Identification of Two Genes Encoding Putative New Members of the ECF Subfamily of Eubacterial RNA Polymerase Sigma Factors in *Clostridium acetobutylicum*

Susanne Behrens², Uta Meyer¹, Holger Schankin¹, Michael A. Lonetto³, Ralf-Jörg Fischer¹, and Hubert Bahl¹∗

¹Institut für Molekulare Physiologie und Biotechnologie, Universität Rostock, Gertrudenstr. 11a, D-18051 Rostock, Germany
²Institut für Mikrobiologie und Genetik, Georg-August-Universität Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany
³Department of Stomatology, University of California at San Francisco Dental School, San Francisco, California 94143-0512, USA
⁴Present address: SmithKline Beecham Pharmaceuticals, 1250 S. Collegeville Road, UP1345 Collegeville, PA 19426, USA

Abstract

Two genes from *Clostridium acetobutylicum* DSM 792 were identified which are predicted to encode new members of the ECF subfamily of eubacterial RNA polymerase sigma factors. The *sigX* gene has the potential to encode a 184-amino acid protein with a molecular mass of 21,870 Da and with the highest overall similarity to *FecI* of *Escherichia coli* (27 % identical residues). The second gene, which is predicted to encode an alternative sigma factor of the ECF subfamily, is the previously described *orf2* gene (Gerischer and Dürr, 1990) located in the *adc* gene region of *C. acetobutylicum*. The deduced protein of *orf2* has significant similarity to SigX of *C. acetobutylicum* (22 % identical residues) and shares structural features with other alternative sigma factors. Therefore, it is proposed to rename *orf2* as *sigY*. Analysis of the phylogenetic relationship revealed that SigX from *C. acetobutylicum*, together with *αE* from *Streptomyces coelicolor* and SigX from *Bacillus subtilis*, form a Gram-positive cluster within the ECF subfamily and that SigY from *C. acetobutylicum* together with UviA from *Clostridium perfringens*, form a separate cluster located between the Gram-positive cluster and the sporulation sigma factor αH from *B. subtilis*.

Introduction

The ability to respond and to adapt to hostile environmental conditions is fundamental for bacterial survival in natural habitats. A variety of signal transduction and regulatory systems have developed to ensure the coordinate control of stress responses and cellular differentiation processes. These cellular rearrangements are mostly based on changes in gene expression which can be mediated by alternative sigma factors. A new class of eubacterial RNA polymerase sigma factors, the ECF subfamily, has been characterized which appears mainly to be responsible for the regulation of genes involved in extracytoplasmic functions (ECF) in response to extracytoplasmic stimuli (Lonetto et al., 1994).

*Clostridium acetobutylicum*, a strict anaerobe, Gram-positive bacterium is characterized by a very complex adaptation system. Environmental changes, like a drop of the pH in the culture medium below 5, the increase of acids, or the limitation of phosphate, result in a switch from acid to solvent formation (Dürr and Bahl, 1996). Along with this metabolic shift, the cells undergo a series of morphological changes in motility, shape, and granulose content. These processes can finally culminate in the formation of endospores as the stationary phase proceeds (Rogers, 1986). Thus, solventogenesis and sporulation appear to be tightly connected in *C. acetobutylicum*. As in *Bacillus subtilis* (Stragier and Losick, 1990; Errington, 1993), there is genetic evidence that the onset of sporulation and the development of spores in *C. acetobutylicum* is based on cascades of alternative sigma factors (Sauer et al., 1994 and 1995). The involvement of alternate sigma factors in solvent production has not yet been found.

In this paper we report the identification of two genes encoding putative new members of the ECF subfamily of eubacterial RNA polymerase sigma factors in *C. acetobutylicum*.  

Cloning and Sequencing of the *C. acetobutylicum sigX* Gene

During the screening of a sub-genomic library of *C. acetobutylicum* DSM 792 containing HindIII-digested DNA fragments in pUC18 with a probe derived from the N-terminus of P66, a protein associated with the RNA polymerase of this bacterium (unpublished results), a false positive clone was isolated, harboring the plasmid pSB20. Sequence analysis of the HindIII fragment revealed the presence of the 3′ end of an open reading frame (orf) whose deduced amino acid sequence showed significant similarity to the conserved region 4 of αR0-type sigma factors such as in *E. coli*. Further analysis of this gene revealed, that it encodes a protein of 456 amino acids with the highest overall similarity to SigX from *C. acetobutylicum* (22 % identical residues).

Received November 11, 1999; revised February 03, 2000; accepted February 03, 2000. ∗For correspondence. Email hubert.bahl@biologie.uni-rostock.de; Tel. +49-381-494-2247; Fax. +49-381-494-2244.

