The LacI/GalR family transcriptional regulator UriR negatively controls uridine utilization of Corynebacterium glutamicum by binding to catabolite-responsive element (cre)-like sequences

Karina Brinkrolf, Svenja Plöger, Sandra Solle, Iris Brune, Svenja S. Nentwich, Andrea T. Hüser, Jörn Kalinowski, Alfred Pühler and Andreas Tauch

1Institut für Genomforschung und Systembiologie, Centrum für Biotechnologie, Universität Bielefeld, Universitätsstraße 25, D-33615 Bielefeld, Germany
2International NRW Graduate School in Bioinformatics and Genome Research, Centrum für Biotechnologie, Universität Bielefeld, Universitätsstraße 25, D-33615 Bielefeld, Germany
3Lehrstuhl für Genetik, Fakultät für Biologie, Universität Bielefeld, Universitätsstraße 25, D-33615 Bielefeld, Germany

Received 12 October 2007
Revised 17 December 2007
Accepted 25 December 2007

INTRODUCTION

Corynebacterium glutamicum is a Gram-positive soil bacterium, able to sense and utilize a large variety of carbon and energy sources so as to adapt to constantly changing environmental conditions (Eikmanns, 2005; Shen et al., 2005; Yukawa et al., 2007). This ability is mediated by a diverse metabolism that underlies a complex regulation on the transcriptional level (Brinkrolf et al., 2006, 2007; Han et al., 2007). With knowledge of the whole genome sequence of the wild-type strain C. glutamicum ATCC 13032 (Kalinowski et al., 2003), reconstruction of the topology of this transcriptional regulatory network became possible (Baumbach et al., 2006; Brinkrolf et al., 2007). The most important components in gene-regulatory systems are DNA-binding transcriptional regulators that sense internal or external fluctuations via effector molecules and function to control rates of RNA-polymerase-mediated transcriptional initiation, leaving only relevant genes in an active state (Madan Babu & Teichmann, 2003b; Martínez-Antonio et al., 2006; Resendis-Antonio et al., 2005). An important prerequisite in global regulatory network analysis is the detection of the complete set of transcriptional regulatory networks. Reduced information on the transcriptional level is therefore a prerequisite for a comprehensive and integrative understanding of global regulatory networks.

Abbreviations: CCR, carbon catabolite repression; cre, catabolite-responsive element; HTH, helix–turn–helix; MFS, major facilitator superfamily; PTS, phosphotransferase system; RACE, rapid amplification of cDNA ends.

Two supplementary tables, listing oligonucleotide primers and differential expression of genes studied by whole-genome microarray hybridizations, are available with the online version of this paper.

The Cg1547 protein of Corynebacterium glutamicum ATCC 13032 is a member of the LacI/GalR family of DNA-binding transcriptional regulators. A defined deletion in the cg1547 gene, now designated uriR (uridine utilization regulator), resulted in the mutant strain C. glutamicum KB1547. Comparison of gene expression levels in C. glutamicum KB1547 and the wild-type strain revealed enhanced expression of the uriR operon genes cg1546 (ribokinase), cg1545 (uridine transporter) and cg1543 (uridine-preferring nucleoside hydrolase). Gene expression of the uriR operon was stimulated by the presence of either uridine or ribose. Growth assays with C. glutamicum mutants showed that functional Cg1543 and Cg1545 proteins are essential for the utilization of uridine as the sole carbon source. Transcriptional regulation of the uriR operon is mediated by a 29 bp palindromic sequence composed of two catabolite-responsive element (cre)-like sequences and located in between the mapped -10 promoter region and the start codon of uriR. A similar cre sequence was detected in the upstream region of rbsK2 (cg2554), coding for a second ribokinase in C. glutamicum ATCC 13032. DNA band-shift assays with a streptavidin-tagged UriR protein and labelled oligonucleotides including the cre-like sequences of uriR and rbsK2 demonstrated the specific binding of the purified regulator in vitro. Whole-genome DNA microarray hybridizations comparing the gene expression in C. glutamicum KB1547 with that of the wild-type strain revealed that UriR is a pathway-specific repressor of genes involved in uridine utilization in C. glutamicum.
DNA-binding transcriptional regulators in the organism of interest (Herrgård et al., 2004). For the type strain C. glutamicum ATCC 13032, this regulatory repertoire has been determined recently, leading to a set of 158 proteins that presumably act in regulation of gene expression (Brinkrolf et al., 2007; Brune et al., 2005). According to amino acid sequence comparisons and protein structure predictions, the subset of 128 DNA-binding transcriptional regulators were grouped into 24 different protein families. One of these groups is the LacI/GalR family of transcriptional regulators, including seven functional members in C. glutamicum ATCC 13032 (Brune et al., 2005). LacI/GalR family proteins consist of an amino-terminal helix–turn–helix (HTH) motif and a carboxy-terminal domain for oligomerization and effector binding, the latter having a three-dimensional structure similar to those of sugar-specific periplasmic binding proteins (Fukami-Kobayashi et al., 2003; Weickert & Adhya, 1992). Already characterized members of this family in a diverse range of bacteria revealed that these regulators mainly act as repressor proteins of genes involved in carbohydrate metabolism (Weickert & Adhya, 1992). LacI/GalR family proteins therefore may provide key players for the variability of C. glutamicum in the utilization of different carbon and energy sources.

Many micro-organisms have evolved an economical mechanism for repressing the utilization of alternative carbon sources in the presence of glucose and other rapidly metabolizable carbon sources (Brückner & Tingemeyer, 2002; Tingemeyer & Hillen, 2002). This mechanism, designated carbon catabolite repression (CCR), is well understood in Gram-negative bacteria like Escherichia coli, where it is mediated by the phosphorylation state of the phosphotransferase system (PTS) protein II^Gk, the signal molecule cAMP and the transcriptional regulator Crc. In Gram-positive bacteria with low G+C content, CCR is carried out by a completely different mechanism, involving the PTS protein HPr, the HPr kinase/phosphatase and the catabolite control protein CcpA, resulting in global changes of gene expression (Titgemeyer & Hillen, 2002). CcpA-like proteins form a subfamily within the LacI/GalR family of transcriptional regulators. In dependence on the availability of certain low-molecular-mass effectors, CcpA-like proteins recognize a cis-acting operator sequence, termed the catabolite-responsive element (cre) (Kraus et al., 1998; Weickert & Adhya, 1992). A remarkable feature of C. glutamicum, which belongs to the class of high-G+C Gram-positive bacteria, is the ability to co-metabolize different sugars and organic acids in the presence of glucose (Eikmanns, 2005; Wendisch et al., 2000). Hence only a few cases of CCR in this organism have been reported so far, including the repression of glutamate uptake in the presence of glucose or fructose (Parche et al., 2001), the repression of the gluconate genes by glucose, fructose and sucrose (Letek et al., 2006), and the recently published repression of the alcohol dehydrogenase gene in the presence of glucose or acetate (Arndt & Eikmanns, 2007).

In this study, we characterized the LacI/GalR family transcriptional regulator Cg1547 (UriR) of C. glutamicum ATCC 13032 to integrate it by experimental evidence into the constantly growing gene regulatory network of this species (Brinkrolf et al., 2007). The Cg1547 protein turned out to recognize cre-like DNA sequences in the chromosome of C. glutamicum ATCC 13032 to negatively control the expression of genes involved in the utilization of uridine as the sole carbon and energy source, defining Cg1547 as a pathway-specific transcriptional regulator of the LacI/GalR family.

