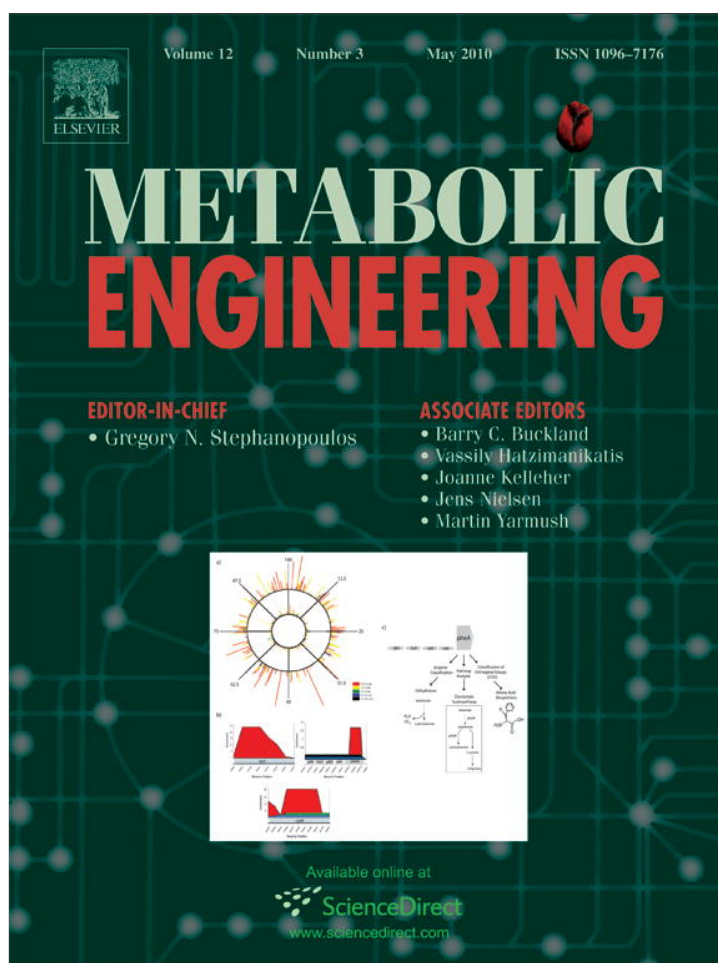


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

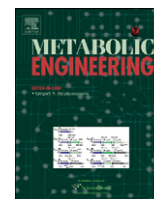


This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



A genomic-library based discovery of a novel, possibly synthetic, acid-tolerance mechanism in *Clostridium acetobutylicum* involving non-coding RNAs and ribosomal RNA processing

Jacob R. Borden^{a,1,2}, Shawn W. Jones^{a,b}, Dinesh Indurthi^b, Yili Chen^b, Eleftherios Terry Papoutsakis^{a,b,*}

^a Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA

^b Department of Chemical Engineering & Delaware Biotechnology Institute, University of Delaware, 15 Innovation Way, Newark, DE 19711, USA

ARTICLE INFO

Article history:

Received 10 September 2009

Received in revised form

30 November 2009

Accepted 21 December 2009

Available online 6 January 2010

Keywords:

Genomic library

rRNA processing

Non-coding RNA

Acid tolerance

Synthetic phenotype

ffh

Signal recognition particle

Sense and antisense libraries

ABSTRACT

We generated a genomic library from sheared *Clostridium acetobutylicum* ATCC 824 DNA, whereby inserts can be expressed in both directions from the thiolase promoter, P_{thi} . Serial transfer of library-bearing *C. acetobutylicum* cultures exposed to increasing butyrate concentrations enriched for inserts containing fragments of rRNA genetic loci. The selected library inserts were placed so that antisense (to the rRNAs) non-coding RNAs (ncRNAs) would be transcribed from P_{thi} . Different enriched inserts imparted similar butyrate-tolerance characteristics. A minimal tolerance fragment (RDNA7) was identified as the 16S-rRNA promoter region. Expressed on plasmid pRD7 off P_{thi} , RDNA7 can produce putative ncRNAs termed ncRNA_{RD7}. *C. acetobutylicum* 824(pRD7) showed superior resistance to butyrate and other carboxylic acids. Transcriptional analysis of butyrate stress identified 120 differentially expressed genes between 824(pRD7) and 824(pSOS95del). The few upregulated genes included the *ffh* gene of the putative signal recognition particle (SRP) system. Northern analysis of ncRNA_{RD7} and corresponding antisense RNAs demonstrated multiple ncRNA_{RD7} molecules in 824(pRD7). Several corresponding antisense RNA molecules were identified both in 824(pRD7) and 824(pSOS95del), but at much higher levels in 824(pRD7). Northern analysis of 16S rRNA expression suggested complex RDNA7-dependent rRNA processing. Our data suggest that by hybridizing against unprocessed rRNA precursors, ncRNA_{RD7} alters rRNA processing, and these alterations result in acid tolerance, possibly through a mechanism involving the Ffh protein.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

For fermentative bacteria, such as *Clostridium acetobutylicum*, substrate-level phosphorylation leading to acid production is the primary means of ATP generation. Acid accumulation during active growth is toxic to cells (Papoutsakis et al., 1987; Russell, 1992; Warnecke and Gill, 2005). Although clostridia are generally viewed as acid tolerant, accumulation of acetate and butyrate lowers the culture pH leading either to a genetic switch to solvent production in solventogenic clostridia (e.g., *C. acetobutylicum* ATCC 824, *C. beijerinckii*, and *C. cellulolyticum*) (Paredes et al., 2005; Zhao et al., 2005) or cessation of metabolism in non-solventogenic clostridia (*C. acetobutylicum* M5, *C. butyricum*, and *C. tyrobutyricum*). Undissociated organic-acid metabolites diffuse freely across the

cell membrane and affect cellular physiology through both free proton and anion interactions (Herrero et al., 1985; Kell et al., 1981; Papoutsakis et al., 1987; Russell, 1992; Russell and Diez-Gonzalez, 1998; Walter and Gutknecht, 1984; Warnecke and Gill, 2005). Excess free protons tend to dissipate the membrane proton gradient, while also potentially impacting purine bases of DNA (Russell, 1992; Russell and Diez-Gonzalez, 1998). Thus, short-chain fatty acids like acetate and butyrate have an antibacterial activity, which likely plays a role in the health of the low-pH human gut environment by protecting against the growth of pathogens; however, bacteria that adapt to high concentrations of these acids may have an advantage in pathogenesis (Russell and Diez-Gonzalez, 1998). In the context of bioprocessing for the production of chemicals and biofuels such as butyrate and butanol, development of acid-tolerant strains is essential for achieving higher cell densities and growth (Papoutsakis, 2008).

A major genetic system contributing to acid tolerance characterized in such diverse organisms as *E. coli*, *C. perfringens*, *Chlamydia pneumoniae*, and *Salmonella enterica* (Foster, 2004) involves the decarboxylation of the amino acids glutamate and arginine. Amino acid decarboxylation results in CO₂ production,

* Corresponding author at: Delaware Biotechnology Institute, University of Delaware, 15 Innovation Way, Newark, DE 19711, USA. Fax: +1 302 831 4841.

E-mail address: epaps@udel.edu (E. Terry Papoutsakis).

¹ The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint first authors.

² Current address: BP America, Naperville, IL, USA.

proton consumption, and generation of either γ -amino butyric acid (GABA) or agmatine (by decarboxylation of glutamate and arginine, respectively) which are exchanged by an antiporter for glutamate or arginine, respectively (Foster, 2004; Richard and Foster, 2004). Therefore decarboxylation coupled with antiport of the decarboxylation products for additional amino acid reactants results in a net efflux of cytoplasmic protons and an increase in cytoplasmic pH. The genes involved include *gadA*, *gadB*, and *gadC*, which code for two decarboxylation isozymes and the cognate glutamate/GABA antiporter, respectively, and *adiA* and *adiC* which code for an arginine decarboxylase and arginine/agmatine antiporter, respectively (Foster, 2004). Although no such decarboxylase/antiporter system has yet been characterized in *C. acetobutylicum*, the gene CAC3285 shows 44% sequence identity with the *E. coli* gene *gadC*, while several *C. acetobutylicum* genes show some homology to *adiA* (CAC0297 [18% identity] and CAC2338 [19% identity]) and *adiC* (CAC0727 [22% identity], CAC0852 [20% identity], CAC2719 [20% identity], CAC3164 [21% identity], and CAC3347 [22% identity]).

An emerging mechanism of acid tolerance has been investigated in *Streptococcus mutans* (Gutierrez et al., 1999; Kremer et al., 2001), and involves the fifty-four homologue (Ffh) protein, a GTPase. Loss of Ffh activity results in loss of acid tolerance (Kremer et al., 2001). Ffh is part of the prokaryotic counterpart to the well-studied mammalian system of signal recognition particle (SRP), which is responsible for the cotranslational membrane targeting of signal-peptide-bearing secretory and membrane proteins to the plasma membrane. In eubacteria, the SRP ribonucleotide–protein complex is composed of a small cytoplasmic RNA (scRNA) and Ffh. SRP functions by direct interaction with the ribosome, although the precise mechanism of this apparently complex interaction is not well known (Gu et al., 2005; Rinke-Appel et al., 2002).

Previously, our lab was able to identify multiple gene fragments from a genomic library (pLib1) that confer improved tolerance to 1-butanol (Borden and Papoutsakis, 2007). Here, we wanted to examine the hypothesis that a similar approach might be useful in identifying genomic loci conferring tolerance to butyrate. In addition to discovering potentially new mechanisms of acid tolerance, genetic loci that confer increased acid tolerance are anticipated to be beneficial for the construction of strains with robust acid-tolerance characteristics that could benefit strain productivity and metabolite titers. Two features of the pLib1 library used in the butanol-tolerance study (Borden and Papoutsakis, 2007) may have potentially limited the diversity of library inserts capable of contributing to a tolerant phenotype. First, the genomic library inserts preferentially taken up upon electroporation of *C. acetobutylicum* were smaller than the average gene size. Second, the pIMP1 cloning vector used to generate the pLib1 library did not contain a strong constitutive-like promoter and terminator flanking the insert site, which would have ensured transcription of all genomic library inserts. In this work, we sought to fix these two potential library deficiencies. An additional advantage of the new library constructed is that it can select for DNA fragments that might downregulate gene expression by antisense or non-coding RNA type of mechanisms. Using this library, we identified and examined genetic fragments that resulted from serial enrichment of the genomic library-bearing *C. acetobutylicum* cultures against increasing butyrate concentrations. All enriched fragments contained some portion of one of the ribosomal RNA-gene clusters, all in an antisense orientation relative to the constitutive plasmid promoter. Furthermore, a minimal insert region was identified as the promoter region upstream of the 16S ribosomal gene, also in the antisense orientation relative to the strong promoter.

2. Materials and methods

2.1. Bacterial strains and plasmids

TOP10 chemically competent *E. coli* (Invitrogen Corp., Carlsbad, CA) were used for routine cloning and library generation. The wildtype (WT) strain *C. acetobutylicum* ATCC 824 was used in this study. Prior to electroporation of *C. acetobutylicum*, plasmid DNA was methylated in *E. coli* ER2275(pAN1) to prevent restriction by a membrane-bound endonuclease (Mermelstein and Papoutsakis, 1993).

