Current Microbiology (2022) 79:55
https://doi.org/10.1007/s00284-021-02747-8

Identification of *Rhodococcus erythropolis* Promoters Controlled by Alternative Sigma Factors Using In Vivo and In Vitro Systems and Heterologous RNA Polymerase

Jan Blumenstein[1,2] · Robert Rädisch[1,2] · Václav Štěpánek[1] · Michal Grulich[1] · Hana Dostálová[1] · Miroslav Pátek[1]

Received: 10 June 2021 / Accepted: 10 December 2021 / Published online: 4 January 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Rhodococcus erythropolis CCM2595 is a bacterial strain, which has been studied for its capability to degrade phenol and other toxic aromatic compounds. Its cell wall contains mycolic acids, which are also an attribute of other bacteria of the Mycolata group, such as *Corynebacterium* and *Mycobacterium* species. We suppose that many genes upregulated by phenol stress in *R. erythropolis* are controlled by the alternative sigma factors of RNA polymerase, which are active in response to the cell envelope or oxidative stress. We developed in vitro and in vivo assays to examine the connection between the stress sigma factors and genes activated by various extreme conditions, e.g., heat, cell surface, and oxidative stress. These assays are based on the procedures of such tests carried out in the related species, *Corynebacterium glutamicum.* We showed that the *R. erythropolis* CCM2595 genes *frmB1* and *frmB2*, which encode S-formylglutathione hydrolases (named corynomycolyl transferases in *C. glutamicum*), are controlled by SigD, just like the homologous genes *cmt1* and *cmt2* in *C. glutamicum*. The new protocol of the in vivo and in vitro assays will enable us to classify *R. erythropolis* promoters according to their connection to sigma factors and to assign the genes to the corresponding sigma regulons. The complex stress responses, such as that induced by phenol, could, thus, be analyzed with respect to the gene regulation by sigma factors.

Introduction

Many *Rhodococcus* strains are distinguished by their properties which are useful in biotechnological processes, such as biosynthesis, biodegradation, bioremediation, and bioconversion (for reviews, see [1–3]). The production of lipids by *Rhodococcus opacus* PD630 [4], degradation of polychlorinated biphenyls (PCB) by *Rhodococcus jostii* RHA1 [5], degradation of hydrocarbons by *Rhodococcus pyridinivorans* 5Ap [6], and conversion of acrylonitrile to acrylamide by *Rhodococcus rhodochrous* K22 [7] are prominent examples. In the course of these processes, rhodococci are affected by various stresses which may impair the efficiency of the corresponding biotechnology. For example, toxic effects of heavy metals were observed to decrease the efficiency of bioremediation of sites polluted with hydrocarbons by *Rhodococcus* strains [8]. A negative effect of oxidative stress on lipid production was also detected in an *R. opacus* PD630 and *R. jostii* RHA1 co-culture grown on lignin as the sole carbon source [4]. Conversely, some compounds which play a role in the defense of bacterial cells are produced due to stress. The carotenoid pigments, which protect cells against oxidative stress, can be counted as such compounds [9].

The enzymes involved in the biodegradation of toxic aromatic compounds and many other biotechnological processes are generally well characterized in rhodococci, whereas the knowledge of regulators and mechanisms controlling the gene expression connected to stress responses is still limited (for a review, see [10]). Sigma factors of RNA polymerase (RNAP) are the regulators which are involved in the control of each gene. In addition to the primary sigma factor (named σ^A in many Gram-positive species), that controls transcription of most genes active during exponential growth, variable number of alternative sigma factors classified into groups 2, 3, and 4 occur in different bacterial species [11]. Five stress sigma factors, classified as members of

✉ Miroslav Pátek
patek@biomed.cas.cz

1 Institute of Microbiology of the CAS, v.v.i., Prague, Czech Republic

2 Department of Genetics and Microbiology, Faculty of Science, Charles University, Prague, Czech Republic

group 4 of sigma factors (called also extracytoplasmic sigma factors, ECF σ), constitute the largest group of these RNAP subunits in *Corynebacterium glutamicum*, closely related to *Rhodococcus* species [12]. *Rhodococcus erythropolis* CCM2595, as an example of a *Rhodococcus* strain, has 19 presumptive alternative sigma factors classified into group 3 or 4 according to analysis of the genome sequences [13]. In *R. jostii* RHA1, as many as 32 out of 34 sigma factors are thought to belong to group 3 or 4 [14]. Despite their importance for the coordination of stress responses, very little is known about their functions and about the genes, which they regulate. Most of the consensus sequences of *Rhodococcus* promoters of various classes specifically recognized by individual sigma factors have not yet been defined.

In *R. jostii* RHA1, the *sigF1* and *sigF3* genes (encoding the sigma factors of group 3) were found to be upregulated during heat, salt, and oxidative stresses [15]. We found that many genes under the control of four stress sigma factors were upregulated in response to phenol stress in the phenol-degrading strain, *R. erythropolis* CCM2595 [10]. However, the functions of each sigma factor are not known.

In comparison to the few data known for the *Rhodococcus* sigma factors, current knowledge of the functions of seven sigma factors in the related strain *C. glutamicum* ATCC13032 is much deeper (for a review, see [16]). For example, the roles of most group 4 sigma factors in stress responses in *C. glutamicum* have been uncovered: SigC is involved in enhancing tolerance to oxidative stress [16], while SigD regulates the expression of the genes involved in the mycolate synthesis and lysozyme stress response [17, 18] and SigH mainly regulates the heat and oxidative stress response [19, 20]. Consensus sequences of the corresponding promoter classes have been well defined [16, 18–20].

We have previously described the in vitro transcription system [21] and in vivo two-plasmid assay [22] for *C. glutamicum* to analyze promoters of the stress-responsive *C. glutamicum* genes and sort them into the specific sigma-controlled classes. In brief, the *sig* genes cloned in the plasmid vector pEC-XT99A were expressed, and the sigma proteins bound to RNAP initiated transcription from promoters cloned in the promoter-probe plasmid vector pEPR1 (which replicated in the same cell) with *gfp*uv as a reporter gene [23]. The level of fluorescence intensity was then determined in the cell-free extracts. We documented that this system is particularly valuable for the analysis of promoters of the stress-responsive genes [22, 24].

In this study, we modified these procedures and used them for the analysis of promoters of the related species *R. erythropolis*. This new approach will also enable us to classify the respective *R. erythropolis* genes into sigma regulons and assess their functions in stress responses. The *R. erythropolis* homologs of the SigD-dependent *C. glutamicum* genes *cmt1* and *cmt2* were chosen for testing this system. These genes encoding corynomycolyl transferases are involved in the synthesis of the mycolate-containing cell wall, which is typical for the Mycolata group of bacteria (e.g., *Mycobacterium*, *Corynebacterium* and *Rhodococcus* species). We, therefore, first searched for the *R. erythropolis* CCM2595 homologs of *cmt1* and *cmt2* that have analogous functions.

Materials and Methods

Strains, Plasmids, Growth Conditions, and Oligonucleotides

Escherichia coli DH5α (Invitrogen, USA), used for cloning, and *E. coli* BL21 (DE3) (Thermo Fisher Scientific, USA), used for expression of the genes encoding *R. erythropolis* sigma factors were cultivated in 500-mL flasks containing 60 mL of 2xYT medium (NaCl, 5 g/L, tryptone 16 g/L, yeast extract 10 g/L) [25] at 37 °C. *Corynebacterium glutamicum* ATCC13032 (referred to here as *C. glutamicum* WT) was used as the host for measuring the activities of promoters inserted into the promoter-probe vector pEPR1 [23]. *Corynebacterium glutamicum rpoC*-His10 was constructed in this study for the isolation of recombinant RNAP. *Rhodococcus erythropolis* CCM2595 was used as a source of chromosomal DNA, which served as a template for PCR amplification of the *sig* genes and promoters. Both *C. glutamicum* WT and *R. erythropolis* CCM2595 were cultivated in 500-mL flasks with 60 mL of 2xYT medium at 30 °C. For the selection, antibiotics were added into the media: ampicillin (Ap; 100 μg/mL), tetracycline (Tc; 10 μg/mL), or kanamycin (Km; 30 μg/mL). The plasmids are shown in Table 1. The oligonucleotides are listed in Supplementary Table S1. Maps of the constructed plasmids are in Supplementary Fig. S1.

Table 1 Plasmid vectors used in this study

Plasmid	Characteristics	Source
pEC-XT99A	*E. coli*–*C. glutamicum* expression vector, Tc^R IPTG-inducible *trc* promoter	[26]
pEPR1	*E. coli*–*C. glutamicum* promoter-probe vector, Km^R, promoter-less *gfp*uv as a reporter	[23]
pRLG770	*E. coli* vector, *rrnB* terminator, Ap^R, used for in vitro transcription analysis	[27]

Plasmid constructs for use of CRISPR-Cas9 technique are shown in Supplementary Fig. S1

Designations such as $RNAP_{Cg}$, $RNAP_{Re}$, $sigA_{Re}$, $sigD_{Re}$, $sigH_{Re}$, or $cmt2_{Cg}$ are only used to avoid confusion.

DNA Manipulations

DNA isolation, DNA digestion by restriction enzymes, DNA ligation, PCR, and the transformation of *E. coli* were carried out using the standard techniques [25]. Promoter DNA fragments were generated by annealing synthetized complementary oligonucleotides. This synthetic DNA (Sigma-Aldrich, Germany) produced double-stranded fragments around 70-nt in length, with overhangs ready for ligation with digested pRLG770 and pEPR1. The respective oligonucleotide sequences are shown in Supplementary Table S1.

Engineering the CRISPR-Cas9 Genome Editing System for *C. glutamicum*

To construct a suitable *C. glutamicum* strain for the easy isolation and purification of RNAP, we introduced a 10xHis tag at the C-end of RpoC by modifying the *C. glutamicum* chromosome using the CRISPR-Cas9 genome editing technique. We chose the single-plasmid approach to achieve this genome editing system in *C. glutamicum*. The *E. coli*/*C. glutamicum* shuttle vector pEC-XT99A [26] was used as a basis for the construction of the all-in-one CRISPR-Cas9 genome editing plasmid. To achieve the desired construct named pEC-XKCgrpoC-H10, several cloning steps were done: insertion of the DNA fragments encoding the Cas9 protein, the aminoglycoside phosphotransferase gene (*aph*) as a kanamycin resistance marker and a sequence determining the guide RNA (gRNA). A sequence for modifying the chromosomal *rpoC* gene by homology-directed repair was included. The templates for PCR amplification were chosen as follows.

The Cas9 gene was amplified from the plasmid CAS9P-1EA (Sigma-Aldrich). The *aph* gene was amplified from pEPR1 [23]. The guide RNA targeting *rpoC* in the chromosome was chosen via CRISPOR [28]. The *rpoC* gene was amplified using pKSAC45rpoC-His8 [21] as a template.

Plasmid pKSAC45rpoC-His8 [21] was modified with a Q5 Site-Directed Mutagenesis Kit (New England BioLabs; NEB, USA) with the primers M-His-F and M-His-R to insert a 10xHis tag at the C-end of the RpoC. This modification resulted in the construct pKSAC45-rpoC-H10.

The Cas9 gene was amplified with the primers Cas-RBS-F and Cas-RBS-R, and the plasmid CAS9p-1EA (Sigma-Aldrich) as a template. In this step, the weak RBS (AAAGGTTCTAAAG) [29] was created as part of the fragment RBS+*cas9*. To introduce the PlacUV5 promoter upstream of the *cas9* gene, the fragment RBS+*cas9* was amplified with the primers Cas9-F0 and Cas9-R0 containing PlacUV5 sequence, thus giving rise to the fragment PlacUV5+RBS+*cas9*.

In parallel, the oligonucleotides 0gRNA-F and 0gRNA-F were hybridized to form an empty 0gRNA double-stranded DNA fragment, which was then cloned into the restriction site *Eco*RI/*Xba*I of pEC-XT99A to construct the intermediate plasmid pEC-XT99A/0gRNA.

The empty plasmid pEC-XKCg for the general use of CRISPR-Cas9 genome editing system in *C. glutamicum* (containing all components of CRISPR-Cas9) was designed and constructed as follows: the vector backbone (carrying the *C. glutamicum* replicon) with 0gRNA was amplified from pEC-XT99A/0gRNA with the primers V-F and V-R. The *aph* gene was amplified from pEPR1 with the primers aph-F and aph-R, the *cas9* gene was amplified from the intermediate fragment PlacUV5+RBS+*cas9* with the primers Cas9-F1 and Cas9-R1, and the last fragment carrying the *E. coli* replicon, which originally came from pBR322, was amplified from pEC-XT99A with the primers pBR322-F and pBR322-R. All four fragments were fused using the Gibson Assembly method [30] to construct pEC-XKCg.

The vector backbone of the final specific construct with gRNA targeting *rpoC* was amplified from pEC-XKCg with the primers V-rpoC-H-F and V-rpoC-H-R, *cas9* was amplified from the same plasmid pEC-XKCg with the primers Cas9-F2 and Cas9-R2, and the last fragment containing modified rpoC-H10 was amplified from pKSAC45rpoC-H10-CRISPR with the primers rpoC-F and rpoC-R. All 3 fragments were fused together using the Gibson Assembly technique to construct pEC-XKCgrpoC-H10.

All the primers used (synthetized by Eurofins Scientific, Luxemburg) are listed in Supplementary Table S1. DNA polymerases, restriction enzymes, Gibson Assembly Master Mix and the Q5 Site-Directed Mutagenesis Kit were purchased from New England BioLabs.

Chromosome Modification with the CRISPR-Cas9 Technique

The final plasmid pEC-XKCgrpoC-H10 was transferred by electroporation [31] into *C. glutamicum* cells. The colonies only appeared after 4 days of incubation at 30 °C. 2xYT medium was inoculated with the chosen clones and induced with 0.1 mM IPTG. A culture without the addition of IPTG was used as a control. After ON cultivation, the cells were plated. The grown colonies were picked for DNA isolation, PCR amplification and verification by sequencing. Three resulting control clones (without IPTG induction) exhibited no insertion of 10xHis triplets at the 3-end of *rpoC* according to the sequencing. In contrast, 50% of clones (2/4) resulting from the cultivation with IPTG carried the sequence encoding the poC protein with a 10xHis tag at its C terminus within the chromosome.

This fact confirmed that the final aim of modifying *rpoC* in *C. glutamicum* WT to produce the *C. glutamicum rpoC*-His10 strain was achieved.

Isolation and Purification of the RNA Polymerase Core from *C. glutamicum rpoC*-His10

The protocol for isolating the *C. glutamicum* His10-tagged RNAP was based on the procedure used for *Bacillus subtilis* RNAP [32] and modified for *C. glutamicum* [21]. *Corynebacterium glutamicum rpoC*-His10 cells were cultivated in a 10-L Biostat MD stirred bioreactor (B. Braun Biotech International, Germany) with an initial working volume of 6 L of 2xYT and 0.5% glucose to $OD_{600} = 16$, harvested by centrifugation, washed, resuspended in phosphate buffer (300 mM NaCl, 20 mM Na_2HPO_4, 20 mM NaH_2PO_4, 3 mM 2-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride, pH 8.0) and disrupted by sonication. Cell debris was removed by centrifugation, and the supernatant was mixed with 6 mL of TALON Metal Affinity Resin (TaKaRa, China). The mixture was incubated for 1.5 h at 4 °C with gentle shaking, the TALON pellet was collected by centrifugation, transferred into three TALON 2-mL Disposable Gravity Columns (TaKaRa) and washed with a phosphate buffer gradient of imidazole (20–500 mM) in phosphate buffer, and the protein content of each fraction was assessed by Bradford assay and SDS-PAGE (10% polyacrylamide). The highest content of RNAP core subunits was detected in the fractions eluted with phosphate buffer containing 100 mM imidazole. These fractions were combined and dialyzed in a D-Tube Dialyzer Maxi, 6–8 kDa (Merck, Germany) against storage buffer (50 mM Tris, 100 mM NaCl, 50% glycerol, 3 mM 2-mercaptoethanol, pH 8.0) overnight. The RNAP core preparation was stored at – 80 °C. The designation $RNAP_{Cg}$ or $RNAP_{Re}$ is only used when confusion is likely.

