Transcriptome analysis of temporal regulation of carbon-metabolism by CcpA in *Bacillus subtilis* reveals additional target genes

Andrzej Lulko, Girbe Buist, Jan Kok and Oscar Kuipers
ABSTRACT

The pleiotropic regulator of carbon metabolism in Gram-positive bacteria, CcpA, regulates gene expression by binding to so-called cre elements, which are located either upstream or in promoter regions, or in open reading frames. In this study we compared the transcriptomes of *B. subtilis* 168 and its ccpA deletion mutant during growth in glucose-containing rich medium. Although growth was similar, glucose was completely consumed by the wild-type strain in the stationary phase, while it was still present in the culture of the mutant. At that stage direct and indirect effects on gene expression were observed. During exponential growth CcpA mainly influences the carbohydrate and energy metabolism, whereas from transition phase onwards its function expands on a broader range of physiological processes including nucleotide metabolism, cell motility and protein synthesis. A genome wide search revealed new putative cre sites, which could function in vivo according to our transcriptome data. Comparison of our data with published transcriptome data of ccpA mutant analysis in the exponential growth phase confirmed earlier identified CcpA regulon members. It also allowed identification of potential new CcpA-repressed genes, amongst others ycgN and the ydh operon. Novel activated members include opuE and the opuAABC, yhb and man operons, which all have a putative cre site that appears to be dependent on helical topology. A comparative analysis of these genes with the known activated genes *i.e.* ackA and pta revealed the presence of a possible upstream activating region (UAR) as has been shown to be functional for the activation of ackA. The data suggest that at later growth phases CcpA may regulate gene expression by itself or complexed with other, yet unknown cofactors.
Introduction

The CcpA protein [98,106,174] is a master regulator of catabolite control in many low-GC Gram-positive bacteria [261]. CcpA is a member of the LacI/GalR family of transcriptional regulators [265] which can act as a repressor in carbon catabolite repression (CCR) and as an activator in carbon catabolite activation (CCA) [104,235]. CcpA forms a dimer [219] and its DNA-binding activity is stimulated by complex formation with HPr-Ser-P [52]; [75] or HPr-like protein Crh-Ser-P [77]. Next to these phosphoproteins also low molecular weight molecules such as NADP, glucose-6-phosphate (G6P) and fructose-1,6-bisphosphate (FBP) can modulate either DNA-binding properties or the interaction with the transcriptional machinery of the CcpA-(HPr-Ser-P) complex [79,125,126,223]. CcpA binds a DNA regulatory sequence known as catabolite-responsive element (cre) of which the consensus sequence TGWAANCGNTNWCA (N=any base; W=A or T) was first detected in the promoter region of amyE [105,112,173,266]. Later it was discovered that the presence of cre either upstream or in the promoter region, or inside a gene does not necessarily imply regulation by CcpA [173]. On basis of in vivo operational cre sites the following consensus sequence, WWTGNAARCGNWWWCAWW, was proposed. This consensus suggests a high level of degeneracy and indeed, depending on the search method and consensus sequence used, different numbers of cre sites can be predicted in the B. subtilis genome [50].

The position of the cre site(s) relative to the transcription start (TS) of a transcriptional unit determines the regulatory effects upon HPr-Ser-P/CcpA complex binding. In general, promoters with a cre site present upstream of the hexameric -35 sequence undergo transcriptional activation as in case of ackA and pta [198,247]. Interestingly, both genes contain an additional conserved sequence upstream of the CcpA binding site, the presence of which appears to be crucial for transcriptional activation of ackA [176]. Recently, a direct positive regulation by CcpA was also demonstrated for the ilvB promoter [225,244]. Transcriptional repression by binding of CcpA to a cre site downstream of the TS, which blocks elongation by RNA polymerase (RNAP) has been shown. This transcriptional roadblock mechanism has been proposed for the xyl, ara and gnt operons [110,112,281] and for sigL [37] and acsA [282]. In the latter case affinity of CcpA for the cre site was also dependent on the composition of the nucleotides flanking cre. Prevention of binding of RNAP to the promoter sequence has been shown for the acuABC and bgI PH operons where a cre site overlaps with parts of the promoter region which may prevent transcription initiation [82,135]. Recently, it has been suggested that CcpA does not prevent RNAP from binding to the promoter of amyE but that it interacts directly with the RNAP complex already bound to its operator site [127]. In addition, activation or repression by CcpA binding to cre is helix-
face dependent since non-integral turns of helix insertions caused relief of CCR of \textit{amyE} [127] or lack of activation as in case of \textit{ackA} [247].

Previously, different transcriptomics technologies have been used, such as Sigma-Genosys macroarrays [19,181] and in-house spotted glass microarrays [276] both containing PCR products of whole open-reading frame (ORF) sequences or commercially available Affymetrix chips consisting of multiple 25-mer oligonucleotides per ORF [150]. However, the common characteristic of these analyses was that they were carried out on mid-exponentially growing cells. The aim of this research was to define the CcpA regulon in time. To this end, we performed whole transcriptome analysis of a \textit{ccpA} mutant strain at four different stages of growth in rich glucose-containing medium and compared it to the well-defined CcpA regulon in mid-logarithmic phase described in literature. We conclude that the CcpA regulon dynamically develops in time and that the impact of the \textit{ccpA} mutation spreads out and intensifies during growth, affecting various cellular processes in growing \textit{B. subtilis} cells. The approach allowed us to further dissect the CcpA regulon over time and permitted identification of new (putative) CcpA targets.

\textbf{Materials ans methods}

\textbf{Bacterial strains, growth conditions and preparation of cells for RNA isolation}

Chromosomal DNA of the \textit{B. subtilis} QB5407 \textit{ccpA} mutant strain [65] was isolated and transformed to \textit{B. subtilis} 168 to ensure the same genetic background of the wild-type and the \textit{ccpA} mutant. Cultures were grown in 250 ml Erlenmeyer flasks by inoculation with “synchronized” cells at OD$_{600}$ 0.01 in the 25 ml TY medium supplemented with 0.5% glucose. Synchronized stocks for inoculation were prepared by diluting overnight (ON) cultures to OD$_{600}$ 0.01 in fresh medium and growing the cells to the mid-exponential phase and diluting again to OD$_{600}$ 0.01. After three cycles of growing to mid-exponential phase synchronized stocks were frozen in glycerol at -80°C for future use. All growth experiments were performed at 37°C with 250 rpm shaking in an Innova 4000 incubator (New Brunswick Scientific). The reason synchronization was to avoid variable lag-phases of the \textit{ccpA} mutant, which were observed when direct inoculation from ON cultures was attempted. Cells for RNA isolation and subsequent microarray analysis were harvested by centrifugation (1 min, 12,000 rpm, 4°C) of early-, mid- and late-exponential phase as well as stationary phase cells corresponding to approximate OD$_{600}$ values of 0.12, 0.8, 3.4 and 5.5, respectively. Cell pellets were immediately frozen in liquid nitrogen and stored at -80°C until RNA isolation.

