Journal of Industrial Microbiology & Biotechnology (2001) 27, 307–313
© 2001 Nature Publishing Group 1367-5435/01 $17.00

www.nature.com/jim

Orf5/SolR: a transcriptional repressor of the *sol* operon of *Clostridium acetobutylicum*?[1]

K Thormann and P Dürre

Mikrobiologie und Biotechnologie, Universität Ulm, D-89069 Ulm, Germany

The gene of Orf5 (SolR) of *Clostridium acetobutylicum* DSM 792 was subcloned and overexpressed in *Escherichia coli*. The protein was purified with Ni-NTA agarose and used for DNA binding assays. No DNA binding of Orf5 to regions upstream of the *sol* operon from *C. acetobutylicum* was observed. Overexpression of Orf5 in *C. acetobutylicum* led to a change in the organism's pattern of glycosylated exoproteins. The Orf5 protein was localized in the cell membrane fraction and to a small extent in the supernatant medium. Based on these results Orf5 (SolR) appears not to act as a transcriptional repressor in *C. acetobutylicum*, but instead may be an enzyme involved in glycosylation or deglycosylation. *Journal of Industrial Microbiology & Biotechnology* (2001) **27,** 307–313.

Keywords: *Clostridium acetobutylicum*; glycosylation; Orf5/SolR; repressor; solventogenesis

Introduction

While during exponential growth of *Clostridium acetobutylicum* acids (mainly acetate and butyrate) are the predominant products, the organism switches to formation of solvents (with butanol and acetone as most characteristic products) at the onset of stationary phase. Several factors such as low pH, concentration of butyrate, acetate, and substrate, temperature, and nutrient limitations [33,39] are responsible for the shift. In recent years, most of the enzymes directly involved in solvent formation have been described. Their genes are regulated at the transcriptional level [3,11] and induced at the onset of solventogenesis. However, the molecular mechanism leading to their activation/derepression is not well understood.

The polycistronic *sol* operon consists of the genes *adhE*, *ctfA* and *ctfB* encoding proteins that are involved in formation of both butanol and acetone. Like the *adc* gene, encoding acetoacetate decarboxylase that catalyzes the final step in acetone formation, the *sol* operon is located not on the chromosome, but on the megaplasmid pSOL1 [9]. Its expression starts early at the onset of solventogenesis. Sequencing of the region upstream of the *adhE* gene in strain DSM 792 revealed an open reading frame, designated *orf5*, 954 bp in length in the same orientation of transcription, encoding a protein characterized by a hydrophobic N-terminus with homology to signal peptides [12,13] (EMBL data base accession number X72831). Studies on the homologous open reading frame in *C. acetobutylicum* ATCC 824 [30] indicated that inactivation of the gene results in an enhanced expression of the *sol* operon accompanied by a significant increase of solvent production. On the other hand, overexpression of the encoded protein led to a complete loss of solvent formation. Thus, a role of this protein, designated SolR, as a transcriptional repressor for the *sol* operon was postulated.

In this study we report the overexpression and purification of Orf5 of *C. acetobutylicum* DSM 792, allowing the protein's characterization, localization, and elucidation of its DNA binding. The present evidence suggests that Orf5 does not act as a transcriptional repressor in solventogenesis but rather is involved in protein glycosylation/deglycosylation.

Materials and methods

Bacterial strains, plasmids, and growth conditions

C. acetobutylicum DSM 792 was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (Braunschweig, Germany). The strain was grown anaerobically in 2× YT medium [32] at 37°C; for product analyses the cultivation was carried out in MES-buffered mineral medium [5]. For primary vector construction and heterologous overexpression in *Escherichia coli* XL1-B [8] was used and for *in vivo* plasmid methylation *E. coli* ER2275 (pAN1) [26] was used. Both strains were routinely grown aerobically in Luria Bertani broth (LB) at 37°C. When applicable, media were supplemented with ampicillin (100 μg/ml), chloramphenicol (30 μg/ml), or erythromycin (50 μg/ml). The plasmids used in this study are listed in Table 1.

Gas chromatography

Cell samples were centrifuged at 20 000×*g* for 10 min. One milliliter of the supernatant fluid was acidified with 0.1 ml 2 N HCl containing 110 mM isobutanol (final concentration in the sample: 10 mM) as internal standard. Subsequently, 1 μl was used for detection and quantification of fermentation products using a Chrompack CP9001 gas chromatograph equipped with a flame ionization detector (Chrompack, Frankfurt, Germany). The following products were measured: acetone, butanol, ethanol, butyrate, and acetate. Separation took place in a Chromosorb 101 column (2 m; 80–100 mesh) at 155–197°C (9°C/min), using N_2 as carrier gas (30 ml/min). The injector temperature was 195°C

Correspondence: Dr P Dürre, Mikrobiologie und Biotechnologie, Universität Ulm, D-89069 Ulm, Germany

[1]This article is dedicated to Gerhard Gottschalk, University of Göttingen, Germany, who contributed enormously to our knowledge on solventogenic clostridia by numerous publications, on the occasion of his 65th birthday.

