Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biotechnology 104 (2003) 5–25
JOURNAL OF Biotechnology
www.elsevier.com/locate/jbiotec

The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins

Jörn Kalinowski [a,*], Brigitte Bathe [b], Daniela Bartels [a], Nicole Bischoff [a], Michael Bott [c], Andreas Burkovski [d], Nicole Dusch [a], Lothar Eggeling [c], Bernhard J. Eikmanns [e], Lars Gaigalat [f], Alexander Goesmann [a], Michael Hartmann [f], Klaus Huthmacher [b], Reinhard Krämer [d], Burkhard Linke [a], Alice C. McHardy [a], Folker Meyer [a], Bettina Möckel [b,1], Walter Pfefferle [b], Alfred Pühler [f], Daniel A. Rey [f], Christian Rückert [f], Oliver Rupp [a], Hermann Sahm [c], Volker F. Wendisch [c], Iris Wiegräbe [g], Andreas Tauch [a]

[a] *Institut für Genomforschung, Universität Bielefeld, Universitätsstraße 25, D-33615 Bielefeld, Germany*
[b] *Degussa AG, Kantstraße 2, D-33788 Halle-Künsebeck, Germany*
[c] *Institut für Biotechnologie 1, Forschungszentrum Jülich, D-52425 Jülich, Germany*
[d] *Institut für Biochemie, Universität zu Köln, Zülpicher Straße 47, D-50674 Köln, Germany*
[e] *Abteilung Mikrobiologie und Biotechnologie, Universität Ulm, D-89069 Ulm, Germany*
[f] *Lehrstuhl für Genetik, Universität Bielefeld, Universitätsstraße 25, D-33615 Bielefeld, Germany*
[g] *Institut für Innovationstransfer an der Universität Bielefeld GmbH, Geschäftsbereich BioTech, Universitätsstraße 25, D-33615 Bielefeld, Germany*

Received 17 March 2003; received in revised form 8 May 2003; accepted 12 May 2003

Abstract

The complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032, well-known in industry for the production of amino acids, e.g. of L-glutamate and L-lysine was determined. The *C. glutamicum* genome was found to consist of a single circular chromosome comprising 3 282 708 base pairs. Several DNA regions of unusual composition were identified that were potentially acquired by horizontal gene transfer, e.g. a segment of DNA from *C. diphtheriae* and a prophage-containing region. After automated and manual annotation, 3002 protein-coding genes have been identified, and to 2489 of these, functions were assigned by homologies to known proteins. These analyses confirm the taxonomic position of *C. glutamicum* as related to Mycobacteria and show a broad metabolic diversity as expected for a bacterium living in the soil. As an example for biotechnological application the complete genome sequence was used to

* Corresponding author. Fax: +49-521-106-5626.
E-mail address: joern.kalinowski@genetik.uni-bielefeld.de (J. Kalinowski).
[1] Present address: Qiagen AG, Max-Volmer-Straße 4, D-40724 Hilden, Germany.

0168-1656/03/$ - see front matter
doi:10.1016/S0168-1656(03)00154-8

reconstruct the metabolic flow of carbon into a number of industrially important products derived from the amino acid L-aspartate.

Keywords: *Corynebacterium glutamicum*; Amino acid production; Vitamin production; Genome sequence; Bioinformatics

1. Introduction

In the mid-1950s, Kinoshita and co-workers in Japan isolated a bacterium, which was shown to excrete large quantities of L-glutamic acid into the culture medium (Kinoshita et al., 1957). This bacterium, *Corynebacterium glutamicum*, was described as a short, aerobic, gram-positive rod capable of growing on a variety of sugars or organic acids. Under optimal conditions, this organism converted glucose into high yields of L-glutamic acid within a few days. Currently about 1×10^6 tons of this amino acid are produced with this microorganism annually and used as a flavoring agent (Leuchtenberger, 1996). During the past 40 years, various mutants of *C. glutamicum* have been isolated with the capacity to produce significant amounts of different L-amino acids. Today, L-lysine is produced with mutants deregulated in the biosynthetic pathway on a scale of 4.5×10^5 tons per year. This amino acid is mainly used as a feed additive.

The common practice of developing amino acid-overproducing strains by mutagenesis and selection is a very well established technique (Rowlands, 1984). Mutagenic procedures were optimized in terms of the mutagen used and the dose applied. Selection procedures were designed to allow maximum expression and detection of the desirable mutant types. So far the improvement of amino acid-producing *C. glutamicum* strains has mainly been carried out by an iterative procedure of mutagenesis and selection. However, the precise genetic and physiological changes resulting in an increased overproduction of amino acids in various *C. glutamicum* strains remained unknown. Future success in attempts to further increase the productivity and yield of already highly productive strains will depend on the availability of detailed information on the metabolic pathways, their regulations, and their mutations. In recent years, genetic engineering has become a fascinating alternative to mutagenesis and random screening procedures (Sahm et al., 1995). Overexpression or deletion of genes in microorganisms via recombinant DNA techniques is the most powerful method for the construction of strains with the desired genotype. Furthermore, this approach avoids the complication of uncharacterized mutations that are often obtained with classical mutagenesis.

Since the mid-1980s, several genes from the biosynthetic pathways leading to the aspartate-derived amino acids L-lysine, L-threonine, and L-isoleucine, as well as to the vitamin D-pantothenate in *C. glutamicum* have been cloned and analyzed (Sahm et al., 2000). These genes were mainly identified by heterologous complementation of *Escherichia coli* mutants, and occasionally, in the homologous system by conferring an amino acid-analog resistance. These studies already led to a general understanding of metabolic pathways, but a complete picture of the complex interactions could not be achieved due to the lack of detailed genetic information. Genomic sequencing followed by automatic and manual annotation turned out to represent the ideal method to obtain the missing genetic information for the development of industrial *C. glutamicum* strains. For this reason, we decided in 1998 to sequence the genome of *C. glutamicum* (Hodgson, 1998), sometimes also referred as *Brevibacterium divaricatum*, *B. flavum*, *B. lactofermentum*, or *C. melassecola* (Liebl et al., 1991; Kämpfer and Kroppenstedt, 1996). The sequencing strategy was to use large-insert libraries, e.g. cosmid- and BAC-clones for establishing the complete genome sequence (Tauch et al., 2002a). We now report on the completed genomic sequence of the type strain *C. glutamicum* ATCC 13032. The genome data provide a rich source for metabolic reconstruction of the pathways leading

to industrially important products derived from the amino acid L-aspartate.

During our sequencing work, we learned that due to its outstanding biotechnological relevance, the genome of *C. glutamicum* was sequenced independently by different groups. The Japanese company Kyowa Hakko Kogyo Co., Ltd. established a sequence independently from our project and put it into the public databases (GenBank NC_003450). Its market competitor, Ajinomoto Co. sequenced a close relative, *C. efficiens*, an organism isolated by researchers of this company (Fudou et al., 2002). The sequence of this strain was released recently in the GenBank database (NC_004369).

2. Assembly and annotation of the *C. glutamicum* ATCC 13032 genome sequence

The complete genome sequence of *C. glutamicum* ATCC 13032 was determined from 116 overlapping genomic clones. Of these, 95 were isolated from an ordered SuperCos I cosmid library (Bathe et al., 1996), and 21 were selected from a set of 2304 bacterial artificial chromosomes (BACs) upon mapping to cosmid contig ends by colony hybridization and terminal BAC sequencing (Tauch et al., 2002a). The cosmid library alone covered only 86.6% of the *C. glutamicum* genome and the ordered BAC library was generated to span 18 physical gaps, which were obviously not clonable in the cosmid vector. The largest of these physical gaps had a size of 213 kb and later turned out as a region showing similarity to proteins encoded by prophages (see below). The systematic sequencing of ordered cosmids and BACs instead of a whole-genome shotgun approach avoided assembly problems caused by repetitive sequences. One important deviation from this strategy was due to the duplication of the ribosomal RNA operons *rrnD* and *rrnE*, which are separated by only 3.3 kb (Tauch et al., 2002a). In this case, both operons were cloned as *Pst*I DNA fragments and sequenced separately by a primer-walking strategy using the nucleotide sequence of the *rrnA* operon for primer design.

The nucleotide sequences of the individual cosmid and BAC clones were finally assembled and ambiguities were resolved using the gap4 computer program (Staden, 1996). The assembled genomic sequence was then aligned to all available *C. glutamicum* nucleotide sequence data deposited in public databases. Apparent differences between the sequences, which could be due to mutations introduced during the cloning procedure, were resequenced from PCR products generated from chromosomal DNA. Therefore, a high-quality nucleotide sequence of the *C. glutamicum* chromosome was obtained.

The edited *C. glutamicum* genome sequence was up-loaded into a GenDB database version 1.0.5 (Meyer et al., 2003) and annotated. Gene finding was performed by combining two bioinformatics tools: CRITICA (Badger and Olsen, 1999) was used to define a gene set, which was subsequently used by GLIMMER (Delcher et al., 1999) to construct a training model and to perform the final gene finding. This combination makes effective use of the selectivity of CRITICA (very few false positives) and the sensitivity of GLIMMER (very few true negatives). Genes were validated and coding sequence starts were checked by visual inspection after TBLASTN comparisons of the protein sequences deduced from all *C. glutamicum* ORFs against three other genome sequences from the *Actinomycetales* phylogenetic lineage, comprising *C. diphtheriae* (ftp://ftp.sanger.ac.uk/pub/pathogens/cdip/), *Mycobacterium tuberculosis* (NC_000962), and *Streptomyces coelicolor* (NC_003888). Additional databases used for gene function analysis were the non-redundant protein sequence database (nr), SWISSPROT, and INTERPRO including several protein pattern databases. Additionally, SignalP (Nielsen et al., 1999) and TMHMM (Krogh et al., 2001) were used to identify proteins that are potentially secreted or located in the cytoplasmic membrane, respectively.

3. The structure of the *C. glutamicum* ATCC 13032 genome

General features of the *C. glutamicum* genome sequence are shown in Table 1 and Fig. 1. The *C.*

Table 1
General features of the *C. glutamicum* ATCC 13032 chromosome

Features of the chromosome	Property
Total size	3 282 708 bp
G+C content	53.8%
Coding sequences (CDS)-total	3002 (100%)
CDS encoding annotated proteins	2489 (83%)
CDS encoding putative cytosolic proteins	1518 (51%)
CDS encoding putative membrane proteins	660 (22%)
CDS encoding putative secreted proteins	311 (10%)
CDS encoding conserved hypothetical proteins	250
CDS encoding hypothetical proteins	263
Coding density	87%
Average gene length	952 bp
Ribosomal RNAs	6 operons (16S-23S-5S)
Transfer RNAs	42 different/60 genes
Other stable RNAs	2

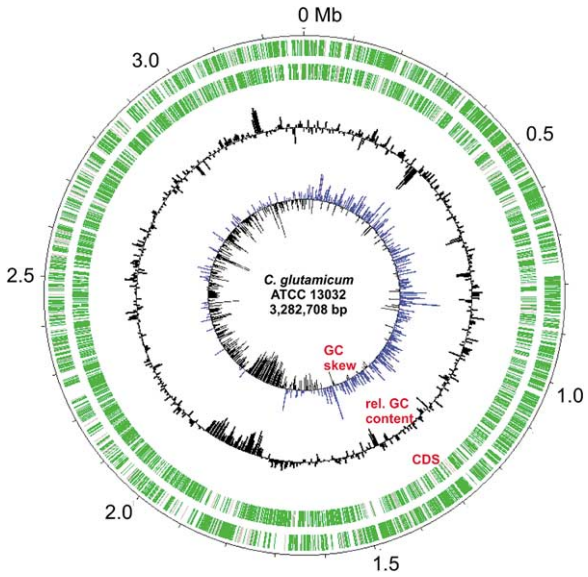

Fig. 1. Circular representation of the *C. glutamicum* ATCC 13032 chromosome. The concentric circles denote (from outward to inward): coding sequences (CDS) transcribed clockwise and counter-clockwise, relative G+C content, and GC skew. A positive deviation in G+C content from the average is shown by bars pointing outward and a negative deviation by bars pointing inward. The same holds for the GC skew plot where positive skew values are shown in blue color and negative values in black. The *C. glutamicum* ATCC 13032 genome sequence was deposited in the EMBL database.

glutamicum genome is represented by a circular chromosome of 3 282 708 bp, which is smaller than the genome of the taxonomically related bacterium *M. tuberculosis* (4.2 Mb), but larger than that of its close relative *C. diphtheriae* (2.5 Mb). The G+C content of the genome is 53.8%, which is close to that of *E. coli* and rather unusual for the taxonomic class of the *Actinobacteria* referred to as 'high G+C gram-positive bacteria'.

