Differential Regulation of Two Thiolase Genes from Clostridium acetobutylicum DSM 792

Klaus Winzer1, Karin Lorenz2, Brigitte Zickner, and Peter Dürré*

Mikrobiologie und Biotechnologie, Universität Ulm, D-89069 Ulm, Germany
1present address: School of Pharmaceutical Sciences, University Park, University of Nottingham, Nottingham NG7 2RD, UK
2present address: Centeon Pharma GmbH, Postfach 1230, D-35002 Marburg, Germany

Abstract

Thiolase of Clostridium acetobutylicum is an important enzyme involved in both acid and solvent fermentation. Two thiolase genes (thlA and thlB) have been cloned and sequenced from Clostridium acetobutylicum DSM 792, showing high homology to each other and to thiolases of PHA-synthesizing bacteria. The thlA gene is identical to the gene already cloned and sequenced from strain ATCC 824 (Stim-Herndon et al., 1995, Gene 154: 81-85). Using primer extension and S1 nuclease analysis a transcriptional start site was identified 102 bp upstream of the thlA start codon. This site was preceded by a region that exhibits high similarity to the σ70 consensus promoter sequences of Gram-positive and -negative bacteria. Regulation of thlA and thlB was studied at the transcriptional level to elucidate the specific function of each gene. Non-radioactive primer extension analysis using fluorescein-labelled oligonucleotides and Northern blot analysis revealed high levels of thlA transcripts in acid- and solvent-producing cells. During an induced shift of a continuous culture from acid to solvent formation, the transcript level transiently decreased to a minimum, 3 to 7 h after induction. The thlA transcript length is about 1.4 kb, indicating a monocistronic organisation, whereas genetic organization and reverse transcription (RT)-PCR analysis indicated that thlB forms an operon with two other adjacent genes, thlR and thlC. Transcription and regulation of the thlB operon was studied using RT-PCR and showed a very low expression in acid- and solvent-producing cells. Heterologously expressed clostridial ThlB showed high thiolase activity in Escherichia coli. The N-terminal part of ThlR possesses a potential helix-turn-helix motif and shows significant homology to regulatory proteins belonging to the TetR/AcrR family of transcriptional regulators. ThlR possibly acts as a transcriptional repressor of thlB operon expression. The data provide strong evidence that ThlA is involved in the metabolism of both acid and solvent formation, whereas the physiological function of ThlB has yet to be elucidated.

Introduction

Clostridium acetobutylicum, a Gram-positive, obligate anaerobe, spore-forming bacterium has been used for the industrial production of acetone and butanol for a long time. During growth in batch culture the organism typically shows a biphasic fermentation pattern: after producing considerable amounts of acetate and butyrate from sugars in the exponential growth, the organism switches to the production of mainly acetone and butanol shortly before entering the stationary phase (for reviews, see Jones and Woods, 1986; 1989; Dürré and Bahl, 1996). Thiolase (acetyl-coenzyme A [CoA] acetyltransferase, EC 2.3.1.19) from C. acetobutylicum is involved in both the metabolism of acids and solvents, as it catalyzes the condensation of two molecules of acetyl-CoA to one molecule of acetoacetyl-CoA. During acidogenesis the enzyme controls the ratio of acetate and butyrate formation and, thus, the regulation of thiolase activity effects the energy level of the cells in this metabolic state. In solvent-producing cells, acetoacetyl-CoA is converted to acetoacetate in a CoA transferase reaction, leading to the reactivation of previously produced acids. As the level of acetoacetyl-CoA in the cells is low (Boynton et al., 1994), the rate of acid reutilization depends on the rate of acetoacetyl-CoA formation. Important regulatory factors in thiolase activity are thought to be the intracellular concentrations of CoA and acetyl-CoA (Wiesenburg et al., 1988) and regulation might occur mainly at the enzymatic level. It has been shown, however, that activity is maximal in batch cultures after growth ceases (Hartmanis and Gateneck, 1984), suggesting additional regulation at the genetic level.

Two thiolases have been described in C. pasteurianum (Berndt and Schlegel, 1975), Escherichia coli (Jenkins and Nunn, 1987; Yang et al., 1990), and three enzymes have been identified in Ralstonia eutropha (Haywood et al., 1988; Peoples and Sinskey, 1989; Slater et al., 1998; formerly designated Alcaligenes eutrophus). Two thiolase genes have been partially cloned and sequenced from a strain of the acid- and solvent-producing C. beijerinckii (formerly C. acetobutylicum strain NCIMB 8052) (sequence information gratefully received from N. P. Minton, CAMR, Porton Down, UK). In addition, genome-sequencing projects have revealed that Bacillus subtilis, E. coli, Mycobacterium tuberculosis and other bacteria contain three or more thiolase gene homologues. However, it has been suggested that only one thiolase exists in C. acetobutylicum (Wiesenburg et al., 1988; Petersen and Bennett, 1991). In this paper, we present the molecular analysis of two distinct thiolase genes from C. acetobutylicum and their regulation during the metabolic shift from acid to solvent production.
Further Reading

Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at caister.com

• MALDI-TOF Mass Spectrometry in Microbiology
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof

• Aspergillus and Penicillium in the Post-genomic Era
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2