Received 12 September 2000; accepted 26 March 2001

Table 1 Plasmids used in this study

Plasmid	Relevant characteristics[a]	Source or reference
pUC18	Ap^R	[41]
pAN1	Cm^R Φ3T I	[26]
pIMP1	MLS^R Ap^R	[25]
pMM40	Ap^R	[16]
pK9	Ap^R *orf5 adhE ctfA ctfB*	[13]
pUG80	Ap^R *ctfB adc*	[14]
pSP1	Ap^R	This study
pMMORF5N	Ap^R *orf5*	This study
pMMORF5H	Ap^R *orf5x6His*	This study
pORF5N	MLS^R *orf5*	This study
pORF5H	MLS^R *orf5x6His*	This study

[a]Ap^R, ampicillin resistance; Cm^R, chloramphenicol resistance; MLS^R, macrolide, lincosamide, and streptogramin B resistance; Φ3T I, Φ3T methylase. Genes listed encode the following proteins:, Orf5; Orf5x6His (carrying an additional His6-tag); AdhE, aldehyde/alcohol dehydrogenase; CtfA/B, acetoacetyl coenzyme A:acetate/butyrate:coenzyme A-transferase subunits; and Adc, acetoacetate decarboxylase.

and that of the detector was 230°C. Signal analysis was performed using the program Maestro II (V. 2.1).

DNA isolation, transformation and manipulation

Molecular biology reagents were obtained from MBI Fermentas (St. Leon-Rot, Germany) and New England Biolabs (Schwalbach, Germany) and were used according to the manufacturer's instructions with the buffers provided. Chromosomal DNA of *C. acetobutylicum* was isolated by the method described by Bertram and Dürre [4]. Plasmids were isolated from *E. coli* using the GFX Micro Plasmid Prep Kit (Amersham Pharmacia Biotech Europe, Freiburg, Germany), large-scale plasmid isolation was performed with the Qiagen Midi Kit (Qiagen, Hilden, Germany). Plasmid DNA was desalted and concentrated using Microcon 100 microconcentrators (Amicon, Witten, Germany). Methods to electrotransform *E. coli* and *C. acetobutylicum* were described previously [10,31]. Before electroporation of *C. acetobutylicum*, plasmids pORF5N and pORF5H were methylated *in vivo* in *E. coli* ER2275 (pAN1) protecting the DNA from restriction by clostridial endonuclease activity [26].

PCR

PCR amplifications were performed in 100-μl volumes containing the relevant primers (100 pM each), deoxyribonucleoside triphosphates (200 μM each), DNA template (10–20 ng), 2.5 mM $MgSO_4$, tetramethylammoniumchloride (up to 50 mM), and DeepVent polymerase (2 U; NEB, Schwalbach, Germany) in the recommended buffer. A Peltier Thermal Cycler (Biozym Diagnostik, Hess. Oldendorf, Germany) was used with the following conditions: 95°C for 45 s, primer-dependent annealing temperature [38] for 45 s, 72°C for 1 min per 1 kb for 30–35 cycles. Purification of PCR products and removal of enzymes and nonincorporated nucleotides and primers were done using the NucleoSpin Extract-Kit (Macherey-Nagel, Düren, Germany).

Construction of plasmids

(i) *orf5* overexpression plasmids pMMORF5N and pMMORF5H: The DNA fragment containing the *orf5* reading frame was amplified by PCR, using plasmid pK9 [13] as template DNA. The upstream primer ORF5-RBS-UP (5′-AAAAGGAA**TTC**-*GAGGA*ATTTAGCATGAATTTATTAAATC-3′; bp 555–593 [[13], EMBL data base accession number X72831]) was designed to introduce an *Eco*RI cleavage site (underlined) and a ribosome binding site (italics) to enhance the protein's translation by nucleotide substitution (bold). The downstream primer for plasmid pMMORF5N ORF5-NAT-RV (5′-TAAG**CT**G**C**AGCTTCTTT-TATACTAAAAATTTTCC-3′; bp 1516–1550 [[13], EMBL data base accession number X72831] was generated on the complementary strand to provide a *Pst*I site (underlined), exchanged bases are marked in bold. The downstream primer for plasmid pMMORF5-H, ORF5-HIS-RV (5′-**GTAACTGCAGTTA*AT-GATGATGGTGATGATG***TACTAAAAATTTTCCGTTAAGTAT-TTTTTTATCATCGATTTC-3′; bp 1490–1563 [[13], EMBL data base accession number X72831]) additionally introduces six histidine codons (italics) at the 3′-end of the gene. The noncomplementary 5′-part of the primer is marked in bold. The amplified PCR products were subsequently digested with *Eco*RI and *Pst*I, and the resulting fragments were cloned into *Eco*RI–*Pst*I-digested pMM40 vector to yield the 6.4-kb plasmids pMMORF5N and pMMORF5H.

(ii) Shuttle plasmids pORF5N and pORF5H: the DNA amplification performed was similar to that described for the plasmids pMMORF5N/H, with the exception that the upstream primer was replaced by ORF5-P-UP (5-GGCGT**G**AAT**T**CGT-GAACAATTG-3′; bp 351–372 [[13], EMBL data base accession number X72831]), in which by two nucleotide substitutions (marked in bold) an *Eco*RI site was constructed (underlined). The DNA fragments obtained and shuttle vector pIMP1 were digested with *Eco*RI and *Pst*I and subsequently ligated, yielding the plasmids pORF5N and pORF5H. Both plasmids were 5.9 kb in length and carried the *orf5*(*x6His*) gene with its natural promoter.

