

**Characterisation of Bacteriophage-derived
Tailspike and Tail Fibre Proteins**

**Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor der Naturwissenschaften
*Dr. rer. nat.***

genehmigte Dissertation

von

**Dipl.-Biochem. David Schwarzer
geboren am 16. Juli 1977 in Peine**

2008

Referentin: Prof. Dr. Rita Gerardy-Schahn

Korreferent: Prof. Dr. Ralf Ficner

Tag der Promotion: Mittwoch, 06.02.2008; 15:00 Uhr

Schlagworte: Endosialidasen, Bakteriophagen, Polysialinsäure

Key words: endosialidasen, bacteriophages, polysialic acid

„Schläft ein Lied in allen Dingen,
Die da träumen fort und fort,
Und die Welt hebt an zu singen,
Triffst du nur das Zauberwort.“

Joseph Freiherr von Eichendorff (1788-1857)

Table of Contents

Chapter 1 – Zusammenfassung	1
Chapter 2 – Abstract	2
Chapter 3 – General Introduction	3
3.1 – Bacteriophages	3
3.1.1 – Tailed Bacteriophages – the Order of <i>Caudovirales</i>	3
3.1.2 – Life Cycle of Bacteriophages and Pathogenicity	6
3.1.3 – Evolution of Bacteriophages	6
3.1.4 – Tailspike and Tail Fibre Proteins	7
3.2 – <i>Escherichia coli</i> K1 and Polysialic Acid	8
3.3 – Anti-K1 Bacteriophages and Endosialidases	9
3.4 – Crystal Structure of Endosialidase F	13
3.5 – Biological and Biomedical Applications of Endosialidases	15
3.6 – Objectives	16
Chapter 4 – Evolution of Bacteriophages Infecting Encapsulated Bacteria: Lessons from <i>Escherichia Coli</i> K1 Specific Phages	17
4.1 – Supplemental Material	31
Chapter 5 – Characterization of a Novel Intramolecular Chaperone Domain Conserved in Endosialidases and Other Bacteriophage Tailspike and Fibre Proteins	34
5.1 – Supplemental Material	47
Chapter 6 – Proteolytic Maturation of Endosialidase F Is Essential to Allow Efficient Binding to Polysialic Acid	48
Chapter 7 – Synthesis and Biological Evaluation of a New Polysialic Acid Hydrogel as Enzymatically Degradable Scaffold Material for Tissue Engineering	67
7.1 – Supplemental Material	81
Chapter 8 – General Discussion	83
8.1 – Evolution of <i>Escherichia coli</i> K1 specific Bacteriophages	83
8.2 – Characterisation of the C-terminal Chaperone Domain of Endosialidases	84
8.3 – The Potential of Endosialidases in Medical Applications	90
Chapter 9 – References	91
Appendix 1 – Abbreviations	98
Appendix 2 – Curriculum Vitae and Publications	99
Appendix 3 – Danksagungen	102

Chapter 1 – Zusammenfassung

Ein gemeinsames Hauptmerkmal von Bakteriophagen, die *Escherichia coli* K1 infizieren, sind ihre *tailspike* Proteine, die Endosialidasen. Diese Enzyme, über die K1-Phagen ihren Wirt spezifisch erkennen, geben ihnen gleichzeitig die Fähigkeit, die dichte bakterielle Kapsel, die aus α 2,8-verknüpfter Polysialinsäure besteht, abzubauen und so den Infektionszyklus einzuleiten. In der vorliegenden Arbeit wurden drei Ziele verfolgt: (i) die Untersuchung der Evolution K1-spezifischer Phagen, (ii) die Eigenschaften zu definieren, welche die ungewöhnlich hohe Stabilität und Spezifität der Polysialinsäure degradierenden Enzyme bedingen und (iii) die Charakterisierung der Fähigkeit der Endosialidasen in Bezug auf den Abbau chemisch modifizierter Polysialinsäure.

Die Evolution K1-spezifischer Phagen wurde in einer vergleichenden Genomanalyse der lytischen Phagen K1E und K1F sowie des temperenten anti-K1 Phagen CUS-3 untersucht. In dieser ersten Studie konnte gezeigt werden, dass K1-Phagen nicht aus einem einzigen Vorläufer entstanden sind, sondern enge Verwandtschaft zu unterschiedlichen Vorläuferphagen aufweisen. Es wurde gezeigt, dass die Wirtsspezifität für *Escherichia coli* K1 hauptsächlich durch lateralen Transfer des Endosialidase-Gens erworben wurde.

Die funktionelle Faltung und Oligomerisierung der Endosialidasen hängt essentiell von der Präsenz einer kurzen C-terminalen Domäne (CTD) ab, die im Verlaufe der Proteinreifung abgespalten wird. Im zweiten Teil dieser Arbeit konnte eine homologe CTD in weiteren Phagenproteinen identifiziert werden, die zu drei unterschiedlichen Proteinfamilien gehören: die *neck appendage* Proteine verschiedener Bacillusphagen, die *L-shaped tail fibre* Proteine des Coliphagen T5, sowie verschiedene K5-Lyasen, die *tailspike* Proteine von *Escherichia coli* K5 infizierenden Phagen. Um die Bedeutung der CTD für die Faltung und Funktion dieser Proteine zu untersuchen, wurde ein Vertreter jeder Familie ausgewählt und einer Struktur- und Funktionsanalyse unterzogen. Diese führte zu dem Ergebnis, dass (i) die CTD eine eigenständige Domäne mit hohem Anteil an α -helikalen Strukturen darstellt; (ii) die proteolytische Freisetzung durch Spaltung an einem hoch konservierten Serinrest erfolgt und der Austausch dieses Restes gegen Alanin die Spaltung verhindert; (iii) die isolierten C-terminalen Domänen zu Hexameren assemblieren; (iv) hoch konservierte Aminosäurereste innerhalb der CTD essentiell für die Faltung und Komplexbildung der Trägermoleküle sind und (v) die CTD zwischen Proteinen der oben genannten Proteinfamilien ausgetauscht werden kann. Somit konnte die CTD als intramolekulare Chaperon-Domäne identifiziert werden. Ein Vergleich der thermischen Stabilität von ungespaltenen und proteolytisch prozessierten Endosialidase-Varianten deutete darauf hin, dass die Freisetzung der CTD die Entfaltungs-Energiebarriere des Enzymkomplexes erhöht, wodurch das prozessierte Trimer in einer kinetisch stabilen Konformation gehalten wird.

