



# Overlap of Promoter Recognition Specificity of Stress Response Sigma Factors SigD and SigH in *Corynebacterium glutamicum* ATCC 13032

Hana Dostálová<sup>1</sup>, Tobias Busche<sup>2†</sup>, Jiří Holátko<sup>1</sup>, Lenka Rucká<sup>1</sup>, Václav Štěpánek<sup>1</sup>, Ivan Barvík<sup>3</sup>, Jan Nešvera<sup>1</sup>, Jörn Kalinowski<sup>2</sup> and Miroslav Pátek<sup>1\*</sup>

<sup>1</sup> Institute of Microbiology of the CAS, v. v. i., Prague, Czechia, <sup>2</sup> Centrum für Biotechnologie, Universität Bielefeld, Bielefeld, Germany, <sup>3</sup> Institute of Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czechia

## OPEN ACCESS

### Edited by:

Daniela De Biase,  
Sapienza University of Rome, Italy

### Reviewed by:

Andreas Burkovski,  
Friedrich-Alexander-Universität  
Erlangen-Nürnberg, Germany  
Saori Kosono,  
The University of Tokyo, Japan

### \*Correspondence:

Miroslav Pátek  
patek@biomed.cas.cz

### † Present address:

Tobias Busche,  
Institute of Biology and Microbiology,  
Freie Universität Berlin, Berlin,  
Germany

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

Received: 13 September 2018

Accepted: 17 December 2018

Published: 09 January 2019

### Citation:

Dostálová H, Busche T, Holátko J,  
Rucká L, Štěpánek V, Barvík I,  
Nešvera J, Kalinowski J and Pátek M  
(2019) Overlap of Promoter  
Recognition Specificity of Stress  
Response Sigma Factors SigD  
and SigH in *Corynebacterium  
glutamicum* ATCC 13032.  
Front. Microbiol. 9:3287.  
doi: 10.3389/fmicb.2018.03287

*Corynebacterium glutamicum* ATCC 13032 harbors five sigma subunits of RNA polymerase belonging to Group IV, also called extracytoplasmic function (ECF)  $\sigma$  factors. These factors  $\sigma^C$ ,  $\sigma^D$ ,  $\sigma^E$ ,  $\sigma^H$ , and  $\sigma^M$  are mostly involved in stress responses. The role of  $\sigma^D$  consists in the control of cell wall integrity. The  $\sigma^D$  regulon is involved in the synthesis of components of the mycomembrane which is part of the cell wall in *C. glutamicum*. RNA sequencing of the transcriptome from a strain overexpressing the *sigD* gene provided 29 potential  $\sigma^D$ -controlled genes and enabled us to precisely localize their transcriptional start sites. Analysis of the respective promoters by both *in vitro* transcription and the *in vivo* two-plasmid assay confirmed that transcription of 11 of the tested genes is directly  $\sigma^D$ -dependent. The key sequence elements of all these promoters were found to be identical or closely similar to the motifs -35 GTAAC<sup>A/G</sup> and -10 GAT. Surprisingly, nearly all of these  $\sigma^D$ -dependent promoters were also active to a much lower extent with  $\sigma^H$  *in vivo* and one (Pcg0607) also *in vitro*, although the known highly conserved consensus sequence of the  $\sigma^H$ -dependent promoters is different (-35 GGAA<sup>T/C</sup> and -10 GTT). In addition to the activity of  $\sigma^H$  at the  $\sigma^D$ -controlled promoters, we discovered separated or overlapping  $\sigma^A$ - or  $\sigma^B$ -regulated or  $\sigma^H$ -regulated promoters within the upstream region of 8 genes of the  $\sigma^D$ -regulon. We found that phenol in the cultivation medium acts as a stress factor inducing expression of some  $\sigma^D$ -dependent genes. Computer modeling revealed that  $\sigma^H$  binds to the promoter DNA in a similar manner as  $\sigma^D$  to the analogous promoter elements. The homology models together with mutational analysis showed that the key amino acids, Ala 60 in  $\sigma^D$  and Lys 53 in  $\sigma^H$ , bind to the second nucleotide within the respective -10 promoter elements (GAT and GTT, respectively). The presented data obtained by integrating *in vivo*, *in vitro* and *in silico* approaches demonstrate that most of the  $\sigma^D$ -controlled genes also belong to the  $\sigma^H$ -regulon and are also transcribed from the overlapping or closely located housekeeping ( $\sigma^A$ -regulated) and/or general stress ( $\sigma^B$ -regulated) promoters.

**Keywords:** *Corynebacterium glutamicum*, stress response, RNA-seq, *in vitro* transcription, sigma factor, promoter

## INTRODUCTION

Most bacterial genes are transcribed by an RNA polymerase (RNAP) holoenzyme that includes a primary sigma subunit ( $\sigma$  factor) responsible for the transcription of housekeeping genes. In addition to the primary  $\sigma$  factor, a number of alternative sigma factors control various functions of the bacterial cells under specific nutritional, growth and environmental conditions. *Corynebacterium glutamicum* is a Gram-positive bacterium, which is applied in industrial biotechnology and also considered a model species for other corynebacteria as well as related actinobacterial genera such as *Mycobacterium* and *Rhodococcus*. The *C. glutamicum* genome encodes 7  $\sigma$  factors: A primary  $\sigma^A$ , a primary-like  $\sigma^B$  and five  $\sigma$  factors of Group IV (Gruber and Gross, 2003) also called extracytoplasmic function (ECF)  $\sigma$  factors ( $\sigma^C$ ,  $\sigma^D$ ,  $\sigma^E$ ,  $\sigma^H$ , and  $\sigma^M$ ), which are involved in various stress responses.

Gene expression in bacteria is controlled by a complex regulatory network composed of protein transcription regulators, regulatory RNAs as well as low-molecular-weight metabolites. Sigma subunits of RNAP that are responsible for transcription initiation play an important role in this network. Groups of genes ( $\sigma$  factor regulons, called also sigmulons) controlled by different  $\sigma$  factors were found to constitute a modular network in *Pseudomonas aeruginosa* (Schulz et al., 2015). Individual separated sigmulons controlled by more than 25  $\sigma$  factors display limited overlaps which probably allows this versatile bacterium to efficiently regulate cell processes and adapt to various environments as well as to the lifestyle of an opportunistic pathogen. The functions of the regulons of alternative  $\sigma$  factors in *P. aeruginosa* are significantly connected by the action of global transcriptional regulators which modulate transcription levels from the respective promoters (Binder et al., 2016). According to this hypothesis, modularity of regulatory networks may ensure rapid evolution (Binder et al., 2016). In contrast to previous notion concerning *P. aeruginosa*, considerable overlaps of  $\sigma$  factor regulons observed in many bacteria display regulatory redundancies and the ability to fine-tune responses to various environmental stimuli and combined stress conditions. Some genes belong to two or more sigmulons, since they are either transcribed from several promoters recognized by different alternative  $\sigma$  factors, or are transcribed with two or more different holoenzymes (RNAP core+ $\sigma$ ) which drive transcription from the same promoter (Luo and Helmann, 2009). Microarray analysis revealed extensive coregulation of the genes involved in many cell processes, which are co-operatively controlled by four alternative  $\sigma$  factors ( $\sigma^B$ ,  $\sigma^C$ ,  $\sigma^H$ , and  $\sigma^L$ ) in *Listeria monocytogenes* (Chaturongakul et al., 2011). Significant regulatory overlaps between the target genes of three  $\sigma$  factors ( $\sigma^M$ ,  $\sigma^W$ , and  $\sigma^X$ ) involved in cell surface stress responses were also found in *Bacillus subtilis* (Jordan et al., 2008). A single promoter of the *sigB* gene encoding an alternative  $\sigma$  in *Mycobacterium tuberculosis* was recognized by as many as three different ECF  $\sigma$  factors ( $\sigma^E$ ,  $\sigma^H$ , and  $\sigma^L$ ) (Dainese et al., 2006). Overlapping  $\sigma$  factor specificity in promoter recognition can be therefore considered to be a common feature of regulatory networks in bacteria.

Various types of stress conditions which bacteria encounter in the environment (e.g., heat, cell surface and oxidative stresses) often result in damaged proteins, which are then degraded by proteases or re-folded by chaperones during the stress response. Expression of some genes encoding proteases (*clpB*) and chaperones (*dnaK* and *dnaJ2*) was found to be activated by both  $\sigma^H$  (involved in heat and oxidative stress response) and  $\sigma^E$  (involved in cell surface stress response) in *C. glutamicum* (Šilar et al., 2016). The *clgR* gene encoding a heat-stress-responsive regulator was also found to be  $\sigma^H$  and  $\sigma^E$ -dependent (Šilar et al., 2016). ClgR controls expression of several genes involved in protein quality control, and the function of ECF  $\sigma$  factors and the transcriptional regulator is thus integrated into a network. The same  $\sigma^H/\sigma^E$  overlap was found in the transcription initiation of the *C. glutamicum sigB* gene, encoding the general stress response sigma factor  $\sigma^B$  (Dostálová et al., 2017). The possible activity of both  $\sigma^H$  and  $\sigma^M$  (involved in oxidative stress response) at the same *C. glutamicum* promoters was suggested, but has not yet been proven (Nakunst et al., 2007; Šilar et al., 2016; Dostálová et al., 2017), although the respective promoter consensus sequences seem very similar. Promoters of  $\sigma^C$ -regulated genes in *C. glutamicum* were found to be different from other promoters recognized by ECF  $\sigma$  factors, and  $\sigma^C$  regulon, which includes genes involved in the aerobic respiratory chain in *C. glutamicum*, seems to be separated from the other  $\sigma$  factor regulons (Toyoda and Inui, 2016).

Many stress response genes of *C. glutamicum* were found to be transcribed not only from stress promoters, but also from overlapping or closely located housekeeping promoters controlled by  $\sigma^A$  and/or  $\sigma^B$  (Engels et al., 2004; Ehira et al., 2009; Busche et al., 2012; Šilar et al., 2016). Such transcription of stress-response genes from multiple promoters also seems to be frequent in other bacteria (Seo et al., 2012; Cho et al., 2014). The activity of  $\sigma^A$  and/or  $\sigma^B$  under unstressed conditions probably ensures a basal level of expression of the respective genes which also play important roles in fast-growing bacterial cells.

We have recently found by using RNA-seq that  $\sigma^D$ -dependent genes are involved in the synthesis of components of the mycomembrane, which is a part of the cell wall in *C. glutamicum* ATCC 13032. The  $\sigma^D$  regulon, including 29 genes, thus contributes to the maintenance of cell wall integrity (Taniguchi et al., 2017). Similar group of  $\sigma^D$ -dependent genes was also detected in the strain *C. glutamicum* R using different techniques (Toyoda and Inui, 2018). In *C. glutamicum* R, the *sigD* overexpression conferred lysozyme resistance and the *lppS* gene (encoding L,D-transpeptidase) and probably some other  $\sigma^D$ -dependent genes were found to contribute to lysozyme resistance. No clear  $\sigma^D$  regulon induction was observed in this strain under any other stress conditions tested (treatment with sodium dodecyl sulfate, cetyl trimethylammonium bromide, ethambutol or ampicillin) (Toyoda and Inui, 2018).

Phenol was previously described to cause a significant oxidative stress at subtoxic concentrations. At higher concentrations, phenol affects bacterial cells by releasing cell wall components and causing non-specific increase in cell

permeability (Denyer, 1995). At the level of the cytoplasmic membrane, phenol induces a loss of structural integrity by the leakage of potassium ions and various organic compounds and can also displace phospholipid molecules in the cell envelope (Denyer, 1995). Such damage to the cell wall might induce a similar stress response as lysozyme treatment. However, effects of toxic aromatic compounds, such as phenol, vanillin and ferulic acid, which are degraded by *C. glutamicum* (Chen et al., 2016, 2017, 2018), on  $\sigma^D$  regulon induction have not yet been tested.

Although the number of precisely localized promoters driving expression of  $\sigma^D$ -dependent genes was not high, it was possible to propose their consensus sequence. Considering the difference between  $\sigma^D$ -controlled promoters and other promoter classes, the  $\sigma^D$  regulon seemed to also be an insulated gene group.

In this study, we discovered that nearly all of the  $\sigma^D$ -dependent promoters found in *C. glutamicum* ATCC 13032 are also to a much lower extent active with RNAP+ $\sigma^H$ , although the known highly conserved consensus sequence of the  $\sigma^H$ -dependent promoters differs from that of the  $\sigma^D$ -dependent promoters. We have found that phenol in the cultivation medium acts as a stress factor, inducing the expression of some  $\sigma^D$ -dependent genes. To achieve reliable results, we combined the genome-wide technique (RNA-seq) and single-promoter analysis (*in vitro* transcription and an *in vivo* two-plasmid system for the assignment of  $\sigma$  factors to genes/promoters). *In silico* homology modeling completed the different approaches to analyzing the sigma-promoter relationships.

