Expression systems for Industrial Gram-positive bacteria with low guanine and cytosine content
Willem M de Vos*†, Michiel Kleerebezem* and Oscar P Kuipers*

Recent years have seen an increase in the development of gene expression systems for industrial Gram-positive bacteria with low guanine and cytosine content that belong to the genera Bacillus, Clostridium, Lactococcus, Lactobacillus, Staphylococcus and Streptococcus. In particular, considerable advances have been made in the construction of inducible gene expression systems based on the capacity of these bacteria to utilize specific sugars or to secrete autoinducing peptides that are involved in quorum sensing. These controlled expression systems allow for present and future exploitation of these bacteria as cell factories in medical, agricultural, and food biotechnology.

Addresses
*Department of Biophysical Chemistry, NIZO, PO Box 20, Kernhemseweg 2, 6710 BA Ede, The Netherlands
†Laboratory for Microbiology, Department of Biomolecular Sciences, Wageningen Agricultural University, Wageningen, The Netherlands; e-mail: willem.devos@algemeen.micr.wau.nl
Correspondence: Willem M de Vos

Current Opinion in Biotechnology 1997, 8:547–553
http://biomednet.com/locartef/0958166900800547
© Current Biology Ltd ISSN 0958-1669

Abbreviations
ABE acetone-butanol-ethanol
CmpA catabolite control protein
GC guanine and cytosine
LAB lactic acid bacteria
LGB low GC Gram-positive bacteria
RCR rolling circle replication
RR response regulator

Introduction
Gram-positive bacteria are well known for their important contributions to medical, agricultural and food biotechnology and as production organisms for industrial enzymes. The latter capacity is largely based on the common architecture of their cell envelope, which allows for direct secretion of proteins into the fermentation medium, which in Bacilli has reached a high level of efficiency [1]. Based on the guanine and cytosine (GC) content of their DNA, Gram-positive bacteria can be divided into two major groups, the low and high GC groups. This dichotomy is confirmed by phylogenetic analysis based on 16S rRNA sequence comparisons that emphasize the separated evolutionary positions of the two groups [2]. As a consequence, the industrially relevant but highly diverse actinomycetes, which belong to the high GC group, are not discussed here, although Streptomyces spp. have been developed into efficient hosts for heterologous protein production, in spite of their GC content of more than 70% [3]. This review will focus on the low GC Gram-positive bacteria (LGB) that have established biotechnological applications and belong to the genera Bacillus, Clostridium, Lactococcus, Lactobacillus, Staphylococcus and Streptococcus. Phylogenetic analysis indicates that these genera, in particular those belonging to the lactic acid bacteria (LAB) that include Lactococcus, Lactobacillus and Streptococcus spp., form a homogeneous group and are relatively closely related in evolutionary terms (Figure 1). This is in line with the widely spread presence, or functionality upon introduction, of highly conserved plasmids and transposons in LGB that are excellent tools for the development of genetic systems. Furthermore, recent findings, discussed below, indicate that common mechanisms of initiation and density-dependent control of gene expression are operating in the LGB. These common genetic characteristics are highly relevant for the further development of LGB to become expression hosts. As a consequence, these genetic characteristics are briefly summarized below before recent progress on controlled and efficient gene expression in the individual genera is discussed.

Figure 1

Unrooted phylogenetic tree based on 16S rRNA sequences showing the relative positions of the genera Bacillus, Clostridium, Staphylococcus, Streptococcus, Lactobacillus, and Lactococcus.

Common characteristics of LGB: mechanisms and applications
Conserved plasmids and transposons
Plasmids a with highly conserved unit of replication, or 'replicon', are found in all the genera of LGB discussed in the introduction and replicate by two well-studied mechanisms, theta replication or rolling circle replication (RCR) [4,5]. Large plasmids usually replicate via theta replication and belong either to the pAMβ1 family of broad host-range conjugative plasmids or to the pC1305
family, the members of which have a limited host-range. A variety of useful vectors have been based on the replicon of pAMβ1, which is functional in all studied genera of LGB [5,6]. These vectors have been equipped with *Escherichia coli* replicons to allow efficient shuttling between Gram-positive and Gram-negative bacteria [7]. The detailed analysis of its replicon allowed for the development of plasmid vectors for LGB which can switch from high to low copy number, and hence are useful for modulating the expression level of cloned genes [8*]. This was achieved by introducing an inactivating linker cassette in the represor gene, copF, that regulates the repE gene coding for the rate-limiting replication enzyme. Removal of the cassette by a simple digestion with the infrequently-cutting restriction enzyme KpnI reconstitutes a functional copF gene, resulting in a change from high to low copy number.