(iii) Gel retardation plasmid pSP1: a DNA fragment containing the *sol* promoter region was amplified by PCR with plasmid pK9 [13] as DNA template. By nucleotide substitutions (bold) in upstream primer FW-ADHE (5′-CATAAATATAAACT**G-C**AGTCTATTTATGCTCC-3′, bp 1699–1730 [[13], EMBL data base accession number X72831]) and downstream primer RV-ADHE (5′-CATC**G**AATTCCTTTACTGTTGTGAC-3′, bp 2204–2228 [[13], EMBL data base accession number X72831]) restriction sites (underlined) for *Pst*I and *Eco*RI were introduced. The fragments obtained were ligated into *Eco*RI–*Pst*I-digested pUC18 vector, resulting in plasmid pSP1 (3.1 kb).

Heterologous expression and enrichment of Orf5x6His

E. coli XL1-B transformed with pMMORF5H was grown aerobically in LB medium. On reaching an optical density (600 nm) of 0.7–0.8, expression of Orf5x6His was induced by adding 1 mM isopropyl-β-thiogalactopyranoside (IPTG). After a further 3 h of growth, the cells were harvested by centrifugation (5000$\times g$, 10 min, 4°C). A crude extract was prepared by washing and suspending the cells in imidazole buffer (50 mM Na_2HPO_4, pH 8.0; 300 mM NaCl; 10 mM imidazole), followed by four passages through a French Press (SLM Instruments, Urbana, IL) at 12.5 mPa and centrifugation (30 min, 30,000$\times g$, 4°C). The 6xHis-tagged Orf5 protein was purified using Ni-NTA agarose with the buffers recommended by the manufacturer (Qiagen) at 4°C. Protein elution occurred at 250 mM imidazole. The

Orf5x6His-containing fractions were dialyzed against 20 mM Tris–HCl (pH 7.5), 10% (v/v) glycerol, 15% (w/v) PEG 20 000 and stored at −70°C.

Protein determination, electrophoresis, and precipitation

Protein determination was carried out according to the method of Bradford [7]. Bio-Rad protein assay solution (Bio-Rad Laboratories, München, Germany) was used as the dye reagent and bovine serum albumin as standard protein. Polyacrylamide gel electrophoresis (PAGE) was performed under denaturing conditions as described by Laemmli [19]. As molecular mass standard the broad range prestained protein marker (NEB, Schwalbach, Germany) was used, consisting of the following proteins: maltose-binding protein (MBP) fused β-galactosidase (175 kDa), maltose-binding protein (MBP) fused paramyosin (83 kDa), glutamic dehydrogenase (62 kDa), aldolase (47.5 kDa), triose-phosphate isomerase (32.5 kDa), β-lactoglobulin A (25 kDa), lysozyme (16.5 kDa), and aprotinin (6.5 kDa). The gels were silver-stained by the method of Blum *et al* [6].

Concentration of proteins was performed by trichloroacetic acid (TCA) precipitation. TCA (100%) was added to the samples to a final concentration of 10%, and after overnight incubation on ice the proteins were sedimented by centrifugation (15 000×*g*, 4°C, 10 min). After neutralization by washing with 5% (w/v) Na-acetate in 95% (v/v) ethanol and a final washing step with 95% (v/v) ethanol, the sediment was resolved in sample buffer (50 mM Tris–HCl, pH 6.8; 100 mM DTT; 2% (w/v) SDS; 0.1% (w/v) bromophenolblue; 10% (v/v) glycerol).

Membrane preparation

Preparations of washed membranes from *E. coli* and *C. acetobutylicum* were carried out as described by Siebers and Altendorf [36].

DNA binding assays

All linear DNA fragments for band shift assays were amplified by PCR with plasmid pK9 [13] for the *sol* operon and plasmid pUG80 [14] for the *adc* operon as templates. The *sol* operon promoter region was amplified in two parts, sol-I (319 bp; position −591 to −273 to the ATG start codon of the first structural gene *adhE*) and sol-II (324 bp; position −297 to +26 to the ATG start codon of the first structural gene *adhE*), using the primer pairs sol1F (5′-TAAGTTTTATATTTAGACCCTGGGG-3′)/sol1R-IRD (5′-AGGTCAAAAATATAACAGCTGTGT-3′) and sol2F (5′-TACACAGCTGTTATATTTTTGACC-3′)/sol2R-IRD (5′-AATTCCTTTACTGTTGTGACTTTC-3′). Amplification of the *adc* promoter (390 bp) was performed with the primer pair adcF (5′-GGAATTGTTTATAGTGTTTGTGAG-3′)/adcR-IRD (5′-TCATCCTTTAACATAAAAGTCACC-3′). The downstream primers were labeled with the fluorescent marker IRD-800. The band shift assay was carried out in a 10-μl reaction mixture consisting of 1 μl binding buffer (500 mM Tris–HCl, pH 7.4; 500 mM KCl; 50 mM $MgCl^2$; 10 mM EDTA; 10 mM DTT; 0.1% (w/v) BSA); 1 μl glycerol; 1 μl poly-[d(I-C)] (20 mg/ml); the labeled DNA fragment (0.5–2 ng), and the protein fraction. The binding reaction was performed at room temperature for up to 30 min. Subsequently, 1 μl of the mixture was separated on an automatic sequencer LI-COR 4000 L (Licor, Lincoln, NE) using a 5% polyacrylamide gel at 400 V and room temperature. Band shift assays with plasmid pSP1 were performed similarly as described for linear DNA fragments; 10 ng of DNA were used and the separation was carried out with the complete amount of the reaction mixture in a 2% (w/v) agarose gel with subsequent ethidium bromide staining.