• The Bacteriocins: Current Knowledge and Future Prospects
  Edited by: RL Dorfl, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins

• Omics in Plant Disease Resistance
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr

• Acidophiles: Life in Extremely Acidic Environments
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles

• Climate Change and Microbial Ecology: Current Research and Future Trends
  Edited by: J Marxsen (2016)
  www.caister.com/climate

• Biofilms in Bioremediation: Current Research and Emerging Technologies
  Edited by: G Lear (2016)
  www.caister.com/biorem

• Microalgae: Current Research and Applications
  Edited by: MN Tsaioglou (2016)
  www.caister.com/microalgae

• Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma

• Virus Evolution: Current Research and Future Directions
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusvol

• Arboviruses: Molecular Biology, Evolution and Control
  Edited by: N Vaslikas, DJ Gubler (2016)
  www.caister.com/arbo

• Shigella: Molecular and Cellular Biology
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella

• Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment
  Edited by: AM Roman, H Guasch, MD Balagu (2016)
  www.caister.com/aquaticbiofilms

• Alphaviruses: Current Biology
  Edited by: S Mahalingam, L Herrero, B Herrin (2016)
  www.caister.com/alphav

• Thermophilic Microorganisms
  Edited by: F Li (2015)
  www.caister.com/thermophile

• Flow Cytometry in Microbiology: Technology and Applications
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow

• Probiotics and Prebiotics: Current Research and Future Trends
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics

• Epigenetics: Current Research and Emerging Trends
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015

• Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications
  Edited by: A Burkovski (2015)
  www.caister.com/cory2

• Advanced Vaccine Research Methods for the Decade of Vaccines
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines

• Antifungals: From Genomics to Resistance and the Development of Novel Agents
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals

• Bacteria-Plant Interactions: Advanced Research and Future Trends
  www.caister.com/bacteria-plant

• Aeromonas
  Edited by: J Graf (2015)
  www.caister.com/aeromonas

• Antibiotics: Current Innovations and Future Trends
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics

• Leishmania: Current Biology and Control
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2

• Acanthamoeba: Biology and Pathogenesis (2nd edition)
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2

• Microarrays: Current Technology, Innovations and Applications
  Edited by: Z He (2014)
  www.caister.com/microarrays2

• Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications
  Edited by: D Marco (2014)
  www.caister.com/n2
Results

Identification, Cloning, and Sequencing of Two Thiolarase Genes

Heterologous thiolarase gene probes from C. beijerinckii NCIMB 8052 were generated in PCR reactions using NCIMB 8052 chromosomal DNA and the primer pairs T2A/T2B and T1A/T1B for the thiolase A and the thiolase B probe, respectively (sequence information gratefully received from N. P. Minton). In Southern hybridization experiments, with HindIII-digested total DNA from different C. acetobutyllicum strains and C. beijerinckii NCIMB 8052, both radiolabelled probes produced identical signal patterns. The intensity of individual signals did vary, however, depending on the probe used. For strains ATCC 824, DSM 792, and DSM 1731 two strong signals at 2.6 and 2.0 kbp were observed, whereas very weak signals for the strain P262 and just one signal for C. beijerinckii NCIMB 8052 could be detected, which is in accordance with the observed strain differences already described (Wilkinson and Young, 1993; Wilkinson et al., 1995; Keis et al., 1995; Johnson et al., 1997). For the DSM and ATCC strains the 2.0 kbp signal was dominant with the thiolase A probe, whereas the 2.6 kbp signal was strongest with the thiolase B probe, indicating the existence of two different thiolase genes in these strains. An SspI digest revealed two bands for NCIMB 8052 with both probes, whereas in P262 two signals could only be observed with the thiA gene probe (Figure 1).

Genomic libraries of C. acetobutyllicum DSM 792 containing HindIII- (Sauer et al., 1994) and partially Sau3AI-digested (Gerischer and Dürr, 1990; Fischer et al., 1993) DNA fragments were screened by hybridization using the same probes, and several positive clones were detected. Again, the clones identified with both probes were identical. The thiolase A probe gave strongest signals with three clones containing Sau3AI fragments in vectors pEcoR252 (designated pKS5 and pKS10) and in pUC9 (designated pH12), whereas the strongest signals were obtained with the thiolase B probe for four clones containing HindIII (designated pKH2, pKH13, and pKH16) and Sau3AI fragments (designated pUG1) in pUC9.

