J Ind Microbiol Biotechnol (2006) 33: 298–308
DOI 10.1007/s10295-005-0050-7

ORIGINAL PAPER

Leighann Sullivan · George N. Bennett

Proteome analysis and comparison of *Clostridium acetobutylicum* ATCC 824 and Spo0A strain variants

Received: 18 July 2005 / Accepted: 1 October 2005 / Published online: 25 November 2005

Abstract The proteomic profiles of several *Clostridium acetobutylicum* strains were compared by two-dimensional gel electrophoresis and mass spectroscopy. The proteomic profile of *C. acetobutylicum* wild type strain ATCC 824 with and without a commonly used control plasmid and with a *spo0A* overexpression plasmid pMSPOA was compared. A total of 2,081 protein spots were analyzed; 23 proteins were chosen to be identified of which 18 were unique and 5 were proteins located in more than one location. The proteins identified were classified into heat shock stress response, acid and solvent formation, and transcription and translation proteins. Spo0A was identified and its protein expression was confirmed to be absent in the *spo0A* knockout SKO1 strain as expected, as was the protein Adc, which is known to be regulated by Spo0A. The expression of six proteins was not detected in strain SKO1 indicating these proteins require Spo0A. Spo0A overexpression affected the abundance of proteins involved in glycolysis, translation, heat shock stress response, and energy production. Two features were identified: five of the 23 proteins identified were located in more than one position and clusters of protein spots resembled fingers of a straightened hand. Normally a protein localizes to only one spot on the gel; localization of a protein to more than one spot is indicative of post-translational modifications, suggesting that such modification of proteins may be a more prevalent mechanism in *C. acetobutylicum* than previously thought. The clusters of protein spots resembling fingers of a straightened hand were in the acidic high molecular weight areas. Two such protein spots were identified as variants of the same protein, GroEL.

L. Sullivan · G. N. Bennett (✉)
Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77005, USA
E-mail: gbennett@bioc.rice.edu
Tel.: +1-713-3484920
Fax: +1-713-3485154

Keywords Clostridium · Spo0A · SKO1 · Proteome · Mass spectroscopy

Introduction

Clostridium acetobutylicum ATCC 824 is a Gram positive, sporulating, obligate anaerobic bacterium whose genome has been sequenced [29]. The solvents produced by *C. acetobutylicum* (acetone, butanol, and ethanol) are commercially important and have been synthesized from petroleum since the mid-1950s [21]. Increased concern over depletion of non-renewable resources has led to renewed interest in producing solvents via microbial fermentative processes. The major obstacle to national or global microbial solvent production has been its cost, the cost of the raw material and the cost of solvent production and recovery. The use of alternative fermentation substrates, such as apple pomace and paper industry waste, and the use of solvent tolerant and hyperproducing strains have decreased the cost of microbial fermentation. Solvent production is a late growth stage physiological change, as are sporulation, granulose accumulation, and heat shock stress response induction. Solvent production is correlated with the late stage abundance of Spo0A [4, 21, 35].

Spo0A is a transcriptional regulator important in controlling the late stage physiological changes identified first in *Bacillus subtilis* [20] and subsequently in *C. acetobutylicum* [9]. In *C. acetobutylicum*, Spo0A activates the transcription of acetoacetate decarboxylase (*adc*) [35], alcohol/aldehyde dehydrogenase (*aad* also known as *adhE)*, and acetoacetyl coenzyme A: acetate/butyrate: coenzyme A transferase (*ctfA* and *ctfB*) [14]. Spo0A downregulates the transition state regulator *abrB* [35]. Spo0A is necessary for the normal solventogenic process by controlling these solvent formation genes.

In order to determine which genes are affected at the transition to solventogenesis, the transcriptional profile has been examined [11], especially as the transition relates to stress [1, 43] and has been found to be effected by

Spo0A [1, 19]. However, transcription is not the only mode of controlling gene expression. Gene expression controlled at the translational level may be caused by altered RNA stability or translational efficiency. Once the RNA has been translated into protein, the activity of the protein may be altered by many means: post-translational modification, the presence of metabolic small molecule effectors, and/or differences in protein stability caused by targeted degradation. The objective of this study is to examine Spo0A-regulated expression at the protein level in *C. acetobutylicum*. A proteomic examination of Spo0A regulation will also reveal regulatory events occurring after transcription of a gene. Protein profiles were compared under three situations: between the acidogenic and solventogenic phases; between a control strain and a *spo0A* null strain; and differences between a control strain and a *spo0A* overexpression strain.

Materials and methods

Bacterial strains and culture conditions

The bacterial strains used are listed in Table 1. *C. acetobutylicum* ATCC 824 (pIMP1) and SKO1 strains were grown in pH 5 controlled fermentations, and the growth was monitored with $OD_{600\ nm}$ [19]. Samples of 15 ml were taken at early-mid exponential, mid-late exponential, early stationary, and late stationary phases. The phases were defined according to the hour of growth and the culture $OD_{600\ nm}$ when the sample was harvested. The two parameters are listed below for each strain. After the samples were harvested they were diluted to 10 ml with an $OD_{600\ nm}$ of 1.0, centrifuged, and stored at –70°C as cell pellets. The time and $OD_{600\ nm}$ of samples when harvested for 824(pIMP1), SKO1, and 824(pMSPOA) are listed in Table 2.

Two-dimensional electrophoresis

Cell pellets were thawed, heat denatured, and disrupted by sonication. Crude protein extracts were separated from cellular debris by centrifugation (7,000*g*, 10 min) and the protein content of the supernatant fraction was determined [7]. Duplicate samples of 200 μg of protein were resuspended in immobilized pH gradient (IPG) reswelling solution containing urea/thiourea rehydration buffer (pH 3–7) overnight and isoelectric focusing separation was carried out. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in the second dimension was performed and the gels were fixed and stained with Coomassie Blue (Genomic Solutions Inc., Ann Arbor, MI, USA).

Image analysis and protein identification

Duplicate two-dimensional gels were obtained for each cell time point sample and imaged on a flat bed scanner (HT analyzer software, 200 dpi, Genomic Solutions Inc.). The gel with the highest resolution was used for analysis. Protein features were detected (Spot Detection Wizard software, Genomic Solutions Inc.) and matched among the gels. The background was subtracted and each spot was normalized to the total spot volume of each individual gel (Genomic Solutions Inc.). The intensity of each spot was normalized to its corresponding spot in the reference gel that is the gel of the first time point in each strain.