The GC skew analysis (Grigoriev, 1998), which is generally applicable to identify the leading and the lagging strand in DNA replication, indicated a bi-directional replication that starts at the proposed *oriC* sequence near the *dnaA* gene (*cg0001*) and ends near to the calculated replication terminus at around 1.6 Mb (Fig. 1). It has to be noted that several regions of the *C. glutamicum* genome deviate significantly in G+C content from the median (Fig. 1). A closer inspection identified two larger genomic regions as exceptionally GC-poor. The first region has a size of approximately 25 kb and covers 20 coding regions (*cg0415–cg0443*) with G+C-contents of 41–49%. The genes that are located in this region are involved in some aspects of murein formation (*murA*, *cg0422*; *murB*, *cg0423*) and lipopolysaccharide synthesis. It is interesting to note that also in other organisms, low-GC regions carrying genes involved in cell wall formation have been found (e.g. *Bacillus subtilis*, NC_000964). However, whether this indicates a horizontal gene transfer or some adaptive pressure for a low G+C content implied by the functional context of these genes is currently unknown.

The second region of low G+C content is much larger and spans more than 200 kb (1.8–2.0 Mb, Fig. 1). The region covers approximately 180 coding regions, most of them are without any significant similarities to known bacterial genes. However, there are a few exceptions including the three genes of the already known restriction-modification system (*cg1996–cg1998*; Schäfer et al., 1997), genes encoding transposases, putative recombination enzymes, and a number of homo-

logs to known bacteriophage proteins, especially a phage primase (*cg1959*) and a putative phage-type integrase (*cg2071*). It is interesting to note that the left border of this region is formed by a cluster of tRNA genes, whereas the putative phage-type integrase is near to the right border of the insertion. These observations might be explained by the integration of one or more prophage-like elements at a specific tRNA locus, a mechanism which is common for phages and integrative plasmids as demonstrated in streptomycetes (Gabriel et al., 1995). This presumed prophage region provided enormous cloning problems and the whole 213-kb region was not represented in the cosmid library. Also in the BAC library of *C. glutamicum*, the presumed prophage region was significantly underrepresented (Tauch et al., 2002a).

In contrast to these regions exceptional in having a lower G+C content, there is one region of 14 kb in size, which deviates significantly to a higher G+C content. The genes of this region (*cg3280–cg3295*) have G+C contents up to 66% and are flanked by defective insertion sequences. The leftward 7 kb of this region are more than 95% identical at the nucleotide level to a segment from the *C. diphtheriae* genome and contain a putative copper transport system and a two-component sensor-regulator system. The high similarity of this region can only be explained by assuming recent horizontal gene transfer from diphtheroid corynebacteria to the soil bacterium *C. glutamicum*.

Beside this clear indication for a recent horizontal gene transfer between rather distantly related corynebacterial species, an evaluation of gene order conservation revealed an astonishing amount of synteny between the *Corynebacterium* species *C. glutamicum*, *C. efficiens*, and *C. diphtheriae* (Fig. 2). This reciprocal BLASTP analysis of all coding regions derived from the three genomes also clearly showed that the putative bacteriophage insertion region is of alien origin. Additionally, there are several smaller regions carrying genes with no counterpart in one of the other species. A closer inspection of these regions might reveal genes, which are also either horizontally transferred or only necessary in a certain ecological niche.

The lack of detectable inversions between the three corynebacterial genomes is striking, although there is a considerable number of repeated DNA elements present in the *C. glutamicum* genome, which are known to be involved in genomic rearrangements by homologous or illegitimate recombination. Beside six copies of the *rrn* genes encoding ribosomal RNA (gene order: *rrs-rrl-rrf*), 24 different insertion sequences were identified (Table 2). These IS elements can be grouped into nine different families according to the classification by Mahillon and Chandler (1998). Although 11 of these copies are partially deleted, and therefore, apparently defective, most of the insertion sequences appear to have functional copies.

Most of the IS elements in the *C. glutamicum* genome seem not to influence its functionality. There are only two obvious cases of gene disruption by transposition of an IS element. The first case is the insertion of IS*Cg13* (copy b, *cg0695*) into an ORF similar to genes encoding sarcosine oxidase. This insertion is found also in the NC_003450 sequence. The second case is the insertional inactivation of IS*Cg1* (copy c, *cg1782*) into a gene similar to *groEL* encoding a chaperone. Interestingly, this insertion is not present in the NC_003450 sequence indicating a relatively recent transposition event. However, there is no obvious phenotype associated with this insertion, since the *groEL-groES* gene cluster is duplicated in the fully sequenced corynebacterial genomes and in *M. tuberculosis*.

Transpositional recombination of IS*Cg1* (IS*31831*) was previously analyzed in detail (Vertès et al., 1994) and the exceptional target site specificity of IS*Cg2*, which is located adjacent to genes involved in aspartate and glutamate metabolism of the *C. glutamicum* genome, was already reported (Quast et al., 1999). Virtually identical copies of IS*Cg3* (IS*1628*), IS*Cg11* (IS*1870*), and IS*Cg13* (IS*1677*) were identified previously during nucleotide sequence analysis of the resistance plasmids pAG1 and pTET3 from *C. glutamicum* (Tauch et al., 2000, 2002b), indicating that genetic exchange occurred between the chromosome of *C. glutamicum* and its potential plasmid replicons. Furthermore, it is noteworthy that both IS*Cg15* elements are part of a cryptic

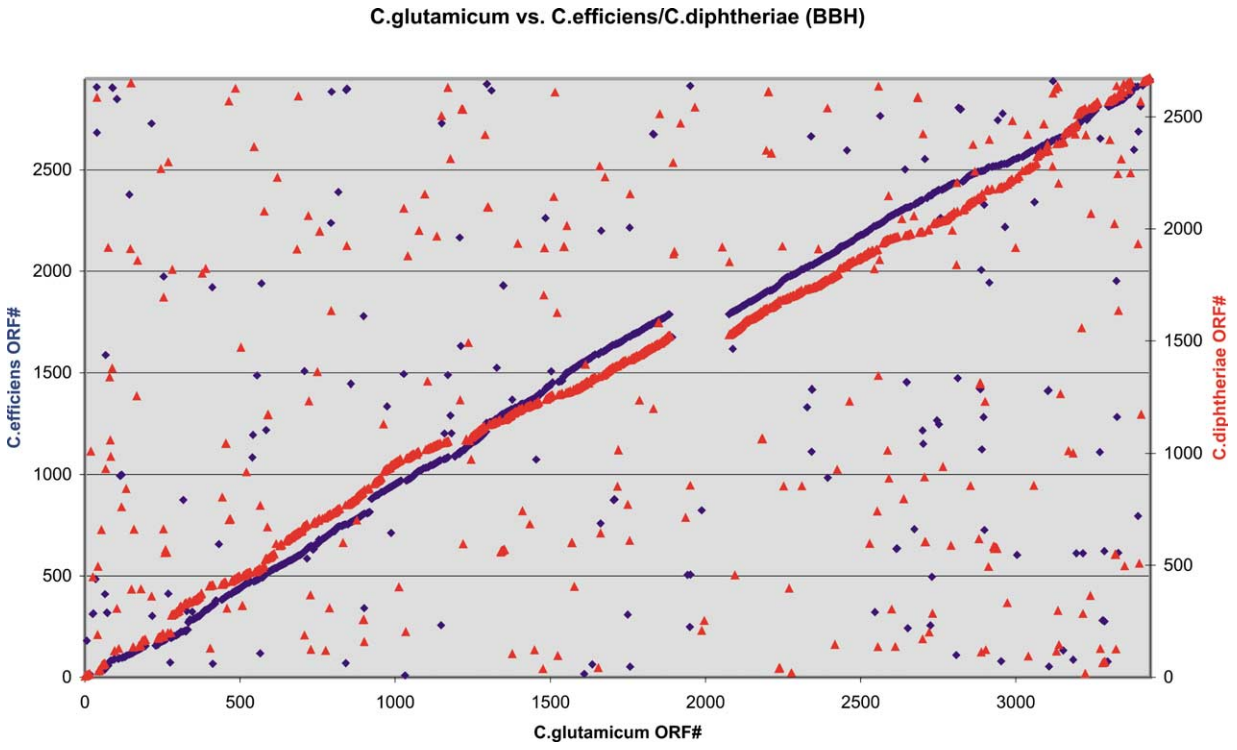

Fig. 2. Synteny between the *Corynebacterium* species *C. glutamicum*, *C. efficiens* and *C. diphtheriae*. The diagram shows X–Y plots of dots forming syntenic regions between the genomes. Each dot represents a reciprocal BLASTP best hit of amino acid sequences from automatically extracted *C. glutamicum* ORFs (3432 ORFs total) with *C. diphtheriae* ORFs (2671 ORFs total, red triangles) or *C. efficiens* ORFs (2950 ORFs total, blue squares), respectively. The *C. efficiens* YS-314 data were obtained from GenBank database (NC_004369).

composite transposon, Tn*Cg1*, carrying the hypothetical gene *cg2758* within the central region (Table 2). Transposon Tn*Cg1* has a length of 3148 bp and generated an 8-bp target duplication at the integration site. Mobile genetic elements also represent the main differences between the two available *C. glutamicum* genome sequences. Both genomes differ not only by the number of IS*Cg1* and IS*Cg2* copies (Table 2) but also by an approximately 31.4-kb region, which is characterized by a terminal gene encoding a prophage-type integrase (*cg2071*) and which is only present in the GenBank sequence NC_003450. Obviously, highly active insertion sequences and probably bacteriophages contribute to the rapid divergence of *C. glutamicum* strains.

4. Annotation of coding regions

Gene finding tools in conjunction with homology searches in databases and an additional expert annotation with the genome annotation tool GenDB (Meyer et al., 2003) revealed 3002 potential protein-coding genes in the *C. glutamicum* genome sequence (Table 1). To 2489 of these, at least putative functions or localizations could be assigned by similarity analyses. Of the remaining predicted genes, 250 are similar to hypothetical proteins in other organisms (conserved hypothetical proteins) and only 9% (263) of the predicted genes remain hypothetical or specific for *C. glutamicum*.