Promoter Activity Determination by the In Vivo Two-Plasmid System

The activity of the promoters cloned in the promoter-probe vector pEPR1 was assayed in the two-plasmid system as described previously [22]. *Corynebacterium glutamicum* WT cells harbored the expression vector pEC-XT99A with the inserted genes encoding the tested sigma factors, and the promoter-probe vector pEPR1 [23] carrying the tested promoter (approx. 70-nt) and the *gfp*uv reporter gene. The intensity of the fluorescence was measured with a Spark multimode microplate reader (Tecan, Austria) with excitation at 397 nm and emission at 509 nm.

In Vitro Transcription

The homologous and heterologous in vitro transcription reactions were carried out in principle in the same way as originally designed [21]. The promoter DNA (approx. 70-nt fragments) in the vector pRLG770 served as a template for PCR. The PCR fragments (350–400 bp) were amplified with Q5 polymerase (New England BioLabs) with the primers 30F and CM3, purified by phenol extraction and concentrated with an Amicon Ultra 0.5 mL s NMWL 30,000 (Merck). The RNAP holoenzyme was produced by mixing the RNAP core (100 nM) isolated from *C. glutamicum rpoC*-His10 and the sigma factors SigA, SigD or SigH (from *C. glutamicum* or *R. erythropolis*) isolated as His-tagged proteins from *E. coli* in a molar ratio of 1:30. The transcription reaction with [α-^{32}P]UTP and unlabeled ATP, CTP, GTP (in final concentration 0.5 mM each NTP) was run for 10 min at 37 °C. The transcripts labeled with [α-^{32}P]UTP were subjected to polyacrylamide gel electrophoresis, and the RNA signals were quantified by phosphorimager analysis. The signals were scanned with a Typhoon Scanner (GE Healthcare, USA) and analyzed with the software ImageQuant TL. In vitro transcription was performed twice with results showing the bands of similar strength and in closely similar positions in the electrophoretograms representing specific transcript.

Results

Assignment of *R. erythropolis* Sigma Factors to Promoters by the In Vivo Two-Plasmid System

We tried to develop a system to assign particular *R. erythropolis* sigma factors to promoters in vivo analogous to such system, which we developed for *C. glutamicum* [22]. Such assay for *Rhodococcus* strains could be used in the studies of the *Rhodococcus* stress response (for a review, see [10]). However, the transformation of *Rhodococcus* cells was much more laborious, and cells harboring two plasmids appeared unstable (data not shown). We therefore decided to construct the heterologous *Corynebacterium*–*Rhodococcus* system. *Corynebacterium glutamicum* cells harbored the same vectors (pEC-XT99A and pEPR1) as in the previous study [22], but both the cloned *sig* genes and promoters came from *R. erythropolis* CCM2595.

To test this system, we first went through the list of the already described *C. glutamicum* stress-responding genes with defined promoters and identified potential analogous *R. erythropolis* CCM2595 genes and promoters. Then, we selected two genes controlled by the SigD-dependent promoters of *C. glutamicum*, *cmt1* and *cmt2*, which we analyzed previously [24]. The homologous genes *frmB1* and *frmB2*

were detected in the *R. erythropolis* CCM2595 genome. The potential promoters P*frmB1* and P*frmB2* with identical or closely similar key sequence elements − 35 and − 10 of P*cmt1* and P*cmt2* (Fig. 1) were also found by sequence inspection. Corynomycolyl transferases (Cmt1 and Cmt2) and S-formylglutathione hydrolases (FrmB1 and FrmB2) are the synonymous names for their homologous protein products (Table 2).

To test the designed two-plasmid system, we used the *C. glutamicum* WT cells transformed with both pEC-XT99A$sigD_{Re}$ (or pEC-XT99A$sigH_{Re}$) and pEPR1-P$frmB1_{Re}$ (or pEPR1-P$frmB2_{Re}$). The presence of the plasmids was checked by restriction enzyme analysis and PCR as described previously [22]. We then determined promoter activity by using the designed two-plasmid system producing $SigD_{Re}$ or $SigH_{Re}$ (Fig. 2B, C), whereas the cells carrying the empty vector pEC-XT99A were used as a control (gray bars in Fig. 2). The promoters P$frmB1_{Re}$ and P$frmB2_{Re}$ were only active with SigD (Fig. 2B, C). SigD dependence was also observed when the *C. glutamicum* WT promoter of the $cmt2_{Cg}$ gene and *C. glutamicum* WT sigma factors were used (Fig. 2A).

Table 2 Amino acid identity levels of Cmt (corynomycolyl transferases) from *C. glutamicum* and FrmB (S-formylglutathione hydrolases) from *R. erythropolis* evaluated by BLASTP search (non-homologous N-ends of proteins not included)

Protein	Identity (%)
Cmt1 vs. FrmB1	30
Cmt1 vs. FrmB2	30
Cmt2 vs. FrmB1	48
Cmt2 vs. FrmB2	45
Cmt1 vs. Cmt2	30
FrmB1 vs. FrmB2	77

Construction of Heterologous System with *C. glutamicum* RNA Polymerase Core for In Vitro Transcription

Since our aim was to use the *C. glutamicum* RNAP to also prove its activity with *R. erythropolis* sigma factors and promoters in vitro, we first isolated the $RNAP_{Cg}$ core from *C. glutamicum rpoC*-H8, which was constructed previously [21]. However, $RNAP_{Cg}$ could not be purified satisfactorily, and moreover the strain *C. glutamicum rpoC*-H8 was not stable enough (data not shown). We therefore constructed a new strain, *C. glutamicum rpoC*-His10 using the CRISPR-Cas9 genome editing technique. The $RNAP_{Cg}$ core containing the His10-tagged β′ subunit (RpoC) was isolated by affinity chromatography and used for in vitro transcription. The sigma factors $SigA_{Re}$, $SigD_{Re}$ and $SigH_{Re}$ were isolated from the *E. coli* strains (Fig. 3) overexpressing the $sigA_{Re}$, $sigD_{Re}$ and $sigH_{Re}$ genes in a similar way as the *C. glutamicum* sigma factors [21].

The RNAP holoenzyme was reconstituted using the $RNAP_{Cg}$ core isolated from *C. glutamicum rpoC*-His10 and *R. erythropolis* CCM2595 sigma factors. We only focused on $SigA_{Re}$, $SigD_{Re}$ and $SigH_{Re}$ in this study. The optimization of the RNAP activity showed that the molar ratio of the RNAP core and the respective sigma factors 1:30 provided the most reliable results (data not shown).

In Vitro Transcription with *C. glutamicum* RNAP Core and *R. erythropolis* Sigma Factors

The activity of the chosen *R. erythropolis* promoters, P$frmB1_{Re}$ and P$frmB2_{Re}$, was assayed for the ability to initiate transcription in vitro in the designed heterologous system (Fig. 4). The signals representing the specific transcripts were always only obtained with $SigD_{Re}$, whereas no signal was observed without the addition of any sigma factor, nor with $SigA_{Re}$ nor $SigH_{Re}$. As a control, in vitro transcription was also run with the proven SigD-dependent *C. glutamicum* promoter P*cmt2* [24] of the *cmt2* gene (Fig. 4B, C). In this case, a homologous system (both the RNAP core and sigma factors from *C. glutamicum*) was used. The specific signal was only observed with $SigD_{Cg}$, whereas other assays only provided unspecific bands (Fig. 4A). Alternatively, these bands might represent

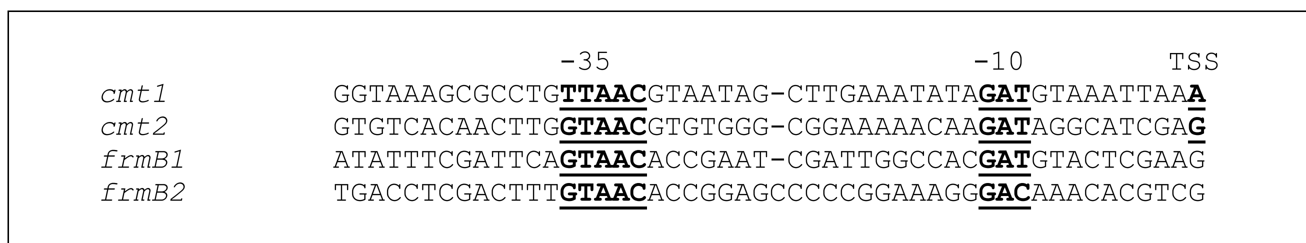
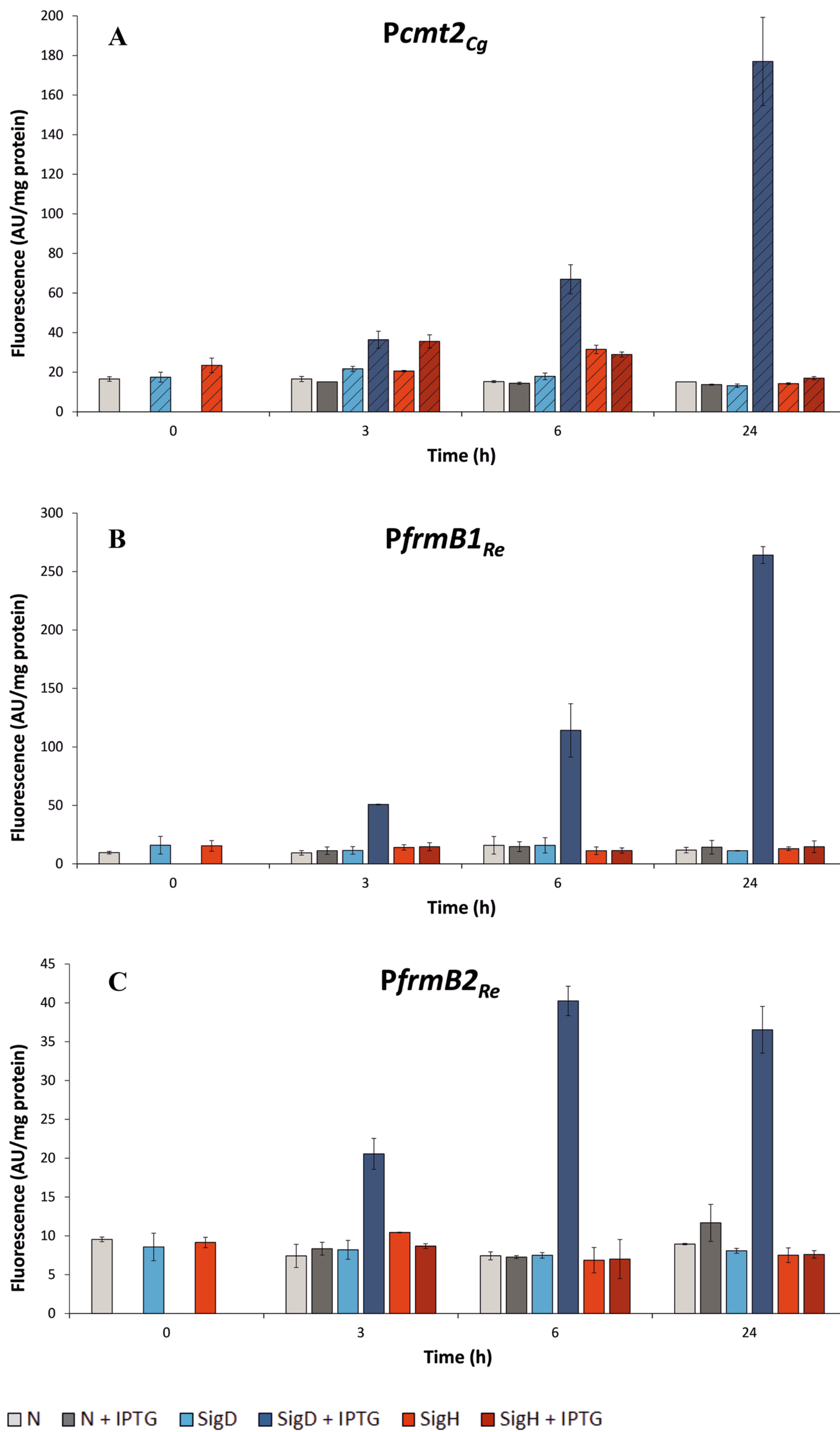


Fig. 1 Alignment of promoter sequences of *C. glutamicum* WT genes *cmt1* and *cmt2* and homologous *R. erythropolis* CCM2595 genes *frmB1* (O5Y_RS01090) and *frmB2* (O5Y_RS25565). Transcriptional start sites + 1 localized previously [32] and the potential key promoter sequence elements − 10 and − 35 are in bold and underlined. The sequences were found in the GenBank database: *cmt1* and *cmt2* (in the complete genome nucleotide sequence of *C. glutamicum* ATCC13032, Acc. Number BX927147), *frmB1* and *frmB2* (in the complete genome nucleotide sequence of *R. erythropolis* CCM2595; Acc. Number NC_022115)



A
Pcmt2Cg
Fluorescence (AU/mg protein)
200
180
160
140
120
100
80
60
40
20
0
0
3
6
24
Time (h)
B
PfrmB1Re
Fluorescence (AU/mg protein)
300
250
200
150
100
50
0
0
3
6
24
Time (h)
C
PfrmB2Re
Fluorescence (AU/mg protein)
45
40
35
30
25
20
15
10
5
0
0
3
6
24
Time (h)
N
N + IPTG
SigD
SigD + IPTG
SigH
SigH + IPTG

◂**Fig. 2** Determination of promoter activity by in vivo two-plasmid test. The *C. glutamicum* WT cells harbored the vector pEC-XT99A carrying the *sigH* or *sigD* genes from *C. glutamicum* (**A**) and *sigH* or *sigD* from *R. erythropolis* (**B**, **C**), respectively. Expression of the *sig* genes was induced with IPTG at time point 0. As a second plasmid, cells harbored the promoter probe vector pEPR1 (carrying the *gfp*uv reporter gene) and promoter P$cmt2_{Cg}$, P$frmB1_{Re}$, or P$frmB2_{Re}$. Open bars show the fluorescence of the clones with sigma factors from *R. erythropolis* CCM2595, and hatched bars represent the fluorescence levels of the clones with sigma factors from *C. glutamicum* WT. Gray bars show the values for control cells without the cloned *sig* genes. *AU* arbitrary units. The standard deviations of three biological replicates are depicted with error bars

longer transcripts driven from *C. glutamicum* promoters. This explanation cannot be, however, applicable for the lane N, where no sigma was added (Fig. 4A). We concluded that both in vivo and in vitro assays proved that the tested *R. erythropolis* promoters are SigD-dependent. In conclusion, both designed heterologous systems (in vivo and in vitro) proved that P$frmB1_{Re}$ and P$frmB2_{Re}$ are SigD controlled. These two independent methods seem to be useful for the analysis of *R. erythropolis* promoters.

Discussion

Sigma factors of RNAP play a major role in the responses of bacterial cells to various stresses and changing growth conditions. Each sigma factor is responsible for the expression of a gene group (regulon), which enables cells to cope with specific changes in the environment [11].