Selected gel plugs of individual protein spots were analyzed by mass spectroscopy for protein identification. The excised protein spots were trypsin digested in the gels. The digested supernatant fluid was mixed with MALDI matrix (α-cyano-4 hydroxy cinnamic acid), spotted onto a MALDI target, and then inserted into the MALDI-TOF mass spectrometer (Voyager DE-STR, Applied Biosystems). Two peptide mass fingerprints were analyzed (MS-Fit software of Protein Prospector and Proteometrics software of ProFound). The digested peptide masses were matched with between 9 and 30 peptide sequences in the database to identify the protein with genome encoded ORFs (Genomic Solutions Inc.).

0A box determination

The intergenic regions upstream of the genes encoding the proteins or upstream of the first gene of a known or putative operon, as indicated in Table 3, were examined for putative 0A boxes [42].

Table 1 Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
C. acetobutylicum		
ATCC 824	Wild type strain	ATCC
SKO1	ATCC 824 *spo0A*::MLS^R	[19]
Plasmids		
pIMP1	Ap^R MLS^R; ColE1 origin	[25]
pMSPOA	Ap^R MLS^R; carries *spo0A*	[19]

MLS^R macrolide-lincosamide-streptograminB resistance, *ATCC* American type culture collection, Manassas, VA, *Ap^R* ampicillin resistance

Table 2 Time and $OD_{600\ nm}$ of samples when harvested

Growth phase	824(pIMP1)		SKO1		824(pMSPOA)	
	Hour	$OD_{600\ nm}$	Hour	$OD_{600\ nm}$	Hour	$OD_{600\ nm}$
Early-mid exponential	9.5	1.8	10.0	3.0	9.5	3.4
Mid-late exponential	13.5	4.9				
Early stationary	18.8	6.2	31.0	6.0	29.5	10.0
Late stationary	100.3	5.4	59.5	4.7	61.5	8.0

Results

The reasons for this study are twofold: (1) in order to investigate how the wild type *C. acetobutylicum* 824 proteome changed over time, especially as the culture transitions from exponential to stationary phase and (2) in order to investigate how Spo0A modulates that global protein pattern. We employed two-dimensional electrophoresis and mass spectroscopy analysis to identify 23 protein of interest. These proteins were chosen based on their differential protein expression. The differential expression of the 23 proteins was examined and identified in one of the three following comparisons: between the acidogenic and solventogenic phases or between a control strain and a *spo0A* null strain or between a control strain and a *spo0A* over-expression strain.

Table 3 Potential role for Spo0A

Gene, operon (first gene of cluster)	0A/R0A[a]	Protein present in SKO1[b,c]	Gene included on microarray[d,e]			Protein abund[g]
			Exp[f]	Tran[f]	Stat[f]	
Upstream of gene located in a known multi-cistronic operon[h]						
atpD, ATP synthase operon	3 0A, 1 R0A[i]	+				−
bcd, BCS operon	0A	+	+, −	−, ++	−−[e], −−[d]	−/NC[j]
groES, GroE operon	R0A	+	++, +−	−−, ++	+[e], −−[d]	+
groEL, GroE operon	R0A	+	++	−−	+[e]	+
Upstream of gene located in a putative multi-cistronic operon[k]						
rpoA (CAC3107)	0A	+	++	+	−−[e]	NC
rplL (CAC3151)	R0A	+				−
chw14 (CAC1532)	0A	−				+
efp (CAC2095)	No	+				NC
Upstream of gene located in a monocistronic operon[l]						
fba	0A	+	++, −+	+, −	−−[e],−−[d]	NC
spo0A	0A	−	++	++	++[e]	NC
hsp18	2 0A	−1 variant	−−, +−	−, ++	++[e], −−[d]	+
pspA	2 0A	+				+
pgi	2 0A, 1 R0A	+	−+	−	−−[d]	−
adc	2 0A, 1 R0A	−	++, −	++, ++	++[e], −[d]	+
chw16/17	0A	−				+
gapC	No	+	−+	−	−−[d]	NC
ppi	R0A	−	−−	−	++[e]	+
tpi	No	+	++, −+	+, −	−−[e], −−[d]	−/NC

[a]0A, 0A box; R0A, reverse 0A box; determined by consensus search of intergenic regions [42]; *adc* 0A boxes have been experimentally demonstrated [35]
[b]+/− in "Protein present in SKO1" columns represent yes/no, respectively
[c]When protein is not present in strain SKO1 this indicates that the gene requires Spo0A for expression
[d]Gene is included on microarray and transcript expression profile examined in 824 (pMSPOA) strain compared to 824(pIMP1) [1]. + and − in "Gene included in microarray" indicate relative transcript levels detected in microarray experiments: ++, very high expression; +, moderately high expression; − moderately low expression; −− very low expression. When the transcript levels of a gene was examined in both microarray experiments, the relative expression levels for both are indicated and separated by comma
[e]Gene is included on microarray and transcript expression profile examined in wild type 824 strain compared to SKO1 strain [42]
[f]Exp, exponential phase, Tran, Transition, Stat, stationary phase
[g]+/− is up-/downregulated, respectively; NC is no change, thus approximately constitutive levels of the overall protein accumulation trend in *C. acetobutylicum* 824 (pIMP1)
[h]References for known operons: [6, 12, 28]
[i]Number in front of 0A or R0A indicates the number of non-overlapping boxes present, if no number is indicated, a single box is present
[j]Each abundance refers to one variant of the protein identified
[k]Based on gene clustering with other genes of similar function in the same orientation
[l]May be demonstrated or putative operon structure

Growth of strain under selective media

Clostridium acetobutylicum 824 (pIMP1) is a control strain for SKO1, *C. acetobutylicum* 824 (pMSPOA), and *C. acetobutylicum* 824 wild type strains. The strains *C. acetobutylicum* 824 (pIMP1), *C. acetobutylicum* 824 (pMSPOA), and SKO1 were grown under either erythromycin or clarithromycin selection. The use of otherwise identical culture conditions allows for comparison between the two strains. *C. acetobutylicum* 824 without a plasmid was also grown without selection and was compared to *C. acetobutylicum* 824 (pIMP1) (data not shown). The resulting proteomic profile of *C. acetobutylicum* 824 was similar *C. acetobutylicum* 824 (pIMP1), thus neither growth under selective media nor the maintenance of a plasmid distorted the proteome pattern in any obvious fashion.

