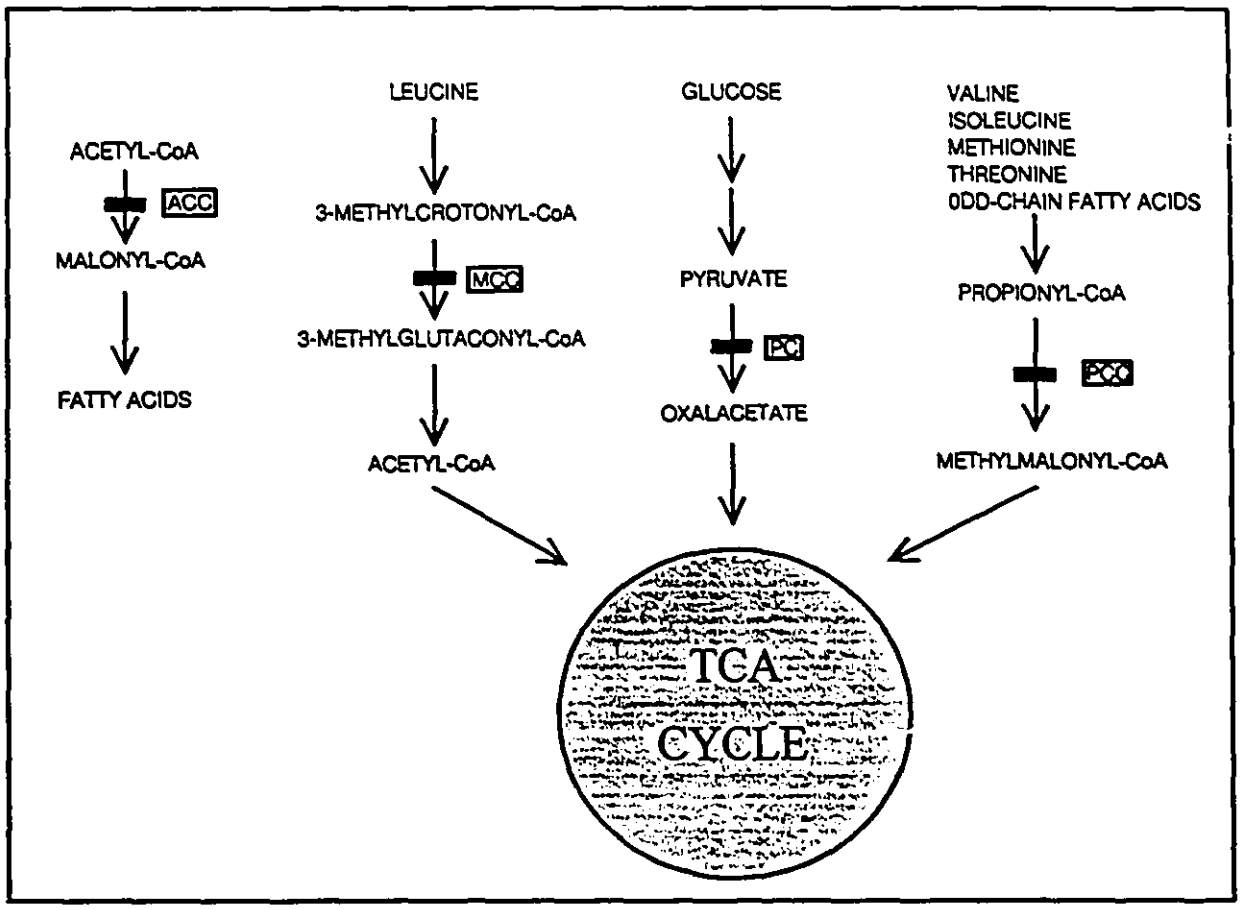
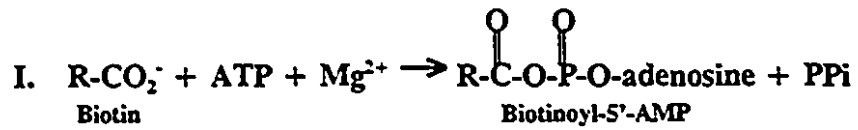
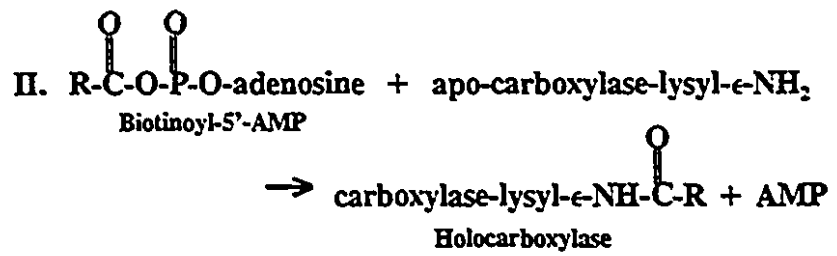


Fig. 3. Role of biotin-dependent carboxylases in human metabolism.

rat liver. They found that biotinylation of PCC requires ATP and a divalent ion, such as Mg^{2+}. The previous isolation of biocytin (ϵ-N-biotinyl-L-lysine) from yeast extract by Wright et al. (85) suggested that the site of attachment was a lysine residue in the enzyme (see section E). Lane et al. (67), using purified preparations of the synthetase and methylmalonyl-CoA- pyruvate apotranscarboxylase from *Propionibacterium shermanii*, demonstrated that the synthetase reaction involves two steps.





In the first part, biotin is activated using ATP with formation of biotinoyl-5'-AMP. During the second step, biotin is attached to the ϵ-amino group of a lysine residue of the acceptor apocarboxylase rendering it an active enzyme. Similar results were obtained by McAllister and Coon (69) with the corresponding synthetase for apo-PCC from rabbit liver. Biotinylation in *E.coli* is catalyzed by the BirA protein through two partial reactions (Fig.4B). In the first part, biotin is activated by ATP with the formation of biotinyl-5'-AMP. In the second half reaction, the biotin moiety is transferred to the apo-BCCP subunit yielding holo-BCCP. Although these observations demonstrated that bacterial and animal synthetases have a common enzymatic mechanism, recent characterization of these two systems has shown that bacterial and animal synthetases, later designated as BirA protein and HCS, respectively, are structurally and functionally

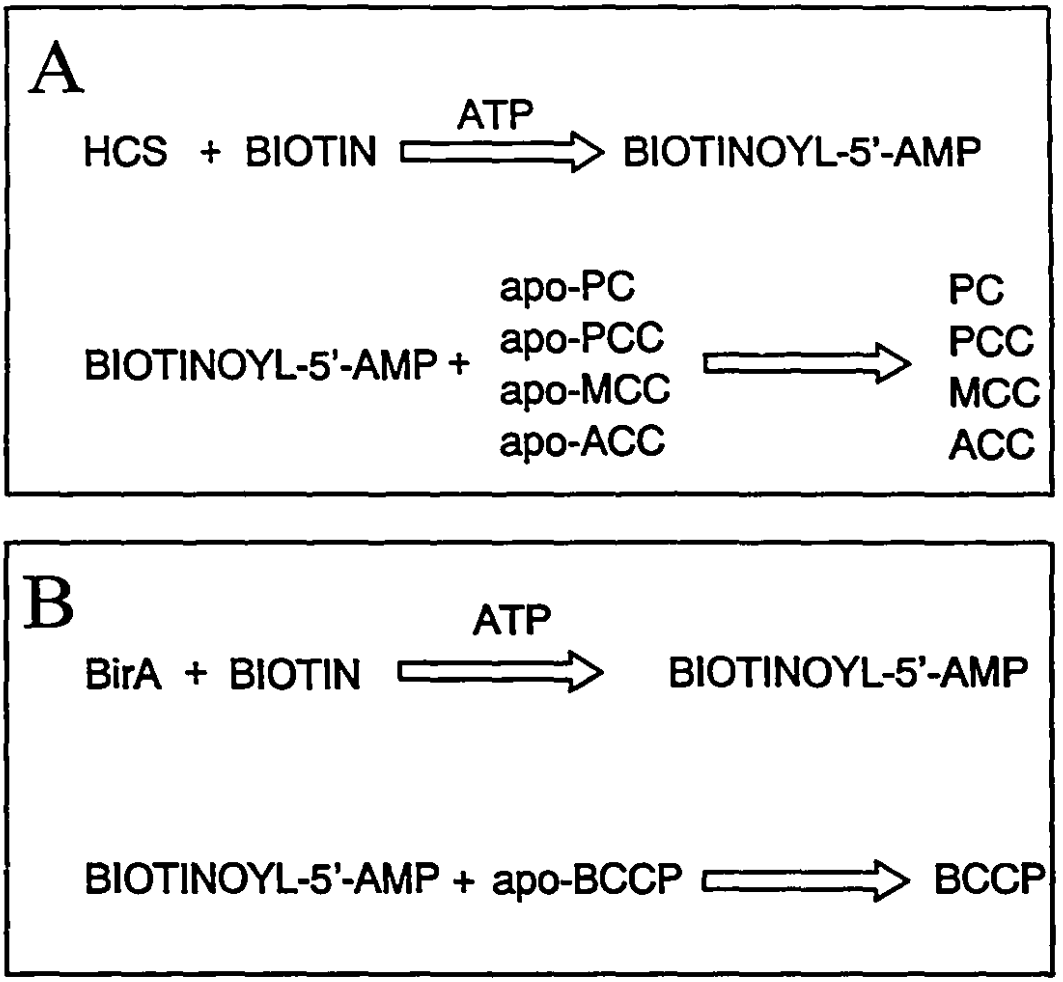


Fig. 4. Biotinylation reaction. Biotinylation of apocarboxylases is catalyzed through a two step reaction by mammalian HCS (A) and bacterial biotin ligase (B).

very different proteins.

BirA Protein and the Biotin Operon. *E. coli* is among 30 different species of prokaryotes able to synthesize biotin *de novo* (25). It occurs through a pathway from pimelic acid to 7,8-substituted pelargonic acid and desthiobiotin, with a subsequent incorporation of a sulfur atom to form biotin (26,27). In these species, the endogenous synthesis ensures an adequate supply of biotin for the activation of essential ACC. In this way, biotin, regardless of the nutrient environment, is not growth limiting.

In one of the most exquisite mechanisms of gene regulation, the BirA protein acts both as the biotin ligase and as the repressor of biotin synthesis (Fig.5), (28,29). The genes that encode the enzymes responsible for biotin synthesis (*bioABCDEF*), with one exception, are located as a cluster near the attachment site of phage lambda on the bacterial chromosome constituting the biotin operon (BIO) (27,28). Transcription of these genes occurs from two partially overlapping promoters. In this system, BirA functions as a switch turning off the transcription of the biotin operon by binding the *bio* operator and obstructing the transcription of *bioBCDEF*. Under these circumstances, maximal transcription (derepression) occurs because BirA is unable to bind the operator site when it is not occupied by its corepressor, biotinyl-5'-AMP (Fig.5A). However, when the supply of biotin exceeds the amount needed to biotinylate the existing apo-BCCP, the BirA-biotinoyl-5'-AMP accumulates and binds the *bio* operator thereby repressing transcription of the operon (Fig.5B). In this beautifully orchestrated system, the rate of biotin operon transcription is sensitive not only to the intracellular

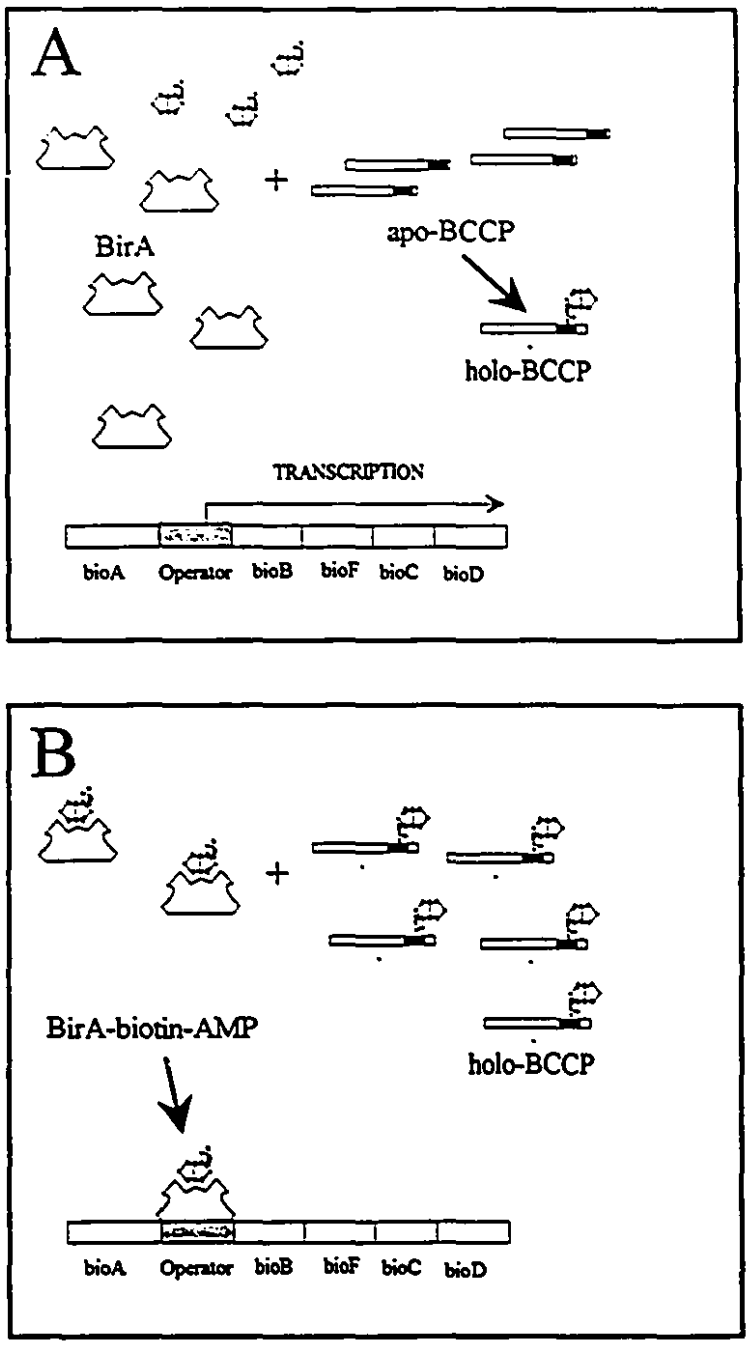


Fig. 5. **The biotin ligase of E. coli acts as the repressor of the biotin operon.** When biotin is present in low concentrations (A) all biotinyl-5'-AMP produced by BirA is used to biotinylate BCCP. When biotin is in excess with respect to BCCP (B), biotinyl-5'-AMP accumulates and the BirA-biotinyl-5'-AMP complex is able to bind the operator of the Bio operon.

concentration of biotin, but also to the amount of BCCP subunit to which biotin must be attached (29). The gene encoding the BirA protein has been cloned and the protein purified (17,27,30,31). These data and the resolution of the structure of BirA by X-ray crystallography have revealed that in this 321 amino acid protein the DNA-binding and enzymatic functions are physically separated. A helix-turn-helix motif situated at positions 22-46 (86) is part of the 80 amino acid domain responsible for the DNA binding activity, while the ligase function is contained at the carboxyl terminus of the protein at residues 83-235 (32,33,86).

Holocarboxylase synthetase. Biotinylation of apocarboxylases in eukaryotes is catalyzed by HCS. In a similar fashion to BirA, biotinylation requires the formation of biotinoyl-5'-AMP as an intermediate product (Fig.4A). In mammalian cells, HCS activity has been located both in cytoplasm and in mitochondria corresponding to the cellular localization of biotin-dependent-carboxylases. Interestingly, in human cells, biotinylation of PCC can occur in the cytoplasm, with HCS acting on isolated α subunits, or within the mitochondria, after multisubunit assembly of PCC has been accomplished (42,43).

To explain the presence of HCS activity in mitochondria and cytosol, I have proposed two alternative hypotheses (Chapter 2). In the first model, biotinylation of all apocarboxylases is catalyzed by a single HCS (Fig.6A). This particular model requires that HCS is targeted in a dual fashion to mitochondria and cytosol. The second hypothesis assumes that there are two different HCS encoded by separate genes (Fig.6B).

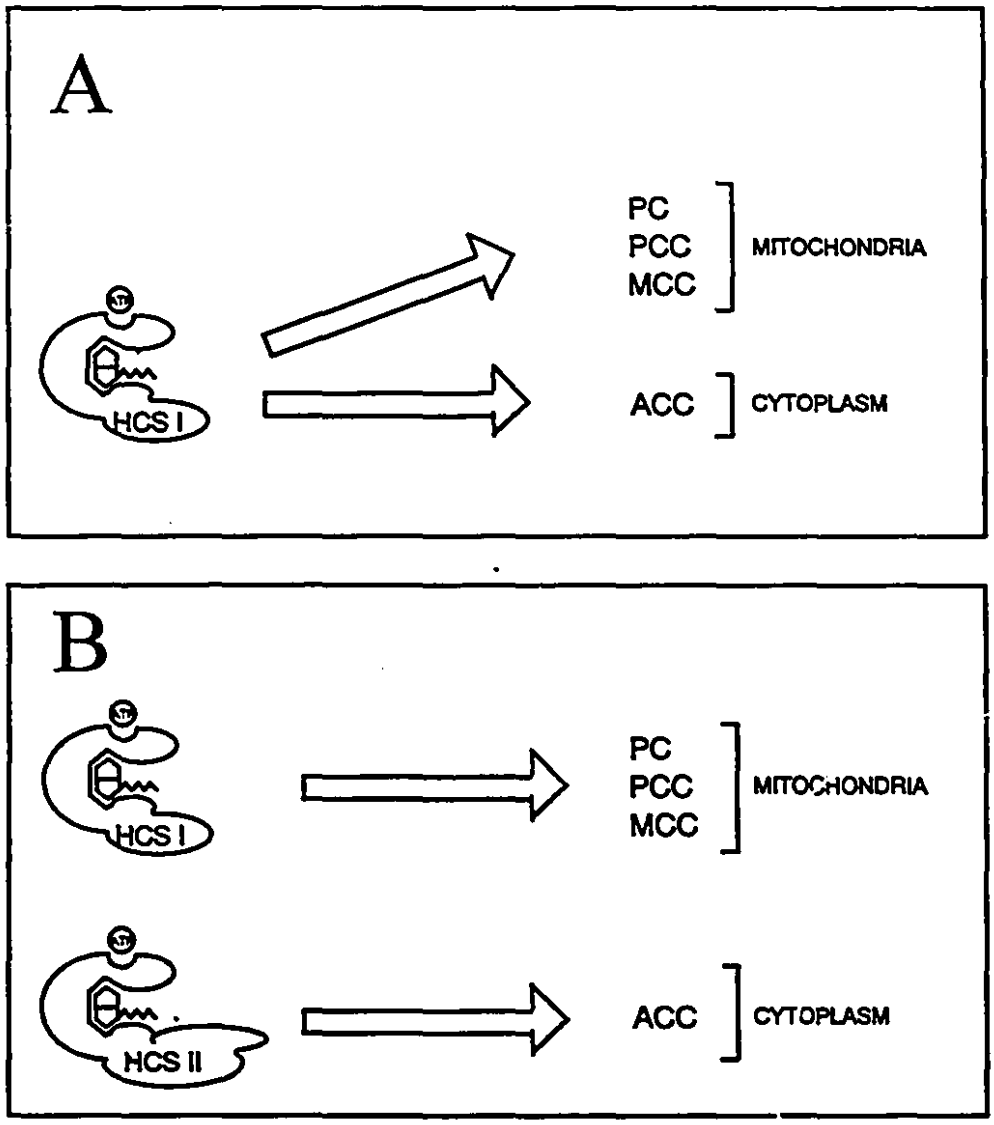


Fig. 6. Two different models are proposed to explain biotinylation of mitochondrial and cytoplasmic carboxylases.

One of these genes would encode a mitochondrial HCS responsible for the biotinylation of PC, MCC and PCC. The second gene would encode a cytoplasmic HCS that would carry out ACC biotinylation.

E. Specificity of Bacterial and Eukaryotic Biotin Ligases

In the synthesis of active holocarboxylases, both BirA protein and HCS show high specificity towards biotin. Some biotin analogues can be used by the synthetases to catalyze the acylation of the correct lysine residue of apocarboxylases but the resulting holoenzyme is inactive (67). The specificity of HCS and BirA protein for acceptor apocarboxylases is not as rigorous as for biotin. The results from various laboratories show that synthetases from different species can catalyze biotinylation of apocarboxylases from different organisms. The cross-species activity of synthetases includes biotinylation of bacterial apocarboxylases by mammalian HCS and biotinylation of mammalian carboxylases by the BirA protein (69,70). These observations suggested that biotin-dependent enzymes, despite their structural differences, must share a common recognition domain for biotinylation (Fig.2). This hypothesis is supported by the strong amino acid sequence conservation around the biotin-binding site among biotin-dependent carboxylases. Three structural features are of particular interest for this study (Fig.7):

1) All carboxylases are biotinylated at a specific lysine residue located within the conserved sequence Met-Lys-Met, with the sole exception of ACC from *Anabaena sp.* where the biotinylated Lys residue is within the sequence Met-Lys-Leu (87). The biotinylation site is located 35 amino acids from the carboxyl terminus of the proteins

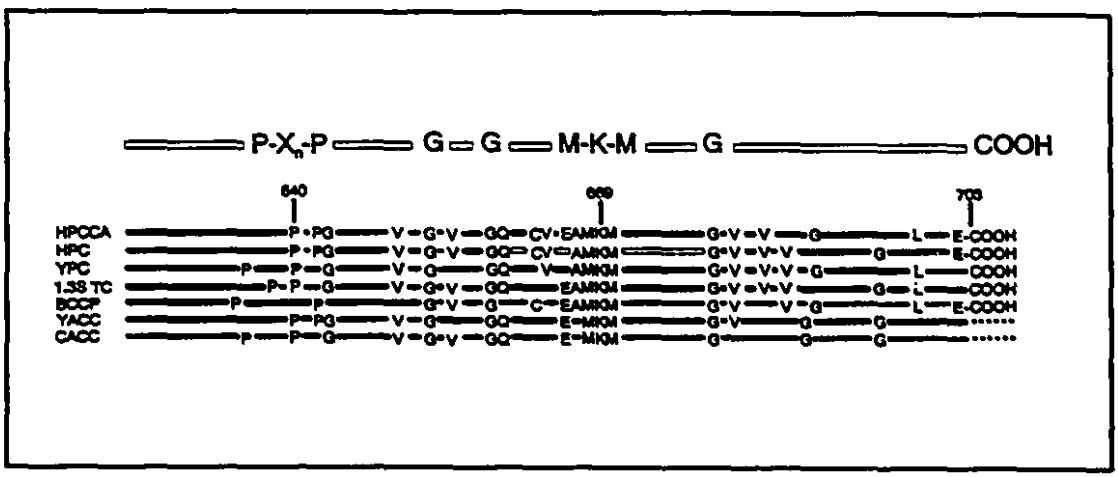


Fig. 7. **Amino acid sequence conservation among biotin-dependent enzymes.** The sequence of different biotin enzymes are aligned with respect to the biotin binding site to show the amino acid conservation around this region. The Pro region described in the text and the conserved Gly residues are shown in the schematic representation of the biotin domain above the protein sequences. The numbers represent the amino acid position in the α subunit of PCC. Protein sequences are as follows: HPCCA, human propionyl-CoA carboxylase, HPC, human pyruvate carboxylase; YPC, yeast pyruvate carboxylase; 1.3S TC, 1.3 S subunit of transcarboxylase of *P. shermanii*; YACC, yeast acetyl-CoA carboxylase; CACC, chicken acetyl-CoA carboxylase.

except in ACC where the Met-Lys-Met motif is located on the amino terminus of the protein.

