

2008

# ENZYMOLGY AND MOLECULAR BIOLOGY OF BILE ACID 7 $\alpha$ - AND 7 $\beta$ - DEHYDROXYLATION BY THE INTESTINAL BACTERIA CLOSTRIDIUM SCINDENS AND CLOSTRIDIUM HYLEMONAE

Jason Michael Ridlon  
*Virginia Commonwealth University*

Follow this and additional works at: <http://scholarscompass.vcu.edu/etd>

 Part of the [Medicine and Health Sciences Commons](#)

© The Author

---

Downloaded from

<http://scholarscompass.vcu.edu/etd/736>

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact [libcompass@vcu.edu](mailto:libcompass@vcu.edu).

© Jason Ridlon

2008

---

All Rights Reserved

ENZYMOLGY AND MOLECULAR BIOLOGY OF BILE ACID 7 $\alpha$ - AND 7 $\beta$ -  
DEHYDROXYLATION BY THE INTESTINAL BACTERIA *CLOSTRIDIUM*  
*SCINDENS* AND *CLOSTRIDIUM HYLEMONAE*

A dissertation submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy at the  
Virginia Commonwealth University

By

Jason Michael Ridlon  
B.S., Bridgewater College, 2002

Major Director: Dr. Phillip B. Hylemon  
Professor of Microbiology and Immunology and Medicine

Virginia Commonwealth University  
Richmond, Virginia  
May, 2008

**DEDICATION**

To my loving wife Liz for her constant support and encouragement to succeed at this tremendous undertaking. To my mother and grandmother for all their love and support.

## **ACKNOWLEDGEMENTS**

I would like to express my utmost and sincere gratitude to Dr. Phillip B. Hylemon for his guidance, patience and support before and during this work. I cherish our many wonderful conversations on science, life and current events. I would also like to extend many thanks to the members of my graduate committee, Drs. D. Peterson, C. Cornelissen, R. Marconi and W.M. Pandak for all of their encouragement and valuable suggestions.

I would also like to thank the Hylemon lab, particularly Pat Bohden, Emily Gurley and Elaine Studer for keeping the lab running and for their friendship and support.

I must give special thanks and appreciation to DJ Kang for his friendship, constant help and support and his expertise on protein expression and purification.

I would also like to thank Dr. Hylemon and the Department of Microbiology and Immunology whose financial support allowed completion of this project.

Man's liver is a brownish blob  
That does a most prodigious job.  
It manufactures gall, or bile  
And normally keeps some on file  
Stored neatly in a pear-shaped sac.  
From there the liver's yields attack  
The food man eats, to change its state  
By methods man can't duplicate,  
Or even halfway understand.  
He ought to treat this outsize gland,  
With due respect and loving care  
To keep it in top-notch repair,  
Because to get along at all  
Man needs an awful lot of gall.

*Irene Warsaw, JAMA, 1975, v. 231, p. 1260*

The known is finite, the unknown infinite;  
intellectually we stand on an islet in the midst of an  
illimitable ocean of inexplicability. Our business in  
every generation is to reclaim a little more land.

- T.H. Huxley, 1887

## Table of Contents

	Page
Acknowledgements.....	ii
List of Tables.....	viii
List of Figures.....	ix
Abstract.....	xvii
Section	
1 Introduction .....	1
The Enterohepatic Circulation of Bile Acids .....	2
Deconjugation of Bile Salts .....	11
Bile Acid Dehydrogenation by Stereospecific	
Hydroxysteroid Dehydrogenases.....	26
Biochemistry and molecular biology of	
bile acid 7 $\alpha$ / $\beta$ -dehydroxylation .....	40
Secondary Bile Acids and Disease .....	60
Current trends and Future Issues .....	65
Statement of research problem and objectives ....	70
2 Materials and methods .....	74
Bacterial strains, culture media,	
and conditions .....	74
General PCR and recombinant DNA methods .....	75
Polyacrylamide gel electrophoresis and	



Immunoblot Analysis .....	79
Enzyme Assays .....	80
Expression and purification of recombinant <i>baiH</i> Gene Product .....	83
Expression and purification of <i>baiJ</i> gene product .....	85
Enzymatic and biological preparation of bile acid Substrates .....	89
Chemical synthesis of bile acid substrates .....	91
Purification and mass spectrometry of BaiCD And BaiH Reaction products .....	94
Genome-walking by polymerase chain reaction .....	95
SMART RACE polymerase chain reaction .....	97
Sequence analysis and assembly .....	99
Nucleotide sequence accession numbers .....	101
3 Results .....	103
Characterization of the BaiCD and BaiH .....	103
Isolation of <i>bai</i> genes from <i>Clostridium</i> <i>hylemonae</i> TN271 .....	125
Identification and characterization of novel genes involved in bile acid metabolism in <i>Clostridium hylemonae</i>	

	vii
and <i>Clostridium scindens</i> .....	166
4 Discussion .....	261
References.....	289
Vita.....	333

## LIST OF TABLES

Page	
Table 1: Characteristics of bile salt hydrolases (BSH) from intestinal bacteria.....	13
Table 2: Characteristics of hydroxysteroid dehydrogenases (HSDH) from intestinal bacteria.....	29
Table 3: Bile acid Inducible ( <i>bai</i> ) genes characterized from <i>Clostridium scindens</i> VPI 12708.....	59
Table 4: Characteristics of selected BaiH homologues.....	104
Table 5: Purification scheme of <i>baiH</i> gene protein in <i>E.coli</i> BL21(DE3)-RIL(pSportI- <i>baiH</i> ).....	112
Table 6: Oligonucleotides Used in isolating <i>baiBCDEFGHI</i> operon from <i>Clostridium hylemonae</i> TN271.....	158
Table 7: Comparison of <i>Clostridium hylemonae</i> TN271 and <i>Clostridium scindens</i> VPI 12708 <i>bai</i> operon DNA Sequences.....	159
Table 8: Oligonucleotides used in isolation of <i>baiA</i> and flanking regions in <i>Clostridium hylemonae</i> TN271.....	177
Table 9: Oligonucleotides used in isolation of <i>baiJKL</i> operon and flanking sequence from <i>Clostridium scindens</i> VPI 12708.....	200

## LIST OF FIGURES

Figure	Page
Figure 1: Bacterial bile salt-biotransforming reactions in the human intestinal tract.....	3
Figure 2: Anatomy, physiology, and microbiology of the gastrointestinal tract.....	7
Figure 3: Composition of bile acids in the gallbladder and feces of healthy individuals.....	9
Figure 4: Multiple sequence alignment of cholylglycine hydrolases.....	14
Figure 5: Proposed catalytic mechanism of bile acid 7 $\alpha$ - Dehydrogenation.....	34
Figure 6: Accumulation of [24- <sup>14</sup> C] CA intermediates during 7 $\alpha$ -dehydroxylation in cell extracts of <i>C. scindens</i> VPI 12708.....	43
Figure 7: Gene organization of the bile acid-inducible (bai) 7 $\alpha$ / $\beta$ -dehydroxylation operons characterized in <i>C. scindens</i> VPI 12708.....	46
Figure 8: Proposed bile acid 7 $\alpha$ -dehydroxylation pathway in <i>C. scindens</i> VPI 12708.....	51
Figure 9: Model of the bile acid 7 $\alpha$ -dehydratase.....	54
Figure 10: Putative binding/active site pocket of bile acid 7 $\alpha$ -dehydratase.....	56
Figure 11: Relationship between the percentage of CDCA and DCA in bile.....	61

Figure 12: Schematic representation of cloning strategy for <i>baiH</i> gene from <i>Clostridium scindens</i> VPI 12708.....	76
Figure 13: Cloning strategy for <i>baiL</i> gene from <i>C. scindens</i> VPI 12708.....	86
Figure 14: Amino acid sequence alignment of the BaiH from <i>C. scindens</i> VPI 12708 with protein homologues.....	106
Figure 15: SDS-PAGE and immunoblot analysis of <i>baiH</i> gene product(NADH:FOR) overexpression and purification.....	110
Figure 16: TLC autoradiograph of BaiCD and BaiH catalyzed NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity assays.....	114
Figure 17: HPLC and mass spectrometry analysis of BaiH catalyzed bile acid reaction products.....	116
Figure 18: SDS-PAGE of cell-extract of recombinant <i>E. coli</i> BL21(DE3) overexpressing <i>baiCD</i> gene.....	119
Figure 19: HPLC and mass spectrometry analysis of BaiCD catalyzed bile acid reaction products.....	121
Figure 20: Proposed reactions catalyzed by the <i>baiCD</i> and <i>baiH</i> gene product.....	123
Figure 21: Region of partial <i>baiCD</i> gene from <i>C. hylemonae</i> TN271 used to design upstream and downstream genome-walking PCR primers.....	127
Figure 22: Preparation of genomic DNA for genome-walking	

by PCR restriction libraries.....	130
Figure 23: Genome-walking by PCR Upstream of partial <i>baiCD</i> gene in <i>Clostridium hylemonae</i> TN271.....	132
Figure 24: Second upstream genome-walking by PCR from partial <i>baiCD</i> gene in <i>Clostridium hylemonae</i> TN271.....	135
Figure 25: Boxshade of CLUSTALW alignment of <i>bai</i> promoters.....	137
Figure 26: SMART RACE PCR analysis to determine mRNA initiation site of <i>baiB</i> transcript in <i>Clostridium hylemonae</i> TN271.....	139
Figure 27: Determination of transcription initiation site for <i>baiB</i> gene from <i>Clostridium scindens</i> VPI 12708.....	142
Figure 28: ClustalW alignment of bile acid regulatory A gene ( <i>barA</i> ).....	144
Figure 29: Downstream genome-walk (DS1) from <i>baiCD</i> gene from <i>Clostridium hylemonae</i> TN271.....	146
Figure 30: Boxshade of CLUSTALW alignment of 7 $\alpha$ -dehydratase gene ( <i>baiE</i> ).....	149
Figure 31: Second downstream genome-walk (DS2) from <i>baiCD</i> gene from <i>Clostridium hylemonae</i> TN271.....	151
Figure 32: Third downstream genome-walk (DS3) from <i>baiCD</i> gene of <i>Clostridium hylemonae</i> TN271.....	154
Figure 33: Schematic representation and comparison of the	

<i>bai</i> oxidative operons of <i>C. hylemonae</i> TN721 and <i>C. scindens</i> VPI 12708.....	156
Figure 34: Autoradiographs of TLC of 7 $\alpha$ -dehydroxylation and 7 $\beta$ -dehydroxylation activity assay reaction products.....	161
Figure 35: Design of degenerate PCR oligonucleotides for isolation of the <i>baiA</i> gene in <i>C. hylemonae</i> TN271.....	164
Figure 36: Isolation of partial <i>baiA</i> gene from <i>C. hylemonae</i> TN271.....	167
Figure 37: Isolation of <i>baiA</i> gene from <i>C. hylemonae</i> TN271 and characterization of upstream promoter elements.....	169
Figure 38: CLUSTALW alignment of amino acid sequence of <i>baiA</i> genes from <i>C. hylemonae</i> TN271, <i>C. scindens</i> VPI 12708 and <i>C. hiranonis</i> sp. strain T0931.....	171
Figure 39: Genome-walking PCR upstream of partial <i>baiA</i> gene of <i>Clostridium hylemonae</i> TN271.....	173
Figure 40: Downstream genome-walking PCR (DS1) from <i>baiA</i> gene of <i>Clostridium hylemonae</i> TN271.....	175
Figure 41: Second genome-walking PCR upstream of <i>baiA</i> from <i>C. hylemonae</i> TN271.....	179
Figure 42: Boxshade of CLUSTALW amino acid sequence alignment of <i>baiF</i> gene product of <i>C. hylemonae</i> and a predicted Type III CoA transferase which	

named the <i>baiK</i> gene.....	182
Figure 43: Third upstream genome-walking PCR from <i>baiA</i> gene of <i>C. hylemonae</i> TN271.....	184
Figure 44: Fourth upstream genome-walking PCR from <i>baiA</i> gene of <i>Clostridium hylemonae</i> TN271.....	186
Figure 45: Identification of conserved <i>bai</i> promoter Elements upstream of the putative <i>baiJ</i> gene of <i>Clostridium hylemonae</i> TN271.....	189
Figure 46: Determination of transcription initiation site for <i>baiJ</i> gene from <i>Clostridium hylemonae</i> TN271 based on SMART RACE PCR analysis.....	191
Figure 47: Amino acid sequence of <i>baiJ</i> gene from <i>Clostridium hylemonae</i> TN271.....	194
Figure 48: Codon usage table for <i>Clostridium scindens</i> VPI 12708 and design of redundant oligonucleotides for isolation of the <i>baiJ</i> gene from <i>C. scindens</i> VPI 12708.....	196
Figure 49: Isolation of partial <i>baiJ</i> gene from <i>Clostridium</i> <i>scindens</i> VPI 12708.....	198
Figure 50: Genome-walking PCR upstream of partial <i>baiJ</i> gene of <i>Clostridium scindens</i> VPI 12708.....	202
Figure 51: BOXSHADE of CLUSTALW alignment of upstream conserved <i>bai</i> promoter elements located upstream of <i>baiJ</i> genes of <i>Clostridium scindens</i> VPI 12708 and <i>Clostridium hylemonae</i> TN271.....	204
Figure 52: Verification of bile acid induction and 5'	



SMART RACE PCR analysis of <i>baiJ</i> transcript from <i>C. scindens</i> VPI 12708.....	206
Figure 53: Second upstream genome-walking PCR (UP2) from partial <i>baiJ</i> gene of <i>Clostridium scindens</i> VPI 12708.....	209
Figure 54: Downstream genome-walking PCR (DS1) from Partial <i>baiJ</i> gene of <i>Clostridium scindens</i> VPI 12708.....	212
Figure 55: Boxshade representation of CLUSTALW amino acid sequence alignment of putative 3-oxo- $\Delta^4$ -reductases ( <i>baiJ</i> ) between <i>Clostridium scindens</i> VPI 12708 and <i>Clostridium hylemonae</i> TN271.....	214
Figure 56: Third genome-walking PCR downstream (DS3) of <i>baiJ</i> gene product in <i>Clostridium</i> <i>scindens</i> VPI 12708.....	217
Figure 57: Genome-walking PCR downstream (DS4) of partial <i>baiJ</i> gene of <i>Clostridium scindens</i> VPI 12708.....	219
Figure 58: Schematic representation of putative reductive genes from <i>Clostridium scindens</i> VPI 12708 and <i>Clostridium hylemonae</i> TN271.....	221
Figure 59: Boxshade of CLUSTALW alignment between the <i>baiJ</i> gene product from <i>Clostridium scindens</i> VPI 12708 and flavoprotein homologues.....	225
Figure 60: Codon Usage Table for <i>Clostridium hiranonis</i> sp. strain T0931.....	227
Figure 61: Isolation of putative <i>baiJ</i> gene from	

Figure 60: Codon Usage Table for <i>Clostridium hiranonis</i> sp. strain T0931.....	227
Figure 61: Isolation of putative <i>baiJ</i> gene from <i>Clostridium hiranonis</i> sp. strain T0931.....	229
Figure 62: Boxshade of CLUSTALW amino acid sequence alignment of <i>baiL</i> genes.....	232
Figure 63: Cloning strategy of <i>baiL</i> gene from <i>C. scindens</i> VPI 12708.....	234
Figure 64: Overexpression of recombinant <i>baiL</i> gene from <i>Clostridium hylemonae</i> TN271 in <i>E. coli</i> BL21 (DE3).....	236
Figure 65: Western blot analysis of <i>baiL</i> expression.....	239
Figure 66: SDS-PAGE and Western immunoblot of elution fraction after biotin binding chromatography of overexpression of recombinant BaiJ-SBP fusion protein.....	242
Figure 67: Measurement of radioactivity in both aqueous and organic phase before and after sodium hydroxide treatment.....	245
Figure 68: Autoradiograph of a thin layer chromatography of BaiJ reaction products.....	248
Figure 69: Codon usage table for <i>Clostridium hylemonae</i> TN271.....	251

Figure 70: Design of degenerate oligonucleotides for isolation of the <i>gyrA</i> gene from <i>C. hylemonae</i> TN271...	253
Figure 71: PCR amplification of <i>gyrA</i> gene from <i>Clostridium</i> <i>hylemonae</i> TN271.....	256
Figure 72: Boxshade alignment of <i>gyrA</i> gene product from <i>C.</i> <i>scindens</i> VPI 12708 with other clostridial and bacterial <i>gyrA</i> genes.....	258
Figure 73: Logo representation of <i>bai</i> promoter regions...	276
Figure 74: Boxshade alignment of <i>baiF</i> and <i>baiK</i> genes from <i>C. scindens</i> VPI 12708 and <i>C. hylemonae</i> TN271 with homologues in the Type III CoA transferase family.....	278

## **ABSTRACT**

### **ENZYMOLGY AND MOLECULAR BIOLOGY OF BILE ACID 7 $\alpha$ - AND 7 $\beta$ - DEHYDROXYLATION BY THE INTESTINAL BACTERIA *CLOSTRIDIUM* *SCINDENS* AND *CLOSTRIDIUM HYLEMONAE***

Jason Michael Ridlon

A dissertation submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy at the  
Virginia Commonwealth University

Virginia Commonwealth University, 2008.

Director: Dr. Phillip B. Hylemon  
Professor of Microbiology and Immunology  
and Medicine.

The collective microbial genomes within our gut (microbiome) represent a powerful metabolic force, leading many authors to call our GI flora an "organ within an organ", and the metagenomic sequencing of our microbiome, "the second human genome project". Bile acids, endogenously produced by the host liver, represent both a strong selective pressure for potential colonizers, as well as substrates for microbial metabolism. Indeed, microbes have evolved enzymes to deconjugate bile salts, epimerize bile acid hydroxyl groups, and 7 $\alpha$ -dehydroxylate primary bile acids. The products of microbial 7 $\alpha$ -

dehydroxylation, secondary bile acids, are suggested by numerous lines of evidence to be involved in promoting colon carcinogenesis. 7 $\alpha$ -dehydroxylating activity is a multi-step pathway, genes of which have only been identified in a small number of organisms within the genus *Clostridium*. The biochemistry of this pathway has been largely worked out. The third step in the pathway is introduction of a  $\Delta^4$ -double bond; however, the gene product(s) responsible have not been identified. The *baiCD* and *baiH* genes were cloned, expressed and shown to have NAD<sup>+</sup>-dependent 3-oxo- $\Delta^4$ -steroid oxidoreductase activity showing stereospecificity for 7 $\alpha$ -hydroxy and 7 $\beta$ -hydroxy bile acid, respectively.

In addition, *bai* genes were isolated from *C. hylemonae* TN271 by bidirectional genome-walking by PCR. This represents the first report of *bai* genes from a "low activity" 7 $\alpha$ -dehydroxylating bacterium. The gene organization and sequence of the *baiBCDEFGHI* operon was highly conserved between *C. hylemonae* TN271 and the "high activity" 7 $\alpha$ -dehydroxylating bacterium *C. scindens* VPI 12708. The *baiA* gene was located by PCR using degenerate oligonucleotides. Bi-directional genome-walking revealed what appears to be several novel genes involved in bile acid metabolism which were also located in *C. scindens* VPI 12708. Expression of a 62 kDa flavoprotein and reaction with [24-<sup>14</sup>C] 3-oxo-DCA and NADP resulted in a product of greater hydrophilicity than deoxycholic acid. The identity of this product was not determined. A second gene appears to share a common evolutionary origin with

the *baiF* gene. A hypothesis is offered regarding the function of these homologues as Type III CoA transferases recognizing 5 $\beta$ -bile acids, or 5 $\alpha$ -bile acids (allo-bile acids). A third gene encodes a putative short chain reductase, similar in size and predicted function to the *baiA* gene, which may be involved in the final reductive step in the pathway. These novel genes also contained a conserved upstream regulatory region with the *bai* oxidative genes. Finally, two genes were identified which may serve as potential drug targets to inhibit bile acid 7 $\alpha$ -dehydroxylation. The first is an ABC transporter which may be co-transcribed with the other novel bile acid metabolizing genes, and what appears to be a bile acid sensor/regulator similar to the Tryptophan-rich sensory protein (TspO)/mitochondrial peripheral benzodiazepine receptor (MBR) family of proteins.

## INTRODUCTION

### **Bile salt metabolism by intestinal bacteria**

The human large intestine harbors a complex microbial flora (1). Bacterial density in the human colon is among the highest found in nature, approaching  $10^{12}$  bacteria/ gram wet weight feces (2, 3). In contrast, the host suppresses significant bacterial colonization of the small intestine by a variety of mechanisms including rapid transit times, antimicrobial peptides, proteolytic enzymes, and bile (4). Failure of these mechanisms leads to bacterial overgrowth of the small intestine, resulting in malabsorption as bacteria compete with the host for nutrients. Under normal conditions, bacterial fermentation in the colon represents an important salvage mechanism. Complex carbohydrates, which are intrinsically indigestible, or which escape digestion and absorption in the proximal gut, are fermented by colonic bacteria to yield short chain fatty acids. It has been estimated that these short-chain fatty acids constitute 3-9% of our daily caloric intake (4). Colonic bacteria also contribute to salvage of bile salts that escape

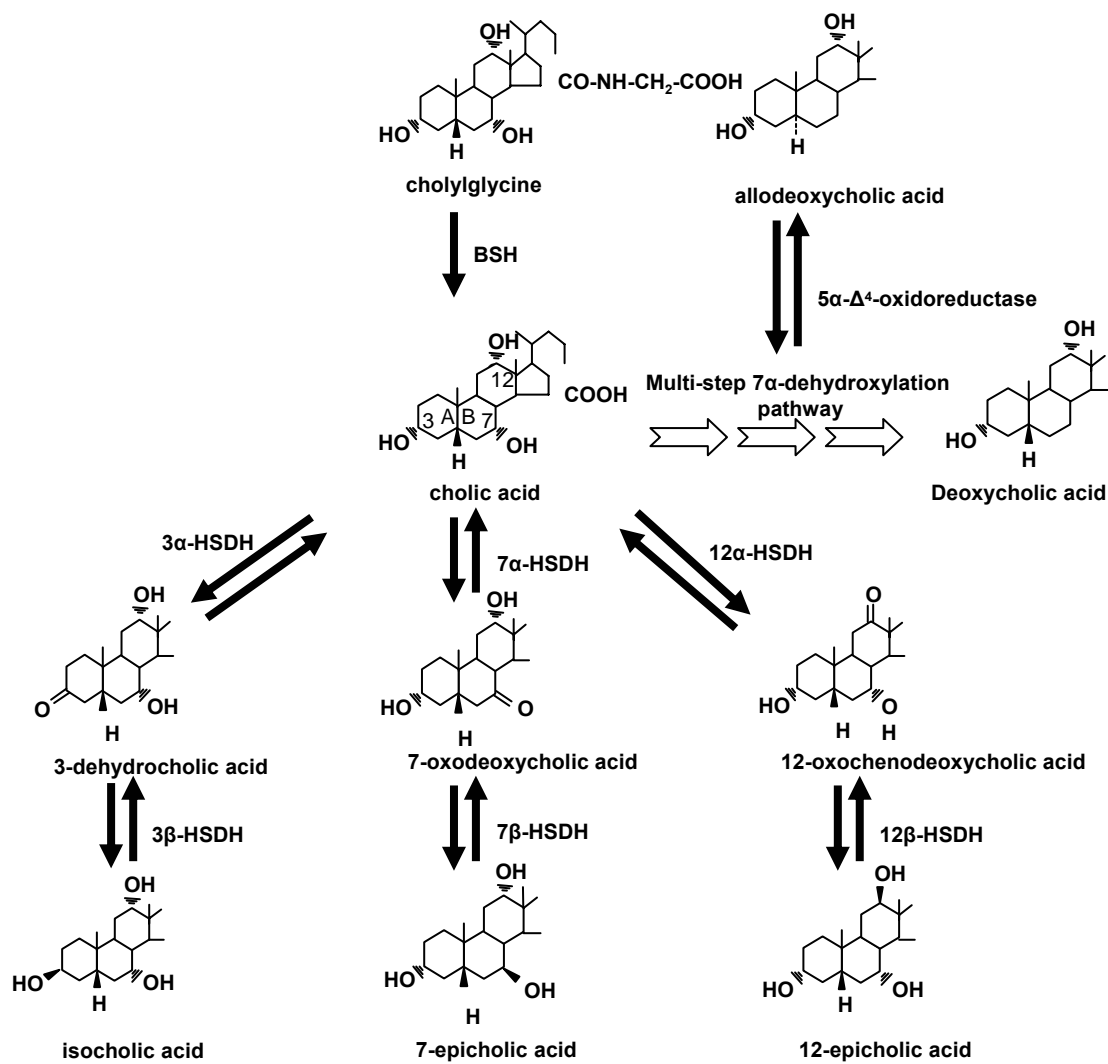
active transport in the distal ileum. The major bile salt modifications in the human large intestine include deconjugation, oxidation of hydroxy groups at C-3, C-7 and C-12, and 7 $\alpha$ / $\beta$ -dehydroxylation (Figure 1). Deconjugation and 7 $\alpha$ / $\beta$ -dehydroxylation of bile salts increases their hydrophobicity and raises their pKa, thereby permitting their recovery via passive absorption across the colonic epithelium. However, the increased hydrophobicity of the transformed bile salts also is associated with increased toxic and metabolic effects. Elevated concentrations of secondary bile acids in feces, blood and bile have been linked to the pathogenesis of cholesterol gallstone disease and colon cancer (5). The following is a current review of the microbiology of bile acid metabolism in the human gastrointestinal tract focusing on understanding the biochemical mechanisms and physiological consequences of such metabolism on both the bacterium and the human host.

### **The enterohepatic circulation of bile acids**

Bile acids are saturated, hydroxylated C-24 cyclopentanepheneanthrene sterols synthesized from cholesterol in hepatocytes. The two primary bile acids



Figure 1. Bacterial bile salt-biotransforming reactions in the human intestinal tract. Hydroxy group carbons of cholate are numbered and the AB rings are identified. The 3, 7, and 12 carbons of cholic acid (CA) are numbered. Nomenclature is that of Hofmann et al. (160). BSH, bilesalt hydrolase; HSDH, hydroxysteroid dehydrogenase.

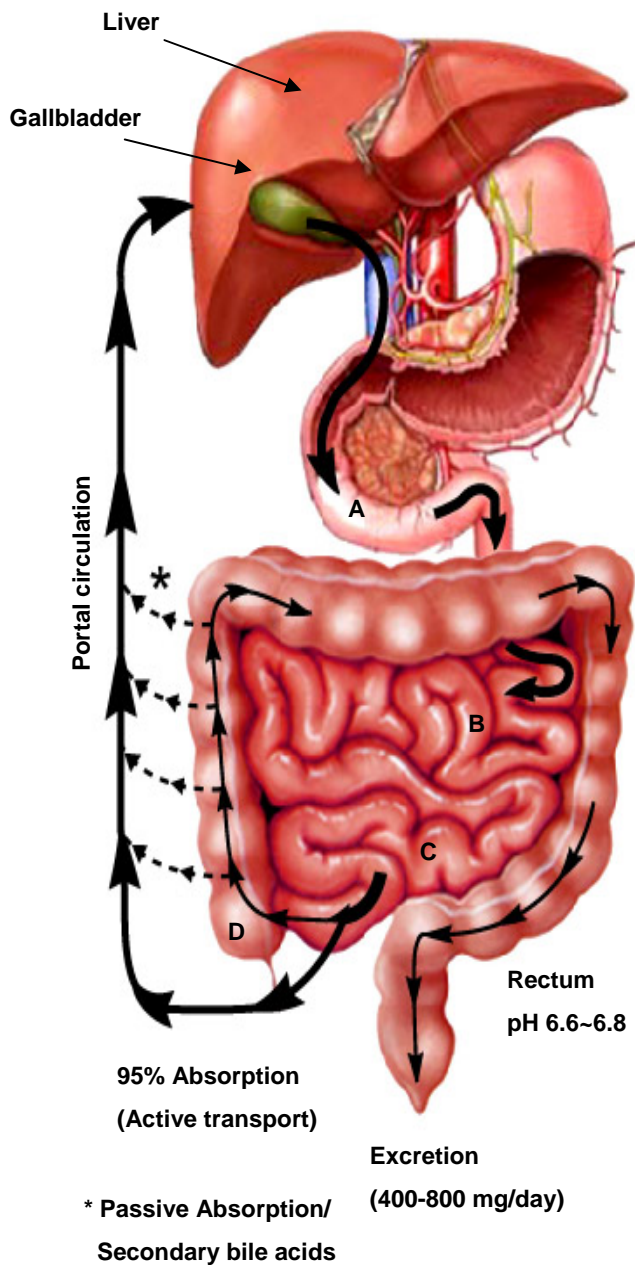


synthesized in the human liver are cholic acid (CA) ( $3\alpha$ ,  $7\alpha$ ,  $12\alpha$ -trihydroxy- $5\beta$ -cholan-24-oic acid) and chenodeoxycholic acid (CDCA) ( $3\alpha$ ,  $7\alpha$ -dihydroxy- $5\beta$ -cholan-24-oic acid). Bile acids are further metabolized by the liver via conjugation (N-acyl amidation) to glycine or taurine, a modification that lowers the pKa to around 5. Thus, at physiological pH, conjugated bile acids are almost fully ionized, and may be termed bile salts (6). Bile salts are secreted actively across the canalicular membrane and are carried in bile to the gallbladder where they are concentrated during the inter-digestive period. Following a meal, release of cholecystokinin from the duodenum stimulates the gallbladder to contract, causing bile to flow into the duodenum (7). Bile salts are highly effective detergents that promote solubilization, digestion and absorption of dietary lipids, and lipid soluble vitamins throughout the small intestine. High concentrations of bile salts are maintained in the duodenum, jejunum and proximal ileum, where fat digestion and absorption take place. Bile salts are then absorbed through high-affinity active transport in the distal ileum (6). Upon entering the bloodstream, bile salts are

complexed to plasma proteins and returned to the liver. Upon reaching the liver they are cleared efficiently from the circulation by active transporters on the sinusoidal membrane of hepatocytes and rapidly secreted into bile. This process is known as the enterohepatic circulation. Figure 2 depicts the enterohepatic circulation in the context of the gastrointestinal anatomy and also indicates the relative numbers and genera of the predominant bacteria inhabiting each section of the GI tract.

During the enterohepatic circulation, bile salts encounter populations of facultative and anaerobic bacteria of relatively low numbers and diversity in the small bowel. Bile salt metabolism by small bowel microbes consists mainly of deconjugation and hydroxy group oxidation. Ileal bile salt transport is highly efficient (~95%), but approximately 400-800 mg of bile salts escape the enterohepatic circulation daily and become substrate for significant microbial biotransforming reactions in the large bowel (6). Comparison of bile acid composition in the gallbladder and feces illustrates the extent of microbial bile acid metabolism in the large intestine (Figure 3). Secondary bile acids deoxycholic acid (DCA)

Figure 2. Anatomy, physiology, and microbiology of the gastrointestinal tract. Large arrows denote the enterohepatic circulation of bile acids, which begins with contraction of the gallbladder, releasing bile into the duodenum. Small arrows denote the passive absorption of bile acids that escape active transport. \* Secondary bile acids produced by 7 $\alpha$ -dehydroxylation are passively absorbed in the large intestine and returned to the liver (see Fig 3). The genera of predominant bacteria isolated from each region of the lower gastrointestinal tract are listed.



### Small Intestine

**A: Duodenum (25cm)**  
**pH 5.7~6.4**  
 $\sim 10^3$  bacteria/ml  
*Lactobacillus*  
*Streptococcus*

**B : Jejunum (1.0m)**  
**pH 5.9~6.8**  
 $\sim 10^4$  bacteria/ml  
*Lactobacillus*  
*Streptococcus*  
*Staphylococcus*  
*Veillonella*

**C : Ileum (2.0m)**  
**pH 7.3~7.7**  
 $10^6 \sim 10^8$  bacteria/ml  
*Enterobacteria*  
*Enterococcus*  
*Bacteroides*  
*Clostridium*  
*Lactobacillus*  
*Veillonella*

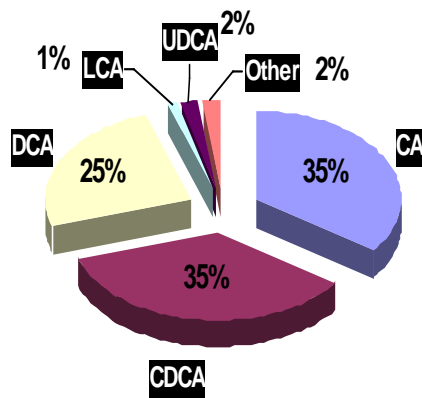
### Large Intestine

**D: Cecum/Colon (150cm)**  
**pH 5.7-6.8**  
 $\sim 10^{11}$  bacteria/g

*Bacteroides*, *Eubacterium*,  
*Bifidobacterium*,  
*Ruminococcus*,  
*Peptostreptococcus*,  
*Propionibacterium*,  
*Clostridium*, *Lactobacillus*,  
*Escherichia*,  
*Streptococcus*,  
*Methanobrevibacter*

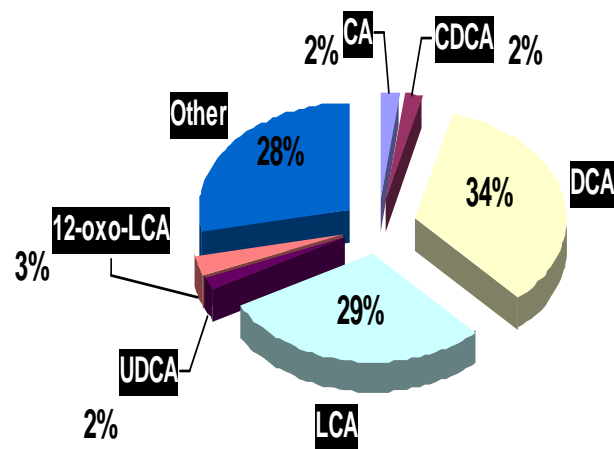
Figure 3. Composition of bile acids in the gallbladder and feces of healthy individuals. ``Other'' bile acids refer to oxo- and  $3\beta$ -hydroxy derivatives of secondary bile acids. Values were derived from published sources (6, 93, 161, 162). CDCA, chenodeoxycholic acid; DCA, deoxycholic acid; LCA, lithocholic acid; UDCA, ursodeoxycholic acid.

### Biliary Bile acid Composition



Bacterial enzymes

### Fecal Bile acid composition





(3 $\alpha$ , 12 $\alpha$ -dihydroxy-5 $\beta$ -cholan-24-oic acid) and lithocholic acid (LCA) (3 $\alpha$ -hydroxy-5 $\beta$ -cholan-24-oic acid) are produced solely by microbial biotransforming reactions in the human large intestine. DCA accumulates in the bile acid pool (LCA to a much lesser extent) due to passive absorption through the colonic mucosa and the inability of the human liver to 7 $\alpha$ -hydroxylate DCA and LCA to their respective primary bile acids. LCA is sulfated in the human liver at the 3-hydroxy position, conjugated at C-24 and excreted back into bile (6). The resultant bile acid sulfate is poorly reabsorbed from the gut. Even though 3-sulfo-LCA glycine and taurine conjugates are deconjugated and to some extent desulfated by intestinal bacteria, 3-sulfo-LCA/LCA is lost in feces and does not normally accumulate in the enterohepatic circulation (8).

### **Deconjugation of Bile Salts**

#### **Characteristics of bile salt hydrolase(s) (BSH)**

Deconjugation refers to the enzymatic hydrolysis of the C-24 N-acyl amide bond linking bile acids to their amino acid conjugates. This reaction is substrate limiting and goes to completion in the large bowel. BSH

are in the choloylglycine hydrolase family (EC 3.5.1.24) and have been isolated and/or characterized from several species of intestinal bacteria (Table 1). The importance of the position, charge, shape, and chirality of various analogues of taurine/glycine conjugates on the rate of hydrolysis by BSH has also been investigated (9). BSH differ in subunit size and composition, pH optimum, kinetic properties, substrate specificity, gene organization and regulation. BSH do, however, share in common several conserved active site amino acids (Cys2, Arg18, Asp21, Asn175, and Arg228) and share a high degree of amino acid sequence similarity with the penicillin V amidase (PVA) of *Bacillus sphaericus* (Figure 4). The conservation of Tyr82 in PVA and Asn82 in BSH are likely a result of differing steric requirements for their respective substrates (10). Recently, a *bsh* from *Clostridium perfringens* has been crystallized both in the apo-enzyme form and in complex with TDCA (hydrolyzed product) at resolutions of 2.1 and 1.7 Å, respectively (11). The structure revealed the Cys2 residue is in position for nucleophilic attack of the N-acyl amide bond. Site directed mutagenesis of the Cys2 residue from the BSH

**Table 1. Characteristics of bile salt hydrolases (BSH) from intestinal bacteria**

Organism	MW (kDa) Native	MW (kDa) Subunits	Apparent $K_m$ (mM) <sup>a</sup>						pH optimum	Reference
			(TCA)	(TCDCA)	(TDCA)	(GCA)	(GCDCA)	(GDCA)		
<i>Bacteroides fragilis</i>	250	32.5	0.45	0.29	0.17	0.35	0.26	0.2	4.2-4.5	25
<i>Bacteroides vulgatus</i>	140	36	+	+	+	-	-	-	ND	162
<i>Clostridium perfringens</i> MCV 185	250	56	37	3	3.5	3.6	14	1.2	5.8-6.4	12
<i>Clostridium perfringens</i> 13	147	36.1	+	ND	+	+	ND	ND <sup>b</sup>	4.5-5.5	14
<i>Lactobacillus johnsonii</i> 100-100		42( $\alpha$ ); 38( $\beta$ )								16,17, 23
isozyme A	115	42	0.76	+	+	+	ND	+	4.2-4.5	
isozyme B	105	42,38	0.95	+	+	+	ND	+	4.2-4.5	
isozyme C	95	42,38	0.45	ND	ND	ND	ND	ND	4.2-4.5	
isozyme D	80	38	0.37	ND	ND	ND	ND	ND	4.2-4.5	
<i>Lactobacillus plantarum</i> 80	ND	37.1 <sup>c</sup>	TR	TR	TR	+	+	+	4.7-5.5	15
<i>Lactobacillus acidophilus</i>	ND	35.0 <sup>c</sup>	+	ND	ND	+	ND	ND	ND	17, 27
<i>Bifidobacterium longum</i> BB536	250	40	0.875	1.61	0.516	1.33	1.4	1.37	5.5-6.5	13
<i>Bifidobacterium longum</i> SBT2928	125-130	35	1.12	0.33	0.79	0.16	0.13	0.28	5.0-7.0	10
<i>Bifidobacterium bifidum</i> ATCC 11863	150	35	+	+	+	++	++	++	ND	18
<i>Bifidobacterium adolescentis</i>	ND	35 <sup>c</sup>	ND	ND	ND	ND	ND	ND	ND	19
<i>Listeria monocytogenes</i>	ND	36.8 <sup>c</sup>	ND	ND	ND	ND	ND	+	ND	20,163

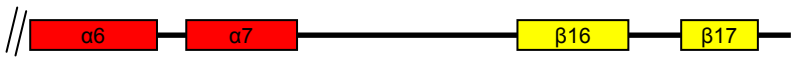

<sup>a</sup> TCA, taurocholate; GCA, glycocholate; TDCA, taurodeoxycholate; GDCA, glycodeoxycholate; TCDCA, taurochenodeoxycholate; GCDCA, glycochenodeoxycholate; + activity detected; - no activity detected; TR, trace of activity.

<sup>b</sup> ND, not determined.

<sup>c</sup> Value derived from Protparam program (<http://www.expasy.ch/tools/protparam.html>) using deduced amino acid sequence.

Figure 4. Multiple sequence alignment of cholyglycine hydrolases. Protein sequences were obtained from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Alignments were made with the ClustalW program (<http://www.ebi.ac.uk/clustalw/>) using the GONNET 250 matrix. Residues highlighted in yellow are predicted active site amino acids based on the crystal structure of the BSH from *C. perfringens* (11) as well as on site-directed mutagenesis and biochemical data (10, 12-14). Residues highlighted in gray correspond to residues involved in substrate binding in the BSH from *C. perfringens* (11). The secondary structural elements, which are based on the conjugated bile acid hydrolase from *C. perfringens* (CBAH-1) crystal structure, are shown above the alignment. The a and b designations of *Lactobacillus johnsonii* refer to the two isoforms of the genes found in this bacterium.

<i>Cl.perfringens</i>	M--CTGLELETKDGLHLFGNNMIEYSNQSIIFIPRNFKCVNKS--KELTITKYAVLGM	57
<i>Li.monocytogenes</i>	M--CTSITYTTK--DHYFGRNFDYELSYKEVVVVTTPKNYPFHFRK--VEDIEKHIALIGI	54
<i>En.faecalis</i>	M--CTAITYVSK--DHYFGRNFDYELSYNEVVITTPRNYKFSFRE--VGNLDHHAFAIIGI	54
<i>La.plantarum</i>	M--CTAITYQSY--NNYFGRNFDYELSYNEMVITTPRKYPLVFRK--VENLDHHAFAIIGI	54
<i>La.gasseri</i>	M--CTSILYSPK--DHYFGRNFDYELSYGQKVITTPKNYEFETD--LPAEKSHYAMIGV	54
<i>La.johnsonii</i> (α)	M--CTSIVYSSNN--HHYFGRNLDLEISFGEHPVITTPRNYEFQYRK--LPNKKAKYAMVGM	55
<i>Bi.longum</i>	M--CTGVRFSDEGNTYFGRNLDWSFSYGETILVTPRGYHYDTVFG--AGGKAKPNAVIGV	57
<i>Bi.adolescentis</i>	M--CTGVRFSDEEGNMYFGRNLDWSFSYGESILATPRGYHYDNVFG--AERKATPNAVIGV	57
<i>Bi.bifidum</i>	M--CTGVRFSDEGNTYFGRNLDWSFSYGETILVTPRGYQYDYEYG--AEGKSEPNNAVIGV	57
<i>La.johnsonii</i> (β)	M--CTGLRFTDDQGNLYFGRNLDVGQDYGEGVITTPRNYPLPYKF--LDNTTTKKAVIGM	56
<i>La.acidophilus</i>	M--CTGLRFTDDQGNLYFGRNLDVGQDYGEGVITTPRNYPLPYKF--LDNTTTKKAVIGM	56
<i>Ba.sphaericus</i>	MLGSSSLIRTTDDKSLFARTMDFTMEEDSKVIVPRNYGIRLLEKENVVNNNSYAFVGM	60
<i>Cl.perfringens</i>	GTIEDDYPTFADGMNEKGLGCAGLNFPVYVSYSKEDIEGKTNIPVYNFLWLVLANFSSVE	117
<i>Li.monocytogenes</i>	AAVMENYPLYDATNEKGLSMAGLNFSGNADYKDFAE--KDNVTPFEFIPWILGQCATVK	113
<i>En.faecalis</i>	AAGTADYPLYDAINEKGLGMAGLNFSGYADYKKIEEG--KENVSPFEFIPWVLGQCSTVD	113
<i>La.plantarum</i>	TADVESYPLYDAMNEKGLCIAGLNFGAGYADYKKYDAD--KVNITPFELIPWLLGQFSSVR	113
<i>La.gasseri</i>	AAVADNTPLYCDAINEKGLGVAGLSFAGQGYFPNAAD--KKNIASFEFISYLLATYETVD	113
<i>La.johnsonii</i> (α)	AIVEDNYPLYFDASNEEGLGIAGLNFDGPGCHYFENAE--KNNVTPFELIPYLLSQCTTVA	114
<i>Bi.longum</i>	GVVMADRPYFDCANEHGLAIAGLNFGYASVHEPVEGTENVATFEFPLWVARNFDSVD	117
<i>Bi.adolescentis</i>	GVVMADRPYFDCANEHGLAIAGLNFGYAEFVHEPVEGTENVATFEFPLWVARNFDSVD	117
<i>Bi.bifidum</i>	GVVMADRPYFDCANEHGLAIAGLNFGYASFAHEPVEGTENVATFEFPLWVARNFDSVD	117
<i>La.johnsonii</i> (β)	GIVVDGYPSYFDCYNEDGLGIAGLNFPFHAKFS--DGPIDGKINLASYEIMLWVTQNFTHVS	116
<i>La.acidophilus</i>	GIVVDGYPSYFDCFNEDGLGIAGLNFPFHAKFS--DGPIDGKINLASYEIMLWVTQNFTHVS	116
<i>Ba.sphaericus</i>	GSTIITSPVLYDGVNEKGLMGAMLYATFATYADEPKKGTGINPVYVISQVLGNCVTV	120
<i>Cl.perfringens</i>	EVKEALKNANIVDIPSENIPNTLHWMISDITGKSIVVEQTKE--KLNVDNNIGVLTNS	176
<i>Li.monocytogenes</i>	EARRLLQRINLVNISFSENPLSPLHWMADQN--ESIVVECVKD--GLHIYDNPVGVLTTN	171
<i>En.faecalis</i>	EAKKLLKNLNLVNINFSDELPLSPLHLLADKE--QSIVVESTKE--GLRVFDNPVGVLTTN	171
<i>La.plantarum</i>	EVKKNIQKLNLVNINFSDELPLSPLHLLVADKQ--ESIVIESVKE--GLKIYDNPVGVLTTN	171
<i>La.gasseri</i>	QVKESLTNANISNVSAKNTAPSELHVLVGDKTGKSIVVESDEK--GLHVYNNPVNALTNA	172
<i>La.johnsonii</i> (α)	EVKDALKDVSILVNINFSDELPLSPLHWMADKTGESIVVESTLS--GLHVYDNPVHVLTTN	173
<i>Bi.longum</i>	EVEEALRNVTLVSQIVP--CQQLSLLHWFIDGDK--RSIVVEQMAD--GMHVHDDVDVLTNQ	174
<i>Bi.adolescentis</i>	EVEEALKNVTLVSQIVP--CQQLSLLHWFIDGDK--RSIVVEQMAD--GMHVHDDVDVLTNQ	174
<i>Bi.bifidum</i>	EVEEALKNVTLVSQIVP--CQQLSLLHWFIDGDK--RSIVVEQMAD--GMHVHDDVDVLTNQ	174
<i>La.johnsonii</i> (β)	EVKEALKNVNLVNEAINTSFAPLHWIISDSD--EAIIVEVSKQYGMKVFDRLGVLTNS	175
<i>La.acidophilus</i>	DVKEALKNVNLVNEAINTSFAPLHWIISDSD--EAIIVEVSKQYGMKVFDRLGVLTNS	175
<i>Ba.sphaericus</i>	DVIEKLTSYTLTNEANIILGFAPPLHYTFTDASGESIVIEPKT--GITIHRKTIGVMTNS	179
<i>Cl.perfringens</i>	PTFDHVAANLNQYVGLRYNQVPEFKLDQSLTALGQGTGLVGLPGDFTPASR--FIRVAFRL	236
<i>Li.monocytogenes</i>	PTFDYQLFNLNNYRVLSSETPENNFSEIDLDAYS--SRGMGGIGLPGDLSSMSR--FVKATFTK	231
<i>En.faecalis</i>	PTFDYQLFNLNNYRVLSSTRTPKNNFSDQIELDIYS--SRGMGGIGLPGDLSSMSR--FVKATFTK	231
<i>La.plantarum</i>	PNFDYQLFNLNNYRALSNTSPQNSFSEKVDLDSYS--SRGMGGIGLPGDLSSMSR--FVRAAFTK	231
<i>La.gasseri</i>	PLFPFQLTNLANYASVVPGEPPDNNFLPGVNLKLYS--SLGTHHLP--GGMDSES--FVKVCFAL	232
<i>La.johnsonii</i> (α)	PEFPFQLRNLNLANYSNIAPAQPKNTLPGVDNLNYS--SLGTHHLP--GGMDSAS--FVKIAFVR	233
<i>Bi.longum</i>	PTFDHFMENLRNYMCVSNEMAEP--TSWGKASLTAWGAGVGMHGIPGDVSSPSR--FVRVAYTN	234
<i>Bi.adolescentis</i>	PTFGFHMENLRNYMCVSNEMAEPATWGKASLSAWGAGVSMHGIPGDVSSPSR--FVRVAYAN	234
<i>Bi.bifidum</i>	PTFDHFMENLRNYMCVSNEMAEP--TWGKAELS--AWGAGVSMHGIPGDVSSPSR--FVRVAYTN	234
<i>La.johnsonii</i> (β)	PDFNWHLTNLGNITGLNPHDATAQSWNGQKVAPW--GVGTGSLGLPGDISPADR--FVKAAYLN	235
<i>La.acidophilus</i>	PDFNWHLTNLGNITGLNPHDATAQSWNGQKVAPW--GVGTGSLGLPGDISPADR--FVKAAYLN	235
<i>Ba.sphaericus</i>	PGYEWHTNLRLAYIGVTPNPPQDIMMGDLDTLTPFGQAGGLGLPGDFTPASR--FLRVAYWK	239

		
<i>Cl.perfringens</i>	DAMIKNDKDSIDLIEFFHILNNVAMVRGSTRTVEEKSDLTQYTSCMCLEKGIYYNTYEN	296
<i>Li.monocytogenes</i>	LNSVSGDSESESIGQFFHILGSVEQQKGLCDVGGEKYEHTIYSSCCNIDKGIYYRTYGN	291
<i>En.faecalis</i>	LNSVSRSEYESISQFFHILSSVEQQKGLCDVGGEKYEHTIYSSCCNLEKGIYYRTYDN	291
<i>La.plantarum</i>	LNSLSMQTESGSVSQFFHILGSVEQQKGLCEVTDGKYEHTIYSSCCMDKGVIYYRTYDN	291
<i>La.gasseri</i>	NHAPKDSDEVENVTNFFHILESVEQAKGMDQIGPNSFEYTMYTSCMNLEKGILYFNCDYD	292
<i>La.johnsonii</i> (α)	AHSPQGNNELSSVTNYFHILHSVEQPKGTDEVGPNSYEHTIYSDGTNLETGTFFYTNYEN	293
<i>Bi.longum</i>	AHYPQQNDEAANVSRLFHTLGSVMVDGMAKMGDGGQFERTLFTSGYSSKTNTYYMNTYDD	294
<i>Bi.adolescentis</i>	THYPQQEGEAAANVSRLFHTLGSVMVDGMAKMSNGQFERTLFTSGYSSKTNTYYMNTYDD	294
<i>Bi.bifidum</i>	THYPQQNNEAANVSRLFHTLVSVQMVDMGMSKMGNGQFERTLFTSGYSGKTNTYYMNTYED	294
<i>La.johnsonii</i> (β)	ANYPTVKGEKANVAKFFNILKSVAMIKGSVVNDQGSDEYTVYTACYSSGSKTYCNCFEDD	295
<i>La.acidophilus</i>	VNYPTVKGKKANVAKFFNILKSVAMIKGSVVNKQGSNEYTVYTACYSAATKTYCNCFEND	295
<i>Ba.sphaericus</i>	KYTEKAKNETEGVTNLFHILSSVNIPKGVVLTNEGKTDYTIYTSAMCAQSKNYYFKLYDN	299
		
<i>Cl.perfringens</i>	NQINAI DMNKENLDGNEIKYKYNTLSINHVN-----	329
<i>Li.monocytogenes</i>	SQITGVDMHQEDLESKELAIYPLVNEQQNLNIVNK----	325
<i>En.faecalis</i>	SQITAVDMNKENLEKDSLIVYPMVETQQINYAN-----	324
<i>La.plantarum</i>	SQINSVSLNHEHLDTELSYPLRSEAQYYAVN-----	324
<i>La.gasseri</i>	SRISAVDMNKEDLDSSDLVVYDLFKKQDISFIN-----	325
<i>La.johnsonii</i> (α)	NQINAI ELNKENLNGDELTDYKLIKQTINYQN-----	326
<i>Bi.longum</i>	PAIRSYAMADYDMDSELISVAR-----	317
<i>Bi.adolescentis</i>	PAIRSYAMADFDMDSELITAA-----	316
<i>Bi.bifidum</i>	PAIRSFAMSDFDMDSELITAD-----	316
<i>La.johnsonii</i> (β)	FELKTYKLD DHTMNSTSLV TY-----	316
<i>La.acidophilus</i>	FELKTYKLDDET MNADKLITY-----	316
<i>Ba.sphaericus</i>	SRISAVSLMAENLNSQDLITFEWDRKQDIKQLNQVNVMS	338

of *Bifidobacterium longum* and *Bi. bifidum* (9, 17) as well as sulfhydryl inhibition of several BSH have shown the importance of this residue in catalysis (10, 12, 13). Alignment of amino acid sequences from BSH shows the Cys2 residue is conserved in all BSH characterized to date (Figure 4). The broad substrate specificities reported (Table 1) are potentially a function of a lack of conservation observed in residues making up the substrate binding pocket of the conjugated bile acid hydrolase gene product (CBAH-1) of *C. perfringens* and the corresponding residues predicted in amino acid multiple sequence alignment with other BSH (Figure 4). The sterol moiety is bound primarily through hydrophobic interactions in the CBAH-1 of *C. perfringens* (residues highlighted in grey in Figure 4) as well as hydrogen bonds to the carboxylate group. While the crystal structure of CBAH-1 did not reveal specific recognition of the taurine/glycine moiety, kinetic data from several BSH suggest the conjugates are important in substrate specificity (Table 1). Therefore, additional crystallization and site-directed mutagenesis (preferably with mutagenesis of Cys2) of BSH from

different species will be helpful in explaining kinetic observations of substrate specificity.

### **Distribution, genetic organization, and regulation of BSH**

Genes encoding BSH have been cloned from *C. perfringens* (14), *Lactobacillus plantarum* (15), *La. johnsonii* (16, 17), *Bi. longum* (10), *Bi. bifidum* (18), *Bi. adolescentis* (19), and *Listeria monocytogenes* (20, 21). Homologues and putative *bsh* genes have also been identified recently through microbial genome analysis. The organization and regulation of genes encoding BSH differs between species and genera. Monocistronic BSH genes have been reported in *La. plantarum* (15), *La. johnsonii* (17), *Li. monocytogenes* (21), and *Bi. bifidum* (18). A gene encoding BSH (CBAH-1) cloned from *C. perfringens* (14) differed significantly in size and amino acid sequence from a BSH purified from a different strain of *C. perfringens* (12). The inactivation of the gene encoding CBAH-1 resulted in only partial reduction in BSH activity (BSH activity 86% of WT) suggesting multiple BSH genes in *C. perfringens*. Furthermore, the crystal structure showed that the enzyme encoded by the CBAH-1 gene forms an active



homotetramer (11). These observations, coupled with detection of both intracellular and extracellular BSH provides further evidence for multiple isoforms though organization and regulation of the *bsh* gene(s) from *C. perfringens* are not presently known (22). Polycistronic operons encoding three genes involved in bile salt deconjugation (*cbst1*, *cbst2*, *cbsH* $\beta$ ) have been characterized in *La. johnsonii* and *La. acidophilus* (17). Genes *cbst1* and *cbst2* appear to be gene duplications each of which encode TCA/CA antiport proteins of the major facilitator superfamily while *cbsH* $\beta$  encodes the BSH  $\beta$ -isoform (23). In addition, an uncharacterized extracellular factor has been detected in *La. johnsonii* 100-100 which stimulates BSH activity and uptake of conjugated bile salts during stationary growth phase (17, 24). BSH expression is also growth phase dependent. Stationary phase expression has been reported in *Bacteroides fragilis* (25) while exponential phase expression was reported for *Bi. longum* (10).

#### **Benefits of BSH to the bacterium**

BSH appear to enhance the bacterial colonization of the lower GI tract of higher mammals. The physiological advantages of BSH are not fully understood and may vary between bacterial species and genera. It has been hypothesized that deconjugation may be a mechanism of detoxification of bile salts. De Smet et al (1995) observed significantly higher rates of deconjugation of GDCA over TDCA in *La. plantarum* (26). Mutants lacking functional BSH (*bsh*<sup>-</sup>) exhibited pH and concentration dependent toxicity of GDCA as compared to wild type cells; whereas, this effect was not demonstrated with TDCA. De Smet et al. (1995) hypothesized that the difference in dissociation constants between GDCA and TDCA resulted in the collapse of cellular proton motive force by intracellular deprotonation with the glycine-conjugate (26). The presence of a functional BSH results in the intracellular accumulation of free bile acids which become protonated in a stoichiometric fashion decreasing energy dependent H<sup>+</sup> ATPase driven proton efflux. BSH from human intestinal lactobacilli generally have higher affinity for glycine conjugates (26-29). This observation may lend weight to the hypothesis of De Smet et al. (1995) or the

higher affinity of BSH for glycine conjugates may have evolved because glycine conjugates are generally higher in proportion (3:1) than taurine conjugates in human bile (6). Tannock et al. (1989) argued against the hypothesis of deconjugation as a means of detoxification in lactobacilli, pointing to the fact that free bile acids are more cytotoxic than their conjugates (30). However, when the free bile acids become 7-dehydroxylated by other intestinal bacteria *in vivo*, the resultant secondary bile acids tend to precipitate (extent depending on luminal pH) and bind to insoluble fiber or may be absorbed through the colonic membrane, and may exist in low concentrations in the bacterium's microenvironment. Therefore, additional studies comparing various characteristics of *bsh* knockouts with their isogenic parent strain will be needed to determine the function of deconjugation in *Lactobacillus* spp.

Strategies to resist bile salt toxicity have been observed in pathogens that colonize the intestinal tract (31-33). Recently, a BSH from *Li. monocytogenes* was shown to be a novel virulence factor (21). Comparative genome analysis revealed the absence of a *bsh* gene in the closely

related non-virulent *Li. innocua* (20). The *bsh* gene is positively regulated by PrfA, which is a transcriptional activator of numerous virulence genes in *Li*.

*monocytogenes*. Deletion of the *bsh* gene results in decreased resistance to bile salts, and significantly reduced infectivity *in vivo*. These results demonstrate the importance of BSH activity for survival *in vivo* and infection in the intestinal and hepatic phases of listeriosis. The mechanism by which BSH activity in *L. monocytogenes* enhances survival and virulence is currently unknown.

Deconjugation may provide a means of obtaining cellular carbon, nitrogen, and sulfur for some bacterial species. This has been demonstrated in bacteroides (34) and is suggested in *Bi. longum* (10). In fact, the *bsh* gene from *Bi. longum* is co-transcribed with the gene encoding glutamine synthetase adenylyltransferase (*glnE*); a component of the nitrogen regulation cascade (10). In this regard, hydrolysis of the conjugated bile acid may provide amino nitrogen; providing a possible explanation for coordinate regulation of these seemingly physiologically unrelated genes (10). Taurine utilization

is also widespread and can serve as an energy source under both aerobic and anaerobic conditions (35). Glycine can be utilized as an energy source by certain clostridia by the Stickland reaction (36). The Stickland reaction is a form of amino acid fermentation where one amino acid donates electrons which are accepted by another amino acid distinct from the electron donor. Another hypothesis suggests BSH are detergent shock proteins enabling survival during stress (37). De Smet et al. (1995) found no evidence for this in lactobacilli after growth with various detergents (26).

The widespread distribution of BSH across gram negative and gram positive intestinal bacteria coupled with a wide range of substrate specificities, genetic regulation and the occurrence of multiple isoforms in certain strains have created conflicting reports regarding the physiological benefit to the bacterium in hydrolyzing bile acid conjugates. Determining the mechanism(s) by which BSH aid bacteria in colonization of the mammalian intestine will be of great interest, especially in regards to bacterial pathogenesis.

**Taurine, hydrogen sulfide production and colon cancer**

The bile acid conjugates glycine and taurine serve as substrate in microbial metabolism. Unlike glycine, taurine contains a sulfonic acid moiety which is reduced and dissimilated to hydrogen sulfide following deconjugation (38, 39). Hydrogen sulfide is highly toxic, and has been shown to increase colonocyte turnover (40). Activation and upregulation of the ERK 1/2 signaling pathway has been suggested as a possible mechanism for sulfide induced colonocyte proliferation (41). Hydrogen sulfide also inhibits butyrate metabolism in colonocytes, a key nutrient and regulator of cell turnover in the gut (40). Levitt et al (1999) demonstrated that colonocytes have evolved a highly efficient mechanism to detoxify volatile reduced sulfides through oxidation to thiosulfate (42). Defects in this detoxification system are suggested to play a role in the etiology of ulcerative colitis, a known risk factor for colon cancer (42, 43). Recently, sulfide has been implicated in preventing apoptosis in the adenocarcinoma cell line HCT116 after exposure of cells to  $\beta$ -phenylethyl isothiocyanate (PEITC), a phytochemical

found in cruciferous vegetables, which has been shown to prevent colon carcinogenesis (44, 45).

A diet high in meat has been shown to significantly increase the levels of both taurine conjugation to bile acids (46, 47), and the production of hydrogen sulfide in the colon (48). A relationship exists between generation of hydrogen sulfide in the colon and chronic gastrointestinal illness such as inflammatory bowel disease and colon cancer (5, 49). Populations such as native black Africans with low incidence of colon cancer consume low meat diets (50). Native black Africans also have low ratios of taurine to glycine conjugation (1:9) and low hydrogen sulfide production when compared to populations consuming a "Western diet" (46, 47). In human fecal slurries obtained from individuals consuming a "Western diet", taurine addition generated some of the highest sulfide levels of any organic or inorganic sulfur source added (43). Taurine addition to a co-culture of a species of bacteroides and an unidentified 7 $\alpha$ -dehydroxylating bacterium resulted in significant sulfide production, which stimulated increased rates of DCA production (34).

While the extent to which taurine metabolism contributes to total colonic sulfide production has yet to be established, several key points have been made: 1) the extent of taurine conjugation in the bile acid pool is largely affected by diet; 2) the same dietary factors which increase taurine conjugation are hypothesized to increase colon cancer risk; 3) taurine metabolism by intestinal bacteria results in hydrogen sulfide generation; 4) sulfide generation is linked to the carcinogenesis process through enhanced cell proliferation, inhibition of butyrate metabolism and activation of cell signaling pathways; 5) sulfide generation may enhance DCA formation in the gut through stimulation of the microbial bile acid 7 $\alpha$ -dehydroxylation pathway.

### **Microbial bile acid hydroxysteroid dehydrogenase(s) (HSDH)**

#### **Oxidation and epimerization**

Oxidation and epimerization of the 3-, 7-, and 12-hydroxy groups of bile acids in the gastrointestinal tract are carried out by HSDH expressed by intestinal bacteria (Figure 1). Epimerization of bile acid hydroxy groups is



the reversible change in stereochemistry from  $\alpha$  to  $\beta$  configuration (or vice versa) with the generation of a stable oxo-bile acid intermediate. Epimerization requires the concerted effort of two position-specific, stereochemically distinct HSDH of intraspecies or interspecies origin. For example, the presence of both  $7\alpha$ - and  $7\beta$ -HSDH in *C. absonum* allow epimerization by a single bacterium (51); whereas epimerization also can be achieved in co-cultures of intestinal bacteria, one possessing  $7\alpha$ -HSDH and the other  $7\beta$ -HSDH (52, 53).

The extent of the reversible oxidation and reduction of bile acid hydroxy groups by HSDH depends in part on the redox potential of the environment. Addition of oxygen to the culture medium increases the accumulation of oxo-bile acids (51). Generation of oxo-bile acids may be more favorable under the higher redox potentials found on the mucosal surface (4); while reduction of oxo-bile acids may be more favorable under the low redox potential (-200 to -300mV) in the large intestinal lumen. Thus, while the redox potential of the colon is net reductive, microenvironments at the mucosa may provide oxidizing conditions favorable for certain microbial reactions.