Overall features of the protein pattern

Six overall features of the protein pattern were noted and are discussed. The first feature is shown in Fig. 1, which is a representative two-dimensional gel from late solventogenic phase. A complete collection of two-dimensional gel images is available at http://www.bioc.rice.edu/bennetlab. The second feature is the classification of the 18 unique proteins identified. The features can be sorted into the following classifications: stress, acid and solvent formation, transcription and translation, and other functions (Fig. 2, Table 4). The third feature was that the calculated molecular weights and isoelectric points generally agree with the apparent values on the two-dimensional gel. The fourth feature examined the abundance of proteins; several identified proteins were upregulated or downregulated while some remained at a relatively constant level of abundance regardless of growth phase. The fifth feature examined Spo0A as a requirement for maximal expression of some identified proteins (Fig. 3) and when overexpressed altered many processes. Lastly, two additional features of the *C. acetobutylicum* proteome were identified; five proteins were found in more than one location and protein clusters resembling fingers of a hand were identified.

Agreement between molecular weights and isoelectric points

The molecular weights and isoelectric points for most of the 18 unique proteins matched well between the calculated and apparent values. In each instance, at least one protein spot matched the calculated values well, while the suspected modified form had differing MW and pI values than those calculated. The few identified proteins that did not match well either have suspected

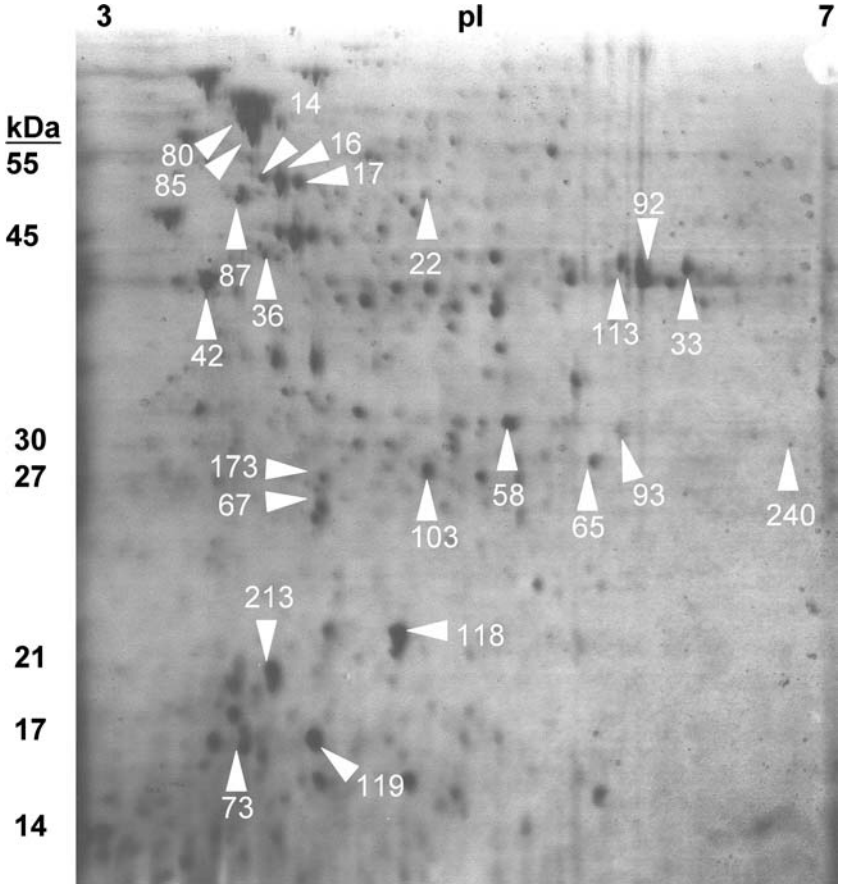

Fig. 1 A representative two-dimensional gel. A representative two-dimensional Coomassie-stained protein gel from late stationary phase of strain 824 (pIMP1). The protein spot numbers shown were identified by mass spectroscopy. The locations are indicated on the gel and the characteristics are listed in Table 4

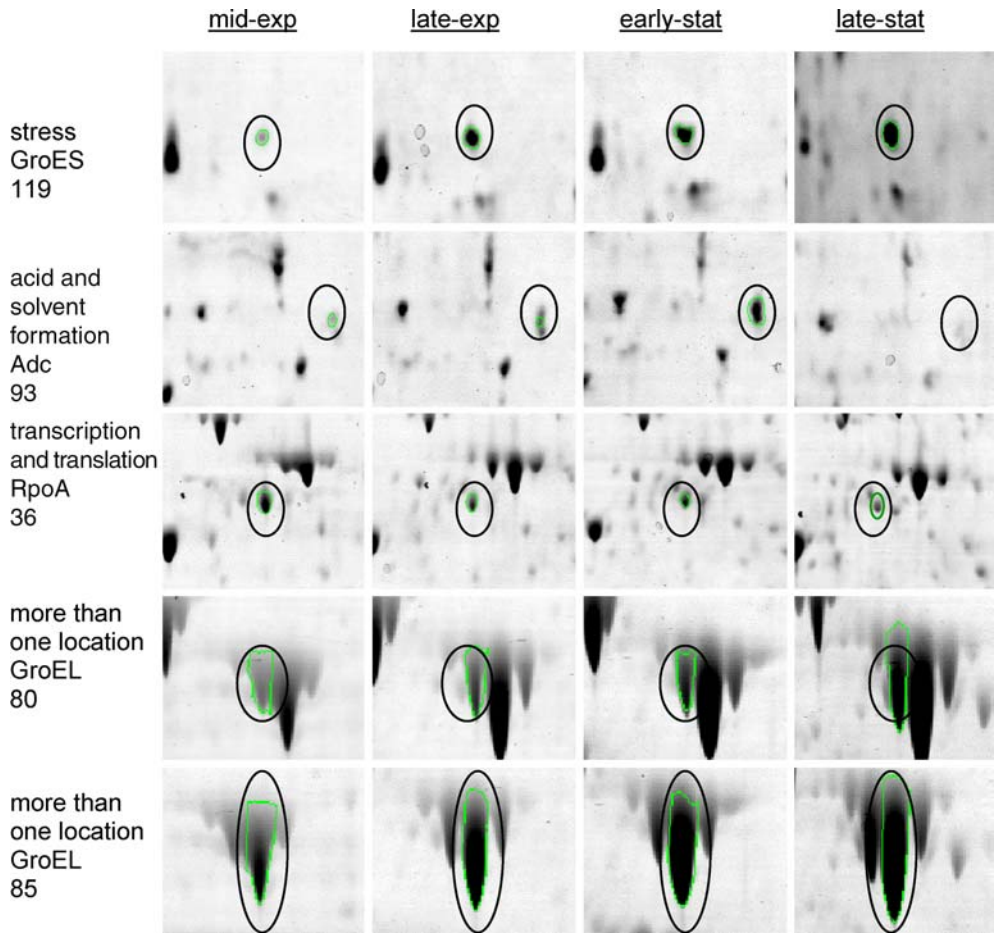

Fig. 2 Growth phase-dependent protein level. The protein expression profile for four representative proteins, one from each functional classification, are shown on corresponding sections of Coomassie-stained two-dimensional gels

post-translational modifications, as will be discussed below, or may be potential degradation products.