© 2000 Horizon Scientific Press


JMMB Communication
<table>
<thead>
<tr>
<th><strong>Further Reading</strong></th>
</tr>
</thead>
</table>
| **MALDI-TOF Mass Spectrometry in Microbiology**  
Edited by: M Kostrzewa, S Schubert (2016)  
www.caister.com/malditof |
| **Aspergillus and Penicillium in the Post-genomic Era**  
Edited by: RP Vries, IB Gelber, MR Andersen (2016)  
www.caister.com/aspergillus2 |
| **The Bacteriocins: Current Knowledge and Future Prospects**  
Edited by: RL Dorfl, SM Roy, MA Riley (2016)  
www.caister.com/bacteriocins |
| **Omics in Plant Disease Resistance**  
Edited by: V Bhadauria (2016)  
www.caister.com/opdr |
| **Acidophiles: Life in Extremely Acidic Environments**  
Edited by: R Quatrini, DB Johnson (2016)  
www.caister.com/acidophiles |
| **Climate Change and Microbial Ecology: Current Research and Future Trends**  
Edited by: J Marxsen (2016)  
www.caister.com/climate |
| **Biofilms in Bioremediation: Current Research and Emerging Technologies**  
Edited by: G Lear (2016)  
www.caister.com/biorem |
| **Microalgae: Current Research and Applications**  
Edited by: MN Tsialoglou (2016)  
www.caister.com/microalgae |
| **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**  
Edited by: H Shintani, A Sakudo (2016)  
www.caister.com/gasplasma |
| **Virus Evolution: Current Research and Future Directions**  
Edited by: SC Weaver, M Denison, M Roosainck, et al. (2016)  
www.caister.com/virusvol |
| **Arboviruses: Molecular Biology, Evolution and Control**  
Edited by: N Vasilakis, DJ Gubler (2016)  
www.caister.com/arbo |
| **Shigella: Molecular and Cellular Biology**  
Edited by: WD Picking, WL Picking (2016)  
www.caister.com/shigella |
| **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**  
Edited by: AM Romani, H Guasch, MD Balaguer (2016)  
www.caister.com/aquaticbiofilms |
| **Alphaviruses: Current Biology**  
Edited by: S Mahalingam, L Herrero, B Herring (2016)  
www.caister.com/alphha |
| **Thermophilic Microorganisms**  
Edited by: F Li (2015)  
www.caister.com/thermophile |
| **Flow Cytometry in Microbiology: Technology and Applications**  
Edited by: MG Wilkinson (2015)  
www.caister.com/flow |
| **Probiotics and Prebiotics: Current Research and Future Trends**  
Edited by: K Venema, AP Carro (2016)  
www.caister.com/probiotics |
| **Epigenetics: Current Research and Emerging Trends**  
Edited by: BP Chadwick (2015)  
www.caister.com/epigenetics2015 |
| **Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications**  
Edited by: A Burkovski (2015)  
www.caister.com/cory2 |
| **Advanced Vaccine Research Methods for the Decade of Vaccines**  
Edited by: F Bagnoli, R Rappuoli (2015)  
www.caister.com/vaccines |
| **Antifungals: From Genomics to Resistance and the Development of Novel Agents**  
Edited by: AT Coste, P Vandeputte (2015)  
www.caister.com/antifungals |
| **Bacteria-Plant Interactions: Advanced Research and Future Trends**  
www.caister.com/bacteria-plant |
| **Aeromonas**  
Edited by: J Graf (2015)  
www.caister.com/aeromonas |
| **Antibiotics: Current Innovations and Future Trends**  
Edited by: S Sánchez, AL Demain (2015)  
www.caister.com/antibiotics |
| **Leishmania: Current Biology and Control**  
Edited by: S Adak, R Datta (2015)  
www.caister.com/leish2 |
| **Acanthamoeba: Biology and Pathogenesis (2nd edition)**  
Author: NA Khan (2015)  
www.caister.com/acanthamoeba2 |
| **Microarrays: Current Technology, Innovations and Applications**  
Edited by: Z He (2014)  
www.caister.com/microarrays2 |
| **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**  
Edited by: D Marco (2014)  
www.caister.com/n2 |
as the alternative stress sigma factor $\sigma^\beta$ ($\sigma^{37}$) or the flagellar sigma factor $\sigma^\alpha$ ($\sigma^{93}$) of *B. subtilis*. This orf, which was tentatively designated sigX, overlapped (eight nucleotides) with another orf of unknown function (orf36). We are interested in the regulation of the stress response in *C. acetobutylicum*, in which new sigma factors might be involved. The entire gene of the putative alternate sigma factor SigX was therefore isolated. SigX is not identical with the known alternative sigma factors of *C. acetobutylicum* involved in sporulation (Sauer et al., 1994 and 1995). A sigX-specific DNA probe (168 bp) was generated by PCR and used as a probe for the screening of a genomic library of *C. acetobutylicum* DSM 792 containing partially Sau3A-digested DNA fragments cloned in pUC9. One clone could be isolated whose recombinant plasmid pSBS21 carried the complete sigX and orf36 genes and the 5' end of orf18 located upstream of sigX in opposite orientation. The nucleotide sequence of the 2.2-kbp Sau3A fragment of *C. acetobutylicum* was determined. The sequence data have been submitted to the EMBL data bank and have been assigned the accession number U58131. The sequence of sigX predicts a protein of 184 amino acids with a molecular weight of 21,870 Da.

