Volume 238, number 1, 31–34 FEB 06320 September 1988

Cloning and sequencing of a *Clostridium perfringens* sialidase gene

Peter Roggentin, Bernd Rothe, Friedrich Lottspeich* and Roland Schauer

Biochemisches Institut, Christian-Albrechts-Universität, Olshausenstr. 40, D-2300 Kiel and **Max-Planck-Institut für Biochemie, D-8033 Martinsried, FRG*

Received 25 July 1988

Escherichia coli was transformed with pUC vectors containing Sau3A restriction fragments (RF) of *Clostridium perfringens* DNA. Two clones expressed sialidase activity when assayed with the fluorogenic substrate 4-methylumbelliferyl-α-D-*N*-acetylneuraminic acid. A synthetic oligonucleotide representing the N-terminus of the expressed enzyme hybridized with the clostridial insert and with a corresponding 2.1 kb Sau3A RF of the *C. perfringens* genome. The insert reduced to 1.4 kb, which still encoded active sialidase, has been sequenced. The structural gene encodes 382 amino acids representing an M_r of 42 770. A hydrophobic leader sequence is absent. Upstream from the initiation codon ATG, a GA-rich region is found and considered as the Shine-Dalgarno sequence. Homology with the N-terminus of the *Vibrio cholerae* sialidase gene and with viral sialidase sequences was not found.

Sialidase; Gene sequence; Gene expression; *(Clostridium perfringens)*

1. INTRODUCTION

Turnover of cell surface glycoconjugates in metazoan animals of the deuterostomate lineage is initiated by the cleavage of terminal sialic acids. The enzyme responsible (sialidase = neuraminidase, EC 3.2.1.18), obligatory in this phylogenetic branch, is also expressed in some viruses, bacteria and protozoa, which have close contact with animal hosts [1]. This may be due to the observation that sialidase-altered cell surfaces allow more direct contact of pathogens with the host and are advantageous to further damage of the cells by lytic enzymes [2]. In order to obtain information on the phylogeny and relatedness of sialidases, we started investigations on bacteria of the genus *Clostridium*, which secrete large amounts of soluble enzymes. The sialidase genes might be easily accessible using DNA probes which represent the N-terminus of isolated and sequenced enzyme proteins, or by looking for sialidase expression of *E. coli* clones transformed with clostridial DNA RF. The latter was observed after cloning of Sau3A RF of a *C. perfringens* strain DNA in *E. coli*. Unexpectedly, the highly expressed gene product does not represent the previously described sialidase of this species [3], although it is encoded by the cloned clostridial insert. The functionally complete gene and the predicted amino acid sequence of this sialidase are presented here.

Correspondence address: R. Schauer, Biochemisches Institut, Christian-Albrechts-Universität, Olshausenstr. 40, D-2300 Kiel, FRG

Abbreviations: RF, restriction fragments; MU-Neu5Ac, 4-methylumbelliferyl-α-D-*N*-acetylneuraminic acids

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00963

2. MATERIALS AND METHODS

C. perfringens strain A99, isolated from human gas gangrene infection, was a gift from Dr R. Hobrecht (Untersuchungsinstitut I der Bundeswehr, Kronshagen, FRG). *C. perfringens* type strain DSM 756 was purchased from Deutsche Sammlung für Mikroorganismen (Göttingen). Restriction enzymes, T_4 kinase, and Bal31 nuclease for insert reduction were from Gibco-BRL (Eggenstein, FRG). Vectors for cloning (pUC18, pUC19) and sequencing (M13mp18, mp19) as well as

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

E. coli strain JM101 were purchased from Pharmacia-LKB (Freiburg). Radiochemicals for hybridization ([γ-^{32}P]ATP) and sequencing ([α-^{35}S]thio-dATP), and nylon sheets (Hybond^{N}) for Southern blotting [4] were obtained from Amersham Buchler (Braunschweig).