\textbf{Design and production of DNA microarrays}

The oligonucleotide design was based on the coding sequences of the GenBank RefSeq file NC_000964 for \textit{B. subtilis} 168. The design was performed with OligoArray 2.1-software [211] with the following parameters: oligonucleotide length 70-72 bps, Tm 88-93°C, secondary
structure melting-point 65°C and GC content 42-52%. In the cases where OligoArray 2.1 could not design a suitable oligo, Picky (Complex Computation Laboratory, Iowa state university) was used with less stringent parameters. For 39 ORFs no unique oligo could be designed due to gene duplication within the *B. subtilis* genome and therefore, when possible, oligos were designed manually. For highly homologous paralogues only one oligo was designed. Next to sequences representing the ORFs, additional sequences of antibiotic markers, plasmids and *gfp/yfp/cfp* probes were spotted (supplementary material). All 4150 sequences were spotted in duplicate on aminosilane glass slides (ez-rays™, Matrix Technologies Corp.). Slide spotting, slide treatment after spotting and slide quality control were performed as described previously [251].

**DNA microarray analysis**

The general approach for DNA microarray experiments was implemented as described previously [136,251]. Total RNA was isolated with the High Pure RNA Isolation Kit (Roche Applied Science) according to the manufacturer's instructions. RNA quantity and quality were assessed with a Nano Drop ND-1000 spectrophotometer (NanoDrop Technologies) and an Agilent Bioanalyzer 2100 with RNA 6000 LabChips (Agilent Technologies Netherlands BV).

cDNA was synthesized with the Superscript III Reverse Transcriptase kit (Invitrogen) using 25 μg of total RNA as template and 400U of SuperScript™ III RT. The reaction contained 0.5 mM dATP, dCTP, dGTP, 0.3 mM dTTP and 0.2 mM of amino allyl-modified dUTP. The reaction was incubated for 16 hrs at 42°C. The amino allyl-modified cDNA was purified with the CyScribe GFX purification kit (Amersham Biosciences) using 80% ethanol as wash buffer and 0.1M sodium carbonate solution (Sigma-Aldrich) pH 9.0 as elution buffer. The purified cDNA was labelled with Cy3- or Cy5 mono-reactive Dye (Amersham Biosciences) and incubated at room temperature in the dark for 90 min. Labelled cDNA was purified with the CyScribe GFX purification kit again as described by manufacturer. Dye incorporation and cDNA concentration were assessed with a Nano Drop ND-1000 spectrophotometer (NanoDrop Technologies). The labelled cDNA was hybridized to oligonucleotide microarrays in Ambion Slidehyb #1 buffer (Ambion Europe Ltd.) at 45°C for 18-20 hrs. After hybridization, slides were washed for 5 min. at 37°C in 2 x SSC with 0.5% SDS and 2 x 5 min. at 37°C in 1 x SSC with 0.25% SDS, dipped five times in 1 x SSC 0.1% SDS and then dried by centrifugation (2 min, 2,000 rpm).

The microarrays were scanned with a GeneTac LS V confocal laser scanner (Genomic solution Ltd.). Determination of the individual intensities of each spot was done with ArrayPro 4.5 (Media Cybernetics inc., Silver Spring, MD) with a local corners background correction method and the resulting expression levels were processed and normalized (Lowess method) with MicroPrep [253]. The In-transformed ratios of the expression levels of *ccpA* mutant versus wild-type were subject to a *t*-test using the Cyber-T tool [149]. For each time point four independent biological replicates (four growth curves with four slides each)
were performed resulting in eight measurements per gene, since each slide contains two duplicate spots for all genes. Genes were selected which had an expression ratio $ΔccpA/wt ≥ ±2$ at least in one of the investigated time points (values below one were converted by the formula $[-1] × [1/ratio value]$) and a Cyber-T (Bayes) $p$-value below 0.01. For these genes the values of three time points were added when the expression ratio $ΔccpA/wt$ was at least 1.5 times and had a $p$-value below 0.05. If these criteria were not met the expression ratio was set to one. The final gene list selected in this manner is presented in Table S1 (supplementary material). The data was further analyzed with FIVA (Functional Information Viewer and Analyzer; Blom et al. submitted) to identify overrepresented functional categories in clusters of up and down-regulated genes per individual time point. Two programs were employed for clustering of the time series data; Genesis [236] was used for hierarchical clustering of known members of the CcpA regulon and STEM (Short Time-series Expression Miner; [60]). All the clustering analyses were performed on the log2 transformed gene expression ratios.

**cre site prediction**

A weight matrix (Fig. S1) was generated in Genome 2D [10] based on all known cre’s from DBTBS (DataBase of Transcriptional Regulation in *Bacillus subtilis*; [162]) and thereafter it was fed into the MotifLocator tool [1] to search the whole genome of *B. subtilis* for the potential CcpA binding sites. This approach yielded 311 cre’s with a score above 0.85 (Table S1). The position of cre site was always related to a known or deduced TS site of a gene.

**Glucose utilisation**

D-glucose depletion from medium during growth was determined for the wild-type and the *ccpA* mutant strain by means of the D-Glucose enzymatic bioanalysis/food analysis kit (R-biopharm). Samples were collected at one hour intervals from the inoculation point till eight hours of growth. Supernatants were separated from cells by centrifugation and kept at 80°C for 15 min and the D-glucose concentration was measured according to the manufacturer’s instructions.

**Results**

The *B. subtilis* CcpA modulon in glucose-rich medium was analyzed at different stages of growth by comparing the transcriptional profiles of *B. subtilis* 168 cells with its isogenic $ΔccpA$ strain. Samples for DNA-microarray analysis were collected in early (E), middle (M) and late exponential (transition; T) phase, as well as three hours upon entry into stationary phase (S). The difference in growth rate between the strains was small (Fig. 1), which allowed a relatively easy estimation of appropriate sampling times.
Global impact of the ccpA mutation at various stages of *B. subtilis* growth