**METHODS**

**Bacterial strains, plasmids and growth conditions.** C. glutamicum strains and plasmids used and constructed in this study are listed in Table 1. E. coli DH5α/MCR (Grant et al., 1990) was grown at 37 °C in Luria–Bertani (LB) broth (Sambrook et al., 1989) or on LB agar and was used for standard cloning procedures. E. coli strains carrying plasmids were selectively grown with kanamycin (50 µg ml⁻¹) or ampicillin (200 µg ml⁻¹). Induction of gene expression on pASK-IBA5+ (Table 1) was carried out by using 200 ng tetracycline ml⁻¹. C. glutamicum strains were grown at 30 °C in CGXII minimal medium (Keilhauer et al., 1993) containing 30 µg protocatechuc acid l⁻¹ and 420 µg thiamine l⁻¹ on LB agar plates. C. glutamicum strains carrying plasmids were selectively grown with kanamycin (25 µg ml⁻¹). Growth of C. glutamicum strains in CGXII minimal medium supplemented with different carbon sources was monitored at intervals of 1 h with a Nephelostar Galaxy nephelometer (BMG Laboratories).

**DNA isolation, manipulation and transfer.** Preparation of plasmid DNA from E. coli cells by an alkaline lysis technique was performed using the QiAprep Spin Miniprep kit (Qiagen). Chromosomal C. glutamicum DNA was prepared as previously described (Tauch et al., 1995). Modification of DNA, analysis by agarose gel electrophoresis and ligation were performed by standard procedures (Sambrook et al., 1989). Transformation of E. coli and C. glutamicum cells was performed by electroporation (Tauch et al., 1994, 2002).

**PCR methods and oligonucleotides.** DNA amplification by PCR was carried out with Sahara Taq DNA polymerase (Bioline) or Pwo DNA polymerase (Roche Diagnostics) and chromosomal DNA as template in a PTC-100 thermocycler (MJ Research). PCR products were purified with the QIAquick PCR Purification kit (Qiagen). All oligonucleotides used in this study were purchased from Operon Biotechnologies; they are listed in Supplementary Table S1, available with the online version of this paper.

**Construction of defined mutant strains in C. glutamicum.** Defined deletions were established within the C. glutamicum genes cg1547, cg1546, cg1545, cg1543 and cg2554 by means of the gene SOEing procedure (Horton et al., 1989). The resulting pK18mobsacB derivatives (Table 1) were suitable for performing allelic exchanges in the chromosome of C. glutamicum (Schafer et al., 1994). For genetic complementation of the cg1547 deletion mutant C. glutamicum KB1547, the complete coding region including the native promoter of cg1547 was amplified by PCR and cloned in E. coli into the shuttle vector pEC-K18mobs2 (Table 1).

**Preparation of total RNA and identification of differential gene expression.** C. glutamicum strains were grown in CGXII minimal medium (Keilhauer et al., 1993) to purify total RNA as described previously (Brune et al., 2007). Real-time reverse transcription
The resulting plasmid pASK-IBA5 was purified from the protein crude extract with a Strep-Tactin protease inhibitors; pH 8.0). The cells were disrupted twice in a French centrifugation and resuspended in 10 ml buffer W (1 M Tris/HCl, 10 m\text{M} \text{KCl}, 5 \text{mM} \text{MgCl}_2, 1 \text{mM} \text{DL-ribose, ribose 5-phosphate, uridine and uracil), to a concentration of 10^9 \text{mg} \text{ml}^{-1} \text{E. coli K18 plasmid vector IBA Tagnologies).}

**Table 1. Bacterial strains and plasmids used in this study**

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. glutamicum ATCC 13032</td>
<td>Wild-type strain</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>KB1547</td>
<td>ATCC 13032 with defined deletion in cg1547</td>
<td>This study</td>
</tr>
<tr>
<td>KB1546</td>
<td>ATCC 13032 with defined deletion in cg1546</td>
<td>This study</td>
</tr>
<tr>
<td>KB1545</td>
<td>ATCC 13032 with defined deletion in cg1545</td>
<td>This study</td>
</tr>
<tr>
<td>KB1543</td>
<td>ATCC 13032 with defined deletion in cg1543</td>
<td>This study</td>
</tr>
<tr>
<td>KB2554</td>
<td>ATCC 13032 with defined deletion in cg2554</td>
<td>This study</td>
</tr>
<tr>
<td>RS8K12</td>
<td>ATCC 13032 with defined deletions in cg1546 and cg2554</td>
<td>This study</td>
</tr>
</tbody>
</table>

**Plasmids**

<table>
<thead>
<tr>
<th>plasmid</th>
<th>Relevant characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pK18mobacB</td>
<td>sacB lacZs Km\text{\textsuperscript{\text{\textdagger}}} \text{\textsuperscript{\textdagger}}; cloning vector for allelic exchange</td>
<td>Schäfer et al. (1994)</td>
</tr>
<tr>
<td>pK18mobacB_A1547</td>
<td>pK18mobacB carrying a modified cg1547 gene region</td>
<td>This study</td>
</tr>
<tr>
<td>pK18mobacB_A1546</td>
<td>pK18mobacB carrying a modified cg1546 gene region</td>
<td>This study</td>
</tr>
<tr>
<td>pK18mobacB_A1545</td>
<td>pK18mobacB carrying a modified cg1545 gene region</td>
<td>This study</td>
</tr>
<tr>
<td>pK18mobacB_A1543</td>
<td>pK18mobacB carrying a modified cg1543 gene region</td>
<td>This study</td>
</tr>
<tr>
<td>pK18mobacB_A2554</td>
<td>pK18mobacB carrying a modified cg2554 gene region</td>
<td>This study</td>
</tr>
<tr>
<td>pEC-K18mob2</td>
<td>lacZs Km\text{\textsuperscript{\text{\textdagger}}}; C. glutamicum cloning vector</td>
<td>Kirchner &amp; Tauch (2003)</td>
</tr>
<tr>
<td>pEC-K18mob2 cg1547</td>
<td>pEC-K18mob2 carrying the C. glutamicum cg1547 gene</td>
<td>This study</td>
</tr>
<tr>
<td>pASK-IBA5+</td>
<td>P\text{\textsuperscript{\text{\textdagger}}} \text{\textsuperscript{\textdagger}}; E. coli expression vector</td>
<td>IBA Tagnologies</td>
</tr>
<tr>
<td>pASK-IBA5+ cg1547</td>
<td>pASK-IBA5+ carrying the C. glutamicum cg1547 gene</td>
<td>This study</td>
</tr>
</tbody>
</table>

**Mapping of transcriptional start sites by RACE-PCR.** For determining transcriptional start sites, total RNA was isolated from C. glutamicum KB1547 cultures grown in LB medium (Sambrook et al., 1989). Specific RACE-PCR primers (Table S1) and 1 \text{mg} total RNA were used for cDNA synthesis. The experimental set-up followed the protocol of the 5'/3' RACE kit, second generation (Roche Diagnostics).

**Construction, expression and purification of a streptavidin-tagged Cg1547 protein.** The cg1547 gene was amplified by means of PCR primers constructed with the PrimerDesigner software (IBA Tagnologies) to generate a Cg1547 protein with an N-terminal streptavidin (Strep) tag (Cg1547\textsubscript{\text{\text{\textdagger}}} \text{\textsuperscript{\textdagger}}). The respective PCR product was digested with BsaI and cloned into the vector pASK-IBA5+ (Table 1). The resulting plasmid pASK-IBA5+ cg1547 was transferred to E. coli DH5\text{\text{\textdagger}} \text{\text{\textdagger}} \text{\text{\textdagger}} MCR cells by electroporation. The corresponding E. coli strain was grown for 16 h at 24 °C in selective LB medium containing 200 \text{\textmu}g ampicillin ml\textsuperscript{-1}. Expression of the recombinant Cg1547_N protein was induced with 200 ng tetracycline ml\textsuperscript{-1} a-g ampicillin ml\textsuperscript{-1} E. coli expression vector IBA Tagnologies).

