Proteolytic Processing and Oligomerization of Bacteriophage-derived Endosialidases*

Received for publication, November 26, 2002, and in revised form, January 21, 2003
Published, JBC Papers in Press, January 29, 2003, DOI 10.1074/jbc.M212048200

Martina Mühlenhoff‡§, Katharina Stummeyer‡, Melanie Grove‡, Markus Sauerborn¶¶, and Rita Gerardy-Schahn‡

From the ‡Abteilung Zelluläre Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Carl-Neuberg-Straße 1, 30625 Hannover and the ¶Institut für Mikro- und Molekularbiologie, Justus-Liebig-Universität Giessen, Ludwigstrasse 23, D-35390 Giessen, Germany

Bacteriophages infecting the neuroinvasive pathogen Escherichia coli K1 require an endosialidase to penetrate the polysialic acid capsule of the host. Sequence information is available for the endosialidases endoNE, endoNF, and endoN63D of the K1-specific phages φK1E, φK1F, and 63D, respectively. The cloned sequences share a highly conserved catalytic domain but differ in the length of the N- and C-terminal parts. Although the expression of active recombinant enzyme succeeded in the case of endoNE, it failed for endoNF. Protein alignments of all three endosialidase sequences gave rise to the assumption that inactivity of the cloned endoNF is caused by a C-terminal truncation. By reinvestigation of the respective gene locus in the φK1F genome, we identified an extended open reading frame of 3195 bp, encoding a 119-kDa protein. Full-length endoNF contains the C-terminal domain conserved in all endosialidases, which may act as an intramolecular chaperone. Comparative studies carried out with endoNE and endoNF demonstrate that endosialidases are proteolytically processed, releasing the C-terminal domain. Using a mutational approach in combination with protein analytical techniques we demonstrate that (i) the C-terminal domain is a common feature of endosialidases and other tail fiber proteins; (ii) the integrity of the C-terminal domain and its presence in the nascent protein are crucial for the formation of active enzymes; (iii) proteolytic processing is not essential for enzymatic activity; and (iv) functional folding is a prerequisite for trimerization of endoNF.

The capsule of the neuroinvasive pathogen Escherichia coli K1 (E. coli K1), causing meningitis and sepsis in neonates, is composed of α2,8-linked sialic acid with up to 200 residues (polysialic acid; polySia) (1, 2). This large homopolymer is an important virulence factor protecting the bacterium from the immune system, but it also serves as an attachment site for lytic bacteriophages. Several coliphages specific for E. coli K1 have been isolated from sewage samples (3–6). Interestingly, administration of anti-K1 phages to E. coli K1-infected mice and chicken was shown to prevent septicemia and meningitis-like infections, demonstrating their potential for antibacterial therapy (5, 7). Anti-K1 phages are lytic linear double-stranded DNA viruses of different morphology. Although most of the isolated phages have an isometric head and a short tail, some are provided with a long and flexible tail (6) similar to that of bacteriophage λ. Common to all anti-K1 phages is an endosialidase (endo-N-acetylneuraminidase; endoN) highly specific for α2,8-linked polySia (8–10) which was observed as a tail fiber protein (6, 11). In contrast to exosialidases, phage-borne endosialidases require oligomers of up to eight α2,8-linked sialic acid residues for binding, and distinct cleavage patterns have been observed for individual enzymes isolated from different anti-K1 phages. Although sialyl dimers were found as the main product for some endosialidases (6), most enzymes release sialyl oligomers of three to seven residues (12).

Because of high substrate specificity, endoN is of particular importance for studying the functional role of polySia in different organisms. In mammals, polySia is found exclusively as a dynamically regulated post-translational modification of the neural cell adhesion molecule (13). The presence of the large polyanionic carbohydrate moiety attenuates the binding properties of the neural cell adhesion molecule and promotes plasticity of cell-cell interactions during cell migration and axonal pathfinding (14). Specific removal of polySia by endoN demonstrated the importance of polySia for synaptic activity involved in learning and memory (15, 16) and its role in promoting metastasis and invasion of polySia-positive tumor cells (17).

In contrast to bacterial and viral exosialidases, little is known about the structure of the phage-derived endosialidases. SDS-resistant trimers, composed of three identical subunits of 105 kDa each, have been described for endoNF of coliphage φK1F (9). For endoNE of φK1E, hetero-oligomeric structures of 208 and 325 kDa were described, composed of endoNE and a not as yet identified 38-kDa protein (8, 18). So far, genes encoding endosialidases have been cloned from four different phages: φK1F (11), φK1E (18, 19), 63D (GenBank accession number AB015437), and the dual specificity phage φK1-5 that can infect E. coli K1 and K5 (20). However, recombinant expression of active enzyme has been reported only in the case of endoNE (18, 19).

Cloning of endoNE revealed that the gene encodes a 90-kDa protein, whereas for the purified phage-borne and the recombinantly expressed endoNE an apparent molecular mass between 74 and 76 kDa was observed (8, 18, 19), indicating post-translational processing. In contrast, no discrepancy be-
enzymatically active part and an 18-kDa C-terminal fragment translation product is proteolytically cleaved into a 101-kDa active endoNF was recombinantly expressed. The primary endoNF encoding a 119-kDa protein, and for the first time, activity was detected for the recombinantly expressed 102-kDa protein (11).