The pSOS95 plasmid, which contains the thiolase promoter and a rho-independent terminator (Supplementary Fig. 1), was modified for use as a destination vector in the Gateway Recombination system (Invitrogen Corp.). First, pSOS95 was digested sequentially with *Bam*HI and then *Kas*I (New England Biolabs, Ipswich, MA) per manufacturer's recommendations to remove the *ctfA-ctfB-adc* gene cassette. The ~5 kb fragment containing the Gram-positive and Gram-negative replication origins, MLS and ampicillin resistance markers, the thiolase promoter (P_{thi}) and rho-independent terminator was gel-excised, was blunt-ended using Klenow fragment (New England Biolabs) and dephosphorylated with calf intestinal phosphatase (CIP—New England Biolabs). Finally, the blunt-ended fragment was ligated to the Gateway® reading frame cassette A (RfA) using the Quick Ligase Kit (New England Biolabs) and transformed into DB3.1 *E. coli*. Individual clones of pSOS_ENT and pSOS_ENT_REV (differing in the direction of the reading frame cassette, and therefore the direction of a library insert upon recombination) were selected and the direction of the RfA cassette was confirmed by sequencing as well as restriction digest analysis.

2.2. Microbial cultures

E. coli cultures were grown aerobically at 37 °C in liquid Luria–Bertani (LB) media or solid LB agar plates, supplemented with ampicillin (100 μ g/mL), spectinomycin (100 μ g/mL), or chloramphenicol (35 μ g/mL), as necessary. *C. acetobutylicum* was revived from –85 °C storage by plating on 2xYTG (pH 5.8) solid media under anaerobic conditions at 37 °C. Individual colonies at least 5 days old were transferred to 10 mL tubes of liquid CGM (Wiesenborn et al., 1988) buffered with 30 mM acetate and adjusted to pH 7.0, and heat shocked at 70–80 °C for 10 min to induce germination. Recombinant strains were supplemented with erythromycin as necessary (40 μ g/mL for plates and 100 μ g/mL for liquid cultures).

2.3. Analytical methods

Cell growth was monitored by A_{600} using a Thermo Spectronic (Rochester, NY) Biomate3 spectrophotometer. Supernatant samples from *C. acetobutylicum* cultures were analyzed for glucose, acetate, butyrate, acetoin, ethanol, acetone, and butanol using a Waters (Milford, MA) HPLC (Tomas et al., 2003b). DNA concentration and purity were measured at 260 and 280 nm, respectively, using a NanoDrop Spectrophotometer (NanoDrop Technologies, Wilmington, DE).

2.4. DNA and RNA isolation

Plasmid DNA was isolated from *E. coli* using Hurricane plasmid mini and maxi-prep kits (Gerard Biotech, Oxford, OH). Isolation of genomic and plasmid DNA from *C. acetobutylicum* strains were carried out as described (Harris et al., 2002). Sampling and isolation of *C. acetobutylicum* RNA for Northern and microarray

analysis have been previously described (Tomas et al., 2003a). RNA used for Q-RT-PCR was processed differently. Instead of adding 500 μL of isopropanol to the aqueous phase from the TRIzol and chloroform separation, 500 μL of RNase-free 70% ethanol was added to the aqueous phase and mixed by vortexing. This solution was then purified using the miRNeasy Mini Kit (Qiagen, Valencia, CA) following the manufacturer's instructions, including the optional DNase treatment steps.

2.5. Genomic library generation

C. acetobutylicum genomic DNA was fragmented by sonication using a Branson Sonifier 150 (Branson Ultrasonics, Danbury, CT) at a power setting of 7 to generate an average fragment size of ~ 3 kb. Sonicated DNA was size-selected on a 0.7% agarose gel to include only fragments larger than 2 kb and smaller than 5 kb and purified using a Qiagen gel-purification column (Qiagen, Valencia, CA). One microgram of size-selected DNA was then blunt-ended by combining it with 1 μL of T4 DNA polymerase (New England Biolabs, Ipswich, MA), 5 μL of 100 $\mu\text{g}/\text{mL}$ BSA, 5 μL of New England Biolabs Buffer #2 (100 mM tris-HCl, 100 mM MgCl₂, 500 mM NaCl), and 1 μL of 10 mM dNTPs in 50 μL total volume and incubating for 2 h at 12 °C. T4 DNA polymerase was next deactivated by incubating at 75 °C for 20 min, and the blunt-ended DNA was dephosphorylated by addition of 0.5 μL of CIP (New England Biolabs) directly to the blunt-ending reaction mixture followed by incubation for 1 h at 37 °C. Size-selected, blunt-ended, dephosphorylated DNA was then purified using a Qiagen purification column, and finally single-A overhangs were added to the purified DNA by incubation with 1 μL of Taq polymerase (New England Biolabs), 6 μL of New England Biolabs Standard Taq polymerase buffer (100 mM tris-HCl, 500 mM KCl, 15 mM MgCl₂) and 1 μL of 10 mM dNTPs for 30 min at 72 °C. Finally, DNA with single-A overhangs was reacted with TOPO vector pCR8/GW/TOPO (Invitrogen) as recommended by the manufacturer, and ten TOP10 *E. coli* aliquots (Invitrogen) were transformed with 2 μL of the TOPO-isomerization reaction. After the 1 h outgrowth period recommended by the manufacturer, the 10 transformation reactions were combined into 250 mL of LB containing 100 $\mu\text{g}/\text{mL}$ of spectinomycin and incubated for 8 h at 37 °C. Representative plates streaked with 100 μL of the pooled transformation mixtures indicated approximately 78,000 total insert-bearing library colonies. The number of colonies (N) required to ensure a coverage probability (P) is based on the fraction (f) of the insert size relative to the entire genome (Clarke and Carbon, 1976) and is $N = \ln(1-P)/\ln(1-f)$. For a conservative average of 2 kb inserts, the low end of the 2–5 kb range of DNA fragments used, the number of clones for 99% coverage is

$$N = \frac{\ln(1-P)}{\ln(1-f)} = \frac{\ln(1-0.99)}{\ln\left(1 - \frac{2000}{3,950,000}\right)} = \frac{\ln(0.01)}{\ln(0.999)} \approx 9100 \text{ clones.}$$

Thus, the 78k clones obtained in the initial library give > 8-fold coverage.

TOPO-based library plasmids were isolated using a Hurricane mini-prep kit (Gerard Biotech), and the genomic inserts were transferred into the clostridia entry vectors pSOS_ENT and pSOS_ENT_REV using the Gateway Recombination system (Invitrogen) according to the manufacturer's recommendations. Each insert can go in either of the two destination vectors randomly, so one vector should have sufficed for both sense and antisense orientation/expression off the P_{thi} promoter of all possible library inserts. Nevertheless, just in case there is any bias as to the full representation of the whole genome in both the sense and antisense orientation and on similar size fragments, we decided to use both destination vectors. This is why the libraries from the

two destination vectors were combined to generate an even larger coverage of the whole genome in both orientations if all inserts could be expressed off the P_{thi} promoter. As before, ten TOP10 *E. coli* aliquots were transformed with 2 μL of the recombination reaction, combined into 250 mL of LB containing 250 $\mu\text{g}/\text{mL}$ ampicillin, and grown for 8 h before sampling for minipreps and frozen stocks. Representative plates indicate approximately 80,000 and 34,000 genomic inserts were recombined into the pSOS_ENT and pSOS_ENT_REV vectors, respectively, generating the pLib2thl genomic library. The number of clones obtained (the 80,000 and 34,000 genomic inserts for each vector) are > 8.5 and > 3.7 fold higher, respectively, than what is needed to provide genome coverage with 99% probability. Thus, the difference in the number of inserts between the two libraries is not very important since each contains sufficient genomic DNA to screen the entire genome. The combined library then provides a very large coverage of the genome in both orientations.

2.6. Serial enrichment of library cultures

Serial enrichment of *C. acetobutylicum* 824(pLib2thl) cultures against increasing butyrate concentrations was performed as described previously for library enrichment with 1-butanol (Borden and Papoutsakis, 2007). Electroporation of *C. acetobutylicum* with methylated pLib2thl yielded $\sim 3.2 \times 10^5$ total transformants, which were grown to an A_{600} of 0.6, and 50 μL frozen stock aliquots were prepared using CGM containing 15% glycerol. To initiate a serial-enrichment experiment, two frozen-stock aliquots were thawed and used to inoculate 10 mL tubes of CGM, grown to mid-exponential growth phase (A_{600} of 0.6), and used to inoculate 50 mL flasks containing CGM, 100 $\mu\text{g}/\text{mL}$ erythromycin, and 0.6% (v/v) butyric acid. The butyric-acid solution was adjusted to pH 6.7 with 10 M NaOH prior to challenge. Serial transfers were then conducted into media containing increasing butyrate concentrations, with transfers approximately every 24 h using a 5 mL inoculum into each fresh challenge flask. Plasmid library samples were collected from each challenge flask for identification of enriched genomic library inserts.

2.7. Growth inhibition and acid tolerance assays

Growth inhibition by butyrate was determined as described (Borden and Papoutsakis, 2007). Briefly, 400 mL of CGM, inoculated with 10 mL from an overnight culture ($A_{600} \sim 0.8$), with 100 $\mu\text{g}/\text{mL}$ clarithromycin were grown to an A_{600} of 1.0, in duplicate for each strain tested. Ten mL of culture was then aliquoted into individual test tubes containing butyrate (pH ~ 6.7) to generate a final concentration of 0–2% (v/v) butyrate, with four independent test tubes per challenge level (i.e., four technical replicates for each of two biological replicates per strain). Challenge tubes were sampled for growth (A_{600}) after 12 h of incubation, and replicate values averaged at each challenge level.

Tolerance of a recombinant strain to acetic, lactic, butyric, and isovaleric acid was also measured in a similar method. Briefly, 800 mL of CGM, inoculated with 10 mL from an overnight culture ($A_{600} \sim 0.8$), with 100 $\mu\text{g}/\text{mL}$ clarithromycin were grown to an A_{600} of 1.0, in duplicate for each strain tested. Ten mL of culture was aliquoted into individual test tubes containing pH-adjusted acid (pH of 5.3, 5.0, 6.7, and 6.7 for acetic, lactic, butyric, and isovaleric acid, respectively) to generate final concentrations, with two independent test tubes per challenge level. The pH values were chosen in order to maintain approximately the same molar ratio of protonated/unprotonated acid as in the original butyrate enrichment experiments. After 12 and 24 h of incubation in

challenged media, all tubes were sampled for growth (A_{600}) and metabolite analysis by HPLC, and replicate values averaged at each challenge level. For lactic acid, a racemic mixture of both L-(+)-lactic acid and D-(–)-lactic acid was used because it has previously been found that both are present for *C. acetobutylicum* (Meyer et al., 1986).

