Journal of Industrial Microbiology & Biotechnology (2002) 28, 118–126

www.nature.com/jim

Molecular characterization and utilization of the CAK1 filamentous viruslike particle derived from *Clostridium beijerinckii*

Y Li[a] and HP Blaschek

Biotechnology and Bioengineering Group, Department of Food Science and Human Nutrition, University of Illinois, 1207 West Gregory Drive, Urbana, IL 61801, USA

An examination of the replication origin and stability determinant associated with the CAK1 filamentous viruslike particle recovered from *Clostridium beijerinckii* NCIMB 6444 was carried out. Seven deletion derivatives, pCKE, pCEP1, pDT5, pCKP, pDTH102, pYL102E and pYL102, were constructed and transformed into *C. beijerinckii* NCIMB 8052. The successful transformation of pCKE, pDT5, pCKP, pDTH102, pYL102E and pYL102 into *C. beijerinckii* 8052, together with the corresponding recovery of single-stranded DNA from *Escherichia coli* indicated that the double- and single-stranded replication origins are located on a 0.4-kb CAK1 DNA fragment. Sequence analysis of the putative 0.4-kb replication origin region of CAK1 reveals a nick site containing 22 base pairs that has homology with plasmids pC194 and pUB110 and suggests the presence of a 2.0-kb DNA region involved in stability. The putative Rep protein of CAK1 contains three conserved motifs and three essential residues of the catalytic site in agreement with Rep proteins associated with the pC194 family. The utility of the developed CAK1-derived phagemid designated pYL102E was evaluated by using it to examine heterologous expression of: (1) the *manA* gene derived from *Thermoanaerobacterium polysaccharolyticum* in *E. coli* and *C. beijerinckii* NCIMB 8052 and (2) the *sol* operon derived from *Clostridium acetobutylicum* DSM 792 in *C. beijerinckii* SA-2.
Journal of Industrial Microbiology & Biotechnology (2002) **28,** 118–126 DOI: 10.1038/sj/jim/7000225

Keywords: *Clostridium beijerinckii*; shuttle vectors; CAK1; viruslike particle; DNA

Introduction

The genus *Clostridium* contains a diverse group of gram-positive, anaerobic, endospore-forming bacteria. In recent years, there has been considerable industrial and biotechnological interest in the clostridia due to their recognized solvent producing potential [18,21,35]. This has greatly motivated the development of genetic systems for this genus in order to understand the metabolic pathway for solvent formation during the fermentation process [3]. Gene transfer systems for the solventogenic clostridia have been developed using vectors containing replication origins from other gram-positive bacteria. However, these plasmids are unstable and may be lost following continuous subculture of the host microorganism [36]. The development of stable indigenous clostridial plasmid- or phage-based vector systems is a prerequisite for successful metabolic engineering and the expression of cloned genes in this group of microorganisms.

The search for indigenous extrachromosomal DNA, plasmids and phage, is an important step towards the development of gene transfer system in the clostridia. The CAKI filamentous viruslike particle was first isolated as a cryptic single-stranded (ss) plasmid pDM6 from *Clostridium beijerinckii* NCIMB 6444 (formerly *Clostridium acetobutylicum* NCIB 6444;), and later characterized as a filamentous viruslike particle [20]. The most advantageous aspects of a clostridial genetic system based on a filamentous viruslike particle, which is similar to the M13 filamentous phage series utilized in *Escherichia coli*, include: (1) the production of ssDNA and dsDNA during different stages of the phage life cycle, (2) the phages do not lyse the host during the propagation of progeny phages and (3) the recombinant phage system can be used as a versatile cloning vector for the study of site-directed mutagenesis and sequencing of clostridial DNA.

A study of the CAK1 filamentous viruslike particle cannot only provide a native shuttle vector system for the genetic study of clostridia, but it may also lead to an understanding of the replication mechanism of the first filamentous viruslike particle recovered from a gram-positive bacterium. In an earlier study we focused on the construction and characterization of a bacteriophage–plasmid hybrid (phagemid), pCAK1, which contained the replicative form (RF) of the CAK1 filamentous viruslike particle [21]. The pCAK1 phagemid was constructed by ligating the RF of CAK1 to the pAK102 plasmid containing ampicillin (Amp^R) and erythromycin (Erm^R) resistance genes and a ColE1 *E. coli* replication origin. The pCAK1 phagemid vector was present as dsDNA in both *C. beijerinckii* NCIMB and *E. coli* strains. pCAK1 was transformed into *C. beijerinckii* NCIMB 8052 and *Clostridium perfringens* strain 13 at an average efficiency of 10^3 transformants/μg DNA. The successful replication of pCAK1 in clostridial strains indicated the presence of a functional complementary strand replication origin. Moreover, a ssDNA protein complex was recovered from *E. coli* DH11S supernatant suggesting that the viral strand replication origin is functional in pCAK1. These studies demonstrated the potential of the CAK1 viruslike particle for developing a unique genetic system for the industrially significant solventogenic clostridia.

The objective of this study was to identify and characterize the replication origin and stability determinant associated with the CAK1 filamentous viruslike particle and to examine the utility of

Correspondence: Dr HP Blaschek, Biotechnology and Bioengineering Group, Department of Food Science and Human Nutrition, University of Illinois, 1207 West Gregory Drive, Urbana, IL 61801, USA
[a]Current address: Alexon-Trend Inc., 160 South Wolfe Road, Sunnyvale, CA 94086, USA
Received 12 September 2000; accepted 23 October 2001

the CAK1-derived phagemid pYL102E for heterologous gene expression.

Materials and methods

Bacterial strains, phage, plasmid and phagemids

The bacterial strains, phage, plasmid and phagemids used in this study are listed in Table 1.

Growth conditions and maintenance

All *E. coli* strains were grown at 37°C in Luria Bertani (LB) medium [27]. All clostridial strains were grown at 35°C in TGY medium under anaerobic conditions in a Coy anaerobic chamber [6]. For solid agar plates, 1.5% bactoagar was added to the liquid medium. Ampicillin was added at a concentration of 100 μg/ml for selection of *E. coli* transformants, and erythromycin was added at 25 μg/ml for selection of clostridial transformants. Spore cultures of *C. beijerinckii* NCIMB 8052 were prepared in cooked meat medium (CMM) and stored at 4°C. Spore suspensions were heat shocked at 80°C for 10 min before inoculation. For long-term storage, *E. coli* and clostridial strains were stored at −80°C in 20% glycerol (vol/vol).