2) All carboxylases contain a Pro-Xn-Pro motif 30 to 35 amino acids from amino terminal side of the biotin binding site, where n= 1 to 5 hydrophobic amino acids. It has been suggested that this motif may be involved in providing carboxylases with a flexible point that would turn the biotin domain into a swinging arm. The arm with biotin located at one end would be able to oscillate, facilitating the transfer of a carboxyl group between the carboxyl donor and acceptor molecules.

3) In all biotin-dependent enzymes sequenced so far, there are several glycine and valine residues conserved at specific positions in relation to the biotinylation site.

The extreme conservation of these motifs among biotin containing proteins implies a functional or structural role for them. It has been suggested that these sequences might be involved in recognition of the biotinylation site by the synthetase, in the mechanism of carboxyl group transfer or in the recognition of biotinylated peptides by biotinidase (29).

To investigate the role of some of these sites, researchers have used the 1.3S subunit (123 amino acids in length) from *P. shermanii* transcarboxylase as a model of a biotin-dependent enzyme. *E. coli* expression of a mutant 1.3S protein, in which the methionine residues that flank the biotin-binding site were substituted by leucine, resulted

in normal biotinylation of the protein by BirA (72,73). However, these changes affected the enzymatic activity of the protein by impairing the transfer of the carboxyl group from the donor molecule to biotin (74). The interpretation of these latter results has become difficult by the recent cloning of *Anabaena sp.* ACC where the biotin binding site has the sequence Met-Lys-Leu. It is possible that structural differences between the biotin carboxylase subunit of *P. shermanii* transcarboxylase and *Anabaena sp.* ACC may impose different restrictions for the recognition and carboxylation of biotin.

The importance of the distance of the biotin-binding site from the carboxyl terminus in carboxylases was examined by expressing mutant versions of the 1.3S subunit in which the length to the C-terminus was varied. These experiments demonstrated that lengthening the 1.3S subunit by 2 residues or shortening it by 1 did not affect biotinylation. However, substitution or deletion of the penultimate amino acid abolished biotinylation of the protein (75).

Recently, Cronan (76) showed that partial sequences of the 1.3S subunit can be biotinylated efficiently by *E. coli*. He showed that the minimum size of the 1.3S protein that allows biotinylation is 75 amino acids. Similarly, Yamano et al. (77) reported that fragments of the 1.3S subunit of *P. freudenreichii* expressed in *E. coli* containing residues 18-123 were efficiently biotinylated. Truncated proteins composed of residues 84-123 were not biotinylated, demonstrating the presence of essential sequences for biotinylation of apocarboxylases upstream from the biotin-binding site. Similar studies have not been done for eukaryotic carboxylases.

The high conservation among biotin enzymes suggests that the findings described above could apply to all biotin-dependent enzymes. However, the biotin domains in eukaryotic carboxylases are part of much longer polypeptides and it is not clear whether the requirements for their biotinylation are more complex or involve more distant residues. An objective of this thesis was to isolate human peptides able to act as substrate for HCS to permit the study of biotinylation reactions in eukaryotic cells.

II. BIOTIN METABOLISM AND DEFICIENCY SYNDROMES

A. Biotin Metabolism

During the course of evolution, some prokaryotes and all eukaryotes lost the ability to synthesize biotin and now depend entirely on diet as the only source of the vitamin. However, biotin is present in very low concentrations in nature and this puts at risk the metabolic homeostasis of the cell. To deal with their requirement for biotin, mammals have evolved a very efficient and complex biotin cycle to ensure an adequate supply and utilization of the vitamin (Fig.8).

Most biotin present in food is not readily available to the organism because it is protein bound (34); that is, it remains attached to carboxylases and must be released before it can be re-utilized in further biotinylation reactions. During protein digestion in the small intestine, the carboxylases are broken down to peptides that can later be absorbed through the intestinal ephithelia (35). Through this process, biotin proteins are catabolyzed to their mono-amino acid derivative, biotinyl-lysine, also known as biocytin (7). Biotin is

released from biocytin through the action of biotinidase, which cleaves specifically the quasi-peptide bond between biotin and the ϵ-amino group of lysine (36). In the pancreas of some organisms, there is a high activity of biotinidase which suggests that biotin release occurs mostly in the intestinal lumen (37,38). At this site, the vitamin is actively absorbed in a sodium-dependent fashion against a concentration gradient through the brush-border membrane of enterocytes (37,39,40,41). Some species are able to transport both biotin and biocytin across the small intestine (37), and this could represent an additional mechanism to maximize the utilization of dietary biotin. Where biocytin is transported across the intestinal epithelium, serum biotinidase would become responsible for releasing the vitamin to permit its re-utilization.

Human serum biotinidase is a glycoprotein composed of a single polypeptide with a molecular weight of 76,000 daltons (44,45). The enzyme shows a higher affinity for biocytin (Km range from 5-10 μM) than for larger biotinylated peptides (46), suggesting this compound is the natural substrate of the enzyme. A cDNA encoding human biotinidase was cloned based on the sequences of tryptic peptides derived from the purified serum enzyme (47). The predicted protein is composed of 543 amino acid residues including a potential 41 amino acid signal peptide. In recent studies, it was suggested that biotinidase may have other distinct enzymatic functions besides its role in biotin recycling. These include lipoamidase activity and enkephalin hydrolysis (48,49). However, the physiological significance of these other activities of biotinidase remain to be demonstrated.

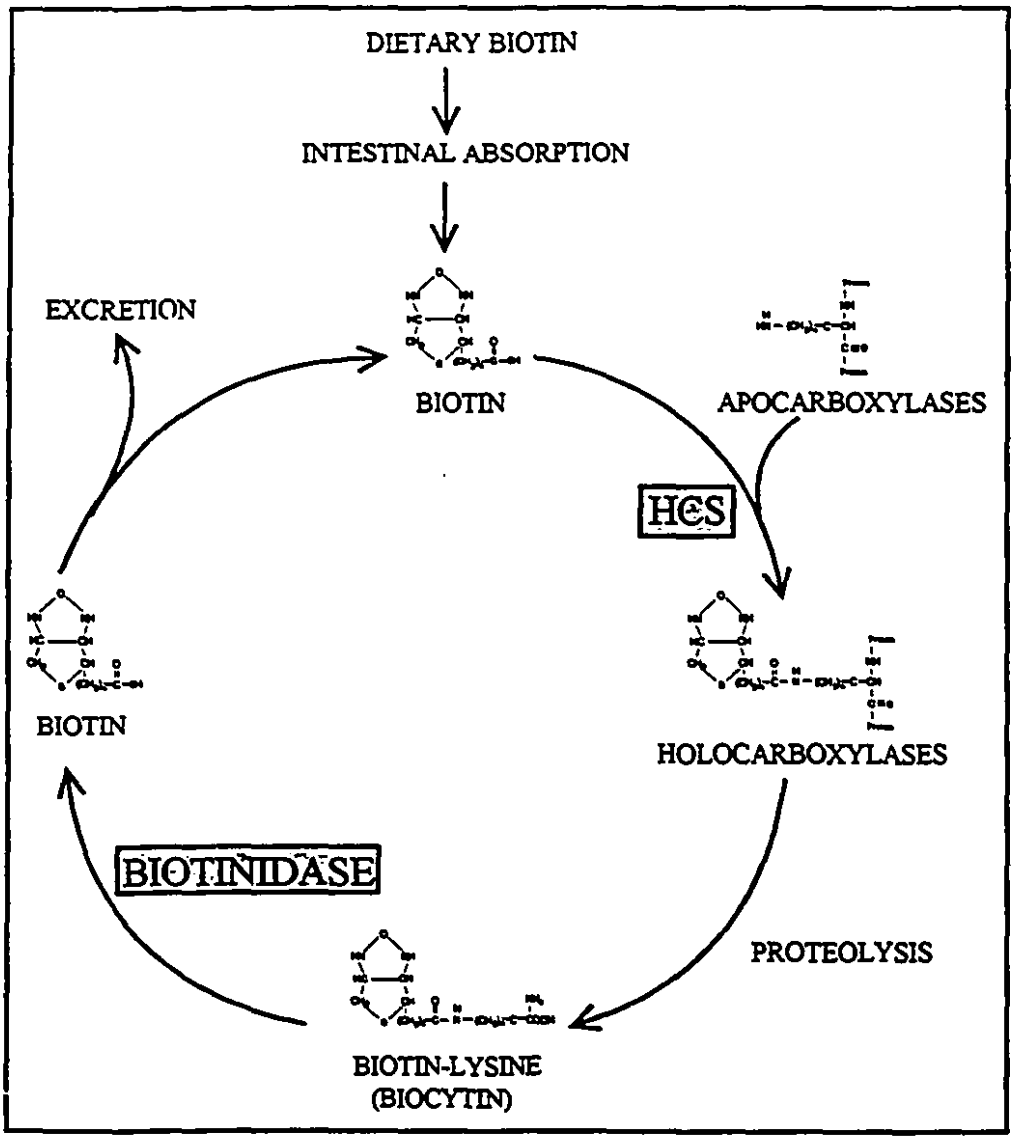


Fig. 8. Biotin cycle in mammals.

B. Disorders of Biotin Metabolism

Acquired biotin deficiency. Biotin deficiency is characterized by a reduction in the activity of the four biotin-dependent carboxylases. Clinically, affected patients present with alopecia, dermatitis, erythematous periorificial rash, dryness and fungal infections (7). Several neurological disorders have been documented such as mild depression, somnolence, muscular pain, hyperesthesia and paraesthesia. Biotin deficiency in humans is extremely rare and generally it is associated with severe malnutrition, prolonged parenteral nutrition (88) or to the consumption of large quantities of avidin, usually in the form of raw eggs (51,52). The highly efficient system for recycling endogenous biotin most likely accounts for the long period of time it takes for human patients to develop the first symptoms of biotin deficiency. In contrast, it takes only 1 to 2 weeks of treatment of biotin-deficient individuals with biotin supplementation for their clinical and biochemical condition to improve.

Deficiency of single carboxylase activities. Inherited metabolic disorders causing deficient activity of each of the mitochondrial carboxylases have been reported in numerous patients (53) and only one case has been documented for ACC deficiency (54). Patients affected with isolated mitochondrial carboxylase deficiency present with vomiting, lethargy, hypotonia, seizures and developmental delay. In cases where PCC or PC activities are impaired, patients may fall into coma and the disease is potentially fatal. Biochemically, deficiency of any of the three mitochondrial carboxylases is associated with metabolic ketoacidosis, and the profile of organic acids in urine is so specific that it can be used to diagnose the particular defect with great accuracy (7).

PCC deficiency results in the accumulation of propionic acid in blood and in the excretion of methylcitrate, tiglylglycine and ß-hydroxypropionate among other compounds. In PC deficiency there is a strong elevation in the excretion of ketone bodies, lactate and alanine. MCC deficient patients excrete high amounts of beta-hydroxyvalerate, beta-methylcrotonate and methylcrotonylglycine.

Individual carboxylase deficiencies are caused by mutations that disrupt the function or structure of the enzyme. The disorders are associated with a failure to respond to biotin treatment (55,56). However these deficiencies can be treated by restricting protein intake or specific amino acids in the diet, thereby preventing the accumulation of toxic intermediates.

Biotin-responsive inborn errors of biotin metabolism. In humans, there are autosomal recessive disorders of biotin metabolism that result from the disruption of the activity of biotinidase or HCS. The main outcome of such deficiencies is the marked reduction in the activity of the four biotin-dependent carboxylases. Clinically and biochemically, affected individuals show a combination of all the isolated carboxylase deficiencies. However, unlike for isolated carboxylase defects, these patients respond successfully to pharmacological doses of biotin. Based on the age of onset of the symptoms, HCS and biotinidase deficiencies are generally known as neonatal and juvenile forms of multiple carboxylase deficiency (MCD), respectively.

Biotinidase deficiency. A defect of biotinidase activity blocks the release of biotin

from food or its recycling after carboxylase proteolysis (see biotin cycle). This disorder results in a secondary biotin deficiency, disrupting the activity of all carboxylases. In these patients the mechanisms of intestinal absorption of biotin and biotinylation of apocarboxylases remain unaltered which explains why affected children respond rapidly to dietary biotin supplementation.

The main clinical and biochemical features in children with biotinidase deficiency include alopecia, developmental delay, organic aciduria, seizures, skin rash, mild hyperammonemia and breathing problems. The onset of the disease varies from two weeks to two years of age. The biotinidase activity in serum of affected individuals is 0 to 9 percent of the mean normal activity (7). The biochemical and most of the clinical manifestations of the disease can be reversed successfully with pharmacological doses of biotin (5 to 20 mg per day).

Holocarboxylase synthetase deficiency. Biotin-responsive multiple carboxylase deficiency caused by mutations in the holocarboxylase synthetase gene is a much more severe metabolic disorder. In this case, the ability of the cells to biotinylate apocarboxylases is directly affected while the intestinal biotin release and the endogenous recycling of the vitamin are normal. Biochemical and clinical manifestations of this disorder can include ketolactic acidosis, organic aciduria, hyperammonemia, skin rash, feeding problems and vomiting, hypotonia, seizures, developmental delay, alopecia and coma (57). The onset of the disease occurs generally within a few hours after birth to fifteen months of age. The urine of these patients show relatively high concentrations

of ß-hydroxyvalerate, ß-methylcrotonylglycine, ß-hydroxypropionate, methylcitrate, lactate and tiglylglycine (7). The excretion of all these metabolites is clear evidence of a generalyzed carboxylase failure. This disorder produces a metabolic catastrophe because it compromises gluconeogensis, synthesis and degradation of fatty acids and the catabolism of several branched chain amino acids. The disease is potentially fatal; however, all its clinical and biochemical manifestations can be reversed with pharmacological doses of biotin (10-80 mg per day).

Fibroblasts from patients with this form of MCD show deficient activity of the three mitochondrial carboxylases when cells are incubated in a medium containing a low concentration of biotin (58,60). When the medium is supplemented with a much higher biotin concentration, the activities of these enzymes increase rapidly to within normal levels (Fig.9). This effect was observed even in the presence of cycloheximide indicating that restoration of enzymatic activity is due to activation rather than to synthesis de novo of carboxylases (61).

The biotin-responsiveness observed in patients and cultured fibroblasts led to the hypothesis that the mutations present in patients with MCD cause a decrease in the affinity of HCS for biotin. Burri et al (62,63) using semipure apocarboxylases from rat as substrate showed that the Km for biotin of HCS from patients with MCD was increased up to 70-fold compared to the normal value from control individuals. These results were substantiated by stdies done on fibroblasts from normal and MCD affected patients. These cells were made biotin deficient by incubating them in a biotin-free

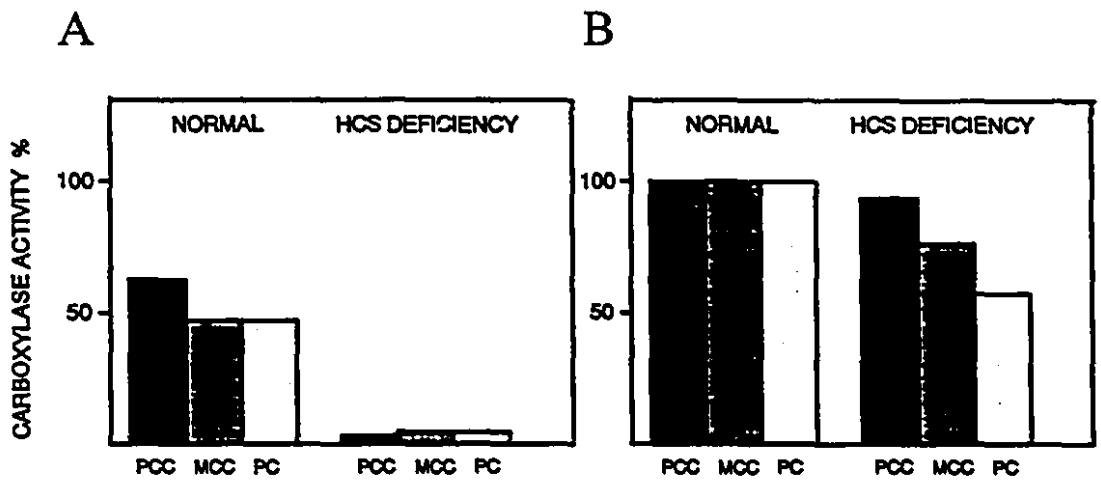


Fig. 9. **Effect of biotin concentration on the activity of carboxylases.** The activity of the three mitochondrial carboxylases was determined in human fibroblasts incubated in low (A), or high (B), biotin concentrations.

medium. Mitochondria prepared from these cells were incubated with ATP, biotin and propionyl-CoA and the incorporation of radioactive bicarbonate into methylmalonyl-CoA was measured. These studies reported a 5000-fold increase in the HCS Km for biotin in MCD patients compared to the value observed in control cells (98). The Km values reported for HCS must be taken with caution because the assays used are indirect, based on the enzymatic activity of an activated carboxylase rather than in the accumulation of biotinylated products. Despite this limitation, these sudies provided the first indication of the mechanism of biotin responsiveness in neonatal MCD. Unfortunately it has not been possible to use the coupled assay to determine heterozygosity in parents of affected children using the coupled assay for HCS. This reflects the limitations of the enzymatic assay.

As described before, it is not completly clear whether there is one or two synthetases in the cell. This has an important impact on the interpretation of the clinical and biochemical findings in multiple carboxylase deficiency. In the single HCS model, mutations affecting the Km of the enzyme for biotin would have a negative impact on the activities of all carboxylases (Fig.6A). If there are two HCS proteins encoded by different genes, the multiple carboxylase deficiency would be the result of mutations in the gene encoding the mitochondrial synthetase. In this case, the patients would show a deficiency of mitochondrial carboxylases only (Fig.6B). Differentiating between these posibilities has been difficult because it has not always been possible to assess the reduction in the activity of all four biotin-dependent carboxylases in MCD. However, some patients have been shown to be deficient in both the mitochondrial and the

cytoplasmic carboxylase activities (64). This observation supports the hypothesis for a single HCS in human cells.

C. Bacterial Model for Multiple Carboxylase Deficiency

Little is known about the structure and physicochemical properties of human HCS because the protein has not been purified or its gene cloned. However, different observations suggest that it may share several functional and possibly structural features with the bacterial BirA protein. First, the biotinylation of apocarboxylases by HCS or BirA proceeds in a similar fashion, via a biotinyl-adenylate intermediate (65,66,67,68,69). Second, *E. coli* cells deficient in biotin-ligase activity are biotin auxotrophs resembling the phenotype of MCD patients in their requirement for high concentrations of exogenous biotin (31,32,33). The mutations that produce a BirA protein with an elevated K_m for biotin are clustered in a discrete region of the gene which encodes the biotin-binding pocket of the protein (31,32,33). It is possible that this domain has been conserved in human HCS. If so, mutations responsible for the biotin-responsiveness of MCD may show a similar distribution pattern.

The conservation of the biotin domain between prokaryotic biotin-ligase and eukaryotic HCS would explain the extraordinary correspondence between the bacterial and human apocarboxylase synthetase deficiencies. This relationship suggested the use of *E. coli* *birA*$^-$ as a model for the study of multiple carboxylase deficiency and human holocarboxylase synthetase. An objective of this thesis was to use *birA*$^-$ cells to isolate a cDNA encoding human HCS based on the cross-species activity exhibited by biotin

ligase and HCS.

PROPOSAL

HCS is responsible for the activation by biotinylation of the four biotin-dependent carboxylases present in human cells. Genetic deficiency of HCS activity results in the neonatal form of the biotin-responsive multiple carboxylase deficiency (MCD), with the consequent impact on fatty acid metabolism, gluconeogenesis and amino acid catabolism. Although it appears that MCD results from a reduction of the affinity of HCS towards biotin, the molecular mechanism underlying the biotin-responsiveness of the disease is not known. The enzymes involved in biotinylation in bacteria have been well documented. However, the lack of a reliable method to determine HCS activity and the fact that no eukaryotic HCS has been purified or cloned has made difficult the analysis of biotinylation in humans.

The objective of this thesis is to determine the molecular basis of neonatal MCD in humans. To achieve this goal, I proposed to develop an accurate assay to determine HCS activity, to isolate the gene or an expressible cDNA encoding human HCS, and to determine mutations and their impact on HCS deficiency in humans.

In Chapter 1, I undertook a study to determine the amino acid sequence requirements of human apocarboxylases for their biotinylation by the BirA protein of *E. coli*. Bacterial expression of truncated fragments of the α subunit of human propionyl-CoA carboxylase, containing a normal or artificially mutated biotin-binding domain, allowed identification

of the smallest peptide able to act as a substrate for biotinylation.

Chapter 2 contains the experiments directed towards the cloning of human HCS. Mutant *E. coli birA*$^{-}$ cells exhibiting biotin auxotrophy and temperature sensitive growth were used as recipients of an expressible human cDNA library. This experiment permitted the isolation of several candidate cDNA clones based on their ability to complement the bacterial phenotype. The peptide substrate for HCS characterized in Chapter 1 was used to identify the enzymatic activity encoded by the cDNA clones.

Finally, in Chapter 3 in a collaborative work with Ms. Lucie Dupuis, the HCS assay was used for the direct analysis of HCS deficiency and biotin-responsiveness in cells from patients with MCD. DNA samples from these cell lines were analyzed by single strand conformation polymorphism (SSCP) for the identification of mutations responsible for the phenotype exhibited by affected individuals.