Upregulated proteins observed

Many proteins identified were upregulated during latter stage growth: stress proteins, the known solvent phase inducible Adc and others. The stress proteins upregulated are GroEL, GroES, heat shock protein 18 (Hsp18), and phage shock protein A (PspA). Stress proteins are under complex and multiple controls [36]. GroEL and its cofactor GroES prevent misfolding and promote refolding and proper initial folding of unfolded polypeptides [40]. GroEL was identified in two locations on the two-dimensional gel, directly juxtaposed to one another. GroES was identified in one location on the gel. *groEL* and *groES* form a two gene operon [28], and each is present in one copy in the *C. acetobutylicum* genome. Hsp18 is induced by heat shock and solventogenesis. Hsp18 is induced by different environmental conditions and this induction suggests that it is a general stress inducible protein [37]. Hsp18 was also identified in two locations on the two-dimensional gel, yet one copy of *hsp18* is present in the genome. One Hsp18 variant, spot 213, was not detectable until late solvent phase, when it was highly abundant. This result is consistent with two protein variants identified previously [38]. The *Escherichia coli* PspA homolog is synthesized in response to several stresses (heat, ethanol, and osmotic shock) [8], and it is involved in maintaining the proton motive force and membrane integrity under stress conditions [17, 23]. PspA is upregulated and does not require Spo0A, as its expression is detected in strain SKO1.

The gene encoding *adc* was expected to be induced at solventogenesis [2, 15] and expected to require Spo0A for expression [35]. Adc is essential for solvent production, because it catalyzes the decarboxylation of acetoacetate to acetone [33]. Adc was identified in one location, was noticeably induced at solventogenesis, and its protein expression is absent in strain SKO1. Contrary to the current results, two protein spots were previously identified [38]. The spot corresponding to the other location may not have been analyzed in this study.

Among the other upregulated proteins are the glycogen-binding regulatory subunit of serine/threonine phosphatase I (Ppi), and two proteins containing repeats of a novel protein domain, Clostridial hydrophobic with a conserved tryptophan (ChW), ChW14 and ChW16/17. The function of Ppi is unknown, but may have a function similar to phosphatase-I of eukaryotes because a gene encoding a cAMP-dependent protein kinase is present in the genome. Phosphatase-I downregulates the transcriptional response caused by a rise in cAMP,

Table 4 Characteristics of identified *C. acetobutylicum* proteins

Protein, short name (Annotation ORF #)	Spot #	% Mass[a]	MW calc[b]	MW app[c]	pI calc	pI app
Stress proteins						
GroEL (CAC2703)	80	67	58.1	60–58[d]	4.70	3.9
GroEL (CAC2703)	85	83	58.1	60–56	4.70	4.0
GroES (CAC2704)	119	52	10.4	16	4.81	4.1
Heat shock protein 18, Hsp18 (CAC3714)	213	50	17.7	20–19	5.09	4.0
Heat shock protein 18, Hsp18 (CAC3714)	118	66	17.7	22	5.09	5.0
Phage shock protein A, PspA (CAC0313)	103	42	25.0	27	5.16	5.0
Acid and solvent formation proteins						
Acetoacetate decarboxylase, Adc (CAP0165)	93	38	27.5	30	5.77	6.0
Fructose-bisphosphate aldolase, Fba (CAC0827)	58	69	30.4	31	5.36	5.4
Glucose-6-phosphate isomerase, Pgi (CAC2680)	22	29	49.8	47	5.21	4.8
Glyceraldehyde-3-phosphate dehydrogenase, GapC (CAC0709)	92	46	35.9	42	5.87	6.0
Triosephosphate isomerase, Tpi (CAC0711)	42	42	26.5	42	5.68	3.6
Triosephosphate isomerase, Tpi (CAC0711)	65	42	26.5	26	5.68	6.0
Butyryl-CoA dehydrogenase, Bcd (CAC2711)	113	36	41.4	42	5.92	6.0
Butyryl-CoA dehydrogenase, Bcd (CAC2711)	33	56	41.4	42	5.92	6.3
Transcription and translation proteins						
DNA-dependent RNA polymerase α subunit, RpoA (CAC3104)	36	49	35.4	43	4.75	4.0
Spo0A, Spo0A (CAC2071)	240	60	31.5	30	6.42	6.9
Ribosomal protein L7/L12, RplL (CAC3145)	73	43	12.6	16	4.66	4.0
Translation elongation factor P, Efp (CAC2094)	67	51	21.1	25	4.78	4.4
Other proteins						
Glycogen-binding regulatory subunit of S/T protein phosphatase I, Ppi (CAP0129)	173	61	28.0	27	6.09	4.4
ATP synthase subunit δ, AtpD (CAC2865)	87	69	51.1	50	4.68	3.8
Protein containing ChW-repeats, ChW14 (CAC1532)	14	45	53.6	54	4.85	4.0
Protein containing ChW-repeats, ChW16/17 (CAC2584)	16	48	52.9	54–51	5.02	4.1
Protein containing ChW-repeats, ChW16/17 (CAC2584)	17	44	52.9	54–52	5.02	4.2

[a]Percent of the total mass of the protein identified as the protein indicated based on 9–30 peptide sequences analyzed and compared
[b]Calc, Calculated MWs and pIs were calculated using the statistical analysis of protein sequence and isoelectric point determination programs, respectively, located at http://www.workbench.sdsc.edu
[c]App, Apparent MWs and pIs as observed on the two-dimensional gel
[d]A range of MWs is given for proteins where their spots were elongated lengthwise

including the downregulation of cAMP-stimulated glycogen breakdown. Because *C. acetobutylicum* accumulates intracellular granules of a reserve polyglucan material (granulose) before sporulation, and because granulose serves as a carbon and energy source by being degraded during endospore formation, Ppi may have a role in granulose metabolism. Ppi is not detectable until solventogenesis and is absent in strain SKO1. ChW14 and ChW16/17 are suggested to be novel extracellular or cell surface system proteins related to those used in cell adhesion [29]. ChW14 appears upregulated until late solventogenesis, when it is no longer detected. ChW16/17 was identified in two locations, spots 16 and 17, which is detected at all growth phases. Both ChW14 and ChW16/17 are absent in strain SKO1.