Deletion of pCKA1

pCAK1 deletion derivatives were generated by removing DNA fragments following restriction enzyme digestion of CAK1. Restriction enzymes, Klenow fragment, calf intestine alkaline phosphatase (CIAP) and T4 DNA ligase used in this study were purchased from Gibco BRL (Bethesda, MD), New England Biolabs (Beverly, MA) or Promega (Madison, WI) and used according to the manufacturer's recommendations. All chemicals were obtained from either Sigma (St. Louis, MO) or Fisher Scientific (Hanover Park, IL). DNA fragments were recovered from agarose gel using the BIO-101 Gene Clean Kit (Bio 101, Vista, CA) following the manufacturer's instructions. Removal of some restriction enzymes and modifying enzymes was carried out using Amicon (Beverly, MA) Micropure-EZ enzyme removing columns.

DNA isolation

Isolation of plasmid DNA from *E. coli* strains was carried out using the DNA mini prep kit from Promega. Isolation of plasmid DNA from clostridial strains was carried out using a modified alkaline lysis protocol described previously [19].

The isolation of ssDNA from phage particles produced from *E. coli* JM109 was carried out according to the protocols provided by Promega. ssDNA was isolated from the *E. coli* cell pellet and culture supernatant, respectively, and examined by agarose gel electrophoresis and Southern hybridization.

DNA transformation

Transformation of *E. coli* strains was carried out using competent cells as per manufacturer's instructions (Promega or Gibco BRL). Transformation of *C. beijerinckii* NCIMB 8052 was carried out using a modification of the electroporation procedures described for the clostridia [4].

Southern hybridization

The ssDNA samples were transferred from an agarose gel to a positively charged nylon membrane (Schleicher & Schuell, Keene, NH) without denaturation. After the membrane was

Table 1 Bacterial strains, phage, plasmids and phagemids used in this study

Strains, phages, plasmid/phagemids	Relevant characteristics	Source or reference
Strains		
E. coli		
DH5α	F$^-$080d*lacZ*ΔM15 Δ(l*acZYA*-argF)U169 *end*A1 *rec*A1 *hsd*R17(r_k-m_k+)*deo*R *thi*-1 *sup*E44 λ^- *gyr*A96 *rel*A1	BRL
JM109	*end*A1 *rec*A1 *gyr*A96 *thi hsd*R17(r_k-m_k+) *rel*A1 *sup*E44 Δ(*lac*-*pro*AB) [F′, *tra*D36, *pro*AB,*lac*^{Iq}Z ΔM15]	Promega
C. beijerinckii		
NCIMB 8052	Solvent-producing, wild type	Laboratory stock
SA-2	Butanol-tolerant, mutant	2
Helper phage		
R408	Mutant f1 filamentous phage	Promega
Plasmid/phagemids		
pAK102	5 kb, Amp^R, Erm^R	20
pMTL500E	6.4 kb, Amp^R, Erm^R	35
pAK1.1	7.8 kb, Amp^R, *sol* operon	10
pGEM-5Zf(+)	3 kb, Amp^R, f1 ori	Promega
pCAK1	11.6 kb, pAK102 with 6.6-kb CAK1	20
pCKE	8.9 kb, pAK102 with CAK1 deleted 2.7-kb *Eco*RV fragment from pCAK1	This study
pCEP1	7.7 kb, pAK102 with CAK1 deleted 1.2-kb *Pst*I fragment from pCKE	This study
pDT5	6.6 kb, pAK102 with CAK1 deleted 5.5-kb *Eco*Rv–*Pst*I fragment from pCAK1	This study
pCKP	8.6 kb, pAK102 with CAK1 deleted 3.0-kb *Pst*I fragment from pCAK1	This study
pDTH102	8.0 kb, pAK102 with CAK1 deleted 0.6-kb *Eco*RV fragment from pCKP	This study
pYL102E	7.4 kb, pAK102 with CAK1 deleted 0.7-kb *Eco*RI fragment from pDTH102	This study
pYL102	7 kb, pAK102 with CAK1 deleted 0.4-kb *Eco*RI–*Eco*RV fragment from pYL102E	This study
pMTL-*manA*	9.8 kb, pMTL500E with *Eco*RI-*manA* fragment from pBK-CMV	This study
pYL-*manA*	10.8 kb, pYL102E with *Eco*RI-*manA* fragment from pBK-CMV	This study
pYL-sol1	12.3 kb, pYL102E with *Ava*I–*Bgl*I-*sol* operon fragment from pAK1.1	This study
pYL-sol2	12.3 kb, pYL102E with *Bgl*I–*Ava*I-*sol* operon fragment from pAK1.1	This study

heated at 80°C to immobilize the nucleic acids, the membrane was hybridized with a biotinylated probe generated from the ampicillin gene in pAK102. A 1-kb amp-probe was obtained by *Rca*I digestion of pAK102. The probe was labeled with biotin-14-dCTP using a BioPrime DNA labeling Kit (BRL). Hybridization with the amp-probe was used in order to detect ssDNA that had undergone replication in *E. coli* JM109. The prehybridization and hybridization steps were performed according to the membrane manufacturer's instructions and a nonisotopic detection procedure was employed according to Ambion's instructions (Ambion, Austin, TX). The membrane was exposed to an X-ray film to visualize DNA bands.

Stability assay

The method described by Manen *et al* [26] was used. The *C. beijerinckii* NCIMB 8052 transformants were grown overnight in 10-ml TGY medium containing 25 μg erythromycin/ml. The culture was diluted 10^4-fold every 12 h with fresh TGY medium without added antibiotic. The procedure was carried out over 100 generations. At each dilution point, 100 μl of diluted culture was plated onto TGY and TGY plates containing 25 μg erythromycin/ml.

DNA sequencing and analysis

DNA sequencing of the two discontinuous fragments of CAK1 DNA in pYL102E was carried out at the UIUC Biotechnology Center sequencing facility. An Applied Biosystems (Model 373A) (Foster City, CA) automatic sequencing unit was employed. M13 universal primer and reverse primer were used for the first ssDNA sequencing of CAK1 DNA in pYL102E. The dsDNA sequencing was accomplished by primer walking and the data were analyzed using the BLAST algorithm [1], GeneWorks 2.5 (Intelligenetics, Mountain View, CA), and Sequencher 3.0 (Gene Codes Cooperation, Ann Arbor, MI).

Nucleotide sequence and accession number

The CAK1 DNA sequence reported here has been submitted to the GenBank database and has been assigned the accession number AF135472.