Zusätzlich konnte in einer dritten Studie gezeigt werden, dass die proteolytische Spaltung von Endosialidasen auch direkt die Interaktion mit Polysialinsäure beeinflusst, da das Vorhandensein der CTD die Substratbindung vermindert.

Der letzte Teil dieser Studie unterstreicht das biotechnologische Potential von bakteriophagen-assoziierten *tailspike* Proteinen, wiederum am Beispiel von Endosialidasen. Hier konnte gezeigt werden, dass Hydrogele aus quervernetzter Polysialinsäure mit Hilfe von Endosialidasen kontrolliert abgebaut werden können. Diese neuartigen, synthetischen und unlöslichen Materialien könnten als biodegradierbare Gerüste in neuroregenerativen Anwendungen zum Einsatz kommen.

Chapter 2 – Abstract

Bacteriophages infecting *Escherichia coli* K1 share one common feature, represented by the endosialidase tailspike proteins. These enzymes allow the K1-phages to specifically bind to and cleave the capsular polysaccharide of the host bacterium composed of α 2,8-linked polysialic acid. The present work was focussed on: (i) the evolution of K1-specific phages, (ii) the unusual stability and specificity of the polysialic acid degrading enzymes, and (iii) the ability of endosialidases to degrade chemically modified polysialic acid.

Genome sequencing of two lytic K1-phages and comparative analyses including a K1-prophage revealed that K1-phages did not evolve from a common ancestor. By contrast, the genomic organisation of the investigated K1-phages indicated a close relationship to different progenitor types, namely T7-, SP6-, and P22-like phages, and strongly suggested that lateral acquisition of the endosialidase tailspike gene was the main evolutionary event in gaining host specificity for *Escherichia coli* K1.

As described previously, folding and assembly of the trimeric central catalytic part of endosialidases crucially depend on a short C-terminal domain (CTD), which is released during proteolytic maturation. In the second part of this study, homologous CTDs were identified in phage proteins belonging to three different protein families: neck appendage proteins of several *Bacillus* phages, L-shaped tail fibres of coliphage T5, and K5-lyases, the tailspike proteins of phages infecting *Escherichia coli* K5. The analysis of a representative of each family demonstrated that in all cases the CTD is cleaved off at a strictly conserved serine residue and alanine substitution prevented cleavage. Further structural and functional analyses revealed that: (i) CTDs are autonomous domains with a high α -helical content; (ii) proteolytically released CTDs assemble into hexamers which are most likely dimers of trimers; (iii) highly conserved amino acids within the CTD are indispensable for CTD-mediated folding and complex formation; (iv) CTDs can be exchanged between proteins of different families; and (v) proteolytic cleavage is essential to stabilise the native protein complex. Data obtained for full-length and proteolytically processed endosialidase variants suggest that release of the CTD increases the unfolding barrier, trapping the mature trimer in a kinetically stable conformation. In summary, the CTD was characterized as a novel C-terminal chaperone domain, which assists folding and assembly of unrelated phage proteins.

Moreover, the comprehensive analysis of processed and unprocessed endosialidase variants presented in the third part of this work revealed that in the full-length protein the CTD interferes with efficient substrate binding by affecting a polysialic acid binding site located outside the active center.

The last part of the present study emphasises the biotechnological importance of bacteriophage tailspike proteins taking endosialidases as an example. It was demonstrated that endosialidases can be applied to degrade hydrogels of cross-linked polysialic acid in an induced and controlled manner. These novel chemically synthesised and insoluble materials might represent promising basic modules for biodegradable scaffolds in neuroregenerative tissue engineering approaches.

Chapter 3 – General Introduction

3.1 – Bacteriophages

Bacteriophages, or in short ‘phages’, are viruses that infect bacteria. They were discovered in the early 20th century independently by Frederick William Twort and Felix Hubert d’Herelle (Twort, 1915; d’Herelle, 1917). With an estimated population size of more than 10^{31} phage particles (virions), bacteriophages are likely to be the numerically most prominent organism in the biosphere (Brüssow and Hendrix, 2002; Chibani-Chennoufi *et al.*, 2004; Brüssow *et al.*, 2004; Ackermann, 2007). Phages infect many bacterial strains of eubacteria and archaea. Since life emerged 3.8 billion years ago, phages have followed prokaryotes into all known habitats including soil, marine and lake waters, animals, plants, and also into the extreme environments occupied by hyperthermophilic and halophilic archaea. Most phages have been found in easily cultivated bacteria of industrial or medical importance, including enterobacteria, lactococci, bacilli, clostridia, pseudomonads, staphylococci, and streptococci (Ackermann, 1998; Hambly and Suttle, 2005; Ackermann, 2007).

Phages contain at least two biomacromolecules: proteins and nucleic acid. The proteins mainly provide the structure of the virion and enclose the nucleic acid. Some phages also contain lipids as internal vesicles or are surrounded by an envelope. Hitherto, more than 5,500 structures of virions have been examined by electron microscopy. Phages thus represent the largest known virus group in biological systems. By their morphology and nucleic acid type the phages are subdivided into 20 families. About 96% of known phages are ascribed to three families of the most well characterised group of bacteriophages, the order of *Caudovirales* or also referred to as tailed phages (Ackermann, 2006; Ackermann, 2007).

3.1.1 – Tailed Bacteriophages – the Order of *Caudovirales*

Tailed bacteriophages are composed of a head and a tail that names the *Caudovirales* (Figure 1-1). They are exclusively composed of proteins and linear, double-stranded DNA. The outer protein part of the head – the capsid – basically exhibits an icosahedral symmetry that can either be isometric or elongated (prolate) and is symmetrically constructed of a single layer of homo- or heterooligomeric capsomer proteins. The radii of the capsids vary between 10 and 100 nm. The DNA is enclosed in the capsid and its size is found in the range of 17 to 725 kb (Ackermann, 1998).

Capsid assembly is initiated by the formation of a prohead by polymerisation of the major capsid proteins on scaffold-core components associated with a head-tail-connector. The scaffold proteins are removed during the further maturation of the capsid. The dodecameric, torus-shaped head-tail connector (or portal protein) is inserted in one of the 12 vertices of the capsid, named the portal vertex. The DNA is processively packed into the prohead by a terminase holoenzyme or a packaging ATPase after being attached to the portal protein. Since DNA is negatively charged and self-repulsive, the packaging mechanism requires energy that is provided by ATP hydrolysis (Tao *et al.*, 1998; Ibarra *et al.*, 2001; Jardine and Anderson, 2006).