Tolerance to butyric acid was also tested after preconditioning in butyrate in a similar growth assay. Briefly, 220 mL of CGM, inoculated with 10 mL from an overnight culture ($A_{600} \sim 0.8$), with 100 $\mu\text{g}/\text{mL}$ clarithromycin and 0.6% (v/v) butyrate (pH 6.7) were grown to an A_{600} of 1.0, in duplicate for each strain. Ten mL of culture was then aliquoted into individual test tubes containing additional butyrate (pH 6.7) to raise the final level of butyrate to the desired amount, namely 0.7%, 0.9%, 1.0%, 1.1%, 1.2%, 1.4%, and 1.65%, with three independent test tubes for each biological replicate. A_{600} measurements were taken 12, 24, and 48 h after aliquoting into the test tubes, and at 48 h, supernatants were collected and used for HPLC analysis.

2.8. Quantitative (real-time) reverse transcription PCR (Q-RT-PCR)

Reverse transcription of 2 μg of purified total RNA was carried out using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA) following the manufacturer's instructions. After reverse transcription, 80 μL of RNase-free H_2O was added to the reactions to give a cDNA concentration of 20 ng/ μL . Each PCR reaction contained 1 μL of cDNA, 1 μM of each primer, and the Power SYBR[®] Green PCR Master Mix (Applied Biosystems) in a total volume of 25 μL . Reactions for each sample were performed on an iCycler Thermal Cycler (Bio-Rad, Hercules, CA) using the iQ5 Real-Time PCR Detection System (Bio-Rad) and the following parameters: 2 min incubation at 50 °C, 10 min incubation at 95 °C, 40 cycles of 15 s at 95 °C, and 1 min at 60 °C. A melt curve was constructed after the 40 cycles by starting at 55 °C and increasing the temperature 0.5 °C every 10 s until 95 °C. CAC3571 was used as the housekeeping gene, and the primer sequences used are in Supplementary Table 1.

2.9. Northern analysis

Single stranded γ -³²P end labeled oligo (ssOligo) probes were used for probing sense RDNA7, antisense RDNA7, and 16S RNA transcripts. Oligo sequences used are listed in Supplementary Table 1. Ten pmol of ssOligo probes were end labeled with γ -³²P (7000 Ci/mmol; MP Biomedicals, Solon, OH) using OptiKinase[™] Kit (USB, Cleveland, OH) following the manufacturer's instructions. Unincorporated radio nucleotides were removed using Micro Bio-Spin 30 columns (Bio-Rad, Hercules, CA). In each lane for the Northern blot, 20 μg of total RNA were loaded and electrophoretically resolved on a 1.5% denaturing MOPS-agarose gel. A Millennium[™] Marker (0.5–9 kb) (Ambion, Austin, TX) and another RNA marker (0.1–1 kb) (USB, Cleveland, OH) were used as size standards. RNA was transferred onto positively charged nylon membranes (Roche Applied Science, Indianapolis, IN) by overnight-capillary transfer with 10X SSC. Northern blots were baked at 80 °C for two hours to crosslink the RNA to the membrane. Pre-hybridization, hybridization, and washing were all carried out at 42 °C. Positively charged nylon membranes (Roche Applied Science, Indianapolis, IN) were pre-hybridized for 2 h with ULTRAhyb[®]-Oligo Hybridization Buffers (Ambion). Membranes were hybridized in the same buffer as used for pre-hybridization for 16–20 h. Membranes were washed twice with 2 \times SSC, 0.5% SDS for 15 min. Membranes were exposed on to a phosphor screen and imaged using a phosphor imager (GE Healthcare, Piscataway, NJ).

2.10. Transcriptional analysis using DNA microarrays

Design and validation of the Agilent-platform (Agilent Technologies, Santa Clara, CA) based *C. acetobutylicum* oligonucleotide microarrays employed in this study have been described (Paredes et al., 2007). Within one week of RNA purification, fluorescently labeled cDNA was generated by random hexamer-primed reverse transcription in the presence of amino-allyl (aa) dUTP (Sigma-Aldrich Corp., St. Louis, MO) using SuperScript II (Invitrogen Corp., Carlsbad, CA) reverse transcriptase as described (Alsaker et al., 2005). Labeled DNA was quantified by measuring A_{260} , and dye incorporation estimated by measuring A_{549} and A_{650} for Cy3 and Cy5, respectively. Agilent 22K arrays (GEO accession # GPL4412) were hybridized, washed, and scanned per Agilent's recommendations. Spot quantitation employed Agilent's eXtended Dynamic Range technique with gains of 100% and 10%. Normalization and slide averaging was carried out as described (Yang et al., 2003) except that ratios and mean intensities for spots of the same ORF were averaged after normalization. Microarray data have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE14433.

3. Results

3.1. All genomic library inserts enriched by serial transfer in the presence of butyrate contain at least part of the 5' region upstream of three 16S ribosomal RNA genes

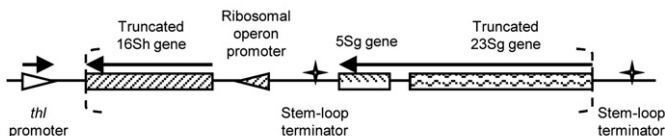
Two biological replicates of library-bearing *C. acetobutylicum* (pLib2thl) cultures were enriched in the presence of butyrate. Both replicate cultures were able to survive 10 transfers into media containing up to 1.3% (v/v) butyrate. When the cultures stopped growing, within 24 h post inoculation, plasmid DNA from the final surviving *C. acetobutylicum* cultures was used to transform TOP10 *E. coli* (Invitrogen), and plasmid DNA from individual *E. coli* colonies was analyzed by sequencing individual plasmid inserts. The sequence identities of these 18 *C. acetobutylicum* library insert colonies (9 from each biological replicate) are shown in Supplementary Table 2. All 18 inserts contained portions of the genome involved in synthesis of 16S ribosomal RNA. All bacteria contain multiple ribosomal RNA operons for the synthesis of 5S, 16S, and 23S ribosomal RNA molecules, and *C. acetobutylicum* contains 11 such operons, representing 0.8% of the genome (Nölling et al., 2001). Each of the 18 sequenced library inserts contained at least some portion of one of these operons, and these 18 inserts could be divided into 4 sub-classes. A cartoon of the contents of each insert sub-class is shown in Fig. 1. It can be seen that all of the sequenced inserts contain the 5' region upstream of three 16S ribosomal RNA genes (16Sg, 16Sh, and 16Sj) which contains the promoter for the corresponding entire ribosomal operon (*rrnG*, *rrnH*, *rrnJ*).

3.2. Growth-inhibition assays of library-bearing cultures

We desired to verify that the inserts isolated after the library enrichment process did in fact confer tolerance to butyrate, in excess of a plasmid control strain. Therefore, three sequenced inserts with different insert identities were selected from the 18 sequenced inserts (Supplementary Table 2) for sub-cloning back into WT *C. acetobutylicum*. Single recombinant *C. acetobutylicum* colonies were then grown and plasmid inserts verified before generating frozen stocks. Growth-inhibition assays of actively growing cultures (OD \sim 1.0) were performed with each of the

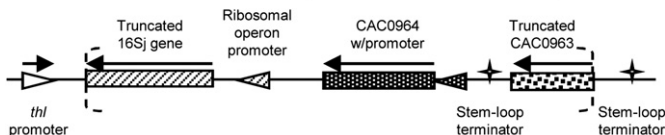
rDNA insert sub-type (I):

Sequenced inserts B, C, and I (Supplementary Table 2)



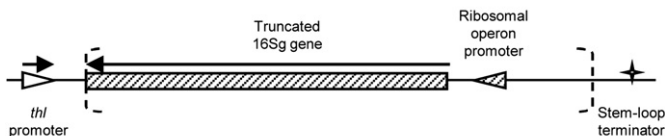
rDNA insert sub-type (II):

Sequenced inserts A and H (Supplementary Table 2)



rDNA insert sub-type (III):

Sequenced inserts F (Supplementary Table 2)



rDNA insert sub-type (IV):

Sequenced inserts G (Supplementary Table 2)

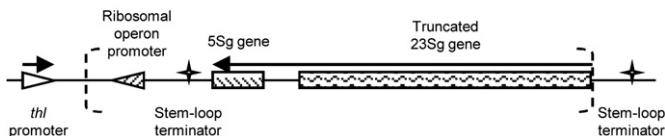


Fig. 1. Cartoon of butyrate-enriched genomic fragments from the library. Regions between the dashed brackets indicate the library inserts, which are flanked by the *thl* promoter and a rho-independent terminator. All inserts contain part, if not all, of the 5' upstream region of the 16S ribosomal RNA gene, which includes the promoter for the ribosomal operon (16S–23S–5S).

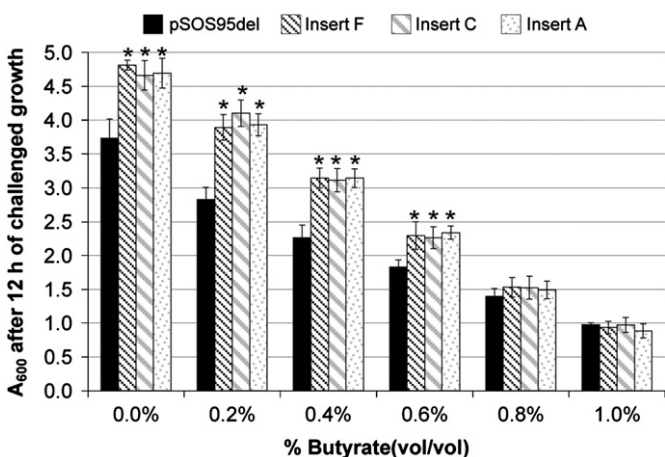


Fig. 2. Optical density measurements (A_{600}) of butyrate-challenged cultures of *C. acetobutylicum* harboring enriched genomic library inserts A, C, and F and plasmid control strain pSOS95del. All optical density measurements were taken after 12 h of growth following the butyrate stress. Measurements were taken from two biological replicates, with four technical replicates for each biological replicate (total of eight measurements). Values shown are the average of all eight measurements and the standard deviation between those eight measurements. Asterisks indicate statistically significant differences between the genomic library inserts and the plasmid control (p -values ≤ 0.05).

three strains carrying individual library inserts and the plasmid control strain *C. acetobutylicum* 824(pSOS95del) (Tomas et al., 2003b). The results of these growth-inhibition assays are shown in Fig. 2. At all butyrate concentrations, except for 0.8% and 1.0% butyrate, the strains harboring library inserts grew to higher cell densities than the plasmid-control strain. For example, at the different butyrate concentrations (0.0%, 0.2%, 0.4%, and 0.6%), insert F had a 29%, 38%, 39%, and 25%, respectively, higher cell density than the plasmid control strain. This growth advantage was seen across all the different library inserts. Insert C grew to 25%, 45%, 37%, and 24% higher cell densities, respectively, and insert A grew to 26%, 39%, 39%, and 28% higher cell densities, respectively. For all inserts, the greatest difference (37–45%) between the inserts and the plasmid control strain are seen at 0.2% and 0.4% butyrate. The 25% growth advantage in the extent of growth (higher cell densities) in the absence of butyrate stress suggests that the library-insert containing cells grow faster than the plasmid control strain.