## MATERIALS AND METHODS

### Bacterial Strains, Plasmids, Oligonucleotides and Growth Conditions

*Escherichia coli* DH5 $\alpha$  (Hanahan, 1985) was cultivated aerobically in 500-ml flasks containing 70–100 ml of 2xYT medium (Green and Sambrook, 2012) at 37°C. Wild-type (WT) *C. glutamicum* ATCC 13032, its deletion derivative *C. glutamicum*  $\Delta sigD$  (Taniguchi et al., 2017) and *C. glutamicum* *sigD*-overexpressing strain (Taniguchi et al., 2017) were used for DNA and RNA isolations, and as hosts for testing the activities of promoters cloned in the promoter-test vector pEPR1. *C. glutamicum* was cultivated aerobically in 500-ml flasks with 70–100 ml of complete 2xYT medium or in minimal CGXII medium (Keilhauer et al., 1993) at 30°C. Media were supplemented with antibiotics, when necessary: kanamycin (Km; 30  $\mu$ g/ml), tetracycline (Tc; 10  $\mu$ g/ml) or ampicillin (Ap; 100  $\mu$ g/ml). The plasmid vectors used are listed in **Table 1**. The oligonucleotides used are listed in **Supplementary Table S1**.

### DNA Manipulations

DNA isolation, PCR, cutting by restriction enzymes, ligation and transformation of *E. coli* were done using the standard techniques (Green and Sambrook, 2012). *C. glutamicum* cells were transformed by electroporation (van der Rest et al., 1999). DNA fragments for cloning promoters in pRLG770 and pEPR1 were assembled from the synthesized complementary

oligonucleotides of around 75 nt in length, with overhangs ready for ligation. Their sequences are shown in **Supplementary Table S1**. Mutations in *sigH* were constructed using a Q5<sup>®</sup> Site-Directed Mutagenesis Kit (New England BioLabs<sup>®</sup> Inc.) according to the recommendations of the manufacturer. The specific oligonucleotide primers for mutagenesis were designed by NEBaseChanger<sup>™</sup> v1.2.8 (New England BioLabs<sup>®</sup> Inc.).

### RNA Isolation, cDNA Library Preparation and Sequencing

*Corynebacterium glutamicum* ATCC 13032 was cultivated in minimal medium CGXII with glucose (2%) or phenol (3.4 mM). The *sigD*-overexpressing *C. glutamicum* strain was used as described previously (Taniguchi et al., 2017). Total RNA was isolated from 3 biological replicates of *C. glutamicum* cells grown to the exponential phase using a Quick-RNA Miniprep Plus kit according to the manufacturer's instructions (Zymo Research). After additional DNase treatment, RNA samples were purified with an RNA Clean&Concentrator-5 kit (Zymo Research) and quantified with a DropSense 16 (Trinean). The quality of total RNA was controlled with an RNA 6000 Nano kit in an Agilent 2100 Bioanalyzer (Agilent Technologies). To construct the whole transcriptome cDNA library, 2.5  $\mu$ g total RNA (RIN > 9) was used for the depletion of rRNA with a RiboZero rRNA Removal Kit (Bacteria) according to manufacturer's instructions (Illumina). The rRNA removal was checked with an Agilent RNA Pico 6000 kit and the Agilent 2100 Bioanalyzer (Agilent Technologies). The mRNA obtained was converted to a cDNA library according to the TruSeq Stranded mRNA Sample Preparation guide (Illumina). The quality and quantity of the cDNA library was checked with an Agilent High Sensitivity DNA kit and the Agilent 2100 Bioanalyzer (Agilent Technologies). Sequencing was performed on an Illumina HiSeq 1500 using 70 bases read length (Illumina).

A primary 5'-end specific cDNA library was synthesized using 2  $\times$  2.5  $\mu$ g total RNA (RIN > 9) as described previously (Kranz et al., 2018). Briefly, after rRNA depletion and quality control, as described for the whole transcriptome cDNA libraries, a terminator 5'-phosphate-dependent exonuclease treatment (TEN, Illumina) was carried out to digest non-primary transcripts. The remaining non-digested non-primary transcripts were tagged by the ligation of the RNA 5'-index adapter 5'-CCCUACACGACGCUCUCCGAUCGAG-**UACCCUAG** (index in bold) to the 5'-monophosphorylated ends. The primary 5'-triphosphate ends were converted to 5'-mono-phosphate by RNA 5'-polyphosphatase (RPP) treatment (Epicenter) to ligate the 5'-adapter to the 5'-ends of primary transcripts. Finally, reverse transcription with a stem-loop DNA adapter and library amplification was performed. The quality and quantity of the cDNA library was checked in the same way as for the whole transcriptome cDNA library. Prior to sequencing, primary 5'-end cDNA libraries were purified and size-selected for fragments approximately 100–1000 bases long via gel electrophoresis, and quantified again. Paired end sequencing was performed in an Illumina MiSeq using MiSeq Reagent Kits v3 with a read length of 2  $\times$  75 bases (Illumina).

**TABLE 1** | Plasmid vectors used in this study.

Plasmid	Characteristics	Source
pRLG770	<i>E. coli</i> vector, <i>rnnB</i> terminator, Ap <sup>R</sup> , used for <i>in vitro</i> transcription analysis	Ross et al., 1990
pEPR1	<i>E. coli</i> - <i>C. glutamicum</i> promoter-test vector, Km <sup>R</sup> , promoter-less <i>gfpuv</i> as a reporter	Knoppová et al., 2007
pEC-XT99A	<i>E. coli</i> - <i>C. glutamicum</i> expression vector, Tc <sup>R</sup> , IPTG-inducible <i>trc</i> promoter	Kirchner and Tauch, 2003

## Read Processing, Mapping and Visualization

Paired-end reads were mapped to the *C. glutamicum* ATCC 13032 reference genome sequence accession number BX927147 (Kalinowski et al., 2003) with bowtie2 v2.2.7 (Langmead and Salzberg, 2012) using the default settings for paired-end read mapping.

False-positive primary transcriptome cDNA reads containing the barcode sequence TACCCTAG at their 5'-ends were discarded. The remaining R1 cDNA reads were mapped to the *C. glutamicum* ATCC 13032 reference genome sequence accession number BX927147 (Kalinowski et al., 2003) with bowtie2 v2.2.7 (Langmead and Salzberg, 2012) using the default settings for single-end read mapping. Read Explorer v.2.2 (Hilker et al., 2014) was used for the visualization of short read alignments, transcription start site (TSS) detection and differential gene expression analysis.

## Identification of Transcription Start Sites (TSS)

Transcription start sites (TSS) were detected essentially as described by Wittchen et al. (2018). Primary 5'-end cDNA library data were analyzed with the software ReadXplorer v2.2 (Hilker et al., 2016) using the *Transcription analysis* function. The parameters for TSS detection were chosen by ReadXplorer using its automatic parameter estimation. TSS were detected with at least 10 read starts and a minimal coverage increase of 100%. The resulting list of predicted TSS was manually checked for false-positives.

## Differential Gene Expression Analysis

Differential gene expression analysis of *C. glutamicum* grown with and without phenol (3.4 mM), including normalization, was performed using the whole transcriptome data and Bioconductor package DESeq2 (Love et al., 2014) included in the software ReadXplorer v2.2 (Hilker et al., 2016). The signal intensity value (*a*-value) was calculated by the log<sub>2</sub> mean of normalized read counts, and the signal intensity ratio (*m*-value) by log<sub>2</sub> fold-change. The evaluation of the differential RNA-seq data was performed using an adjusted *p*-value cut-off of  $P \leq 0.01$  and a signal intensity ratio (*m*-value) cut-off of  $\geq 1$  or  $\leq -1$ .

## *In vitro* Transcription

The *in vitro* transcription assay was carried out essentially as described previously (Holátko et al., 2012). The holo-RNAP was reconstituted from the RNAP core enzyme isolated from *C. glutamicum* and individual *C. glutamicum* sigma factors isolated as His-tagged recombinant proteins from *E. coli*. The RNAP core (100 nM) was mixed with the respective  $\sigma$  factor ( $\sigma^C$ ,  $\sigma^D$ ,  $\sigma^E$ ,  $\sigma^H$ , or  $\sigma^M$ ) in a molar ratio of 1:30. The holo-RNAP was assembled for 10 min at 37°C. The transcription mixture was incubated for 15 min at 37°C. The transcripts labeled with [ $\alpha$ -<sup>32</sup>P]UTP were separated in 5.5% polyacrylamide gel. *In vitro* transcription assays were done 2 or 3 times for each promoter, with essentially the same results.

## Promoter Activity Measurements Using the Two-Plasmid System

Sigma factors were assigned to individual promoters *in vivo* using the two-plasmid system for *C. glutamicum* described recently (Dostálová et al., 2017) and similar that developed for the identification of  $\sigma^E$ -dependent promoters in *E. coli* (Rezuchova and Kormanec, 2001). In principle, the sigma factors overproduced using the expression vector pEC-XT99A drove transcription from the individual promoters transcriptionally fused to the *gfpuv* reporter gene in the other plasmid present in the cell, the promoter-test vector pEPR1. Promoters were cloned to pEPR1 as approximately 75-nt fragments using the restriction sites *Pst*I and *Bam*HI. The same pEC-XT99A constructs as described previously (Dostálová et al., 2017) were used to overexpress the *sig* genes after the addition of isopropyl- $\beta$ -thiogalactopyranoside (IPTG). Two-plasmid strains were cultivated in 2xYT medium (containing Km and Tc) for 24 h and the cell samples were collected. The cells were disrupted with a FastPrep homogenizer (MP Biomedicals). The fluorescence of the cell-free extracts was determined with a Sapphire2 microplate spectrophotometer (Tecan; excitation wavelength 397 nm; emission wavelength 509 nm). Cells harboring the pEPR1 construct and the expression vector without a *sig* gene were used as a control. To determine the background fluorescence of *C. glutamicum* cells, the strain only carrying the promoter-less vector pEPR1 was used. Protein concentration in the extract was determined by Bradford assay, and promoter activity was expressed in arbitrary units/mg protein.

## Homology Modeling and Molecular Dynamics (MD) Simulations

The homology models of the  $\sigma^D$  and  $\sigma^H$  domains which recognize the -10 and -35 sequences of the respective promoters were produced by using the Swiss-Model server (Waterhouse et al., 2018). The crystal structures of *E. coli*  $\sigma^E$ , PDBid: 4LUP (for -10 element GTC) (Campagne et al., 2014) and PDBid: 2H27 (for -35 element GGAAC) (Lane and Darst, 2006) were used as templates. The nucleotides within the *E. coli*  $\sigma^E$  consensus were replaced to match the consensus for *C. glutamicum*  $\sigma^H$  or  $\sigma^D$ , where necessary. Molecular dynamics simulations were done using the software package AMBER (Salomon-Ferrer et al., 2013)

**TABLE 2** | Promoter sequences of  $\sigma^D$ -dependent genes in *C. glutamicum* ATCC 13032 determined by RNA sequencing specific for transcription start sites.

Gene	Promoter sequence <sup>A</sup>	Distance TSS-start codon
<i>cmt1</i>	AGGTAAGCGCCTG <b><u>TTAACG</u></b> TAATAG-CTTGAATATAG <b><u>GAT</u></b> GTAATTTAA	173 nt
<i>cg0420</i>	CATCTGAGCAGTTA <b><u>GAAACG</u></b> GATATGT-CGGTAGTAACCG <b><u>GAT</u></b> ACGATTTATT <b><u>G</u></b>	33 nt
<i>lpd</i>	ATTCGGCAGAGTG <b><u>CTAACG</u></b> GTTAGG-CACTATTTTC <b><u>GTT</u></b> AGTTCCTTT <b><u>G</u></b>	360 nt
<i>cg0607</i>	TTGCGCGTTAATAA <b><u>GGAA</u></b> CAATATCG-GTGTGATTCGCG <b><u>GAT</u></b> ATATTAATCA	49 nt
<i>rsdA</i>	GGTCAGCGATGGAA <b><u>GTAACA</u></b> GAGTTA-GGGAACTTCTC <b><u>GAT</u></b> CTACTGAGT <b><u>G</u></b>	158 nt
<i>cmt3</i>	TTTGGGGGAGGTT <b><u>GTTACA</u></b> AAAACATA CGTCTGTGA <b><u>GAT</u></b> ATGACG <b><u>AGT</u></b> <b><u>G</u></b>	44 nt
<i>cg1056</i>	TTTTGACGACGCA <b><u>GTAACG</u></b> CAATCGGGG ATTGTGGT <b><u>GAT</u></b> TCTTTAAGCA	84 nt
<i>cg2047</i>	CATTTTTAGTGAC <b><u>GTAACA</u></b> TCAAAG-AAGTATTCAC <b><u>TGAT</u></b> GTAAGTAGT <b><u>G</u></b>	120 nt
<i>lppS</i>	AAGGCATTCTGGAC <b><u>GTAACG</u></b> CTCCGG-CATCTACAAGG <b><u>GAT</u></b> GATCAAATA	225 nt
<i>fadD2</i>	GTTAGTTTGCAAA <b><u>GTAATA</u></b> AAATGT-TCATCTTTGT <b><u>GAT</u></b> GGTCACA <b><u>ATA</u></b>	30 nt
<i>cmt2</i>	AGTGTCAACA <b><u>CTTGTAACG</u></b> TGTGGG-CGGAAAACAA <b><u>GAT</u></b> AGGCATCGAG	59 nt

<sup>A</sup>The predicted  $-35$  motif **GTAAC<sup>A</sup>/G** and  $-10$  motif **GAT** are in bold, underlined. Elements identical to the consensus of  $\sigma^H$ -controlled promoters are highlighted in red. The gaps were introduced to align the key motifs.

and Linux computer nodes with powerful NVIDIA GPUs that enable the accumulation of 50-ns MD trajectories at 280 K.