HSDH differ in their reductive and oxidative pH optima, NAD(H) or NADP(H) requirements,  $M_r$ , and gene regulation (Table 2).

### **3 $\alpha$ - and 3 $\beta$ -HSDH**

3 $\alpha$ /3 $\beta$ -HSDH specifically catalyze the reversible, stereospecific oxidation/reduction between 3-oxo-bile acids and 3 $\alpha$ - or 3 $\beta$ -hydroxy bile acids. 3 $\alpha$ -HSDH have been detected in some of the most prevalent intestinal bacteria including *C. perfringens* (54), *Peptostreptococcus productus* (55) and *Eggerthella lenta* (formerly *Eubacterium lentum*) (56, 57) as well as intestinal bacteria present in lower numbers ( $\leq 10^5$ /g wet weight feces) including *C. scindens* (58), and *C. hiranonis* (59), and in non-intestinal bacteria including *Pseudomonas testosteroni* (60, 61). 3 $\beta$ -HSDH activity has been described in species of *Clostridium* and *Rumminococcus* (62-64). It appears that intraspecies 3-epimerization favors the 3 $\alpha$ -position. In fact, growing cultures of *C. perfringens* in the presence of 3-oxo-CDCA formed CDCA (84%) preferentially over iso-CDCA (16%) under anaerobic conditions (65).

**Table 2. Characteristics of hydroxysteroid dehydrogenases (HSDH) from intestinal bacteria**

organism	Stereo Specificity	cofactor	pH optimum	Native MW (kDa)	Induction	Reference
<i>Eggerthella lenta</i>	3 $\alpha$ (CE) <sup>a</sup>	NAD(H)	11.3	205	Repressed by BA	56, 57
	12 $\alpha$ (CE)	NAD(H)	8.0-10.5	125	Repressed by BA	
<i>Clostridium perfringens</i>	3 $\alpha$ (CE)	NAD P(H)	11.3	ND	Non-inducible	54, 65
	12 $\alpha$ (CE)	NAD(H)	10.5	ND	Non-inducible	
<i>Peptostreptococcus productus</i>	3 $\alpha$ (CE)	NAD(H)	8.5	132	Non-inducible	55, 170, 171
	3 $\beta$ (CE)	NAD(H)	9.5	95	Non-inducible	
	7 $\beta$ (PP) <sup>†</sup>	NAD P(H)	9.8	53	induced by 7-oxo-bile acids	
<i>Bacteroides fragilis</i>	7 $\alpha$ (P) <sup>†</sup>	NAD(H)	8.5 (oxid); 6.5 (red)	110 (tetramer)	induced by CA	72,80,165
	7 $\alpha$ (PP)	NAD P(H)	7.0-9.0	127	induced by CA	
<i>Bacteroides thetaiotaomicron</i>	7 $\alpha$ (PP)	NAD(H)	ND	320	ND	75
<i>Escherichia coli</i>	7 $\alpha$ (P)	NAD(H)	8.5	120 (tetramer)	Non-inducible	74
<i>Clostridium</i> sp. 25.11.C	3 $\beta$ (CE)	NAD P(H)	7.5	104	Non-inducible	62
	7 $\alpha$ (CE)	NAD P(H)	8.6	82	Induced by CA, CDCA, 7-oxo-CA	
	7 $\beta$ (CE)	NAD P(H)	8.6	115	Induced by CA, CDCA, 7-oxo-CA	
<i>Clostridium absconum</i>	7 $\alpha$ (PP)	NAD(H)	9.5-11.5	ND	induced by CDCA and DCA/repressed by UDCA	77, 83
	7 $\beta$ (PP)	NAD P(H)	9.0-10.0	200 (tetramer)	induced by CDCA and DCA/repressed by UDCA	
<i>Clostridium sordellii</i>	7 $\alpha$ (P)	NAD P(H)	ND	120 (tetramer)	Induced by CDCA, CA, DCA	71
<i>Clostridium innocuum</i>	3 $\beta$ (PP)	NAD(H)	10-10.2	56	Non-inducible	63
<i>Clostridium scindens</i> VPI 12708	3 $\alpha$ (P)	NAD(H)/NAD P(H)	5.0-9.0	27 (tetramer)	Induced by CA-CoA and CDCA-CoA	58, 81
	7 $\alpha$ (P)	NAD P(H)	ND	124 (tetramer)	Non-inducible	
<i>Clostridium bifermians</i>	7 $\alpha$ (PP)	NAD P(H)/NAD(H)	11	ND	Induced by 7- $\alpha$ -LCA, CDCA, DCA	84
<i>Clostridium limosum</i>	7 $\alpha$ (CE)	NAD(H)/NAD P(H)	10.5	ND	Induced by CA, CDCA, UDCA	76
	7 $\beta$ (CE)	NAD P(H)	10.5	ND	Induced by CA, CDCA, UDCA	
<i>Clostridium leptum</i>	12 $\alpha$ (PP)	NAD P(H)	8.5-9.0	225	Non-inducible	90
<i>Clostridium</i> group P strain C48-50	12 $\alpha$ (PP)	NAD P(H)	7.8	100	Repressed by CDCA, CA	91
<i>Clostridium paraputrificum</i>	12 $\beta$ (CE)	NAD P(H)	7.8(oxid); 10(red)	126	Non-inducible	92

<sup>a</sup> (CE) cell extract; (PP) partially purified; (P) purified; ND, not determined; BA, bile acids

Pyridine nucleotide cofactor requirements differ between 3 $\alpha$ / $\beta$ -HSDH. 3 $\alpha$ -HSDH require NAD(H), with the exception of the enzyme purified from *C. perfringens*, which uses NADP(H) and that purified from *C. scindens* which can use either NAD(H) or NADP(H) (54, 56-58, 61). 3 $\beta$ -HSDH have been shown to preferentially require NADP(H), with the exception of *C. innocuum* which uses NAD(H) (62-64). Dihydroxy bile acids (DCA, CDCA and UDCA) are generally better substrates than trihydroxy bile acids (CA) (62, 65).

3 $\alpha$ / $\beta$ -HSDH characterized to date are constitutively expressed with the exception of *C. scindens* and *C. hiranonis* which are induced by primary bile acids CA and CDCA. In fact, three copies of 3 $\alpha$ -HSDH genes (*baiA1*, *baiA2* and *baiA3*) have been identified from *C. scindens* (66-67) and the *baiA1* gene has been expressed in *E. coli* and characterized (58). The *baiA* gene products are unique among 3 $\alpha$ / $\beta$ -HSDH due to their high specificity toward CA-CoA and CDCA-CoA conjugates and relatively low activity toward free bile acids (58).

### **7 $\alpha$ - and 7 $\beta$ -HSDH**

$7\alpha/\beta$ -HSDH catalyze the reversible, stereospecific oxidation/reduction of the  $7\alpha$ - and  $7\beta$ -hydroxyl groups of bile acids. While  $7\alpha/\beta$ -HSDH are common among intestinal bacteria, the extent of  $7\alpha/\beta$ -dehydrogenation in the intestine, or in mixed fecal suspensions is difficult to interpret due to the competing, and irreversible  $7\alpha/\beta$ -dehydroxylation of bile acids (see section 6) (68).

$7\alpha/\beta$ -HSDH are widespread among the bacteroides and clostridia as well as *Escherichia coli*, and *Ruminococcus* sp. (Table 2) (54, 64, 69-74). In addition, several intestinal clostridia express both  $7\alpha$ - and  $7\beta$ -HSDH and have been shown to epimerize the  $7\alpha/\beta$ -hydroxy group (75, 76-78).  $7\alpha/\beta$ -HSDH have been partially purified from intestinal bacteria including *Ba. fragilis* (71, 79), *Ba. thetaiotaomicron* (74), *C. scindens* (80), *C. sordellii* (70), *E. coli* (73), as well as the soil isolates *Xanthomonas maltophilia* (81), *C. absonum* (82), and *C. bifermentans* (83). Interestingly, several intestinal bacteria with  $7\alpha/\beta$ -HSDH activity also possess BSH including: *Ba. fragilis* (25), *C. sordellii* (70), *C. perfringens* (84), *C. innocuum* (63), and the soil isolate *C. bifermentans* (83).

7 $\alpha$ -HSDH generally utilize NADP(H) as a cofactor, with the exception of *E. coli* (85), and *Ba. thetaiotaomicron* (74). *C. bifermentans*, *C. absonum*, and *Ba. fragilis* 7 $\alpha$ -HSDH can utilize either NAD(H) or NADP(H) as a cofactor (76, 79, 83). 7 $\beta$ -HSDH enzymes characterized to date utilize NADP(H) as a cofactor (53, 62, 75, 76). 7 $\alpha/\beta$ -HSDH enzymes have higher affinity for dihydroxy bile acids (CDCA, 7-oxo-LCA) than trihydroxy (CA, 7-oxo-DCA). *C. limosum* cell extracts utilized both free and conjugated bile acids, while whole cells only oxidized free bile acids (75, 85). This is due to intracellular location of the 7 $\alpha$ -HSDH, and the inability of the organism to take up a conjugated bile salt. The genes encoding 7 $\alpha$ -HSDH have been cloned from *E. coli* (73), *C. scindens* (69), and *C. sordellii* (70). Sequence similarity suggests these enzymes belong to the short-chain polyol dehydrogenase family. Regulation of 7 $\alpha/\beta$ -HSDH expression is generally growth phase dependent and inducible by bile acid substrates (Table 2). *C. scindens* and *E. coli* constitutively express 7 $\alpha$ -HSDH and are non-inducible (69, 85). Unexpectedly, the non-substrate DCA can induce 7 $\alpha$ -HSDH expression in *C. absonum* and *C. sordellii* though the

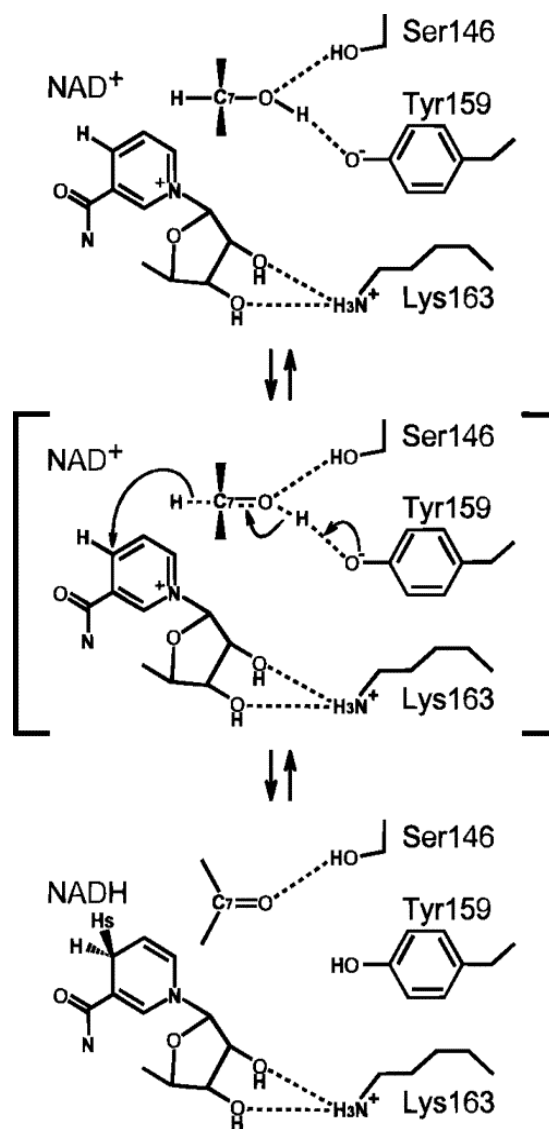
reason for this induction remains unclear (70, 76).

Macdonald et al. (1983) observed the expression of five soluble and two membrane polypeptides upon exposure of *C. absonum* to DCA and CDCA in the culture medium although the functions of these additional polypeptides have not been determined (82).

Crystal structures of the *E. coli* 7 $\alpha$ -HSDH binary and ternary complexes have been solved and a mechanism for 7 $\alpha$ -dehydrogenation proposed (Figure 5) (87). Binding of 7 $\alpha$ -hydroxy bile acids elicits major conformational changes at the substrate binding loop and C-terminal domain. A two step mechanism is proposed in which Tyr159 acts as a catalytic base removing the C-7 hydroxy hydrogen. Hydrogen bonding by Ser146 is hypothesized to stabilize the intermediate. Regeneration of the catalyst occurs through transfer of the acquired hydride from the phenolic group on Tyr159 to Lys163 to position 4 of NAD<sup>+</sup>. Lys163 serves two important roles: anchoring NAD<sup>+</sup> through bifurcated hydrogen bonding, and indirect hydride transfer from Tyr159 to position 4 of NAD<sup>+</sup> (87). Site directed mutagenesis confirmed the role of these amino acids in 7 $\alpha$ -HSDH catalysis (88). Analysis of the 7-oxo-GLCA bile acid

Figure 5. Proposed catalytic mechanism of bile acid 7 $\alpha$ -dehydrogenation based on the crystal structure and site-directed mutagenesis of active site amino acids of the 7 $\alpha$ -dehydrogenation from *E. coli*. See text for a description of catalysis. Reprinted with permission from Tanaka et al. (87). Copyright © 1996 American Chemical Society.





substrate/enzyme complex revealed tight binding of the sterol with loose association for the glycine conjugate. Binding of glycine and taurine conjugates of CDCA were not significantly different from the free bile acid (87).

### **12 $\alpha$ - and 12 $\beta$ -HSDH**

12 $\alpha$ / $\beta$ -HSDH have been detected mainly among members of the genus *Clostridium*. NADP-dependent 12 $\alpha$ -HSDH have been detected in *C. leptum* (89), *Clostridium* group P (89), while NAD-dependent 12 $\alpha$ -HSDH activity was reported in *Eg. lenta* (56), and *C. perfringens* (54). 12 $\beta$ -HSDH have been detected in *C. tertium*, *C. difficile* and *C. paraputrificum* (91, 92). 12 $\alpha$ / $\beta$ -HSDH characterized to date are constitutively expressed, and non-inducible, with the exception of the 12 $\beta$ -HSDH from *C. paraputrificum* which is induced by 12-oxo-bile acid substrates (92). 12 $\alpha$ / $\beta$ -HSDH generally have higher affinity for dihydroxy (DCA) than trihydroxy bile acids (CA, iso-CA) and free vs. conjugated bile acids. The 12 $\alpha$ -HSDH from *C. leptum* is an exception; demonstrating higher affinity for CA conjugates than free CA (89). 12 $\alpha$ -HSDH appear to be repressed by the addition of bile acid substrates (DCA>CDCA>CA) to the growth medium

at 1 mM concentrations. It has been suggested that these enzyme activities should be repressed in bacteria colonizing the large intestine (56, 92); although 12-oxo-bile acids have been detected at low levels in the feces of healthy individuals (Figure 3) (93-94).

### **Benefits of bile acid hydroxysteroid oxidoreductases to the bacterium**

The oxidation of bile acid hydroxyl groups generates reducing equivalents for cellular biosynthetic reactions and possibly electron transport phosphorylation. Bile acid dehydrogenation is hypothesized to generate energy in *Ba. thetaiotaomicron* (74). Sherod and Hylemon (1977) suggested that reduced pyridine nucleotides generated from 7 $\alpha$ -hydroxy oxidation serve to generate ATP via a cytochrome-linked electron transport chain in the presence of electron acceptors (i.e., fumarate) (74).

Bile acids are potent antimicrobial agents provided the proper concentration and proportion of hydrophobic bile acids (CDCA, LCA, DCA) are present (95). Alteration of hydroxy group stereochemistry has marked influence on the physiochemical properties of bile acids (96, 97). The

epimerization of the 7 $\alpha$ -hydroxy group of CDCA lowers the hydrophobicity and toxicity of the bile acid (97). Macdonald, White and Hylemon (1983) observed that *C. absonum* grew on plates containing 1 mM UDCA, though was unable to grow on plates containing 1 mM CDCA (82). Furthermore, when cultured in the presence of 7-oxo-bile acids only low concentrations of CA and CDCA were formed while the majority of 7-oxo-bile acids were reduced to 7-epi-CA and UDCA, respectively, by log phase *C. absonum* (51). UDCA has also been shown to act as a repressor of 7 $\alpha$ / $\beta$ -HSDH production in *C. absonum* suggesting UDCA is an end-product (76). The enzyme also displays markedly higher affinity for CDCA than CA, the former being more toxic. In summary, dehydrogenation may serve functions related to energy generation as well as attempts to maintain low concentrations of more hydrophobic bile acids in the bacterium's microenvironment.

### **Interplay between HSDH enzymes in human liver and intestinal bacteria**

The co-evolution between host and gut flora is evident when observing the interplay between liver and

bacterial biotransforming reactions. The liver synthesizes bile acids in which the hydroxy groups are in the  $\alpha$  orientation. In the  $\alpha$ -hydroxy orientation, one face of the molecule is hydrophobic while the other side is hydrophilic. This translates to efficient solubilization of lipid molecules through formation of mixed micelles capable of efficient emulsification while remaining soluble in aqueous environments. Generation of  $\beta$ -hydroxy-bile acids by microbial enzymes alters the efficiency of micelle formation due to hydrophilic groups on both faces of the sterol molecule. The differences observed between the composition of bile acids in serum and bile are a result of the continual interplay between liver and bacterial enzymes. Exposure of bile salts to intestinal bacteria results in roughly 50% of bile acids requiring reconjugation and low levels of bile acids returned to the liver in the  $3\beta$ -hydroxy orientation (98). Without a means of epimerizing bile acid hydroxy groups in the human liver,  $\beta$ -hydroxy bile acids would accumulate in the bile acid pool. Interestingly, the liver seems to "allow" the accumulation of UDCA ( $7\beta$ -hydroxy) in the biliary pool as in the case of therapeutic administration of CDCA (99).

The protective effects of UDCA observed in clinical studies (100), as well as cell culture studies (97, 101) may provide an evolutionary explanation for this phenomenon.

## **The biochemistry and molecular biology of bile acid 7 $\alpha$ / $\beta$ -Dehydroxylation**

### **Introduction**

Secondary bile acids (DCA, LCA) predominate in human feces (Figure 3). Therefore, 7 $\alpha$ -dehydroxylation is the most quantitatively important bacterial bile salt biotransformation in the human colon. The rapid rate of conversion of primary to secondary bile acids is surprising given that current estimates suggest this metabolic pathway is found in roughly 0.0001% of the total colonic flora (102-104). Human intestinal bacteria capable of bile acid 7 $\alpha$ -dehydroxylation have been isolated (104, 105) and 16s rDNA phylogenetic analysis has led to their classification to the genus *Clostridium* (106-108).

Unlike bile acid oxidation and epimerization, 7 $\alpha$ / $\beta$ -dehydroxylation appears restricted to free bile acids. Removal of glycine/taurine bile acid conjugates via BSH enzymes is thus a prerequisite for 7 $\alpha$ / $\beta$ -dehydroxylation by

intestinal bacteria (109-112). Some intestinal bacteria are capable of both  $7\alpha/\beta$ -dehydroxylating activities (113), while  $7\beta$ -dehydroxylation activity is absent in other intestinal  $7\alpha$ -dehydroxylating bacteria (113). Epimerization of UDCA ( $7\beta$ -hydroxy) to CDCA ( $7\alpha$ -hydroxy) via  $7\beta$ -HSDH produced by members of the gut flora result in subsequent  $7\alpha$ -dehydroxylation. In this regard, it would appear the presence of  $7\beta$ -dehydroxylation activity is more of a luxury than a necessity.

#### **Elucidating the bile acid $7\alpha/\beta$ -dehydroxylation pathway**

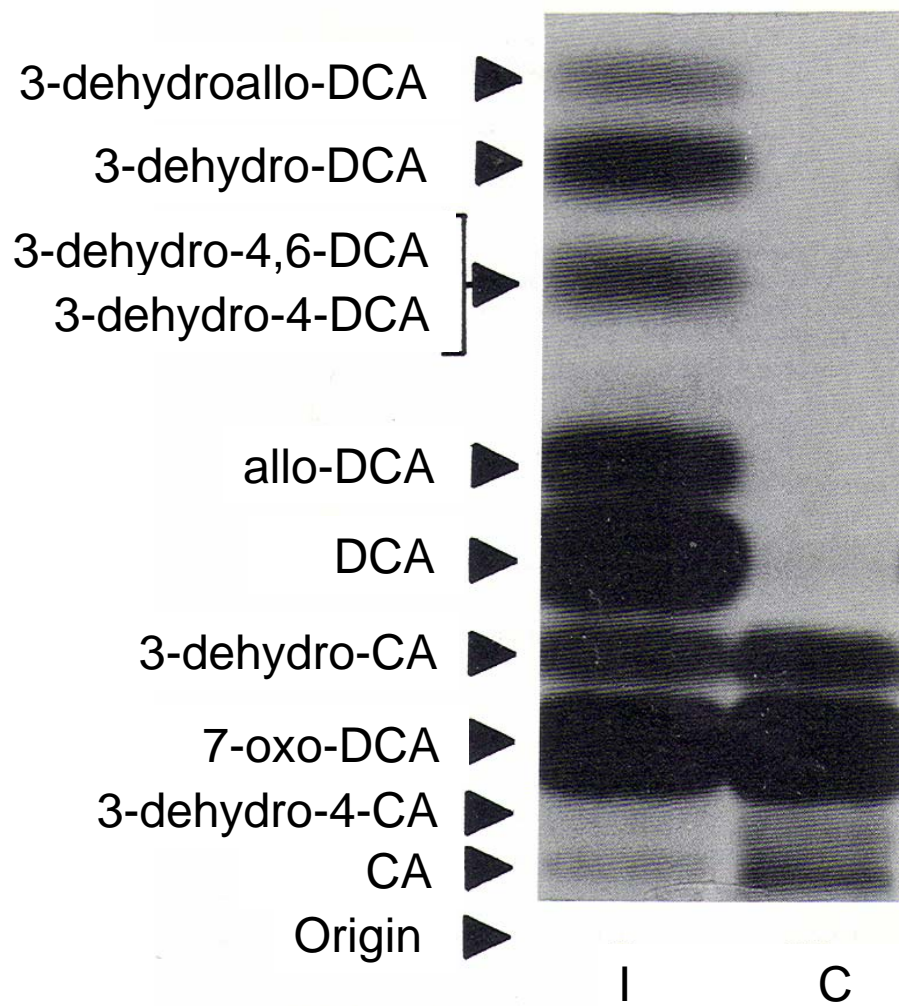
Samuelsson (1960) administered [ $6\alpha$ - $^3\text{H}$ ,  $6\beta$ - $^3\text{H}$ ,  $8\beta$ - $^3\text{H}$ ]-[24  $^{14}\text{C}$ ] CA to bile duct cannulated rabbits and rats (114). Analysis of the products recovered following exposure to intestinal bacteria revealed a differential loss of the [ $6\beta$ - $^3\text{H}$ ] during  $7\alpha$ -dehydroxylation of CA. Previous work showed complete retention of the  $7\beta$ - $^3\text{H}$  in [ $7\beta$ - $^3\text{H}$ ]-[24- $^{14}\text{C}$ ]-CA during  $7\alpha$ -dehydroxylation in the rat intestine (115, 116). These data led Samuelsson (1960) to propose a mechanism for CA  $7\alpha$ -dehydroxylation involving two steps: diaxial trans-elimination of the  $7\alpha$ -hydroxy group and  $6\beta$ -hydrogen atom followed by reduction through trans-hydrogenation of the  $6\beta$  and  $7\alpha$  positions of the cholen-6-

oic acid intermediate forming DCA (114). Björkhem et al. (1989) showed the formation of a 3-dehydro-4-cholenoic acid intermediate following the differential loss of the  $5\beta\text{H}$  *in vitro* and *in vivo* using a  $[3\beta\text{-}^3\text{H}]\text{-}[24\text{-}^{14}\text{C}]$  and  $[5\beta\text{-}^3\text{H}]\text{-}[24\text{-}^{14}\text{C}]$ -labeled CA (117). Hylemon et al. (1991) subsequently observed accumulation of multiple bile acid intermediates in cell extracts of *Clostridium scindens* induced by CA (Figure 6) (118). These radiolabeled CA intermediates were identified by mass spectrometry, then chemically synthesized and added to cell extracts of CA induced *C. scindens*. Each  $[24\text{-}^{14}\text{C}]$  CA intermediate was converted into  $[24\text{-}^{14}\text{C}]$  DCA in cell extracts prepared from CA induced cultures of *C. scindens*. These observations suggested that the  $7\alpha$ -dehydroxylation mechanism was more complex than the two-step mechanism proposed by Samuelsson et al (1960). Furthermore, these data demonstrated that bile acid  $7\alpha$ -dehydroxylation was a multi-step pathway in *C. scindens* and suggested the presence of multiple bile acid inducible genes (*bai*).

The induction of  $7\alpha$ -dehydroxylation activity in *C. scindens* by unconjugated  $\text{C}_{24}$  primary bile acids resulted in the appearance of several new polypeptides as observed on



Figure 6. Accumulation of [24- C]CA intermediates during 7 $\alpha$ -dehydroxylation in cell extracts of *C. scindens* VPI 12708. Cell extracts were prepared from either CA-induced (I) or control (C) cells. This figure was modified from Hylemon et al. (118). Copyright© 1992 American Society for Biochemistry and Molecular Biology, Inc.

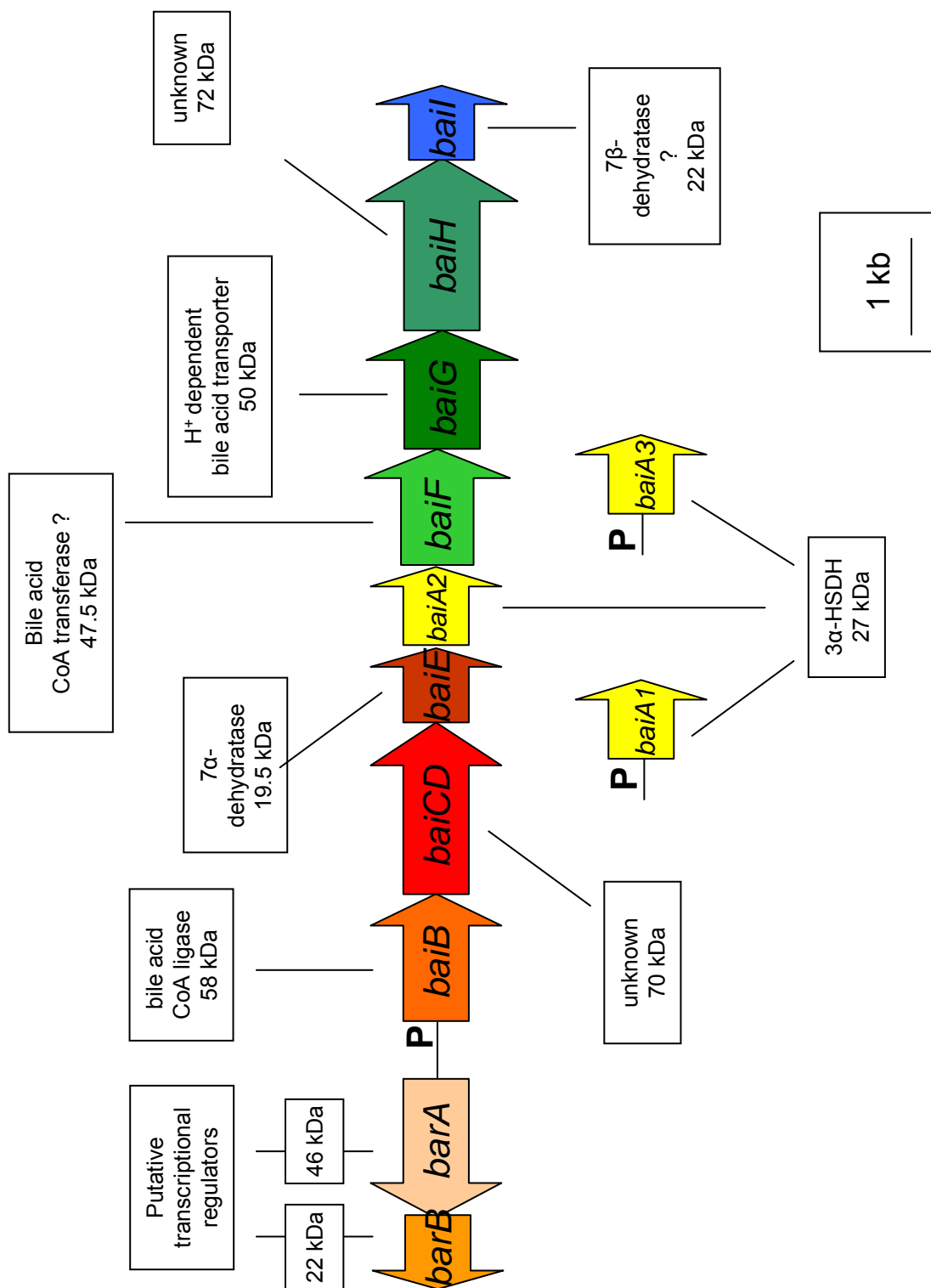


one and two-dimensional SDS-PAGE (119, 120). Purification and N-terminal sequencing of these *bai* polypeptides facilitated cloning of *bai* genes through design of degenerate probes (121-123). Northern blot analysis indicated the presence of a large CA inducible ( $\geq 10$  kb) mRNA transcript and a smaller transcript ( $< 1.5$  kb) in *C. scindens* (121, 124). These studies led to the discovery of a *bai* regulon encoding at least 10 open reading frames (Figure 7). Individual *bai* genes have been subcloned into *E. coli* and the functions of many determined (58, 67, 123, 125-130, Hylemon P.B., unpublished data). The proposed bile acid  $7\alpha/7\beta$ -dehydroxylation pathway in *C. scindens* is shown in Figure 8. A *bai* operon has also been characterized from *C. hiranonis* (59), though the following discussion of the  $7\alpha/7\beta$ -dehydroxylation pathway will center on *C. scindens* from which the functions of the gene products have been determined.

#### ***bai* genes: a regulon for $7\alpha/7\beta$ -dehydroxylation**

The transport of unconjugated primary bile acids into *C. scindens* is facilitated by the *baiG* gene product which belongs to a major pump/facilitator superfamily of protein

Figure 7. Gene organization of the bile acid-inducible (bai) 7 $\alpha$ / $\beta$ -dehydroxylation operons characterized in *C. scindens* VPI 12708. P indicates the promoter region.



transporters (126). The *baiG* gene has been cloned into *E. coli* and shown to encode a 50 kDa H<sup>+</sup>-dependent bile acid transporter (126). The BaiG facilitates transport of unconjugated CA and CDCA but not the secondary bile acids DCA or LCA (126). Computer-aided modeling suggests the *baiG* polypeptide contains 14 membrane spanning domains (126).

Following transport, ligation to coenzyme A (CoA) is the first step in activating CA and CDCA for 7 $\alpha$ -dehydroxylation as several subsequent enzymatic steps are specific for CoA conjugates. The *baiB* gene was shown to encode a 58 kDa bile acid CoA ligase (125). CoA ligation is ATP, CoA, Mg<sup>2+</sup> dependent and also requires a free carboxyl group on C<sub>24</sub> bile acids (125). CoA ligation may function both to sterically hinder the constitutive 7 $\alpha$ -HSDH committing the bile acid to 7 $\alpha$ -dehydroxylation and trapping the bile acid inside the cell.

The 3 $\alpha$ -hydroxy group is oxidized following CoA ligation. Oxidation of the 3-hydroxy group inhibits 7 $\alpha$ -hydroxy group dehydrogenation; favoring 7-dehydroxylation over constitutively expressed 7 $\alpha$ -HSDH in *C. scindens* (58, 80). The *baiA* gene products encode 27 kDa polypeptides

that have significant similarity with the short-chain alcohol/polyol dehydrogenase gene family (58, 124). Amino acid multiple sequence alignment and comparison between *baiA* gene products and other members of the short-chain alcohol dehydrogenase family revealed a possible NAD(P) binding site and catalytic active site (58). Three *baiA* genes have been cloned from *C. scindens*; the *baiA1* and *baiA3* genes are monocistronic while the *baiA2* gene is part of the polycistronic *bai* operon (66, 67, 124). The *baiA* genes from *C. scindens* were cloned in *E. coli* and shown to encode 3 $\alpha$ -HSDH (58). The enzymes only recognize bile acid CoA conjugates and can utilize either NAD<sup>+</sup> or NADP<sup>+</sup> as electron acceptors (58). Interestingly, the different 3 $\alpha$ -HSDH of *C. scindens* share 92% amino acid sequence identity with one another suggesting gene duplication. The physiological importance of multiple *baiA* genes remains unclear.

The *baiH* gene encodes a 72 kDa polypeptide containing 661 amino acids (121). The BaiH protein exists as a homotrimer with NADH: flavin oxidoreductase activity (121). The *baiH* gene has been subcloned into *E. coli*, purified and shown to contain 1 mol FAD, 2 mol iron, and 1

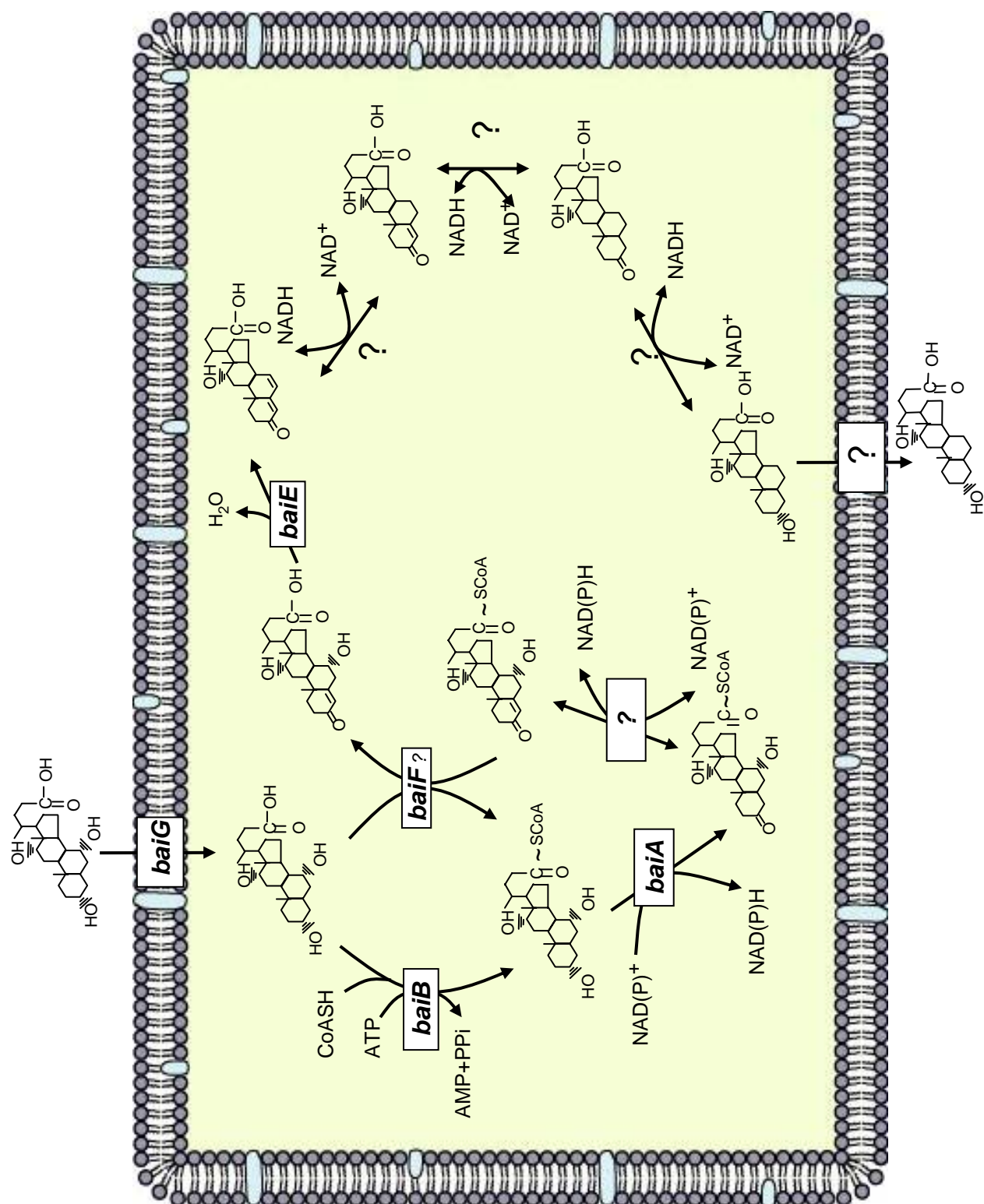
mol copper per mol polypeptide subunit (129). The BaiCD polypeptide shows considerable amino acid sequence identity with the 'Old Yellow Enzyme' family, a putative NADH oxidase from *Li. monocytogenes*, several 2,4-dienoyl CoA reductases, as well as the *baiH* from *C. scindens* (121, 129). The physiological function of the *baiCD* and *baiH* gene products is currently unknown.

The *baiF* gene encodes a 47.5 kDa polypeptide containing 426 amino acids which was shown to have bile acid CoA hydrolase activity (122, 130). However, the *baiF* gene is hypothesized to encode a CoA transferase due to reasons for energy conservation (Fig. 8) and homology to the type III family of CoA transferases (131). The first few cycles of 7 $\alpha$ -dehydroxylation would require ATP hydrolysis (Figure 8); though ATP independent recycling of the thioesterase intermediates via transfer of CoA from an accumulated downstream intermediate, to incoming primary bile acids would significantly conserve energy. However, this hypothesis remains to be tested.

7 $\alpha$ -Dehydration of 3-dehydro-4-cholenoic acid and 3-dehydro-4-chenodeoxycholenoic acid results in the generation of a conjugated double bond in rings A and B



Figure 8.      Proposed bile acid 7 $\alpha$ -dehydroxylation  
pathway in *C. scindens* VPI 12708 (see text).



forming a stable 3-dehydro-4,6-deoxycholdienoic acid and 3-dehydro-4,6-lithocholdienoic acid intermediates, respectively. The  $7\alpha$ -dehydration step results in the largest calculated energy change (-9.4 kcals/mol) of any reaction in this pathway, and the reverse reaction was not detected *in vitro* (128). The 19.5 kDa bile acid  $7\alpha$ -dehydratase is encoded by the *baiE* gene (128). This enzyme showed no activity with 3-dehydro-4-ursodeoxycholenoic acid (128). The *baiE* gene product was modeled based on the crystal structure of protein homologues (secondary structure) ketosteroid isomerase (KSI) and scytalone dehydratase (STD) (Figure 9) (Woodford K et al. unpublished). The putative active site/binding pocket is shown in Figure 10. A catalytic mechanism has been proposed for the  $7\alpha$ -dehydration step based on the conservation in secondary structure and site directed mutagenesis of the key active site amino acids. Tyr30 acts as a general acid withdrawing electron density from the 3-oxo-group. This shift in electron density is hypothesized to make the  $6\beta$ -hydrogen labile for removal by the general base His83 (assisted by Asp35). His83 is thought to donate its hydrogen to the  $7\alpha$ -hydroxy forming the water

Figure 9. Model of the bile acid 7 $\alpha$ -dehydratase (*baiE*) based on structural data from scytalone dehydratase, nuclear transport factor 2, and steroid D5-isomerase. The  $\alpha$ -helices are denoted by white strands,  $\beta$ -sheets are denoted by the yellow backbone, and the blue structure represents a steroid molecule modeled into the binding/active site pocket. Model kindly provided by Dr. Alexey G. Murzin (Cambridge University).

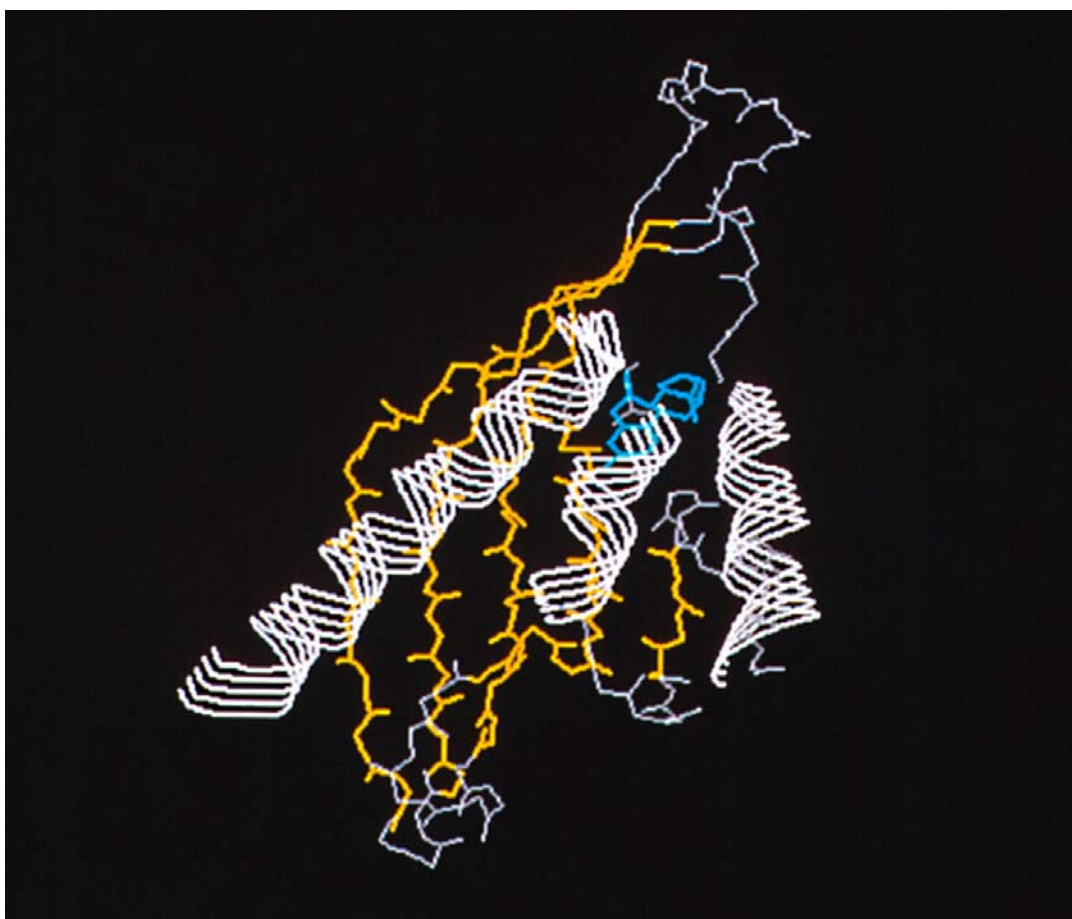
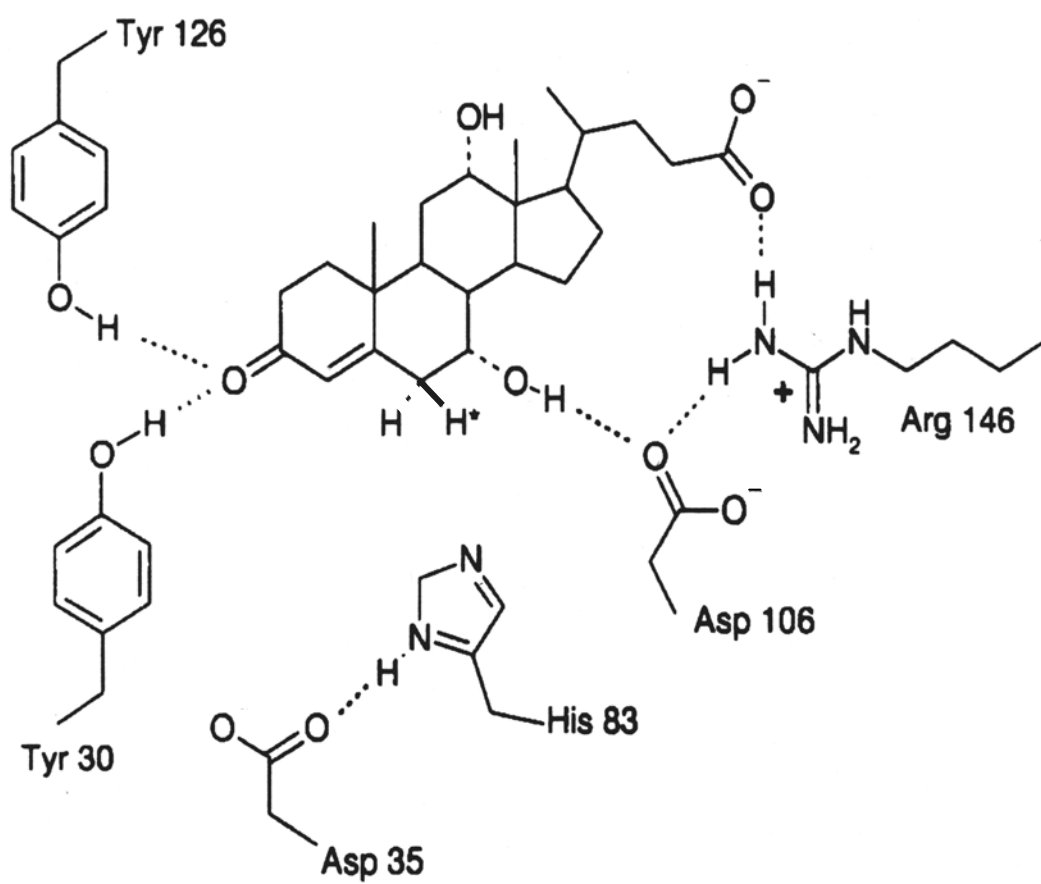


Figure 10. Putative binding/active site pocket of bile acid 7 $\alpha$ -dehydratase. Putative active site amino acid residues are shown in relation to the bile acid substrate. See text for discussion of the proposed catalytic mechanism.



leaving group which is stabilized by Asp106. Site-directed mutagenesis based on the proposed mechanism supports the important role of the putative active site amino acids in enzyme catalysis. It is hypothesized that the *baiI* gene encodes a bile acid 7 $\beta$ -dehydratase due to amino acid sequence homologies between the *baiE* and *baiI* gene products.

Genes involved in the reductive arm of the 7 $\alpha/\beta$ -dehydroxylation pathway have not been isolated. These genes should encode oxidoreductases catalyzing the reduction of 3-dehydro-4,6-deoxycholdienoic acid  $\rightarrow$  3-dehydro-4-deoxycholenoic acid  $\rightarrow$  3-dehydro-deoxycholic acid  $\rightarrow$  DCA as well as a bile acid exporter to remove secondary bile acid end-products from the bacterium (Figure 8). Genes encoding putative transcriptional regulators have been detected upstream of the bile acid inducible promoter region (Table 3) (Mallonee DH and Hylemon PB; unpublished data). Additional studies will be required to determine the mechanism of induction/repression of this pathway and to identify additional *bai* genes.

### **The benefits of 7 $\alpha/\beta$ -dehydroxylation to the bacterium**



**Table 3. Bile acid Inducible (*bai*) genes characterized from *Clostridium scindens* VPI 12708**

<i>bai</i> gene	MW (kDa)	Catalytic Activity/function	Gene Family	Reference
<i>baiA</i>	27	3 $\alpha$ -HSDH	short-chain alcohol/polyol dehydrogenase	58
<i>baiB</i>	58	bile acid CoA ligase	AMP-binding	127
<i>baiCD</i>	70	3-dehydro-4-CDCA/CA oxidoreductase	Pyridine nucleotide-disulphide oxidoreductase; NADH: Flavin Oxidoreductase	<sup>a</sup>
<i>baiH</i>	72	3-dehydro-4-UDCA/7-epiCA oxidoreductase	Pyridine nucleotide-disulphide oxidoreductase; NADH: Flavin Oxidoreductase	131, <sup>a</sup>
<i>baiE</i>	19.5	7 $\alpha$ -dehydratase	COG 4875	130
<i>baiI</i>	22	7 $\beta$ -dehydratase <sup>b</sup>	COG 4876	166
<i>baiF</i>	47.5	bile acid CoA hydrolase bile acid CoA transferase <sup>b</sup>	Type III CoA transferase	132
<i>baiG</i>	50	H <sup>+</sup> dependent bile acid Transporter	major facilitator superfamily	128
<i>barA</i>	46	transcriptional regulation <sup>b</sup>	AraC/XylS	<sup>a</sup>
<i>barB</i>	21	transcriptional regulation <sup>b</sup>	RpoB; Permeases of the major facilitator superfamily	<sup>a</sup>

<sup>a</sup> unpublished data (Hylemon P. B., et al. )

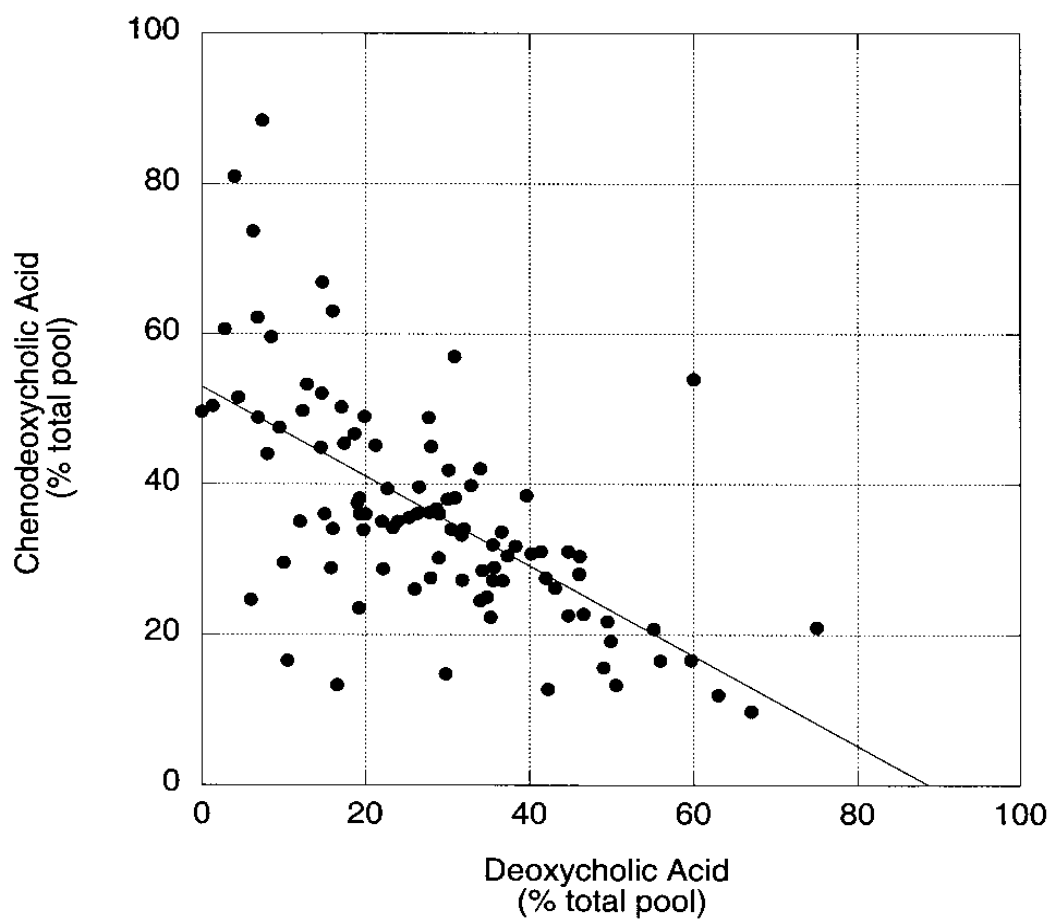
<sup>b</sup> hypothesized function

The ability to utilize bile acids as electron acceptors is an important niche for 7 $\alpha$ -dehydroxylating bacteria in the human colon. The 7 $\alpha$ / $\beta$ -dehydroxylation pathway requires multiple oxidative and reductive steps with a net 2 electron reduction (Figure 8). The hypothesized energy benefits of this pathway assume, however, that the *baiF* gene encodes a CoA-transferase and the toxic end-products, the secondary bile acid, are removed from the microenvironment *in vivo* (precipitation and binding to insoluble fiber). The generation of secondary bile acids may also function to exclude bacteria sensitive to these hydrophobic molecules.

### **Secondary bile acids and disease**

In humans, DCA accumulates in the bile acid pool to high levels in some individuals. An increase in DCA in the bile acid pool is associated with a decrease in CDCA (Figure 11). Unlike rodents, the human liver can not 7 $\alpha$ -hydroxylate DCA forming CA. Hence, under normal physiological conditions, there is no metabolic pathway for removing DCA from the bile acid pool in man. The amount of DCA in the bile acid pool is a function of at

Figure 11. Relationship between the percentage of CDCA and DCA in bile of patients at McGuire VA Hospital (Richmond, VA) (P. B. Hylemon et al., unpublished data).



least three variables: 1) rate of formation and absorption of DCA through the colon (input) (132); 2) colonic transit time (133) and 3) colonic pH (134).

High levels of DCA in blood, bile and feces have been correlated with an increased risk of cholesterol gallstone disease and colon cancer, two major diseases of "Western Society" (5, 135). High levels of CA 7 $\alpha$ -dehydroxylating fecal bacteria have been correlated with increased amounts of DCA in bile of a subset of cholesterol gallstone patients (132). Treatment of these cholesterol gallstone patients (high DCA group) with antibiotics significantly decreased the levels of fecal CA 7 $\alpha$ -dehydroxylating bacteria, DCA in bile and the cholesterol saturation index in bile (132). Early studies by Low-Beer and coworkers (1978) reported that treating control individuals with metronidazole, an antibiotic effective against anaerobic bacteria, significantly decreased the cholesterol saturation index of bile (135). Moreover, excess DCA in bile has been reported to decrease the nucleation time for cholesterol crystallization (136, 137). In total, these results suggest a possible link between intestinal

bacteria, DCA, and the risk of cholesterol gallstone disease in some patients.

DCA and LCA have been linked to colon carcinogenesis in a number of laboratory animal models and human epidemiological studies (For reviews see 5, 138). Most animal studies conclude that DCA is a promoter of the carcinogenesis process (139-142). However, some researchers argue that bile acids may cause DNA damage and act as carcinogens in humans (138). Higher levels of DCA are found in the blood of colon cancer patients as compared to control patients (98, 143). Moreover, DCA is a logical candidate for promoting colon carcinogenesis for the following reasons: 1) it is found in fecal water in high concentrations ( $> 100 \mu\text{M}$ ) (138); 2) it can cross biological membranes via passive diffusion; 3) it can activate mammalian cell signaling pathways that are known to be involved in promoting the carcinogenesis process. In this regard, cell signaling pathways activated by DCA in mammalian epithelial cells include: protein kinase C (144); ERK 1/2 via the epidermal growth factor receptor (101, 145, 146);  $\beta$ -catenin (147) and JNK 1/2 (148). Secondary bile acids have been shown to cause apoptosis in

colonic epithelial cells, and high concentrations of DCA and LCA in stool may promote carcinogenesis by exerting selective pressure for emergence of epithelial cell mutants which are resistant to apoptosis (for example, via loss of p53) (138). LCA has been found to be an excellent activator of the vitamin D receptor (149, 150). Activation of this receptor in intestinal epithelial cells activates genes that metabolize LCA (150). This may be a protective mechanism that evolved to limit LCA toxicity to intestinal epithelial cells.

### **Current trends and Future Issues**

Bile salt metabolism is a widespread and fundamental property of the gastrointestinal microflora encompassing the most commonly isolated species of intestinal bacteria including, but not limited to the genera *Bacteroides*, *Clostridium*, *Lactobacillus*, *Bifidobacterium*, *Eubacterium*, and *Escherichia*. However, our current understanding of the intestinal microbiology of bile salt modifications is limited to cultivated species. Discovering novel genes in the gut microbiome (collective genomes of the GI flora) encoding BSH, HSDH and enzymes involved in the  $7\alpha/\beta$ -dehydroxylation pathway through molecular techniques will

facilitate a much greater understanding of the diversity and complexity of these reactions in the human colon. Techniques such as PCR-denaturing gradient gel electrophoresis (PCR-DGGE) can be utilized to measure the diversity of organisms based on specific phylogenetic markers such as 16s rDNA as well as functional genes, potentially bile salt modifying enzymes (151, 152). Measuring the true diversity of bile salt modifying bacteria is crucial in studying the relationship between the levels and activity of these bacteria and disease risk.

Determining the conditions in which secondary bile acids are formed in significant quantities and retained in the enterohepatic circulation of certain individuals is suggested to be important in the etiology of cholesterol gallstone disease and colon cancer. Since secondary bile acids are formed exclusively through bacterial enzymatic reactions, the study of microbes capable of bile acid  $7\alpha/\beta$ -dehydroxylation is important in understanding these chronic gastrointestinal illnesses. The goal of such research is to find ways to block the source of secondary bile acid production. Long-term use of antibiotics to



prevent 7 $\alpha$ -dehydroxylation of bile acids would be impractical. The design of pharmaceuticals to block the 7 $\alpha$ -dehydroxylation pathway is a possibility. However, this approach requires targeting microbial enzymes and runs the risk of eventual drug resistance as with antibiotics. Alternatively, reducing secondary bile acid production may be achieved through administering specialized CA accumulating probiotic bacteria (153). Studies with bifidobacteria and lactobacilli isolated from human feces have shown their ability to assimilate CA spontaneously *in vitro* (153, 154). The mechanism for CA uptake in lactic acid bacteria appears to be diffusion of a hydrophobic weak acid through the membrane via the transmembrane proton motive force (153, 154). The higher intracellular pH causes the bile acids to become trapped due to ionization.

The use of specific bacteria as "drugs" to treat chronic gastrointestinal illnesses caused in part by other intestinal bacteria certainly has potential, though rigorous studies are needed to demonstrate the effectiveness of such therapies. Overall, clinical trials using probiotics to solve specific health disorders have

been met with mixed results. Discrepancies between studies are due in part to different methodologies, choice of probiotic strains, colony forming units (CFU) administered/day and patient characteristics. One particular variable that would have to be addressed when determining bile acid assimilation by lactic acid bacteria in clinical trials is determining viability *in vivo* after passage through gastric juice and bile. A bacterium must be alive in order to create a membrane potential capable of accumulating bile acids. Targeted delivery of probiotic bacteria to the intestine in microencapsulated form has shown promise in improving viability in the presence of gastric conditions and bile (155, 156). Another issue is whether the introduction of billions of probiotic bacteria to the small bowel will have an effect on bile acid input into the colon. Significant bile salt hydrolysis proximal to the terminal ileum reduces the efficiency of bile salt uptake through high affinity transport, allowing enhanced excretion of bile acids in feces. This principle is behind attempts at lowering serum cholesterol using probiotics with bile salt hydrolytic activity (157-159). Thus, the choice in

probiotic delivery mechanisms is important in addressing both issues of bacterial viability as well as preventing increased bile salt hydrolysis proximal to the terminal ileum. The potential impact of probiotic bacteria in reducing secondary bile acid formation in the human colon through sequestering bile acids would be greatly reduced if these same bacteria cause an increase in bile acid input into the large intestine greater than their capacity for sequestering bile acids.

## STATEMENT OF RESEARCH PROBLEMS AND OBJECTIVES

*Clostridium scindens* is able to  $7\alpha$ - and  $7\beta$ -dehydroxylate primary bile acids through an inducible multi-gene regulon. Several of these gene products have been identified, cloned and characterized. However, a number of remaining questions exist that will provide a more complete understanding of bile acid metabolism in the human colon. First, while it has been observed that *Clostridium scindens* is able to  $7\beta$ -dehydroxylation ursodeoxycholic acid to lithocholic acid, genes involved in this process have yet to be identified. In addition, no physiological role has been given to the homologous *baiCD* and *baiH* gene products located together on the polycistronic oxidative operon of both *Clostridium scindens* and *Clostridium hiranonis*. Second, while the gene organization for the oxidative operons in the "high activity" CA  $7\alpha$ -dehydroxylating activity strains *Clostridium scindens* and *Clostridium hiranonis* have been

well characterized, the gene organization for the “low activity” CA  $7\alpha$ -dehydroxylating activity strain *Clostridium hylemonae* has yet to be determined. Finally, genes in the “reductive arm” of the bile acid  $7\alpha$ -dehydroxylating pathway have yet to be identified. These genes are predicted to encode flavoproteins which reduce the 3-oxo- $\Delta^{4,6}$ -DCA intermediate to 3-oxo- $\Delta^4$ -DCA and finally to 3-oxo-DCA. We predict the existence of a hydroxysteroid dehydrogenase which will reduce the 3-oxo-DCA intermediate to the secondary bile acid DCA. Finally, we predict an ABC permease which will act to remove DCA from the bacterial cell.

### **The objectives of these studies include:**

#### **Objective 1: Characterization of the *baiCD* and *baiH* gene products**

- 1) Perform bioinformatic and secondary structural predictions of *baiCD* and *baiH* gene products with other proteins in the databases to predict the functions of these gene products

- 2) Subclone the *baiCD* and *baiH* genes from *Clostridium scindens* into *E. coli* for overexpression of recombinant protein.
- 3) Synthesize appropriate substrates for enzyme assay and determine reaction product identity by thin layer chromatography and mass spectrometry

**Objective 2: Isolate and characterize *bai* genes from *Clostridium hylemonae* TN271**

- 1) Isolate a fragment of the *bai* oxidative operon from *Clostridium hylemonae* via polymerase chain reaction, clone and sequence this product.
- 2) Perform bi-directional genome-walking by PCR to obtain the remainder of the genes and cis-elements in this operon.
- 3) Characterize operon "in silico"
- 4) Determine mRNA initiation site(s).

**Objective 3: Identify novel *bai* genes in intestinal clostridia**

- 1) Locate reductive gene(s) in *Clostridium scindens* and *Clostridium hylemonae*

- 2) Clone, overexpress and purify recombinant protein of one reductive gene in *E. coli*
- 3) Synthesize bile acid substrate for characterization of reductive gene product
- 4) Determine mRNA initiation site(s).