**Identification of the sigX and orf2 (sigY) Gene Products as Proposed Members of the ECF Subfamily of Sigma Factors**

The nucleotide sequence and the deduced protein were analyzed using DNA Strider Software on a Macintosh LCII computer. Further sequence analysis and sequence comparisons were carried out by using the Wisconsin Genetics Computer Group sequence analysis software package, Version 8.0-Unix (University of Wisconsin Biotechnology Center, Madison). The databases searched were EMBL, GenBank, SwissProt, PIR, and NBRF. The similarity search using the BLAST network service revealed an overall similarity to SigX, which was also reflected in their similar lengths and predicted molecular weights. Eight of these proteins, including $\sigma^E$ and Fecl of *Escherichia coli*, $\sigma^E$ of *Streptomyces coelicolor*, and SigX of *B. subtilis* (Table 1), have previously been classified as members of the new ECF subfamily of $\sigma^{70}$-type sigma factors, according to their common structural and functional features (Lonetto et al., 1994). All these proteins, with the exception of the *B. subtilis* SigX, have been identified as positive regulators of extracytoplasmic functions and as effector molecules responding to extracytoplasmic stimuli. Furthermore, there is biochemical evidence that the $\sigma^E$ proteins of *E. coli* and *S. coelicolor* represent RNA polymerase sigma factors (Buttner et al., 1990; van Howe et al., 1990; Lonetto et al., 1994). Fecl of *E. coli*, which positively regulates the citrate-dependent iron (III) transport system (van Howe et al., 1990), and SigX, which is suggested to play a similar role in *B. subtilis* (Sorokin et al., 1993), share the highest overall similarities with the SigX protein of *C. acetobutylicum*.

Surprisingly, SigX of *C. acetobutylicum* also showed a significant similarity to the deduced protein of the previously described orf2 from the same organism. This gene is located in the *adc* (acetocacetate decarboxylase) gene region which is important for the metabolic switch from acidogenesis to solventogenesis in *C. acetobutylicum* (Gerischer and Dürré, 1990). Furthermore, similarity of Orf2 was found to the early sporulation sigma factor $\sigma^D$ of *B. subtilis* (52% similar residues) and to the UviA protein of *C. perfringens* (50% similar residues) (Table 2). The function of the UviA protein has not been characterized to date. Due to the similarity of the orf2 product to sigma factors (see also below) we propose to designate this gene sigY.

The multi-alignment of the SigX and SigY proteins of *C. acetobutylicum* with some proposed ECF subfamily members, the alternative sigma factors $\sigma^B$, $\sigma^D$, and $\sigma^H$ of *B. subtilis*, and UviA of *C. perfringens*, clearly demonstrates that both proteins share similarity across all four conserved regions of the $\sigma^{70}$ family (Figure 1). As for all members of the ECF subfamily (2), the strongest resemblance can be found in region 4, which has been shown to participate in -35 promoter sequence recognition (Dombroski et al., 1992). The conservation also includes the putative helix-turn-helix (HTH) DNA binding motif in subregion 4.2. However, in contrast to other ECF family members, there is a good conservation of the basic cluster downstream of the HTH in SigY and, to a lesser extent, in SigX. In common with other members of the ECF family of sigma factor peptides, much of region 3 is missing in the SigX protein of *C.

Table 2. Similarities of the SigY Protein of *C. acetobutylicum* to Members of the $\sigma^{70}$ Family

<table>
<thead>
<tr>
<th>Organism</th>
<th>Protein</th>
<th>Identity/ Similarity [%]*</th>
<th>Length [aa]</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. perfringens</em></td>
<td>UviA</td>
<td>27 / 50</td>
<td>185</td>
<td>[M32882]</td>
</tr>
<tr>
<td><em>C. acetobutylicum</em></td>
<td>SigX</td>
<td>22 / 38</td>
<td>184</td>
<td>[U58131]</td>
</tr>
<tr>
<td><em>B. subtilis</em></td>
<td>SigF</td>
<td>22 / 45</td>
<td>255</td>
<td>[P07860]</td>
</tr>
<tr>
<td><em>C. acetobutylicum</em></td>
<td>SigG</td>
<td>22 / 45</td>
<td>257</td>
<td>[P36558]</td>
</tr>
<tr>
<td><em>B. subtilis</em></td>
<td>SigH</td>
<td>21 / 52</td>
<td>218</td>
<td>[M29693]</td>
</tr>
</tbody>
</table>