From a soil bacterium, it can be expected that its genome has to encode all necessary functions for

Table 2
Insertion sequences in the *C. glutamicum* ATCC 13032 genome

IS element	IS family	Copy number	Coding region[a]	Structure	Size (bp)	IR[b] (bp)	DR[c] (bp)	Closest similarity to IS element; microorganism	GenBank No.
IS*Cg1a*	IS*L3*	4	*cg1213*	Complete	1453	24	8	IS*31831*; *C. glutamicum* ATCC 31831	D17429
IS*Cg1b*			*cg2725*	Complete	1453	24	8		
IS*Cg1c*			*cg0692*	Complete	1453	24	8		
IS*Cg1d*			*cg2600*	Complete	1453	24	8		
IS*Cg2b*	IS*30*	5	*cg3151*	Complete	1636	26	3	IS*1513*; *Corynebacterium striatum*	AF024666
IS*Cg2c*			*cg2854*	Complete	1636	26	3		
IS*Cg2d*			*cg2426*	Complete	1636	26	3		
IS*Cg2e*			*cg2353*	Complete	1636	26	3		
IS*Cg2f*			*cg0226*	Complete	1636	26	3		
IS*Cg3a*	IS*6*	2	*cg1094*	Complete	841	22/21	N.D.	IS*1628*; *C. glutamicum* 22243	AF121000
IS*Cg3b*			*cg1757*	Complete	841	22/21	8		
IS*Cg4*	IS*5*	1	*cg2463*	Complete	870	21	N.D.	IS*1421*; *Ralstonia solanacearum*	AL646079
IS*Cg5a*	IS*3*	3	*cg0824*	Complete	1720	73/71	12	ISA0963; *Archaeoglobus fulgidus*	AE001007
IS*Cg5b*			*cg2915*	Complete	1720	73/71	12		
IS*Cg5c*			*cg3266*	3′-end	1641	–/71	N.A.		
IS*Cg6a*	IS*3*	2	*cg1030/31*	Complete	1299	26/25	N.D.	IS*3*; *Vibrio vulnificus*	AF499932
IS*Cg6b*			*cg1022/23*	Complete	1299	26/25	N.D.		
IS*Cg7*	IS*L3*	1	*cg1024*	3′-end	347	–/24	N.A.	IS*31831*; *C. glutamicum* ATCC 31831	D17429
IS*Cg8*	IS*3*	1	*cg3058/59*	Complete	1295	15	N.D.	IS*3502*; *Corynebacterium jeikeium*	AY033500
IS*Cg9*	IS*110*	1	*cg1178*	Complete	N.D.	N.D.	N.D.	IS*110*; *S. coelicolor*	Y00434
IS*Cg10*	IS*L3*	1	*cg1184–87*	3′-end	979	–/24	N.A.	IS*31831*; *C. glutamicum* ATCC 31831	D17429
IS*Cg11*	IS*630*	1	*cg2807*	3′-end	371	–/20	N.A.	IS*1870*; *C. glutamicum* LP-6	AJ420072
IS*Cg12*	IS*256*	1	*cg2652/54*	Complete	1426	21/20	8	IS*1249*; *Corynebacterium striatum*	AF024666
IS*Cg13a*	IS*4*	2	*cg2808*	Complete	1688	18	5	IS*1677*; *C. glutamicum* LP-6	AJ420072
IS*Cg13b*			*cg1782*	Complete	1688	18	5		
IS*Cg14*	IS*3*	1	*cg1950/51*	Complete	1293	26	3	IS*1206*; *C. glutamicum* Bl15	X69104
IS*Cg15a*[d]	IS*6*	2	*cg2757*	Complete	843	20	8/–	IS*1674*; *C. glutamicum* LP-6	AJ420072
IS*Cg15b*			*cg2759*	Complete	846	20	–/8		
IS*Cg16a*	IS*3*	2	*cg0292*	Complete	N.D.	N.D.	N.D.	IS*1141*; *Mycobacterium intracellulare*	L10239
IS*Cg16b*			*cg1716*	Complete	N.D	N.D.	N.D.		
IS*Cg17*	IS*L3*	1	*cg0426–28*	3′-end	842	–/21	N.A.	IS*31831*; *C. glutamicum* ATCC 31831	D17429
IS*Cg18*	IS*5*	1	*cg0919*	5′-end	N.D.	N.D.	N.A.	IS*1502*; *Leptospira interrogans*	AF434658
IS*Cg19*	IS*6*	1	*cg3296–98*	3′-end	676	–/14	N.A.	IS*1673*; *C. glutamicum* ATCC 31830	AF164956
IS*Cg20*	IS*6*	1	*cg3277*	5′-end	270	14/–	N.A.	IS*1674*; *C. glutamicum* LP-6	AJ420072
IS*Cg21*	IS*4*	1	*cg2804*	Complete	1653	18	4	IS*1677*; *C. glutamicum* LP-6	AJ420072
IS*Cg22*	IS*256*	1	*cg0037*	3′-end	N.D.	N.D.	N.A.	IS*1249*; *Corynebacterium striatum*	AF024666
IS*Cg23*	IS*3*	1	*cg1513*	3′-end	N.D.	N.D.	N.A.	IS*1206*; *C. glutamicum* Bl15	X69104
IS*Cg24*	IS*3*	1	*cg1515*	5′-end	N.D.	N.D.	N.A.	IS*1206*; *C. glutamicum* Bl15	X69104

Abbreviations: N.D., not determined; N.A., not applicable.

[a] Coding regions underlined are absent from the *C. glutamicum* genom sequence deposited with GenBank Accession number NC_003450.

[b] IR indicates the length(s) of the terminal inverted repeat in base pairs. A unique number refers to two IRs with the same length. A hyphen indicates that an IR is absent due to a partial sequence. The lengths of IRs from partial IS elements was deduced from closely related elements.

[c] DR indicates the number of target base pairs duplicated on insertion.

[d] IS*Cg15a* and IS*Cg15b* are part of the composite transposon Tn*Cg1* characterized by a length of 3148 bp and an 8-bp target duplication.

primary metabolism, for catabolism of a wide variety of different nutrients and for optimal adaptation to changes in the environment. *C. glutamicum* is the first completely sequenced gram-positive soil bacterium from the CMN-group of the *Actinobacteria* (Barksdale, 1981). The other members of this group whose genomes are known are *C. diphtheriae*, *M. tuberculosis*, and *M. leprae*, all of them important human pathogens. Since non-pathogenic model systems are necessary, *C. glutamicum* may serve as an ideal system for studying the cell wall and especially, mycolic acid synthesis.

As expected for a soil bacterium, *C. glutamicum* is capable of growing in a simple mineral salts medium, i.e. it is able to synthesize from simple precursors all cell constituents including metabolites, cofactors and vitamins, except for D-biotin. This defect is most probably due to the fact that the gene *bioF*, encoding the biotin biosynthetic enzyme 7-keto-8-aminopelargonic acid synthetase, is missing in *C. glutamicum* (Hatakeyama et al., 1993a,b).

All of the genes already described for different *C. glutamicum* strains and represented as nucleotide sequences in public databases were found also in the sequence obtained here with one important exception. The gene for the paracristalline surface-layer protein *cspB* (Peyret et al., 1993) from *C. glutamicum* ATCC 17965, which is synthesized in extremely large amounts and has a possible function in protecting the bacterium in soil against rough conditions, is missing in both *C. glutamicum* ATCC 13032 sequences. It is not clear why this gene is absent, but it can be speculated that bacterial strains in laboratories adapt to the specific growth conditions by loosing functions that provide a heavy metabolical load carrying out a protective function unnecessary under optimal growth conditions (Fujita et al., 1997).

5. Metabolic reconstruction of the biosyntheses of aspartate-derived amino acids and vitamins from glucose

A number of metabolites of biotechnological importance are derived from the amino acid L-aspartate. These are L-lysine, L-threonine, L-methionine and L-isoleucine. Two others compounds, the amino acid L-valine and the vitamin D-pantothenate, are strongly interconnected to the synthesis of aspartate-derived amino acids and were, therefore, included into this study. For the reconstruction of the formation of all these compounds from glucose, several functional complexes have to be considered. These most important ones are: sugar uptake, central metabolism, amino acid and vitamin biosynthesis as well as their transport processes (Fig. 3).

5.1. Sugar uptake system in C. glutamicum

C. glutamicum uses the PTS system for the uptake of glucose, fructose, mannose, and sucrose (Dominguez and Lindley, 1996). The PTS system functions as follows: The general phosphotransferase enzyme I (EI; *ptsI*) becomes autophosphorylated by phosphoenolpyruvate and transfers its phosphoryl group to the second general component HPr (*ptsH*). HPr in turn phosphorylates a number of sugar specific permeases, the so-called enzyme II-sugar complexes (Saier and Reizer, 1992; Lengeler et al., 1994). These complexes transport their substrates by concomitant phosphorylation. Up to now, the PTS system in *C. glutamicum* was characterized mainly by mutational analyses and phenotypic descriptions (Parche et al., 2001a). On the molecular level, only two genes were described, one encoding a mannose-specific enzyme II (*ptsM*; Lee et al., 1994) and the other encoding enzyme I (*ptsI*; Kotrba et al., 2001). However, the genome sequence of the closely related *C. diphtheriae* was already investigated by similarity analyses to describe the PTS components in this organism (Parche et al., 2001b).

A similarity search in the *C. glutamicum* genome for the general PTS components EI and HPr revealed *ptsI* (*cg2117*) and close to it, a *ptsH*-like ORF (*cg2121*) transcribed in opposite orientation. Between *ptsI* and *ptsH*, three genes were identified that encode a putative transcriptional regulator (*cg2118*), a fructose 1-phosphate kinase (*pfkB*, *cg2119*), and a putative fructose-specific enzyme II (*ptsF*, *cg2120*), giving the gene order *ptsI-cg2118-*

Table 3
Genes and proteins involved in the synthesis of aspartate-derived amino acids and vitamins

CDS	Gene	Protein encoded	Functional complex	Reference
cg0148	*panC*	Pantoate-β-alanine ligase	Pantothenate synthesis	Sahm and Eggeling, 1999
cg0149	*panB*	3-Methyl-2-oxobutanoate hydroxymethyltransferase	Pantothenate synthesis	Sahm and Eggeling, 1999
cg0172	*panD*	Aspartate-α-decarboxylase	Pantothenate synthesis	Dusch et al., 1999
cg0306	*lysC*	Aspartate LysCα and LysCβ subunits	Lysine biosynthesis	Kalinowski et al., 1991
cg0307	*asd*	Aspartate semialdehyde dehydrogenase	Lysine biosynthesis	Kalinowski et al., 1990
cg0314	*brnF*	Isoleucine exporter, component 2	Isoleucine transport	Kennerknecht et al., 2002
cg0315	*brnE*	Isoleucine exporter, component 1	Isoleucine transport	Kennerknecht et al., 2002
cg0441	*lpd*	Lipoamide dehydrogenase	Central carbon metabolism	Schwinde et al., 2001
cg0445	*sdhC*	Succinate dehydrogenase CD	Tricarboxylic acid cycle	AX113263
cg0446	*sdhA*	Succinate dehydrogenase A	Tricarboxylic acid cycle	AX113263
cg0447	*sdhB*	Succinate dehydrogenase B	Tricarboxylic acid cycle	AX113263
cg0482	*gpm*	Phosphoglycerate mutase	Glycolysis	This work
cg0644	*ppsA*	Phosphoenolpyruvate synthase	Central carbon metabolism	This work
cg0754	*metX*	Homoserine *O*-acetyltransferase	Methionine biosynthesis	Park et al., 1998
cg0755	*metY*	*O*-acetylhomoserine sulfhydrylase	Methionine biosynthesis	Hwang et al., 2002
cg0763	*mdh*	Malate dehydrogenase	Tricarboxylic acid cycle	This work
cg0766	*icd*	Isocitrate dehydrogenase	Tricarboxylic acid cycle	Eikmanns et al., 1995
cg0790	*lpdA*	Lipoamide dehydrogenase	Central carbon metabolism	This work
cg0791	*pyc*	Pyruvate carboxylase	Central carbon metabolism	Peters-Wendisch et al., 1998
cg0798	*prpC1*	Citrate synthase/methylcitrate synthase	Tricarboxylic acid cycle	Claes et al., 2002
cg0897	*pdxR*	Pyridoxine biosynthesis regulator involved in valine biosynthesis	Valine biosynthesis	McHardy et al., 2003
cg0949	*gltA*	Citrate synthase	Tricarboxylic acid cycle	Eikmanns et al., 1994
cg0973	*pgi*	Glucose-6-phosphate isomerase	Glycolysis	AX253248
cg1069	*gapX*	Putative glyceraldehyde-3-phosphate dehydrogenase	Glycolysis	This work
cg1075	*prsA*	Phosphoribosyl pyrophosphate synthase isoenzyme 2 precursor	Pentose phosphate cycle	This work
cg1105	*lysI*	Lysine permease	Lysine transport	Seep-Feldhaus et al., 1991
cg1111	*eno*	2-Phosphoglycerate dehydratase, 2-phospho-D-glycerate hydrolyase	Glycolysis	AX136862
cg1133	*glyA*	Serine hydroxymethyltransferase	C1 metabolism	Simic et al., 2002
cg1145	*fum*	Fumarate hydratase	Tricarboxylic acid cycle	This work
cg1239	–	(Falsely) predicted ketopantoate reductase	Pantothenate synthesis	Merkamm et al., 2003
cg1253	*dapC*	*N*-succinyl diaminopimelate aminotransferase	Lysine biosynthesis	Hartmann et al., 2003
cg1256	*dapD*	Tetrahydrodipicolinate succinylase	Lysine biosynthesis	Wehrmann et al., 1998
cg1260	*dapE*	*N*-succinyl diaminopimelate desuccinylase	Lysine biosynthesis	Wehrmann et al., 1994
cg1280	*sucA (odhA)*	2-Oxoglutarate dehydrogenase	Tricarboxylic acid cycle	Usuda et al., 1996