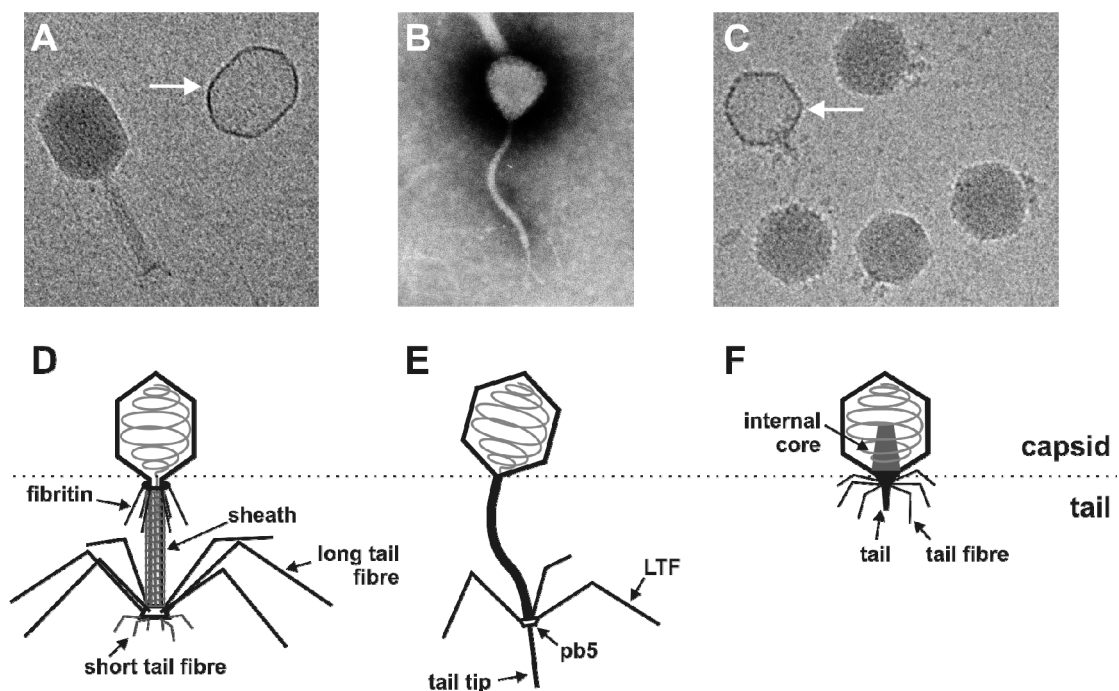


FIGURE 1-1: Structure of tailed bacteriophages A–C, electron micrographs and D–F, schematic representation of tailed bacteriophages with the capsids depicted as *black hexagons*. The DNA is shown as *grey spool*. Characteristic tail structure proteins are indicated exemplarily for the respective phages in different colours. A and D, coliphage T4 of the family of *Myoviridae* with the six *fibrinitins*, the contractile *sheath*, the six *long* and the six *short tail fibres* at the base plate (adapted from Mesyanzhinov *et al.*, 2004). B and E, coliphage T5 of the family of *Siphoviridae* with the three L-shaped tail fibre proteins (*LTF*), the *tail tip*, and protein *pb5* (adapted from Ackermann, 1998). C and F, coliphage K1E and T7, respectively of the family of *Podoviridae* with the *internal core* proteins (*dark grey cone*), the *tail* proteins and the six *tail fibres* (adapted from Molineux, 2006). Arrows in Figures A and C show phage heads lacking the DNA. Cryo-electron micrographs of Figures A and C are courtesies of Dr. Petr G. Leiman, and the phosphotungstate stained electron micrograph in Figure B is courtesy of Prof. Hans-Wolfgang Ackermann.

In the mature capsid, the DNA shows a nearly crystalline density and is packed in concentric rings around the central axis of the phage (Cerritelli *et al.*, 1997; Leiman *et al.*, 2007). The DNA packaging motor is removed after finishing its job and the bacteriophage tail is assembled at the portal protein. The tail, essentially a hollow tube with six-fold symmetry, can contain additional structures, namely a base plate with six copies of tailspike proteins, tail fibres, or collars. According to microscopically observable differences of the tail morphology, the *Caudovirales* are subdivided into three different families: the *Myoviridae*,

the *Siphoviridae*, and the *Podoviridae*, reproducing Bradley groups A, B, and C, respectively (Fig. 1-1; Bradley, 1967; Ackermann, 1998).

3.1.1.1 – The Family of *Myoviridae*

The tail of *Myoviridae* comprises a long, rigid tail tube that is surrounded by a contractile sheath. The sheath is separated from the head by a neck. One well-studied myovirus is phage T4 of the group of T-even coliphages, given in *Figures 1-1A* and *D*. The six-fold symmetric base plate of T4, consisting of at least 16 different gene products (gp's), includes the short tail fibres and exhibits the assembling platform of the tail. In addition, six kinked long tail fibre proteins are attached to the base plate. The fibrinins or 'whiskers' act as assembling proteins for the long tail fibres and stabilise the fibres in unfavourable conditions. The long and short tail fibres are required to attach the phage to the outer membrane of the bacterial host. Upon binding of the fibres to their cell receptors, the base plate undergoes a large conformational switch, which initiates sheath contraction and culminates in transfer of the phage DNA from the capsid into the host cell through the tail tube (Leiman *et al.*, 2003; Mesyanzhinov *et al.*, 2004; Rossmann *et al.*, 2004; Mosig and Eiserling, 2006).

3.1.1.2 – The Family of *Siphoviridae*

Siphoviridae contain a long and flexible tail structure that is not contractile. One member of this family is coliphage T5 (*Figure 1-1B* and *E*). In contrast to phage T4, little is known about the structures of phage T5 (Sayers, 2006). It contains three L-shaped tail fibre (LTF) proteins that are attached to the base plate. Interestingly, LTF accelerates binding to the bacterial cell surface but is dispensable in the infection cycle. After the tail tip protein has formed a pore through the bacterial membrane for DNA injection, irreversible adsorption is triggered by binding of the pb5 protein to the FhuA receptor protein of the host cell (Heller, 1984; Heller and Schwarz, 1985; Feucht *et al.*, 1990; Guihard *et al.*, 1992; Ackermann, 1998; Sayers, 2006).