DNA construction

The *manA* gene from *Thermoanaerobacterium polysaccharolyticum* was previously cloned into phagemid pBK-CMV (Strategene, Menasha, WI) at the *Eco*RI site and transformed into *E. coli* XLOLR [5]. A 3.4-kb *Eco*RI fragment containing *manA* gene was recovered from pBK-CMV and cloned into pMTL500E and pYL102E at the *Eco*RI site (Fig. 1). Plasmid pAK1.1 (7.8 kb) containing the *sol* operon [11] derived from *C. acetobutylicum* DSM 792 kindly provided by P. Durre (Universität ULM, ULM, Germany) was digested using the restriction enzymes *Ava*I and *Bgl*I [10]. Following Klenow treatment, the DNA fragments were separated by agarose gel electrophoresis. The 4.9-kb blunt-ended *Ava*I–*Bgl*I fragment containing *sol* operon was eluted and concentrated. Phagemid pYL102E was digested with *Eco*RI and blunt ended by Klenow. The blunt-ended *sol* operon fragment and pYL102E DNA were ligated with T4 DNA ligase at 14°C overnight.

Enzyme activity assay

Transformant *E. coli* and *C. beijerinckii* strains exhibiting endoglucanase activity were analyzed by plate and colorimetric

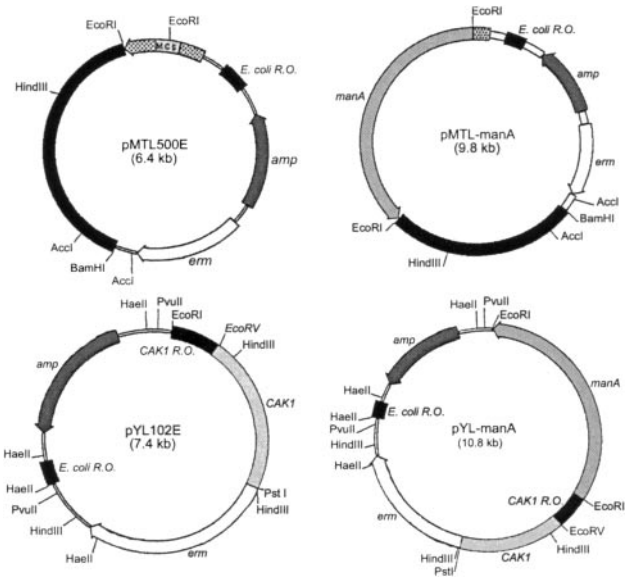

Figure 1 Schematic representations of the vectors, pMTL500E, pYL102E and their corresponding constructs, pMTL-*manA* and pYL-*manA*, respectively.

assays. The ManA enzyme activity of *E. coli* strains was detected using a plate assay [5,22]. Endoglucanase activity of the cell pellet and supernatant from *Clostridium* transformants was determined colorimetrically by measuring the amount of reducing sugars released from CMC [5]. Single colonies of *C. beijerinckii* NCIMB 8052 and its transformants containing pYL102E and pYL-*manA* were picked from nonselective and selective TGY plates and inoculated into 10 ml of TGY medium with or without 50 μg/ml erythromycin. The culture was incubated at 35°C in an anaerobic chamber until the optical density at 600 nm reached 1.2. The culture was centrifuged at 6,000×*g* for 10 min at 4°C. The supernatant was collected and concentrated using an Amicon concentrator (S1). The cell pellet was suspended in 5 ml 0.1 M phosphate buffer (pH 7.0) and passed through a French pressure cell at 15,000 psi to disrupt cells. The cell debris was centrifuged at 17,000×*g* for 10 min at 4°C. The supernatant (S2) from the cell pellet was collected. Both S1 and S2 were used for colorimetric assay.

CoA transferase activity was measured by following the disappearance of acetoacetyl-CoA at 310 nm [6]. One unit of enzyme activity is defined as the disappearance of 1 μmol acetoacetyl CoA per minute.

Protein concentration was determined using the dye-binding assay (Bio-Rad Laboratories, Hercules, CA) according to the manufacturer's instructions and using bovine serum albumin as a standard.

Batch fermentation and solvent analysis

Fermentation experiments were conducted in 50 ml P2 medium containing 6% glucose. A 5% inoculum was used. Solvent in culture supernatants was determined using a 6890 gas chromatograph (Hewlett-Packard, Avondale, PA) equipped with a flame ionization detector (FID) and a capillary column (30 m by 0.53 mm [inner diameter]) packed with FFAP (polyethylene glycol-TPA modified, Hewlett-Packard). The oven temperature was programmed from 60°C to 160°C. The injector and detector temperatures were set at 260°C. A 1-μl sample was used.

Results and discussion

Construction of pCAK1 deletion derivatives

A linear map of the pCAK1phagemid and its deletion derivatives is shown in Figure 2. pCKE was obtained by *Eco*RV digestion of pCAK1, pCEP1 was obtained by *Pst*I digestion of pCKE, pDT5 was obtained by *Eco*RV–*Pst*I digestion of pCEP1, pCKP was obtained by *Pst*I digestion of pCAK1, pDTH102 was obtained by *Eco*RV digestion of pCKP and pYL102E was obtained by *Eco*RI digestion of pDTH102. These six deletion derivatives of CAK1 DNA all contain a common 0.4-kb region. Phagemid pYL102 obtained following *Eco*RI–*Eco*RV digestion of pYL102E does not contain the 0.4-kb CAK1 region, which is present in the other deletion derivatives. Phagemid pCAK1 and its deletion derivatives were transformed into *E. coli* DH5α or JM109 and analyzed by agarose gel electrophoresis of *Hin*dIII digests (data not shown).

Identification of the double-stranded and single-stranded DNA replication origins

In order to examine the functionality of the replication origin associated with pCAK1, phagemid pCAK1 and the corresponding deletion derivatives isolated from *E. coli* transformants were transformed into *C. beijerinckii* NCIMB 8052. The transformation efficiency (T. E.) defined as transformants per microgram DNA is shown in Figure 2. No transformants were recovered in *C. beijerinckii* NCIMB 8052 when plasmid pAK102 was used as control. Clostridial transformants were screened for resistance to erythromycin at 25 μg/ml. With the exception of pYL102, the CAK1 deletion derivatives were able to transform *C. beijerinckii* NCIMB 8052, suggesting that a functional complimentary strand (CS) replication origin is located on the 0.4-kb CAK1 fragment. A transformation efficiency of 10^3 to 10^4 transformants/μg DNA for CAK1-based phagemids is higher than that observed when plasmid pMTL500E (10^1 to 10^3 transformants/μg DNA) was used to transform *C. beijerinckii* NCIMB 8052 [36].