## RESULTS

### Mapping of Transcription Start Sites of $\sigma^D$ -Dependent Genes by RNA-seq

RNA-seq with a *C. glutamicum* WT strain, a strain overexpressing *sigD* and a *sigD* deletion ( $\Delta sigD$ ) strain was performed previously (Taniguchi et al., 2017). Using this approach, the 29 genes were identified, which showed the increased transcription in *C. glutamicum* overexpressing *sigD* and decreased transcription in the  $\Delta sigD$  strain. The genes selected may be those which are directly under the control of  $\sigma^D$ , or the genes whose transcription is influenced by an indirect effect of *sigD* overexpression or deletion.

To further investigate these  $\sigma^D$ -dependent genes and to map their transcriptional start sites (TSS), RNA-seq of 5'-enriched primary transcripts of *C. glutamicum* overexpressing *sigD* was used. The sequencing of primary 5'-end specific cDNA library allows the exact mapping of TSS. To detect these TSS, the data obtained from the primary 5'-end cDNA sequencing were mapped to *C. glutamicum* ATCC 13032 reference genome sequence, accession number BX927147 (Kalinowski et al., 2003) and analyzed via ReadXplorer (Hilker et al., 2016). To distinguish between the background and actual TSS signal, we considered a position to be +1 (TSS) if the number of read starts here was 10 times higher than at position  $-1$ . In this way the 5'-ends of the transcripts of the above mentioned 29 putative  $\sigma^D$ -dependent genes were detected. Since several genes were cotranscribed in operons, the total number of TSS was 23 (Supplementary Table S2). The 50-nt sequences ( $-49$  to  $+1$ ) were aligned at the TSS, and the consensus motifs within the region of the supposed promoters were searched using the software Improbizer (Ao et al., 2004) as described previously (Albersmeier et al., 2017). The consensus sequence or closely similar motifs within the  $-35$  and  $-10$  regions were found in

11 sequences (Table 2). The products of the respective genes are involved in the synthesis of corynomycolic acid (*fadD2*, *cmt1*, *cmt2*, *cmt3*), peptidoglycan formation (*lppS*), inhibition of  $\sigma^D$  activity (*rsdA*) or have hypothetical functions (Table 3). The  $\sigma^D$  is thus involved in the control of cell wall integrity and cell envelope stress response, as described previously in the strains *C. glutamicum* ATCC 13032 (Taniguchi et al., 2017) and *C. glutamicum* R (Toyoda and Inui, 2018).

The sequences were used to generate a sequence logo with Weblogo3 (Figure 1). The consensus sequences **GTAAC<sup>A</sup>/G** as the  $-35$  motif and **GAT** as the  $-10$  motif are apparent.

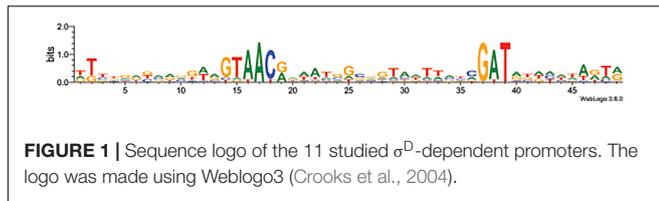
### In vitro Transcription From the $\sigma^D$ -Controlled Promoters

The activity of the 11 proposed  $\sigma^D$ -dependent promoters, which were deduced from the TSS determination by RNA-seq, was tested using the *in vitro* transcription system (Holátko et al., 2012). The promoter sequence elements of the proposed  $\sigma^D$ -dependent promoters were apparently different from the consensus sequence of  $\sigma^A/\sigma^B$ -driven promoters (Pátek and

**TABLE 3** | Products of  $\sigma^D$ -dependent genes in *C. glutamicum* ATCC 13032.

Gene ID <sup>A</sup>	Gene name	Function
<i>cg0413</i>	<i>cmt1</i>	Trehalose corynomycolyl transferase
<i>cg0420</i>		Hypothetical glycosyltransferase
<i>cg0441</i>	<i>lpd</i>	Dihydroliipoamide dehydrogenase
<i>cg0607</i>		Hypothetical secreted protein
<i>cg0697</i>	<i>rsdA</i>	Anti-sigma D
<i>cg1052</i>	<i>cmt3</i>	Trehalose corynomycolyl transferase
<i>cg1056</i>		Hypothetical membrane protein
<i>cg2047</i>		Hypothetical secreted protein
<i>cg2720</i>	<i>lppS</i>	L,D-transpeptidase
<i>cg3179</i>	<i>fadD2</i>	Long-chain fatty-acid-CoA ligase
<i>cg3186</i>	<i>cmt2</i>	Trehalose corynomycolyl transferase

<sup>A</sup>Genes are listed according to CoryneRegNet (<http://www.coryneregnet.de/>).

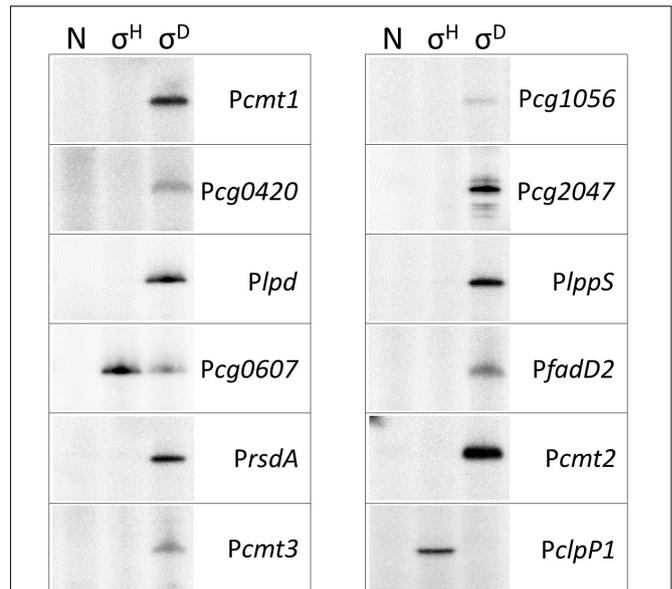


Nešvera, 2011; Albersmeier et al., 2017) and therefore ECF  $\sigma$  factors ( $\sigma^C$ ,  $\sigma^D$ ,  $\sigma^E$ ,  $\sigma^H$ , and  $\sigma^M$ ) were only used. This assay may directly prove whether a promoter is specifically recognized by RNAP associated with the chosen  $\sigma$  factor, and thus distinguish between direct and indirect control by the tested  $\sigma$ . As shown in **Figure 2**, all but one promoter were found to be exclusively  $\sigma^D$ -specific. The only exception was the promoter of *cg0607* (*Pcg0607*), which was active with both  $\sigma^D$  and  $\sigma^H$ . The *in vitro* transcript level with  $\sigma^H$  was even higher than with  $\sigma^D$  (**Figure 2**). It is worth mentioning that *Pcg0607* is the only one out of the 11 analyzed promoters that has the  $-35$  sequence element GGAAC which is a consensus for  $\sigma^H$ -dependent promoters (Busche et al., 2012). The ECF  $\sigma$  factors  $\sigma^C$ ,  $\sigma^E$ , and  $\sigma^M$  did not provide any signal in the assays (data not shown). The  $\sigma^H$ -specific *clpP1* promoter (Busche et al., 2012) was used as a control.

### Determination of Activity of $\sigma^D$ -Controlled Promoters *in vivo* Using Two-Plasmid System

To confirm the results of *in vitro* transcription, the activity of  $\sigma^D$ -dependent promoters was determined *in vivo* using the two-plasmid *C. glutamicum* cells carrying a *sig* gene in the expression vector pEC-XT99A and the analyzed promoter in the promoter-test vector pEPR1 (Dostálová et al., 2017). The promoter activity was measured as the fluorescence intensity of the Gfp reporter protein. The results of the assays with 7 most felicitous example promoters which confirmed the activity of  $\sigma^H$  in transcription are presented in **Figure 3**. As shown in the left part of **Figure 3**, all promoters exhibited high activity with the overexpressed *sigD*. Results with *Pcg1056* were included to document that not all  $\sigma^D$ -controlled promoters are also recognized by  $\sigma^H$ , although there is no apparent reason found by inspection of the promoter sequence only. Activity of *Pcg2047* with  $\sigma^H$  was approximately twice higher after induction at time point 3 (T3) and time point 6 (T6) although it was at the same level of the control. It is an unusual feature of this two-plasmid system in some cases, which we described in our previous paper (Dostálová et al., 2017). We concluded, therefore, that there was an activity of *Pcg2047* initiated by  $\sigma^H$ .

The level of Gfp reporter fluorescence was lowest 3 h (T3) after the addition of the inducer (IPTG) in all cases and highest after 24 h (T24) in 10 of the 11 tested promoters. The promoters *Pcmt2*, *PfadD2*, *PrsdA*, *Pcg2047*, *PlppS*, *Plpd*, and *Pcg0607* were also found to be active with  $\sigma^H$ , as shown in the right part of **Figure 3**. Although the maximal  $\sigma^H$ -induced activities were 1.5- to 64-fold lower than those with  $\sigma^D$  under the conditions used, the effect was apparent (**Figure 3**, right part). The promoter *Pcg1056* is an example of a promoter which did not demonstrate any

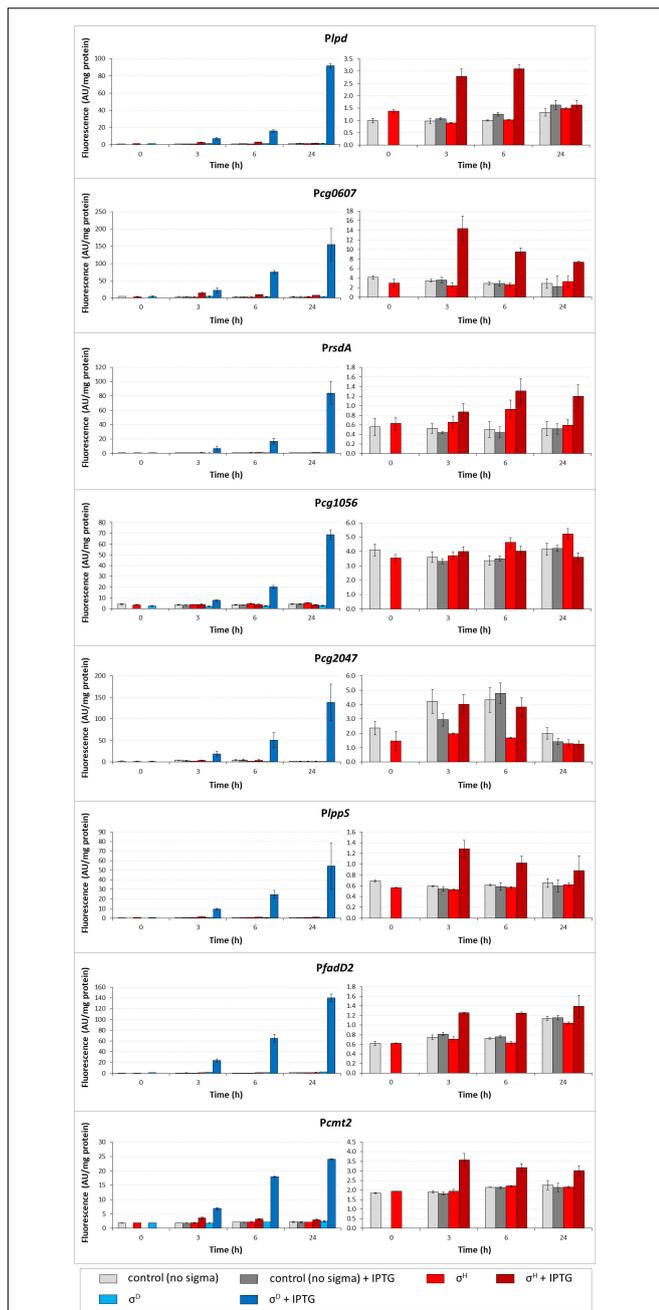


**FIGURE 2** | *In vitro* transcription driven from  $\sigma^D$ -dependent promoters with reconstituted *C. glutamicum* holo-RNAPs (autoradiogram of the SDS-PAGE gel). The lanes with no sigma (N),  $\sigma^H$  and  $\sigma^D$  are shown. *In vitro* transcription was carried out as described previously (Holátko et al., 2012). The derivatives of the plasmid vector pRLG770 with inserted promoter fragments (approximately 75 nt) were used as DNA templates, recombinant RNAP core was isolated from *C. glutamicum*, and sigma factors were isolated as His-tagged proteins from *E. coli* using the pET-22b(+) constructs carrying the respective inserted *sig* genes (Holátko et al., 2012; Šilar et al., 2016). The  $\sigma^H$ -specific promoter *PclpP1* was used as a control. Representative results of 2–3 assays are shown.