To identify differentially expressed genes between the wild-type and the ccpA mutant we subjected the transcriptome data in the form of Lowess-normalized expression ratios [253] to the Cyber-T statistical package [149]. This program was used to evaluate the significance of the ΔccpA/wt expression ratios. The gene expression ratios from the four individual time points was analyzed by the FIVA software (Functional Information Viewer and Analyzer; Blom et al., submitted) to identify which general functional classes (categories) are affected upon ccpA mutation and to distinguish the temporal aspect of these changes. Examination of clusters of orthologous groups (COG) of proteins [240] showed that the effects of the mutation in the early exponential phase concern exclusively ‘carbohydrate transport and metabolism’ as well as ‘general energy production and conversion’ processes (Fig. 2). The majority of these genes are up-regulated in the ccpA mutant. The same two COG categories become even more strongly influenced by CcpA throughout the exponential phase. In the stationary phase gene regulation in the ‘energy production and conversion’ category shows a swap from up- to down-regulation. The transcriptional profile alters in time since already in the mid-exponential phase additional COG’s are significantly influenced by CcpA, namely ‘cell motility and secretion’ as well as ‘nucleotide and amino acid transport and metabolism’. When the clusters of the mid-logarithmic phase are compared to the clusters of the transition phase (Fig. 2).
state the same COG’s stay affected but with enhanced effects, as more genes fall into each individual cluster. There is an apparent switch in scope of the ccpA mutation after the cells enter the stationary phase of growth. Transcriptional profiles of two categories, namely ‘nucleotide transport and metabolism’ as well as ‘energy production and conversion’,
undergo reversion. In addition three new categories become influenced; (i) ‘translation and ribosomal structure and biogenesis’ with a majority of genes showing up-regulation; (ii) ‘secondary metabolites biosynthesis, transport and catabolism’ as well as (iii) ‘inorganic ion transport and metabolism’ with a majority of genes falling into cluster ‘down’. The time-course approach indicates the dynamics of the CcpA modulon revealing an increasing range of CcpA regulation throughout the whole life cycle.

**Clustering of time series data (STEM)**
To identify expression profiles during the investigated time points the log₂ transformed ratios from table S1 (supplementary material) were subjected to clustering using STEM [60] which implements a unique method of clustering of short time series. The data was divided into 15 arbitrary clusters with the maximum unit change set to one (ratio change of 2 in log₂ scale) in model profiles between the time points. The obtained profiles were sorted based on the number of genes assigned (Fig. 3). A gene table with assigned profiles is available in the supplementary material (Table S2). Profiles 12 and 13, the former containing the highest number of 167 members, represent genes which are first up-regulated in the *ccpA* mutant in the logarithmic phase and down-regulated in the stationary phase. Profiles 1 and 5-9 depict genes which respond to lack of CcpA either by elevated (7, 8, 9) or by decreased (1, 5, 6) expression levels at specific stages of growth.

**Fig. 3. STEM clustering profiles.** 15 clusters contain the genes filtered on basis of the expression ratio change \( \Delta_{ccpA/wt} \geq 1 \) (in log₂ scale) between the time points. The dashed line indicates no change in expression between the strains. Growth phases E, M, T and S as described in Fig. 1.
Clusters 14 and 0 contain genes which respond to the absence of CcpA by constantly increasing or decreasing ratios, respectively. The remaining profiles (2, 3, 4, 10, 11) contain genes whose expression fluctuates in time. The general tendency of repression by CcpA, either direct or indirect, emerges from the distribution of genes, as the first four most densely populated clusters all have genes with higher expression levels in the ccpA mutant.

**Regulation of the CcpA regulon members in time**

We extracted known members of the CcpA regulon from DBTBS (DataBase of Transcriptional Regulation in *Bacillus subtilis*; [162]) and compared them to genes found to be regulated in our time series transcriptome data. The outcome of hierarchical clustering (complete linkage method) on this set of genes is presented in Fig. 4.

![Fig. 4. Hierarchical clustering of regulation of the CcpA regulon members. Complete linkage clustering as an agglomeration rule was used in GENESIS programme. Red and green colour indicate higher or lower expression in the ccpA mutant strain, respectively. Growth phases E, M, T and S as described in Fig. 1. The three major trends emerging from the clustering are marked as I, II, III.](image-url)
A few genes from this list were not regulated within our data set: \textit{araE}, \textit{bgI}S, \textit{citM}, \textit{mmgA}, \textit{yxjC}, \textit{yxkJ}, \textit{acuA}, \textit{licB} and \textit{gntR}. In case of the latter three, they were included for clustering since the other neighbouring operon member was significantly regulated (\textit{acuB}, \textit{licC} and \textit{gntK}). In previous transcriptome studies [18,150,181,276] even more of the known CcpA regulon members, as listed in DBTBS, were missing. Three major trends can be distinguished within the clustering profiles including the down-regulated (I) and up-regulated (II) genes or up-regulated in exponential growth and down-regulated in stationary phase (III) (see Fig. 4). Genes involved in acetoin metabolism (\textit{alsS}, \textit{acoA}) seem to be CcpA-activated in later phases of growth, as well as another gene of overflow metabolism, \textit{ackA}, which showed strong activation throughout the whole exponential phase. The \textit{ilvB} gene, the first member of \textit{ilv-leu} operon was specifically CcpA-activated at the transition state. On the other hand, \textit{pta} was first activated in mid- and late exponential phase and subsequently repressed in the stationary phase. Genes from cluster III demonstrated an opposite expression pattern than the \textit{pta} expression pattern. A few genes were repressed during all four growth phases, including, amongst others, \textit{glpF}, \textit{kdgA}, \textit{rbsR} and \textit{ydhO}.

**Comparison of CcpA transcriptome data**

Since four reports are already published describing the mid-exponential transcriptome of CcpA [18,150,181,276], we compared these data to ours (Table S1, supplementary material). The comparison of five transcriptome data sets allowed us to identify a group of 53 genes (Table S3, supplementary material) common in at least three of the studies indicating that these genes likely belong to the CcpA regulon, although they are not present yet in the DBTBS database. 23 of these genes have a predicted \textit{cre} site which makes them plausible candidates as direct CcpA targets. In case of \textit{sucC}, \textit{sucD}, \textit{rocA} and \textit{rocF} their protein products were also identified to be CcpA-dependent by a proteomics approach [276]. The comparison also provided an extraction of more than 200 genes not found in other whole-transcriptome studies [18,150,181,276] (Table S4, supplementary material). Computational searches performed by Miwa et al. [173] resulted in the identification of 126 putative \textit{cre} sites present in the \textit{B. subtilis} 168 genome. Because different predictions for \textit{cre} sites have been performed on the genome sequence of \textit{B. subtilis} we decided to use a weight matrix (see supplementary material) based on the known \textit{cre} elements of from the DBTBS database. With our search we detected 311 putative \textit{cre} like sequences (Table S1).

**Differentially regulated genes lacking a \textit{cre} site**

Analysis of the data in Table S1 revealed the presence of many genes which are higher or lower expressed in the \textit{ccpA} mutant but lack the presence of a putative \textit{cre} site. Expression of the whole \textit{alb-sbo} operon coding for proteins required for the synthesis and maturation of subtilosin [287,288], the \textit{dhb} operon (dihydroxybenzoate siderophore biosynthesis) [212] and \textit{tdh} operon are strongly elevated in the transition phase. Also for \textit{yerA}, the first gene of the
**yeCAyerABC** operon, no predicted cre was found although it was found to be repressed at all time points in our studies as well as in two other transcriptome analyses [150,276]. On the other hand the expression of manAyjDF and yqIX, mngE was decreased throughout the exponential phase and for some of these genes also in the stationary phase. These are just a few examples representing a large group of genes showing differential expression in the ccpA mutant but not possessing an identifiable cre sequence.

