Cloning, Sequencing and Expression of a Sialidase Gene from *Clostridium sordellii* G12

By BERND ROTHE,1 PETER ROGGENTIN,1 RONALD FRANK,2 HELMUT BLÖCKER2 AND ROLAND SCHAUER1*

1 Biochemisches Institut, Christian-Albrechts-Universität, Olshausenstr. 40, D-2300 Kiel, Federal Republic of Germany
2 DNA-Synthesegruppe der Gesellschaft für Biotechnologische Forschung mbH, Mascheroder Weg 1, D-3300 Braunschweig, Federal Republic of Germany

(Received 14 March 1989; revised 21 June 1989; accepted 9 August 1989)

A 4.3 kb *XbaI* restriction fragment of DNA from *Clostridium sordellii* G12 hybridized with a synthetic oligonucleotide representing the N-terminus of the sialidase protein secreted by *C. sordellii*. This cloned fragment was shown to encode only part of the sialidase protein. The sialidase gene of *C. sordellii* was completed by a 0.7 kb *RsaI* restriction fragment overlapping one end of the *XbaI* fragment. After combining the two fragments and transformation of *Escherichia coli*, a clone that expressed sialidase was obtained. The nucleotide sequence of the sialidase gene of *C. sordellii* G12 was determined. The sequence of the 18 N-terminal amino acids of the purified extracellular enzyme perfectly matched the predicted amino acid sequence near the beginning of the structural gene. The amino acid sequence derived from the complete gene corresponds to a protein with a molecular mass of 44735 Da. Upstream from the putative ATG initiation codon, ribosomal-binding site and promoter-like consensus sequences were found. The encoded protein has a leader sequence of 27 amino acids. The enzyme expressed in *E. coli* has similar properties to the enzyme isolated from *C. sordellii*, except for small differences in size and isoelectric point. Significant homology (70%) was found with a sialidase gene from *C. perfringens*.

**INTRODUCTION**

Sialidases (= neuraminidases, EC 3.2.1.18) hydrolyse O-glycosidic linkages between sialic acids and glycosyl residues of oligosaccharides, glycoproteins and glycolipids. These enzymes are widely distributed among micro-organisms, i.e. viruses, bacteria and protozoa, and are ubiquitous in the deuterostomate branch of metazoa (Corfield & Schauer, 1982). Interestingly, most of the bacteria known to produce sialidase are symbionts or pathogens of animals. Sialidases have been suggested to be pathogenic factors in microbial infections (Müller, 1974; Schauer, 1983), although their specific role still remains uncertain.

A sialidase of *Clostridium sordellii* G12, an isolate from human gas gangrene, has been purified (Roggentin *et al.*, 1987). This enzyme was used to start investigations on the molecular biology of bacterial sialidases. Although several viral sialidase genes have been studied (Colman *et al.*, 1987; Jorgenson *et al.*, 1987), the relationship and phylogeny of viral, bacterial and eukaryotic sialidase genes is still unknown. The sequence of a sialidase gene of *C. perfringens* (Roggentin *et al.*, 1988b) and part of the sialidase gene of *Vibrio cholerae* (Vimr *et al.*, 1988) have been published. In this communication we present the complete sequence of a sialidase gene of *C. sordellii* G12.

**Abbreviations**: MU-NeuSAc, 4-methylumbelliferyl-α-D-N-acetylneuraminic acid.

0001-5469 © 1989 SGM
**METHODS**

**Bacterial strains and vectors.** *C. sordellii* G12 was a gift of Dr Rainer Hobrecht, Untersuchungsinstitut I des Sanitättdienstes der Bundeswehr, Kronshagen, FRG. For cloning and sequencing, the *Escherichia coli* strains JM101 and JM109 (Messing, 1983) were used as hosts, transformed with the vectors pUC18, M13mp18 or M13mp19 (Yanisch-Perron et al., 1985).