**Hybridization of whole-genome DNA microarrays.** For the amplification of cDNA, 8 \text{\textmu}g total RNA from C. glutamicum ATCC 13032 and C. glutamicum KB1547 cultures was used. Labelling of probes and hybridization of the C. glutamicum whole-genome DNA microarray were carried out as described previously (Brune et al., 2006). DNA microarray hybridization was performed in duplicate, using label swapping. Each DNA microarray contains four replicates per C. glutamicum gene, making a total of eight spots per gene available for calculating differential gene expression. During image acquisition and data processing, mean signal and mean local background intensities were determined for each spot of the microarray images by using the ImageGene 6.0 software (BioDiscovery) for spot detection, image segmentation and signal quantification. After subtraction of local background intensities from the signal intensities, the log\textsubscript{2} value of the ratio of intensities was calculated for each spot according to the formula M=log\textsubscript{2}(R/G). In particular, R=I_{\text{his}}-B_{\text{his}} and G=I_{\text{abs}}-B_{\text{abs}}, where I_{\text{his}} or I_{\text{abs}} is the intensity of a spot in channel 1 or channel 2 and B_{\text{his}} or B_{\text{abs}} is

Electrophoretic mobility shift assays with recombinant Cg1547 protein. DNA band shifts were performed with Cy3-labelled oligonucleotides (Table S1) that were annealed with the corresponding complementary oligonucleotides to double-stranded DNA fragments by heating for 5 min at 94 °C and annealing at room temperature for 15 min. Binding assays were performed in a final volume of 15 \text{\textmu}l containing 30 pmol Strep-tagged Cg1547_N protein mixed with 0.04–0.05 pmol labelled DNA fragments, 20%, v/v, glycerol, herring sperm DNA (0.06 \text{\mu}g) and binding buffer (20 mM Tris/HCl, 5% (v/v) glycerol, 1 mM MgCl\textsubscript{2}, 40 mM KCl, 1 mM DL-1,4-dithiothreitol; pH 8.3). The binding buffer was modified where appropriate by adding rising amounts of potential effector (glucose, ribose, ribose 5-phosphate, uridine and uracil), to a concentration of up to 20 mM, to the shift assay. Shift assays were incubated at room temperature for 20 min and were subsequently separated in a 2% agarose gel prepared in gel buffer (20 mM Na\textsubscript{2}HPO\textsubscript{4}, pH 7.0). A voltage of 70 V was supplied for 1 h. The agarose gel was scanned with a Typhoon 8600 Variable Mode imager (Amersham Biosciences Europe). Dose-dependent binding of the purified Cg1547_N protein was tested in the range from 5 to 50 pmol of the Strep-tagged protein, leading to optimal protein–DNA interactions with concentrations greater than 30 pmol.

**Determination of transcriptional start sites by RACE-PCR.** For determining transcriptional start sites, total RNA was isolated from C. glutamicum KB1547 cultures grown in LB medium (Sambrook et al., 1989). Specific RACE-PCR primers (Table S1) and 1 \text{\textmu}g total RNA were used for cDNA synthesis. The experimental set-up followed the protocol of the 5'/3' RACE kit, second generation (Roche Diagnostics).

**Construction, expression and purification of a streptavidin-tagged Cg1547 protein.** The cg1547 gene was amplified by means of PCR primers constructed with the PrimerDesigner software (IBA Tagnologies) to generate a Cg1547 protein with an N-terminal streptavidin (Strep) tag (Cg1547\textsubscript{\text{\textdagger}}} \text{\textsuperscript{\textdagger}}). The respective PCR product was digested with BsaI and cloned into the vector pASK-IBA5+ (Table 1). The resulting plasmid pASK-IBA5+ cg1547 was transferred to E. coli DH5\text{\text{\textdagger}}} \text{\text{\textdagger}} \text{\text{\textdagger}} MCR cells by electroporation. The corresponding E. coli strain was grown for 16 h at 24 °C in selective LB medium containing 200 \text{\textmu}g ampicillin ml\textsuperscript{-1}. Expression of the recombinant Cg1547_N protein was induced with 200 ng tetracycline ml\textsuperscript{-1} a-g ampicillin ml\textsuperscript{-1} E. coli expression vector IBA Tagnologies).

**Hybridization of whole-genome DNA microarrays.** For the amplification of cDNA, 8 \text{\textmu}g total RNA from C. glutamicum ATCC 13032 and C. glutamicum KB1547 cultures was used. Labelling of probes and hybridization of the C. glutamicum whole-genome DNA microarray were carried out as described previously (Brune et al., 2006). DNA microarray hybridization was performed in duplicate, using label swapping. Each DNA microarray contains four replicates per C. glutamicum gene, making a total of eight spots per gene available for calculating differential gene expression. During image acquisition and data processing, mean signal and mean local background intensities were determined for each spot of the microarray images by using the ImageGene 6.0 software (BioDiscovery) for spot detection, image segmentation and signal quantification. After subtraction of local background intensities from the signal intensities, the log\textsubscript{2} value of the ratio of intensities was calculated for each spot according to the formula M=log\textsubscript{2}(R/G). In particular, R=I_{\text{his}}-B_{\text{his}} and G=I_{\text{abs}}-B_{\text{abs}}, where I_{\text{his}} or I_{\text{abs}} is the intensity of a spot in channel 1 or channel 2 and B_{\text{his}} or B_{\text{abs}} is
the background intensity of a spot in channel 1 or channel 2, respectively. The average intensity in both channels was determined for each spot according to the formula \( A_i = \log_2(R_{ij}) \). Normalization and evaluation of the hybridization data were accomplished by the EMMA 2.2 microarray data analysis software (Dondrup et al., 2003), using a signal intensity ratio (M-value) cut-off of ±1, which corresponds to relative expression changes equal to or greater than 2-fold, applying an error probability (P-value) of less than 5% for the Student’s t test.

RESULTS

The regulatory protein Cg1547 is a member of the LacI/GalR family of transcriptional regulators

The proposed regulatory protein Cg1547 of *C. glutamicum* ATCC 13032 consists of 346 amino acids and has an estimated molecular mass of 36.6 kDa (Brune et al., 2005). Protein structure predictions using the SUPERFAMILY server (Madera et al., 2004) identified an amino-terminal HTH domain of the \( \lambda \) repressor-like superfamily. This DNA-binding domain comprises amino acids 10–29 of the protein and represents a LacI-type HTH domain signature (PROSITE documentation PS00356) (Hulo et al., 2006). The carboxy-terminal domain of Cg1547 is homologous to the superfamily of periplasmic binding proteins, especially to the L-arabinose-binding-protein-like family (Madera et al., 2004). This composition of an N-terminal HTH domain and a C-terminal substrate-binding domain is typical for LacI/GalR family proteins (Weickert & Adhya, 1992) and was apparently acquired by an early ancestral family member, when a substrate-binding protein fused to an HTH DNA-binding protein (Fukami-Kobayashi et al., 2003).

The transcriptional regulator Cg1547 is encoded by the cg1547 gene in the chromosome of *C. glutamicum* ATCC 13032 (Fig. 1a). cg1547 is the first gene in a cluster of four genes, including cg1546 (rbsK1) encoding a putative ribokinase (Kalinowski et al., 2003), cg1543 encoding a transport protein of the major facilitator superfamily (Winnen et al., 2005) and cg1543 encoding a putative nucleoside hydrolase (Kalinowski et al., 2003). The genes located downstream of the cg1547 cluster encode a putative membrane protein (cg1542) and a hypothetical protein (cg1541). These genes are separated from the cg1547 cluster by a rho-independent transcriptional terminator (Fig. 1a).

On the opposite DNA strand, the cg1547 gene cluster is flanked by the cg1548 gene, which encodes a conserved hypothetical protein (Fig. 1a).