It is also noteworthy that in the growth-inhibition assay no growth was observed for any of the tested strains past 1.0% (v/v) butyrate, whereas serially enriched cultures were able to tolerate up to 1.3% butyrate, although the growth assay conditions were different from the enrichment conditions. This difference in tolerance levels may be due to the conditioning of the culture to butyrate through slowly increasing the butyrate concentration during the serial enrichment. We address this possibility with additional assays described below.

3.3. Determination of a minimal DNA insert size

As mentioned above, all library inserts harbor at least one common region—the 5' upstream region of the 16S ribosomal coding region which contains the promoter region of the entire ribosomal operon. It was next desired to identify if a subsection of these similar library inserts was responsible for the observed tolerant phenotype. A sub-library of the enriched and sequenced genomic inserts was synthesized whereby PCR products were generated using the enriched library plasmids as template to generate smaller sub-fragments of the sequenced inserts (Supplementary Fig. 2). These PCR sub-fragments, 7 in total, were designed so as to cover all relevant permutations of key features contained in the butyrate-enriched library inserts (compare Supplementary Table 2 with Fig. 1 and Supplementary Fig. 2). Partial or intact open reading frames (CAC0964 and the 5S ribosomal gene) were cloned for constitutive expression from P_{thl} , and intergenic regions (i.e., the ribosomal operon promoter region found in all sequenced inserts and the intergenic region between the 5S and 23S ribosomal genes) were also cloned because these regions may impart tolerant phenotypes by a mechanism such as by titrating regulatory elements.

The set of 7 sub-fragments were then cloned into the TOPO-TA cloning vector, recombined into both pSOS_ENT and pSOS_ENT_REV (so that each fragment will express both the sense and antisense transcripts from the P_{thl}), methylated, and transformed into *C. acetobutylicum*. This microbial library of minimal fragments was then subjected to serial butyrate enrichment as before, to identify which fragment(s) provided the best survival under increasingly stringent selection. After the 14th transfer into media containing 1.25% (v/v) butyrate, plasmid DNA was isolated and inserts identified by sequencing. Fifteen of the 18 sequenced inserts contained the 353 bp intergenic region upstream of the 16Sh ribosomal gene and downstream of the 5Sg ribosomal gene (rDNA PCR Product 7 in Supplementary Fig. 2), while the other three sequenced inserts contained a shorter, 240 bp region upstream of the 16Sg ribosomal gene and downstream of the

CAC0290 gene (rDNA PCR Product 6 in Supplementary Fig. 2). These two sequences shared a large degree of similarity with 60.4% identity between them. Based on these data we conclude that butyrate tolerance associated with the enriched library fragments is largely due to the intergenic region upstream of the 16S ribosomal genes. Hereafter, this minimal tolerance-conferring promoter DNA fragment is referred to as RDNA7 (and the corresponding plasmid carrying this fragment as pRD7), so named due to its origination as a ribosomal operon promoter region and being the seventh insert as labeled and shown in Supplementary Fig. 2.

3.4. RDNA7 imparts broader general acid tolerance

To examine whether the pRD7 plasmid provided general acid tolerance, the ability to grow in media containing acetic, lactic, butyric, or isovaleric acid was also examined. Acetic acid and lactic acid were chosen because they are primary metabolites. While isovaleric acid is not a primary metabolite, it has been shown to be produced from leucine in some clostridia (Britz and Wilkinson, 1982; Britz and Wilkinson, 1983). As with the previous growth-inhibition assays, 824(pRD7) was able to grow to higher optical densities in the presence of any of the acids tested (Fig. 3A and B, Supplementary Table 3). Both acetic and lactic acid demonstrated some stimulatory effect leading to higher optical densities under acid stress than unstressed cultures, but

824(pRD7) still grew to higher A_{600} values than the plasmid control 824(pSOS95del) (Supplementary Table 3). In addition to just comparing A_{600} values under butyric and isovaleric acid, A_{600} measurements were converted into biomass concentrations using a mass index coefficient (Roos et al., 1985), and yield coefficients ($Y_{X/S}$, g biomass/mol glucose) for the two strains were calculated (Fig. 3C and D). 824(pRD7) exhibited greater yield coefficients, meaning expression of the insert stimulated greater biomass formation with less glucose consumption. This growth advantage can be particularly seen after 24 h of growth under isovaleric acid stress (Fig. 3D). These data indicate that the pRD7 plasmid provides robust tolerance to a variety of primary-metabolite and other carboxylic acids.

3.5. Preconditioning culture in sub-lethal level of butyrate imparts greater butyrate tolerance in 824(pRD7)

Though 824(pRD7) generally outperformed 824(pSOS95del) at all butyrate levels up to 0.75% (Fig. 3A), it was desired to demonstrate this ability at higher levels of butyrate. As mentioned earlier, the initial library survived up to 1.3% butyrate, but when the enriched inserts were tested again, they were barely surviving the 1.0% butyrate stress (Fig. 2). It was hypothesized that the initial library could survive higher levels of butyrate because it was conditioned in butyrate and the level of stress slowly increased. If the benefit of the insert involves induction of gene expression,

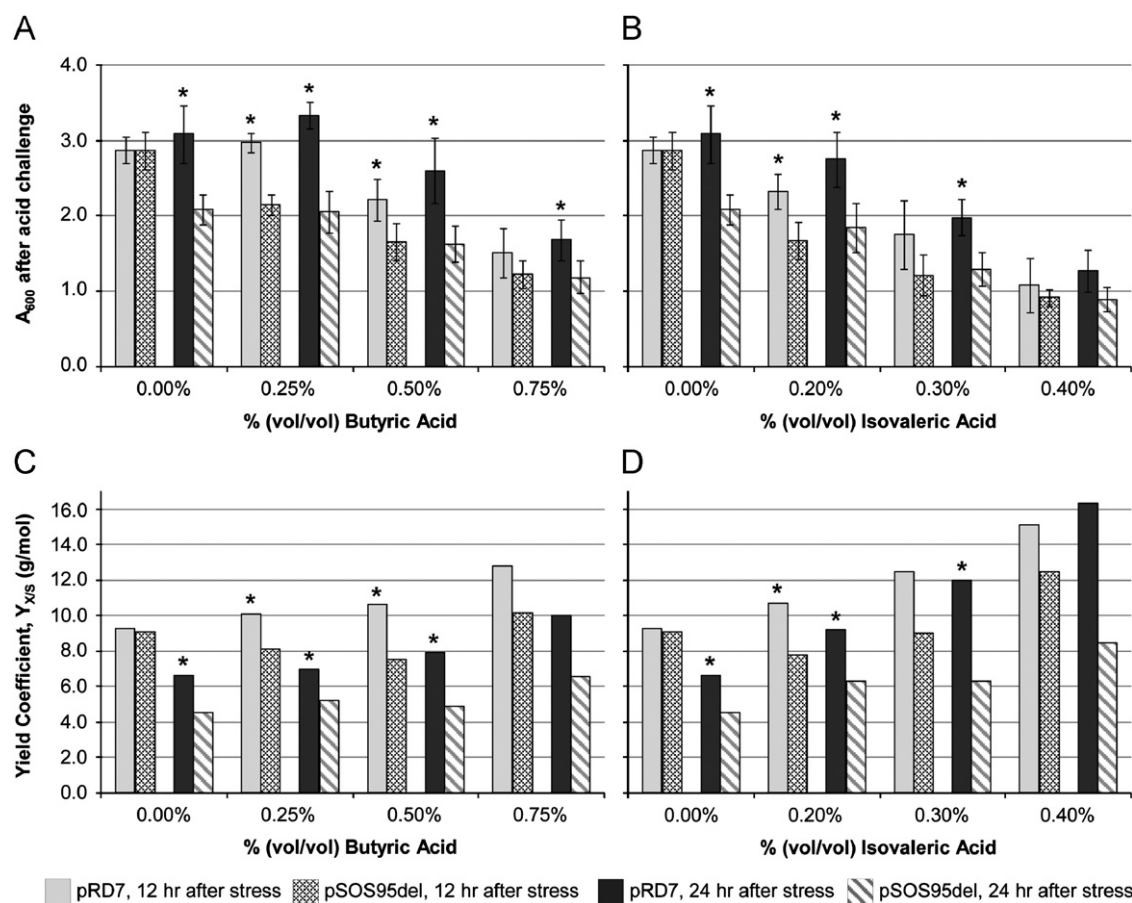


Fig. 3. Optical density measurements (A_{600}) and yield coefficients, $Y_{X/S}$ (g biomass/mol glucose), of *C. acetobutylicum* (pRD7) and (pSOS95del) challenged with 0.00%, 0.25%, 0.50%, and 0.75% (v/v) of butyric acid (pH 6.7) (A and C) and 0.00%, 0.20%, 0.30%, and 0.40% (v/v) of isovaleric acid (pH 6.7) (B and D). A_{600} measurements were taken 12 and 24 h following acid stress (A and B). The initial A_{600} measurements (~ 1.0) were subtracted off the measurements after 12 and 24 h so that only growth after acid challenge is shown. Measurements were taken from two biological replicates, with two technical replicates for each biological replicate (total of four measurements). Error bars show the standard deviation between the measurements. For the yield coefficients (C and D), the A_{600} values were converted into dry cell weight per volume (g/L) and divided by the amount of glucose consumed (mol/L). Raw data for the yield coefficients is in Supplementary Table 4. Asterisks indicate statistically significant differences between 824(pRD7) and 824(pSOS95del) (p -values ≤ 0.05).