**Sialidase assays and purification.** *E. coli* clones were sprayed with the synthetic substrate 4-methylumbelliferyl-

**DNA methods and manipulations.** *C. sordellii* G12 was grown anaerobically in Todd–Hewitt broth (Difco) for 6 h at 37 °C after 1:1 inoculation. Cells from 5 litres of culture were harvested by centrifugation (10 000 g, 20 min, 4 °C), resuspended in saline EDTA (0-15 M-NaCl, 0-1 M-EDTA, pH 6-8; 10 ml per g cells) containing 10 mg lysozyme ml⁻¹ and 0-1 mg proteinase K ml⁻¹ (Boehringer). Cells were lysed by the addition of SDS to a final concentration of 2% (w/v). Chromosomal DNA was isolated according to Marmur (1961). Preparation of plasmid DNA, ligatation, transformation, agarose gel electrophoresis and electroelution were done as described by Maniatis et al. (1982). All enzymes for DNA technology were obtained from Gibco BRL or Boehringer.

**Synthesis and labelling of oligonucleotide probes.** The oligonucleotides corresponding to the N-terminus of the sialidase secreted by *C. sordellii* were synthesized simultaneously on cellulose discs as segmental supports (Frank et al., 1987) by established phosphotriester procedures (Ohtsuka et al., 1982), and purified by anion-exchange HPLC (Gait et al., 1982). Otherwise, DNA probes were synthesized by Dr Werner Liesack, Institut für Allgemeine Mikrobiologie, University of Kiel, using an automated DNA synthesizer (380 B, Applied Biosystems). The oligonucleotides were labelled with [γ-³²P]ATP by T4 polynucleotide kinase in 0-03 M-mercaptopethanol, 0-17 M-spermidine, 0-02 M-MgCl₂, and 0-05 M-Tris/HC1, pH 8-0.

**Blotting procedures.** DNA was transferred from agarose gels to nylon sheets as described by Southern (1975). *E. coli* colonies were transferred to nylon filters (Davis et al., 1986). Adherent cells were lysed, and DNA was denatured and immobilized on the filter by 5 min treatment with each of the following solutions: 0-5 M-NaOH; 1-5 M-NaCl in 0-5 M-Tris/HC1, pH 7-4; 1-5 M-NaCl in 2 × SSC buffer (1 × SSC is 0-15 M-NaCl, 0-015 M-trisodium-citrate). After drying, the sheets were heated for 2 h at 80 °C.

**Hybridization.** This was done as described by Wallace et al. (1981). Nylon sheets with immobilized DNA were prehybridized in 6 × SSC buffer, containing 5 × Denhardt’s solution (Maniatis et al., 1982), 0-5% SDS and 200 μg denatured salmon sperm DNA ml⁻¹, at 41 °C for at least 1 h. For overnight hybridization, the same buffer and temperature were used. The filters were washed three times in 6 × SSC at 41 °C and exposed to X-ray film (Kodak).

**DNA sequencing.** Nucleotide sequencing was done by the dideoxynucleotide method of Sanger et al. (1977) using [α-³⁵S]dATP (Amersham). For primer extension the universal primer or a synthetic oligonucleotide were used with a M13 sequencing kit (Boehringer). Electrophoresis was performed in 6% polyacrylamide wedge-shaped gels.
RESULTS

Design of synthetic oligonucleotides

The following sequence of 18 amino acids of the N-terminus of the purified sialidase from \textit{C. sordellii} was obtained by automated Edman degradation: S N L N T T N E P Q K T T V F N K N. A sequence of six amino acids (numbers 7 to 12) was selected, which shared minor codon variability. A comparison of clostridial protein sequence data so far published (Beguin \textit{et al.}, 1985; Eisel \textit{et al.}, 1986; Graves \textit{et al.}, 1985; Grepinet \\& Beguin, 1986) revealed a strongly biased codon usage. A and T were used more frequently at the third position. Therefore a mixture of only the four following oligonucleotides, all 17 bases long, were synthesized:

\begin{align*}
5' &\text{ AAT GAA CCA CAA AAA AC } 3' \\
&\text{G T N E P Q K T}
\end{align*}