Data on the sigma factors in rhodococci are somewhat scarce. Generally, we assume that the functions of the individual sigma factors are similar to those in other bacterial genera of the Mycolata group, *Mycobacterium*, *Corynebacterium* and *Nocardia.* However, the numbers of sigma factors are very different in these genera, and also differ markedly between the *Rhodococcus* species. The most variable are the sigma factors of groups 3 and 4, which are mostly involved in various stress responses [10]. We studied the phenol stress response in *C. glutamicum* and *R. erythropolis*, as an example of the response to the presence of toxic aromatic compounds, which are degraded by these bacteria (our unpublished results). According to the preliminary results, hundreds of their genes were upregulated or downregulated in the presence of phenol [24]. In *C. glutamicum*, SigD was found to control genes which are involved in oxidative, chemical and cell envelope stress responses [17, 18, 24]. The *C. glutamicum* WT genes *cmt1* and *cmt2* (encoding corynomycolyl transferases), which contribute to the synthesis of mycolic acids in the cell wall, were also found to be members of the SigD regulon [24]. We supposed that the homologous genes in *R. erythropolis*, whose cell wall also contains mycolic acid, are suitable for testing the techniques for assigning sigma factors to promoters and genes in *R. erythropolis*. We therefore analyzed the *R. erythropolis* CCM2595 genes *frmB1* and *frmB2*, which encode the enzyme called S-formylglutathione hydrolase, which has the same or similar activity to Cmt1 and Cmt2. The level of aa-identity between Cmt and FrmB was 30–45%, which is similar to the identity level between Cmt1 and Cmt2 (Table 2).

To make use of the developed in vivo and in vitro systems for *C. glutamicum* sigma factors and promoters, we decided to modify them and apply them to *R. erythropolis*. We have shown previously that the RNAP core from *B. subtilis* or *E. coli* can be functional with the *C. glutamicum* sigma factors in vitro [21]. In two-plasmid in vivo systems, the activity of the hybrid holo-RNAP composed of *E. coli* RNAP and sigma factors from *Mycobacterium tuberculosis* [33] or *Staphylococcus aureus* [34] was proved. This is in agreement with the fact that RNAP is highly conserved in bacteria. In this study, we proved that *R. erythropolis* sigma factors may function with the *C. glutamicum* RNAP core. Our further experiments combining various *R. erythropolis* CCM2595 sigma factors with *C. glutamicum* RNAP to define promoter classes and sigma regulons in *R. erythropolis* are in progress. There could be an advantage of such a heterologous system that the expressed *R. erythropolis* sigma factor gene is not present in the *C. glutamicum* WT genome, and possible interference in the assay is thus avoided.

Both in vivo and in vitro assays proved that the genes *frmB1* and *frmB2* are controlled by SigD in *R. erythropolis*. The functions of SigD in *C. glutamicum* and *R. erythropolis* are thus probably analogous. − 35 and − 10 sequence elements identical or closely similar to the consensus of promoters recognized by SigD in *C. glutamicum* (GTAAC-$N_{18\text{-}19}$-GAT) were found within the predicted and confirmed promoter fragments (Fig. 1). This suggests that the genes under control of at least some sigma factors may be driven from promoters belonging to the same class in *C. glutamicum* and *R. erythropolis*. The conclusion is that these genes may be controlled by the respective SigD, and their expression is part of the response to the stress inflicted by cell envelope damage.

SigD function has not been described in any *Rhodococcus* or *Nocardia* species. By database search, we found two *R. jostii* RHA1 genes with protein products with similarity to mycolyltransferases. Sequence motifs (GTAAC-N_{18}-GAT or GCAAC-N_{18}-GAT) identical or closely similar to the − 35 and − 10 elements of *C. glutamicum* SigD-dependent promoters could be recognized closely upstream of these genes.

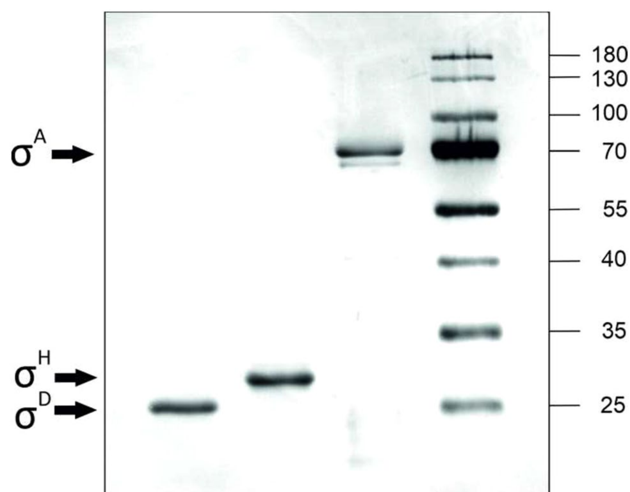


Fig. 3 Isolated *Rhodococcus erythropolis* CCM2595 sigma subunits of RNA polymerase (silver-stained SDS-polyacrylamide gel). His-tagged sigma factors SigA, SigD, and SigH were isolated from *E. coli*, which expressed the corresponding *R. erythropolis sig* genes, using affinity chromatography. Protein size marker is on the right. Molecular sizes in kilodaltons are indicated

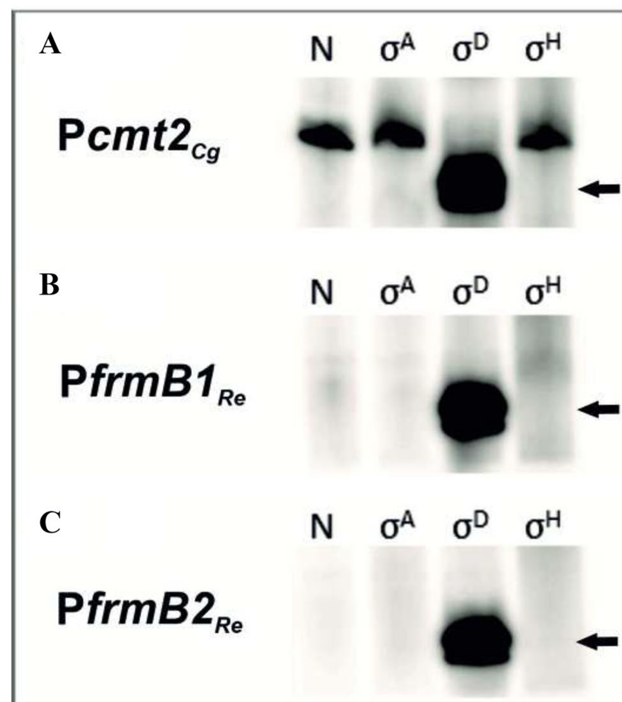


Fig. 4 In vitro transcription with recombinant *C. glutamicum* RNA polymerase core and sigma factors from *C. glutamicum* (**A**) and *R. erythropolis* (**B**, **C**). Sigma factors and the promoter came from *C. glutamicum* in **A**, whereas the sigma factors and promoter came from *R. erythropolis* CCM2595 in **B** and **C**. RNA polymerase core is from *C. glutamicum* in all cases

The search within the *Nocardia brasiliensis* ATCC700358 HUJEG-1 genome revealed the motifs GTAAC-N_{18}-GAC upstream of two genes which putatively encode mycolyltransferases. The role of SigD in *M. tuberculosis* was found to be connected to the pathogenicity and expression of ribosomal genes in the stationary growth phase. However, the putative promoter sequences of the respective genes regulated by SigD were very diverse, and a consensus sequence could not be easily defined [35]. Since the transcriptional starts of these SigD-dependent genes were not experimentally detected by these authors, and potential promoter sequences were chosen according to their similarity to *B. subtilis* promoters recognized by SigW, we suppose that many of these sequences do not function as promoters.

Three SigD-dependent *M. tuberculosis* promoters of genes with experimentally determined transcriptional start sites were proved in another study [36]. The consensus GTA ACG in the – 35 region of these promoters was deduced, whereas no consensus was found in the – 10 region. This result indicated that at least the – 35 sequence of some SigD-dependent *M. tuberculosis* promoters is nearly identical (5/6) to the – 35 motif in the two proved SigD-dependent *R. erythropolis* promoters.

When analyzing the genomes of other *Mycobacterium* strains, *M. leprae* TN and *M. ulcerans* Agy99, we also discovered several genes encoding putative mycolyltransferases. Upstream of several *M. leprae* TN and *M. ulcerans* Agy99 genes, GTAAC-N_{18}-GAT sequence elements or highly similar motifs were detected. These similarities in gene functions and putative promoter sequences tempted us to hypothesize that the enzymes related to mycolate synthesis are encoded by genes which are members of the SigD regulon in some Mycolata strains.

We improved the strain for the isolation and purification of RNAP from *C. glutamicum* using the CRISPR-Cas9 technique in this study. The constructed plasmid pEC-XKCg is ready for any further use in editing the *C. glutamicum* genome. The recombinant RNAP carrying a 10xHis tag at the C-end could be used for in vitro transcription in both *C. glutamicum* and *R. erythropolis*. The detection of transcriptional starts by RNA-sequencing [37] that localizes promoters in the genome-wide range can be thus combined with two independent techniques analyzing single promoters and genes in *Rhodococcus* strains. This combination may finally provide consistent and reliable results.

The development and use of molecular techniques in the Mycolata genera (*Corynebacterium, Mycobacterium*, *Rhodococcus*, *Nocardia*) is still on its beginning in some aspects. Description of promoters, definition of sigma regulons, roles of sigma factors in stress responses, and progress in genome editing will facilitate advances in synthetic/systems biology of these bacteria important in biotechnology and medical studies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00284-021-02747-8.

Acknowledgements We wish to thank Jan Nešvera and Jiří Matějů (Institute of Microbiology, Prague) for critical reading of the manuscript and helpful comments.

Author Contributions HD and MP conceived the study and wrote the initial draft. HD performed in vitro transcription, JB performed cloning and in vivo assays, RR performed the CRISPR-Cas9 constructions, VŠ carried out mutagenesis and database searches, and MG isolated and analyzed the proteins. All the authors read, edited, and approved the final manuscript.

Funding This work was supported by Grant 18-13254S from the Czech Science Foundation, Grants from Charles University (Projects GAUK 452120 and 338321), and Institutional Research Project RVO61388971 from the Institute of Microbiology of the CAS.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Martínková L, Uhnákova B, Pátek M, Nešvera J, Křen V (2009) Biodegradation potential of the genus *Rhodococcus*. Environ Int 35:162–177. https://doi.org/10.1016/j.envint.2008.07.018
2. Donini E, Firrincieli A, Cappelletti M (2021) Systems biology and metabolic engineering of *Rhodococcus* for bioconversion and biosynthesis processes. Folia Microbiol 66:701–713. https://doi.org/10.1007/s12223-021-00892-y
3. Cappelletti M, Presentato A, Piacenza E, Firrincieli A, Turner RJ, Zannoni D (2020) Biotechnology of *Rhodococcus* for the production of valuable compounds. Appl Microbiol Biotechnol 104:8567–8594. https://doi.org/10.1007/s00253-020-10861-z
4. Li X, He Y, Zhang L, Xu Z, Ben H, Gaffrey MJ, Yang Y, Yang S, Yuan JS, Qian WJ, Yang B (2019) Discovery of potential pathways for biological conversion of poplar wood into lipids by co-fermentation of *Rhodococci* strains. Biotechnol Biofuels 12:60–75. https://doi.org/10.1186/s13068-019-1395-x
5. Wang M, Chen J, Yu H, Shen Z (2018) Improving stress tolerance and cell integrity of *Rhodococcus ruber* by overexpressing small-shock-protein Hsp16 of *Rhodococcus*. J Ind Microbiol Biotechnol 45:929–938. https://doi.org/10.1007/s10295-018-2066-9
6. Bukliarevich HA, Charniauskaya MI, Akhremchuk AE, Valentovich LN, Titok MA (2019) Effect of the structural and regulatory heat shock proteins on hydrocarbon degradation by *Rhodococcus pyridinivorans* 5Ap. Microbiology 88:573–579. https://doi.org/10.1134/s0026261719050023
7. Kobayashi M, Yanaka N, Nagasawa T, Yamada H (1992) Primary structure of an aliphatic nitrile-degrading enzyme, aliphatic nitrilase, from *Rhodococcus rhodochrous* K22 and expression of its gene and identification of its active site residue. Biochemistry 31:9000–9007. https://doi.org/10.1021/bi00152a042
8. Kuyukina MS, Krivoruchko A, Ivshina IB (2018) Hydrocarbon- and metal-polluted soil bioremediation: progress and challenges. Microbiol Aust 39:133–136. https://doi.org/10.1071/MA18041
9. Zheng YT, Toyofuku M, Nomura N, Shigeto S (2013) Correlation of carotenoid accumulation with aggregation and biofilm development in *Rhodococcus* sp. SD-74. Anal Chem 85:7295–7301. https://doi.org/10.1021/ac401188f
10. Pátek M, Grulich M, Nešvera J (2021) Stress response in *Rhodococcus* strains. Biotechnol Adv. https://doi.org/10.1016/j.biotechadv.2021.107698
11. Gruber TM, Gross CA (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol 57:441–466. https://doi.org/10.1146/annurev.micro.57.030502.090913
12. Pátek M, Dostálová H, Nešvera J (2020) Sigma factors of RNA polymerase in *Corynebacterium glutamicum*. In: *Corynebacterium glutamicum*, biology and biotechnology. pp 89–112. https://doi.org/10.1007/978-3-030-39267-3_4
13. Strnad H, Pátek M, Fousek J, Szokol J, Ulbrich P, Nešvera J, Pačes V, Vlček Č (2014) Genome sequence of *Rhodococcus erythropolis* strain CCM2595, a phenol derivatives degrading bacterium. Genome Announc. https://doi.org/10.1128/genomeA.00208-14
14. McLeod MP, Warren RL, Hsiao WWK, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJM, Holt R, Brinkman FSL, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103:15582–15587. https://doi.org/10.1073/pnas.0607048103
15. Ekpanyaskun P (2006) Transcriptomic analysis of *Rhodococcus* sp. RHA1 responses to heat shock and osmotic stress. Master Thesis, University of British Columbia, Vancouver
16. Toyoda K, Inui M (2016) The extracytoplasmic function σ factor σ^C regulates expression of a branched quinol oxidation pathway in *Corynebacterium glutamicum*. Mol Microbiol 100:486–509. https://doi.org/10.1111/mmi.13330
17. Taniguchi H, Busche T, Patschkowski T, Niehaus K, Pátek M, Kalinowski J, Wendisch VF (2017) Physiological roles of sigma factor SigD in *Corynebacterium glutamicum*. BMC Microbiol 17:158–168. https://doi.org/10.1186/s12866-017-1067-6
18. Toyoda K, Inui M (2018) Extracytoplasmic function sigma factor σ^D confers resistance to environmental stress by enhancing mycolate synthesis and modifying peptidoglycan structures in *Corynebacterium glutamicum*. Mol Microbiol 107:312–329. https://doi.org/10.1111/mmi.13883
19. Busche T, Šilar R, Pičmanová M, Pátek M, 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. https://doi.org/10.1186/1471-2164-13-445
20. Ehira S, Teramoto H, Inui M, 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. https://doi.org/10.1128/JB.00112-09
21. Holátko J, Šilar R, Rabatinová A, Šanderová H, Halada P, Nešvera J, Krásny L, Pátek M (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. https://doi.org/10.1007/s00253-012-4336-1
22. Dostálová H, Holátko J, Busche T, Rucká L, Rapoport A, Halada P, Nešvera J, Kalinowski J, Pátek M (2017) Assignment of sigma factors of RNA polymerase to promoters in *Corynebacterium*