Here we describe the isolation of the full-length gene of endoNF. However, in the latter case no enzymatic activity was detected for the recombinantly expressed 102-kDa protein (11).

Restriction sites introduced for subcloning are underlined.

| Mutation       | Nucleotide exchange | Primer pair       |
|----------------|---------------------|------------------|
| **endoNE-P702A** | 236 CTC → GCG      | s 5′-TCAGGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |
| **endoNE-T705A** | 231 ACT → GCT      | as 5′-GCACGATTAAGGAGCTATCCATCTTCAGATGC-3′ |
| **endoNE-S706A** | 2115 TCA → GCA     | s 5′-CCCTGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |
| **endoNE-H747A** | 2239 CAC → GCC     | as 5′-ACGTTGACATTAAGGAGCTCTTCAGATGC-3′ |
| **endoNF-P907A** | 2719 CCT → GCA     | s 5′-CCCTGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |
| **endoNF-T910A** | 2729 ACT → GCT     | s 5′-CCCTGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |
| **endoNF-S911A** | 2739 TCT → GCT     | s 5′-CCCTGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |
| **endoNF-H954A** | 2866 CAC → GCC     | s 5′-CCCTGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |
| **endoNF-L903L** | 2707 CTC → GTA     | s 5′-CCCTGCTAGTTACCAAGAGCTCTTCAGATGC-3′ |

**TABLE I**

Oligonucleotides used for the introduction of amino acid exchanges and C-terminal truncations

Single amino acid exchanges in endoNE and endoNF were introduced by site-directed mutagenesis in a PCR approach using the sense (s) and antisense (as) primers given in the upper part of this table. In the case of endoNF-L903L, an AvrII restriction site was introduced by a silent mutation. C-terminal truncated forms of endoNE and endoNF were generated by PCR using the primer pairs given in the lower part of this table.

Restriction sites introduced for subcloning are underlined.

| Mutation       | Nucleotide exchange | Primer pair       |
|----------------|---------------------|------------------|
| **endoNE wild type** |                    | s 5′-CATGCCATGGTTATCCAAAGCATCAGCTTCTTC-3′ |
| **endoNE Δ38**   |                     | as 5′-GGGACGATCTTAAACTGCGAGTCGCT-3′ |
| **endoNE Δ103**  |                     | s 5′-CATGCCATGGTTATCCAAAGCATCAGCTTCTTC-3′ |
| **endoNE Δ104**  |                     | as 5′-GGGACGATCTTAAACTGCGAGTCGCT-3′ |
| **endoNE Δ105**  |                     | s 5′-CATGCCATGGTTATCCAAAGCATCAGCTTCTTC-3′ |
| **endoNE Δ106**  |                     | as 5′-GGGACGATCTTAAACTGCGAGTCGCT-3′ |
| **endoNE Δ107**  |                     | s 5′-CATGCCATGGTTATCCAAAGCATCAGCTTCTTC-3′ |
| **endoNE Δ108**  |                     | as 5′-GGGACGATCTTAAACTGCGAGTCGCT-3′ |
| **endoNE Δ109**  |                     | s 5′-CATGCCATGGTTATCCAAAGCATCAGCTTCTTC-3′ |
| **endoNE Δ153**  |                     | as 5′-GGGACGATCTTAAACTGCGAGTCGCT-3′ |
| **endoNF Δ154**  |                     | s 5′-GGGACGATCTTAAACTGCGAGTCGCT-3′ |

Maturation of Bacteriophage Endosialidases

Maturation of Bacteriophage Endosialidases

Here we describe the isolation of the full-length gene of endoNF. However, in the latter case no enzymatic activity was detected for the recombinantly expressed 102-kDa protein (11).

Here we describe the isolation of the full-length gene of endoNF. However, in the latter case no enzymatic activity was detected for the recombinantly expressed 102-kDa protein (11).

Here we describe the isolation of the full-length gene of endoNF. However, in the latter case no enzymatic activity was detected for the recombinantly expressed 102-kDa protein (11).
propagation of bacteriophages $\phi$K1E and $\phi$K1F (3). Phage DNA was isolated from purified phage particles as described previously (18).

Cloning of EndoNF—Based on the endoNF sequence published by Petter and Vimr (11) a primer binding nucleotides 2652–2669 (5'-TAACGATAACGTCACTGC-3') was generated and used for direct sequencing of purified $\phi$9278K1F phage DNA. An in-frame stop codon was found by primer walking using a primer binding nucleotides 2848–2865 (5'-GATGCTCGTATTCACTTC-3'). The newly identified 3195-bp open reading frame (ORF) was confirmed by sequencing. Full-length endoNF (nucleotides 1–3192 lacking the stop codon) was amplified by PCR using purified phage DNA as a template and the primer pair MM97/MM98 (5'-CGGGATCCATGTCCACGATTACACAATTTGCTAGCAGAAGAAGGAAAGTCCATACGTTATTTGCTTAGACGAGGATG-3', containing BamHI and XhoI restriction sites (underlined) on the 5'- and 3'-ends, respectively, and ligated into pET23a. The resulting expression plasmid pT7-endoNF-His encodes full-length endoNF with an N-terminal T7 epitope tag (MASMTGGQQMG) and a C-terminal His$_6$ tag. The identity of this construct was confirmed by sequencing.