## MATERIALS AND METHODS

### Bacterial strains, culture media, and conditions

*Clostridium scindens* VPI 12708 (formerly *Eubacterium* sp. strain VPI 12708) was routinely cultivated under anaerobic conditions in Brain Heart Infusion (BHI) broth containing sodium cholate (100  $\mu$ M) and maintained as previously described (111). *Clostridium hylemonae* TN271 was a generous gift of Dr. Fusae Takimine, University Riken, Japan. *C. hylemonae* was cultivated under anaerobic conditions in BHI containing 0.02 % fructose. *Clostridium absonum* (ATCC 27555) was purchased from ATCC. *C. absonum* was initially grown in cooked meat broth for 24 hrs at 37°C and subsequently stored at 4°C and used for no more than 3 weeks as starter culture source. The cells were cultured in BHI supplemented with 0.2 % fructose (w/v). *E. coli* BL21-CodonPlus (DE3)-RIL and ArcticExpression Competent *E. coli* were purchased from Stratagene (La Jolla, CA) and used for expression of recombinant proteins. Recombinant *E. coli* cells were grown with shaking in modified TYGPN (tryptone 20 g, yeast extract 10

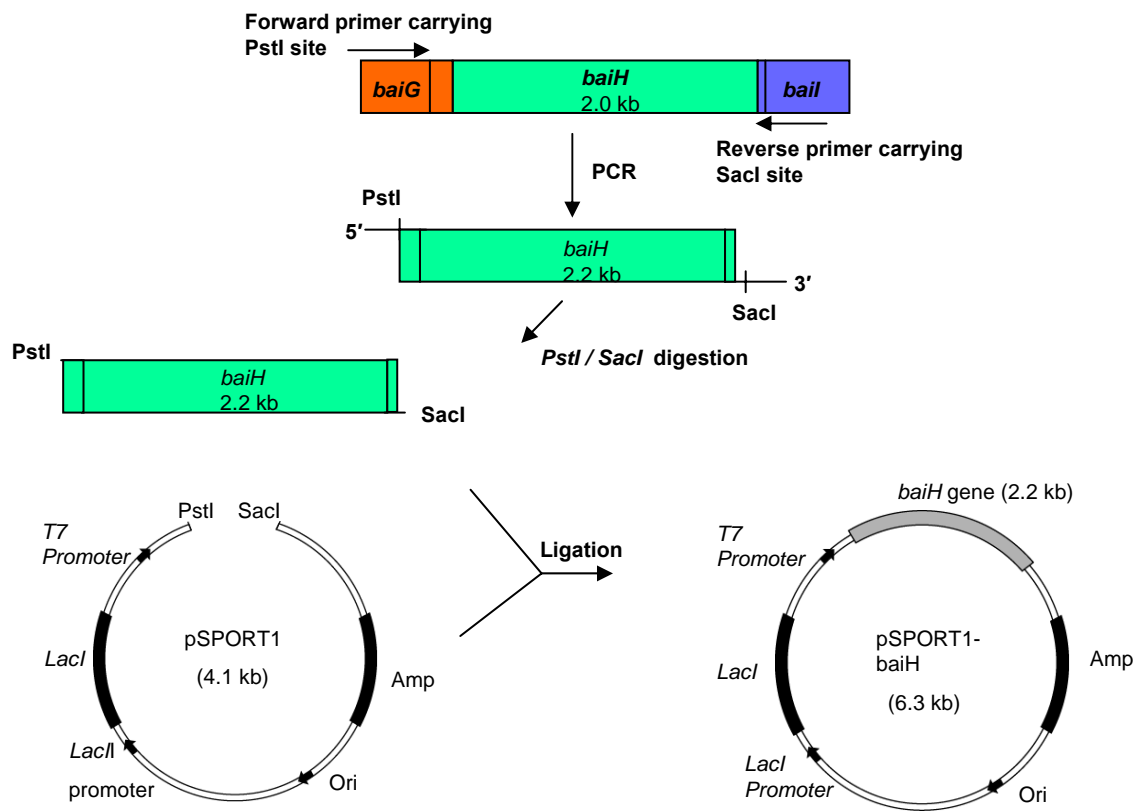


g, 0.8 % glycerol, 5 g Na<sub>2</sub>HPO<sub>4</sub>, 10 g KNO<sub>3</sub>, and 3 µM hematin per liter; pH 7.0) containing 100 µg/ml ampicillin.

### General PCR and recombinant DNA methods

Total genomic DNA from *C. scindens* VPI 12708 was purified as described previously (169) and used as a template for PCR amplification of the *baiH* gene using primers baiHF-PstI (5'-TAGCA**CTGCAG**GCCTATGTT-3') containing *PstI* restriction site and baiHR-SacI (5'-ATCAG**AGCTC**ATTGCCTTCAC-3') containing *SacI* restriction site (**Figure 12**). The expression vector for the *baiCD* gene from *Clostridium scindens* VPI 12708 was constructed in pSPORT1 (Invitrogen, Carlsbad, CA). The *baiCD* gene was cloned using primers baiCDF-SalI (5'-CCGTCGACGGAGGAGTAACAAAATGAGTTACG-3') containing *SalI* restriction site and baiCDR-BglII (5'-CGTAAGCTTCTAGATTGCCATTCCTGCGTC-3') containing *BglII* restriction site. The *baiJ* gene was PCR amplified using primers 12708baiJ-F (5'-TGCATCTCC**CTCGAG**GGCACATTATGTTCCAGGC-3') containing *XhoI* restriction site (bold) and 12708baiJ-R (5'-TACAG**GTACCA**AGCATTGACTGTCCCGATGCT-3') containing *KpnI*

Figure 12. Schematic representation of cloning strategy for *baiH* gene from *Clostridium scindens* VPI 12708. Primers baiHF-PstI and baiHR-SacI (see text) with internally designed restriction sites were used to amplify a 2.2 kb stretch encoding the *baiH* gene. Following restriction digestion of both the insert and vector, and GENECLAN purification of electrophoresed restriction products, the 2.2 kb *baiH* insert was ligated with T4 DNA ligase to the 4.1 kb pSport1 vector with *PstI* and *SacI* "sticky ends" generating the 6.3 kb vector pSport1-baiH.



restriction site (bold). The *baiL* gene was PCR amplified using forward primer *baiLpETF* (5'-GTCAG**GAATTC**CAGTTTGCCATGGAGG-3') containing *EcoRI* restriction site (bold) and reverse oligonucleotide *baiLpETR* (5' ATT**CTCGAG**CATATTCAGGAAGCCGGCGTCG-3') containing *XhoI* site (bold) for pET-21a(+) vector.

The PCR reaction mixture contained 100 ng of *C. scindens* VPI 12708 genomic DNA, 100 pmol each primer, 0.3 mM dNTPs, Pfu polymerase buffer, and 2.5 U Pfu DNA polymerase in a total volume of 50 µl. Cycling parameters included an initial 3 min denaturation at 94°C followed by 30 cycles of amplification 94°C for 35 s, 52°C for 30 s and 68°C for 2 min followed by one additional 7 min elongation at 68°C. Amplification of the *baiH* PCR product was performed on a DNA Thermal cycler (MJ research, Watertown, MA).

The *baiJ* gene and all subsequent PCRs were performed on an Eppendorf Mastercycler Gradient PCR thermocycler (Eppendorf, Westbury, NY, USA). The *baiH* PCR product and pSport1 (Invitrogen, Carlsbad, CA) vector were double-digested with *PstI* and *SacI*, gel purified (GENECLEAN spin kit) and ligated with T4 DNA ligase (New England Biolabs).

The *baiJ* PCR product and pT7-SBP-2 Expression vector (Sigma, St. Louis, MO) were double-digested with *XhoI* and *KpnI*, gel purified (GENECLEAN spin kit) and ligated with T4 DNA ligase (New England Biolabs, Ipswich, MA, USA). Plasmids were transformed into *E. coli* DH5 $\alpha$  (Bio-line, Randolph, MA) by electroporation using an Eppendorf 2510 electroporator at 2500V. The recombinant plasmids (pSport1-baiH, pT7SBP2baiJ, p12708baiL) were purified from *E. coli* DH5 $\alpha$  and transformed into *E. coli* BL21-CodonPlus (DE3)-RIL or ArcticExpress *E. coli* for overexpression. The sequence was verified at the Virginia Commonwealth University Nucleic Acid Core Facility using standard primers.

### **Polyacrylamide gel electrophoresis and immunoblot analysis**

Protein samples were separated and analyzed on a 10% SDS-polyacrylamide gel electrophoresis (PAGE) (170). Protein samples were prepared for SDS-PAGE by heating for 5 min at 100°C in an equal volume of sample buffer (0.1 M Tris-HCl, 5 % SDS, 0.9 % 2-mercaptoethanol (2ME), 20 % glycerol, pH 6.8). Gels were stained with 0.2 % (w/v)

Coomassie brilliant blue R-250 (Sigma, St. Louis, MO) in ethanol: acetic acid: water (30:10:60, v/v/v). A broad-range protein weight marker (Bio-Rad, Hercules, CA) was used for protein size determinations. For western blot analysis, proteins were similarly separated by SDS-PAGE, and electrophoretically transferred to a nitrocellulose membrane by electroblotting using a Bio-Rad Immuno-Blot assay kit.

Antiserum against purified *baiH* gene product was raised in rabbits as described previously (121). Detection of the *baiH* gene product on nitrocellulose membranes was performed using a goat anti-rabbit alkaline phosphatase conjugate. Detection of the *baiJ* gene product was performed using anti-Strep-Tactin Horse radish peroxidase (HRP) conjugate (IBA, Gottingen, Germany). Detection of the *baiL* gene product was performed using Novagen HRP Anti-his tag monoclonal antibody (Novagen, Madison, WI, USA).

### **Enzyme Assays**

NADH:FOR activity was determined at 25°C by monitoring the oxidation of NADH at 340 nm ( $\epsilon_{340}=6.22 \text{ mM}^{-1}$

$^1\text{cm}^{-1}$ ) using a Beckman DU-640 spectrophotometer (Fullerton, CA). The standard assay mixture (1 ml) contained 100 mM sodium phosphate buffer (pH 6.8), 150  $\mu\text{M}$  NADH, 150  $\mu\text{M}$  FAD and an appropriate amount of enzyme added to initiate the reaction. Assays were performed under aerobic conditions in a 1 ml quartz cuvette with a 1 cm pathlength. One unit of NADH:FOR activity is defined as the amount of enzyme catalyzing the oxidation of 1  $\mu\text{mol}$  of NADH  $\text{min}^{-1}\text{mg}^{-1}$  of protein. Protein concentrations were determined according to Bradford (171) using the Bio-Rad protein assay kit, with BSA as standard for routine assays.

BaiH and BaiCD oxidoreductase activity assays were performed in reaction mixtures containing the following: 50 mM sodium phosphate buffer (pH 6.8), 1.5 mM  $\text{NAD}^+$ , 225  $\mu\text{M}$   $[24\text{-}^{14}\text{C}]\text{-3-dehydro-UDCA}$  (3-oxo-7 $\beta$ -hydroxy) or  $[24\text{-}^{14}\text{C}]\text{-3-dehydro-CDCA}$  (3-oxo-7 $\alpha$ -hydroxy) (specific activity of 0.01  $\mu\text{Ci } \mu\text{mole}^{-1}$  substrate) and purified recombinant BaiH protein (100  $\mu\text{g}$ ) or 1 mg cell extract containing overexpressed BaiCD.

BaiJ oxidoreductase activity assay was performed in the following mixture: 100 mM sodium phosphate buffer (pH 6.8), 1.5 mM  $\text{NAD}^+$  (or  $\text{NADP}^+$ ), 90  $\mu\text{M}$   $[24\text{-}^{14}\text{C}]$  3-dehydro-

DCA(~SCoA) (specific activity 550 DPM/nmole). The reaction mixture (500  $\mu$ l) was incubated 37 °C for 30 min. BaiJ reaction was initially brought to pH 10 with 5N NaOH and boiled 30 min to hydrolyze CoA ester (125). The pH was then lowered to ~3 with 5 N HCl and bile acid metabolites were extracted three times with 1 vol ethyl acetate, dried under a stream of nitrogen gas, dissolved in 15  $\mu$ l MeOH, spotted, separated by thin layer chromatography on Baker flex silica gel IB-2 plates (J.T Baker Chemical Co, Philipsburg, NJ) using a solvent system S1 (benzene:dioxane:glacial acetic acid 75:20:2 (v/v/v)) and radiolabeled bile acid metabolites were visualized by autoradiography (Kodak MS film, Rochester, N.Y) (172). Extracts prepared from *E. coli* containing pSport plasmids without *baiH* or *baiCD* gene inserts served as negative controls. Sections of TLC silica gel containing bile acid reaction products were scraped and radioactivity quantified by liquid scintillation spectrometry (Packard, Meriden, CT).



## **Expression and purification of recombinant baiH gene product**

*E. coli* BL21-CodonPlus(DE3)-RIL harboring pSport1-baiH were grown in 5 ml of modified TYGPN medium containing 100 µg of ampicillin per ml at 37°C for 8 hrs. The culture was stored at 4°C overnight. The next day, cells were collected by centrifugation (3,000 × g, 10 min) and suspended in 5 ml of TYGPN medium containing 100 µg/ml ampicillin. This cell suspension was used to inoculate 2 liter flasks containing 600 ml TYGPN broth medium containing 100 µg/ml ampicillin. The cells were cultivated at 37°C with shaking at 250 rpm for about 3 hr to reach an A<sub>600</sub> of about 0.4 at which point the cultures were induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to 1 mM final concentration. Cultures were grown at 25°C for an additional 4 hr at which time the cells were harvested by centrifugation at 5,500 × g, 15 min at 4°C. The cells were suspended in 400 ml of buffer A (50 mM sodium phosphate buffer; pH 6.8, 2% (w/v) glycerol, 15 mM of 2-mercaptoethanol (2ME)) cells were lysed by sonicating twice for 10 min intervals and cell debris removed by

centrifugation at  $21,000 \times g$  for 30 min. The supernatant fluid containing soluble recombinant BaiH was removed for further purification.

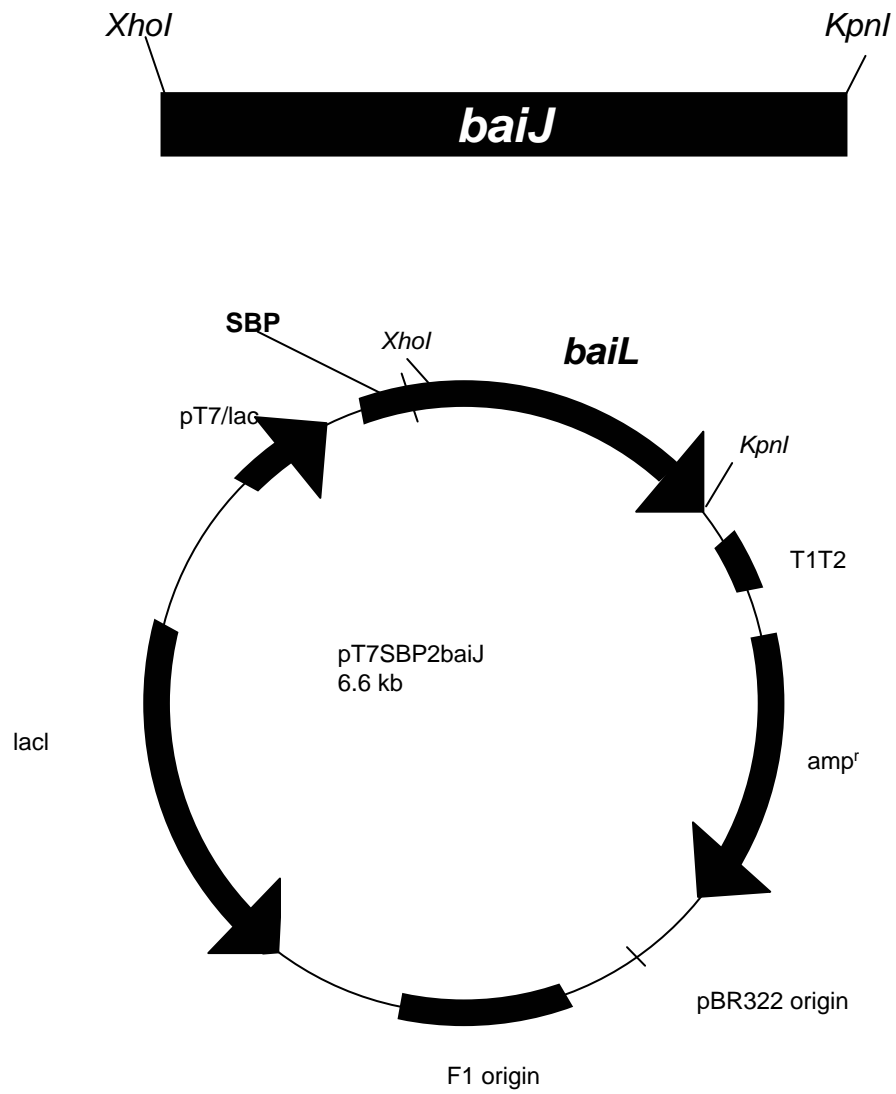
All purification methods were followed essentially as described previously with some modifications (129). Frozen cells were suspended in buffer A, disrupted by sonication on ice, and centrifuged at  $21,000 \times g$  for 30 min. Protein was concentrated and protein under 50 kDa removed using Centriprep YM50 (50 kDa cut-off) concentrators (Millipore, Billerica, MA). The protein solution was then applied to a  $2 \times 10$  cm Protein-Pak DEAE 40HR column (Waters, Milford, MA) equilibrated with buffer A. After the column was washed with 60 ml of buffer A, NADH:FOR was eluted with a 500 ml linear gradient of NaCl (0 to 1.0 M) in buffer A at a flow rate of  $2 \text{ ml min}^{-1}$ . Fractions with a yellow color and NADH:FOR activity were pooled and dialyzed overnight against 2 L of the same buffer. The dialyzed solution was loaded onto a  $1.0 \times 5.5$  cm Cibacron blue 3GA agarose column (Sigma, St. Louis, MO) equilibrated with buffer A. The column was then washed with 100 ml of buffer A, and protein partially eluted with 20 ml of buffer A containing 1 mM NADH, 1 mM FAD, and 20 mM dithiothreitol (DTT). The

remaining protein was eluted with buffer A containing 1M KCl. Active protein fractions eluted with NAD and FAD were pooled and loaded onto a 1.0 × 6.4 cm Bio-Scale CHT5-1 hydroxyapatite column (Bio-Rad, Hercules, CA) equilibrated with buffer C (5 mM potassium phosphate buffer pH 7.0, 2% glycerol, 15 mM 2-ME). After the column was washed with buffer C at 1ml min<sup>-1</sup>, active fractions were eluted with a linear gradient of 0 to 0.5 M potassium phosphate (pH 7.0). The purified protein from this step was dialyzed with buffer A, concentrated by a Centriprep YM50 concentrator and stored in buffer A containing 50% glycerol (w/v) at -20 °C. Protein remained active for 4 weeks under these conditions.

### **Expression and purification of the *baiJ* and *baiL* gene product**

*E. coli* Arctic ExpressBL21-CodonPlus(DE3)-RIL harboring pT7SBP-*baiJ* (Figure 13) was grown in 5 ml of LB medium (tryptone 10 g; yeast extract 5g; NaCl 10g) with 100 µg of ampicillin per ml at 37°C for 8 hrs. The culture was transferred to 2 Liter flasks containing 500 ml

Figure 13. Cloning strategy for *baiJ* gene from *C. scindens* VPI 12708. The *baiJ* gene was PCR amplified using primers 12708baiJ-F (5'-TGCATCTCC**CTCGAG**GGCACATTATGTTCCAGGC-3') containing *XhoI* restriction site (bold) and 12708baiJ-R (5'-TACAG**GTACCA**AAGCATTGACTGTCCCGATGCT-3') containing *KpnI* restriction site (bold). Double digestion of PCR product and vector pT7SBP-2 with *XhoI* and *KpnI*, gel purification and ligation resulted in vector pT7SBP2baiJ.



Modified TYGPN broth medium containing 100 µg/ml ampicillin. The cells were cultivated at 18°C with shaking at 250 rpm for about 18 hr to reach an  $A_{600}$  of about 0.4 at which point the cultures were induced by addition of IPTG to 0.8 mM final concentration. Cultures were grown for an additional 12 hr at which time the cells were harvested by centrifugation at  $5,500 \times g$ , 15 min at 4°C. The cells were suspended in 40 ml of Strep buffer (50 mM sodium phosphate buffer; pH 7.9, 15 mM of 2-mercaptoethanol (2ME) and 1M NaCl) and the cells were lysed by sonicating twice for 15 min intervals. Cell debris was removed by centrifugation at  $21,000 \times g$  for 30 min. The supernatant fluid containing soluble recombinant BaiJ was removed for further purification.

All purification methods using IBA Strep-tag purification were followed according to the manufacturer (IBA, Gottingen, Germany). Eluents were concentrated by Centriprep YM10 (10 kDa cut-off) concentrators (Millipore, Billerica, MA).

The *baiL* gene was expressed in *E. coli* BL21 CodonPlus(DE3)-RIL in a similar manor to the *baiJ* gene, except at 25°C and 3 hour growth following induction with

IPTG. Protein concentration was determined by Bradford protein assay.

### **Enzymatic and biological preparation of bile acid substrates**

[24-<sup>14</sup>C]-CDCA was purchased from American Radiolabeled Chemicals, Inc. (St. Louis, Mo), and determined to be 99.9% pure by TLC. *C. absonum* expresses both 7 $\alpha$ -HSDH and 7 $\beta$ -HSDH and can be used to efficiently epimerize CDCA to UDCA (76, 77). [24-<sup>14</sup>C]-UDCA was biologically synthesized from [24-<sup>14</sup>C]-CDCA using *Clostridium absonum* ATCC 27555 grown in BHI. *C. absonum* was induced at O.D<sub>600nm</sub> 0.4 by addition of 200  $\mu$ M sodium chenodeoxycholate (CDCA) plus 2  $\mu$ Ci of [24-<sup>14</sup>C]-CDCA for production of [24-<sup>14</sup>C]-UDCA (0.01  $\mu$ Ci  $\mu$ mole<sup>-1</sup>). After 24 hr growth following induction, the pH of the culture medium was lowered to ~pH 3 and bile acid metabolites were extracted twice with ethyl acetate, dried under a stream of nitrogen gas and dissolved in 100  $\mu$ l of absolute methanol (MeOH). [24-<sup>14</sup>C] CDCA, [24-<sup>14</sup>C] 7-oxo-LCA and [24-<sup>14</sup>C] UDCA were separated by TLC as described above. Non-radiolabeled CDCA and UDCA standards

(Sigma, St. Louis, Mo, USA) were also chromatographed along with the radiolabeled reaction products and this portion of the plate stained with phosphomolybdic acid for detection. Radiolabeled bands co-migrating with UDCA standard (Sigma) were scraped and extracted with ethyl acetate, dried, and suspended in MeOH.

[24-<sup>14</sup>C] UDCA was converted to [24-<sup>14</sup>C] 3-dehydro-UDCA (3-oxo-7 $\beta$ -hydroxy) enzymatically using commercially available 3 $\alpha$ -HSDH (Sigma) in the following reaction mixture: 0.5 ml total volume 50 mM sodium phosphate buffer (pH 6.8), 1 U 3 $\alpha$ -hydroxysteroid dehydrogenase, 1.5 mM NAD<sup>+</sup>, 100  $\mu$ M unlabeled UDCA and 1  $\mu$ Ci [24-<sup>14</sup>C]-UDCA. The reaction mixture was allowed to incubate for 4 h at 37 °C. Bile acids were separated by TLC as described above and radiolabeled bile acid reaction products were detected by autoradiography.

[24-<sup>14</sup>C] 3-oxo-DCA was produced from commercially available [24-<sup>14</sup>C] CA and non-labeled CA (Sigma). 1  $\mu$ Ci [24-<sup>14</sup>C] CA and 5  $\mu$ l 100 mM CA (500 nmoles) was added to 5 ml BHI and inoculated with *C. scindens* VPI 12708 and grown 48 hrs at 37 °C. The reaction products were extracted in the usual manner. The concentrated reaction products were



separated by TLC in solvent system S1 (172). The band corresponding to DCA standard was scraped and [24-<sup>14</sup>C] DCA was extracted with ethyl acetate and concentrated under nitrogen gas. The pellet was resuspended in 50 mM sodium phosphate buffer and reacted with commercially available 3 $\alpha$ -HSDH in the presence of 100  $\mu$ M NAD<sup>+</sup> as co-substrate. Substrate and product were separated by TLC and band corresponding to [24-<sup>14</sup>C] 3-oxo-DCA ( $R_f$ =1.74; Solvent SI) was extracted (118).

### **Chemical synthesis of bile acid substrates**

3-oxo-DCA-CoA is hypothesized to be a common substrate for both the *baiJ* and *baiL* gene products. Thus, 3-oxo-DCA was synthesized from commercially available DCA essentially as described by Jones et al. (1949) and Danielsson et al. (1962) with some modifications (173, 174). Toluene was substituted for benzene in the reflux step as well as the chromatography. Briefly, 5 grams of deoxycholic acid was methylated overnight in 250 ml

absolute methanol and 3 ml concentrated sulfuric acid. The solvent was evaporated under nitrogen gas and the crude methyl ester was resuspended in chloroform and chromatographed over 100 g silicic acid with a step-wise gradient of acetone in chloroform up to 25%. The DCA-methyl ester was then oxidized in 140 ml toluene and 55 ml acetone with 5.3 g aluminum *tert.*-butoxide under reflux for 18 hours. After cooling, the reaction mixture was poured into an equal volume 2N sulfuric acid and emulsified. The toluene layer was separated and washed once again with 2N sulfuric acid followed by equal volumes of 7% sodium hydrogen carbonate (twice), and water (once). The toluene layer was then dried under  $\text{CaCl}_2$  and evaporated. The methyl ester was hydrolyzed overnight in 2 M methanolic potassium hydroxide. The crude acid was added slowly to 0.5 N HCl and filtered under vacuum. The crude acid was dissolved in ethyl acetate, dried with sodium sulfate, concentrated under nitrogen gas and chromatographed on 100 g silicic acid with an increasing gradient of acetone in toluene. Fractions eluted with 15% acetone in toluene yielded semi-pure 3-oxo-DCA as observed on thin layer chromatography. Identity of 3-oxo-DCA was

confirmed by comparison of  $R_f$  value after reaction with 3 $\alpha$ -HSDH in the presence of 150  $\mu$ M NADH. Commercially available deoxycholic acid (Sigma) was used as a standard. Material used in enzyme reactions was scraped from TLC plate and purified by ethyl acetate extraction.

The sodium hydrogen carbonate fraction was acidified with dilute (10% v/v) sulfuric acid and allowed to precipitate overnight. The precipitate was vacuum filtered and worked up and purified as described above.

Pure 3-oxo-DCA was esterified to coenzyme A by the method of Killenberg and Dukes (1976) which was a slight modification of Shah and Staple (1968) (175, 176). Tetrahydrofuran and methylene chloride were distilled prior to use and  $\gamma$ -collidine was dried over sodium sulfate. 3.5 mg [24- $^{14}$ C] 3-oxo-DCA (550 DPM/nmole) was added to 700  $\mu$ l dry methylene chloride and incubated ten minutes at room temperature. 250  $\mu$ l of methylene chloride containing 1.4  $\mu$ l ethyl chloroformate was added and the solution was magnetically stirred on low for one hour. The solvent was then evaporated under nitrogen gas and the pellet was suspended in 600  $\mu$ l tetrahydrofuran. The tetrahydrofuran solution was slowly added to 600  $\mu$ l solution of 62.5 M CoA

in water (pH 8.0). After fifteen minute incubation, the pH was lowered to 5 with 1% perchloric acid and the organic solvent evaporated under nitrogen gas. The aqueous layer containing 3-oxo-deoxycholyl-CoA was washed three times with 2 vol diethyl ether. The aqueous solution was acidified with 170 ml 10% perchloric acid and the 3-oxo-deoxycholyl-CoA was collected by centrifugation at 15,000 x g for 30 minutes. The pellet was stored at -80°C overnight and used the next day.

#### **Purification and mass spectrometry of BaiCD and BaiH reaction products**

The analysis and identification of BaiH catalyzed bile acid reaction product was carried out on a Shimadzu HPLC (Kyoto, Japan) with a SIL-HTC Autosampler by reverse-phase HPLC using a Brownlee RP-300 (Perkin Elmer, Wellesley, MA) C8 7m 2.1 i.d. × 100 mm column. The mobile phase consisted of 10 mM ammonium acetate (solvent A) and 10 mM ammonium acetate in acetonitrile (solvent B). The column was equilibrated in solvent A. After injection of a 5 ml sample of the bile acid, solvent B was increased from 0-100% at 10% min<sup>-1</sup>. The column was regenerated by

switching back to solvent A and re-equilibrating for 5 min. The flow rate was  $0.2 \text{ ml min}^{-1}$  and the eluted material was passed into the electrospray ionization interface of a MDS Sciex/Applied Biosystems API 4000 Q Trap (Foster city, CA) operating in the negative ion mode. Mass spectra were recorded over a 350-700  $m/z$  range.

### **Genome-walking by polymerase chain reaction**

Genome-walking DNA libraries were prepared using the Clontech Universal Genome-Walker Kit according to the manufacturer. DNA was prepared essentially as described previously (177) with a few alterations. To produce high molecular weight DNA (~50 kb), pipetting steps during phenol/phenol:chloroform:iso-amyl alcohol were changed to pouring the aqueous phase, which prevents substantial DNA sheering. An 18 gauge needle was used to pierce the very bottom of 50 ml falcon conical tubes following centrifugation and the organic phase was emptied. DNA was then precipitated by addition of 10% 3N sodium acetate and 3 volumes of absolute ethanol, washed and resuspended in TE buffer, ph 7.6, overnight with agitation at 45 °C.

DNA quality was determined on a 0.5% agarose gel containing 75 µg/ml ethidium bromide. Prior to electrophoresis, the DNA sample was heated to 65°C. The gel was run 18 hours at 25 V overnight along with a High Molecular Weight DNA Marker (Invitrogen). DNA >48 kb was restriction digested and adapter ligated using GenomeWalker kit. Genomic DNA libraries were stored at -20°C. Oligonucleotides were designed using Primer3 software with  $T_m$  values between 60°C and 70°C with  $\geq 45\%$  G+C and  $\geq 25$  mer in length.

PCR reactions were prepared using Advantage cDNA Polymerase Kit or Advantage 2 Polymerase Kit (Clontech) as follows: 40 µl ddH<sub>2</sub>O, 5 µl 10x Advantage PCR buffer, 1 µl dNTP mix (10 mM each), 1 µl adaptor primer AP1 (10 µM), 1 µl gene-specific primer (10 µM), 1 µl Advantage Polymerase mix (50X) and 1 µl adaptor-ligated DNA restriction library template. Reaction mixtures were mixed briefly and tubes quick spun. "Touchdown" PCR amplification was performed on Eppendorf Mastercycler Gradient PCR machine using the following parameters: 95°C for 25 s, 72°C for 3 min (7 cycles); 94°C for 25 s, 67°C for 3 min (32 cycles); 67°C for 7 min, hold 4°C. Products (5 µl) were separated by

electrophoresis on a 1% agarose/EtBr gel and visualized on a Fotodyne analyzer. PCR products of interest were GENECLAN purified from the gel and ligated into a pCR8 GW TOPO TA Cloning vector (Invitrogen) and transformed into TOP 10 Chemically Competant *E. coli* according to the manufacturer (Invitrogen). Colonies were cultivated in LB broth containing 100 µg/ml spectinomycin. Plasmids were purified using Qiagen Miniprep Spin Kit according to the manufacturer. Sequencing was performed at the Nucleic Acid Research Facility, School of Medicine, Virginia Commonwealth University on ABI 3700 Sequencer using oligonucleotides synthesized by Integrated DNA Technologies (Coralville, IA, USA).

### **SMART-RACE Polymerase Chain Reaction**

*C. scindens* VPI 12708 as well as *C. hylemonae* were cultivated in 400 ml BHI containing 0.02% fructose to an O.D.<sub>600nm</sub> 0.3 (*C. scindens*) or 0.2 (*C. hylemonae*) at which point the cultures were induced 3 times at hourly intervals with 50 µM CA. One hour after the third induction, 1 ml culture was transferred to a serum bottle containing [24-<sup>14</sup>C] CA (0.5 µCi/500 nmoles) and incubated

37°C 30 minutes. A non-induced control was also tested for 7 $\alpha$ -dehydroxylation activity. Cells were centrifuged 13,000 x g for 30 min at 4°C. The cell pellet was suspended in 1 ml TE buffer and cells transferred to 1.5 ml microcentrifuge tube and spun 1 min at 13,600 x g. The supernatant was decanted and bacterial cells were stored in *RNAlater* solution (Ambion, Austin, TX, USA) and stored at - 70°C until further processing.

Once 7 $\alpha$ -dehydroxylating activity, and thus induction, was confirmed by presence of [24-<sup>14</sup>C] DCA in [24-<sup>14</sup>C] activity assay, total RNA was extracted using the Ribopure Bacteria RNA extraction kit (Ambion) essentially according to the manufacturer with the exception that cells were broken on a Mini-bead beater (Biospec Products Inc, Bartlesville, OK, USA) at 3,000 rpm for two cycles of 1 min. Optional DNase treatment was performed for 1 hour at 37°C. Some of the RNA was poly-A tailed prior to cDNA synthesis using the *A-Plus* Poly(A) Polymerase Tailing Kit (Epicentre Biotechnologies, Madison, WI) according to the manufacturer with a 15 minute extension time. cDNA was generated using the SMART-RACE cDNA synthesis kit (Clontech) according to the manufacturer starting with 1



µg total RNA. 5'CDS primer and 3'CDS primer (Clontech) were used to make 5' RACE ready cDNA and 3' RACE ready cDNA, respectively, when poly-A RNA was used as the template; whereas gene-specific primers replaced 5'CDS primer and 3' CDS primer when total bacterial RNA was used as template. Reaction was carried out for 1.5 hrs at 42°C and terminated by addition of 50 µl TE buffer pH 7.6 and heating at 70°C for 7 minutes. SMART RACE PCR of 5' ends was performed using the Advantage-GC 2 PCR Kit (Clontech) according to the manufacturer's instruction. 0.5 µl GC-Melt was added to each 50 µl reaction volume along with 2.5 µl cDNA sample. A nested PCR was performed when product concentration was low or to reduce non-specific products. PCR products were excised from agarose gel and purified using the GENECLAN Spin Kit (MP Biomedicals, Solon, Ohio). The purified PCR product was then cloned into a pCR8 GW TOPO TA cloning vector (Invitrogen), transformed into TOP 10 Chemically Competent *E. coli* (Invitrogen) and selected on LB agar containing 100 µg/ml spectinomycin. Plasmids containing insert were sequenced.

### **Sequence analysis and Assembly**

Sequence ABI files were edited and assembled using Vector NTI software (Invitrogen, Carlsbad, CA, USA), as well as CAP3 Sequence Assembly Program <http://pbil.univ-lyon.fr/cap3.php>. Pairwise alignments were performed to determine percent identity/similarity using EMBOSS Needle program <http://www.ebi.ac.uk/emboss/align/>. Homology searches were made against the UniProt database. Multiple sequence alignments were prepared using T-COFFEE <http://tcoffee.vital-it.ch/cgi-bin/Tcoffee.cgi/index.cgi> (178) or CLUSTALW2 <http://www.edi.ac.uk/Tools/clustalw2/index.html> with some manual editing based on structural data. Sequence conservation within multiple alignments was emphasized using the BOXSHADE utility [http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html). Protein  $M_r$  calculations were performed using the Compute pI/MW tool on the ExPASy Proteomics Server <http://www.expasy.ch/tools/pitool.html>. Open reading frames (ORF) were detected using the ORF Finder Program

<http://www.ncbi.nlm.nih.gov/gorf/gorf.html> and searched against Genbank by BLAST analysis (179).

*Clostridium scindens* VPI 12708 *baiH* (AAC45417) and *Clostridium hiranonis* *baiCD* (AAF22846) genes were sequenced previously (59). Secondary structural predictions were carried out with the PSIPRED program (<http://bioinf.cs.ucl.ac.uk/psipred/>) (180, 181). Secondary structural predictions of the BaiH gene product were compared with the known secondary structure from the crystal structure of 2, 4-dienoyl CoA reductase (*dcr*) from *Escherichia coli* (182).

### **Nucleotide sequence accession number**

The *baiC* and *baiD* genes from *C. scindens* were originally reported as separate genes, however, after discovery that the *baiCD* is a single gene product in *C. hiranonis*, the *baiC* and *baiD* gene region was resequenced from *C. scindens* and found to encode a single gene (D.H. Mallonee and P.B. Hylemon, unpublished data). Also, a *baiH* gene has been located downstream of the *bai* operon in *C. hiranonis* and has not been reported previously (J.E.

Wells and P.B. Hylemon, unpublished data). GenBank accession numbers for the *baiCD* gene from *C. scindens* VPI 12708 is U57489 and the *baiH* gene from *C. hiranonis* T0931 is AF210152. Genbank accession numbers for the *baiBCDEFGHI* operon from *C. hylemonae* TN271 is (EU675332). The *baiA* gene from *C. hylemonae* TN271 can be found using accession number (EU675329). Genbank accession numbers for the *baiJKL* genes from *C. scindens* VPI 12708 and *C. hylemonae* TN721 are (Eu675330) and (EU675331), respectively.

## RESULTS

### Characterization of the BaiCD and BaiH

#### Amino acid sequence alignment and secondary structural prediction of the BaiH.

Protein-protein BLAST searches of the *baiH* gene product suggests homology with a family of pyridine-nucleotide dependent flavoproteins (Table 4) including NADH oxidase from *Thermoanaerobacter brockii* (CAA47660) (183), 2-enoate reductase from *Clostridium tyrobutyricum* (CAA71086) (184) and *Moorella thermoacetica* (CAA76082) (184), 2, 4-dienoyl-CoA reductase from several bacterial species, and Old yellow enzyme (OYE) from *Thermoanaerobacter tengcongensis* (NP\_623692) (185). *E. coli* 2, 4-dienoyl CoA reductase (*dcr*) (AAC76116) is a 73 kDa protein which uses NADPH as an electron donor and contains a 4Fe-4S cluster, 1 mol FMN, and 1 mol FAD per mole subunit (182). The *dcr* gene product from *E. coli* shares 32 % identity and 49 % similarity at the amino acid sequence level with the *baiH* gene product from *C. scindens*

**Table 4. Characteristics of selected BaiH homologues**

Enzyme	Organism	Subunit M (kDa)	Quaternary Structure	Cofactor	Electron donor	Electron Acceptor	Percent Identity/Similarity to BaiH	Accession Number	Reference
BaiH	<i>Clostridium scindens</i>	72	homotrimer	1 FAD, Fe-S <sub>y</sub> *	3-dehydro-UDCA	NAD <sup>+</sup>	100/100	AAC45417	This study, 2, 15
BaiCD	<i>Clostridium scindens</i>	70	ND	ND	3-dehydro-CDCA	NAD <sup>+</sup>	32/46	EF539210	This study
Enoate reductase	<i>Clostridium tyrobutyricum</i>	73	homododecamer	1 FAD, 1 FMN 1 4Fe-4S	NADH	α, β-unsaturated carbonyls	29/48	CAA71086	51
2,4-dienoyl CoA Reductase	<i>Escherichia coli</i>	73	monomer	1 FAD, 1 FMN 1 4Fe-4S	NADPH	unsaturated fatty acyl-CoA	29/49	1PS9A	21
Old Yellow Enzyme (OYE2)	<i>Saccharomyces cerevisiae</i>	49.9	homodimer	1 FMN	NADPH	α, β unsaturated carbonyls	23/42†	Q03558	57, 61
YqjM (OYE)	<i>Bacillus subtilis</i>	37.4	homotetramer	1 FMN	NADPH	α, β unsaturated carbonyls and nitroaromatics	29/49‡	BAA12619	14, 26
NoxB1	<i>Archaeoglobus fulgidus</i>	69	monomer	FAD	NADH	unknown	35/55	NP_069291	25

\* Data inconclusive (2)

† alignment between full length OYE2 and amino acids 1-457 of BaiH

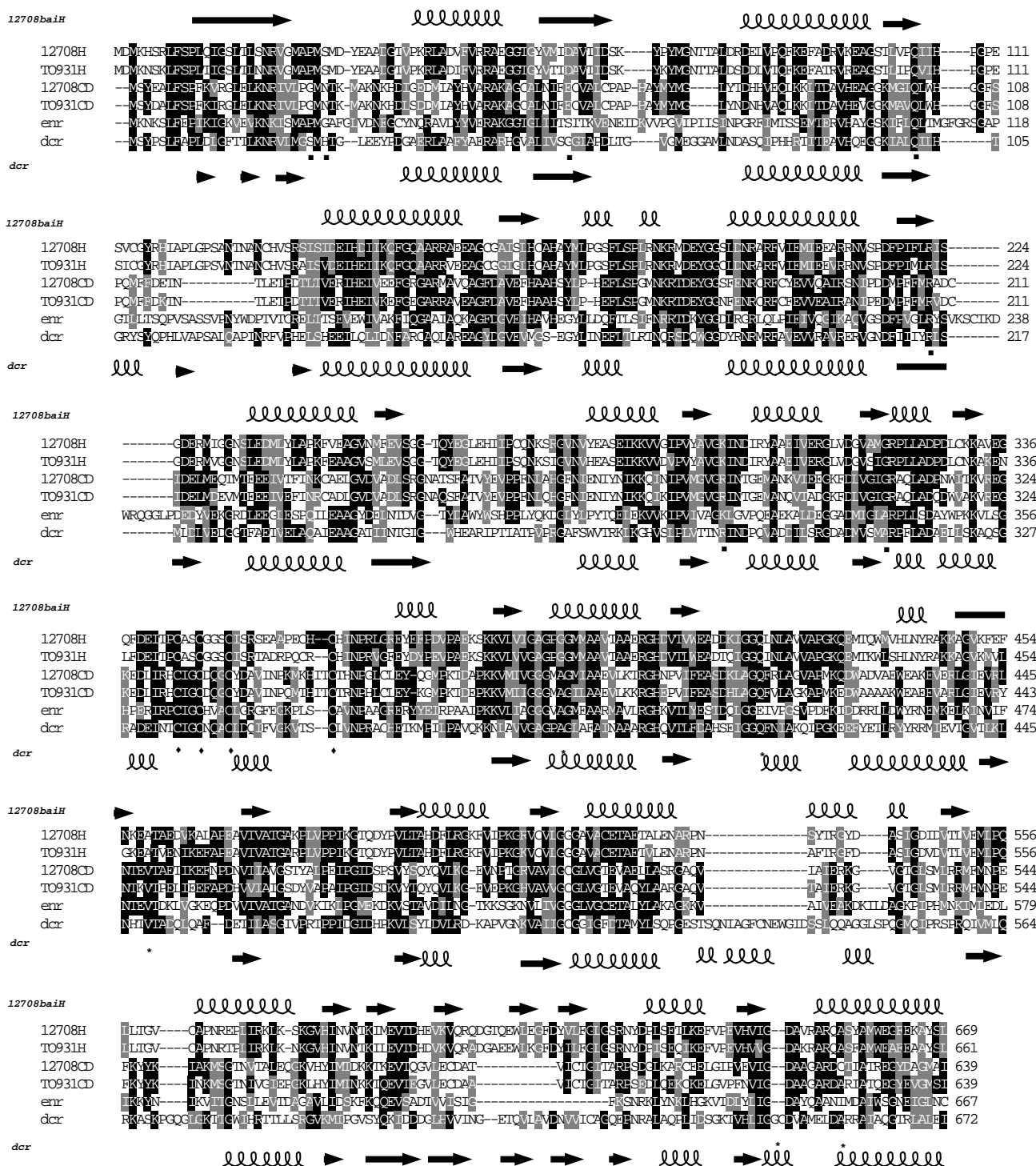
‡ alignment between full length YqjM and amino acids 1-351 of BaiH

VPI 12708. A 2.4 Å structure has been solved for the *E. coli dcr* gene product allowing a basis of comparison for secondary structure prediction of the *baiH* gene product (Figure 14). The predicted secondary structure of the BaiH is very similar to that of the *dcr* gene product suggesting a similar TIM-barrel fold. It is interesting to note the conserved cysteine consensus pattern of C-(2X)-C-(2-3X)-C-(11-12X)-C between the BaiH, *E. coli* Dcr and other homologues suggesting that the BaiH and BaiCD contain 4Fe-4S centers.

Several *baiCD* and *baiH* homologues introduce carbon-carbon double bonds. This suggests that the *baiCD* and *baiH* gene products encode flavoproteins which may introduce carbon-carbon double bonds into the bile acid structure. Previously, we identified a 3-oxo- $\Delta^4$ -cholic acid intermediate in the bile acid 7 $\alpha$ -dehydroxylation pathway (117). Because there are two homologous genes and two stereochemistries differing about the 7-hydroxyl between CA, CDCA (7 $\alpha$ -hydroxy) and UDCA (7 $\beta$ -hydroxy), we hypothesize that the *baiCD* and *baiH* introduce the  $\Delta^4$ -bond in bile acids, recognizing one stereochemistry about the 7-hydroxyl or the other.

Figure 14. Amino acid sequence alignment of the BaiH (AAC45417) from *C. scindens* VPI 12708 with protein homologues: 2,4-dienoyl-CoA reductase (*dcr*) (AAC76116) from *E. coli*, 2-enoate reductase (CAA71086) from *C. tyrobutyricum*, the *baiCD* gene product from *C. scindens* (EF539210) and *C. hiranonas* (AAF22846) and the *baiH* gene product from *C. hiranonis* (EF539211). Secondary structural elements based on the crystal structure of DCR (182) are shown (below alignment). Predicted secondary structural elements of the BaiH protein are also shown (above alignment). Arrows correspond to  $\beta$ -sheets and loops to  $\alpha$  helices. Residues corresponding to FMN binding (■), FAD binding (\*), and cysteine residues involved in a 4Fe-4S center (♦) from the crystal structure of DCR are also shown. The alignment was generated with the T-COFFEE program (178) and sequence similarities/identities highlighted with Boxshade 3.1.





### **Purification of recombinant BaiH**

Initially, the *baiH* gene was cloned into a pET24a (+) vector with C-terminal hexa-histidine tag to make a BaiH-His recombinant. The BaiH-His showed NADH:FOR activity in crude extract prior to nickel affinity chromatography. However, following nickel affinity chromatography, the purified enzyme lost detectable NADH:FOR activity. Problems with His-tag use has also been reported for the YqjM protein from *Bacillus subtilis*; a *baiH* homologue (186). The quaternary structure of recombinant His-tagged YqjM was reported to be altered by the His-tag resulting in expression of an inactive enzyme. This is not the case with the *baiH* gene product as the enzyme showed NADH:FOR activity in crude extract similar to previously reported in recombinant BaiH lacking affinity tag (129) before His-binding chromatography (data not shown). The BaiH is known to contain 1 mol copper per subunit (129). While the mechanism for this loss of activity is not known, we speculate that the nitriloacetic

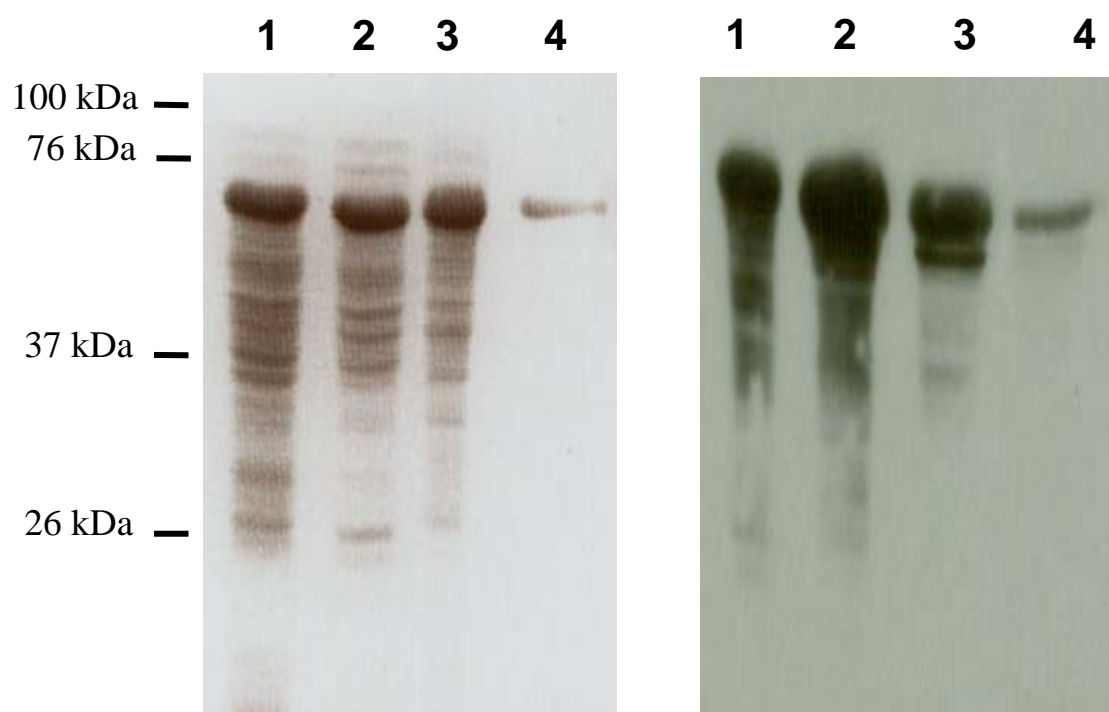
acid, which chelates  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ , and  $\text{Co}^{2+}$ , and in turn binds the hexa-histidine tail, may have chelated copper from the enzyme causing inactivation (187, 188). Addition of various concentrations of  $\text{Cu}^{2+}$  to the pure enzyme did not restore activity. Therefore, we opted to clone the gene into the pSport1 vector without an affinity tag.

SDS-PAGE analysis of the active fractions after CHT5-1 hydroxyapatite chromatography revealed a highly enriched protein with a subunit molecular mass of 72 kDa that was reactive against anti-BaiH antiserum (**Figure 15**). The specific activity and purification fold were dramatically increased by Cibacron Blue 3GA and hydroxyapatite chromatography (**Table 5**). The *baiH* gene product was purified 23-fold over the crude extract. The final enzyme preparation had a NADH:FOR specific activity of  $16.4 \text{ U mg}^{-1}$ . NADH:FOR enzyme activity existed in bright yellow fractions as observed previously (129).

**Stereospecific NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity.**

The hypothesis that the *baiH* gene product encodes an NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase was

Figure 15. SDS-PAGE and immunoblot analysis of *baiH* gene product (NADH:FOR) overexpression and purification (a.) Proteins were separated on a 10 % SDS-polyacrylamide gel and stained with Coomassie Blue R-250. Lane 1, cell free extract after IPTG induction (20 µg loaded); Lane 2, flow-through from DEAE 40HR (20 µg loaded); Lane 3, flow-through from Cibacron Blue 3GA agarose (20 µg loaded); lane 4, flow-through from Hydroxyapatite CHT5-1 column (5 µg of purified enzyme loaded). (b.) Immunoblot of SDS-PAGE using anti-BaiH antiserum.



**Table 5. Purification scheme of *baiH* gene protein in *E.coli* BL21(DE3)-RIL(pSport1-baiH)**

Step	Total Volume (mL)	Total Protein (mg/mL)	Total activity (U)	Specific activity (U/mg)	Purification (fold)	Recovery (%)
Cell free extract	35	410	93.3	0.7	1.0	100
Centriprep treatment	20	175	80	1.2	1.7	86
DEAE 40 HR Cellulose	11	92	66	3	4.3	71
Cibacron blue 3GA agarose	5.2	63	55.5	8	11.4	60
Hydroxyapatite CHT	2.3	28	35.1	16.4	23.4	38

tested by addition of substrates [24-<sup>14</sup>C] 3-dehydro-UDCA or [24-<sup>14</sup>C] 3-dehydro-CDCA with NAD<sup>+</sup> and purified BaiH. A single reaction product of more hydrophilic character was detected by autoradiography following TLC when [24-<sup>14</sup>C] 3-dehydro-UDCA and NAD<sup>+</sup> were substrates (**Figure 16**).

Previous studies have shown the specificity of this enzyme for NAD(H) (129) and the BaiH did not show 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity with NADP(H) (data not shown). The bile acid metabolite generated following BaiH catalyzed reaction with NAD<sup>+</sup> was scraped from TLC plates, pooled, further purified by ion exchange chromatography and identified by electrospray mass spectrometry (Figure 17). The deprotonated molecular ion given by the product was 2 Da lighter than that given by the 3-dehydro-CDCA substrate, consistent with a loss of 2 hydrogens from the substrate and potentially indicating a 3-dehydro- $\Delta^4$ -CDCA structure. Previous observations showed loss of the 5 $\beta$  hydrogen during incubation of [5 $\beta$ -<sup>3</sup>H] [24-<sup>14</sup>C] CA to CA induced whole cells and cell extracts of *C. scindens* VPI 12708 (118). Taken together, these observations strongly suggest the formation of a 3-dehydro- $\Delta^4$ -UDCA structure.

Figure 16. TLC autoradiograph of BaiCD and BaiH catalyzed NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity assays. The reactions were carried out in a 500  $\mu$ l volume consisting of 50 mM sodium phosphate buffer (pH 6.8) containing 225  $\mu$ M of each respective substrate with a specific activity of 0.01  $\mu$ Ci  $\mu$ mole<sup>-1</sup>. A: [24-<sup>14</sup>C] 3-dehydro-CDCA standard, B: [24-<sup>14</sup>C] 3-dehydro-UDCA, C: [24-<sup>14</sup>C] 3-dehydro-CDCA + 100  $\mu$ g cell-free extract of *E. coli* (plasmid, no insert) D: [24-<sup>14</sup>C] 3-dehydro-CDCA + 1.6 U purified BaiH, E: [24-<sup>14</sup>C] 3-dehydro-CDCA + 100  $\mu$ g *E. coli* BL21(DE3) expressing *baiCD*. F: [24-<sup>14</sup>C] 3-dehydro-UDCA + 100  $\mu$ g *E. coli* BL21(DE3) (plasmid, no insert), G: [24-<sup>14</sup>C] 3-dehydro-UDCA + 100  $\mu$ g cell-free extract of *E. coli* BL21(DE3) expressing *baiCD*, H: [24-<sup>14</sup>C] 3-dehydro-UDCA + 1.6 U purified BaiH. All reactions were performed with 1.5 mM NAD<sup>+</sup> as co-substrate.



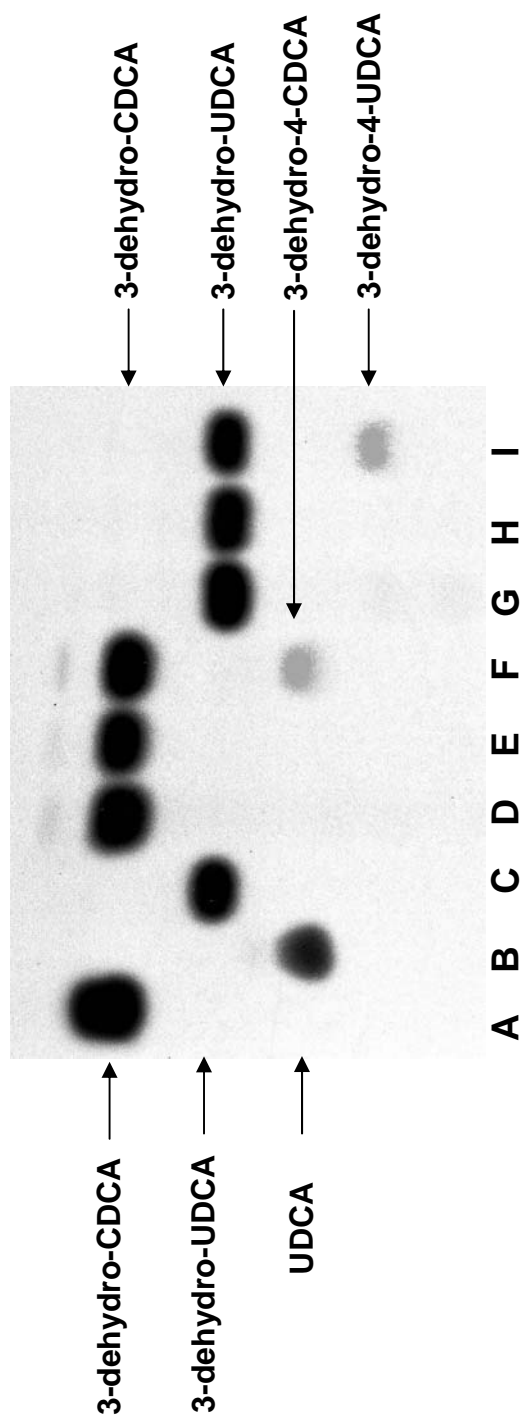
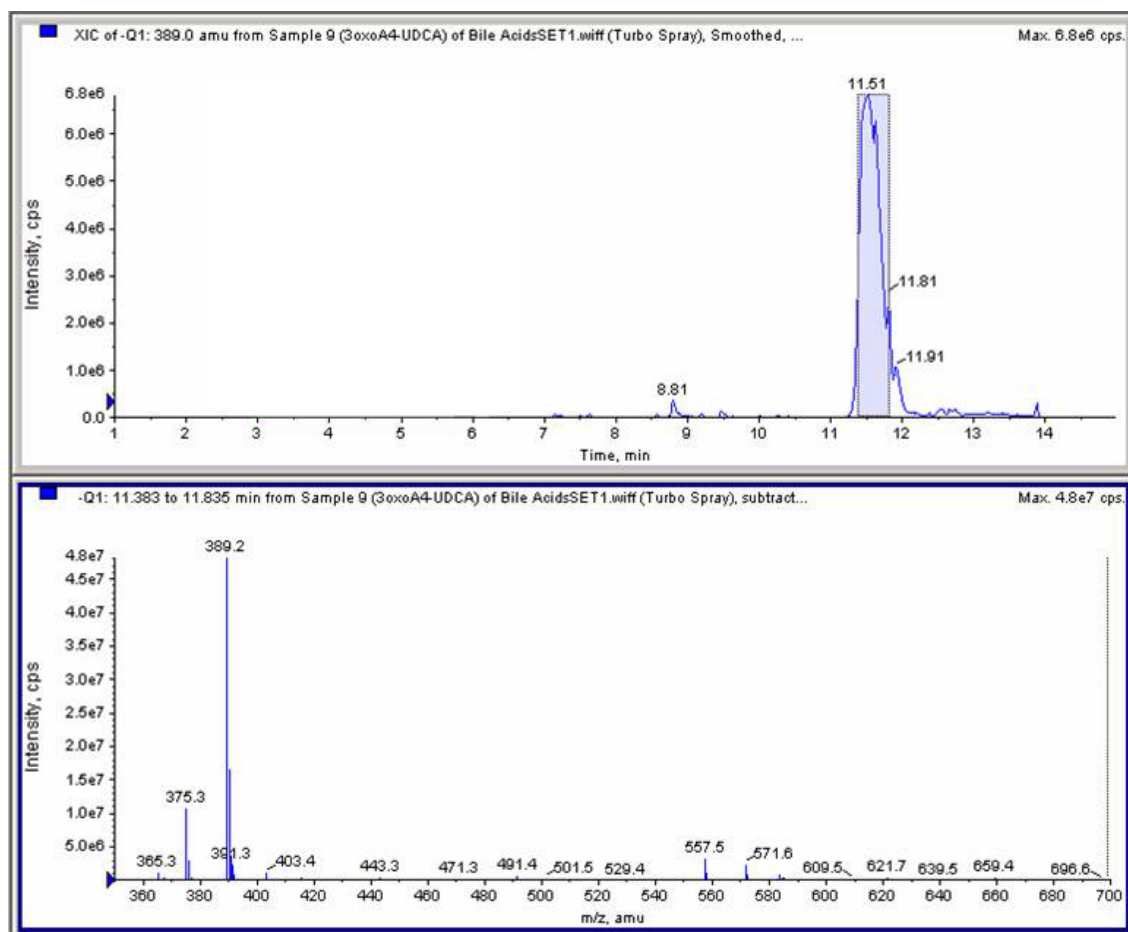


Figure 17. HPLC and mass spectrometry analysis of BaiH catalyzed bile acid reaction products. The bile acid metabolites produced by purified baiH gene protein were separated by TLC, scraped and further purified by reverse-phase HPLC-ESI-MS. The initial mobile phase was 10 mM  $\text{NH}_4\text{OAc}$ ; after injection bound materials were eluted by a 0-100 % linear gradient of acetonitrile in 10 mM  $\text{NH}_4\text{OAc}$  at  $10\% \text{ min}^{-1}$ . The eluate was passed into the electrospray interface operating in the negative ion mode. Ions eluting from  $m/z$  350-700 were recorded and analyzed. The upper trace is a selected ion chromatogram for  $m/z$  389. The mass spectra for the peak at 11.52 min were summed and presented in the bottom trace. The  $m/z$  389.2 ion corresponds to the expected  $[24\text{-}^{14}\text{C}]$  3-dehydro-4-UDCA product.



In addition, the *baiCD* gene was overexpressed in *E. coli* BL21 (DE3) and tested for 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity. The *baiCD* gene product was detected by SDS-PAGE at the expected molecular mass of 70 kDa (Figure 18). Addition of *baiCD* overexpressed *E. coli* cell extract to reaction mixture containing 1.5 mM NAD<sup>+</sup> resulted in a product of greater hydrophilicity than the starting substrate [24-<sup>14</sup>C] 3-dehydro-CDCA (Figure 16). This reaction product was not obtained using cell-extract of *E. coli* BL21(DE3) containing pSport1 vector without *baiCD* gene insert. BaiCD did not recognize the substrate [24-<sup>14</sup>C] 3-dehydro-UDCA. The product derived from reaction with [24-<sup>14</sup>C] 3-dehydro-CDCA was scraped from TLC plates, pooled and subjected to mass-spectrometry (Figure 19). The major mass ion was 2 amu less than the 3-dehydro-CDCA substrate and thus consistent with the loss of 2 hydrogens from the substrate bile acid and potentially a 3-dehydro- $\Delta^4$ -CDCA structure. Figure 20 illustrates the reactions catalyzed by the *baiCD* and *baiH* gene products.

Figure 18. SDS-PAGE of cell-extract of recombinant *E. coli* BL21(DE3) overexpressing *baiCD* gene. (M) Marker (UI) uninduced, (I) induced 1 mM IPTG. 30 µg protein loaded and gel stained with Coomassie brilliant blue R-250. A large band at ~ 70 kDa was detected in IPTG induced cell-extracts indicating overexpression of *baiCD* gene.

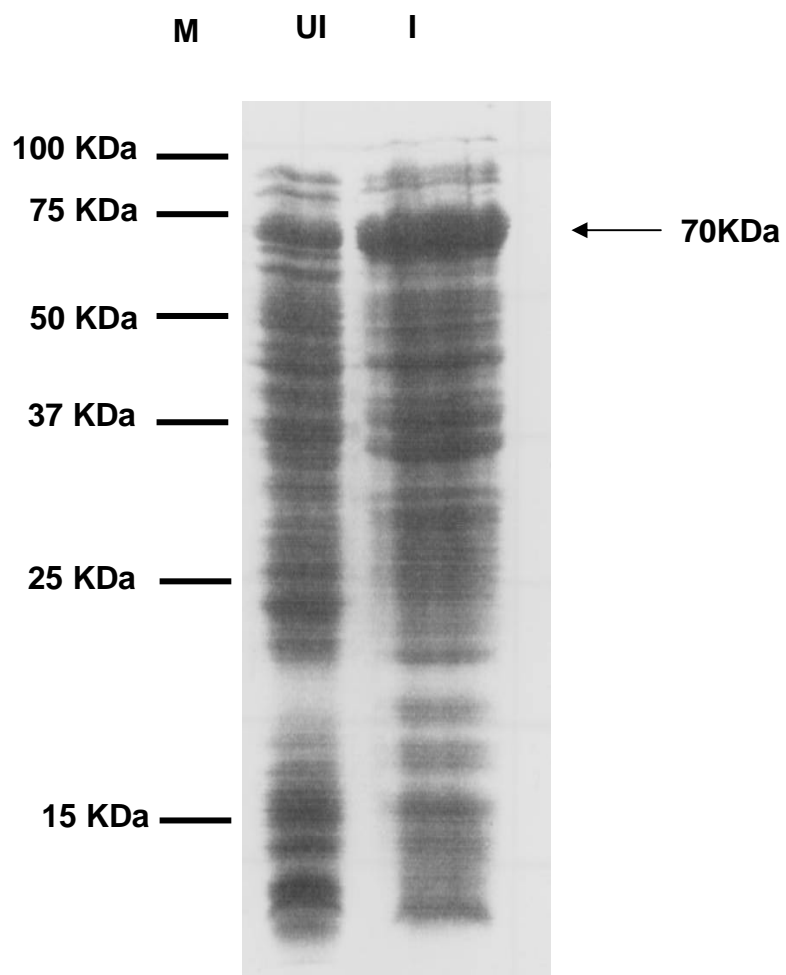


Figure 19. HPLC and mass spectrometry analysis of BaiCD catalyzed bile acid reaction products. The bile acid metabolites produced by cell-extract of *E. coli* overexpressing *baiCD* gene were separated by TLC, scraped and further purified by reverse-phase HPLC-ESI-MS. The initial mobile phase was 10 mM NH<sub>4</sub>OAc; after injection bound materials were eluted by a 0-100 % linear gradient of acetonitrile in 10 mM NH<sub>4</sub>OAc at 10 % min<sup>-1</sup>. The eluate was passed into the electrospray interface operating in the negative ion mode. Ions eluting from m/z 350-700 were recorded and analyzed. The upper trace is a selected ion chromatogram for m/z 389. The mass spectra for the peak at 11.59 min were summed and presented in the bottom trace. The m/z 389.3 ion corresponds to the expected [24-<sup>14</sup>C] 3-dehydro-4-CDCA product.