Deletion of cg1547 results in enhanced expression of the downstream genes cg1546, cg1545 and cg1543

The most promising way to characterize the functional role of a putative transcriptional regulator is to compare relevant gene expression levels in a knockout mutant with those in a wild-type control. Therefore, a defined deletion mutant was constructed by an allelic exchange procedure (Scharer et al., 1994), resulting in the mutant strain *C. glutamicum* KB1547, which has an in-frame deletion of 558 bp within the cg1547 coding region. Differential expression of the neighbouring genes cg1548, cg1546, cg1543 and cg1543 was measured by real-time RT-PCR assays, comparing the expression levels in *C. glutamicum* KB1547 with those in the wild-type control, *C. glutamicum* ATCC 13032.

![Fig. 1. The cg1547 gene region of *C. glutamicum* ATCC 13032. (a) Genetic organization of the cg1547 gene region and proposed physiological functions of the encoded proteins. The hairpin structure downstream of the cg1543 gene denotes a transcriptional terminator. (b) Detailed view of the cg1547 upstream region. The deduced promoter regions (−35 and −10) are indicated. The transcriptional start site (+1) is marked with an arrow, a putative ribosome-binding site (RBS) is depicted in bold, and the translational start codon (ATG) of cg1547 is underlined. The 29 bp palindromic sequence is marked by arrows; the cre-like sequences are boxed. (c) The cre-like sequences in front of cg1547. The nucleotide sequences of the cre sites are compared with the cre consensus sequence of *Bacillus subtilis* (Weickert & Chambliss, 1990) and a cre site from *Lactobacillus pentosus* (Mahr et al., 2000). Letters in bold type indicate conserved residues.](http://mic.sgmjournals.org)
ATCC 13032. For this assay, total RNA was purified from two exponentially growing cultures of each strain, cultivated in CGXII minimal medium with 40 g glucose l\(^{-1}\) as the sole carbon source. This comparative approach on the transcriptional level revealed an enhanced expression of cg1546 (34.5-fold), cg1545 (34.4-fold) and cg1543 (16.2-fold) in C. glutamicum KB1547, whereas no differential expression of cg1548 was detected.

The apparent deregulation of gene expression in C. glutamicum KB1547 was complemented by electrotransformation with plasmid pEC-K18mob2, carrying a functional copy of the cg1547 gene of the wild-type strain (Table 1). Real-time RT-PCR experiments were carried out, analysing the gene expression of cg1546, cg1545 and cg1543 in the plasmid-carrying mutant C. glutamicum KB1547. C. glutamicum ATCC 13032 and C. glutamicum KB1547, both carrying the empty cloning vector pEC-K18mob2 (Table 1), served as controls. The cg1546, cg1545 and cg1543 genes showed no differential expression in the complemented derivative of C. glutamicum KB1547 when compared with the wild-type control (data not shown). These data confirmed that the deletion of cg1547 was responsible for the enhanced expression of cg1546, cg1545 and cg1543 in C. glutamicum KB1547. The Cg1547 protein can therefore be regarded as a negative transcriptional regulator of the downstream genes cg1546, cg1545 and cg1543 (Fig. 1). The physiological role of these genes in C. glutamicum was investigated by generating targeted gene deletions.

The cg1543 gene, encoding a nucleoside hydrolase, is essential for the utilization of uridine as the sole carbon source

The Cg1543 protein consists of 312 amino acids and has a calculated molecular mass of 33.9 kDa (Kalinowski et al., 2003). Amino acid sequence comparison with the experimentally characterized nucleoside hydrolase Rih1 from Corynebacterium ammoniagenes (Kim et al., 2006) revealed 48% sequence identity between the two proteins. The general function of nucleoside hydrolases is to catalyse the irreversible hydrolysis of ribose and free nucleic bases (Miles et al., 1999). To test the ability of C. glutamicum to metabolize different nucleosides as the sole carbon source for growth, the wild-type strain ATCC 13032 was grown in CGXII minimal medium supplemented either with 15 mM of adenosine, thymidine, cytidine, uridine or inosine, or with 3 mM of guanosine (due to its lower solubility). Growth assays with 15 mM glucose and without any supplementation of carbon served as positive and negative controls, respectively. Growth of the C. glutamicum cultures was automatically monitored by nephelometry (Fig. 2a). These assays clearly demonstrated that C. glutamicum can utilize uridine as the sole carbon source, while there was no significant growth with adenosine, thymidine, guanosine, cytidine or inosine when compared to the control assays. C. glutamicum ATCC 13032 cultures grew with a doubling time of 2.6 h in CGXII minimal medium supplemented with up to 50 mM uridine as the sole carbon source, whereas higher concentrations of uridine progressively reduced doubling times and biomass production. To verify whether the ability to degrade uridine is dependent on the enzymic function of Cg1543, 295 bp within the coding region of cg1543 were deleted by an allelic-exchange procedure, resulting in the deletion mutant C. glutamicum KB1543. This mutant strain was subsequently used in a growth assay with CGXII minimal medium either supplemented with 50 mM of glucose, ribose or uridine, or without any supplementation of carbon as a control. By carrying the defined deletion in the cg1543 gene, C. glutamicum KB1543 was unable to grow in CGXII minimal medium supplemented with uridine (Fig. 2b). Therefore, a functional Cg1543 protein is essential for the utilization of uridine as the sole carbon source, suggesting that uridine-preferring nucleoside hydrolase. The cg1543 gene of C. glutamicum ATCC 13032 was therefore designated uriH (uridine-preferring nucleoside hydrolase).

The cg1545 gene, encoding a transporter of the major facilitator superfamily, is also essential for the utilization of uridine as the sole carbon source

The predicted transport protein Cg1545 of C. glutamicum ATCC 13032 consists of 483 amino acids and has a calculated molecular mass of 50.1 kDa (Kalinowski et al., 2003). A BLASTP search against the NCBI database revealed numerous homologies to permeases of the major facilitator superfamily. The Cg1545 protein possesses 14 transmembrane helices and is therefore classified into the DHA2 (14-spanner) subgroup (Winnen et al., 2005). For the analysis of the cg1545 gene, the mutant strain C. glutamicum KB1545 was constructed by introducing an in-frame deletion of 1305 bp within the coding region of cg1545. Growth of the mutant strain was compared to that of C. glutamicum ATCC 13032 under the same conditions as described above. C. glutamicum KB1545 was unable to grow with uridine as the sole carbon source (Fig. 2b), as observed for the deletion of the cg1543 gene. A functional copy of the cg1545 gene is thus essential for the utilization of uridine as the sole carbon source, suggesting that Cg1545 might be involved in the transport of uridine into the C. glutamicum cell. The cg1545 gene was preliminarily named uriT (uridine transporter).