**Comparison of glucose utilisation between the wild-type and ccpA mutant strain**

It has been suggested that glucose uptake is slower in ccpA mutants [152,225,244] which could influence the glucose dependent regulation by CcpA especially in later stages of growth. Analysis of the concentration of glucose present in the medium showed that it was fully consumed by the wild-type strain at stationary growth, while it was still present in the culture of the ccpA mutant. With this knowledge in mind, we expected mainly CcpA-independent effects during stationary phase. In the stationary phase genes coding for proteins involved in the metabolism of nucleic acids and ribosomal proteins were strongly up-regulated in the ccpA mutant (Table S5), showing that most of these genes should be CcpA-independently regulated. However, some of these genes do contain a predicted cre sequence (for example ggt, lcfA, sigY, ycdl, yknW, ysbA, yvdG, yxeB) rendering them possible candidates of CcpA-dependent regulation independent of glucose presence. In these cases CcpA may interact with an unknown cofactor which is present intracellularly in the stationary phase or CcpA may bind to DNA without any associated cofactors.

**Identification of new putative CcpA activated genes**

The CcpA target ackA is one of the few genes that are activated throughout exponential growth but not in the stationary phase. Activation of ackA expression during growth in the presence of glucose is known to be dependent on the cre CcpA binding site (centred at -57 relative to the TS site) and an upstream activating region (UAR). Downstream of ackA, the gene for the molybdopterin precursor biosynthesis protein B (moaB) is located. moaB has an identical expression pattern as ackA suggesting it is in an operon with ackA.

A known target of CcpA, pta, is activated by CcpA during mid-exponential growth and in the transition phase, while repression was observed for the stationary phase. Its cre site is centred at -57-bp upstream of the TS. yhbl codes for a transcriptional regulator of the MarR family and it is most likely the first gene of the yhbl[yhcABCDEFGHI] operon. yhbl is activated during exponential growth, while the downstream located genes yhbl[yhcABC] are also activated in the transition phase of growth. Moreno et al also observed activation of yhbl at mid-exponential growth in the presence of glucose [181]. A putative cre site was predicted around 57-bp upstream of a putative TS site. manR, which codes for the putative transcriptional activator of the mannose operon, is activated during exponential phase, while the putative manPA-yjDF operon is regulated not only at these time points but also during
stationary phase. Comparison of the promoter sequence of *manR* with those of the other CcpA activated genes we detected a putative cre site around 57-bp upstream of the putative TS. A second putative cre could be found which overlaps the TS site.

The gene for the proline transporter (*opuE*) is the only gene that was lower expressed in *B. subtilis* 168 compared to its *ccpA* mutant throughout growth. We detected a putative cre 104-bp upstream of the TS. The genes for the glycine betaine transport system components (*opuAABC*) [121] are activated at the first three time points. Two different TS sites were determined for this operon. A putative cre was found 143-bp upstream of the first TS site and another one 181-bp upstream of the second TS.

### Identification of new putative CcpA repressed genes

Among the CcpA targets that were found to be repressed at all time points investigated were the known targets *glpFK, glpTQ, ycsA, yagY, rocG*, the *ydh* operon and *sdhCAB*. Although Miwa et al. identified a putative cre located in the ORF of *sdhC* [173], we detected a putative cre site that overlaps the -35 sequence of the *sdhCAB* promoter. Interestingly *ysbA, yxjO* and the operons *yknWYZ* and *ptsGHI* were found to be repressed at the stationary phase only. All other were lower expressed in the wild-type at one of the exponential growth phases. For more than 50% of the cases we detected a yet unidentified cre site. The location of cre sites of all the exponential phase targets varied in overlapping either the -35, -10, TS or sequences upstream of the TS.

The *ydhMNOPQRST* and *kdgRKAT* operons each contain two cre sites. The first one is in both cases located between the TS and the start codon. The second one was found in *ydhT* and *kdgA*, respectively. Novel identified targets that are directly repressed by the presence of CcpA are the *ydhMNOPQRST* operon for which a putative cre site was detected present upstream of *ydhM* 7-bp downstream of *ydgN* for which a cre site was identified within the ORF.

<table>
<thead>
<tr>
<th>gene/operon</th>
<th>cre position</th>
<th>cre sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>sdhCAB</em></td>
<td>-40</td>
<td>ATGTACGGCTTTCTT</td>
</tr>
<tr>
<td><em>glpFK</em></td>
<td>-27</td>
<td>TTGACACGGCTTTTAC</td>
</tr>
<tr>
<td><em>kdgA</em></td>
<td>ORF</td>
<td>ATGGAAAGCGCTGACAT</td>
</tr>
<tr>
<td><em>glpTQ</em></td>
<td>4</td>
<td>AAGAAAGCGCTATCAT</td>
</tr>
<tr>
<td><em>ydh</em></td>
<td>7</td>
<td>ATGTAAAGCGTTTTAAT</td>
</tr>
<tr>
<td><em>ydhT</em></td>
<td>ORF</td>
<td>TTGGAAGCGGTATCAT</td>
</tr>
<tr>
<td><em>phoPR</em></td>
<td>22</td>
<td>ATGGAAAGCGTATCAT</td>
</tr>
<tr>
<td><em>rocG</em></td>
<td>45</td>
<td>TTAAAGCGCTTACAT</td>
</tr>
<tr>
<td><em>bglP</em></td>
<td>-36</td>
<td>ATGAAAGCGTTGACAT</td>
</tr>
</tbody>
</table>

*Fig. 5.* Alignment of the cre sites of genes that are CcpA regulated throughout growth. cre position reflects the distance from the centre of the cre to the transcription start site.
**Location and composition of cre sites**

Comparison of the promoter sequences of the CcpA-activated gene clusters showed that *yhbI*yhcAB share a number of nucleotides suggesting a possible cre site and a UAR similar to *ackA* and *pta* [176] (Fig. 6). The distance from the centre of the cre towards the TS site, -57 bp, is identical in the promoter regions of *ackA* and *pta*, while for *yhbl* it could be located at -57 or 74-bp upstream of the putative TS. In the promoter region of *manR* two cre elements were identified of which one is located 57-bp upstream of the putative TS, determined on basis of the canonical -10 and hexameric -35 sequences. For the other activated genes/operons no such UAR was found.

![Fig. 6. Alignment of the promoter regions of the CcpA-activated genes.](image)

The cre sites of the CcpA-repressed targets were found to partly overlap the -35, -10, TS site or are located in the ORF (Table 1). Notably, genes repressed throughout growth (Fig. 5) share a conserved cre sequence. *bgIP* is an exception as it has a very similar cre consensus, although it is repressed only in the early exponential phase.