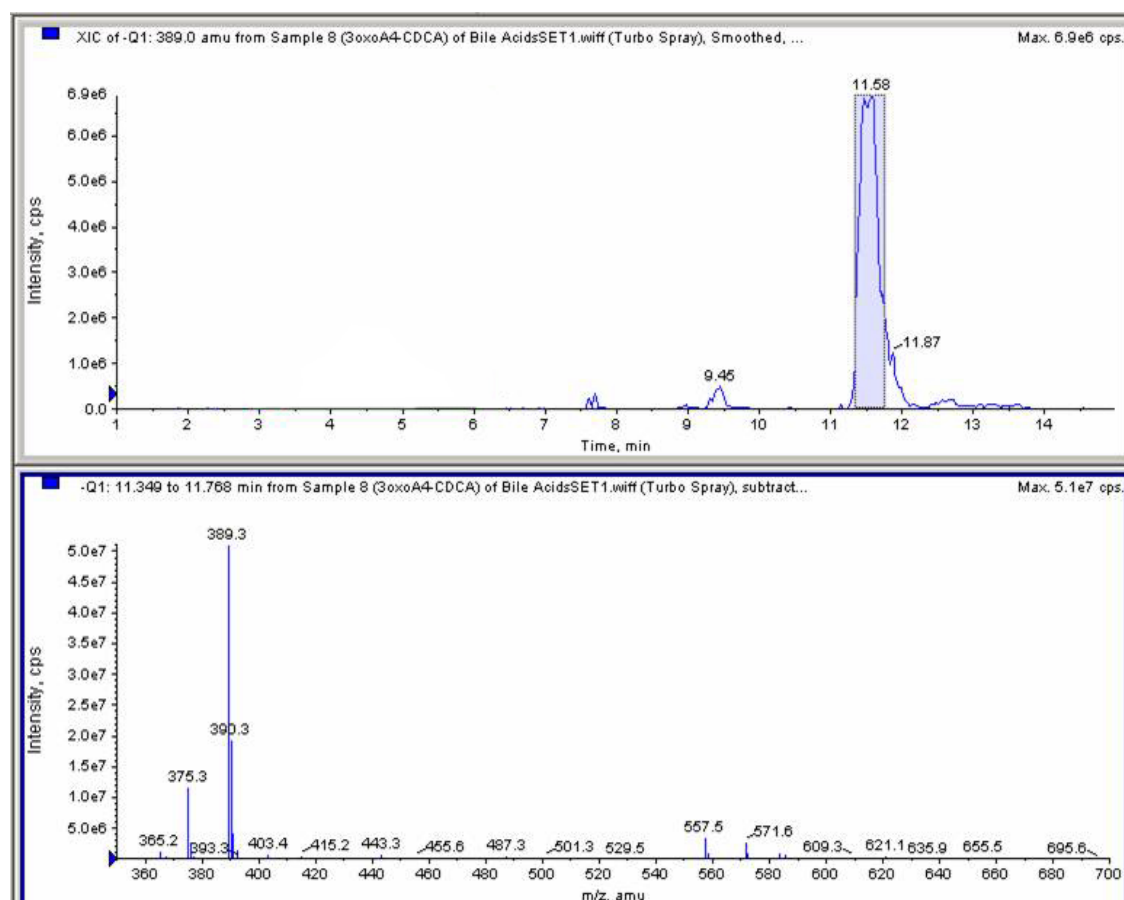
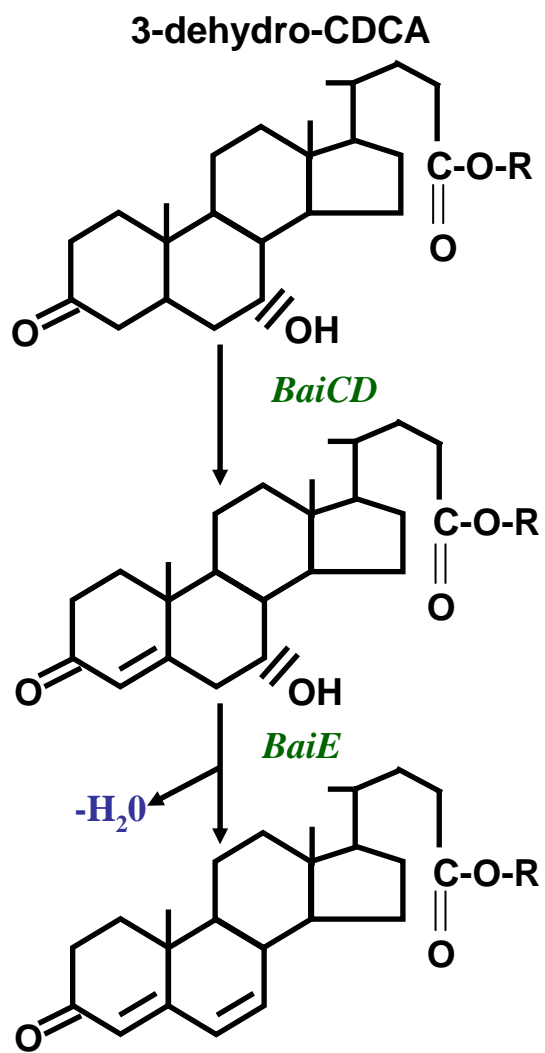
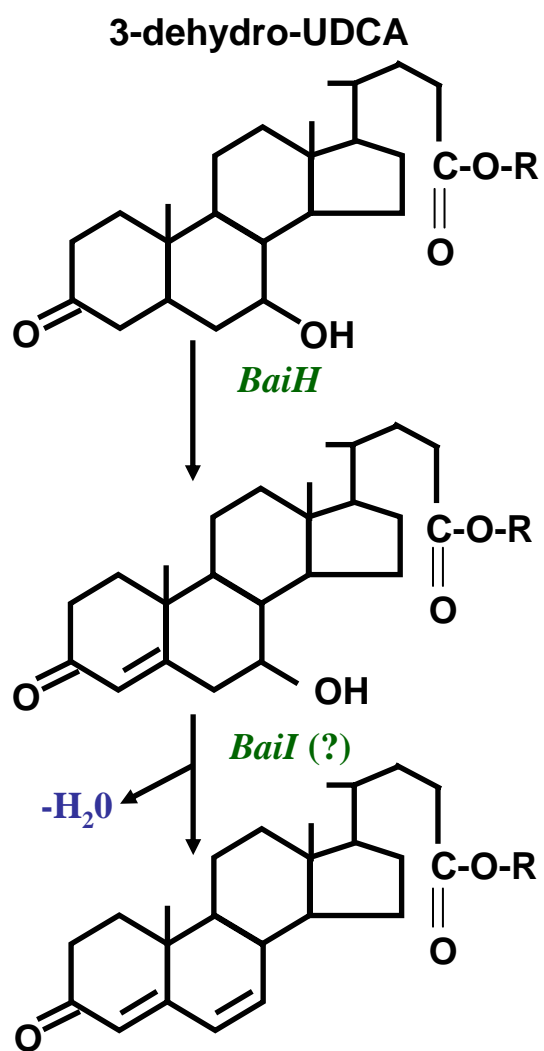




Figure 20. Proposed reactions catalyzed by the *baiCD* and *baiH* gene product. R group may be the carboxylic acid or the coenzyme A conjugate.



## Isolation of *bai* genes from *Clostridium hylemonae*

### TN271

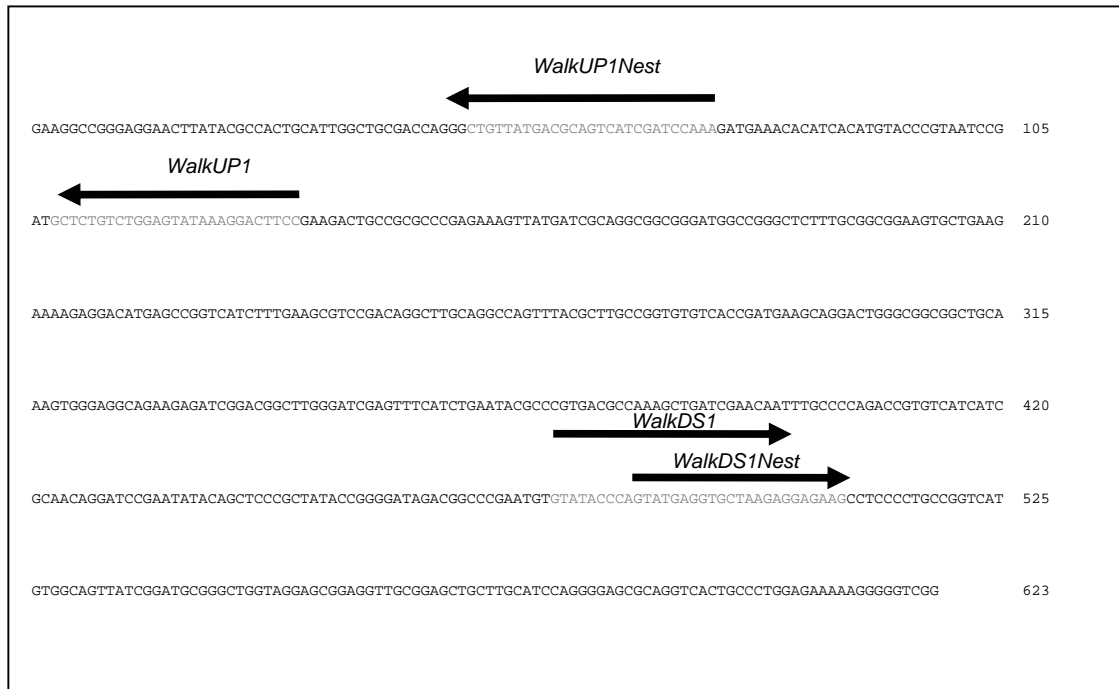
#### Identification and Characterization of the *baiBCDEFHGI* operon

Apart from basic biochemical characteristics and 16S rDNA sequence, little is known about *C. hylemonae*. Previously, our laboratory performed activity assays on cultured 7 $\alpha$ -dehydroxylating bacteria and found that these organisms clustered into two groups based on relative rate of conversion of [24-<sup>14</sup>C] CA to [24-<sup>14</sup>C] DCA (189). Much effort has been focused on the "high activity strains" *C. scindens* and *C. hiranonas*, however, organisms such as *C. hylemonae*, described as a "low activity strain" has received little characterization (See Intro). Recent data indicates that *C. hylemonae* may play a significant role in bile acid metabolism in the human colon based on *in vivo* studies where this species was found to be a numerically predominant 7 $\alpha$ -dehydroxylating bacterium (190). Therefore, characterization of the bile acid 7-dehydroxylating pathway in this organism is important. Genes involved in the bile acid 7 $\alpha$ -dehydroxylating pathway

have yet to be isolated and sequenced from any "low activity strain". We were thus interested in elucidating the gene organization, gene and promoter sequence, as well as upstream regulatory genes and comparing these elements to "high activity strains" to see whether there are obvious differences that might underly the ten fold differences in relative rate of bile acid 7 $\alpha$ -dehydroxylation (189).

Wells and Hylemon (2000) designed redundant PCR primers targeting the *baiCD* gene which are aimed at detecting 7 $\alpha$ -dehydroxylating bacteria in the human colon (59). Wells and Hylemon (2000) reported detection of the *baiCD* gene of *C. hylemonae* TN271 using these oligonucleotides. Therefore, these primers were used to amplify a 1.2 kB segment of the *baiCD* gene from *C. hylemonae* which was subsequently cloned into a pCR8 GW TOPO TA vector and sequenced (Figure 21). BLAST analysis of the sequence data and CLUSTALW alignment of the nucleotide and translated amino acid sequence verified the identity of a *baiCD* gene in *C. hylemonae*. We hypothesized that the *bai* genes involved in the oxidative arm of the bile acid 7 $\alpha$ -dehydroxylation pathway (*baiA* - *baiI*) would

Figure 21. Region of partial *baiCD* gene from *C. hylemonae* TN271 used to design upstream and downstream genome-walking PCR primers. Sequence represents sense-strand of *baiCD* gene, arrows denote direction of amplification and grey font represents regions of primer design.



be located on a polycistronic operon flanking the *baiCD* gene. Therefore, a bidirectional genome-walking by PCR strategy was used to isolate the remaining *bai* genes.

A slight modification of the method of Ridlon et al. (2005) was utilized to prepare genomic DNA (>50 kb) from both *C. scindens* VPI 12708 and *C. hylemonae* TN271 for genome-walking by PCR (see Materials and Methods) (Figure 22) (177). Restriction libraries derived from this DNA sample yielded PCR products averaging 2 kb with a range of 0.7 kb to 4 kb. The limit of this technique, according to the manufacturer (Clontech GenomeWalker Universal Kit) is around 6 kb. Upstream genome-walking PCR using primer WALKUP1 and subsequent nested amplification using primer WALKUP1Nest resulted in 2,139 bp of sequence data which overlapped the *baiCD* gene data by 673 bp resulting in a 2,690 bp contig (Figure 23). Analysis of this data revealed the 5' end of the *baiCD* gene. A second open reading frame (ORF) was detected 19 bp upstream of the *baiCD* gene. BLAST results suggest this ORF encodes an AMP-dependent coenzyme A ligase with highest identity to the *baiB* gene from *C. hiranonis*. Continued upstream genome-walking resulted in a PCR product of 1,964 bp

Figure 22. Preparation of genomic DNA for genome-walking by PCR restriction libraries. Representative 0.4% agarose/EtBr gel showing 1, 2, and 3  $\mu$ g genomic DNA from *C. scindens* VPI 12708. Size comparisons made with Invitrogen's High Molecular Weight DNA ladder. Electrophoresis was performed at 25 volts for 18 hours.



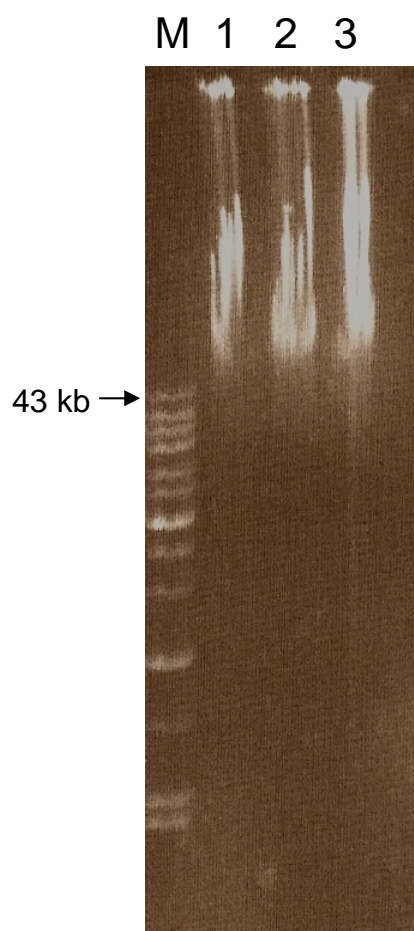
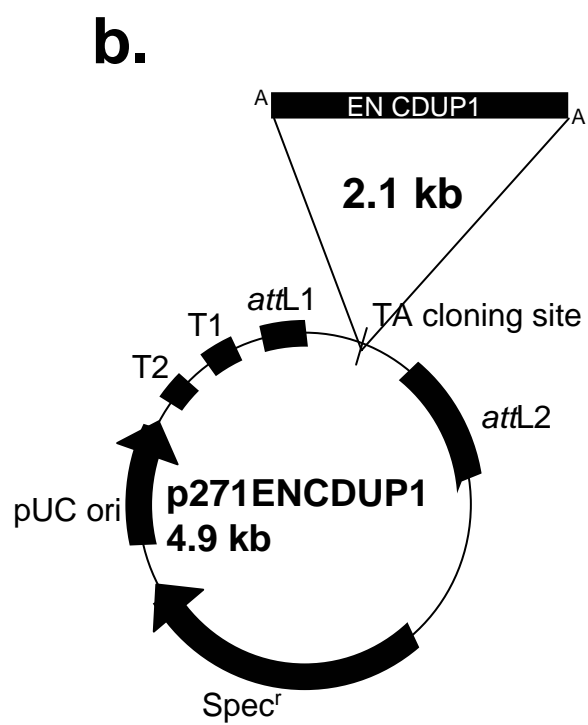
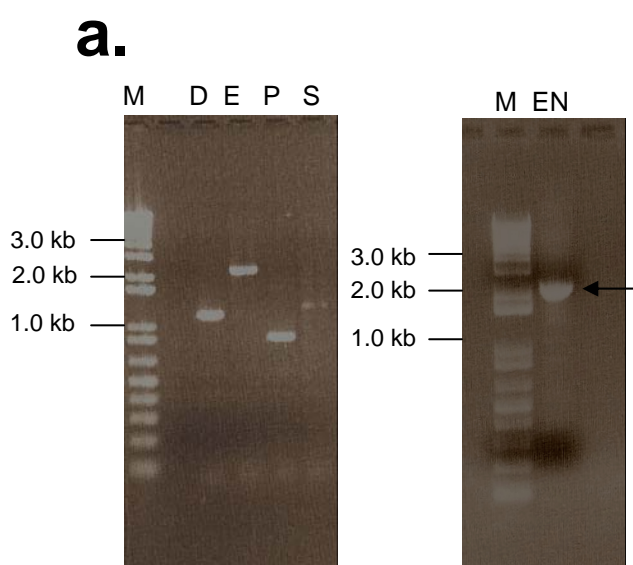


Figure 23. Genome-walking by PCR Upstream of partial *baiCD* gene in *Clostridium hylemonae* TN271. PCR products amplified using WALKUP1 and WALKUP1NEST PCR primers (Table 6). "M" is marker "D", "E", "P", "S" and "EN" represent genome-walking restriction library templates digested with *DraI*, *EcoRV*, *PvuII*. and *StuI*, respectively. "EN" represents nested PCR amplification of *EcoRV* primary PCR product diluted 1:50. The PCR product generated from nested amplification was purified by GENECLAN and ligated into a pCR8 GW TA TOPO cloning vector (2.8 kb) to yield vector p271CDUP1 (4.9 kb).



(Figure 24). Assembly of this data resulted in a 4,586 bp contig containing the complete *baiB* gene. The 1.5 kb *baiB* gene is predicted to encode a 57.7 kDa polypeptide similar in size to the reported 58 kDa subunit recombinant BaiB polypeptide from *C. scindens* VPI 12708 (125). Two additional interesting features were identified within this sequence data. First, a conserved bile acid inducible promoter was located on the sense strand 49 bp upstream of the *baiB* gene (Figure 25) (59, 191). SMART-RACE PCR analysis of the cDNA from a CA induced culture of *C. hylemonae* led to the identification of the mRNA start site which appears in a short purine rich stretch corresponding to the ribosome binding site (Figure 26). Interestingly, the distance between the mRNA start site in the *baiB* polycistronic operon in *C. hylemonae* is only 11 bp from the initiation codon while in the "high activity strains" there is a distance of 108 bp in *C. hiranonis* (59) and 68 bp in *C. scindens* (191). As proof of principle, SMART RACE PCR was also performed on the *baiB* gene of *C. scindens* VPI 12708, whose mRNA initiation site was previously determined by primer extension (191). The results obtained in the present study agree with those of

Figure 24. Second upstream genome-walking by PCR from partial *baicD* gene in *Clostridium hylemonae* TN271. a.) PCR products amplified using WALKUP2 and WALKUP2NEST PCR primers (Table 6). Nested PCR product (DN=DraI template, nested primer amplification product) purified from 1% agarose gel by GENECLEAN Spin Kit and ligated into pCR8 GW TOPO TA cloning vector to make vector p271CDUP2 (b.).

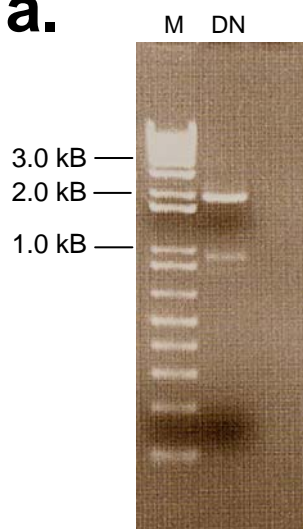
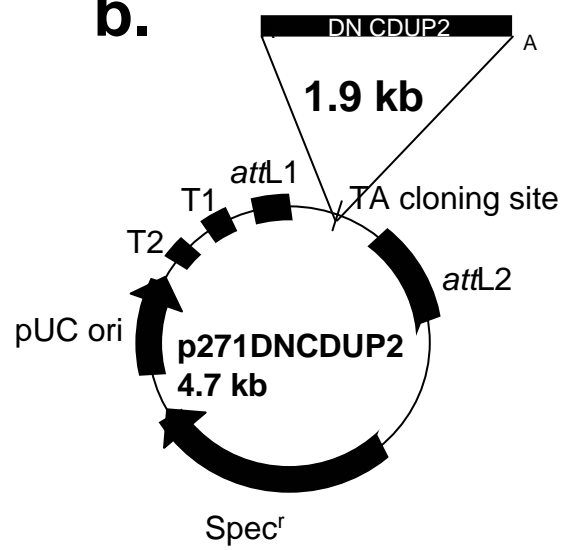
**a.****b.**

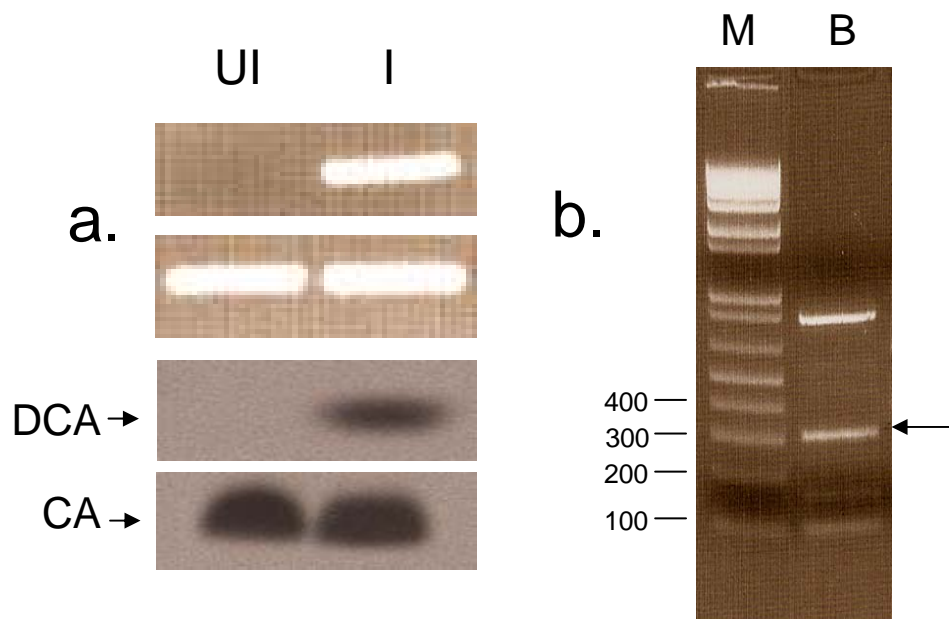
Figure 25. Boxshade of CLUSTALW alignment of *bai* promoters upstream of *baiB* gene between *C. hylemonae* TN271, *C. scindens* VPI 12708 and *C. hiranonis* sp. strain T0931. Putative -10 element is indicated below alignment. Transcription initiation sites (TIS) highlighted in yellow (See Fig 26 for TIS data for *C. hylemonae*).

TO931baiB	ATATTGTACCGGCAAAAAATGTTAAGATATATCAAGAT
12708baiB	ATATTGTATCTCAAAAAATGTTAAAGTTATATCAAGCA
TN271baiB	ATGTCATAGACTATGTAAGTGTTAAAGTTATATCAAAAG

- 10



Figure 26. SMART RACE PCR analysis to determine TIS for the *baiB* transcript in *Clostridium hylemonae* TN271. (a.) RT-PCR analysis of *C. hylemonae* TN271 cDNA generated from CA induced (I) and uninduced (UI) cultures. Primers against *baiB* gene (above) and *gyrA* gene (below) were used to determine whether *baiB* transcript was expressed. Below RT-PCR is the corresponding [24-<sup>14</sup>C] CA activity assay performed prior to RNA isolation of I and UI cultures of *C. hylemonae*. [24-<sup>14</sup>C] DCA is detected only in CA induced (50  $\mu$ M) culture. (b.) 5' SMART RACE PCR of *baiB* transcript using RACEB325 oligonucleotide against cDNA from induced culture of *C. hylemonae* TN271. PCR product was purified by GENECLEAN Spin Kit and cloned into a pCR8 GW TOPO TA cloning vector and sequenced. (c.) TIS for *baiB* transcript (underlined; direction of transcription designated by arrow). "\*" designates conserved *bai* promoter elements (See Figure 25), putative ribosome binding site in bold, "+" designates mRNA start site determined for *baiB* in *C. scindens*, "‡" designates mRNA start site for *baiB* in *C. hiranonis* sp. strain T0931, italics indicates anticodons with corresponding amino acids (below).



**c.**

†‡  
 ATATTATCCAATGTCATAGAGTATGTAAGTGTTAAAGTTATATCAAAGGAGGGCGAAATATGGACTTG  
 \*\* \* \*\*                      \* \* \* \* \*                      \* \* \* \* \*                      M D L

Mallonee et al. (1990) providing proof of principle and additional evidence for the TIS of the *baiB* operon in *C. scindens* VPI 12708 (Figure 27) (191).

The second notable feature is a conserved 47 kDa AraC/XylS transcription factor located on the anti-sense strand upstream of the *bai* promoter. Figure 28 illustrates the conserved helix-turn-helix motifs identified in the amino acid sequence of the "bile acid regulatory A" gene (*barA*). Additional genome-walking upstream of the *barA* did not result in ORFs predicted to be involved in bile acid metabolism.

Downstream genome-walking from the *baiCD* fragment yielded a PCR product of 1,131 bp which increased contig length to 5,625 bp (Figure 29). This stretch of sequence resulted in completion of the 1.9 kb *baiCD* gene predicted to encode a 70 kDa polypeptide. In the first section, overexpression of the recombinant *baiCD* gene from *C. scindens* VPI 12708 in *E. coli* resulted in a 70 kDa polypeptide as estimated by SDS-PAGE (Figure 18). In addition, 46 bp downstream of the *baiCD* gene we located a 510 bp ORF predicted to encode the 19.8 kDa bile acid 7 $\alpha$ -dehydratase (*baiE* gene). Dawson et al (1996) reported

Figure 27. Determination of transcription initiation site for *baiB* gene from *Clostridium scindens* VPI 12708 based on SMART RACE PCR analysis. Alignment shows determination of transcription initiation site (TIS) (arrow, underline) based on alignment of SMART RACE sequence data (SR7, SR8) with *baiB* gene and upstream region from *Clostridium scindens* VPI 12708 (12708*baiB*). TIS is determined based on the first base following 3' end of SMART II A oligo sequence (boxed region) (5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3'). This nucleotide (underlined, arrow) must also correspond to the known sequence of the upstream region of the gene. The TIS determined through SMART RACE analysis agreed with primer extension analysis performed by Mallonee et al (1990) (191).

↓

SRB 7	GGTATNA-CGCAGAGTACGCGGGAGC	AAAAGAAGAAAATGTTGAAAAAATGTACAAATA	98
SRB 8	GGTATCAACGCAGAGTACGCGGGAGC	AAAAGAAGAAAATGTTGAAAAAATGTACAAATA	120
12708b <i>a</i> iB	---AAAATGTTAAAGTTATATCA	AGC AAAAGAAGAAAATGTTGAAAAAATGTACAAATA	56
	* * *	*****	
SRB 7	GCACGAAGAAAAGATATTAAAGCATTAA	AGAAATGCACAAAAAATCAGCGTGTGAGAGGGA	158
SRB 8	GCACGAAGAAAAGATATTAAAGCATTAA	AGAAATGCACAAAAAATCAGCGTGTGAGAGGGA	180
12708b <i>a</i> iB	GCACGAAGAAAAGATATTAAAGCATTAA	AGAAATGCACAAAAAATCAGCGTGTGAGAGGGA	116
	*****		

M

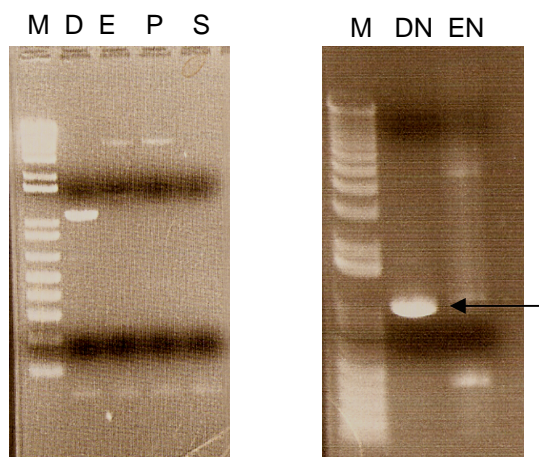
Figure 28: ClustalW alignment of bile acid regulatory A gene (*barA*) between *C. hylemonae* TN271, *C. scindens* VPI 12708 and *C. hiranonis* sp. strain T0931. Regions highlighted in red are predicted helix-turn-helix motifs shared between members of the Arac/XylS family of transcription factors.



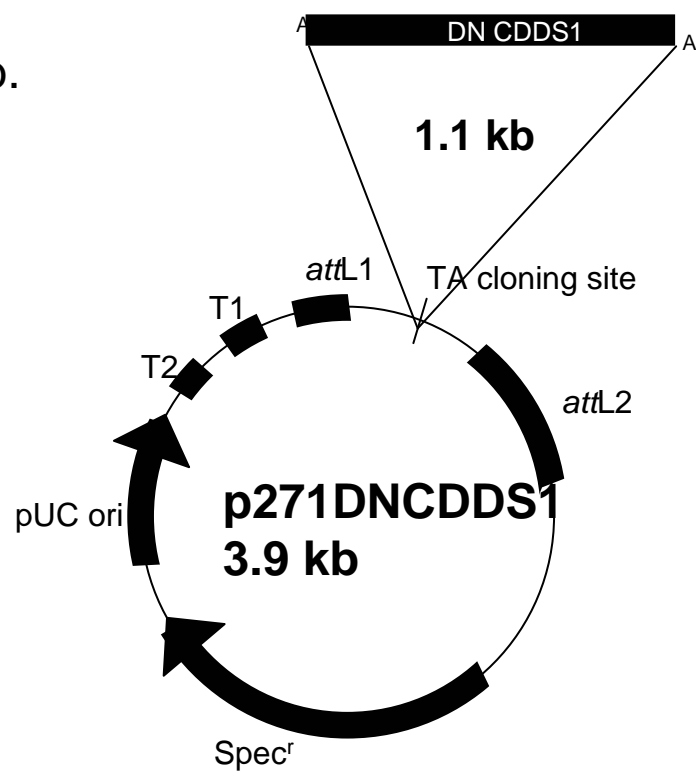
Figure 29. Downstream genome-walk (DS1) from *baiCD* gene from *Clostridium hylemonae* TN271. a.) Agarose gel of DS1 products using *DraI* (D), *EcoRV* (E), *PvuII* (P), *StuI* (S) restriction libraries of *C. hylemonae* TN271 genomic DNA as template with adaptor primer AP1 and gene-specific primer WALKDS-1 (See Table 6). b.) "DN" and "EN" represent nested PCR amplification using nested adaptor primer AP2 and gene-specific primer WALKDSNEST1 (See Table 6). c.) Schematic representation of TA cloning of GENE CLEAN purified "DN" product (arrow) into pCR8/GW/TOPO TA vector (2.8 kb) yielding 3.9 kb vector p271DNCDDS1.



a.



b.



overexpression of the 19 kDa *baiE* gene from *C. scindens* VPI 12708 (128). Multiple sequence alignment of this ORF with the *baiE* genes from *C. scindens* VPI 12708 and *C. hiranonis* sp. strain T0931 revealed >80% amino acid sequence identity and conservation of active site amino acids predicted through computer modeling and site-directed mutagenesis data (Figure 30) (Hylemon et al unpublished data). Continued downstream GW-PCR resulted in a 3,856 bp PCR product which when added to the previous alignment results in a contig of 9,291 bp (Figure 31). An ORF was located 25 bp downstream of the *baiE* gene 1,275 bp in length predicted to encode a 47 kDa polypeptide in the Type III CoA transferase family of proteins. BLAST analysis suggests this gene encodes the *baiF* gene product. Ye et al (1999) previously determined the  $M_r$  of the recombinant *baiF* from *C. scindens* VPI 12708 to be 47.5 kDa (130). A second ORF, 1.4 kb in length, was located 53 bp downstream of the *baiF* gene whose deduced amino acid sequence suggests this gene encodes a 49.5 kDa AraJ-like efflux permease. BLAST analysis suggests this ORF encodes the *baiG* gene product. Mallonee et al (1996) overexpressed the 50 kDa *baiG* gene product in *E. coli* and showed that it

Figure 30. Boxshade of CLUSTALW alignment of 7 $\alpha$ -dehydratase gene (*baiE*) between *C. hylemonae* TN271, *C. scindens* VPI 12708, and *C. hiranonis* sp. strain T0931. Asterisks denote conserved active site amino acids predicted by computer modeling and site-directed mutagenesis (Hylemon et al. unpublished data).

```

                                *   *
12708baiE  MTLLEERVEALEKELECEMKDIEAIKELKGYFRCLDGKMWDELETTLSPNIVTSYSNGKLV  61
T0931baiE  MTLLEARIEALEKEIQRLNDIEAIKQLKIKYFRCLDGKLUDELETTLSPNIEQTSYSDGKLV  61
TN271baiE  MSIEERLEALEKEIQKMKDIEEIKKLKGYFRCLDGKFWDELETTLSPNIVTSYSNGKLV  61

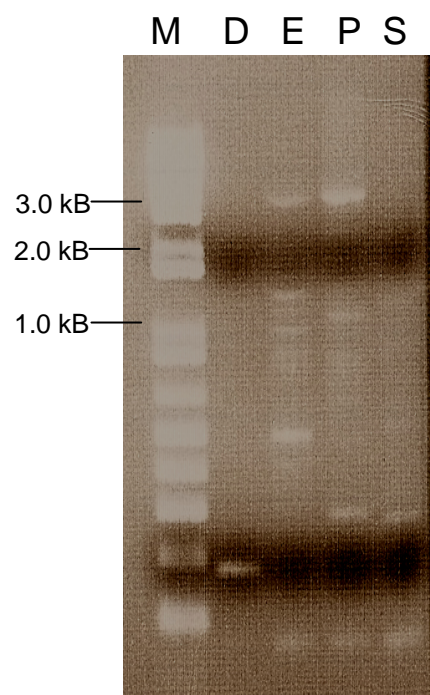
                                *           *
12708baiE  FHSPKEVTDYLKSSMPKEEISMHHMGHTPEITIDSETTATGRWYLEDRLIFTDG-KYRDVG  119
T0931baiE  FHSPKEVTIYLAAAMPKEEISMHHMGHTPEITIDSENTATGRWYLEDNLIFTDG-KYRNWG  119
TN271baiE  FHGPKEVTDYFKKAMPKEEISMHHMGHTPEITIDSETTATGRWYLEDKLIFTESKYAGSG  120

                                *
12708baiE  INGGAFTDKYEKIDGQWYILETGYVRIYEEHFMRDPKIHITMNMHK--  166
T0931baiE  INGGAFTDKYEKIDGQWYIKETGYVRIYEEHFMRDPKIHITSNMHKEK  168
TN271baiE  VNGGAFTDKYEKVDGRWYILETGYLRVYEEHFMRDPKIKITMNMHKK  169

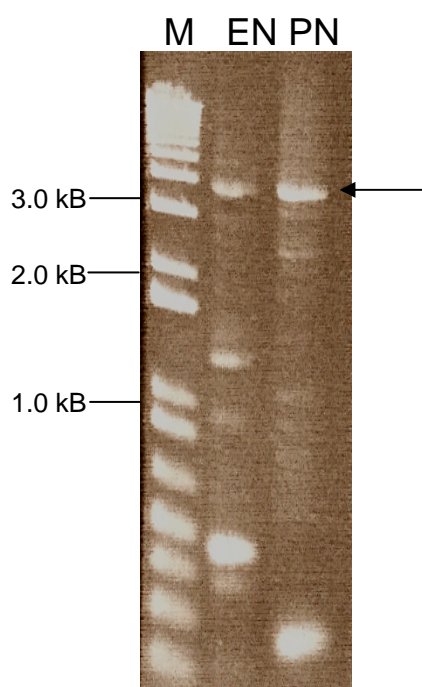
```

Figure 31. Second downstream genome-walk (DS2) from *baicD* gene from *Clostridium hylemonae* TN271. a.) Agarose gel of DS2 products using *DraI* (D), *EcoRV* (E), *PvuII* (P), *StuI* (S) restriction libraries of *C. hylemonae* TN271 genomic DNA as template with adaptor primer AP1 and gene-specific primer WALKDS-2 (See Table 6). b.) "EN" and "PN" represent nested PCR amplification using nested adaptor primer AP2 and gene-specific primer WALKDSNEST2 (See Table 6). c.) Schematic representation of TA cloning of GENE CLEAN purified "PN" product (arrow) into pCR8/GW/TOPO TA vector (2.8 kb) yielding 6.6 kb vector p271PNCDDS2.

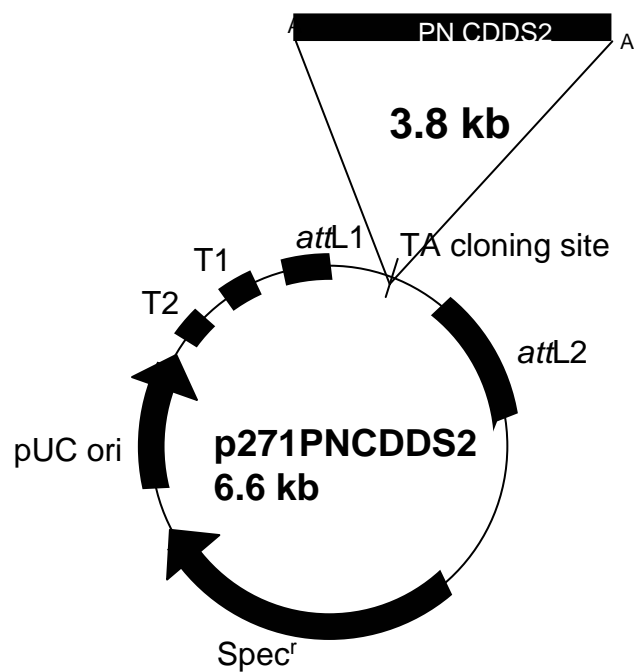
a.



b.



c.



is involved in transporting primary bile acids across the bacterial membrane (126). Further analysis revealed a partial ORF 19 bp downstream of the *baiG* gene encoding an "Old Yellow Enzyme" like FMN oxidoreductase with the *baiH* gene from *C. hiranonis* sharing the highest amino acid identity. The third downstream GW-PCR resulted in completion of the *bai* oxidative operon (Figure 32). The 2,019 bp *baiH* gene was completed and the deduced amino acid sequence suggested this gene encodes a 71.7 kDa polypeptide. The *baiH* gene from *C. scindens* VPI 12708 has been shown to encode a 72 kDa polypeptide (Figure 5) (121, 129). Finally, the 549 bp *baiI* gene was located 38 bp downstream of the *baiH* gene and predicted to encode a 21.8 kDa polypeptide. The complete *bai* oxidative operon from *C. hylemonae* is represented schematically in Figure 33 and compared to the gene organization in *C. scindens* and *C. hiranonis*. Table 6 lists oligonucleotides used in isolation and sequencing of the *baiBCDEFGHI* operon in *C. hylemonae*. Table 7 compares amino acid sequence identity/similarity between *C. hylemonae* and *C. scindens*. Clearly, there is a high degree of conservation both between the genes and several of the intergenic regions;

Figure 32. Third downstream genome-walk (DS3) from the *baiCD* gene of *Clostridium hylemonae* TN271. a.) Agarose gel of DS3 products using *DraI* (D), *EcoRV* (E), *PvuII* (P), *StuI* (S) restriction libraries of *C. hylemonae* TN271 genomic DNA as template with adaptor primer AP1 and gene-specific primer WALKDS-2 (See Table 6). b.) "DN" represents nested PCR amplification using nested adaptor primer AP2 and gene-specific primer WALKDSNEST2 (See Table 6). c.) Schematic representation of TA cloning of GENECLAN purified "DN" product (arrow) into pCR8/GW/TOPO TA vector (2.8 kb) yielding 5.9 kb vector p271DNCDDS3.



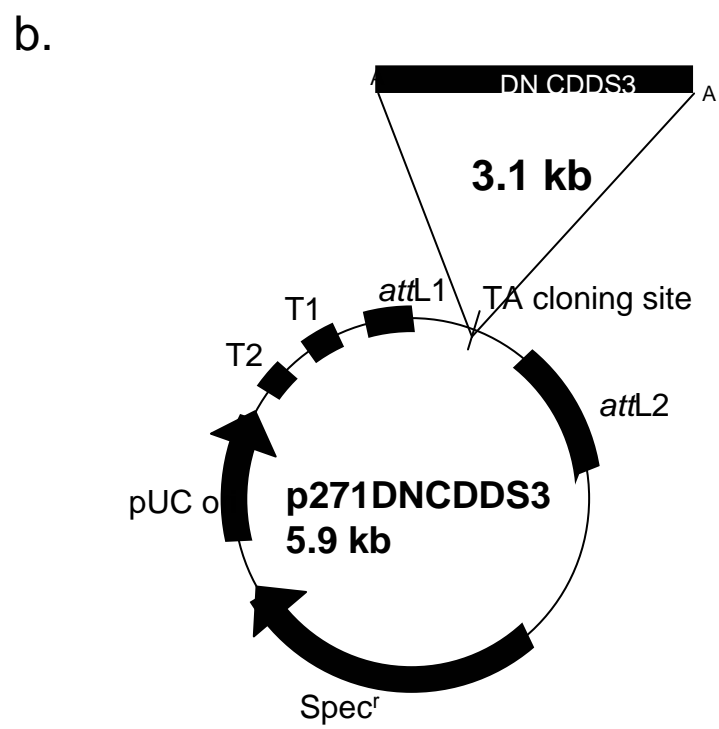
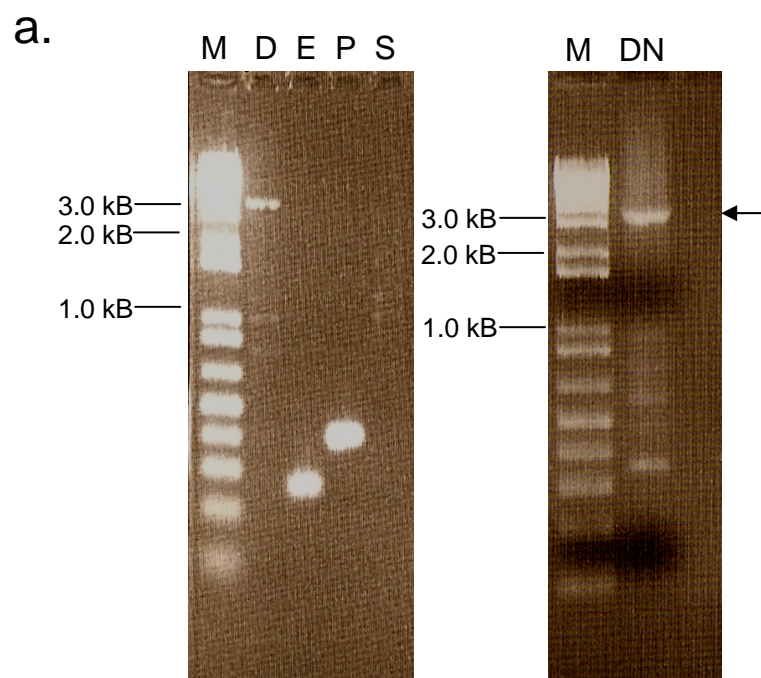
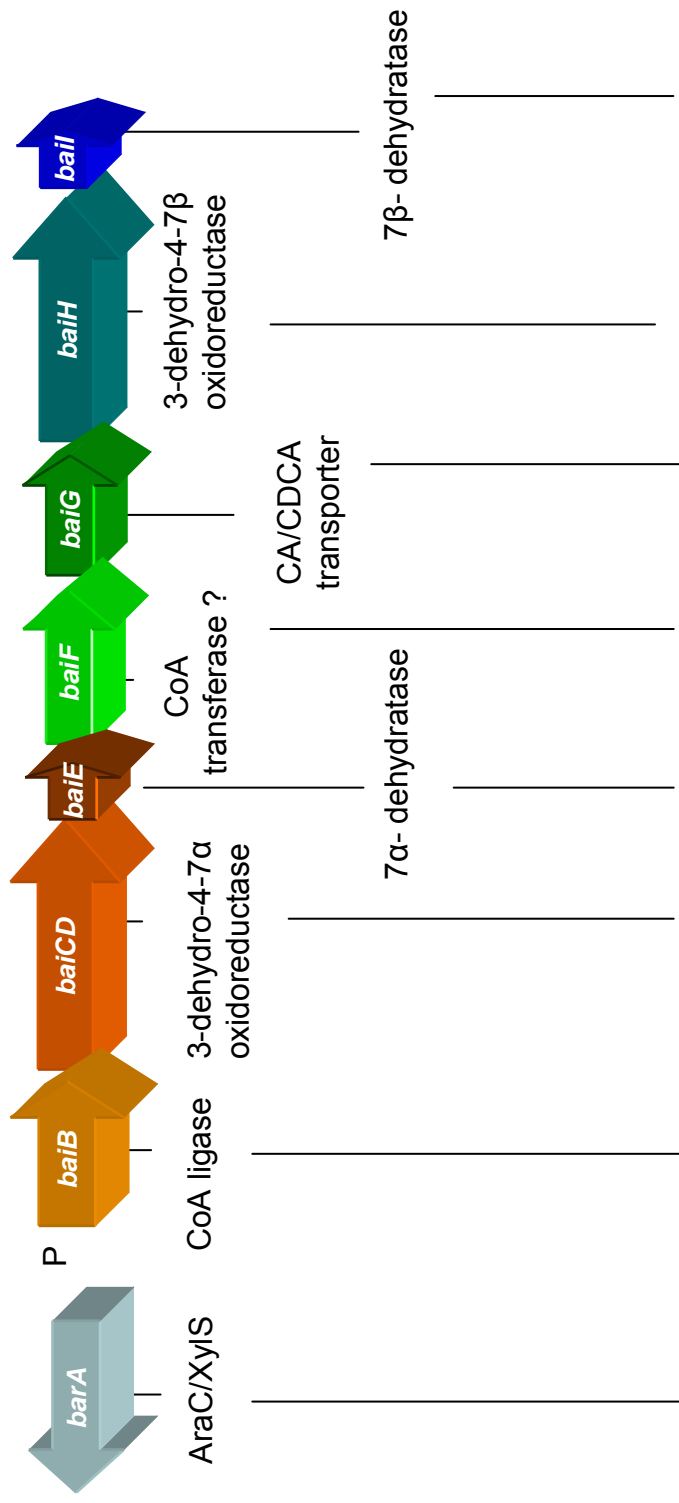
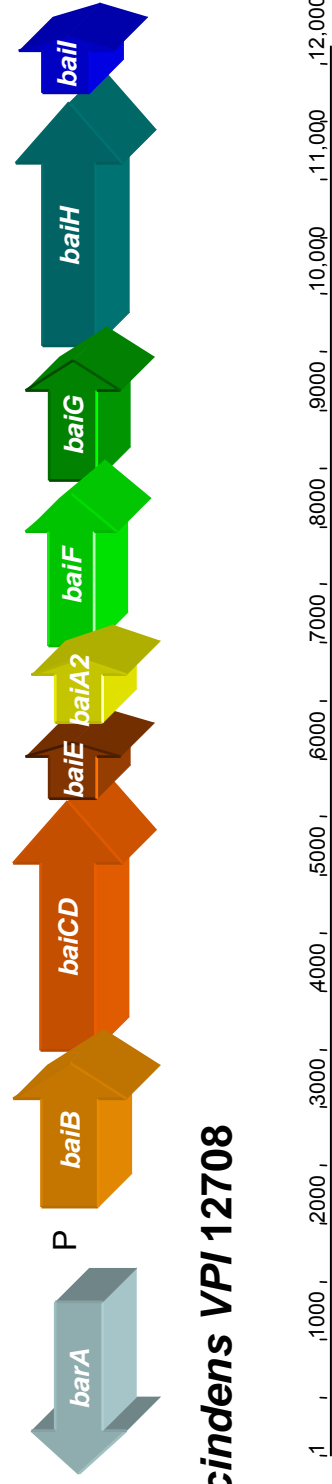


Figure 33. Schematic representation and comparison of the *bai* oxidative operons of *C. hylemonae* TN721 and *C. scindens* VPI 12708. See Table 7 for comparison of amino acid sequence identity and other parameters. Notice the high degree of conservation in gene organization, with the exception of the *baiA* gene, present in both *C. scindens* and *C. hiranonis* though conspicuously missing from the *bai* oxidative operon of *C. hylemonae* TN271.

## *C. hylemonae* TN271



## *C. scindens* VPI 12708



**Table 6. Oligonucleotides Used in Isolating *baiBCDEFGHI* Operon from *Clostridium hylemonae* TN271**

Primer name	Sequence (5'-3')	Reference
Redundant oligonucleotides		
ba1CD-F	GGATTGAGGCTTCAGATGTTCTTG	0e11s
ba1CD-R	GAATTTTGGGTTTATGAACATTTCTCAAG	0e11s
Genome-walking oligonucleotides		
API		
AP2	GTATATACGACTCATATATAGGGC	Clontech
WALKUP-1	ACTATAGGCGACGCTGCT	Clontech
WALKUPDOWN-1	GAATCTCTTTATATCTTCAGACAGAGC	This study
WALKUP2	TTTGATTCGATGACGCTGCTATATACAC	This study
WALKUP2Me st	AATCAAGAGCTGTCTCTCTCAAGATTC	This study
WALKD3-1	GTATATCTCTCAATGAGGCTGATGAC	This study
WALKDOWNest-1	CTATATGAGGCTGTATAGAGGAGAGGC	This study
WALKD3-2	TATCAAGAGCAATTTATCTGAC	This study
WALKDOWNEST2	CAGAGATCAAGATCACTATGATATCTG	This study
WALKD33	GAATTTCTGGAGGATATGCTTATCTG	This study
WALKD33NE3T	CTTAACTTTGAGAGCGGAGGCTG	This study
Primer-walking oligonucleotides		
271-100-F	ACATCGGGTATGTAAGACAGAGAGC	This study
271-001-F	GTTCTGGATGAACTGGAGACTACATCTTC	This study
271-100-R	GCTATGACTTCAGAGCGAGAACTGAGC	This study
CGAP1	CTCTCTCTCTCTATGTTCTC	This study
EGAP1	CTCTACACGCTTGAAGAGACTGG	This study
EGAPUP	GTTTCATCTGATCTTCGATAC	This study
EGAPF1	GTTCTCACTATCTGAACTC	This study
EGAP1	TTTACCATCTGCTGAC	This study
EGAP1	TGATGAGACGAGACACTGG	This study
271ba1F	ACATCTCATCTGAAAGGAGAC	This study
271ba1R	AGATAGAGGCTATCTCTCAAG	This study
271ba1HR2	CTTAACTGTGACGAGTC	This study
EGAP1	ACTTTCTTTCTCTCTG	This study
271ba1F2	GATATATAGACGAGAGATGACG	This study
ED32271ba1F2	ACTGGATCTCGAGAGAGAC	This study
ED32271ba1HR	GTTAACTGTAACTGCTATG	This study
ED32HGAP2	GTTCTGTAGGATCAAGAGAC	This study
ED32HR2F	TCAAGAGCAGAGAGAGATGAC	This study
ED32HR2R	CTTATGAGAGTCTGCTCTC	This study
ED32F2R	CTATCATCTTTCTATGCTGAG	This study
ED32F2F	GTTCTATGAGGATATCTGAGAC	This study
D32BA1HF2	CATATATGTACCTGACACAC	This study
D32CLUSYMF1	CAGAAATATATGATCTCTCTG	This study
D32CLUSYMR2	AGAAATCTCTGAGGAGCTAGC	This study
ba1HTR271	ACCTGCTGAGATGTAGTCTATGACG	This study
ba1IR	ACTTCTCTCTTATATATCATAGC	This study

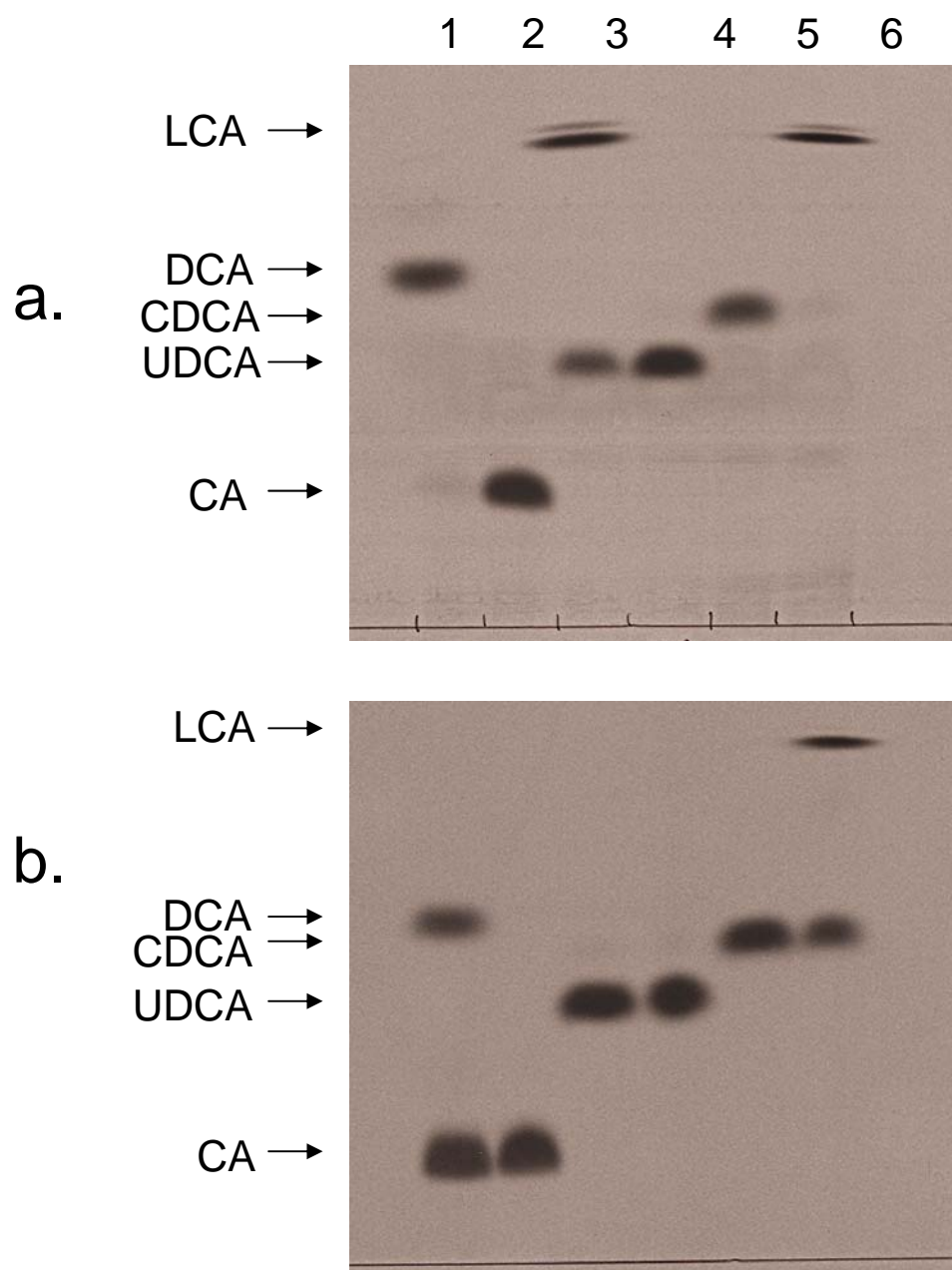
Table 7. Comparison of *Clostridium hylemonae* TN271 and *Clostridium scindens* VPI 12708 *bai* operon DNA sequences

Gene or region	<i>C. scindens</i> vs. <i>C. hylemonae</i>	Length (bp)	Identity %	Enzyme Function	Identity %	Similarity %
<i>baiB</i>		1 563 vs. 1 545	65.7	bile acid-CoA ligase	64.4	78.9
<i>baiB-CD</i> (intergene)		45 vs. 44	52			
<i>baiCD</i>		1 920 vs. 1 923	75	3-dehydro-4-steroid oxidoreductase	80.6	89.2
<i>baiCD-baiE</i> (intergene)		43 vs. 46	51.1			
<i>baiE</i>		501 vs. 510	76.6	Bile acid 7alpha-dehydratase	82.8	92.3
<i>baiA</i>		747 vs. 750	76.2	3 alpha-hydroxysteroid dehydrogenase	85.5	92.4
<i>baiF</i>		1 281 vs. 1 275	78.7	Bile acid-CoA hydrolase/transferase?	88	94.8
<i>baiF-baiG</i> (intergene)		5 vs. 25	5			
<i>baiG</i>		1 434 vs. 1 428	71.1	Bile acid transporter	80.3	88.9
<i>baiG-baiH</i> (intergene)		28 vs. 19	55.6			
<i>baiH</i>		1 986 vs. 2 019	74.2	3-dehydro-4-steroid oxidoreductase	80.8	90.0
<i>baiH-baiI</i> (intergene)		17 vs. 38	41.9			
<i>baiI</i>		555 vs. 549	68.2	Bile acid 7beta-dehydratase?	61.4	76.6

not surprising given the close phylogenetic relation between these species as determined by neighbor-joining of 16s rDNA sequences of 7 $\alpha$ -dehydroxylating bacteria and other gut microbes (106). One of the most striking observations is the fact that the *baiA* gene is not found within this operon in *C. hylemonae* as it is in *C. scindens* and *C. hiranonis*. We predict that the *baiA* gene will be found as a single cistron with a *bai* promoter region as observed in *C. scindens* with the *baiA1* and *baiA3* genes (67). Taken together, these data indicate that genes located on the 9.5 kb *bai* oxidative operon of the "low activity strain" *C. hylemonae* TN271 are highly conserved both in organization as well as amino acid sequence with the homologous *bai* polycistronic operons identified in the "high activity" species *C. scindens* and *C. hiranonis* with the exception of the *baiA* gene.

The ability of *C. hylemonae* TN271 to 7 $\beta$ -dehydroxylate UDCA has not been reported in the literature. Therefore, *C. hylemonae* and *C. scindens* were induced with 50  $\mu$ M CA for 2 hrs and tested for the ability to 7 $\alpha$ -dehydroxylate [24-<sup>14</sup>C] CA and [24-<sup>14</sup>C] CDCA as well as 7 $\beta$ -dehydroxylate [24-<sup>14</sup>C] UDCA for 30 minutes (Figure 34) at 37°C as well as

Figure 34. Autoradiographs of TLC of 7 $\alpha$ -dehydroxylation and 7 $\beta$ -dehydroxylation activity assay reaction products. (a.) *C. scindens* VPI 12708 induced with 50  $\mu$ M CA during mid-log phase in BHI. 1 ml culture inoculated in 1 ml activity assay consisting of BHI containing 0.1  $\mu$ Ci/2  $\mu$ mole bile acid incubated 37°C for 30 min. Lane 1: [24-<sup>14</sup>C] CA activity assay of CA induced culture of *C. scindens*. Lane 2: [24-<sup>14</sup>C] CA activity assay of uninduced culture of *C. scindens*. Lane 3:[24-<sup>14</sup>C] UDCA activity assay of CA induced culture of *C. scindens*. Lane 4: [24-<sup>14</sup>C] UDCA activity assay of uninduced culture of *C. scindens*. Lane 5:[24-<sup>14</sup>C] chenodeoxycholic activity assay of uninduced activity assay. Lane 6 represents [24-<sup>14</sup>C] CDCA activity assay of cholic acid induced culture of *C. scindens*. The same order applies in (b.) for cultures of *C. hylemonae*. Arrows depict positions of non-radiolabelled bile acid standards. Induced cultures of *C. scindens* and *C. hylemonae* were able to 7 $\alpha$ -dehydroxylate both CA and CDCA (lanes 1 and 6) while only *C. scindens* was able to 7 $\beta$ -dehydroxylate UDCA (lane 3). Uninduced cultures showed no activity against bile acid substrates. Solvent system was S1 according to Eneroth (172).






overnight incubation (Data not shown). Both *C. scindens* and *C. hylemonae* were able to 7 $\alpha$ -dehydroxylate both CA and CDCA resulting in [24-<sup>14</sup>C] DCA and [24-<sup>14</sup>C] LCA, respectively. However, while *C. scindens* was able to 7 $\beta$ -dehydroxylate UDCA, we were unable to detect 7 $\beta$ -dehydroxylation activity by *C. hylemonae* despite sequence data that indicated the presence of genes encoding enzymes suggested to be involved in 7 $\beta$ -dehydroxylation. It is currently unclear why this activity is not present.