Restriction enzyme mapping of pKS5, pKS10, and pH12 revealed insert lengths of app. 7.8 kbp, 7.4 kbp, and 7.2 kbp, respectively. Both ends of the inserts were sequenced and shown to possess identical nucleotide sequences at one end (sequence data not shown), but started at different sites within an open reading frame at the other end. This open reading frame was identical to a thiolase gene already cloned and sequenced from the strain ATCC 824 (Petersen and Bennett, 1991; Stim-Herndon et al., 1995) and was therefore designated as thiA gene (Figure 2). As revealed by sequence comparison, the thiA gene encoded by pKS5 was not complete, and lacked 43 bp at the 5' end. Using HindIII-digested and religated DNA from C. acetobutyllicum DSM 792 as a circular template, inverse PCR was performed with oligonucleotides 2U1 and 2U2 as primers. A 2.0 kbp DNA fragment was obtained and cloned into pUC18 after digestion with HindIII and XbaI. The cloned fragment contained the thiA 5' end and an XbaI-digested (designated pKH13, pKH16) and in pUC9 (designated pUG1) at the amino terminal. This open reading frame was designated as thiA gene (Figure 2), which exhibited a significant similarity to the NfrC protein of E. coli (53 % identity, [SwissProt P27828]), OrlX of B. subtilis (53 % identity, [SwissProt P39131]), Cap5p and Cap8p of Staphylococcus aureus (51 %, [GenBank U81973 and U73374]), EspC of Burkholderia solanacearum (50 % identity, [SwissProt P52641]), and RfbC of Salmonella borreze (48 % identity, [SwissProt P52642]) at the amino acid level. This open reading frame was designated as nfrC (in this context 1 indicates truncated genes or open reading frames). The complete sequence of the thiA gene region is available under accession no. AF072734 in the GenBank data base.

6 bp upstream of the AUG start codon of thiA, a putative ribosome-binding site (5'-AGAGG-3') was identified. A stem-loop structure and a stretch of U's resembling a ribosome-binding site (5'-AGGAGG-3') was identified. A stem-loop structure and a stretch of U's resembling a ribosome-binding site (5'-AGGAGG-3') was identified. A stem-loop structure and a stretch of U's resembling a ribosome-binding site (5'-AGGAGG-3').
frames (Figs. 2 and 3). One of the open reading frames was shown to possess a high similarity to other thiolase genes, but clearly differed from the thlA gene of C. acetobutylicum, and was therefore designated thlB. The deduced amino acid sequences of thlA and thlB from C. acetobutylicum DSM 792 exhibited 76% identity and 86% similarity to each other, and showed highest similarities to other bacterial thiolases, most of which are known to be involved in biosynthesis, rather than degradation, of acetocetyl-CoA (70% identity of ThlA (64% of ThlB) at the amino acid level to ThlA of Thermoanaerobacterium thermosaccharolyticum [EMBL Z82038], 64% (61%) identity to PhbA of Pseudomonas sp. 61-3 [DDBJ AB014757], 63% (63%) identity to AtoB of Haemophilus influenzae [SwissProt P44873], 63% (61%) identity to PhbA of R. eutropha [GenBank J044987], 62% (59%) identity to PhaA of Paracoccus denitrificans [SwissProt P54810], 62% (61%) identity to AtoB of E. coli [SwissProt P76461], 60% (59%) identity to PhbA of Allochromatium vinosum [SwissProt P45369]). Two other open reading frames were located directly upstream and downstream of thlB in the same orientation, and were designated thlR and thlC', respectively. The N-terminal region of ThlR showed significant similarity to proteins belonging to the TetR/AcrR family of transcriptional regulators. A typical helix-turn-helix-motif according to the convention of Pabo and Sauer (1984) is represented by the amino acid sequence INDILSGT (helix) GLP (turn) KGSFYFHFA (helix). The deduced amino acid sequence of the truncated thlC' showed no homology to other sequences available in EMBL and GeneBank data bases, but was highly similar to a protein encoded by a gene downstream of the thiolase B gene from C. beijerinckii (sequence information gratefully received from N. P. Minton). A divergently transcribed open reading frame of unknown function, orfX', was located upstream of thlR. The insert of pUG1 did not represent the original arrangement of Sau3AI fragments, as revealed by comparison of PCR fragments amplified from plasmid and chromosomal DNA (data not shown). Therefore, this plasmid contained only the part of the thlB gene region which is indicated in Figure 2. The complete sequence of the thlB gene region is available under accession no. AF072735 in the GenBank data base.

Putative ribosome-binding sites (5'-AGGAGG-3', 5'-AGGGAG-3', and 5'-AGGAGG-3') were located 8 bp upstream of thlB, 8 bp upstream of thlR, and 6 bp upstream of thlC', respectively. However, no such site was found in front of orfX' (Figure 3).

Determination of Transcription Start Points
The transcription start point of the thlA gene was determined by primer extension (oligonucleotides 2U5 and 2U6) used for RT-PCR are marked by thick solid bars and the resulting PCR fragments are indicated by lines.
Figure 3. Nucleotide sequence of the inserts of pKH2, pKH13, pKH16, and the analyzed part of pUG1 representing the DNA region from \textit{C}. acetobutylicum DSM 792 that contains the \textit{thlB} operon. The non-coding strand is shown. The \textit{thlR}, \textit{thlB}, \textit{thlC}', and \textit{orfX} genes have been translated using the one letter amino acid code and amino acid symbols are written below the first nucleotide of the corresponding codon. Putative ribosome binding sites are underlined; potential promoter regions are marked by thick solid bars below the -10 and -35 regions joined by a line; putative secondary structures are marked by arrows above the sequence indicating the length and orientation of the palindromic sequence, disregarding short mismatches.
mRNA Analysis of the \(thlA\) and \(thlB\) Genes