The cg1546 gene, encoding a ribokinase, is dispensable for the utilization of uridine as the sole carbon source

The last gene to be characterized within the cg1547 gene cluster was cg1546 (rbsK1), annotated to code for a ribokinase, generally catalysing the conversion of D-ribose into D-ribose 5-phosphate (Sigrell et al., 1997). The
Fig. 2. Growth of *C. glutamicum* in minimal medium supplemented with different nucleosides and sugars as carbon sources. (a) Growth of the wild-type *C. glutamicum* ATCC 13032 in CGXII minimal medium supplemented with 15 mM of glucose (Glc), uridine (Uri) or inosine (Ino). A non-supplemented culture served as control (Cont). Growth with adenosine, thymidine, guanosine and cytidine was similar to that with inosine and is not shown in the graph for clarity. (b) Growth assays of the wild-type *C. glutamicum* ATCC 13032 and the derived deletion mutants *C. glutamicum* KB1543, KB1545, KB1546, KB2554 and RBSK12. Cultures were grown in CGXII minimal medium supplemented with 30 mM of glucose (Glc), ribose (Rib) or uridine (Uri). All growth curves were recorded using a nephelometer and five parallel cultures. Data are given as relative turbidity.
corresponding protein comprises 307 amino acids, and has a molecular mass of 30.8 kDa. In a first approach, the function of RbsK1 in uridine utilization was tested for \textit{C. glutamicum} by constructing a mutant strain with a defined in-frame deletion of 768 bp within the cg1546 coding region, resulting in the mutant strain \textit{C. glutamicum} KB1546. This mutant was used in a growth assay in CGXII minimal medium either supplemented with 50 mM of glucose, ribose or uridine, or without any supplementation of carbon as a control. No significant effect on the growth behaviour of \textit{C. glutamicum} KB1546 was detected when compared to the wild-type control (Fig. 2b). The RbsK1 protein is apparently dispensable for the utilization of uridine and ribose as the sole carbon sources for growth. This experimental result can be explained by the presence of a second ribokinase protein that is homologous with Cg1545 and encoded by cg2554 (rbsK2) in the genome sequence of \textit{C. glutamicum} ATCC 13032 (Kalinowski et al., 2003). A pairwise amino acid sequence comparison between RbsK1 and RbsK2 revealed 57% identical amino acids.

To clarify the functional role of the ribokinase RbsK2 in uridine utilization, the single-knockout mutant \textit{C. glutamicum} KB2554 and the ribokinase double mutant \textit{C. glutamicum} RBSK12 were constructed by introducing a defined deletion into the coding region of cg2554 in the wild-type strain and in \textit{C. glutamicum} KB1546, respectively. Both mutants were analysed in growth assays using CGXII minimal medium supplemented with 50 mM of glucose, ribose or uridine, and without any supplementation of sugar as a control. The single deletion of the \textit{rbsK2} gene in \textit{C. glutamicum} KB2554 caused no obvious growth deficiencies (Fig. 2b). On the other hand, the \textit{rbsK1/rbsK2} double mutant \textit{C. glutamicum} RBSK12 was unable to grow with uridine or ribose as the sole carbon source (Fig. 2b). These results provided evidence that at least one functional ribokinase is necessary for the utilization of uridine or ribose as carbon sources.

**Differential expression of cg1547, cg1546, cg1545 and cg1543 is stimulated by uridine and ribose**

To monitor the effect of exogenous carbon sources on the expression level of the cg1547 gene cluster, \textit{C. glutamicum} ATCC 13032 was grown in minimal medium supplemented with 30 mM of glucose, ribose or uridine. Cells were harvested at mid-exponential growth phase, total RNA was isolated from the samples, and gene expression levels were measured by real-time RT-PCR. The comparison of gene expression levels between samples obtained from uridine- and glucose-grown cultures showed an enhanced expression of cg1547 (17.7-fold), cg1546 (9.8-fold), cg1545 (8.7-fold) and cg1543 (9.8-fold) in the presence of uridine. An even higher effect on the expression level of cg1547 (73.4-fold), cg1546 (36.6-fold), cg1545 (34.8-fold) and cg1543 (26-fold) was measured in the presence of ribose, when comparing data generated from RNA samples obtained from ribose- and glucose-grown cultures. To investigate a potential effect of uracil on the gene expression of the cg1547 cluster, the \textit{C. glutamicum} wild-type was grown with a mixture of 15 mM glucose and 15 mM uracil, since \textit{C. glutamicum} is unable to utilize uracil as the sole carbon source. This assay revealed no differential expression of the cg1547 gene cluster (data not shown). These data therefore demonstrated that the expression of the cg1547 gene cluster is significantly stimulated by the presence of uridine or ribose in the growth medium. The same set of experiments was performed with the cg1547 mutant strain \textit{C. glutamicum} KB1547 (Table 1).

When comparing ribose-, uridine- and uracil-grown \textit{C. glutamicum} KB1547 cultures with a glucose-grown control, no significant differences in gene expression levels were observed (data not shown). Due to the lack of a functional Cg1547 protein, the cg1547 gene cluster is apparently expressed at the same level in \textit{C. glutamicum} KB1547 irrespective of the carbon source tested.

We then addressed the question whether this stimulation of differential gene expression is dependent on the function of the cg1545 (\textit{uriT}) and cg1543 (\textit{uriH}) genes. To elucidate this question, the expression of cg1547 was measured in \textit{C. glutamicum} ATCC 13032 and the mutant strains \textit{C. glutamicum} KB1545 and KB1543 grown in CGXII medium with either 30 mM glucose or a mixture of 15 mM glucose and 15 mM uridine. The latter medium was used since \textit{C. glutamicum} KB1545 and \textit{C. glutamicum} KB1543 are unable to grow in the presence of uridine as the sole carbon source (Fig. 2b). The expression of cg1547 in the \textit{C. glutamicum} wild-type was enhanced 9.9-fold in cultures grown in the presence of both uridine and glucose. \textit{C. glutamicum} ATCC 13032 grown with glucose served as control during these assays. In the case of \textit{C. glutamicum} KB1545 and KB1543, RT-PCR experiments revealed no significant deregulation of cg1547 expression during growth with uridine and glucose when compared with a glucose control (data not shown). On the other hand, stimulation of gene expression was observed in \textit{C. glutamicum} KB1545 and KB1543 cultures grown in a mixture of glucose and ribose, indicating that functional \textit{uriT} and \textit{uriH} genes are dispensable for the utilization of ribose as carbon source. These results clearly showed that the stimulation of cg1547 gene expression by uridine occurs as long as the \textit{C. glutamicum} strain possesses functional \textit{uriT} and \textit{uriH} genes. Stimulation of differential gene expression is therefore directly linked to the functionality of both the major facilitator superfamily (MFS) permease and the nucleoside hydrolase. This implies that the stimulus of the cg1547 gene cluster is not primarily given by uridine, but by a degradation product of this nucleoside.

The stimulation of cg1547 gene expression was also measured in the ribokinase double mutant \textit{C. glutamicum} RBSK12, grown in CGXII minimal medium with varying carbon sources as described above. RNA samples were analysed by RT-PCR, measuring the relative expression of
the cg1547 gene. Expression of cg1547 was elevated 7.1-fold in C. glutamicum RBSK12 grown with glucose and uridine and 52.6-fold in cultures grown with glucose and ribose, when compared to a glucose-grown control. This result indicated that expression of the cg1547 gene cluster is stimulated by uridine and ribose in the rbsK1/rbsK2 double mutant and that a functional ribokinase is dispensable for generating the cognate effector of the Cg1547 repressor.

The upstream region of the cg1547 operon is characterized by a cre-like DNA sequence

The genetic organization of the cg1547 gene cluster and, in particular, the short intergenic regions between the respective genes are indicative of an operon structure (Fig. 1a). This view is supported by results obtained with the operon prediction tool vims, which also denoted cg1547, cg1546, cg1545 and cg1543 as an operon (Price et al., 2005). To provide further evidence for the presence of a polycistronic mRNA molecule, total RNA was isolated from an exponentially growing C. glutamicum KB1547 culture and used in a real-time RT-PCR experiment with DNA primers that overlap the intergenic regions between cg1547 and cg1546, cg1546 and cg1545, and cg1545 and cg1543. The PCR products were separated in an agarose gel, showing amplicons for each of the three intergenic regions, suggesting that the four genes are transcribed into a polycistronic mRNA (data not shown). No amplicons were obtained when reverse transcriptase was omitted from the assay, demonstrating the utilization of a pure RNA sample.