#### **Isolation and characterization of a *baiA* gene in *Clostridium hylemonae* TN271**

The 27 kDa *baiA* gene product from *C. scindens* VPI 12708 catalyzes the first oxidative step in the bile acid 7 $\alpha$ -dehydroxylating pathway following ATP-dependent CoA ligation (125). As noted in the previous section, and illustrated in Figure 33, the *baiA* gene was not found in the oxidative polycistronic *bai* operon in *C. hylemonae*. Therefore, redundant primers were designed based on a multiple sequence alignment of *baiA* genes from *C. scindens* and *C. hiranonis* (Figure 35). A PCR product of 722 bp was

Figure 35. Design of degenerate PCR oligonucleotides for isolation of the *baiA* gene in *C. hylemonae* TN271. Clustal alignment of *baiA1*, *baiA2* and *baiA3* genes from *C. scindens* VPI 12708 and the *baiA* gene from *C. hiranonis* sp. strain T0931. Arrows indicate direction of PCR primers, the sequence of sense and antisense primer is shown below alignment.



```


baiA1      ATGAACTTGTACAGGACAAAATTACAATTATCACAGGCGGAACCCGTGGAATCGGATTC 60
baiA3      ATGAACTTGTACAGGACAAAATTACAATTATCACAGGCGGAACCCGTGGAATCGGATTC 60
baiA2      ATGAATCTCGTACAAGACAAAGTTACGATCATCACAGGCGGCACAAGAGGTATTGGATTC 60
baiAT0931  ATGAACTTAGTACAGGACAAAATAGTTATAATAACAGGTGGAACAAGTGGTATAGGTCTT 60
          *****
          * * * * *

baiA1      GCAGCAGCAAACTCTTTATTGAGAATGGAGCAAAAGTCTCCATATTTGGCGAGACCCAG 120
baiA3      GCAGCAGCAAACTCTTTATTGAGAATGGAGCAAAAGTCTCCATATTTGGCGAGACCCAG 120
baiA2      GCCGCTGCCAAAATATTTATCGACAATGGCGCAAAAGTATCCATCTTCGGAGAGACGCAG 120
baiAT0931  TGCGCAGCAAAATATTCATGGATAACGGTGCAACAGTTTCTATATTCGGAAAACTCAG 120
          * * * * *

baiA1      GATCGCGAATGTATATCTGTTTCCTTGCATCCGACCTGGCTAGCGGCATCACGGCTACGAC 719
baiA3      GATCGCGAATGTATATCTGTTTCCTTGCATCCGACCTGGCTAGCGGCATCACGGCTACGAC 719
baiA2      GATCGCTAATGTATACCTGTTTCCTTGCATCTGACTTGGCAAGCGGCATTACGGCTACTAC 719
baiAT0931  AATAGCTAACACTTACTTATCTTACTGCTTCTGACTTAGCTAGTGGTATAACAGCTACAAC 719
          * * * * *

baiA1      GATCAGCGTAGATGGGGCTTACAGGCCATAG--- 750
baiA3      GATCAGCGTAGATGGGGCTTACAGGCCATAG--- 750
baiA2      GGTCAGCGTAGACGGGGCTTACAGACCATAA--- 750
baiAT0931  TGTAAGCGTTGACGGTGCTTATAGACCATCATAA 753
          * * * * *

```



baiA271F: 5'-CAG GYG GMA CMM GWG GWA THG G-3'

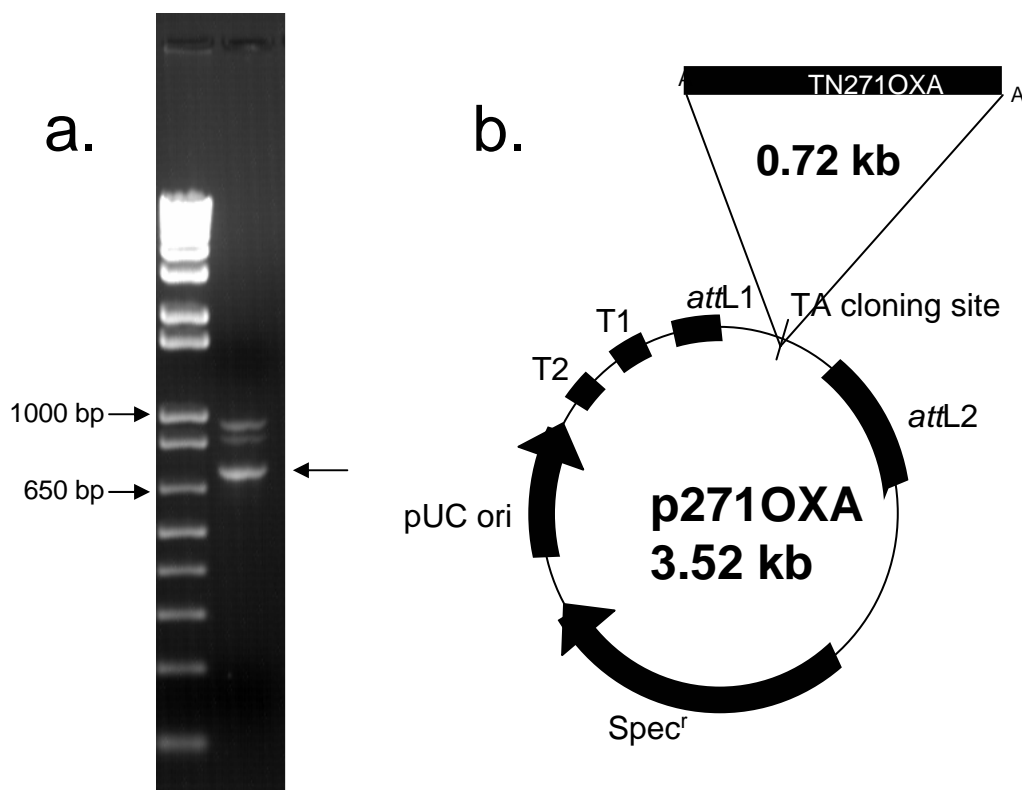
baiA271R: 5'-GGR CTR TAA GCM CCR TCW TCG C-3'

obtained, sequenced and predicted to encode the *baiA* gene by BLAST search (Figure 36). The complete 789 bp gene encoding a predicted 27.9 kDa protein was obtained following bidirectional genome-walking from the initial sequence data (Figure 37). Sequence analysis upstream of the *baiA* gene revealed a conserved *bai* promoter and RBS (Figure 37). Regions identified as highly conserved within *bai* promoters showed 100% identity between the upstream *baiB* promoter and that of the *baiA* promoter. 5' SMART RACE analysis suggests the mRNA initiation site is 56 bp upstream of the *baiA* gene (Figure 37). Figure 38 displays an alignment of the *baiA* gene from *C. hylemonae* with *baiA* genes from *C. scindens* and *C. hiranonis*.

**Identification and characterization of novel genes suggested to be involved in bile acid metabolism in *Clostridium hylemonae* and *Clostridium scindens***

The 3.3 kb contig containing the *baiA* and upstream *bai* promoter was further analyzed for additional ORF potentially involved in the reductive arm of the bile acid 7 $\alpha$ /7 $\beta$ -dehydroxylation pathway (Figure 39; Figure 40) (Table

Figure 36. Isolation of partial *baiA* gene from *C. hylemonae* TN271 using degenerate PCR oligonucleotides designed from conserved regions of *C. scindens* and *C. hiranonis baiA* genes. Arrow indicates PCR product corresponding to partial *baiA* gene as shown in nucleotide sequence alignment. PCR product was cloned into a pCR8 GW TOPO TA cloning vector resulting in vector p2710XA.



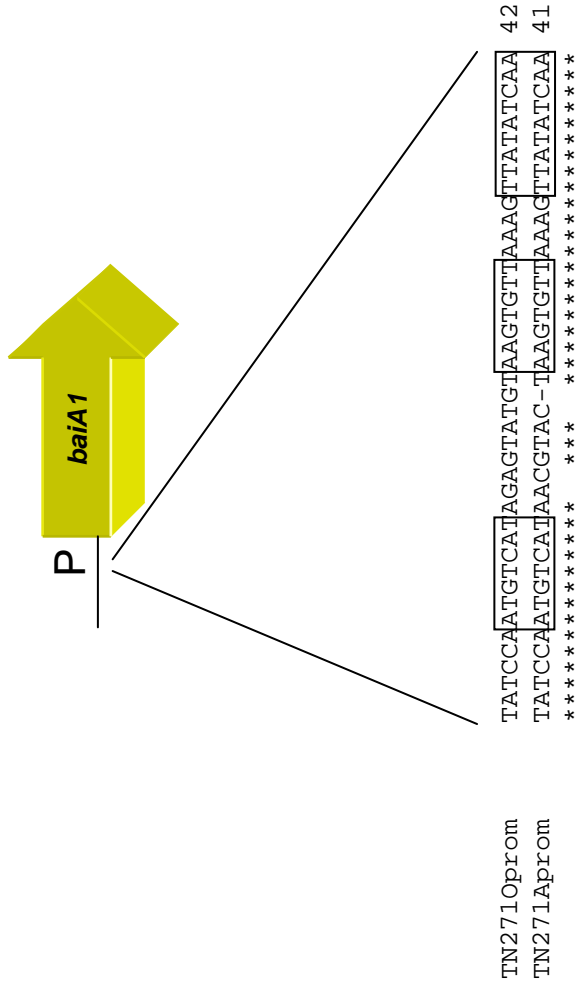
**C.**

12708b <i>aia2</i>	MDLVQDKTIIITGGTGGIGTAAAKIINDNGAKOSIFGPTQEEVDLSAQLKELYP EEEVL	50
TN271b <i>aia</i>	MDLVQDKTIIITGGTGGIGTAAAKIINDNGAKOSIFGPTQEEVDLSAQLKELYP EEEVL	14
T0931b <i>aia</i>	MDLVQDKTIIITGGTGGIGTAAAKIINDNGAKOSIFGPTQEEVDLSAQLKELYP EEEVL	50
12708b <i>aia2</i>	GFAPDLTSRDAMMAAVGQVAQKYGRLDUMINWAGITSSNUTSRVSEEEFFKHIMDINUTGV	120
TN271b <i>aia</i>	GFAPDLTSRDAMMAAVGQVAQKYGRLDUMINWAGITSSNUTSRVSEEEFFKHIMDINUTGV	74
T0931b <i>aia</i>	GFAPDLTSRDAMMAAVGQVAQKYGRLDUMINWAGITSSNUTSRVSEEEFFKHIMDINUTGV	120
12708b <i>aia2</i>	ENGAWAAYQCMKCPQQGVIIINTASUTGINGSLSGUGYPASKSSVGLTQGLGREIIRKN I	180
TN271b <i>aia</i>	ENGAWAAYQCMKCPQQGVIIINTASUTGINGSLSGUGYPASKSSVGLTQGLGREIIRKN I	134
T0931b <i>aia</i>	ENGAWAAYQCMKCPQQGVIIINTASUTGINGSLSGUGYPASKSSVGLTQGLGREIIRKN I	179
12708b <i>aia2</i>	RVUGVAPGVUNTDMTMCMPDEIEGVLKALPMKRMLEPEEIAMVYLFLASDLASGITATT	240
TN271b <i>aia</i>	RVUGVAPGVUNTDMTMCMPDEIEGVLKALPMKRMLEPEEIAMVYLFLASDLASGITATT	194
T0931b <i>aia</i>	RVUGVAPGVUNTDMTMCMPDEIEGVLKALPMKRMLEPEEIAMVYLFLASDLASGITATT	239
12708b <i>aia2</i>	VSUDGAYRE- 249	
TN271b <i>aia</i>	VSDDGAYS-- 202	
T0931b <i>aia</i>	VSUDGAYRES 249	

Figure 37. Isolation of *baiA* gene from *C. hylemonae* TN271 and characterization of upstream promoter elements. (a.) Schematic representation of sequence alignment and location of *baiA* gene relative to PCR product alignment. (b.) CLUSTALW alignment of *C. hylemonae* TN271 oxidative *baiB* promoter (TN271Oprom) with *baiA* promoter (TN271Aprom). Boxed regions correspond to conserved *bai* promoter elements. (c.) Results of 5' SMART RACE PCR analysis of *baiA* cDNA from CA induced culture (See Figure 20). mRNA initiation site (bold, underlined) indicated by arrow which points in direction of transcription. Italicized letters indicate 5' end of *baiA* gene with corresponding amino acids (below).



a.



b.



c.



Figure 38. CLUSTALW alignment of amino acid sequence of *baiA* genes from *C. hylemonae* TN271, *C. scindens* VPI 12708 and *C. hiranonis* sp. strain T0931. The *baiA* gene encodes a 3 $\alpha$ -hydroxysteroid dehydrogenase specific for choly-CoA (58).

12708	MNLVQDKVTII	TGGTRGIGFAAAKIFIDNGAKVSI	FGETQEEVD	TALAQLKELY	PEEEVL	60
TN271	MKLIVQDKITVI	TGGTRGIGFTAAKIFIENGAKVSI	FGETKEEVD	TALAEELKELY	PEEEVL	60
TO931	MNLVQDKIVII	TGGTSGIGLCAAKIFMDNGATVSI	FGKTQEEVD	AAKAEELKETH	PDKEVL	60
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
12708	GFAPDLTSRDAVMAA	VGQVAQKYGR	LDVMINNAGITS	NNVFSRVSEEE	FKHIMDINVTGV	120
TN271	GFAPDLTSRDAVMAA	VRVAEKEYGR	LDVMINNAGIT	TMTVFSRVTEEA	FKHMMDINVIGV	120
TO931	GFAPDLTNRDEVMAA	VGVAEKEYGR	LDVMINNAGVT	SSNVFSRVSP	EFTYLM	120
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
12708	FNGAWCAYQCMKDA	KKGVIINTASVTG	IFGSLSGVGY	PASKASVIGL	THGLGREI	180
TN271	FNGAWAAYQCMKGA	QQGVINTASVTGI	YGSLSGIGYP	ASKAGVIGL	TQGLGREI	180
TO931	FHGAWAAHCLKGE	KK-IIINTASVTGI	HGSLSGVGYPT	SKSAVVGFT	QALGREI	179
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
12708	RVVGVAPGVVNTDM	TNGNPPEIMEGYL	KALPMKRMLE	PEEIANVYL	FLASDLASGITATT	240
TN271	RVVGVAPGVVDTDM	TKGLPPEILEDY	LKSFPMKRML	KAEEIANVYL	FLASDLASGITATT	240
TO931	RVVGVAPGVVNTPM	VGNIPDEILDGYL	SSFPMKRMLE	PEEIANTYL	FLASDLASGITATT	239
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
12708	VSV	DGAYRP	-----	249		
TN271	VSD	DGAYS	PRANSTQLFVQVGI	262		
TO931	VSV	DGAYRPS	-----	249		
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	
	*****:	*****:	*****:	*****:	*****:	

Figure 39. Genome-walking PCR upstream of partial *baiA* gene of *Clostridium hylemonae* TN271. A nested PCR product derived from *PvuII* restriction library template (PN) ~2.5 kb was purified by GENECLEAN and ligated into pCR8 GW TOPO TA vector to produce vector p271PNUP1A.

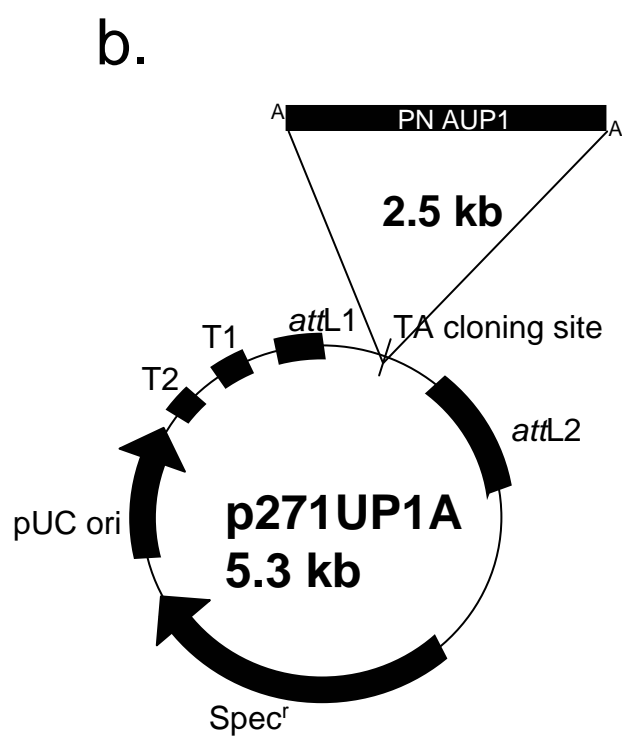
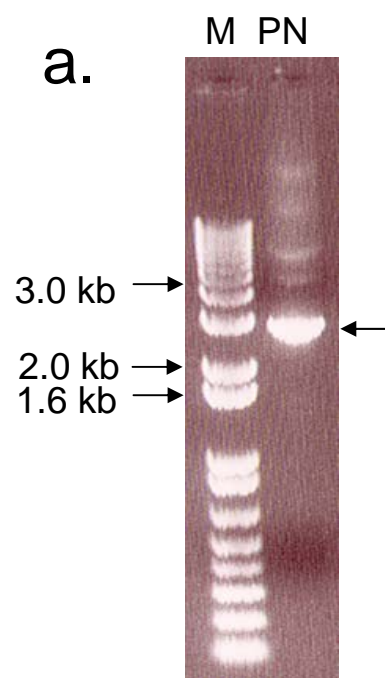


Figure 40. Downstream genome-walking PCR (DS1) from *baiA* gene of *Clostridium hylemonae* TN271. A nested PCR product derived from *EcoRV* (EN) restriction library template ~2.3 kb was purified by GENECLAN and ligated into pCR8 GW TOPO TA vector to produce vector p271ENDS1A.

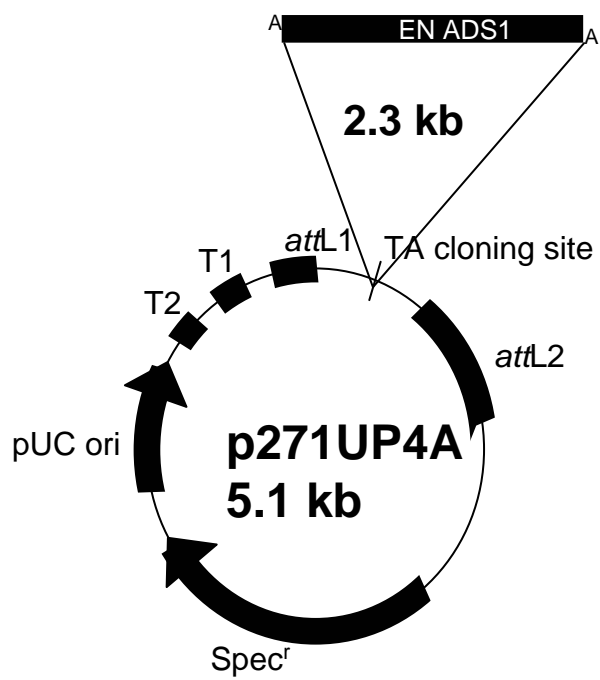
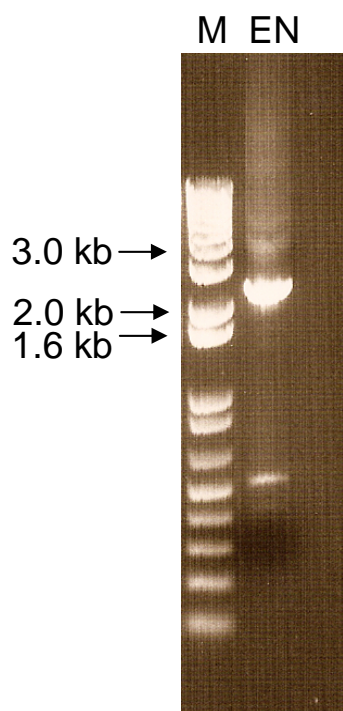


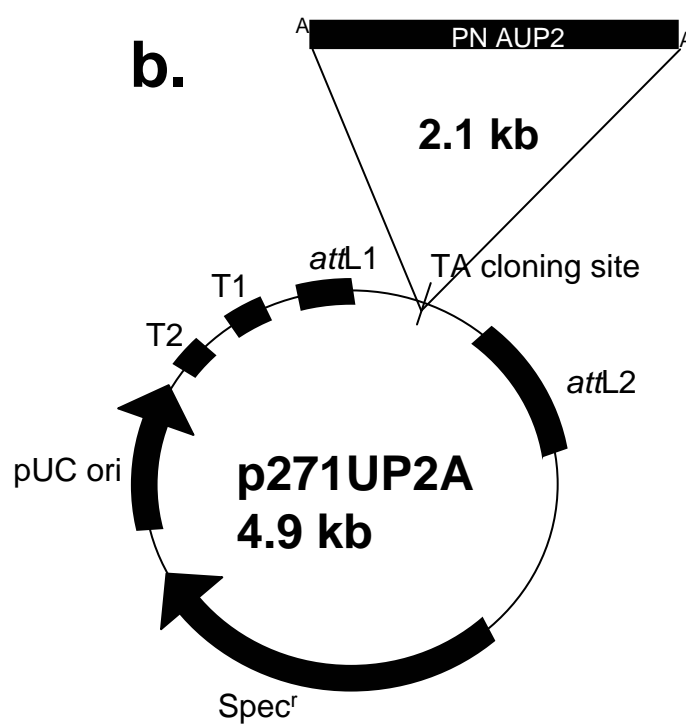
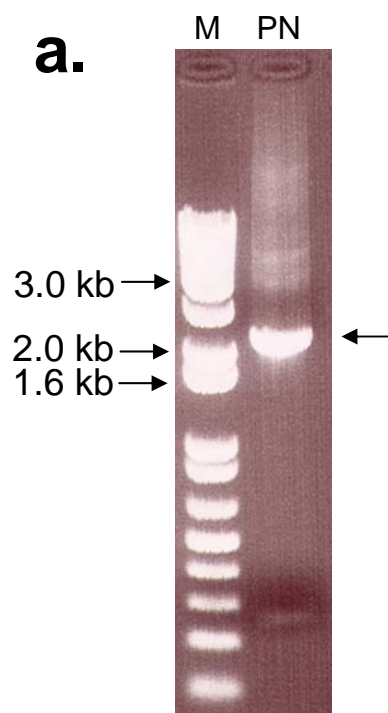
Table 8. Oligonucleotides used in isolation of *baiA* and flanking regions in *Clostridium hylemonae* TN271

Primer name	Primer Sequence (5'-3')	Reference
<b>Redundant oligonucleotides</b>		
baiΔ27LF	CAGCTGCGACCGGGGCGGGAATGCG	This study
baiΔ27LR	GCGCTTTAACTTCCTTCTGTCC	This study
<b>Genome-walking oligonucleotides</b>		
API	GTAAATC GACTCATATAGGC	Clontech
AP2	ACTATAAGGCACTGCTGCTGCT	Clontech
baiΔ03LKD3 L	GAATAGCGTATCTTGCGAATAAGG	This study
baiΔ03LKP1	ATACCTTATTC CGACAAGAACC	This study
baiΔUP1NEST	ACAAA CCGTA GATGCC TGTAC AGATGC	This study
baiΔDS1NEST	GTCGATTCGCTCTTACACAGAGTCTCG	This study
SUP2TR87 L	GCACTATAS GA AAA CTGAACAGAGCA	This study
SUP2TR87? LNEST	CTTGGAAGTCTTATATASCTGTGC	This study
SUP2TR87 L	GCACTGCTGCTATGATCTCAC	This study
SUP2TR87? LNEST	TGCAAGTCTGCTA CTGTGATTC	This study
UP4A	TTAAAGCTTCGAGTCTGACATTCC	This study
UP4BET	CATCTTCCTAGACCTGTTACATTACG	This study
<b>Primer walking oligonucleotides</b>		
G01	GTTCGAAACA AATGATGACGAAATCC	Invitrogen
G02	GTTCGAAACA AATGATGACGAAATTA	Invitrogen
baiΔR1	CTGATCGTTC GC GTTCTTC	This study
baiΔUPAR2	ATGTCTGTCAA TGGAAC	This study
baiΔUPAR3	CAGCTGTTCCCTGTTGAGG	This study
baiΔUPARR	GAATATC GGTATC CTGCACTGC	This study
baiΔUPDUFF2	GAAGTTGTGTCCGAGCTATAC C	This study
baiΔUPDUFFE	CAGCATATCTTTC CACTTC	This study
SUP2 G02R	GTCCTATATCTG CAAAC CATCG	This study
SUP2 G01F	CTCAAGAT GA AAAAT GGTGC	This study
UP4P1	GA AAATACCTGAGAACCTGATCC	This study
UP4R1	CACATATCTGTTGTTG CAAAGC	This study
UP4AF2	CTGATTC GCTTTA GAAGAAAC	This study
UP4AR2	CATATTC ACTGCTGC CTGTTC C	This study
UP4AF3	GGAAGTCACTCGAC GAGATCTGAGG	This study
UP4AR3	AAGTTCGTTTCCTCTGCAGC	This study

8). An ORF was located 235 bp upstream of the *baiA* gene. The putative 732 bp gene was predicted to encode a 26 kDa short chain dehydrogenase/reductase. Members of this family include the *baiA* gene which encodes a 3 $\alpha$ -HSDH. However, these two gene products shared only 19% amino acid sequence identity and 35% similarity. This ORF is of interest because we expect the final reductive step in the pathway to encode a 3 $\alpha$ -HSDH recognizing 3-dehydro-DCA and/or 3-dehydro-LCA. A second partial ORF was located 67 bp upstream of the putative short chain dehydrogenase. Intriguingly, this ORF was suggested to encode a type III CoA transferase whose highest sequence identity was that of the *baiF* gene from *C. scindens* VPI 12708. A second upstream genome-walk resulted in a PCR product 2,107 bp in length which increased the current contig length to 4,949 bp (Figure 41). The *baiF* homologue sequence was completed and found to be 1,332 bp in length and predicted to encode a 49.5 kDa polypeptide. The *baiF* gene product from *C. scindens* VPI 12708 was previously overexpressed in *E. coli* and reported to encode a 47.5 kDa CoA hydrolase (130); however, recent data suggests this gene product belongs to the Type III family of CoA transferases (131). The *baiF*



Figure 41. Second genome-walking PCR upstream of *baiA* from *C. hylemonae* TN271. a.) Nested PCR obtained from *PvuII* restriction digest (PN) of *C. hylemonae* genomic DNA template. b.) 2 kb product GENECLAN purified from agarose gel and ligated into pCR8 GW TOPO TA vector resulting in vector p271UP2A.



and this second ORF share 60% amino acid sequence identity, 75% amino acid similarity. Figure 42 shows an amino acid sequence alignment between the *baiF* gene and the putative CoA transferase.

Directly upstream, 52 bp from the hypothetical CoA transferase gene was another partial ORF with a flavin binding domain. Further genome-walking resulted in 760 bp of sequence data and completion of this 1,722 bp ORF encoding a putative 62 kDa flavoprotein (Figure 43). BLAST analysis suggested this gene product encodes a flavoprotein similar to fumarate reductases and 3-ketosteroid- $\Delta^1$ -dehydrogenases. The former can likely be ruled out due to the lack of four conserved N-terminal heme binding sites (CXXCH) found in fumarate reductases such as that of *Shewanella putrefaciens* MR-1, whose structure has been solved (192). The latter possibility was of particular interest because we expect one of the reductive genes to encode a 3-oxo- $\Delta^4$ -steroid oxidoreductase. A fourth upstream genome-walk resulted in a 2,765 bp of additional sequence data which increased contig length to 7,993 (Figure 44). This additional sequence data allowed analysis upstream of the putative 3-

Figure 42. Boxshade of CLUSTALW amino acid sequence alignment of *baiF* gene product of *C. hylemonae* and a predicted Type III CoA transferase which we have named the *baiK* gene (See text).

**baiK** MEGTGLNNFPQFCVLAGVKILVCGGAIAGPFGATLLGEIGAEVHLESPRNPDSTRGHYG 60  
**baiF** --MAGLKDFPSTFCALAGLKILDGSGNIAGPLCGGLLSECGATVTHFGPRKPDNQRGWYG 58

**baiK** YSQNHRNQISMVADNKTPEGLEIFKLIKWDIFIESSKGGTYEKMGLTDEVLEEVNPGI 120  
**baiF** YPQNHRNQISMVADIKTEEGREIFKLIKWDIIVESSKGGCYDRLGLSDEEIFKVNPKI 118

**baiK** VIVHVSFGGQTCVPEYISRASVDAGQAQFSGYMSFNGTPKEALKISPYLSDYVTALNTCM 180  
**baiF** AIVHVSFGGQTCDEPSYVTRASVDAGQAQFSGYMSLNGT-TEALKTNPYLSDFVCGLTTCM 177

**baiK** TALAAVYHVLFTGRGEAVDVAQYESLARINDTRPIEYFTDGKEFPRTGNKDSQAALFSFY 240  
**baiF** AMLACYVSTVLTGKGESVDIAQYELARINDGRVMQFATDGVKVPRTGNKDGQAALFSFY 237

**baiK** TCKDGGTIFIGMNCYGPVHRGYPLIGLPRPGDGDPEIDQIISGWMADTDLGRLEAAMEK 300  
**baiF** TCKDGRITIFIGTGAEVCKRGFPILIGLPCGSGDPDFPEGFTGMLIDSPVGRMEAAAMEK 297

**baiK** FVSEHTVDEVKIMLENQIPCQKVYSLEDCVDRDPHUKAREIFTEWDDPMMGRVKIGIVM 360  
**baiF** YVSEHTMEVDAMMCENQIPCQRVYELEDCTSDPHUIARETITEWDDPMLGHVTGLGLIM 357

**baiK** KWKNNPGEIKUGAPLFGENNKEVLRDLGYTEETEDLARRGITACLEFECTYETYKLGEF 420  
**baiF** KFKNNPSKIWRGAPLFGMDNRDILRDLGYTDEEIDGLYEKGITNEFIRETIKRYRLKEV 417

**baiK** PHYREGFTEHURTEQEDETYEI 442  
**baiF** -----IPHURDE----- 424

Figure 43. Third upstream genome-walking PCR from *baiaA* gene of *C. hylemonae* TN271. a). Agarose gel of DS2 products using *DraI* (D), *EcoRV* (E), *PvuII* (P), *StuI* (S) restriction libraries of *C. hylemonae* TN271 genomic DNA as template with adaptor primer AP1 and gene-specific primer WALKDS-2 (See Table 6). b.) "EN" and "PN" represent nested PCR amplification using nested adaptor primer AP2 and gene-specific primer WALKDSNEST2 (See Table 6). "M" represents 1 kb DNA ladder. c.) Schematic representation of TA cloning of GENECLLEAN purified "DN" product (arrow) into pCR8/GW/TOPO TA vector (2.8 kb) yielding 3.56 kb vector p271DNCDUP3.

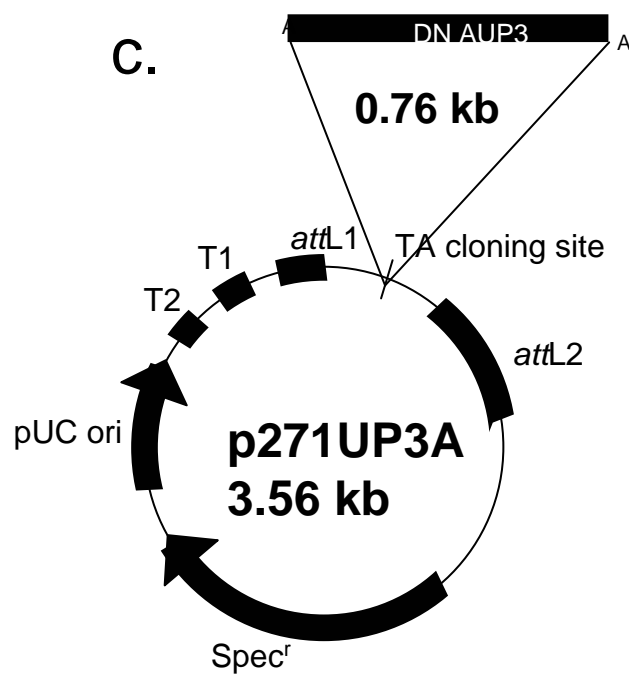
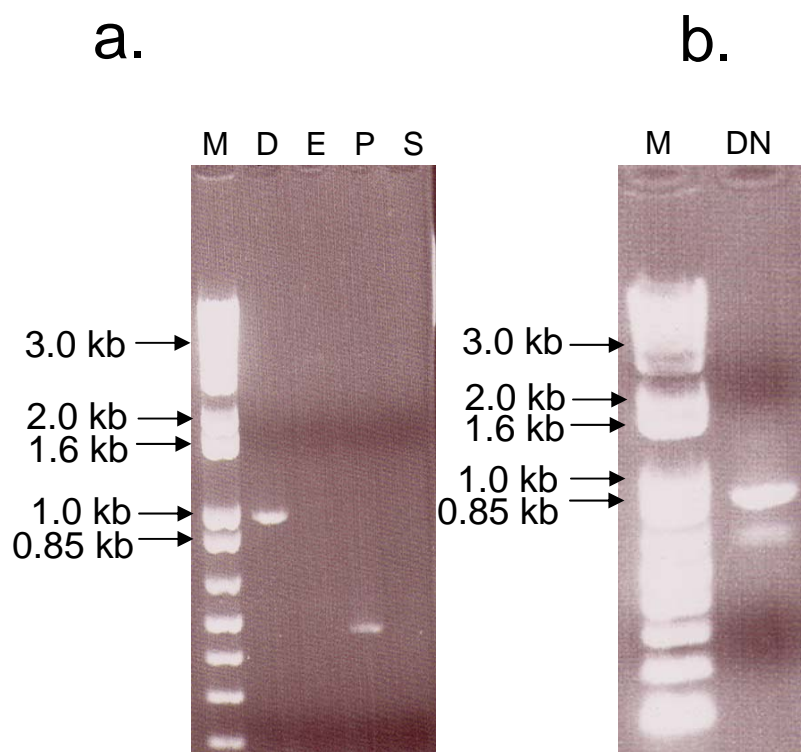
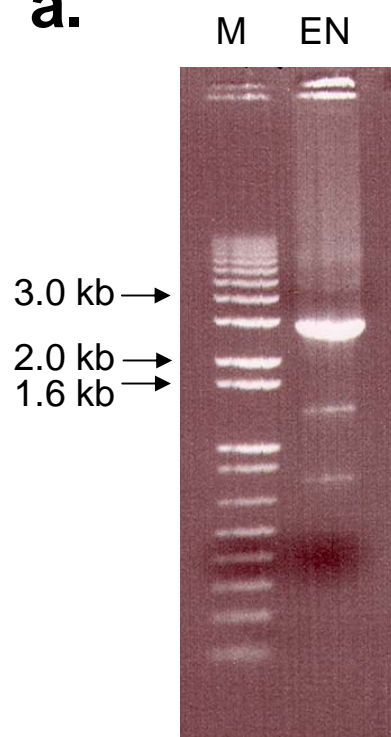
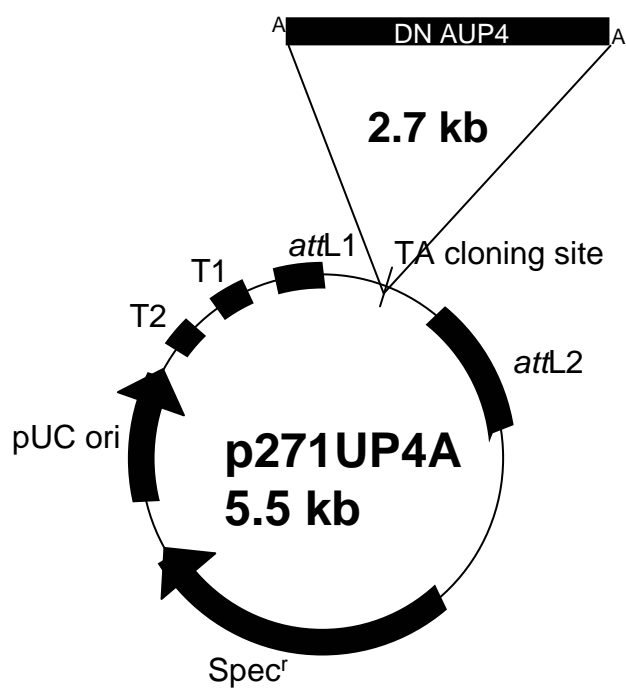


Figure 44. Fourth upstream genome-walking PCR from *baIA* gene of *Clostridium hylemonae* TN271. a.) Nested PCR obtained from *EcoRV* (EN) restriction digest of *C. hylemonae* genomic DNA template. b.) 2.7 kb product GENECLAN purified from agarose gel and ligated into pCR8 GW TOPO TA vector resulting in vector p271UP4A (5.5 kb)



**a.****b.**

oxo- $\Delta^4$ -steroid oxidoreductase gene for a *bai* promoter sequence. Indeed, we located a putative *bai* promoter 24 bp upstream of this ORF (Figure 45). Additional conserved *bai* promoter regions were not found in any other portions of this contig. 5' SMART RACE PCR was performed on the putative 3-oxo- $\Delta^4$ -steroid oxidoreductase gene and the mRNA start site was determined at the very 5' end of the conserved *bai* region suggesting the promoter is located upstream of this conserved region (Figure 46). It is quite possible that this conserved region is a promoter region for both the *baiA* and *baiB* operons; however, this region does not appear to be a promoter for the *baiJKL* genes. The functional significance of this conserved region remains to be elucidated.

For the sake of simplicity in description we will provide a preliminary name to these gene products following the sequence of nomenclature ending with the *baiI* gene on the oxidative operon. This first ORF encoding the putative 3-oxo- $\Delta^4$ -steroid oxidoreductase will be called the *baiJ* gene. The *baiF* homologue will be named the *baiK* gene and the putative short chain dehydrogenase will be referred to as the *baiL* gene.

Figure 45. Identification of conserved *bai* promoter elements upstream of the putative *baiJ* gene of *Clostridium hylemonae* TN271.

TN271redprom  
12708oxidprom

ACGATTAAAGCAGAAAGGCAATGCTTATTAAATATACAA  
ATATTCTATCTCA-AAAATGTTAAGTTATTATTCAG

Figure 46. Determination of transcription initiation site for *baiJ* gene from *Clostridium hylemonae* TN271 based on SMART RACE PCR analysis. Alignment above shows determination of transcription initiation site (TIS) based on alignment of SMART RACE sequence data (SR3, SR4) with *baiJ* gene and upstream region from *Clostridium hylemonae* TN271 (TN271D4). TIS is determined based on the first base following 3' end of SMART II A oligo sequence (*italics*) (5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3'). This nucleotide (          , arrow) must also correspond to the known sequence of the upstream region of the gene. Therefore, the TIS is 47 bp upstream of the ATG start site (          ). RBS=ribosome binding site.



SR3 CTAATACGACTCACTATAGGGCAAGCAGTGGTATCCACGC-AGAGTACGCGGGAGAAAAA  
 SR4 CTAATACGACTC-CTATAGGGCAAGCAGTGGTATCCACGC-AGAGTACGCGGGAGAAAAA 58  
 TN271D4 -----TTAAATTATACAGGCTGATTTTGTAGATGTTAGAATA-AAAGCAGAAAAA 48  
   \*\*      \*\*  \*  \*      \*  \*      \*  \*      \*  \*\*\*\*\*


SR3 CATGTTATTAAAAATAACAAATGATCAGAAAAAGGAGGCAAATGCGATGGCAAATATGTA 119  
 SR4 CATGTTATTAAAAATAACAAATGATCAGAAAAAGGAGGCAAATGCGATGGCAAATATGTA 118  
 TN271D4 CATGTTATTAAAAATAACAAATGATCAGAAAAAGGAGGCAAATGCGATGGCAAATATGTA 108  
   \*\*\*\*\*

SR3 CCAGGCACGTACGAAGGTATCGGTCACGGATACCAGGGCAGGCTTATCGTGAATGTAACA 179  
 SR4 CCAGGCACGTACGAAGGTATCGGTCACGGATACCAGGGCAGGCTTATCGTGAATGTAACA 178  
 TN271D4 CCAGGCACGTACGAAGGTATCGGTCACGGATACCAGGGCAGGCTTATCGTGAATGTAACA 168  
   \*\*\*\*\*

SR3 GTTACGGAAGATGAGATCACGGAGGTTAAAATCATCAAAACATAAGGAAGTGCAGGGGGCTT 239  
 SR4 GTTACGGAAGATGAGATCACGGAGGTTAAAATCATCAAAACATAAGGAAGTGCAGGGGGCTT 238  
 TN271D4 GTTACGGAAGATGAGATCACGGAGGTTAAAATCATCAAAACATAAGGAAGTGCAGGGGGCTT 228  
   \*\*\*\*\*

SR3 GCGTGGGACCTTCCCACATCCCCTGTCGAGGTGATGCCGCCTCAAATTGTCAAATATCAG 299  
 SR4 GCGTGGGACCTTCCCACATCCCCTGTCGAGGTGATGCCGCCTCAAATTGTCAAATA---- 294  
 TN271D4 GCGTGGGACCTTCCCACATCCCCTGTCGAGGTGATGCCGCCTCAAATTGTCAAATATCAG 288  
   \*\*\*\*\*

TN271 baiJ start site:



RBS

AGAAAAACATGTTATTAAAAATAACAAATGATCAGAAAAAGGAGGCAAATG

**Isolation of the *baiJ* gene from *C. scindens* VPI 12708 and elucidation of conserved *baiJKL* genes.**

A necessary question immediately arose: are these genes conserved in other 7 $\alpha$ -dehydroxylating bacteria? To answer this question, the *baiJ* gene from *C. hylemonae* TN271 was analyzed for short 7-8 amino acid stretches with regions of low codon degeneracy. Amino acids 183-189 (EEMCYNY) and 394-401 (NEWSYQYH) were selected as sites of lowest degeneracy, a section whose nucleotide sequence would result in an expected PCR product of 654 bp (Figure 47). A codon usage table was prepared using sequence data from *bai* genes from *C. scindens* VPI 12708 using the codonW program (Figure 48). This table allowed design of least degenerate oligonucleotides for an attempt at isolation of the *baiJ* gene from *C. scindens* VPI 12708 (Table 9). Amplification using forward primer Delta4F1 and each of the reverse primers Delta4R1, Delta4R2 and Delta4R3 each resulted in a product around 650 bp (Figure 49) (See Table 9 for primer sequences). The 650 bp PCR product amplified using Delta4F1 and Delta4R2 primers was excised from the

Figure 47. Amino acid sequence of *baiJ* gene from *Clostridium hylemonae* TN271. Residues highlighted in green have one or two possible codons. Bold regions correspond to amino acids 183-189 (EEMCYNY) and 394-401 (NEWSYQYH). These regions were selected based on low codon degeneracy which will be used to design oligonucleotides to potentially isolate the *baiJ* gene from *Clostridium scindens* VPI 12708.



MANNVPGTVEGIGHGVQGRLLIVNVTVTEDEITTEVKI IKHKEVRGLANDLPTSPVEVMPPPQIVKVCSLNIP  
 LVNGADLTSGAIL EAVEGALRAAGAADDVIEKLKAAPGE APEVKDEFVRTVDVAVFGAGAGGLAAAI EAK  
 EGGADVVL IEKQGITGGSTARSGGKLLCAGTKWQKKQGIYDTEEMCYNYVLMNVGMRRGNFMDASKNRVLY  
 KNLNSTLDMLGTMGYKVQDV EAIHVS MQPVRVHNSMGGGGQTN GGG EITTTPLTRHFAGRLLGGEI IYNTA  
 LKELLTDENGVVNGAVCEKLDGSKLTVYAKKGVILATGGYSRNKEMCARYPVAHYFC TAPKSNVGEGLVA  
 AEKIGARNFVHPG IQVVTSLSCGIGINDESLIVNERGERVWNEWSYQYHVSDALAAAGSNCGWYITSG  
 DEPYGGVQYGFKQAVEGTSRDKAADSIEELAALIKCDPATLRATFDRYQELTARGEDEDFGKPARFLHPI  
 DGPKY AALRLHPCVTVS FGGLE ETDI SARVLDN EGRPI PGLVAAAGEVADTCMF GTEYPTCGTSLGGALFYG  
 RIAGRMASGQSM L \*

Figure 48. Codon usage table for *Clostridium scindens* VPI 12708 and design of redundant oligonucleotides for isolation of the *baiJ* gene from *C. scindens* VPI 12708. a). Codon usage table generated using *bai* genes from *Clostridium scindens* as input (U57489). b). Design of forward and reverse degenerate oligonucleotides based on regions identified in Figure 45.

a.

Phe	UUU	50	0.78	Ser	UCU	25	0.95	Tyr	UAU	70	1.35	Cys	UGU	15	0.45
	UUC	78	1.22		UCC	51	1.94		UAC	34	0.65		UGC	52	1.55
<u>Leu</u>	UUA	18	0.46		UCA	17	0.65	TER	UAA	5	1.88	TER	UGA	2	0.75
	UUG	13	0.33		UCG	8	0.30		UAG	1	0.38	Trp	UGG	33	1.00
	<u>CUU</u>	71	1.80	Pro	CCU	33	0.92	His	CAU	38	1.29	Arg	CGU	29	1.27
	CUC	11	0.28		CCC	6	0.17		CAC	21	0.71		CGC	33	1.45
	CUA	1	0.03		CCA	35	0.98	Gln	CAA	1	0.03		CGA	4	0.18
	<u>CUG</u>	123	3.11		CCG	69	1.93		CAG	72	1.97		CGG	8	0.35
Ile	AUU	41	0.54	Thr	ACU	20	0.48	Asn	AAU	61	1.07	Ser	AGU	7	0.27
	AUC	164	2.17		ACC	55	1.33		AAC	53	0.93		AGC	50	1.90
	AUA	22	0.29		ACA	49	1.18	Lys	AAA	61	0.66	Arg	AGA	44	1.93
Met	AUG	124	1.00		ACG	42	1.01		AAG	124	1.34		AGG	19	0.83
Val	GUU	44	0.74	Ala	GCU	44	0.64	Asp	GAU	68	0.89	Gly	GGU	12	0.16
	GUC	42	0.70		GCC	51	0.74		GAC	84	1.11		GGC	127	1.73
	GUA	102	1.71		GCA	114	1.65	Glu	GAA	128	1.12		GGA	136	1.86
	GUG	51	0.85		GCG	68	0.98		GAG	100	0.88		GGG	18	0.25

3152 codons in 12708 (used Universal Genetic code)

b.

Forward Primer

TN271 sequence : GAG GAG ATG TGC TAC AAC TAC CTG ATG GAA  
 Amino Acid Seq : E E M C Y N Y L M E

(Redundant) 5'GA(G/A)GA(G/A)ATGTG(T/C)TA(T/C)AA(T/C)TA(T/C)(C/T)T(A/G/T/C)ATGGA(A/G)-3'

	E	E	M	C	Y	N	Y	L	M	E
Delta4F1	5'GA(G/A)	GA(G/A)	ATG	TG(T/C)	TA(T/C)	AA(T/C)	TA(T/C)	(C/T)T(G/T)	ATG	GA(A/G)
Delta4F2	5'GA(G/A)	GA(G/A)	ATG	TG	TA(T/C)	AA(T/C)	TA(T/C)	(C/T)T(G/T)	ATG	GA(A/G)
Delta4F3	5'GA(G/A)	GA(G/A)	ATG	TG	TA(T)	AA(T/C)	TA(T/C)	(C/T)T(G/T)	ATG	GA(A/G)

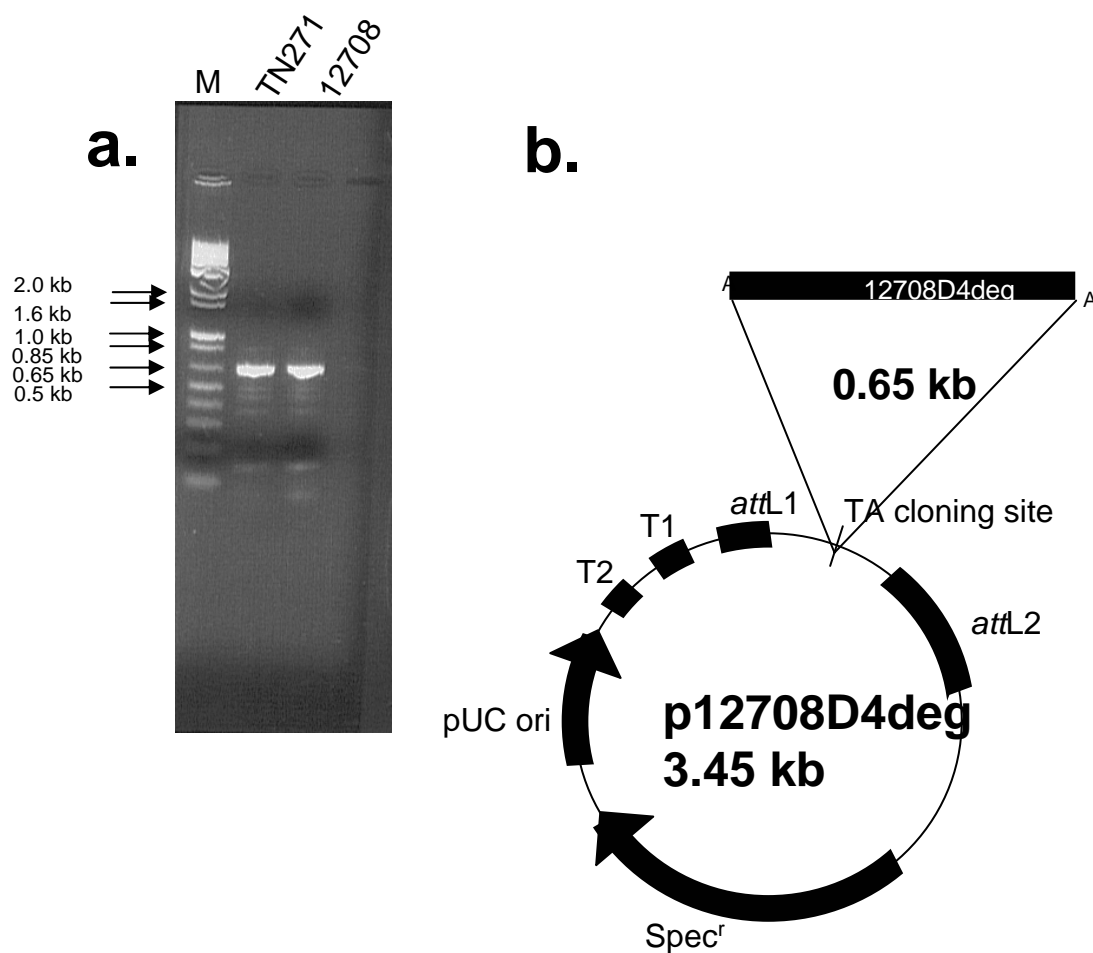
Reverse Primer

TN271 sequence : AAT GAA TGG AGC TAC CAG TAT CAC  
 Amino Acid seq : N E W S Y Q Y H

GAA(T/C)GA(A/G)TGGAG(T/C/A)TA(T/C)CA(A/G)TA(T/C)CA(T/C)

	H	Y	Q	Y	S	W	E	N
Delta4R1	5'-(G/A)TG	(A/G)TA	(T/C)TG	(A/G)TA	(T/A/G)CT	CCA	(T/C)TC	(G/A)TT C-3'
Delta4R2	5'-(G/A)TG	(A/G)TA	(T/C)TG	(A/G)TA	(G)CT	CCA	(T/C)TC	(G/A)TT C-3'
Delta4R3	5'-(G/A)TG	(A/G)TA	(C)TG	(A/G)TA	(T/A/G)(G/C)(T/A)	CCA	(T/C)TC	(G/A)TT C-3'

Figure 49. Isolation of partial *baiJ* gene from *Clostridium scindens* VPI 12708. a). PCR products generated from amplification with primers Delta4F1 and Delta4R2 from both *Clostridium hylemonae* TN271 and *Clostridium scindens* VPI 12708 genomic DNA templates. b). 0.65 kb PCR product from *Clostridium scindens* VPI 12708 was GENECLAN purified and cloned into a pCR8/GW/TOPO TA vector generating a 3.45 kb vector p12708D4deg. c.) CLUSTALW alignment of sequence data obtained from p12708D4deg insert against *baiJ* gene from *Clostridium hylemonae* TN271.



**c.**

12708 <i>baiJ</i>	-----MDASKNRYLVEHLNETLDWLGTMGYQVQDVEA IHVSLQPW	40
TN271 <i>baiJ</i>	DTEEMCYNYLMEVGNRRGNFMDASKNRYLVEHLNSTLDWLGTMGYQVQDVEA IHVSMQPW	240
	*****; **; *****; *****; ***	
12708 <i>baiJ</i>	RUHNSMGGGGQTNGQGGIITPLTHHYVDKLGGELIYNTALKELLTDENGTVNGAVCEKL	100
TN271 <i>baiJ</i>	RUHNSMGGGGQTNGQGGIITPLTRHFAGRLGGELIYNTALKELLTDENGTVNGAVCEKL	300
	*****; *; _; *****; *****; *****; *****	
12708 <i>baiJ</i>	DGSKLTUYAKKGUILATGGYSRNKEMCARYPUAHYFSTTPKSNVGEGLIAAEKIGARNFV	160
TN271 <i>baiJ</i>	DGSKLTUYAKKGUILATGGYSRNKEMCARYPUAHYFCTAPKSNVGEGLVAAEKIGARNFV	360
	*****; *; *****; *****	
12708 <i>baiJ</i>	HPGIQVUYTSLTCGIGINDESGLIUMERGERUV-----	193
TN271 <i>baiJ</i>	HPGIQVUYTSLSCGIGINDESGLIUMERGERUVNEWSYQYHVS DALAASGSNCGWYITSG	420
	*****; *****	

Table 9. Oligonucleotides used in isolation of *baiJKL* operon and flanking sequence from *Clostridium scindens* VPI 12708

Primer name	Primer Sequence (5'-3')	Reference
<b>Redundant oligonucleotides</b>		
De1c4F1	GAR GARA TCTGTTAAAYAAATAYATATGAG	This study
De1c4F2	GAR GARA TCTGTTAAAYATAYATATGAG	This study
De1c4F3	GAR GARA TCTGTTAAAYATAYATATGAG	This study
De1c4F4	ATGATATGCTGATGCTTCAATCTCTTC	This study
De1c4F5	ATGATATGCTGATGCTTCAATCTCTTC	This study
De1c4F6	ATGATATGCTGATGCTTCAATCTCTTC	This study
<b>Genome-walking oligonucleotides</b>		
API	CTATATAC GACTCATATATAGGCG	Clontech
AP2	ACTATAGCGGACGCGGCGT	Clontech
12708walKUP1	GTTATGTATCTTCTCATGCTGAGCG	This study
12708walKUP1M	CTTCCACGTTCTGAGGCTTGCTAGC	This study
12708walKUP2M	GCTTCTTATATAGGACGATCTTATATCC	This study
12708walKUP2	CAGACTATAGACAGGCTTCTTCTGATCC	This study
12708walKUP3	CTTCTGAGGCTTCTGATGAGGAC	This study
12708walKUP3L	GAGACGATGCTGCTACGAGCTTACG	This study
12708walKUP3LM	CATTTCTAGCTATATGCTTATATCTTC	This study
12708walKUP3Z	CTTCTAGGAGCTATGAGGCTTATCTTC	This study
12708walKUP3ZM	CTGACATATCTGAGGATATCTTCTGAC	This study
12708MD33	CTGCTTCTGCTGCTATATCTTCTGATCC	This study
12708MD33M	AGGAGAGCTCTGCTAGGAGATCTGAGG	This study
DS412708	AGGATTTGAGGAGGACATCTGAGGAGCTG	This study
DS412708M		
<b>Primer-walking oligonucleotides</b>		
12708UP1R1	CTTATGCTTCTTATATCTTCTATCC	This study
12708UP1F1	CGTTTGAATCTCATATGAGATATCTTC	This study
UP2R1	GATTTCTCTATATCTATGCGAGACAGC	This study
UP2F1	GCGCTATATCTATCTTCTCTAGC	This study
12708D3LR1	CTGCTCTTAAATCTCATATAGC	This study
12708D3LF1	TCACTGAGCTATCATATCTATCT	This study
DS12F	TTCCTCTGATGAGCTCATCTTCTGAC	This study
DS12FM	ATTTCTTCTTCTATATGATGATCTTC	This study
DS12R	ATGTCACGCTCTTATATATCTTCTGCA	This study
12708D32R1	GTCGAGAGCTCTGATATATCTTCTTC	This study
12708D32F1	TCTTATATGAGCTGATGATGAG	This study
DS22F	TCTTATATGAGCTGAGCTGCTTCTG	This study
DS22R	TGCTGCTTCTTCTTCTCATATGAC	This study
DS400F	AAAGCTACCTGAGAGGAGGCT	This study
DS23FM	TAACTGCTTAACTGAG	This study
DS23RM	GCTATCATTTCTCTGCA	This study
601	CTTCTAACTATGATGAGCTATCTC	Invitrogen
602	CTTCTAACTATGATGAGCTATCTC	Invitrogen

agarose gel, GENECLAN purified, cloned into a pCR8 TOPO TA cloning vector and both strands sequenced. An alignment of the partial *baiJ* sequence obtained from *C. scindens* against the *baiJ* gene from *C. hylemonae* TN271 shows the high degree of amino acid sequence conservation (Figure 49). The partial *baiJ* gene from *C. scindens* was used to design bidirectional genome-walking PCR oligonucleotides to both complete the gene sequence and obtain information about any flanking ORF.

The first upstream genome-walk from the partial *baiJ* gene resulted in a 1,330 bp product (Figure 50). Addition of this sequence data to the partial *baiJ* sequence resulted in a contig of 1.8 kb. This sequence data contained the 5' end of the *baiJ* gene. A conserved "bai promoter region" was located directly upstream of the *baiJ* gene and shared a high degree of sequence identity with the putative *bai* promoter region found upstream of the *baiJ* gene of *C. hylemonae* (Figure 51). SMART RACE analysis of the 5' end of the *baiJ* gene revealed the TIS upstream of the putative conserved *bai* "promoter" region (Figure 52). These results were similar, but not identical to the TIS of the *baiJ* gene from *C. hylemonae*

Figure 50. Genome-walking PCR upstream of partial *baiJ* gene of *Clostridium scindens* VPI 12708. a.) Nested PCR products obtained from *DraI* (DN) and *EcoRV* (EN) restriction digest of *C. scindens* genomic DNA template. Due to larger size, product "EN" was isolated, cloned and sequenced. b.) 1.3 kb product GENECLAN purified from agarose gel and ligated into pCR8 GW TOPO TA vector resulting in vector p12708D4ENUP1 (1.3 kb).



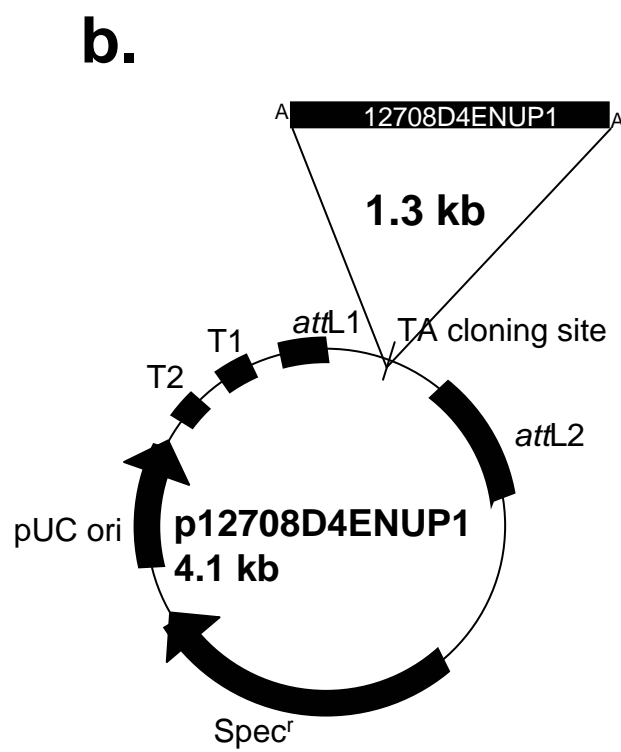
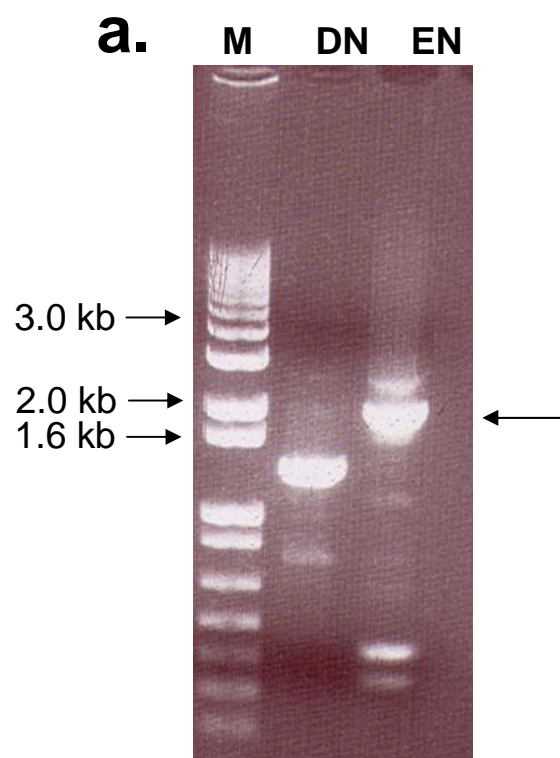


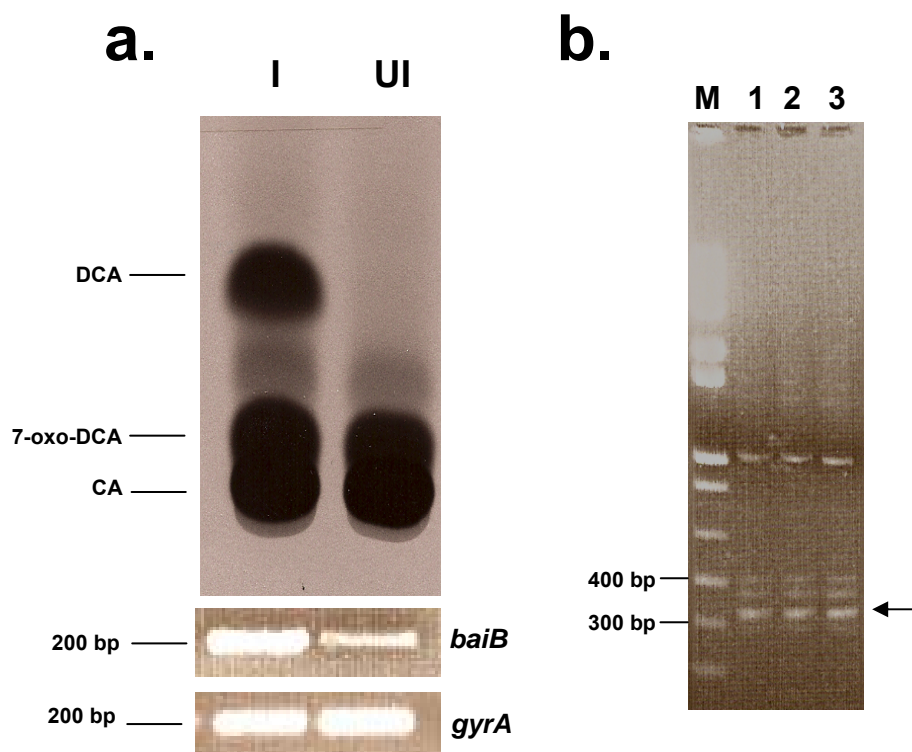
Figure 51. BOXSHADE of CLUSTALW alignment of upstream conserved *bai* promoter elements located upstream of *baiJ* genes of *Clostridium scindens* VPI 12708 and *Clostridium hylemonae* TN271.

TN271redprom	AGAAATTAAGCAAGCAAAACATCGTTATTAAATTAACAAAT
12708redprom	AGAAATGAACTTAATTAATAAAATCGTTATTAAATTAACAAAT
consensus	AGAAATaAAXXXAgxAAAAAATCGTTATTAAATTAACAAAT

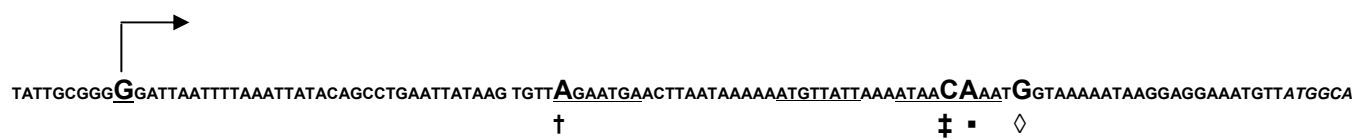
---

-10

Figure 52. Verification of bile acid induction and 5' SMART RACE PCR analysis of *baiJ* transcript from *C. scindens* VPI 12708. (a). TLC autoradiograph (Top) between CA induced (I) and CA uninduced (UI) cultures of *C. scindens* VPI 12708. RT-PCR (Below) of *baiB* gene comparing cDNA produced from RNA isolated from CA I and CA UI cultures of *C. scindens* VPI 12708 using primers *baiB*12708RTF (5'-GAT GAG CAG GGC TTA TCT GTA CTT T-3'), and *baiB*12708RTR (5'-GTA TCT CTT TCC CTG TCT CAA TGA C-3') and compared to *gyrA* gene. (b). 1% Agarose gel stained with ethidium bromide showing 5' RACE PCR products using gene-specific primer SR12708D1 (5'-TTC TAT GTC CTC ATC CGT TGC CCC TG-3'). Lane "M" is a 1 kB DNA marker, lanes 1, 2, and 3 represent PCR products obtained at various annealing temperatures (62°C, 64°C, 66°C, respectively). Arrow indicates band that was isolated, and sequenced. (c). Schematic representation of TIS obtained from 5'RACE sequence data. "†"= TIS position obtained from 5'RACE of *C. hylemonae* TN271 *baiJ* gene (Figure 46). "‡"= TIS position determined by primer extension (Mallonee et al. 2000) and 5'RACE (Figure 27) for the *baiB* gene from *C. scindens* VPI 12708. "■" represents TIS position determined by primer extension analysis of the *baiB* gene from *C. hiranonis* sp. strain T0931. "◇" represents TIS position determined by 5'RACE PCR of *baiB* gene from *C. hylemonae* TN271 (Figure 26).