In order to examine the functionality of the CAK1 viral strand (VS) synthesis replication origin associated with pCAK1 and its deletion derivatives, ssDNA secretion was examined in *E. coli* JM109. Phage DNA was isolated from both the *E. coli* cell pellet and supernatant in the presence of helper phage R408. Since the genes encoding proteins for ssDNA production in pCAK1 deletion derivatives could be disrupted as a consequence of the deletion process, R408 helper phage was used to produce ssDNA from pCAK1 and its deletion derivatives containing a functional ssDNA synthesis origin. R408 helper phage contains all gene products required for packaging of ssDNA but it is unable to efficiently package itself. Samples were transferred from an agarose gel to a positively charged nylon membrane in the absence of denaturation. The membrane was probed with an ampicillin resistance gene (amp-probe) and Southern blotting carried out using a nonisotopic chemiluminescent detection system (Figure 3). Phagemid pGEM5Zf(+) was used as a positive control, while plasmid pAK102 served as a negative control. Hybridization of ssDNA from CAK1-derived deletion derivatives pCAK1, pDTH102 and pYL102E was observed both intracellularly as well as in the supernatant derived from *E. coli* JM109 (Figure 3B, lanes 4, 5, 6 and 11, 12, 13 and 3B, lanes 4, 5, 6 and 11, 12, 13, respectively).

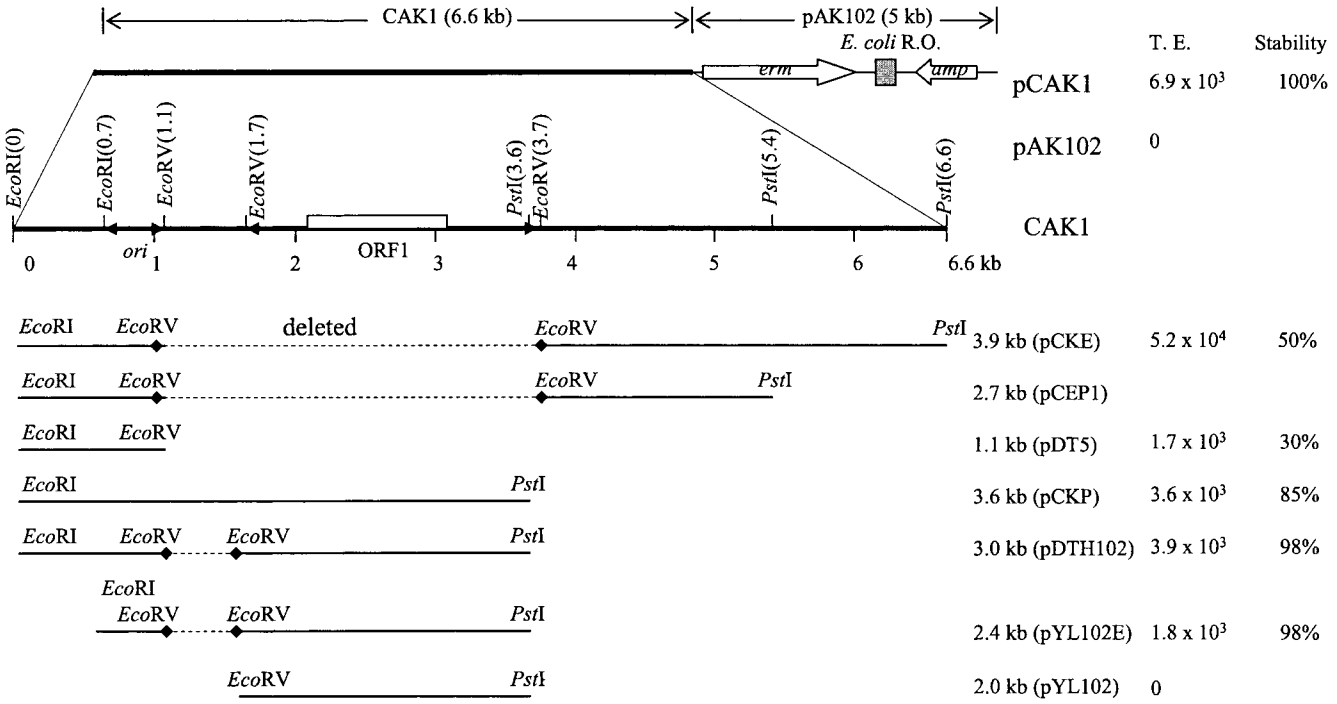

Figure 2 Deletion map of CAK1 DNA. The dashed lines with solid diamond in the endpoints indicate the deleted region of CAK1. The length of remaining CAK1 associated with each deletion derivative is indicated to the left of the phagemid name. The sequenced portions of CAK1 are indicated by solid lines with headed arrows. The short sequenced region is the putative *ori* of CAK1. The putative Rep protein, ORF1, is shown by an open box. Restriction sites are shown with their locations (kb) indicated in parentheses. The pAK102 plasmid is intact in all the deletion derivatives (not shown on each deletion map). *E. coli* R.O. is indicated by filled square in pAK102. The transformation efficiency (T. E.) and the stability results of pCAK1 and its deletion derivatives are shown on the right.

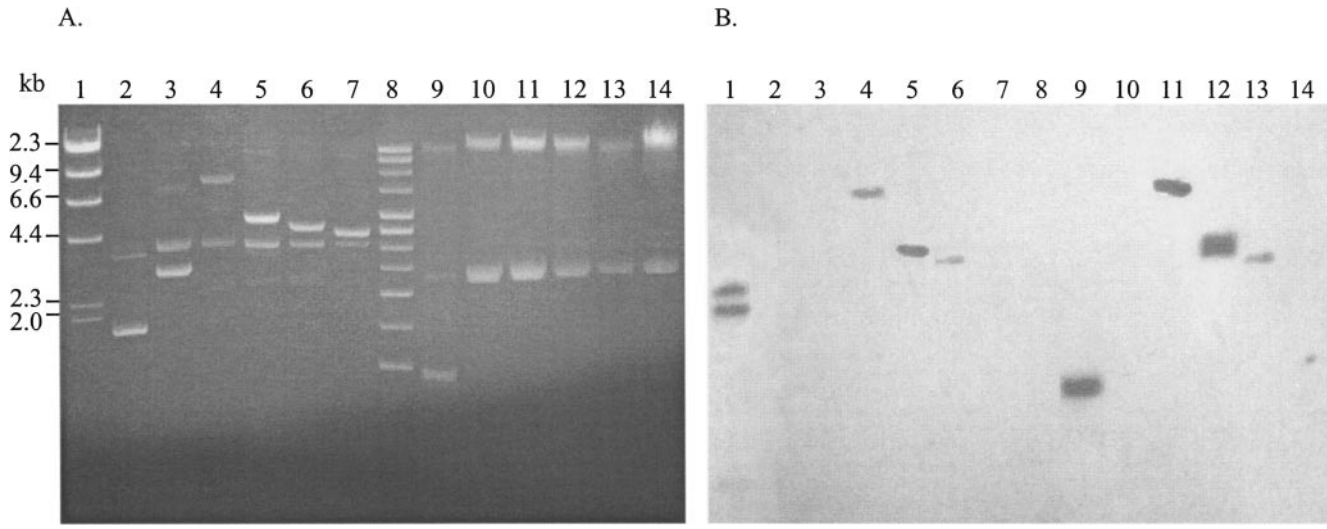

Figure 3 Agarose gel (A) and Southern hybridization (B) analysis of CAK1-derived phagemids. The DNA samples were isolated from *E. coli* JM109 cell pellet and supernatant in the presence of helper phage R408. Lane 1 contains biotinylated λDNA/*Hin*dIII fragments and lane 8 supercoiled DNA marker. Lanes 2–7 contain the DNA samples pGEM5Zf(+), pAK102, pCAK1, pDTH102, pYL102E and pYL102 isolated from cell pellet, respectively. Lanes 9–14 contain the corresponding DNA samples isolated from supernatant. ssDNA from both the *E. coli* cell pellet and supernatant containing pCAK1, pDTH102 and pYL102E was detected by hybridization with the amp-probe.