His-tag detection

Proteins were transferred onto nitrocellulose membrane (Hybond ECL, Amersham Buchler, Braunschweig, Germany) as described by Kyhse-Andersen [18], using a Multiphor II NovaBlot unit (Amersham Pharmacia Biotech Europe). Subsequently, detection of His-tagged proteins was performed with Ni-AP conjugate following the manufacturer's (Qiagen) instructions.

Glycoprotein detection

Periodate oxidation and subsequent digoxigenin-succinyl-amidocaproic acid hydrazide labeling of the glycoproteins were carried out on the nitrocellulose membrane with a glycoprotein detection kit according to the description of the manufacturer (Roche Molecular Biochemicals, Mannheim, Germany).

Computer programs

Homology searches were performed using BLAST (release 2.0, September 1997) [1] on the WWW Blast Server (www.ncbi.nlm.-nih.gov). Searching the nr peptide sequence database (all non-redundant GenBank CDS translations, PDB, SwissProt, and Pir) was done using the program BLASTP.

Results

Subcloning of orf5 and purification of the gene product

For characterization of its proposed function as a repressor of solventogenesis, purified Orf5 protein was required. Therefore, the *orf5* gene was amplified *via* PCR. Due to the possible detrimental effect of Orf5 overproduction to *E. coli* cells (the protein carries a hypothetical signal peptide and has potential DNA binding activity), expression vector pMM40 was chosen. This vector is characterized by a hybrid *trplac* promoter under control of the $lacI^q$ allele of the *lac* repressor gene carried on the vector, providing a tight repression of gene expression before induction. The native ribosome binding site of the *orf5* gene, located 11–16 bases upstream of the start codon, resulted in a low expression of the protein. By introducing a ribosome binding-similar sequence [35] 7–11 bases upstream instead of the natural sequence, the amount of the heterologously expressed protein was enhanced (data not shown). An additional C-terminal His-tag fusion allowed easy purification of the protein by affinity chromatography.

Using Ni-NTA agarose, Orf5x6His was purified to >99% homogeneity from a crude extract of induced *E. coli* cells (Figure 1). The protein was also detected using Ni-NTA AP conjugate following Western blot membrane transfer. In a denaturating gel electrophoresis Orf5x6His migrated as a single band on a position equivalent to about 33 kDa in size, somewhat smaller than calculated from the sequence data (37.7 kDa). This

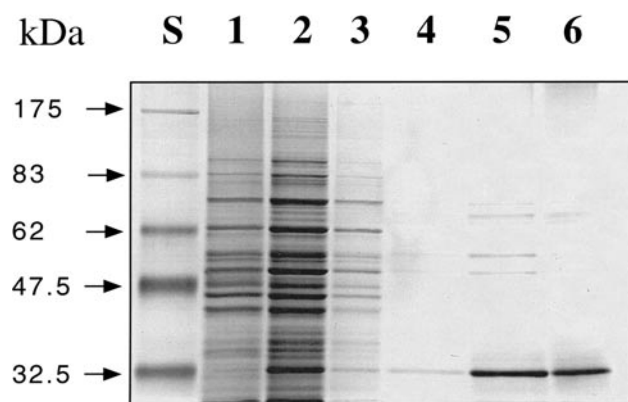

Figure 1 SDS-PAGE analysis (10% acrylamide) of *C. acetobutylicum* Orf5×6His overproduced in *E. coli* and purified *via* Ni-nitrolotriacetate (Ni-NTA) agarose. Lane 1: 10 μg crude extract of *E. coli* XL1-B (pMM40). Lane 2: 10 μg crude extract of *E. coli* XL1-B (pMMORF5H). Lanes 3 and 4: washing fractions. Lane 5: eluting fraction. Lane 6: 3 μg of Orf5×6His after dialysis. S: standard proteins. The gel was silver-stained.

difference may be due to the hydrophobic N-terminus of the protein.

DNA binding studies

Band shift assays were carried out for the upstream regions of the *sol* operon and the *adc* gene, the first genes to be induced upon onset of solventogenesis [34]. The region upstream of the *sol* operon was amplified in two overlapping fragments, both of them containing the distal of two postulated promoters [13]. The fragments were incubated with up to 4 μg Orf5x6His per nanogram probe DNA. With none of the fragments tested, neither solI/II nor adcI, was gel retardation observed. To rule out a possible effect of the His-tag fusion, the incubation was also carried out with various concentrations (up to 100 μg protein per nanogram probe DNA) of crude extract obtained from *E. coli* cells overexpressing Orf5 without His-tag. Again, no change of the probe mobility was observed. As a control, assays were carried out with the same amount of crude extract from cells harboring vector pMM40 without insert.

Because the *sol* operon is located on a plasmid [9] and recent studies have shown that the DNA topology has an effect on expression of the *sol* operon [40], it should be elucidated whether potential DNA binding of Orf5 was dependent on supercoiled DNA. Therefore, band shift assays were performed with plasmid pSP1, carrying the region upstream of the *adhE* gene of the *sol* operon. However, as with linear DNA fragments no gel retardation was observed with pSP1 when the plasmid was incubated with purified Orf5x6His, or crude extracts from *E. coli* harboring pMMORFN and pMMORFH.