This study provides a description of the mechanisms for the recognition of human apocarboxylases during the biotinylation reaction. The use of carboxylase fragments as a substrate for biotinylation constitutes the first direct assay for HCS activity and for the diagnosis of the neonatal form of MCD. This thesis contributes to the understanding of biotinylation reaction in human cells by the characterization of the predicted sequence of HCS and the description of possible models for its intracellular fate. The preliminary results on mutation analysis in MCD provide insight into the structure and function of human HCS.

CHAPTER 2

SEQUENCE REQUIREMENTS FOR THE BIOTINYLATION OF CARBOXYL TERMINAL FRAGMENTS OF HUMAN PROPIONYL-COA CARBOXYLASE α SUBUNIT EXPRESSED IN *E. coli*.

PREFACE TO CHAPTER 2

One of the main objectives of this thesis was to develop a substrate to study HCS deficiency and biotin-responsiveness in neonatal MCD. The cross-species activity exhibited by prokaryotic and eukaryotic biotin ligases suggested the use of *E. coli* as the host to develop the HCS substrate. I proposed the expression of fragments of the α subunit of human PCC in this bacteria to study their biotinylation by the BirA protein. These experiments in conjunction with site-directed mutagenesis allowed me to determine the amino acid sequence requirements for the biotinylation of human PCC by biotin ligase and to establish the basis for a direct enzymatic assay to study HCS activity in vitro and in vivo.

I would like to acknowledge Mr. Rino Clarizio, head of the Oligonucletide Synthesis Facility of the Montreal Children's Hospital-Research Institute for helpful technical comments and for his assistance in the synthesis of the oligonucletide primers used in this study.

The results of this chapter have been published under the title "Sequence requirements for the biotinylation of carboxyl-terminal fragments of human propionyl-CoA carboxylase α subunit expressed in *E. coli*" by Alfonso Leon-Del-Rio and Roy A. Gravel (J. Biol. Chem. 269:22964-22968, 1994).

INTRODUCTION

Biotin-dependent enzymes are widely distributed among living organisms where they are involved in decarboxylation and carboxylation reactions (1). Their biotinylation occurs by the addition of a biotin molecule to a specific lysine residue on the apoenzyme. This reaction is catalyzed by biotin ligase in bacteria and by holocarboxylase synthetase (HCS) in mammalian cells (2,3). In one case, the biotin ligase has more than one function. The BirA protein of *E. coli* catalyzes the biotinylation of the biotin carboxyl carrier protein (BCCP) of the enzyme, acetyl-CoA carboxylase, and also acts as the repressor of the biotin operon (3,4). This enzyme, unlike mammalian HCS, has been purified and its gene cloned (5,6), but little is known about the way in which the biotin ligase recognizes BCCP (7,8). It has been shown that the synthetase from one organism can biotinylate apoenzymes from other sources. This cross-species activity, which includes biotinylation of bacterial apocarboxylases by mammalian HCS (9) or mammalian apocarboxylases by bacterial ligases (9,10,11), reveals a molecular mechanism common to all these enzymes.

The biotin containing proteins show a high degree of sequence conservation surrounding the biotin binding site (Fig.1). The biotinylated lysine, called biocytin, is located within the conserved tetrapeptide, Ala-Met-Lys-Met, 35 residues from the carboxyl terminus and about 39 amino acids from a Pro-X_n-Pro sequence (X_n= one to 5 amino acids involving Met, Ala, Ser, Ile, Leu, or Arg) on the amino terminal side. Also, there are several

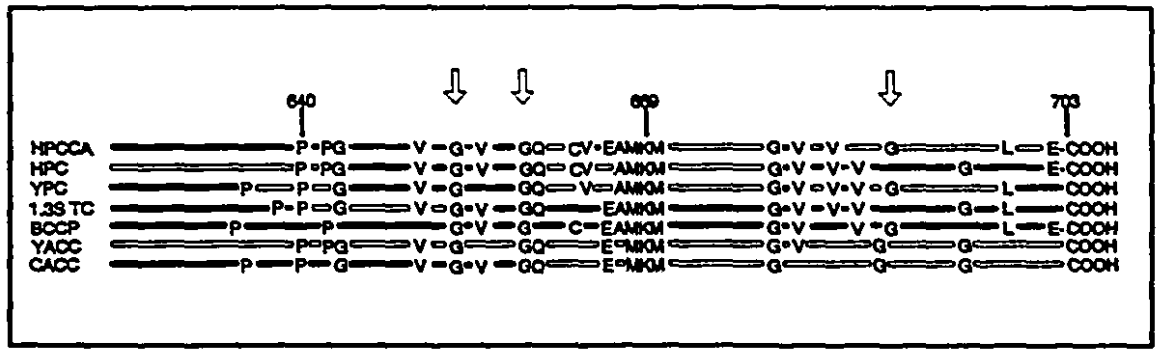


Fig. 1. **Amino acid sequence conservation among biotin enzymes.** The sequence of different biotin enzymes are aligned with respect to the biotin binding site to show the amino acid conservation around this region. Only the amino acid residues that are conserved in at least four of the seven protein fragments are shown in the figure. The numbers above the sequences represent the amino acid position in the α subunit of human PCC. Arrows indicate putative flexible points along the protein fragments. Protein sequences are as follows: HPCCA, human propionyl-CoA carboxylase; HPC, human pyruvate carboxylase; YPC, yeast pyruvate carboxylase; 1.3S TC, 1.3S subunit of *P. shermanii* transcarboxylase; BCCP, biotinyl carboxyl carrier protein of *E. coli* acetyl-CoA carboxylase; YACC, yeast acetyl-CoA carboxylase; CACC, chicken acetyl-CoA carboxylase.

conserved glycine and valine residues at identical positions in relation to the biocytin site. These residues might be involved in apoprotein recognition by the HCS or in the activation of biotin during the carboxylation reaction (12).

In vitro mutagenesis has been used to determine the role of conserved sequences on the biotinylation reaction. In the best studied case, the 1.3S subunit of *Propionibacterium shermanii* transcarboxylase, it was shown that the *E. coli* biotin ligase can biotinylate large protein fragments containing the biotin attachment site, as well as the full length protein (13). Shortening of the carboxyl terminus by two residues or deletion of sequences located within 30 residues on the amino terminal side of the biocytin site prevented biotinylation of the peptide by the *E. coli* biotin ligase (13,14). In this protein, substitution of either methionine flanking the biocytin residue, despite the conservation of these residues across species, had no evident effect on biotinylation efficiency (15). Other studies suggest a role for the methionine residues in the carboxylation reaction (16).

This investigation focuses on the identification of amino acid residues required for the recognition and biotinylation of human apocarboxylases. The apocarboxylase model used is the α subunit of human PCC (PCC-α). PCC (E.C.6.4.1.3) is a heteropolymeric protein, Mr= 560,000, with the tentative structure $\alpha_4\beta_4$ (17,18). The biotin containing α subunit contains 703 amino acids and carries a single biotin molecule at position 669, which is used by the enzyme to transfer a CO_2 molecule from bicarbonate to propionyl-CoA.

We have expressed carboxyl terminal fragments of PCC-α in *E. coli* to study the effect of peptide chain length and sequence on biotinylation by the bacterial biotin ligase. Our studies show that a carboxyl terminal peptide containing the last 67 amino acids of the α subunit of PCC carries enough information to be recognized by the biotin ligase. We also used site directed mutagenesis of the PCC-α fragment to investigate the role of specific sites in the biotinylation reaction.

MATERIALS AND METHODS

Reagents. [8,9-^{3}H]biotin (45 Ci/mmol) was purchased from Dupont, Boston, MA, and [^{14}C]biotin (57 mCi/mmol) was from Amersham, UK. Anti-flag monoclonal antibody and anti-flag affinity gel were from International Biotechnologies, Inc. New Haven, CT and the Doubletake double-stranded mutagenesis kit was from Stratagene, La Jolla, CA. Sequencing of DNA constructs was done by the dideoxy chain termination method using Sequenase from United States Biochemical Corporation, Cleveland, OH. Isopropylthio-ß-D-galactoside (IPTG) was purchased from Sigma, St. Louis, MO. All other reagents were of analytical grade.

Strains and plasmids. For plasmid propagation, we used *E. coli* strain DH5-α F'/endA1, hsdR17($r_k^-m_{k+}$), supE44, thi-1, recA1, gyrA (Nal^r), relA1, Δ(lacZYA-argF)$_{U169}$, (m80lacZΔM15). Site directed mutagenesis was performed on cDNA fragments cloned in plasmid Bluescript from Stratagene. Expression studies were done with the expression vector pFLAG-1 from International Biotechnologies Inc., The host

strain was HB101 Δ(gpt-proA)62, leuB6, thi-1, lacY1, hsdS_B.20, recA, rpsL20 (Str^r), ara-14, galK2, xyl-5, mtl-1, supE44, $mcrB_B$. Expression of unbiotinylated PCC-α fragments was done using *E. coli* C-124, generously provided by Michel Goudry (Universite Pierre et Marie Curie, France).

Construction of expression vectors encoding carboxyl terminal PCC-α fragments. Specific oligonucleotides were designed for PCR amplification of 3' terminal fragments of the full length PCC-α cDNA (19). All these fragments contained the region coding for the biotin binding site but varied in the length of their amino or carboxyl extension. The amplification primers contained new restriction sites that facilitated the subcloning of the fragments. After sequencing, the PCR products were subcloned into the expression vector pFLAG-1 which allowed their inducible expression as fusion proteins in the presence of 2 mM IPTG. These proteins contained the fusion peptide (FP), Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys, at their amino terminus. Detection and purification of fusion proteins was done using a monoclonal antibody (anti-Flag antibody) directed against the FP fragment.

Construction of mutant PCC-α sequences. cDNA fragments coding for the last 63 to 80 amino acids of PCC-α subcloned in Bluescript were subjected to site-directed mutagenesis performed by recombinant PCR (20) or using the Doubletake mutagenesis kit. The mutant constructs were sequenced and subcloned into the pFLAG-1 vector and expressed in HB101 cells as above.

Immunological detection of fusion proteins. For in vivo biotinylation of fusion peptides, *E. coli* cells were grown in L-Broth medium until an $O.D._{600}$ of 0.4 at 37°C was reached. At this point, the synthesis of fusion proteins was induced by the addition of IPTG to 2mM and growth was continued for 2 h. Ten ml aliquots were removed from the bacterial cultures and the cells were collected by centrifugation. The pellets were resuspended in 100 μl of buffer containing 50 Mm Tris pH 8.0, 1 mM EDTA, 0.1 mM DL-dithiothreitol, 2 mM phenylmethylsulfonyl fluoride and 5% glycerol. The cells were disrupted by sonication and the proteins were fractionated by electrophoresis in duplicate 12% SDS-polyacrylamide gels. After electrophoresis, the proteins of both gels were electrophoretically transferred to nitrocellulose membranes. One of the filters was incubated with anti-Flag antibody followed by incubation with goat anti-mouse IgG-alkaline phosphatase (IgG-AP) conjugate. The second membrane was incubated with streptavidin-AP to detect biotinylation of the protein fragments. Both membranes were color developed with AP substrate reagents. For in vitro biotinylation of PCC-α fragments, pFLAG-1 vectors encoding terminal fragments of PCC-α were used to transform biotin-dependent *E. coli* strain C-124. This strain contains a mutation that specifically blocks the synthesis of dethiobiotin, an essential intermediate in the production of biotin in *E. coli*. Log-phase C-124 cultures in L-Broth medium were transferred to a biotin-free medium (7.5 mM $(NH_4)_2SO_4$, 33 mM KH_2PO_4, 60 mM K_2HPO_4, 1.7 mM sodium citrate, 1 mM $MgSO_4$, 0.2% dextrose, 0.1% casamino acids, 2mM IPTG) and grown at 37°C for 2 h. The cells were pelleted and the unbiotinylated fusion proteins were purified with the anti-Flag affinity gel according to the manufacturer. After removal of the Flag peptide with enterokinase, PCC-α fragments

(1 μg) were incubated with 10-50 μl of *E. coli* HB101 cell extracts for 2 h at 37°C. The samples were loaded into 12% SDS-polyacrylamide gels and biotinylation of the fragments was determined using streptavidin-AP as above.

Time-course of [^{3}H]biotin incorporation into PCC-α fragments. *E. coli* HB101 cells expressing mutant carboxyl terminal fragments of PCC-α were grown in biotin-free medium supplemented with [^{3}H]biotin (1.0 μCi/ml) and 2 mM IPTG. The proteins were fractionated by SDS-polyacrylamide gel electrophoresis. After drying the gel, the exact position of biotinylated proteins was determined by comparison with a duplicate gel that was developed with streptavidin-AP. The regions of the gel containing the radioactive proteins were cut, dissolved in scintillation cocktail and the radioactivity incorporated was analyzed using a scintillation counter.

Quantitation of biotinylation of mutant PCC-α fragments. *E. coli* cells expressing mutant PCC-α terminal fragments were initially examined immunologically using the anti-Flag antibody/streptavidin-AP as above. This experiment allowed examination of the level of fusion protein synthesis and a crude estimate of their biotinylation. To quantitatively compare the efficiency of biotinylation, the cells were grown in a biotin-free medium supplemented with [^{14}C]biotin (0.5-1.0 μCi/ml) and 2mM IPTG. The proteins (50 μg per sample) were fractionated by SDS-polyacrylamide gel electrophoresis. After electrophoresis, the gel was dried at 80°C and exposed for 12 h to an imaging plate and the incorporated radioactivity was quantitated using a Fujix BAS 2000 Bio-Imaging Analyzer (Fuji Medical Systems USA Inc, Stamford, CT).

Biotinylation results were normalized with respect to biotinylation of a control peptide containing the unmodified carboxyl terminal sequence of PCC-α.

RESULTS

Effect of peptide chain length on biotinylation. The PCC-α mutant fusion proteins expressed in *E. coli* were visualized with the monoclonal anti-Flag antibody or streptavidin-AP to test for the presence of the FP sequence or biotin, respectively (Fig.2A-B). When the Western blot membranes were incubated with streptavidin-AP, one band of 18.4 kDa was present in all the samples (Fig.2B). This band corresponds to the BCCP of acetyl-CoA carboxylase, the only biotin dependent protein in *E. coli* (3). We used this protein as an internal control in all the expression experiments.

Representative results of the effect of protein chain length on biotinylation of PCC-α fragments are shown in Fig.2, while the structure of the PCC-α peptides expressed and summary of the data are given in Fig.3. Fusion proteins carrying the last 80 amino acids (p-80; residues 624-703) or 67 amino acids (p-67; residues 637-703) of PCC-α were efficiently biotinylated by the bacterial ligase. However, we did not find evidence of biotinylation when we expressed a peptide containing the last 57 amino acids of PCC-α (p-57; residues 647-703). The data show that sequences located upstream of Val636 are dispensable, whiie the region spanning residues 637-647 includes sequences critical for biotin ligation. Comparison of this region with sequences of other biotin-dependent proteins (Fig.1) showed that the only conserved segment is the Pro-X_n-Pro motif.

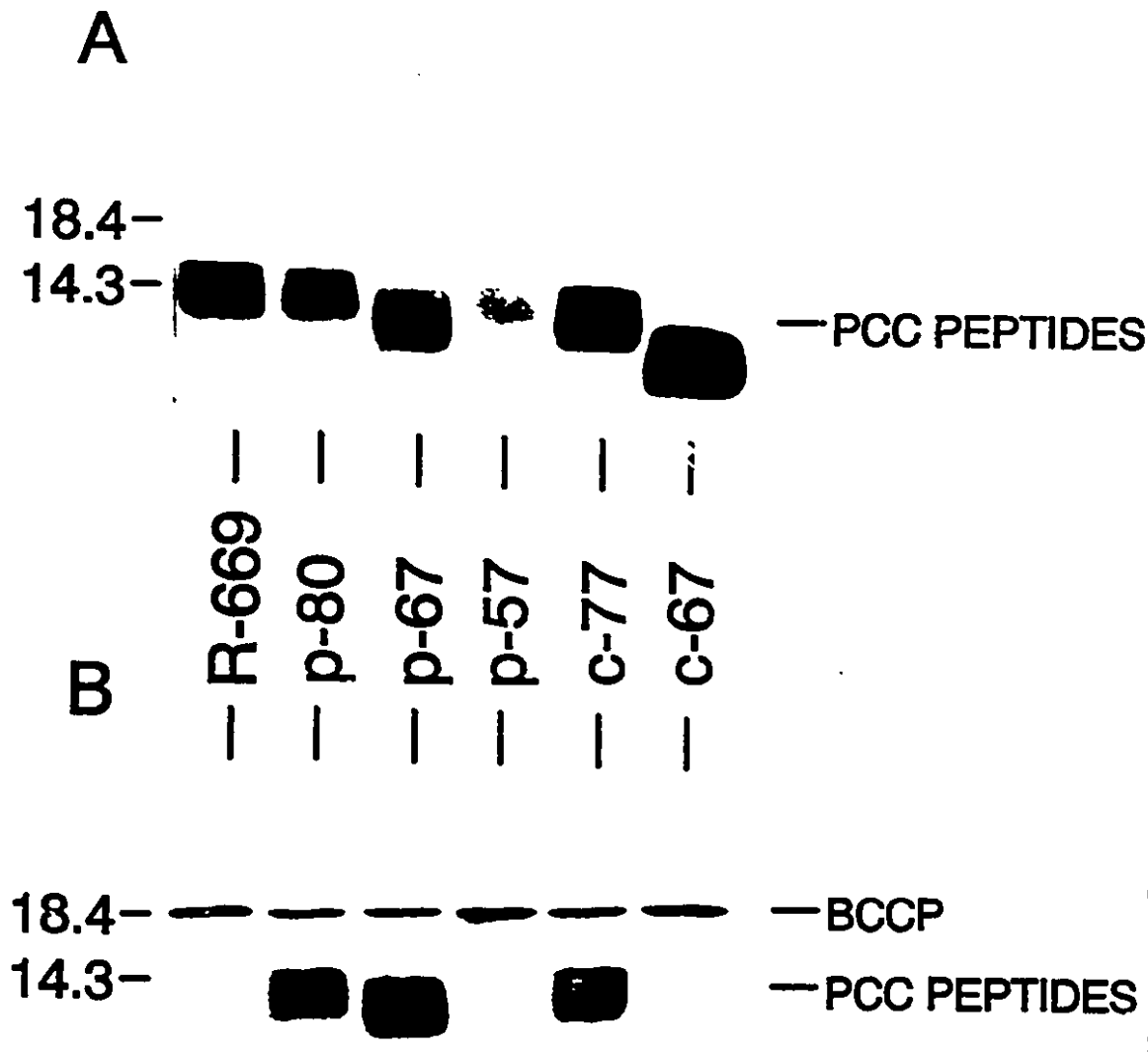


Fig. 2. **Immunoblot analysis of carboxyl terminal fragments of PCC-α expressed in *E. coli*.** Carboxyl terminal fragments of PCC-α expressed in *E. coli* were separated by SDS-polyacrylamide gel electrophoresis (12% gel). **A**, immunoblot analysis with monoclonal anti-Flag antibody to verify the expression of fusion proteins; **B**, immunoblot analysis using streptavidin-AP for detection of the biotin moiety. The numbers at the left of the blots indicate the molecular weight in kDa. Detailed procedures for sample preparation are described under "Experimental Procedures". The peptides expressed in each lane are defined in Fig.3.

A fusion protein containing the last 60 amino acids of PCC-α (p-60; residues 644-703) which lacked the proline sequence was not biotinylated (Fig.3), suggesting that the sequence, Pro-Met-Pro, is essential for biotinylation of the peptide. However, the presence of the tripeptide as a terminal sequence was not sufficient to promote biotinylation of the mutant fragments (Fig.3, p-63). To eliminate the possibilty that the lack of biotinylation of p-63 was due to the Flag peptide at the amino terminal side of the Pro-Met-Pro sequence, we studied its biotinylation in vitro after removal of the Flag octapeptide. This PCC-α fragment (Fig.3 p-63B), unlike purified p-80 or p-67, was not biotinylated by *E. coli* HB101 cell extracts under the experimental conditions tested. These results showed that the last 67 amino acids of PCC-α contain the information required to be recognized and biotinylated by the *E. coli* biotin ligase. Increasing the length of the α fragment from 67 to 80 amino acids did not affect the apparent accumulation of biotinylated protein under the experimental conditions (see below).