No hybridization was observed for the CAK1 deletion derivative, pYL102 (Figure 3B, lanes 7 and 14). The presence of intracellular and extracellular ssDNA in *E. coli* JM109 indicates that a functional VS synthesis replication origin exists within the 0.4-kb CAK1 DNA region.

Stability of pCAK1 derivatives

Stability testing of the pCAK1 phagemid derivatives was carried out in *C. beijerinckii* NCIMB 8052 cultures in the absence of antibiotic. The percentage of cells retaining erythromycin resistance at 50 generations is shown in Figure 2. The results indicated that pCAK1, pCKP, pDTH102 and pYL102E are more stable than pCKE and pDT5 in the absence of antibiotic pressure in *C. beijerinckii* NCIMB 8052. A 2.0-kb *Eco*RV–*Pst*I fragment of CAK1 (Figure 2) which is present in pCAK1, pCKP, pDTH102 and pYL102E, but is absent from pCKE and pDT5 may be responsible for phagemid stability.

Sequencing analysis of two CAK1 DNA segments within pYL102E

Two regions, representing the putative 0.4-kb replication origin and the 2.0-kb stability determinant derived from pYL102E were sequenced. The DNA sequence was submitted to the GenBank (AF135472) and analyzed using GeneWorks, BLAST and Sequencher computer programs. DNA alignment of the 0.4-kb CAK1 region with known *ori* sequences in gram-positive and gram-negative bacteria did not reveal any homology. Also, the sequences for *palA*, *palB* and *palT*-like sites that are typically associated with single-stranded origins [26,31] were not found within the 0.4-kb CAK1 sequence. However, a 22-bp region at positions 81–102 shows ca. 64% identity with pC194 and pUB110 after DNA alignment with the *nic* sequences of known rolling circle (RC) replicons [8,9,13,14,33] (Table 2). Two loci, the *bind* and *nic* regions, have been defined at the *dso* for RC plasmids. The *bind* locus is required for the Rep protein to bind to plasmid DNA, whereas the *nic* locus is where the Rep protein introduces the initial nick. Typically, the *nic* regions are highly conserved among replicons of the same family, whereas the *bind* regions are divergent [8]. The 22-bp region may serve as a nick site for initiation of CAK1 replication.

The 2.0-kb CAK1 DNA region contains six open reading frames termed ORF1 to ORF6. All ORFs were examined by comparison with other known sequences contained in the National Center for Biotechnology Information (NCBI) database. The large ORF (ORF1) (342 amino acids) demonstrated approximately 40% identity to the replication (Rep) proteins of *Bacillus pumilus* plasmids pSH1452 [34] (Figure 4) and pPL10 [25], *Bacillus popilliae* plasmid pBP614 [24], *B. subtilis* plasmids pTA series [28] and *Streptococcus thermophilus* plasmids pST1 [17] and pCI65st [32]. These plasmids replicate by an RC mechanism. The three conserved motifs, motif 1: FLTLTIRN; motif 2: HPHFCIMM; motif 3: STKYTVKDNA and three essential residues of the catalytical sites, E (glutamate), E and Y (Tyrosine) were identified in ORF1. This finding is agreement with Rep proteins encoded by RC plasmids, especially for the pC194 family (Table 3; Refs. [15,30]). A coherent arrangement of the three conserved motifs N-1-2-3-C in the initiator proteins for RC DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria has been observed [15]. Motif 2 consists of the amino acid sequence HisHydrHisHydrHydrHydr and the two conserved His

Table 2 Comparison of nucleotide sequences of nick sites associated with the putative 0.4-kb CAK1 replication origin and plasmids pC194 and pUB110

Plasmid	Sequence of nick sites[a]
0.4-kb CAK1	A**AT**GA**ATCA**CT**AT**C**AAGAAA**TT
pC194	T**AT**–T**ATCA**AG**AT**–**AAGAAA**GA
pUB110	**AT**GT**ATCA**AG**AT**–**AAGAAA**GA

[a]Bold and underlined letters indicate conserved bases; dashes indicate no bases.