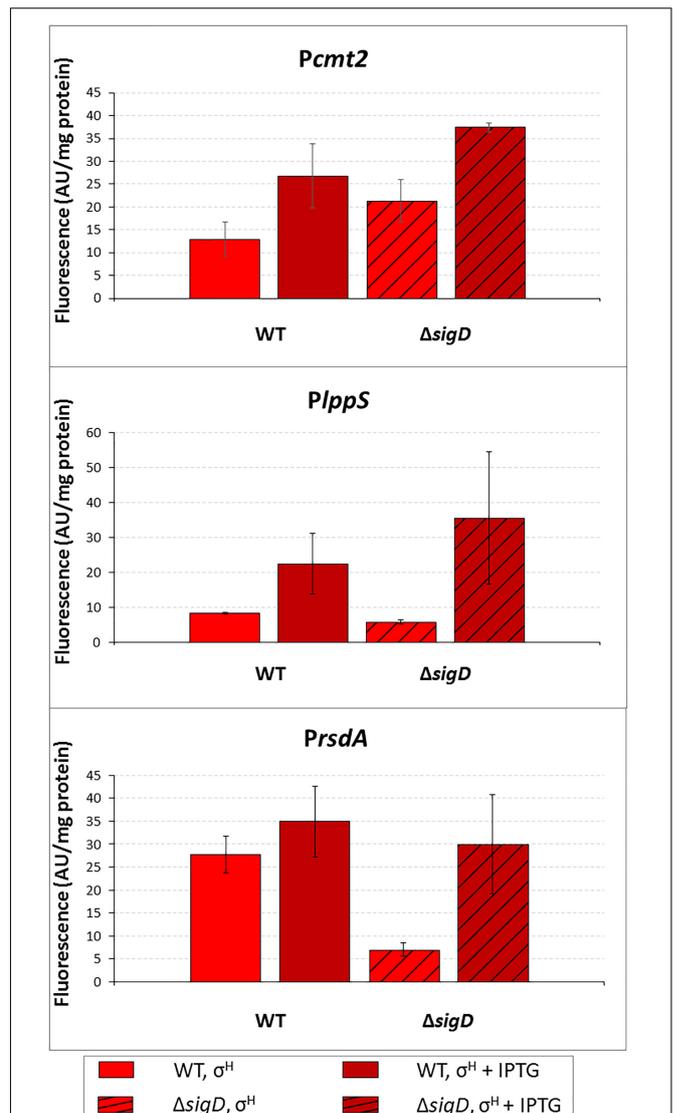
increase in activity with *sigH* overexpression. Neither of the other overproduced  $\sigma$  factors supported a promoter activity increase with any of the promoters tested (data not shown). Interestingly, the trend of changes in the  $\sigma^H$ -directed promoter activity was mostly the opposite of that with  $\sigma^D$ : the activity with  $\sigma^H$  was highest 3 h after the addition of IPTG, then mostly decreased at T6, and the difference in Gfp level between the cells with IPTG-induced and non-induced  $\sigma^H$ -overproduction nearly vanished at T24 in most cases (**Figure 3**, right part). Due to this trend,  $\sigma^D$ - and  $\sigma^H$ -induced promoter activities were nearly comparable at T3. This effect was most apparent in *Pcg0607*, which was also found to be active with  $\sigma^H$  by *in vitro* assay (**Figure 2**): this ratio between promoter activity with  $\sigma^D/\sigma^H$  was 1.5 at T3 but 20.5 at T24 (**Figure 3**). The overall conclusion is that most of the  $\sigma^D$ -controlled promoters are also to a lesser extent  $\sigma^H$ -activated.

### Promoter Activity in the *sigD* Deletion Strain ( $\Delta sigD$ )

To prove that  $\sigma^H$  is directly responsible for the minor activity of  $\sigma^D$ -controlled promoters *in vivo*, we tried to overexpress *sigH* in the  $\Delta sigD$  strain also harboring pEPR1 with the target promoters. The  $\Delta sigD$  strain grew slightly slower than the wild type strain (Taniguchi et al., 2017). However, the  $\Delta sigD$  clones carrying two vector constructs grew extremely poorly. Therefore, we only succeeded in testing the clones with *Pcmt2*, *PlppS*, and *PrsdA*



**FIGURE 3 |** Determination of promoter activity under *sigD* or *sigH* overexpression conditions using the two-plasmid assay. Promoter activity was measured as Gfpuv fluorescence intensity of cell extracts shown as bars in colors representing respective sigma factors. The *C. glutamicum* strains carried the pEC-XT99A constructs overexpressing *sigD* (blue bars) or *sigH* (red bars) after IPTG addition (added at time point 0, dark bars) and the promoter-test vector pEPR1 carrying the reporter *gfpuv* gene downstream of the target promoter (*Plpd*, *Pcg0607*, *PrsdA*, *Pcg1056*, *Pcg2047*, *Plpps*, *PfdD2*, or *Pcmt2*). The strains harboring the pEPR1 with the target promoters and empty pEC-XT99A were used as the controls (gray bars). Only the  $\sigma^H$ -induced activity and controls are shown in the right panels to emphasize the minor recognition of  $\sigma^D$ -controlled promoters by  $\sigma^H$ . The background fluorescence intensity with control cells only carrying the empty vector pEPR1 was  $0.25 \pm 0.09$  in all cases. AU, arbitrary units. The standard deviations of three biological replicates are depicted by error bars.



**FIGURE 4 |** Determination of promoter activity in *C. glutamicum* WT and  $\Delta sigD$  strains under *sigH* overexpression conditions using the two-plasmid assay. The *C. glutamicum* strains carried the vector pEC-XT99A construct overexpressing *sigH* with IPTG and the promoter-test vector pEPR1 with a target promoter *Pcmt2*, *Plpps* or *PrsdA*. The promoter activity was measured as the Gfpuv fluorescence intensity of the cell extracts after 6-h growth on complete agar plates with or without IPTG. Fluorescence of cultures without IPTG are shown as light bars; fluorescence of cultures with IPTG induction are shown as dark bars. Open bars depict fluorescence of the strains with WT background, hatched bars depict fluorescence of the strains with the  $\Delta sigD$  background. AU, arbitrary units. The standard deviations of three biological replicates are depicted by error bars.

(cloned in pEPR1), moreover, only on a solid agar medium. The increase in promoter activity, (induction ratio of WT and  $\Delta sigD$  strains with *sigH* overexpressed), was 2.1 and 1.8, respectively, for *Pcmt2*, 2.2 and 3.5, respectively, for *Plpps* and 1.1 and 6.1, respectively, for *PrsdA* (Figure 4). Thus, RNAP+  $\sigma^H$  was apparently directly active in transcription from all three tested  $\sigma^D$ -dependent promoters.

## Analysis of *C. glutamicum* $\sigma^D$ and $\sigma^H$ Protein Sequence Similarity

To explain the observed function of  $\sigma^H$  in transcription from some  $\sigma^D$ -dependent promoters, the similarity of the amino acid sequences of *C. glutamicum*  $\sigma^D$  and  $\sigma^H$  proteins was analyzed. The level of overall similarity of the  $\sigma^D$  and  $\sigma^H$  sequences is very low (<25% identical AA), in contrast to *C. glutamicum* ECF sigma factors  $\sigma^H$ ,  $\sigma^E$ , and  $\sigma^M$ , which exhibit a 25–38% mutual identity of AAs. However, when the regions which are thought to interact with the –35 and –10 promoter sequences (Lane and Darst, 2006; Campagne et al., 2014) were compared, some AAs were found to be identical or similar in  $\sigma^D$  and  $\sigma^H$  (shown in green and yellow, respectively, in **Figure 5**). As shown in **Figure 5**, majority of these AAs were also found to be conserved in the model ECF sigma factor,  $\sigma^E$  from *E. coli*, whose crystal structure was previously used for studies on sigma-promoter DNA interactions (Lane and Darst, 2006; Campagne et al., 2014).

## Computer Modeling of $\sigma^D$ and $\sigma^H$ in Complexes With –10 and –35 Elements of $\sigma^D$ - and $\sigma^H$ -Controlled Promoters

We have shown that the consensus sequences of the *C. glutamicum* –10 and –35 elements recognized by  $\sigma^D$  (this study) and  $\sigma^H$  (Busche et al., 2012) differ (i.e., GAT vs. GTT at –10 and GTAAC<sup>A</sup>/<sub>G</sub> vs. GGAA<sup>T</sup>/<sub>C</sub> at –35). To identify the key AAs responsible for recognizing different nucleotides at the second position of these consensus sequences and to understand why  $\sigma^H$  is, to a certain extent, able to initiate transcription from the  $\sigma^D$ -dependent promoters, we used computer modeling tools. First, we created homology models of those  $\sigma^D$  and  $\sigma^H$  domains, which recognize the –10 and –35 sequences of the respective promoters. The structure of *E. coli*  $\sigma^E$  based on the crystal protein in complex with promoter DNA was used as a template (Lane and Darst, 2006; Campagne et al., 2014). It should be noted that the DNA promoter sequences recognized by *E. coli*  $\sigma^E$  are closer to the *C. glutamicum* consensus of the  $\sigma^H$ -specific –10 promoter element (GTC in *E. coli*  $\sigma^E$ , i.e., T at the second position in both cases) as well as the –35 element (GGAAC in *E. coli*  $\sigma^E$ , i.e., G in the second position in both cases). This also corresponds to the similarity of the AA sequences of *E. coli*  $\sigma^E$  and *C. glutamicum*  $\sigma^H$  which are supposed to interact with these nucleotides (Lane and Darst, 2006; Campagne et al., 2014) (shown in bold in **Figure 5**).

As for the –10 element, there is no X-ray or cryo-electron microscopy structure of the RNAP complex with any of the group IV (ECF) sigma factors described, and a model of the complete DNA transcription bubble is not yet available. The key nucleotide at the second position of the *C. glutamicum* –10 elements (i.e., A in  $\sigma^D$  and T in  $\sigma^H$ ) is in close contact with the surface of the  $\sigma$  subunit in our homology models based on the *E. coli*  $\sigma^E$  crystal structure 4LUP (**Figure 6**). According to these models, the presence of Ala 60 in  $\sigma^D$  and Lys 53 in  $\sigma^H$  seems to be crucial for the recognition of the nucleotide at the second position of the –10 element. Ala 60 with the minimal side chain creates space for a larger adenine at the second position in the –10 element at  $\sigma^D$  (**Figure 6A**). In contrast, Lys 53 with a larger

side chain fills the free space made by a smaller base, thymine, at the second position of the –10 consensus element at  $\sigma^H$  (**Figure 6B**).

As for the –35 element, template and non-template DNA strands clearly form a duplex. There is no potential conformational polymorphism that could appear with the –10 element, where the transcription bubble opens. Therefore, our homology models of  $\sigma^D$  and  $\sigma^H$  in complex with the respective –35 elements can be considered highly plausible. To verify the stability and reliability of these homology models, we introduced a water envelope into the models and carried out 50-ns MD simulations. It was found, that the hydrophobic interactions of the methyl group of T at the second position of the  $\sigma^D$ -consensus of the –35 element with the side chains of Val 169 and Ala 166 in  $\sigma^D$  are crucial for its recognition (**Figure 6D**). Conversely, a hydrogen bond with the side chain of Ser 171 is most important for the recognition of G at the second position of the  $\sigma^H$ -consensus of the –35 element (**Figure 6E**). The hydrogen bond between the hydroxyl group of Thr 168 and the phosphate group of the sugar-phosphate DNA backbone provides another stabilizing interaction.

We used the established models to investigate why  $\sigma^H$  is, to some degree, capable of recognizing  $\sigma^D$ -controlled promoters.

### –10 Element

The larger side chain of Lys 53 in  $\sigma^H$  must interact with the larger adenine at the second position of the GAT consensus instead of the small side chain of Ala 60 in  $\sigma^D$ , if we consider the recognition of the –10 GAT motif of a  $\sigma^D$ -controlled promoter by  $\sigma^H$ . This is possible since the Lys 53 side chain is highly flexible and therefore probably does not represent a major barrier to the larger adenine (**Figure 6C**). To test whether Ala 60 in  $\sigma^D$  and Lys 53 in  $\sigma^H$  play key roles in recognizing the –10 element in  $\sigma^D$ -controlled promoters, we designed the mutant  $\sigma^H$ Lys-53-Ala for *in vivo* analysis.

### –35 Element

With the –35 element within the  $\sigma^D$ -dependent promoter, thymine at the second position of the  $\sigma^D$ -promoter consensus (GTAAC), whose methyl group interacts with the side chains of the hydrophobic AAs Val 169 and Ala 166 in  $\sigma^D$  (**Figure 6D**), interacts with Thr 168 and Ser 171 in  $\sigma^H$  (**Figure 6F**). To assess these interactions in complex ( $\sigma^H$ /–35 element of  $\sigma^D$ -dependent promoter) in detail, we again performed an extensive 50-ns MD simulation. It was found that the side chain of Thr 168 in  $\sigma^H$  can rotate, and its methyl group then interacts with the methyl group of the thymine within GTAAC, which undoubtedly leads to a hydrophobic stabilization. Moreover, the hydroxyl group from the Ser 171 side chain forms a hydrogen bond with the oxygen atom of the thymine. To test whether the replacement of the AA sequence 168-ThrValMetSerArg-172 with AlaValArgValAla from  $\sigma^D$  can improve the recognition of  $\sigma^D$ -controlled promoters by  $\sigma^H$ , we designed the mutant  $\sigma^H$ 168-AlaValArgValAla-172 for the *in vivo* analysis.