**c.**



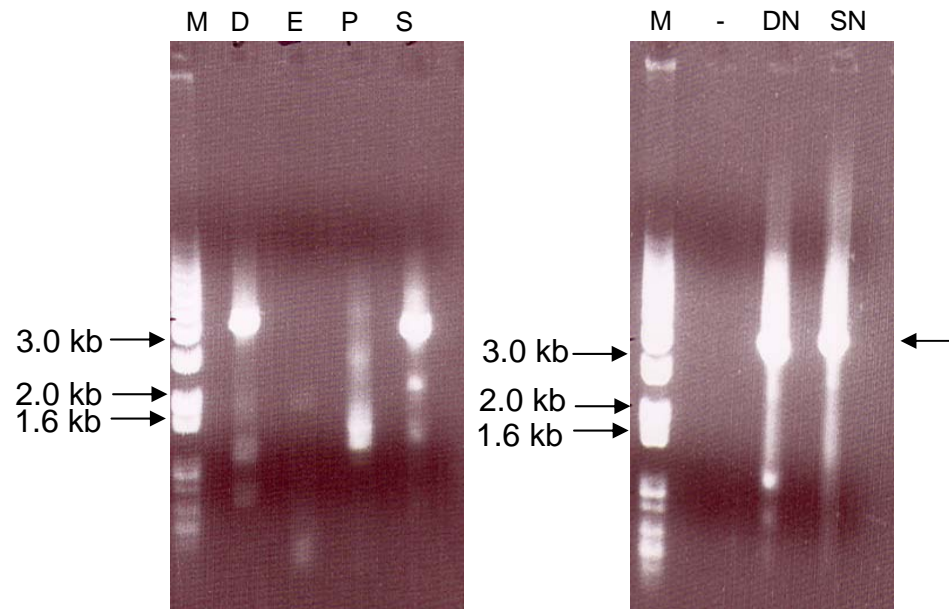
TN271. The TIS, located upstream of this conserved region, effectively rules this region out as the *bai* promoter.

In addition, a partial ORF was found 128 bp upstream of the *baiJ* gene on the antisense strand. A BLAST search of the ORF suggested this gene belonged to the LysR family of transcription factors. Due to the proximity of the TIS of the *baiJ* gene to the LysR like gene on the opposite strand, it is very likely that the *baiJ* promoter overlaps the LysR-like gene. A second upstream genome-walk resulted in a 4.0 kb PCR product which increased the contig to 5.7 kb (Figure 53). This sequence data completed the remainder of a 960 bp gene encoding a deduced 36.2 kDa polypeptide predicted to encode a transcriptional regulator in the LysR family. A second ORF located on the anti-sense strand relative to the *baiJ* gene was found 73 bp upstream of the putative transcription factor. This 465 bp ORF was predicted to encode a 17.8 kDa polypeptide in the Tryptophan-rich sensory protein (TspO)/ mitochondrial benzodiazepine receptor (MBR) homolog family of signal transduction integral membrane proteins. Further upstream, and on the

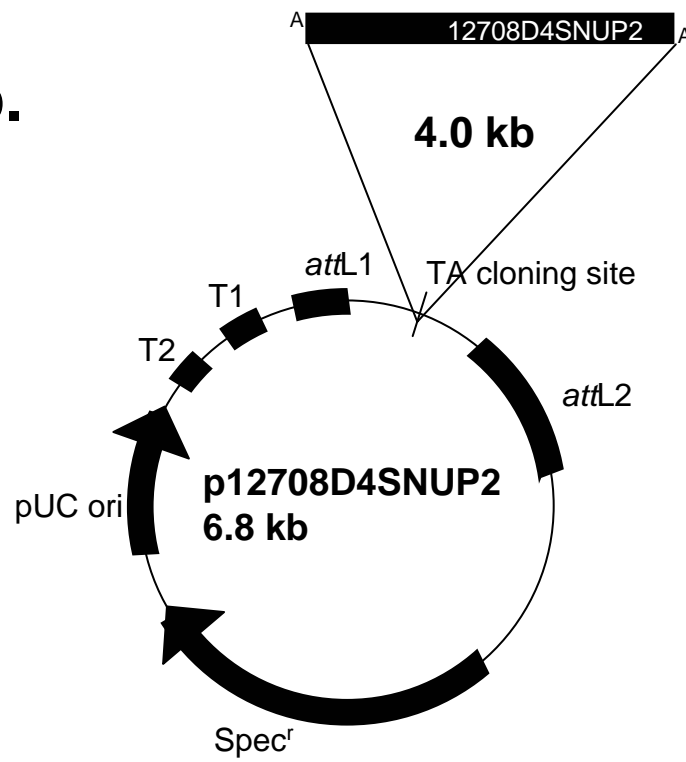
Figure 53. Second upstream genome-walking PCR (UP2) from partial *baiJ* gene of *Clostridium scindens* VPI 12708. a). Agarose gel of UP2 products using *DraI* (D), *EcoRV* (E), *PvuII* (P), *StuI* (S) restriction libraries of *C. scindens* VPI 12708 genomic DNA as template with adaptor primer AP1 and gene-specific primer 12708WalkUP2 (See Table 6). b.) "DN" and "SN" represent nested PCR amplification using nested adaptor primer AP2 and gene-specific primer 12708WalkUP2N (See Table 6). A positive clone was obtained for product "SN", so this was used for sequencing. "M" represents 1 kb DNA ladder. c.) Schematic representation of TA cloning of GENECLEAN purified "SN" product (arrow) into pCR8/GW/TOPO TA vector (2.8 kb) yielding 6.8 kb vector p12708D4SNUP2.

**a.**

210



**b.**





same strand relative to the *baiJ* gene, we located the *baiA1* gene and promoter region which has been previously reported (122). This gene was located directly upstream of the TspO homologue. It is interesting to note that the *baiA* gene and *baiA* promoter located in *C. hylemonae* was found in proximity to this putative reductive *bai* operon.

The initial genome-walk downstream of the partial *baiJ* gene resulted in a PCR product 1,929 bp in length (Figure 54). Alignment with the previous sequence data resulted in a contig of 7,114 bp. This additional sequence data resulted in completion of the 1,722 bp *baiJ* gene predicted to encode a 61.2 kDa polypeptide. Amino acid sequence identity of the *baiJ* gene is highly conserved between *C. scindens* and *C. hylemonae* (89%) (Figure 55). In addition, a partial *baiK* gene was located 37 bp downstream of the *baiJ* gene in *C. scindens*. A second downstream genome-walk resulted in a fragment 3,366 bp in length increasing the contig length to 10.13 kb. The 1,314 bp *baiK* was completed and predicted to encode a 49.1 kDa polypeptide. The *baiL* gene, 732 bp in length, was located 79 bp downstream of the *baiK* gene. The

Figure 54. Downstream genome-walking PCR (DS1) from partial *baiJ* gene of *Clostridium scindens* VPI 12708. a.) Nested PCR products obtained from *DraI* (DN) and *EcoRV* (EN) and *StuI* (SN) adaptor ligated restriction libraries of *C. scindens* VPI 12708 genomic DNA templates. Due to larger size, product "EN" was cloned. b.) 1.9 kb product GENECLAN purified from agarose gel and ligated into pCR8 GW TOPO TA vector resulting in vector p12708D4ENDS1 (4.7 kb).

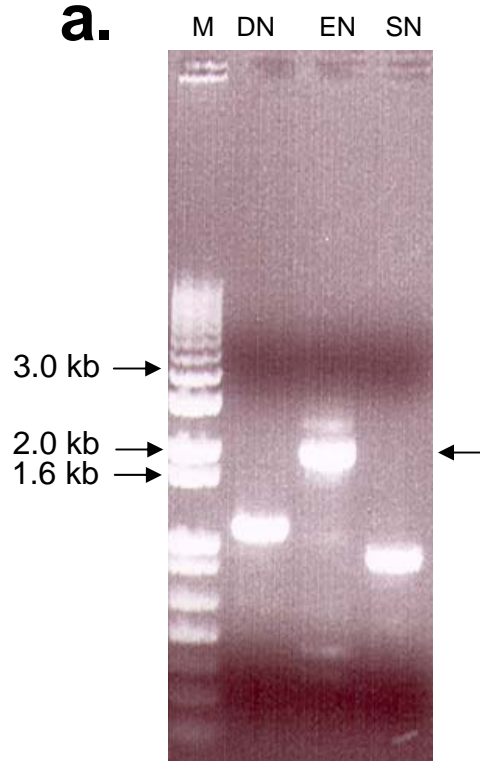
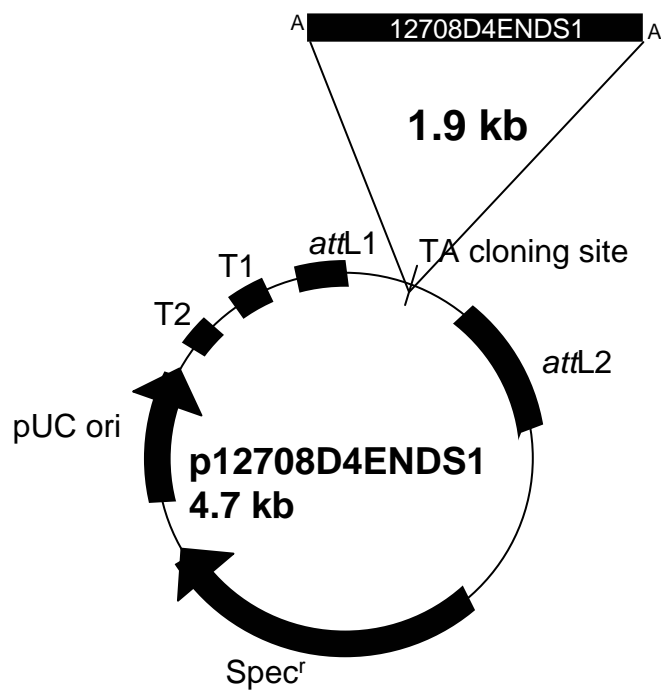
**a.****b.**

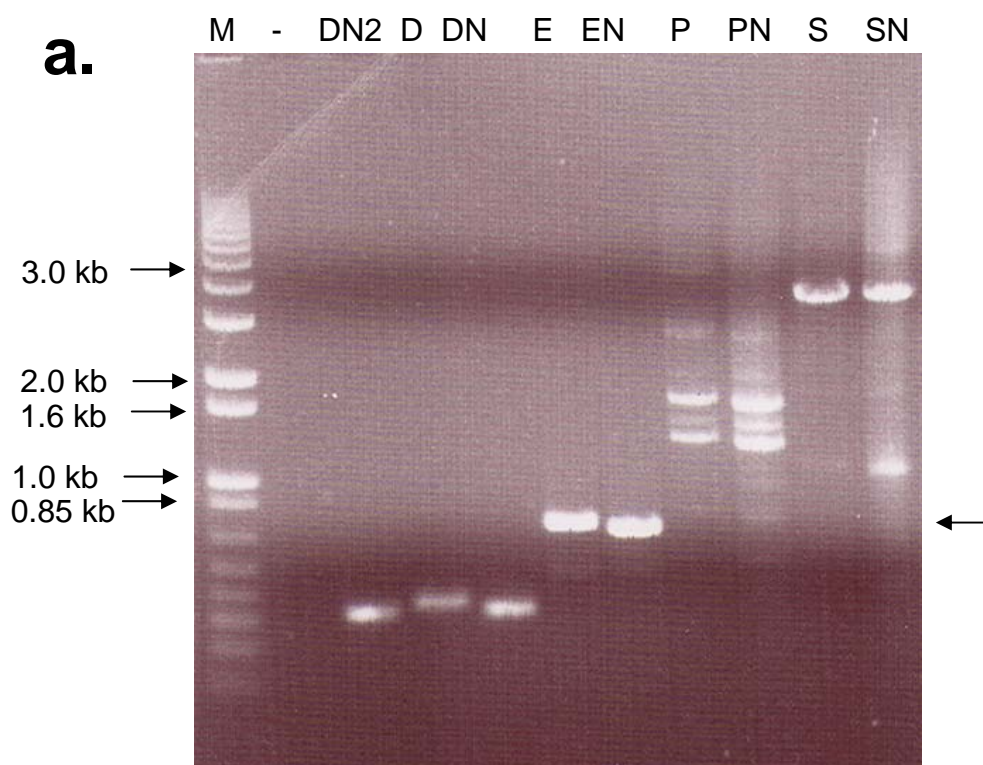
Figure 55. Boxshade representation of CLUSTALW amino acid sequence alignment of putative 3-oxo- $\Delta^4$ -reductases (*baiJ*) between *Clostridium scindens* VPI 12708 and *Clostridium hylemonae* TN271.

delta412708	MAHYVPGAYEGIGRGYRGRLIVNVTVTEDEIEKIQIVKHKEVRGLAWDLPTSPIEVIPPQ	61
delta4TN271	MAHYVPGTYEGIGRGYQGRLIVNVTVTEDEITEWKIIVKHKEVRGLAWDLPTSPVEVMPPQ	61
delta412708	IIIEYQSLNIPLVNGADLTSAAILDAVAALKAAGATDEDEIEQLRCAPGPEAPEPKDEVRT	120
delta4TN271	IVKYQSLNIPLVNGADLTSGAILEAVEGALRAAGADDVIEKLKAAPGPEAPEWKDEVRT	120
delta412708	VDVAVFGAGAGGLAAAIEAKEGGADVILIEKQGITGGSTARSGGKLLGAGTKWQKEQGIY	180
delta4TN271	VDVAVFGAGAGGLAAAIEAKEGGADVILIEKQGITGGSTARSGGKLLGAGTKWQKKQGIY	180
delta412708	DTKEHCYDYLMEVGNRRGDFMDASKNRYLVEHLNETLDWLGTMGYQVQDVEAIHVSLOPW	240
delta4TN271	DTEHCYNYLMEVGNRRGNFMDASKNRYLVKNLNETLDWLGTMGYKVQDVEAIHVSLOPW	240
delta412708	RVHNSMGGGGQTNGQGGEITTPLTRHHVDEKLGGELIYNTALKELLTDENGTVNGAVCEKL	300
delta4TN271	RVHNSMGGGGQTNGQGGEITTPLTRHFAAGELGGELIYNTALKELLTDENGTVNGAVCEKL	300
delta412708	DGSKLTVYAKKGVILATGGYSRNKEMCARYPVAHYFSTTPKSNVGEGLIAAEKIGARNFV	360
delta4TN271	DGSKLTVYAKKGVILATGGYSRNKEMCARYPVAHYFCTAPKSNVGEGLIAAEKIGARNFV	360
delta412708	HPGIQVVYTSLSGIGINDESGLIVNERGERVVNEWSYQYHVSDALAASGSNCGWYITSG	420
delta4TN271	HPGIQVVYTSLSGIGINDESGLIVNERGERVVNEWSYQYHVSDALAASGSNCGWYITSG	420
delta412708	DEPYSGVQYGFKAQAVEGTSRDKAADSIEELAAMIKCDPAVLRATFDRYSELVDRKGVEDDF	480
delta4TN271	DEPYGCVQYGFKAQAVEGTSRDKAADSIEELAAMIKCDPATLRATFDRYCELTARGEDEDF	480
delta412708	GKPSRFLHPINGPKYAALRLHPCVTVTFGGLETDMARVLDEGRPIPGLYAAGEVADTG	540
delta4TN271	GKPARFLHPIDGPKYAALRLHPCVTVTFGGLETDMARVLDEGRPIPGLYAAGEVADTG	540
delta412708	MFGTEYPTCGTSIGGALFYGRIAGRVASGQSML	573
delta4TN271	MFGTEYPTCGTSIGGALFYGRIAGRVASGQSML	573

deduced amino acid sequence suggests this gene encodes a 26.4 kDa polypeptide. In addition, a partial ABC transport protein was located 85 bp downstream of the *baiL* gene in *C. scindens*. This short intergenic distance suggests this gene may be found on the same operon. Indeed, we hypothesize that an additional ABC transporter will be involved in translocating DCA and LCA into the extracellular environment (Figure 8). A short 850 bp genome-walk (Figure 56), and a fourth genome-walk of 3 kb (Figure 57) resulted in completion of the ABC transport gene 1854 bp in length, encoding a deduced 70.3 kDa polypeptide. In addition, two ORF were located directly downstream of the ABC transport protein. BLAST analysis suggests this ORF encodes a flavoprotein with the highest similarity to guanine deaminases. The second ORF is predicted to encode a flavoprotein with greatest similarity to xanthine dehydrogenases and 4-hydroxybenzoyl-CoA reductases.

Figure 58 is a schematic representation of the putative reductive *bai* operons in both *C. scindens* VPI 12708 and *C. hylemonae* TN271. Some notable features include the similar organization of the *baiJKL* genes,

Figure 56. Third genome-walking PCR downstream (DS3) of *baiJ* gene product in *Clostridium scindens* VPI 12708. a.) Agarose gel of DS3 products using *DraI* (D), *EcoRV* (E), *PvuII* (P), *StuI* (S) restriction libraries of *C. scindens* VPI 12708 genomic DNA as template with adaptor primer AP1 and gene-specific primer 12708WDS3 (See Table 6). b.) "DN" and "SN" represent nested PCR amplification using nested adaptor primer AP2 and gene-specific primer 12708WDS3N (See Table 6). Attempts to clone product "SN" were unsuccessful. While "EN" was smaller in size, the specificity was high and attempted cloning yielded several positive clones. "M" represents 1 kb DNA ladder. b.) Schematic representation of TA cloning of GENECLAN purified "EN" product (arrow) into pCR8/GW/TOPO TA vector (2.8 kb) yielding 3.65 kb vector p12708D4ENDS3.



**b.**

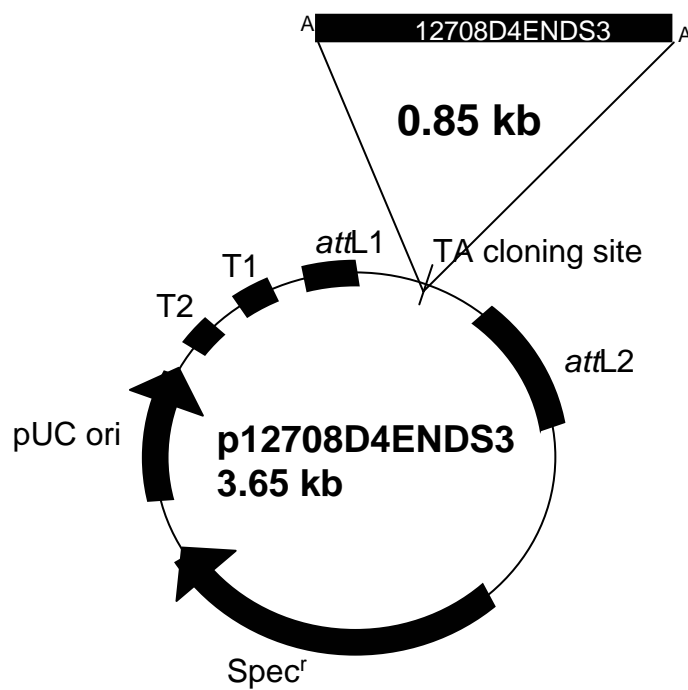
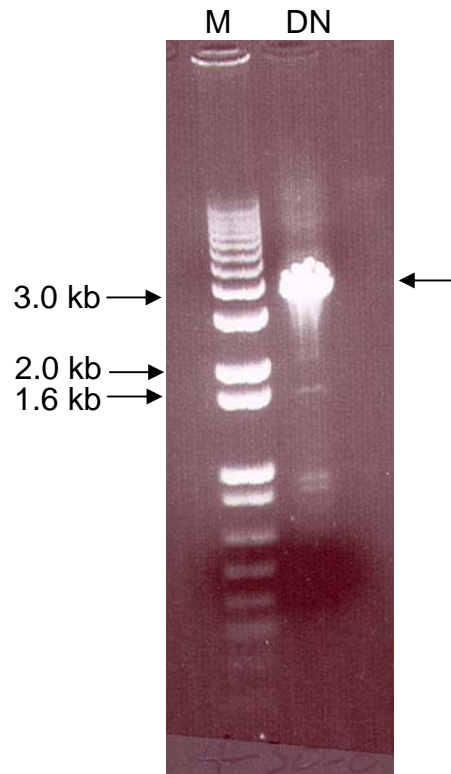




Figure 57. Genome-walking PCR downstream (DS4) of partial *baiJ* gene of *Clostridium scindens* VPI 12708. a.) Nested PCR products obtained from *DraI* (DN) restriction digest of *C. scindens* VPI 12708 genomic DNA template. b.) 3.0 kb product was GENECLAN purified from agarose gel and ligated into pCR8 GW TOPO TA vector resulting in vector p12708D4DNDS4 (5.8 kb).

a.



220

b.

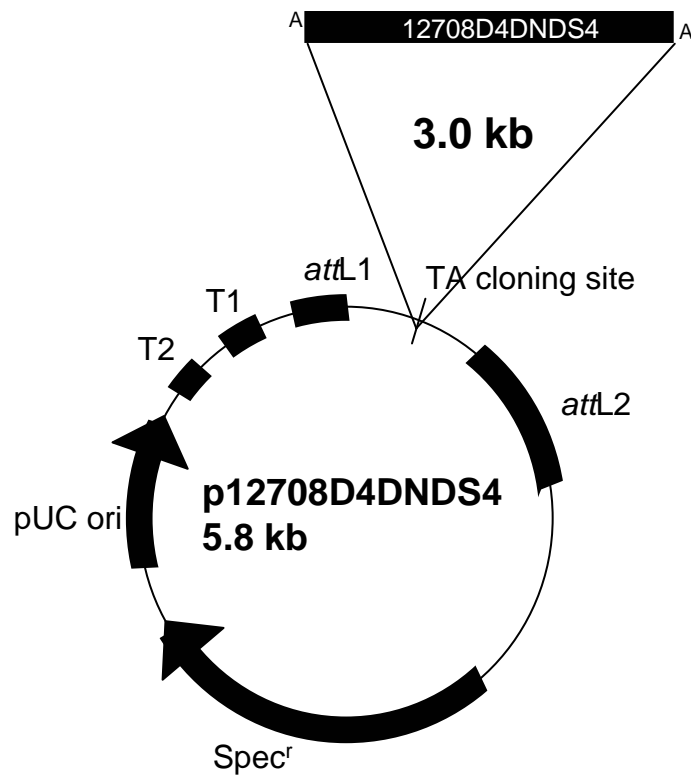
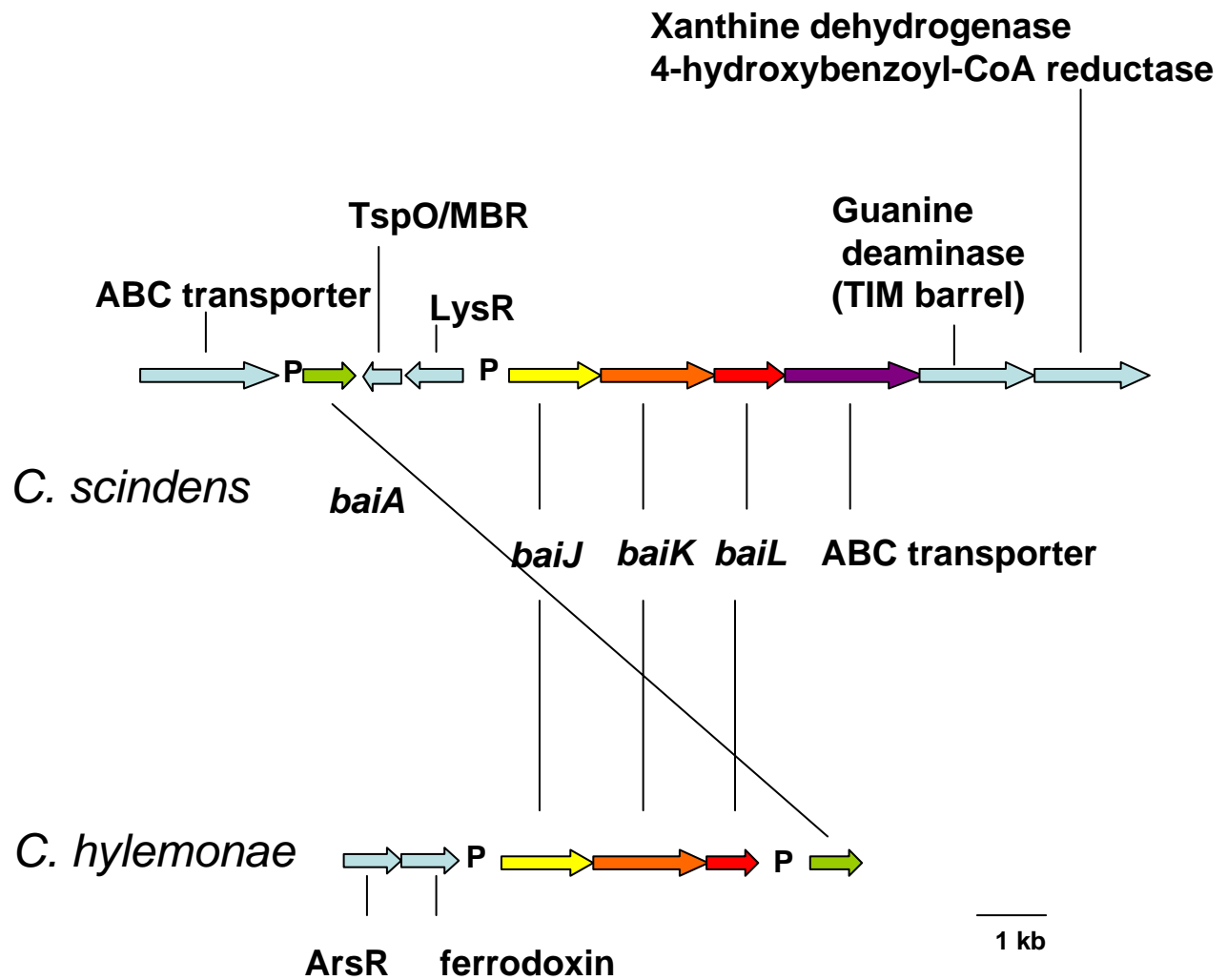


Figure 58. Schematic representation of putative reductive genes from *Clostridium scindens* VPI 12708 and *Clostridium hylemonae* TN271. The *baiA* genes encode 27 kDa 3 $\alpha$ -hydroxysteroid dehydrogenases. The *baiJ* genes encode 62 kDa flavoproteins. The *baiK* genes encode a predicted 49 kDa Type III CoA transferase similar to the *baiF*. The *baiL* genes encode 27 kDa proteins of the short chain reductase family. Finally, the *baiM* encodes a predicted ABC transport protein hypothesized to function as a secondary bile acid efflux protein. "P" represents conserved *bai* "promoter" region.



upstream conserved elements with the oxidative *bai* upstream region, and proximity of TIS. Interestingly, the upstream ORFs are not conserved between *C. scindens* and *C. hylemonae*. Each region includes an ORF predicted to encode a transcription factor of a different family of proteins on opposite strands. In addition, the ORF immediately upstream of the *baiJ* gene from *C. hylemonae* TN271 is predicted to encode an 18 kDa 4Fe-4S ferredoxin presumably involved in shuttling electrons. This ORF was not located adjacent to the *baiJKL* genes from *C. scindens*. Finally, the *baiA* gene and conserved upstream "promoter" appear to cluster around the *baiJKL* genes.

**Isolation of a partial *baiJ* gene in *Clostridium hiranonis* sp. strain T0931**

The *baiJ* gene from *C. hylemonae* TN271 and *C. scindens* VPI 12708 are very similar at the amino acid sequence level. Designing oligonucleotides for isolation of the same gene from more distantly related species can be difficult with such high sequence identities. Attempts were made to isolate the *baiJ* gene from *C. hiranonis* using

redundant primers in Table 9 without success. Therefore, a multiple sequence alignment was prepared with the amino acid sequence of the *baiJ* gene product against several homologues in the databases. This way, regions could be detected that would normally be highly conserved within this family of flavoproteins. Two gene products were selected for the alignment: 1). ZP\_01362869.1, a fumarate reductase flavoprotein subunit precursor from *Clostridium* sp. OhILAS (2) ZP\_00802728.1, a Flavocytochrome C from *Alkaliphilus metalliredigenes* QYMF. These protein sequences were then aligned with the 12708 *baiJ* using ClustalW (Figure 59). Two regions of high amino acid sequence identity with low codon degeneracy spanning a 282 bp stretch within the *baiJ* gene were identified (Figure 59). A codon usage table was prepared using *bai* gene sequences from *C. hiranonis* sp. strain T0931 (accession no. AF210152) (Figure 60). These primers were used to isolate a 657 bp gene product corresponding to what appears to be a partial *baiJ* gene of *C. hiranonis* sp. strain T0931 (Figure 61). A genome-walking library will have to be constructed for *C. hiranonis* sp. strain T0931 to elucidate the remaining gene and flanking regions;

Figure 59. Boxshade of CLUSTALW alignment between the *baiJ* gene product from *Clostridium scindens* VPI 12708 and flavoprotein homologues. ZP\_01362869.1 represents the GenBank accession number for the fumarate reductase flavoprotein subunit precursor from *Clostridium* sp. OhILAs and ZP\_00802728.1 represents the GenBank accession number for the Flavocytochrome c from *Alkaliphilus metalliredigenes* QYMF. Arrows represent regions of high amino acid conservation and low codon degeneracy which will be used to design oligonucleotides to attempt to isolate the *baiJ* gene from more distantly related 7 $\alpha$ -dehydroxylating clostridia.

ZP\_01362869.1 MKKTQKILSVLLQFIMMLSFVGC3KPASNGGFGALPKAGTSTAKAKGLGDUUEDEMILTE 61  
 ZP\_00802728.1 -----SREIPILMLGCQ-3GPTTETTDEALQAGTYAGU3PGHSGEITMDTOME 51  
 12708b.aij -----MNNVUPCTREGICRGYQGLIDMTUTE 29

ZP\_01362869.1 TTIESVKKLSRSSTP---GISDPAIEQIPANIOASQGLKIDAVSGATUTSEAILAAVID 116  
 ZP\_00802728.1 MEIQSONILERMESQ---GISDPAIDRIPNADIDGQTLANDAISGATITSDAILAAVIN 106  
 12708b.aij DEITEVKIKKRPURGLAWDEPTSPDEUAPPQIOKYOGLNIPLONGADLTSGAILEADEG 88

ZP\_01362869.1 CLTQAGAN---IDELKTKG-LKAQOLEDETIEDTUUVUGSGIAGLSAAVSAAEBCAKOVL 172  
 ZP\_00802728.1 ALBQAGAN---IDALSOKGENVURUGGETOEYITDOUIIGGGGAGLAAAVSAHQMGQOIIV 163  
 12708b.aij ALRAAGSANDVIEKLKAAPGPEAPEOKDEORTVDVADFGAGAGELAAIEAKEGGDUVL 148

ZP\_01362869.1 LEKMGSGGASINC GGELAGGSEMQAAGIEDTAEGLKQYULGCKGK---INEEIIISF 229  
 ZP\_00802728.1 LEKMPRLGGMTILAGGALMAAGSPROAAGIEDSOKKHFTQTYEGGDEL---GATDILIQT 220  
 12708b.aij IEKQGITGGSTARS GGKLLGAGTKWQKKQGIIYITEEMCYNVLMVCGMERGNFMDASKNRY 208

ZP\_01362869.1 IAKKSADTOKMLAERGOKPRNUTFSYYPTQSOYR---MNTETFTSGADFILPLQEKPK 285  
 ZP\_00802728.1 LVENAYPADEMLESGLAMENDEUFTOLG--GLWPR---EKKPSFPLGTGFDITYQNIIN 274  
 12708b.aij LVKMLNSTLDWLGTMGYKQVDVEAIHQSMQPWRQHN3MGCGGQNGQGGEITTPLTRHFA 268

ZP\_01362869.1 S-LGUEFRFETP8U3LINDKGTIKG--U08KN8G8TUTONAKS-TILATGGYANKRELVE 341  
 ZP\_00802728.1 ANEGIEILLDTAEIHELVEDGRUUG--UKAAGLSGDUIOHNMNGVIMAAAGGFGAMTEHIE 332  
 12708b.aij GRLGCEIINYNTALKELTIDENGUONGAVCEKIDG3ELTOYAKKGVILATGGYSRMEHCA 328

ZP\_01362869.1 QY--MPNKGKCSALGEALNGDGLINAREUGAEI0ANGGCI0APNDLGSTRYUD----- 393  
 ZP\_00802728.1 KYNQ3WPS8ANUKSTNPPGATGGLAI AEAUGASLIGLEHLQLLPLGDPMTGS3GNIEQ 392  
 12708b.aij RY----PVARYCNAPKS-NUGELG8AEEKIGARNFORPCIQ0VYT3L3CGIGIND---- 379

ZP\_01362869.1 --AGGIFLN-----UTPECKRFANENEYAFKRA8TLVHDLGFSNYFAITD8KLDN---- 441  
 ZP\_00802728.1 GVENRIFUNKSGNRIFUDEGARDUMTKALLEED8FLVUULDQKTYPTPETTNMFM---- 448  
 12708b.aij --ESGLI0N-----BRGERUUNWSYQ8HUS08LAASGSNCGWYITS8DEPYG6UQYG 429

→

ZP\_01362869.1 ENDEKADPMGTAFKADTIEEL8KFIGAKPEILKATOTRYNGF8PTGK0EDFGKP8UGPVG 501  
 ZP\_00802728.1 ETIEELI8QDRAFKGDTIEEL8GKIGONP8MLUKADEEHMKUVE8GGSD8FGRTL-FD-- 505  
 12708b.aij FKQAVEGT8RD-KARDSIEEL8ALIKCDPAT8RATFDRYQELT8G8ED8DFGKP8RT--- 486

←

ZP\_01362869.1 NREDNLKELQLLMA0Q8PYPAIAFNTTG0TGTFGGPKIDMNAQIISTT8DUIPGLYAAG 561  
 ZP\_00802728.1 -----TKIDT8PY8G8R0PT-0HHTMG8IEIM8VE8LD8NC8DUIPGLYAAG 552  
 12708b.aij -----LHPIDGPKY88RLHPC-0TV8FG8LETDI88RULDNE8RP8IPGLYAAG 534

ZP\_01362869.1 EV8NGELFYEQYEC8G88IQMSTRMGIIQ8GK888888-- 598  
 ZP\_00802728.1 EUTGG--IHG8NRLGGNAL8D8AVYGR8AG888888-- 587  
 12708b.aij EV8DTGMD8TEYETCGTSIG88FYGR8AG88888888 573

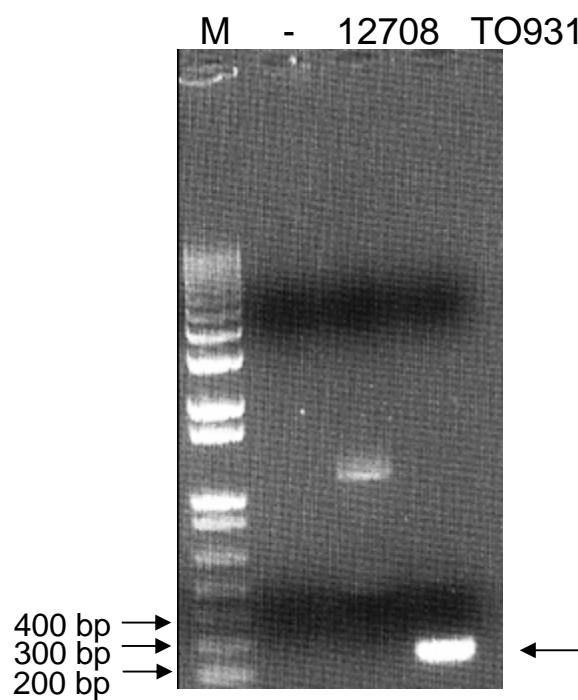


Figure 60. Codon Usage Table for *Clostridium hiranonis* sp. strain T0931. The CodonW program was used with *bai* gene sequences from *C. hiranonis* sp. strain T0931 as input (GenBank accession no. AF210152).

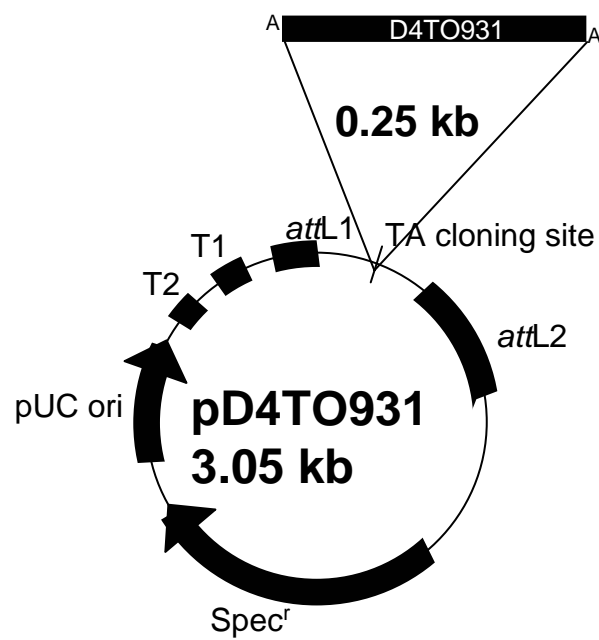
Phe	UUU	34	0.62	Ser	UCU	34	1.73	Tyr	UAU	27	0.71	Cys	UGU	30	1.20
	UUC	76	1.38		UCC	0	0.00		UAC	49	1.29		UGC	20	0.80
Leu	UUA	115	3.73		UCA	40	2.03	TER	UAA	5	2.50	TER	UGA	0	0.00
	UUG	1	0.03		UCG	0	0.00		UAG	1	0.50	Trp	UGG	28	1.00
	CUU	39	1.26	Pro	CCU	31	1.13	His	CAU	25	0.98	Arg	CGU	18	1.27
	CUC	0	0.00		CCC	0	0.00		CAC	26	1.02		CGC	0	0.00
	CUA	29	0.94		CCA	79	2.87	Gln	CAA	2	0.08		CGA	0	0.00
	CUG	1	0.03		CCG	0	0.00		CAG	50	1.92		CGG	0	0.00
Ile	AUU	8	0.13	Thr	ACU	63	1.87	Asn	AAU	44	0.88	Ser	AGU	27	1.37
	AUC	14	0.22		ACC	0	0.00		AAC	56	1.12		AGC	17	0.86
	AUA	168	2.65		ACA	72	2.13	Lys	AAA	146	1.78	Arg	AGA	67	4.73
Met	AUG	99	1.00		ACG	0	0.00		AAG	18	0.22		AGG	0	0.00
Val	GUU	109	2.29	Ala	GCU	108	2.10	Asp	GAU	84	1.34	Gly	GGU	103	1.84
	GUC	0	0.00		GCC	2	0.04		GAC	41	0.66		GGC	4	0.07
	GUA	76	1.60		GCA	89	1.73	Glu	GAA	162	1.88		GGA	115	2.05
	GUG	5	0.11		GCG	7	0.14		GAG	10	0.12		GGG	2	0.04

Figure 61. Isolation of putative *baiJ* gene from *Clostridium hiranonis* sp. strain T0931. a.) Agarose gel of PCR products derived from amplification with primers T0931D4F (5'-GCW GCW GAY TCW ATA GAA GAA TTA GC-3'), and T0931D4R (5'-TTC WCC WGC WGC TAA WCC WGG TAT-3') against (-) no DNA, (12708) *Clostridium scindens* VPI 12708 genomic DNA, (T0931) *Clostridium hiranonis* sp. strain T0931. A single PCR product was obtained from *C. hiranonis* sp. strain T0931. b.) PCR product (arrow) was cloned into a pCR8/GW/TOPO TA vector to obtain a 3.05 kb vector pD4T0931. c). Boxshade alignment of multiple sequence alignment of *C. hylemonae* TN271 *baiJ* (TN271baiJ), *C. scindens* VPI 12708 *baiJ* (12708baiJ) and partial sequence of *C. hiranonis* sp. strain T0931 *baiJ*. Asterisks signify amino acids corresponding to gene region where primers anneal.

a.



b.



c.

	*****	
TN271b <i>aiJ</i>	KAADSIEELAALIKCDPATLRATFDRYBELTAGGEDDFGKPARFL	487
12708b <i>aiJ</i>	KAADSIEELAAMIKCDPAVLRAFDRYBELTAGGEDDFGKPSRFL	487
TO931 <i>delta4</i>	LAADSIEELAEKIDUKPEULKATODRYNELCDAGGEDDYETPASEM	46
	*****	
TN271b <i>aiJ</i>	HPIDGPKYAALRLHPC'UTVTFGGLETDSAROLDNEGRP IPGLYAAG	534
12708b <i>aiJ</i>	HPIDGPKYAALRLHPC'UTVTFGGLETDSAROLDTEGRP IPGLYAAG	534
TO931 <i>delta4</i>	RAIRGDIYVAVFLRPATSVTFGGLETDSAROLDNMKKI IPGLYAAG	94

however, this data represents a significant step in possibly recovering these putative reductive genes from a third 7 $\alpha$ -dehydroxylating clostridial species present in the human colon.

### **Cloning and Overexpression of *baiL* gene in *Escherichia coli*.**

The *baiL* gene is predicted to encode a 26 kDa polypeptide in the 3-ketoacyl-acyl carrier protein reductase (*fabG*) family (Figure 62). We predict this protein will convert 3-oxo-DCA(~SCoA) and/or 3-oxo-LCA(~SCoA) to the secondary bile acids DCA(~SCoA) and LCA(~SCoA), respectively. Therefore, the strategy for characterizing this gene product was to clone the *baiL* gene from *C. hylemonae* TN271 into a pET21(a+) vector, overexpress the recombinant protein in *E. coli* codon plus BL21(DE3) and test 3 $\alpha$ -hydroxysteroid oxidoreductase activity with substrates NAD(P)H and [24-<sup>14</sup>C] 3-oxo-DCA or NAD(P) and [24-<sup>14</sup>C] DCA (See Material and Methods).

After cloning (Figure 63) and sequencing of the vector insert to verify proper reading frame, the *BaiL* was overexpressed in *E. coli* (Figure 64a). The *BaiL* protein

Figure 62. Boxshade of CLUSTALW amino acid sequence alignment of *baiL* genes from *Clostridium scindens* VPI 12708 and *Clostridium hylemonae* TN271 with similar short chain dehydrogenases/3-keto acyl reductases.

ZP\_01696223.1 -----MALLDGRVAVUTGAASGMCKAIALFAKEGAKVVASDLNLESUQQTJAEIDAAAG--GKALAVQAMQASBDDV  
 ZP\_02688472.1 -----MGLITGRVAVUTGAASGMCKQIALFAKEGAKVVASDLNLEAAQKTQELQEKED--GTGLAVVAMTKQED I  
 ZP\_01373200.1 -----MALLMNVAVUTGAASGMCKSIALLFAKEGAKVVASDLNLECAHKVAEEITSAAG--GTALAKTNQAVTED I  
 ZP\_01722898.1 -----MELEGKVAIUTGAASGMCKAIAEGFAKEGAKVVASDLNLDCAQAQDEGQQAAG--GTATATQTNQASTED I  
 ZP\_02463417.1 -----MUDLSGRVAVUTGGEIGIGIAISRTLSAFAGARTIIGGILDEAGEKAARDITEEG--RQATFQKTDURFEDQ  
 ZP\_02629629.2 MSKIIDGIFSDKVMITGAGGCI GKAAALRFAKEGAKVVASDLNLEADPKKEMSKETLKEDEIT--SNVEILICDREGEMC  
 TN271b.a1l -----MEGAKVVASARRLDRLEATRAEDMEGGGAGSILBLQTDUTUPKQI  
 12708b.a1l -----MEGGKVVASARRLLEELKAEDEAAKGAGCILBLQTDURIPKQI

ZP\_01696223.1 QHLIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 ZP\_02688472.1 PMINQKQAFGLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 ZP\_01373200.1 QHLIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 ZP\_01722898.1 QHLIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 ZP\_02463417.1 EHLIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 ZP\_02629629.2 KKVIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 TN271b.a1l DRVIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG  
 12708b.a1l DRVIDTAKKENGTLIDLUMNAGIMINAGPAGDLTDALWEKVFQ-INTTTPMRTIEKSLPIFIEKGN-GVIUMIASAG

ZP\_01696223.1 GLNGSRAGAAAYTASKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 ZP\_02688472.1 GLNGSRAGAAAYTASKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 ZP\_01373200.1 GLNGSRAGAAAYTASKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 ZP\_01722898.1 GLNGSRAGAAAYTASKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 ZP\_02463417.1 GLNGSRAGAAAYTASKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 ZP\_02629629.2 GLNGSRAGAAAYTASKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 TN271b.a1l 3IRGLKGGTTTATKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---  
 12708b.a1l 3IRGLKGGTTTATKHAUIGLTKVUGHQYULAGIRCNAIAPGGQETNIGTTMSEPN---AGLERAMSGSINP---

ZP\_01696223.1 -----RSGQEEIATIALFLASDQAS-FUNGAVITADGGWTAY 253  
 ZP\_02688472.1 -----RSGQEEIATIALFLASDQAS-FUNGAVITADGGWTAY 253  
 ZP\_01373200.1 -----RSGQEEIATIALFLASDQAS-FUNGAVITADGGWTAY 252  
 ZP\_01722898.1 -----RSGQEEIATIALFLASDQAS-FUNGAVITADGGWTAY 252  
 ZP\_02463417.1 YKRMVPAARRKGLQEVADUAFILASDQAS-KVSGQ2IADGGWTAT 263  
 ZP\_02629629.2 GKUTSPQMRVARRKGLQEVADUAFILASDQAS-KVSGQ2IADGGWTAT 272  
 TN271b.a1l PUEGAKPIL-GQPEDCASITIALFLASDQAS-KVSGQ2IADGGWTAT 243  
 12708b.a1l HTEGAKPIL-GQPEDCASITIALFLASDQAS-KVSGQ2IADGGWTAT 243

Figure 63. Cloning strategy of *baiL* gene from *C. scindens*  
VPI 12708.



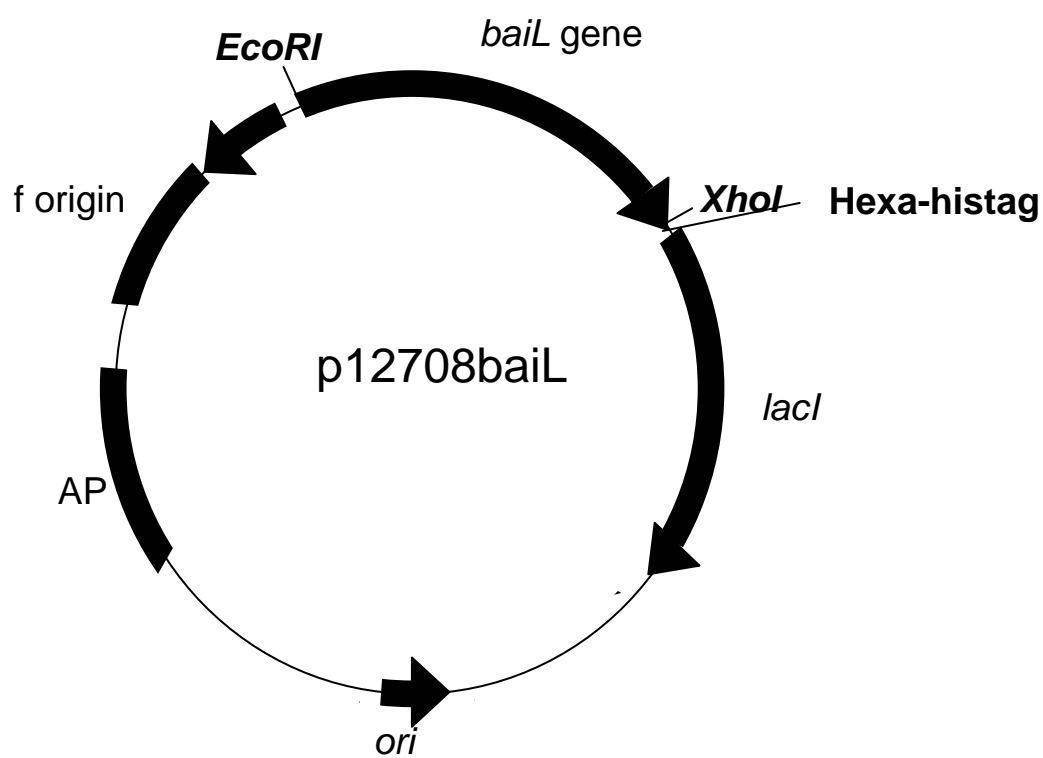
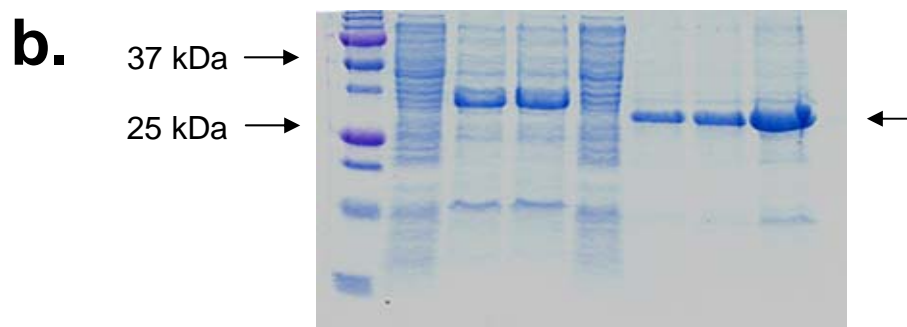
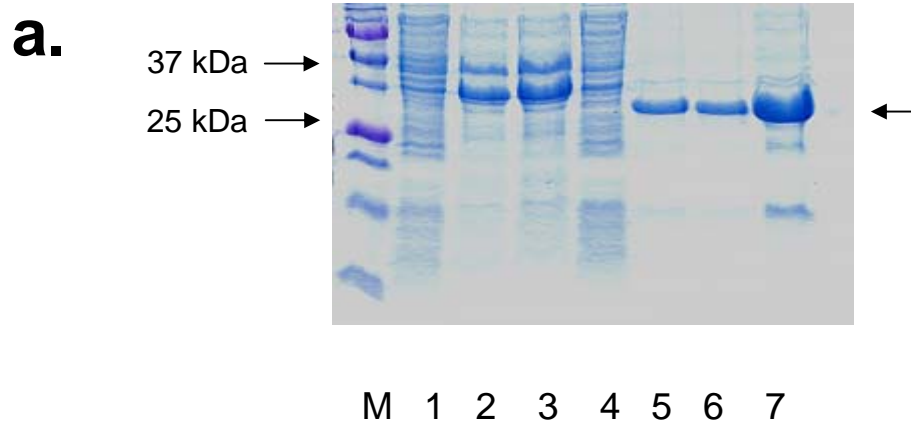


Figure 64. Overexpression of recombinant *baiL* gene from *Clostridium hylemonae* TN271 in *E. coli* BL21 (DE3). "M" stands for Kaleidograph protein standard, Lane 1= 30 µg soluble protein from IPTG uninduced *E. coli*, Lane 2= 25 µg insoluble protein from IPTG uninduced *E. coli*, Lane 4= 30 µg soluble protein from 1 mM IPTG induced *E. coli*, Lane 5= 10 µg insoluble fraction 1 mM IPTG induced *E. coli*, Lane 6= 15 µg insoluble fraction IPTG induced *E. coli*, Lane 7= 30 µg insoluble fraction 1 mM IPTG induced *E. coli*. a). Cultivation in LB broth b). Cultivation in LBS broth (See Text).

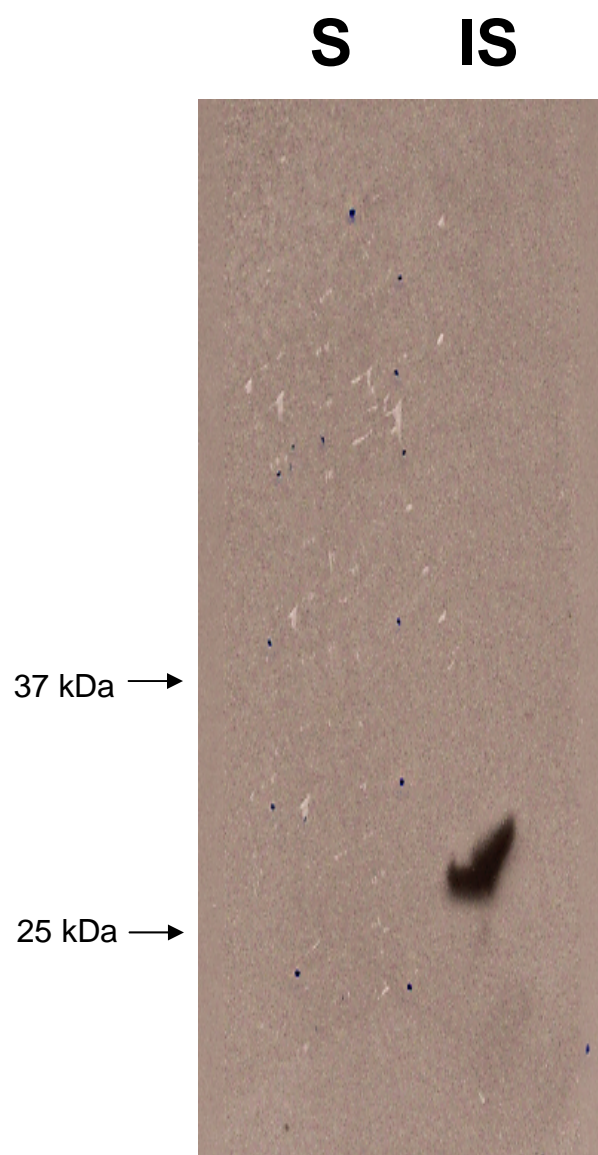


was highly expressed, though found in inclusion bodies. Several attempts were made with different colonies, ranges of IPTG concentration (0.1-1mM), temperatures (18°C to 37°C) and media (LB, TYGPN). Indeed, we attempted the sorbitol/betaine method (medium LBS) of Blackwell and Horgan (1991) although results were similar between LB and LBS (Figure 64b) (193). We verified expression of His-tag using anti-hexahistidine antibodies in the insoluble, but were unable to detect any soluble recombinant BaiL-His (Figure 65).

The *baiL* gene was also cloned in a pET24(a+) vector as well as a pSport1 vector from both the SP6 and T7 promoters and overexpressed. In each case, the BaiL was located in the insoluble fraction. Attempts at refolding by solubilization in 8M urea followed by step-down buffer exchanges with urea in the presence of DTT and arginine were unsuccessful in producing soluble protein.

**Cloning and overexpression of the *baiJ* gene product in *Escherichia coli*.**

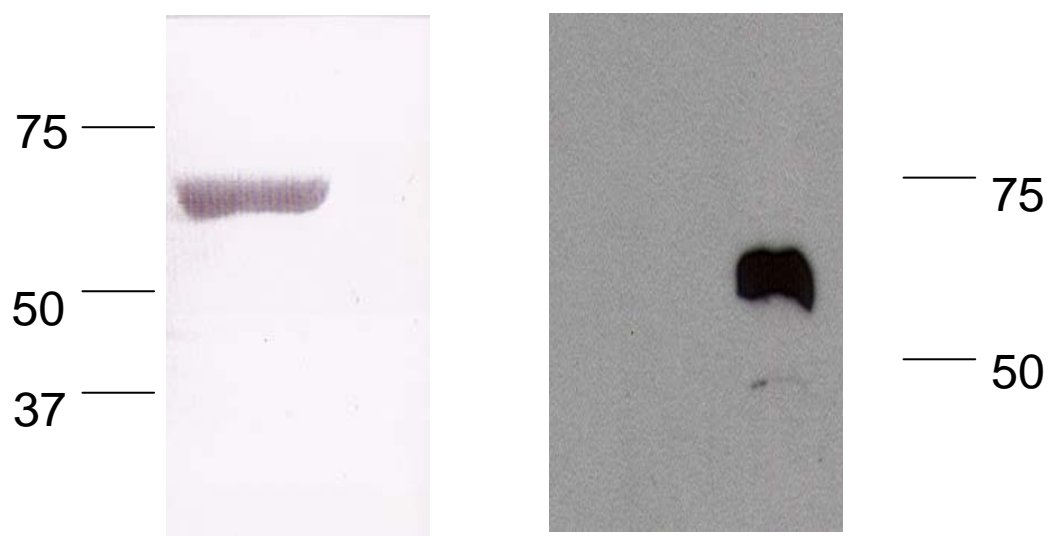
Figure 65. Western blot analysis of *baiL* expression in soluble and insoluble fractions of IPTG induced *E. coli* cell extracts. Protein samples were electrophoresed on an SDS-PAGE gel (12%) and electroblotted onto nitrocellulose. The blot was probed with anti-hexahistidine antibody. Lane S=soluble fraction from 1 mM IPTG induced *E. coli* whole cells; Lane IS= insoluble fraction solubilized in 8.5 M urea. Each lane contains 20 µg protein. Expected size of recombinant *baiL* gene product is 27 kDa. Position of molecular weight markers shown to left of blot.



The *baiJ* gene is predicted to encode a 3-oxo- $\Delta^4$ -DCA(~CoA) oxidoreductase. To test this hypothesis, the *baiJ* gene was PCR amplified with primers 12708baiJ-F (5'-TGCATCTCC**CTCGAG**GGCACATTATGTTCCAGGC-3') and 12708baiJ-R (5'-TAC**AGGTACCA**AGCATTTGACTGTCCCGATGCT-3'). Double digestion of the pT7-SBP-2 expression vector and PCR product with *XhoI* and *KpnI* restriction endonucleases and ligation resulted in the pT7SB2baiJ vector (Figure 13). The vector was initially transformed into *E. coli* DH5 $\alpha$  and subsequently cloned into expression host *E. coli* ArcticExpress BL21(DE3). After verifying the sequence and reading frame of the insert, the BaiJ gene was overexpressed by IPTG induction (0.8 mM). The ArcticExpress strain was chosen so that growth and thus expression could be slowed to allow for less inclusion body formation which was observed for the *baiL* gene. Figure 66 shows the purification and immunoblot detection of the recombinant BaiJ gene product. The strep-tag was chosen based on results with the BaiH protein which lost activity following Ni<sup>+</sup> affinity chromatography. The strep-tag recognizes biotin, which should not result in chelation of metals or flavins. The BaiJ polypeptide was

Figure 66. SDS-PAGE and Western immunoblot of elution fraction after biotin binding chromatography of overexpression of recombinant BaiJ-SBP fusion protein. 10  $\mu$ g protein was loaded onto a 12% polyacrylamide gel. MW (kDa) marker represented to left of SDS-PAGE (a) and right of immunoblot (b). Expected size of recombinant fusion protein: ~66 kDa.

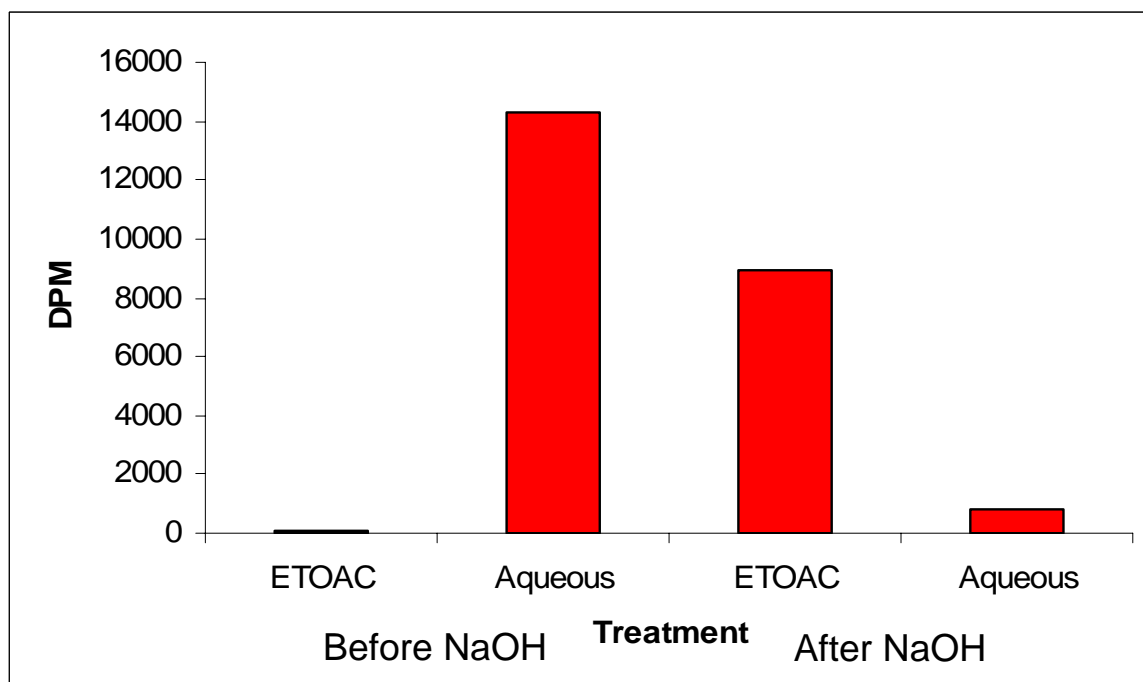




purified to about 95% based on visual inspection of SDS-PAGE of eluent fraction. The eluent was concentrated by centrifugation on Centriprep YM10 (10 kDa cut-off) concentrators. The concentrated material was yellow in color; indicative of bound flavins such as observed for the BaiCD and BaiH proteins (121).

Synthesis of bile acid CoA thioester was tested by differential solubility. CoA esters of bile acids are water soluble at low pH, while free bile acids precipitate and can be extracted into organic solvents; particularly relatively hydrophobic bile acids such as 3-oxo-DCA. 30  $\mu$ l CoA conjugate sample (90  $\mu$ M; 25,000 DPM) was added to 500  $\mu$ l reaction buffer (100 mM sodium phosphate, pH 6.8) and the pH was lowered to 3 with 1N HCl. The sample was extracted three times with two volumes ethyl acetate. The ethyl acetate fraction and 10  $\mu$ l aqueous phase were counted by liquid scintillation. The aqueous phase was then brought to pH 10 with 5N NaOH and boiled 30 min. The pH was then lowered to 3 with 10 N HCl and extracted three times with two volumes ethyl acetate. Again, the ethyl acetate fraction and the aqueous phase were counted. Figure 67 shows that radioactivity is extracted only after

Figure 67. Measurement of radioactivity in both aqueous and organic phase before and after sodium hydroxide treatment. This assay was used to provide evidence for the presence of a CoA conjugate based on relative solubility. Ethyl acetate (ETOAC) extraction "before NaOH" represents organic extraction of the aqueous phase containing bile acid CoA conjugates. The pH was lowered to 3 with 1N HCl. Free bile acids are expected to partition to the organic phase, while conjugates are expected to remain in the aqueous phase. "After NaOH" represents the DPM value in the organic and aqueous phase once the aqueous phase from the first extraction was increased to pH 10 with 5 N NaOH, boiled for 30 min, and lowered to pH 3 with 5 N HCl. Extraction of the aqueous phase resulted in the majority of radioactivity partitioning in the organic phase, suggesting hydrolysis of CoA thioester.