In order to study the effect of sequences located at the very carboxyl terminus of the peptide on biotinylation, we expressed truncated carboxyl terminal versions of the biotinylation-competent p-80 fragment (Fig.3). A peptide lacking the last 3 amino acids (c-77) was biotinylated as efficiently as the full-length 80 amino acid fragment (Fig.2). However, deletion of the last 27 amino acids (residues 677-703) abolished biotinylation of the peptide (Fig.3, c-53). The failure of the ligase to biotinylate two shorter fusion proteins lacking the last 17 or 13 amino acids (c-63 and c-67), localized sequences essential for biotinylation between residues 690-700 (c-67, illustrated in Fig.2). In contrast to the amino terminal truncated versions of the p-80

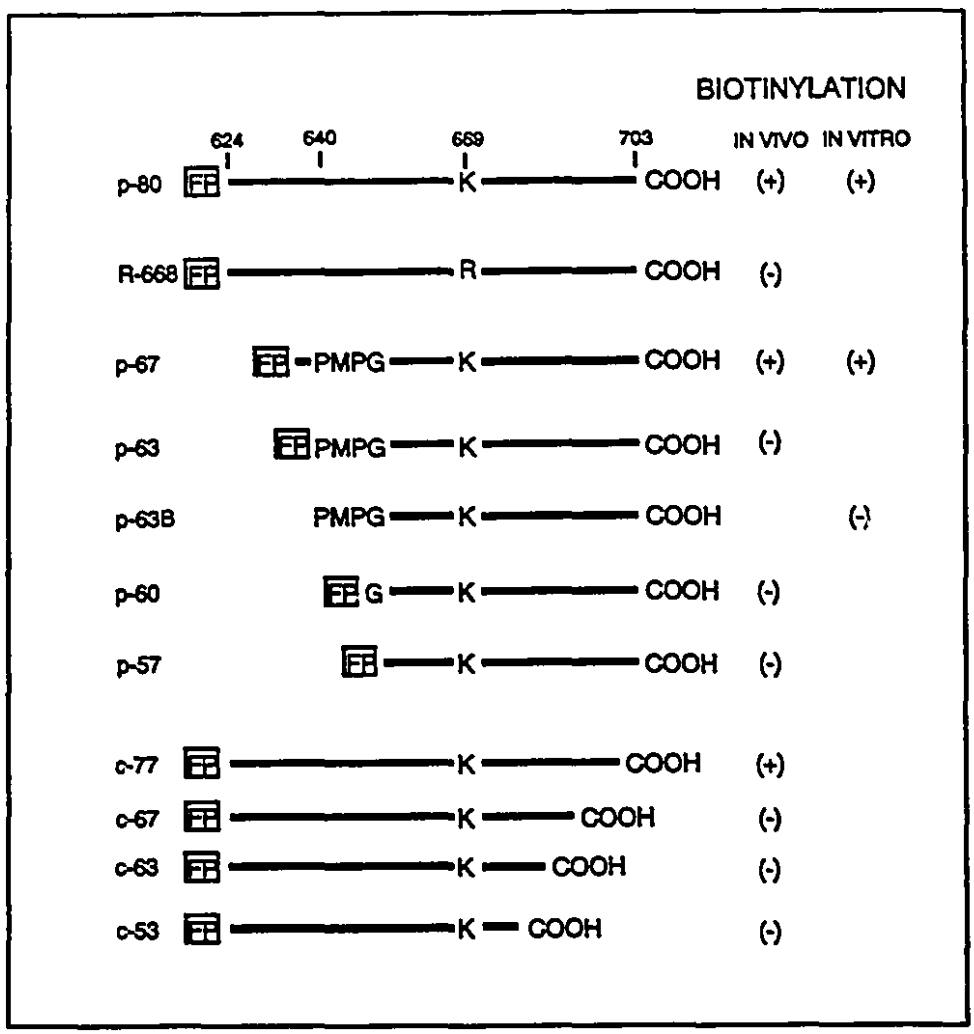


Fig. 3. **Effect of amino acid chain length on biotinylation of PCC-α peptides.** Amino terminal or carboxyl terminal truncated versions of p-80 were expressed in *E. coli* and their biotinylation studied as described under "Experimental Procedures". The position of the fusion peptide (FP), the sequence Pro-Met-Pro and the biotin binding site are indicated for each fragment. A plus or a minus sign, under the biotinylation column, indicates when a fragment was or was not biotinylated by *E. coli* (lane completely blank when blots were probed with streptavidin-AP). In vitro biotinylation studies were carried out using purified fusion proteins incubated with E. coli HB101 cell extracts. The numbers in the names of the PCC-α fragments indicate their length in amino acids. The numbers at the top of the figure indicate the amino acid position in the full length PCC-α.

fragment, comparison of the sequence 690-700 with the carboxylase consensus sequence revealed no extensively conserved residues at these positions. These experiments and the results obtained from peptide p-67 showed that the minimum α-PCC fragment that can act as a substrate for the biotin ligase is between 64 and 67 amino acids in length.

Time course of biotinylation of PCC-α fragments. *E. coli* cells carrying the constructs coding for the p-80 or p-67 fragments were grown in the presence of [^{3}H]biotin and 2mM IPTG for 1 to 22 hrs (Fig.4). Maximum biotin incorporation by the two protein fragments occurred at 2 h after IPTG induction, being higher for the p-67 than the p-80 fragment. After 22 h the amount of accumulated p-67 was also higher. When we quantitated the accumulation of biotin, as [^{14}C]biotin, in p-67 and p-80 using a phosphor image analyzer, we found that the biotin incorporation by p-67 was 33% higher than that observed by p-80 (Fig.5). Our results suggest that the smaller p-67 peptide is a better substrate for the biotin ligase or is more stable within the cell. During the period tested the biotinylation of the BCCP remained constant, demonstrating the specificity of the results.

Effect of amino acid substitution on biotinylation. Based on the above experiments, the p-67 fragment was chosen for study of the effect of amino acid substitutions on biotinylation. The targets for site-directed mutagenesis were Lys669 (the biocytin site), Met668, Met670, Pro640, Met641, Pro642, Gly653, Cys663 and Gly680 (Fig.5). These amino acids were selected due to their high conservation among biotin enzymes and on the chain length experiments described above. Immunological detection of the mutant

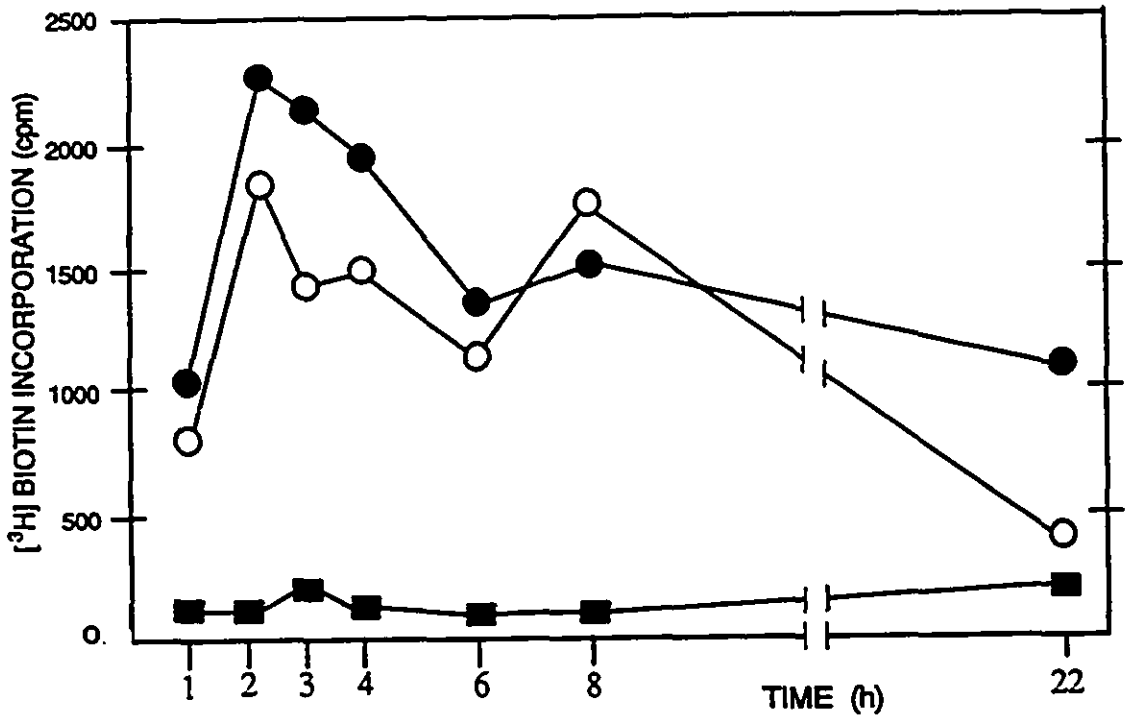


Fig. 4. **Time course of *in vivo* biotinylation of PCC-α peptides.** *E. coli* cells, expressing PCC-α peptides p-67 O, or p-80 O, were incubated in the presence of [^{3}H]biotin for different times. Crude extracts from these cultures were analyzed by SDS-polyacrylamide gel electrophoresis and the biotinylation of the peptides was quantitated as described under "Experimental Procedures" using a scintillation counter. Biotinylation of BCCP ■ was used as an internal control for every experiment. The experiment was done on two independent occasions; The data shown are from one of the experiments.

fusion proteins with anti-flag antibody (data not shown), showed that the amino acid substitutions did not affect the amount of mutant peptide synthesized by *E. coli* compared with the unmutated p-67 control. Substitution of Lys669 by Arg abolished biotinylation completely (Fig.5, R-669). This mutant peptide was used as a negative control in all the mutagenesis experiments. Substitution of Met668 by Leu had no apparent effect on the biotinylation of the mutant peptide; however, substitution of Met670 by Leu increased biotinylation by 36% (Fig.5, L-K-M and M-K-L, respectively). Substitution of Cys663, located only five residues upstream of the modified lysine, by Trp, reduced the amount of biotinylated peptide by 62% (Trp663).

From the results of the expression of amino terminal truncated versions of p-80, it seemed that the sequence Pro-Met-Pro at position 640 might be important for biotinylation of the peptides, even though a PCC-α fragment carrying the tripeptide as the amino terminal sequence failed to be biotinylated (Fig.3, p-63). To explore this matter further, we evaluated derivatives of p-67 containing mutations in this region (Fig.5). Substitution of the sequence, Pro-Met-Pro, by the tripeptide Ala-Ala-Ala (A-A-A in Fig.5) caused a decrease of 46% in the biotinylation of the peptide. A greater decrease (76%) was observed when the hydrophobic methionine residue of the Pro sequence was substituted by lysine (P-K-P). Finally, deletion of the entire Pro-Met-Pro sequence with retention of upstream residues abolished biotinylation of the mutant peptide (Fig.5 Δ-Δ-Δ).

In a different set of experiments, we focused on the role of some of the conserved

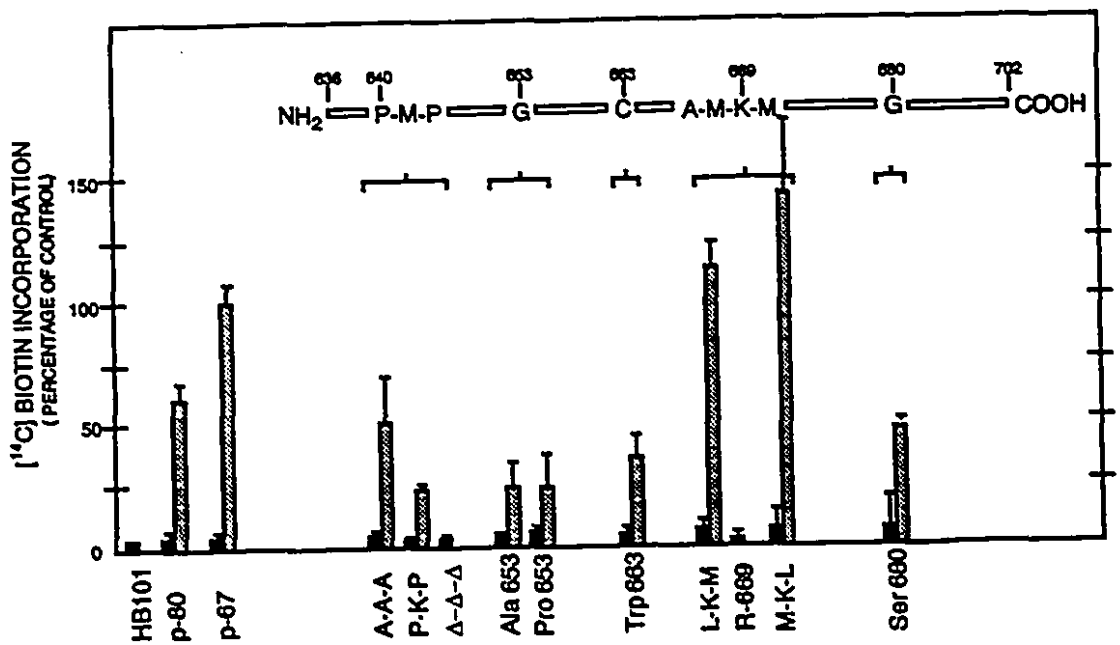


Fig. 5. **Effect of amino acid substitution on biotinylation of PCC-α peptides.** Cultured *E. coli* cells expressing mutant versions of p-67 were labelled with [^{14}C]biotin. Crude extracts from these cultures were analyzed by SDS-polyacrylamide gel electrophoresis and biotinylation of the mutant peptides was quantitated using a phosphor image analyzer as described under "Experimental Procedures". The target amino acids for site directed mutagenesis, and their position in PCC-α, are shown in the schematic representation at the top of the figure. The brackets below the targeted amino acids indicate the specific group of results obtained from their substitution. The nature of the amino acid substitutions introduced into p-67 are indicated in the bottom of the graph and described in the text. HB101 represents biotin incorporation by *E. coli* HB101 cells that were not transformed with any expression vector; p-80 and p-67 represent the biotinylation of fusion proteins containing the unaltered last 80 and 67 amino acids of of PCC-α, respectively. Results are expressed as percent of control (p-67) biotinylation. Black bars represent biotinylation of BCCP, grey bars represent biotinylation of PCC-α peptides. The grey bar value for Δ-Δ-Δ and R-668 is "0". Each bar represents the mean ± S.E.M. of three experiments (two in the case of Ala-652 and Trp-662).

glycine residues on biotinylation. In this respect, an *E. coli* strain containing a mutation changing Gly100 (position 680 in Fig.1) to Ser in the BCCP subunit of ACC is responsible for deficient ACC activity when the cells are grown at 42°C (21). The mutation was also found to cause a less efficient biotinylation of the BCCP subunit (22). To test the importance of this residue in the human enzyme, we reproduced the *E. coli* mutation at the same position on p-67, changing Gly680 to Ser. We found a 53% reduction in the biotinylation of the mutant peptide compared with the original p-67 fragment (Fig.5, Ser680). However, when the cells expressing the mutant peptide were grown at 42°C we did not notice any further decrease in the biotinylation of the fragment (data not shown).

Next, we mutated another highly conserved glycine residue located at position 653 on PCC-α (Fig.1). This amino acid was of particular interest because it is located between the proline-rich region and the biotin binding site. Two different changes were introduced at this position. The first one was a conservative change that substituted Gly to Ala (Ala633). The second change substituted Gly to Pro (Pro633), a change that was predicted to affect the secondary structure of the fragment by disrupting the putative conformation of this region. Both mutations reduced the biotinylation of the fragment by 74% (Fig.5).

DISCUSSION

The successful *in vivo* biotinylation in *E. coli* of carboxyl terminal fragments of human PCC-α subunit shows that the carboxyl terminus of this protein can function as an independent domain. The required sequence is defined on the amino terminal side by the tripeptide, Pro640-Met-Pro, and on the carboxyl terminal side, very close to the C-terminus. The shortest peptide that supported biotinylation was 67 amino acids long. This compares favorably to the 75 amino acids required to support biotinylation of the 1.3S subunit of *P. shermanii* (13). These are much larger than the 13 amino acid peptides that can be biotinylated when selected from a synthetic peptide library (23). The size of the latter structures maybe related to the minimum of constrains associated with a requirement only for ligand binding in the peptides (23).

The lack of biotinylation of the apocarboxylase fragments in which the Pro-Met-Pro tripeptide was deleted suggests that it may be required for the recognition and biotinylation of the peptides by the biotin ligase. However, it is possible that the tripeptide might be only the most characteristic feature of a longer recognition sequence, since peptides carrying the tripeptide at their very amino terminus were not biotinylated. If so, it may be that the ligase recognizes this sequence only when it is located within a particular amino acid environment and not when it is situated at the end of the protein. Alternatively, the Pro-Met-Pro sequence may act only to add length to an extended biotinylation domain. We believe this to be unlikely because substitution of the Pro-Met-Pro sequence by three consecutive Ala residues, which preserved the original peptide

length, reduced the biotinylation of the fragment. Similarly, substitution of the central Met641 residue by Lys, while leaving the prolines intact, also reduced biotinylation suggesting that the non-polar hydrophobic nature of the tripeptide may play a role in the recognition of the apoenzymes by the ligase.

It has been suggested that the sequence Pro-X_n-Pro might turn the carboxyl terminus of biotin dependent enzymes into a flexible "extended arm" allowing biotin to oscillate between donor and acceptor sites of carboxylation (12). This hypothesis was explored using a computer program (24) to predict the amino acid chain flexibility for all the carboxylase sequences shown in Fig.1. The analysis showed the biotin binding site to be located between three putative flexible points in all the proteins (Fig.1, arrows). These flexible regions are located at or near one of the highly conserved Gly residues and not on the Pro-Met-Pro sequence. Although these results do not exclude the involvement of the tripeptide on chain flexibility, we believe that this sequence may have a dual role: (a) due to the tendency of Pro residues to bend protein chains, the sequence Pro-Met-Pro might act to isolate the carboxyl terminus of biotin enzymes from the rest of the protein, turning this region into an independent domain; and (b) it may be required by the biotin ligase to identify the lysine residue that has to be biotinylated.

On the carboxyl terminal side of the domain, mutations that removed up to three amino acids from the carboxyl terminus of p-80 did not affect the efficiency of biotinylation. However, longer deletions at the carboxyl terminus of p-80 demonstrated that the amino acids located between residues 690-700 in PCC-α are necessary for modification of the

apopeptide. Our results differ significantly from those obtained with the 1.3S subunit from *P. shermanii* expressed in *E. coli*. In this protein, the penultimate amino acid is essential for biotinylation (14). Although comparison of the sequence 690-700 with the analogous region in other biotin enzymes shows no amino acid sequence conservation, this region is rich in hydrophobic residues in all these enzymes. It is possible that the hydrophobic domain is important in preserving the secondary structure of the peptide or it is involved in the interaction with the biotin-ligase.

The importance of the conserved glycine residues on the biotinylation or catalytic activity of biotin-dependent enzymes may be illustrated by the substitution of Gly680 to Ser. In this case the peptide mimics a mutation in the BCCP subunit that results in temperature-sensitive growth dependence of the mutant *E. coli* (21). When we expressed the mutation in the PCC-α fragment, we found a reduction in its biotinylation by the biotin ligase. When Gly653, another highly conserved amino acid, was mutated, we also found a strong decrease on biotinylation of mutant peptides. Unexpectedly, the effect on biotinylation of the mutant fragments was the same whether the mutation introduced a conservative amino acid, substitution to alanine, or introduced a change that disrupted the putative secondary structure, substitution to proline. It is possible that these residues are recognition signals for the ligase or they might be important for the correct folding of the biotin domains.

For years, researchers have speculated that the absolute conservation of the tripeptide Met-Lys-Met among biotin enzymes implies a functional role (12). It has been suggested

that this sequence might be important for the identification of the Lys residue to be modified by the ligase or for the activation of biotin in the carboxylation reaction (12,15,16). Our results showed that the conserved Met residues are not necessary for biotinylation of PCC-α peptides confirming the findings reported by Shenoy et al (15) for the 1.3S subunit from *P. shermanii* transcarboxylase. In this latter protein, substitution of the Met residues had no effect on biotinylation. However, the mutations resulted in deficient biotin carboxylation (25). Computer modelling of the structure of yeast PC (26) suggests that the biotin domain forms a ß-hairpin loop containing a single ß-turn in which the biotin binding site is located. This model proposes that the biotin binding site is isolated and is not important for the folding of the rest of the domain. This might explain why substitution of the Met residues in the tetrapeptide, Ala-Met-Lys-Met, had no effect on biotinylation of PCC-α fragments.

REFERENCES

1. Wood H.G. (1977). Ann. Rev. Biochem. 46, 385-413
2. Burri, B.J.,Sweetman, L. and Nyhan, W.L. (1981). J. Clin. Invest. 68, 1491-1495
3. Cronan, J.E. (1989) Cell 58, 427-429
4. Cronan, J.E. (1988) J. Mol. Chem. 263, 10332-10336
5. Eisenberg, M.A., Prakash, O. and Hsiung, S.C. (1982) J. Biol. Chem. 257, 15167-15173
6. Howard, P.K., Shaw, J. and Otsuka, A.J. (1985) Gene 35, 321-331
7. Buoncristiani, M.R., Howard, P.K. and Otsuka A.J. (1986) Gene 44, 255-261
8. Wilson, K.P., Shewchuk, L.M., Brennan, R.G., Otsuka, A.J. and Matthews, B.W. (1992) Proc. Natl. Acad. Sci. USA 89, 9257-9261
9. McAllister, H.C. and Coon, M.J. (1966) J. Biol. Chem. 241, 2855-2861
10. Kosow, D.P., Huang, S.C. and Lane, M.D. (1962) J. Biol. Chem. 237, 3633-3639
11. Lane, M.D., Young, D.L. and Lynen, F. (1964) J. Biol. Chem. 239, 2858-2864
12. Samols, D., Thornton, C.G., Murtif, V.L., Kumar, G.K., Haase, F.C. and Wood, H.G. (1988) J. Biol. Chem. 263, 6461-6464
13. Cronan, J.E. (1990) J. Biol Chem. 265, 10327-10333
14. Murtif, V.L., and Samols. D. (1987) J. Biol. Chem. 262, 11813-11816
15. Shenoy, B.C., Paranjape, S., Murtif, V.L., Kumar, G.K., Samols, D. and Wood, H.G. (1988) FASEB J. 2, 2505-2511

16. Kondo, H., Uno, S., Komizo, Y., Sunamoto, J. (1984) J. Biol. Chem. 23, 559-564.

17. Gravel, R.A., Lam, K.F., Mahuran, D., Kronis, A. (1983) Arch. Biochem. Biophys. 201, 669-673

18. Kalousek, F., Darigo, M.D., Rosenberg, L.E. (1980) J. Biol. Chem. 255, 60-65

19. Lamhonwah, A.M., Mahuran, D. and Gravel, R.A. (1989) Nuc. Acids Res. 17, 4396

20. Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (1990) PCR Protocols, Academic Press, Inc., San Diego, California.

21. Silbert, D.F., Pohlman, T., and Chapman, A. (1976) J. Bacteriol. 126, 1351-1354

22. Li., S.J. and Cronan, J.E. (1992) J. Biol. Chem. 267, 855-863

23. Schatz, P.J. (1993) Bio/Technology 11, 1138-1143

24. Karplus, P.A., Schulz, G.E. (1985) Natrwissenchaften 72, 212-213

25. Wood, H.G., Shenoy, B.C., Kumar, K.G., Paranjape, S., Murtif, V.L. and Samols, D. (1987) Federation Proc. 46, 2124

26. Brocklehurst S.M. and Perham, R.N. (1993) Protein Science 2, 626-639

CHAPTER 3

ISOLATION OF A cDNA ENCODING HUMAN HOLOCARBOXYLASE SYNTHETASE BY FUNCTIONAL COMPLEMENTATION OF A BIOTIN-AUXOTROPH OF *Escherichia Coli*

PREFACE TO CHAPTER 3

In the previous chapter, I demonstrated the autonomy of the biotin domain of human PCC-α as a substrate for biotinylation by the biotin ligase of *E. coli*. This result suggested the possibility of cloning the human HCS gene by performing the reciprocal experiment: using the human synthetase to biotinylate the bacterial carboxylase from a biotin ligase-deficient *E. coli* strain. These cells resemble the phenotype of patients with neonatal MCD by requiring of high doses of exogenous biotin to grow. The success of this approach required that human HCS take over the role of the mutant biotin ligase and rescue the phenotype of this bacterial strain. Coexpression of fragments of human PCC-α in survivor *E. coli* cells would allow confirmation of the enzymatic activity encoded by the candidate clones.