Downregulated proteins observed

The identified proteins that appear downregulated are the glucose metabolism isomerases, glucose 6-phosphate isomerase (Pgi), and triosephosphate isomerase (Tpi), energy metabolism ATP synthase subunit δ (AtpD), the core solventogenic butyryl-CoA dehydrogenase (Bcd), and the translation regulating ribosomal protein L7/L12 (RplL). Pgi is a housekeeping enzyme of metabolism that catalyzes the interconversion of glucose 6-phosphate and fructose 6-phosphate in the glycolytic pathway [10]. Decreasing levels of Pgi may be a result of glycolysis being most active in the exponential growth phase. Tpi is a ubiquitous glucose metabolism enzyme that catalyzes the conversion between dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate. One gene for Tpi is present in the *C. acetobutylicum* genome. Tpi was identified in two locations on the two-dimensional gel; the two locations have differing MWs and pIs and thus are not near each other. One Tpi variant, spot 65, appears downregulated. Tpi from various clostridia have different heat stabilities [13], and the Tpi from the three examined clostridia all had the same MW of 53 kDa as dimers, but with different pIs [41]. ATP synthase catalyzes the formation of ATP from ADP and inorganic phosphate using a proton or sodium ion gradient [12]. AtpD abundance decreases during exponential phase and is undetectable by late solventogenesis. Bcd catalyzes the conversion of crotonyl-CoA to butyryl-CoA at which point the pathway diverges to production of butyrate or butanol. Its single encoded gene is clustered

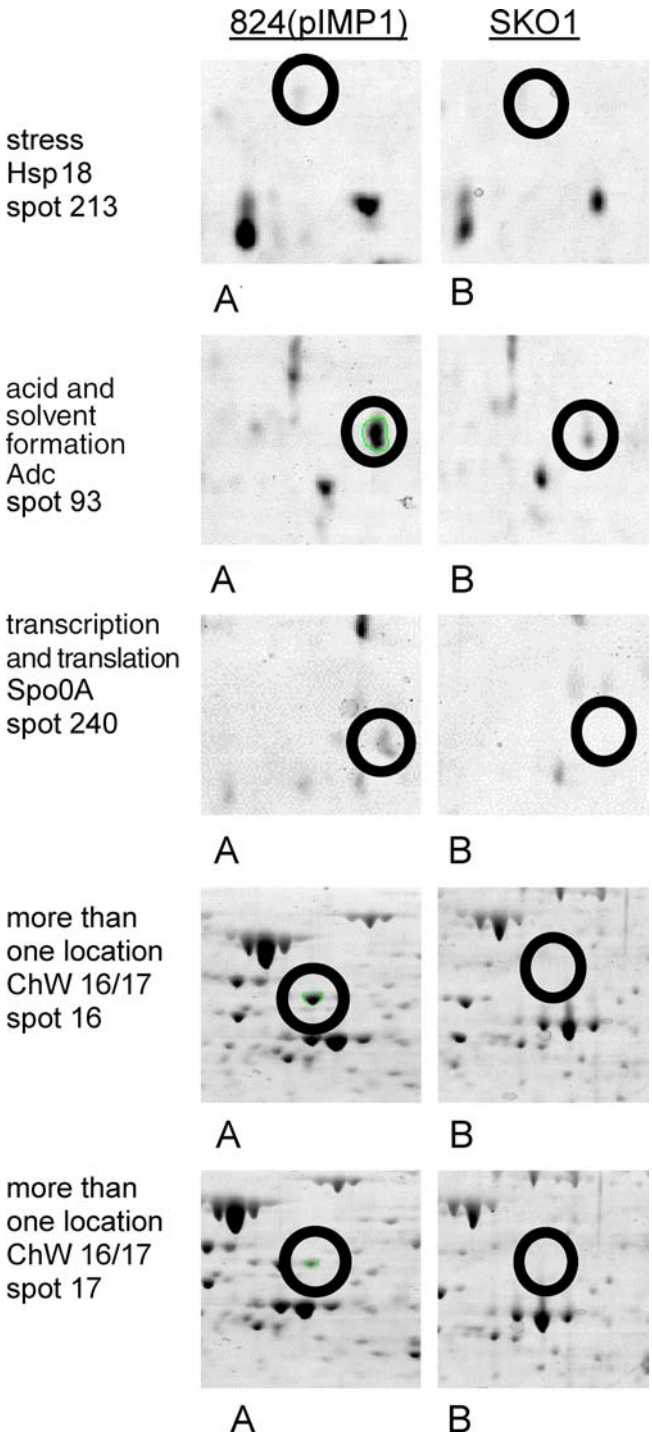

Fig. 3 Profile of several proteins requiring Spo0A for regulation. Six proteins identified require Spo0A for protein expression. Corresponding sections of Coomassie stained two-dimensional gels the six proteins from early stationary phase in **a** wild type (pIMP1) and **b** SKO1 strains are illustrated. Spo0A (**b**) protein itself is absent in strain SKO1

with β-hydroxybutyryl-CoA, crotonase, and electron transfer flavoprotein α and β subunits to form an operon. Bcd was identified in two distal locations, one of which, spot 33, appeared downregulated. The *E. coli* homolog of RplL is a ribosomal protein central to the translocation step of translation [27]. RplL is the heterodimer formed from two proteins, L7 and L12, and is anchored to the large ribosomal subunit L10 [16]. Unlike the usual situation where subunits for a protein complex are encoded by different transcriptional units, the subunits L7 and L12 are encoded by the same gene, and L12 is converted to L7 by a post-translational acetylation at its N-terminus [22]. In this study we identified only one RplL but cannot determine which subunit form was detected. Ribosomal protein L7/L12 is highly abundant in mid and late exponential and decreases by late solvent phase.