**A**

**-10 region**

SigD *Cg* 48 **TAEDVAQEICLAVAT**SIRNFV DQGRPF**MAFVYGLAS**NKVADAH**RAMS**RD

SigH *Cg* 41 **DAEDLVQDTYIKAYQAFAS**F KPGT-NLKAWLYRIMTNTYINMYR**KKQ**RQ

SigE *Ec* 44 **DVPDVVQEAFTKAYRALDS**F RGDS-AFYTWLYRLAVNTAKNYLVA**QGR**R

**B**

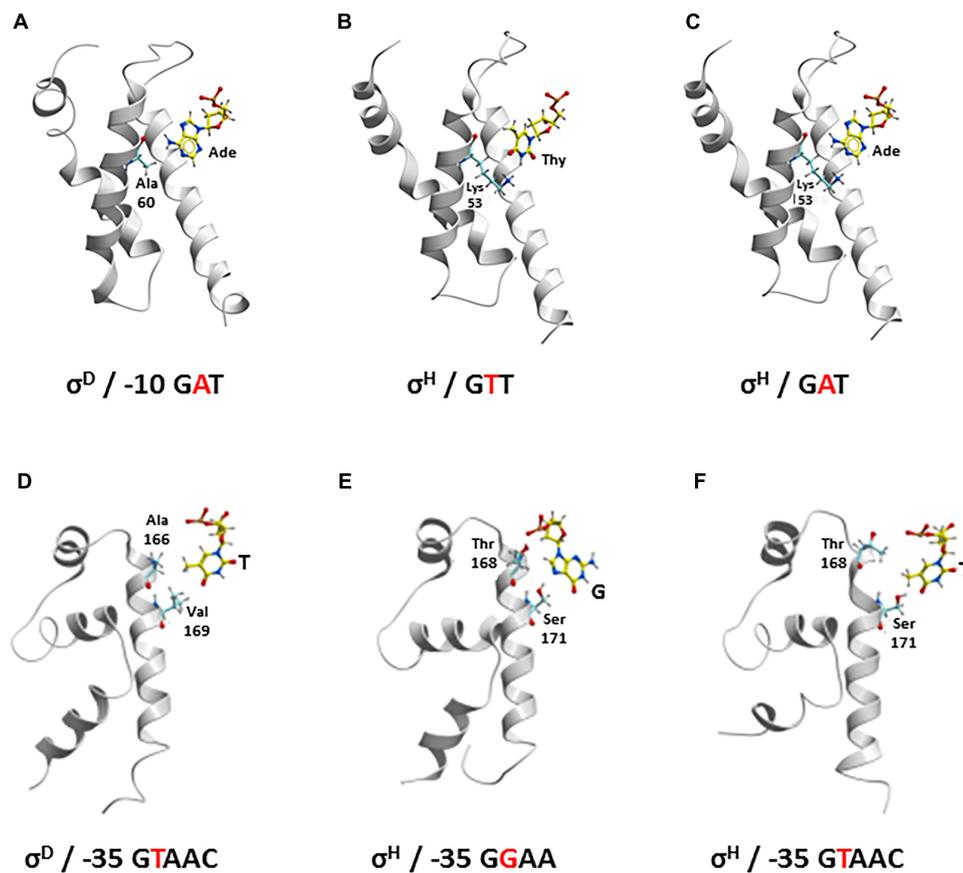
**-35 region**

SigD *Cg* 159 **VGSTPGAVRVAQHR**RALTTLRSTL

SigH *Cg* 161 **MDVPLGTVM**SRLLHRGRKQL**RGML**

SigE *Ec* 162 **MDCPVGT**VRSHIFAREAI**DNKV**

**FIGURE 5** | Alignment of the  $\sigma$  AA sequences binding promoter  $-35$  and  $-10$  elements in *C. glutamicum*. *E. coli*  $\sigma^E$ , which was analyzed using its crystal structure (Lane and Darst, 2006; Campagne et al., 2014), is aligned for comparison. The AAs which are supposed to interact with the  $-10$  (**A**) and  $-35$  (**B**) elements (Lane and Darst, 2006; Campagne et al., 2014) are in bold. The  $\sigma^H$  AAs which we replaced with the corresponding AAs from  $\sigma^D$  by mutagenesis are underlined. Identical AAs are shown in green, similar in yellow; *Cg*, *C. glutamicum*; *Ec*, *E. coli*.

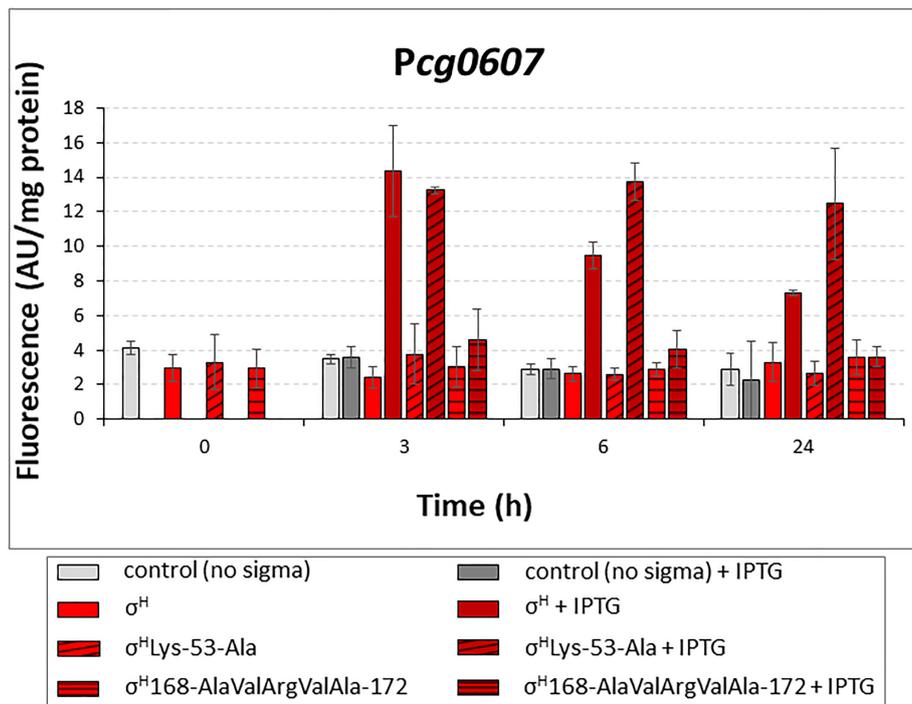


**FIGURE 6** | Recognition of the nucleotide at second position of  $-10$  element (yellow represents carbon; red, oxygen; blue, nitrogen; white, hydrogen, Ade, adenine; Thy, thymine) (**A–C**) by Ala 60 in  $\sigma^D$  and Lys 53 in  $\sigma^H$  (cyan, carbon) and at second position of  $-35$  element (yellow, carbon) (**D–F**) by Ala 166 and Val 169 in  $\sigma^D$  and Thr 168 and Ser 171 in  $\sigma^H$  (cyan, carbon).  $-10$ : (**A**)  $\sigma^D$  and  $\sigma^D$ -dependent promoter; (**B**)  $\sigma^H$  and  $\sigma^H$ -dependent promoter; (**C**)  $\sigma^H$  and  $\sigma^D$ -dependent promoter  $-35$ : (**D**)  $\sigma^D$  and  $\sigma^D$ -dependent promoter; (**E**)  $\sigma^H$  and  $\sigma^H$ -dependent promoter; (**F**)  $\sigma^H$  and  $\sigma^D$ -dependent promoter.

## Activity of $\sigma^H$ - $\sigma^D$ Mutants With the $\sigma^D$ -Controlled Promoters

Most of the  $\sigma^D$ -controlled promoters were to a lesser extent also activated by  $\sigma^H$ , although the highly conserved sequences of the  $\sigma^H$ -specific promoters ( $-35$  GGAA<sup>T/C</sup>,  $-10$  GTT) differ

from those of the identified  $\sigma^D$ -specific promoters ( $-35$  GTAAC,  $-10$  GAT). Computer modeling showed that  $\sigma^H$  can bind to the key regions of promoter DNA in a similar manner as  $\sigma^D$  to the analogous elements. Based on the results of the sequence similarity analysis and computer modeling, we designed



**FIGURE 7 |** Determination of promoter activity in *C. glutamicum* WT with overexpressed *sigH* or its mutants  $\sigma^H$ Lys-53-Ala and  $\sigma^H$ 168-AlaValArgValAla-172 by using two-plasmid assay. The *C. glutamicum* strains carried the vector pEC-XT99A overexpressing *sigH* (open bars),  $\sigma^H$ Lys-53-Ala (cross-hatched bars) or  $\sigma^H$ 168-AlaValArgValAla-172 (horizontally hatched bars) after IPTG addition and the promoter-test vector pEPR1 with the target promoter *Pcg0607*. The strain with the empty expression vector pEC-XT99A was used as a control (gray bars). The promoter activity was measured as the Gfpuv fluorescence intensity of the cell extracts. Fluorescence of cultures without IPTG is shown as light bars; fluorescence of cultures with IPTG induction is shown as dark bars. AU, arbitrary units. The standard deviations of three biological replicates are depicted by error bars.

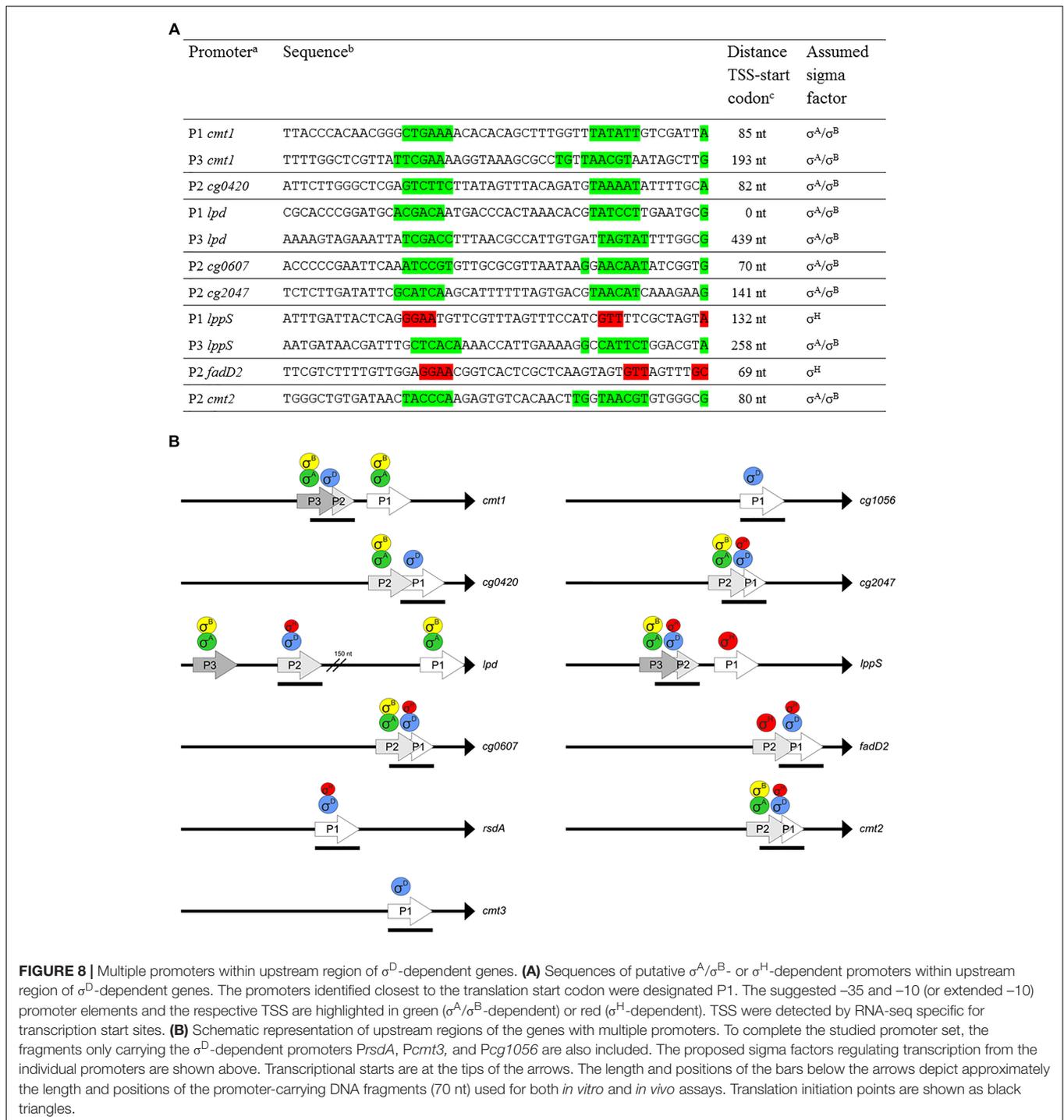
mutations to make  $\sigma^H$  more  $\sigma^D$ -like and possibly more efficient at transcription from  $\sigma^D$ -dependent promoters. We introduced the mutations into the *sigH* gene cloned in the expression vector pEC-XT99A, which resulted in the production of the mutant proteins  $\sigma^H$ Lys-53-Ala or  $\sigma^H$ 168-AlaValArgValAla-172. These mutant genes were used in the two-plasmid assays with the promoters which exhibited the highest activity with  $\sigma^H$ : *Pcg0607*, *Plpps*, *Pcmt2*, and *PrsdA*. We assumed that the overproduced  $\sigma^H$ Lys-53-Ala and  $\sigma^H$ 168-AlaValArgValAla-172 would trigger higher promoter activity in two-plasmid assays with the chosen  $\sigma^D$ -dependent promoters than  $\sigma^H$ . This assumption was only confirmed with  $\sigma^H$ Lys-53-Ala and *Pcg0607* (Figure 7). This effect was particularly apparent at T24, when promoter activity with  $\sigma^H$ Lys-53-Ala was 1.7-fold higher than that with  $\sigma^H$ . This higher activity compared to other promoters could be explained by the unparalleled presence of the GGAA sequence in the  $-35$  region of *Pcg0607*, which is a consensus sequence of  $-35$  elements in  $\sigma^H$ -dependent promoters. This result indicated that Lys 53 in  $\sigma^H$  and Ala 53 in  $\sigma^H$ Lys-53-Ala (and Ala 60 in  $\sigma^D$ ) most likely interact with the  $-10$  promoter elements.