Cloning of \textit{C. sordellii} DNA fragments

Genomic DNA from \textit{C. sordellii} was digested with several restriction enzymes, electrophoresed on 0.7-7% agarose gel, transferred to nylon sheets and hybridized with the labelled oligonucleotide mixture. One single band each was obtained after treatment of DNA with \textit{HindIII} (3.9 kb), \textit{XbaI} (4.3 kb) or \textit{Sau3A} (1.1 kb). The 4.3 kb \textit{XbaI} fragment was selected for cloning, as follows. \textit{XbaI}-digested DNA (100 \mu g) of \textit{C. sordellii} was electrophoresed on a preparative agarose gel. The fragments ranging from 3.5 to 5.0 kb were electroeluted, ligated into the \textit{XbaI} site of the pUC18 vector, and used for transformation of \textit{E. coli}. About 600 clones were screened by colony hybridization. One positive clone was identified and shown by Southern blot hybridization to carry the 4.3 kb \textit{XbaI} fragment.

To map the position which encodes the N-terminus of the sialidase on the 4.3 kb \textit{XbaI} fragment, the insert was isolated and successively reduced by \textit{Bal31} exonuclease digestion from both ends. Only fragments larger than 2.8 kb hybridized with the N-terminal DNA probe; on the (possibly over-simplified) assumption that \textit{Bal31} digests at the same rate at both ends, it was tentatively concluded that the sequence near the 5' end of the mRNA of the sialidase gene is located 700–800 bp from one end of the \textit{XbaI} fragment. Furthermore, a 2.7 kb \textit{HindIII}/\textit{XbaI} fragment (see Fig. 1) obtained from the \textit{XbaI} fragment hybridized with the probe, indicating that the starting point maps about 700–800 bp from the right-hand \textit{XbaI} site in Fig. 1.

To determine the orientation of the sialidase gene on the cloned 4.3 kb fragment, the 2.7 kb \textit{HindIII}/\textit{XbaI} fragment was subcloned into M13mp18 and into mp19. By hybridization against single-stranded DNA of both M13 subclones, only M13mp18 clones gave positive signals. These results clarified that the sialidase gene runs in the direction of the right-hand \textit{XbaI} site in Fig. 1. From this and the result that the hybridization site maps 700–800 bp from this site, it was concluded that the \textit{XbaI} fragment contains only part of the sialidase gene.

The sequence stretching about 1000 bp from the right-hand \textit{XbaI} site was determined from M13 phage clones shown in Fig. 1. Subclones carrying restriction fragments of the \textit{XbaI} fragment, and also clones obtained after successive \textit{Bal31} exonuclease deletion of the \textit{XbaI} insert from one side, were used for overlapping sequencing.

An oligonucleotide (GTATTGGCATCATATT) was synthesized, which maps right from the first \textit{RsaI} site indicated by an asterisk on the sequenced region (Fig. 1). Hybridization of restricted DNA of \textit{C. sordellii} with this labelled oligonucleotide revealed a signal with a 0.7 kb \textit{RsaI} fragment. The \textit{RsaI} fragment was cloned in pUC18 using methods as for the \textit{XbaI} fragment and was subcloned in M13 for sequencing. In addition, restriction fragments of the \textit{RsaI} fragment were subcloned in M13 and used for sequencing (Fig. 1). The \textit{RsaI} fragment and the
Fig. 1. Map of restriction sites on the cloned fragments and sequencing strategy for the sialidase gene of C. sordellii G12. , Synthetic oligonucleotides used for screening: --, sequence determined from M13mp18/19 clones carrying restriction fragments of the XbaI or Rsal fragment; ----, sequence determined from M13mp18 clones carrying reduced HindIII-XbaI fragments obtained after successive Bal31 digestion from one end.