Relatively constitutive abundant proteins observed

The proteins constitutively abundant are the solvent-controlling Spo0A, the metabolic enzymes glyceraldehyde-3-phosphate dehydrogenase (GapC), one variant of Bcd, one variant of Tpi, and fructose bisphosphate aldolase (Fba), the transcription regulating RNA polymerase α (RpoA), and the translation elongation factor P (Efp). All of the proteins identified with relatively constitutive abundance are acid and solvent formation. These proteins may be more stable compared to the acid and solvent formation, which appear to be downregulated. Spo0A is a transcriptional regulator that positively controls sporulation and solvent production [18]. As expected, no Spo0A protein is detected in strain SKO1 because *spo0A* is disrupted in this strain. GapC catalyzes the reversible phosphorylation of glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate [39]. GapC was identified in one location, but the gene encoding the enzyme is present in the genome twice. The previously identified GapC [38] agrees in MW and pI with the one identified in this study, but the abundances of the protein in the two studies are disparate. This study indicates that GapC is constitutively abundant, while the previous study [38] observed an upregulation in solventogenesis. The second variant of Bcd, spot 113, is constitutively abundant at a low level. The second variant of Tpi, spot 42, is twice as abundant as the spot 65 variant. Fba is responsible for the interconversion of fructose 1,6-bisphosphate with glyceraldehyde 3-phosphate during glycolysis and gluconeogenesis. Fba class II has one gene in the genome that could encode this protein, and one protein spot was identified. The *E. coli* and *B. subtilis* Fba homologs are post-translationally modified [26, 32] suggesting that *C. acetobutylicum* may have another Fba variant, as yet unidentified. RpoA is part of the core RNA polymerase that transcribes genes specified by σ factor promoter recognition and thus is important for regulating gene expression [34]. RpoA is a moderately abundant protein. RpoA MW corresponds to that previously reported [34]. In *E. coli*, Efp stimulates efficient translation and peptide-bond synthesis [3]. In *B. subtilis*, Efp is maintained at a high level at sporulation and is required for the sporulation

developmental program [31]. Efp is a protein present in low levels in all growth phases in *C. acetobutylicum*.

Many genes have putative Spo0A binding sites

Spo0A is a transcription factor that directly binds to a known 7 bp region [35], designated 0A box, in the upstream sequence of the genes it regulates. The intergenic regions of the genes encoding the proteins identified in this study were examined for possible Spo0A binding sites. The genes with putative 0A boxes are listed in Table 3.

Overexpression of *spo0A* alters many processes

In order to examine what processes Spo0A affects at the protein level, the proteome of *C. acetobutylicum* 824 (pMSPOA) was compared to that of 824 (pIMP1). As compared to the control strain, protein abundances were perturbed for many proteins of varying functions in *C. acetobutylicum* 824 (pMSPOA). Three analyzed proteins had lower levels. Two of the three were glycolytic proteins, Pgi, and Tpi, and showed levels decreasing during growth, while Tpi showed an accelerated decrease in protein levels during growth. The third was a ribosomal protein, RplL. Likewise, three proteins had higher levels in *C. acetobutylicum* 824 (pMSPOA) in addition to the expected increased level of Spo0A. Spo0A showed an accelerated increase in abundance as did one variant of Hsp18. A primary metabolism protein, Bcd, and an energy production protein, AtpD, showed an increase compared to the control strain.

Features of the *C. acetobutylicum* proteome

This study identified two features of the *C. acetobutylicum* proteome. Several proteins were identified in more than one location on the two-dimensional gel: GroEL, Hsp18, Tpi, Bcd, and ChW16/17. Proteins clustering in orientations reminiscent of straightened fingers on a hand were prominent in the larger MWs and more acidic pIs. One protein obviously displaying this feature was the adjacent two variants of GroEL. Two other proteins formed this orientation to a lesser degree, ATP synthase subunit δ and ChW16/17. In the case of ATP synthase subunit δ, the middle protein of the cluster was identified, but the spot on either side was not analyzed. ChW16 and ChW17 were adjacent proteins in the same cluster.

Discussion

Post-translational modification

The appearance of proteins at more than one location on the two-dimensional gel is indicative of post-translational modification [36]. Post-translational phosphorylation and glycosylation modifications of *C. acetobutylicum* proteins have been previously reported [5, 24]. Of the 23 proteins identified, five localized in more than one position. They are GroEL, Hsp18, Tpi, Bcd, and ChW16/17. More than one location for GroEL was expected because *E. coli* GroEL is phosphorylated as a result of the heat shock stress response [40]. GroEL was previously identified albeit in one location [38]. GroEL may be present in more than one location in the previous study and might perhaps be identified upon further protein spot identification. The MWs and pIs agree in both studies. Two locations for Hsp18 were detected and one variant was highly induced in the solventogenic phase in both the previous study [38] and this study, thus these data are consistent.

While post-translational modifications may be intended by the organism to confer altered protein activities, such as phosphorylation and glycosylation, unintended modifications may also occur. Longer growth conditions that result in stasis and a stressed environment induce oxidative modifications to *E. coli* proteins [30]. The selective pressure of certain antibiotics can also increase the level of carbonylation, an oxidative protein modification [30]. Additionally, erroneous incorporation of amino acids into a protein under such conditions may result in a protein migrating to a spot on a two-dimensional gel with a similar molecular weight to the authentic protein but with subtle differences in isoelectric points [30]. If unintended oxidative changes are incorporated into the proteins by stress conditions then these modifications could account for the two isoforms of GroEL and ChW16/17. These two proteins have both isoforms migrating to the same molecular weights with slightly differing isoelectric points. Protein oxidation could not account for the protein isoforms that migrate to spots distal from each other, such as Hsp18, Tpi, and Bcd.

Perhaps, different culture conditions may induce a more stressful environment for the cells and thus oxidative changes to the proteins. The culture conditions between the previous study [38] and this study were different. This study used the standard clostridia growth media, which has been routinely used for solvent production in our previous work, rather than phosphate-limited mineral media [38]. Moreover, it is likely that sample handling may have also differed between the two studies. The more time that has elapsed between the sample being taken and its electrophoresis may inadvertently increase the risk of oxidative protein changes.

Comparison between transcriptome and proteome expression

The expression level of many genes has been reported and where possible we have examined the correlation between protein level and gene expression. Many protein levels correlate well with their RNA profiles in the wild

type and control strains. The protein and RNA levels examined were the gene expression protein (RpoA); Ppi; the Spo0A-controlled proteins (Spo0A and Adc), the heat shock stress response proteins (GroEL, GroES, and Hsp18), and the acid and solvent formation (Fba, GapC, Tpi, Pgi, and Bcd). RpoA and Ppi both correlate well with their respective RNA profiles [42].