Clostridial cells were harvested by centrifugation (10000 × *g*, 20 min, 4°C) after 1:10 inoculation and multiplication for 6 h at 37°C in Todd-Hewitt broth (Difco, Detroit), supported by 0.1 mM sialoglycopeptides isolated from edible bird nest substance [3]. The sedimented cells were treated with 10 mg lysozyme (Serva, Heidelberg) and 0.1 mg proteinase K (Boehringer, Mannheim) per ml saline-EDTA (0.15 M NaCl, 0.1 M EDTA; pH 6.8) for 30 min at 37°C. Cells were lysed by addition of SDS to a final concentration of 2%. DNA was purified according to [5]. Cloning, plasmid isolation, electroelution, blotting and hybridization were performed as in [6].

Clones expressing sialidase activity were detected by spraying the colonies with 0.1 mM MU-Neu5Ac in 0.1 M acetate buffer, pH 6. Positive clones were visualised by the blue-white fluorescence of released MU under UV light of 360 nm.

Purified sialidase protein (unpublished) from one active clone was applied onto a 10% PAGE and electroblotted onto siliconized glass fiber (Glassybond, Biometra, Göttingen) for 3.5 h at 200 mA and 4°C. The 40 kDa band was excised and sequenced on a 470A gas-phase sequencer (Applied Biosystems, Pfungstadt, FRG) as described [7].

DNA probes were synthesized applying an automated DNA synthesizer (380B Applied Biosystems).

Nucleotide sequences were determined by the chain termination method [8] using an M13 sequencing kit (Boehringer). The sialidase sequence obtained was compared with sequences of the MicroGenie program (Microsoft available from Beckman, München).

3. RESULTS AND DISCUSSION

Two of about 3000 clones of *E. coli* JM101, transformed with pUC18 bearing Sau3A RF from *C. perfringens* A99-DNA, expressed sialidase activity as determined with MU-Neu5Ac. The responsible inserts have a size of 2.1 kb which could be reduced by digestion with *Hin*dIII (in the direction of the promotor) and Bal31 (in the direction of the C-terminus of the encoded protein) to a 1.4 kb fragment, which still encoded an active sialidase. During insert reduction, the expressed enzyme protein remained constant in size, indicating that the procedure had not influenced the structural gene.

Subcloning of the reduced insert in pUC18 and pUC19 demonstrated that the expression is independent of the lacZ′ promotor of the vector.

A DNA probe (17 nucleotides) representing the N-terminus of the cloned sialidase hybridizes with the insert and with 2.1 kb Sau3A fragments of DNA from both *C. perfringens* strains, but not with restricted *E. coli* DNA.

Fig.1 presents the restriction map of the reduced insert and the strategy for the subcloning of *Rsa*I- or Bal31-reduced insert fragments in M13 vectors for subsequent sequencing. The strategy resulted in high diversity of M13 subclones which frequently overlapped, thus assuring us of the determined sequence. The nucleotide sequence and the predicted amino acid sequence are depicted in fig.2. The nucleotide composition shows a low mol% G+C of 29 confirming its origin from *C. perfringens*, a species with a genomic mol% G+C of 24–27 [9]. The region upstream from ATG exhibits dyad symmetry between −82 and −119 including a TATAAT and a CAAT box which may represent the promotor. The Shine-Dalgarno sequence (−9

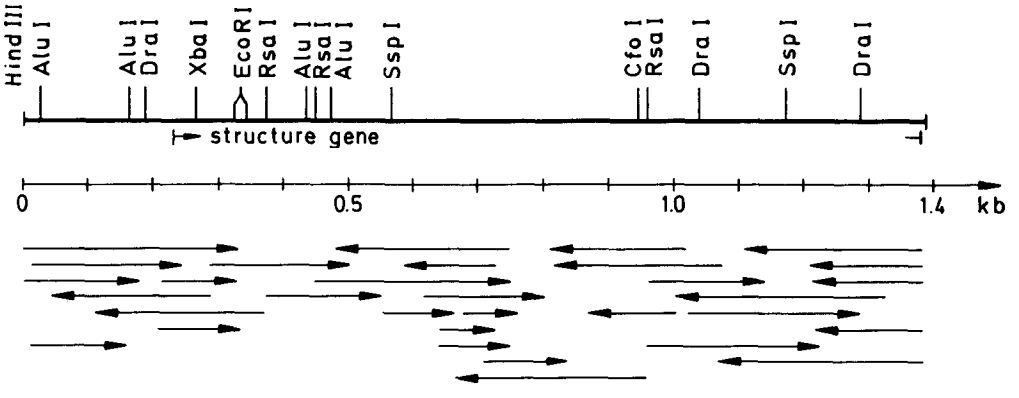

Fig.1. Restriction map of a *Clostridium perfringens* DNA fragment after maximum reduction but retaining sialidase activity in *E. coli*. Arrows indicate the direction and length of sequences obtained from insert fragment subclones.