Expression Plasmids—Using purified $\phi$K1E phage DNA as a template, a 2.4-kb PCR fragment containing full-length endoNE with a C-terminal factor Xa cleavage site was amplified with primers AH18 (5'-GCGGATCCATGATTCAAAGACTAGG-3') and AH26b (5'-GAGGATCCTTAACTGATTTTATTAGTGGC-3'), including BamHI restriction sites (underlined) and the coding sequence for a factor Xa cleavage site (italics) in the case of AH26b. The PCR product was digested with BamHI and ligated in the BamHI/BglII sites in front of the His$_6$ tag of the pQE16 vector (Qiagen). The 2.4-kb BamHI/HindIII fragment (encompassing full-length endoNE with a C-terminal factor Xa cleavage site followed by a His$_6$ tag) obtained from this plasmid was then subcloned into pET23d. For coexpression experiments, the 3'-end of the endoNE gene encoding amino acids 706–811 was amplified with primers MM72 (5'-CATGCCATGGTTATTCAAAGACTAGGTTCTTC-3') and AH25 (5'-GAGGATCCTTAACTGATTTTATTAG-3'), digested with NcoI and BamHI, and ligated into pET23d. The resulting expression plasmid pET-endoNE encoding full-length endoNE was constructed in pET23d. Starting with purified $\phi$9278K1E phage DNA, a 2.4-kb fragment was amplified by PCR with the primers MM72 and AH25, digested with NcoI and BamHI, and ligated into pET23d.
plified by PCR using the primers given in Table I, and the PCR product was ligated into the NdeI/XhoI fragment of pET23a. The 492-bp BglII/BamHI fragment of this construct, containing the T7 promoter, the ribosomal binding site, and the His6-tagged C-terminal domain of endoNE, was then cloned in pMCL200 with a PA15 origin of replication. The identity of all constructs was confirmed by sequencing.

Site-directed Mutagenesis and Generation of C-terminal Truncated and Chimeric Endosialidases—Site-directed mutagenesis was performed by PCR using the QuikChange site-directed mutagenesis kit (Stratagene) following the manufacturer’s guidelines with the primers given in Table I. For the generation of chimeric endosialidases an AvrII restriction site was introduced in pT7-endoNF-His by the silent mutation C2709A, maintaining a triplet coding for Leu-903 (primers used for PCR are given in Table I). EndoNE contains an endogenous AvrII site at the corresponding position, and this site was used to exchange the gene fragments encoding the C-terminal domains. C-terminal truncated constructs were generated by PCR using the primers given in Table I. PCR products were ligated into the NcoI and BamHI sites of pET23d and BamHI and XhoI sites of pET23a for endoNE and endoNF constructs, respectively. The identity of all constructs was confirmed by sequencing.

Expression of Recombinant Endosialidases—Freshly transformed E. coli BL21(DE3) were cultivated at 37 °C in Luria-Bertani medium containing 200 μg/ml carbenicillin. At an absorbance (A600) of 0.6 expression was induced by adding 1 mM isopropyl-1-thio-β-D-galactopyranoside, and bacteria were harvested 2 h after induction. Cells were resuspended in 50 mM Tris-HCl, pH 8.0, and 2 mM EDTA and incubated for 15 min at 30 °C in the presence of 100 μg/ml lysozyme. Bacteria were lysed by sonication, and soluble fractions were obtained after centrifugation (12,000 × g, 15 min).

Affinity Purification of Recombinant EndoNF—Bacteria expressing T7-endoNF-His were harvested 2 h after induction, resuspended in

FIG. 2. Proteolytic cleavage of endosialidases. A, endoNE and endoNF with N-terminal T7 and C-terminal His6 tags were expressed in E. coli BL21(DE3). 2 h after induction soluble fractions of bacterial lysates were analyzed by 14% ProSieve SDS-PAGE and immunoblotting using anti-His6 tag antibody or a combination of anti-His6 and anti-T7 antibodies. B, schematic representation of fragments obtained after proteolytic processing. Molecular masses of N-terminal and C-terminal fragments were calculated for enzymes with T7 and His6 tags. The amino acid sequences around the cleavage sites are shown on the right. For endoNF we propose cleavage after the conserved Ser-911 followed by spontaneous degradation of Asn-912 (for details, see “Results”). C, the soluble fraction of bacterial lysates containing T7- and His6-tagged endoNF was passed through a Ni2+-chelating column. The flow-through (lanes 1, 4, and 7) and all fractions of the eluted protein peak were analyzed by 12% ProSieve SDS-PAGE and immunoblotting with the indicated antibodies. The first (lanes 2, 5, and 8) and the last (lanes 3, 6, and 9) peak fraction are shown exemplarily. Bands corresponding to full-length protein and N- and C-terminal fragments are indicated with arrows. D, endosialidase activity was analyzed in the flow-through of the Ni2+-chelating column, containing exclusively the N-terminal fragment, and in the last peak fraction, containing only the C-terminal fragment.
binding buffer (50 mM Na₂HPO₄, NaH₂PO₄, pH 8.0, 500 mM NaCl) and lysed by sonication. The soluble fraction was applied to a 5-ml Ni²⁺/H⁺-loaded HiTrap chelating column (Amersham Biosciences) according to the manufacturer's instructions. After washing with 25 ml of binding buffer containing 10% glycerol, the protein was eluted with a linear imidazol gradient (0–500 mM imidazol in binding buffer containing 10% glycerol). For purification of the T7 epitope tagged N-terminal fragment the flow through of the chelating column was adsorbed to 1 ml of T7 tag antibody agarose (Novagen) according to the manufacturer's protocol. Protein was eluted with 100 mM glycine, pH 2.7, and the fractions were neutralized with 1 M Tris, pH 9.0.