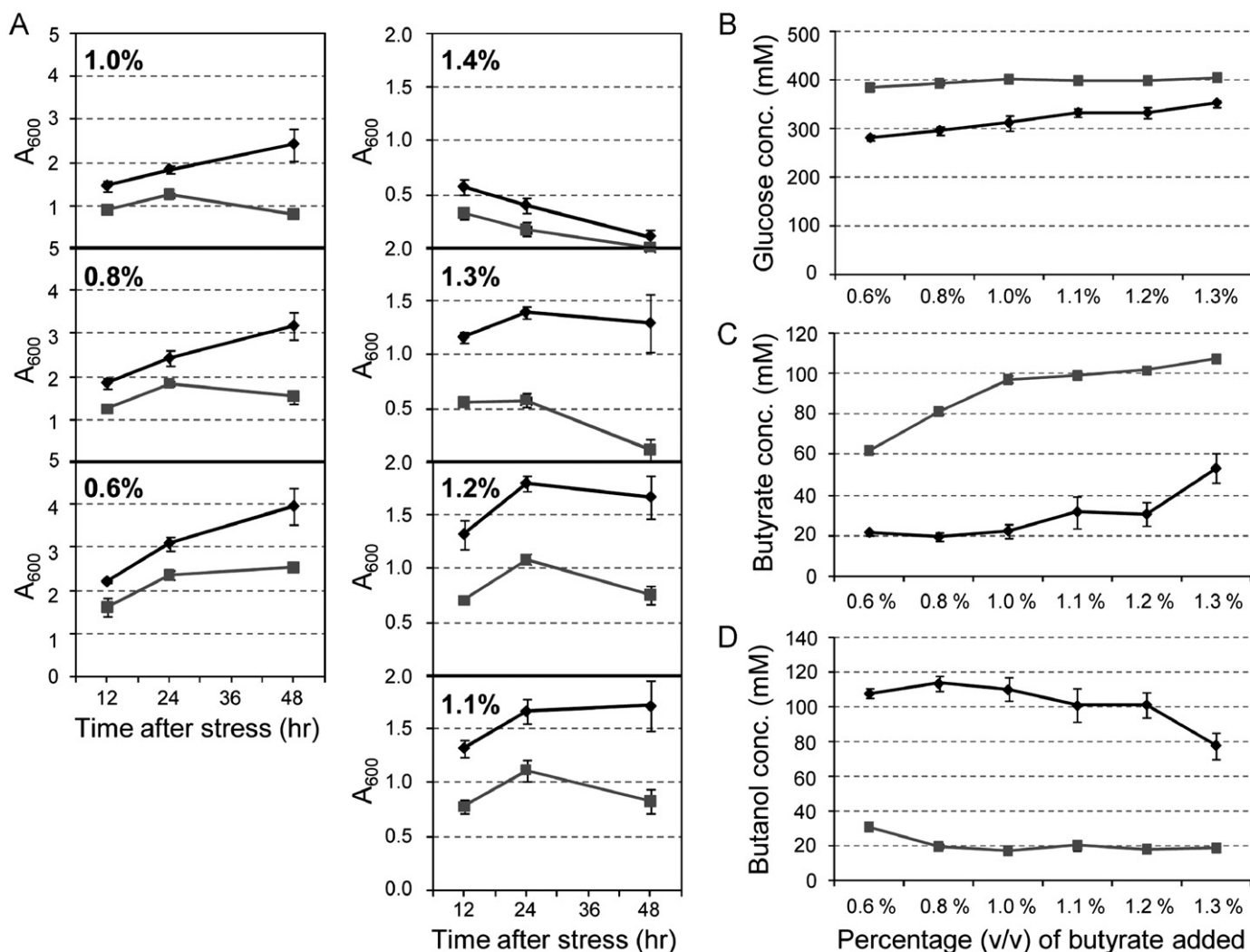


Fig. 4. Preconditioning the culture at sub-lethal levels of butyrate demonstrates the large tolerance advantage of *C. acetobutylicum* (pRD7) (black diamond) against the control strain (gray square). Two biological replicate cultures were grown to an A_{600} of 1.0 under 0.6% (v/v) butyrate stress, and then split into three technical replicates each and subjected to the different butyrate stress levels. (A) Optical density measurements from the two strains were taken 12, 24, and 48 h after stress. The initial A_{600} measurements (~ 1.0) were subtracted off the measurements so that only growth after acid challenge is shown. Residual glucose (B), butyrate (C) along with butanol production (D) were also measured after 48 h of growth under stress. A_{600} values of 1 correspond to one culture doubling and 3 to two doublings after the additional stress. The glucose levels in the cultures prior to the additional stress were $451 (\pm 3.2)$ mM, while the butanol and butyrate levels were ca. $2.8 (\pm 0.3)$ mM, and $53 (\pm 0.9)$ mM, respectively. Error bars show the standard deviation between the measurements.

stressing the cells right away with a high level of butyrate would not allow this induction to occur. To test this, biological replicates of 824(pRD7) and (pSOS95del) were grown up in media containing 0.6% (v/v) of butyrate (pH 6.7), and once the cultures reached an A_{600} of 1.0, they were stressed with additional levels of butyrate until the final concentrations were reached. A summary of the data are shown in Fig. 4. Significantly, 824(pRD7) maintained positive growth through 48 h up to 1.0% butyrate and at 1.1%, 1.2% and 1.3% was able to sustain its optical density. In contrast, 824(pSOS95del) could only achieve positive growth through 48 h with no additional stress only at 0.6% butyrate. Even though neither culture could sustain their optical density at 1.4%, 824(pRD7) still grew to a greater A_{600} than the plasmid control (Fig. 4A). In addition, at all stress levels 824(pRD7) consumed more glucose (Fig. 4B) and produced more butanol (Fig. 4D). Before the additional stress, glucose levels were at $451 (\pm 3.2)$ mM and the butanol concentration was $2.8 (\pm 0.3)$ mM, for both strains. Butyric acid levels were also lower for the 824(pRD7) cultures, indicating that a significant amount of butyrate was reassimilated to produce butanol (Fig. 4C). For both

strains, the level of butyrate before the additional stress was $53 (\pm 0.9)$ mM, or ca. 0.5% (v/v). Overall, 824 (pRD7) could survive, grow, and produce solvents at significantly higher concentrations of butyric acid than the plasmid control strain. Under unstressed culture conditions the two strains produced similar levels of solvents. Indeed, cultures of *C. acetobutylicum* 824(pRD7) ($n=2$) and 824 (pSOS95del) ($n=3$) were grown for 168–192 h and analyzed by HPLC for end-point metabolite levels. Strains 824(pRD7) and (pSOS95del) strains produced an average of 171 mM (± 0.7) and 154 mM (± 6.9) butanol, respectively, while consuming 358 mM (± 1.4) and 319 mM (± 15.2) glucose, respectively.

3.6. Analysis of non-coding RNA expression from RDNA7 in the tolerant 824 (pRD7) and the plasmid control strain 824(pSOS95del)

In order to confirm expression of transcripts from pRD7 and to determine if similar transcripts are also produced in the plasmid control strain, we used both Q-RT-PCR and Northern blot analysis. For

(Fig. 5B) and high (Ct values lower than ca. 15 are indicative of high expression in this organism; data not shown). This high level of expression in the plasmid control strain indicates that the Q-RT-PCR primers could be amplifying pre-processed 16S mRNA in addition to any native ncRNA_{RD7} transcript. In contrast, expression of ncRNA_{RD7} transcripts in 824(pRD7) decreased significantly (Fig. 5B) under both stress and no stress in a virtually identical manner, likely reflecting the activity of the *P_{thl}* promoter. For comparison, we also examined expression by Q-RT-PCR of the mature 16S rRNA that is coded by the adjacent *rrn* operon (Fig. 5C). 16S RNA expression, by the PCR probe in the middle of the mature 16S RNA, was somewhat lower in 824(pRD7) compared to 824(pSOS95del), and in both strains there was a small initial stress response, but overall, 16S RNA levels remained largely invariant in 824(pSOS95del), but increased slightly in 824(pRD7) (Fig. 5C). Notice that at time 0, the Ct values of 16S RNA are quite similar to the Ct values for the ncRNA_{RD7} transcripts in 824(pRD7). As a control, the Q-RT-PCR was repeated using RNA not reverse transcribed with each set of primers. The Ct values ranged from 24 to 26, more than 10 Ct values higher than the greatest Ct value using reverse transcribed RNA, indicating no significant genomic DNA contamination. Though Q-RT-PCR demonstrated high expression of putative transcripts corresponding to RDNA7, the size of the transcripts and, significantly, the orientation of the transcript(s) cannot be determined with Q-RT-PCR. Thus, we employed Northern analysis.

Initial Northern analysis using double-stranded DNA probes demonstrated the presence of multiple bands. To determine the nature of transcripts that could be transcribed from the RDNA7 fragment, we used single-stranded DNA probes (Fig. 6A). RNA from

0, 15, 120, and 360 min post butyrate stress from both 824(pRD7) and 824(pSOS95del) were prepared (Fig. 6B). This RNA was then probed with radio-labeled single-stranded oligonucleotide probes for both expression orientations of RDNA7: probe II to capture transcripts in the direction of 16S rRNA expression and probe I to capture putative ncRNA_{RD7} transcripts off the *P_{thl}* (Fig. 6A). When probe I was used, no transcripts were observed in any of the 824(pSOS95del) samples (Fig. 6C). In the 824(pRD7) strain, there was a 400-nt faint band at *t*=0 (more visible in a biological replicate shown as Supplementary Fig. 3), and up to four bands after butyrate stress. The ca. 400 nt band would correspond to the 353 nt transcript that could be produced from pRD7 using the *P_{thl}* (Fig. 6C). After stripping, probe II was hybridized to the same membrane to visualize any possible transcripts in the same orientation as 16S rRNA. In this case, both the plasmid control and the recombinant strain displayed a number of different bands, from 1700 to less than 100 nts, most of which appeared perfectly matched between the two strains, although they were significantly more intense in 824(pRD7) (Fig. 6D). We also note that the pRD7 strain displays some bands not detectable in the control, of which most notable are the bands below the 1400 and 400 nt bands (Fig. 6D). The 1700 nt molecule just above the mature 16S band (Fig. 6B) is probably a precursor 16.3S or 17S molecule similar to what has been previously reported for *E. coli* (Li et al., 1999). Among the larger molecules, there was a 1400-nt molecule just below the 16S mature molecule, as well as a slightly smaller molecule more visible in the 824(pRD7) lanes. There was also a prominent 900-nt molecule (and a slightly larger molecule better seen in the 824(pRD7) strain). The 700-nt molecule corresponds to the rRNA band of the same size observed in