3.1.1.3 – The Family of *Podoviridae*

In contrast to *Myo-* or *Siphoviridae*, the *Podoviridae* have a short non-contractile tail (*Figure 1-1C* and *F*). The tail forms a rigid structure and is involved in binding to a cellular component. Coliphage T7 is the prototype member of the T7 supergroup. Interestingly, these phages contain three internal core proteins, which, after attachment of the tail to the cell surface, are ejected to generate a tunnel into the periplasm of the host for DNA transfer. The host specificities of the *Podoviridae* coliphages T7 and K1E are determined by six copies of tail fibres (*Fig. 1-1F*) or tailspike proteins (*Fig. 1-1C* and *1-4A*), respectively (Molineux, 2006; Leiman *et al.*, 2007).

3.1.2 – Life Cycle of Bacteriophages and Pathogenicity

Depending on their infection cycle, bacteriophages can be subdivided into virulent and temperate phages. The infection cycle is generally initiated by the attachment of the bacteriophage to the bacterial surface and the injection of the DNA. The genes of virulent (or lytic) bacteriophages are controlled by strong viral promoters that yield high expression of the genes. The nucleic acid is replicated in high copies and new bacteriophage virions are assembled. Finally, the host cell is disrupted and a new generation of 100–200 phage particles is released to start a new lytic cycle.

In contrast, the DNA of temperate phages is integrated into the host genome at a specific insertion site, preferably into a host tRNA gene. This prophage persists in the so called lysogenic host and is replicated with the host genome. The lysogenic cells can be induced to re-enter the lytic cycle by UV light or other exogenous stress factors (Ackermann, 1998; Campbell, 2003; Canchaya *et al.*, 2004).

Prophages can encode several proteins that alter the host properties (lysogen conversion genes). Factors that modify the host surface to prevent superinfection by homologous lytic phages yield advantages primarily for the prophage. On the other hand, some prophages harbour genes allowing the host to survive in otherwise lethal environments, like in mammals with an adaptive immune system. Examples of prophage encoded virulence factors are the diphtheria toxin, *Vibrio cholerae* toxins, the Shigatoxin, *Streptococcus pyogenes* virulence factors, and botulinum neurotoxins C1 and D. As a consequence, an otherwise beneficial bacterium can be converted into a pathogen by infection with a prophage harbouring such toxins (Desiere *et al.*, 2001; Brüssow *et al.*, 2004; Canchaya *et al.*, 2004; Tinsley *et al.*, 2006; Brüssow, 2006).

3.1.3 – Evolution of Bacteriophages

Structural similarities between bacteriophages and adenoviruses or reoviruses argue that prokaryotic and eukaryotic viruses share common origins and did not emerge independently by convergent evolution (Campbell, 2003; Duda *et al.*, 2006). However, recent studies demonstrated that some structural elements like the tail tube found in tailed bacteriophages and tectiviruses might have developed by convergent evolution (Ackermann, 2007). Otherwise, there is no evidence for a polyphyletic origin of tailed phages: They rather seem to share a common gene pool. More than 600 phage and prophage genomes are sequenced to date and display a mosaic composition of genes and gene clusters that has emerged by horizontal gene transfer. In line with the concept of modular evolution of bacteriophages, it was suggested that the transfer of gene fragments, whole genes or gene clusters can occur by

homologous recombination between a prophage and a lytic phage during superinfection of a host. This mechanism is restricted to predetermined recombination sites. Another model proposed that the transfer of fragmented DNA stretches which carry novel gene elements may also occur by random, illegitimate recombination at non-homologous sites. Additionally, for temperate phages, it has been described that novel genes are located close to the integration site, suggesting that they may have been acquired by imprecise prophage excision. (Botstein, 1980; Ackermann, 1998; Hendrix *et al.*, 2000; Juhala *et al.*, 2000; Hendrix, 2002; Brüssow *et al.*, 2004; Casjens, 2005; Brüssow and Desiere, 2006).

3.1.4 – Tailspike and Tail Fibre Proteins

Tailed phages have evolved tailspike and fibre proteins for efficient virus-host-interactions. These specialised adhesins mediate the recognition and attachment to the bacterial surface and are the key determinants for host specificity. In mature phage particles, tailspikes and tail fibres are exposed structures, which require high stability to maintain their functional conformation even under extreme environmental conditions, such as high salt concentrations, the presence of extracellular proteases, and drastic variations in pH and temperature. Interestingly, many spikes and fibres are composed of homotrimers containing stretches of intertwined subunits like coiled-coil, triple β -helix or triple β -spiral folds, leading to protein complexes which remain stable even in the presence of sodium dodecyl sulphate (SDS) (Steinbacher *et al.*, 1994; Chappell *et al.*, 1997; van Raaij *et al.*, 2001; Kanamaru *et al.*, 2002; Freiberg *et al.*, 2003; Weigele *et al.*, 2003).

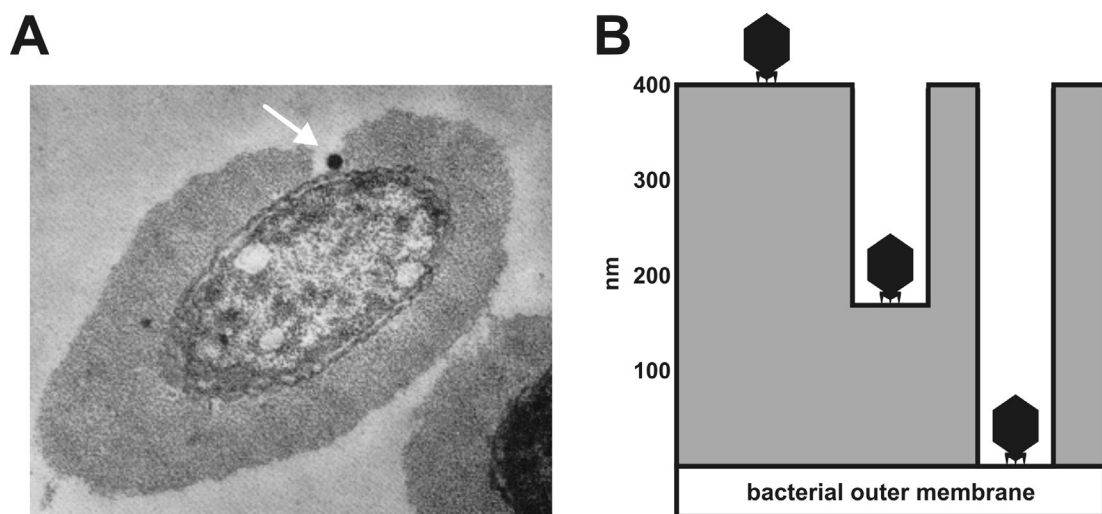


FIGURE 1-2: A, a bacteriophage is attached to the outer membrane of its host (*arrow*). The polysaccharide capsule was enzymatically removed shaping a narrow path towards the outer membrane (modified from Sutherland, 1977). B, schematic representation of a podovirus (black) that penetrates the capsular polysaccharide (depicted in grey) of an encapsulated bacterium. The thickness of the capsule is indicated on the left (adapted from Hughes *et al.*, 1998).