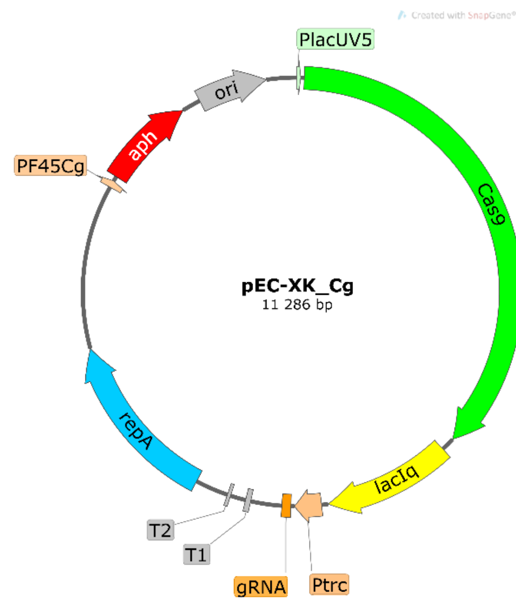
glutamicum. AMB Express 7:133–146. https://doi.org/10.1186/s13568-017-0436-8
23. Knoppová M, Phensaijai M, Veselý M, Zemanová M, Nešvera J, Pátek M (2007) Plasmid vectors for testing *in vivo* promoter activities in *Corynebacterium glutamicum* and *Rhodococcus erythropolis*. Curr Microbiol 55:234–239. https://doi.org/10.1007/s00284-007-0106-1
24. Dostálová H, Busche T, Holátko J, Rucká L, Štěpánek V, Barvík I, Nešvera J, Kalinowski J, Pátek M (2019) Overlap of promoter recognition specificity of stress response sigma factors SigD and SigH in *Corynebacterium glutamicum* ATCC 13032. Front Microbiol 9:3287–3304. https://doi.org/10.3389/fmicb.2018.03287
25. Green MR, Sambrook J (2012) Molecular cloning: a laboratory manual, 4th edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
26. Kirchner O, Tauch A (2003) Tools for genetic engineering in the amino acid producing bacterium *Corynebacterium glutamicum*. J Biotechnol 104:287–299. https://doi.org/10.1016/S0168-1656(03)00148-2
27. Ross W, Gourse RL (2009) Analysis of RNA polymerase-promoter complex formation. Methods 47:13–24. https://doi.org/10.1016/j.ymeth.2008.10.018
28. Concordet JP, Haeussler M (2018) CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res 2:242–245. https://doi.org/10.1093/nar/gky354
29. Zhang B, Zhou N, Liu YM, Liu C, Lou CB, Jiang CY, Liu SJ (2015) Ribosome binding site libraries and pathway modules for shikimic acid synthesis with *Corynebacterium glutamicum*. Microb Cell Factories 14:1–14. https://doi.org/10.1186/s12934-015-0254-0
30. Gibson DG, Young L, Chuang RY, Venter JC, Hutchinson CA III, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345. https://doi.org/10.1038/nmeth.1318
31. van der Rest ME, Lange C, 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. https://doi.org/10.1007/s002530051557
32. Qi Y, Hulett FM (1998) PhoP~P and RNA polymerase σ^A holoenzyme are sufficient for transcription of Pho regulon promoters in *Bacillus subtilis*: PhoP~P activator sites within the coding region stimulate transcription *in vitro*. Mol Microbiol 28:1187–1197. https://doi.org/10.1046/j.1365-2958.1998.00882.x
33. Homerova D, Surdova K, Mikusova K, Kormanec J (2007) Identification of promoters recognized by RNA polymerase containing *Mycobacterium tuberculosis* stress-response sigma factor σ^F. Arch Microbiol 187:185–197. https://doi.org/10.1007/s00203-006-0185-6
34. Homerova D, Bischoff M, Dumolin A, Kormanec J (2004) Optimization of a two-plasmid system for the identification of promoters recognized by RNA polymerase containing *Staphylococcus aureus* alternative sigma factor σ^B. FEMS Microbiol Lett 232:173–179. https://doi.org/10.1016/S0378-1097(04)00063-1
35. Calamita H, Ko C, Tyagi S, Yoshimatsu T, Morrison NE, Bishai WR (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. https://doi.org/10.1111/j.1462-5822.2004.00454.x
36. Raman S, Hazra R, Dascher CC, Husson RN (2004) Transcription regulation by the *Mycobacterium tuberculosis* alternative sigma factor SigD and its role in virulence. J Bacteriol 186:6605–6616. https://doi.org/10.1128/JB.186.19.6605-6616.2004
37. Albersmeier A, Pfeifer-Sancar K, Rückert C, 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. https://doi.org/10.1016/j.jbiotec.2017.04.008

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Materials:

Table S1: Oligonucleotides used

Oligonucleotide	Sequence
0gRNA-F	AATTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTTT
0gRNA-R	CTAGAAAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACG
30F	CCACCTGACGTCTAAGAAACC
Cas-RBS-F	AAAAAACCTGCAGGTAAAAAGGTTCTAAAGATGGACAAAAAGTACAGCATCG
Cas-RBS-R	AAAAAACGATCGTTACACTTTGCGCTTTTTCTTGG
Cas9-F0	AATGCATTTACACTTTGCGC
Cas9-R0	GTGCAATTGTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGACCTGCAGGTAAAAAGGTTCTAAAG
Cas9-F1	TCGGTACCACCGGTCTTTACACTTTATGCTTCCG
Cas9-R1	GGTGTCAACGTAAATGCATTTACACTTTGCGCTTTTTC
Cas9-F2	CGTCGTTTGCGGTGTGGTGTCGGTACCACCTTTACACTTTATGCTTCCGGCTCG
Cas9-R2	CCATCCAATGGGACGGATGCGAATTCCATGGTCTGTTTCCTGTG
Cm3	CGGCGGATTTGTCCTACT
V-F	GAAAAAGCGCAAAGTGTAAATGCATTTACGTTGACACCATCGAATG
V-R	CGAATGAGGATCAGATCAAATCTAAGATAAAGTTATCGATGATAAAGTGCAAC
aph-F	GATAACTTTATCTTAGATTTGATCTGATCCTCATTCG
aph-R	CGCTCAGTGGAACGAAAACTAAAACGATTCCGAAGCCC
pBR322-F	CGGAATCGTTTTAGTTTTCGTTCCACTGAGCGT
pBR322-R	CGGAAGCATAAAGTGTAAAGACCGGTGGTACCGACGTCCATA
V-rpoC-H-F	ATGGAATTCGCATCCGTCCCATTGGATGGTTTTAGAGCTAGAAATAGC
V-rpoC-H-R	GCTGACGTCCATATGCGGTGTGAAATAC
rpoC-F	CACCGCATATGGACGTCAGCCTGCAGAGCTGCGTTC
rpoC-R	GGTGGTACCGACACCACACCGCAAACGACG
M-His-F	CACCACCACCACCACTAGAGCAAGGAAACCTTAAATC
M-His-R	GTGGTGGTGGTGGTGGAAAATTTCATCCAATGGGACGG



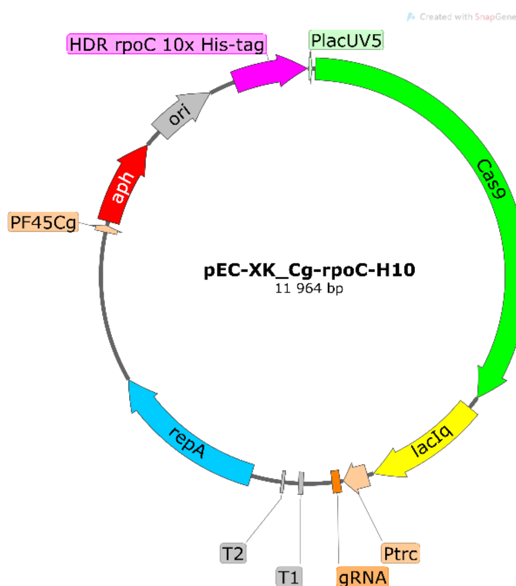


Fig. S1: Plasmid constructs for modification of the *rpoC* gene by CRISPR-Cas9 technique. (A) Plasmid pEC-XKCg. The empty plasmid pEC-XKCg can be generally used for CRISPR-Cas9-genome editing in *C. glutamicum.* (**B)** Plasmid pEC-XKCgrpoc-H10 for construction of 10xHis-tagged C-end of RpoC. Fragment rpoC-10xHis for modifying (tagging) the chromosomal *rpoC* gene by homology-directed repair, and specific guide RNA targeting *rpoC* in the chromosome. Construction of the plasmids in detail is described in the main text Engineering the CRISPR-Cas9 Genome Editing System for *C. glutamicum*. Abbreviations: aph = aminoglycoside phosphotransferase gene conferring kanamycine resistance; cas9 = gene for CRISPR associated protein 9**;** gRNA = guide RNA; lacIq = lac repressor gene; ori = origin of replication in *Escherichia coli*; PF45Cg = constitutive promoter; PlacUV5 = IPTG inducible promoter; Ptrc = IPTG inducible promoter repA – origin of replication in *C. glutamicum*; T1, T2 – terminators

OXFORD FEMS

FEMS Microbiology Letters, 2022, **369**, 1–10
DOI: 10.1093/femsle/fnac014
Advance access publication date: 15 February 2022
Research Letter – Physiology, Biochemistry & Genetics

Sigma regulatory network in *Rhodococcus erythropolis* CCM2595

Václav Štěpánek[1], Hana Dostálová[1], Tobias Busche[2], Jan Blumenstein[1,3], Michal Grulich[1], Lukáš Plašil[1,4], Lenka Rucká[1], Jan Nešvera[1] and Miroslav Pátek[1,*]

[1]Institute of Microbiology of the CAS, v.v.i., Videňská 1083, CZ-14220 Prague 4, Czech Republic
[2]Center for Biotechnology, Bielefeld University, Universitätsstrasse 27, D-33615 Bielefeld, Germany
[3]Genetics and Microbiology, Faculty of Science, Charles University, Viničná 5, CZ-128 44 Praha 2, Czech Republic
[4]Faculty of Food and Biochemical Technology, University of Chemistry and Technology, Technická 5, CZ-166 28 Prague 6, Czech Republic
*__Corresponding author:__ Institute of Microbiology, v.v.i., Czech Academy of Sciences, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic. Tel: (+420)241062398; Fax: +420241722257; E-mail: patek@biomed.cas.cz

One sentence summary: The subject of this study is regulatory network formed by sigma subunits of RNA polymerase that is the major enzyme which starts gene expression in the bacterium *Rhodococcus erythropolis*.
Editor: Olga Ozoline

Abstract

The aim of this investigation was to discover the promoters that drive expression of the *sig* genes encoding sigma factors of RNA polymerase in *Rhodococcus erythropolis* CCM2595 and classify these promoters according to the sigma factors which control their activity. To analyze the regulation of major sigma factors, which control large regulons that also contain genes expressed under exponential growth and non-stressed conditions, we used the *R. erythropolis* CCM2595 culture, which grew rapidly in minimal medium. The transcriptional start sites (TSSs) of the genes *sigA*, *sigB*, *sigD*, *sigE*, *sigG*, *sigH*, *sigJ*, and *sigK* were detected by primary 5′-end-specific RNA sequencing. The promoters localized upstream of the detected TSSs were defined by their −35 and −10 elements, which were identical or closely similar to these sequences in the related species *Corynebacterium glutamicum* and *Mycobacterium tuberculosis*. Regulation of the promoter activities by different sigma factors was demonstrated by two independent techniques (*in vivo* and *in vitro*). All analyzed *sig* genes encoding the sigma factors with extracytoplasmic function (ECF) were found to be also driven from additional housekeeping promoters. Based on the classification of the *sig* gene promoters, a model of the basic sigma transcriptional regulatory network in *R. erythropolis* was designed.

Keywords: *Rhodococcus erythropolis*, transcriptional regulatory network, sigma factor, *in vitro* transcription, RNA-seq

Introduction

Transcriptional regulation is a major level of bacterial adaptation to changing growth conditions. The respective mechanisms switching the alternative cell programs, which are mainly controlled by specific sigma (σ) factors of RNA polymerase (RNAP) and other transcriptional regulators, enable cells to survive under various adverse conditions. Knowledge of such regulatory mechanisms in *Rhodococcus* species, which form a large group of bacteria potentially applicable in biotechnology, is still rather limited.

Rhodococci are Gram-positive, aerobic, non-sporulating bacteria that belong to the Actinobacteria phylum. The presence of mycolic acids in their cell wall places the genus *Rhodococcus* in a Mycolata group that also includes members of the genera *Corynebacterium*, *Mycobacterium*, *Nocardia*, and *Gordonia* (Cole *et al.* 2005).

Although many studies of various regulatory mechanisms which control gene expression in rhodococci have been published, data on the σ factors which are also involved in cell responses to changing conditions are still quite limited. Analysis of the complete genome sequences of 22 strains of the genus *Rhodococcus* (KEGG Gene Database) revealed that rhodococci encode many alternative σ factors. According to the GenBank database, rhodococci possess a diverse number of various σ factors (from 12 in *Rhodococcus* sp. PBTS 1 to 37—including 3 encoded by a plasmid—in *R. jostii* RHA1). The strain which this study focuses on is the soil isolate *R. erythropolis* CCM2595 which degrades various aromatic compounds (Čejková *et al.* 2005). The genes encoding 19 different σ factors have been annotated in its genome sequence (GenBank: accession number NC_022115). We named these predicted σ factors according to the sequentially closest σ factors of *Mycobacterium tuberculosis*. The highest degree of similarity of the *R. erythropolis* σ factors to those of *M. tuberculosis* was found in a primary σ^A (group 1; Gruber and Gross 2003), a primary-like σ^B (group 2), a σ^F-like sigma factor (group 3) and seven extracytoplasmic function (ECF) σ factors (group 4): σ^D, σ^E, σ^G, σ^H, σ^J, σ^K, and σ^M (Pátek *et al.* 2021).

The regulatory functions of individual σ factors in rhodococci have not yet been thoroughly described. It is assumed that σ^A, the principle σ factor in rhodococci as in many bacteria, is mainly expressed during exponential growth. Expression of the *sigB* gene was found to be increased during the transition growth phase in *R. opacus* B4, and was also induced after heat shock, ethanol stress and in response to specific organic solvents (Kita *et al.* 2009). Thus, σ^B is probably the general stress response σ in *R. opacus* B4. We hypothesize that the major ECF σ factors (σ^D, σ^E, and σ^H) in rhodococci have the same or similar functions as their homologs in *M. tuberculosis* or *Corynebacterium glutamicum*. The σ^H factor in *M. tuberculosis* (Manganelli *et al.* 2004) and *C. glutamicum* (Busche *et al.* 2012) controls general heat shock response. As many as 100 σ^H-

Received: September 24, 2021. **Revised:** January 13, 2022. **Accepted:** February 13, 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of FEMS. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

dependent *C. glutamicum* genes are also involved in thiol homeostasis, protein quality control, ribosome modulation, and DNA repair (Busche *et al.* 2012). The σ^E factor is involved in the response to surface stress and heat shock in both *M. tuberculosis* (Sachdeva *et al.* 2010) and *C. glutamicum* (Park *et al.* 2008).

As for the other *R. erythropolis* CCM2595 σ factors, which are missing in *C. glutamicum*, it is known that σ^G is activated during starvation in *R. jostii* RHA1 (Patrauchan *et al.* 2012), while it was found to play a role in SOS response or macrophage infection in *M. tuberculosis* (Sachdeva *et al.* 2010). Other σ factors from *R. jostii* RHA1, namely σ^{F1} and σ^{F3}, are involved in the response to desiccation stress (LeBlanc *et al.* Mohn 2008), whereas in *M. tuberculosis* σ^F regulates the biosynthesis of the mycobacterial cell envelope and immunopathology phenotype (Sachdeva *et al.* 2010).

In this study, we analyzed the transcriptional relationship between the genes encoding the primary σ^A, a primary-like σ^B and three major σ factors of the ECF type, σ^D, σ^E, and σ^H in *R. erythropolis* CCM2595. We detected and cloned the respective promoters and classified them according to the conserved promoter elements using two independent methods. Transcriptional relationships between the σ factors and the *sig* genes were used to construct the basic σ regulatory network of *Rhodococcus erythropolis*.