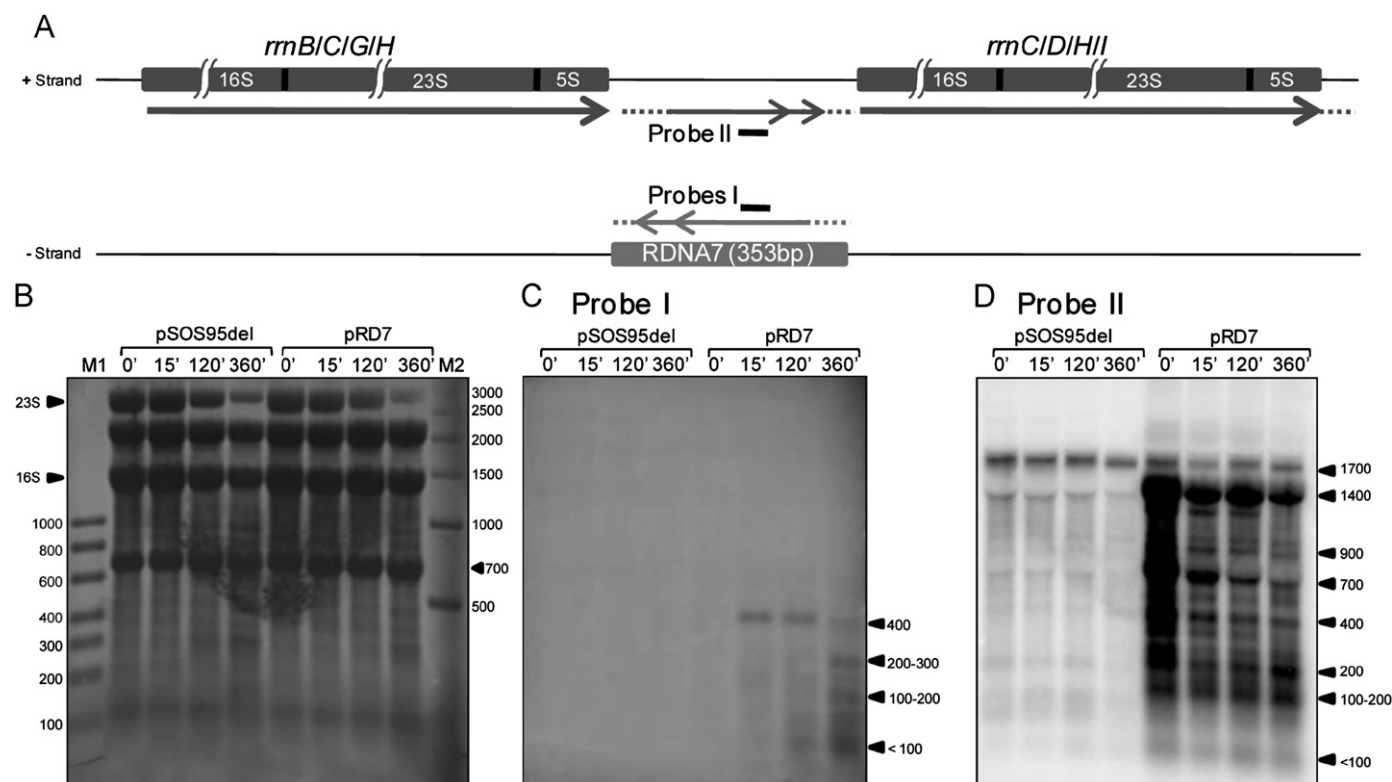


Fig. 6. Northern blot analysis of RDNA7 and its complementary transcript. Oligonucleotide radio-labeled probes were prepared for both the RDNA7 transcript (probe I) and its complementary transcript (probe II) (A). The RDNA7 probe was designed for the transcript produced from pRD7 and the potential native transcript on the negative strand of the intergenic region between RNA operons *rrnB/C*, *rrnC/D*, *rrnG/H*, and *rrnH/I*. A probe was also designed for the complementary RDNA7 sequence to probe the positive strand between the RNA operons. Since the length of any potential native transcript in the intergenic region is not known, dotted lines and double arrows indicate this potential transcript. Total RNA was extracted at 0, 15, 120, and 360 min after butyrate stress from pRD7 and pSOS95del (plasmid control). (B) After staining the membrane with methylene blue, four prominent bands are seen: 2800 nt (23S), 2400 nt, 1600 nt (16S), and 700 nt. M1 and M2 are RNA markers from USB (0.1–1 kb) and Ambion (0.5–9 kb), respectively. (C) After probing with probe I, complementary to positions 120–147 nt of RDNA7, four prominent bands are seen in pRD7. Membrane was exposed for 48 h. (D) After stripping probe I, complementary probe II was used to probe the same membrane. Membrane was exposed for 3 h.

methylene-blue stained RNA gels (Fig. 6B). We note that unlike *E. coli* and most other prokaryotes, instead of the usual two large bands of rRNA (16S and 23S), *C. acetobutylicum* RNA gels display 4 major distinct bands such as those shown in Fig. 6B, namely: the 23S (ca. 3000 nts), a 2000+ nt band, the 16S (ca. 1500 nts), and the aforementioned ca. 700-nt band. We assume that these are processed rRNA from the *rrn* operons, and the data of Fig. 6 as well other data (see below) support this assumption. Fig. 5A shows the putative promoters and transcriptional start site of the rRNA molecules (typically first as a single long 30S molecule (Britton et al., 2007)) transcribed from these *rrn* operons (namely *rrnC*, *D*, *H*, and *I*). These transcripts, as in other prokaryotes (Britton et al., 2007), start over 100 nts upstream of the DNA sequences coding for the mature 16S molecule, and contain several ribonuclease processing sites, whose action leads to the formation of the trimmed, mature 16S and other rRNA molecules. For 16S rRNA, based on promoter predictions in *C. acetobutylicum* (Paredes et al., 2004) and the phylogenetically closely related lactobacilli (de Vries et al., 2006), as well as on folding models of precursor 16S RNA molecules in *E. coli* and *Bacillus subtilis* (Britton et al., 2007), a folding model of the precursor 16S rRNA in *C. acetobutylicum* is shown in Fig. 5D. The long double-stranded stalk of this precursor molecule presumably contains several RNase processing sites, which are used to generate the mature 16S rRNA molecule through a sequence of cleavage steps that make up the rRNA processing pathway (Britton et al., 2007). The predicted unprocessed rRNA would hybridize against probe II of Fig. 6A and also against a large portion of ncRNA_{RD7} molecules produced from the RDNA7 fragment on pRD7 off the P_{thl} . From the data of Fig. 6, we conclude that in strain 824(pRD7) there are 4 short ncRNA_{RD7} molecules produced in a butyrate-stress dependent manner. If there are any such molecules in the control strain, their levels are not detectable in this assay. These data suggest that expression of ncRNA_{RD7} molecules from pRD7 alters rRNA expression and processing, leading to the accumulation of significantly higher levels of precursor rRNA molecules mostly of size lower than 1700 nt, but likely also of higher sizes as the faint bands, up to 3000 nts, of Fig. 6D (strain 824(pRD7)) suggest. These data also suggest that rRNA processing is more complex than anticipated and strongly affected by pRD7. We wanted to probe the rRNA processing complexity further by using probes against the mature 16S rRNA.

Since the RDNA7 fragment is just upstream of the 16S rRNA gene and dramatic differences in the expression of rRNA precursor molecules were observed between the two strains (Fig. 6D), Northern blots were also prepared to visualize the 16S rRNA. Two radio-labeled single-stranded oligonucleotide probes were designed to probe the 16S rRNA: one in the 5' region of the mature 16S rRNA (probe III) and a second in the middle of the mature 16S rRNA (probe IV). The hybridization patterns on the two membranes of Fig. 7 (C and E; we used different membranes to avoid potential artifacts due to probe stripping) are somewhat similar between the strains but different between probes. These unexpected patterns suggest unusual complexity of rRNA processing in both strains, which is the main reason for showing these gels. The methylene blue stains (Figs. 7B and D) were used to assess that the overall loading of RNA in the various lanes was approximately equal. Between the two strains, overall there are few discernable differences in the expression of the mature 16S rRNA. One difference though is observed when probing for the middle of the 16S (Fig. 7C): expression of the mature 16S in 824(pRD7) drops off more dramatically at 360 min than in the plasmid control strain, but this same drop is not seen when probed for the first half of the 16S. Another is the difference at 120 and 360 min on the gel of Fig. 7E in the rRNA species between 700 and just above 1100 nts: these lanes are less distinct or nonexistent in 824(pRD7) (see also details in the overexposed image of Supplementary Fig. 4). Both membranes (Fig. 7C and E)

share a band of about 700 nts, which corresponds to a ca. 700 nts rRNA band seen on the methylene blue stained membranes (Fig. 7B and D). As discussed above (Fig. 6B and D), this is a previously unidentified band seen on all RNA gels of this organism with the 16S rRNA and 23S rRNA bands when stained with methylene blue or ethidium bromide. The membrane of Fig. 7C displays bands (at ca. 900, 500, and 300 nts) not seen in the membrane of Fig. 7E, which displays bands (at ca. 1100, 200–300, 100–200, and the less than 100 nt bands; see also Supplementary Fig. 4 for these small size bands in an overexposed image of the membrane). Based on these images and the images of Fig. 6D, we conclude that rRNA processing is more complex than was anticipated based on published reports in clostridia and other organisms, and that the ncRNA_{RD7} molecules produced from plasmid pRD7 dramatically impact the expression levels and processing of the precursor rRNA molecules, and this is associated with the observable acid-tolerance phenotype.

3.7. Microarray analysis of the tolerant *C. acetobutylicum* 824(pRD7) strain

In order to obtain a better understanding of the differences in the transcriptional program imparted by pRD7 for a butyrate-tolerant phenotype, microarray analysis was carried out on replicate cultures of the 824(pRD7) and the plasmid control 824(pSOS95del) strains using the same type of experiments as those used for the Northern analysis of Figs. 6 and 7. Briefly, the strains were grown to the mid-exponential growth phase (OD~1.0), at which time a bolus of pH 6.7 butyrate was added to each flask to a final concentration of 0.6% (v/v) butyrate. Supernatant and RNA samples were taken from each culture to track the metabolic and transcriptional patterns of the cells. The first RNA sample was taken immediately prior to butyrate addition, when each culture was at an OD~1.0. RNA and supernatant samples were then taken 15, 40, 120, 240, and 360 min after butyrate addition in order to identify the short-term transcriptional differences between the two strains. The short-term response was chosen in order to capture the immediate rather than the secondary patterns of differential gene expression.

As before with the growth-inhibition assays on actively growing cultures (Figs. 2–4), the strain harboring pRD7 was better able to utilize glucose and grow following butyrate addition, as indicated by the higher OD and total glucose consumption post-butyrate challenge. 824(pRD7) began producing butanol very early in the culture, at an OD less than 0.2, corresponding to the very early exponential and acidogenic growth phase prior to external butyrate addition. Initiation of solventogenesis in batch culture has been correlated to critical levels of intracellular undissociated butyric acid (Huesemann and Papoutsakis, 1986, 1988; Zhao et al., 2005). Following butyrate addition, strain pRD7 continued to produce butanol more rapidly than the plasmid control culture. After isolating RNA from all cultures and samples, RNA was reverse transcribed to cDNA and fluorescently labeled. DNA microarrays were then hybridized for each timepoint, with one biological replicate of 824(pRD7) cDNA hybridized against cDNA from the same timepoint pre or post-challenge but from one of the 824(pSOS95del) cultures. To eliminate potential dye bias, all timepoints were hybridized in duplicate with a dye-swap configuration. In order to eliminate false positives due to biological variability, two independent sets of biological replicate experiments were analyzed. In total, 24 microarrays were hybridized—12 for each biological replicate challenge experiment.