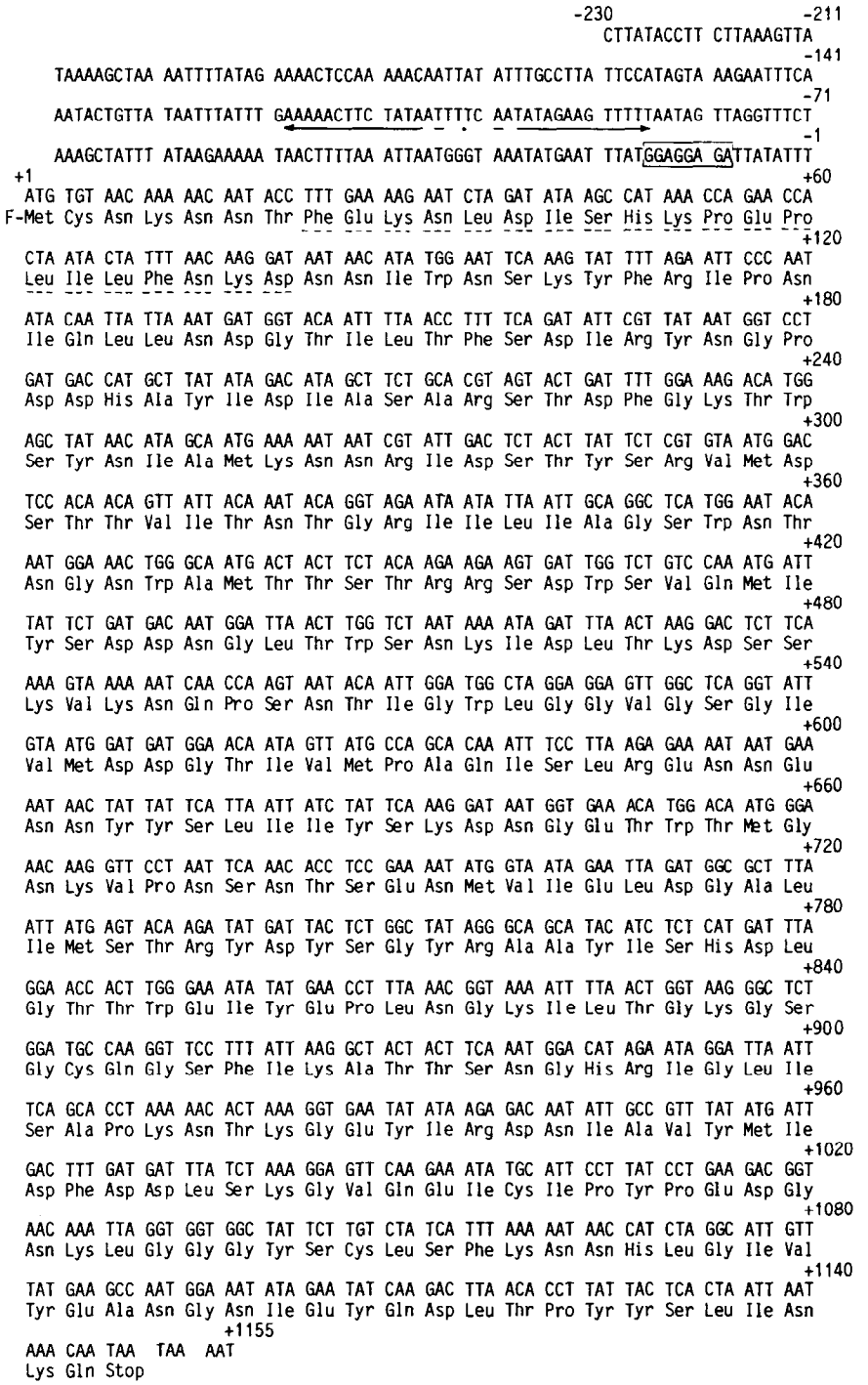

Fig.2. The nucleotide sequence and the predicted amino acid sequence of a *Clostridium perfringens* DNA fragment encoding an active sialidase, as confirmed by gene expression in *E. coli*. The arrows denote a region with dyad symmetry probably representing the promotor. The predicted Shine-Dalgarno sequence is framed. The amino acid sequence obtained from N-terminal sequencing of the expressed enzyme is indicated by a broken line.

to −16) is comparable with the ribosome-binding site of *C. thermocellum* cel C gene [10], *C. tetani* toxin gene [11], or *C. pasteurianum* Mo-pterin binding protein gene [12].

The sequence of 20 amino acids obtained from the N-terminus of the expressed sialidase protein was found to be encoded by the insert. Although three ATG codons are in correct reading frame, the ATG nearest to the known amino acid sequence probably represents the start codon which typically follows a Shine-Dalgarno sequence. The six amino acids between this ATG and the protein sequence do not show the common hydrophobicity of a leader. This indicates the absence of a leader including the possibility that the sequence obtained by Edman degradation does not represent the real N-terminus, as the enzyme protein may have been partially degraded during purification.

The predicted insert region encoding a 42770 Da protein of 382 amino acids shows no highly significant homology with sequences presented by the MicroGenie program. Surprisingly, this was also valid for virus sialidase genes [13]. No homology was found with the N-terminal gene regions of *V. cholerae* sialidase [14], and *C. sordellii* sialidase (unpublished). This indicates that the N-terminus of bacterial sialidases is not highly conserved.