Materials and methods

Strains, growth conditions, and oligonucleotides

Escherichia coli TOP10 (Grant *et al.* 1990), used for cloning, and *E. coli* BL21 (DE3) (Studier and Moffatt 1986), used for expression of the genes encoding *Rhodococcus erythropolis* sigma factors, were cultivated in 500-mL flasks containing 60 mL of 2 × YT medium (Green and Sambrook 2012) at 37°C.

Corynebacterium glutamicum ATCC 13032 was used as the host strain for measuring the activities of promoters inserted into the promoter-probe vector pEPR1 (Knoppová *et al.* 2007). *Corynebacterium glutamicum* rpoC-His10 (Holátko *et al.* 2012) was used for the isolation of recombinant RNA polymerase. *R. erythropolis* CCM2595 (Veselý *et al.* 2003) was used as a source of chromosomal DNA, which served as a template for PCR amplification of the *sig* genes and promoters. All strains were cultivated in 500-mL flasks with 60 mL of 2 × YT medium at 30°C.

For selection, antibiotics were supplemented into the media: ampicillin (100 μg mL^{-1}), tetracycline (10 μg mL^{-1}), or kanamycin (30 μg mL^{-1}). All oligonucleotides used are listed in Supplementary Table S1.

DNA manipulations

DNA isolation, DNA digestion by restriction enzymes, DNA ligation, PCR, transformation of *E. coli* and electroporation of the host strain *C. glutamicum* were performed by the standard techniques (Green and Sambrook 2012). Promoter DNA fragments were prepared by annealing synthetized complementary oligonucleotides. This synthetic DNA (Sigma-Aldrich, Germany) generated double-stranded fragments around 70 nt in length, with overhangs ready for ligation with pRLG770 (Ross *et al.* 1990) and pEPR1.

RNA isolation

RNA extraction and purification essentially followed the previously developed protocol for *Corynebacterium diphtheriae* (Wittchen *et al.* 2018). *R. erythropolis* CCM2595 was cultivated in minimal BSM medium with glycerol (Veselý *et al.* 2007) and harvested at $OD_{600} = 1.2$. The culture (1 mL) was centrifuged for 30 s at 20 000 × g, and the pellet was immediately frozen with liquid nitrogen. RNA was isolated with a Quick-RNA Miniprep Plus kit (Zymo Research, USA). Crude RNA was purified with a RNA Clean&Concentrator-5 kit (Zymo Research) and its quality was checked with an Agilent RNA Nano 6000 Kit with an Agilent 2100 Bioanalyzer (Agilent Technologies, Germany). Ribosomal RNA was removed with a Ribo-Zero rRNA Removal Kit (Bacteria) from Illumina (San Diego, USA).

Primary 5′-end cDNA library preparation and RNA-seq

Construction of the primary 5′-end cDNA library and RNA sequencing was in principle done using previously described procedures (Pfeifer-Sancar *et al.* 2013) with some modifications (Wittchen *et al.* 2018). RNA samples were treated with terminator exonuclease (Epicentre Biotechnologies, USA) at 30°C for 60 min and then at 42°C for 30 min. RNA fragments were then transcribed to cDNA with a ThermoScript RT-PCR system (Life Technologies, Germany). A random stem-loop DNA adapter 5′ AGATCGGAAGAGAGACGTGTGCTCTTCCGATCTNNNNNNN 3′ was used for cDNA synthesis which hybridizes via a 3′-NNNNNNN-tail preferentially to the 3′-end of the RNA fragments (Pfeifer-Sancar *et al.* 2013). The reaction was carried out at 16°C for 30 min followed by 50°C for 1 h. The reaction was stopped by heat inactivation, and cDNA was amplified by the commercial Illumina RNA PCR primers with Phusion High-Fidelity DNA Polymerase (New England BioLabs, Germany). The produced cDNA library was purified and size-selected by gel electrophoresis for fragment sizes between 100 and 1000 bp. The samples were sequenced in a single read mode with a 75 nt read length with an Illumina MiSeq.

Detection of transcription start sites and bioinformatics analysis of the promoter regions

Transcriptional start sites (TSSs) were automatically detected by the Transcription Analyses function of ReadXplorer (Hilker *et al.* 2016). Possible TSSs were defined by the values T = 14 (number of reads considered as background) and R = 5 (i.e. at least 5-fold difference in the number of read starts between positions –1 and + 1). The thresholds R = 5 and T = 14 were chosen after manual inspection of the predicted TSSs to avoid the effect of noise ratio, and the resulting list of predicted TSSs was manually checked for false positives. The promoter regions, which were apparent just upstream of the TSSs were aligned (70-nt sequences) and analyzed by the software Improbizer (Ao *et al.* 2004) as described previously (Albersmeier *et al.* 2017).

Promoter activity determination by the *in vivo* two-plasmid system

The activity of the promoters cloned in the promoter-test vector pEPR1 was assayed in the heterologous two-plasmid system (a modified method of Dostálová *et al.* 2017). The *C. glutamicum* ATCC 13032 cells harbored the expression vector pEC-XT99A (Kirchner and Tauch 2003) with the inserted genes from *R. erythropolis* CCM2595 encoding the tested σ factors and the promoter-test vector pEPR1 carrying the analyzed promoter (70 nt) from *R. erythropolis* and the *gfp*uv reporter gene. The cells were disrupted with a FastPrep homogenizer (MP Biomedical) and the intensity of the fluorescence of the cell-free extract was detected with a Spark multimode microplate reader (Tecan, Austria) with excitation at 397 nm and emission at 509 nm. Protein concentration in the ex-

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

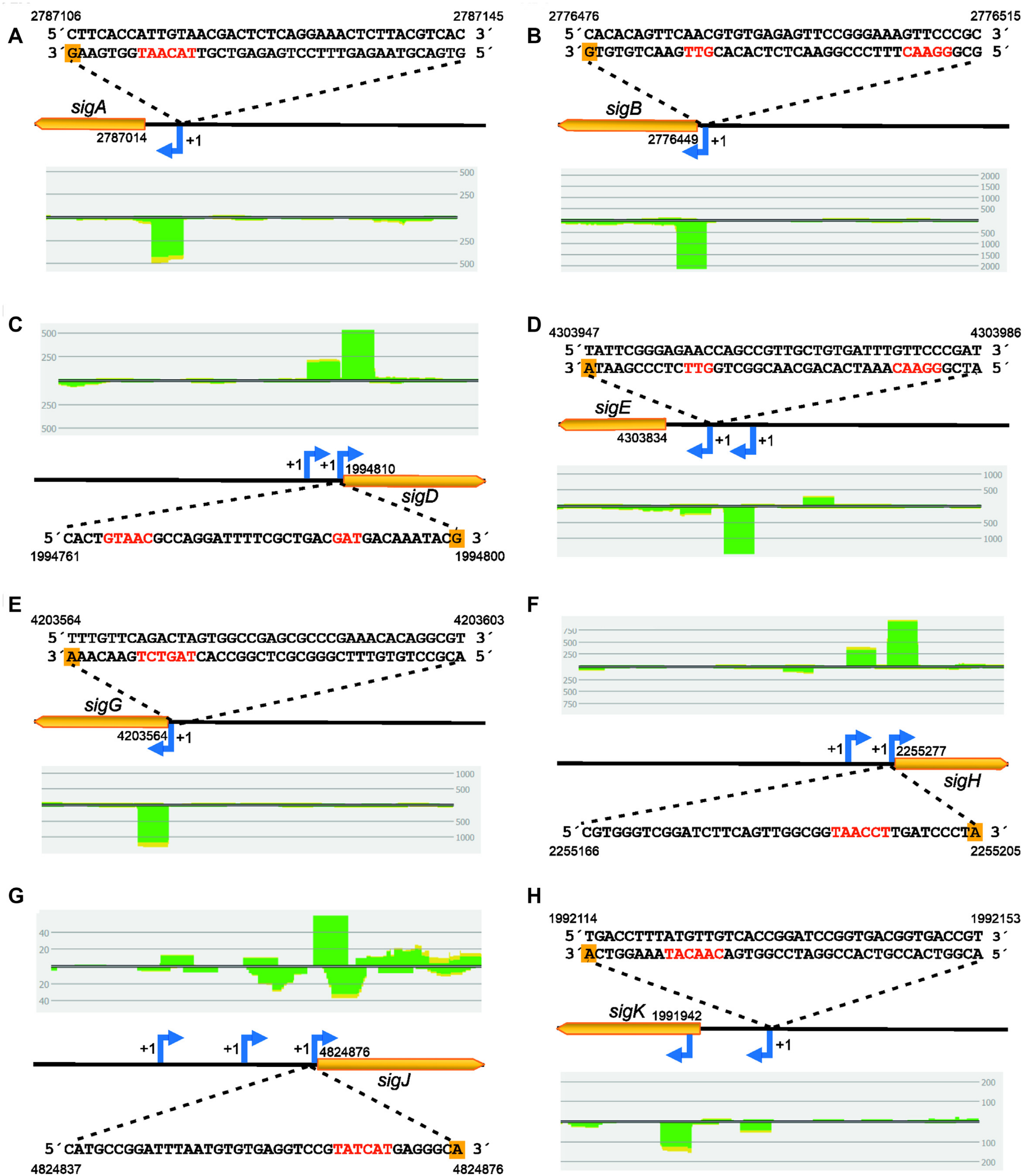


Figure 1. Transcriptional start site (TSS) determination and localization of the corresponding promoters based on the results of 5′-end-specific RNA sequencing. Representative examples of 5′-end RNA-seq data for the detection of TSSs of the *sig* genes are shown. Bent blue arrows indicate the positions of TSSs, based on the RNA-seq signals, which met the defined requirements. The sequences of the promoters closest to the genes are shown. The respective potential –35 and –10 promoter elements (in red; –10 elements only for the housekeeping promoters) were deduced at the appropriate distance from TSSs. Genomic coordinates of the 5′ends and 3′ends of the presented promoter sequences of *R. erythropolis* CCM2595 and positions of the initiation codons according to GenBank RefSeq NC_022115 are shown. The scale for number of reads (y axis) of the detected transcripts is shown at the left or right side and the respective signals are in green.

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

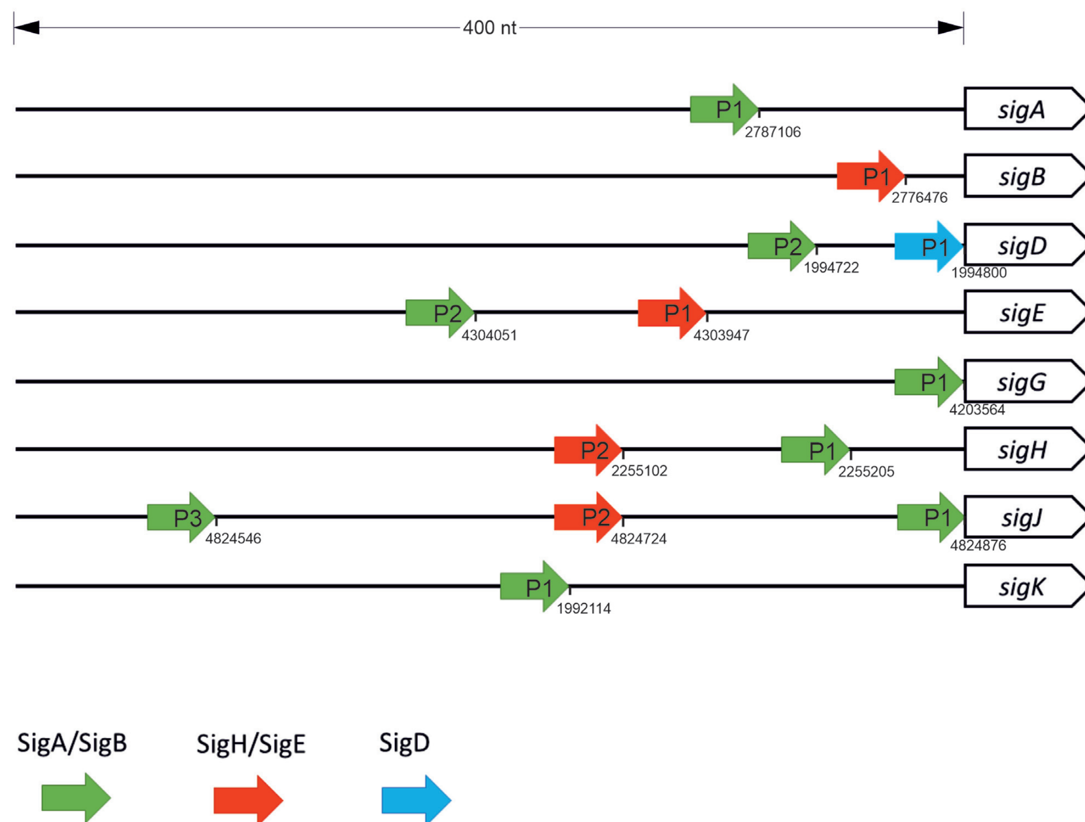


Figure 2. Location of potential promoters of the *sig* genes in *Rhodococcus erythropolis* CCM2595 based on bioinformatic analysis of the region upstream of transcriptional start sites localized precisely by 5′-end-specific RNA-seq. Green arrows represent potential σ^A/σ^B-dependent promoters, the stress promoters of the *sig* genes are shown in red (σ^H/σ^E-dependent) and blue (σ^D-dependent). The determined TSSs are depicted as short vertical lines and their genomic coordinates (GenBank RefSeq NC_022115) are appended below. Coordinates of the translation initiation codons are shown in Fig. 1. 400-nt regions upstream of the initiation codons is marked above the scheme.

tract was determined and promoter activity was expressed in arbitrary units/mg protein (Dostálová *et al.* 2017).

In vitro transcription

The heterologous *in vitro* transcription reactions (RNAP core from *C. glutamicum rpoC*-His10 derived from *C. glutamicum* ATCC 13032 and corresponding σ-subunit from *R. erythropolis* CCM2595) were carried out in principle in the same way as the reaction that was originally designed for homologous *in vitro* transcription (Holátko *et al.* 2012). The reliability of the heterologous arrangement was confirmed recently (Blumenstein *et al.* 2022). The promoter DNA region (70-nt fragments) in the vector pRLG770 served as a template for PCR. The PCR fragments (350–400 bp) were amplified with Q5 polymerase (New England Biolabs) with the primers 30F and CM3, purified by phenol extraction and concentrated with Amicon Ultra 0.5 mL s NMWL 30000 (Merck, Germany). The RNAP holoenzyme was produced by mixing the RNAP core (100 nM) and sigma factors σ^A, σ^B, σ^D, σ^E, or σ^H isolated as His-tagged proteins from *E. coli* in a molar ratio of 1:30. The transcription reaction with [α-^{32}P]UTP and unlabeled ATP, CTP, GTP (in final concentration 0.5 mM each NTP) was run for 10 min at 37°C. The transcripts labeled with [α-^{32}P]UTP were subjected to polyacrylamide gel electrophoresis, and RNA signals were quantified by phosphorimager analysis. The signals were scanned with a Typhoon Scanner (GE Healthcare, USA) and analyzed with ImageQuant TL software. *In vitro* transcription was performed twice with highly similar results.