Contrary to the expectation,  $\sigma^H$ 168-AlaValArgValAla-172 recognized *Pcg0607* less efficiently. To assess how the mutant  $\sigma^H$ 168-AlaValArgValAla-172 could interact with the  $\sigma^D$ -controlled promoters, we analyzed the homology model with this mutant  $\sigma$  factor in more detail. Using a 50-ns MD simulation,

we found that the complex of  $\sigma^H$ 168-AlaValArgValAla-172 with the  $-35$  element of the  $\sigma^D$ -controlled promoter is probably not stable. This was something of a surprise given that the native  $\sigma^H$  was able to efficiently bind  $\sigma^D$ -controlled promoter DNA in previous MD simulations. A detailed analysis of the MD trajectories showed that this instability appeared to be due to the absence of Arg at the final position of the AA sequence AlaValArgValAla (ThrValMetSerArg in  $\sigma^H$ ). The positively charged Arg 172 side chain of  $\sigma^H$  was found to create a significant stabilizing interaction (a salt bridge) with the negatively charged backbone of promoter DNA. This hypothesis, which might explain the low efficiency of the mutant  $\sigma^H$ 168-AlaValArgValAla-172, may be tested in the future by creating additional mutations suggested by the *in silico* analysis.

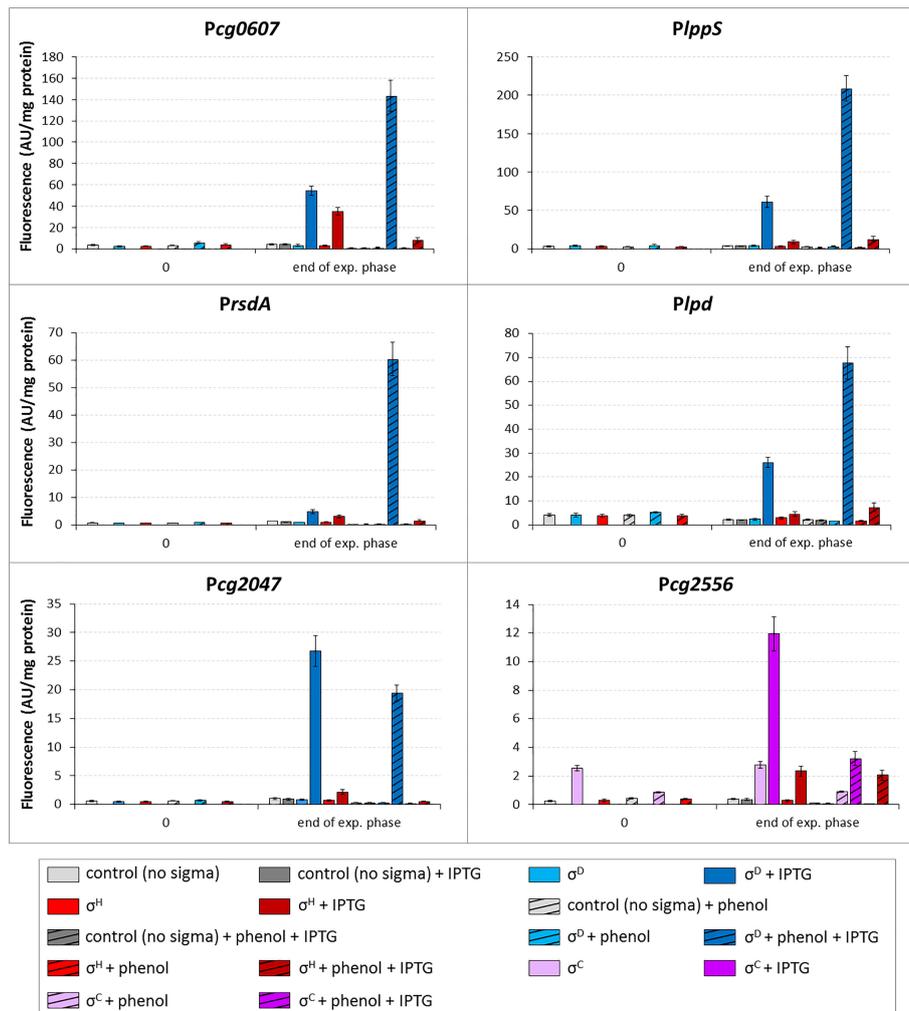
## Multiple Overlapping Promoters Upstream of the $\sigma^D$ -Controlled Genes

In addition to the TSS ascribed to the  $\sigma^D$ -controlled promoters (Table 2), start sites of transcription driven from different overlapping or nearby promoters were detected by RNA-seq for 8 of the 11 analyzed genes. We suggest that most of these promoters are activated by  $\sigma^A$  and/or  $\sigma^B$ , and two are activated by  $\sigma^H$ , according to the upstream sequences (Figure 8A, in green and red, respectively). The putative housekeeping promoters



are always located upstream of the  $\sigma^D$ -dependent promoters if there is a single additional  $\sigma^A$ - and/or  $\sigma^B$ -dependent promoter (upstream of the *cg0420*, *cg0607*, *cg2047*, and *cmt2* genes). Promoters which we consider to be  $\sigma^H$ -dependent according to the sequences of the putative  $-35$  (GGAA) and  $-10$  (GTT) elements were found upstream of *lppS* and *fadD2* genes. The detection of all TSS in the upstream regions of the  $\sigma^D$ -dependent genes was important to avoid effects of transcription

started from the additional promoters. The position of the cloned promoter fragments for *in vivo* and *in vitro* assays could thus be chosen in such a way, that only the activity of the supposed  $\sigma^D$ -controlled promoters was determined. The organization of the multiple promoters within the upstream regions of the 8 genes is shown in **Figure 8B**. In the case of *fadD2*, the studied  $\sigma^D$ -dependent P1<sub>*fadD2*</sub> and the additional  $\sigma^H$ -dependent P2<sub>*fadD2*</sub> partially overlap and the activity of the



**FIGURE 9 |** Effect of phenol on activity of  $\sigma^D$ -dependent promoters. The *C. glutamicum* strains carried the vector pEC-XT99A overexpressing *sigH* (red bars), *sigD* (blue bars) or *sigC* (purple bars) after the addition of IPTG (added at  $OD_{600} = 1$ ; dark bars) and the promoter-test vector pEPR1 constructs carrying the tested promoters (above the graphs). The strain overexpressing *sigC* and harboring pEPR1 with the  $\sigma^C$ -dependent promoter *Pcg2556* (Dostálová et al., 2017), which is not activated by phenol, was used as a control (purple bars). Phenol (3.4 mM) was present in media from the beginning of cultivation (hatched bars). Plasmids carrying the pEPR1 construct and empty pEC-XT99A were used as controls (gray bars). The samples were taken at  $OD_{600} = 1$  (0) and at the end of exponential growth phase. AU, arbitrary units. The standard deviations of three biological replicates are depicted by error bars.

additional  $\sigma^H$ -specific promoter might thus contribute to the overall  $\sigma^H$ -dependent activity of the fragment. However, such activity was not detected by *in vitro* transcription. *In vivo*, the  $\sigma^H$ -dependent activity was similarly low as in other cases. The  $\sigma^A$ - and/or  $\sigma^B$ -dependent promoters *P3cmt1*, *P2cg0420*, *P2cg0607*, *P2cg2047*, *P3lppS*, and *P2cmt2* (Figures 8A,B) were found to partially (mostly by the  $-10$  region) reside on the 70-nt fragments with the analyzed  $\sigma^D$ -dependent promoters. However, their possible activity most probably did not interfere with the *in vivo* measurements of the effects of  $\sigma^D$  and  $\sigma^H$  overexpression.

## Phenol as a Stressor

We have tested a number of stress conditions (heat shock, SDS, penicillin G, glycine, mitomycin C, phenol and limitation

by glucose) to detect a slower growth of the *C. glutamicum*  $\Delta sigD$  strain or to induce transcription from *PsigD* or from the promoters found to be  $\sigma^D$ -dependent (data not shown). Out of the conditions tested, only phenol exhibited a stronger inhibitory effect on the growth of the  $\Delta sigD$  strain carrying the vector pEC-XT99A than on the growth of the analogous WT strain. We found that growth rate of *C. glutamicum* WT culture in 2xYT medium after the addition of 7.5 mM phenol decreased from 0.88 to 0.46  $h^{-1}$  (1.9-fold), whereas the growth of the *C. glutamicum*  $\Delta sigD$  decreased from 0.68  $h^{-1}$  to 0.22 (3.1-fold). A difference in the effect of phenol was not significant in plasmidless strains.

To analyze the transcriptional response of *C. glutamicum* to phenol stress, we performed RNA-seq using the RNA isolated from the *C. glutamicum* WT cells grown in minimal medium with glucose (2%) or phenol (3.4 mM). Differential gene expression

analysis including normalization, was performed using the whole transcriptome data and Bioconductor package DESeq2 (Love et al., 2014) included in the software ReadXplorer v2.2 (Hilker et al., 2014). The evaluation of the differential RNA-seq data was performed using an adjusted  $p$ -value cut-off of  $P \leq 0.01$  and a signal intensity ratio ( $m$ -value) cut-off of  $\geq 1$  or  $\leq -1$ . According to the results (Supplementary Tables S3, S4), five genes previously found to be  $\sigma^D$ -dependent were more highly expressed on phenol than on glucose: *cg0607*, *lppS*, *rsdA*, *lpd*, and *cg2047* (Supplementary Table S4). The respective TSS identified by RNA-seq of 5'-enriched primary transcripts using the cultures grown on phenol were identical with those found using the *sigD* overexpressing strain.

The effect of phenol on the activity of the promoters *Pcg0607*, *PlppS*, *PrsdA*, *Plpd*, and *Pcg2047* was tested *in vivo* with the two-plasmid system. The strains carrying *sigD* or *sigH* inserted in pEC-XT99A and the promoters in pEPRI were cultivated in minimal medium with glucose (2%) or phenol (3.4 mM). The promoter activity was measured in the extract of the cells before the addition of IPTG and at the end of exponential growth phase.

The activity of four promoters (*Pcg0607*, *PlppS*, *PrsdA*, and *Plpd*) in the strains with overexpressed *sigD* was clearly higher after cultivation with phenol (Figure 9). The activity of the  $\sigma^C$ -dependent *Pcg2556* used as a control was much lower in the medium with phenol under *sigC* overexpression conditions. The overexpression of *sigH* in the presence of phenol only resulted in a small increase in the *Plpd* activity. These results confirmed the data from RNA-seq (with the exception of *Pcg2047* which was induced on phenol according to RNA-seq, but not in the two-plasmid assay). This is one of a few examples of the stress response to the action of phenolic compounds regulated by the ECF sigma factors in *C. glutamicum* (Chen et al., 2016, 2017). The phenol stress response mediated by ECF sigma factors will be further studied in *C. glutamicum* on the basis of the large amount of data produced by RNA-seq.

## DISCUSSION

We recently began to study the function of *C. glutamicum*  $\sigma^D$  and its regulon for the first time (Taniguchi et al., 2017). The ECF sigma factor  $\sigma^D$  was found to be crucial for the envelope stress response, the synthesis of the mycomembrane and formation of the cell wall (Taniguchi et al., 2017; Toyoda and Inui, 2018). The genes of the  $\sigma^D$  regulon are involved in the synthesis of mycolic acids, the modification of peptidoglycan and other cell envelope-related functions. A subgroup of these genes which control mycolate metabolism are specific to the Mycolata group of bacteria, which include *Corynebacterium*, *Mycobacterium*, *Rhodococcus*, and *Nocardia* as the most important genera.

Sigma factors exhibiting a high level of similarity to *C. glutamicum*  $\sigma^D$  have been found in many bacteria. The highest level of AA identity with *C. glutamicum*  $\sigma^D$  was detected in  $\sigma^D$  proteins deduced from the respective genes in species closely related to *C. glutamicum* [*C. deserti* (94%), *C. callunae* (93%) and *C. efficiens* (90%)]. A lower level of identity (70–79%) was found

with  $\sigma^D$  proteins from most of the phylogenetically more distant *Corynebacterium* species. The identity of *C. glutamicum*  $\sigma^D$  with  $\sigma^D$  proteins from other bacteria of the Mycolata group (e.g., genera *Mycobacterium*, *Rhodococcus*, and *Nocardia*) exhibited a range of 54–57%. All bacteria with a  $\sigma^D$  identity level higher than 45% (e.g., streptomycetes) belong to *Actinobacteria*. However, the functions of the genes of the  $\sigma^D$  regulons in these related bacteria are partially different. In addition to mycolate synthesis, the *M. tuberculosis*  $\sigma^D$  regulon is involved in virulence, lipid metabolism and gene expression under starvation conditions (Raman et al., 2004; Calamita et al., 2005).