Cellular localization of Orf5 in C. acetobutylicum

When the *orf5* gene was cloned and sequenced [12,13] (EMBL data base accession number X72831), it was noted that the N-terminus of the deduced gene product was very hydrophobic, due to the presence of 21 amino acids with lipophilic side chains within the first 30 residues. A signal peptide sequence was indicated by the presence of three characteristic domains, (i) a short N-terminal sequence of basic amino acids, (ii) a central domain of hydrophobic amino acids, and (iii) a region with small amino acids at the processing site [43]. These features were found in the *orf5* gene product [12] and the presence of a signal peptide was supported by the computer programm “signalP” (http://www.cbs.dtu.dk/services/signalP/). Since such a feature would be somewhat inconsistent to a proposed role as a cytoplasmic transcriptional repressor, the cellular localization of Orf5 was determined. For this purpose, shuttle plasmid pORF5H was constructed by cloning the gene of Orf5x6His with its native promoter into vector pIMP1. Due to the plasmid’s pIM13 origin of replication, a copy number of about 7–10 per cell was maintained [21], leading to a moderate overexpression. The introduced His-tag allowed a sensitive colorimetric detection of the expressed protein *via* Ni-NTA AP conjugate.

C. acetobutylicum harboring (i) pORF5H and (ii) pIMP1 (as a negative control) were grown in mineral medium until midlog phase. The cells were harvested and subsequently crude extracts and cell membranes were prepared. Proteins excreted into the medium supernatant were precipitated with TCA. The different fractions were separated by SDS-PAGE, followed by Western blot membrane transfer and His-tag detection (Figure 2). Strong signals indicating the presence of Orf5x6His were observed exclusively in the membrane fraction, while no Orf5x6His protein was detected in the cytoplasmic fraction. Similar results were obtained with membrane preparations of *E. coli* cells overexpressing Orf5x6His (data not shown). To a smaller extent, the protein was detected in the medium supernatant of *C. acetobutylicum*. Because silver-stained SDS-PAGE separations of the two fractions revealed an absence of the crude extract’s predominant protein bands in the supernatant fraction, it appears less likely that this amount stems from lysed cells. These findings are consistent with a function of the N-terminus as a signal peptide as well as a membrane anchor. In *C. acetobutylicum*, the protein appears to be located attached to the outer side of the cell membrane, part of it being released into the medium. The Orf5x6His protein detected in the supernatant appears to be slightly larger than that in the crude

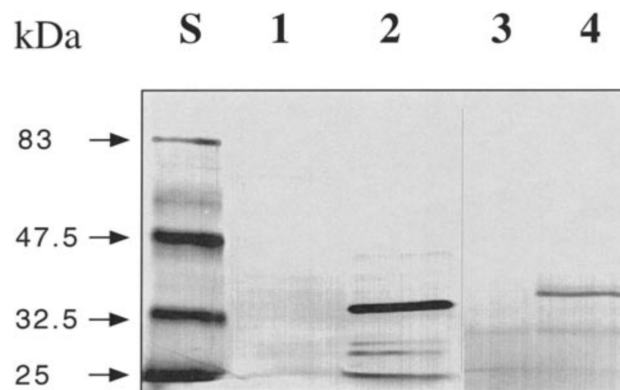

Figure 2 Localization of Orf5×6His stained by Ni-nitrilotriacetate (Ni-NTA) AP conjugate after SDS-PAGE (10% acrylamide) and Western blot membrane transfer. Lane 1: 50 μg crude extract of *C. acetobutylicum* (pIMP1). Lane 2: 50 μg membrane fraction of *C. acetobutylicum* (pORF5H). Lane 3: 50 μg cytoplasmic fraction of *C. acetobutylicum* (pORF5H); 10 μg of TCA-precipitated exoproteins of *C. acetobutylicum* (pORF5H). S: standard proteins.

extract. This is probably caused by a loss of the N-terminal signal sequence.

Effect of Orf5 overproduction in C. acetobutylicum

It was reported [30] that overexpression of the Orf5 homolog SolR in *C. acetobutylicum* ATCC824 leads to a complete decrease of solvent formation. Batch fermentation assays were performed to determine whether this effect of Orf5 overproduction on solvent production could be confirmed with *C. acetobutylicum* DSM 792. Therefore, *C. acetobutylicum* DSM 792 harboring pIMP1 (as a negative control), pORF5N, or pORF5H were grown for 72 h. Subsequently, the fermentation products were quantified (Table 2). While the acetone level decreased about 3.5-fold, no reduction of the butanol concentration was observed. Also, the level of ethanol was slightly increased (about 20%). Orf5 protein's overexpression led to a lower concentration of both acetic and butyric acids. With the exception of the lowered acetone production these findings are in sharp contrast to those reported by Nair *et al* [30] who used *C. acetobutylicum* ATCC 824 for their experiments. A renewed BLAST search with the Orf5 sequence revealed sequence homology to proteins with known or postulated functions as O-linked UDP-*N*-acetylglucosamine-peptide *N*-acetylglucosaminyltransferases (O-linked GlcNac transferases) of various eukaryotic organisms and archaea. At the amino acid level Orf5 is 23% identical (44% similarity) to an O-linked GlcNAc transferase of *Caenorhabditis elegans* [23], 24% identical (41% similarity) to the same enzyme of *Rattus norvegicus* [17], and 22% identical (43% similarity) to an O-linked GlcNAc transferase of *Homo sapiens* [23]. Among archaea highest scores were obtained with the putative O-linked GlcNAc transferase of *Methanobacterium thermoautotrophicum* [37] (25% identical, 48% similarity). These data indicate that Orf5 may be involved in protein glycosylation rather than being a transcriptional repressor. Consequently, overexpression of the protein might be expected to change the protein glycosylation pattern of the organism.