Several phages have developed tailspike proteins with an enzymatic activity in order to penetrate the thick layers of lipo- or capsular polysaccharides possessed by many pathogenic bacteria. These capsule-specific depolymerases (endoglycosidases or lyases) are required to gain access to and to fix the phage at the bacterial outer membrane (Figure 1-2; Stirn and Freund-Mölbart, 1971; Lindberg, 1977; Sutherland, 1977; Scholl *et al.*, 2005).

3.2 – *Escherichia coli* K1 and Polysialic Acid

A prominent example of an encapsulated bacterium is *Escherichia coli* K1, a gram-negative neuroinvasive bacterium that causes meningitis and sepsis in neonates. The pathogen is surrounded by a thick capsule of polysialic acid (polySia), a linear homopolymer that is composed of up to 200 α 2,8-linked 5-*N*-acetylneuraminic acid residues per chain (Figure 1-3). Identical capsules were found in *Neisseria meningitidis* serogroup B, *Moraxella nonliquefaciens*, and *Mannheimia (Pasteurella) haemolytica* A2 (Barry and Goebel, 1957; McGuire and Binkley, 1964; Furowicz and Orskov, 1972; Bhattacharjee *et al.*, 1975; Jennings *et al.*, 1985; Aalto *et al.*, 2001).

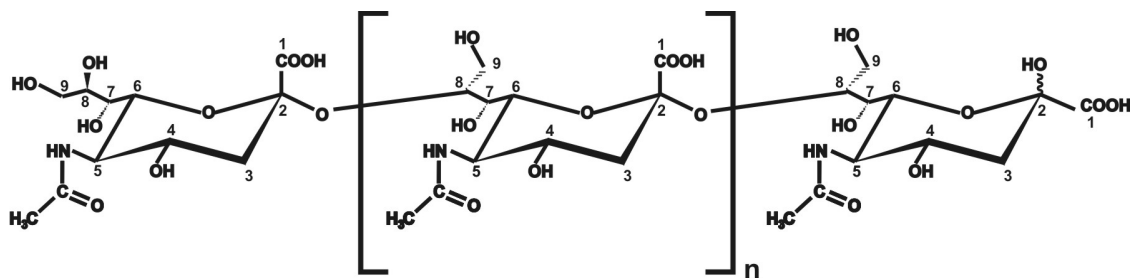


FIGURE 1-3: α 2,8-linked polysialic acid.

In vertebrates, polySia is primarily found as a posttranslational modification of the neural cell adhesion molecule NCAM. It was also described as a major glycan structure of the CD36 scavenger receptor in human milk, the α -subunit of the voltage-gated sodium channel in rat brain, and recently identified as a posttranslational modification of neuropilin-2 on maturing human dendritic cells (Finne *et al.*, 1983; Zuber *et al.*, 1992; Yabe *et al.*, 2003; Moebius *et al.*, 2007; Curreli *et al.*, 2007). PolySia is widely expressed during ontogenetic development and remains an important modulator of neuronal plasticity in the adult brain (Roth *et al.*, 1988; Rutishauser and Landmesser, 1996; Kleene and Schachner, 2004; Weinhold *et al.*, 2005; Conchonaud *et al.*, 2007). By mimicry of this host structure, polySia encapsulated bacteria evade the human immune system. The thick capsule protects the prokaryotes from complement-mediated lysis and opsonophagocytosis, and was found to enable the bacterium to cross the blood-brain barrier (Frosch *et al.*, 1987; Cross, 1990; Prasadarao *et al.*, 1996; Jódar *et al.*, 2002; Taylor and Roberts, 2005).

3.3 – Anti-K1 Bacteriophages and Endosialidases

The polySia capsule is not only an essential virulence factor of *E. coli* K1; it also provides an attachment site for specialised anti-K1 bacteriophages. More than 20 lytic bacteriophages specific for *E. coli* K1 have been isolated, mainly from sewage samples, and were found to exhibit different morphologies (*Table 1*).

TABLE 1: Overview of isolated bacteriophages specific for *E. coli* K1 and cloned endosialidase genes

Phage	Virus family	Host strain	Phage isolation	Endosialidase gene	Enzyme
ΦK1A	<i>Podoviridae</i>	<i>E. coli</i> K1	(Gross <i>et al.</i> , 1977)	(Jakobsson <i>et al.</i> , 2007)	endoNA
ΦK1B, ΦK1C, ΦK1D	<i>Podoviridae</i>	<i>E. coli</i> K1	(Gross <i>et al.</i> , 1977)	-	-
ΦK1E	<i>Podoviridae</i>	<i>E. coli</i> K1	(Gross <i>et al.</i> , 1977)	(Gerardy-Schahn <i>et al.</i> , 1995; Long <i>et al.</i> , 1995)	endoNE
ΦK1F	<i>Podoviridae</i>	<i>E. coli</i> K1	(Vimr <i>et al.</i> , 1984)	(Petter and Vimr, 1993; Mühlenhoff <i>et al.</i> , 2003)	endoNF
Φ63D	<i>Siphoviridae</i>	<i>E. coli</i> K1	(Miyake <i>et al.</i> , 1997)	(Machida <i>et al.</i> , 2000a)	endoN63D
ΦK1-5	<i>Podoviridae</i>	<i>E. coli</i> K1, <i>E. coli</i> K5	(Scholl <i>et al.</i> , 2001)	(Scholl <i>et al.</i> , 2001)	endoNK1-5
Φ92	<i>Myoviridae</i>	<i>E. coli</i> K92, <i>E. coli</i> K1	(Kwiatkowski <i>et al.</i> , 1983)	-	-
Φ1.2	<i>Podoviridae</i>	<i>E. coli</i> K1	(Kwiatkowski <i>et al.</i> , 1982)	-	-
B, C, D, F, G, K, L, P, and R	n.d.	<i>E. coli</i> K1	(Smith and Huggins, 1982)	-	-
3, 9, Φ63A, and Φ63E	<i>Podoviridae</i>	<i>E. coli</i> K1	(Miyake <i>et al.</i> , 1997)	-	-
a and d	<i>Siphoviridae</i>	<i>E. coli</i> K1	(Miyake <i>et al.</i> , 1997)	-	-
CUS-3	n.d. (temperate)	<i>E. coli</i> K1	(Deszo <i>et al.</i> , 2005; this study)	(this study)	endoNK1