To obtain mRNA from cells during the transition from acid to solvent production, samples of cell suspension from a continuous culture of \(C.\) acetobutylicum DSM 1731 were taken at various time points before and after induction of solventogenesis. DSM 1731 was chosen because of its very reproducible behavior in continuous culture experiments. Both strains, DSM 792 and 1731, have been shown to be virtually identical based on available transcription signals, gene sequences, biotyping, DNA fingerprint analysis, and DNA-DNA hybridization (Gerischer and Dürr, 1992; Keis et al., 1995; Johnson et al., 1997). pH and product formation were monitored throughout the shift and showed the typical pattern described by Bahl et al. (1982): the pH drop was accompanied by an immediate and drastic decrease in acetate and butyrate concentrations, and then after several hours acetone and butanol were produced. Total RNA was isolated from the cell samples and used for the preparation of Northern blots. Hybridizations, with both radiolabelled \(thlA\) (\(Bgl\|\)HindIII-fragment of plasmid pKS5; bp position 1101-2011 in Figure 2) and \(thlB\) probes (\(HaeIII\|\)HindIII-fragment of plasmid pH2; bp position 1598-2664 in Figure 2), revealed very similar signal patterns (Figure 5). A dominant signal corresponding to a transcript of 1.4 kb was detected in mRNA preparations of acid- and solvent-producing cells. This size is in accordance with a monocistronic organization of \(thlA\). The intensity of this signal decreased to a minimum shortly after induction of solventogenesis (3-7h), but increased again concurrently with the establishment of solvent production. In mRNA preparations from solvent-producing cells another strong signal of 0.9 kb was obtained with both probes. Additional, but less intensive signals of 1.7 and 2.7 kb were also present, particularly with the thiolase B probe (Figure 5), but were also visible with the thiolase A probe after extended exposure of the autoradiogram (not shown). Thus, it was impossible to assign specific signals to the two probes by Northern experiments due to their high sequence homology.

PrA and S1 nuclease analysis (oligonucleotide 2U5). Two signals in positions bp -102 and bp -29, relative to the \(thlA\) start codon were obtained by primer extension experiments using the oligonucleotide 2U5. However, only one transcription start point (position -102) was confirmed by S1 nuclease analysis (Figure 4). The additional signal in position -29 might be due to nonspecific hybridization of 2U5, caused by less stringent hybridization conditions in primer extension analysis, since it could not be verified using the oligonucleotide PrA (data not shown). The identified transcription start point deviated from the start point published by Stim-Herndon et al. (1995) by 3 bp (position -105).

Primer extension analysis (oligonucleotides 1U6, 1U7, 1U8, 1U9, 1U10), and S1 nuclease analysis (oligonucleotides 1U6 and 1U7) were performed to determine the \(thlB\) and \(thlR\) gene transcription start points, but no signals were detected. Since evidence for transcription of the \(thlB\) gene was obtained by RT-PCR (see below), the failure of start point determinations might be attributed to a very low \(thlB\) transcript level in the mRNA preparations from both, acid- and solvent-producing cells.

Figure 5. Northern hybridizations using radiolabelled restriction fragments complementary to the \(thlA\) gene (upper panel) and \(thlB\) gene (lower panel). Each lane contained 10 \(\mu\)g total RNA from \(C.\) acetobutylicum DSM 1731, isolated from cells of a continuous culture, taken at various time points before (0 h = acid-producing cells) and after turning off the pH control. Fragment sizes, as judged by comparison with size markers, are indicated to the right.

Figure 4. Mapping of the 5’ end of \(thlA\) transcripts by S1 nuclease analysis (A) and primer extension analysis (B). (A): \(32\)S-radiolabelled DNA was synthesized using the oligonucleotide PrA complementary to the mRNA of the \(thlA\) gene (upper panel) and \(thlB\) probe after extended exposure of the autoradiogram (not shown). Thus, it was impossible to assign specific signals to the two probes by Northern experiments due to their high sequence homology.
To differentiate between \textit{thlA} and \textit{thlB} transcripts, non-radioactive primer extension analysis was performed, using fluorescein-labelled oligonucleotide PrA, which resulted in the formation of \textit{thlA}-specific cDNA. As the amount of cDNA was proportional to the concentration of \textit{thlA} transcripts in the various mRNA preparations, and the intensity of fluorescence depends on the amount of fluorescein, quantification was possible using an A.L.F. sequencer (Porcher et al., 1992). To verify the typical course of the shift experiment at the transcriptional level, the solventogenesis-associated \textit{adc} and \textit{adhE} transcripts were also analyzed, using the oligonucleotides PrADC and PrADHE. \textit{adc} and \textit{adhE} transcript levels both increased after the induction of solvent formation (Figure 6). The \textit{thlA} transcript level at different time points during the shift corresponded well to the pattern of signal intensity observed at 1.4 kb in the Northern experiments. As the deduced transcript length of \textit{thlA} is about 1.35 kb, it can be concluded that this signal indeed represents the \textit{thlA} transcript and that \textit{thlA} is monocistronically organized.