C. acetobutylicum harboring pIMP1 (as a negative control) or pORF5N were grown to midlog phase in mineral medium. Subsequently, crude extracts and membrane proteins were prepared, and the media supernatant fluids were TCA-precipitated. The fractions were separated by SDS-PAGE, followed by Western blot membrane transfer and glycoprotein detection. The result (Figure 3) revealed that indeed Orf5 overexpression is accompanied by a significant decrease of the amount of glycosylated proteins. Almost all of the signals obtained with the organism's exoenzymes disappeared. Even in the crude extract the most distinct signal at about 40 kDa was no longer detected.

Table 2 Product formation of *C. acetobutylicum* overexpressing Orf5

	Fermentation products (mM)				
	Solvents			Acids	
Plasmids	Ethanol	Butanol	Acetone	Acetate	Butyrate
pIMP1 (control)	31.5	61.0	9.3	48.5	60.5
pORF5N	39.8	70.0	2.6	30.3	43.1
pORF5H	37.8	70.9	2.8	27.7	44.0

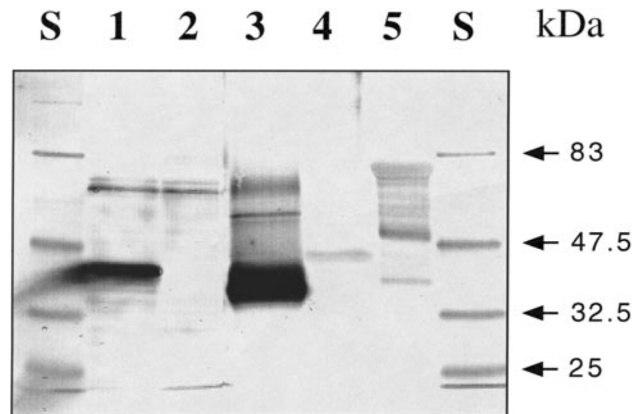

Figure 3 Effect of Orf5 overexpression of glycosylation: detection of glycoproteins using alkaline phosphatase-coupled digoxigenin antibodies after SDS-PAGE (10% acrylamide), Western blot membrane transfer, treatment with periodate and subsequent coupling with digoxigenin. Lane 1: 25 μg crude extract of *C. acetobutylicum* (pIMP1). Lane 2: 25 μg crude extract of *C. acetobutylicum* (pORF5N). Lane 3: 5 μg of TCA-precipitated exoproteins of *C. acetobutylicum* (pIMP1). Lane 4: 5 μg of TCA-precipitated exoproteins of *C. acetobutylicum* (pORF5N). Lane 5: positive control (transferrin). S: standard proteins.

Discussion

The suggestion that the gene product of *orf5* (*solR*) might represent a transcriptional repressor for the *sol* operon was based mainly on studies at the transcriptional level [30]. Therefore, the gene was subcloned and a His-tag-encoding sequence was added to allow easy purification and subsequent studies on the protein and its localization in the cell.

Orf5 did not exhibit *in vitro* DNA binding to the region upstream of the *adhE* gene, the first structural gene of the *sol* operon, nor to the promoter region upstream of the *adc* gene. Although even under disadvantageous conditions at least a weak protein–DNA interaction was expected, it cannot be completely ruled out that a cofactor essential for binding is missing under *in vitro* conditions. Overexpression of Orf5 in *C. acetobutylicum* DSM 792 did not lead to a loss of solvent production, as reported by Nair *et al* [30] for SolR in *C. acetobutylicum* ATCC 824. This was a surprising result, since both constructs for Orf5/SolR overexpression in *C. acetobutylicum* are based on the gene's native promoter and a vector with the same (pIM13) origin of replication was used, indicating a similar copy number and level of overexpression. It remains unknown whether this difference is due to potential strain differences.

Analysis of the deduced Orf5 amino acid sequence indicated the presence of a signal peptide [12] (EMBL data base accession number X72831), suggesting that this protein may be a membrane or exoprotein. Experimental data confirmed this assumption (Figure 2), since in both organisms, *C. acetobutylicum* and *E. coli*, the protein was located in the membrane. This does not completely rule out the possibility that Orf5 might act as a transcriptional repressor, e.g., in a way as has been demonstrated for the cholera toxin transcriptional activator ToxR [28]. However, the observation that a small amount of Orf5 can also be detected in the

supernatant medium suggests that the protein may be attached to the extracellular side of the membrane. This location would exclude interaction with the DNA and a function as a transcriptional repressor.