Most of the characterised phages belong to the *Podoviridae*, three isolates were found to be *Siphoviridae* (Miyake *et al.*, 1997), whereas the isolate Φ92 contained a contractile tail – the characteristic feature of *Myoviridae* (Kwiatkowski *et al.*, 1983). In parallel to this work, the first temperate K1-specific phage (CUS-3) was identified. By analysing the partially sequenced genome of *E. coli* K1 strain RS218, the CUS-3 prophage was found to be integrated into the genome of the bacterium (Deszo *et al.*, 2005). A common feature of all known *E. coli* K1 specific bacteriophages is the presence of specialised tailspike proteins, the endosialidases. These enzymes specifically bind to and cleave α 2,8-linked polysialic acid. Interestingly, myovirus Φ92 was isolated from *E. coli* K92 that contains a capsular polymer composed of sialic acid in alternating α 2,8/ α 2,9-ketosidic linkages. However, Φ92 is also able to infect *E. coli* K1 and further studies revealed, that the endosialidase solely cleaves the α 2,8-linkages (Kwiatkowski *et al.*, 1983).

Endosialidases were also termed ‘endo-*N*-acetyl-neuraminidases’ (endoN). Even though Cabezas later on suggested to employ the term ‘endosialidases’ to follow the recommendations made for other (exo-)sialidases (Cabezas, 1991), the abbreviation ‘endoN’ has been maintained. For example, the endosialidase of coliphage K1F is termed ‘endosialidase F’ or ‘endoNF’.

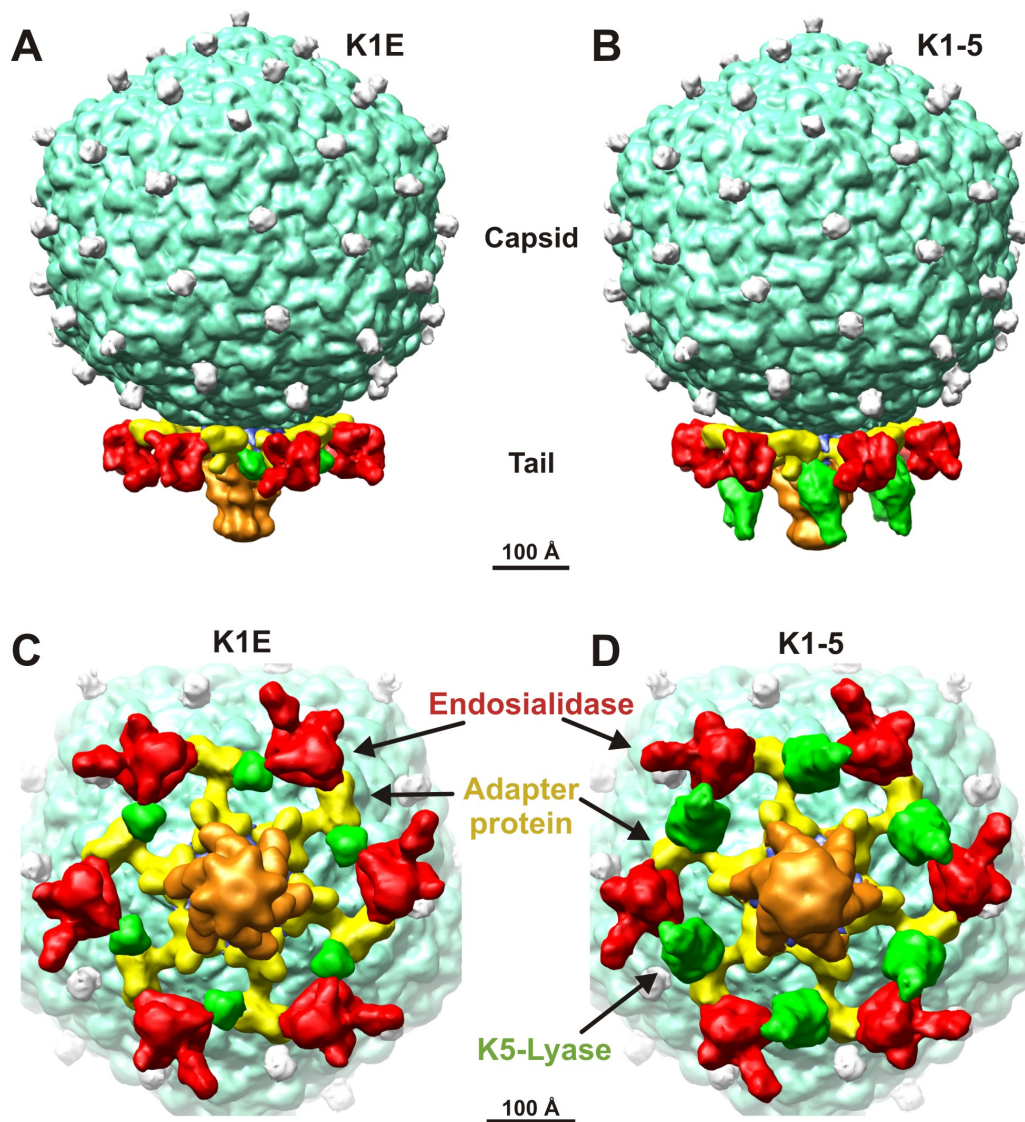


FIGURE 1-4: Structure of K1E (left-hand panel) and K1-5 (right-hand panel) in side view (*A* and *B*) or end-on view of the tail (*C* and *D*). A combined cryo-EM map, containing the 5-fold symmetric capsid and 6-fold symmetric tail, is shown contoured at a density level of 0.8 σ . The endosialidases are shown in *red*, the adapter protein in *yellow*, the K5-lyase and a small protein of unknown function in Φ K1E are depicted in *green*. The tail structure is shown in *orange* (with permission modified from Leiman *et al.*, 2007).