Table 3 (*Continued*)

CDS	Gene	Protein encoded	Functional complex	Reference
cg1290	*metE*	Homocysteine methyltransferase	Methionine biosynthesis	Rückert et al., 2003
cg1334	*lysA*	Diaminopimelate decarboxylase	Lysine biosynthesis	Yeh et al., 1988
cg1337	*hom*	Homoserine dehydrogenase	Threonine biosynthesis	Peoples et al., 1988
cg1338	*thrB*	Homoserine kinase	Threonine biosynthesis	Peoples et al., 1988
cg1409	*pfkA*	Phosphofructokinase A	Glycolysis	This work
cg1424	*lysE*	Lysine efflux permease	Lysine transport	Vrljic et al., 1996
cg1425	*lysG*	Lysine export regulator protein	Lysine transport	Bellmann et al., 2001
cg1432	*ilvD*	Dihydroxy-acid dehydratase	Isoleucine/valine biosynthesis	Radmacher et al., 2002
cg1435	*ilvB*	Acetohydroxy acid synthase, large subunit	Isoleucine/valine biosynthesis	Keilhauer et al., 1993
cg1436	*ilvn*	Acetohydroxy acid synthase, small subunit	Isoleucine/valine biosynthesis	Keilhauer et al., 1993
cg1437	*ilvc*	Acetohydroxy acid isomeroreductase	Isoleucine/valine biosynthesis	Keilhauer et al., 1993
cg1537	*ptsG (ptsM)*	Glucose-specific PTS enzyme II	Sugar uptake	Lee et al., 1994
cg1546	*rbsK1*	Ribokinase sugar family kinase	Pentose phosphate cycle	This work
cg1643	*gnd*	6-Phosphogluconate dehydrogenase	Pentose phosphate cycle	AX253243
cg1701	*metH*	Homocysteine methyltransferase	Methionine biosynthesis	Rückert et al., 2003
cg1737	*acn*	Aconitase	Tricarboxylic acid cycle	This work
cg1774	*tkt*	Transketolase	Pentose phosphate cycle	Ikeda et al., 1999
cg1776	*tal*	Transaldolase	Pentose phosphate cycle	AX076274
cg1778	*zwf*	Glucose-6-phosphate 1-dehydrogenase	Pentose phosphate cycle	AX074270
cg1780	*devB*	Putative 6-phosphogluconolactonase	Pentose phosphate cycle	This work
cg1787	*ppc*	Phosphoenolpyruvate carboxylase	Central carbon metabolism	O'Regan et al., 1989; Eikmanns et al., 1989
cg1789	*tpi*	Triosephosphate isomerase	Glycolysis	Eikmanns, 1992
cg1790	*pgk*	Phosphoglycerate kinase	Glycolysis	Eikmanns, 1992
cg1791	*gap*	Glyceraldehyde-3-phosphate dehydrogenase	Glycolysis	Eikmanns, 1992
cg1801	*rpe*	Ribulose-5-phosphate-3-epimerase	Pentose phosphate cycle	This work
cg2117	*ptsI*	PTS enzyme I	Sugar uptake	Kotrba et al., 2001
cg2118	–	Transcriptional regulator, DeoR family	Sugar uptake	This work
cg2119	*pfkB*	Phosphofructokinase B	Glycolysis	This work
cg2120	*ptsF*	Fructose-specific PTS enzyme II	Sugar uptake	This work
cg2121	*ptsH*	Phosphocarrier protein HRP	Sugar uptake	This work
cg2129	*dapF*	Diaminopimelate epimerase	Lysine biosynthesis	Hartmann et al., 2003
cg2161	*dapA*	Dihydrodipicolinate synthase	Lysine biosynthesis	Cremer et al., 1990
cg2163	*dapB*	Dihydrodipicolinate reductase	Lysine biosynthesis	Cremer et al., 1990
cg2192	*mqo*	Malate:quinone oxidoreductase	Tricarboxylic acid cycle	Molenaar et al., 1998
cg2291	*pyk*	Pyruvate kinase	Glycolysis	Jetten et al., 1994
cg2334	*ilvA*	Threonine dehydratase	Isoleucine biosynthesis	Möckel et al., 1992

Table 3 (*Continued*)

CDS	Gene	Protein encoded	Functional complex	Reference
cg2374	*murE*	UDP-*N*-acetylmuramoyl-tripeptide synthetase	Cell wall synthesis	Wijayarathna et al., 2001
cg2383	*metF*	5,10-Methylenetetrahydrofolate reductase	Methionine biosynthesis	Rückert et al., 2003
cg2399	*glk*	Glucokinase	Central carbon metabolism	Park et al., 2000
cg2418	*ilvE*	Branched-chain amino acid aminotransferase	Isoleucine/valine biosynthesis	Radmacher et al., 2002
cg2421	*sucB*	Dihydrolipoamide succinyltransferase	Tricarboxylic acid cycle	This work
cg2437	*thrC*	Threonine synthase	Threonine biosynthesis	Han et al., 1990
cg2466	*aceE*	Pyruvate dehydrogenase E1 component	Central carbon metabolism	This work
cg2536	*aecD*	Cystathionine β-lyase	Methionine biosynthesis	Rossol and Pühler, 1992
cg2537	*brnQ*	Isoleucine uptake carrier	Isoleucine transport	Tauch et al., 1998
cg2554	*rbsK2*	Ribokinase sugar family kinase	Pentose phosphate cycle	This work
cg2559	*aceB*	Malate synthase	Glyoxylate cycle	Reinscheid et al., 1994b
cg2560	*aceA*	Isocitrate lyase	Glyoxylate cycle	Reinscheid et al., 1994a
cg2613	*mdh*	Malate dehydrogenase	Tricarboxylic acid cycle	AJ303072
cg2658	*rpi*	Possible phosphopentose isomerase	Pentose phosphate cycle	This work
cg2687	*metB*	Cystathionine γ-synthase	Methionine biosynthesis	Hwang et al., 1999
cg2800	*pgmA*	Phosphoglycerate mutase	Central carbon metabolism	This work
cg2836	*sucD*	Succinyl-CoA synthetase α subunit	Tricarboxylic acid cycle	This work
cg2837	*sucC*	Succinyl-CoA synthetase β subunit	Tricarboxylic acid cycle	This work
cg2891	*poxB*	Pyruvate oxidase	Central carbon metabolism	AX253251
cg2900	*ddh*	*Meso*-diaminopimelate dehydrogenase	Lysine biosynthesis	Ishino et al., 1987
cg2905	*thrE*	Threonine export carrier	Threonine transport	Simic et al., 2001
cg2925	*ptsS*	Sucrose-specific PTS enzyme II	Sugar uptake	This work
cg2926	*scrB*	Sucrose-6-phosphate hydrolase	Sugar metabolism	This work
cg3068	*fda*	Fructose-bisphosphate aldolase	Glycolysis	von der Osten et al., 1989
cg3169	*pck*	Phosphoenolpyruvate carboxykinase	Central carbon metabolism	Riedel et al., 2001
cg3335	*malE (mez)*	Malic enzyme	Central carbon metabolism	Gourdon et al., 2000

pfkB-ptsF-ptsH. The *ptsF* gene product is similar to several enzyme II components of the fructose/mannitol family. In addition, the gene order in this locus is the same as in *C. diphtheriae* (Parche et al., 2001b).

A screening with the presumed mannose-specific enzyme II sequence of *C. glutamicum* (Lee et al., 1994) led to an almost identical gene that encodes an enzyme II of the glucose/sucrose transporter family (*ptsG*, *cg1537*). From genome analysis, it became clear that *ptsM* and *ptsG* are the same gene and the designation *ptsG* is more appropriate since it encodes the only candidate for the glucose transporting enzyme II in *C. glutamicum*. Beside *ptsF* and *ptsG*, a third gene encoding an enzyme II was detected. This protein encoded by *cg2925*