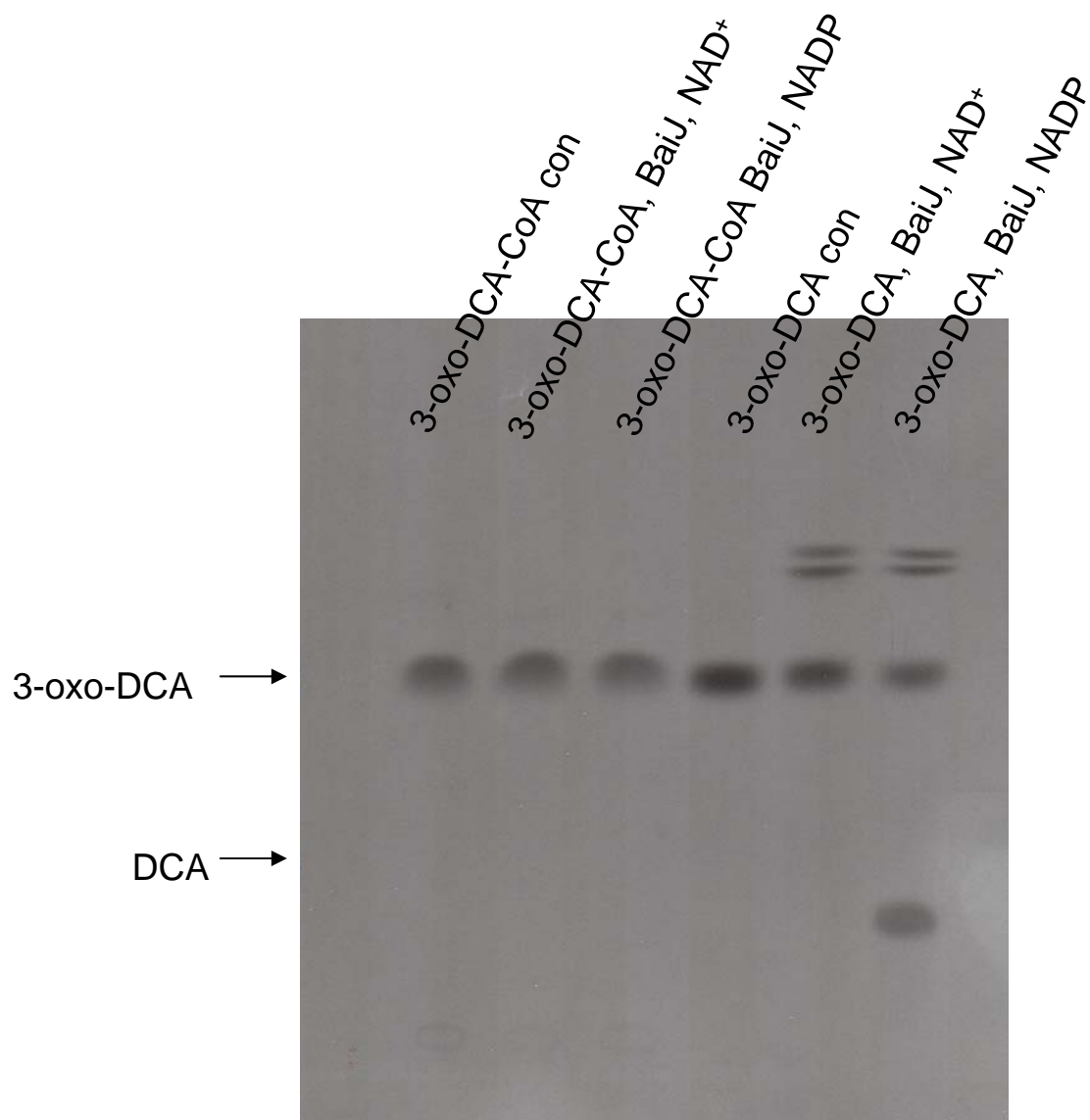


boiling at pH 10, suggesting that the CoA thioester is being hydrolyzed.

Initial enzyme assays were run for 30 min at 37°C in a 500 µl volume of 100 mM sodium phosphate buffer containing 90 µM [24-<sup>14</sup>C] 3-oxo-deoxycholy-CoA (or the free acid) (550 DPM/nmole), 150 µM NAD(P)(H) and 30 µg purified BaiJ. When 3-oxo-DCA and NADP were substrates, a product of greater hydrophilicity than the substrate was obtained whose  $R_f$  value was even lower than the commercial DCA standard (Figure 68). This strongly suggests the 3 $\alpha$ -hydroxyl group is reduced in addition to an oxidation reaction. Additional characterization will be required to determine the function of this enzyme. In summary, the 62 kDa BaiJ gene was expressed and purified and initial enzyme assays suggest this polypeptide metabolizes bile acids; however, this enzyme does not appear to have 3-oxo- $\Delta^4$ -DCA oxidoreductase activity.

**Isolation of the *gyrA* gene: housekeeping gene for normalizing RT-PCR in *C. scindens* and *C. hylemonae*.**

Figure 68.      Autoradiograph of a thin layer chromatography of BaiJ reaction products. Standards (not shown) are represented by arrows. 3-oxo-DCA and 3-oxo-DCA~SCoA (90  $\mu$ M; 25,000 DPM) were used as substrates in a 500  $\mu$ l reaction volume. The reaction buffer, 100  $\mu$ M sodium phosphate (pH 6.8), containing 150  $\mu$ M pyridine nucleotide (labeled) and 30  $\mu$ g purified recombinant BaiJ-SBP.



Future characterization and studies of *bai* gene expression by RT and qRT PCR requires normalization using "housekeeping genes". Several studies have used the *gyrA* subunit in this capacity (194, 195). At the time of study there was no genomic data for any 7 $\alpha$ -dehydroxylating bacterium in the databases, which has since changed (currently *Clostridium scindens* VPI 12708 and *Clostridium scindens* ATCC 35704 by our lab and Jeffrey Gordon's lab, respectively). The strategy for isolating the *gyrA* subunit from *Clostridium hylemonae* TN271 was to generate a multiple sequence alignment of several *gyrA* genes from various species of *Clostridia* in addition to two *gyrA* genes from *Alkaliphilus oremlandii* OhILAs. After a multiple alignment was generated and searched for regions of low codon degeneracy based on a codon usage table generated for *C. hylemonae* TN271 (Figure 69) two regions of high sequence identity and low codon degeneracy were identified. Figure 70 illustrates these regions in Logo form along with redundant oligonucleotide variations based on these sequence regions. PCR amplification using sixteen combinations of forward and reverse redundant oligonucleotides resulted in five positive reactions

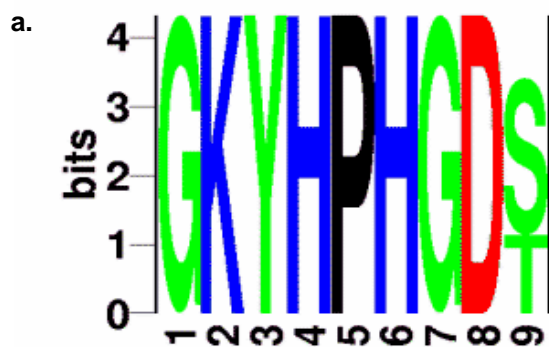


Figure 69. Codon usage table for *Clostridium hylemonae* TN271. The CodonW program was used to generate table using *baiJ*, *baiK* and *baiL* gene sequences as input.

Phe	UUU	25	1.02	Ser	UCU	9	0.79	Tyr	UAU	39	1.32	Cys	UGU	10	0.83
	UUC	24	0.98		UCC	23	2.03		UAC	20	0.68		UGC	14	1.17
Leu	UUA	5	0.27		UCA	10	0.88	TER	UAA	1	0.75	TER	UGA	1	0.75
	UUG	8	0.43		UCG	5	0.44		UAG	2	1.50	Trp	UGG	18	1.00
	CUU	32	1.73	Pro	CCU	12	0.76	His	CAU	19	1.41	Arg	CGU	16	1.22
	CUC	26	1.41		CCC	7	0.44		CAC	8	0.59		CGC	23	1.75
	CUA	0	0.00		CCA	6	0.38	Gln	CAA	4	0.19		CGA	1	0.08
	CUG	40	2.16		CCG	38	2.41		CAG	39	1.81		CGG	10	0.76
Ile	AUU	22	0.63	Thr	ACU	8	0.33	Asn	AAU	35	1.09	Ser	AGU	7	0.62
	AUC	61	1.76		ACC	18	0.74		AAC	29	0.91		AGC	14	1.24
	AUA	21	0.61		ACA	35	1.44	Lys	AAA	45	1.11	Arg	AGA	17	1.29
Met	AUG	43	1.00		ACG	36	1.48		AAG	36	0.89		AGG	12	0.91
Val	GUU	11	0.38	Ala	GCU	22	0.60	Asp	GAU	27	0.72	Gly	GGU	29	0.66
	GUC	35	1.21		GCC	40	1.10		GAC	48	1.28		GGC	55	1.25
	GUA	26	0.90		GCA	35	0.96	Glu	GAA	62	1.03		GGA	60	1.36
	GUG	44	1.52		GCG	49	1.34		GAG	58	0.97		GGG	32	0.73

1567 codons in baiAF2delta43HSD (used Universal Genetic code)

Figure 70. Design of degenerate oligonucleotides for isolation of the *gyrA* gene from *C. hylemonae* TN271. A CLUSTALW alignment was prepared from the *gyrA* gene from the following organisms: *Alkaliphilus oremlandii* OhILAs (ABW17577.1; EAT31502.1), *C. perfringens* str. 13 (NP\_560923.1), *C. beijerinckii* NCIMB 8052 (YP\_001307153.1), *C. tetani* E88 (NP\_780806.1), *C. acetobutylicum* ATCC 824 (AAK77994.1). Short stretches of conserved amino acids were examined manually within this alignment for low codon degeneracy. Two regions were found spanning 235 amino acids which would give a PCR product of 705 bp. a). Logo representation of conserved residues 77-84 (*C. tetani gyrA*) identified in multiple sequence alignment of *gyrA* genes as highly conserved and low in codon degeneracy. Lowest degenerate forward primers (*gyrAF1*, *gyrAF2*) are listed. b.) Logo representation of residues 299-307 (*C. tetani gyrA*) identified as highly conserved and containing residues with low codon degeneracy. The sense strand reverse translated is shown below amino acid residues (bold) with wobble based listed as "N". "Rev Comp" is the reverse complementary strand for reverse oligonucleotides (*gyrAR-gyrAR8*).

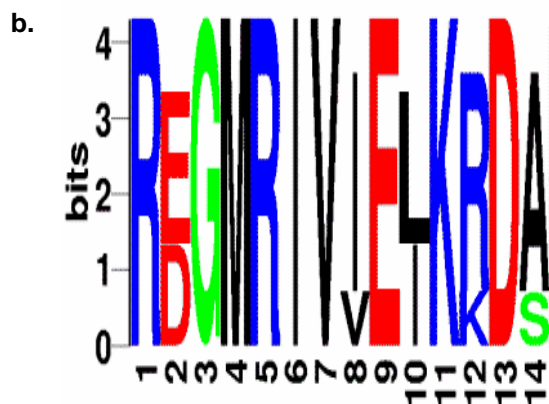


## Primer Name

## Sequence

G K Y H P H G D  
 5'-GGN AAN TAN CAN CCN CAN GGN GAN-3'

gyrAF1 GGV AAR TAY CAY CCG CAY GGV GAY  
 gyrAF2 GGV AAR TAY CAY CCH CAY GGV GAY



## Primer Name

## Sequence

R E G M R I C I E  
 5' CGN GAN GGN ATG CGN ATN TGN ATN GAN 3'  
 Rev 5' TCN ATN CAN ATN CGN ATN CCN TCN CGN 3'  
 Comp

gyrAR TCK ATR CAK ATS CGC ATK CCY TCS CG  
 gyrAR2 TCK ATR CAK ATT CTC ATK CCY TCT CT  
 gyrAR3 TCK ATR CAK ATT CTC ATK CCY TCS CG  
 gyrAR4 TCK ATR CAK ATS CGC ATK CCY TCT CT  
 gyrAR5 TCK ATS ACC ATS CGC ATK CCY TCS CG  
 gyrAR6 TCK ATS ACC ATT CTC ATK CCY TCT CT  
 gyrAR7 TCK ATS ACC ATT CTC ATK CCY TCS CG  
 gyrAR8 TCK ATS ACC ATS CGC ATK CCY TCT CT

(Figure 71). Cloning of PCR product derived from primer pair *gyrAF1/gyrAR4*, sequencing of this product and alignment of the deduced amino acid sequence suggests the *gyrA* gene has been isolated. Primers *gyrAF1* and *gyrAR8* were used to isolate the partial *gyrA* gene from *C. scindens* VPI 12708 (data not shown) which was cloned into a pCR8/GW/TOPO, sequenced and aligned with section of *gyrA* gene from *C. hylemonae* TN271. Following this data, the entire *gyrA* gene from *C. scindens* VPI 12708 was detected by local BLAST search of the 13X coverage of the *C. scindens* VPI 12708 genome (Ridlon JM, Alves J, Kang D, Hylemon PB and Buck G unpublished data). A multiple sequence alignment of the *gyrA* genes from *C. hylemonae* TN271, *C. scindens* VPI 12708 and *gyrA* genes from clostridial genome data available from GenBank (Figure 72). A set of RT-PCR primers were designed to detect the *gyrA* gene of *C. hylemonae* TN271 (*gyrATN271F* 5'-ACTTTGACGAGACAGAAAAGAACC-3'; *gyrATN271R* 5'-CAGTATATGTTTCGATCGTCGT) (primers used in Figure 26) and *C. scindens* VPI 12708 (*gyrA12708F* 5'-ATCTATGGAGCACTGGTAAATATGG-3'; *gyrA12708R* 5'-AGTTAGGAGCAAAATCAACCGTATC-3') which showed high

Figure 71. PCR amplification of *gyrA* gene from *Clostridium hylemonae* TN271. Combinations of oligonucleotides listed in Figure 65 were used against 50 ng *C. hylemonae* TN271 genomic DNA template. Combinations of primers used in PCR reaction are as follows: Lane M= 1kb DNA ladder; Lane 1= *gyrAF1*, *gyrAR*; Lane 2=*gyrAF1*, *gyrAR2*; Lane 3=*gyrAF1*, *gyrAR3*; Lane 4=*gyrAF1*, *gyrAR4*; Lane 5=*gyrAF1*, *gyrAR5*; Lane 6=*gyrAF1*, *gyrAR6*; Lane 7=*gyrAF1*, *gyrAR7*; Lane 8=*gyrAF1*, *gyrAR8*; Lane 9=*gyrAF2*, *gyrAR*; Lane 10=*gyrAF2*, *gyrAR2*; Lane 11=*gyrAF2*, *gyrAR3*; Lane 12=*gyrAF2*, *gyrAR4*; Lane 13=*gyrAF2*, *gyrAR5*; Lane 14=*gyrAF2*, *gyrAR5*; Lane 13=*gyrAF2*, *gyrAR5*; Lane 14=*gyrAF2*, *gyrAR6*; Lane 15=*gyrAF2*, *gyrAR7*; Lane 16=*gyrAF2*, *gyrAR8*. 750 bp PCR product from Lane 4 was cloned into a pCR8/GW/TOPO TA vector generating a ~3.5 kb vector p271gyrA.

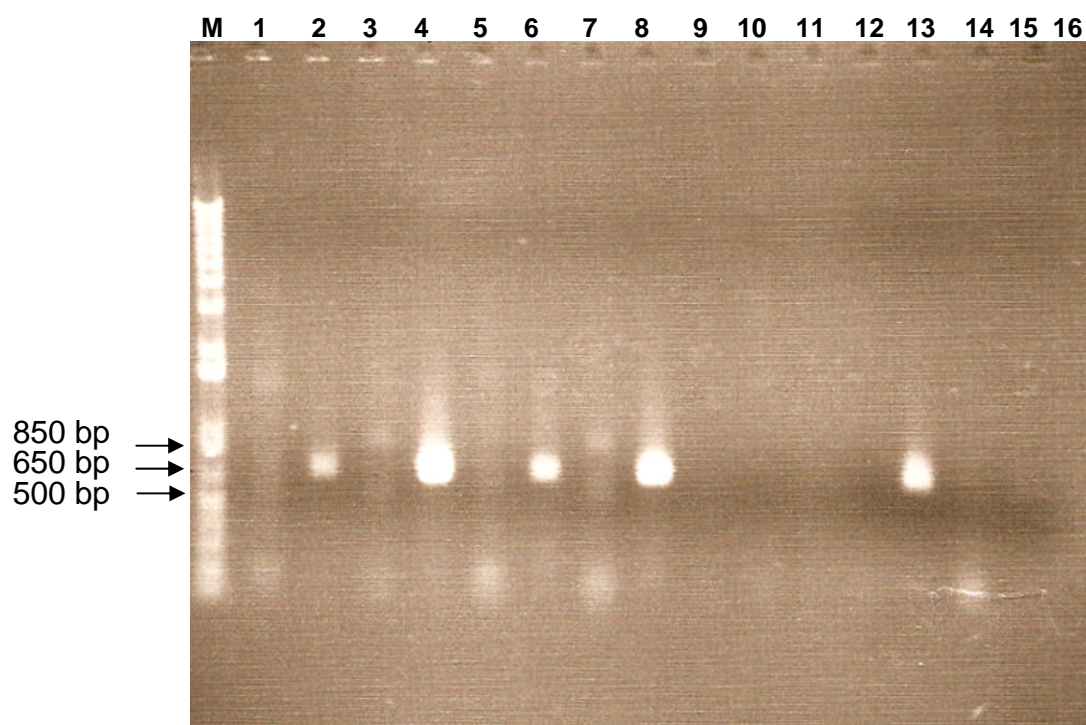


Figure 72. Boxshade alignment of *gyrA* gene product from *C. scindens* VPI 12708 with other clostridial and bacterial *gyrA* genes. *Alkaliphilus oremlandii* OhILAs (ABW17577.1; EAT31502.1 ), *C. perfringens* str. 13 (NP\_560923.1), *C. beijerinckii* NCIMB 8052 (YP\_001307153.1), *C. tetani* E88 (NP\_780806.1), *C. acetobutylicum* ATCC 824 (AAK77994.1).



ABW17577.1 ---MLEKRNIVPNIIEKEMKNSYIDYMSVIVGRALPDVRDGLKPVHRRILYSMEELGLTPEKPHKKSARIVGDVLGKYHHPHGDGA 85  
 EAT31502.1 ---MLEKRNIVPNIIEKEMKNSYIDYMSVIVGRALPDVRDGLKPVHRRILYSMEELGLTPEKPHKKSARIVGDVLGKYHHPHGDGA 85  
 NP\_560923.1 ---MSLNIEGKVIPIDINKEMKKCYIDYAMSVIVGRALPDVRDGLKPVHRRILYSMEELGLTPEKGYRKCARIVGDLGKYHHPHGDSS 85  
 YP\_001307153.1 ---WDFNIEGKVIPVDIKNEMKKCYIDYAMSVIVGRALPDVRDGLKPVHRRILYSMEELGLTPEKGYRKCARIVGDLGKYHHPHGDSS 85  
 NP\_780806.1 MNYINNIEGKVLEVDIKQEMKKCYIDYAMSVIVGRALPDVRDGLKPVHRRILYSMEELGLTPEKGYRKCARIVGDLGKYHHPHGDSS 88  
 AAK77994.1 ---MLNIEGKVLEVDISSEMKKKCYIDYAMSVIVGRALPDVRDGLKPVHRRILYSMEELGLTPEKGYRKCARIVGDLGKYHHPHGDSS 84  
 12708gyrA --MEDNIFIEKVLEVDIKKFMETS YIDYAMSVIASRALPDVRDGLKPVHRRILYSMEELANGPEKPHKKSARIVGDVLGKYHHPHGDSS 86

ABW17577.1 VYDAMVLAQDFSRYMLVDGHGNGFSVDGCAAAAMRYTEARLSKNAMEIRDINKNTVDVFPNFDPELKEPSVLPSPFPNLLVNGS 171  
 EAT31502.1 VYDAMVLAQDFSRYMLVDGHGNGFSVDGCAAAAMRYTEARLSKNAMEIRDINKNTVDVFPNFDPELKEPSVLPSPFPNLLVNGS 171  
 NP\_560923.1 VIDALVRMAQDFSRYMLVDGHGNGFSVDGDSAAAMRYTEAKMNKIAEMLRDINKNTVDVFPNFDGEKEPEVVLPSRFPNLLVNGS 171  
 YP\_001307153.1 VIDALVRMAQDFSRYMLVDGHGNGFSVDGDSAAAMRYTEAKMNKIAEMLRDINKNTVDVFPNFDGEKEPEVVLPSRFPNLLVNGS 171  
 NP\_780806.1 VIDALVRMAQDFSRYMLVDGHGNGFSVDGDSAAAMRYTEAKMCKTLEMRLDINKNTVDVFPNFDGEKEPEVVLPSRFPNLLVNGS 174  
 AAK77994.1 VYDALVLAQDFNRYMLVDGHGNGFSVDGDSAAAMRYTEAKMNKIAEMLRDINKNTVDVFPNFDGEKEPEVVLPSRFPNLLVNGS 170  
 12708gyrA TYDALVLAQDFNRYMLVDGHGNGFSVDGCAAAAMRYTEARLSKNAMEIRDINKNTVDVFPNFDPELKEPEVVLPSRFPNLLVNGS 172

ABW17577.1 NGIAVGMATNIPPNNLGEVIDCGVTLIDN----SPAQIEDLMEHIKGPDPFPTAGTIMCTEGIEAYITGRCRIRKRAKASIEEVNK 253  
 EAT31502.1 NGIAVGMATNIPPNNLGEVIDCGVTLIDN----SPAQIEDLMEHIKGPDPFPTAGTIMCTEGIEAYITGRCRIRKRAKASIEEVNK 253  
 NP\_560923.1 SGIAVGMATNIPPNNLGEVIDCGVTLIDN----PBLTLELMTYIKGPDPFPTAGTIMGKSIRAAAYETGRKIIIVRAKAEIEEE-N 253  
 YP\_001307153.1 SGIAVGMATNIPPNNLGEVIDCGVTLIDN----PBLTLELMTYIKGPDPFPTAGTIMGKSIRAAAYETGRKIIIVRAKAEIEEE-N 253  
 NP\_780806.1 SGIAVGMATNIPPNNLGEVIDCGVTLIDN----PBLTLELMTYIKGPDPFPTAGTIMGKSIRAAAYETGRKIIIVRAKAEIEEE-N 256  
 AAK77994.1 AGIAVGMATNIPPNNLGEVIDCGVTLIDN----PBLTLELMTYIKGPDPFPTAGTIMGKSIRAAAYETGRKIIIVRAKAEIEEVN 252  
 12708gyrA SGIAVGMATNIPPNNLGEVIDCGVTLIDN----PBLTLELMTYIKGPDPFPTAGTIMGKSIRAAAYETGRKIIIVRAKAEIEEVN 259

ABW17577.1 BEVNKGROQIIITEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 335  
 EAT31502.1 BEVNKGROQIIITEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 335  
 NP\_560923.1 EEE-NGRHRIIVTEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 334  
 YP\_001307153.1 EEE-NGRHRIIVTEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 334  
 NP\_780806.1 EEE-NGRHRIIVTEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 337  
 AAK77994.1 EEE-NGRHRIIVTEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 334  
 12708gyrA EEE-NGRHRIIVTEIPYQVNKARLIEKIAELVRDCKKVEGISDLRDESREGMRVIELKRDANANVVLNLYKHKTCLQDTFGVIMLA 341

ABW17577.1 IVNEBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 422  
 EAT31502.1 IVNEBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 422  
 NP\_560923.1 LVDNBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 421  
 YP\_001307153.1 LVNNEBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 421  
 NP\_780806.1 LVDNBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 424  
 AAK77994.1 LVDNBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 421  
 12708gyrA LVNNEBPVNLNKEVLYHYLDHOKDIITRRTOFDLNKAERAHILEGLKIALDNIDEVIRLIRGSENQIAKTGLMDFEILSEKQAQA 428

ABW17577.1 ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 509  
 EAT31502.1 ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 509  
 NP\_560923.1 ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 508  
 YP\_001307153.1 ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 508  
 NP\_780806.1 ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 511  
 AAK77994.1 ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 508  
 12708gyrA ILDMRLQRLTGLEREKIEEYEDLILKLNHPEILANERLMLNLIKDELLEIKSKYGDERRTEITRALGEINIEDMITEEVEVITLT 515

ABW17577.1 HPGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 596  
 EAT31502.1 HPGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 596  
 NP\_560923.1 HPGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 595  
 YP\_001307153.1 HPGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 595  
 NP\_780806.1 NSGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 598  
 AAK77994.1 HPGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 595  
 12708gyrA KLGYIKRLPADTYTSQKRGKGKIAALTTRDEDFVENLTSSSHDYLLFFTNGRUVYKLNVEIPEARROAKGTAINVNLPLSPNETV 602

ABW17577.1 AATTPVGRQFQSKVLVLATKNGIIKKTQLSQESTRKSGLIAISIREDELKVKITDDESEIILITAGMSIRFREQDVRPMGRSA 683  
 EAT31502.1 AATTPVGRQFQSKVLVLATKNGIIKKTQLSQESTRKSGLIAISIREDELKVKITDDESEIILITAGMSIRFREQDVRPMGRSA 683  
 NP\_560923.1 QTVLSFKVQEDGFLFMGTGKLGIVKKTPLKDFKNIRKNGLIAIALNREGDELAKKITYGDAIITFVTDGNAIRFNEQDVRPMGRSA 682  
 YP\_001307153.1 QTVLSFKVQEDGFLFMGTGKLGIVKKTPLKDFKNIRKNGLIAIALNREGDELAKKITYGDAIITFVTDGNAIRFNEQDVRPMGRSA 682  
 NP\_780806.1 REVLAQKDLDEGVLITGKGLIKKTSNFKNIRKNGLIAIALNREGDELAKKITYGDAIITFVTDGNAIRFNEQDVRPMGRSA 685  
 AAK77994.1 QAVLSFKVQEDGFLFMGTGKLGIVKKTPLKDFKNIRKNGLIAIALNREGDELAKKITYGDAIITFVTDGNAIRFNEQDVRPMGRSA 682  
 12708gyrA SAVLTPDEYKGEVLFMATKGLVKKTPLETAHVRKTGLAIALNREGDELAKKITYGDAIITFVTDGNAIRFNEQDVRPMGRSA 689

ABW17577.1 MGVKGINLSPKDEHVSMEIVBENKDLLVSEKGFGRKTDTSYRLCHRGGKGKITYNVTKTGELVGAKIVDDEDELLINNYETVI 768  
 EAT31502.1 MGVKGINLSPKDEHVSMEIVBENKDLLVSEKGFGRKTDTSYRLCHRGGKGKITYNVTKTGELVGAKIVDDEDELLINNYETVI 768  
 NP\_560923.1 SEVKATSLREGDIAVQMDIAVEDEKLIVSENGGKRTPLAEYKLVQNGGGLITYKISDKTGKLVGATVCKEDDLMLHYK----- 761  
 YP\_001307153.1 SEVKATSLREGDIAVQMDIAVEDEKLIVSENGGKRTPLAEYKLVQNGGGLITYKISDKTGKLVGATVCKEDDLMLHYK----- 769  
 NP\_780806.1 SEVKATSLREGDIAVQMDIAVEDEKLIVSENGGKRTPLAEYKLVQNGGGLITYKISDKTGKLVGATVCKEDDLMLHYK----- 772  
 AAK77994.1 MGVKGINLSPKDEHVSMEIVBENKDLLVSEKGFGRKTDTSYRLCHRGGKGKITYNVTKTGELVGAKIVDDEDELLINNYETVI 769  
 12708gyrA MGVRGINLSPKDEHVSMEIVBENKDLLVSEKGFGRKTDTSYRLCHRGGKGKITYNVTKTGELVGAKIVDDEDELLINNYETVI 776

ABW17577.1 -----  
 EAT31502.1 RNNNSNLSKLGENTKGVTLMRMEEDQSITIALASTNEEVSEGLTEE----- 819  
 NP\_560923.1 -----  
 YP\_001307153.1 RNNNADSVTSESAMGVTLMRNEDDKVIAIAKLLSSDQETSGDEA--ESESEINNIEE 828  
 NP\_780806.1 RNCVSDSVTSESAMGVTLMRNEDDKVIAIAKLLSSDQETSGDEA--ESESEINNIEE 813  
 AAK77994.1 RNNNDSVTSSESAMGVTLMRNEDDKVIAIAKLLSSDQETSGDEA--ESESEINNIEE 830  
 12708gyrA RNCVSDSVTSESAMGVTLMRNEDDKVIAIAKLLSSDQETSGDEA--ESESEINNIEE 828

specificity against genomic DNA from their respective target species. While preliminary studies suggest *gyrA* gene expression is not affected by bile acid induction, additional work must be done to verify the *gyrA* gene as an appropriate internal control for qRT-PCR. This represents an important step toward development of a quantitative assay to measure *bai* gene expression in *C. scindens* VPI 12708 and *C. hylemonae* TN271. In addition, these probes have yet to be tested on *C. hiranonis* sp. strain T0931 genomic DNA and may result in isolation of this gene from this organism as well.

## Discussion

The first life on earth evolved in an anoxic environment. For several billion years, microorganisms have successfully exploited niches devoid of oxygen (196). In fact, it is interesting to note that one of the most densely inhabited environment on earth is the mammalian large intestine ( $>10^{11}$  bacteria/gram wet weight feces) (197, 198). To put this in perspective, the adult human body contains an order of magnitude more prokaryotic cells ( $10^{14}$ ) than mammalian cells ( $10^{13}$ ). The collective bacterial genomes (termed "microbiome") of the human microflora encode an estimated 2-4 million genes, surpassing the human genome by a staggering 100-fold (199). Science continues to change the way we view ourselves, and our place in nature. Advances in genomics and GI microbiology suggest that the human body should be regarded as a superorganism—a complex ecosystem composed of all three domains of life (eukaryote, prokaryote and archaea) balanced by selective pressures from both the top-down and the bottom-up (200).

Intestinal bacteria and archaea can carry out hundreds, if not thousands, of enzymatic reactions not carried out by host cells. Relman and Falkow (2001) have called for the task of sequencing this microbiome, which they have called 'the second human genome project' (201). Intestinal microbes actively metabolize molecules of endogenous and exogenous origin. The products of this metabolism then interact with host cells with both beneficial and harmful consequences. Bile acid metabolism represents a particularly important aspect of microbial metabolism of endogenous substrates due to enterohepatic circulation, which allows accumulation of bacterial products of bile acid metabolism. Bile acids also represent a strong top-down selection pressure on potential colonizers (95). A microbe must be prepared to cope with mmol concentrations of bile salts and bile acids. Bile resistance is a major factor used to screen potential "probiotic" species (202). In addition, numerous studies have shown that bile acids induce expression of virulence genes in intestinal pathogens (95). Host-microbe co-evolution has resulted in widespread acquisition of

bile-salt metabolizing enzymes in all major gut colonizing microbial species (See introduction) (95).

Interestingly, it appears that genes involved in bile acid 7 $\alpha$ / $\beta$ -dehydroxylation are found in the tiniest fraction of *Firmicutes* classified within various clusters of the heterogenous genus *Clostridium* (189). 7 $\alpha$ -dehydration of bile acids represents a microbial pathway which may significantly alter host physiology. Bile acids, once regarded as simple detergents involved in lipid digestion, are now recognized as potent signaling molecules (203). A full understanding of the biology and biochemistry of bile acid metabolism by intestinal bacteria represents an important and challenging area of research. The current work represents a significant step in understanding the biochemistry of bile acid 7 $\alpha$ / $\beta$ -dehydroxylation by intestinal clostridia.

The current work provides strong evidence that the *baiCD* and *baiH* gene products encode NAD<sup>+</sup>-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductases recognizing 7 $\alpha$ -hydroxy and 7 $\beta$ -hydroxy bile acids, respectively (Figure 5). Previously, we demonstrated loss of the 5 $\beta$ -hydrogen during 7 $\alpha$ -dehydroxylation of [5 $\beta$ -<sup>3</sup>H] + [24-<sup>14</sup>C]-CA both *in vivo*

and from cell extracts of CA induced *C. scindens* whole cells suggesting introduction of a C<sub>4</sub>-C<sub>5</sub> double-bond (118). Addition of [24-<sup>14</sup>C] 3-oxo-Δ<sup>4</sup>-cholenoic acid to CA induced whole cells and cell-extracts of *C. scindens* VPI 12708 resulted in conversion to [24-<sup>14</sup>C] DCA (117). Subsequent studies elucidated the first two steps of the bile acid 7α/β-dehydroxylation pathway; CoA ligation (125) and oxidation of the 3α-hydroxy group producing a 3-dehydro-CA-CoA intermediate (58). Collectively, these data suggested the presence of an enzyme catalyzing the region-selective oxidation at C<sub>4</sub>-C<sub>5</sub> yielding a 3-dehydro-Δ<sup>4</sup>-CA-CoA intermediate. Others have speculated that the *baiH* and *baiCD* genes were involved in C-C double bond oxidation/reduction based on homology to enzymes with similar function (204, 205). The fact that there are two homologous flavin oxidoreductases of similar M<sub>r</sub> (70-72 kDa) with 32 % identity and 46 % similarity that are co-expressed from a *bai* operon suggests a possible similar function.

This hypothesis was tested by incubation of a crude cell extract of *E. coli* overexpressing the *baiCD* gene with [24-<sup>14</sup>C] 3-dehydro-UDCA or [24-<sup>14</sup>C] 3-dehydro-CDCA, and

indeed the reaction yielded a 3-dehydro- $\Delta^4$ -CDCA metabolite as suggested by mass spectrometry (Figure 19). 3-dehydro-UDCA did not appear to be a substrate of the *baiCD* gene product. Bile acid 7 $\alpha$ -hydroxy groups are axial while 7 $\beta$ -hydroxy groups are equatorial. These data suggest that the key to substrate recognition between the *baiCD* and *baiH* is the stereochemistry of the 7-hydroxy group, whose orientation apparently requires a different architecture in the substrate binding pocket. Taken together these results suggest these homologous enzymes act at an analogous step in the 7-dehydroxylation pathway, and therefore, two distinct branches of the oxidative arm of the pathway exist in *C. scindens* and *C. hiranonis*.

The BaiH is a homotrimer composed of 72 kDa subunits, each subunit containing 1 mol of FMN and 1 mol of FAD per subunit. BaiH is homologous to a class of flavoenzymes involved in reduction of unsaturated fatty acids and aldehydes (Table 2) (129). Enoate reductase (*enr*) from *Clostridium tyrobutyricum* catalyzes the reduction of C-C double bonds of several  $\alpha/\beta$ -unsaturated aldehydes as well as cyclic ketones and methylketones (184). Enr is a dodecamer with subunit  $M_r$  of 73 kDa each containing a 4Fe-

4S cluster and 1 mol of FMN and 1 mol of FAD (206). 2, 4-dienoyl-CoA reductase from *E. coli* is a 73 kDa monomeric enzyme that reduces C-C double bonds with NADPH and contains FMN, FAD as well as a 4Fe-4S cluster (182, 207). The *noxB* gene from *Archaeoglobus fulgidus* encodes an NADH oxidase of unknown physiological function (204). The *noxB* gene product is a 69 kDa monomer containing 1 mol FAD and shares the same conserved ferredoxin domain (C-2X-C-2-3X-C-11-12X-C) with each of these proteins including the *baiH* and *baiCD* gene products (204). The *noxB-2* gene shares 98.9 % identity with the *noxB-1* gene from *A. fulgidus* and is found upstream of medium-chain acyl-CoA ligase in this bacterium suggesting these genes are also involved in fatty acid metabolism (204). The more distantly related Old Yellow Enzyme (OYE) family homologues are found both in eukaryotes such as *Saccharomyces cerevisiae* (OYE1, OYE2) (205) and *S. carlsbergensis* (OYE1) (208), as well as the prokaryotes including *Bacillus subtilis* (YqjM) (209), and *Shewanella oneidensis* (SYE1-SYE4) (210). OYE genes have been extensively characterized both biochemically and structurally (206, 209, 211). Numerous compounds act as electron acceptors including various quinones, as well as



$\alpha$ ,  $\beta$ -unsaturated aldehydes and ketones which are reduced at the olefinic bond (211); however, only recently have physiologically relevant electron acceptors been identified. The *oye2* and *oye3* gene from *Sac. cerevisiae*, *syel* from *She. oneidensis*, and the *yqjM* gene from *B. subtilis* were shown to be inducible by small  $\alpha/\beta$ -unsaturated aldehydes during oxidative stress responses to lipid peroxidation (186, 210, 212) and actin cytoskeletal oxidative damage in yeast (213). Detoxification does not appear to be the function of the bile acid  $7\alpha/\beta$ -dehydroxylating pathways, but rather the substrates appear to serve as electron sinks to enhance fermentation (95).

Electron flow in the *E. coli* DCR as determined by redox titrations (214), and crystallography (182) show that two reducing equivalents are supplied to FAD by direct hydride transfer and follow an electron flow from FAD to a 4Fe-4S cluster to FMN and finally to the oxidized substrate. OYEs reduce C=C double bonds through a ping-pong bi-bi mechanism in which FMN is reduced by NADPH, followed by substrate binding and hydride transfer from FMNH<sub>2</sub> to the C <sub>$\beta$</sub>  carbon of the substrate (215). Additional

studies will be required to determine the mechanism of the reaction catalyzed by the *baiCD* and *baiH* gene products.

We observed low NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity based on the amount of products formed, as observed by autoradiography of TLC (Figure 16). However, we would predict that the physiological substrates for these enzymes are the CoA conjugates of 3-oxo-cholanoic acids. This is based on the observations that the *baiA* gene product encodes a 3 $\alpha$ -hydroxysteroid dehydrogenase that has high specificity for coenzyme A conjugated bile acids (58). However, large quantities of the CoA derivatives of 3-dehydro-CDCA and 3-dehydro-UDCA will have to be chemically synthesized to test this hypothesis. In this regard, we have detected NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductase activity using small quantities of enzymatically generated radiolabeled CoA conjugates of 3-dehydro-UDCA and 3-dehydro-CDCA (D. Kang, J.M. Ridlon, and P.B. Hylemon, unpublished data). Moreover, other members of this gene family recognize substrates linked to CoA (Table II). Future mechanistic and structural studies will require

chemical synthesis of CoA conjugated bile acid intermediates.

This is the first report of a gene product specifically involved in 7 $\beta$ -dehydroxylation of UDCA. We predict that the *baiI* gene encodes a bile acid 7 $\beta$ -dehydratase, based on its amino acid sequence identity (20 %) and similarity (40 %) with the *baiE* gene product as well as co-expression of this gene product during CA induction in *C. scindens* VPI. We have previously demonstrated that the *baiE* gene encodes a bile acid 7 $\alpha$ -dehydratase (128). Determining the function of the *baiI* as well as identifying genes involved in reducing the 3-oxo-4, 6-choladienoic structure to LCA will be the focus of future of research.

In summary, these data indicate that the *baiCD* and *baiH* genes encode NAD(H)-dependent 3-oxo- $\Delta^4$ -cholenoic acid oxidoreductases. These results can account for the ability of this organism to use UDCA as a substrate for 7 $\beta$ -dehydroxylation, and provide additional support for formation of a 3-oxo- $\Delta^4$ -cholenoic acid intermediate during bile acid 7-dehydroxylation.

Doerner et al (1997) demonstrated that  $7\alpha/\beta$ -dehydroxylating bacteria cluster into two groups based on relative rates of conversion of CA to DCA (189). Our current understanding of the molecular biology and enzymology of bile acid  $7\alpha/\beta$ -dehydroxylation come from study of *C. scindens* VPI 12708 and *C. hiranonis* sp. strain TO931 which are "high activity strains" ( $1-7 \text{ nmol mg}^{-1} \text{ h}^{-1}$ ). Genes involved in bile acid  $7\alpha/\beta$ -dehydroxylation have not been identified in a "low activity" strain ( $0.05-0.16 \text{ nmol mg}^{-1} \text{ h}^{-1}$ ). Wells et al. (2003) detected the *baiCD* gene from the low activity strain *C. hylemonae* TN271 which facilitated design of bidirectional genome-walking PCR primers (106). The genome-walking by PCR technique is useful for amplification of unknown stretches of DNA provided at least a small amount of adjacent sequence data is available for design of gene specific primers. Adaptor sequences are engineered and ligated to restriction libraries of total genomic DNA. These adaptor molecules provide the second stretch of known sequence, allowing PCR amplification between adaptor-specific and gene-specific primers (216). PCR products up to ~4 kb were obtained in

this study, with an average of 2 kb using a starting genomic template >48 kb (Figure 21).

Two upstream genome-walking PCR products (Figures 22, 23) and three downstream genome-walking PCR products (Figures 28, 30, 31) were cloned, sequenced and assembled, along with the partial *baiCD* gene into a single contiguous sequence of ~9.5 kb. ORFs were identified by the ORF Finder program and BLAST searches were used to determine the putative function of these ORFs. The gene organization of the *baiBCDEFGHI* operon from *C. hylemonae* TN271 was almost identical to the *baiBCDEAFGHI* operon from *C. scindens* VPI 12708 (Figure 32). Table 7 compares both nucleotide and amino acid sequence identity/similarity between *bai* genes from *C. hylemonae* TN271 and *C. scindens* VPI 12708. Nucleotide sequence identity was >65% and amino acid sequence identities ranging from 61.4 to 88% with similarity ranging from 76.6% to 94.8%. The upstream putative "promoter" regions were also found to be highly conserved along with relative location of TIS (Figures 24, 25). Finally, a conserved transcriptional regulator of the AraC/XylS family (*barA*) was located upstream of the *C. hylemonae* TN271 *baiB* gene and

"promoter" region on the antisense strand (Figures 27, 32). It is unclear at present the reason for the 10 fold difference in levels of bile acid 7 $\alpha$ -dehydroxylating activity between *C. hylemonae* TN271 and "high activity" strains. However, it is very likely that this regulation is at the transcriptional level. Indeed, design of qRT-PCR oligonucleotides against genes in the *baiBCDE(A)FGHI* operons from *C. hylemonae* and *C. scindens* VPI 12708 and quantitation of message following bile acid induction should provide the relevant information. Isolation and sequencing of the *gyrA* gene in the current study and current genomic data for *Clostridium scindens* will allow normalization of qRT-PCR results. Finally, it is interesting to note that *C. hylemonae* TN271, despite having a *baiH* and *baiI* gene, were unable to 7 $\beta$ -dehydroxylate UDCA over a 30 min or overnight incubation following induction by CA (Figure 33). It is unclear at present the reason for this loss-of-function. Importantly, this data represents the first characterization of *bai* genes from a "low activity" 7 $\alpha$ -dehydroxylating bacterium.

The *baiA* gene from *C. hylemonae* TN271 was isolated by PCR amplification using redundant oligonucleotides whose design was based on a nucleotide multiple sequence alignment of *baiA* genes isolated from *C. scindens* VPI 12708 and *C. hiranonis* sp. strain T0931 (Figures 34, 35, 37). 5' SMART RACE PCR analysis of the *baiA* transcript revealed a TIS 56 bp upstream of the *baiA* gene and downstream of the putative *bai* promoter region (Figure 36). Gopal-Srivastava et al (1990) identified multiple copies of the *baiA* gene in *C. scindens* VPI 12708, two of which were monocistrons with conserved *bai* "promoters" (67).

Completion of the *baiA* gene sequence, facilitated by bidirectional genome-walking from the partial sequence obtained with primers *baiA*271F and *baiA*271R (Table 8) resulted in isolation of genes with predicted function suggested to be involved in bile acid 7 $\alpha$ / $\beta$ -dehydroxylation. *In silico* (ORF Finder, BLAST) analysis upstream of the *baiA* gene revealed a deduced 26 kDa short chain dehydrogenase/reductase (*baiL*), a 49.5 kDa *baiF* homologue (*baiK*) and a putative 62 kDa 3-keto-steroid  $\Delta^1$ -dehydrogenase (*baiJ*). Upstream of the *baiJ* gene was a

conserved *bai* "promoter" region (Figure 44). 5' SMART RACE analysis suggests the TIS is at the very 5' end of this conserved region while the TIS is downstream of this conserved region in the *baiB* and *baiA* operons of both *C. scindens* VPI 12708 and *C. hylemonae* TN271 (Figure 45). This would suggest that the *baiJ* promoter is upstream of this conserved region.

In order to determine whether these novel genes were conserved in *C. scindens* VPI 12708, the *baiJKL* genes from *C. hylemonae* were analysed for 7-8 amino acid stretches of low codon degeneracy (Figures 46, 47). A region within the *baiJ* gene allowed design of redundant oligonucleotides which allowed isolation of a partial *baiJ* gene from *C. scindens* VPI 12708. Several upstream and downstream genome-walks resulted in identification of the *baiJKL* genes in *C. scindens* VPI 12708 which were highly conserved both in sequence and organization to the *baiJKL* genes from *C. hylemonae* TN271 (Figure 56). Elucidation of the TIS for the *baiJ* gene provided similar results to those obtained for TIS of the *baiJ* gene from *C. hylemonae* in respect to the upstream position relative to the conserved "promoter" region identified previously for the *baiB* and



*baiA* genes from *C. scindens* VPI 12708 (124, 191), *Clostridium hiranonis* sp. strain T0931 (59) and here for the *baiB* gene of *Clostridium hylemonae* TN271 (Figure 26). This brings into question whether this conserved region is indeed a bile acid inducible promoter rather than a repressor or operator sequence. Figure 73 is a Weblogo representation of a CLUSTALW alignment of these conserved upstream regions between *C. scindens* VPI 12708, *C. hiranonas* sp. strain T0931 and *C. hylemonae* TN271. Further experimental work is necessary to determine the function of these regions.

A particularly interesting find in this study is a predicted Type III CoA transferase which appears to share a common evolutionary history with the *baiF* gene (Figure 42). This gene has been given the preliminary name *baiK* (Figure 58). Alignment of characterized type III CoA transferases such as the crotonoyl-CoA (*caiB*) and *yfdW* gene products from *E. coli* as well as the formyl CoA transferase from *Oxalobacter formigenes* with the *baiF* and *baiK* gene products from *C. scindens* VPI 12708 and *C. hylemonae* TN271 demonstrates a remarkable degree of conservation (Figure 74). Numerous authors have

Figure 73. Logo representation of *bai* promoter regions from *Clostridium scindens* VPI 12708, *Clostridium hiranonis* sp. strain T0931 and *Clostridium hylemonae* TN271. Promoter sequences from Figure 24 and Figure 49 were aligned using T-COFFEE and FASTA format was run in the Weblogo program (<http://weblogo.berkeley.edu/>). Large letters represent based of potential biological relevance to promoter function.

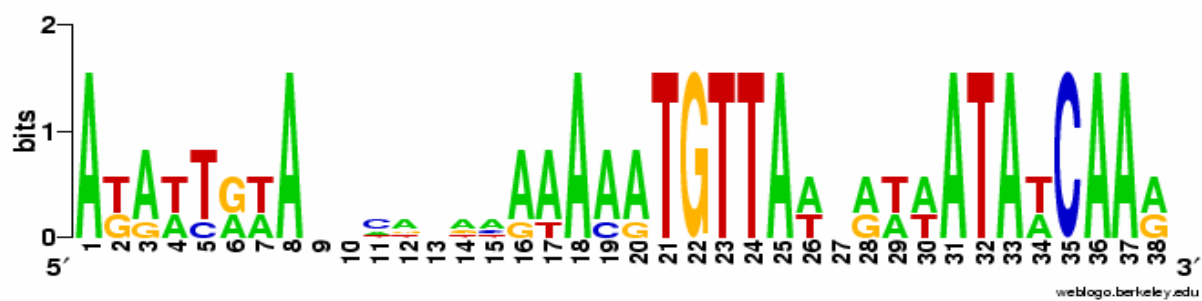


Figure 74. Boxshade alignment of *baiF* and *baiK* genes from *C. scindens* VPI 12708 and *C. hylemonae* TN271 with homologues in the Type III CoA transferase family. Proteins used in alignment include: *E. coli* Crotonobetainyl-CoA:carnitine CoA-transferase (*CaiB*) (P31572), *E. coli* Formyl-CoA transferase (*YfdWA*) (P69902), *Oxalobacter formigenes* formyl-CoA transferase (*fCTOx*) (AAC45298).

12708baik MKGTGNNFPQFGVMEGVKILVCGGATAGPFGATLLGETGAEVVHLESPPKNPD<sup>\*</sup>SVRGHYG 61  
 TN271baik MEGTGNNFPQFGVLAGVKILVCGGATAGPFGATLLGETGAEVVHLESPPKNPD<sup>\*\*</sup>STRGHYG 61  
 12708baif --MAGIKDFPKFGALAGLKILDSGSNIAGPLGGGLLECGATVIHFEGPKKPDNQRGNYG 59  
 TN271baif --MAGIKDFPSPFGALAGLKILDSGSNIAGPLGGGLLECGATVIHFEGPKKPDNQRGNYG 59  
 caiBEco ---MDHLPMPKFGPLAGLRVVFSGIETAGPFAQMFAGVGA<sup>\*</sup>FIWLENVAVADTIRVQPN 58  
 YfdwA -----MSTPLQGTIKVLDFTGVQSGPSCTQMLAWFGADV<sup>\*</sup>IKLERPVGVDVTRHQLR 51  
 fCTOx -----MTKPLDGINVLDFTHVQAGPACTQMMGFLGANV<sup>\*</sup>IKLERRGSGDMTRGWLQ 51

12708baik -----YSON-HRNQLSMVADMKTEEGLEIFKKLIKWT<sup>\*</sup>DIPIESSKGGTYEKMGLTDE 111  
 TN271baik -----YSON-HRNQLSMVADMKTEEGLEIFKKLIKWT<sup>\*</sup>DIPIESSKGGTYEKMGLTDE 111  
 12708baif -----YPON-HRNQLSMVADIKTEEGRKIFLDLIKWADIN<sup>\*</sup>VESSKGGQYDRIGLSDE 109  
 TN271baif -----YPON-HRNQLSMVADIKTEEGREIFMKLIKWADIN<sup>\*</sup>VESSKGGQYDRIGLSDE 109  
 caiBEco -----YPQLSRNLHALSLNIFKDEGREAFKLMTETD<sup>\*</sup>PIEASKGPAFAARGITDE 109  
 YfdwA DIPDIDALYFTMLNSNKRSTIELNTKTAEGKEVMEKLI<sup>\*</sup>READILVENFHPGAIDEMCGFTWE 110  
 fCTOx DKPNVDSL<sup>\*</sup>YFTMFNCKRSIELDMKTEEGKELLEQMIKKADVMVENFGPCALDRMGFTWE 110

12708baik VLWEINPRLAIVHVS<sup>\*</sup>GFGQTGVPEYIDRAS<sup>\*</sup>YDAVGQA<sup>\*</sup>AFSGYMSFN<sup>\*</sup>GT<sup>\*</sup>PK<sup>\*</sup>EAMKVS-PYLS 170  
 TN271baik VLWEVNPGLAIVHVS<sup>\*</sup>GFGQTGVPEYISRAS<sup>\*</sup>YDAVGQA<sup>\*</sup>AFSGYMSFN<sup>\*</sup>GT<sup>\*</sup>PK<sup>\*</sup>EALKIS-PYLS 170  
 12708baif VIWEVNP<sup>\*</sup>KIAIVHVS<sup>\*</sup>GFGQTGDSYVTRASYDAVGQA<sup>\*</sup>AFSGYMSLN<sup>\*</sup>GT-TEALKTN-PYLS 167  
 TN271baif EIEKVNPKIAIVHVS<sup>\*</sup>GFGQTGDSYVTRASYDAVGQA<sup>\*</sup>AFSGYMSLN<sup>\*</sup>GT-TEALKTN-PYLS 167  
 caiBEco VLWOHNPKLVIAHLSGFGQYGT<sup>\*</sup>EEYTNLPAYNTIAQA<sup>\*</sup>AFSGYLIQNGD<sup>\*</sup>VDQPM<sup>\*</sup>PAF-PYTA 168  
 YfdwA HIEEINPRLIFGSLKGFDEC--SPYINVKAYENVAQA<sup>\*</sup>AGGAATTEFWDGPPLVSAALG 168  
 fCTOx YIEELNPRVILASVKG<sup>\*</sup>YAE<sup>\*</sup>G--HAN<sup>\*</sup>EHLKVYENVAQA<sup>\*</sup>CGS<sup>\*</sup>GAAATTEFWDGPPTVSGAALG 168

12708baik DVVTALNTCTALAA<sup>†</sup>YVHLRTGKGESVDVAQY<sup>\*</sup>ESLARIL<sup>\*</sup>DT<sup>\*</sup>RPMEYFTDGKE----- 223  
 TN271baik DVVTALNTCTALAA<sup>†</sup>YVHLRTGKGESVDVAQY<sup>\*</sup>ESLARIL<sup>\*</sup>DT<sup>\*</sup>RPMEYFTDGKE----- 223  
 12708baif DFVCGLITTCWAMLAC<sup>†</sup>YVSTILTGKGESVDVAQY<sup>\*</sup>EALARIM<sup>\*</sup>GRMIQYATDGVK----- 220  
 TN271baif DFVCGLITTCWAMLAC<sup>†</sup>YVSTILTGKGESVDVAQY<sup>\*</sup>EALARIM<sup>\*</sup>GRVMQYATDGVK----- 220  
 caiBEco DFVSGLTATTAALAA<sup>†</sup>LHVRETGKGESID<sup>†</sup>AMYEVM<sup>\*</sup>LRMGQYFMM<sup>\*</sup>DFNGSEM----- 221  
 YfdwA DSNTGMHLLIGLLAALLREK<sup>†</sup>TGRGQ<sup>†</sup>RV<sup>†</sup>MSMODAVLN<sup>\*</sup>CRVKLRQQR<sup>\*</sup>LDKLGYLEEYP 228  
 fCTOx DSN<sup>†</sup>SGHLMIGILAAL<sup>†</sup>EMRHKTGRGQKVA<sup>†</sup>VAMODAVLN<sup>\*</sup>VR<sup>†</sup>IKLRQQR<sup>\*</sup>LEBTGILAEYP 228

12708baik -----FPRTGN-KDTQAALFSFY<sup>\*</sup>TCKDGGGEIFIGMNG<sup>\*</sup>GPV<sup>\*</sup>RR 260  
 TN271baik -----FPRTGN-KDSQAALFSFY<sup>\*</sup>TCKDGGGEIFIGMNG<sup>\*</sup>GPV<sup>\*</sup>RR 260  
 12708baif -----MPRTGN-KDAQAALFSFY<sup>\*</sup>TCKDGR<sup>\*</sup>TIFIGMTCAEV<sup>\*</sup>CKR 257  
 TN271baif -----VPRTGN-KDQAALFSFY<sup>\*</sup>TCKDGR<sup>\*</sup>TIFIGMTCAEV<sup>\*</sup>CKR 257  
 caiBEco -----CPRMSKGD<sup>\*</sup>PYYAGCGLYKCA<sup>\*</sup>DG-YIVMEIVG<sup>\*</sup>ITQIEE 258  
 YfdwA QY-PNGTFG-----DAVPRG<sup>\*</sup>GN-AGGGQPGWILKCKGWE<sup>\*</sup>TDPNAYIYETIQEQ 275  
 fCTOx QAQPNFAFDRDGNPLSFDNITSVPRG<sup>\*</sup>GN-AGGGQPGWMLKCKGWE<sup>\*</sup>TDADSYVYETTHAAN 287

12708baik GYPLIGLPKPGDGDPEID<sup>†</sup>ILSG-WMADTD<sup>†</sup>GRRL<sup>†</sup>EAA<sup>†</sup>MEK<sup>†</sup>FVSEHTVDEVEKIMLEN<sup>†</sup>OI 319  
 TN271baik GYPLIGLPKPGDGDPEID<sup>†</sup>QIIISG-WMADTD<sup>†</sup>GRRL<sup>†</sup>EAA<sup>†</sup>MEK<sup>†</sup>FVSEHTVDEVEKIMLEN<sup>†</sup>OI 319  
 12708baif GEPLIGLPVPGTGD<sup>†</sup>DPFEGFTG-WMIYTPV<sup>†</sup>GORMEKAMEK<sup>†</sup>FVSEHTMEEVEAE<sup>†</sup>MQAH<sup>†</sup>OI 316  
 TN271baif GEPLIGLPQPGSGD<sup>†</sup>DPFEGFTG-WLIDSPV<sup>†</sup>GRMEAA<sup>†</sup>MEK<sup>†</sup>FVSEHTMEEVEAV<sup>†</sup>MQEN<sup>†</sup>OI 316  
 caiBEco CEKDIGLAHL-LGTEPEI<sup>†</sup>PGTQLIHRLECPY<sup>†</sup>GPLVEEK<sup>†</sup>LDANLATHTIAEVKERFAEL<sup>†</sup>NI 317  
 YfdwA NMENTCKAIGKPEWITDPAYSTAHAROPHIFD--IFAETEK<sup>†</sup>TVTIDKHEAVAY<sup>†</sup>ITQFDI 333  
 fCTOx MNPQICDMIDKPEWKDDPAYNTEFGRVDKLD--IFSFE<sup>†</sup>ETKFADKDKFEVTEWAAQYGI 345

12708baik PCLKVYTLKPCAKDPHWKARDIFVEWDDPMMGRVKG<sup>\*</sup>GLG<sup>\*</sup>LINK<sup>\*</sup>MKN<sup>\*</sup>NPGEIKWGAPLFGEN 379  
 TN271baik PCQKVYSLEDCVRODPHWKAREIFTEWDDPMMGRVKG<sup>\*</sup>GIGIVNKMKN<sup>\*</sup>NPGEIKWGAPLFGEN 379  
 12708baif PCQRVYELEDCLNDPHWKARGTITTEWDDPMMGHITGLGLINK<sup>\*</sup>FKRNPSEIWRGAPLFGMD 376  
 TN271baif PCQRVYELEDQISDPHWIARETITTEWDDPMLGHVITGLGLINK<sup>\*</sup>FKNNPSKIWRGAPLFGMD 376  
 caiBEco ACAKVLTVPLESNPQYVARE<sup>\*</sup>SITQWOTMDGR<sup>\*</sup>TC<sup>\*</sup>CPN<sup>\*</sup>IMP<sup>\*</sup>KFKNN<sup>\*</sup>PGQIWRGMPHSGMD 377  
 YfdwA PCAPVLSMK<sup>\*</sup>ISLDESLRQSGSVVEVEOP<sup>\*</sup>IRGKYLTVC<sup>\*</sup>PMK<sup>\*</sup>SAFTPD<sup>\*</sup>IKALAPLLGEH 392  
 fCTOx PCGPVMSMK<sup>\*</sup>LAHDBSLQKVGTVVEVVD<sup>\*</sup>TRGNH<sup>\*</sup>LTVC<sup>\*</sup>APFK<sup>\*</sup>SGFQPEITR-APLLGEH 404

12708baik NEEVLKDLGYTEEEIEDFAKRGITASDFDQTY  
 TN271baik NKEVLKDLGYTEEEIEDLARRGITACLD<sup>†</sup>FQTY  
 12708baif NRDILKDLGYDDAKIDELYEQGI<sup>†</sup>VNEFDL<sup>†</sup>TTI  
 TN271baif NRDILKDLGYTDEEIDGLYEKGITNEFD<sup>†</sup>RTTI  
 caiBEco TAAILKNIGYSEN<sup>†</sup>DIQELVSKGLAKVED-----  
 YfdwA TAAVLQELGYSDDETAAMKQNHAI-----  
 fCTOx TDEVLKELGLDDAKIKELHAKQV-----

speculated that CoA transfer proceeds through Asp<sup>169</sup> acylation (217-219). Mutagenesis of Asp<sup>169</sup> (D169A, D169S, D169E) in the formyl-CoA transferase (FRC) from *O. formigenes* resulted in abolition of, or appreciable decrease in specific activity despite no significant structural alterations (217). The proposed catalytic mechanism for the FCR suggests that formyl-CoA forms an acid anhydride with Asp<sup>169</sup> resulting in elimination of <sup>-</sup>SCoA. The second substrate, oxalate, then nucleophilically attacks the Asp<sup>169</sup>-formate anhydride resulting in release of formate and formation of Asp<sup>169</sup>-oxalate anhydride. The anion <sup>-</sup>SCoA then attacks the carbonyl carbon of oxalate resulting in release of oxalyl-CoA and regeneration of Asp<sup>169</sup> (217). Asp<sup>169</sup> is conserved in the *baiF* and *baiK* genes of *C. scindens* VPI 12708 and *C. hylemonae* TN271 suggesting that these proteins act through a similar mechanism (Figure 74). Crystallization with CaiB yielded crystals with only CoA bound suggesting hydrolysis occurs in the absence of the CoA accepting substrates (220). This would provide a plausible explanation for observed CoA hydrolase activity shown for the *baiF* gene product in the absence of co-substrate

(130). We speculate that a similar mechanism is involved in bile acid CoA transfer by the *baiF* and possibly the *baiK* genes to that proposed for the FRC. Isolation of a 3-oxo- $\Delta^4$ -DCA~CoA intermediate from cell extracts of cholic acid induced cultures of *C. scindens* VPI 12708 would suggest that CoA transfer would occur at this step or possibly following formation of 3-oxo-DCA or DCA. Therefore, these three metabolites would substitute for formate in the FRC model and cholyl-CoA would substitute for oxalate in the FRC model.

There are several gene homologues in the oxidative arm of the  $7\alpha/\beta$ -dehydroxylation pathway performing a similar function but recognizing a distinct stereoisomer. Indeed, in the present work, the *baiCD* and *baiH* gene products were shown to act as 3-oxo- $\Delta^4$ -steroid oxidoreductases recognizing  $7\alpha$ -hydroxy and  $7\beta$ -hydroxy bile acids, respectively. Additionally, the *baiE* was shown to  $7\alpha$ -dehydrate bile acids while the *baiI* is predicted to  $7\beta$ -dehydrate bile acids. In each case, stereospecificity is directed at the orientation of the 7-hydroxyl group. However, in the reductive arm of the  $7\alpha/\beta$ -dehydroxylation the only differences include hydroxylation at C12, and

stereochemical differences about the A/B ring. Assuming the *baiK* gene is involved in bile acid CoA transfer, there are a few possibilities: (1) one transferase recognizes LCA (not hydroxylated at C12) while the other recognizes DCA (hydroxylated at C12) (2) one transferase recognizes DCA and LCA (A/B ring in "cis" orientation) while the other transferase recognizes alloDCA and alloLCA (A/B ring in "trans" orientation) (3) both transferases recognize the same substrate(s); similar to the multiple copies of the *baiA* genes (124). Testing these possibilities will require acquisition of alloDCA (not available commercially; highly concentrated in rabbit gallstones (221)) and synthesis of 3-oxo-alloDCA(~SCoA).