In parallel to this work, the cryo-EM structure of coliphages K1E (*Figure 1-4A*) and K1-5 (*Figure 1-4B*) were solved (Leiman *et al.*, 2007). Notably, coliphage K1-5 exhibits dual host specificity due to a second tailspike protein, which is a K5-lyase. This enzyme specifically cleaves the *E. coli* K5-capsular polymer composed of the disaccharide repeating unit of *N*-acetyl-heparosan -4)-GlcA-(β 1,4)-GlcNAc-(α 1- (Legoux *et al.*, 1996; Scholl *et al.*, 2001).

The cryo-EM structure revealed that both tailspikes, the endosialidase (*red*) and the K5-lyase (*green*), are attached to the phage base plate via the same small adapter protein (*yellow*) that was further characterised in the course of this study.

In contrast to exosialidases that cleave only sialic acid residues from the non-reducing end of sialylated glycoconjugates, endosialidases exclusively cleave within the polySia chain. Several studies on different purified anti-K1 phages and endosialidases unravelled minimum substrate lengths ranging from a trimer to an octamer (Hallenbeck *et al.*, 1987; Pelkonen *et al.*, 1989; for an overview: Miyake *et al.*, 1997). Interestingly, Miyake *et al.* described that the endosialidase of phage 63D also cleaves trimeric sialic acid and releases a sialic acid residue from the non-reducing end. The data presented argue for an exo-mechanism of this enzyme (Miyake *et al.*, 1997; Machida *et al.*, 2000a; Kataoka *et al.*, 2006).

All endosialidases purified from phage particles show oligomeric complexes. EndoNF was found as an SDS-resistant homotrimer of 103 kDa subunits (Hallenbeck *et al.*, 1987; Mühlenhoff *et al.*, 2003), whereas for endoNE an SDS-resistant heterooligomeric complex with an unknown 38 kDa protein was reported (Tomlinson and Taylor, 1985; Gerardy-Schahn *et al.*, 1995). This protein could presumably resemble the adapter protein mentioned above (*Figure 1-4*). EndoN63D was described to form homotetrameric complexes linked via disulfide bonds (Machida *et al.*, 2000a; Machida *et al.*, 2000b).

So far, five endosialidase genes of the lytic phages K1A, K1E, K1F, 63D, and K1-5 have been cloned (Petter and Vimr, 1993; Gerardy-Schahn *et al.*, 1995; Long *et al.*, 1995; Machida *et al.*, 2000a; Scholl *et al.*, 2001; Mühlenhoff *et al.*, 2003; Jakobsson *et al.*, 2007). The endosialidase genes of Φ K1E and Φ K1-5 are nearly identical with about 98% sequence similarity, and the recently published sequence of endoNA exhibits 90% similarity to endoNE. Moreover, the size of these three endosialidases is almost identical (Jakobsson *et al.*, 2007); therefore endoNE is considered as a representative for this group in the following section.

All known endosialidases share a modular architecture (*Fig. 1-5*). The central region of about 650 amino acids shows a high sequence similarity (50-81%) and harbours the enzymatic activity (Mühlenhoff *et al.*, 2003). Conversely, the enzymes differ in their N-terminal regions. While endoNE contains a short N-terminal stretch only, endoNF and endoN63D are N-terminally extended by 200 and 120 amino acids, respectively (Machida *et al.*, 2000a). The N-terminal parts of endoNF and endoN63D display high sequence similarities to the N-terminal region of the tail fibre protein gp17 of coliphages T3 and T7, and of yersiniophage YeO3-12 (Pajunen *et al.*, 2000). For gp17 of coliphage T7 it has been

shown that this tailspike protein is attached to the phage particle via the N-terminal domain (Steven *et al.*, 1988). A similar function of the homologous part of endoNF and endoN63D has been suggested previously (Petter and Vimr, 1993). Attempts to prove this hypothesis have been carried out in this study (Chapter 4).

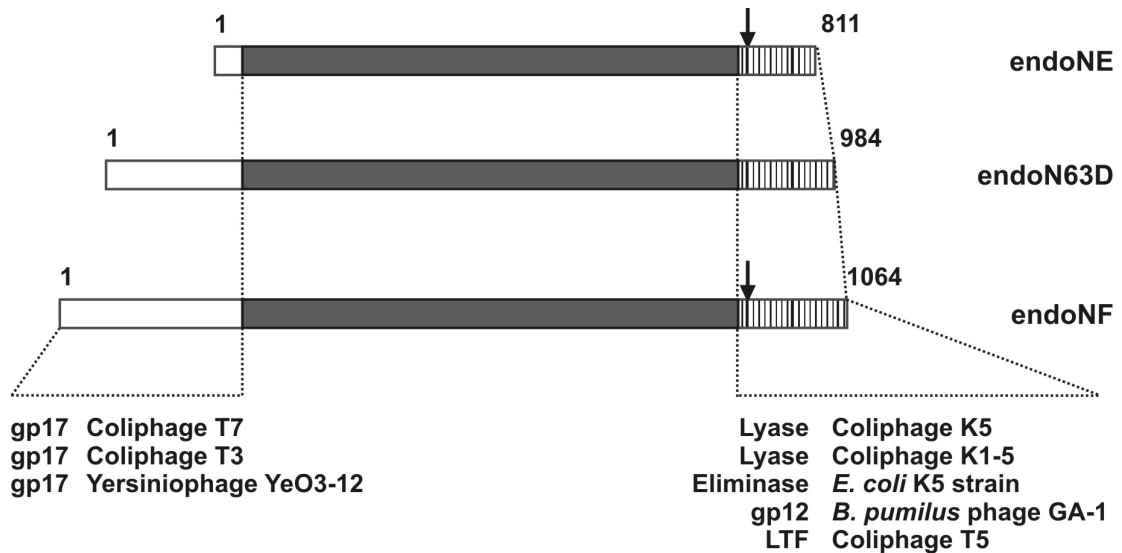


FIGURE 1-5: Schematic representation of cloned endosialidases. The highly conserved central domain common to all endosialidases is depicted in *grey*. The variable N-terminal parts are shown in *white*. Proteolytic cleavage sites identified in Mühlenhoff *et al.*, 2003 are indicated by *arrows*, whereas conserved stretches within the C-terminal part are schematically represented by *black bars*. The N-terminal domains of endoN63D and endoNF shows sequence similarities to the N-terminal part of the tail fibre protein gp17 of coliphages T3 and T7, and of yersiniophage YeO3-12. The C-terminal part exhibits sequence similarities to the C-terminal parts of K5-specific lyases of Coliphages K5 and K1-5, the K5-eliminase of an *E. coli* strain, the neck-appendage protein precursor of *Bacillus pumilus* phage GA-1, and the L-shaped tail-fibre protein (LTF) of coliphage T5.