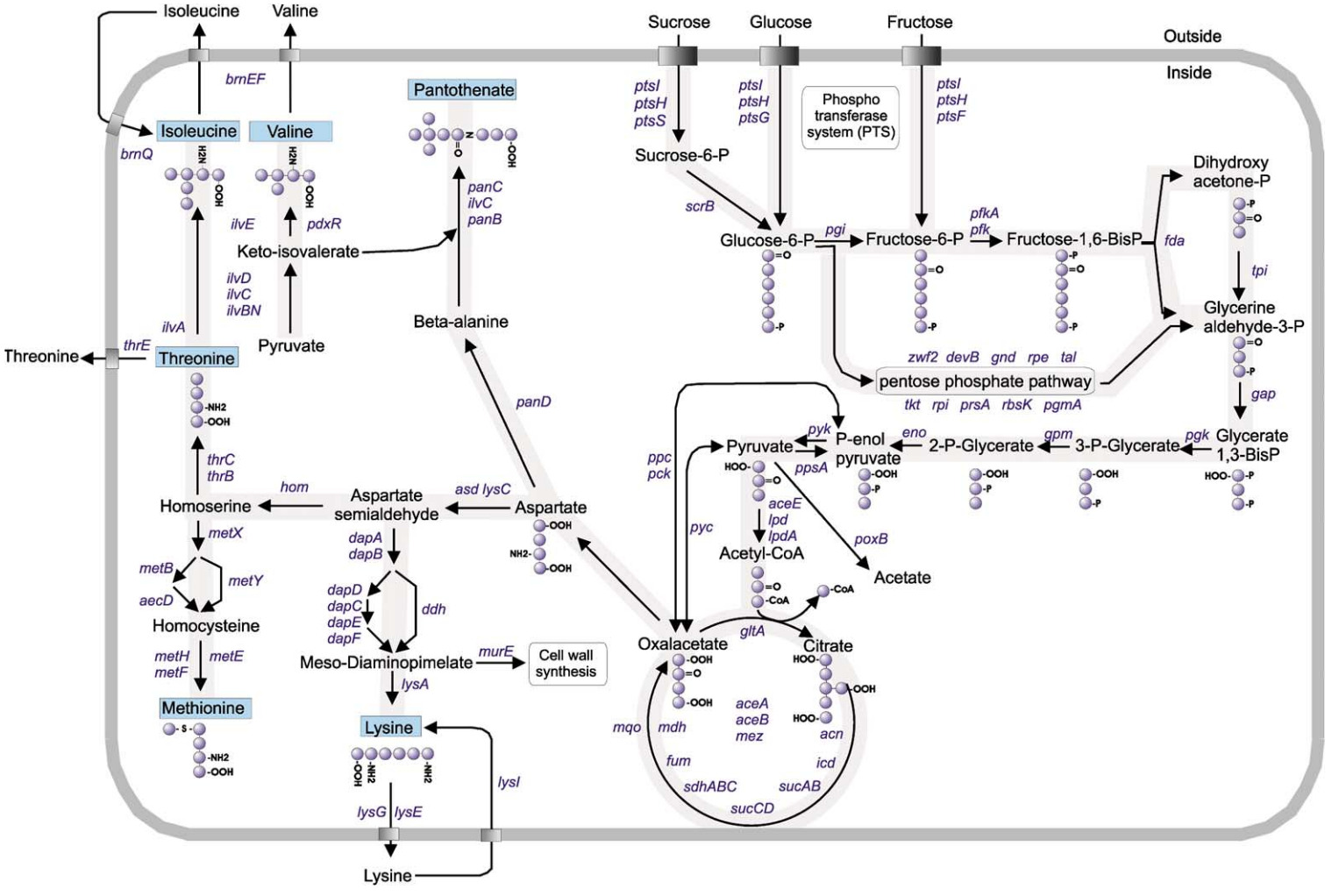

Fig. 3. Metabolic reconstruction of biosynthetic and transport pathways important for the overproduction of L-amino acids and vitamins derived from aspartate or pyruvate. A number of key metabolites are shown in structural formulas with bullets representing methyl groups. Details on the genes and their gene products are given in the text and in Table 3.

(*ptsS*) is similar to sucrose-specific enzyme II in several other bacteria and is located next to a gene encoding a sucrose-6-phosphate hydrolase (*scrB*, *cg2926*), producing glucose-6-phosphate and fructose from sucrose-6-phosphate. Whereas *ptsF* and *ptsG* homologous genes were also found in *C. diphtheriae*, a *ptsS* locus was not detected in this organism (Parche et al., 2001b). Although the existing homologies of the PTS components in *C. glutamicum* gave clear hints for the involvement of PtsG in the uptake of glucose, as well as PtsF in that of fructose and PtsS in the uptake of sucrose, the exact substrate spectra of the enzyme II components remain to be determined experimentally.

5.2. The central carbon metabolism in C. glutamicum

C. glutamicum can metabolize a variety of carbon and energy sources such as carbohydrates, organic acids and alcohols. Several genes for central metabolic pathways have been found and characterized before, however, genome analysis now established that the genes encoding all the enzymes for glycolysis (Embden-Meyerhof-Parnas pathway), the pentose phosphate pathway, the tricarboxylic acid cycle, the glyoxylate cycle and other anaplerotic enzymes and for gluconeogenesis are present in *C. glutamicum*. The only gene not found so far is an *aceF* homologue encoding the E2p subunit of the pyruvate dehydrogenase complex (PDHC). Since activity of this central enzyme complex has been observed (Shiio et al., 1984), it seems that E2p in *C. glutamicum* is different from those deposited in present data banks. It is noteworthy that also the 2-oxoglutarate dehydrogenase complex (OGDHC) in *C. glutamicum* may be different from that in other bacteria (Usuda et al., 1996). Based on the characterization of an unusual *odhA* gene (*cg1280*), these authors speculate that the E1o and E2o subunits of this complex represent a single bifunctional enzyme in *C. glutamicum*. The presence of an unusual PDHC in this organism might be of fundamental importance for amino acid production as it has been shown that a mutant with low PDHC activity produces much more lysine than the parental strain (Shiio et al., 1984).

The genes coding for central metabolic pathways are sometimes clustered (e.g. *gap-pgk-tpi-ppc*, *aceA-aceB* or *sdhC-sdhA-sdhB*), however, these clusters and also the single genes are scattered over the whole chromosome. Thus, the chromosomal organization of these genes is quite similar to that in *M. tuberculosis* (NC_000962).

Although the central metabolic pathways in *C. glutamicum* have been known before, genome analysis revealed interesting new information. It becomes evident that the metabolic flexibility of *C. glutamicum* has been underestimated since obviously a number of isoenzymes within the fundamental pathways are present. Also genes for enzymes hitherto not known to be present in *C. glutamicum* have been detected. There are genes encoding phosphofructokinases A and B (*cg1409* and *cg2119*), which in other organisms are known to be regulated differently by allosteric activators. Genes for two different glyceraldehyde-3-phosphate dehydrogenases (*cg1069* and *cg1791*), two putative phosphoglycerate mutases (*cg0482* and *cg2800*), two putative lipoamide dehydrogenases (*cg0441* and *cg0790*), and two citrate synthases (*cg0949* and *cg0798*) were detected. A gene with significant identity to the *E. coli* pyruvate oxidase gene (*poxB*, *cg2891*) has been identified. The presence of this enzyme further increases the complexity of the *C. glutamicum* C-flux at the pyruvate/oxaloacetate node, which just recently has been shown to be of pivotal significance for glutamate and lysine production (Peters-Wendisch et al., 2001; Riedel et al., 2001). However, the actual function and significance of the isoenzymes and of pyruvate oxidase for growth and amino acid production is not clear.

5.3. L-Lysine biosynthesis and lysine excretion in C. glutamicum

C. glutamicum synthesizes the amino acid L-lysine via a split pathway (Schrumpf et al., 1991), where L-piperideine-2,6-dicarboxylate is converted to the ultimate lysine precursor D,L-diaminopimelate either by a single step reaction catalyzed by D-diaminopimelate dehydrogenase or by the succi-

nylase branch of the pathway, consisting of four reactions catalyzed by the enzymes DapD, DapC, DapE, and DapF. Genetic (Schrumpf et al., 1991) and flux analyses (Sonntag et al., 1993) showed that both branches are used by the bacterium in a contribution of about 30–70%. This ratio varies in a great range depending on cultivation time and other parameters (e.g. availability of ammonium ions).

Since *C. glutamicum* is used for industrial production of L-lysine and possesses, therefore, a high economic value, great efforts were attempted to optimise the biosynthesis with regard to higher efficiencies of producing strains. During the last 25 years nearly all of the genes involved in L-lysine biosynthesis in *C. glutamicum* could be identified (Cremer et al., 1990; Ishino et al., 1987; Kalinowski et al., 1990, 1991; Wehrmann et al., 1994, 1998; Yeh et al., 1988), but the genes for *N*-succinyl-aminoketopimelate aminotransferase (*dapC*) and diaminopimelate epimerase (*dapF*) remained unknown. Based on the complete genome sequence a coding region with high homology to *dapF* genes from different organisms such as *E. coli*, *S. coelicolor*, *M. tuberculosis* could be found. Phenotypic analysis and enzyme activity measurements of a *C. glutamicum* mutant with a deletion of *dapF* resulted in the unequivocal identification of this gene (Hartmann et al., 2003). A *dapC* gene encoding aminoketopimelate aminotransferase is only identified and experimentally verified in *Bordetella pertussis* (Fuchs et al., 2000). Database mining in the *C. glutamicum* genome sequence revealed an ORF (*cg1253*) with significant homology to the *dapC* gene from *Bordetella* and a large number of hypothetical genes from a great variety of organisms (mycobacteria, *E. coli*, *Streptomyces*), which all include an aminotransferase signature. In the meantime, it was shown that a strain with overexpressed *cg1253* revealed a 9-fold increase of specific DapC enzyme activity, proving that *cg1253* is in fact the *dapC* gene (Hartmann et al., 2003). Most surprisingly, even a strain with deleted *dapC* and *ddh* genes could be constructed (Hartmann et al., 2003), although the interruption of both branches of the lysine biosynthesis should be lethal for the cell, since *meso*-diaminopimelate is an essential component for cell wall synthesis. Therefore, another aminotransferase in *C. glutamicum* must exist to substitute for the DapC activity.

L-Lysine uptake in *C. glutamicum* is catalyzed by a lysine/alanine exchange carrier (Bröer and Krämer, 1990) encoded by the *lysI* gene (Seep-Feldhaus et al., 1991). The overall activity of this uptake system, however, is extremely low [about 0.15 nmol min^{-1} (mg dry weight)$^{-1}$] and thus not relevant for L-lysine production.

L-Lysine export is required when *C. glutamicum* cells grow on lysine-containing peptides as a carbon source (Erdmann et al., 1993). Due to the lack of L-lysine-degrading enzymes in *C. glutamicum*, lysine would accumulate in the cytoplasm in the absence of an excretion system. The synthesis of the export carrier is effectively controlled. When the gene coding for the lysine excretion carrier, *lysE*, was isolated (Vrljic et al., 1996), a putative regulatory gene, *lysG*, was identified, localized immediately adjacent to *lysE*. The protein encoded by *lysG* displays all the typical structural features of an autoregulatory transcriptional regulator of the LysR family. Overexpression of *lysEG* in *C. glutamicum* wildtype strain ATCC 13032 resulted in export rates exceeding those of various lysine production strains (Vrljic et al., 1996). These results emphasize the importance of excretion transport systems and their regulatory properties for biotechnological amino acid production.

5.4. L-Threonine biosynthesis and excretion in *C. glutamicum*

L-Threonine synthesis proceeds in three steps, starting from aspartate semialdehyde. The corresponding biosynthesis genes *hom* and *thrB*, which form an operon, were cloned (Peoples et al., 1988), as was *thrC* (Han et al., 1990). The *hom* gene codes for homoserine dehydrogenase, and alleles of this gene, such as HomG378E [*hom*(Fbr)] that code for a dehydrogenase, which is no longer feedback-inhibited by L-threonine, have been identified (Reinscheid et al., 1991). The overexpression of *hom* and *thrB* with high-copy-number plasmids is possible (Eikmanns et al., 1991; Morinaga et al., 1987), whereas *hom*(Fbr) *thrB* can only be expressed at low levels (Reinscheid et

al., 1991). This is due to the resulting high internal L-threonine concentration of up to 100 mM, versus less than 1 mM in the wildtype. Increased internal L-threonine concentration is associated with increased glycine formation (Colón et al., 1995; Reinscheid et al., 1991). Furthermore, the very high internal concentration of L-threonine indicates that its export is limited.

Beside intracellular L-threonine accumulation by overproduction, limited catabolism is another mechanism explaining the biological origin of a threonine excretion system. When threonine-containing peptides are used as nutrients, L-threonine accumulates to physiologically unfavorable concentrations. In this case, an Na^+-coupled secondary export system with considerable activity specific for L-threonine and L-serine is activated (Palmieri et al., 1996; Simic et al., 2001). The corresponding gene, *thrE*, was identified by transposon mutagenesis (Simic et al., 2001). With nine predicted transmembrane segments and homologues found in other species, ThrE, like LysE, is a prototype for a new transporter family.