The Spo0A-controlled proteins Spo0A and Adc were compared with their RNA levels. A lower level of Spo0A RNA is present in *C. acetobutylicum* SKO1 than in the parent strain *C. acetobutylicum* 824 at all time points [42]. This result is expected since the *spo0A* gene is disrupted in strain SKO1. The Spo0A RNA levels [19] and protein levels correlate well in *C. acetobutylicum* 824 and *C. acetobutylicum* 824 (pIMP1), respectively. A lower level of Adc RNA is present in strain SKO1 than in *C. acetobutylicum* 824 at all time points. This finding is consistent with previous studies that show Spo0A controls the expression of *adc* [35, 42]. The Adc RNA levels [19] and protein levels correlate well in *C. acetobutylicum* 824 and *C. acetobutylicum* 824 (pIMP1), respectively. The RNA profile for Adc does not change when *spo0A* is overexpressed as compared to *C. acetobutylicum* 824; both peak at the transition [1, 19].

The heat shock stress response RNA profiles for GroEL and GroES are identical [42] as expected because they are part of a two gene operon and thus co-transcribed [28]. A higher abundance of GroEL and GroES RNA is present in *C. acetobutylicum* 824 than in strain SKO1 in exponential phase, and then at the transition RNA levels decrease. During stationary phase, RNA levels again increase. The overexpression of the *groEL* and *groES* genes can increase solvent production [42] and tolerance [43]. In *C. acetobutylicum* 824 (pMSPOA), GroES RNA levels peak at higher levels compared to *C. acetobutylicum* 824 (pIMP1) and at earlier times, and then the RNA levels decrease during stationary phase [1]. Taken together, these data suggest that Spo0A positively regulates GroEL and GroES to attain maximal expression at the transition. Both variants of GroEL protein and the GroES protein have the same protein profile. They are characterized by an increase at the transition and are highly expressed during stationary phase.

Hsp18 RNA was present in lower levels in *C. acetobutylicum* 824 than in strain SKO1 until late exponential phase when RNA levels increase [42]. In *C. acetobutylicum* 824 (pMSPOA), Hsp18 RNA peaks at higher levels compared to *C. acetobutylicum* 824 (pIMP1) and at earlier times, and then the RNA levels decrease during stationary phase [1]. The Hsp18 RNA levels and protein levels of variant 213 correlate well. The pattern of Hsp18 is different from those of GroEL and GroES in the absence of Spo0A, perhaps representing a different stress alleviation function or perhaps representing different regulatory controls for Hsp18 than for GroEL and GroES.

RNA levels of many glycolytic genes (*fba, gapC, tpi, pgi*) and the primary metabolism gene (*bcd*) are lower at the transition and during stationary phase. Thus, glycolysis is strongly downregulated in stationary phase [1]. The RNA and protein levels of Fba and Tpi variant 65 correlate well and exhibit the expected molecular weights based on the gene sequence [42].

Instances in which the protein and transcript levels of Spo0A-affected expression do not agree may indicate more complex controls, which have yet to be elucidated. If the relative level of change is minor, the difference may be caused by experimental variation or measurement errors. Disagreement in correlation between RNA and protein levels may result from the instability of the RNA or protein, as this study did not examine their respective stabilities.

In conclusion, this study provided evidence of how the proteomic profile changes at the transition to solventogenesis. The heat shock stress proteins, proteins with ChW-repeats, Adc, and Ppi increased in protein abundance during growth. Some acid and solvent formation proteins (one variant of Bcd, Pgi, and one variant of Tpi), AtpD, and rplL had decreasing protein levels while other acid and solvent formation proteins (one variant of Bcd, Fba, GapC, and one variant of Tpi) and the transcription and translation proteins (RpoA, Efp, and Spo0A) demonstrated no significant change in protein levels during growth.

Moreover, the *spo0A* disrupted and overexpression strains were characterized and Spo0A implicated additional processes it may affect. Analysis of the *spo0A* knockout indicated six protein spots that were present in wild type *C. acetobutylicum* but absent in the SKO1 strain: Ppi, ChW14, Adc, Spo0A, ChW16, and ChW17. Thus, these proteins require Spo0A based on protein level determination. The finding that Adc requires Spo0A confirms previous studies [34]. Spo0A overexpression affected the abundance of proteins involved in glycolysis, translation, heat shock stress response, and energy production. Of the 23 proteins identified, five were located in more than one position indicating a post-translational modification. Some protein isoforms of the same protein, such as GroEL, clustered into arrangements resembling fingers of a straightened hand. Finally, this study serves as a basis for further explorations in comparing the proteome pattern with the results obtainable from microarrays in *C. acetobutylicum*.

Acknowledgements We thank Latonia Harris for sample collection, Beth Allen for two-dimensional gel electrophoresis and mass spectroscopy, and Carlos Paredes for 0A box determination. This work was supported by National Science Foundation grant # BES-0418289.

References

1. Alsaker KV, Spitzer TR, Papoutsakis ET (2004) Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. J Bacteriol 186:1959–1971