**N-terminal Amino Acid Analysis**—Affinity-purified His₅-tagged C-terminal fragments of endoNE and endoNF were separated by 10% SDS-PAGE and immunoblotting. A polyclonal anti-endoNE guinea pig serum was used for detection. A faint 75 kDa band that is also visible in the first lane showing lysate of mock transformed bacteria is visualized by cross-reactivity of the polyclonal serum. Bands corresponding to full-length and cleaved N-terminal endoNE are indicated with arrows. Bands containing an N-terminal T7 tag were analyzed by 10% SDS-PAGE and immunoblotting using an anti-T7 antibody.

**RESULTS**

**Cloning of the Full-length Endosialidase Gene from Bacteriophage øK1F**—Comparison of the protein sequences of endosialidases from bacteriophages øK1E (18, 19), øK1F (11), and ø63D (accession no. AB015437) revealed an extended region of about 650 amino acids with high sequence similarity (indicated as gray boxes in Fig. 1A). This core domain shows 81% amino acid identity between endoNE and endoNF, 52% identity between endoNF and ø63D, and 50% identity between en-
Maturation of Bacteriophage Endosialidases

12639

doNE and endoN63D. However, the N- and C-terminal parts differ in length, and the N-terminal domains were found to possess only very low sequence similarity. Of all cloned endosialidases, endoNF has the shortest C-terminal part, lacking the amino acid clusters shared by endoNE and endoN63D (indicated as black boxes in Fig. 1A). Moreover, no enzymatic activity was observed for the recombinantly expressed endoNF (11). We knew from earlier studies that endoNE lacking the last 38 amino acids is completely inactive (18) and asked whether the recombinantly expressed endoNF was inactive because of a truncated C terminus. By direct sequencing of DNA isolated from bacteriophage φK1F, we reinvestigated the size and sequence of the endoNF gene. In contrast to the published data describing a 2783-bp ORF (11), we identified a 3195-bp ORF. Although the first 2740 bp are identical in both sequences, the insertion of two nucleotides (G2741 and G2744) that caused a premature stop codon could not be confirmed. As shown in Fig. 1B, lack of these two nucleotides resulted in an elongation of the previously described ORF by 432 bp. The newly identified ORF encodes a potential 118.9-kDa protein, containing all conserved amino acid clusters in the C-terminal part found in endoNE and endoN63D (see Fig. 1A).

Proteolytic Cleavage of Endosialidases—Recombinant endoNE and endoNF containing an N-terminal T7 epitope and a C-terminal His6 tag (see Fig. 2B) were expressed in E. coli BL21(DE3), and enzyme expression was analyzed by Western blotting. For each enzyme three bands were detected (Fig. 2A). The top bands, visible with anti-His6 antibody, indicate the expression of full-length proteins with molecular masses corresponding to the calculated mass of 93 and 121 kDa for epitope-tagged endoNE and endoNF, respectively. In addition, low molecular weight fragments of 13 (endoNE) and 19 kDa (endoNF) were detected with anti-His6 tag antibody, demonstrating that these fragments contain the C terminus. The corresponding N-terminal fragments were detected exclusively with the anti-T7 antibody. These results clearly demonstrate proteolytic cleavage of the full-length endosialidases into a small C-terminal and a large N-terminal fragment.

To identify the cleavage sites, the small fragments were isolated by affinity purification on a Ni2+-chelating column and subjected to N-terminal sequencing. Two peptide sequences (DADXXGYGISS and GerkTepVVF for endoNE and endoNF, respectively) were identified, corresponding to 707DADHKG-GISSL in endoNE and 911GerkTepVVF922 in endoNF. Based on this sequence information, the cleavage sites were located between Ser-706 and Asn-912 (endoNE) and Asp-707 and Gly-913 (endoNF), respectively (see Fig. 2B). As indicated in Fig. 1A and shown in detail in the protein alignment given in Fig. 7, the identified cleavage site is located at a similar site in both proteins with a highly conserved serine residue in front of the cleavage site (Ser-706 and Ser-911 in endoNE and endoNF, respectively). It is very likely that endoNF is also cleaved directly after Ser-911, releasing a C-terminal fragment containing Asn-912 at the N terminus. Asparagine residues, particularly in front of a glycine, are very labile and can undergo spontaneous degradation by intramolecular succinimide-forming reactions, leading to deamidation, isomerization, and peptide bond cleavage (24, 25). As a consequence, Gly-913 but not Asn-912 was the first amino acid identified by Edman degradation of the C-terminal fragment.