As a further step in improving L-threonine overproduction in *C. glutamicum*, the reduction in the amount of the by-product glycine was obtained by placing the essential *glyA* gene encoding serine hydroxymethyltransferase (SHMT) in the chromosome under the control of P_{tac}, making *glyA* expression isopropyl β-D-1-thiogalactopyranoside-dependent (Simic et al., 2002). In this way, the SHMT activity in an L-threonine producer was reduced to 8% of the initial activity, which led to a 41% reduction in glycine, while L-threonine was simultaneously increased by 49%. In addition, the intracellular availability of L-threonine to aldol cleavage was also reduced by overexpressing the L-threonine exporter gene *thrE*. In *C. glutamicum* DR-17, which overexpresses *thrE*, accumulation of 67 mM instead of 49 mM L-threonine was obtained (Simic et al., 2002).

5.5. L-Methionine biosynthesis in C. glutamicum

The biosynthesis of L-methionine branches from threonine biosynthesis at the comitted intermediate L-homoserine. The genes responsible for this pathway are described by Rückert et al. (2003). In this work, the *C. glutamicum* genome sequence was analyzed by bioinformatics tools to identify candidate genes. These candidate genes were evaluated by gene deletion experiments and all members of this branched pathway were identified. The identification of all biosynthetic genes prepare the ground for pathway engineering by genetic techniques in order to construct efficient overproducing strains for this biotechnologically very important amino acid, which is nowadays exclusively synthesized by chemical procedures. At present, it is not known which genes are responsible for specifying L-methionine uptake in *C. glutamicum* and whether this organism has the potential to excrete this amino acid into the medium. As in other cases, it is expected that the annotated whole genome in conjunction with post-genome analyses will answer these questions in the foreseeable future.

5.6. L-Isoleucine biosynthesis and transport in C. glutamicum

The genes and the enzymes that synthesize isoleucine from threonine in *C. glutamicum* have been well characterized. The biosynthetic pathway has L-threonine as its precursor and consists of the enzymes IlvA (Möckel et al., 1994), IlvBN, IlvC (Keilhauer et al., 1993), IlvD (Radmacher et al., 2002), and a transaminase encoded by the *ilvE* gene (Radmacher et al., 2002). Isoleucine has been overproduced by introducing excess threonine dehydratase (encoded by *ilvA*) into threonine-producing strains (Colón et al., 1995). Threonine dehydratase is normally feedback inhibited by isoleucine. Mutant derivatives of threonine dehydratase with reduced sensitivity to isoleucine have been an additional dividend in this isoleucine production system (Hashiguchi et al., 1997; Morbach et al., 1995). In a second approach, the *ilvD* and *ilvE* genes have been combined with *ilvBNC* and a deregulated *ilvA* gene to increase L-isoleucine accumulation to 42 mM (Radmacher et al., 2002). Despite these gains, it appears that amino acid export has seriously limited the effectiveness of amino acid production (Kelle et al., 1996).

An uptake system for isoleucine was characterized in *C. glutamicum* (Ebbighausen et al., 1989a) transporting isoleucine by a secondary Na^+-coupled symport mechanism into the cell. The corresponding carrier is encoded by the *brnQ* gene (Tauch et al., 1998). An isoleucine excretion carrier system has already been described biochemically (Ebbighausen et al., 1989b; Hermann and Krämer, 1996). Excretion of isoleucine seems to be advantageous for the bacterial cell when this amino acid accumulates to high cytoplasmic concentrations, because of the putative detrimental effects of amphiphilic isoleucine (Hermann and Krämer, 1996; Eggeling et al., 1997). Isoleucine-producing strains of *C. glutamicum* are characterized by a deregulated biosynthesis pathway (Eggeling et al., 1997) and thus accumulate isoleucine to significant intracellular concentrations.

The system responsible for L-isoleucine export was identified recently by transposon mutagenesis to isolate *C. glutamicum* mutant strains sensitive to the peptide isoleucyl-isoleucine (Kennerknecht et al., 2002). In one such mutant, strong peptide sensitivity resulted from insertion into a gene designated *brnF* encoding a hydrophobic protein predicted to possess seven transmembrane spanning helices. The *brnE* gene is located downstream of *brnF* and encodes a second hydrophobic protein with four putative membrane-spanning helices. A mutant deleted of both genes no longer exports L-isoleucine, whereas an overexpressing strain exports this amino acid at an increased rate (Kennerknecht et al., 2002). The combination of an elevated export carrier with the isoleucine-producing biosynthetic mutants will be the next step in the construction of a competitive industrial production strain.

5.7. L-Valine biosynthesis and transport

Although the amino acid L-valine is not a member of the aspartate family of amino acids, its synthesis is strongly interconnected with that of isoleucine and was, therefore, included into this study. L-Valine originates from two pyruvate molecules that were processed by the same enzymes (IlvBN, IlvC, IlvD, and IlvE) already involved in isoleucine biosynthesis. The gene encoding the second final transaminase (*pdxR*) was discovered recently (McHardy et al., 2003). The genes for the first three enzymes, *ilvBNCD*, overexpressed in an *ilvA* deletion mutant, which is unable to synthesize L-isoleucine, increased the concentration of valine to 58 mM (Radmacher et al., 2002). A further increase was obtained when the *panBC* genes were deleted, making the resulting mutant auxotrophic for D-pantothenate. When *C. glutamicum* 13032*ilvApanBC* with overexpressed *ilvBNCD* genes was grown under limiting conditions, it accumulated 91 mM L-valine. Also in this case, the identification of the excretion system is expected to contribute significantly to the optimization of valine-overproducing strains. The valine excretion system is identical to that of the chemically similar amino acid isoleucine. It has been found that the excretion carrier BrnEF is also responsible for excretion of valine (Radmacher et al., 2002).

5.8. D-Pantothenic acid biosynthesis and transport

Also most of the genes of the D-pantothenate biosynthetic pathway were known before the genome was completely sequenced. Pantothenate synthesis branches from the synthesis of L-valine at the intermediate 2-ketoisovalerate. The specific biosynthesis pathway of this vitamin consists of only four steps. The first reaction, catalyzed by the ketopantoate hydroxymethyltransferase (PanB), uses 2-ketoisovalerate to generate ketopantoate, which is reduced to D-pantoic acid (PanE in *E. coli*; Elischewski et al., 1999). An aspartate-α-decarboxylase activity (PanD) generates β-alanine, which is ligated with pantoic acid to yield D-pantothenate (PanC). The corresponding genes *panBC* (Sahm and Eggeling, 1999) and *panD* (Dusch et al., 1999) have already been identified.

The only unknown gene from the pantothenate biosynthetic pathway was *panE* encoding ketopantoate reductase. In this case, the identification of *panE* in the genome sequence of *C. glutamicum* by similarity analyses provided a surprise: the closest homolog to the *E. coli panE* gene (*cg1239*) turned out not to encode ketopantoate reductase, but the *ilvC* gene could be identified to close the pathway (Merkamm et al., 2003).

In a development of a D-pantothenate overproducing strain *of C. glutamicum*, the *ilvBNC* genes were used together with *panBC* to make D-pantothenate (Sahm and Eggeling, 1999). It is expected, that the recent identification of the genes for the branched-chain transaminases (Radmacher et al., 2002; McHardy et al., 2003) in addition to the hitherto unknown pantothenate transport genes will provide further targets for a metabolic engineering towards D-pantothenate overproduction in *C. glutamicum*.

6. Conclusions

The establishment of a completely annotated *C. glutamicum* genome sequence is a big leap forward to the understanding of the biology of this organism and will boost metabolic engineering to overproduce compounds of biotechnological relevance. It helped to identify missing genes to close the respective biosynthetic pathways directly or by providing a limited number of candidate genes to be tested. The complete genome sequence is the basis for extensive expression analyses by proteome and transcriptome technologies, which will lead to a comprehensive systemic understanding of gene expression and regulatory networks. With the help of the already well-established genetic engineering techniques for this organism (Kirchner and Tauch, 2003), *C. glutamicum* is an ideal candidate for a model system not only representing corynebacteria of biotechnological and of medical importance, but also other important organisms of the so-called high- G+C gram-positives, e.g. the closely related genera *Nocardia*, *Rhodococcus* and *Mycobacterium*.

Acknowledgements

The *C. diphtheriae* sequence data were produced by the *C. diphtheriae* Sequencing Group at the Sanger Institute and obtained from ftp://ftp.sanger.ac.uk/pub/pathogens/cdip/. The work was supported by grant 031U213D of the *Bundesministerium für Bildung und Forschung*.

References

Badger, J.H., Olsen, G.J., 1999. CRITICA: coding region identification tool invoking comparative analysis. Mol. Biol. Evol. 16, 512–524.

Barksdale, L., 1981. The genus *Corynebacterium*. In: Starr, M.P., Stolp, H., Trüper, H.G., Balows, A. (Eds.), The Prokaryotes. Springer, New York, pp. 1827–1837.

Bathe, B., Kalinowski, J., Pühler, A., 1996. A physical and genetic map of the *Corynebacterium glutamicum* ATCC 13032 chromosome. Mol. Gen. Genet. 252, 255–265.

Bellmann, A., Vrljic, M., Patek, M., Sahm, H., Krämer, R., Eggeling, L., 2001. Expression control and specificity of the basic amino acid exporter LysE of *Corynebacterium glutamicum*. Microbiology 147, 1765–1774.

Bröer, S., Krämer, R., 1990. Lysine uptake and exchange in *Corynebacterium glutamicum*. J. Bacteriol. 172, 7241–7248.

Claes, W.A., Pühler, A., Kalinowski, J., 2002. Identification of two *prpDBC* gene clusters in *Corynebacterium glutamicum* and their involvement in propionate degradation via the 2-methylcitrate cycle. J. Bacteriol. 184, 2728–2739.

Colón, G.E., Nguyen, T.T., Jetten, M.S.M., Sinskey, A.J., Stephanopoulos, G., 1995. Production of isoleucine by overexpression of *ilvA* in a *Corynebacterium lactofermentum* threonine producer. Appl. Microbiol. Biotechnol. 43, 482–488.

Cremer, J., Eggeling, L., Sahm, H., 1990. Cloning of the *dapA dapB* cluster of the lysine-secreting bacterium *Corynebacterium glutamicum*. Mol. Gen. Genet. 220, 478–480.

Delcher, A.L., Harmon, D., Kasif, S., White, O., Salzberg, S.L., 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. 27, 4636–4641.

Dominguez, H., Lindley, N.D., 1996. Complete sucrose metabolism requires fructose phosphotransferase activity in *Corynebacterium glutamicum* to ensure phosphorylation in liberated fructose. Appl. Environ. Microbiol. 62, 3878–3880.

Dusch, N., Pühler, A., Kalinowski, J., 1999. Expression of the *Corynebacterium glutamicum panD* gene encoding L-aspartate-α-decarboxylase leads to pantothenate overproduction in *Escherichia coli*. Appl. Environ. Microbiol. 65, 1530–1539.

Ebbighausen, H., Weil, B., Krämer, R., 1989a. Transport of branched-chain amino acids in *Corynebacterium glutamicum*. Arch. Microbiol. 151, 238–244.

Ebbighausen, H., Weil, B., Krämer, R., 1989b. Isoleucine excretion in *Corynebacterium glutamicum*: evidence for a specific efflux carrier system. Appl. Microbiol. Biotechnol. 31, 184–190.

Eggeling, L., Morbach, S., Sahm, H., 1997. The fruits of molecular physiology: engineering the L-isoleucine biosynthesis pathway in *Corynebacterium glutamicum*. J. Biotechnol. 56, 167–182.

Eikmanns, B.J., 1992. Identification, sequence analysis, and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzyme glyceraldehyde-3-

phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase. J. Bacteriol. 174, 6076–6086.

Eikmanns, B.J., Follettie, M.T., Griot, M.U., Sinskey, A.J., 1989. The phosphoenolpyruvate carboxylase gene of *Corynebacterium glutamicum*: molecular cloning, nucleotide sequence, and expression. Mol. Gen. Genet. 218, 330–339.