The transcriptional start site of the cg1547 operon was mapped by means of a RACE-PCR experiment. The 5’ end of the transcript was amplified, cloned in E. coli TOP10 (Invitrogen) and sequenced. The DNA sequence of the PCR product allowed the detection of the transcriptional start site 66 nucleotides upstream of the translational start codon of cg1547 (Fig. 1b). Based on these experimental data, a potential promoter region was deduced, taking into consideration (Pátek et al., 2003). The identified −35 (TTGCCA) and −10 (TAGGTT) regions are separated by an 18 bp spacer (Fig. 1b). The predicted −10 region is positioned 8 nucleotides upstream of the transcriptional start site and agrees in 3 out of 6 base positions with the consensus −10 sequence TA(C/T)AAT of corynebacterial promoters, suggesting that the four genes are transcribed into a polycistronic mRNA (data not shown). The predicted −35 region perfectly matches the typical −35 sequence TTGCCA (Pátek et al., 2003). The 3’ end of the transcript is specified by a 39 bp terminator structure downstream of the cg1543 gene (Fig. 1a). This hairpin structure was revealed with a ΔG value of −26.1 kcal mol⁻¹ (−109.2 kJ mol⁻¹) (Zuker, 2003).

To elucidate the transcriptional regulation of the cg1547 operon, the upstream region of cg1547 was searched for potential DNA binding sites, using the repeter tool (Kurtz et al., 2001). By this means, a 29 bp palindromic sequence was detected 32 nucleotides upstream of the cg1547 start codon (Fig. 1b). A more detailed analysis of the palindromic sequence revealed striking similarities to catabolite-responsive elements (cre) of diverse catabolite control proteins (CcpA), representing a prominent subgroup of LacI/GalR-like transcriptional regulators (Hueck et al., 1994; Kraus et al., 1998). A consensus sequence of the cre motif has been defined for Bacillus subtilis (Weickert & Chambliss, 1990). The palindromic sequence in front of the cg1547 gene includes two copies of the cre sequence (Fig. 1b), both matching the B. subtilis consensus with one nucleotide difference, whereas the second sequence additionally lacks the first consensus base (Fig. 1c). Both sequence motifs of C. glutamicum are also similar to the cre of Lactobacillus pentosus (Mahr et al., 2000). The position of these cre-like sequences in between the transcriptional and translational start sites of cg1547 is indicative of a repressor-binding site (Madan Babu & Teichmann, 2003a).

A streptavidin-tagged Cg1547 protein binds to cre-like DNA sequences in front of the cg1547 operon

To provide experimental evidence that the cre-like sequences are integral parts of the cognate operator for Cg1547, electrophoretic mobility shift assays were performed. The native cg1547 gene was amplified by PCR and cloned in E. coli DH12aMCR into the tetracycline-inducible expression vector pASK-IBA5+ (Table 1). The resulting E. coli strain was grown in LB medium and induced to express the recombinant Cg1547 protein fused to an N-terminal streptavidin tag (Cg1547_N). Strep-Tactin affinity chromatography was applied to purify a homogeneous protein that was identified as Cg1547 by MALDI-TOF mass spectrometry (data not shown).

The 29 bp palindromic sequence identified in the upstream region of cg1547, containing two cre sites, was selected as potential operator of Cg1547. The DNA band-shift assays were performed with 30 pmol of the recombinant Cg1547_N protein and 0.04 to 0.05 pmol of Cy3-labelled, double-stranded DNA fragments of varying length. The 29 bp sequence was always located in the centre of the respective DNA fragments, while the flanking regions were filled with nucleotides of the native chromosomal context (Fig. 3a). Initial DNA band-shift assays revealed that the recombinant Cg1547_N protein is able to shift a DNA fragment of 59 bp in length (59 bp oligo) that possesses flanking sequences of 15 bp on each side of the cre sequences (Fig. 3a). To narrow the native flanking sequences required for a DNA band shift with Cg1547_N, DNA fragments of decreasing sizes were tested, revealing that a border sequence of 7 bp to each side of the cre sequences (43 bp oligo) is still sufficient to change the electrophoretic mobility of the selected DNA fragments (Fig. 3a). On the other hand, a DNA band-shift assay with a 40-mer fragment revealed no differences in electrophoretic mobility. DNA band-shift assays were also performed with 59 bp oligonucleotides and modified binding buffer including up to 20 mM glucose, ribose, ribose...
Fig. 3. Electrophoretic mobility shift assays with recombinant Cg1547_N protein and the cre sequences in front of cg1547.

(a) Band-shift assays with DNA fragments of decreasing sizes of the native flanking sequences. The band-shift assays were carried out with Cy3-labelled, double-stranded DNA fragments ranging from 59 to 40 bp in size. The cre sequences present in the selected DNA fragments are boxed and shown in bold. The 29 bp palindrome is indicated by arrows. The assays were separated in a 2% agarose gel. Lanes 1, assay without Cg1547_N as negative control; lanes 2, shift assays that included the purified Cg1547_N protein. (b) Band-shift assays with modified DNA fragments. The selected DNA fragments included either the first cre sequence or the second one. A mutated version of the cre sites included transversions for all conserved base positions of the cre sequence (italic letters). The 59 bp oligonucleotide (59 bp), also shown in (a), served as control. The arrows mark differences in electrophoretic mobility. (c) Band-shift assays with DNA fragments covering the intergenic region between cg1548 and cg1547. The region was covered by six oligonucleotide fragments (oligo a to oligo f). The positions of the cre motifs and the promoter regions (−10 and −35) are indicated.

K. Brinkrolf and others
5-phosphate, uridine and uracil in order to analyse differences in electrophoretic mobility in the presence of potential effector molecules of Cg1547_N. DNA band shifts were also detected in the presence of these compounds, indicating that they do not prevent binding of Cg1547_N to the 29 bp sequence in vitro (data not shown).

Furthermore, the regulatory role of the two cre sites within the potential operator sequence of cg1547 was examined. Band-shift assays with DNA fragments lacking either the first half of the palindromic sequence or the second one were compared to band shifts with the 59 bp oligonucleotide including both cre motifs. These assays revealed differences in the electrophoretic mobility between single and double cre sequences (Fig. 3b). Moreover, the first cre sequence was shifted completely by Cg1547_N, whereas the second cre motif led to only an incomplete shift in vitro (Fig. 3b). This limited ability of the second cre sequence to be shifted by the recombinant Cg1547_N protein could be due to the lack the conserved thymidine in the first position of the motif (Fig. 1c).

To verify the regulatory function of the 29 bp palindromic sequence as operator of the cg1547 operon, a 59 bp oligo was designed that possesses a mutated version of the cre sequence (Fig. 3b), but contains the native flanking regions. Transversions were introduced for all conserved bases of the cre consensus sequence by means of an exchange of purine bases to pyrimidine bases and vice versa. In this case, the Cg1547_N protein was unable to shift the modified operator sequence (Fig. 3b). The flanking regions of the 29 bp palindrome can therefore be excluded from constituting the DNA-binding motif of Cg1547. In addition, the lack of a DNA band shift supported the view that binding of the recombinant Cg1547_N protein is sequence-specific in the in vitro assay used (Fig. 3b). Furthermore, the intergenic region between cg1548 and cg1547 was completely covered by DNA oligonucleotides that were used in DNA band-shifts assays (Fig. 3c).

Interaction of the Cg1547_N protein was detected only with oligonucleotide e, which covered the detected cre sequences in front of cg1547. No binding was observed with DNA fragment d, which includes the promoter region of the cg1547 operon (Fig. 3c). These data provided clear evidence that the 29 bp palindromic sequence upstream of cg1547 is part of the Cg1547 operator.