**Table 1 (next page).** Genes and operons putatively regulated by CcpA by binding to a (predicted) cre sequence (in grey when predicted in this study). Genes located within one operon are listed from the promoter onwards. Gene expression ratios ΔccpA/wt of activated (minus values) and repressed (plus values) genes are in bold when they meet the cut off of 2 and are indicated by 1 when the data did not meet the statistical criteria as described in Experimental procedures. The distance/location of the centre of the cre sites to the (in grey when predicted) transcription start is listed (Pos. cre1 or 2). Promoter characteristics such as the -35 and -10 sequences (both in capitals and italics), transcription start sites (in capitals, bold and underlined) and start codon (capitals and bold) are indicated in the (putative) cre sequences. Genes for which to promoter sequence could be predicted are indicated in light-grey at the end of the table. Growth phases E, M, T and S as described in Fig. 1.
<table>
<thead>
<tr>
<th>Genes</th>
<th>E</th>
<th>M</th>
<th>T</th>
<th>S</th>
<th>pos. cre 1</th>
<th>pos. cre 2</th>
<th>Seq. cre 1</th>
<th>Seq cre 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>opuAA</td>
<td>-2.1</td>
<td>-2.4</td>
<td>-1.8</td>
<td>1</td>
<td>-143</td>
<td>-181</td>
<td>aataaaagcgttttcaaa</td>
<td></td>
</tr>
<tr>
<td>opuAB</td>
<td>-2.6</td>
<td>-3.7</td>
<td>-1.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>opuAC</td>
<td>-2.1</td>
<td>-2.6</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>opuE</td>
<td>-2.2</td>
<td>-2.4</td>
<td>-1.5</td>
<td>-1.9</td>
<td>-104</td>
<td>tttgaaaagcgttttatt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ilvB</td>
<td>1</td>
<td>1</td>
<td>-3</td>
<td>1</td>
<td>-88</td>
<td>atgaaaagcgtataca</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ilvH</td>
<td>1</td>
<td>1</td>
<td>-4.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ilvC</td>
<td>1</td>
<td>1</td>
<td>-8.3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leuA</td>
<td>1</td>
<td>1</td>
<td>-2.8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leuB</td>
<td>1</td>
<td>1</td>
<td>-4.0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leuC</td>
<td>1</td>
<td>1</td>
<td>-4.2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leuD</td>
<td>1</td>
<td>1</td>
<td>-2.7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yhbl</td>
<td>-2.6</td>
<td>-3.6</td>
<td>1</td>
<td>1</td>
<td>-74</td>
<td>caagaaaaacgcttcgaa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yhbj</td>
<td>-2.0</td>
<td>-2.7</td>
<td>-3.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yhCA</td>
<td>-2.4</td>
<td>-2.6</td>
<td>-2.3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yhCB</td>
<td>-2.0</td>
<td>-2.8</td>
<td>-1.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ackA</td>
<td>-3.3</td>
<td>-2.9</td>
<td>-3.4</td>
<td>1</td>
<td>-117</td>
<td>-57</td>
<td>attgtaacgtttcatcaat</td>
<td>attgtaacgtttcatcaat</td>
</tr>
<tr>
<td>moaB</td>
<td>-1.9</td>
<td>-1.8</td>
<td>-3.3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>manR</td>
<td>1</td>
<td>-2.4</td>
<td>1</td>
<td>1</td>
<td>-57</td>
<td>1</td>
<td>ctttttccgaagcttcg</td>
<td>aatgtaaacGtttctta</td>
</tr>
<tr>
<td>manP</td>
<td>-2.1</td>
<td>-3.8</td>
<td>1</td>
<td>-1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>manA</td>
<td>-6.8</td>
<td>-9.9</td>
<td>1</td>
<td>-2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yjDF</td>
<td>1</td>
<td>-1.6</td>
<td>1</td>
<td>-1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pta</td>
<td>1</td>
<td>-2.7</td>
<td>-1.7</td>
<td>2.1</td>
<td>-55</td>
<td>tatgaaacgcctataatg</td>
<td>aaAGaaacgttttgtta</td>
<td></td>
</tr>
<tr>
<td><strong>Repressed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ywcJ</td>
<td>1</td>
<td>1</td>
<td>1.9</td>
<td>1</td>
<td>-143</td>
<td></td>
<td>taagaatagcgttcatc</td>
<td></td>
</tr>
<tr>
<td>citB</td>
<td>1</td>
<td>1.9</td>
<td>4.9</td>
<td>-2.6</td>
<td>-106</td>
<td></td>
<td>tgagagacgttaacagt</td>
<td></td>
</tr>
<tr>
<td>levD</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td>1</td>
<td>-44</td>
<td>aatgaaacgcctaaca</td>
<td></td>
<td></td>
</tr>
<tr>
<td>levF</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>-1.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>levG</td>
<td>1</td>
<td>1</td>
<td>2.0</td>
<td>-2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sacC</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sdhC</td>
<td>2.5</td>
<td>3.4</td>