For identification of \textit{thlB} transcripts and regulatory studies, RT-PCR was performed using the oligonucleotide...
cThioB2 to synthesize the cDNA and various primer pairs were used (3B4/5B2, 3B5/5B2, IRA/5B2, IRB/5B2, and 1R6/5B2) for subsequent amplification. To ensure that the resulting PCR products were amplified from cDNA instead of contaminating chromosomal DNA, control experiments were performed in which reverse transcriptase was omitted. In the controls no PCR fragments were detected. Amplification of cDNA was possible with the primer pair IRB/5B2, but no PCR products were observed with primers 1R6/5B2 (Figure 7), suggesting a 3′ end of the cDNA between positions 850 bp and 1050 bp within the putative thlR gene (Figure 2). Northern hybridization experiments were performed to identify thlR transcripts and to determine if thlR and thlB formed a single transcriptional unit and were responsible for the 1.7 kb and 2.7 kb signals observed with the thiolase B probe. A radiolabelled PCR fragment, synthesized using primers 1U7 and IRB, was used as probe, but no signals were detected. Regulation of the thlB gene was, therefore, studied using RT-PCR. cDNA was amplified in 18 to 20 cycles with the 3B4 and 5B2 primers, resulting in fluorescein-labelled PCR fragments. Figure 7 shows the relative amount of PCR products detected on an A.L.F. sequencer. Quantification, achieved by integrating the peak areas, showed that transcription of thlB was maximal at about 10 to 15 h after induction of solvent formation. However, the level of thlB transcripts was very low, compared to the level of thlA transcripts, as PCR amplification of thlB-specific cDNA was required for detection.

**Enzymatic Activity of the thlB Gene Product**

In order to obtain conclusive results for the function of the thlB gene product, this gene was subcloned in several E. coli expression plasmids. A 1302-bp TaqI-HindIII fragment (nt positions 1359-2660 in Figure 3) was cut and isolated from pKH2 and subsequently ligated into XbaI-HindIII-digested pUC18 and pUC19, yielding plasmids pUC18-thlB and pUC19-thlB. From the former, an app. 1.3-kbp XbaI-HindIII thlB fragment was cut, isolated, and ligated into XbaI/HindIII-digested pET21, yielding plasmid pET21-thlB. Recombinant E. coli strains carrying the three newly constructed plasmids and a pET21 control were checked for expression of the ThlB protein by SDS polyacrylamide gel electrophoresis (PAGE) with and without induction by isopropyl-β-D-thiogalactoside (IPTG). The plasmids carrying the thlB insert showed in all cases an additional protein band at the expected size of app. 42 kDa, which was missing in the pET21 control (Figure 8). High specific activities of more than 4 U/mg protein could be detected in strains carrying pUC18-thlB and pET21-thlB, whereas the respective plasmids without inserts yielded values of only 0-0.05 U/mg protein (Table 1). pUC19-thlB with reverse insertion of the fragment with respect to the lac promoter showed a low activity (0.4 U/mg). The low expression from this plasmid and the two other recombinant vectors without IPTG induction is most probably due to the AT-rich region upstream of the thlB gene. This will mimic a -10 region of a typical E. coli promoter. Although a -35 region is definitely missing, such a promoter half site might initiate the weak transcription observed. In summary, the subcloning of thlB and enzyme assays of ThlB provide conclusive evidence for its function as a thiolase.

**Discussion**

Two thiolase genes have been identified in C. acetobutylicum DSM 792, thlA and thlB, which are very similar to each other and show high homology to other biosynthetic bacterial thiolases. Two or more different thiolases have also been identified in other bacteria (Berndt and Schlegel, 1975; Jenkins and Nunn, 1987; Haywood et al., 1988; Peoples and Sinskey, 1989; Yang et al., 1990; Slater et al., 1998), but it has been assumed that only one is present in C. acetobutylicum (Wiesenborn et al., 1988; Petersen and Bennett, 1991). Previous Southern hybridization results (Petersen and Bennett, 1991), performed under rather stringent conditions, revealed no evidence of additional thiolase genes and Western blot analysis of whole-cell extracts revealed only one band with a thiolase-specific antibody (Petersen and Bennett, 1991). In addition, no indications of a second thiolase were found during purification of the enzyme by Wiesenborn et al. (1988). In this work the existence of two thiolase genes has been demonstrated using heterologous thiolase probes from C. beijerinckii in Southern blot hybridizations and subsequent cloning and sequencing of the respective genes. Since ThlA and ThlB are very similar in molecular size and amino acid sequence, both enzymes might have been indistinguishable in SDS gel electrophoresis and in Western blot analysis, leading to the assumption that only one enzyme is present in this organism.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Enzymatic activity [U/mg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC18</td>
<td>0.04</td>
</tr>
<tr>
<td>pUC18-thlB</td>
<td>4.73</td>
</tr>
<tr>
<td>pUC19</td>
<td>0.0</td>
</tr>
<tr>
<td>pUC19-thlB</td>
<td>0.39</td>
</tr>
<tr>
<td>pET21</td>
<td>0.05</td>
</tr>
<tr>
<td>pET21-thlB</td>
<td>4.44</td>
</tr>
</tbody>
</table>
The thlA gene of *C. acetobutylicum* DSM 792 is identical to the thiolase gene already cloned and sequenced from the strain ATCC 824 and codes for the enzyme purified by Wiesenborn *et al.* (1988), as deduced from the N-terminal amino acid sequence (Wiesenborn *et al.*, 1988; Petersen and Bennett, 1991). An incomplete open reading frame, designated as nfrC, was identified upstream of thlA, which is in agreement with the results described by Stirm-Herndon *et al.* (1995) for *C. acetobutylicum* ATCC 824. NfrC from *C. acetobutylicum* shows high similarity to NfrC from *E. coli* (Kiino *et al.*, 1993) and OrfX (Soldo *et al.*, 1993) from *B. subtilis* as well as other proteins which have been postulated to act as UDP-N-acetyl-glucosamine-2-epimerase. Therefore, a similar function might be expected for NfrC from *C. acetobutylicum*.