2. Andersch W, Bahl H, Gottschalk G (1983) Level of enzymes involved in acetate, butyrate, acetone and butanol formation in *Clostridium acetobutylicum*. Eur J Appl Microbiol Biotechnol 18:327–332
3. Aoki H, Adams SL, Chung DG, Yaguchi M, Chuang SE, Ganoza MC (1991) Cloning, sequencing and overexpression of the gene for prokaryotic factor EF-P involved in peptide bond synthesis. Nucleic Acids Res 19:6215–6220
4. Bahl H, Muller H, Behrens S, Joseph H, Narberhaus F (1995) Expression of heat shock genes in *Clostridium acetobutylicum*. FEMS Microbiol Rev 17:341–348
5. Balodimos IA, Rapaport E, Kashket ER (1990) Protein phosphorylation in response to stress in *Clostridium acetobutylicum*. Appl Environ Microbiol 56:2170–2173
6. Boynton ZL, Bennet GN, Rudolph FB (1996) Cloning, sequencing, and expression of clustered genes encoding beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. J Bacteriol 178:3015–3024
7. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254
8. Brissette JL, Russel M, Weiner L, Model P (1990) Phage shock protein, a stress protein of *Escherichia coli*. Proc Natl Acad Sci USA 87:862–866
9. Brown DP, Ganova-Raeva L, Green BD, Wilkinson SR, Young M, Youngman P (1994) Characterization of *spo0A* homologues in diverse *Bacillus* and *Clostridium* species identifies a probable DNA-binding domain. Mol Microbiol 14:411–426
10. Davies C, Muirhead H (2003) Structure of native phosphoglucose isomerase from rabbit: conformational changes associated with catalytic function. Acta Crystallogr D Biol Crystallogr 59:453–465
11. Durre P, Bohringer M, Nakotte S, Schaffer S, Thormann K, Zickner B (2002) Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum*. J Mol Microbiol Biotechnol 4:295–300
12. Externbrink T, Hujer S, Winzer K, Durre P (2000) Sequence analysis of the *atp* operon of *Clostridium acetobutylicum* DSM 792 encoding the F0F1 ATP synthase. DNA Seq 11:109–118
13. Finne G, Matches JR, Liston J (1975) A comparative study on the heat stability of triosephosphate isomerase in psychrophilic, psychrotrophic, and mesophilic clostridia. Can J Microbiol 21:1719–1723
14. Fischer RJ, Helms J, Durre P (1993) Cloning, sequencing, and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. J Bacteriol 175:6959–6969
15. Gerischer U, Durre P (1992) mRNA analysis of the *adc* gene region of *Clostridium acetobutylicum* during the shift to solventogenesis. J Bacteriol 174:426–433
16. Griaznova O, Traut RR (2000) Deletion of C-terminal residues of *Escherichia coli* ribosomal protein L10 causes the loss of binding of one L7/L12 dimer: ribosomes with one L7/L12 dimer are active. Biochemistry 39:4075–4081
17. Hankamer BD, Elderkin SL, Buck M, Nield J (2004) Organization of the AAA(+) adaptor protein PspA is an oligomeric ring. J Biol Chem 279:8862–8866
18. Harris LM, Desai RP, Welker NE, Papoutsakis ET (2000) Characterization of recombinant strains of the *Clostridium acetobutylicum* butyrate kinase inactivation mutant: need for new phenomenological models for solventogenesis and butanol inhibition? Biotechnol Bioeng 67:1–11
19. Harris LM, Welker NE, Papoutsakis ET (2002) Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J Bacteriol 184:3586–3597
20. Ireton K, Rudner DZ, Siranosian KJ, Grossman AD (1993) Integration of multiple developmental signals in *Bacillus subtilis* through the Spo0A transcription factor. Genes Dev 7:283–294
21. Jones DT, Woods DR (1986) Acetone-butanol fermentation revisited. Microbiol Rev 50:484–524
22. Kitaura H, Kinomoto M, Yamada T (1999) Ribosomal protein L7 included in tuberculin purified protein derivative (PPD) is a major heat-resistant protein inducing strong delayed-type hypersensitivity. Scand J Immunol 50:580–587
23. Kleerebezem M, Crielaard W, Tommassen J (1996) Involvement of stress protein PspA (phage shock protein A) of *Escherichia coli* in maintenance of the protonmotive force under stress conditions. EMBO J 15:162–171
24. Lyristis M, Boynton ZL, Peterson D, Kan Z, Bennett GN, Rudolph FB (2000) Cloning, sequencing and characterization of the gene encoding flagellin, *flaC*, and the post-translational modification of the flagellin from *Clostridium acetobutylicum* ATCC 824. Anaerobe 6:69–79
25. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET (1992) Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. Biotechnology (N Y) 10:190–195
26. Mitchell C, Morris PW, Lum L, Spiegelman G, Vary JC (1992) The amino acid sequence of a *Bacillus subtilis* phosphoprotein that matches an orfY-tsr coding sequence. Mol Microbiol 6:1345–1349
27. Montesano-Roditis L, Glitz DG, Traut RR, Stewart PL (2001) Cryo-electron microscopic localization of protein L7/L12 within the *Escherichia coli* 70 S ribosome by difference mapping and Nanogold labeling. J Biol Chem 276:14117–14123
28. Narberhaus F, Bahl H (1992) Cloning, sequencing, and molecular analysis of the *groESL* operon of *Clostridium acetobutylicum*. J Bacteriol 174:3282–3289
29. Nolling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR (2001) Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. J Bacteriol 183:4823–4838
30. Nystrom T (2004) Stationary-phase physiology. Annu Rev Microbiol 58:161–181
31. Ohashi Y, Inaoka T, Kasai K, Ito Y, Okamoto S, Satsu H, Tozawa Y, Kawamura F, Ochi K (2003) Expression profiling of translation-associated genes in sporulating *Bacillus subtilis* and consequence of sporulation by gene inactivation. Biosci Biotechnol Biochem 67:2245–2253
32. Packman LC, Berry A (1995) A reactive, surface cysteine residue of the class-II fructose-1,6-bisphosphate aldolase of *Escherichia coli* revealed by electrospray ionisation mass spectrometry. Eur J Biochem 227:510–515
33. Petersen DJ, Bennett GN (1990) Purification of acetoacetate decarboxylase from *Clostridium acetobutylicum* ATCC 824 and cloning of the acetoacetate decarboxylase gene in *Escherichia coli*. Appl Environ Microbiol 56:3491–3498
34. Pich A, Bahl H (1991) Purification and characterization of the DNA-dependent RNA polymerase from *Clostridium acetobutylicum*. J Bacteriol 173:2120–2124
35. Ravagnani A, Jennert KC, Steiner E, Grunberg R, Jefferies JR, Wilkinson SR, Young DI, Tidswell EC, Brown DP, Youngman P, Morris JG, Young M (2000) Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol Microbiol 37:1172–1185
36. Rosen R, Ron EZ (2002) Proteome analysis in the study of the bacterial heat-shock response. Mass Spectrom Rev 21:244–265
37. Sauer U, Durre P (1993) Sequence and molecular characterization of a DNA region encoding a small heat shock protein of *Clostridium acetobutylicum*. J Bacteriol 175:3394–3400
38. Schaffer S, Isci N, Zickner B, Durre P (2002) Changes in protein synthesis and identification of proteins specifically induced during solventogenesis in *Clostridium acetobutylicum*. Electrophoresis 23:110–121
39. Schreiber W, Durre P (1999) The glyceraldehyde-3-phosphate dehydrogenase of *Clostridium acetobutylicum*: isolation and purification of the enzyme, and sequencing and localization of the gap gene within a cluster of other glycolytic genes. Microbiology 145(Pt 8):1839–1847

40. Sherman M, Goldberg AL (1992) Heat shock in *Escherichia coli* alters the protein-binding properties of the chaperonin *groEL* by inducing its phosphorylation. Nature 357:167–169
41. Shing YW, Akagi JM, Himes RH (1975) Psychrophilic, mesophilic, and thermophilic triosephosphate isomerases from three clostridial species. J Bacteriol 122:177–184
42. Tomas CA, Alsaker KV, Bonarius HP, Hendriksen WT, Yang H, Beamish JA, Paredes CJ, Papoutsakis ET (2003) DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5. J Bacteriol 185:4539–4547
43. Tomas CA, Beamish J, Papoutsakis ET (2004) Transcriptional analysis of butanol stress and tolerance in *Clostridium acetobutylicum*. J Bacteriol 186:2006–2018