Small plasmids of several kilobases in general use RCR, are ubiquitous in all studied genera, and also belong to a limited number of families [4,5,9]. The replicons of these plasmids have been developed in a series of highly useful cloning vectors that replicate in a broad range of genera, notably those based on the related *Lactococcus lactis* plasmids pWV01/pSH71, which also replicate in *E. coli* [5]. A thermosensitive derivative of pWV01, termed pG+host, has been generated and has shown its potential as a delivery system in LGB [10]. By introduction of the lactococcal insertion sequence ISS1 element in pG+host, a random transposition system has been constructed that is functional in a broad range of genera [11*] and may be used to generate mutants or introduce foreign genes. This system may also be used as an alternative to the naturally occurring transposons that are found in LGB [12]. These include the transposons related to Tn917 that have been used for insertional mutagenesis in Bacillus and several LAB [12,13]. Recently, a Tn917 derivative was optimized for promoter screening in *Lactococcus* [14]. In addition, conjugative transposons have been found in several genera [15]; however, most use has been made of the well-characterized Tn916, which is able to generate insertions in all LGB [12,15]. The application of the described tools into suitable cloning, integration and expression vectors has been the subject of several recent reviews, with specific attention being paid to LAB [5,8*,9,12].

**Initiation of transcription and translation**

The composition of core RNA polymerases in LGB resembles that of *E. coli* [16]. In addition, while the number of identified sigma factors may vary between the different genera, the main sigma factor, sigma A, is also conserved [5,16]. This is in line with the observation that most sigma A-dependent promoters in LGB contain the canonical −35 and −10 sequences [5,9,16]. Recently, a vital −16 region was identified that affects promoter efficiency in *Bacillus subtilis* and *E. coli* [17]. In addition, sequence analysis of *B. subtilis* sigma A-dependent promoters indicated that there are extended contacts in this region and also upstream as far as −70 [18]. This confirms previous analysis of LGB other than *B. subtilis* that show conservation of a thymine–guanine dinucleotide in the −16 region and an excess of adenine and thymine sequences upstream of the −35 consensus [5,9]. Furthermore, it may explain why some *E. coli* promoters are not efficiently used in LGB; however, within LGB the efficiency of transcription initiation may also vary in a species-dependent way, as was observed when transcriptional gusA fusions were studied in different LAB, suggesting that optimization for each expression host is advisable [19].

Translation initiation in LGB shows remarkable features that discriminate it from that in other Gram-positive and Gram-negative bacteria, confirming its phylogenetic position [20]. Only LGB have been found to lack the ribosomal protein S1, which is coupled to a high translational selectivity not observed with other bacteria, including the high GC Gram-positive ones. This is reflected in a more extensive complementarity between the Shine-Dalgarno sequences and the 3' end of the 16S rRNA than found in other bacteria [5,9,20]. The requirement for a stringent Shine-Dalgarno region for high translational efficiency has been observed in many cases and explains previous observations that *E. coli* genes, which do not generally have Shine-Dalgarno sequences with such complementarity, are not well expressed in several LGB [21].

**Catabolite repression**

While catabolite repression in *E. coli* and other Gram-negative bacteria has been well characterized and involves the catabolite repressor protein and its interaction with cyclic AMP, these compounds have not been found to a similar extent in LGB [22]. In contrast, it has been shown that catabolite repression in LGB is mediated via a catabolite control protein (CcpA) and the ε-acting sequence cre [22,23]. CcpA interacts with the serine-phosphorylated form of the heat stable protein (HPr), a phospho-carrier involved in the phosphotransferase system mediating the active uptake of sugars [23,24]. The serine-phosphorylation of HPr, which seems exclusive to LGB, is part of the signalling pathway, since it is catalyzed by a serine kinase when the intracellular fructose-1,6-diphosphate concentration is high. Immunological studies have shown the presence of a CcpA-like protein in all genera of LGB discussed here [25]. Recently, this has been supported by genetic data on the sequence and functionality of the *ccpA* genes [25,26]. While some *ccpA* genes seem constitutively expressed, the *ccpA* gene in *Staphylococcus xylosus* is transcribed from two promoters, one of which contains a cre sequence and is subject to CcpA-mediated autoregulation [27]. Further insight in the global control system exerted by CcpA will allow for optimal exploitation of sugar inducible expression systems that are often used to control gene expression in LGB. Well-established
examples are the xylose inducible systems studied in *Bacillus*, *Staphylococcus*, and *Lactobacillus*, which are subject to specific repression by a homologous repressor XylR and subject to negative regulation via CcpA [28–30]. Finally, analysis of the CcpA mediated control may contribute to eliminating catabolite repression that is often observed during industrial mixed substrate fermentations.