Genes were identified that were differentially regulated with greater than 95% confidence in at least three time points of both

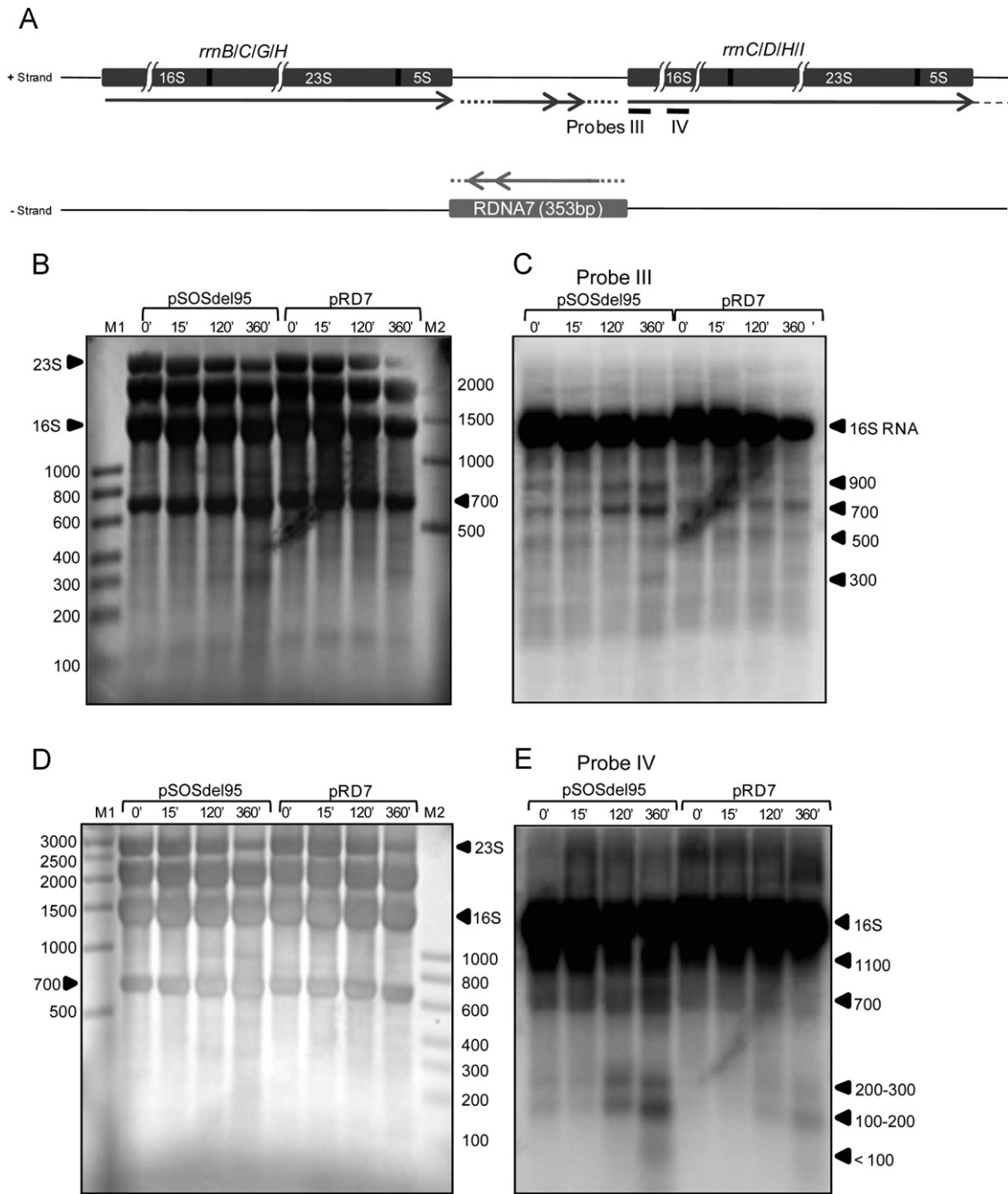


Fig. 7. Northern blot analysis of 16S rRNA. Two radio-labeled oligonucleotide probes were designed for the 16S ribosomal RNA gene: one for 16–43 nt of the 16S gene (probe III) and the other for 819–919 nt of the 16S gene (probe IV) (A). Total RNA was extracted at 0, 15, 120, and 360 min after butyrate stress from pRD7 and pSOS95del (plasmid control). (B and D) After staining the membrane with methylene blue, four prominent bands are seen: 2800 nt (23S), 2400 nt, 1600 nt (16S), and 700 nt. M1 and M2 are RNA markers from USB (0.1–1 kb) and Ambion (0.5–9 kb), respectively. (C) Visualization of the membrane in (B) probed with probe III. (E) Visualization of the membrane in (D) probed with probe IV. Both membranes were exposed for 10 min.

replicate butyrate challenge experiments. The expression ratios of ten genes with consistently increased expression in 824(pRD7) are shown in Fig. 8, and 110 genes with generally lower expression in 824(pRD7) are shown in Supplementary Fig. 5. An additional set of 55 genes that showed consistent temporal patterns of differential regulation, but were not differentially expressed with greater than 95% confidence in at least three time points of one of the two replicate experiments, are shown in Supplementary Fig. 6. Among the upregulated genes (Fig. 8 and Supplementary Fig. 6) with an identifiable role in acid tolerance is CAC1754, the *ffh* gene. As discussed, the Ffh protein is a component of the SRP system for transport of proteins to the cytosolic

membrane and interacts with an scRNA. This scRNA was identified in *C. acetobutylicum* based on similarity to those found in *E. coli* and *B. subtilis*, and its expression was probed and verified by Northern analysis (data not shown). Furthermore, its expression pattern in 824(pRD7) and 824(pSOS95del) after butyrate stress was probed with Q-RT-PCR along with the *ffh*, to confirm its upregulation in 824(pRD7). Expression of the scRNA was similar in both 824(pRD7) and 824(pSOS95del) with no significant up or down regulation. Expression of *ffh* by Q-RT-PCR displayed a similar 3-fold higher expression in 824(pRD7) over 824(pSOS95del) as seen in the microarray data. CAC2473, a predicted transcriptional regulator of

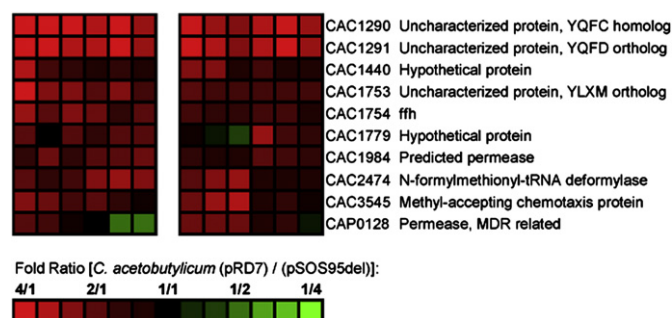


Fig. 8. Genes upregulated following butyrate challenge. Genes shown were upregulated with 95% confidence in three of six timepoints from two biological replicate butyrate challenge experiments. Color scheme shows red for genes with higher expression in the pRD7 strain and green for lower expression in the pRD7 strain.

the xenobiotic-responsive element (XRE) family, with roles in acid tolerance in *Lacobacillus acidophilus* (Azcarate-Peril et al., 2004) and solvent tolerance in *C. acetobutylicum* (Borden and Papoutsakis, 2007), was also among the upregulated genes. Also, several genes with homology to transporter genes of the *E. coli* acid decarboxylation/transport system of acid tolerance were consistently upregulated, including CAC1984, CAP0128, CAC0297, and CAC3285 (Fig. 8 and Supplementary Fig. 6). As discussed earlier, CAC3285 has high similarity to *gadC*, which codes for a cognate glutamate/GABA antiporter in *E. coli*, and CAC0297 has similarity to *adiA*, which codes for an arginine decarboxylase in *E. coli*. The upregulation of these genes may also play a role in the acid tolerance phenotype of 824(pRD7), but how they relate back to the overexpression of ncRNA_{RD7} is not known. Among downregulated genes (Supplementary Figs. 5 and 6), genes in the operon coding for the KdpATPase complex (CAC3679–3682), which is an essential osmosensor responsible for potassium ion transport and homeostasis in many Gram-positive and Gram-negative bacteria (Ballal et al., 2007; Wood, 2006), and the associated two-component signal transduction system (CAC3677 and CAC3678) showed the largest fold-decrease in expression.

Three findings are worthy of note. First, we note that the relatively few differentially expressed genes (Fig. 8 and Supplementary Figs. 5 and 6) are enriched in genes coding for membrane-associated proteins. Of the 10 genes with increased expression, half (CAC1291, CAC1779, CAC1984, CAC3545, and CAP0128) are predicted to be membrane-associated proteins. Second, the majority of differentially expressed genes are organized in a rather small number of operons (transcriptional units), many made up of hypothetical proteins. This would suggest that these operons belong to the regulon of a transcriptional regulator, whose activity is altered in the 824(pRD7) strain. Third, by far the largest fraction of differentially expressed genes were downregulated in the 824(pRD7) strain; this would suggest a mechanism of transcriptional suppression typical of non-coding regulatory RNAs (Bouvier et al., 2008).

4. Discussion

4.1. The pLib2thl library generates double diversity by expressing both sense and antisense transcripts off the thiolase promoter (P_{thl})

Both the pLib2thl library and subfragment libraries were constructed such that any genomic fragment used in the original library construction would be inserted in both possible orientations relative to the constitutive thiolase promoter (P_{thl}) (see Supplementary Fig. 1). The intent of such a design was to double the diversity of possible effects, by expressing both sense and antisense transcripts,

from a single library of cloned genomic fragments. Such a design makes it possible to explore the impact of not only overexpressing a gene or neighboring genes in the sense orientation, by significantly also whether an insert contributes to tolerance (or, generally, any desirable and testable phenotype) through generation of antisense transcripts (downregulation of gene expression or protein synthesis similar to mechanisms employed by small non-coding RNAs (Bouvier et al., 2008)), or by titration of regulatory elements that may indirectly impart the tolerant phenotype. For instance, if the inserts identified after butyrate enrichment existed in both the sense and antisense direction (i.e., independent of insert orientation), this would suggest an indirectly derived phenotype based on titration of regulator elements. Interestingly, both in the original library enrichment experiment, as well as the experiment using sub-fragments of the enriched library DNA, all resulting enriched fragments were oriented such that the P_{thl} generated an antisense transcript, and no fragments were enriched with the opposite, sense orientation. We conclude that the butyrate-tolerant phenotype is sensitive to library insert orientation, and such an effect might have been captured by a library that would have expressed only naturally occurring transcripts.

4.2. The unexpected diversity of ribosomal RNAs and their processing

We demonstrated that *C. acetobutylicum* displays a larger number of large-size rRNA bands beyond the expected 16S and 23S, namely those of ca. 2000 and 700 nts (Figs. 6 and 7). We also demonstrated that rRNA processing appears considerably more complex and dynamic under butyrate stressed culture conditions (Figs. 6D and 7C and E), although there appears to be very little change in the concentration of the predicted mature 16S rRNA (Fig. 5C). Strain 824(pRD7) appears to produce higher concentrations of precursor molecules of rRNAs smaller than the mature 16S rRNA (Fig. 6D). The physiological significance of these molecules and of the apparent complexity of rRNA processing is unknown. It has been suggested however that at least some molecules of processed rRNA, which are different from the mature rRNA molecules, may still be functional in ribosome assembly (Li et al., 1999). While the diversity of rRNA processing has been suggested in the literature (e.g., Song et al., 1999), the physiological significance of that diversity is not known.

4.3. Natural or synthetic mechanism of acid tolerance?

The impact of the library inserts from the rRNA-coding chromosomal loci and of the insert in the pRD7 plasmid in producing a robust acid tolerant phenotype is clear and reproducible in all assays employed in this study. It is significant to note the enhanced growth and biomass yields ($Y_{X/S}$) of 824(pRD7) (Fig. 4) (and of selected library clones; Fig. 2) even in the absence of external acid stress, which would indicate the enhanced ability of these strains to cope with the acid stress due to the accumulation of butyrate and acetate as a result of cell growth.