Enzymatic Activity Is Associated with the N-terminal Fragment—For identification of the enzymatically active fragment, bacterial lysates containing N-terminal T7- and C-terminal His6-tagged endoNF were passed through a Ni2+-chelating column to remove the His6-tagged full-length protein and the small C-terminal fragment. As shown in Fig. 2C (lane 1), the flow-through contained exclusively the N-terminal fragment of 103 kDa. Although the first fractions of the protein peak eluted from the Ni2+-chelating Sepharose contained full-length and cleaved protein (lane 2, Fig. 2C), later fractions contained exclusively the C-terminal fragment (lane 3, Fig. 2C). As shown in Fig. 2D, no enzymatic activity was detected for the purified
C-terminal fragment. In contrast, high enzymatic activity was observed in the flow-through (Fig. 2D), demonstrating that the large N-terminal fragment contains the catalytic domain. The fact that a small part of the N-terminal fragment lacking the C-terminal His6 tag was adsorbed to the Ni$^{2+}$/H$^{11001}$-chelating Sepharose and coeluted with the full-length protein can be explained by oligomerization of N-terminal domains (see below). By the formation of hetero-oligomers containing His$_{6}$-tagged full-length endoNF and the cleaved N-terminal domain, the latter fraction can be bound to the column.

The C-terminal Domain Is Essential for the Formation of an Active Protein—To examine the role of the small fragment, C-terminally truncated forms of endoNE and endoNF (for schematic representation, see Fig. 3, A and B) were expressed in E. coli BL21(DE3), and the enzymatic activity was monitored in soluble fractions of the bacterial lysates. As shown in Fig. 3, C and D, all truncated forms of endoNE and endoNF were expressed as soluble proteins with the expected molecular mass. For truncated endoNF, several degradation products became visible, indicating reduced stability of the truncated proteins. Although the C-terminal domain is dispensable after cleavage of the full-length protein, none of the truncated enzymes was active (data not shown), suggesting that the C-terminal domain plays an important role in the nascent protein. Interestingly, a short C-terminal truncation of only 38 amino acids (endoNE H$_{9004}$A38) resulted in a noncleaved protein of 86 kDa (Fig. 3C, third lane), indicating that the complete C-terminal domain is essential for proper cleavage.

To address directly whether the truncated forms can be rescued by the presence of the C-terminal fragment, truncated forms of endoNE were coexpressed with a His$_{6}$-tagged C-terminal domain encompassing amino acids 706–811 (Fig. 3A). Although both fragments, encoded on two different plasmids, were expressed and detected in the soluble fraction of bacterial lysates, coexpression of the C-terminal fragment could not restore the activity of the truncated forms (data not shown). This result clearly demonstrates that in a first step the C-terminal domain must be part of the primary translation product to fulfill its function.

Expression of Chimeric Endosialidases—Assuming a similar function for the C-terminal domain of all endosialidases, we next analyzed whether the C-terminal part of endoNE can be replaced by the corresponding domain of endoNF and vice versa (Fig. 4A). Both chimeras, endoNE-F and endoNF-E, were expressed in E. coli BL21(DE3) with an N-terminal T7 and a C-terminal His$_{6}$ tag. As shown in Fig. 4B, chimeric proteins...
were proteolytically processed, and fragments of 78 and 19 kDa (endoNE-F) and 103 and 13 kDa (endoNF-E) were released, indicating structural integrity of the cleavage site. Although endoNE-F showed partial activity (70% activity of wild type endoNE), no activity was detected for endoNF-E (see Fig. 4C).

Because the C-terminal domain of endoNE is 49 amino acids shorter than the corresponding domain in endoNF, the positioning of functional important amino acid residues may not fulfill the requirements of endoNF.

Proteolytic Cleavage Is Not Essential for Enzymatic Activity—To investigate whether proteolytic processing is required to generate active endosialidases, several single amino acid exchanges were introduced in endoNE and endoNF by site-directed mutagenesis. After expression in E. coli BL21(DE3), cleavage and enzymatic activity were analyzed in bacterial lysates. The results are summarized in Fig. 5, for endoNE and in Fig. 5, for endoNF. Compared with the corresponding wild type proteins, the amino acid exchanges P702A and P907A in endoNE and endoNF, respectively, had no significant effect on proteolytic cleavage or enzymatic activity. Similar results were obtained for endoNF-T910A, but the activity of the corresponding endoNE mutant (T705A) was reduced to 24% of the wild type activity. In contrast, exchange of the serine residue that is part of the cleavage site (S706A and S911A in endoNE and endoNF, respectively) prevented proteolytic processing (Fig. 5, and B). The expressed full-length proteins were active (Fig. 5, D and C), demonstrating that cleavage of the C-terminal part is not a prerequisite for the formation of active enzymes.

If, however, a conserved histidine residue located in the center of the C-terminal domain (see Fig. 7) was replaced by alanine, proteolytic cleavage was prevented, and enzymatic activity was completely abolished (see mutants H747A and H954A in Fig. 5). The expressed full-length proteins were detected predominantly in the insoluble fraction, indicating accumulation of misfolded protein. These latter mutants demonstrate an involvement of the C-terminal domain in the folding of endosialidases and in combination with the previous data suggest a potential chaperone function for the C-terminal domain.