**Quorum sensing**

Cell density-dependent gene expression, also known as quorum sensing, appears to be widely spread in microbes. In Gram-negative bacteria, this form of cell-cell communication has been studied extensively and involves diffusible N-acetyl homoserine lactones as autoinducing signals, which interact with LuxR-like transcriptional activators [31]. Recent studies have indicated that LGB also communicate with each other but use peptide pheromones as signal molecules [32*] (Figure 2). These peptide pheromones are made as precursors that are post-translationally processed and exported via dedicated ATP binding cassette transporters. In addition, the extracellular peptide pheromones function as input signals for a two component signal transduction system consisting of a membrane located sensor histidine protein kinase and a cytoplasmatic response regulator (RR), which activates transcription. It is assumed that binding of the peptide pheromone to histidine protein kinase results in a phosphorylation cascade resulting in formation of a phosphorylated RR, which subsequently binds and transcriptionally activates one or more dedicated promoters, including the promoter of an operon of the peptide pheromone precursor gene and the *hpk* and *rr* genes, encoding HPR and RR, respectively (Figure 2).

Examples of this novel quorum sensing involving autoinduction of peptide pheromones have been found in all members of LGB described here and include the development of competence in *Bacillus* [33,34] and *Streptococcus* spp. [35,36], the development of virulence in *Staphylococcus* spp. [37,38], and the production of antimicrobial peptides by *Lactobacillus* spp [39,40] or *Lactococcus* [41]. The latter system concerns the production of the lantibiotic nisin, an extensively post-translationally modified peptide that is used in the food industry because of its activity against a wide range of LGB, including pathogens and food spoilage bacteria [42]. The nisin autoinduction system was the first example of quorum sensing involving antimicrobial peptide production. It has now been developed into an efficient and industrially applicable controlled gene expression system by introduction of a gene of interest (*gene X*) under control of a nisin-inducible promoter either in a nisin producing strain (Figure 2a) or in a strain harboring only the genes for the signal transduction pathway (Figure 2b) [43**]. In the first case, expression of *gene X* is dependent on the cell density and occurs at the end of the growth (Figure 2a). In the second configuration, *gene X* expression can be induced at will by the addition of the inducing peptide pheromone (Figure 2b). Examples of these systems in *L. lactis* have been described [43**]. Similar approaches could be applied to the quorum sensing modules that are presently known in LGB (see above) or may be found in future studies.
Controlled and efficient expression systems in LGB: new developments and applications

**Bacilli**

One of the first controlled expression systems in *B. subtilis* was based on the *E. coli lacI* repressor gene, equipped with an appropriate promoter and ribosome binding site in conjunction with the inducible *spac-I* promoter, consisting of the promoter of phage SPO-1 coupled to the *lac operon* [44]. This system, however, could only be used for specific, small-scale applications because of the need to add the gratuitous inducer isopropyl-β-D-thiogalactoside. A variety of other systems have been reported in the meantime but it was only recently that another successful *E. coli* system, the T7 system, was implemented in *B. subtilis* [45].

The use of the T7 system was achieved in *B. subtilis* by inserting the T7 RNA polymerase gene *rpoT7* into the chromosome under control of a xylose-inducible promoter (see above). In conjunction with a plasmid carrying the T7 promoter driving expression of the gene of interest, this heterologous expression system resulted in high induction and production levels of both intracellular and extracellular enzymes upon addition of xylose and may be suitable for larger scale applications. The xylose-inducible system was also used recently — without the T7 amplification step — in a set of cassettes that could be integrated in the *amyE* locus of the *B. subtilis* chromosome [46]. High induction levels were also obtained here upon xylose addition and only repression by glucose, not by other rapidly metabolizable compounds, was observed. This versatile system is not only suitable for fundamental studies aimed at comparing expression efficiencies but may also have potential for further applications, since it may be used in other, industrially more relevant *Bacilli* and eliminate the need for unstable plasmid vectors and their selection. These may include hosts such as *Bacillus licheniformis* and *Bacillus amyloliqufaciens* [1]. In addition, in the past year significant progress has been reported with *Bacillus brevis*, which produces large amounts of surface layer proteins. Based on the appropriate signals of these surface layer proteins, expression-secretion vectors were constructed that allowed for the extracellular production of large amounts (more than 1 g/l) of functional human epidermal growth factor [47*]. Similar, although less efficient, production was obtained in another study with animal erythropoetin and the extracellular domain of its receptor [48]. Finally, hyperproduction of the *Clostridium* alpha-toxin was observed in a similar approach in *B. brevis* which showed ten times more production than *B. subtilis* [49].