How might a ribosomal RNA promoter region, when transcribed in an antisense orientation, impart an acid-tolerant phenotype, such profound changes in the expression and processing of rRNAs, and, finally, the consistent alteration of 179 mRNAs? One possibility is that the inclusion of the 16S ribosomal promoter element within a multi-copy plasmid may impact the regulation of ribosomal RNA molecules by titration of transcription factors that bind rRNA promoter regions and regulate rRNA transcription. This is however unlikely due to the fact that the observable effect is orientation sensitive, i.e., only the antisense-expression orientation was enriched in four independent enrichment experiments, which argues against a DNA-titration effect

and in favor of an effect deriving from the ncRNA_{RD7} transcripts, which apparently hybridize to unprocessed nascent rRNAs (e.g., Fig. 5D) to alter their processing by ribonucleases. One can then hypothesize the formation of ribosome variants based on differently processed rRNAs, and such variants might affect the function and possibly the targeting of the ribosome to the membrane or the mother vs. endospore compartments (Paredes et al., 2005) of the cells.

Transcriptionally, the improved metabolic robustness of the 824(pRD7) strain was associated with the differential expression of osmotic responsive elements, amino acid decarboxylation components, and significantly the key protein (Ffh) of the SRP system. In *S. mutants*, the *ffh* gene is the second gene of the quaternary *sat* (secretion and acid tolerance) operon. The first gene of the *sat* operon is an ortholog to the *ylxM* *B. subtilis* genes and is apparently a transcriptional regulator affecting genes of the SRP system (Kremer et al., 2001). We note here that the *ylxM* ortholog is also strongly upregulated together with the *ffh* gene (Fig. 8). The SRP ribonucleotide–protein complex requires the so-called small cytoplasmic RNA (scRNA), which in *E. coli* is known as 4.5S RNA. Some Gram-positive bacteria have a longer SRP RNA than the 4.5S (ca. 100 nts) RNA in *E. coli*. For example, the size of scRNA in *B. subtilis* is 271 nucleotides (Nakamura et al., 1992). In *C. acetobutylicum*, we have predicted the scRNA to be 312–315 nt long. Northern analysis verified its expression and determined its size to be ca. 400 nt long (data not shown). Interestingly, the expression of *ffh* along with CAC1753, the *ylxM* ortholog, reached peak expression during acid production for a batch culture of *C. acetobutylicum* (Supplementary Fig. 7) (Jones et al., 2008). These data would suggest that the observable tolerant phenotype might be related to a SRP-dependent tolerance mechanism similar to that in *S. mutants*.

We would have expected higher levels of expression of ncRNA_{RD7} molecules in Fig. 6C due to the fact that the *P_{thi}* is strong promoter (Tummala et al., 1999). The reasons we likely observe relatively low levels of such transcripts is the relatively low sensitivity of the Northern-blot assay using single-stranded DNA probes, but more significantly because hybridization of these transcripts against the complimentary rRNA transcripts likely results in unstable heteroduplexes that are readily degraded.

Acknowledgements

This work was supported by the National Science Foundation Grants BES-0331402 and CBET-0756451, Department of Energy Grant DE-FG36-03GO13160, and an NIH/NIGMS biotechnology training Grant T32-GM08449 fellowship to Jacob R. Borden.

We thank the Northwestern University Biotechnology Core Laboratory for assistance with sequencing of plasmid library inserts. We also thank Bryan Tracy and Haw Siang Ang for assistance in generating the library, Sergios Nicolaou and Ryan Sillers for assistance in HPLC analysis of multiple acid stresses, and Dr. Carlos Paredes for assistance with promoter predictions.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at [Doi:10.1016/j.ymben.2009.12.004](https://doi.org/10.1016/j.ymben.2009.12.004).

References

Alsaker, K.V., et al., 2005. Design, optimization and validation of genomic DNA microarrays for examining the *Clostridium acetobutylicum* transcriptome. *Biotechnol. Bioprocess Eng.* 10, 432–443.

Azcarate-Peril, M.A., et al., 2004. Identification and inactivation of genetic loci involved with *Lactobacillus acidophilus* acid tolerance. *Appl. Environ. Microbiol.* 70, 5315–5322.

Ballal, A., et al., 2007. The Kdp-ATPase system and its regulation. *J. Biosci.* 32, 559–568.

Borden, J.R., Papoutsakis, E.T., 2007. Dynamics of genomic-library enrichment and identification of solvent tolerance genes for *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 73, 3061–3068.

Bouvier, M., et al., 2008. Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol. Cell.* 32, 827–837.

Britton, R.A., et al., 2007. Maturation of the 5' end of *Bacillus subtilis* 16S rRNA by the essential ribonuclease YkqC/RNase J1. *Mol. Microbiol.* 63, 127–138.

Britz, M.L., Wilkinson, R.G., 1982. Leucine dissimilation to isovaleric and isocaproic acids by cell suspensions of amino acid fermenting anaerobes: the Stickland reaction revisited. *Can. J. Microbiol.* 28, 291–300.

Britz, M.L., Wilkinson, R.G., 1983. Partial purification and characterization of two enzymes involved in isovaleric acid synthesis in *Clostridium bif fermentans*. *J. Gen. Microbiol.* 129, 3227–3237.

Clarke, L., Carbon, J., 1976. Colony bank containing synthetic Col EI hybrid plasmids representative of entire *Escherichia coli* genome. *Cell* 9, 91–99.

de Vries, M.C., et al., 2006. Comparative and functional analysis of the rRNA-Operons and their tRNA gene complement in different lactic acid bacteria. *Syst. Appl. Microbiol.* 29, 358–367.

Foster, J.W., 2004. *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat. Rev. Microbiol.* 2, 898–907.

Gu, S.Q., et al., 2005. Conformation of 4.5S RNA in the signal recognition particle and on the 30S ribosomal subunit. *RNA* 11, 1374–1384.

Gutierrez, J.A., et al., 1999. *Streptococcus mutans ffh*, a gene encoding a homologue of the 54 kDa subunit of the signal recognition particle, is involved in resistance to acid stress. *Microbiology* 145 (Pt 2), 357–366.

Harris, L.M., et al., 2002. Northern, morphological, and fermentation analysis of *spoOA* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol.* 184, 3586–3597.

Herrero, A.A., et al., 1985. Growth-inhibition of *Clostridium thermocellum* by carboxylic-acids-A mechanism based on uncoupling by weak acids. *Appl. Microbiol. Biotechnol.* 22, 53–62.

Huesemann, M., Papoutsakis, E.T., 1986. Effect of acetoacetate, butyrate, and uncoupling ionophores on growth and product formation of *Clostridium acetobutylicum*. *Biotechnol. Bioeng.* 8, 37–42.

Huesemann, M., Papoutsakis, E.T., 1988. Solventogenesis in *Clostridium acetobutylicum* fermentations related to carboxylic acid and proton concentrations. *Biotechnol. Bioeng.* 32, 843–852.

Jones, S.W., et al., 2008. The transcriptional program underlying the physiology of clostridial sporulation. *Genome Biol.* 9, R114.

Kell, D.B., et al., 1981. On the permeability to weak acids and bases of the cytoplasmic membrane of *Clostridium pasteurianum*. *Biochem. Biophys. Res. Commun.* 99, 81–88.

Kremer, B.H., et al., 2001. Characterization of the *sat* operon in *Streptococcus mutans*: Evidence for a role of Ffh in acid tolerance. *J. Bacteriol.* 183, 2543–2552.

Li, Z., et al., 1999. RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.* 18, 2878–2885.

Mermelstein, L.D., Papoutsakis, E.T., 1993. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. *Appl. Environ. Microbiol.* 59, 107710–107781.

Meyer, C.L., et al., 1986. Carbon-monoxide gasing leads to alcohol production and butyrate uptake without acetone formation in continuous cultures of *Clostridium acetobutylicum*. *Appl. Microbiol. Biotechnol.* 24, 159–167.

Nakamura, K., et al., 1992. Small cytoplasmic RNA of *Bacillus subtilis*—functional-relationship with human signal recognition particle 7S RNA and *Escherichia coli* 4.5S RNA. *J. Bacteriol.* 174, 2185–2192.

Nölling, J., et al., 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol.* 183, 4823–4838.

Papoutsakis, E.T., 2008. Engineering solventogenic clostridia. *Curr. Opin. Biotechnol.* 19, 420–429.

Papoutsakis, E.T., et al., 1987. Transport of substrates and metabolites and their effect on cell metabolism (in butyric-acid and methylotrophic fermentations). *Ann. N Y Acad. Sci.* 506, 24–50.

Paredes, C.J., et al., 2005. A comparative genomic view of clostridial sporulation and physiology. *Nat. Rev. Microbiol.* 3, 969–978.

Paredes, C.J., et al., 2004. Transcriptional organization of the *Clostridium acetobutylicum* genome. *Nucleic Acids Res.* 32, 1973–1981.

Paredes, C.J., et al., 2007. A general framework for designing and validating oligomer-based DNA microarrays and its application to *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 73, 4631–4638.

Richard, H., Foster, J.W., 2004. *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *J. Bacteriol.* 186, 6032–6041.

Rinke-Appel, J., et al., 2002. Crosslinking of 4.5S RNA to the *Escherichia coli* ribosome in the presence or absence of the protein Ffh. *RNA* 8, 612–625.

Roos, J.W., et al., 1985. The effect of pH on nitrogen supply, cell lysis, and solvent production in fermentations of *Clostridium acetobutylicum*. *Biotechnol. Bioeng.* 27, 681–694.

Russell, J.B., 1992. Another explanation for the toxicity of fermentation acids at low pH—Anion accumulation versus uncoupling. *J. Appl. Bacteriol.* 73, 363–370.

- Russell, J.B., Diez-Gonzalez, F., 1998. The effects of fermentation acids on bacterial growth. *Adv. Microb. Physiol.* 39, 205–234.
- Song, X.M., et al., 1999. Fragmentation heterogeneity of 23S ribosomal RNA in *Haemophilus* species. *Gene* 230, 287–293.
- Tomas, C.A., et al., 2003. DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5. *J. Bacteriol.* 185, 4539–4547.
- Tomas, C.A., et al., 2003. Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and large changes in the cell's transcriptional program. *Appl. Environ. Microbiol.* 69, 4951–4965.
- Tummala, S.B., et al., 1999. Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. *Appl. Environ. Microbiol.* 65, 3793–3799.
- Walter, A., Gutknecht, J., 1984. Monocarboxylic acid permeation through lipid bilayer membranes. *J. Membr. Biol.* 77, 255–264.
- Warnecke, T., Gill, R.T., 2005. Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. *Microb. Cell. Fact.* 4, 25.
- Wiesenborn, D.P., et al., 1988. Thiolase from *Clostridium acetobutylicum* ATCC 824 and its role in the synthesis of acids and solvents. *Appl. Environ. Microbiol.* 54, 2717–2722.
- Wood, J.M., 2006. Osmosensing by bacteria. *Sci. STKE.* 2006, pe43.
- Yang, H., et al., 2003. A segmental nearest neighbor normalization and gene identification method gives superior results for DNA-array analysis. *Proc. Natl. Acad. Sci. USA* 100, 1122–1127.
- Zhao, Y., et al., 2005. Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation. *Appl. Environ. Microbiol.* 71, 530–537.