Results

Analysis of potential *sig* genes in the *R. erythropolis* CCM2595 genome sequence

The complete nucleotide sequence of the *R. erythropolis* CCM2595 genome was previously determined (Strnad *et al.* 2014) and deposited in DDBJ/EMBL/GenBank under the accession numbers CP003761 (chromosome) and CP003762 (plasmid pRECF1). Analysis of the annotated genome *R. erythropolis* CCM2595 sequence (GenBank RefSeq NC_022115) revealed 19 proteins deduced from the nucleotide sequences of the genes potentially encoding a σ subunit of RNAP. These proteins are at least 30% identical to various σ factors from related *M. tuberculosis*. The σ factors σ^A, σ^B, σ^D, σ^E, σ^F, σ^G, σ^H, σ^J, and σ^K were found to exhibit the highest sequence similarity to the corresponding *M. tuberculosis* σ factors (at least 54% identity).

Transcriptional start sites and promoters of the *sig* genes

R. erythropolis cells were cultivated in a minimal medium with glycerol that is readily utilized by this strain. RNA was isolated and purified, and a primary 5′-end cDNA library was sequenced. We focused on the determination of TSSs of all the *sig* genes, which were transcribed under the conditions used. Potential TSSs of the *sig* genes in the *R. erythropolis* CCM2595 genome were mapped by analysis with the software ReadXplorer (Fig. 1). First, automatically detected TSSs were found

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

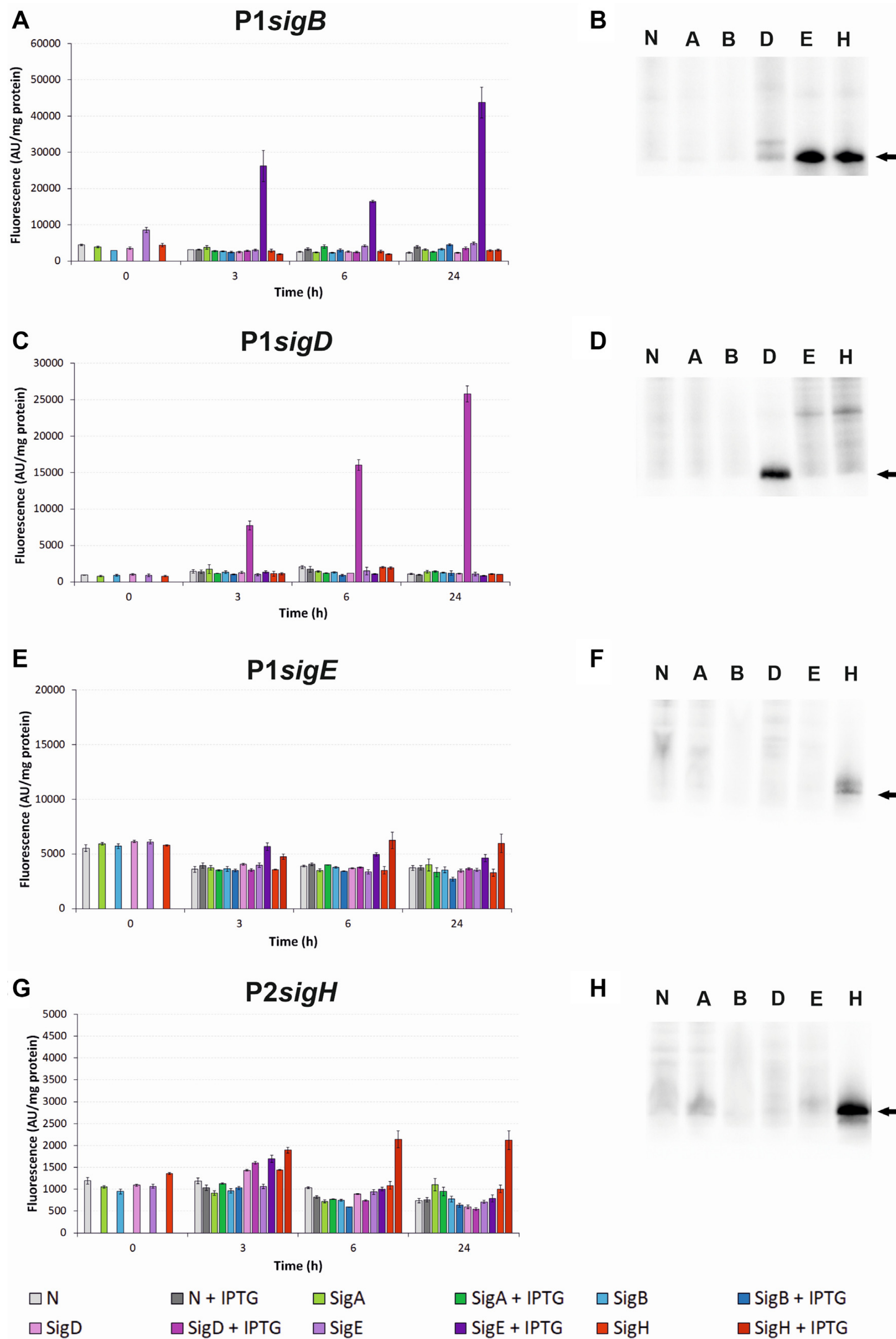


Figure 3. Assignment of rhodococcal sigma factors to analyzed promoters of *Rhodococcus erythropolis* CCM2595 and heterologous *in vitro* transcription with recombinant *C. glutamicum* RNA polymerase core and sigma factors from *R. erythropolis* CCM2595. For the *in vivo* fluorescence assay **(A, C, E, G)**, the *C. glutamicum* ATCC13032 cells harbored the vector pEC-XT99A carrying corresponding *sig* genes from *R. erythropolis* and the promoter-probe vector pEPR1 carrying the *gfpuv* reporter gene expressed from the tested *sig* promoter. The expression of all *sig* genes was induced with IPTG at time point 0; grey bars show the values for control cells without the cloned *sig* genes; Promoter activity was measured as Gfpuv fluorescence intensity of cell extracts and is shown as bars in colors representing respective sigma factors. AU/(mg protein), arbitrary units normalized to protein concentration; the standard deviations of three biological replicates are depicted with error bars. For the *in vitro* transcription **(B, D, F, H)**, the lanes with no sigma (N), σ^A (A), σ^B(B), σ^D (D), and σ^H (H) are shown; arrows indicate the specific transcripts.

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

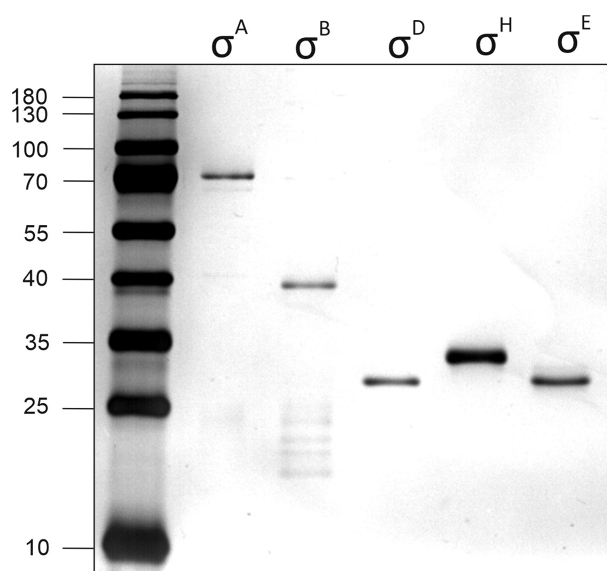


Figure 4. Purified *Rhodococcus erythropolis* CCM2595 sigma subunits of RNA polymerase (silver-stained SDS-polyacrylamide gel). His-tagged sigma factors SigA, SigB, SigD, SigE, and SigH were isolated using an *E. coli* pET system expressing the corresponding *R. erythropolis sig* genes. The protein size marker is on the left; molecular sizes are indicated in kilodaltons.

and in another round, further TSSs were manually assigned. From one to three TSSs were defined upstream of the analyzed *sig* genes encoding σ^A (O5Y_RS12855), σ^B (O5Y_RS12800), σ^D (O5Y_RS09160), σ^E (O5Y_RS19300), σ^G (O5Y_RS18880), σ^H (O5Y_RS10370), σ^J (O5Y_RS21805) and σ^K (O5Y_RS09140) based on the results of the RNA-seq.

The TSSs mapped by the sequencing of the primary 5′-end-specific cDNA library were used to localize the upstream promoter regions. The promoter regions upstream of the TSSs were aligned (70-nt sequences) and analyzed by the software Improbizer (Ao *et al.* 2004) as described by Albersmeier *et al.* 2017. Bioinformatics analysis showed that the majority of the potential promoters are most likely σ^A-dependent (Fig. 2). Alternatively, these genes could be σ^B-dependent (or σ^A/σ^B), since it is very difficult to differentiate between σ^A and σ^B-specific genes in many bacteria (Typas *et al.* 2007, Dostálová *et al.* 2017). The housekeeping promoters were always recognized by both σ^A and σ^B in *in vitro* transcription assays that we conducted (Šilar *et al.* 2016). This phenomenon was also reported in *E. coli*. Upstream of all analyzed *sig* genes, with the exception of *sigB*, one or two σ^A/σ^B-dependent promoters were detected (Fig. 2). The analyzed sequences of the respective promoters are shown in Supplementary Table S2.

Upstream of the *sigG* and *sigK*, housekeeping promoters were only detected and identified according to the sequences of the –10 elements (TAGTCT and CAACAT, respectively) at an appropriate distance from TSSs. Interestingly, the TSS1 of *sigG* (–10 element TAGTCT) and TSS1 of *sigJ* (TATCAT) were found at the nt A belonging in both cases to the ATG initiation codon. This implies that the genes *sigG* and *sigJ* are leaderless. Leaderless transcripts were also found in 22% of *C. glutamicum* housekeeping genes (Albersmeier *et al.* 2017). Similarly, nearly 25% of mycobacterial transcripts are also leaderless (Shell *et al.* 2015).

Upstream of the other 11 potential *sig* genes, no TSSs were detected under the conditions used for *R. erythropolis* cultivation and RNA-seq, which suggests that these genes are only expressed under specific conditions. We concentrated our analysis on the putative stress promoters recognized by ECF sigma factors.

Assignment of *R. erythropolis* sigma factors to promoters by *in vivo* two–plasmid system

The DNA fragments covering the positions +5 to –65 relative to the chosen TSSs, which carried the assumed promoters, were cloned in the promoter-test vector pEPR1. These fragments included the promoters P1*sigB*, P1*sigD*, P1*sigE* and P2*sigH* (Fig. 2; Supplemetary Table S3). The *R. erythropolis* genes encoding σ^A, σ^B, σ^D, σ^E and σ^H cloned in the expression vector pEC-XT99A were used for the construction of the two-plasmid *C. glutamicum* clones. The developed *in vivo* two-plasmid system (Dostálová *et al.* 2017) was applied to define the class of the promoters (their activity with a specific σ factor). In this particular analysis, a heterologous *Rhodococcus/Corynebacterium* two-plasmid system (Blumenstein *et al.* 2022) was used as a variation of this assay. The activity of the promoters with a specific *R. erythropolis* σ factor was measured using the fluorescence intensity of the Gfpuv reporter protein (Fig. 3, left part). The promoter activity assay showed that P1*sigB* is strongly σ^E-dependent (Fig. 3A), P1*sigD* is strictly σ^D-dependent (Fig. 3C), while the other two promoters belong to the σ^H/σ^E-dependent group (Fig. 3E and G).

Determination of sigma factors required for the initiation of transcription from individual promoters by *in vitro* transcription assay

To verify the results of the two-plasmid assay by a different technique, an *in vitro* transcription system (Holátko *et al.* 2012) was used. The *R. erythropolis* promoters P1*sigB*, P1*sigD*, P1*sigE* and P2*sigH* were assayed for their ability to initiate transcription *in vitro* in a heterologous *Corynebacterium/Rhodococcus* system (RNAP core from *C. glutamicum* and σ factors from *R. erythropolis*) (Blumenstein *et al.* 2022) (Fig. 3, right part). The *R. erythropolis* σ factors σ^A, σ^B, σ^D, σ^E and σ^H (Fig. 4) were isolated from the corresponding *Escherichia coli* pET systems overexpressing the *R. erythropolis sigA*, *sigB*, *sigD*, *sigE* and *sigH* genes in a similar way to the *C. glutamicum* sigma factors in the previous studies (Holátko *et al.* 2012, Dostálová *et al.* 2019).

Strong signals representing the specific transcripts were obtained for the promoter P1*sigB* with σ^E and σ^H (Fig. 3B), for P1*sigD* with σ^D (Fig. 3D), or for P2*sigH* with σ^H (Fig. 3H). A weaker signal was detected for P1*sigE* with σ^H (Fig. 3F), whereas no specific signals were observed with σ^A or σ^B or without the addition of any σ factor as a negative control sample.

Thus, the *in vivo* and *in vitro* assays gave the same results for the promoters for P1*sigD* and P2*sigH*, whereas the assays came out differently for P1*sigB* and P1*sigE*. The P1*sigB* promoter was only active with σ^E *in vivo*, but gave strong signals with both σ^E and σ^H *in vitro*. P1*sigE* initiated transcription with both σ^H and σ^E *in vivo*, whereas a signal was only visible with σ^H *in vitro*.

Sequence comparison of promoter regions of *sigB*, *sigD*, *sigE*, and *sigH* genes in *Rhodococcus* strains

To identify the conserved potential promoter sequences of the *sig* genes from *R. erythropolis* CCM2595, we performed a comparative analysis of the upstream regions of TSSs of the genes encoding sigma factors σ^B, σ^D, σ^E, and σ^H using various *Rhodococcus* strains. Nucleotide sequences of representatives of the genus *Rhodococcus*

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

P1*sigB* promoter

Strain	Sequence
R. erythropolis CCM2595	GGCGCGGGAACTTTCCCGGAACTCTCACACGTTGAACTGTGTG
R. aetherivorans IcdP1	ACCCCGGGAACTTTCCCAGGGGTCCTGCGCGTTGAACGCATCG
R. equi 103S	GTCTCGGGAACTTTCCCGGACCGTCGAAGCGTTGAGGGATATG
R. fascians D188	GGCGTGGGAACTTATCGGAGTGTCCCGAGCGTTGAACCTGTGA
R. jostii RHA1	CGGCCGGGAACTTTCCCAGGCCTTCGGTGCGTTGATGCATATG
R. opacus B4	CGGCCGGGAACTTTCCCAGGCCTTCGGTGCGTTGATCCATATG
R. rhodochrous ATCC BAA870	CGGACGGGAACTCCACCGCTTCGCCGGT-CGTTGAACCCTGTG
R. ruber P14	ACCTCGGGAACTTTCCCAGGGGCCTGGCGCGTTGAACGCTTCG

P1*sigD* promoter

Strain	Sequence
R. erythropolis CCM2595	TGCCACTGTAACGCCAGGATTTTCGCTGACGATGACAAATACG
R. aetherivorans IcdP1	TCACTTTGTAACGCCAGGACTTGCGCAATCGATGACTGTCATG
R. equi 103S	TGCCACTGTAACGCCAGGATCCGCTCTGACGATGTCAAACACG
R. fascians D188	CGTACCTGTAACGCGGGGAAATTTGTCGACGATGAACAGTGCG
R. jostii RHA1	TGCCACTGTAACGCCAGGATATTCAGCAACGATGACTAATACG
R. opacus B4	TGCCACTGTAACGCCAGGATATTCAGCAACGATGACTAACACG
R. rhodochrous ATCC BAA870	GCCGTCTGTAACGCCAGGACTTGCGCAATCGATGACTGTTATG
R. ruber P14	TCACTTTGTAACGCCAGGACTTGCGCAATCGATGACTGTCATG