The XbaI fragment overlapped in a sequence of 182 bp. On combining the sequences obtained from both fragments the complete sequence of the sialidase gene, ending about 50 bp from the end of the Rsal fragment, was determined.

The nucleotide sequence and the deduced amino acid sequence of the sialidase gene of C. sordellii are shown in Fig. 2. Translation of nucleotides +82 to +135 gives an amino acid sequence identical to the sequence obtained by Edman degradation of the N-terminus of the extracellular C. sordellii sialidase. The putative ATG initiation codon is assigned to position +1, which follows a Shine–Dalgarno sequence AAAAGGGGAGAG. Two promoter consensus sequences are found upstream from the ATG codon. Between –195 and –172, the sequences TTGACA (–35 region) and TAATAT (–10 region) are part of the more extended one of the two inverted repeats in front of the gene. A second putative promoter sequence, ATGACA and TATTAT, was found between positions –86 and –58.

The reading frame from residues +1 to +1212 encodes a protein of 404 amino acids with a calculated molecular mass of 44735 Da. The sequence from the first methionine to the N-terminus determined by Edman degradation comprises 27 amino acids. A triplet sequence Met-Lys-Lys represents positively charged amino acids, followed by a stretch of mostly hydrophobic residues ending with Ala at the assumed cleavage site of the leader peptide. Discounting the leader peptide, the molecular mass of the encoded protein can be calculated as 41900 Da, which is close to the 40000 Da estimated by SDS-PAGE for the C. sordellii sialidase (Roggentin et al., 1987). Sequences showing dyad symmetry were not found downstream from the stop codon.

The 2.7 kb HindIII-XbaI fragment was subcloned in pUC18. The insert-bearing vector was digested with XbaI and SsrI. The linearized plasmid was separated from the excised part of the polylinker by agarose gel electrophoresis and electroelution. The plasmid carrying the 0.7 kb Rsal fragment was digested with XbaI and SsrI, resulting in a 0.5 kb fragment which still encodes the C-terminus of the sialidase. After isolation, this 0.5 kb fragment was inserted
Fig. 2. Nucleotide sequence of the sialidase gene of C. sordellii G12 and derived amino acid sequence. Two inverted repeats are indicated by arrows. The putative ribosomal binding site and the sequence obtained by Edman degradation are underlined. Two hypothetical promoter sequences are framed.
Table 1. Codon usage in the sialidase gene of C. sordellii

<table>
<thead>
<tr>
<th>Codon</th>
<th>Amino Acid</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>TTT 9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTC 1</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>TTA 15</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>TTG 4</td>
<td>Stop</td>
</tr>
<tr>
<td></td>
<td>CTT 2</td>
<td>His</td>
</tr>
<tr>
<td></td>
<td>CTC 0</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>CTA 3</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td>CGT 0</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>ATT 16</td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td>ACT 14</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>ATA 13</td>
<td>Lys</td>
</tr>
<tr>
<td>Met</td>
<td>ATG 13</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>GTT 12</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>GTC 0</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>GTA 8</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>GTG 0</td>
<td></td>
</tr>
</tbody>
</table>

Codon usage

Table 1 shows the codon usage of the sialidase gene of C. sordellii. A clear preference for A and T residues at the third position of the triplets was found. In addition, fourteen triplets are not used for coding.