* based on BestFit alignments using the GCG software package

* proposed ECF subfamily sigma factors (Lonetto et al., 1994)

+ based on the alignment of the C-terminal (region 4) of the proteins
Figure 1. Multiple alignment of the SigX and SigY proteins of C. acetobutylicum with members of the ECF subfamily, and other alternative sigma factors of $\sigma_{70}$-type sigma factors. Multiple alignments derived from automated alignment tools were edited by hand to reflect the results produced by BLAST (Altschul et al., 1990) as well as earlier alignments (Lonetto et al., 1992 and 1994). Regions of conservation (Helmann and Chamberlin, 1988; Lonetto et al., 1992) are shown below the alignments, as are putative helix-turn-helix motifs (Buttner et al., 1990). Residues which are conserved across the full set of sigmas are highlighted where 5 or more residues match. Identical residues and the following sets of residues are considered matched: DE, NQ, RK, ST, FYW, and ILMV. The accession numbers (Genbank EMBL) of proteins compared are given in parentheses: C. a. SigX: C. acetobutylicum SigX (U58131); B.s. SigX: B. subtilis SigX (L09228); P. a. AlgU: P. aeruginosa AlgU (L02119); E.c. SigE: E. coli $\sigma_E$ (D13169); M.x. CarQ: M. xanthus CarQ (X71062); S.c. SigE: S. coelicolor $\sigma_E$ (L29636); B.s. SigH: B. subtilis $\sigma_H$ (M29693); B.s. SigB: B. subtilis $\sigma_B$ (M34995); B.s. SigD: B. subtilis $\sigma_D$ (M20144); C.p. UviA: C. perfringens UviA (M32882); C.a. SigY: C. acetobutylicum SigY (M55392).
The length of region 3 in SigY corresponds to the one of region 3 of the early sporulation sigma factor $\sigma^H$ of B. subtilis and is in between the length of ECF and other $\sigma^{70}$ family members. Region 2 represents the most highly conserved segment of sigma factors. However, some of these strongly conserved features appear to differ in ECF sigma factors. Most notably, the sequence pattern of subregion 2.4, proposed to form an alpha-helix and to be involved in -10 promoter recognition, was shown to have a different structure in ECF sigmas but at the same time to be highly conserved within this subfamily (Lonetto et al., 1994). This is also the case for SigX of C. acetobutylicum but not for SigY of C. acetobutylicum and UviA of C. perfringens. The latter show a higher similarity of subregion 2.4 to other $\sigma^{70}$ family members than to the ECF group (Figure 1).

To support the sequence comparison data which suggest that SigX and SigY are sigma factors we have further examined the relationships between SigX, SigY and other sigma factors by calculating phylogenetic trees by both distance and parsimony methods (Figure 2). Both methods agree on the placements of SigX and SigY. As expected, SigX of C. acetobutylicum clusters more directly with the ECF group than SigY or UviA of C. perfringens. The latter two proteins form a separate cluster located between the Gram-positive cluster and the alternative sporulation factor $\sigma^H$ of B. subtilis. On the basis of the similarity and identity values (Table 1), the C. acetobutylicum and B. subtilis SigX proteins are closely related and together with S. coelicolor $\sigma^E$ form the Gram-positive cluster. This cluster and the iron-responsive group, including E. coli Fecl, appear to represent a separate branch from the extensive branching node formed by HrpL, CnrH, CarQ, and the $\sigma^E$-cluster, comprising AlgU and the E. coli and S. typhimurium $\sigma^E$ proteins. Finally, it is notable...
with a molecular mass of 22 kDa, as estimated by SDS-polyacrylamide gel electrophoresis (data not shown). Furthermore, as a prerequisite to provide biochemical evidence that SigX and SigY function as sigma factors, both proteins were overexpressed in *E. coli* and affinity purified via a His-tag (data not shown). A specific interaction of SigX bound to a Ni-NTA agarose affinity column, with the core enzyme of RNAPolymerase of *E. coli* was observed and is in agreement with the function of SigX as a sigma factor (Figure 4).

In conclusion, the structural features of SigX and SigY of *Clostridium acetobutylicum*, evident from the alignment data, together with their phylogenetic relationship strongly suggest that these two proteins represent new members of, or are closely related to, the new ECF subfamily of eubacterial RNA polymerase sigma factors.

Acknowledgements

This work was supported by a grant from the Deutsche Forschungsgemeinschaft. Sequence analysis was carried out using the NCBI BLAST Network Service at NIH, and facilities of the UCSF Computer Graphics Lab.

References