P1*sigE* promoter

Strain	Sequence
R. erythropolis CCM2595	GTGATCGGGAACAAATCACAGCAACGGCTGGTTCTCCCGAATA
R. aetherivorans IcdP1	GGTACCGGGAACAAGCCGGCGCCCCCGGGTGTTGGACGAAGAC
R. equi 103S	ATCGTCGGGAACAATTTCTCGTTCTCGGGAGTTGAACCGGCAG
R. fascians D188	GTGAGCTGGAACATCCGAGGGTAGTCGTCGGTTGCATCAGAAG
R. jostii RHA1	GAAACGGGGAACAAATCCCGGCCCTCGGGAGTTGACGCCAATA
R. opacus B4	GAGTCCGGGAACAAATCCCGGCACCCGGGAGTTGACGCCTATA
R. rhodochrous ATCC BAA870	CTGTGAAGGAATATTCCCTGAAGGTCGCGTGTTGTAGGCCCCG
R. ruber P14	GGGCCCGGGAACAAACCCGGCGTACACGGTGTTGGACCACGAC

P2*sigH* promoter

Strain	Sequence
R. erythropolis CCM2595	AGGCCGGGAACAGAGTCGCAGACTCCCGTGTTGGTACCAGTTG
R. aetherivorans IcdP1	ATGTAGGGAACAAGGGCCGGACCTGTCGTGTTGATTCCGGGTG
R. equi 103S	GGTCCGGGAACAGACTCTCGCCCCGTGCTGTTGATACCTCGCG
R. fascians D188	GCGGCGGGAACACCGCGGCGCGCGGAAGTGTTGTGACAGGGCA
R. jostii RHA1	CGGTCGGGAACACAGCGGACGGGAGGCGTGTTGGTACCTGCAG
R. opacus B4	GGTGCGGGAACACCGCGGGCGGGAGTCGTGTTGGTACCTGCAG
R. rhodochrous ATCC BAA870	AGGTGGGGAACAAGTGCGGCGGGTACGGCGTTGCACACCGACG
R. ruber P14	ATGTAGGGAACAAGCACCCACCCTGTCGTGTTGATTCCGGGTG

Figure 5. Promoter sequences of rhodococcal genes *sigB*, *sigD*, *sigE*, and *sigH*. Transcriptional start sites of *R. erythropolis* CCM2595 genes determined by RNA-seq and the key promoter sequence elements -10 and -35 are underlined. Identical nucleotides in all sequences near the main motifs in the promoters of particular genes are highlighted in yellow. The only differing nt in the -35 element is highlighted in cyan. The promoters P1*sigB*, P1*sigE*, and P2*sigH* are proposed to be SigH/E-dependent, the P1*sigD* promoter is SigD-dependent.

were obtained from the genome RefSeq of *R. aetherivorans* IcdP1 (GenBank Acc. Number NZ_CM002177), *R. equi* 103S (GenBank Acc. Number NC_014659), *R. fascians* D188 (GenBank Acc. Number NZ_CP015235), *R. jostii* RHA1 (GenBank Acc. Number NC_008268), *R. opacus* B4 (GenBank Acc. Number NC_012522), *R. rhodochrous* ATCC BAA870 (GenBank Acc. Number NZ_CP032675) and *R. ruber* P14 (GenBank Acc. Number NZ_CP024315). It was found that both corresponding promoter regions –35 and –10 in the analyzed ECF-dependent promoters are identical in all rhodococcal strains of interest (with one exception, a C→T substitution in the –35 promoter element of *sigE* of *R. rhodochrous* ATCC BAA870; Fig. 5).

Taken together, the results indicate that σ^H and σ^E recognized the promoters with the sequence motifs –35 GGAAC and –10 GTT, whereas σ^D recognized –35 GTAAC and –10 GAT. This knowledge can be used for further analysis of promoters and stress regulons in *Rhodococcus erythropolis* and other *Rhodococcus* species.

Discussion

The prediction of promoter sequences without experimental proofs is notoriously difficult. Although many bioinformatics tools and algorithms have been developed to localize the promoters of both housekeeping and stress genes according to the conserved sequences (Jacques *et al.* 2006), the deduced promoters mostly remain putative, and the number of false-positive promoter sequences is usually uncertain. Finding the association of the suggested promoters with σ factors (classification of the promoter) is another tricky task, which again frequently lacks experimental proof. We combined RNA-seq to precisely localize TSSs, which indicate the positions of the promoters, with two methods which assign the σ factor to a particular promoter. Since the genes encoding ECF σ factors are in many bacteria preceded by the corresponding promoters (i.e. these genes are autoregulated; Staroń *et al.* 2009) we focused on the TSSs upstream of a few major *R. ery-*

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

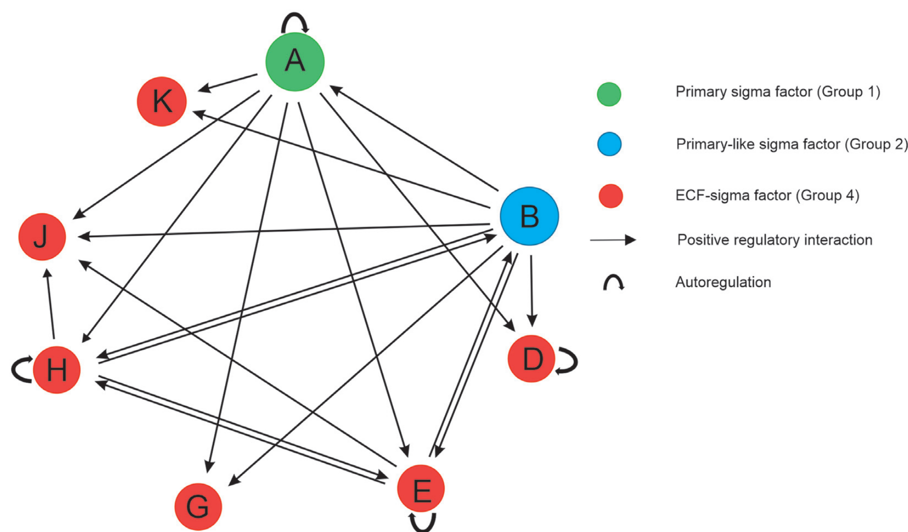


Figure 6. Scheme of partial transcriptional regulatory network in *Rhodococcus erythropolis* CCM2595. The model is based on bioinformatics analysis of the *R. erythropolis* genome sequence and our experimental results. Arrows indicate a positive regulatory interaction (σ factor involved in transcription of other *sig* gene). Since the conditions in the used *in vivo* and *in vitro* assays are not the same as in standard cells and ratios of the influence of the pairs SigA/SigB and SigH/SigE on the expression of the *sig* genes under various conditions cannot be determined, we propose this model as a simplified hypothesis.

thropolis sig genes and thus localized their promoters. Using the *in vivo* two-plasmid assay and *in vitro* transcription enabled us to associate the promoter sequences with particular σ factors.

The –35 and –10 promoter sequences recognized by sigma factors σ^A, σ^B, σ^D, σ^E, and σ^H of rhodococci show identity or a high similarity to the promoter consensus sequences recognized by the homologous σ factors of another member of the Mycolata group, *C. glutamicum* (Dostálová *et al.* 2019) and/or *M. tuberculosis* (Rodrigue *et al.* 2006). The promoter −35 and −10 motifs in the appropriate distance upstream of the *R. erythropolis* CCM2595 *sigD* TSS (GTAACG-N_{17}-GAT) conform to the *C. glutamicum* consensus of σ^D-dependent promoters GTAAC$^A/_G$-N_{17}-GAT (Toyoda and Inui 2018, Dostálová *et al.* 2019). The -35 and -10 elements upstream of the *R. erythropolis* CCM2595 *sigE* and *sigH* genes are identical (GGAAC-N_{18}-GTT) and nearly precisely fit the *C. glutamicum* and *M. tuberculosis* consensus sequences of σ^H-dependent promoters. We have previously identified *C. glutamicum* σ^E-dependent promoters P2*dnaK*, P2*dnaJ*2, and P1*clgR* (Šilar *et al.* 2016), and all of them were also found to be σ^H-dependent. The same is true for the *C. glutamicum* and *M. tuberculosis* P*sigB* that were both σ^H/σ^E-dependent with the identical sequence GGAAC-N_{18}-GTT (Raman *et al.* 2001, Halgašová *et al.* 2002, Dostálová *et al.* 2019). The general consensus sequences of *M. tuberculosis* σ^E- and σ^H-dependent promoters are also GGAAC-$N_{17\text{-}18}$-GTT (Rodrigue *et al.* 2006). It therefore seems that SigE- and SigH-dependent promoters are generally highly similar, and only subtle differences outside the main motifs decide whether transcription with one of the other sigmas is suppressed or not clearly apparent. In fact, the most frequent sequence of σ^H-specific promoters in *C. glutamicum* is GGAA**T**-N_{18}-GTT, although some σ^H-dependent promoters have C or G or A in the last −35 element position (Ehira *et al.* 2009, Busche *et al.* 2012). We were unable to differentiate between the key motifs of the promoters recognized by *R. erythropolis* σ^E and σ^H, because in both cases the -35 sequences were GGAAC. This trend could probably be found in many *Rhodococcus* species, since we found that in all *sigB, sigD, sigE* and *sigH* genes, the -35 promoter element GGAAC was present in all (with a single exception) of the 8 *Rhodococcus* strains analyzed (Fig. 5). The P2*sigJ* promoter, which was apparently σ^H/σ^E-dependent according the key sequence elements, was found to be weak by RNA-seq (Fig. 1). Therefore, we did not include P2*sigJ* into *in vivo* and *in vitro* assays and defined it as σ^H/σ^E-dependent. We suppose that P2*sigJ* may be more active under some stress conditions. In contrast, the housekeeping P1*sigJ* was much stronger under the conditions (optimum growth) used.

The σ^J factor is involved in the survival of *M. tuberculosis* under oxidative stress, and in the virulence of the pathogen. The function of σ^J in rhodococci is not known.

The results of the *in vivo* two-plasmid assay and *in vitro* transcription figures were not always identical (Fig. 3). However, we suppose that the tested promoters are recognized by both σ factors, but with different efficiency under different conditions. Further analysis of the σ^E and σ^H regulons in rhodococci may shed more light on the recognition of the promoter classes and functions of the two σ factors in stress response.

The vegetative P1*sigK* seems to be a weak promoter, but the *sigK* gene may be additionally transcribed from another promoter that is active under stress conditions. We also detected a signal corresponding to TSS inside the *sigK* gene (Fig. 1). This may belong to an alternative transcript and a shorter version of σ^K or a regulatory RNA and was not further studied here. Transcription of the *sig* genes encoding ECF σ factors, which are mostly involved in stress responses, from σ^A- and/or σ^B-dependent promoters was also found in *C. glutamicum* (Busche *et al.* 2012, Pfeifer-Sancar *et al.* 2013, Dostálová *et al.* 2019). The *R. erythropolis* CCM2595 genes *sigD*, *sigE*, *sigH* and *sigJ* were also found to be transcribed from additional vegetative promoters (Fig. 2). Such combined transcription can be found for many stress genes in bacteria (Pátek *et al.* 2013, Cho *et al.* 2014, Dostálová *et al.* 2019). It is thought that transcription from σ^A/σ^B-dependent genes ensures a basal level of expression during exponential and slow growth without the influence of particular stresses.

The key promoter hexamers of the supposed σ^A/σ^B-dependent promoters are shown in Supplementary Table S2. We compared two unusual −10 promoter sequences AAGGCT (P1*sigJ*) and AAAGGT (P2*sigE*) with analogous sequences found in *C. glutamicum* (Albersmeier *et al.* 2017). Identical −10 AAGGCT hexamer

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

found in R. erythropolis was twice detected in C. glutamicum. C. glutamicum vegetative promoters also possess closely similar −10 hexamers AAAGAT, AAAAGT, AAGACT, AAGGAT. Since the –35 regions of vegetative promoters are very variable in Rhodococcus (similar to promoters in Corynebacterium) their position and sequences cannot be easily identified.

Precise localization of the promoters of the sig genes and their classification enabled us to construct the model of the hypothetical basic sigma transcription regulatory network (Fig. 6). This basic scheme shows links between the σ factors, i.e. which σ proteins control the transcription of individual sig genes. Whereas sigD, sigE, and sigH were found to be autoregulated, sigG, sigJ and sigK were expressed from σ^A/σ^B- or σ^H/σ^E-dependent promoters. Moreover, all the analyzed sig genes (with the exception of sigB) were transcribed from at least one σ^A/σ^B-dependent promoter. This may explain why these genes were expressed during exponential growth under non-stressed conditions. The respective σ factors thus play an important role in the standard growth and metabolism of R. erythropolis CCM2595 cells. The expression of the genes encoding ECF σ factors sigC, sigD and sigH from only vegetative promoters was observed in C. glutamicum (Busche et al. 2012, Toyoda and Inui 2016, 2018).

We are currently working on revealing the expression of stress genes under various conditions and regulation of stress response by ECF sigma factors in Rhodococcus strains. The recognized promoter sequences will help us to define the sig regulons.

Supplementary Data

Supplementary data are available at FEMSLE online.

Acknowledgments

We wish to thank Jörn Kalinowski (Bielefeld) for the possibility to perform RNA sequencing in the Center for Biotechnology, Bielefeld, using their facilities and assistance.

Funding

This work was supported by grants [18-13254S] from the Czech Science Foundation, GAUK [338321] from Charles University (JB) and Institutional Research Project [RVO61388971] from the Institute of Microbiology of the CAS.

Conflicts of interest statement. None declared.

References

Albersmeier A, Pfeifer-Sancar K, Ruckert C et al. Genome-wide determination of transcription start sites reveals new insights into promoter structures in the actinomycete Corynebacterium glutamicum. J Biotechnol 2017;**257**:99–109.

Ao W, Gaudet J, Kent WJ et al. Environmentally induced foregut remodelling by PHA-4/FoxA and DAF-12/NHR. Science 2004;**305**:1743–6.

Blumenstein J, Rädisch R, Štěpánek V et al. Identification of Rhodococcus erythropolis promoters controlled by alternative sigma factors using in vivo and in vitro systems and heterologous RNA polymerase. Curr Microbiol 2022;**79**:55

Busche T, Šilar R, Pičmanová M et al. 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 2012;**13**:445.

Čejková A, Masák J, Jirků V et al. Potential of Rhodococcus erythropolis as a bioremediation organism. World J Microbiol Biotechnol 2005;**21**:317–21.

Cho BK, Kim D, Knight EM et al. Genome-scale reconstruction of the sigma factor network in Escherichia coli: topology and functional states. BMC Biol 2014;**12**:4.

Cole JR, Chai B, Farris RJ et al. The ribosomal database project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res 2005;**33**:D294–6.

Dostálová H, Busche T, Holátko J et al. Overlap of promoter recognition specificity of stress response sigma factors SigD and SigH in Corynebacterium glutamicum ATCC 13032. Front Microbiol 2019;**9**:3287.

Dostálová H, Holátko J, Busche T et al. Assignment of sigma factors of RNA polymerase to promoters in Corynebacterium glutamicum. AMB Express 2017;**7**:133.

Ehira S, Teramoto H, Inui M et al. Regulation of corynebacterium glutamicum heat shock response by the extracytoplasmic-function sigma factor SigH and transcriptional regulators HspR and HrcA. J Bacteriol 2009;**191**:2964–72.

Grant SGN, Jessee J, Bloom FR et al. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proc Natl Acad Sci 1990;**87**:4645–9.

Green MR, Sambrook J. Molecular cloning: A laboratory manual, Fourth Edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2012.

Gruber TM, Gross CA. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol 2003;**57**: 441–66.

Halgašová N, Bukovská G, Ugorčáková J et al. The Brevibacterium flavum sigma factor SigB has a role in the environmental stress response. FEMS Microbiol Lett 2002;**216**:77–84.

Hilker R, Stadermann KB, Schwengers O et al. ReadXplorer 2-detailed read mapping analysis and visualization from one single source. Bioinformatics 2016;**32**:3702–8.