A comparison of the promoter sequences of the putative  $\sigma^D$ -dependent genes among various bacteria was limited by the fact that either homologous genes were only found in a small number of *Corynebacterium* species (e.g., the genes *cg0420* and *cg2047* only occur in the genomes of three *C. glutamicum* strains) or no consensus sequences suggested for  $\sigma^D$ -dependent promoters were identified upstream of the homologous genes in the vast majority of other *Corynebacterium* species. Conservation of the putative  $\sigma^D$ -dependent promoter sequences among various bacteria was only found in the genes *cmt1* and *cmt2*, encoding trehalose corynomycolyl transferases.

Almost total conservation (37/38) of the  $-10$  and  $-35$  sequences (GTAAC – 15 nt – CTCGAT) in the *PrsdA* promoter was found in 38 *Corynebacterium* species encoding the anti-sigma factor *RsdA*. In other *Actinobacteria* (genera *Rhodococcus*, *Nocardia*, and *Mycobacterium*), a putative  $\sigma^D$ -dependent promoter is not present upstream of the *rsdA* gene but upstream of the *sigD* gene. Such an arrangement indicates that the *sigD-rsdA* operon is autoregulated by  $\sigma^D$  in these bacteria, in contrast to *C. glutamicum*, where  $\sigma^D$  seems to only control *rsdA* gene expression but not its own synthesis. Only two vegetative promoters were localized upstream of the *C. glutamicum* ATCC 13032 *sigD* gene (Pfeifer-Sancar et al., 2013).

We showed that  $\sigma^D$  played a dominant role in the transcription from all promoters tested in this study, however, low  $\sigma^H$  activity with most of these promoters was detected by the *in vivo* two-plasmid assay. This activity of the promoters was not observed when using *in vitro* transcription (with one exception). The promoter *Pcg0607* was an exception and its efficiency documented by *in vitro* transcription was even higher with  $\sigma^H$  than with  $\sigma^D$  (Figure 2). The reason for this is probably that the  $-35$  sequence GGAA within *Pcg0607* is identical to the consensus of the  $-35$  element of  $\sigma^H$ -controlled promoters. Similarly, the  $-10$  element GTT, typical for  $\sigma^H$ -controlled promoters, was found in *Plpd*. Its activity with  $\sigma^H$  *in vivo* was also higher than that of other promoters (Figure 3). We can consider *Pcg0607* and *Plpd* to be natural  $\sigma^D/\sigma^H$ -dependent hybrid promoters. However, it is apparent that the promoter efficiencies determined *in vivo* and *in vitro* do not correlate perfectly.

*SigH* plays a prominent role in the sigma regulatory network. Promoters controlled by  $\sigma^H$  were found to drive the transcription of *sigA* (Toyoda and Inui, 2015), *sigB* (Dostálová et al., 2017) and *sigM* (Nakunst et al., 2007). *SigH*, which also controls the expression of several genes encoding transcriptional regulators

(*clgR*, *sufR*, and *hspR*), is therefore considered to be a candidate for a global regulatory molecule (Schröder and Tauch, 2010). Overlap of the recognition specificities of  $\sigma^H$  and  $\sigma^E$  was described for some promoters (Šilar et al., 2016). In this study, an unexpected overlap of  $\sigma^H$  and  $\sigma^D$  in recognizing  $\sigma^D$ -dependent promoters has been proven.

To prove that the promoters are recognized directly with RNAP+ $\sigma^H$ , we measured the activities of *Pcmt2*, *P1ppS*, and *PrsdA* with  $\sigma^H$  in the  $\Delta sigD$  strain. All three promoters were found to be active (Figure 4). The transcription driven by  $\sigma^H$  was apparently stronger in the absence of  $\sigma^D$  in the cell, probably because there was no competition with  $\sigma^D$ . Upstream regions of *cmt2* and *rsdA* carried no potential  $\sigma^H$ -specific promoter and the possible  $\sigma^H$ -specific promoter *P1ppS* was not included within the tested 70-nt *P2lppS* fragment (distance between TSS1 and TSS2 was 93 nt). Thus, it was apparently  $\sigma^H$ , which drove the detected activity of the  $\sigma^D$ -dependent promoter *P2lppS* in the  $\Delta sigD$  mutant. Thus,  $\sigma^H$  was able to partially substitute for missing  $\sigma^D$ .

We have investigated the  $\sigma^D$ -dependent genes by different methods in *C. glutamicum* ATCC 13032 than those used in *C. glutamicum* R (Toyoda and Inui, 2018). In our study, differential gene expression analysis based on the obtained RNA-seq data of the whole transcriptome libraries (Illumina TruSeq stranded mRNA libraries) and RNA-seq of the 5'-enriched primary transcripts of *C. glutamicum* ATCC 13032 overexpressing *sigD* enabled us to not only discover the  $\sigma^D$ -regulated genes, but also precisely identify the TSS of the genes studied. Moreover, individual  $\sigma^D$ -dependent promoters were analyzed using *in vitro* transcription and *in vivo* two-plasmid assay. The *in vitro* assay revealed that *Pcg0607* can directly drive transcription with RNAP+ $\sigma^H$  (Figure 2). Two  $\sigma^D$ -dependent genes of *C. glutamicum* ATCC 13032, *cg0420* and *cg2047* (encoding glycosyltransferase and a secreted protein, respectively), are not present on the *C. glutamicum* R chromosome. Another gene, *lpd* was not found to be  $\sigma^D$ -dependent in *C. glutamicum* R. The promoter *P2lpd* was localized 360 nt upstream of the translation initiation codon. There is a possibility that a potential small RNA, which was detected by RNA-seq, is transcribed from this promoter. *P2lpd* has a GTT sequence in the -10 region which is the consensus of  $\sigma^H$ -dependent promoters. The *lpd* gene (encoding dihydrolipoamide dehydrogenase) is known to play a role in the oxidative stress response in bacteria (Krisko et al., 2014). In *M. tuberculosis* *lpd* contributes to survival against host-generated reactive oxygen species (Daly et al., 2007). The activity of *P2lpd* was found to increase with *sigD* and, to a lesser extent, with *sigH* overexpression in the presence of phenol in *C. glutamicum* (Figure 9).

Based on the *in silico* analysis, the *sigH* mutants resulting in the production of the mutant proteins  $\sigma^H$ Lys-53-Ala or  $\sigma^H$ 168-AlaValArgValAla-172 were constructed. The mutation Lys-53-Ala within the -10 sequence-binding region of  $\sigma^H$  improved the  $\sigma^H$ -controlled activity of *Pcg0607*, which is in agreement with the suggested crucial role of Ala in this position in recognizing the -10 element of  $\sigma^D$ -controlled promoters. The low promoter

activity with  $\sigma^H$  mutant carrying 168-AlaValArgValAla-172 was probably due to the missing Arg 172, which may establish important stabilizing salt bridges with the sugar-phosphate backbone of DNA according to the *in silico* homology modeling.

Differential gene expression analysis of the transcriptome of *C. glutamicum* ATCC 13032 grown with or without phenol revealed that some  $\sigma^D$ -dependent genes (*cg0607*, *lppS*, *rsdA*, and *cg2047*) were higher transcribed under the phenol treatment. The promoters *Pcg0607*, *P1ppS*, *Prsd* and in addition *P1pd*, were also found to be induced by phenol in minimal medium by two-plasmid analysis (Figure 9). In *C. glutamicum* R, differential gene expression analyzed by microarrays showed that the  $\sigma^D$ -dependent genes play a role in the response to the lysozyme treatment (Toyoda and Inui, 2018). The changes in gene expression seem to be induced by the cell envelope stress exerted by the action of phenol or lysozyme.

Phenol in complete medium did not significantly slow the growth of the  $\Delta sigD$  strain compared to the WT strain. However, when the replication of a single or even two plasmids placed a burden on the cell growth, these plasmid-harboring  $\Delta sigD$  strains grew slower or did not grow at all. This extremely poor growth of  $\Delta sigD$  strain with two plasmids was probably due to deficiencies in the cell envelope synthesis in combination with detrimental effects of supplemented antibiotics (Km and Tc) necessary for plasmid maintenance.

The function of  $\sigma^D$ -dependent genes and the organization of their transcription suggest that these genes play an important role not only in stress response, but also in cell homeostasis and in building a cell envelope during rapid growth under optimum conditions. Under optimal growth conditions, the transcription of the genes involved in corynomycolate synthesis is probably driven from additional  $\sigma^A$ - and/or  $\sigma^B$ -dependent promoters or  $\sigma^H$ -dependent promoters, which we detected upstream of many  $\sigma^D$ -dependent genes (Figure 8A). The transcription of  $\sigma^D$ -dependent genes from multiple promoters controlled by other sigma factors and the transcription from the  $\sigma^D$ -dependent promoters by  $\sigma^H$  may partially ensure the expression of these genes if the  $\sigma^D$  function is damaged or eliminated. However, these promoters are probably not strong enough to enable the cell to cope with the combined stress conditions if  $\sigma^D$  is missing. The low  $\sigma^H$ -triggered activity of the largely  $\sigma^D$ -controlled promoters, which is probably induced by modulation of the  $\sigma^H$  recognition specificity under particular physiological conditions, may contribute to fine-tuning transcription under various stresses, especially under the cell envelope stress.

Several genes of the *M. tuberculosis*  $\sigma^D$  regulon were found to be active in the stationary growth phase (Raman et al., 2004). Gfpuv fluorescence, which should reflect the activity of *C. glutamicum*  $\sigma^D$ -dependent promoters in our study, increased during cultivation and was found to be maximal at T24 (Figure 3). Interestingly, activity of the same promoters driven by  $\sigma^H$  exhibited the opposite trend. It therefore seems that  $\sigma^H$  may partially substitute for  $\sigma^D$  activity during the exponential phase.

The possible global regulator  $\sigma^H$  has been proven to play a crucial role in the hierarchy of  $\sigma$  factors, since it drives the transcription of at least three of them ( $\sigma^A$ ,  $\sigma^B$ , and  $\sigma^M$ ). Moreover, the regulatory overlaps of  $\sigma^H$  with transcriptional regulators

(Toyoda and Inui, 2015) and with several sigma factors including  $\sigma^D$  seem to control a number of physiological functions in *C. glutamicum* including heat, SOS, oxidative, chemical and cell surface stress responses. The accumulated evidence indicates that overlaps between the regulons of *C. glutamicum* ECF sigma factors are rather a common regulatory strategy how to cope with complex environmental stresses than an unusual exception.

## AUTHOR CONTRIBUTIONS

MP and JK conceived the project and led the studies performed in Prague and Bielefeld, respectively. HD, TB, LR, and VŠ carried out the most experiments. TB and HD carried out RNA-seq and processed the data. JH did the *in vitro* transcription assays. JN performed the sequence analyses. IB carried out *in silico* analyses. All authors analyzed the results. MP and JN drafted the initial manuscript. All authors contributed to writing the final manuscript, reading, and approving the submitted version.

## FUNDING

This work was supported by Grant 17-06991S from the Czech Science Foundation, Mobility Grant DAAD-18-11 from Czech

Academy of Sciences (CAS) and Deutscher Akademischer Austauschdienst.

## ACKNOWLEDGMENTS

Hironori Taniguchi (Wendisch Lab, Bielefeld University) is acknowledged for providing *sigD*-overexpressing *C. glutamicum* cultures for RNA isolation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.03287/full#supplementary-material>

**TABLE S1** | Oligonucleotides used.

**TABLE S2** | List of identified TSSs of SigD-upregulated genes and the respective sequences as detected by RNA-seq of transcriptome of *C. glutamicum* with overexpressed *sigD* cultivated on glucose.

**TABLE S3** | List of identified TSSs of and the respective sequences as detected by RNA-seq of transcriptome of *C. glutamicum* WT cultivated on glucose.

**TABLE S4** | List of identified SigD-dependent TSSs of SigD-upregulated genes and the respective sequences as detected by RNA-seq of transcriptome of *C. glutamicum* WT cultivated on phenol.