The thlA mRNA start point suggested a putative promoter region, TTGATA (N)$_{17}$ TATAAT, with high similarity to the $\sigma^{70}$ consensus sequence of Gram-negative and -positive bacteria. Attempts to determine the start point for thlB failed, possibly due to very low transcript levels. PCR amplifications, using several primer combinations, provided evidence that one cDNA end is located at least several hundred base pairs upstream of thlB in the thlR gene. However, the length of the cDNA does not necessarily represent the full length of the mRNA. Premature termination of reverse transcription can be caused by several factors, including even short RNA stem loop structures (Boorstein and Craig, 1989). For that reason it can be postulated that the mRNA start point is located upstream of thlR, where a putative $\sigma^{70}$ promoter region (TTGACA (N)$_{19}$ TATAAT) was identified on the basis of DNA sequence analysis (Figure 3). Putative promoter regions upstream of thlB and thlC could not be deduced from the nucleotide sequence, whereas ribosome-binding sites were found upstream of thlR, thlB, and thlC, suggesting a polycistronic operon organization of these genes. Nevertheless, the existence of an additional promoter region directly upstream of thlB could not be ruled out by the RT-PCR approach used in this study, since only the length of the longest cDNA product could be estimated.

For investigation of the expression pattern and physiological function of ThlA and ThlB, mRNA was isolated from cells from a continuous culture taken at various time points during the shift from acid to solvent production. From Northern blot analyses almost identical signal patterns were seen with thlA- and thlB-specific probes, probably due to the high sequence similarity of both genes and the abundance of transcripts. Additional primer extension experiments showed that the dominant 1.4 kb signal was specific for thlA transcripts, which is in good agreement with the hypothetical transcript length deduced from the mRNA start point determination and the putative transcription termination structure. In solvent-producing cells another strong signal of 0.9 kb was observed, possibly representing a degradation product of the thlA transcript. Additional weak signals of 1.7 kb and 2.7 kb were, however, preferably detected with the thlB-specific probe. Since these signals could not be reproduced with thlR-specific probes, it is likely that they are nonspecific and, thus, might not represent transcripts of the proposed thlB operon.

Quantification of fluorescein-labelled PCR products on an automatic DNA sequencer has been described previously for the determination of mRNA copy number (Porcher *et al.*, 1992). With this method the range over which the recorded signal is proportional to the amount of labelled product is far greater than the scanning of autoradiograms. It is also less problematic than counting the radioactivity of gel or blot fragments. Since the level of thlA transcripts was high during acid- and solvent-production, fluorescence of thlA-specific cDNA could be detected directly, whereas thlB-specific cDNA had to be amplified. The transient, but marked decrease of the thlA transcript level, with a minimum 3 to 7 h after induction of solventogenesis, can be explained as a result of gene regulation, which has not been observed at the level of enzyme activity determined in crude extracts of batch cultures (Hartmanis and Gatenbeck, 1984). For adhE and adc, which are specifically expressed during solvent formation, the same induction pattern was observed as described previously (Gerischer and Dürr, 1992; Sauer and Dürr, 1995). This indicates that the drop in thlA transcript level was not an artefact due to damaged mRNA preparations or caused by a general decrease in gene transcription during the transition to solvent formation. Since the putative promoter upstream of thlA shows high homology to the $\sigma^{70}$ consensus sequence and the gene is expressed in *E. coli* (Petersen and Bennett, 1991), the observed changes in thlA transcript level can be best explained by the action of a repressor protein. The increase in expression during solventogenesis would then require an inactivation of the repressor or an overruling activity of an activator. Identification of a putative regulator protein and its DNA-binding site is currently attempted. This transcriptional regulation of thlA (high expression during acidogenesis, drastic decrease at initiation of the shift, increase of expression at the onset of solventogenesis) is so far unique in *C. acetobutylicum*. The transcriptional analysis presented in this study clearly demonstrates that ThlA is involved in the metabolism of both, acid and solvent formation.

RT-PCR revealed maximal thlB transcript levels 10 hours after the induction of the shift from acid- to solvent formation, being two- to threefold higher compared to RNA from either acid- or solvent-producing cells. Since extensive PCR amplification was necessary to obtain signal intensities comparable to those obtained directly from thlA-specific cDNA, the level of thlB transcripts must be very low. Thus, thlB does not seem to play a significant role in acid and solvent metabolism under the conditions employed in this work.