Holátko J, Šilar R, Rabatinová A et al. Construction of in vitro transcription system for Corynebacterium glutamicum and its use in the recognition of promoters of different classes. Appl Microbiol Biotechnol 2012;**96**:521–9.

Jacques PE, Rodrigue S, Gaudreau L et al. Detection of prokaryotic promoters from the genomic distribution of hexanucleotide pairs. BMC Bioinf 2006;**7**. DOI: 10.1186/1471-2105-7-423.

Kirchner O, Tauch A. Tools for genetic engineering in the amino acid-producing bacterium Corynebacterium glutamicum. J Biotechnol 2003;**104**:287–99.

Kita A, Takiguchi N, Kato J. Cloning and characterization of sigA and sigB genes from Rhodococcus opacus strain B4: involvement of sigB in organic solvent tolerance. J Environ Biotechnol 2009;**9**:43–50.

Knoppová M, Phensaijai M, Veselý M et al. Plasmid vectors for testing in vivo promoter activities in Corynebacterium glutamicum and Rhodococcus erythropolis. Curr Microbiol 2007;**55**:234–9.

LeBlanc JC, Gonçalves ER, Mohn WW. Global response to desiccation stress in the soil actinomycete Rhodococcus jostii RHA1. Appl Environ Microbiol 2008;**74**:2627–36.

Manganelli R, Provvedi R, Rodrigue S et al. Sigma factors and global gene regulation in Mycobacterium tuberculosis. J Bacteriol 2004;**186**:895–902.

Park SD, Youn JW, Kim YJ et al. Corynebacterium glutamicum sigmaE is involved in responses to cell surface stresses and its activity is controlled by the anti-sigma factor CseE. Microbiology 2008;**154**:915–23.

Pátek M, Grulich M, Nešvera J. Stress response in Rhodococcus strains. Biotechnol Adv 2021;**53**:107698.

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

Pátek M, Holátko J, Busche T *et al.* *Corynebacterium glutamicum* promoters: a practical approach. *Microb Biotechnol* 2013;**6**:103–17.

Patrauchan MA, Miyazawa D, LeBlanc JC *et al.* Proteomic analysis of survival of *Rhodococcus jostii* RHA1 during carbon starvation. *Appl Environ Microbiol* 2012;**78**:6714–25.

Pfeifer-Sancar K, Mentz A, Ruckert C *et al.* Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique. *BMC Genomics* 2013;**14**. DOI: 10.1186/1471-2164-14-888.

Raman S, Song TS, Puyang XL *et al.* The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. *J Bacteriol* 2001;**183**: 6119–25.

Rodrigue S, Provvedi R, Jacques PE *et al.* The sigma factors of *Mycobacterium tuberculosis*. *FEMS Microbiol Rev* 2006;**30**:926–41.

Ross W, Thompson JF, Newlands JT *et al.* *Escherichia coli* fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J* 1990;**9**:3733–42.

Sachdeva P, Misra R, Tyagi AK *et al.* The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *FEBS J* 2010;**277**:605–26.

Shell SS, Wang J, Lapierre P *et al.* Leaderless transcripts and small proteins are common features of the mycobacterial translational landscape. *PLos Genet* 2015;**11**. DOI: 10.1371/journal.pgen.1005641.

Šilar R, Holátko J, Rucká L *et al.* Use of *in vitro* transcription system for analysis of *corynebacterium glutamicum* promoters recognized by two sigma factors. *Curr Microbiol* 2016;**73**:401–8.

Staroń A, Sofia HJ, Dietrich S *et al.* The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol Microbiol* 2009;**74**: 557–81.

Strnad H, Pátek M, Fousek J *et al.* Genome sequence of *Rhodococcus erythropolis* strain CCM2595, a phenol derivative-degrading bacterium. *Genome Announc* 2014;**2**. DOI: 10.1128/genomeA.00208-14.

Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 1986;**189**:113–30.

Toyoda K, Inui M. The extracytoplasmic function σ factor σ^{C} regulates expression of a branched quinol oxidation pathway in *Corynebacterium glutamicum*. *Mol Microbiol* 2016;**100**:486–509.

Toyoda K, Inui M. Extracytoplasmic function sigma factor σ^{D} confers resistance to environmental stress by enhancing mycolate synthesis and modifying peptidoglycan structures in *Corynebacterium glutamicum*. *Mol Microbiol* 2018;**107**: 312–29.

Typas A, Becker G, Hengge R. The molecular basis of selective promoter activation by the σ^{S} subunit of RNA polymerase. *Mol Microbiol* 2007;**63**:1296–306.

Veselý M, Knoppová M, Nešvera J *et al.* Analysis of *catRABC* operon for catechol degradation from phenol-degrading *Rhodococcus erythropolis*. *Appl Microbiol Biotechnol* 2007;**76**:159–68.

Veselý M, Pátek M, Nešvera J *et al.* Host-vector system for phenol-degrading *Rhodococcus erythropolis* based on *Corynebacterium* plasmids. *Appl Microbiol Biotechnol* 2003;**61**:523–7.

Wittchen M, Busche T, Gaspar AH *et al.* 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* 2018;**19**. DOI: 10.1186/s12864-018-4481-8.

Downloaded from https://academic.oup.com/femsle/article/369/1/fnac014/6528914 by Institute of Microbiology/Academy of Sciences user on 23 February 2022

Supplementary data are available:

Table S1: List of oligonucleotides used.

Table S2: Sequences of the vegetative promoters found upstream of the TSSs determined by RNA-seq.

Table S3: Sequences of the analyzed alternative promoters.

Table S1. List of oligonucleotide primers.

Name	Sequence	Purpose
SIGAREpECF2	CGCCGGAATTCCACTCTGG	Cloning of structural gene *sigA* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGAREpECR	TCTTCTAGAAGCTGGTGTGACTA	Cloning of structural gene *sigA* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGBREpECF	GAGAATTCGCCGATAGGGAGG	Cloning of structural gene *sigB* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGBREpECR	CTGACTCTAGAGCTGTGTTGTTTGA	Cloning of structural gene *sigB* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGDREpECF2	GCCACTGAATTCCCAGGATTTTCG	Cloning of structural gene *sigD* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGDREpECR	CACTCTAGACCGCAGCCAAGTCC	Cloning of structural gene *sigD* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGEREpECF	GCACTACGGAATTCGGACAAC	Cloning of structural gene *sigE* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGEREpECR	TCCCTCTAGATGATTGACCAC	Cloning of structural gene *sigE* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGHREpECF	AGACTCCCGTGAGCTCACCAGTTGAT	Cloning of structural gene *sigH* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGHREpECR	CTGTCTAGAAGTCGAGCCTTTC	Cloning of structural gene *sigH* from *R. erythropolis* CCM2595 in pEC-XT99A
SIGAREpETF	GAAAGGGCGCATATGGCAGCC	Cloning of structural gene *sigA* from *R. erythropolis* CCM2595 in pET-22b
SIGAREpETR	GGTGAAGCTTGTCGAGGTAGTCG	Cloning of structural gene *sigA* from *R. erythropolis* CCM2595 in pET-22b
SIGBREpETF	TAGGGAGGCAACATATGACAAGCC	Cloning of structural gene *sigB* from *R. erythropolis* CCM2595 in pET-22b
SIGBREpETR	GTGTTGTTCTCGAGGCTCGCGTAG	Cloning of structural gene *sigB* from *R. erythropolis* CCM2595 in pET-22b
SIGDREpETF	GATTTTCGCTGCATATGACAAATACGAG	Cloning of structural gene *sigD* from *R. erythropolis* CCM2595 in pET-22b
SIGDREpETR	CCTCGCTCGAGGCCAAAGCTCTCACC	Cloning of structural gene *sigD* from *R. erythropolis* CCM2595 in pET-22b
SIGEREpETF	AGTCGGAACATATGACGACGGAAT	Cloning of structural gene *sigE* from *R. erythropolis* CCM2595 in pET-22b
SIGEREpETR	GAACCCGACTCGAGGAAACCG	Cloning of structural gene *sigE* from *R. erythropolis* CCM2595 in pET-22b
SIGHREpETF	AAGGGATCCATATGCTGGAACACGAC	Cloning of structural gene *sigH* from *R. erythropolis* CCM2595 in pET-22b
SIGHREpETR	TGCTCTCGAGTCGTGTGACACCTTCCGTCTC	Cloning of structural gene *sigH* from *R. erythropolis* CCM2595 in pET-22b
P1*sigB*REpEPRF	GGTCACGGCGTTTTAGCGGAAATCGGCGCGGGAACTTTCCCGGAACTCTCACACGTTGAACTGTGTGAGAGG	Cloning of promoter P1*sigB* from *R. erythropolis* CCM2595 in pEPR1
P1*sigB*REpEPRR	GATCCCTCTCACACAGTTCAACGTGTGAGAGTTCCGGGAAAGTTCCCGCGCCGATTTCCGCTAAAACGCCGTGACCTGCA	Cloning of promoter P1*sigB* from *R. erythropolis* CCM2595 in pEPR1
P1*sigD*REpEPRF	GCGGGAGCACACCCGCTCGGGAGGTGCCACTGTAACGCCAGGATTTTCGCTGACGATGACAAATACGAGCGG	Cloning of promoter P1*sigD* from *R. erythropolis* CCM2595 in pEPR1
P1*sigD*REpEPRR	GATCCCGCTCGTATTTGTCATCGTCAGCGAAAATCCTGGCGTTACAGTGGCACCTCCCGAGCGGGTGTGCTCCCGCTGCA	Cloning of promoter P1*sigD* from *R. erythropolis* CCM2595 in pEPR1
P1*sigE*REpEPRF	GCCAAGTTGGTCTTTCGCAGCTCGGTGATCGGGAACAAATCACAGCAACGGCTGGTTCTCCCGAATATCGG	Cloning of promoter P1*sigE* from *R. erythropolis* CCM2595 in pEPR1
P1*sigE*REpEPRR	GATCCCGATATTCGGGAGAACCAGCCGTTGCTGTGATTTGTTCCCGATCACCGAGCTGCGAAAGACCAACTTGGCTGCA	Cloning of promoter P1*sigE* from *R. erythropolis* CCM2595 in pEPR1

P2*sigH*REpEPRF	GATCGCCGACAGGTACATCCGACCAGGCCGG GAACAGAGTCGCAGACTCCCGTGTTGGTACC AGTTGATGAG	Cloning of promoter P2*sigH* from *R. erythropolis* CCM2595 in pEPR1
P2*sigH*REpEPRR	GATCCTCATCAACTGGTACCAACACGGGAGT CTGCGACTCTGTTCCCGGCCTGGTCGGATGT ACCTGTCGGCGATCTGCA	Cloning of promoter P2*sigH* from *R. erythropolis* CCM2595 in pEPR1
P1*sigB*REp770F	AATCCGGTCACGGCGTTTTAGCGGAAATCGG CGCGGGAACTTTCCCGGAACTCTCACACGTT GAACTGTGTGAGAGA	Cloning of promoter P1*sigB* from *R. erythropolis* CCM2595 in pRLG770
P1*sigB*REp770R	AGCTTCTCTCACACAGTTCAACGTGTGAGAG TTCCGGGAAAGTTCCCGCGCCGATTTCCGCT AAAACGCCGTGACCG	Cloning of promoter P1*sigB* from *R. erythropolis* CCM2595 in pRLG770
P1*sigD*REp770F	AATCCGCGGGAGCACACCCGCTCGGGAGGTG CCACTGTAACGCCAGGATTTTCGCTGACGAT GACAAATACGAGCGA	Cloning of promoter P1*sigD* from *R. erythropolis* CCM2595 in pRLG770
P1*sigD*REp770R	AGCTTCGCTCGTATTTGTCATCGTCAGCGAA AATCCTGGCGTTACAGTGGCACCTCCCGAGC GGGTGTGCTCCCGCG	Cloning of promoter P1*sigD* from *R. erythropolis* CCM2595 in pRLG770
P1*sigE*REp770F	AATTCCCAAGTTGGTCTTTCGCAGCTCGGTG ATCGGGAACAAATCACAGCAACGGCTGGTTC TCCCGAATATCGCGA	Cloning of promoter P1*sigE* from *R. erythropolis* CCM2595 in pRLG770
P1*sigE*REp770R	AGCTTCGCGATATTCGGGAGAACCAGCCGTT GCTGTGATTTGTTCCCGATCACCGAGCTGCG AAAGACCAACTTGGG	Cloning of promoter P1*sigE* from *R. erythropolis* CCM2595 in pRLG770
P2*sigH*REp770F	AATCCGATCGCCGACAGGTACATCCGACCAG GCCGGGAACAGAGTCGCAGACTCCCGTGTTG GTACCAGTTGATGAA	Cloning of promoter P2*sigH* from *R. erythropolis* CCM2595 in pRLG770
P2*sigH*REp770R	AGCTTTCATCAACTGGTACCAACACGGGAGT CTGCGACTCTGTTCCCGGCCTGGTCGGATGT ACCTGTCGGCGATCG	Cloning of promoter P2*sigH* from *R. erythropolis* CCM2595 in pRLG770

Table S2. Sequences of the vegetative promoters found upstream of the TSSs determined by RNA-seq

Gene No.	Gene	50-nt promoter sequence[a]	TSS distance
O5Y_RS12855	*sigA*	ACGGCCCGCCGTGACGTAAGAGTTTCCTGAGAGTCGT**TACAAT**GGTGAA**G**	92
O5Y_RS09160	*sigD*	CTGGGCGGGCATGCGGAGACCAGCCGGTCCGGATGGT**TACTCT**TAGAGG**G**	88
O5Y_RS19300	*sigE*	CTCAGGCTCCTCATATGGCGTGCACACCGCGATGAGA**AAAGGT**GTATAC**A**	217
O5Y_RS18880	*sigG*	ATTCTCGCCGACGCCTGTGTTTCGGGCGCTCGGCCAC**TAGTCT**GAACAA**A**	0
O5Y_RS10370	*sigH*	ACCACCGGCCCGTGGGTCGGATCTTCAGTTGGCGG**TAACCT**TGATCCCT**A**	72
O5Y_RS21805	*sigJ*	CGGTTACATTCATGCCGGATTTAATGTGTGAGGTCCG**TATCAT**GAGGGC**A**	0
O5Y_RS21805	*sigJ*	GACGACTTCGTCGGAGGAGCCCGAGCAGGCAACGGC**AAGGCT**GGACAGA**G**	330
O5Y_RS09140	*sigK*	CGAGGGCGGAACGGTCACCGTCACCGGATCCGGTGA**CAACAT**AAAGGTC**A**	172

[a]The proposed -10 hexamers are in bold, underlined

Table S3. Sequences of the analyzed alternative promoters.

Promoter	Nucleotide sequence[a]	TSS distance
P1*sigB*	GTCACGGCGTTTTAGCGGAAATCGGCGCG**GGAAC**TTTCCCGGAACTCTCACAC**GTT**GAACTGTGT**G**AGAG	27
P1*sigD*	CGGGAGCACACCCGCTCGGGAGGTGCCACT**GTAAC**GCCAGGATTTTCGCTGAC**GAT**GACAAATAC**G**AGCG	10
P1*sigE*	CCAAGTTGGTCTTTCGCAGCTCGGTGATCG**GGAAC**AAATCACAGCAACGGCTG**GTT**CTCCCGAAT**A**TCG	113
P2*sigH*	ATCGCCGACAGGTACATCCGACCAGGCCG**GGAAC**AGAGTCGCAGACTCCCGT**GTT**GGTACCAGTT**G**ATGA	175
P2*sigJ*	GTATGCGGCGCTCGGCACAGGCAGGTGTGG**GGAAT**GTCAGTTCAACGACAACA**GTT**GAGCGACAT**G**AGCG	152

[a] -35 and -10 promoter regions and TSSs are in bold and underlined.