**Clostridia**

In spite of their considerable medical and economic importance, genetic systems for *Clostridium* spp. are not very well developed, hampering their use as expression hosts [50]. Considerable attention has been given to the medical aspects of *Clostridia*, which have been reviewed recently together with the present state of its molecular biology [51]. Next to their role in pathogenesis and as producers of exoenzymes, some *Clostridia*, notably *Clostridium acetobutylicum* are well known for the production of acetone–butanol–ethanol (ABE). In an attempt to increase the ABE fermentation, the well-known *C. acetobutylicum* strain ATCC 824 was transformed with plasmids carrying homologous sporulation genes. Transformants were reported not to survive, however, and so an integration approach was followed that resulted in mutants with scrambled chromosomal inserts, some of which showed increased stability and efficiency of ABE fermentation [52]. In addition to its established application areas, a novel potential application for *Clostridia* has been recently described that is based on its potential to target hypoxic regions of tumour cells and specifically germinate at these locations [53*]. This would create a specific delivery system for toxic compounds that could inactivate the cancer cells. This interesting possibility was tested with *Clostridium beijerinckii* expressing the *E. coli* gene for cytosine deaminase, which may convert the nontoxic prodrug 5-fluorocytosine into an active chemotherapeutic compound. *In vitro* conversion of the prodrug was achieved with a clostridial supernatant [53*], illustrating the feasibility of this approach.

**Staphylococci**

Most studies on staphylococcal expression systems have concentrated on *Staphylococcus carnosus* or *Staphylococcus xylosus*, that are non-pathogenic and have a long tradition of safe use as starters for meat fermentations. An improvement of the xylose induction system (see above) from *S. xylosus* has been developed [29]. Using a previously constructed lipase reporter system it was shown that the new expression vector pTX15 carrying the xylose inducible *xylA* promoter showed better induction and lower background activity than a previously described plasmid [54]. The utility of the vector pTX15 was demonstrated in *S. carnosus* with the *Staphylococcus epidermidis* *epiB* gene, which is involved in the production of lanthionine rings in the lantibiotic epidermin. Finally, the T7 system also has been applied to *S. carnosus*, again using the xylose-inducible *xylA* promoter [55]. *S. carnosus* was also further developed for the display of recombinant proteins at its cell surface, exploiting the anchor domain of the staphylococcal protein A [56]. This approach is aimed at developing non-pathogenic *Staphylococci* with a long tradition of safe use as novel vaccine carriers. Similarly, strains of both *S. carnosus* and *S. xylosus* appeared to be excellent hosts for the surface production of functional single-chain Fv antibodies [57]. It is speculated that this approach may lead to the development of these LGB as whole-cell diagnostic devices or as alternatives to filamentous phages for surface display of single chain Fv libraries.

**Streptococci**

Interest in the genetics of a number of different streptococcal species is in many cases supported by their role
as pathogens or their contribution to food-fermentations. Although a variety of plasmid vectors is available for *Streptococci*, integrative gene expression systems were evaluated in two different *Streptococcus* spp., representing each of these application groups. A promoterless chloramphenicol acetyltransferase gene was inserted in between the chromosomally located *lacS* and *lacZ* genes of the yoghurt bacterium *Streptococcus thermophilus*, the expression of which is induced during growth on lactose [58]. Although *LacZ* activity was reduced to about 10%, the chloramphenicol acetyltransferase activities paralleled those of β-galactosidase, indicating that both genes were expressed and regulated by the same control elements. In addition, a host-vector system for heterologous gene expression was developed for the pathogenic *Streptococcus gordonii* (formerly *Streptococcus sanguis*), which included the streptococcal M6 protein to allow for surface expression [59].

**Lactobacilli**

The development of convenient cloning and controlled expression systems in *Lactobacilli* has been recently reviewed [9,60,61]. Increasing attention is being given to *Lactobacillus plantarum*, which has a reasonable capacity to secrete proteins. This was exploited by the secretion of α-amylase and levanase using homologous expression secretion signals both located on the chromosome and on plasmids [62]. In addition, using a similar approach as described above for *B. brevis*, high level production of heterologous proteins was obtained in both *Lactobacillus* and *Lactococcus* with a secretion system based on the signals of the surface layer protein discovered in *Lactobacillus brevis* [63].