<td>1.7</td>
<td>2.4</td>
<td>-40</td>
<td>ORF</td>
<td>tatgtaacgttTIG</td>
<td>TGTAACCGGTATCA</td>
</tr>
<tr>
<td>sdhA</td>
<td>1.6</td>
<td>2.8</td>
<td>1.6</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sdhB</td>
<td>1</td>
<td>2.1</td>
<td>1.8</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yqgY</td>
<td>1.6</td>
<td>1.7</td>
<td>2.4</td>
<td>1.5</td>
<td>-37</td>
<td>aatgaaatgTTTACAtt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genes</td>
<td>E</td>
<td>M</td>
<td>T</td>
<td>S</td>
<td>pos. cre 1</td>
<td>pos. cre 2</td>
<td>Seq. cre 1</td>
<td>Seq. cre 2</td>
</tr>
<tr>
<td>-------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>bgLP</td>
<td>2.5</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>-36</td>
<td></td>
<td>aatgaaagcgTTGACAtc</td>
<td></td>
</tr>
<tr>
<td>yxiE</td>
<td>1.9</td>
<td>1.9</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ysbA</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7.9</td>
<td>-36</td>
<td></td>
<td>cttgtaagcgcttttaTAA</td>
<td></td>
</tr>
<tr>
<td>glpF</td>
<td>3.3</td>
<td>3.9</td>
<td>2.7</td>
<td>4.3</td>
<td>-27</td>
<td>ORF</td>
<td>aTTGACACgctttcatg</td>
<td></td>
</tr>
<tr>
<td>glPK</td>
<td>2.6</td>
<td>3.0</td>
<td>2.7</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acoR</td>
<td>3.7</td>
<td>4.8</td>
<td>5.2</td>
<td>-2.4</td>
<td>-27</td>
<td>gTTGAAAGgcgttttatg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ssPH</td>
<td>3.7</td>
<td>3.2</td>
<td>1</td>
<td>-1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>msmX</td>
<td>1</td>
<td>1</td>
<td>2.6</td>
<td>1</td>
<td>-15</td>
<td></td>
<td>taagaaagcgTACAA T</td>
<td></td>
</tr>
<tr>
<td>sacP</td>
<td>2.0</td>
<td>2.2</td>
<td>1</td>
<td>1</td>
<td>-2</td>
<td></td>
<td>GATgaaagcGtttcatg</td>
<td></td>
</tr>
<tr>
<td>sacA</td>
<td>8.7</td>
<td>3.9</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ywdA</td>
<td>3.7</td>
<td>1.5</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ycsA</td>
<td>1.7</td>
<td>1.6</td>
<td>2.4</td>
<td>-1.6</td>
<td>-1</td>
<td></td>
<td>atagaagaGcctacgat</td>
<td></td>
</tr>
<tr>
<td>sucC</td>
<td>2.3</td>
<td>3.0</td>
<td>2.6</td>
<td>1</td>
<td>-1</td>
<td></td>
<td>AA1gaaagcGcgtctat</td>
<td></td>
</tr>
<tr>
<td>sucD</td>
<td>2.6</td>
<td>2.0</td>
<td>4.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yvfK</td>
<td>1</td>
<td>1.6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
<td>ttcttaAgcgctttcat</td>
<td></td>
</tr>
<tr>
<td>yvfM</td>
<td>2.4</td>
<td>4.2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lacA</td>
<td>3</td>
<td>3.7</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glpT</td>
<td>2.9</td>
<td>3.7</td>
<td>4.4</td>
<td>2.9</td>
<td>4</td>
<td></td>
<td>taagaAgcgctacatg</td>
<td></td>
</tr>
<tr>
<td>glpQ</td>
<td>2.6</td>
<td>4.4</td>
<td>5.9</td>
<td>5.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>amyE</td>
<td>1.7</td>
<td>2.2</td>
<td>3.0</td>
<td>1</td>
<td>4</td>
<td></td>
<td>aatgAagcgttaacaa</td>
<td></td>
</tr>
<tr>
<td>yvfH</td>
<td>1</td>
<td>1</td>
<td>2.9</td>
<td>1</td>
<td>4</td>
<td></td>
<td>atagaAacgcttacat</td>
<td></td>
</tr>
<tr>
<td>ytkA</td>
<td>1</td>
<td>1</td>
<td>5.0</td>
<td>1</td>
<td>5</td>
<td></td>
<td>gcgtTaagcgtttctac</td>
<td></td>
</tr>
<tr>
<td>pmi</td>
<td>1.7</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td>5</td>
<td></td>
<td>atagGgaaagctttctt</td>
<td></td>
</tr>
<tr>
<td>ysfC</td>
<td>1</td>
<td>1</td>
<td>1.8</td>
<td>-2.0</td>
<td>5</td>
<td></td>
<td>atagAagcgtttttga</td>
<td></td>
</tr>
<tr>
<td>rbsR</td>
<td>5.9</td>
<td>3.8</td>
<td>11.3</td>
<td>1</td>
<td>6</td>
<td></td>
<td>taGttaaacggttacata</td>
<td></td>
</tr>
<tr>
<td>rbsA</td>
<td>3.9</td>
<td>2.7</td>
<td>3.6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rbsB</td>
<td>5.5</td>
<td>3.3</td>
<td>6.7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rbsC</td>
<td>5.4</td>
<td>3.7</td>
<td>5.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rbsD</td>
<td>4.5</td>
<td>3.0</td>
<td>7.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rbsK</td>
<td>2.4</td>
<td>2.0</td>
<td>3.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhM</td>
<td>1</td>
<td>1.7</td>
<td>1</td>
<td>4.6</td>
<td>7</td>
<td></td>
<td>aatgtaagcgtttaata</td>
<td></td>
</tr>
<tr>
<td>ydhN</td>
<td>1</td>
<td>1.8</td>
<td>1</td>
<td>-2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhO</td>
<td>3.3</td>
<td>3.3</td>
<td>1</td>
<td>6.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Temporal regulation by CcpA