This work presents the rare case that the existence of an as yet unknown enzyme is disclosed after cloning and expression of the corresponding gene; after this finding a second sialidase with low activity as possible product of this gene could be detected in *C. perfringens* culture medium (work in progress). To our knowledge, these results represent the first description of a complete sequence of a bacterial sialidase gene.

Acknowledgements: We are grateful to Werner Liesack, Dr Guido Krupp and Professor Erko Stackebrandt for DNA probe synthesis and help with the MicroGenie program. We thank Dr Rainer Hobrecht for the human isolate, Marzog El Madani for technical assistance and Elfriede Schauer for secretarial support. This work was supported by Deutsche Forschungsgemeinschaft grant Scha 202/13-1.

REFERENCES

[1] Schauer, R. ed. (1982) Sialic Acids, Cell Biol. Monogr. 10, Springer, Vienna.
[2] Müller, H.E. (1974) Behring Inst. Mitt. 55, 34–56.
[3] Nees, S., Veh, R.W. and Schauer, R. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1027–1042.
[4] Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
[5] Marmur, I. (1961) J. Mol. Biol. 3, 208–218.
[6] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, NY.
[7] Eckerskorn, C., Mewes, W., Goretzki, H. and Lottspeich, F. (1988) Eur. J. Biochem., in press.
[8] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[9] Cato, E.P., George, W.L. and Finegold, S.M. (1986) in: Bergey's Manual of Systematic Bacteriology (Sneath, P.H.A. et al. eds) vol.2, pp.1141–1200, Williams and Wilkins, Baltimore.
[10] Schwarz, W.H., Schimming, S., Rücknagel, K.P., Burgschwaiger, S., Kreil, G. and Staudenbauer, W.L. (1988) Gene 63, 23–30.
[11] Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hudel, M., Habermann, E. and Niemann, H. (1986) EMBO J. 5, 2495–2502.
[12] Hinton, S.M., Slaughter, C., Eisner, W. and Fisher, T. (1987) Gene 54, 211–219.
[13] Jorgensen, E.D., Collins, P.L. and Lomedico, P.T. (1987) Virology 156, 12–24.
[14] Vimr, E.R., Lawrisuk, L., Galen, J. and Kaper, J.B. (1988) J. Bacteriol. 170, 1495–1504.