I would like to acknowledge the contributions of the following persons to the realization of this study: Dr. Daniel Leclerc for his numerous comments and for his help with the computer analysis of the HCS sequence, Dr. Roy A. Gravel for helpful discussions and for proof-reading sequencing gels, Dr. Nobuaki Wakamatsu who isolated a cosmid clone containing the putative 5' end of the HCS gene and Ms. Beverly Akerman who confirmed the sequence of the alternatively spliced regions of HCS mRNA in genomic DNA and cDNA by PCR and RT-PCR, respectively. Finally, I would like to thank Dr. Hai Shiene Chen, head of the Sequencing Core Facility of the Canadian Genetic Diseases Network at the Hospital for Sick Children in Toronto for invaluable sequencing assistance.

The results of this chapter have been published under the title "Isolation of a cDNA encoding human holocarboxylase synthetase by functional complementation of a biotin auxotroph of *Escherichia coli*" by Alfonso Leon Del Rio, Daniel Leclerc, Beverly Akerman, Nobuaki Wakamatsu and Roy A. Gravel (Proc. Natl. Acad. Sci. USA, (1995) 92:4626-4630).

INTRODUCTION

Holocarboxylase synthetase (HCS) catalyzes the biotinylation of three mitochondrial and one cytoplasmic biotin-dependent carboxylase in mammalian cells (1). In contrast, *E. 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 Lys residue located within a highly conserved sequence, Met-Lys-Met, present in all biotin-dependent 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 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 of a cDNA library in *birA* cells. We report here the isolation of human HCS cDNA clones capable of replacing the biotin ligase function of *E. coli*.

MATERIALS AND METHODS.

Materials. Radiolabeled compounds were from Dupont. A human multiple tissue Northern blot and ß-actin cDNA probe kit were from Clontech. cDNA probes were labeled using a random priming kit from Boehringer-Mannheim. Oligonucleotides were synthesized in-house by R. Clarizio. Sequenase was from United States Biochemicals. Temperature-sensitive *E. coli birA* mutants were kindly provided by A.M. Campbell (16,17) and A.J. Otsuka. A 67 amino acid carboxyl terminal fragment (p-67) of the α subunit of human propionyl-CoA carboxylase (PCC-α), in the vector pFLAG67, was used as substrate for HCS (20). A human cDNA library, λYES, from EB virus-transformed B-lymphocytes and *E. coli* strain BNN132 containing λKC were generously provided by S.J. Elledge (21) and E. McIntosh. A cosmid genomic DNA library, in pWE15 and made from human lymphocyte DNA, was from Stratagene.

Cloning of HCS cDNA and genomic DNA clones. The *birA104* mutant was lysogenized with a helper phage (λKC) harbouring the kanamycin-resistance and *cre* genes (*kan-cre*). Kanamycin-resistant colonies showed wild-type growth at 30°C in LB plates with 5 μM biotin but no growth at 42°C. The resulting *E. coli birA104*(λKC) was infected with a human cDNA library in λYES as described (21) and incubated for 2 h at 30°C in non-selective LB medium (5 μM biotin, 2 mM IPTG). The cells were collected by centrifugation at 1500xg, washed twice with LB medium to eliminate biotin and plated in selective LB agar medium (ampicillin, 50 μg per ml; no added biotin) at 30°C for 24 h. 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 *EcoR*I site of Bluescript for sequencing. The cosmid genomic DNA library was used for the isolation of clones containing the 5'-end of the *HCS* gene using a 5'-end, 366 bp *Kpn*I fragment of cDNA clone BL-11 as probe.

Functional characterization of HCS cDNA. The candidate HCS cDNA clones in pSE936 were used to re-transform *birA104*. The cells were plated in biotin-restricted LB agar medium supplemented with 2mM IPTG and evaluated for growth after overnight incubation. To assess the HCS activity, individual colonies were grown to OD_{600} of 0.4-0.5 in LB liquid medium and transferred to biotin-free minimal medium (A-medium (22)) with 2mM IPTG at 30°C or 42°C for 10 min. [^{14}C]biotin, 1 μCi/ml, was added to the cultures, and the incubation was continued for 2 h. Crude cell extracts were prepared and their proteins were separated by electrophoresis in 12% polyacrylamide gels. The proteins were transferred to nitrocellulose membranes and the radioactivity was quantitated using a phosphor image analyzer (FUJI-BAS 2000, Fuji). In some experiments, *birA104* was transformed with pFLAG67, with or without prior transformation with candidate cDNAs, to test the HCS activity of cDNA clones on p-67. Detection of the radiolabelled peptide was determined after electrophoresis as above.

Analysis of RNA. The multiple tissue Northern blot prepared from poly(A^+) RNA (2 μg 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 ß-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 (RT-PCR). Total RNA was isolated from fibroblasts and reverse-transcribed using oligo-dT_{15} 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. They are 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*(λKC) cells were infected with 25 x 10^6 λYES bacteriophage expressing a human lymphocyte cDNA library and plated in selective LB plates without added biotin. After 24 hr 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*KI, which flanked the inserts, and *Kpn*I. Four classes of cDNA inserts were identified of 4.0 kb (3 clones), 3.0 kb (8 clones), 2.2 kb (5 clones) and 2.0 kb (5 clones) in length (Fig. 1A). 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 shared the same internal sequence but were 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 amino acids, M_r=80,759 (Fig. 2). The first in-

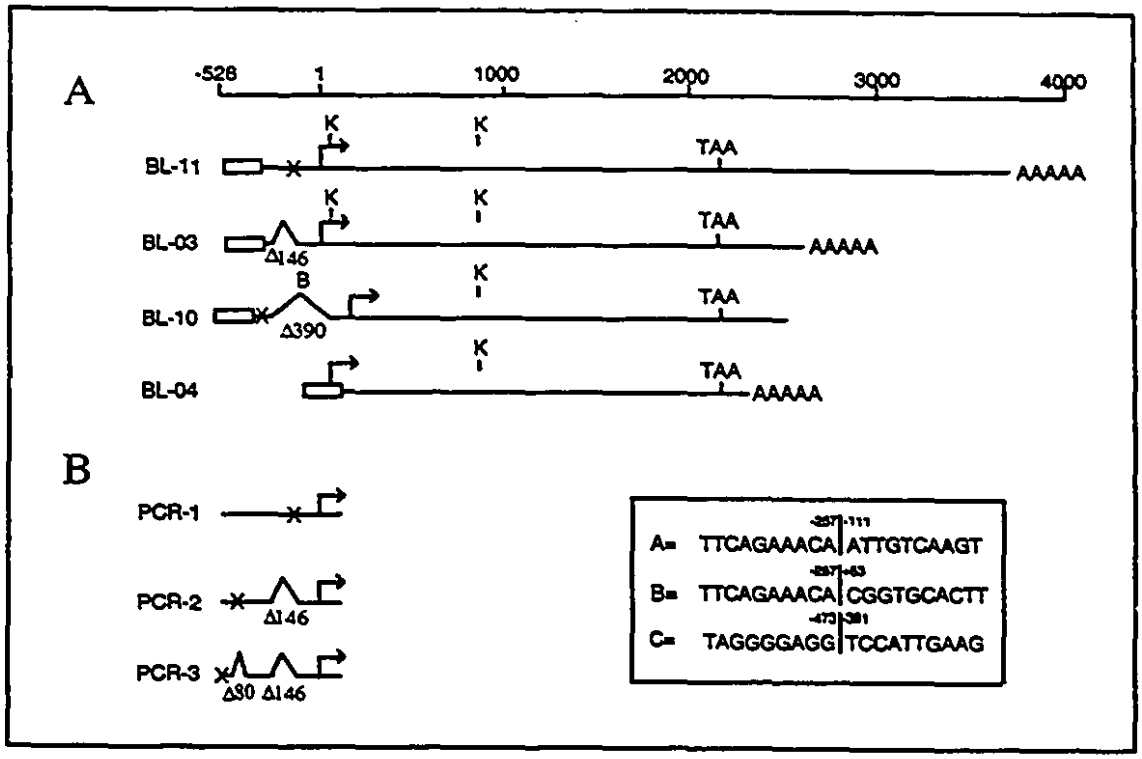


Fig. 1. HCS cDNA classes isolated by complementation of *E. coli birA104*. Alternatively spliced regions relative to BL-11 are indicated as Δ146 and Δ309 (A). In-frame ATG codons suitable for initiation of translation are indicated by arrows. Location of stop codons, polyadenylation tails and *Kpn*I restriction sites used for determining cDNA classes are shown. Nonsense codons are represented by an 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 RT-PCR products obtained using primers S1-A1. Alternatively spliced regions relative to RP-1 are shown as Δ81, Δ146 and Δ309. Arrows indicate the position of the first ATG codon in relation to BL-11. The sequence of the three splice junctions detected in HCS cDNA is shown in the inset.

frame ATG is preceded by a nonsense codon at position -123. At the 3' end, there is an additional ~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 146 or 309 bp sequence (Fig. 1A, Δ150 and Δ290) beginning at -257 compared to 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 polyadenylation 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 preceeds the ATG (BL-11 and BL-10) or 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 (nine amino acids). This cDNA begins at +165 relative to BL-11 and the first ATG encoded by the cDNA is located at +172 (Met58).

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 [^{14}C]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* transformed with any of the cDNA clones could

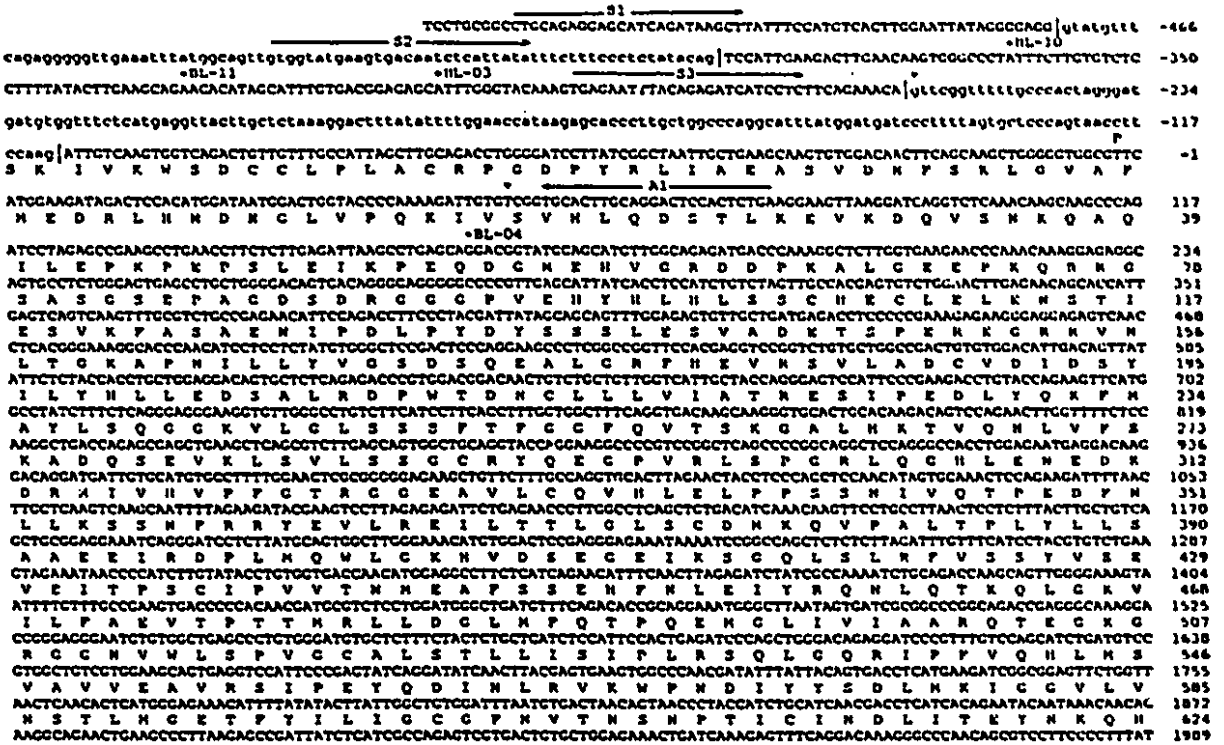


Fig. 2. **DNA sequence and predicted amino acid sequence of human HCS cDNA.** The sequence shown was derived from BL-11 and RP-1 (Fig. 1). Alternatively spliced out sequences Δ81 and Δ146 are shown in small characters flanked by vertical lines. Arrowheads indicate the spliced out sequence in BL-10 (Δ309). The 5'end of each cDNA class is indicated by a filled circle. Oligonucleotide primers used in RT-PCR experiments are overlined above the nucleotide sequence. Numbering at the right of the figure gives positions in bp or amino acid residues accordingly, with counting starting at the first in-frame ATG. Stop codons are indicated by ***.

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* was also unable to carry out the biotinylation of the p-67 peptide at 42°C. On the other hand, when *birA104* cells were co-transformed 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. 1A) 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), with the strongest signals apparent in skeletal muscle, kidney and pancreas. Two fainter bands, possibly degradation products, of 4.6 and 4.0 kb were detected in most lanes. A 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, an RT-PCR experiment was done on fibroblast RNA in which the oligonucleotide primers (S3 and A1) were positioned flanking Δ309, the larger of the two deletions (Fig. 2). Two PCR products were obtained and sequenced. The longer matched BL-11, while the shorter contained a deletion of 146 bp corresponding exactly to the excised sequence of BL-03 (Fig. 1A; Δ146). The Δ309 deletion was not detected in the fibroblast RNA.

Fig. 3. **In vivo [14C]biotin incorporation into BCCP and p-67 by BirA and human HCS in *E. coli birA104*.** The lane labelled BirA104 shows *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 the human PCC-α subunit; +p-67+BL-11, *birA104* cells co-transformed with pFLAG67 and pSE936BL-11. The number under the brackets indicates the temperature used during biotin incorporation by the bacterial cultures, 30°C or 42°C. Detection of the radiolabel in BCCP or p-67 was by overnight exposure of nitrocellulose blots to a phosphor image analyzer. The data are for a representative experiment. In arbitrary units, the band densities for BCCP are (left to right): 344, 12; 465, 290; 32, 3; 278, 101. For p-67, they are 23, 0; 36, 24; 93, 11; 1908, 808.

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 sequences were detected in cDNA libraries. However, we obtained additional 5' genomic DNA sequences using the 5' *EcoRI-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 *EcoR*I 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 Δ146 and Δ309 deletions, the genomic clone continued into an intron with the splice donor sequence TTCAGAAACA | gtgagtacca (| = splice site).

To determine if any additional sequence 5' of -363 was represented in HCS mRNA, RT-PCR experiments were done on fibroblast RNA using oligonucleotides as primers that bridged sequences in the genomic clone and the 3' side of Δ309 (Fig. 1B). Three oligonucleotides (S1-S3 used with A1), which extended the transcribed sequence at least to -528, gave PCR products with longer 5'-end 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, Δ146 and another 81 bp in length (Δ81) beginning at position -473. The sequence of Δ81 has the structure of a conventional intron with consensus donor and acceptor splice motifs (Fig. 2).

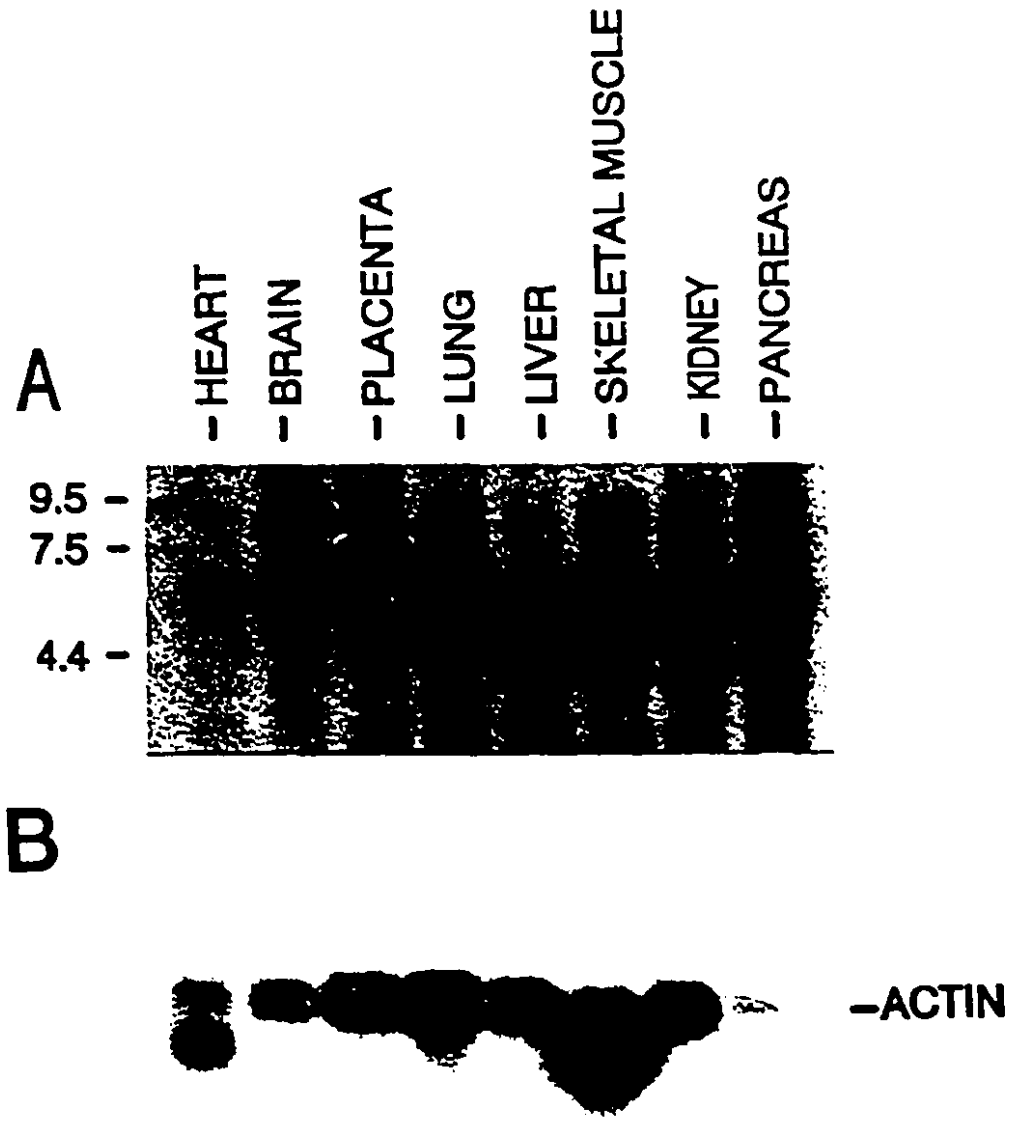


Fig. 4. **Northern analysis of HCS mRNA.** A Northern blot of poly(A+) RNA from the indicated human tissues (2 μg per lane) was hybridized with a KpnI-generated 837 bp fragment of clone BL-11 (A) or with human ß-actin (B). The position of the molecular weight markers is indicated at the left of the figure.

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-α co-expressed in the bacterial strain. During revision of this manuscript, a similar cDNA was reported by Susuki et al. (24) using oligonucleotides deduced from peptides determined from bovine HCS. It has a structure similar to clone BL-11 reported here.

Previous studies have demonstrated HCS activity in both the mitochondria and 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 Susuki et al. (24) was cloned from sequences derived from the cytoplasmic form of HCS. Yet, they identified mutations, one of them a frame-shift, in a patient with multiple carboxylase deficiency. These data point to a single HCS 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 through 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 N-terminal sequences into the encoded 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 (Met1 and Met58) among the four classes of cDNAs. These sites could represent the mitochondrial and cytosolic translation initiation sites, respectively. Inspection of the predicted N-terminal sequence defined by Met1 (or Met58) 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 Met1 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 a yet unidentified upstream sequence could introduce the required mitochondrial import sequence 5' of the currently designated Met1. When the predicted protein sequence upstream of Met1 was evaluated for a potential mitochondrial pre-sequence using the subroutine TRANSPEP from the program PCGENE (Intelligenetics, Inc), the sequence composed of PSKIVKWSDCCLPLACRPG (encoded by nucleotides -120 to -64) was identified. However, the sequence lacks a start

ATG and is preceded by several nonsense codons, the first at -123. The alternate 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 Met residue. Therefore, validation of this model would require the introduction of a still unidentified exon containing the N-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 pre-sequence, 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 RT-PCR sequences are taken into account, there remains over 1 kb to be accounted for when compared to the 5.8 kb major mRNA species detected by Northern blotting. The missing sequence may be present at either end. While we suggest that additional 5' sequences may be required, the presence of several poly-adenylation sites and the absence of consensus poly-adenylation signals also suggest that a poly-adenylation signal could remain to be identified farther downstream. Resolution of the complex relationship between the cDNA and RNA stuctures will require the analysis of a more abundantly expressed tissue 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).

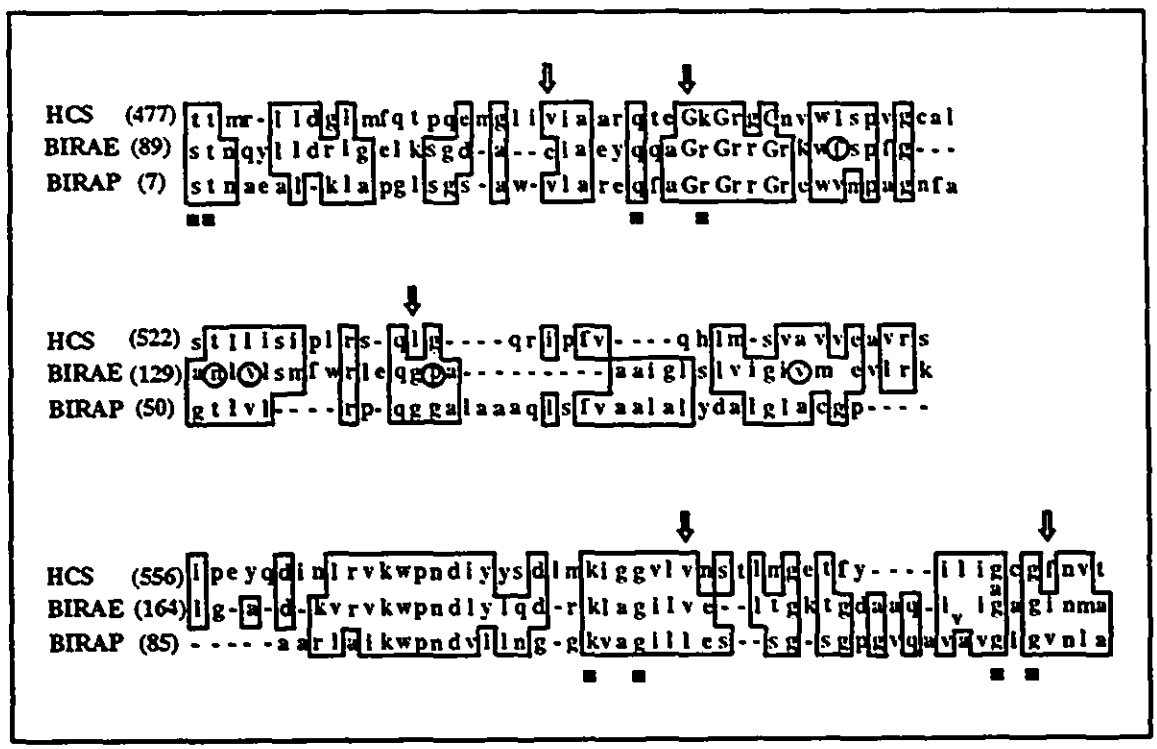


Fig. 5. **Amino acid sequence comparison among biotin ligases.** HCS, human holocarboxylase synthetase; BIRAE, *E. coli* BirA protein; BIRAP, *P. denitrificans* putative biotin-ligase. Conserved and similar amino acid residues (A,G; S,T; D,E; N,Q; V,L,I,M; K,R; F,W,Y) are enclosed in boxes. The position of amino acids that contact biotin in BirA (■), as determined by crystallographic analysis (35), is indicated under the protein sequences. The localization of mutations associated with an increase in the km for biotin in *E. coli birA* strains *birA815*, *birA1*, *birA215* and *birA104* (in order from the N-terminus) are indicated by arrows (13-15). Amino acid position for each protein is shown at the left of the figure.