Comparison of the enzymes expressed by E. coli and C. sordellii

The sialidase gene of C. sordellii is expressed in E. coli, and the enzymes expressed by the two species have similar properties. The recognition by polyclonal antibodies raised against two clostridial sialidases was not affected by cloning in E. coli: antibodies raised against C. sordellii sialidase and antibodies raised against a cloned sialidase of C. perfringens (Roggentin et al., 1988b) inhibited sialidase activity by 70% and 50%, respectively, for both the cloned enzyme and that isolated from C. sordellii. Inhibitors like Hg^{2+}, 2-deoxy-2,3-didehydro-N-acetylneuraminic acid, N-(4-nitrophenoxy)oxamic acid and N-acetylneuraminic acid decreased the activity of both sialidases to the same extent. Substrate specificity and $K_m$ values, and the temperature and pH optima of the C. sordellii sialidase (Roggentin et al., 1987) were not changed by cloning in E. coli. One difference, however, was measured concerning the isoelectric point (pI), which was 0.3 pH units lower for the cloned enzyme as compared to the original one. Additionally, the enzyme expressed in E. coli ran slightly more slowly on SDS-PAGE than the enzyme purified from C. sordellii culture medium (Fig. 3).

DISCUSSION

The codon usage pattern of the sialidase gene (Table 1) reflects the low G + C content (25%) of C. sordellii DNA (Cato et al., 1986). The sequence of 27 amino acids considered to be the leader peptide of the sialidase gene closely resembles the common features of prokaryotic signal sequences (Watson, 1984). In addition, the signal sequence shows similarity with some proteins...
Sialidase gene of Clostridium sordellii

Fig. 3. SDS-PAGE of clostridial sialidases stained with silver. Lane 1, sialidase of *C. perfringens* expressed in *E. coli*; lane 2, sialidase of *C. sordellii* isolated from *C. sordellii* culture medium; lane 3, sialidase of *C. sordellii* isolated from *E. coli* cells.

from Gram-positive organisms, which have been shown to be exported via a lipoprotein intermediate (Nielsen & Lampen, 1982). These proteins are cleaved in front of a cysteine within the leader peptide, forming a glyceride thioether bond which links the protein to the membrane. A second proteolytic cleavage releases the mature extracellular protein. On comparison of five proteins processed and exported in this way, Nielsen & Lampen (1982) proposed the following consensus sequence for the cleavage site in front of the cysteine residue:

\[
\text{LAGC}^{\text{neutral or non-polar}} \text{SN}
\]

The sialidase contains the following sequence in the leader region:

\[
\text{LSACNIIN}
\]

This is similar to the leader peptide of *Staphylococcus aureus* β-lactamase (McLaughlin *et al.*, 1981), which is exported via a lipoprotein intermediate:

\[
\text{LSACNNSN}
\]

It is therefore possible that the sialidase produced by *C. sordellii* undergoes a similar means of export.

The signal sequence seems to be at least partially recognized by *E. coli*, since the sialidase is liberated by lysozyme treatment, indicating its location in the periplasmic space. The sialidase is not transported through the outer membrane of *E. coli* into the extracellular medium, which has
### Comparison of Amino Acid Sequences of C. sordellii and C. perfringens Sialidases