<table>
<thead>
<tr>
<th>Genes</th>
<th>E</th>
<th>M</th>
<th>T</th>
<th>S</th>
<th>pos. cre 1</th>
<th>pos. cre 2</th>
<th>Seq. cre 1</th>
<th>Seq cre 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ydhP</td>
<td>2.5</td>
<td>3.3</td>
<td>1</td>
<td>8.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhQ</td>
<td>2.5</td>
<td>2.0</td>
<td>1.8</td>
<td>7.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhR</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td>5.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhS</td>
<td>1</td>
<td>3.0</td>
<td>1</td>
<td>6.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhT</td>
<td>1</td>
<td>1</td>
<td>1.8</td>
<td>2.8</td>
<td>ORF</td>
<td></td>
<td></td>
<td>attgaagcgcatctg</td>
</tr>
<tr>
<td>yfjG</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1.5</td>
<td></td>
<td>6</td>
<td></td>
<td>ttagaaagcggttaca</td>
</tr>
<tr>
<td>yfjH</td>
<td>1</td>
<td>2.2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ndk</td>
<td>2.3</td>
<td>1.8</td>
<td>4.4</td>
<td>1</td>
<td></td>
<td>15</td>
<td></td>
<td>ttagaaagccctacata</td>
</tr>
<tr>
<td>kdgR</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td>17</td>
<td></td>
<td>ttagaaatcggttca</td>
</tr>
<tr>
<td>kdgK</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kdgA</td>
<td>1</td>
<td>2.3</td>
<td>1</td>
<td>2.2</td>
<td>ORF</td>
<td></td>
<td></td>
<td>ttagaaagcagtac</td>
</tr>
<tr>
<td>kdgT</td>
<td>1</td>
<td>1</td>
<td>2.8</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yxiO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2.1</td>
<td></td>
<td>21</td>
<td></td>
<td>catgtaaacgtaa</td>
</tr>
<tr>
<td>odhA</td>
<td>3.8</td>
<td>2.1</td>
<td>1.9</td>
<td>1</td>
<td></td>
<td>22</td>
<td></td>
<td>gttgtaagcgttta</td>
</tr>
<tr>
<td>odhB</td>
<td>3.9</td>
<td>2.5</td>
<td>1.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phoP</td>
<td>1</td>
<td>1.7</td>
<td>2.5</td>
<td>1.6</td>
<td></td>
<td>22</td>
<td></td>
<td>aatgtaagcgtacata</td>
</tr>
<tr>
<td>phoR</td>
<td>1</td>
<td>1</td>
<td>2.2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yvdG</td>
<td>1</td>
<td>1</td>
<td>1.6</td>
<td>-3.1</td>
<td></td>
<td>30</td>
<td></td>
<td>gttgtaacgccttat</td>
</tr>
<tr>
<td>cstA</td>
<td>1</td>
<td>1</td>
<td>1.8</td>
<td>1</td>
<td></td>
<td>31</td>
<td></td>
<td>aATGaatggcgttca</td>
</tr>
<tr>
<td>rocG</td>
<td>3.4</td>
<td>5.6</td>
<td>10.3</td>
<td>3.2</td>
<td></td>
<td>45</td>
<td></td>
<td>tttaaacgcctcatta</td>
</tr>
<tr>
<td>yrO</td>
<td>1</td>
<td>2.8</td>
<td>3.1</td>
<td>1</td>
<td></td>
<td>44</td>
<td></td>
<td>ttagaaagcgattcat</td>
</tr>
<tr>
<td>araA</td>
<td>1</td>
<td>1</td>
<td>2.9</td>
<td>1.8</td>
<td></td>
<td>69</td>
<td></td>
<td>ttagaaagcgtttaa</td>
</tr>
<tr>
<td>araB</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2.6</td>
<td>ORF</td>
<td></td>
<td></td>
<td>tgaaacgattac</td>
</tr>
<tr>
<td>araL</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>araP</td>
<td>1</td>
<td>1</td>
<td>1.6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abfA</td>
<td>1</td>
<td>1.7</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>citZ</td>
<td>2.2</td>
<td>3.7</td>
<td>3.7</td>
<td>1</td>
<td></td>
<td>86</td>
<td></td>
<td>aatgtaacgctctt</td>
</tr>
<tr>
<td>mdh</td>
<td>1</td>
<td>1.5</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qcrA</td>
<td>2.1</td>
<td>2.1</td>
<td>2.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>ttagtaagcgttatc</td>
</tr>
<tr>
<td>qcrB</td>
<td>1.8</td>
<td>1.8</td>
<td>2.6</td>
<td>1</td>
<td>ORF</td>
<td></td>
<td></td>
<td>ttagtaacgcgttca</td>
</tr>
<tr>
<td>qcrC</td>
<td>2.8</td>
<td>3.2</td>
<td>3.2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>treP</td>
<td>1</td>
<td>4.1</td>
<td>1</td>
<td>1</td>
<td>ORF</td>
<td></td>
<td></td>
<td>ttagaaagcgcgg</td>
</tr>
<tr>
<td>treA</td>
<td>6.1</td>
<td>10.5</td>
<td>1.2</td>
<td>-2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>treR</td>
<td>1</td>
<td>1.8</td>
<td>1</td>
<td>-1.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genes</td>
<td>E</td>
<td>M</td>
<td>T</td>
<td>S</td>
<td>pos. cre 1</td>
<td>pos. cre 2</td>
<td>Seq. cre 1</td>
<td>Seq cre 2</td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
<td>-----------</td>
<td>------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>xylA</td>
<td>1.7</td>
<td>1</td>
<td>1</td>
<td>ORF</td>
<td>TGGAAGCGCAAAACAAA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gntP</td>
<td>2.4</td>
<td>1</td>
<td>-3.5</td>
<td>1</td>
<td>ORF</td>
<td>TGAAGCGGTACCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gntZ</td>
<td>2.3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>ORF</td>
<td>AGCGGTACCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rocD</td>
<td>5.3</td>
<td>3.7</td>
<td>6.7</td>
<td>1</td>
<td>ORF</td>
<td>GAGGAATTCGCTGCTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rocE</td>
<td>1</td>
<td>1.6</td>
<td>3.7</td>
<td>1</td>
<td>ORF</td>
<td>ATCGGAACCGTTTCTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rocF</td>
<td>2.5</td>
<td>1.9</td>
<td>5.0</td>
<td>1</td>
<td>ORF</td>
<td>GAGGAATTCGCTGCTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yneN</td>
<td>1.9</td>
<td>3.7</td>
<td>-2.0</td>
<td>ORF</td>
<td>CTTGAAAACGCTATAGG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acsA</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td>ORF</td>
<td>CTGGAACCAGTCAGC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ydhT</td>
<td>1</td>
<td>1.8</td>
<td>2.8</td>
<td>1</td>
<td>ORF</td>
<td>ATTGAAGCGCATACGTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yknW</td>
<td>1</td>
<td>1</td>
<td>3.7</td>
<td>ORF</td>
<td>ATTGAAGCGCATACGTG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yknY</td>
<td>1</td>
<td>1</td>
<td>3.8</td>
<td>ORF</td>
<td>ATTGAAGCGCATACGTG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yknZ</td>
<td>1</td>
<td>1</td>
<td>4.0</td>
<td>ORF</td>
<td>ATTGAAGCGCATACGTG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>spoVG</td>
<td>2.4</td>
<td>1.6</td>
<td>3.9</td>
<td>-2.6</td>
<td>ORF</td>
<td>CGTGAAGCGATGCAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sunA</td>
<td>3.1</td>
<td>4.1</td>
<td>1</td>
<td>ORF</td>
<td>AGCGGTACCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sunT</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td>ORF</td>
<td>ATCGGAACCGTTTCTT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yoiI</td>
<td>3.4</td>
<td>4.8</td>
<td>1.9</td>
<td>1</td>
<td>ORF</td>
<td>ACGTGAAGCGTTAAAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bdbB</td>
<td>1</td>
<td>6.5</td>
<td>2.5</td>
<td>1</td>
<td>ORF</td>
<td>AGCGGTACCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ypiB</td>
<td>1</td>
<td>1</td>
<td>2.4</td>
<td>1</td>
<td>ORF</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ypiF</td>
<td>1</td>
<td>2.6</td>
<td>4.3</td>
<td>1</td>
<td>ORF</td>
<td>ACCGTGAAAGCGTTAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ycgM</td>
<td>1.1</td>
<td>1</td>
<td>1</td>
<td>-9.7</td>
<td>ORF</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ycgN</td>
<td>6.5</td>
<td>6.0</td>
<td>8.7</td>
<td>-4.2</td>
<td>ORF</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ycgO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1.8</td>
<td>ORF</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yxH</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td>1</td>
<td>intergenic</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yerD</td>
<td>1</td>
<td>3.7</td>
<td>1</td>
<td>intergenic</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pckA</td>
<td>1</td>
<td>1</td>
<td>1.9</td>
<td>-5.3</td>
<td>intergenic</td>
<td>ATTGATAACCCTCTCTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ptsG</td>
<td>1</td>
<td>1</td>
<td>-2.2</td>
<td>2.7</td>
<td>ORF</td>
<td>TCGTAAAGCGTTAACCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ptsH</td>
<td>1</td>
<td>1</td>
<td>2.4</td>
<td>ORF</td>
<td>TCGTAAAGCGTTAACCT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ptsI</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td>ORF</td>
<td>TCGTAAAGCGTTAACCT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dctP</td>
<td>4.9</td>
<td>6.9</td>
<td>1</td>
<td>-11</td>
<td>intergenic</td>
<td>TCGTAAAGCGTTAACCT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