The physiological function of thlB is still unknown, although the gene product was clearly shown to exhibit thiolase activity. The codon usage of thlB is in agreement with those of highly expressed genes involved in the metabolism of acid- and solvent-production (data not shown), indicating a rapid expression under specific conditions. A second thiolase gene was also identified in several other strains of solvent-producing clostridia (Figure 1). Furthermore, thlC is very similar to a gene downstream of thlB in *C. beijerinckii* (sequence information gratefully received from N. P. Minton), thus analogous operons seem to exist in at least some of these strains. Since in *C. acetobutylicum* production of PHA-like storage compounds has not been described, the function of an additional biosynthetic thiolase, which is not involved in acid- or solvent formation, remains obscure, whereas the presence of two genes in aerobic or PHA-synthesizing bacteria can be explained by the specific function of the respective
enzymes: different thiolases may be involved in the degradation of fatty acids via \( \beta \)-oxidation compared to those involved in the biosynthesis of PHA from acetyl-CoA. In *E. coli* two thiolases are involved in the degradation of short and long chain fatty acids (Jenkins and Nunn, 1987; Yang *et al.*, 1990). For *B. subtilis* a \( \sigma^C \)-dependent operon of at least five open reading frames has been identified, coding for gene products homologous to thiolase, 3-hydroxybutyryl-CoA dehydrogenase, acyl-CoA dehydrogenase and a citrate synthase (Bryan *et al.*, 1996).

This operon is transcribed in the mother cell compartment at an intermediate stage of sporulation, but as of yet no specific function has been assigned. For *thlR* and *thlB* no obvious \( \sigma^C \) consensus promoter sequence could be identified, but the differential expression of these genes during sporulation of *C. acetobutylicum* has still to be investigated. Future work will focus on the functional analysis of the *thlB* operon.

Expression of the *thlB* operon is possibly controlled by a regulator encoded by the *thlR* gene. The \( N \)-terminal region of ThlR shows significant similarity to transcriptional regulators of the TetR/ACrR family, the most similar proteins (beside some hypothetical TetR-like proteins) being the A-factor receptor homologues CprA and CprB of *Streptomyces coelicolor* A3(2) (Onaka *et al.*, 1998), and LmrR of *B. subtilis* (Kumano *et al.*, 1997). All TetR-like proteins that have been analyzed for their regulatory function have been shown to act as repressors. Thus, it seems likely that ThlR has DNA-binding properties and possibly acts as a transcriptional regulator of the *thlB* operon. A putative \( \sigma^70 \) promoter region was identified upstream of *thlR* on the basis of sequence comparisons, suggesting that ThlR might regulate the expression of the operon in some sort of autoregulation. Regulators of the TetR/ACrR family have a particular theme in common: they bind a small molecule, an antibiotic or signal molecule, which after entering the cell will bind to and inactivate the respective repressor protein. It is therefore tempting to speculate that the *thlB* operon is expressed in response to the extracellular accumulation of a signal molecule or other compounds. This possibility is currently under investigation.

**Experimental Procedures**

**Bacterial Strains, Plasmids, and Growth Conditions**

*C. acetobutylicum* DSM 792, DSM 1731 (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany), ATCC 824 (kindly supplied by E. T. Papoutsakis, Northwestern University, Evanston, Ill., P 262 (kindly supplied by D. R. Woods, University of Cape Town, Cape Town, South Africa), and C. beijerincki/NCIMB 8052 (National Collection of Industrial and Marine Bacteria, Aberdeen, UK) were used as a source of genomic DNA for Southern hybridizations and PCR amplifications. Total DNA was isolated from continuous cultures of *C. acetobutylicum* DSM 1731. In cloning experiments, E. coli JM109 was used as the host for pUC18- and pUC19-based vectors (Yanisch-Perron *et al.*, 1985), whereas E. coli BL21 DE3 served as host for pET21 (Novagen Inc., Madison, Wis.) and its derivative.

In batch culture, *C. acetobutylicum* was grown under strictly anaerobic conditions at 37 °C in clostridial basal medium (O’Brien and Morris, 1971). Continuous cultures were performed with strain DSM 1731, as described by Bahi *et al.* (1982). Solventogenesis was induced by turning off the pH control of the continuous culture. The production of acids caused a pH drop to 4.3, at which point the pH control was turned on again. E. coli was cultivated at 37 °C in LB medium or terrific broth medium (Sambrook *et al.*, 1989) supplemented with ampicillin (50 \( \mu \)g/ml), isopropyl-\( \beta \)-D-thiogalactopyranoside (IPTG; 50 \( \mu \)g/ml), or 5-bromo-4-chloro-3-indolyl-\( \beta \)-D-galactopyranoside (X-Gal; 40 \( \mu \)g/ml) when required.

**Nucleic Acid Isolation and Manipulation**

DNA of *C. acetobutylicum* was isolated by the method of Marmur (1960) as modified by Bertram and Dürre (1989) or by using Genomic-tip 100/G columns (Qiagen GmbH, Hilden, Germany) according to the manufacturer’s instructions. For isolation of plasmid DNA from *E. coli* the Plasmid-Midi-Kit (Qiagen) was used. Total RNA of *C. acetobutylicum* was isolated by the hot phenol-chloroform procedure described by Oelmüller *et al.* (1990) with the modifications described by Gerischer and Dürre (1992).