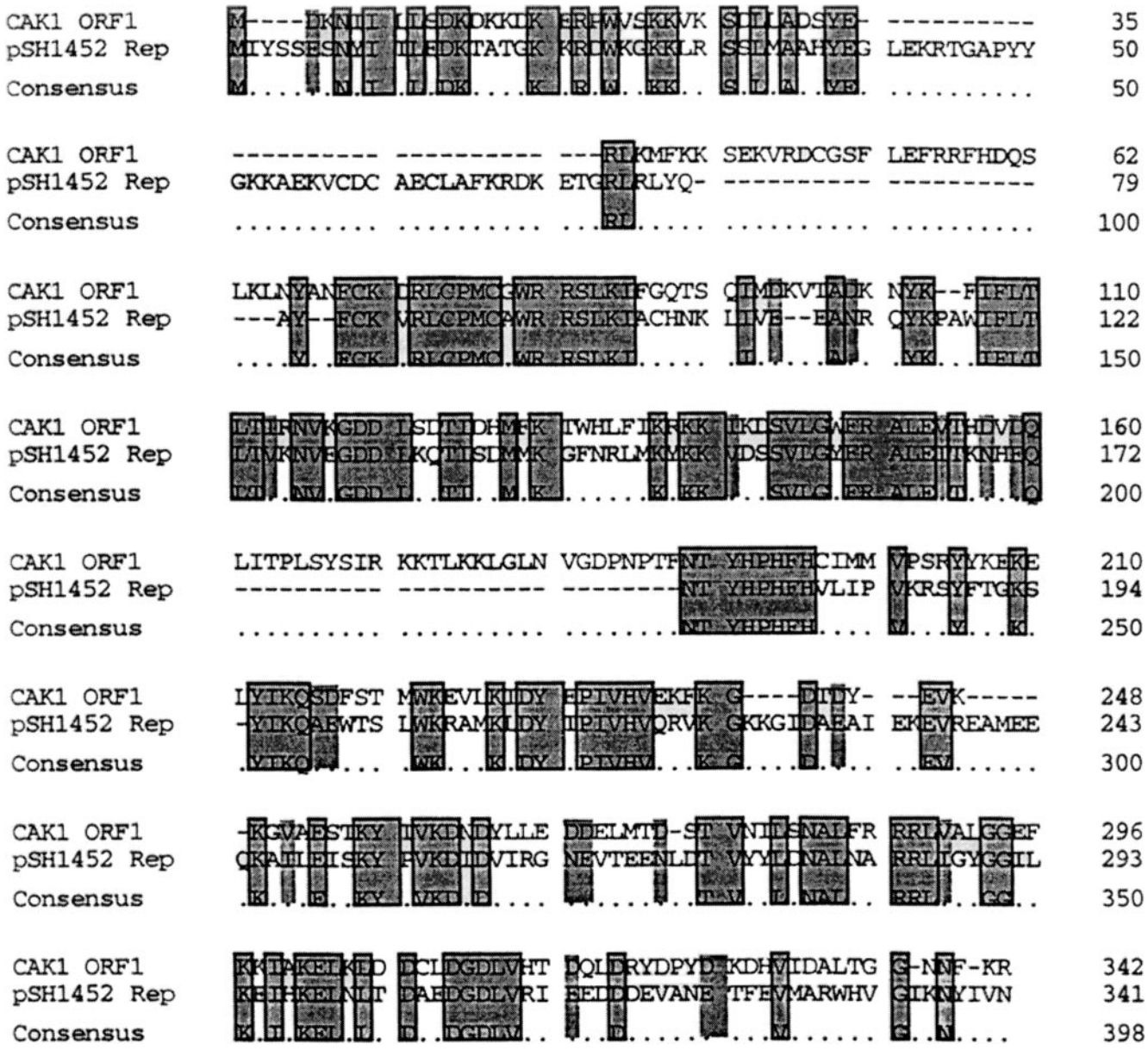

Figure 4 Comparison of the deduced amino acid sequence of ORF1 with the Rep protein of *B. pumilus* plasmid pSH1452. Identical amino acids are boxed and shaded. Conserved residues are shaded. Three conserved motifs in ORF1 are located at positions 108–115 for motif 1, 192–200 for motif 2 and 254–263 for motif 3. The three essential amino acids in ORF1 are located at positions 153, 253 and 257, respectively.

residues in this motif may be involved in metal ion coordination required for the activity of Rep. The C-terminal motif encompassing the tyrosine residue(s) may be involved in the formation of a covalent link with nicked DNA at the initiation step of RC replication. One active site containing Tyr residue(s) is common for Rep proteins. However, Noirot-Gros *et al* [30] found two different catalytic centers, corresponding to a tyrosine and a glutamate in RepA protein in pC194. No significant homologies were found for the other five small ORFs (ORFs 2–6) following BLAST search. The functions of five small ORFs are unknown currently. As per the above sequence analysis and ssDNA data, the possibility exists that CAK1 DNA replication involves a pC194-related *rep* gene which encodes an essential initiator protein and initiates RC replication by introducing a site-specific nick in the *dso* of CAK1. A plasmid-encoded Rep protein, which can maintain plasmid stability by affecting plasmid copy number has been reported [16]. This may explain why the 2.0-kb CAK1 fragment is involved in CAK1-based phagemid stability.

Gene expression in E. coli and C. beijerinckii

The *manA* gene from *T. polysaccharolyticum* was cloned into pYL102E and pMTL500E and the constructs were designated pYL-*manA*, and pMTL-*manA*, respectively (Figure 1). Plasmid pMTL500E, which is based on pAMβ1 originally derived from *Streptococcus faecalis*, is an *E. coli–Clostridium* shuttle vector [35]. The constructs, 10.8-kb pYL-*manA* and 9.8-kb pMTL-*manA* were identified by agarose gel electrophoresis following

Table 3 Conserved motifs in ORF1 of 2.0-kb CAK1 DNA and in initiator proteins of the pC194 family[a]

Plasmid/phage	Motif 1	Motif 2	Motif 3
pC194	FLTLTTPN	NPHFHVLIA	MAKYSGKDSD
CAK1	FLTLTIRN	HPHFHCIMM	STKYTVKDND
pUB110	FLTLTVKN	NQHMHVLVC	TAKYPVKDTD
pBAA1	FLTLTVRN	HPHFHVLIP	ISKYPVKDTD
pFTB14	FLTLTVRN	HPHFHVLLP	ISKYPVKDTD
pLP1	FLTLTVKN	NQHLHVLLF	TAKYEVKSAD
pLAB1000	FLTLTAEN	HQHMHVLLF	TAKYQVKSKD
pKYM	FLTLTVRN	HPHFHCLLM	TLKYSVKPED
Consensus[b]	FLTLTvxN	xxHuHvLUx	xxKYxxKxxD

[a]References: [8,15].
[b]Uppercase letters denote conserved residues; lowercase letters denote residues conserved in at least half of the members; v denotes hydrophobic residues; x denotes any residues.

digestion with appropriate restriction enzymes (data not shown). The ManA of *T. polysaccharolyticum* is a modular protein and contains both mannanase and endoglucanase activities [5]. Examination of the expression of the *manA* gene in *E. coli* was carried out using a plate assay. The transformant strains exhibiting endoglucanase activity were evaluated by examining the presence of zones of clearing around isolated colonies (Figure 5). Among all ampicillin-resistant transformants of *E. coli* DH5α tested, only those which harbored vectors carrying the *manA* gene demonstrated enzyme activity (Figure 5), lower portion of plate); *E. coli* transformants containing pYL102E and pMTL500E did not demonstrate clearing (Figure 5), upper portion of plate). These results indicated that the *manA* gene from *T. polysaccharolyticum* could be expressed in *E. coli*.