The similarities with HCS are restricted to a region of the *E. coli* BirA shown by X-ray crystallography to contain the biotin binding site (35). The two proteins share 36% identity across a 129 amino acid region, with 6 of the 8 residues involved in direct contact with biotin in BirA identical in human HCS and the remaining two showing conservative changes (Thr477 to Ser; Lys506 to Arg). Several bacterial mutations associated with an increase in Km 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 also contains a consensus GXGXXG sequence at amino acid residues 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, as suggested (35), that this might reflect 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.

Acknowledgments. 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 λYES, 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. ALDR is recipient of a scholarship from the Programa Universitario de Investigación en Salud de la Universidad Nacional Autonoma de México.

REFERENCES

1. Wood, H. G. & Barden, R. E. (1977) *Annu. Rev. Biochem.* 46, 385-413.

2. Li, S. J. & Cronan, J. E., Jr. (1992) *J. Biol. Chem.* **267**, 855-863.

3. 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.

5. 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* **1**, 1187-1188.

7. 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.

8. 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.

10. 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.

15. Buoncristiani, M. R., Howard, P. K. & Otsuka, A. J. (1986) *Gene* **44**, 255-261.

16. 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., Susuki, Y. & Narisawa, K. (1991) *Jpn. Soc. Inherit. Metab. Dis.* 183.

19. Chiba, Y., Susuki, Y., Aoki, Y., Ishida, Y. & Narisawa, K (1994) *Arch. Biochem. Biophys.* **313**, 8-14.

20. Leon-Del-Rio, A. & Gravel, R. A. (1994) *J. Biol. Chem.* **269**, 22964-22968.

21. Elledge, S. J., Mulligan, J. T., Ramer, S. W., Spottswood, M. & Davis, R. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1731-1735.

22. 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 Publishing Associates and Wiley-Interscience, New York), pp. 1.1.1-1.1.2.

23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning - A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), p. 7.12-7.15.

24. Susuki, Y., Aoki, Y., Ishida, Y., Chiba, Y., Iwamatsu, A., Kishino, T., Niikawa, N., Matsubara, Y. & Narisawa, K. (1994) *Nat. Genet.* **8**, 122-128.

25. 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.

27. Courchesne Smith, C., Jang, S. H., Shi, Q., DeWille, J., Sasaki, G. & Kolattukudy, P. E. (1992) *Arch. Biochem. Biophys.* **298**, 576-586.

28. Slusher, L. B., Gillman, E. C., Martin, N. C. & Hopper, A. K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9789-9793.

29. Susuki, T., Yoshida, T. & Tuboi, S. (1992) *Eur. J. Biochem.* **207**, 767-772.

30. Kwiatkowski, D. J., Mehl, R. & Yin, H. L. (1988) *J. Cell Biol.* **106**, 375-384.

31. 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.

35. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9257-9261.

36. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) *EMBO J.* **1**, 945-951. 37. Higgins, C. F., Hiles, I. D., Salmond, G. P., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W. & et al, (1986) *Nature* **323**, 448-450.

CHAPTER 4

MOLECULAR BASIS OF BIOTIN-RESPONSIVENESS OF PATIENTS WITH NEONATAL MULTIPLE CARBOXYLASE DEFICIENCY.

PREFACE TO CHAPTER 4

In Chapter 4, the technology and knowledge developed in Chapters 2 and 3 were used to determine the molecular events that lead to the neonatal form of biotin-responsive MCD. First, I developed a direct assay to determine the activity of HCS using a carboxyl terminal fragment of the α subunit of human PCC as substrate for biotinylation. This protocol has been used to demonstrate HCS deficiency and biotin responsiveness of protein extracts of patients affected with MCD. Second, the determination of the sequence of the HCS cDNA allowed the mutation analysis of DNA from affected indivisuals. Preliminary results indicate clustering of mutations in a 129 amino acid region and define the putative biotin-binding domain of human HCS.

I would like to acknowledge Ms. Lucie Dupuis who collaborated with me throughout this study and who characterized the mutations. She has graciously allowed me to use her mutation data for the preparation of this chapter. However, additional studies will be completed by her before it is ready for publication as a collaborative report. I would also like to thank to Steven Baird of the Sequencing Core Facility of the Canadian Genetic Diseases Network at the Children's Hospital of Eastern Ontario for sequencing assistance.

INTRODUCTION

Holocarboxylase synthetase (HCS) catalyzes the biotinylation of the three mitochondrial and the cytoplasmic biotin-dependent carboxylases present in human cells (57,82,83,84). Genetic deficiency of HCS causes the neonatal form of multiple carboxylase deficiency (MCD), a disorder characterized by a reduction in the activity of all carboxylases (2). MCD is potentially fatal because biotin enzymes catalyze key reactions in gluconeogenesis, fatty acid metabolism and amino acid catabolism (57). However, all the biochemical and clinical manifestations of MCD, including vomiting, hypotonia, lethargy, developmental delay, coma, ketoacidosis, hyperammonemia and glycinuria can be successfully reversed with pharmacological doses of biotin. There is a second, less severe, form of MCD caused by a defect of biotinidase, the enzyme responsible for the intestinal release of protein-bound dietary biotin (38) and for the recycling of the vitamin during carboxylase turnover. Patients affected by this disease also respond to biotin treatment but, unlike neonatal MCD, their symptoms appear later in life usually between 1 to 3 months of age.

The biochemical defect in HCS deficiency was demonstrated by incubating cells in a medium with low concentrations of biotin (58,60). Under these conditions, the activity of the three mitochondrial carboxylases was reduced to undetectable levels, while normal cells retained most of their carboxylase function. The activity of the biotin enzymes was restored to normal values after the cells where transferred to a biotin-rich medium. Burri et al. (63), using semipurified apocarboxylases from biotin-starved rats as a substrate for

HCS, went on to show that the synthetase from affected individuals had a reduced affinity for biotin.

Recently, we (99) and Susuki et al. (97) isolated cDNA clones encoding human HCS. Analysis of the predicted amino acid sequence (726 amino acids) revealed that HCS is a member of a family of biotin-ligases that includes the BirA protein from *E. coli* (321 amino acids) (93) and the putative biotin-ligase of *P. denitrificans* (94). The homology among these proteins is restricted to a small region of 129 amino acids which has been identified by X-ray crystallography as the biotin binding domain of the BirA protein (86). Outside this domain, the bacterial biotin ligases and human HCS do not show any significant sequence conservation. This led us to postulate that mutations in HCS involved in the biotin-responsive phenotype might be located in or near this biotin binding domain (99).

In the current study, we developed an in vitro assay to determine HCS activity based on the biotinylation of a carboxyl terminal fragment of a human propionyl-CoA carboxylase. We used the assay to confirm the deficiency of HCS activity in neonatal MCD and that biotin-responsiveness is a property of the mutant enzyme. Futher, we identified several mutations responsible for the enzyme defect and showed that most of them mapped to the biotin-binding region of HCS.

MATERIALS AND METHODS

Materials. [^{14}C]Biotin (57 mCi/mmol) was purchased from Amersham (United Kingdom). Sequenase was from United States Biochemical. Anti-FLAG affinity gel was purchased from International Biotechnologies Inc. All other reagents were of analytical grade. A 67 amino acid carboxyl terminal fragment (p-67) of the α subunit of human propionyl-CoA carboxylase (PCC-α), encoded in the vector pFLAG67, was used as substrate for HCS (71). The skin cell fibroblast cell lines from patients with neonatal MCD were kindly provided by Lawrence Sweetman from the University of California in Los Angeles. *E. coli* C-124 was kindly provided by Dr. Michel Gaudry from the Pierre and Marie Curie University, Paris.

Synthesis of unbiotinylated p-67 substrate (p-67). Production of unbiotinylated p-67 was accomplished by transformation of biotin auxotroph *E. coli* C-124 cells, unable to convert desthiobiotin into biotin, with the plasmid pFLAG67. The resulting cells were grown to an OD_{600} of 0.5-0.7 in LB medium. The cells were collected by centrifugation and the pellet was washed twice with biotin-free minimal A-medium (15). The synthesis of p-67 was induced by resuspending the cells in A-medium containing IPTG to a final concentration of 2 mM at 37°C for 2 hrs. Purification of the fusion protein p-67 using anti-FLAG affinity gel was performed following the recommendations of the manufacturer.

Direct determination of HCS activity. Cell extracts were prepared from normal and

MCD primary fibroblast cultures and an aliquot (50 μg total protein) was combined in a test tube with 1μg of p-67 (7.0 μM), in 20 μl of reaction buffer (300 mM Tris-HCl pH7.4, 100 mM reduced glutathione, 50 mM $MgCl_2$, 10 mM ATP, 2mM PMSF, 50-500 μM [^{14}C]biotin) at 37°C. The proteins in the reaction mixture were separated by electrophoresis in 12% SDS-polyacrylamide gels. Biotinylation of p-67 was quantitated with a phosphor image analyzer (Fujix BAS 2000;Fuji) after the proteins were transferred to nitrocellulose filters (JBC).

Mutation analysis of HCS deficient patients. Mutation screening was done on PCR products derived from reverse-transcribed RNA (RT-PCR). Amplification of HCS cDNA was done as a series of 6 overlapping fragments of 450 to 650 bp in length using [α^{35}S]ATP as a tracer (L. Dupuis, experiments in progress). The PCR products were digested with restriction enzymes to generate smaller fragments and were subjected to single strand polymorphism analysis (SSCP) in 6% polyacrylamide gels (100). The presence of mutations on the cDNA fragments were detected as shifts in the mobility of the DNA bands. The exact location of the mutations was achieved by direct sequencing of the PCR products. The presence of mutations was confirmed on genomic DNA or cDNA from affected patients by establishment of PCR-dependent diagnostic tests (Dupuis, personal communication).

RESULTS

Determination of HCS activity in MCD patients. To ensure that the biotinylation

experiments were done under saturating conditions, we used a biotin concentration of 50 μM in the HCS assay. Previous studies on kinetics of biotinylation in human cells reported a HCS Km for biotin of 15 nM which is 3000 times lower than the biotin concentration used in this assay (63). Biotinylation of p-67 by normal fibroblast extracts showed a linear relationship with respect to incubation time from 0 to 90 min (Fig.1). Biotinylation of the human peptide was linear with respect to the concentration of total proteins (0-60 μg) used as a source of HCS (Fig.2). The concentrations of p-67 and biotin were not limiting for the assay. An incubation time of 1 h and 50 μg of total protein from fibroblast extracts were selected for subsequent experiments.

Comparison of the HCS activity present in normal and MCD fibroblast extracts, showed that synthetase activity was on average 50 times lower in affected patients than in normal individuals, varying between 0.5% and 2.0% of normal (Fig. 3). The lowest and highest values for biotinylation of p-67 were for cell lines JR and MC, respectively.

Effect of biotin concentration on HCS activity in MCD fibroblasts. To determine if the direct HCS assay was able to demonstrate the biotin responsiveness observed in vivo, we studied biotinylation of p-67 by normal fibroblasts and two mutant cell lines after the biotin concentration was elevated from 50 to 500 μM. Normal fibroblast showed an increase of $<10\%$ in the biotinylation of p-67 at 500 μM compared to 50 μM. On the other hand, cell lines JR and MC showed activities that were 2.4 and 3.2 times the HCS activity obtained at 50 μM (Table 1).

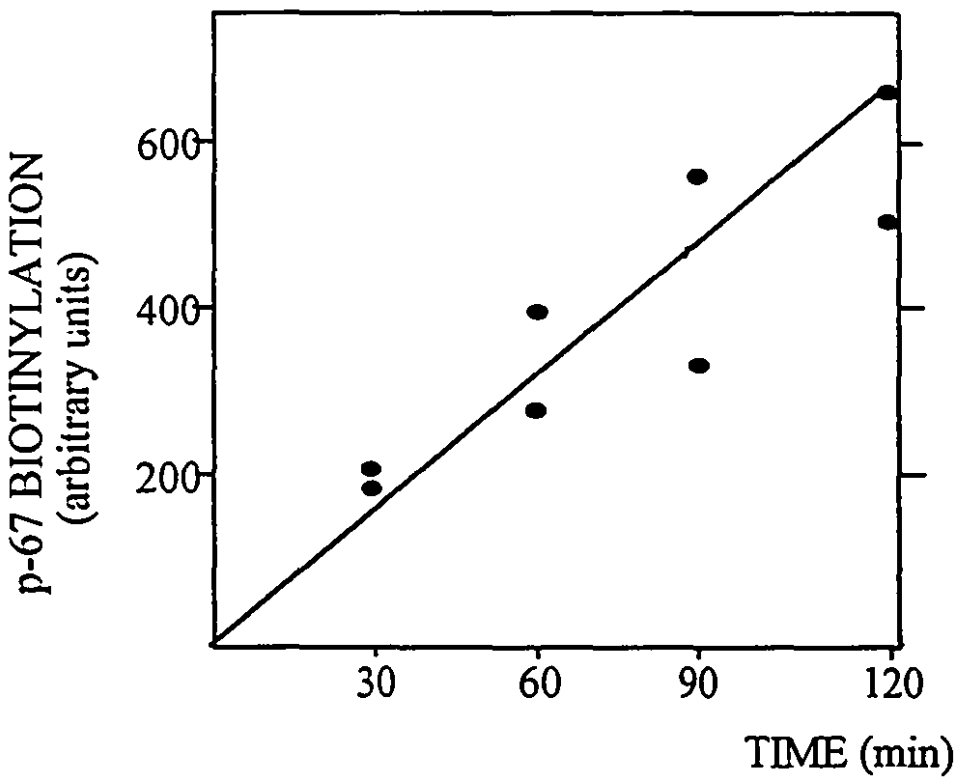


Fig. 1. **Time course of in vitro biotinylation of a human carboxyl-terminal fragment of PCC-α (p-67) by normal cells.** Crude extracts from normal human fibroblasts were incubated in the presence of 1 μg of p-67 and [^{14}C]biotin for different times. The proteins from the cell extracts were analyzed by SDS-polyacrylamide gel electrophoresis, and the biotinylation of p-67 was quantitated as described under "Materials and Methods" using a phosphor image analyzer.

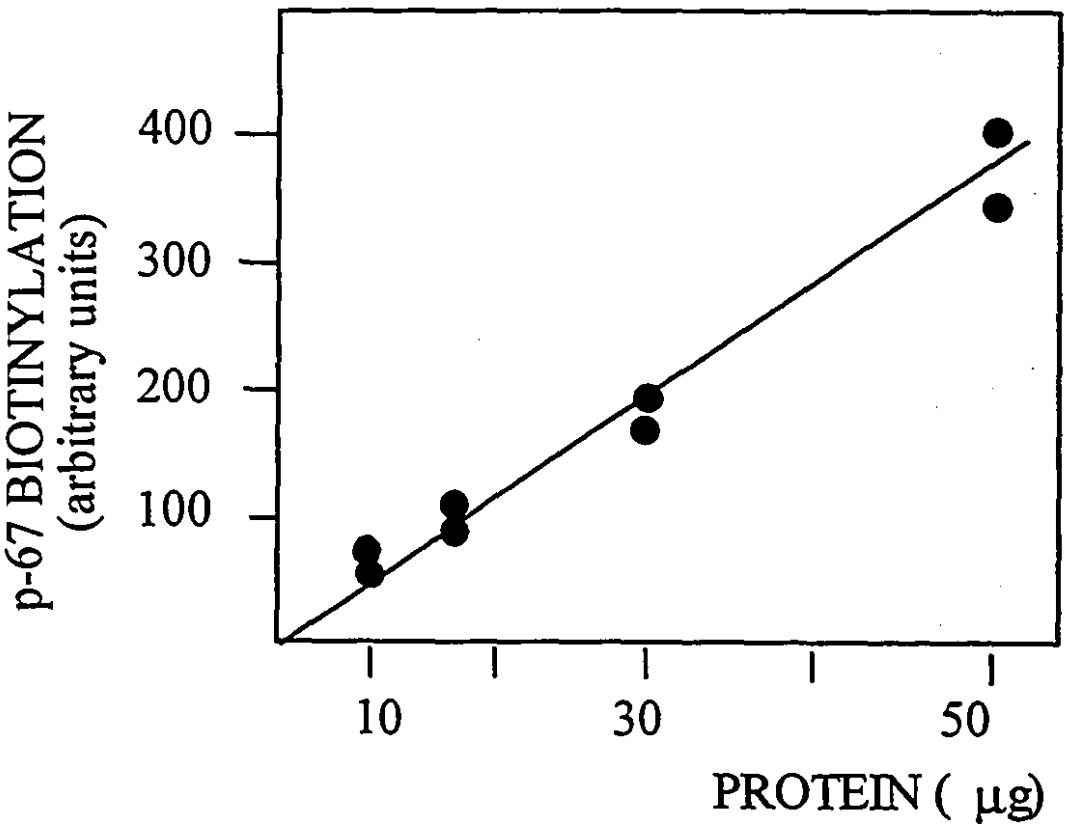


Fig. 2. **Effect of crude extract protein concentration on biotinylation of p-67.** Different amounts of crude extracts from normal fibroblasts (10 to 50 μg of total protein) were incubated in the presence of 1 μg of p-67 and [^{14}C]biotin at 37°C for 1 hr. Biotinylation of p-67 was quantitated as described for figure 1.

Mutation analysis of HCS deficient patients. SSCP analysis and sequencing of cDNA from HCS-deficient cell lines revealed 6 different single-base mutations (Fig.4) All of the mutations resulted in amino acid substitutions in the HCS primary sequence. Two mutations, Arg508Trp and Val550Met, were each present in two different patients. None of the patients had homozygous mutations, since in all cases the normal allele was also identified. However, in two of the six patients two different mutations were identified (PD and MM, Fig. 1), while in the remaining four only one of the two alleles could be identified.

DISCUSSION

We have developed an in vitro assay that utilizes carboxyl terminal fragments of the α subunit of human PCC as a substrate to determine HCS activity in human cells. This system overcame the methodological problems in previous studies that required the purification of apocarboxylases from biotin-starved rats (63). Our method for the determination of HCS activity is direct because it is based on the accumulation of biotinylated products rather than the enzymatic activity of secondary enzymes. Second, it allows control over the concentration of the biotin-acceptor substrate which will permit establishment of kinetic parameters. Third, the synthesis of unbiotinylated p-67 in *E. coli* provides a potentially unlimited, fast and inexpensive source of substrate for HCS.

We have shown that all the patients studied by the direct HCS assay show low but detectable residual synthetase activity. These results could explain why all the HCS

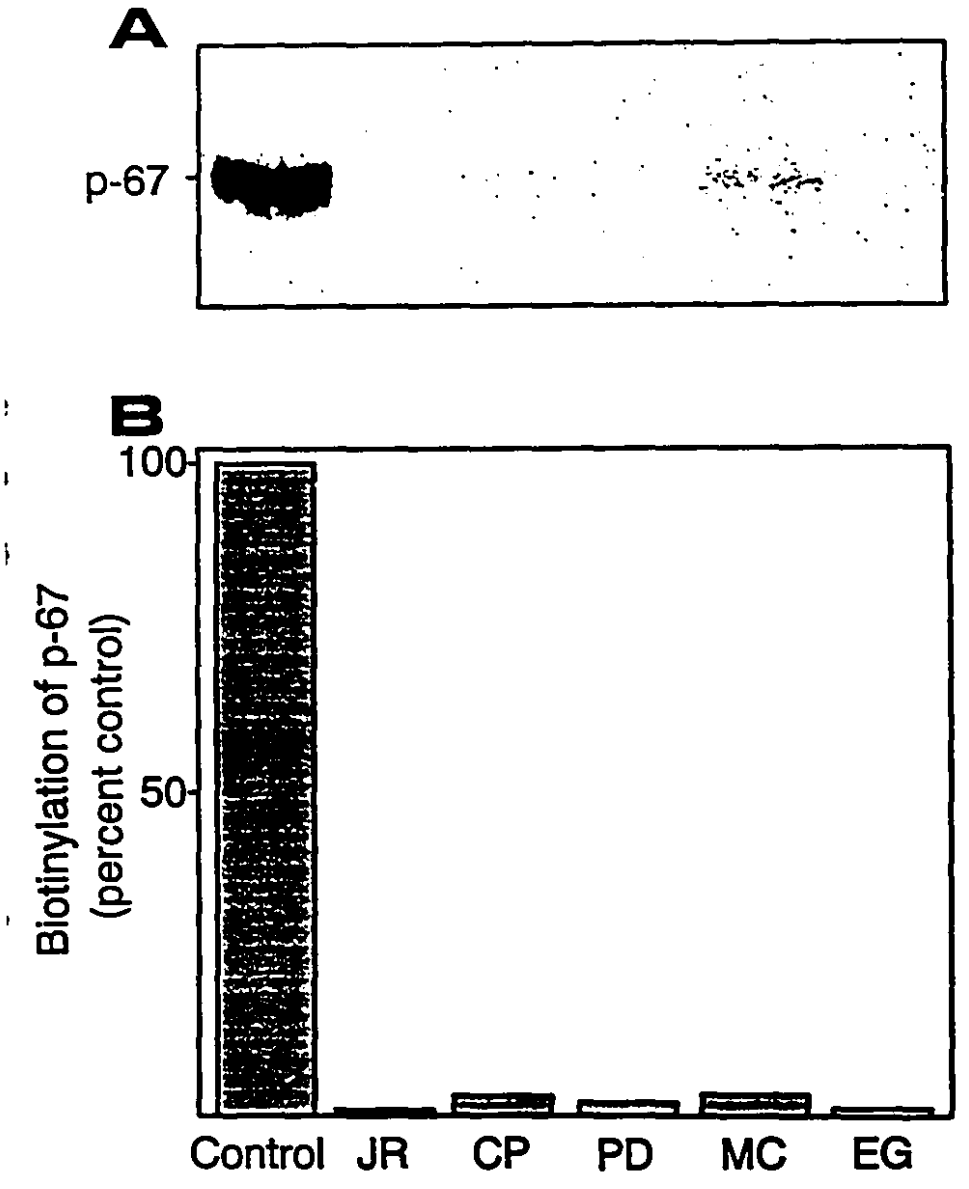


Fig. 3. **HCS activity in fibroblasts from patients with neonatal MCD.** Cell extracts were prepared from normal and MCD fibroblasts and incubated in the presence of ATP, [^{14}C]biotin and p-67 for 1 hr at 37°C. Biotinylation of p-67 was quantitated using a phosphor image analyzer (A), as described in "Material and Methods". (B) Graphical representation of p-67 biotinylation by normal (control) and MCD cells. All samples were run in duplicate. The results are shown as percent of biotinylation in normal cells.