## REFERENCES

- Albersmeier, A., Pfeifer-Sancar, K., Rückert, C., and Kalinowski, J. (2017). Genome-wide determination of transcription start sites reveals new insights into promoter structures in the actinomycete *Corynebacterium glutamicum*. *J. Biotechnol.* 257, 99–109. doi: 10.1016/j.jbiotec.2017.04.008
- Ao, W., Gaudet, J., Kent, W. J., Muttumu, S., and Mango, S. E. (2004). Environmentally induced foregut remodeling by PHA-4/FoxA and DAF-12/NHR. *Science* 305, 1743–1746. doi: 10.1126/science.1102216
- Binder, S. C., Eckweiler, D., Schulz, S., Bielecka, A., Nicolai, T., Franke, R., et al. (2016). Functional modules of sigma factor regulons guarantee adaptability and evolvability. *Sci. Rep.* 6:22212. doi: 10.1038/srep22212
- Busche, T., Šilar, R., Pičmanová, M., Pátek, M., and Kalinowski, J. (2012). Transcriptional regulation of the operon encoding stress-responsive ECF sigma factor SigH and its anti-sigma factor RshA, and control of its regulatory network in *Corynebacterium glutamicum*. *BMC Genomics* 13:445. doi: 10.1186/1471-2164-13-445
- Calamita, H., Ko, C., Tyagi, S., Yoshimatsu, T., Morrison, N. E., and Bishai, W. R. (2005). The *Mycobacterium tuberculosis* SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence. *Cell. Microbiol.* 7, 233–244. doi: 10.1111/j.1462-5822.2004.00454.x
- Campagne, S., Marsh, M. E., Capitani, G., Vorholt, J. A., and Allain, F. H. (2014). Structural basis for -10 promoter element melting by environmentally induced sigma factors. *Nat. Struct. Mol. Biol.* 21, 269–276. doi: 10.1038/nsmb.2777
- Chaturongakul, S., Raengpradub, S., Palmer, M. E., Bergholz, T. M., Orsi, R. H., Hu, Y., et al. (2011). Transcriptomic and phenotypic analyses identify coregulated, overlapping regulons among PrfA, CtsR, HrcA, and the alternative sigma factors  $\sigma^B$ ,  $\sigma^C$ ,  $\sigma^H$ , and  $\sigma^L$  in *Listeria monocytogenes*. *Appl. Environ. Microbiol.* 77, 187–200. doi: 10.1128/AEM.00952-10
- Chen, C., Pan, J., Yang, X., Guo, C., Ding, W., Si, M., et al. (2016). Global transcriptomic analysis of the response of *Corynebacterium glutamicum* to vanillin. *PLoS One* 11:e0164955. doi: 10.1371/journal.pone.0164955
- Chen, C., Pan, J., Yang, X., Xiao, H., Zhang, Y., Si, M., et al. (2017). Global transcriptomic analysis of the response of *Corynebacterium glutamicum* to ferulic acid. *Arch. Microbiol.* 199, 325–334. doi: 10.1007/s00203-016-1306-5
- Chen, C., Zhang, Y., Xu, L., Zhu, K., Feng, Y., Pan, J., et al. (2018). Transcriptional control of the phenol hydroxylase gene phe of *Corynebacterium glutamicum* by the AraC-type regulator PheR. *Microbiol. Res.* 209, 14–20. doi: 10.1016/j.micres.2018.02.001
- Cho, B. K., Kim, D., Knight, E. M., Zengler, K., and Palsson, B. O. (2014). Genome-scale reconstruction of the sigma factor network in *Escherichia coli*: topology and functional states. *BMC Biol.* 12:4. doi: 10.1186/1741-7007-12-4
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Dainese, E., Rodrigue, S., Delogu, G., Provvedi, R., Laflamme, L., Brzezinski, R., et al. (2006). Posttranslational regulation of *Mycobacterium tuberculosis* extracytoplasmic-function sigma factor sigma L and roles in virulence and in global regulation of gene expression. *Infect. Immun.* 74, 2457–2461. doi: 10.1128/IAI.74.4.2457-2461.2006
- Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Leapman, R. D., et al. (2007). Protein oxidation implicated as the primary determinant of bacterial radioresistance. *PLoS Biol.* 5:e92. doi: 10.1371/journal.pbio.0050092
- Denyer, S. P. (1995). Mechanisms of action of antibacterial biocides. *Int. Biodeterior. Biodegr.* 36, 221–225. doi: 10.1016/0964-8305(96)00015-7
- Dostálová, H., Holátko, J., Busche, T., Rucká, L., Rapoport, A., Halada, P., et al. (2017). Assignment of sigma factors of RNA polymerase to promoters in *Corynebacterium glutamicum*. *AMB Express* 7:133. doi: 10.1186/s13568-017-0436-8

- Ehira, S., Teramoto, H., Inui, M., and Yukawa, H. (2009). Regulation of *Corynebacterium glutamicum* heat shock response by the extracytoplasmic-function sigma factor SigH and transcriptional regulators HspR and HrcA. *J. Bacteriol.* 191, 2964–2972. doi: 10.1128/JB.00112-09
- Engels, S., Schweitzer, J. E., Ludwig, C., Bott, M., and Schaffer, S. (2004). *ClpC* and *clpP1P2* gene expression in *Corynebacterium glutamicum* is controlled by a regulatory network involving the transcriptional regulators ClgR and HspR as well as the ECF sigma factor  $\sigma$ H. *Mol. Microbiol.* 52, 285–302. doi: 10.1111/j.1365-2958.2003.03979.x
- Green, M. R., and Sambrook, J. (2012). *Molecular Cloning: A Laboratory Manual*, Fourth Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Gruber, T. M., and Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* 57, 441–466. doi: 10.1146/annurev.micro.57.030502.090913
- Hanahan, D. (1985). “Techniques for transformation of *E. coli*,” in *DNA Cloning. A Practical Approach*, Vol. 1, ed. D. M. Glover (Oxford: IRL), 109–135.
- Hilker, R., Stadermann, K. B., Doppmeier, D., Kalinowski, J., Stoye, J., Straube, J., et al. (2014). ReadXplorer—visualization and analysis of mapped sequences. *Bioinformatics* 30, 2247–2254. doi: 10.1093/bioinformatics/btu205
- Hilker, R., Stadermann, K. B., Schwengers, O., Anisiforov, E., Jaenicke, S., Weisshaar, B., et al. (2016). ReadXplorer 2-detailed read mapping analysis and visualization from one single source. *Bioinformatics* 32, 3702–3708. doi: 10.1093/bioinformatics/btw541
- Holátko, J., Šilar, R., Rabatinová, A., Šanderová, H., Halada, P., Nešvera, J., et al. (2012). Construction of in vitro transcription system for *Corynebacterium glutamicum* and its use in the recognition of promoters of different classes. *Appl. Microbiol. Biotechnol.* 96, 521–529. doi: 10.1007/s00253-012-4336-1
- Jordan, S., Hutchings, M. I., and Mascher, T. (2008). Cell envelope stress response in Gram-positive bacteria. *FEMS Microbiol. Rev.* 32, 107–146. doi: 10.1111/j.1574-6976.2007.00091.x
- Kalinowski, J., Bathe, J., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., et al. (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.* 104, 5–25. doi: 10.1016/S0168-1656(03)00154-8
- Keilhauer, C., Eggeling, L., and Sahm, H. (1993). Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J. Bacteriol.* 175, 5595–5603. doi: 10.1128/jb.175.17.5595-5603.1993
- Kirchner, O., and Tauch, A. (2003). Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. *J. Biotechnol.* 104, 287–299. doi: 10.1016/S0168-1656(03)00148-2
- Knoppová, M., Phensajjai, M., Veselý, M., Zemanová, M., Nešvera, J., and Pátek, M. (2007). Plasmid vectors for testing *in vivo* promoter activities in *Corynebacterium glutamicum* and *Rhodococcus erythropolis*. *Curr. Microbiol.* 55, 234–239. doi: 10.1007/s00284-007-0106-1
- Kranz, A., Busche, T., Vogel, A., Usadel, B., Kalinowski, J., Bott, M., et al. (2018). RNAseq analysis of alpha-proteobacterium *Gluconobacter oxydans* 621H. *BMC Genomics* 19:24. doi: 10.1186/s12864-017-4415-x
- Krisko, A., Copic, T., Gabaldon, T., Lehner, B., and Supek, F. (2014). Inferring gene function from evolutionary change in signatures of translation efficiency. *Genome Biol.* 15:R44. doi: 10.1186/gb-2014-15-3-r44
- Lane, W. J., and Darst, S. A. (2006). The structural basis for promoter -35 element recognition by the group IV sigma factors. *PLoS Biol.* 4:e269. doi: 10.1371/journal.pbio.0040269
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Luo, Y., and Helmann, J. D. (2009). Extracytoplasmic function sigma factors with overlapping promoter specificity regulate sublancin production in *Bacillus subtilis*. *J. Bacteriol.* 191, 4951–4958. doi: 10.1128/JB.00549-09
- Nakunst, D., Larisch, C., Hüser, A. T., Tauch, A., Pühler, A., and Kalinowski, J. (2007). The extracytoplasmic function-type sigma factor SigM of *Corynebacterium glutamicum* ATCC 13032 is involved in transcription of disulfide stress-related genes. *J. Bacteriol.* 189, 4696–4707. doi: 10.1128/JB.00382-07
- Pátek, M., and Nešvera, J. (2011). Sigma factors and promoters in *Corynebacterium glutamicum*. *J. Biotechnol.* 154, 101–113. doi: 10.1016/j.jbiotec.2011.01.017
- Pfeifer-Sancar, K., Mentz, A., Rückert, C., and Kalinowski, J. (2013). Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique. *BMC Genomics* 14:888. doi: 10.1186/1471-2164-14-888
- Raman, S., Hazra, R., Dascher, C. C., and Husson, R. N. (2004). Transcription regulation by the *Mycobacterium tuberculosis* alternative sigma factor SigD and its role in virulence. *J. Bacteriol.* 186, 6605–6616. doi: 10.1128/JB.186.19.6605-6616.2004
- Rezuchova, B., and Kormanec, J. (2001). A two-plasmid system for identification of promoters recognized by RNA polymerase containing extracytoplasmic stress response  $\sigma$ E in *Escherichia coli*. *J. Microbiol. Methods* 45, 103–111. doi: 10.1016/S0167-7012(01)00237-8
- Ross, W., Thompson, J. F., Newlands, J. T., and Gourse, R. L. (1990). *E. coli* Fis protein activates ribosomal RNA transcription in vitro and in vivo. *EMBO J.* 9, 3733–3742. doi: 10.1002/j.1460-2075.1990.tb07586.x
- Salomon-Ferrer, R., Case, D. A., and Walker, R. C. (2013). An overview of the amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* 3, 198–210. doi: 10.1002/wcms.1121
- Schröder, J., and Tauch, A. (2010). Transcriptional regulation of gene expression in *Corynebacterium glutamicum*: the role of global, master and local regulators in the modular and hierarchical gene regulatory network. *FEMS Microbiol. Rev.* 34, 685–737. doi: 10.1111/j.1574-6976.2010.00228.x
- Schulz, S., Eckweiler, D., Bielecka, A., Nicolai, T., Franke, R., Dotsch, A., et al. (2015). Elucidation of sigma factor-associated networks in *Pseudomonas aeruginosa* reveals a modular architecture with limited and function-specific crosstalk. *PLoS Pathog.* 11:e1004744. doi: 10.1371/journal.ppat.1004744
- Seo, J. H., Hong, J. S., Kim, D., Cho, B. K., Huang, T. W., Tsai, S. F., et al. (2012). Multiple-omic data analysis of *Klebsiella pneumoniae* MGH 78578 reveals its transcriptional architecture and regulatory features. *BMC Genomics* 13:679. doi: 10.1186/1471-2164-13-679
- Šilar, R., Holátko, J., Rucká, L., Rapoport, A., Dostálová, H., Kadeřabková, P., et al. (2016). Use of in vitro transcription system for analysis of *Corynebacterium glutamicum* promoters recognized by two sigma factors. *Curr. Microbiol.* 73, 401–408. doi: 10.1007/s00284-016-1077-x
- Taniguchi, H., Busche, T., Patschkowski, T., Niehaus, K., Pátek, M., Kalinowski, J., et al. (2017). Physiological roles of sigma factor SigD in *Corynebacterium glutamicum*. *BMC Microbiol.* 17:158. doi: 10.1186/s12866-017-1067-6
- Toyoda, K., and Inui, M. (2015). Regulons of global transcription factors in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 100, 45–60. doi: 10.1007/s00253-015-7074-3
- Toyoda, K., and Inui, M. (2016). The extracytoplasmic function sigma factor  $\sigma$ C regulates expression of a branched quinol oxidation pathway in *Corynebacterium glutamicum*. *Mol. Microbiol.* 100, 486–509. doi: 10.1111/mmi.13330
- Toyoda, K., and Inui, M. (2018). Extracytoplasmic function sigma factor  $\sigma$ D confers resistance to environmental stress by enhancing mycolate synthesis and modifying peptidoglycan structures in *Corynebacterium glutamicum*. *Mol. Microbiol.* 107, 312–329. doi: 10.1111/mmi.13883
- van der Rest, M. E., Lange, C., and Molenaar, D. (1999). A heat shock following electroporation induces highly efficient transformation of *Corynebacterium*

- glutamicum* with xenogeneic plasmid DNA. *Appl. Microbiol. Biotechnol.* 52, 541–545. doi: 10.1007/s002530051557
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427
- Wittchen, M., Busche, T., Gaspar, A. H., Lee, J. H., Ton-That, H., and Kalinowski, J. (2018). Transcriptome sequencing of the human pathogen *Corynebacterium diphtheriae* NCTC 13129 provides detailed insights into its transcriptional landscape and into DtxR mediated transcriptional regulation. *BMC Genomics* 19:82. doi: 10.1186/s12864-018-4481-8
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Dostálová, Busche, Holátko, Rucká, Štěpánek, Barvík, Nešvera, Kalinowski and Pátek. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.