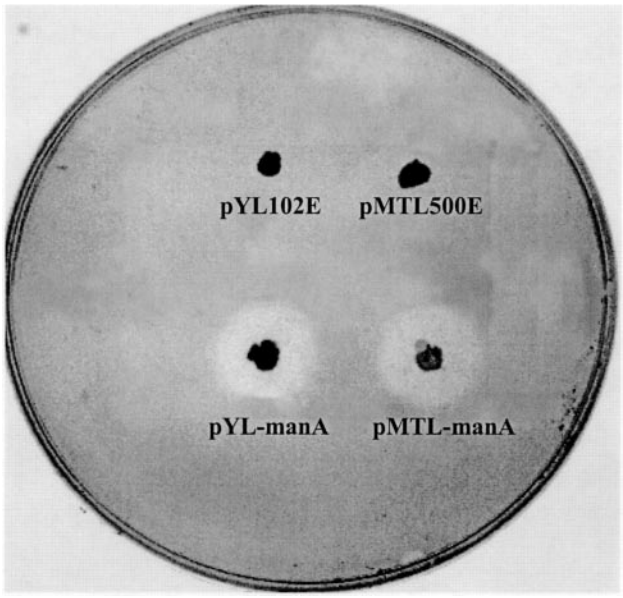

Figure 5 Endoglucanase activities of *E. coli* transformants containing the *manA* gene from *T. polysaccharolyticum* on a LB/Amp plate overlaid with 0.5% carboxymethylcellulose. Single colony isolates represent *E. coli* DH5α transformants containing pYL102E, pMTL500E (upper portion) and the corresponding constructs pYL-*manA* and pMTL-*manA* (lower portion), respectively.

The constructs pMTL-*manA* and pYL-*manA* together with pMTL500E and pYL102E were subsequently transformed into *C. beijerinckii* NCIMB 8052 by electroporation. Transformants of *C. beijerinckii* containing pMTL500E pYL102E and pYL-*manA* were obtained on TGY/Erm^R plates at a transformation efficiency of 102–103 transformants/μg DNA, however transformants of *C. beijerinckii* containing pMTL-*manA* were not observed. Because the plate assay is not sensitive enough for examining the activity of clostridial transformants, the expression of the *manA* gene in *C. beijerinckii* NCIMB 8052 was examined by colorimetric assay employing cell pellets and culture supernatants. The enzyme activities associated with the cell pellet extracts and concentrated supernatants were evaluated. Only the concentrated supernatant from *C. beijerinckii* NCIMB 8052 containing pYL-*manA* was found to have detectable enzyme activity (0.97 ± 0.10 μmol/min/mg of protein). These results suggest that pYL102E can be used for heterologous expression of ManA in *C. beijerinckii* NCIMB 8052.

The *sol* operon containing *ctfA*, *ctfB*, *adhE* genes derived from *C. acetobutylicum* DSM 792 is responsible for butanol formation [11]. Two constructs with the *sol* operon inserted in opposite orientations in pYl102E were designated pYL-*sol*1 and pYL-*sol*2. Both constructs were transformed into *C. beijerinckii* SA-2 (formerly *C. acetobutylicum* SA-2; [12]) by electroporation. *C. beijerinckii* SA-2 is a solvent-tolerant, degenerated strain that produces only trace amounts of butanol and acetone [2]. It was chosen in order to examine the expression of the sol operon in a low solvent-producing host when using the pYL102E phagemid vector. Solvent production associated with the culture supernatant of *C. beijerinckii* SA-2 and the corresponding transformants containing pYL102E, pYL-sol1 and pYL-*sol*2, which are designated SA-YL, SA-sol1 and SA-sol2, are shown in Table 4. As anticipated, *C. beijerinckii* SA-2 produced a negligible amount of solvent. The *C. beijerinckii* SA-YL recombinant (which contains only phagemid pYL102E) demonstrated similar results. However, more butanol was produced by *C. beijerinckii* SA-sol2 than by *C. beijerinckii* SA-2 or by the SA-sol1 strain. These results indicated that the *sol* operon is expressed in *C. beijerinckii* SA-2 and that butanol production is affected by the direction of transcription. There have been reports that acetone and butanol production is synergistic and efficient acid reassimilation will benefit solvent production [6,7,29]. Elevated acetone production was observed for *C. beijerinckii* SA-sol1 and SA-sol2. However, the SA-sol2 strain did not show higher acetone production than SA-sol1.

Table 4 Solvent production by *C. beijerinckii* SA-2 and associated transformant strains in 50 ml P2 medium containing 6% glucose after 48-h fermentation

Strain	Acetone (g/l)	Butanol (g/l)
C. beijerinckii		
SA-2	0.0 ± 0.0	0.2 ± 0.0
SA-YL[a]	0.0 ± 0.0	0.1 ± 0.1
SA-sol1[b]	0.4 ± 0.1	0.6 ± 0.3
SA-sol2[c]	0.4 ± 0.2	1.6 ± 0.1

The samples were collected at 48 h and the data represent the results of triplicate experiments; mean$\pm$standard deviation.
[a]Transformant strains of *C. beijerinckii* SA-2 containing pYL102E.
[b]Transformant strains of *C. beijerinckii* SA-2 containing pYL-*sol*1.
[c]Transformant strains of *C. beijerinckii* SA-2 containing pYL-*sol*2.

The reason for this is unknown. Because acetoacetyl-CoA:acetate/butyrate:CoA transferase (CoAT) is a key enzyme involved in reassimilation of acid to solvent, CoA transferase activity associated with a 24-h culture supernatant of *C. beijerinckii* SA-2 and the *C. beijerinckii* SA-sol2 transformant was measured. *C. beijerinckii* SA-2 demonstrated no detectable CoA transferase activity. However, the CoA transferase activity of *C. beijerinckii* SA-sol2 was 0.24±0.05 μmol/min/mg of protein.

The successful heterologous expression of the *manA* gene from *T. polysaccharolyticum* in *C. beijerinckii* NCIMB 8052 as well as the *sol* operon from *C. acetobutylicum* DSM 792 when using pYL102E suggests that (1) phagemid pYL102E can be used as an *E. coli–C. beijerinckii* shuttle vector and (2) it can be used to produce strains that have a better ability to utilize complex substrates as well as to alter product formation by the solventogenic clostridia. There are several goals that have yet to be achieved via metabolic engineering of the solventogenic clostridia in order to make the acetone–butanol–ethanol (ABE) fermentation commercially attractive again. These include alteration of product formation, increase in substrate range and production of strains with greater butanol tolerance [2,22,29]. To fulfill these tasks, effective shuttle vectors together with an efficient transformation system are required. pYL102E is the first shuttle vector that harbors an indigenous clostridial DNA replicon to be used for gene expression in *C. beijerinckii*. The expression of *manA* gene in *C. beijerinckii* along with the expression of endoglucanase from *C. cellulovorans* in *C. beijerinckii* [22] will ultimately improve this solventogenic organism in order to utilize more complex substrates, such as cellulosic biopolymers, which represent a very cheap carbon source for the ABE fermentation. The expression of *sol* operon from *C. acetobutylicum* in *C. beijerinckii* SA-2 provides useful information regarding the alteration of final product formation in these organisms.

Acknowledgements

We sincerely thank Douglas Tan for the preparation and analysis of pCKP, pCEP1, pDT5 and pDTH102 deletion derivatives and Dr. Peter Durre for providing plasmid pAK1.1. This work was partially supported by USDA competitive grant AG98-35504-6181 and Hatch grant 50-304.