	Biotin	
Cell Line	50 μM	500 μM
CONTROL	3471	4059
MC	155	490
CP	131	320

TABLE 1. Effect of biotin concentration on HCS activity. Biotinylation of p-67 by normal and MCD fibroblasts was quantitated after incubation in the presence of a [^{14}C]Biotin at 50 μM or 500 μM. Biotinylation of p-67 is expressed in arbitrary units.

deficient patients studied so far are biotin-responsive. Previous studies demonstrated the biotin-responsiveness of neonatal MCD based on the rapid recovery of patients given pharmacological doses of biotin (10-80 mg per day) (57) and by the restoration of carboxylase activity in cells cultured in high biotin medium (58,60). Comparison of the increments observed in biotinylation of p-67 by normal and mutant fibroblasts at high versus low biotin concentrations indicate that normal HCS is saturated even at 50 μM biotin, while the mutant synthetase is not. Our observations confirm the hypothesis (6) that neonatal MCD results from a decrease in the affinity of HCS for biotin. However, further studies, with optimization of kinetic parameters, will be required to define in detail the nature and impact of the residual activity of mutant HCS on the onset of MCD.

Most of the mutations in HCS deficient patients were localized to a highly conserved region which conforms to the biotin-binding domain of biotin-ligases (63,86,99). All of the mutations found in the biotin-binding domain affected amino acids that are highly conserved among biotin-ligases (Fig.5). In one of the patients (PD), both mutations were found in this domain. The Arg508Trp is located within the sequence G505-K-G-R-G-G which is similar to the consensus sequence G-X-G-X-X-G associated with ATP binding proteins (86,99). The specific role of this sequence in HCS is not clear since crystallographic analysis of *E. coli* BirA protein revealed that the first Gly residue of the motif is involved in biotin binding. The second mutation, Asp571Asn, affects a protein region which is one of the most conserved sequences between biotin ligase and human HCS. The two mutant alleles for patient MM were also determined but in this case only one, Arg508Trp is present at the biotin-binding domain while the second, Val363Asp,

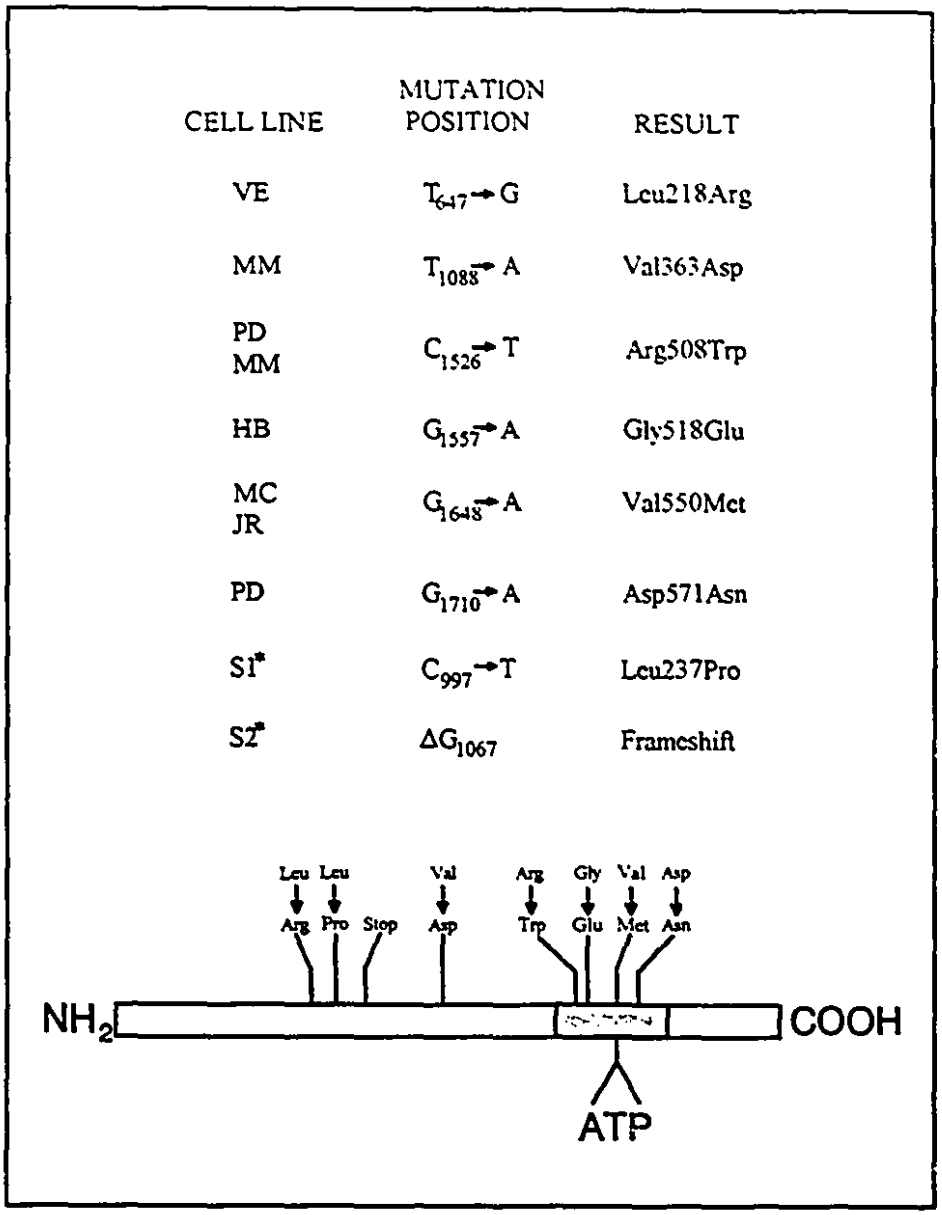


Fig. 4. **Mutation analysis in HCS deficient patients.** The location of mutations present in five MCD patients analyzed by this study are shown at the top of the figure with their corresponding amino acid substitutions. Two different mutations have been previously reported by Suzuki et al. (97) and have been included in this figure (*). The distribution and clustering of mutations is shown in the schematical representation of HCS at the bottom of the figure. The shaded area corresponds to the putative biotin-binding site. The location of a consensus ATP-binding site is shown under the figure.

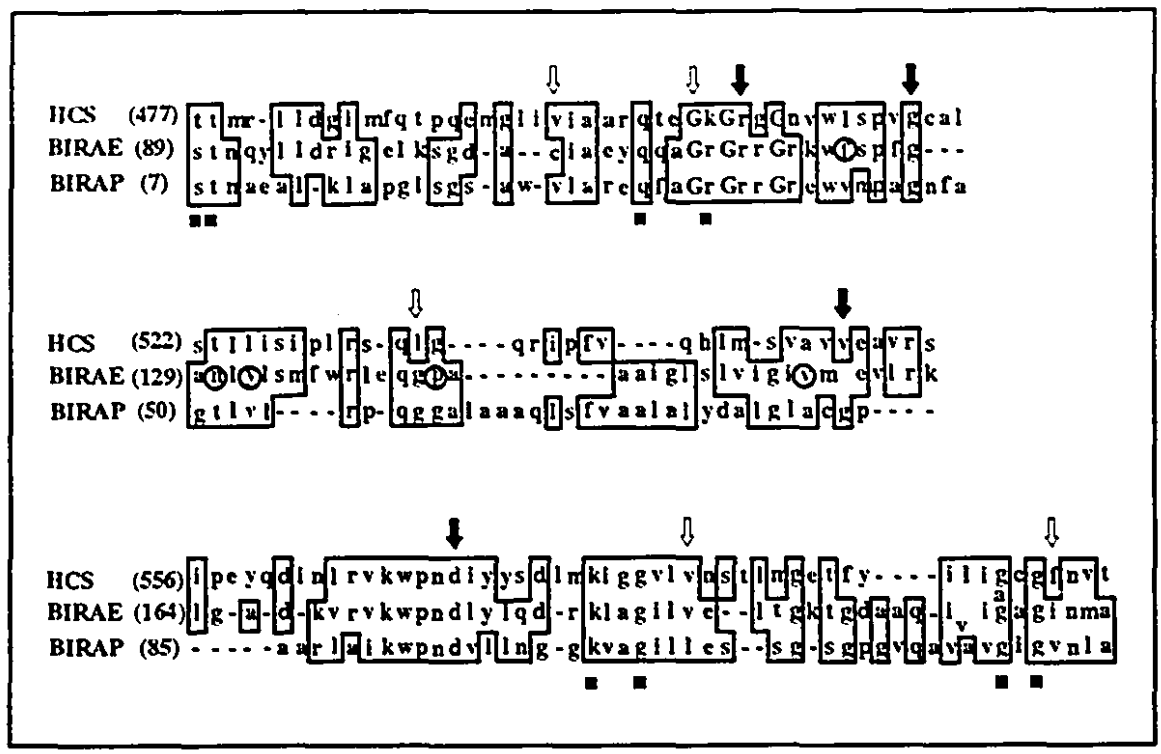


Fig. 5. **Localization of mutations in bacterial BirA protein and human HCS.** Alignment of amino acid sequences of human HCS with *E. coli* BirA protein and the putative biotin ligase from *P. denitrificans* shows that these proteins share a common biotin-binding site. Positions of amino acids involved in contact with biotin in BirA (86), as determined by crystallographic analysis (86), are indicated below the protein sequence. Locations of mutations associated with an increase in the Km for biotin in *E. coli* birA$^-$ are indicated by clear arrows (99). The location of 4 mutations in HCS-deficient patients found by this study is shown by black arrows. Amino acid positions for each protein are shown at left.

was located towards the N-terminal side of HCS. It is possible that the first allele is the one involved in biotin-responsiveness. Most of the mutations found in the biotin domain involve relatively subtle changes which may be asssociated with an increase in the Km of the protein towards biotin and the presence of residual activity of HCS in all patients tested.

The apparent clustering of mutations and the fact that amino acid substitutions in the homologous region of the BirA protein result in biotin auxotrophy in *E. coli* (Fig), support the hypothesis that biotin-responsiveness in neonatal MCD arises from mutations affecting this domain in the human synthetase (99). However, analysis of the DNA from a Japanese patient revealed two mutations that are not located within this region (97). One mutation changes Leu237Pro, while the other, a frameshift mutation, introduces a stop codon at amino acid position 280. In this work, we have characterized two mutations that map outside the biotin domain and one is near the substitution identified by Susuki et al. (97) (Leu216Arg and Val363Asp). It is possible in human HCS that the structure of the biotin-binding domain extends towards the N-terminal side of the protein, suggesting a requirement for sequences which are not conserved in bacteria. Alternatively, this second mutation "hot spot" in HCS may define a distinct site important for enzyme function.

The frameshift mutation (97) present in hemizygous form, would be expected to express an inactive truncated HCS lacking the biotin-binding domain. The patient was reported as biotin-responsive but this might be associated with the second allele, Leu237Pro (97

and Fig. 1), although it is located outside the presumptive biotin-binding region of HCS. It is noteworthy that of the several patients characterized so far, none have had two mutant alleles compatible with a null enzymatic phenotype. Such patients would not be expected to be biotin-responsive. Since such patients have not been described (only one has been described as poorly responsive), it is possible that the complete absence of HCS activity may be lethal in utero.

Our results showed that raising the concentration of biotin to 500 μM increased the activity of the mutant HCS in two MCD cell lines to only 7.0% and 12.0% of the activity in normal cells. Since the patients were fully biotin-responsive (L. Sweetman, personal communication), it suggests that there is a threshold activity of 5% - 10% of normal that may be sufficient to sustain biotinylation of apocarboxylases in HCS-deficient patients. The time it takes these patients to recover after treatment with pharmacological doses of biotin may represent the period required to increase the intracellular concentration of biotin to such levels as to permit a slight increase in HCS activity.

The effect of all the mutations described in this work on biotin-responsiveness will need to be determined through their expression in in vitro-generated mutant HCS. This should be possible using the cellular biotinylation assay involving the co-expression of human HCS and p-67 in *E. coli* (71,99) or, possibly, through the direct determination of HCS activity after extraction of the mutant enzyme expressed in *E. coli*.

CHAPTER 5

GENERAL DISCUSSION

In this thesis, I have used the biotinylation system of *E. coli* for the investigation of the mechanisms of biotinylation in human cells and the characterization of the human biotin disorder, neonatal MCD.

Characterization of the biotin domain of carboxylases. In Chapter 1, I characterized the structure of the biotin domain of PCC-α. Expression of carboxyl terminal fragments of this protein in *E. coli* indicates that the recognition of human carboxylases by BirA is restricted to a discrete region of 67 amino acids in length around the biotin-binding site (71). The biotin domain of PCC-α (p-67) was defined at its amino terminus by the sequence Pro-Met-Pro. I showed that this tripeptide functions as a specific signal for the biotin-ligase and defines the minimum size, at the amino terminal side, that could act as an independent domain for biotinylation. This N-terminal sequence may involve specific structural features, defined by the proline sequence, or it may act as a flexible hinge that turns the biotin domain into a "oscillating arm", enabling biotin to move between donor and acceptor sites during the transfer of a carboxyl group.

This latter hypothesis is strengthened by structural similarities between the biotin domain of carboxylases and the lipoic acid domains of the enzyme dihydrolipoyl acetyltransferase (E_{2p}), a subunit of the *E. coli* pyruvate dehydrogenase complex (43,89). The amino terminal half of the E_{2p} chain consists of three independently folded domains, each of about 80 amino acids in length that contain a lipoic acid molecule covalently attached to a specific Lys residue. These proteins are thought to function in a similar fashion to

biotin enzymes, in the sense that lipoic acid is able to oscillate between donor and acceptor molecules during the transfer of an acetyl group. Limited proteolysis and ^1H NMR spectroscopy (90,91,92) revealed that the movement of the lipoyl domains is facilitated by conformationally flexible protein segments, rich in proline, alanine and charged amino acids, located between them. The Pro-Met-Pro sequence in carboxylases may have a similar role to the Pro-rich regions in E_{2p} proteins. However, prediction of the amino acid chain flexibility of the biotin domains of carboxylases localizes three flexible points near or at three highly conserved Gly residues and not on the Pro-Met-Pro sequence (71).

Recently Blokenhurst and Perham (89) predicted the three dimensional structure of the lipoic acid domain of the H-protein from the pea leaf glycine cleavage system and the biotin domain of yeast pyruvate carboxylase. They are remarkably similar (Fig.1). In both cases, the cofactor binding site is located in a ß-turn at the end of a finger-like structure formed by two ß-sheets. Significantly, both domains have three Gly residues located at similar positions with respect to the cofactor binding sites. Two of these residues, identified as candidate flexible points, are situated at opposite ends at the base of the "cofactor arm". It is possible that the movement of the biotin domain is facillitated by these residues and not by the Pro tripeptide. However, the role of the Pro-Met-Pro sequence or the conserved Gly residues in chain flexibility will have to be determined by NMR or X-ray crystallography. Either one of these possibilities suggests that is the entire domain, not just the biotin cofactor, that oscillates during the carboxylation reaction. According to the predicted structure of the lipoic acid domain,

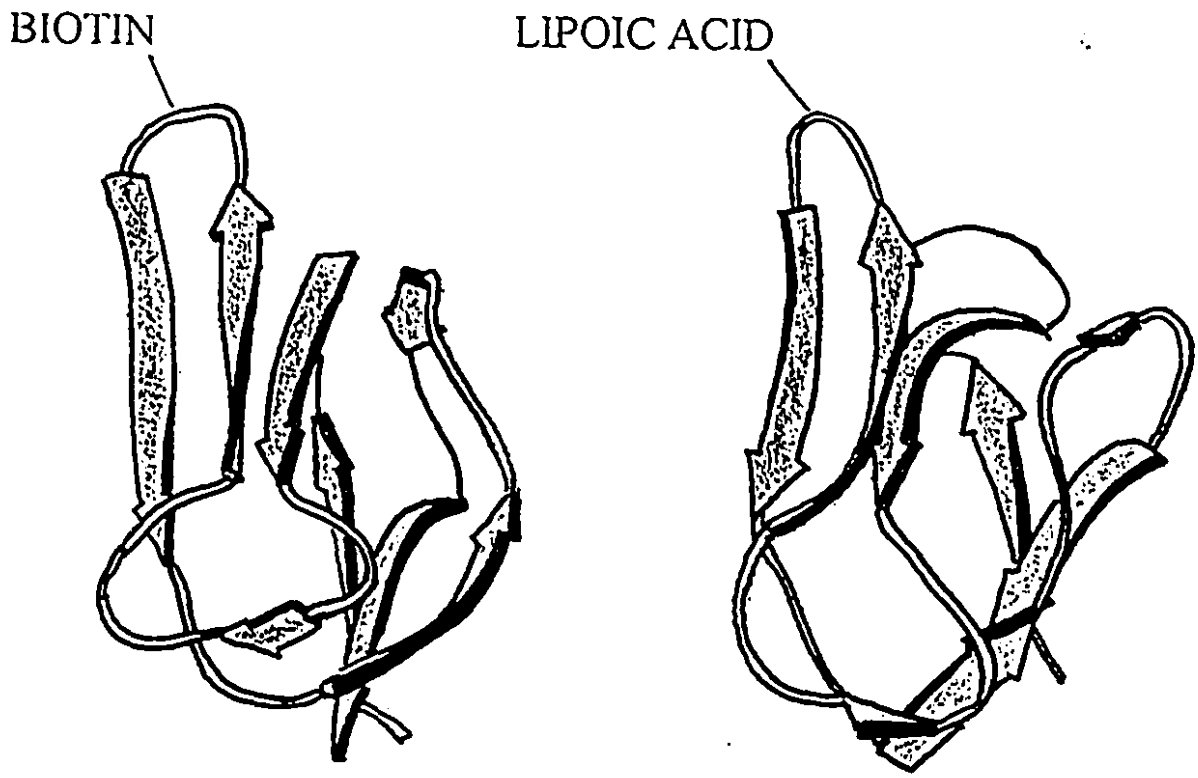


Fig. 1. **Structural similarities between biotin and lipoic acid domains.** Schematic drawing of the predicted three-dimensional structure of the biotin domain from yeast pyruvate carboxylase (A) and the lipoic acid domain from the H-protein from the glycine cleavage system of pea leaf (B), (89). ß-strands are indicated by arrows.

the conserved Gly residues are important for maintaining the integrity of the domain (89). The same may be true for the biotin-dependent carboxylases. This view of their role suggests that the inefficient biotinylation observed by mutant PCC-α peptides containing mutations at the Gly positions could be the result of a disruption in the architecture of the biotin binding site.

It is possible that the homologous sequence between biotin- and lipoic acid-containing proteins represents the minimum structure required for a functional "movable protein arm" involved in chemical group transfer. This structure may have been selected early in evolution and modified since to accommodate the different enzymatic activities that characterize biotin and lipoic acid containing enzymes. In this light, it is interesting that the similarities between these two classes of enzymes extend to their cofactors which are structurally similar (Fig.2).

The characterization of a "universal substrate" (p-67) that is biotinylated by bacterial and eukaryotic biotin ligases suggests an additional application of the results presented in this work. Purification of active eukaryotic proteins expressed in *E. coli* is not always possible because the required posttranslational modification, subunit assembly or correct protein folding do not occur in bacteria. This problem might be solved if the proteins of interest, containing p-67 as a fusion peptide at the amino or carboxyl terminus, are expressed in mammalian cells. The active proteins could then be easily purified using an avidin-sepharose column taking advantage of the endogenous addition of the biotin tag. The introduction of an enteropeptidase recognition sequence between p-67 and the

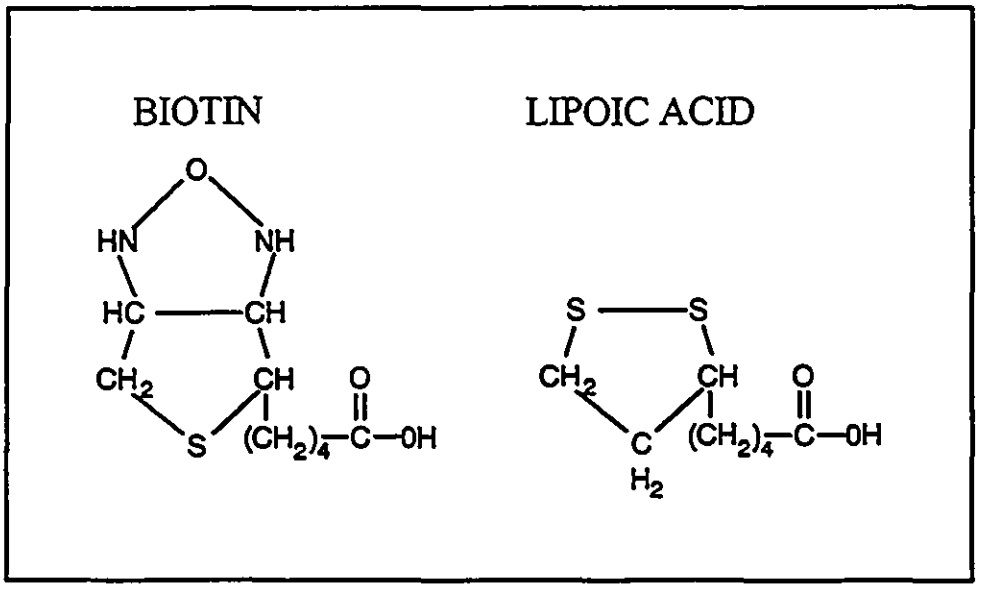


Fig. 2. **Structural similarities between biotin and lipoic acid.**

fusion protein would allow the single step purification of the experimental proteins. Incubation of the column with enteropeptidase would cleave the fusion proteins allowing their elution, leaving p-67 and the four biotin-containing carboxylases bound to the column. Alternatively, p-67 might also be used as a fusion peptide to purify large amounts of ectopic proteins expressed in *E. coli* when enzymatic activity is not a problem. The extraordinary affinity of biotin for avidin ($K_{dis}=10^{-15}M$) would permit a level of purification not generally available with current methods. The expression of p-67 in *E. coli*, showed that this peptide could compete so well with the endogenous carboxylases for biotinylation that overexpression of fusion proteins could result in the death of the host. In mammalian cells, this could lead to growth arrest or an ineficient biotinylation of the fusion protein. To overcome this problem and maximize the expression of fusion proteins, human HCS could be coexpressed from the same plasmid that encodes the p-67 fusion protein eliminating the limiting step presented by the biotinylation reaction.