**Lactococci**

By far the best studied LAB are the *Lactococci*, specifically *L. lactis*, an important starter in cheese production. In the past few years several expression systems, both constitutive and regulated, have been developed [5,6,61]. Recent advances include the isolation and characterization of more than 10 different promoter elements from the chromosome of *L. lactis*, exhibiting a 70-fold range of activity [64]. The strongest promoter, P15, shows unusual spacing between 35 and −10 regions, which could indicate a regulation mechanism acting at the transcriptional level. In addition, several reports have appeared describing the use of the T7 expression system in *L. lactis*, which was based on the inducible *lac* promoter [65]. High, inducible extracellular production of interleukin-2 was achieved in *Lactococcus* [66]. The potential of *L. lactis* as a vaccine carrier was further developed and a protective immune response was obtained with strains secreting the tetanus toxin fragment C [67]. Recently, several different systems for inducible gene expression were developed that are briefly discussed below.

Based on the complete nucleotide sequence of the temperate *L. lactis* bacteriophage φ11, repressor-operator sequences were identified and exploited by developing a system in which gene expression could be activated by the addition of mitomycin C, resulting in an induction factor of about 70 [68]. Recently, this system was further improved by development of a thermosensitive repressor, which facilitates a wider applicability of this system (J Kok, personal communication).

A system for explosive gene expression was developed based on the φ31 middle promoter and origin of replication (ori) [69**]. Essentially, any gene of interest can be cloned under control of the middle promoter combined with a low copy number vector containing the phage ori, which will be induced by infection by φ31, resulting in an increase in plasmid copy number and considerable overproduction of the protein of interest to >1000-fold. The protein(s) of interest will be released into the growth medium since phage infection will lyse the cells within one hour.

A system has been developed for controlled overproduction of proteins, employing the two-component regulatory system for the biosynthesis of the antimicrobial peptide nisin (see above, Figure 2) [42]. It was shown that the nisA promoter could be employed in a series of vectors for expression of several homologous and heterologous genes, which can be induced by the addition of subinhibitory amounts of the signalling molecule nisin [43,70]. A variety of genes have been expressed in *L. lactis* using this system and the hyperproduction (more than 50% protein) of the intracellular aminopeptidase N was reported [42]. Food-grade expression vectors were also developed based on the dominant lacF selection system of *L. lactis*, allowing the overproduction of proteins to be used in the food industry [43]. This study also showed the inducible (Figure 2a) and cell density dependent expression of a reporter gene under control of the nisA promoter (Figure 2b). Recently, two compatible, broad host-range plasmids were constructed, one based on a theta replicon and expressing the nisRK genes and the other based on a high copy number RCR plasmid carrying the nisA promoter controlling the expression of a reporter gene. This dual plasmid system could be introduced and maintained in *Lactococcus, Lactobacillus* and *Streptococcus* spp., and resulted in nisin controlled gene expression (Kleerebezem M, Vaughan EE, de Vos WM, Kuipers OP, unpublished data). As a consequence, the nisin inducible systems hold a high potential for a variety of applications in different LGB. Major advantages are the food-grade nature of the plasmids, strains and inducing agent, the high expression levels that can be reached and the low background activity of the nisA promoter, which enables induction factors of over 1000.

**Conclusion**

A great variety of inducible expression systems have now been developed for LGB. Most of these have been based on the advanced knowledge of gene expression in these hosts and naturally occurring regulatory systems that seem
to have a wide host range in LGB. Various species of LGB, including most of the genera discussed here, are now being analyzed by genomic sequencing. It may be expected that this will, on the one hand, promote the development of new or improved expression systems and, on the other hand, lead to an increased use of these systems in order to exploit LGB as cell factories in medical, agricultural and food biotechnology.

Acknowledgements

We are grateful to Arthur Wolterink for help with Figure 1 and Luis Quadri for handling the Figures. Part of the work in the authors’ laboratory was supported by BIOT contracts from the European Union.

References and recommended reading

Papers of particular interest, published within the annual period of review have been highlighted as:

* of special interest
** of outstanding interest


Expression systems for industrial Gram-positive bacteria de Vos, Klemerezem and Kuipers 553


44. This report discusses the different systems used for enzyme overproduction in lactic acid bacteria and describes the nisin-inducible gene expression system resulting in controlled hyperproduction of intracellular enzymes in L. lactis.


71. The authors describe an elegant system describing coupling of overproduction and lysis of the expression host L. lactis. Induction is not triggered by the addition of a chemical but of a bacteriophage particle.