**The cg2554 (rbsK2) upstream region contains a cre-like DNA sequence that is recognized in vitro by Cg1547**

A catabolite-responsive element with one mismatch to the B. subtilis consensus sequence of cre (Weickert & Chambliss, 1990) was detected 22 bp upstream of the start codon of cg2554 as an integral part of a 20 bp palindromic sequence (Fig. 4a). Transcriptional regulation of cg2554 and the regulatory role of Cg1547 in cg2554 gene expression were therefore subject to further analysis. In a first experiment, the start site of the cg2554 transcript was mapped by means of RACE-PCR. A total RNA sample of a C. glutamicum ATCC 13032 culture was used to amplify and sequence the 5’ end of the cg2554 transcript. Data analysis revealed the transcriptional start site to be located 42 nucleotides upstream of the start codon of cg2554 (Fig. 4a). This mapping was used to deduce potential −35 (TAACCA) and −10 (TAGAAT) promoter regions that are separated by a 19 bp spacer (Fig. 4a). The location of the cre motif between the transcriptional and the translational start site of cg2554 is suggestive of a repressor-binding site (Madan Babu & Teichmann, 2003a).

To examine the binding of the Cg1547 protein to the cre motif in front of cg2554, DNA band-shift assays were performed, resulting in an almost complete mobility shift of the DNA fragment (Fig. 4b). Therefore, the second ribokinase gene cg2554 might be under negative transcriptional control by Cg1547, although this finding required confirmation by in vivo data. For this purpose, RT-PCR

![Fig. 4](image_url)

**Fig. 4.** Regulatory features of the cg2554 gene and DNA band-shift assay with Cg1547_N. (a) Genetic map of the upstream region of cg2554. The predicted ribosome-binding site (RBS), the transcriptional start site (+1) and the deduced −10 and −35 regions are marked. The start codon of the cg2554 gene is indicated by an arrow. The palindromic DNA region covering the cre sequence (box) is given in bold. (b) Electrophoretic mobility shift assays with purified Cg1547_N protein and the cre sequence of cg2554. Shift assays were carried out with a Cy3-labelled, double-stranded 40-mer (underlined in a) and 30 pm of the Cg1547_N protein. Separation in a 2% agarose gel shows the control assay without protein (lane 1) and the DNA band-shift assay with the purified Cg1547_N protein (lane 2).
experiments were used to examine the stimulation of cg2554 gene expression in vivo. C. glutamicum ATCC 13032 was grown to mid-exponential phase in CGXII minimal medium supplemented with 30 mM of glucose, ribose or uridine. Total RNA was isolated from the samples and was used to measure the expression of cg2554 in the uridine- and ribose-grown cultures, in comparison to that in the glucose culture. Relative expression of cg2554 was enhanced 2.4-fold and 4.7-fold in the presence of uridine and ribose, respectively. Differential expression of cg2554 is thus stimulated by the presence of external uridine and ribose, as was already shown for the genes of the cg1547 operon.

Cg1547 is a pathway-specific regulator as deduced from whole-genome DNA microarray hybridizations

To elucidate the regulatory role of Cg1547 on a global scale, whole-genome DNA microarray hybridizations were carried out, comparing the gene expression in the cg1547 mutant C. glutamicum KB1547 with that in the wild-type strain C. glutamicum ATCC 13032. Both strains were cultivated in CGXII minimal medium (with 40 g glucose l⁻¹ as carbon source) to mid-exponential growth phase. RNA was purified from the samples and used for two hybridization assays with label swapping. Evaluation of the transcriptional profiling in C. glutamicum KB1547 revealed eight genes with significant changes in expression (Table 2; see supplementary Table S2 for the full list of genes in the microarray). This set includes four genes with enhanced expression (M-value ≥ 1), which are given by the cg1547 operon genes (cg1546, cg1545 and cg1543) and the cg0672 gene, coding for a conserved hypothetical protein. Four genes with decreased expression in C. glutamicum KB1547 (M-value ≤ −1) were detected (Table 2), including the cg2893 gene for a multidrug efflux permease and the adjacent regulatory gene cg2894 (Itou et al., 2005). Besides the cre sequences in front of cg1547, which regulate the expression of the cg1547 operon genes cg1546, cg1545 and cg1543, no further cre-like motifs were detectable in the upstream regions of the remaining five differentially expressed genes. Thus, transcriptional profiling on the whole-genome level revealed that Cg1547 acts as a pathway-specific regulator, negatively controlling the expression of its own gene and of the operon genes cg1546 (rbsK1), cg1545 (uriT) and cg1543 (uriH). Interestingly, the cg2554 (rbsK2) gene was not detected as differentially expressed in the cg1547 mutant C. glutamicum KB1547. Nevertheless, since all these genes are relevant for the utilization of uridine as the sole carbon source for growth, the Cg1547 protein was named UriR (uridine utilization regulator).

DISCUSSION

The uriR gene region is involved in uridine utilization by C. glutamicum

In the present study, we elucidated the physiological role of the transcriptional regulator UriR of C. glutamicum ATCC 13032. The last gene of the uriR operon codes for the nucleoside hydrolase UriH, which is one of three nucleoside hydrolases encoded in the genome of C. glutamicum (Kalinowski et al., 2003). Nucleoside hydrolases are present in many bacteria, yeasts and protozoa and catalyse the irreversible hydrolysis of the N-ribosidic bond of nucleosides to generate free nucleic bases and D-ribose (Kim et al., 2006; Miles et al., 1999). The nucleoside hydrolases Rih1 and Rih2 from C. ammoniagenes have been purified and characterized (Kim et al., 2006). While Rih2 is a non-specific nucleoside hydrolase with enzymic activities between 23.8 and 5.1 units for the nucleosides inosine, adenosine, uridine, guanosine, xanthosine and cytidine, Rih1 is a pyrimidine-prefering nucleoside hydrolase with much higher activities for the pyrimidine nucleosides uridine (945 units) and cytidine (115.5 units) than for the purine nucleoside inosine (3 units). A comparison on the amino acid sequence level of the two nucleoside hydrolases with C. glutamicum proteins revealed that Rih2 is homologous to the predicted nucleoside hydrolase Cg2168, while Rih1 is similar to UriH (Kim et al., 2006). Growth experiments showed that the nucleoside hydrolase

Table 2. Differentially expressed genes in C. glutamicum KB1547 compared to the wild-type detected by DNA microarray hybridization

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene name</th>
<th>Gene product or deduced function</th>
<th>M-value*</th>
<th>A-value*</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg1543</td>
<td>uriH</td>
<td>Uridine-prefering nucleoside hydrolase</td>
<td>3.0</td>
<td>12.2</td>
<td>7.9 x 10⁻⁸</td>
</tr>
<tr>
<td>cg1545</td>
<td>uriT</td>
<td>Uridine uptake system</td>
<td>2.4</td>
<td>11.4</td>
<td>3.4 x 10⁻⁸</td>
</tr>
<tr>
<td>cg1546</td>
<td>rbsK1</td>
<td>Ribokinase</td>
<td>2.2</td>
<td>11.1</td>
<td>1.9 x 10⁻⁸</td>
</tr>
<tr>
<td>cg0672</td>
<td>−</td>
<td>Conserved hypothetical protein</td>
<td>2.2</td>
<td>10.8</td>
<td>4.3 x 10⁻⁸</td>
</tr>
<tr>
<td>cg0995</td>
<td>rpmF</td>
<td>50S ribosomal protein L32</td>
<td>−1.2</td>
<td>12.5</td>
<td>5.1 x 10⁻⁵</td>
</tr>
<tr>
<td>cg0097</td>
<td>−</td>
<td>Conserved hypothetical protein</td>
<td>−1.5</td>
<td>10.4</td>
<td>3.8 x 10⁻⁵</td>
</tr>
<tr>
<td>cg2894</td>
<td>−</td>
<td>Transcriptional regulator, TetR family</td>
<td>−3.0</td>
<td>11.8</td>
<td>3.2 x 10⁻⁷</td>
</tr>
<tr>
<td>cg2893</td>
<td>−</td>
<td>Multidrug efflux permease, MFS type</td>
<td>−3.5</td>
<td>11.0</td>
<td>2.6 x 10⁻⁷</td>
</tr>
</tbody>
</table>