Eikmanns, B.J., Metzger, M., Reinscheid, D., Kircher, M., Sahm, H., 1991. Amplification of three biosynthesis genes in *Corynebacterium glutamicum* and its influence on carbon flux in different strains. Appl. Microbiol. Biotechnol. 34, 617–622.

Eikmanns, B.J., Thum-Schmitz, N., Eggeling, L., Ludtke, K.U., Sahm, H., 1994. Nucleotide sequence, expression and transcriptional analysis of the *Corynebacterium glutamicum gltA* gene encoding citrate synthase. Microbiology 140, 1817–1828.

Eikmanns, B.J., Rittmann, D., Sahm, H., 1995. Cloning, sequence analysis, expression, and inactivation of the *Corynebacterium glutamicum icd* gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme. J. Bacteriol. 177, 774–782.

Elischewski, F., Pühler, A., Kalinowski, J., 1999. Pantothenate production in *Escherichia coli* K12 by enhanced expression of the *panE* gene encoding ketopantoate reductase. J. Biotechnol. 75, 135–146.

Erdmann, A., Weil, B., Krämer, R., 1993. Regulation of lysine excretion in the lysine producer strain *Corynebacterium glutamicum* MH20-22B. Biotechnol. Lett. 17, 927–932.

Fuchs, T.M., Schneider, B., Krumbach, K., Eggeling, L., Gross, R., 2000. Characterization of a *Bordetella pertussis* diaminopimelate (DAP) biosynthesis locus identifies *dapC*, a novel gene coding for an *N*-succinyl-L, L-DAP aminotransferase. J. Bacteriol. 182, 3626–3631.

Fudou, R., Jojima, Y., Seto, A., Yamada, K., Kimura, E., Nakamatsu, T., Hiraishi, A., Yamanaka, S., 2002. *Corynebacterium efficiens* sp. nov., a glutamic-acid-producing species from soil and vegetables. Int. J. Syst. Evol. Microbiol. 52, 1127–1131.

Fujita, M., Moriya, T., Fujimoto, S., Hara, N., Amoko, K., 1997. A deletion in the *sapA* homologue cluster is responsible for the loss of the S-layer in *Campylobacter fetus* strain TK. Arch. Microbiol. 167, 196–201.

Gabriel, K., Schmid, H., Schmidt, U., Rausch, H., 1995. The actinophage RP3 DNA integrates site-specifically into the putative $tRNA^{Arg}$(AGG) gene of *Streptomyces rimosus*. Nucleic Acids Res. 23, 58–63.

Gourdon, P., Baucher, M.F., Lindley, N.D., Guyonvarch, A., 2000. Cloning of the malic enzyme gene from *Corynebacterium glutamicum* and role of the enzyme in lactate metabolism. Appl. Environ. Microbiol. 66, 2981–2987.

Grigoriev, A., 1998. Analyzing genomes with cumulative skew diagrams. Bioinformatics 14, 252–258.

Han, K.S., Archer, J.A., Sinskey, A.J., 1990. The molecular structure of the *Corynebacterium glutamicum* threonine synthase gene. Mol. Microbiol. 4, 1693–1702.

Hartmann, M., Tauch, A., Eggeling, L., Bathe, B., Möckel, B., Pühler, A., Kalinowski, J., 2003. Identification and characterization of the last two unknown genes, *dapC* and *dapF*, in the succinylase branch of the L-lysine biosynthesis of *Corynebacterium glutamicum*. J. Biotechnol. 104, 199–211.

Hashiguchi, K., Kojima, H., Sato, K., Sano, K., 1997. Effects of an *Escherichia coli ilvA* mutant gene encoding feedback-resistant threonine deaminase on L-isoleucine production by *Brevibacterium flavum*. Biosci. Biotechnol. Biochem. 61, 105–108.

Hatakeyama, K., Kohama, K., Vertès, A.A., Kobayashi, M., Kurusu, Y., Yukawa, H., 1993a. Analysis of the biotin biosynthesis pathway in coryneform bacteria: cloning and sequencing of the *bioB* gene from *Brevibacterium flavum*. DNA Seq. 4, 87–93.

Hatakeyama, K., Kohama, K., Vertès, A.A., Kobayashi, M., Kurusu, Y., Yukawa, H., 1993b. Genomic organization of the biotin biosynthetic genes of coryneform bacteria: cloning and sequencing of the *bioA-bioD* genes from *Brevibacterium flavum*. DNA Seq. 4, 177–184.

Hermann, T., Krämer, R., 1996. Mechanism and regulation of isoleucine excretion in *Corynebacterium glutamicum*. Appl. Environ. Microbiol. 62, 3238–3244.

Hodgson, J., 1998. LION and Degussa apply genomics to fermentation. Nat. Biotechnol. 16, 715.

Hwang, B.-J., Kim, Y., Kim, H.-B., Hwang, H.-J., Kim, J.-H., Lee, H.-S., 1999. Analysis of *Corynebacterium glutamicum* methionine biosynthetic pathway: isolation and analysis of *metB* encoding cystathionine γ-synthase. Mol. Cells 9, 300–308.

Hwang, B.-J., Yeom, H.-J., Kim, Y., Lee, H.-S., 2002. *Corynebacterium glutamicum* utilizes both transsulfuration and direct sulfhydrylation pathways for methionine biosynthesis. J. Bacteriol. 184, 1277–1286.

Ikeda, M., Kamada, N., Takano, Y., Nakano, T., 1999. Molecular analysis of the *Corynebacterium glutamicum* transketolase gene. Biosci. Biotechnol. Biochem. 63, 1806–1810.

Ishino, S., Mizukami, T., Yamaguchi, K., Katsumata, R., Araki, K., 1987. Nucleotide sequence of the *meso*-diaminopimelate D-dehydrogenase gene from *Corynebacterium glutamicum*. Nucleic Acids Res. 15, 3917.

Jetten, M.S., Gubler, M.E., Lee, S.H., Sinskey, A.J., 1994. Structural and functional analysis of pyruvate kinase from *Corynebacterium glutamicum*. Appl. Environ. Microbiol. 60, 2501–2507.

Kalinowski, J., Bachmann, B., Thierbach, G., Pühler, A., 1990. Aspartokinase genes *lysCα* and *lysCβ* overlap and are adjacent to the aspartate β-semialdehyde dehydrogenase gene *asd* in *Corynebacterium glutamicum*. Mol. Gen. Genet. 224, 317–324.

Kalinowski, J., Cremer, J., Bachmann, B., Eggeling, L., Sahm, H., Pühler, A., 1991. Genetic and biochemical analysis of the aspartokinase from *Corynebacterium glutamicum*. Mol. Microbiol. 5, 1197–1204.

Kämpfer, P., Kroppenstedt, R.M., 1996. Numerical analysis of fatty acid patterns of coryneform bacteria and related taxa. Can. J. Microbiol. 42, 989–1005.

Keilhauer, C., Eggeling, L., Sahm, H., 1993. Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. J. Bacteriol. 175, 5595–5603.

Kelle, R., Hermann, T., Weuster-Botz, D., Eggeling, L., Krämer, R., Wandrey, C., 1996. Glucose-controlled L-isoleucine fed-batch production with recombinant strains of *Corynebacterium glutamicum*. J. Biotechnol. 50, 123–136.

Kennerknecht, N., Sahm, H., Yen, M.R., Patek, M., Saier, M.H., Jr, Eggeling, L., 2002. Export of L-isoleucine from *Corynebacterium glutamicum*: a two-gene-encoded member of a new translocator family. J. Bacteriol. 184, 3956–3957.

Kinoshita, S., Udaka, S., Shimono, M., 1957. Studies on the amino acid fermentation. I. Production of L-glutamic acid by various microorganisms. J. Gen. Appl. Microbiol. 3, 193–205.

Kirchner, O., Tauch, A., 2003. Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. J. Biotechnol. 104, 287–299.

Kotrba, P., Inui, M., Yukawa, H., 2001. The *ptsI* gene encoding enzyme I of the phosphotransferase system of *Corynebacterium glutamicum*. Biochem. Biophys. Res. Commun. 289, 1307–1313.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

Lee, J.-K., Sung, M.-H., Yoon, K.-H., Yu, J.-H., Oh, T.-K., 1994. Nucleotide sequence of the gene encoding the *Corynebacterium glutamicum* mannose enzyme II and analyses of the deduced protein sequence. FEMS Microbiol. Lett. 119, 137–145.

Lengeler, J.W., Jahreis, K., Wehmeier, U.F., 1994. Enzymes II of the phosphoenolpyruvate-dependent phosphotransferase systems: their structure and function in carbohydrate transport. Biochim. Biophys. Acta 1188, 1–28.

Leuchtenberger, W., 1996. Amino acids-technical production and use. In: Rehm, H.-J., Reed, G., Pühler, A., Stadler, P. (Eds.), Biotechnology, vol. 6. VCH, Weinheim, Germany, pp. 465–502.

Liebl, W., Ehrmann, M., Ludwig, W., Schleifer, K.H., 1991. Transfer of *Brevibacterium divaricatum* DSM 20297T, '*Brevibacterium flavum*' DSM 20411, '*Brevibacterium lactofermentum*' DSM 20412 and DSM 1412, and *Corynebacterium lilium* DSM 20137T to *Corynebacterium glutamicum* and their distinction by rRNA gene restriction patterns. Int. J. Syst. Bacteriol. 41, 255–260.

Mahillon, J., Chandler, M., 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62, 725–774.

McHardy, A.C., Tauch, A., Rückert, C., Pühler, A., Kalinowski, J., 2003. Genome-based analysis of biosynthetic aminotransferase genes of *Corynebacterium glutamicum*. J. Biotechnol. 104, 229–240.

Merkamm, M., Chassagnole, C., Lindley, N.D., Guyonvarch, A., 2003. Ketopantoate reductase activity is only encoded by *ilvC* in *Corynebacterium glutamicum*. J. Biotechnol. 104, 253–260.

Meyer, F., Goesmann, A., McHardy, A.C., Bartels, D., Bekel, T., Clausen, J., Kalinowski, J., Linke, B., Rupp, O., Giegerich, R., Pühler, A., 2003. GenDB-an open source genome annotation system for prokaryote genomes. Nucleic Acids Res. 31, 2187–2195.

Möckel, B., Eggeling, L., Sahm, H., 1992. Functional and structural analyses of the threonine dehydratase from *Corynebacterium glutamicum*. J. Bacteriol. 174, 8065–8072.

Möckel, B., Eggeling, L., Sahm, H., 1994. Threonine dehydratase of *Corynebacterium glutamicum* with altered allosteric control: their generation and biochemical and structural analysis. Mol. Microbiol. 13, 833–842.

Molenaar, D., van der Rest, M.E., Petrovic, S., 1998. Biochemical and genetic characterization of the membrane-associated malate dehydrogenase (acceptor) from *Corynebacterium glutamicum*. Eur. J. Biochem. 254, 395–403.

Morbach, S., Sahm, H., Eggeling, L., 1995. Use of feedback-resistant threonine dehydratases of *Corynebacterium glutamicum* to increase carbon flux towards L-isoleucine. Appl. Environ. Microbiol. 61, 4315–4320.

Morinaga, Y., Takagi, H., Ishida, M., Miwa, K., Sato, T., Nakamori, S., Sano, K., 1987. Threonine production by coexistence of cloned genes coding homoserine dehydrogenase and homoserine kinase in *Brevibacterium lactofermentum*. Agric. Biol. Chem. 51, 93–100.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1999. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int. J. Neural Syst. 8, 581–599.