The C-terminal Domain of EndoNF Is Essential for Trimerization—For the phage-borne endoNF, SDS-resistant trimers were observed (9) which were also detectable for the recombinantly expressed enzyme (lane 2, Fig. 6A). If samples were separated by SDS-PAGE omitting the boiling step, the appearance of a high molecular mass band of about 300 kDa and the disappearance of the 103-kDa N-terminal fragment were observed, confirming the formation of a trimer. The high molecular mass band was detected using anti-T7, but not with anti-His6 tag antibody, demonstrating that the C-terminal domain is not part of the complex. Nevertheless, neither of the C-terminal truncated forms (Δ153 and Δ154) was able to assemble in an oligomeric complex, and only monomeric proteins were detected (lanes 12 and 14 in Fig. 6A). Similar results were obtained for the endoNE-F and endoNF-E chimeras (see Fig. 6B). Although SDS-resistant complexes have been described for the phage-associated endoNE (8, 18), no oligomerization was observed for the recombinant protein (see lane 8 in Fig. 6B). In contrast to the homotrimer formed by endoNF, endoNE was reported to form hetero-oligomeric complexes together.

![Fig. 6. Trimer formation of endoNF.](http://www.jbc.org/Downloaded from)
with a 38-kDa protein (8, 18). The absence of this protein in the recombinant enzyme fraction (data not shown) might explain the lack of SDS-resistant complexes. Fusion of the C-terminal part of endoNF to the catalytic domain of endoNE did not change the monomeric state of the protein. Interestingly, the enzymatically inactive endoNF-E chimera lost the ability to assemble into trimers. These results confirm the pivotal role of the C-terminal domain for folding and/or oligomerization.

The next question was whether the trimer provides the active form of endoNF. Therefore, catalytically active and inactive forms of endoNF were analyzed by SDS-PAGE as shown in Fig. 6A. Trimer formation was observed for all endoNF mutants with enzymatic activity (Fig. 6A). For endoNF-T910A a slightly increased expression of full-length protein was observed, and part of the unprocessed protein was detected in the complex with anti-His6 tag antibody (see lane 6 in Fig. 6A, lower panel). The formation of mixed trimers containing one or two full-length proteins is additionally indicated by the appearance of several bands above 300 kDa.

In the case of endoNF-S911A, an active mutant lacking the ability for proteolytic cleavage, trimers composed of full-length protein were formed. Although large amounts of monomeric 120-kDa full-length protein were detected in the heat-treated sample (lane 7 in Fig. 6A), the corresponding band is drastically reduced in the nonboiled sample, and a trimer detected with anti-His6 antibody appears. An additional band corresponding to a dimer of 240 kDa became visible, suggesting reduced complex stability for trimers containing full-length protein, exclusively. The enzymatically inactive mutant endoNF-H954A, which is also expressed as a full-length protein, lost its capacity to oligomerize (Fig. 6A, lane 10). In combination, these data demonstrate that trimerization indicates correctly folded and thus active endoNF.

**DISCUSSION**

Bacteriophages specific for encapsulated bacteria have to penetrate the capsular polysaccharide before infection, and therefore the host range is determined by the presence of specific polysaccharide depolymerases. In this study, we have characterized the endosialidases of the E. coli K1-specific phages φK1E (X78310), φK1-5 (AF322019), φK1F (AJ505988), and 63D (AB015437), the neck appendage protein gp12 of the B. subtilis phage GA-1 (NC_002649), the L-shaped tail fiber protein of coli phage T5 (X69480), the K5 lyases of the E. coli K5-specific phages φK5 (Y10025) and φK5-1 (AF222019), and the K5 lyase of E. coli K5 (X96495). Amino acid numbers are shown on the left. Residues that are identical in all sequences are shaded with black. Residues that are conserved in five or more of the nine sequences are enclosed in boxes, showing identical and similar amino acids in bold. The position of the cleavage site identified for endoNE and endoNF is indicated by an arrowhead, and the positions of amino acids that were exchanged in this study by alanine are marked with an asterisk. The multiple sequence alignment was performed using the MultAlin program (37). Alignment of the N-terminal parts of the proteins is not included in this figure because of the very low similarity among endosialidases, gp12, L-shaped tail fiber protein (LTF), and lyases.

![Fig. 7](http://www.jbc.org)
A BLAST (26) search performed with the newly identified C-terminal part of endoNF identified similarities to the C-terminal parts of several tail fiber proteins of other bacteriophages (Fig. 7). The alignment includes the C-terminal domains of the neck appendage protein gp12 of the Bacillus subtilis phage GA-1 (EMBL accession number X96887.2), the L-shaped tail fiber protein of coliphage T5 (27), the K5-specific lyases of coliphages φK1-5 (20) and φK5 (28), and the corresponding K5 lyase of E. coli K5 (29). The protein alignment reveals that the cleavage site identified for endoNE and endoNF (marked with an arrowhead in Fig. 7) is highly conserved, implying that proteolytic processing might be a common feature of all proteins listed in Fig. 7. The serine residue that was shown in this study to be essential for proper cleavage (Ser-706 and Ser-911 in endoNE and endoNF, respectively) is present in all sequences, and the aspartic acid after the cleavage site is highly conserved with the single exception of endoNF, where an exchange by asparagine is found.