Taken together the data presented (no DNA binding of purified Orf5, no effect of its overproduction on solvent formation, probable localization at the outer side of the membrane) rule out a function as a repressor of solventogenesis. In a search for the physiological role of Orf5 BLAST databank searches revealed homology to described or postulated O-linked GlcNAc transferases of various eukaryote organisms. These proteins are responsible for the glycosylation of proteins *via* hydroxyl groups of serine or threonine residues. Protein modification *via* glycosylation is well studied for eukaryotic organisms but less so in bacteria. However, glycosylated proteins are also widespread among eubacteria and archaea [29]. It is postulated that, as in eukaryotic systems, glycosylation may influence protein conformation and stability, provide protection against proteolytic activity, and allow surface or intracellular recognition and cell adhesion. Up to now only a few clostridial proteins have been reported to be glycosylated: the S-layer protein of *Clostridium symbiosum* [27], a cellulosome subunit of *Clostridium thermocellum* [15], and the flagellin of *Clostridium tyrobutyricum* [2]. In *C. acetobutylicum* P262, autolysin, for which a strong association to sugars has been reported [42], has been described as a hypothetical glycoprotein, and recently glycosylation of the flagellin of *C. acetobutylicum* ATCC 824 has been shown [24]. This protein probably represents the dominant band of the crude extract as well as of the medium supernatant fraction (Figure 3). This study provides evidence that *C. acetobutylicum* possesses additional glycosylated proteins, especially exoproteins. So far, little is known on the location of glycosylation processes in prokaryotes, but there are several indications that this reaction occurs at the outer side of the cytoplasmic membrane [22,44]. The probable localization of Orf5 at the outer side of the membrane and the drastic effect of its overexpression on the pattern of glycosylated proteins support a role of Orf5 in this process. Further experiments will address the question of whether this function is rather deglycosylation of proteins and whether there are targets of glycosylation that are not separated by SDS-PAGE. Further support for a glycosylating or deglycosylating role of Orf5 comes from amino acid sequence analysis. In the helix–turn–helix motif postulated by Nair *et al* [30] a typical tetratrico peptide repeat motif [20] could be found (**Y**NN**LG**NN**Y**NNNNNNNN**S**NNN**Y**NN**A**NNNN**P**; consensus positions marked in bold, N stands for not conserved amino acid residues). Such repeats are present in a variety of proteins from bacteria and eukaryotes and seem to be responsible for protein–protein rather than DNA–protein interactions [20]. Thus, the postulated DNA interaction domain might be directly involved in protein target binding. In this respect, it should be noted that the O-linked GlcNAc transferase of *Caenorhabditis elegans* also contains tetratrico peptide repeat motifs [23].

In conclusion, the data presented provide evidence for an involvement of Orf5 in protein glycosylation/deglycosylation and not for a role of Orf5 as a transcriptional repressor of solventogenesis in *C. acetobutylicum*.

Acknowledgements

We are grateful to Brigitte Zickner for excellent technical assistance and to Ruth Schmitz-Streit for generously providing plasmid pMM40. This work was supported by a grant from the Deutsche Forschungsgemeinschaft.

References

1 Altschul SF, TL Madden, AA Schäffer, J Zhang, W Miller and DJ Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.

2 Bédouet L, F Arnold, G Robrau, P Batina, F Talbot and A Binet. 1998. Evidence for an heterogeneous glycosylation of the *Clostridium tyrobutyricum* ATCC 25755 flagellin. *Microbios* 94: 183–192.

3 Bennett GN and DJ Petersen. 1993. Cloning and expression of *Clostridium acetobutylicum* genes involved in solvent production. In: Sebald M (Ed), Genetics and Molecular Biology of Anaerobic Bacteria. Springer-Verlag, New York, pp. 317–343.

4 Bertram J and P Dürre. 1989. Conjugal transfer and expression of streptococcal transposons in *Clostridium acetobutylicum*. *Arch Microbiol* 151: 551–557.

5 Bertram J, A Kuhn and P Dürre. 1990. Tn*916*-induced mutants of *Clostridium acetobutylicum* defective in regulation of solvent formation. *Arch Microbiol* 153: 373–377.

6 Blum H, H Beier and HJ Gross. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8: 93–99.

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 Bullock WO, JM Fernandez and JM Short. 1987. XL1-B: a high efficiency plasmid transforming *recA E. coli* strain with β-galactosidase selection. *BioTechniques* 5: 376–378.

9 Cornillot E, RV Nair, RT Papoutsakis and P Soucaille. 1997. The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *J Bacteriol* 179: 5442–5447.

10 Dower WJ, JF Miller and CW Ragsdale. 1988. High efficiency transformation of *E. coli* by high voltage electrotransformation. *Nucleic Acids Res* 16: 6127–6145.

11 Dürre P, RJ Fischer, A Kuhn, K Lorenz, W Schreiber, B Stürzenhofecker, S Ullmann, K Winzer and U Sauer. 1995. Solventogenetic enzymes of *Clostridium acetobutylicum*: catalytic properties, genetic organization and transcriptional regulation. *FEMS Microbiol Rev* 17: 251–262.

12 Fischer RJ. 1993. Klonierung, Sequenzierung, molekulare Charakterisierung und Mutagenese von Genen der Lösungsmittelbildung aus *Clostridium acetobutylicum*. PhD thesis, Universität Göttingen, Germany.

13 Fischer RJ, J Helms and P Dürre. 1993. Cloning, sequencing and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. *J Bacteriol* 175: 6959–6969.

14 Gerischer U and P Dürre. 1990. Cloning, sequencing, and molecular analysis of the acetoacetate decarboxylase gene region from *Clostridium acetobutylicum*. *J Bacteriol* 172: 6907–6918.

15 Gerwig GJ, JP Kammerling, JFG Vliegenhardt, E Morag, R Lamed and EA Bayer. 1993. The nature of the carbohydrate-peptide linkage region in glycoproteins from the cellulosomes of *Clostridium thermocellum* and *Bacteroides cellulosolvens*. *J Biol Chem* 268: 26956–26960.

16 Kleiner D, P Wyatt and MJ Merrick. 1988. Construction of multicopy expression vectors for regulated overproduction of proteins in *Klebsiella pneumoniae* and other enteric bacteria. *J Gen Microbiol* 134: 1779–1784.

17 Kreppel LK, MA Blomberg and GW Hart. 1997. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J Biol Chem* 272: 9308–9315.

18 Kyhse-Andersen J. 1984. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J Biochem Biophys Methods* 10: 203–209.

19 Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.