References

1 Altschul SF, W Gish, W Miller, EW Myers and DJ Lipman. 1990. Basic local alignment tool. *J Mol Biol* 215: 403–410.
2 Baer SH, HP Blaschek and TL Smith. 1987. Effect of butanol challenge and temperature on lipid composition and membrane fluidity of butanol-tolerant *Clostridium acetobutylicum*. *Appl Environ Microbiol* 53: 2854–2861.
3 Blaschek HP and BA White. 1995. Genetic system development in the clostridia. *FEMS Microbiol Rev* 17: 349–356.
4 Blaschek HP, J Formanek and CK Chen. 1999. *Clostridium* — biotechnology and food technology. In: Eynard N and J Teissie, (eds), Electrotransformation of bacteria — Springer lab manual. pp. 47–55. Springer, Heidelberg, Germany.
5 Cann IKO, S Kocherginskaya, MR King, BY White and RI Mackie. 1999. Molecular cloning, sequencing, and expression of a novel multidomain mannanase gene from *Thermoanaerobacterium polysaccharolyticum*. *J Bacteriol* 181: 1643–1651.
6 Chen C-K and HP Blaschek. 1999. Effect of acetate on molecular and physiological aspects of *Clostridium beijerinckii* NCIMB 8052 solvent production and strain degeneration. *Appl Environ Microbiol* 65: 499–505.
7 Chen C-K and HP Blaschek. 1999. Examination of physiological and molecular factors involved in enhanced solvent production by *Clostridium beijerinckii* BA101. *Appl Environ Microbiol* 65: 2269–2271.
8 Del Solar G, R Giraldo, MJ Ruiz-Echevarria, M Espinosa and R Diaz-Orejas. 1998. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* 62: 434–464.
9 Dotto GP, K Horiuchi and ND Zinder. 1984. The functional origin of bacteriophage f1 DNA replication: its signals and domains. *J Mol Biol* 172: 507–521.
10 Durre P. 1998. Personal communication.
11 Fischer RJ, J Helms and P Durre. 1993. Cloning, sequencing, and molecular analysis of the sol operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. *J Bacteriol* 175: 6959–6969.
12 Formanek JA. 1998. PhD thesis. University of Illinois at Urbana–Champaign, IL.
13 Greenstein D and K Horiuchi. 1987. Interaction between the replication origin and the initiator protein of the filamentous phase f1: binding occurs in two steps. *J Mol Biol* 197: 157–174.
14 Gruss A and SD Ehrlich. 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol Rev* 53: 231–241.
15 Ilyina TV and EV Koonin. 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes, and archaebacteria. *Nucleic Acids Res* 20: 3279–3285.
16 Janniere L, A Gruss and SD Ehrlich. 1993. *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. In: Soneshein AL, R Losick and JA Hoch, (eds), Plasmids. pp. 625–644. American Society for Microbiology, Washington, DC.
17 Janzen T, J Kleinschmidt, H Neve and A Geis. 1992. Sequencing and characterization of pST1, a cryptic plasmid from *Streptococcus thermophilus*. *FEMS Microbiol Lett* 95: 175–180.
18 Johnson JL and J-S Chen. 1995. Taxonomic relationships among strains of *Clostridium acetobutylicum* and other phenotypically similar organisms. *FEMS Microbiol Rev* 17: 233–240.
19 Kim AY and HP Blaschek. 1989. Construction of an *E. coli–Clostridium perfringens* shuttle vectors and plasmid transformation of *Clostridium perfringens*. *Appl Environ Microbiol* 173: 360–365.
20 Kim AY and HP Blaschek. 1991. Isolation and characterization of a filamentous viruslike particle from *Clostridium acetobutylicum* NCIB 6444. *J Bacteriol* 173: 530–535.
21 Kim AY and HP Blaschek. 1993. Construction and characterization of a phage–plasmid hybrid (phagemid), pCAK1, containing the replicative form of viruslike particle CAK1 isolated from *Clostridium acetobutylicum* NCIB 6444. *J Bacteriol* 175: 3838–3843.
22 Kim AY, GT Attwood, SM Holt, BA White and HP Blaschek. 1994. Heterologous expression of endo-β-1,4-D-glucanase from *Clostridium cellulovorans* in *Clostridium acetobutylicum* ATCC 824 following transformation of the *engB* gene. *Appl Environ Microbiol* 60: 337–340.
23 Kim AY, AA Vertes and HP Blaschek. 1990. Isolation of a single-stranded plasmid from *Clostridium acetobutylicum* NCIB 6444. *Appl Environ Microbiol* 56: 1725–1728.
24 Longly M and RTM Poulter. 1996. Direct submission to NCBI database. Accession number 1688311 (U78608).
25 Lovett PS, EJ Duvall and KM Keggins. 1976. *Bacillus pumilus* plasmid pPL10: properties and insertion into *Bacillus subtilis* 168 by transformation. *J Bacteriol* 127: 817–828.
26 Manen D, T Goebel and L Caro. 1990. The *par* region of pSC101 affects plasmid copy number as well as stability. *Mol Microbiol* 4: 1839–1846.
27 Maniatis T, F Fritsch and J Sambrook. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
28 Meijer WJJ, G Venema and S Bron. 1995. Characterization of single strand origins of cryptic rolling-circle plasmids from *Bacillus subtilis*. *Nucleic Acids Res* 23: 612–619.
29 Mermelstein LD, ET Papoutsakis, DJ Petersen and GN Bennett. 1993. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by enhancement of acetone formation

enzyme activities using a synthetic acetone operon. *Biotech Bioeng* 42: 1053–1060.

30 Noirot-Gros MF, V Bidnenko and SD Ehrlich. 1994. Active site of the replication protein of the rolling circle plasmid pC194. *EMBO J* 13: 4412–4420.

31 Novick RP. 1989. Staphylococcal plasmids and their replication. *Annu Rev Microbiol* 43: 537–565.

32 O'Sullivan TF, D van Sinderen and GF Fitzgerald. 1998. Direct submission to NCBI database. Accession number 2921235 (AF027167).

33 Rasched I and E Oberer. 1986. Ff coliphages: structural and functional relationships. *Microbiol Rev* 50: 401–427.

34 Thomas CM and S Hasnain. 1996. Direct submission to NCBI database. Accession number 2444010 (U53767).

35 Young M, NP Minton and WL Staudenbauer. 1989. Recent advances in the genetics of the clostridia. *FEMS Microbiol Rev* 63: 301–326.

36 Young M, WL Staudenbauer and NP Minton. 1989. Genetics of *Clostridium*. In: Minton, NP and DJ Clark, (eds), Clostridia. pp. 63–103. Plenum, New York.