O'Regan, M., Thierbach, G., Bachmann, B., Villeval, D., Lepage, P., Viret, J.F., Lemoine, Y., 1989. Cloning and nucleotide sequence of the phosphoenolpyruvate carboxylase-coding gene of *Corynebacterium glutamicum* ATCC 13032. Gene 77, 237–251.

Palmieri, L., Berns, D., Krämer, R., Eikmanns, M., 1996. Threonine diffusion and threonine transport in *Corynebacterium glutamicum* and their role in threonine production. Arch. Microbiol. 165, 48–54.

Parche, S., Burkovski, A., Sprenger, G.A., Weil, B., Krämer, R., Titgemeyer, F., 2001a. *Corynebacterium glutamicum*: a dissection of the PTS. J. Mol. Microbiol. Biotechnol. 3, 423–428.

Parche, S., Thomae, A.W., Schlicht, M., Titgemeyer, F., 2001b. *Corynebacterium diphtheriae*: a PTS view to the genome. J. Mol. Microbiol. Biotechnol. 3, 415–422.

Park, S.-D., Lee, J.-Y., Kim, Y., Kim, J.-H., Lee, H.-S., 1998. Isolation and analysis of *metA*, a methionine biosynthetic gene encoding homoserine acetyltransferase in *Corynebacterium glutamicum*. Mol. Cells 8, 286–294.

Park, S.-Y., Kim, H.-K., Yoo, S.-K., Oh, T.-K., Lee, J.-K., 2000. Characterization of *glk*, a gene coding for glucose kinase of *Corynebacterium glutamicum*. FEMS Microbiol. Lett. 188, 209–215.

Peoples, O.P., Liebl, W., Bodis, M., Maeng, P.J., Follettie, M.T., Archer, J.A., Sinskey, A.J., 1988. Nucleotide sequence and fine structural analysis of the *Corynebacterium glutamicum hom-thrB* operon. Mol. Microbiol. 2, 63–72.

Peyret, J.L., Bayan, N., Joliff, G., Gulik-Krzywicki, T., Mathieu, L., Schechter, E., Leblon, G., 1993. Characterization of the *cspB* gene encoding PS2, an ordered surface-layer protein in *Corynebacterium glutamicum*. Mol. Microbiol. 9, 97–109.

Peters-Wendisch, P.G., Kreutzer, C., Kalinowski, J., Patek, M., Sahm, H., Eikmanns, B.J., 1998. Pyruvate carboxylase from *Corynebacterium glutamicum*: characterization, expression and inactivation of the *pyc* gene. Microbiology 144, 915–927.

Peters-Wendisch, P.G., Schiel, B., Wendisch, V.F., Katsoulidis, E., Möckel, B., Sahm, H., Eikmanns, B.J., 2001. Pyruvate carboxylase is a major bottleneck for glutamate and lysine production by *Corynebacterium glutamicum*. J. Mol. Microbiol. Biotechnol. 3, 295–300.

Quast, K., Bathe, B., Pühler, A., Kalinowski, J., 1999. The *Corynebacterium glutamicum* insertion sequence IS*Cg2* prefers conserved target sequences located adjacent to genes involved in aspartate and glutamate metabolism. Mol. Gen. Genet. 262, 568–578.

Radmacher, E., Vaitsikova, A., Burger, U., Krumbach, K., Sahm, H., Eggeling, L., 2002. Linking central metabolism with increased pathway flux: L-valine accumulation by *Corynebacterium glutamicum*. Appl. Environ. Microbiol. 68, 2246–2250.

Reinscheid, D.J., Eikmanns, B.J., Sahm, H., 1991. Analysis of a *Corynebacterium glutamicum hom* gene coding for a feedback-resistant homoserine dehydrogenase. J. Bacteriol. 173, 3228–3230.

Reinscheid, D.J., Eikmanns, B.J., Sahm, H., 1994a. Characterization of the isocitrate lyase gene from *Corynebacterium glutamicum* and biochemical analysis of the enzyme. J. Bacteriol. 176, 3474–3483.

Reinscheid, D.J., Eikmanns, B.J., Sahm, H., 1994b. Malate synthase from *Corynebacterium glutamicum*: sequence analysis of the gene and biochemical characterization of the enzyme. Microbiology 140, 3099–3108.

Riedel, C., Rittmann, D., Dangel, P., Möckel, B., Petersen, S., Sahm, H., Eikmanns, B.J., 2001. Characterization of the phosphoenolpyruvate carboxykinase gene from *Corynebacterium glutamicum* and significance of the enzyme for growth and amino acid production. J. Mol. Microbiol. Biotechnol. 3, 573–583.

Rossol, I., Pühler, A., 1992. The *Corynebacterium glutamicum aecD* gene encodes a C-S lyase with α, β-elimination activity that degrades aminoethylcysteine. J. Bacteriol. 174, 2968–2977.

Rowlands, R.T., 1984. Industrial strain improvement: mutagenesis and random screening procedures. Enzymes Microb. Technol. 6, 3–10.

Rückert, C., Pühler, A., Kalinowski, J., 2003. Genome-wide analysis of the L-methionine biosynthetic pathway in *Corynebacterium glutamicum* by targeted gene deletion and homologous complementation. J. Biotechnol. 104, 213–228.

Sahm, H., Eggeling, L., 1999. D-Pantothenate synthesis in *Corynebacterium glutamicum* and use of *panBC* and genes encoding L-valine synthesis for D-pantothenate overproduction. Appl. Environ. Microbiol. 65, 1973–1979.

Sahm, H., Eggeling, L., Eikmanns, B., Krämer, R., 1995. Metabolic design in amino acid producing bacterium *Corynebacterium glutamicum*. FEMS Microbiol. Lett. 16, 243–252.

Sahm, H., Eggeling, L., de Graaf, A.A., 2000. Pathway analysis and metabolic engineering in *Corynebacterium glutamicum*. Biol. Chem. 381, 899–910.

Saier, M.H., Jr, Reizer, J., 1992. Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. J. Bacteriol. 174, 1433–1438.

Schäfer, A., Tauch, A., Droste, N., Pühler, A., Kalinowski, J., 1997. The *Corynebacterium glutamicum cglIM* gene encoding a 5-cytosine methyltransferase enzyme confers a specific DNA methylation pattern in an McrBC-deficient *Escherichia coli* strain. Gene 203, 95–101.

Schrumpf, B., Schwarzer, A., Kalinowski, J., Pühler, A., Eggeling, L., Sahm, H., 1991. A functionally split pathway for lysine synthesis in *Corynebacterium glutamicum*. J. Bacteriol. 173, 4510–4516.

Schwinde, J.W., Hertz, P.F., Sahm, H., Eikmanns, B.J., Guyonvarch, A., 2001. Lipoamide dehydrogenase from *Corynebacterium glutamicum*: molecular and physiological analysis of the *lpd* gene and characterization of the enzyme. Microbiology 147, 2223–2231.

Seep-Feldhaus, A.-H., Kalinowski, J., Pühler, A., 1991. Molecular analysis of the *Corynebacterium glutamicum lysI* gene involved in lysine uptake. Mol. Microbiol. 5, 2995–3005.

Shiio, I., Toride, Y., Sugimoto, S., 1984. Production of lysine by pyruvate dehydrogenase mutants of *Brevibacterium flavum*. Agric. Biol. Chem. 48, 3091–3098.

Simic, P., Sahm, H., Eggeling, L., 2001. L-threonine export: use of peptides to identify a new translocator from *Corynebacterium glutamicum*. J. Bacteriol. 183, 5317–5324.

Simic, P., Willuhn, J., Sahm, H., Eggeling, L., 2002. Identification of *glyA* (encoding serine hydroxymethyltransferase) and its use together with the exporter ThrE to increase L-threonine accumulation by *Corynebacterium glutamicum*. Appl. Environ. Microbiol. 68, 3321–3327.

Sonntag, K., Eggeling, L., de Graaf, A.A., Sahm, H., 1993. Flux partitioning in the split pathway of lysine synthesis in *Corynebacterium glutamicum*. Quantification by ^{13}C- and ^{1}H-NMR spectroscopy. Eur. J. Biochem. 213, 1325–1331.

Staden, R., 1996. The Staden sequence analysis package. Mol. Biotechnol. 5, 233–241.

Tauch, A., Hermann, T., Burkovski, A., Krämer, R., Pühler, A., Kalinowski, J., 1998. Isoleucine uptake in *Corynebacterium glutamicum* ATCC 13032 is directed by the *brnQ* gene product. Arch. Microbiol. 169, 303–312.

Tauch, A., Pühler, A., Kalinowski, J., Thierbach, G., 2000. TetZ, a new tetracycline resistance determinant discovered in gram-positive bacteria, shows high homology to gram-negative regulated efflux systems. Plasmid 44, 285–291.

Tauch, A., Homann, I., Mormann, S., Rüberg, S., Billault, A., Bathe, B., Brand, S., Brockmann-Gretza, O., Rückert, C.,

Schischka, N., Wrenger, C., Hoheisel, J., Möckel, B., Huthmacher, K., Pfefferle, W., Pühler, A., Kalinowski, J., 1988. Strategy to sequence the genome of *Corynebacterium glutamicum* ATCC 13032: use of a cosmid and a bacterial artificial chromosome library. J. Biotechnol. 95, 25–38.

Tauch, A., Götker, S., Pühler, A., Kalinowski, J., Thierbach, G., 2002b. The 27.8-kb R-plasmid pTET3 from *Corynebacterium glutamicum* encodes the aminoglycoside adenyltransferase gene cassette *aadA9* and the regulated tetracycline efflux system Tet 33 flanked by active copies of the widespread insertion sequence IS*6100*. Plasmid 48, 117–129.

Usuda, Y., Tujimoto, N., Abe, C., Asakura, Y., Kimura, E., Kawahara, Y., Kurahashi, O., Matsui, H., 1996. Molecular cloning of the *Corynebacterium glutamicum* ('*Brevibacterium lactofermentum*' AJ12036) *odhA* gene encoding a novel type of 2-oxoglutarate dehydrogenase. Microbiology 142, 3347–3354.

Vertès, A.A., Inui, M., Kobayashi, M., Kurusu, Y., Yukawa, H., 1994. Isolation and characterization of IS*31831*, a transposable element from *Corynebacterium glutamicum*. Mol. Microbiol. 11, 739–746.

von der Osten, C.H., Barbas, C.F., III, Wong, C.H., Sinskey, A.J., 1989. Molecular cloning, nucleotide sequence and fine-structural analysis of the *Corynebacterium glutamicum* fda gene: structural comparison of *C. glutamicum* fructose-1,6-biphosphate aldolase to class I and class II aldolases. Mol. Microbiol. 3, 1625–1637.

Wehrmann, A., Eggeling, L., Sahm, H., 1994. Analysis of different DNA fragments of *Corynebacterium glutamicum* complementing *dapE* of *Escherichia coli*. Microbiology 140, 3349–3356.

Wehrmann, A., Phillipp, B., Sahm, H., Eggeling, L., 1998. Different modes of diaminopimelate synthesis and their role in cell wall integrity: a study with *Corynebacterium glutamicum*. J. Bacteriol. 180, 3159–3165.

Wijayarathna, C.D., Wachi, M., Nagai, K., 2001. Isolation of *ftsI* and *murE* genes involved in peptidoglycan synthesis from *Corynebacterium glutamicum*. Appl. Microbiol. Biotechnol. 55, 466–470.

Vrljic, M., Sahm, H., Eggeling, L., 1996. A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. Mol. Microbiol. 22, 815–826.

Yeh, P., Sicard, A.M., Sinskey, A.J., 1988. Nucleotide sequence of the *lysA* gene of *Corynebacterium glutamicum* and possible mechanisms for modulation of its expression. Mol. Gen. Genet. 212, 112–119.