Data reported on the endosialidase isolated from the K1-specific phage 63D confirm the presence of a common cleavage site. The activity of the phage-derived enzyme was reported to be associated with a 90-kDa protein (10), although the corresponding gene (GenBank accession number AB015437) encodes a protein of 108.3 kDa. Cleavage after Ser-852 at the proposed cleavage site would release fragments of 93 and 15 kDa, suggesting that in endoN63D cleavage occurs at the same position as in endoNE and endoNF.

The L-shaped tail fiber protein of the T5 phage accelerates adsorption to the lipopolysaccharide of E. coli F (30). The ltf gene codes for a 148-kDa protein (27), and in agreement with the cleavage site proposed here, processing of a 150-kDa precursor protein into a 125-kDa protein was reported (30). The presence of a common cleavage site is also supported by results obtained for the recombinantly expressed lyase of E. coli K5 which depolymerizes the capsular polysaccharide composed of repeating units of 4-linked α-N-acetylgalactosamine and β-glucuronic acid (29, 31). In addition to full-length protein of 89 kDa, a 70-kDa fragment lacking the C terminus was observed (29), corresponding to the calculated fragment sizes of 73 and 16.7 kDa. Similar to our results obtained for endoNE lacking the last 38 amino acids, no activity was found for the E. coli K5 lyase truncated by only 15 amino acids (29). Both truncations affect homologous stretches at the very C terminus, and, obviously, the presence of the complete C-terminal domain is essential for proper folding and subsequent cleavage of these proteins.

Whether proteolytic maturation occurs by self-cleavage or catalyzed by E. coli proteases will need further analysis. Coupled in vitro transcription-translation of the endoNE gene in an E. coli lysate system resulted in an unprocessed translation product (19), indicating that the in vitro system might be insufficient in providing a proper folding environment.

Proteolytic processing and complex formation of a phage tail protein were demonstrated recently in molecular detail for the tail lysozyme (gp5) of bacteriophage T4 (32). Cleavage was observed between Ser-351 and Ala-352 when gp5 was incorporated into the phage baseplate or stored at high concentration. Although the released N-terminal part contains the catalytic domain, the C-terminal fragment remains part of the phage particle forming the tip of the cell-puncturing device. Similar to the C-terminal part of endoNF, the C-terminal fragment of gp5 was reported to be necessary for trimerization of three copies of a hetero-oligomeric complex composed of the N- and C-terminal fragment of gp5 and gp27 (32). In contrast to gp5, the C-terminal domain of endoNF is not associated with the endosialidase complex and can be easily separated.

In the case of endoNE, hetero-oligomeric complexes with a tightly associated 38-kDa protein were observed for the phage-derived enzyme (8, 18). This protein might act as an assembly protein and in its absence, no SDS-resistant complexes were observed for the recombinantly expressed protein. Interestingly, the catalytic and the C-terminal domain of recombinant endoNE remain associated after cleavage, 2 and dissociation might occur after assembling with the 38-kDa protein. During the preparation of this manuscript, Leggate et al. (33) also reported on proteolytic cleavage of endoNE. In line with our results, no SDS-resistant oligomers were detected for a recombinantly expressed glutathione S-transferase fusion protein of endoNE. However, under nondenaturing conditions, a complex above 250 kDa became visible, indicating trimer formation. Similar to our results obtained for C-terminally truncated endoNF, only monomeric protein was detectable for a 105-amino acid truncated enzyme (33), supporting our finding that the C-terminal domain of endosialidases is essential for oligomerization of the catalytic domain.

The finding that several tail fiber proteins with different functions contain the same conserved C-terminal domain argues for horizontal transfer of the corresponding gene fragment. Evidence for horizontal transfer among tail fiber genes has been reported (34), and the endosialidase gene of K1F is another example for a mosaic gene composed of at least three different parts. Although the catalytic domain shows no similarities to other known phage proteins, significant similarities have been identified between the first 175 N-terminal amino acids and the N-terminal parts of the tail fiber protein gp17 of the coliphages T7 and T3 and the yersiniophage φYeO3 specific for Yersinia enterocolitica serotype O:3 (35). The N-terminal part of gp17 attaches the protein to the phage tail (36), and in endoNF this part might have a similar function. By identification of a C-terminal domain shared by several tail fiber proteins of distinct phases (B. subtilis phage GA-1 and coliphages T5), we provide further evidence for horizontal transfer of tail fiber gene fragments across different groups of tailed bacteriophages.

Acknowledgments—We thank Dr. Dietmar Linder for protein sequencing and Drs. Françoise Routier and Joe Tiralongo for critical remarks on the manuscript.