20 Lamb JR, S Tugendreich and P Hieter. 1995. Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* 20: 257–259.
21 Lechner J and F Wieland. 1989. Structure and biosynthesis of prokaryotic glycoproteins. *Annu Rev Biochem* 58: 173–194.
22 Lee SY, LD Mermelstein and ET Papoutsakis. 1993. Determination of plasmid copy number and stability in *Clostridium acetobutylicum* ATCC 824. *FEMS Microbiol Lett* 108: 319–324.
23 Lubas WA, DW Frank, M Krause and JA Hanover. 1997. O-linked GlcNAc transferase is a conserved nucleocytoplasmatic protein containing tetratricopeptide repeats. *J Biol Chem* 272: 9316–9324.
24 Lyristis M, ZL Boynton, D Petersen, Z Kann, GN Bennett and FB Rudolph. 2000. Cloning, sequencing and characterization of the gene encoding flagellin, *flaC*, and the post-translational modification of flagellin, FlaC, from *Clostridium acetobutylicum* ATCC 824. *Anaerobe* 6: 60–79.
25 Mermelstein LD, NE Welker, GN Bennett and ET Papoutsakis. 1992. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Bio/Technology* 10: 190–195.
26 Mermelstein LD and ET Papoutsakis. 1993. *In vivo* methylation in *Escherichia coli* by the *Bacillus subtilis* phage Φ3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol* 59: 1077–1081.
27 Messner P, K Bock, R Christian, G Schulz and UB Sleytr. 1990. Characterization of the surface layer proteins of *Clostridium symbiosum* HB25. *J Bacteriol* 172: 2576–2583.
28 Miller VL, RK Taylor and JJ Mekalanos. 1987. Cholera toxin transcriptional activator ToxR is a transmembrane DNA binding protein. *Cell* 48: 271–279.
29 Moens S and J Vanderleyden. 1997. Glycoproteins in prokaryotes. *Arch Microbiol* 168: 169–175.
30 Nair RV, EM Green, DE Watson, GN Bennett and ET Papoutsakis. 1999. Regulation of the *sol* locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 by a putative transcriptional repressor. *J Bacteriol* 181: 319–330.
31 Nakotte S, S Schaffer, M Böhringer and P Dürre. 1998. Electroporation of, plasmid isolation from and plasmid conservation in *Clostridium acetobutylicum* DSM 792. *Appl Microbiol Biotechnol* 50: 564–567.
32 Oultram JD, M Loughlin, TJ Swinfield, JK Brehm, DE Thompson and NP Minton. 1988. Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. *FEMS Microbiol Lett* 56: 83–88.
33 Roos JW, K McLaughlin and ET Papoutsakis. 1985. The effect of pH on nitrogen supply, cell lysis, and solvent production in fermentations of *Clostridium acetobutylicum*. *Biotechnol Bioeng* 27: 681–694.
34 Sauer U and P Dürre. 1995. Differential induction of genes related to solvent formation during the shift from acidogenesis to solventogenesis in continuous culture of *Clostridium acetobutylicum*. *FEMS Microbiol Lett* 125: 115–120.
35 Shine J and L Dalgarno 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA* 7: 1342–1346.
36 Siebers A and K Altendorf. 1988. The K^+ translocating ATPase from *Escherichia coli*. *Eur J Biochem* 178: 131–140.
37 Smith DR, LA Doucette-Stamm, C Deloughery, HM Lee, J Dubois, T Aldredge, R Bashirzadeh, D Blakely, R Cook, K Glibert, D Harrison, L Hoang, P Keagle, W Lumm, B Poithier, D Qiu, R Spadafora, R Vicare, Y Wang, J Wierzbowski, R Gibson, N Jiwani, A Caruso, D Bush, H Safer, D Patwell, S Prabhakar, S McDougall, G Shimer, A Goyal, S Pietrovski, GM Church, CJ Daniels, JI Mao, P Rice, J Nolling and JN Reeve. (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* Δ H: functional analysis and comparative genomics. *J Bacteriol* 179: 7135–7155.
38 Suggs SV, T Hirose, T Miyake, EH Kawashima, MJ Johnson, KI Itakura and RB Wallace. 1981. Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences. In: Brown DD (Ed), Developmental Biology Using Purified Genes. Academic Press, New York, pp. 683–693.
39 Terracciano JS and ER Kashket. 1986. Intracellular conditions required for initiation of solvent production by *Clostridium acetobutylicum*. *Appl Environ Microbiol* 52: 86–91.
40 Ullmann S, A Kuhn and P Dürre. 1996. DNA topology and gene expression in *Clostridium acetobutylicum*: implications for the regulation of solventogenesis. *Biotechnol Lett* 18: 1413–1418.
41 Vieira J and J Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259–268.
42 Webster JR, SR Reid, DT Jones and DR Woods. 1981. Purification and characterization of an autolysin from *Clostridium acetobutylicum*. *Appl Environ Microbiol* 41: 371–374.
43 Wickner WT and HF Lodish. 1985. Multiple mechanisms of protein insertion into and across membranes. *Science* 230: 400–407.
44 Zhu BCR, RR Drake, H Schweingruber and RA Laine. 1995. Inhibition of glycosylation by amphomycin and sugar nucleotide analogs PP36 and PP55 indicates that *Haloferax volcanii* glycosylates both glycoproteins and glycolipids through lipid-linked sugar intermediates: evidence for three novel glycoproteins and a novel sulfated dihexosyl-archaeol glycolipid. *Arch Biochem Biophys* 319: 355–364.