The *baiL* gene is predicted to encode a 3-keto-acyl reductase, similar in function to the *baiA* gene. The predicted function in terms of bile acid 7 $\alpha$ -dehydroxylation is NAD(P)H dependent 3 $\alpha$ -hydroxysteroid dehydrogenation resulting in DCA and LCA. An attempt was made to overexpress the *baiL* gene in *E. coli* as a hexahistidine tagged fusion protein. Unfortunately, and often, the production of biologically active recombinant proteins is prevented by the formation of insoluble



inclusion bodies. Numerous variables were modified to improve solubility including decreased temperature, IPTG concentration, different media, various host strains, and vectors. However, while the protein was highly expressed, I was unable to detect soluble BaiL by Western immunoblot or NAD(P)-dependent 3 $\alpha$ -HSDH activity against DCA. It may be necessary to experiment with additional fusion tags, further decrease in temperature, or attempt expression in a different host. Shuttle vectors have been developed between *E. coli* and gram-positive bacterial genera including *Clostridium* (222).

The *baiJ* gene is predicted to encode a 62 kDa flavoprotein similar to fumarate reductase and 3-ketosteroid- $\Delta^1$ -dehydrogenases. Alignment of the *baiJ* gene product with the fumarate reductase from *Shewanella putrefaciens* MR-1, whose structure has been elucidated, suggests the similarity between these proteins is in the Rossman fold of the FAD-binding motif. The BaiJ lacks the cytochrome domain containing four conserved heme-binding sites common to fumarate reductases (CXXCH) and there was no conservation of active-site amino acids (data not shown). The *baiJ* gene was cloned into a pT7-SBP-2 vector

containing an N-terminal Streptavidin Binding Peptide (SBP) fusion tag. IPTG induced overexpression of the BaiJ polypeptide resulted in a soluble protein ~64 kDa which was detected by anti-SBP antibody (Figure 66). Initial characterization of the enzyme with radiolabelled 3-oxo-DCA and 3-oxo-DCA-CoA resulted in three products formed when NADP was used as co-substrate. The two hydrophobic products were also detected when  $\text{NAD}^+$  was co-substrate. A substrate of greater hydrophilicity than DCA was detected in the presence of NADP. This class of enzyme is predicted to metabolize C=C double bonds. This hydrophilic product is puzzling because the migration suggests the 3-oxo group has been reduced. Further, the reduced migration may indicate oxidation of the bile acid structure as there is no other source of reducing equivalents. It is possible that a double bond is being introduced in the side-chain of the steroid rings. Indeed,  $\Delta^{22}$ -bile acids have been identified in rats (223). Additional studies will be required to determine the identity of these bile acid metabolites. These findings may indicate a novel bile acid metabolic activity in *C. scindens* VPI 12708.

Future studies may include determining which genes are co-transcribed in both the *baiBCDEFGHI* and *baiJKL* operons from *C. hylemonae* TN271 and the *baiJKLM* operon from *C. scindens* VPI 12708. A sensitive method for determining this would be PCR amplification of cDNA fragments containing intergenic regions between adjacent ORF's. Additionally, attempts to isolate *bai* genes from *C. leptum* and *C. sordellii* using degenerate oligonucleotides based on conserved regions in the oxidative *bai* genes from *C. scindens* VPI 12708, *C. hylemonae* TN271 and *C. hiranonis* sp. strain TO931 have failed (Ridlon JM and Hylemon PB, unpublished data). It would be interesting to try and design degenerate oligonucleotides to isolate the *baiJKL* genes, if present, in *C. leptum*. This group of "low activity" organisms was not detected by Southern blot (189) or by PCR using primers against the *baiCD* gene (106). Isolation of *bai* genes from this organism could allow design of gene probes which could be used to quantify levels of a wide group of known 7 $\alpha$ -dehydroxylating bacteria. Additionally, large-scale metagenomic sequencing will provide a wealth of information which will, in the future, allow *in silico*

mining of putative *bai* genes. Another study of interest would be to determine whether the *baiF* and *baiK* genes encode bile acid CoA transferases, and determine physiological substrate(s) through kinetic analysis. The limiting factor in such a study is chemical synthesis of numerous bile acid intermediates as both free acids and CoA conjugates. In addition, comparison of wild-type proteins with site-directed mutagenesis of Asp<sup>169</sup> would provide additional evidence for a common mechanism of the *baiF* and *baiK* with that of the FRC protein.

Finally, the major aim of research into bile acid 7 $\alpha$ -dehydroxylation is to identify targets for development of specific inhibitors aimed at lowering 7 $\alpha$ -dehydroxylation, and thus secondary bile acids in individuals with high levels (Figure 11). A particularly interesting gene was located upstream of the *baiJKL* genes and directly downstream, and on the opposite strand relative to the *baiA* gene in *C. scindens* VPI 12708 (Figure 58). The ORF encodes a protein similar to the TspO/MBR family of cell surface sensors. TspO, or Tryptophan Rich Sensory protein, is a negative regulator of photosynthesis in response to light/oxygen. The tryptophan rich region acts

as part of a hydrophobic pore that acts in pumping porphyrin across the cell wall. Could it be that this protein, in 7 $\alpha$ -dehydroxylating bacteria, may be involved in regulation of 7 $\alpha$ -dehydroxylating gene expression and efflux of hydrophobic secondary bile acids from the cell? MBR, or mitochondrial peripheral benzodiazepine receptor, has been shown to substitute for bacterial TspO (224) and is thought to function in cholesterol transport across the mitochondrial membrane. If this gene product could be shown to regulate bile acid repression/efflux, this could serve as potential drug target; particularly because this polypeptide is predicted to be located in the outer membrane. Also, an ABC transport protein was located directly downstream of the *baiL* gene; possibly co-transcribed. This suggests that this polypeptide may be involved in efflux of secondary bile acids. ABC transporters have been targeted for chemotherapy due to their ability to transport antibiotics and other cytotoxins from the cell (225). This protein, if involved in bile acid efflux, may serve as a potential drug target. A possible approach for studying the function of these genes is insertional inactivation. Recently, a new vector

has been developed specifically for insertional inactivation of clostridial genes, which has been named ClosTron (226). Phenotypic tests could be performed in the presence and absence of bile acids. Ultimately, a goal for future research will be to identify a target for drug development to lower secondary bile acid production in patients with high levels of these cytotoxic, and co-carcinogenic metabolites produced by the human microbiome.

## References

1. Eckburg P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson, and D.A. Relman. 2005. Diversity of the human intestinal flora. *Science*. **308**:1635-1638.
2. Savage D.C. 1977. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Nutr.* **31**:107-133.
3. Whiteman W.B., D. C. Coleman, and W.J. Wiebe. 1998. Prokaryotes: the unseen majority. *Proc. Nat. Acad. Sci. USA*. **95**: 6578-6583.
4. Wilson M. 2005. Microbial Inhabitants of Humans. Cambridge University Press, Cambridge, UK. 1-47; 251-313; 375-392.
5. McGarr S.E., J.M. Ridlon, and P.B. Hylemon. 2005. Diet, anaerobic bacterial metabolism and colon cancer risk: a review of the literature. *J. Clin. Gastroenterol.* **39(2)**: 98-109.
6. Vlahcevic Z.R., D.M. Heuman, and P.B. Hylemon. 1996. Physiology and pathophysiology of enterohepatic circulation of bile acids. In *Hepatology: A Textbook of*

Liver Disease. D. Zakim and T. Boyer, editors. 3<sup>rd</sup> ed, vol. 1. Saunders, Philadelphia, PA. 376-417.

7. Hofmann A.F. 1999. The continuing importance of bile acids in liver and intestinal disease. *Arch. Intern. Med.* **159**: 2647-2658.

8. Cowen A.E., M.G. Korman, A.F. Hofmann, O.W. Cass, and S.B. Coffin. 1975. Metabolism of lithocholate in healthy man. II. Enterohepatic circulation. *Gastroenterology*. **69(1)**: 67-76.

9. Heijghebaert S.M., and A.F. Hofmann. 1986. Influence of the amino acid moiety on deconjugation of bile acid amidates by cholylglycine hydrolase on human fecal cultures. *J. Lipid Res.* **27**:742-752.

10. Tanaka H., H. Hashiba, J. Kok , and I. Mierau. 2000. Bile salt hydrolase of *Bifidobacterium longum* biochemical and genetic characterization. *Appl. Environ. Microbiol.* **66(6)**: 2502-2512.

11. Rossocha M., R. Schultz-Heienbrok, H. von Moeller, J.P. Coleman, and W. Saenger. 2005. Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product. *Biochemistry*. **44(15)**:5739-5748.



12. Elkins C.A., S.A. Moser, and D.C. Savage. 2001. Genes encoding bile salt hydrolases and conjugated bile salt transporters in *Lactobacillus johnsonii* 100-100 and other *Lactobacillus* species. *Appl. Environ. Microbiol.* **147**: 3403-3412.

13. Gopal-Srivastava R., and P.B. Hylemon. 1988. Purification and characterization of bile salt hydrolase from *Clostridium perfringens*. *J. Lipid Res.* **29**: 1079-1085.

14. Grill J.P., F. Schneider, J. Crociani, and J. Ballongue. 1995. Purification and characterization of conjugated bile salt hydrolase from *Bifidobacterium longum* BB536. *Appl. Environ. Microbiol.* **61**(7): 2577-2582.

15. Coleman J.P., and L.L. Hudson. 1995. Cloning and characterization of a conjugated bile acid hydrolase gene from *Clostridium perfringens*. *Appl. Environ. Microbiol.* **61**(7): 2514-2520.

16. Christiaens H., R.J. Leer, P.H. Pouwels, and W.Verstraete. 1992. Cloning and Expression of a conjugated bile acid hydrolase gene from *Lactobacillus plantarum* by using a direct plate assay. *Appl. Environ. Microbiol.* **58**(12): 3792-3798.

17. Elkins C.A., and D.C. Savage. 1998. Identification of genes encoding conjugated bile salt hydrolase and

transport in *Lactobacillus johnsonii* 100-100. *J. Bacteriol.* **180(17)**: 4344-4349.

18. Kim G.B., C.M. Miyamoto, E.A. Meighen, and B.H. Lee. 2004. Cloning and characterization of the bile salt hydrolase genes (*bsh*) from *Bifidobacterium bifidum* strains. *Appl. Environ. Microbiol.* **70(9)**: 5603-5612.

19. Kim G.B., M. Brochet, and B.H. Lee. 2005. Cloning and characterization of a bile salt hydrolase (*bsh*) from *Bifidobacterium adolescentis*. *Biotechnol. Lett.* **27(12)**: 817-822.

20. Glaser P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, F. Chetouani, E. Couve, A. de Daruvar, P. Dehoux, E. Domann, G. Dominguez-Bernal, E. Duchaud, L. Durant, O. Dussurget, K.D. Entian, H. Fsihi, F. Garcia-del Portillo, P. Garrido, L. Gautier, W. Goebel, N. Gomez-Lopez, T. Hain, J. Hauf, D. Jackson, L.M. Jones, U. Kaerst, J. Kreft, M. Kuhn, F. Kunst, G. Kurapkat, E. Madueno, A. Maitournam, J.M. Vicente, E. Ng, H. Nedjari, G. Nordsiek, S. Novella, B. de Pablos, J.C. Perez-Diaz, R. Purcell, B. Rammel, M. Rose, T. Schlueter, N. Simoes, A. Tierrez, J.A. Vazquez-Boland, H. Voss, J. Wehland, and P. Cossart. 2001. Comparative genomics of *Listeria* species. *Science.* **294**:849-852.

21. Dussurget O, D. Cabanes, P. Dehoux, M. Lecuit, C. Buchrieser, P. Glaser, P. Cossart, and the European *Listeria* Genome Consortium. 2002. *Listeria monocytogenes* bile salt hydrolase is a PrfA-regulated virulence factor involved in the intestinal and hepatic phases of listeriosis. *Mol. Microbiol.* **45(4)**: 1095-1106
  
22. Kishinaka M., A. Umeda, and S. Kuroki. 1994. High concentrations of conjugated bile acids inhibit bacterial growth of *Clostridium perfringens* and induce its extracellular cholyglycine hydrolase. *Steroids.* **59**: 485-489.
  
23. Elkins C.A., and D.C. Savage. 2003. CbsT2 from *Lactobacillus johnsonii* 100-100 is a transport protein of the major facilitator superfamily that facilitates bile acid antiport. *J. Mol. Microbiol. Biotechnol.* **6(2)**: 76-87.
  
24. Lundeen S.G., and D.C. Savage. 1992. Multiple forms of bile salt hydrolase from *Lactobacillus* sp. Strain 100-100. *J. Bacteriol.* **174(22)**: 7217-7220.
  
25. Stellwag E.J., and P.B. Hylemon. 1976. Purification and characterization of bile salt hydrolase from *Bacteroides fragilis* subsp. *fragilis*. *Biochim. Biophys. Acta.* **452**: 165-176.

26. De Smet I., L. Van Hoorde, M. Vande Woestyne, H. Christiaens, and W. Verstraete. 1995. Significance of bile salt hydrolytic activities of lactobacilli. *J. Appl. Bacteriol.* **79**: 292-301.
27. Corzo G., and S.E. Gilliland. 1999. Measurement of bile salt hydrolase activity from *Lactobacillus acidophilus* based on disappearance of conjugated bile salts. *J. Dairy. Sci.* **82**: 466-471.
28. De Boever P., and W. Verstraete. 1999. Bile salt deconjugation by *Lactobacillus plantarum* 80 and its implication for bacterial toxicity. *J. Appl. Microbiol.* **87**: 345-352.
29. Grill J.P., C. Cayuela, J.M. Antoine, and F. Schneider. 2000. Isolation and characterization of a *Lactobacillus amylovorus* mutant depleted in conjugated bile salt hydrolase activity: relation between activity and bile salt resistance. *J. Appl. Microbiol.* **89**: 553-563.
30. Tannock G.W., M.P. Dashkevich, and S.D. Feighner. 1989. Lactobacilli and bile salt hydrolysis in the murine intestinal tract. *Appl. Environ. Microbiol.* **55** (7): 1848-1851.
31. Flahaut S, J. Frere, P. Boutibonnes, and Y. Auffray. 1996. Comparison of the bile salts and sodium dodecyl

sulfate stress response in *Enterococcus faecalis*. *Appl. Environ. Microbiol.* **62(7)**: 2416- 2420.

32. Rince A., Y. Le Breton, N. Verneuil, J.C. Giard, A. Hartke, and Y. Auffray. 2003. Physiological and molecular aspects of bile salt response in *Enterococcus faecalis*. *Int. J. Food Microbiol.* **88**: 207-213.

33. Lin J., O. Sahin, L.O. Michel, and Q. Zhang. 2003. Critical role of multidrug efflux pump CmeABC in bile resistance and in vivo colonization of *Campylobacter jejuni*. *Infect Immun.* **71(8)**:4250-4259.

34. Van Eldere J., P. Celis, G. De Pauw, E. Lesaffre, and H. Eyssen. 1996. Tauroconjugation of cholic acid stimulates 7 $\alpha$ -dehydroxylation by fecal bacteria. *Appl. Environ. Microbiol.* **62(2)**: 656-661.

35. Cook A.M., and K. Denger. 2002. Dissimilation of the C<sub>2</sub> sulfonates. *Arch. Microbiol.* **179**:1-6.

36. Lengeler J.W., G. Drews, and H.G. Schlegel. 1999. Biology of the Prokaryotes. Blackwell Science, New York, NY. 756-758.

37. Adamowicz M., P.M. Kelley, and K.W. Nickerson. 1991. Detergent (sodium dodecyl sulfate) shock proteins in *Escherichia coli*. *J. Appl. Bacteriol.* **137**: 229-233.

38. Laue H, M. Friedrich, J. Ruff, and A.M. Cook. 2001. Dissimilatory sulfite reductase (desulfovibridin) of the taurine-degrading, non-sulfate-reducing bacterium *Bilophila wadsworthia* RZATAU contains a fused DsrB-DsrD subunit. *J. Bacteriol.* **183(5)**:1727-1733.
39. Lie T.J., M.L. Clawson, W. Godchaux, and E.R. Leadbetter. 1999. Sulfidogenesis from 2-aminoethanesulfonate (taurine) fermentation by a morphologically unusual sulfate-reducing bacterium, *Desulforhopalus singaporensis* sp. nov. *Appl. Environ. Microbiol.* **65(8)**: 3328-3334.
40. Christl S.U., H.D. Eisner, G. Dusel, H. Kasper, and W. Scheppach. 1996. Antagonistic effects of sulfide and butyrate on proliferation of colonic mucosa: a potential role for these agents in the pathogenesis of ulcerative colitis. *Dig. Dis. Sci.* **41(12)**: 2477-2481.
41. Deplancke B., and H.R. Gaskins. 2003. Hydrogen sulfide induces serum-independent cell cycle entry in nontransformed rat intestinal epithelial cells. *FASEB J.* **17(10)**: 1310-1312.
42. Levitt M.D., J. Furne, J. Springfield, F. Suarez, and E. DeMaster. 1999. Detoxification of hydrogen sulfide and methanethiol in the cecal mucosa. *J. Clin. Invest.* **104(8)**: 1107-1114.

43. Levine J., C.J. Ellis, J.K. Furne, J. Springfield, M.D. Levitt. 1998. Fecal hydrogen sulfide production in ulcerative colitis. *Am. J. Gastroenterol.* **93(1)**: 83-87.
44. Rose P., P.K. Moore, S.H. Ming, O.C. Nam, J.S. Armstrong, and M. Whiteman. 2005. Hydrogen sulfide protects colon cancer cells from chemopreventative agent  $\beta$ -phenylethyl isothiocyanate induced apoptosis. *World J. Gastroenterol.* **11(26)**: 3990-3997.
45. Zhang Y., and P. Talalay. 1994. Anticarcinogenic activities of organic isothiocyanates: chemistry and mechanisms. *Cancer Res.* **54**: 1976-1981.
46. Hardison W.G. 1978. Hepatic taurine concentration and dietary taurine as regulators of bile acid conjugation with taurine. *Gastroenterology.* **75(1)**: 71-75.
47. Sjövall J. 1959. Dietary glycine and taurine on bile acid conjugation in man; bile acids and steroids 75. *Proc. Soc. Exp. Biol. Med.* **100(4)**: 676-678.
48. Magee E.A., C.J. Richardson, R. Hughes, and J.H. Cummings. 2000. Contribution of dietary protein to sulfide production in the large intestine: an in vitro and a controlled feeding study in humans. *Am. J. Clin. Nutr.* **72**: 1488-1494.

49. Gibson G.R., J.H. Cummings, and G.T. Macfarlane. 1991. Growth and activities of sulphate reducing bacteria in gut contents of healthy subjects and patients with ulcerative colitis. *FEMS Microbiol. Ecol.* **86**:103-112.
50. O'Keefe S.D.J., M. Kidd, G. Espitalier-Noel, and P. Owira. 1999. Rarity of colon cancer in Africans is associated with low animal product consumption, not fiber. *Amer. J. Gastroenterol.* **94**(5):1373-1380.
51. Sutherland J.D., and I.A. Macdonald. 1982. The metabolism of primary, 7-oxo- and 7 $\beta$ -hydroxy bile acids by *Clostridium absonum*. *J. Lipid Res.* **23**:726-732.
52. Hirano S., and N. Masuda. 1981. Epimerization of the 7-hydroxyl group of bile acids by the combination of two kinds of microorganisms with 7 $\alpha$ - and 7 $\beta$ -hydroxysteroid dehydrogenase activity, respectively. *J. Lipid Res.* **22**:1060-1068.
53. Macdonald I.A., Y.P. Rochon, D.M. Hutchison, and L.V. Holdeman. 1982. Formation of ursodeoxycholic acid from chenodeoxycholic acid by a 7 $\beta$ -hydroxysteroid dehydrogenase-elaborating *Eubacterium aerofaciens* strain when co-cultured with 7 $\alpha$ -hydroxysteroid dehydrogenase-elaborating organisms. *Appl. Environ. Microbiol.* **44**:1187-1195.



54. Macdonald I.A., E.C. Meier, D.E. Mahony, and G.A. Costain. 1976.  $3\alpha$ ,  $7\alpha$ - and  $12\alpha$ -hydroxysteroid dehydrogenases activities from *Clostridium perfringens*. *Biochim. Biophys. Acta.* **450**: 466-476.
55. Edenharder R., A. Pfützner, and R. Hammann. 1989. Characterization of NAD-dependent  $3\alpha$ - and  $3\beta$ -hydroxysteroid dehydrogenase and of NADP-dependent  $7\beta$ -hydroxysteroid dehydrogenase from *Peptostreptococcus productus*. *Biochim. Biophys. Acta.* **1004**: 230-238.
56. Macdonald I.A., D.E. Mahony, J.F. Jellet, and C.E. Meier. 1977. NAD-dependent  $3\alpha$ - and  $12\alpha$ -hydroxysteroid dehydrogenase from *Eubacterium lentum* ATCC 25559. *Biochim. Biophys. Acta.* **489**: 466-476.
57. Macdonald I.A., J.F. Jellet, D.E. Mahony, and L.V. Holdeman. 1979. Bile salt  $3\alpha$ - and  $12\alpha$ -hydroxysteroid dehydrogenases from *Eubacterium lentum* and related strains. *Appl. Environ. Microbiol.* **37**: 992-1000.
58. Mallonee D.H., M.A. Lijewski, and P.B. Hylemon. 1995. Expression in *Escherichia coli* and characterization of a bile acid- inducible  $3\alpha$ -hydroxysteroid dehydrogenase from *Eubacterium* sp. strain VPI 12708. *Curr. Microbiol.* **30**: 259-263.
59. Wells J.E., and P.B. Hylemon. 2000. Identification and characterization of a bile acid  $7\alpha$ -dehydroxylating

operon in *Clostridium* sp. strain T0931, a highly active 7 $\alpha$ -dehydroxylating strain isolated from human feces. *Appl. Environ. Microbiol.* **66**: 1107-1113.

60. Skalhegg B.A. 1974. On the 3 $\alpha$ -hydroxysteroid dehydrogenase from *Pseudomonas testosteroni*: purifications and properties. *Eur. J. Biochem.* **46**:117-125.

61. Skalhegg B.A. 1975. 3 $\alpha$ -hydroxysteroid dehydrogenase from *Pseudomonas testosteroni*: kinetic properties with NAD and its thionicotinamide analogue. *Eur. J. Biochem.* **50**: 603-609.

62. Edenharder R., M. Pfützner, and R. Hammann. 1989. NADP-dependent 3 $\beta$ -, 7 $\alpha$ - and 7 $\beta$ -hydroxysteroid dehydrogenase activities from a lecithinase-lipase-negative *Clostridium* species 25.11.c. *Biochim. Biophys. Acta.* **1002**: 37-44.

63. Edenharder R., and M. Pfutzner. 1989. Partial purification and characterization of an NAD-dependent 3 $\beta$ -hydroxysteroid dehydrogenase from *Clostridium innocuum*. *Appl. Environ. Microbiol.* **55(6)**: 1656-1659.

64. Akao T., T. Akao, M. Hattori, T. Namba, and K. Kobashi. 1987. Enzymes involved in the formation of 3 $\beta$ , 7 $\beta$ -dihydroxy-12-oxo-5 $\beta$ -cholanic acid from dehydrocholic acid by *Ruminococcus* sp. obtained from human intestine. *Biochim. Biophys. Acta.* **921**: 275-280.

65. Macdonald I.A., D.M. Hutchison, T.P. Forrest, V.D. Bokkenheuser, J. Winter, and L.V. Holdeman. 1983. Metabolism of primary bile acids by *Clostridium perfringens*. *J. Steroid Biochem.* **18**: 97-104.
66. Coleman J.P., W.B. White, M. Lijewski, and P.B. Hylemon. 1988. Nucleotide sequence and regulation of a gene involved in bile acid 7-dehydroxylation by *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **170**: 2070-2077.
67. Gopal-Srivastava R., D.H. Mallonee, W.B. White, and P.B. Hylemon. 1990. Multiple copies of a bile acid-inducible gene in *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **172(8)**: 4420-4426.
68. Hylemon P.B., and T.L. Glass. 1983. Bile acid and cholesterol metabolism. In Human Intestinal Microflora in Health and Disease. D.J. Henteges, editor. Academic Press, New York, NY. 189-213.
69. Baron S.F., C.V. Franklund, and P.B. Hylemon. 1991. Cloning, sequencing, and expression of the gene coding for bile acid 7 $\alpha$ -hydroxysteroid dehydrogenase from *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **173(15)**: 4558-4569.

70. Coleman J.P., L.L. Hudson, and M.J. Adams. 1994. Characterization and regulation of the NADP-linked 7 $\alpha$ -hydroxysteroid dehydrogenase gene from *Clostridium sordellii*. *J. Bacteriol.* **176(16)**: 4865-4874.
71. Bennett M.J., S.L. McKnight, and J.P. Coleman. 2003. Cloning and characterization of the NAD-dependent 7 $\alpha$ -hydroxysteroid dehydrogenase from *Bacteroides fragilis*. *Curr. Microbiol.* **47**: 475-484.
72. Warchol M., L. Car, J.P. Grill, and F. Schneider. 2003. Metabolic changes in *Clostridium absonum* ATCC 27555 accompanying induction of epimerization of a primary bile acid. *Curr. Microbiol.* **47**: 425-430.
73. Yoshimoto T., H. Higashi, A. Kanatani, X. Sheng Lin, H. Nagai, H. Oyama, K. Kurazono, and D. Tsuru. 1991. Cloning and sequencing of the 7 $\alpha$ -hydroxysteroid dehydrogenase gene from *Escherichia coli* HB101 and characterization of the expressed enzyme. *J. Bacteriol.* **173(7)**: 2173-2179.
74. Sherod J.A., and P.B. Hylemon. 1977. Partial purification and characterization of NAD-dependent 7 $\alpha$ -hydroxysteroid dehydrogenase from *Bacteroides thetaiotaomicron*. *Biochimic. Biophys. Acta.* **486**: 351-358.

75. Sutherland J.D., and C.N. Williams. 1985. Bile acid induction of  $7\alpha$ - and  $7\beta$ -hydroxysteroid dehydrogenases in *Clostridium limosum*. *J. Lipid Res.* **26**: 344-350.
76. Macdonald I.A., and P.D. Roach. 1981. Bile salt induction of  $7\alpha$ - and  $7\beta$ -hydroxysteroid dehydrogenases in *Clostridium absonum*. *Biochim. Biophys. Act.* **665**: 262-269.
77. Macdonald I.A., D.M. Hutchison, and T.P. Forrest. 1981. Formation of urso- and ursodeoxy-cholic acids from primary bile acids by *Clostridium absonum*. *J. Lipid Res.* **22**: 458-466.
78. Edenharder R., and T. Knafllic. 1981. Epimerization of chenodeoxycholic acid to ursodeoxycholic acid by human intestinal lecithinase-lipase negative clostridia. *J. Lipid Res.* **22**: 652-658.
79. Hylemon P.B., and J.A. Sherrod. 1975. Multiple forms of a  $7\alpha$ -hydroxysteroid dehydrogenase in selected strains of *Bacteroides fragilis*. *J. Bacteriol.* **122 (2)**: 418-424.
80. Franklund C.V., P. de Prada, and P.B. Hylemon. 1990. Purification and characterization of a microbial, NADP-dependent bile acid  $7\alpha$ -hydroxysteroid dehydrogenase. *J. Biol. Chem.* **265 (17)**: 9842-9849.
81. Medici A, P. Pedrini, E. Bianchini, G. Fantin, A. Guerrini, B. Natalini, and R. Pellicciari. 2002.  $7\alpha$ -OH

epimerization of bile acids via oxido-reduction with *Xanthomonas maltophilia*. *Steroids*. **67**: 51-56.

82. Macdonald I.A., B.A White, and P.B. Hylemon. 1983. Separation of 7 $\alpha$ - and 7 $\beta$ -hydroxysteroid dehydrogenase activities from *Clostridium absonum* ATCC# 27555 and cellular response of this organism to bile acid inducers. *J. Lipid Res.* **24**: 1119-1126.

83. Sutherland J.D., C.N. Williams, D.M. Hutchison, and L.V. Holdeman. 1987. Oxidation of primary bile acids by a 7 $\alpha$ -hydroxysteroid dehydrogenase elaborating *Clostridium bifermentans* soil isolate. *Can. J. Microbiol.* **33**: 663-669.

84. Hirano S., N. Masuda, H. Oda, and H. Mukai. 1981. Transformation of bile acids by *Clostridium perfringens*. *Appl. Environ. Microbiol.* **42**: 394-399.

85. Prabha V., M. Gupta, and K.G. Gupta. 1989. Kinetic properties of 7 $\alpha$ -hydroxysteroid dehydrogenase from *Escherichia coli* 080. *Can. J. Microbiol.* **35**: 1076-1080.

86. Sutherland J.D., L.V. Holdeman, C.N. Williams, and I.A. Macdonald. 1984 Formation of urso- and ursodeoxycholic acids from primary bile acids by a *Clostridium limosum* soil isolate. *J. Lipid Res.* **25**:1084-1089.

87. Tanaka N., T. Nonaka, T. Tanabe, T. Yoshimoto, D. Tsuru, and Y. Mitsui. 1996. Crystal structures of the

binary and ternary complexes of 7 $\alpha$ -hydroxysteroid dehydrogenase from *Escherichia coli*. *Biochemistry*. **35**: 7715-7730.

88. Tanabe T., N. Tanaka, K. Uchikawa, T. Kabashima, K. Ito, T. Nonaka, Y. Mitsui, M. Tsuru, and T. Yoshimoto. 1998. Roles of the Ser146, Tyr159, and Lys163 residues in the catalytic action of 7 $\alpha$ -hydroxysteroid dehydrogenase from *Escherichia coli*. *J. Biochem (Tokyo)*. **124(3)**: 634-641.

89. Harris J.N., and P.B. Hylemon. 1978. Partial purification and characterization of NADP-dependent 12 $\alpha$ -hydroxysteroid dehydrogenase from *Clostridium leptum*. *Biochim. Biophys. Acta*. **528**: 148-157.

90. Macdonald I.A., J.F. Jellet, and D.E. Mahony. 1979. 12 $\alpha$ -hydroxysteroid dehydrogenase from *Clostridium* group P. strain C48-50 ATCC 29733; partial purification and characterization. *J. Lipid Res*. **37**: 992-1000.

91. Edenharder R., and J. Schneider. 1985. 12 $\beta$ -dehydrogenation of bile acids by *Clostridium paraputrificum*, *C. tertium*, and *C. difficile* and epimerization at carbon-12 of deoxycholic acid by cocultivation with 12 $\alpha$ -dehydrogenating *Eubacterium lentum*. *Appl. Environ. Microbiol*. **49**: 964-968.

92. Edenharder R., and A. Pfützner. 1988. Characterization of NADP-dependent 12 $\beta$ -hydroxysteroid dehydrogenase from *Clostridium paraputrificum*. *Biochim. Biophys. Acta.* **962**: 362-370.
93. Reddy B.S. 1981. Dietary fat and its relationship to large bowel cancer. *Cancer Res.* **41**: 3700-3705.
94. Ali S.S., A. Kuksis, and J.M.R. Beveridge. 1966. Excretion of bile acids by three men on a fat-free diet. *Can. J. Biochem.* **44**: 957-969.
95. Begley M., C.G.M. Gahan, and C. Hill. 2005. The interaction between bacteria and bile. *FEMS Microbiol. Rev.* **29(4)**: 625-651.
96. Armstrong M.J., and M.C. Carey. 1982. The hydrophobic-hydrophilic balance of bile salts. Inverse correlation between reverse-phase high performance liquid chromatographic mobilities and micellar cholesterol-solubilizing capacities. *J. Lipid Res.* **23(1)**: 70-80.
97. Heuman D.M., W.M. Pandak, P.B. Hylemon, and Z.R.Vlahcevic. 1991. Conjugates of ursodeoxycholate protect against cytotoxicity of more hydrophobic bile salts: in vitro studies in rat hepatocytes and human erythrocytes. *Hepatology.* **14(5)**: 920-926.



98. Bayerdörffer E., G.A. Mannes, T. Ochsenkühn, P. Dirschedl, B. Wiebecke, and G. Paumgartner. 1995. Unconjugated secondary bile acids in the serum of patients with colorectal adenomas. *Gut*. **36**: 268-273.
99. Salen G., G.S. Tint, B. Eliav, N. Deering, and E.H. Mosbach. 1974. Increased formation of ursodeoxycholic acid in patients treated with chenodeoxycholic acid. *J. Clin. Invest.* **53(2)**: 612-621.
100. Hofmann A.F. 1995. Bile acids as drugs: principles, mechanisms of action and formulations. *Ital. J. Gastroenterol.* **27(2)**: 106-113.
101. Im E., and J.D. Martinez. 2004. Ursodeoxycholic acid (UDCA) can inhibit deoxycholic acid (DCA)-induced apoptosis via modulation of EGFR/Raf-1/ERK signaling in human colon cancer cells. *J. Nutr.* **134(2)**: 483-486.
102. Stellwag E.J., and P.B. Hylemon. 1978. Characterization of 7 $\alpha$ -dehydroxylase in *Clostridium leptum*. *Am. J. Clin. Nutr.* **31**: S243-S247.
103. Ferrari A., C. Scolastino, and L. Baretta. 1977. Activity on bile salts of a *Clostridium bifermentans* cell-free extract. *FEBS Lett.* **75**: 166-168.
104. Wells J.E., F. Berr, L.A. Thomas, R.H. Dowling, and P.B. Hylemon. 2000. Isolation and characterization of

cholic acid 7 $\alpha$ -dehydroxylating fecal bacteria from cholesterol gallstone patients. *J. Hepatol.* **32**: 4-10.

105. Hirano S., R. Nakama, M. Tamaki, N. Masuda, and H. Oda. 1981. Isolation and characterization of thirteen intestinal microorganisms capable of 7 $\alpha$ -dehydroxylating bile acids. *Appl. Environ. Microbiol.* **41**: 737-745.

106. Wells J.E., K.B. Williams, T.R. Whitehead, D.M. Heuman, and P.B. Hylemon. 2003. Development and application of a polymerase chain reaction assay for the detection and enumeration of bile acid 7 $\alpha$ -dehydroxylating bacteria in human feces. *Clin. Chim. Acta.* **331**: 127-134.

107. Kitahara M., F. Takamine, T. Imamura, and Y. Benno. 2000. Assignment of *Eubacterium* sp. VPI 12708 and related strains with high bile acid 7 $\alpha$ -dehydroxylating activity to *Clostridium scindens* and proposal of *Clostridium hylemonae* sp. nov., isolated from human feces. *Int. J. Syst. Evol. Microbiol.* **50** (Pt 3): 971-978.

108. Kitahara M., F. Takamine, T. Imamura, and Y. Benno. 2001. *Clostridium hiranonis* sp. nov., a human intestinal bacterium with bile acid 7 $\alpha$ -dehydroxylating activity. *Int. J. Sys. Evol. Microbiol.* **51** (Pt 1): 39-44.

109. Stellwag E.J., and P.B. Hylemon. 1979. 7 $\alpha$ -Dehydroxylation of cholic acid and chenodeoxycholic acid by *Clostridium leptum*. *J. Lipid Res.* **20**: 325-333.

110. Batta A.K., G. Salen, R. Arora, S. Shefer, M. Batta, and A. Person. 1990. Side chain conjugation prevents bacterial 7 $\alpha$ -dehydroxylation of bile acids. *J. Biol. Chem.* **265(19)**: 10925-10928.
111. White B.A., R.H. Lipsky, R.J. Fricke, and P.B. Hylemon. 1980. Bile acid induction specificity of 7 $\alpha$ -dehydroxylase activity in an intestinal *Eubacterium* sp. *Steroids*. **35**: 103-109.
112. Schmassmann A, H.F. Fehr, J. Locher, J. Lillienau, C.D. Schteingart, S.S. Rossi, and A. F. Hofmann. 1993. Cholylsarcosine, a new bile acid analogue: metabolism and effect on biliary secretion in humans. *Gastroenterology*. **104**: 1171-1181.
113. White B.A., R.J. Fricke, and P.B. Hylemon. 1982. 7 $\beta$ -dehydroxylation of ursodeoxycholic acid by whole cells and cell extracts of the intestinal anaerobic bacterium, *Eubacterium* species VPI 12708. *J. Lipid. Res.* **23**: 145-153.
114. Samuelsson B. 1960. On the mechanism of the biological formation of deoxycholic acid from cholic acid. *J. Biol. Chem.* **235(2)**: 361-366.
115. Lindstedt S., and B. Samuelsson. 1959. Bile acids and steroids. LXXXIII. On the inter-conversion of cholic

and deoxycholic acid in the rat. *J. Biol. Chem.* **234**(8): 2026-2030.

116. Bergstrom S., S. Linstedt, and B. Samuelsson. 1959. Bile acids and steroids. LXXXII. On the mechanism of deoxycholic acid formation in the rabbit. *J. Biol. Chem.* **234**(8): 2022-2025.

117. Björkhem I., K. Einarsson, P. Melone, and P.B. Hylemon. 1989. Mechanism of intestinal formation of deoxycholic acid from cholic acid in humans: evidence for a 3-oxo- $\Delta^4$ -steroid intermediate. *J. Lipid Res.* **30**: 1033-1039.

118. Hylemon P.B., P.D. Melone, C.V. Franklund, E. Lund, and I. Björkhem. 1991. Mechanism of intestinal  $7\alpha$ -dehydroxylation of cholic acid: evidence that allo-deoxycholic acid is an inducible side-product. *J. Lipid Res.* **32**: 89-95.

119. Paone D.A.M., and P.B. Hylemon. 1984. HPLC purification and preparation of antibodies to cholic acid-inducible polypeptides from *Eubacterium* sp VPI 12708. *J. Lipid Res.* **25**: 1343-1349.

120. White B.A., A.F. Cacciapuoti, R.J. Fricke, T.R. Whitehead, E.H. Mosbach, and P.B. Hylemon. 1981. Cofactor requirements for  $7\alpha$ -dehydroxylation of cholic and chenodeoxycholic acid in cell extracts of the intestinal

anaerobic bacterium, *Eubacterium* species VPI 12708. *J. Lipid Res.* **22**: 891-898.

121. Franklund C.V., S.F. Baron, and P.B. Hylemon. 1993. Characterization of the *baiH* gene encoding a bile acid-inducible NADH:Flavin Oxidoreductase from *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **175(10)**: 3002-3012.

122. White W.B., J.P. Coleman, and P.B. Hylemon. 1988. Molecular cloning of a gene encoding a 45,000- dalton polypeptide associated with bile acid 7-dehydroxylation in *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **170(2)**: 611-616.

123. Coleman J.P., W.B. White, and P.B. Hylemon. 1987. Molecular cloning of bile acid 7-dehydroxylase from *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **169(4)**: 1516-1521.

124. White W.B., C.V. Franklund, J.P. Coleman, and P.B. Hylemon. 1988. Evidence for a multigene family involved in bile acid 7-dehydroxylation in *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **170(10)**: 4555-4561.

125. Mallonee D.H., J.L. Adams, and P.B. Hylemon. 1992. The bile acid-inducible *baiB* gene from *Eubacterium* sp. strain VPI 12708 encodes a bile acid-coenzyme A ligase. *J. Bacteriol.* **174**: 2065-2071.

126. Mallonee D.H., and P.B. Hylemon. 1996. Sequencing and expression of a gene encoding a bile acid transporter from *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **178**(24): 7053-7058.
127. Mallonee D.H., and P.B. Hylemon. 1999. Use of a short A/T-rich cassette for enhanced expression of cloned genes in *Escherichia coli*. *Mol. Biotech.* **11**: 27-35.
128. Dawson J.A., D.H. Mallonee, I. Björkem, and P.B. Hylemon. 1996. Expression and characterization of a C<sub>24</sub> bile acid 7 $\alpha$ -dehydratase from *Eubacterium* sp. strain VPI 12708 in *Escherichia coli*. *J. Lipid Res.* **37**: 1258-1267.
129. Baron S.F., and P.B. Hylemon. 1995. Expression of the bile acid-inducible NADH: flavin oxidoreductase gene of *Eubacterium* sp. VPI 12708 in *Escherichia coli*. *Biochim. Biophys. Acta.* **1249**: 145-154.
130. Ye H.Q., D.H. Mallonee, J.E. Wells, I. Björkem, and P.B. Hylemon. 1999. The bile acid-inducible *baiF* gene from *Eubacterium* sp. strain VPI 12708 encodes a bile acid-coenzyme A hydrolase. *J. Lipid Res.* **40**: 17-23.
131. Heider J. 2001. A new family of CoA-transferases. *FEBS Lett.* **509**: 345-349.
132. Berr F., G.A. Kullak-Ublick, G. Paumgartner, W. Munzig, and P.B. Hylemon. 1996. 7  $\alpha$ -dehydroxylating

bacteria enhance deoxycholic acid input and cholesterol saturation of bile in patients with gallstones.

*Gastroenterology*. **111(6)**: 1611-1620.

133. Dowling R.H., M.J. Veysey, S.P. Pereira, S.H. Hussaini, L.A. Thomas, J.A. Wass, and G.M. Murphy. 1997. Role of intestinal transit in the pathogenesis of gallbladder stones. *Can. J. Gastroenterol.* **11(1)**: 57-64.

134. Thomas L.A., M.J. Veysey, G. French, P.B. Hylemon, G.M. Murphy, and R.H. Dowling. 2001. Bile acid metabolism by fresh human colonic contents: a comparison of caecal versus faecal samples. *Gut*. **49(6)**: 835-842.

135. Low-Beer T.S., and S. Nutter. 1978. Colonic bacterial activity, biliary cholesterol saturation, and pathogenesis of gallstones. *Lancet*. **2(8099)**: 1063-1065.

136. Marcus S.N., and K.W. Heaton. 1988. Deoxycholic acid and the pathogenesis of gall stones. *Gut*. **29(4)**: 522-533.

137. Hussaini S.H., S.P. Pereira, G.M. Murphy, and R.H. Dowling. 1995. Deoxycholic acid influences cholesterol solubilization and microcrystal nucleation time in gallbladder bile. *Hepatology*. **22(6)**: 1735-1744.

138. Bernstein H., C. Bernstein, C.M. Payne, K. Dvorakova, and H. Garewal. 2005. Bile acids as carcinogens

in human gastrointestinal cancers. *Mutat. Res.* **589(1)**: 47-65.

139. Reddy B.S., T. Narasawa, J.H. Weisburger, and E.L. Wynder. 1976. Promoting effect of sodium deoxycholate on colon adenocarcinomas in germfree rats. *J. Natl. Cancer Inst.* **56(2)**: 441-442.

140. Pereira M.A., W. Wang, P.M. Kramer, and L. Tao. 2004. DNA hypomethylation induced by non-genotoxic carcinogens in mouse and rat colon. *Cancer Lett.* **212(2)**: 145-151.

141. Narisawa T., N.E. Magadia, J.H. Weisburger, and E.L. Wynder. 1974. Promoting effect of bile acids on colon carcinogenesis after intrarectal instillation of N-methyl-N'-nitro-N-nitrosoguanidine in rats. *J. Natl. Cancer Inst.* **53(4)**: 1093-1097.

142. Zusman I., M. Chevion, and N. Kitrosski. 1992. Effects of N'methyl-N'-nitro-N-nitrosoguanidine and deoxycholic acid on the content of free radicals in rat serum. *Exp. Toxicol. Pathol.* **44(4)**: 187-189.

143. Bayerdörffer E., G.A. Mannes, W.O. Richter, T. Ochsenkühn, B. Wiebecke, W. Kopcke, and G. Paumgartner. 1993. Increased serum deoxycholic acid levels in men with colorectal adenomas. *Gastroenterology.* **104(1)**: 145-51.



144. Zhu Y., P. Hua, S. Rafiq, E.J. Waffner, M.E. Duffey, and P. Lance. 2002.  $\text{Ca}^{2+}$  and PKC-dependent stimulation of PGE2 synthesis by deoxycholic acid in human colonic fibroblasts. *Am. J. Physiol. Gastrointest. Liver Physiol.* **283(3)**: G503-10.
145. Rao Y.P., E.J. Studer, R.T. Stravitz, S. Gupta, L. Qiao, P. Dent, and P.B. Hylemon. 2002. Activation of the Raf-1/MEK/ERK cascade by bile acids occurs via the epidermal growth factor receptor in primary rat hepatocytes. *Hepatology*. **35(2)**: 307-314.
146. Qiao L., E. Studer, K. Leach, R. McKinstry, S. Gupta, R. Decker, R. Kukreja, K. Valerie, P. Nagarkatti, W. El Deiry, J. Molkentin, R. Schmidt-Ullrich, P.B. Fisher, S. Grant, P.B. Hylemon, and P. Dent. 2001. Deoxycholic acid (DCA) causes ligand-independent activation of epidermal growth factor receptor (EGFR) and FAS receptor in primary hepatocytes: inhibition of EGFR/mitogen-activated protein kinase-signaling module enhances DCA-induced apoptosis. *Mol. Biol. Cell.* **12(9)**: 2629-2645.
147. Pai R., A.S. Tarnawski, and T. Tran. 2004. Deoxycholic acid activates beta-catenin signaling pathway and increases colon cell cancer growth and invasiveness. *Mol. Biol. Cell.* **15(5)**: 2156-63.

148. Gupta S., R. Natarajan, S.G. Payne, E.J. Studer, S. Spiegel, P. Dent, and P.B. Hylemon. 2004. Deoxycholic acid activates the c-Jun N-terminal kinase pathway via FAS receptor activation in primary hepatocytes. Role of acidic sphingomyelinase-mediated ceramide generation in FAS receptor activation. *J. Biol. Chem.* **279(7)**: 5821-5828.
149. Makishima M., T.T. Lu, W. Xie, G.K. Whitfield, H. Domoto, R.M. Evans, M.R. Haussler, and D.J. Mangelsdorf. 2002. Vitamin D receptor as an intestinal bile acid sensor. *Science*. **296(5571)**: 1313-1316.
150. Adachi R., Y. Honma, H. Masuno, K. Kawana, I. Shimomura, S. Yamada, and M. Makishima. 2005. Selective activation of vitamin D receptor by lithocholic acid acetate, a bile acid derivative. *J. Lipid Res.* **46(1)**: 46-57.
151. Dar S.A., J.G. Kuenen, and G. Muyzer. 2005. Nested PCR-denaturing gradient gel electrophoresis approach to determine the diversity of sulfate-reducing bacteria in complex microbial communities. *Appl. Environ. Microbiol.* **71(5)**: 2325-2330.
152. Mai V., and J.G. Morris Jr. 2004. Colonic bacterial flora: changing understandings in the molecular age. *J. Nutr.* **134(2)**: 459-464.

153. Kurdi P., H. Tanaka, H.W. Van Veen, K. Asano, F. Tomita, and A. Yokota. 2003. Cholic acid accumulation and its diminution by short-chain fatty acids in bifidobacteria. *Microbiology*. **149(Pt 8)**: 2031-2037.
154. Kurdi P., H.W. van Veen, H. Tanaka, I. Mierau, W.N. Konings, G.W. Tannock, F. Tomita, and A.Yokota. 2000. Cholic acid is accumulated spontaneously, driven by membrane  $\Delta pH$ , in many lactobacilli. *J. Bacteriol.* **182(22)**: 6525-6528.
155. Chandramouli V., K. Kailasapathy, P. Peiris, and M. Jones. 2004. An improved method of microencapsulation and its evaluation to protect *Lactobacillus* spp. in simulated gastric conditions. *J. Microbiol. Methods*. **56(1)**: 27-35.
156. Guerin D., J.C. Vuilleumard, and M. Subirade. 2003. Protection of bifidobacteria encapsulated in polysaccharide-protein gel beads against gastric juice and bile. *J. Food Prot.* **66(11)**: 2076-2084.
157. De Smet I., P. De Boever, and W. Verstraete. 1998. Cholesterol lowering in pigs through enhanced bacterial bile salt hydrolase activity. *Br. J. Nutr.* **79(2)**: 185-194.
158. Hepner G., R. Fried, S. St Jeor, L. Fusetti, and R. Morin. 1979. Hypocholesteremic effect of yogurt and milk. *Am. J. Clin. Nutr.* **32**: 19-24.

159. Hlivak P., J. Odraska, M. Ferencik, L. Ebringer, E. Jahnova, and Z. Mikes. 2005. One-year application of probiotic strain *Enterococcus faecium* M-74 decreases serum cholesterol levels. *Bratisl. Lek. Listy*. **106(2)**: 67-72.
160. Hofmann A.F., J. Sjövall, G. Kurz, A. Radomska, C.D. Schteingart, G.S. Tint, Z.R. Vlahcevic, and K.D. Setchell. 1992. A proposed nomenclature for bile acids. *J. Lipid Res.* **33(4)**: 599-604.
161. Reddy B.S., and E.L. Wynder. 1977. Metabolic epidemiology of colon cancer: fecal bile acids and neutral sterols in colon cancer patients and patients with adenomatous polyps. *Cancer*. **39**: 2533-2539.
162. Paumgartner G., and A. Stiehl. 1976. Falk Symposium 24: Bile Acid Metabolism in Health and Disease. University Park Press, Baltimore, MD. 167-172.
163. Kawamoto K., I. Horibe, and K. Uchida. 1989. Purification and characterization of a new hydrolase for conjugated bile acids, chenodeoxycholytaurine hydrolase, from *Bacteroides vulgatus*. *J. Biochem.* **106**: 1049-1053.
164. Sue D., K.J. Boor, and M. Wiedmann. 2003.  $\sigma^B$ -dependent expression patterns of compatible solute transporter genes *opuCA* and *Imo1421* and the conjugated

- bile salt hydrolase gene *bsh* in *Listeria monocytogenes*. *Microbiology*. **149**: 3247-3256.
165. Hirano S., and N. Masuda. 1982. Characterization of NADP-dependent 7 $\beta$ -hydroxysteroid dehydrogenase from and *Eubacterium aerofaciens*. *Appl. Environ. Microbiol.* **43(5)**: 1057-1063.
166. Masuda N., H. Oda, and H. Tanaka. 1983. Purification and characterization of NADP-dependent 7 beta-hydroxysteroid dehydrogenase from *Peptostreptococcus productus* strain b-52. *Biochim. Biophys. Acta.* **755(1)**:65-9.
167. Macdonald I.A., C.N. Williams, D.E. Mahony, and W.M. Christie. 1975. NAD- and NADP-dependent 7 $\alpha$ -hydroxysteroid dehydrogenase from *Bacteroides fragilis*. *Biochim. Biophys. Acta.* **384**: 12-24.
168. Hylemon P.B., and J. Harder. 1999. Biotransformation of monoterpenes, bile acids, and other isoprenoids in anaerobic ecosystems. *FEMS Microbiol. Rev.* **22**: 475-488.
169. Marmur J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J. Mol. Biol.* **3**:208-218.

170. U.K. Laemmli. 1970 Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**:680-685.
171. M. M. Bradford. 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
172. Eneroth, P. 1963. Thin-layer chromatography of bile acids. *J. Lipid Res.* **4**:11-16.
173. Jones, A.S., Webb, H., Smith, F. 1949. *J. Chem. Soc.* 2164.
174. Danielsson, H., Eneroth, P., Hellström, K., Sjövall, J. 1962. *J. Biol. Chem.* **237**:3657.
175. Killenberg, P.G., and D.F. Duke 1976. Coenzyme A derivatives of bile acids-chemical synthesis,

purification, and utilization in enzymic preparation of taurine conjugates. *J. Lipid Res.* **17(5)**:451-55.

176. Shah, P. P., and E. Staple. 1968. Synthesis of coenzyme A esters of some bile acids. *Steroids*. **12**:571-576.

177. Ridlon, J.M., McGarr, S.E., and P.B. Hylemon. 2005. Development of methods for the detection and quantification of 7 $\alpha$ -dehydroxylating clostridia, *Desulfovibrio vulgaris*, *Methanobrevibacter smithii*, and *Lactobacillus plantarum* in human feces. *Clin. Chim. Acta.* **357**:55-64.

178. Notredame, C., Higgins, D., and J. Heringa. 2000. T-Coffee: A novel method for multiple sequence alignments. *J. Mol. Biol.* **302**:205-217.

179. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410.

180. D.T. Jones. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**:195-202.
181. McGuffin, L.J, Bryson, K., and D.T. Jones. 2000. The PSIPRED protein structure prediction server. *Bioinformatics.* **16**:404-405.
182. Hubbard, P.A., Liang, X., Schulz, H., and J.P. Kim. 2003. The crystal structure and reaction mechanism of *Escherichia coli* 2,4-dienoyl-CoA reductase. *J. Biol. Chem.* **278 (39)**:37553-37560.
183. Liu, X.L., and R. K. Scopes. 1993. Cloning, sequencing and expression of the gene encoding NADH oxidase from the extreme anaerobic thermophile *Thermoanaerobium brockii*. *Biochim. Biophys. Acta* **1174 (2)**:187-190.
184. Rohdich, F., Wiese, A., Feicht, R., Simon, H., and A. Bacher. 2001. Enolate reductases of *Clostridia*. Cloning,



sequencing, and expression. *J. Biol. Chem.* **276(8)**:5779-5787.

185. Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Xu, J., and H. Yang. 2002. A complete sequence of the *T. tengcongensis* genome. *Genome Res.* **12(5)**:689-700.

186. Fitzpatrick, T.B., Auweter, S., Kitzing, K., Clausen, T., Amrhein, N., and Macheroux. 2004. Structural and functional impairment of an Old Yellow Enzyme homologue upon affinity tag incorporation. *Protein Expr. Purif.* **36**:280-291.

187. Li, Y., Chen, J., Liu, J., Yang, X., and K. Wang. 2004. Binding of  $\text{Cu}^{2+}$  to S-adenosyl-L-homocysteine hydrolase. *J. Inorg. Biochem.* **98**:977-983.

188. E. Sulkowski. 1985. Purification of proteins by IMAC. *Trends Biotechnol.* **3**:1-7.

189. Doerner, K.C., Takamine, F., LaVoie, C.P., Mallonee, D.H., and P.B. Hylemon. 1997. Assessment of fecal bacteria with bile acid 7 $\alpha$ -dehydroxylating activity for the presence of *bai*-like genes. *Appl. Environ. Microbiol.*

**63(3):**1185-1188.

190. Narushima, S., Itoha, K., Miyamoto, Y., Park, S.H., Nagata, K., Kuruma, K., and K. Uchida. 2006. Deoxycholic acid formation in gnotobiotic mice associated with human intestinal bacteria. *Lipids*. **41(9)**:835-43.

191. Mallonee, D.H., White, W.B., and P.B. Hylemon. 1990. Cloning and sequencing of a bile acid-inducible operon from *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.*

**172**:7011-7019.

192. Leys, D., Tsapin, A.S., Nealson, K.H., Meyer, T.E., Cusanovich, M.A., and J.J. Van Beeumen. 1999. Structure and mechanism of the flavocytochrome c fumarate reductase of *Shewanella putrefaciens* MR-1. *Nat. Struc. Biol.*

**6(12)**:1113-1117.

193. Blackwell, J.R., and R. Horgan. 1991. A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS lett.* **295**:10-12.

194. Takle, G.W., Toth, I.K., and M.B. Brurberg. 2007. Evaluation of reference genes for real-time RT-PCR expression studies in the plant pathogen *Pectobacterium atrosepticum*. *BMC Plant Biol.* **7**:50-58.

195. Theis T., Skurray R.A., and M.H. Brown. 2007. Identification of suitable internal controls to study expression of a *Staphylococcus aureus* multidrug resistance system by quantitative real-time PCR. *J Microbiol Methods.* **70 (2)**:355-362.

196. Brocks J.J., Logan, G.A., Buick, R., and R.E. Summons. 1999. Archean molecular fossils and the early rise of eukaryotes. *Science* **285**:1033-1036.

197. Moore, W.E.C., and L.V. Holdeman. 1974. Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl. Microbiol.* **27**:961-979.

198. Whitman W.B., Coleman, D.C., and W.J. Wiebe. 1998. Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. USA* **95**:6578-6583.

199. Bäckhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and J.I. Gordon. 2005 Host-bacterial mutualism in the human intestine. *Science* **307**:1915-1920.

200. Ley, R.E., Peterson, D.A., and J.I. Gordon. 2006. Ecology and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**:837-848.

201. Relman, D.A., and S. Falkow. 2001. The meaning and impact of the human genome sequence for microbiology. *Trends Microbiol.* **5**:206-208.

202. Begley, M., Hill, C., and C.G. Gahan. 2006. Bile salt hydrolase activity in probiotics. *Appl Environ Microbiol.*

**72 (3)**:1729-38.

203. Scotti, E., Gilardi, F., Godio, C., Gers, E., Krneta, J., Mitro, N., De Fabiani, E., Caruso, D., and M.

Crestani. 2007. Bile acids and their signaling pathways: eclectic regulators of diverse cellular functions. *Cell Mol Life Sci.*

**64 (19-20)**:2477-91.

204. Kengen, S.W., van der Oost, J., and W. M. de Vos.

2003. Molecular characterization of H<sub>2</sub>O<sub>2</sub>-forming NADH oxidases from *Archaeoglobus fulgidus*. *Eur. J. Biochem.*

**270**:2885-2894.

205. Stott, K., Saito, K., Thiele, D.J., and V. Massey.

1993. Old yellow enzyme: the discovery of multiple isozymes and a family of related proteins. *J. Biol. Chem.*

**268 (9)**:6097-6106.

206. Kuno, S., Bacher, A., and H. Simon. 1985. Structure of enoate reductase from a *Clostridium tyrobutyricum* (C. spec. Lal). *Biol. Chem. Hoppe-Seyler*. **366**:463-472.
207. Dommes V., and W. H. Kunau. 1984. 2,4-Dienoyl coenzyme A reductases from bovine liver and *Escherichia coli*. Comparison of properties, *J. Biol. Chem.* **259**:1781-1788.
208. Saito, K., Thiele, D.J., Davio, M., Lockridge, O., and V. Massey. 1991. The cloning and expression of a gene encoding Old Yellow Enzyme from *Saccharomyces carlsbergensis*, *J. Biol. Chem.* **266(31)**:20720-20724.
209. K. Kitzing, T. B. Fitzpatrick, C. Wilken, J. Sawa, G. P. Bounekov, P. Macheroux, T. Clausen, The 1.3 Å crystal structure of the flavoprotein YqjM reveals a novel class of Old Yellow Enzymes, *J. Biol. Chem.* 280(30) (2005) 27904-27913.
210. Brigé, A., van den Hemel, D., Carpentier, W., De Smet, L., and J. J. Van Beeumen. 2006. Comparative

characterization and expression analysis of the four Old Yellow Enzyme homologues from *Shewanella oneidensis* indicate differences in physiological function. *Biochem. J.* **394**:335-344.

211. Vaz, A., Chakraborty, D.N.S., and V. Massey. 1995. Old yellow enzyme: aromatization of cyclic enones and the mechanism of a novel dismutation reaction. *Biochemistry* **34**:4246-4256.

212. Trotter, E.W., Collinson, E.J., Dawes, I.W., and C. M. Grant. 2006. Old yellow enzymes protect against acrolein toxicity in the yeast *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **72** (7):4885-4892.

213. Haarer, B.K., and D. C. Amberg. 2004. Old yellow enzyme protects the actin cytoskeleton from oxidative stress. *Mol. Biol. Cell.* **15**:4522-4531.

214. Liang, X., Thorpe, C., and H. Schulz. 2000. 2,4-Dienoyl-CoA reductase from *Escherichia coli* is a novel

iron-sulfur flavoprotein that functions in fatty acid  
beta-oxidation. *Arch. Biochem. Biophys.* **380 (2)**:373-379.

215. Brown, B.J., Deng, Z., Karplus, P.A., and V. Massey.  
1998. On the active site of Old Yellow Enzyme. Role of  
histidine 191 and asparagine 194. *J. Biol. Chem.*  
**273 (49)**:32753-32762.

216. Shyamala, V., and G.F. Ames. 1989. Genome walking by  
single-specific-primer polymerase chain reaction: SSP-PCR.  
*Gene*. 1989 **84 (1)**:1-8.

217. Jonsson, S., Ricagno, S., Lindqvist, Y., and N.G.J.  
Richards. 2004. Kinetic and mechanistic characterization  
of the formyl-CoA transferase from *Oxalobacter formigenes*.  
*J. Biol. Chem.* **279 (34)**:36003-36012.

218. Ricagno, S., Jonsson, S., Richards, N., and Y.  
Lindqvist. 2003. Formyl-CoA transferase encloses the CoA  
binding site at the interface of an interlocked dimer.  
*EMBO J.* **22 (13)**:3210-3219.



219. Gogos, A., Gorman, J., and L. Shapiro. 2004. Structure of *Escherichia coli* YfdW, a type III CoA tranferase. *Acta Cryst.* D**60**:507-511.
220. Rangarajan, E.S., Li, Y., Iannuzzi, P., Cygler, M., and A. Matte. 2005. Crystal structure of *Escherichia coli* crotonobetainyl-CoA: Carnitine CoA-transferase (CaiB) and its complexes with CoA and Carnitinylnl-CoA. *Biochemistry* **44**:5728-5738.
221. Carlisle, V.F., and C. Tasman-Jones. 1977. Prevention of gallstone formation in rabbits by the oral administration of kanamycin. *Surg. Gynecol. Obstet.* **144(2)**:195-198.
222. Morimoto, K., Kimura, T., Sakka, K., and K. Ohmiya. 2005. Overexpression of a hydrogenase gene in *Clostridium paraputrificum* to enhance hydrogen gas production. *FEMS Microbiol Lett.* **246(2)**:229-234.

223. Rodrigues, C.M., Kren, B.T., Steer, C.J., and K.D. Setchell. 1996. Formation of delta 22-bile acids in rats is not gender specific and occurs in the peroxisome. *J Lipid Res.* **37(3)**:540-550.
224. Zeng, X., and S. Kaplan. 2001. TspO as a modulator of the repressor/antirepressor (PpsR/AppA) regulatory system in *Rhodobacter sphaeroides* 2.4.1. *J. Bacteriol.* **183(21)**:6355-6364.
225. Tannert, A., Pohl, A., Pomorski, T., and A. Herrmann. 2003. Protein-mediated tranbilayer movement of lipids in eukaryotes and prokaryotes: the relevance of ABC transporters. *Int. J. Antimicrob. Agents.* **22**:177-187.
226. Heap, J.T., Pennington, O.J., Cartman, S.T., Carter, G.P., and N.P. Minton. 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J. Microb. Meth.* **70**:452-464.

**VITA**

Jason Michael Ridlon was born on May 8, 1980 in Vienna, Virginia and is a citizen of the United States of America. Mr. Ridlon graduated from Peninsula Catholic High School, Newport News, Virginia in 1998. He received a Bachelor of Science degree (with honors) from Bridgewater College in Bridgewater, Virginia in 2002. Mr. Ridlon worked as a laboratory technician in the lab of Dr. Phillip Hylemon from Summer 2002 until entering the PhD program in 2005.

*baiCD* and *baiH* data reprinted from Biochimica et Biophysica Acta, 1781(1-2), Dae-Joong Kang, **Jason M. Ridlon**, Doyle Ray Moore II, Stephen Barnes and Phillip B. Hylemon, *Clostridium scindens baiCD* and *baiH* genes encode stereo-specific 7 $\alpha$ /7 $\beta$ -hydroxy-3-oxo- $\Delta^4$ -cholenoic acid oxidoreductases, 16-25, Copyright (2008), with permission from Elsevier.

Introduction reprinted from The Journal of Lipid Research, 47(2), **Jason M. Ridlon**, Dae-Joong Kang, and Phillip B. Hylemon. Bile salt biotransformations by human intestinal bacteria, 241-259, Copyright (2005), with permission from ASBMB Journals.