*Abbreviations: M, signal intensity ratio; A, signal intensity; P, error probability.
UriH is necessary for the utilization of uridine as the sole carbon source in *C. glutamicum*, whereas utilization of cytidine and inosine was not observed. We suggest that UriH is a pyrimidine-prefering nucleoside hydrolase with high activity for uridine conversion like the homologous counterpart in *C. ammoniagenes*, thereby enabling growth of *C. glutamicum* with this nucleoside as the sole carbon and energy source. The enzymic conversion by UriH of other nucleosides such as cytidine and inosine may be insufficient to support growth of *C. glutamicum* cultures. Alternatively, uridine may be the only nucleoside imported sufficiently into the *C. glutamicum* cell to allow growth. A candidate protein for the import of external uridine into the *C. glutamicum* cell is the secondary carrier UriT of the MFS (Winnen et al., 2003). A comparable arrangement of a nucleoside hydrolase gene and an MFS permease gene is also observed for rhi1 in *C. ammoniagenes* (Kim et al., 2006) and cel1482 in the genome sequence of *Corynebacterium efficiens* (Nishio et al., 2003). The import of external uridine into the *C. glutamicum* cell is followed by its irreversible hydrolysis into uracil and D-ribose by one of the two ribokinases. This function is only lost in the double mutant *C. glutamicum* RBSK12, which is unable to utilize uridine and ribose as carbon sources for growth. This result implies that the ribokinase functions of RbsK1 and RbsK2 can be replaced by one another.

The repressor UriR controls the expression of the uridine utilization genes

The UriR protein is one of seven functional members classified into the LacI/GalR family of transcriptional regulators in *C. glutamicum* (Brune et al., 2006). In corynebacteria, regulatory proteins of this family have not been characterized so far, but homologous proteins of other bacteria have revealed that many LacI/GalR proteins are involved in the control of genes associated with carbohydrate metabolism or de novo nucleotide biosynthesis (Weickert & Adhya, 1992). The regulatory gene uriR is organized in an operon structure along with genes necessary for uridine utilization. Expression of these genes is jointly upregulated in the *uriR* mutant *C. glutamicum* KB1547, suggesting negative regulatory control of *rbsK1*, *uriT* and *uriH* by the UriR protein. A common feature of LacI/GalR regulators is the composition of an amino-terminal HTH motif and a carboxy-terminal domain for oligomerization and effector binding with a three-dimensional structure similar to those of sugar-specific periplasmic binding proteins (Fukami-Kobayashi et al., 2003; Weickert & Adhya, 1992). LacI/GalR proteins bind small-molecular-mass effectors such as the carbohydrates allo-lactose (LacI), ribulose (RbtR), galactose and fructose (GalR), nucleosides like cytidine and adenosine (CytR), or the nucleobases hypoxanthine and guanine (PurR) (Daber et al., 2007; Weickert & Adhya, 1992). Binding of these effector molecules to the C-terminal domain of the protein usually results in a conformational change of the protein followed by the dissociation of the transcriptional regulator from its target DNA (Nguyen & Saier, 1995). In the case of UriR of *C. glutamicum*, the cognate effector molecule could not be detected by *in vitro* DNA binding assays. Nevertheless, expression of the uridine utilization genes is stimulated by the presence of uridine or ribose in the growth medium, indicating a relation between the regulatory function of UriR and a degradation product of uridine. However, the uridine stimulus is lost with the deletion of *uriH* or *uriT*, but it is still present in the ribokinase double mutant *C. glutamicum* RBSK12, suggesting that the D-ribose moiety released from uridine may function as effector molecule of the UriR repressor *in vivo*.

The UriR repressor binds to catabolite-responsive elements upstream of the *uriR* operon

DNA-binding transcriptional regulators can bind to their target DNA via an HTH motif that is generally able to recognize a maximum of 5–6 base pairs (Dong et al., 1999). The specificity of this protein–DNA interaction is increased by the formation of dimers or higher-order aggregates, because longer recognition sequences occur at a lower frequency by chance within a given genome sequence (Dong et al., 1999). LacI/GalR regulators usually form homodimers, with each HTH motif recognizing one half of a palindromic DNA target sequence (Weickert & Adhya, 1992). UriR of *C. glutamicum* exhibits an amino-terminal HTH motif of the λ repressor-like DNA-binding domain type (Brune et al., 2005). DNA band-shift assays with recombinant protein revealed binding to a *cis*-active 29 bp palindromic sequence originating from the intergenic region of the *uriR* gene. The operator sequence includes two cre sites, which caused reduced electrophoretic mobility of DNA fragments in conjunction with recombinant UriR protein. UriR showed a lower affinity to the second *cre* site of the palindrome than to the first one. This difference in DNA binding may be due to the fact that the second *cre* site lacks the first position of the motif (Fig. 1c). With this proof of functionality of both cre sequences *in vitro* and the detection of differences in electrophoretic mobility between single and double cre motifs (Fig. 3b), we suggest that binding of two homodimers of UriR to the operator sequence upstream of the *uriR* gene is possible. Moreover, the existence and functionality of these cre motifs seems to link the transcriptional regulator UriR to the CcpA-like proteins, which form a subgroup within the LacI/GalR family and use cre sequences to interact with DNA. CcpA-like proteins were defined by the conservation of 67 specific amino acids in comparison with non-CcpA-like family members (Kraus et al., 1998). In general, CcpA-like proteins function to direct CCR in low-G+C Gram-positive bacteria like *B. subtilis* (Lulko et al., 2007) and *L. pentosus* (Mahr et al., 2000) involving components of the PTS. This mechanism is different in *C. glutamicum*, where appropriate components are missing (Mahr et al., 2000) and CCR is rarely observed at the regulatory level (Arndt & Eikmanns, 2007; Letek et al., 2006; Parche et al., 2001). Moreover, the amino acid sequence of UriR shares only 12
out of the 67 characteristic amino acids with CcpA-like proteins and is rather similar to classical LacI/GalR family proteins. Global transcriptional profiling revealed that UriR acts as a pathway-specific regulator of uridine metabolism and is not involved in the CCR of *C. glutamicum*. It is known that *C. glutamicum* can utilize AMP and UDP-glucose as sources of phosphorus, a process requiring the extracellular nucleotidase UshA (Rittmann et al., 2005). The activity of UshA allows *C. glutamicum* to access nucleotides in the soil environment as sources for phosphorus for growth. The released nucleoside uridine can additionally serve as carbon source via the pathway encoded by the cg1547 gene cluster, providing a physiological link between phosphorus and carbohydrate metabolism.

**The UriR protein binds to the promoter region of the ribokinase gene rbsK2**

Electrophoretic mobility shift assays revealed binding of the recombinant UriR protein to the deduced operator sequence of *rbsK2* in vitro, although transcriptional profiling on the whole-genome level failed to detect differential expression of *rbsK2* in the *uriR* mutant *C. glutamicum* KB1547. We therefore suggest that transcriptional control of *rbsK2* is more complex and occurs in conjunction with another regulatory protein. A candidate regulator for this purpose with proposed involvement in the ribose metabolism of *C. glutamicum* is Cg1410, which is also classified as a LacI/GalR family member (Brune et al., 2005). Previous studies analysing the transcriptional profile of *C. glutamicum* cultures grown with glucose or ribose revealed differential expression of the ribose transporter genes *rbsABCD* (Wendisch, 2003), which are clustered with the regulatory gene cg1410 (Kalinowski et al., 2003). A potential regulatory role of Cg1410 in *rbsK2* gene expression is currently under investigation.

**ACKNOWLEDGEMENTS**

The authors thank Susanne Götker and Manuela Meyer for excellent technical assistance and Andreas Albersmeier for helpful discussions regarding the RACE-PCR experiments.

**REFERENCES**


Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B. J. & other