Cloning and characterization of the cDNA and predicted amino acid sequence of the human HCS. The sequence of HCS indicates that it belongs to a family of biotin ligases which includes the BirA protein of *E. coli* (93) and a candidate BirA sequence from *P. denitrificans* (94). The sequence homology between these proteins is restricted to a fragment of 129 amino acids in length on BirA which corresponds to the biotin binding site. All the mutations associated with biotin auxotrophy in biotin ligase deficient *E. coli* cells were clustered within this region of the BirA protein (86).

BirA is a 321 amino acid protein that can be divided into two major domains. An N-terminal region of about 80 amino acids, which is responsible for its DNA binding activity and a C-terminal domain containing the biotin ligase activity (86). The facts that the 129 amino acid biotin binding domain represents 40% of the total length of the protein and that no other region seem to have been conserved in biotin ligases suggest that all of the HCS functional requirements may be encoded by this small protein domain. The high sequence conservation of this region would explain the cross-species activity shown by bacterial and eukaryotic synthetases. It will be possible to test this hypothesis by determining if truncated versions of HCS are able to rescue the phenotype of BirA deficient *E. coli* cells and promote the biotinylation of the endogenous BCCP or the human p-67 peptide as described in Chapter 2. This will also determine if the site implicated by the mutations Leu216Arg and Val363Asp has a role in biotinylation. It too may need to be included in the analysis of truncated proteins.

The structure of isolated cDNA clones indicate that the HCS mRNA is alternatively spliced at the 5' end. I discussed the possibility that alternative splicing is responsible for the expression of synthetase activity in mitochondria versus cytoplasm. The exact mechanism used for the targeting of HCS remains to be confirmed, whether by the utilization of the different candidates for translation initiation that were identified or by the splicing of a still unidentified exon containing the rest of a mitochondrial leader sequence linked to the sequence 5' of Met1. One approach to assaying the targeting that I began but was unable to complete was to express a chimeric protein containing candidate amino terminal regions of HCS linked to the sequence of the green fluorescent

protein (GFP) from *Aequorea victoria* at the C-terminus (95). The fluorescent moiety would be used as a marker to facilitate the identification of the cellular fate of the fusion proteins. Part of this work should include the use of a procedure such as rapid amplification of cDNA ends (RACE) or exon trapping of genomic DNA sequences to determine if there is additional sequence extending the open reading frame at the 5' end of the HCS transcript. If such sequences exist, their competence as mitochondrial import signals could be evaluated as part of the fluorescent fusion protein expression experiments.

The presence of different forms of HCS mRNA in human cells was confirmed by Northern blot analysis. It is possible that the different mRNA forms observed in this study are involved in the tissue or temporal regulation of the expression of HCS. The recent cloning of mouse HCS cDNA and the finding of similar mRNA forms in this organism (M. Loyer, R.A Gravel, J. Trasler, personal communication) provides an accessible source of tissues for characterizing the structure of these mRNA species and determining their role in the expression of HCS in development.

Determination of HCS activity and mutations present in patients affected wit neonatal MCD. The bacterial biotinylation of human PCC-α fragments and the possibility of synthesizing them in an unbiotinylated form in *E. coli* suggested their use for the study of human MCD. The method of determininig HCS activity by biotinylation of carboxyl terminal fragments of PCC described in Chapter 3 contrasts with the coupled assay of Burri et al (63). The use of p-67 to quantitate HCS activity allowed the direct

determination of the product of the reaction, radiolabelled biotin-p67, rather than through the activity of the activated carboxylases. Further, the ability to synthesize unbiotinylated p-67 in *E. coli* provides a rapid, virtually unlimited, inexpensive source of HCS substrate for the diagnosis of MCD. This method will facilitate the determination of the real kinetic constants of HCS towards biotin, ATP and apocarboxylases, something that is not possible with the indirect assay described above.

The suitability of this approach to study HCS deficiency was demonstrated by the analysis of several normal and HCS deficient cell lines. The results show direct evidence that MCD patients have a defective HCS activity that is unable to biotinylate carboxylases at low biotin concentration. The increase in p-67 biotinylation by the mutant but not normal HCS when the biotin concentration was increased from 50 μM to 500 μM supports the hypothesis that a decrease in the affinity of the enzyme for biotin is responsible for MCD. All the patients studied with the p-67 assay show residual activity of HCS suggesting that the total loss of HCS activity may be lethal in utero. The cloning of the mouse HCS gene will make it possible to develop an animal model to study the effects of the total absence of HCS activity on mammalian development.

The use of the p-67 assay showed that the biotinylation of carboxylases is catalyzed by only one synthetase in human cells. The almost complete absence of p-67 biotinylation by cell extracts from MCD patients, where cytoplasmic and mitochondrial enzymes have been mixed, abolishes the possibility of the existence of a second cytoplasmic HCS. The likelihood that the cytoplasmic synthetase may not recognize p-67 as a substrate is

eliminated by the fact that full length PCC-α is biotinylated efficiently both in mitochondria and in cytosol (96). I proposed that the HCS activity reported in both cell compartments is the result of a differential targeting of the same protein to both sites. HCS targeting may be mediated through alternative splicing of mRNA as was described in Chapter 2.

Preliminary mutation analysis by SSCP from patients affected with MCD has resulted in the identification of 6 point mutations that result in amino acid substitutions. Four of these changes are in the predicted biotin-binding domain of HCS. These results confirmed the hypothesis presented in Chapter 2 that mutations causing MCD should map to the conserved 129 amino acid region in HCS. After the article that communicated the cloning of HCS was submitted for publication, Susuki et al. (97), published the cloning of a cDNA, similar to our clone BL-11. In this study they reported two different mutations occurring outside from the predicted biotin-binding domain, towards the N-terminal half of HCS in several Japanese patients. Two additional mutations described in Chapter 3 map near this location, suggesting that this region may also be involved in biotin binding or in other enzyme function. These results make it necessary to investigate the effect of all these mutations on biotin-binding by their expression in *E. coli birA*$^-$ cells to evaluate their effect on biotinylation of p-67 under different concentrations of biotin.

SUMMARY

Biotin-dependent enzymes are widely distributed in all living organisms where they are involved in carboxylation and decarboxylation reactions. Biotinylation of these proteins is catalyzed by HCS in eukaryotes and biotin ligases in prokaryotes. The ability of these enzymes to recognize apocarboxylases from different organisms and the high sequence conservation between biotin-enzymes suggested that the mechanism of biotinylation has been conserved through evolution from bacteria to humans. The biotinylation domain is a highly conserved region among biotin-dependent enzymes. Expression in *E. coli* of a 67 amino acid peptide containing the biotinylation domain of PCC-α showed its ability to act as an independent domain for biotinylation. Successful biotinylation of PCC-α fragments required of the presence of a Pro-Met-Pro recognition sequence at the N-terminus of the peptide. Three highly conserved Gly residues within the biotinylation region may be involved in determining the structure of the domain and in providing the carboxylase with a flexible hinge to facilitate the movement of biotin between donor and acceptor sites of carboxylation. In humans, deficient biotinylation of carboxylases due to an impairment of HCS activity results in MCD, a disease characterized by deficient activity of all biotin-dependentarboxylases. The use of the biotin domain of PCC-α as a biotinylation substrate confirmed HCS deficiency and biotin-responsiveness in cell extracts from patients affected with MCD. cDNA cloning by functional complementation of biotin ligase-deficient *E. coli* cells proved that human HCS belongs to a family of biotin ligases that includes the BirA protein of *E. coli* and the putative biotin ligase of *P. denitrificans*. Amino acid sequence conservation between these proteins is restricted

to the biotin-binding domain. Analysis by RT-PCR showed that most of the mutations present in HCS-deficient patients are clustered in this region of the protein. I propose that neonatal MCD and the biotin-responsiveness that characterizes the disorder arise from mutations that reduce the affinity of the biotin pocket of HCS towards biotin.

CLAIMS TO ORIGINALITY

CHAPTER ONE

1. This thesis contains a comprehensive analysis of the requirements for the biotinylation of a human biotin-dependent carboxylase (Leon Del Rio and Gravel 1994).

2. I determined that the functional biotin domain of human PCC-α is 67 amino acids in length and can act as an independent domain during its recognition and biotinylation by *E. coli* biotin ligase.

3. I demonstrated the use of C-terminal fragments of human PCC-α to study the biotinylation reaction in a quantitative fashion.

CHAPTER TWO

4. This chapter describes the isolation and characterization of a cDNA encoding a eukaryotic (human) holocarboxylase synthetase by functional complementation in *E. coli* (Leon-Del-Rio at al. 1995).

5. I showed that human HCS, like prokaryotic biotin ligases, is able to biotinylate apocarboxylases from different organisms.

6. I demonstrated the presence of multiple forms of HCS mRNA in human tissues.

7. I documented alternative splicing of HCS mRNA as a possible mechanism to target the synthetase to both mitochondria and cytosol.

8. I demonstrated that the biotin binding site of biotin ligases is 129 amino acids in length and has been conserved during evolution from bacteria to humans.

9. My results provide an explanation, at the molecular level, for the cross-species activity shown by bacterial and eukaryotic holoenzyme synthetases and for the biotin-responsiveness in MCD.

CHAPTER THREE

10. I developed a direct assay for the determination of HCS activity.

11. I determined the biochemical basis of biotin-responsiveness in cells from MCD patients using biotinylation as an indicator of HCS activity.

12. I presented evidence that supports the hypothesis for the existence of a single HCS in human cells.

13. I showed evidence that mutations responsible for a decrease in the affinity of HCS

towards biotin are clustered in the 129 biotin domain of HCS. The mutations were determined by Lucie Dupuis, a graduate student in the same laboratory.

GENERAL REFERENCES

1. Du Vigneaud V., Hoffmann K., Melville D.B. (1942), J. Am. Chem. Soc. 64:188.
2. Melville D., Moyer A.W. Hoffman K, Du Vigneaud V. (1942) J. Biol. Chem. 146:487.
3. Kogl F. (1935) Chem. Ber. 68:A16
4. Gyorgy P., (1939) J Biol Chem 131:733
5. Gyorgy P., (1931) Z Aerztl Forbild 28:377
6. Harding M.G., Crooks H. (1961) J Am Diet Assoc 38:204.
7. Wolf B., in The Metabolic and Molecular Bases of Inherited Disease. Scriver C., Beaudet A.L., Sly W.S. Valle D. (eds) (1985) p. 3151-3177 McGraw-Hill, Inc. New York.
8. Lane M.D., Lynen F. (1964) J. Biol. Chem. 739:2858-2864.
9. Dakshinamurti K., Bhagavan H.N. (eds) (1985) Ann. N.Y. Acad. Sci 447, 1-441.
10. Moss J., Lane M.D. (1971) Adv. Enzymol. 35,321-442.
11. Alberts A.W., Vagelos P.R. (1972) in The Enzymes (Boyer, P.D., ed) Vol 6 pp, 37-82, Academic Press, Orlandp, Fl.
12. Wood, H.G., Barden, R.E. (1977) Annu. Rev. Biochem. 46, 385-413.
13. Christner J.E., Schlesinger M.J., Coon M.J. (1964) J. Biol. Chem. 238:3997
14. Chiang G.S., Mistry S.P., (1975) J. Biochem (Tokyo) 6:527.
15. Barker, D.F., Campbell A.M. (1981) J. Mol. Biol. 146, 451-467.
16. Barker, D.F., Campbell A.M. (1981) J. Mol. Biol. 146, 469-492.

17. Eisenberg M.A., Prakash O., Hsiung S.C. (1982) J. Biol. CHem. 257,15167-15173.

18. Samols D., Thorton C.G., Murtif V., Kumar G.K., Haase F.C., Wood H.G. (1988) J. Biol Chem 263,6461-6464.

19. Knowles J.R. (1989) Annu. Rev. Biochem. 58,195-221.

20. Alberts A.W., Nervi A.M., Vagelos P.R. (1969) Proc. Natl. Acad. Sci. USA 63,1319-1326.

21. Guchhait R.B., Polakis S.E., Dimroth P., Stoll E., Moss J., Lane M.D.(1974) J. Biol. Chem. 249, 6633-6642.

22. Li S.J., Cronan J.E. (1992) J. Biol. Chem. 267,16841-16847.

23. Li S.J., Cronan J.E. (1992) J. Biol. Chem. 267,855-863

24. Li S.J., Cronan J.E. (1993) J. Bacteriol. 175,332-340

25. Gyorgy P., Langer B.W. (1968) Biotin VI, in Sebrell W.H., Harris R.S. (eds): The Vitamins, Chemistry, Physiology, Pathology Methods. New York, Academic Press, pp292.

26. Eisenberg M.A. (1973) Adv. Enzymol. 38:317.

27. Otsuka A.J., Buoncristiani M.R., Howard P.K., Flamm J., Johnson C., Yamamoto R., Uchida K., Cook C., Ruppert J., Matsuzaki J. (1988) 263:19577-19585.

28. Cronan J.E. (1989) Cell 58:427-429

29. Cronan J.E. (1988) J. Biol. Chem. 263: 10332-10336

30. Buoncristiani M.R., Otsuka A.J. (1988) J. Biol. Chem. 263, 1013-1016.

31. Howard P.K., Shaw J., Otsuka A.J. (1985) Gene 35,321-331.

32. Wilson K.P., Shewchuk L.M., Brennan R.G., Otsuka A.J., Matthews B.W. (1992) Proc. Natl. Acad. Sci. USA 898,9257-9261.

33. Buoncristiani M.R., Howard P.K., Otsuka A.J. (1986) Gene 44,255-261.

34. Friedrich W. (1988) Vitamins 1:753.

35. Matthews D.M., Addison J.M., Burston D. (1974) Clin. Sci. Mol. Med. 46:693-705.

36. Thoma R.W., Peterson W.H. (1954) J. Biol. Chem. 210,569.

37. Leon-Del-Rio A., Velazquez A., Vizcaino G., Robles-Diaz G., Gonzalez-Noriega A. (1990) Ann. Nutr. Metab. 34;266-272.

38. Wolf B., Heard G.S., McVoy J.S., Raetz H.M. (1984) J. Inherited Metab. Dis. 7 (Suppl 2): 121.

39. Spencer R.P., Brody K., (1964) Am. J. Physiol. 266;653-657

40. Berger E., Long S., Semenza G. (1972) Biochim. Biophys. Acta 255;873-877.

41. Gore J., Hoinard C. (1987) J. Nutr. 117;527-532.

42. Landman A.D. (1976) Life Sci. 19,1377.

43. Browner M.F., Taroni F., Setul E., Rosenberg L.E. (1989) J. Biol. Chem. 264;12680-12685.

44. Chauhan J., Dakshinamuti K. (1992) J. Biol. Chem.

45. Hayakawa K., Oizumi J. (1988) J. Biochem 773.

46. Craft D.V., Goss N.H., Chandramouli N. Wood H.G. (1985) Biochem 24,2471.

47. Cole H., Reynolds T.R., Lockyer J.M., Buck G.A., Denson T., Spence J.E., Hymes J., Wolf B. (1994) J.Biol. Chem. 269, 6566-6570.

48. Nilsson L., Kagedal B. (1993) Biochem. J. 292,545-551.

49. Oizumi J., Hayakawa K. (1991) Biochim. Biophys. Acta. 1074,433-438.

50. Chauhan J., Dakshinamurti K. (1988) Biochem. J. 256,265.

51. Velazquez A., Zamudio S., Baez A., Munguia-Corral R., Rangel-Peniche B., Carrasco A. (1990) Eur. J. Clin. Nutr. 44:1116

52. Dixon M., Weeb E. Enzymes third edition (1979). Academic Press p 507

53. Bonjour J.P. (1981) Wld. Rev. Nutr. Diet. 38,1.

54. Blom W., Scholte H.R. (1981) N. Engl. J. Med. 305,465.

55. Robinson B.H., Oci J., Saunders M., Gravel R.A. (1983) J. Biol. Chem. 268,6660.

56. Lamhonnwah A., Kam K.F., Tsui F., Robinson B.H. Saunders M., Gravel R.A. (1983) Am. J. Hum. Genet. 35,889.

57. Wolf B., Rosenberg R., Prusiner S.B., Di Mauro S., Barchi R.L., Kunkel L.M. (eds), Stoneham, Mass, Butterworth Publishers, 1992.

58. Saunders M., Sweetman L., Robinson B.H., Roth K., Cohn R., Gravel R.A. (1979) J. Clin. Invest. 64,1695.

59. Feldman G.L., Hsia Y.E., Wolf B. (1981) Am. J. Hum. Genet. 33, 692.

60. Bartlett K., Dale G., Green A., Leonard J.V. (1981) J. Inherited. Metab. Dis. 4,183.

61. Bartlett K., Gombertz D. (1976) Lancet 2,804.

62. Burri B.J., Sweetman L., Nyhan W.L. (1985) Am. J. Hum. Genet. 37,326-337.

63. Burri B.J., Sweetman L., Nyhan W.L. (1981) J. Clin. Invest. 68,1491-1495.

64. Fuchshuber A., Suormala T., Rcth B., Duarn M., Michalk D., Baumgartner E.R. (1993) Eur. J. Pediatr. 152,446-449.

65. Siegel L., Foote J.L., Christner J.E., Coon M.J. (1963) Biochem. Biophys. Res. Comm. 3,307.

66. Siegel L., Foote J.L., Coon M.J. (1965) J. Biol. Chem. 240,1025.

67. Lane M.D., Rominger K.L., Young D.L., Lynen F. (1964) J. Biol. Chem. 239,2865

68. Cazzulo J.J., Gundaram T.K., Dilks S.N., Kornberg H.L. (1971) Biochem. J. 122,663.

69. McAllister H.C., Coon M.J. (1966) J. Biol. Chem 241,2855.

70. Chiba Y., Susuki Y., Aoki Y., Ishida T., Narisawa K. Arch. Biochem. Biophys. (in press).

71. Leon Del Rio A., Gravel R.A. (1994) J. Biol. Chem. 269,22964-22968.

72. Shenoy B.C., Murtif V., Samols D., Wood H.G. (1987) Fed. Proc. 46,2124 (abstr.).

73. Shenoy B.C., PAranjape S., Murtif V., Kumar G.K., Samols D., Wood H.G. (1988) FASEB J. 2,2505-2511.

74. Wood H.G., Shenoy B.C., Kumar G.K., Paranjape S., Murtif V., Samols D. (11987) Fed Proc 46, 2124 (abst.)

75. Murtif V., Samols D. (1987) J. Biol. Chem. 262,11813-11816.

76. Cronan J.E. (1990) J. Biol. Chem. 265,10327-10333.

77. Yamano N., Kawata Y., Kojima H., Yoda K., Yamasaki M. (1992) Bioxc. Biotech. Biochem. 56,1017-1026.

78. Lardy H.A., Peanasky R. (1953) Physiol. Rev 33,560.

79. Wakil S.J., Titchener E.B., Gibson D.M. (1958) Biochim. Biophys. Acta 29:225.

80. Lynen F. (1979) Critical Reviews in Biochemistry 103

81. Obermayer M., Lynen F. (1976) TIBS 169-171

82. Kosow D.P., Lane M.D. (1961)Biochem. Biophys. Res. Commun. 5;191.

83. Kosow D.P., Lane M.D. (1962) Federation Proc. 21;286.

84. Kosow D.P., Lane M.D. (1962) Biochem. Biophys. Res. Commun. 7;439.

85. Wright L.D., Cresson E.L., Skegs H.R., Wood T.R., Peck R.L., Wolf D.E., Folkers K.J. (1951) J. Amer. Chem. Soc. 72;1048.

86. Wilson K.P., Shewchuck L.M., Brennan R.G., Otsuka A.J., Matthews B.M. (1992) Proc. Natl. Acad. Sci. USA. 89;9257-9261.

87. Gornicki P, Sacappino L.A., Haselkorn R. (1993) J. Bacteriol. 175,5268-5272.

88. Velazquez A., Zamudio S., Baez A., Munguia-Corral R., Rangel-Peniche B., Carrasco A. (1990) Eur. J. Clin. Nutr. 44:1116

89. Brockehurst S.M., Perham R.N. (1993) Protein Science 2,626-639.

90. Perham R.N., Duckworth H.W., Roberts G.C. (1981) Nature 272,474-477.

91. Texter F.L., Radforf S.E., Lave E.D., Perham R.N., Miles J.S., Guest J.R. (1988) Biochemistry 27;289-296.

92. Radford S.E., Lave E.D., Perham R.N., Martin S.R., Appella E. (1989) J. Biol. Chem. 264; 767-775.

93. Howard P.K.,Shcow J., Otsuka A.J. (1985) Gene 35;321-331.

94. Xu X., Matsuno Y.A., Yogi T. (1993) Biochemistry 32,968-981.

95. Chalfie M., Tu Y., Euskirchen G., Ward W.W., Prasher D.C. (1994) Science 263;802-805.

96. Taroni F., Rosenberg L.E. (1991) J. Biol. Chem. 266;13267-13271.

97. Susuki Y., Aoki Y., Ishida Y., Chiba Y., Iwamatsu A., Kishino T., Niikawa N., Matsubara Y., Narisawa K. (1994) Nature Genetics 8:122-128.

98. Ghneim H.K., Bartlett K. (1982) Lancet 1:1187.

99. Leon-Del-Rio A., Leclerc D., Akerman B., Wakamatsu N., Gravel R.A. (1995) Proc. Natl. Acad. Sci. USA (in press).

100. Gravel R.A., Akerman B., Lamhonwah A.M., Loyer M., Leon-Del-Rio A., Italiano I. (1994) Am. J. Hum. Genet. 55:51-88.