<table>
<thead>
<tr>
<th>CsS</th>
<th>CpS</th>
</tr>
</thead>
<tbody>
<tr>
<td>M K K F I K I L K V L S H A I V L S A C</td>
<td>M C</td>
</tr>
<tr>
<td>N I N G I F A S N L N T T N E P Q K T T</td>
<td></td>
</tr>
<tr>
<td>N H U N T E K N L D I S H K F E P L I</td>
<td></td>
</tr>
<tr>
<td>V F N K N D N T W N A Q Y F R I P S L G</td>
<td></td>
</tr>
<tr>
<td>L T N K D N N I W N S K Y F R I P N I C</td>
<td></td>
</tr>
<tr>
<td>T L A D G T M L A F S D I R Y N G A E D</td>
<td></td>
</tr>
<tr>
<td>L N D G T I L T F S D I R Y N G P D D</td>
<td></td>
</tr>
<tr>
<td>H A Y I D I G A A K S T D N G Q T W D V</td>
<td></td>
</tr>
<tr>
<td>H A Y I D I A S A R S T D F C T W S Y</td>
<td></td>
</tr>
<tr>
<td>K T V K E N D R I D S T F S R V K D S T</td>
<td></td>
</tr>
<tr>
<td>N I A M K N R I D S T Y S R V K D S T</td>
<td></td>
</tr>
<tr>
<td>T V T T D T G R I I L I A G S W N K N G</td>
<td></td>
</tr>
<tr>
<td>T V I T N T G R I I L I A S V N T N G</td>
<td></td>
</tr>
<tr>
<td>N V A S S T T S L R S D W S V Q N V Y S</td>
<td></td>
</tr>
<tr>
<td>N V A M T T S T R S D W S V Q N Y S</td>
<td></td>
</tr>
<tr>
<td>D D N G E T W S D K X D L T T N K A R I</td>
<td></td>
</tr>
<tr>
<td>D D N C L T W S N X I D L T K D S S K V</td>
<td></td>
</tr>
<tr>
<td>K N Q P S N T I G W L A C V G S G I V M</td>
<td></td>
</tr>
<tr>
<td>K N Q P S N T I G W L G C V G S G I V M</td>
<td></td>
</tr>
<tr>
<td>S D C T I N P I Q Z A L R E N N A N N</td>
<td></td>
</tr>
<tr>
<td>D D C T I V M P A Q I S L R E N N K E N N</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4 (continued on facing page).** Alignment of amino acid sequences of *C. sordellii* (CsS) and *C. perfringens* (CpS) sialidases. : denotes amino acids with similar properties.

...also been observed with other proteins cloned and expressed in this bacterium (Oliver, 1985; Pugsley & Schwarz, 1985).

The sialidases expressed in *E. coli* and *C. sordellii* have almost identical properties with regard to molecular and kinetic parameters. The slight differences in pI and mobility measured by SDS-PAGE cannot yet be explained. They may be due to different processing of the leader peptide in *E. coli* and/or to some proteolytic degradation at the C-terminus.

Comparison of the complete sequence of the *C. sordellii* sialidase gene with sequences of the MicroGenie gene bank revealed only weak homologies. Significant homology was recognized between the sialidase gene of *C. sordellii* and a cloned sialidase gene of *C. perfringens* (Roggentin et al., 1988). In the coding regions, 73.1% of the nucleotide residues and 67.4% of the amino acid residues are identical. Furthermore, amino acids with similar properties occur at many other positions (Fig. 4). Both sialidases have similar molecular masses: 41900 Da for the processed *C. sordellii* and 42818 Da for the cloned *C. perfringens* sialidase, as determined from the DNA sequences. For the *C. perfringens* sialidase the difference in mobility on SDS-PAGE (Fig. 3) might be explained by partial degradation during purification (Roggentin et al., 1988b). Furthermore, enzymic and molecular properties of the *C. sordellii* sialidase (Roggentin et al., 1987) and the cloned *C. perfringens* sialidase (unpublished results) are similar. The lack of a hydrophobic leader sequence in the latter sialidase represents the main difference in the primary structure between the two enzymes. Interestingly, the *C. perfringens* sialidase has a cysteine...
residue at the N-terminus, followed by NKN, which is similar to a corresponding part of the consensus sequence discussed above for the C. sordellii sialidase.

The broad distribution of sequence similarities over the whole length of both clostridial sialidase genes does not allow us to pinpoint any conserved regions, which could be essential for function. Such regions may be discovered on investigation of the sequences of further sialidase genes.

We thank Professor Dr Friedrich Lottspeich, Max-Planck-Institut für Biochemie, Martinsried, for N-terminal sequencing, Dr Werner Liesack, Institut für Allgemeine Mikrobiologie, Universität Kiel, for synthesizing DNA probes, Dr Beate Joos and Dr Telse Roggentin for the preparation of polyclonal antibodies, Professor Dr E. Stackebrandt, Institut für Allgemeine Mikrobiologie, Universität Kiel, for his cooperation, and Marzog El Madani for technical assistance.

This work was supported by the Deutsche Forschungsgemeinschaft (grant Scha 202/13-1) and Fonds der Chemischen Industrie.

REFERENCES