We compared the transcriptomes of wild-type and a ccpA mutant strain of *B. subtilis* 168 at four different time points of growth in rich medium supplemented with glucose. The lack of the CcpA protein is sensed by the cells from early stages of growth as the expression of genes involved in carbohydrate metabolism was clearly influenced in the mutant. The effect developed further in time involving other functional categories of proteins when cells entered the later stages of exponential growth as shown by FIVA analysis. These effects are caused by direct or indirect CcpA action and that is why a distinction should be made between the CcpA regulon and the CcpA-modulon, respectively. The CcpA-modulon would encompass the genes whose regulation results, for example, from the altered cell physiology due to slower glucose uptake caused by lack of functional CcpA. Indeed, many genes with no apparent cre site were differentially expressed in our, as well as in other transcriptome studies [18,150,181,276] which make them likely candidates for the CcpA-modulon. However, another yet unknown mode of regulation by CcpA could be present which would allow classifying at least some of these genes as direct CcpA targets. The *tdh* and *yerA* genes are just two interesting examples. They have no predicted cre sites but show a clear repression throughout time in our experiments.

In addition, the cre sequence itself is still not precisely defined yet, since the search performed with the WTGNAANCGNWNCCW consensus [173] retrieved only 126 genes with possible cre’s but still novel functional or putative cre’s were reported [110,171] and an altered cre consensus WTGAAARCGYTTWW (W=A or T; N=any base; R=A or G; Y=C or T) was proposed [171]. Taking this into account and the fact that DNA binding properties of CcpA may differ at different stages of growth, we performed our own cre search within the whole *B. subtilis* genome using known cre sites from DBTBS and applying a non-stringent, 0.85 score cut-off. This search resulted in more than 300 hits. *In vitro* studies [79] have shown that CcpA binding to cre sites is dependent on the acidity of the buffer used. It is possible that during growth on a rich glucose medium internal pH changes in time, which could influence recognition and CcpA-binding to its target sequence. Moreover, the availability and concentrations of various cofactors change in time, which would also affect CcpA-DNA binding or the CcpA interactions with the transcription machinery. The latter mechanism has been already suggested [126,127] and a NADP molecule appears to be one of the putative cofactors in this process [126,127]. Under our experimental conditions the CcpA protein in the wild-type strain was present at all stages of growth (Western blot analysis, data not shown), even in the stationary phase when glucose was already absent from the medium. This demonstrates that CcpA is constitutively expressed and corresponds well with previously published data [174]. Most likely the temporal regulation of different groups of genes by CcpA results from the difference in quantitative and qualitative availability of cofactors, the presence of which would orchestrate alternative modes of CcpA action. In the
exponential phase of growth the derivatives of glucose metabolism probably modulate CcpA interaction and when glucose is depleted other coeffectors could take over the role in CcpA-dependent gene regulation. The interplay of two distinct coeffectors, G6P and HPt-Ser-P, has been already shown for the gnt operon [172] and the xyl operon [79]. In the latter case the acidity played a role in selective determination of the CcpA interaction with cofactors. Besides, both operons contain multiple cre sites present in the promoter regions and structural genes. We have identified several genes and operons also containing more than one cre site, whose expression was altered when comparing the wild-type and the ccppA mutant strain. These included: ptsG, sdhC, glpF, kdg, rbs and the ydh operon. The ydhMNO genes are involved in the transport of oligo-β-mannosides [207].

The ilvBHC-levABCD has been also shown to be activated by CcpA [225,244]. This operon is repressed by CodY and TnrA, but during transition phase the operon is activated in the presence CcpA in the wild-type strain. This is likely due to the shortage of amino acids or GTP, which results in a loss of the repression by CodY. This shows another aspect of temporal regulation by CcpA, namely interplay of distinct regulators of gene expression whose presence differs in the course of time and bacterial growth and depends on the availability of energy sources.

Another interesting observation emerging from our time series data concerns the strong activation of the yvd gene cluster and the malA-yfiA-malP operon exclusively in the stationary phase. This coincides with the lack of repression of amyE, the expression of which was repressed in the exponential phase. The α-amylase, coded by the amyE gene, hydrolyzes extracellular polysaccharides to smaller sugars, such as dextrans, maltose and glucose which can be transported through the cell membrane. The mal operon and the yvdE-pgcM cluster (yvdGHl=mdxEFG) are responsible for the maltose and maltodextrin utilisation, respectively [218,272]. Glucose repression of these genes has been shown to be mediated by CcpA and cre [272]. Interestingly, the yvdE-pgcM gene cluster contains a cre in the yvdG (mdxE) intergenic region and shows the same expression pattern in time as the mal operon, suggesting that CcpA is also involved in the regulation of the maltodextrin utilisation.

With our approach we have identified three new gene clusters, namely yhb, man and opu, and one gene, manR, that are activated by CcpA. Alignment of the cre site and its UAR of the well-studied CcpA target ackA showed a certain degree of homology with the promoter region of yhbl (Fig. 6), suggesting a similar type of regulation. No homology with the UAR was obtained with the promoter region of manR, which has a putative cre located at the same location as ackA. Assuming that the cre site of yhbl is located at -57 bp from the putative TS site, all these activated genes would follow the model proposed by Kim et al. [127]. These authors have shown that the CcpA binding site has to be located at the same site of the helix as the TS site. Except for opuE, activation of expression of all the CcpA activated genes was lost in the stationary growth suggesting that activation might be strictly glucose dependent.

In this study, we show dynamic regulation by CcpA in time. Analysis of the data has identified potential new members of the CcpA regulon and hints at potential differential
binding specificity that may depend on the relative presence or absence of perhaps yet unknown cofactors. We show that some transcriptional effects might be attributed to a difference in glucose consumption, though also glucose-independent regulation by CcpA may occur.

**Supplementary material**
Supplementary tables and graphs mentioned in this manuscript as well as the slide images and raw data are available from http://molgen.biol.rug.nl/publication/ccpa_bsub_data/

**Acknowledgements**
The authors thank Evert-Jan Blom for providing the possibility of using the FIVA software before publication and Sacha van Hijum for his support in the microarray data handling process and the preparation of the website with supplementary material. We are also grateful to Aldert Zomer for the skilful support with MotifLocator, Wiep Klaas Smits for helpful discussions and Siger Holssappel for technical assistance in determining the glucose concentrations.

This work was supported by IOP genomics grant IGE01018 from the ‘Innovatiegerichte Onderzoeksprogramma’s’ of the Dutch Ministry of Economic Affairs.