DNA was manipulated by standard methods (Sambrook *et al.*, 1989). Restriction enzymes (MBI Fermentas, St. Leon-Rot, Germany) were used in accordance to the manufacturer’s instructions. Transformation was achieved by electroporation using a Gene Pulser (BioRad Laboratories GmbH, Munich, Germany). Conditions were as follows: Electro-competent cells were prepared as described by Sambrook *et al.* (1989). 2 \( \mu \)l of DNA solution were mixed with 40 \( \mu \)l of competent cells, transferred into 0.2-cm electroporation cuvettes (Bio-Rad), and transformed using the following conditions: 2.5 kV, 200 \( \Omega \), 25 \( \mu \)F. Cells were then suspended in 1 ml LB, incubated for 1 h at 37 °C and plated on appropriate selective media.

**Hybridization**

Total DNA of *C. acetobutylicum* was digested to completion with the respective restriction enzymes, separated on agarose gels and transferred to nylon membranes (GeneScreen Plus; NEN Research Products, Dreieich, Germany) or, for non-radioactive hybridizations, Hybond-N*" (Amerham Buchler GmbH & Co KG, Braunschweig, Germany) by capillary transfer in 10x SSC (1x SSC is 0.15 M NaCl plus 15 mM sodium citrate). Prehybridization and hybridization were performed as described previously (Winzer *et al.*, 1997). Non-radioactive DNA labelling and detection was achieved using an ECL direct nucleic labeling and detection systems kit (Amerham Buchler) according to the manufacturer’s instructions. Total RNA for Northern (RNA) blots was separated in denaturing formaldehyde gels and transferred to nylon membranes (GeneScreen Plus) as described by Sambrook *et al.* (1989). Fragment sizes were estimated by comparison with an RNA ladder ( Gibco/BRL GmbH, Eggenstein, Germany). Hybridization and washing were performed as described (Winzer *et al.*, 1997).

**DNA Sequencing**

For radioactive sequencing a T7 sequencing kit (Amerham Pharmacia Biotech Europe GmbH, Freiburg, Germany) was used. The deoxy-terminated fragments were separated on 55-cm wedge-shaped thickness-gradient gels (0.2 to 0.4 mm; 6% [wt/vol] polyacrylamide) with a Macrophor gel grade. Hybridization and washing were performed as described (Winzer *et al.*, 1997).

**Oligonucleotides**

Synthesis of oligonucleotides was performed on a Gene Assembler Plus (Amerham Pharmacia Biotech) on 0.2 \( \mu \)mol-capacity columns as recommended by the manufacturer. 5'-fluorescent-labelled oligonucleotides were synthesized by using fluorescein amine (Fluorprime®, Amerham Pharmacia Biotech). After deprotection oligonucleotides were purified by column purification on Sephadex G-25 (NAP™ 5; Amerham Pharmacia Biotech). The sequence of oligonucleotides is given in the 5’ to 3’ direction, with the “F” at the 5’ end indicating a fluorescein group. Oligonucleotides for primer extension and S1 nuclease analysis: PrA (FGCGTTGTTCTAGCGACAG, nt positions 1076-1098 in AF072734), PrAD (FAATTGCGGCTGATATTGTTT, nt positions 3557-3576 in M53532), PrAc (GGTGCATTTTTACCATGAG, nt positions 1076-1098 in X72831), PrB (FTACTCTGAGAAGATGAC, nt positions 894-913 in AF072735), 2U5 (TTTATACTGCTGAGTCC, nt positions 1145-1161 in AF072734), 1U6 (TTAAGCTGCCAG, nt positions 1610-1626 in AF072735), 1U7 (GGTGGTTATTCGAT, nt positions 1357-1373 in AF072735), 1U8 (CTTCTATATGGCAA, nt positions 1193-1209 in AF072735), 1U9 (TTTATTTTATCCATC, nt positions 933-949 in AF072735), and 1U10 (GCGCTTGGTTATCTGCG, nt positions 728-744 in AF072735). Oligonucleotides for PCR: T1A (ATGTTAGGGATGGATC, based on sequence information from C. beijerincki) and T1B (AATGATTACCTCACCTG, based on sequence information from C. beijerincki) were used. General primer information from C. beijerincki (Bearden, N. P. Minton), T1B (ATGATTACCTCACCTG, based on sequence information from C. beijerincki) were used. General primer information from C. beijerincki (Bearden, N. P. Minton), T1B (ATGATTACCTCACCTG, based on sequence information from C. beijerincki) were used.
**Sequencing**

The sequence data reported here have been submitted to the GenBank database and assigned accession no. AF072734 (thiA) and AF072735 (thiB gene region).

**Acknowledgements**

We thank J. D. Santangelo and B. Farthmann for stimulating discussions. N. P. Minton for C. beijerincki/thiilogen genes sequence information, and A. Kuhn for excellent technical assistance in continuous culture experiments.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft and travel support by the DAAD within the British-German ARC program.

**References**


Gerischer, U., and Dürre, P. 1992. mRNA analysis of the *ade* gene region of *Clostridium acetobutylicum* during the shift to solventogenesis. J.

**Computer Programs**

Sequence analysis and database searches were performed with Intelligenetics (Menlo Park, Calif.) and Genetics Computer Group (Madison, Wis.) software packages and the NCBI BLAST (http://www.ncbi.nlm.nih.gov) server. For sequence comparisons the programs Gap (complete protein sequences) or BestFit (for truncated protein sequences) were used. The resulting peaks of non-nanopore active primer extension and RT-PCR experiments were integrated by using the Fragment Manager program from Amersham Pharmacia Biotech.