REFERENCES

1. Rohr, T. E., and Troy, F. A. (1980) J. Biol. Chem. 255, 2332–2342
2. Pelkonen, S., Hayrinen, J., and Finne, J. (1988) J. Bacteriol. 170, 2646–2653
3. Gross, R. J., Cheasty, T., and Rowe, B. (1977) J. Clin. Microbiol. 6, 548–550
4. Kwaistowski, B., Boscheck, B., Thiele, H., and Stirm, S. (1982) J. Virol. 43, 697–704
5. Smith, H. W., and Huggins, M. B. (1982) J. Gen. Microbiol. 128, 307–318
6. Miyake, K., Muraki, T., Hattori, R., Machida, Y., Watanabe, M., Kawase, M., Yoshida, Y., and Iijima, S. (1997) J. Gen. Microbiol. 84, 90–93
7. Barrow, P., Lovell, M., and Berchieri, A., Jr. (1998) Clin. Diag. Lab. Immunol. 5, 294–298
8. Tomlinson, S., and Taylor, P. W. (1985) J. Virol. 55, 374–378
9. Halleenbeck, P. C., Vinar, E. E., Yu, F., Bassler, B., and Troy, F. A. (1987) J. Biol. Chem. 262, 3553–3561
10. Machida, Y., Miyake, K., Hattori, K., Yamamoto, S., Kawase, M., and Iijima, S. (2000) FEMS Microbiol. Lett. 182, 333–337
11. Peter, J. G., and Vimar, E. R. (1993) J. Bacteriol. 175, 4354–4363
12. Pelkonen, S., Pelkonen, J., and Finne, J. (1989) J. Virol. 63, 4409–4416
13. Mühlenhoff, M., Eckhardt, M., and Gerardy-Schahn, R. (1996) Curr. Opin. Struct. Biol. 6, 558–564
14. Rutsikhauser, U., and Landmesser, L. (1996) Trends Neurosci. 19, 422–427
15. Becker, C. G., Artola, A., Gerardy-Schahn, R., Becker, T., Welzl, H., and Schaefer, M. (1996) J. Neurosci. Res. 45, 143–152
16. Muller, D., Djebbara-Hannas, Z., Jourdain, P., Vutskits, L., Durbec, P., Rougon, G., and Kiss, J. Z. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 4315–4320

2 M. Mühlenhoff, M. Sauernborn, and R. Gerardy-Schahn, unpublished results.
17. Daniel, L., Durbec, P., Gautherot, E., Rouvier, E., Rougon, G., and Figarella-Branger, D. (2001) Oncogene 20, 997–1004
18. Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Zesing, S., Lottspeich, F., and Frosch, M. (1995) Mol. Microbiol. 16, 441–450
19. Long, G. S., Bryant, J. M., Taylor, P. W., and Luzio, J. P. (1995) Biochim. J. 309, 543–550
20. Scholl, B., Rogers, S., Adhya, S., and Merril, C. R. (2001) J. Virol. 75, 2509–2515
21. Nakano, Y., Yoshida, Y., Yamashita, Y., and Koga, T. (1995) Gene (Amst.) 162, 157–158
22. Lifely, M. R., Nowicka, U. T., and Moreno, C. (1986) Carbohydr. Res. 156, 125–135
23. Skoza, L., and Mohos, S. (1976) Biochem. J. 159, 457–462
24. Voorter, C. E., Haard-Hoekman, W. A., van den Oetelaar, P. J., Bloemendal, H., and de Jong, W. W. (1988) J. Biol. Chem. 263, 19020–19023
25. Stephenson, R. C., and Clarke, S. (1989) J. Biol. Chem. 264, 6164–6170
26. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
27. Kalman, A. V., Kulshin, V. E., Shlyapnikov, M. G., Kaenzenko, V. N., and Kryukov, V. M. (1995) FEBS Lett. 366, 46–48
28. Clarke, B. R., Esumeh, F., and Roberts, I. S. (2000) J. Bacteriol. 182, 3761–3766
29. Legoux, R., Lelong, P., Jourde, C., Feuillerat, C., Capdevielle, J., Sure, V., Ferran, E., Kaghad, M., Delpech, B., Shire, D., Ferrara, P., Loison, G., and Salonen, M. (1996) J. Bacteriol. 179, 7260–7264
30. Heller, K. J., and Krauel, V. (1986) J. Bacteriol. 167, 1071–1073
31. Hänfling, P., Shashkov, A. S., Jann, B., and Jann, K. (1996) J. Bacteriol. 178, 4747–4750
32. Kanamaru, S., Leiman, P. G., Kostyuchenko, V. A., Chipman, P. R., Mesyanzhinov, V. V., Arisaka, F., and Rossmann, M. G. (2002) Nature 415, 553–557
33. Leggate, D. R., Bryant, J. M., Redpath, M. B., Head, D., Taylor, P. W., and Luzio, J. P. (2002) Mol. Microbiol. 44, 749–760
34. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E., and Hatfull, G. F. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 2192–2197
35. Pajunen, M. I., Kiljunen, S. J., Soderholm, M. E., and Skurnik, M. (2001) J. Bacteriol. 183, 1928–1937
36. Steven, A. C., Trus, B. L., Maizel, J. V., Unser, M., Parry, D. A., Wall, J. S., Handfield, J. F., and Studier, F. W. (1988) J. Mol. Biol. 200, 351–365
37. Corpet, F. (1988) Nucleic Acids Res. 16, 10881–10890
Proteolytic Processing and Oligomerization of Bacteriophage-derived Endosialidases
Martina Mühlenhoff, Katharina Stummeyer, Melanie Grove, Markus Sauerborn and Rita Gerardy-Schahn

J. Biol. Chem. 2003, 278:12634-12644.
doi: 10.1074/jbc.M212048200 originally published online January 29, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M212048200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 37 references, 21 of which can be accessed free at http://www.jbc.org/content/278/15/12634.full.html#ref-list-1