Additionally, at the C-terminus, all endosialidases contain a short part characterised by the presence of highly conserved amino acid clusters. For endoNE and endoNF it has already been demonstrated that this C-terminal region is released from the mature protein, after the larger N-terminal portion had formed the SDS-resistant trimeric complexes (Leggate *et al.*, 2002; Mühlenhoff *et al.*, 2003). For endoNE and endoNF, a highly conserved serine residue was identified as the proteolytic cleavage site (Leggate *et al.*, 2002; Mühlenhoff *et al.*, 2003). Exchange of this highly conserved serine residue to alanine resulted in non-cleavable mutants that also formed SDS-resistant complexes. Since the mutants were enzymatically active, it could be concluded that proteolytic cleavage is not a prerequisite for the formation of an active enzyme. Further investigations revealed that truncation of the C-terminal part by 32 amino acids (endoNE) or 153 amino acids (endoNF), as well as exchange of a highly conserved histidine residue in the C-terminal part to alanine resulted in exclusively insoluble protein (Petter and Vimr, 1993; Gerardy-Schahn *et al.*, 1995; Mühlenhoff *et al.*, 2003). In addition, the C-terminal part was found to be required in the primary translation product to support the formation of soluble and enzymatically active trimeric complexes. Taken

together, these findings indicate that the C-terminal part has an important function for the proper folding or complex formation of the catalytic N-terminal part.

Database searches using the primary sequence of the C-terminal part of endoNF revealed sequence similarities to the C-terminal region of other bacteriophage tailspike proteins which are completely unrelated to endosialidases in their N-terminal portions (*Figure 1-5*; Mühlhoff *et al.*, 2003). These include the neck-appendage protein precursor of *Bacillus pumilus* phage GA-1, the L-shaped tail-fibre protein (LTF) of coliphage T5, and K5-specific lyases including the eliminase of an *E. coli* K5 strain as well as the lyases of Coliphages K5 and K1-5, respectively. One question addressed in this study was to find out, whether the chaperone like function suggested previously for the C-terminal part of endoNE and NF (Mühlhoff *et al.*, 2003) is also preserved in the newly identified phage proteins.

3.4 – Crystal Structure of Endosialidase F

In parallel to this work, the first crystal structure of an endosialidase was solved by my colleague Dr. Katharina Stummeyer in collaboration with the laboratory of Prof. Ralf Ficner, Göttingen. Crystallisation was achieved for the central catalytic domain of the enzyme cloned from Φ K1F (endoNF) (Stummeyer *et al.*, 2005). The homotrimer resembles the cap and the stalk outline of a mushroom and contains, apart from a short α -helix at the N-terminus, predominantly β -fold structures (*Figure 1-6*).

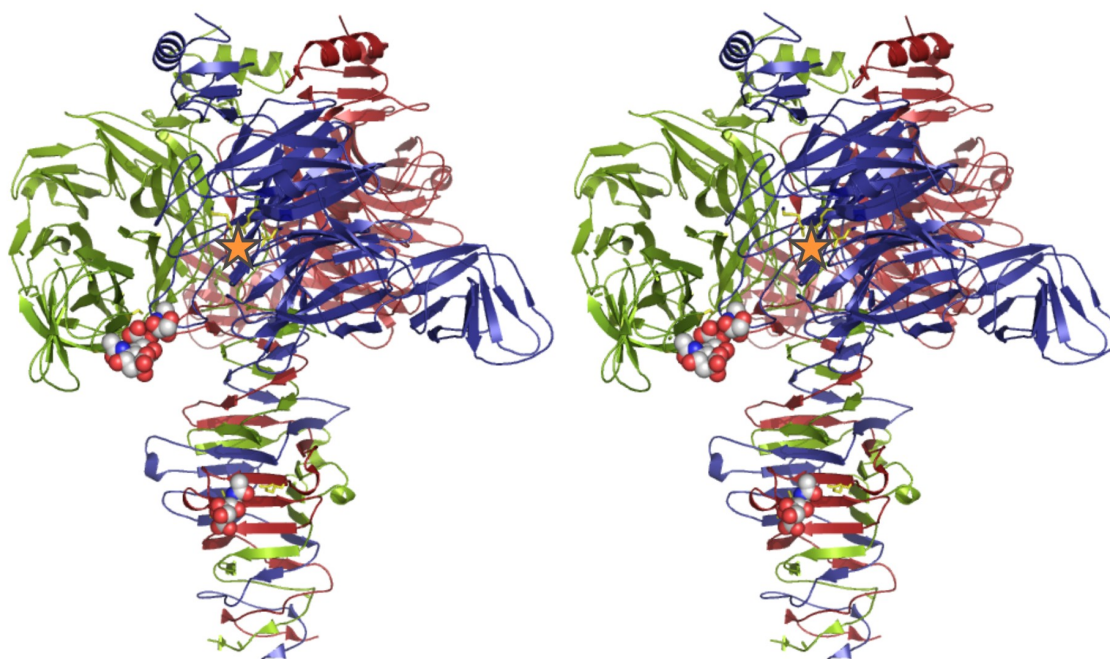


FIGURE 1-6: Stereo ribbon diagram of endosialidase F (Stummeyer *et al.*, 2005). The three monomers are red, green, and blue. The active site (a) of the *blue* subunit is schematically depicted by an *orange asterisk*. *Spheres* represent di-sialic acid and sialic acid bound in the binding sites in the beta-barrel domain of the *green* subunit and in the beta-prism domain of the *red* subunit, respectively. Amino acids involved in enzymatic activity or polySia binding as determined in Stummeyer *et al.* are depicted as *yellow sticks*.

The N-terminal ‘cap’ of the mushroom contains a six-bladed β -propeller that is typical for sialidases, and a lectin-like nine-stranded β -barrel domain that is inserted in the third blade of the propeller. This dual domain structure is found independently in each subunit interacting with each other along the three-fold symmetry axis passing through the centre of the trimer. The β -propeller motif was found to share a higher structural similarity with bacterial and eukaryotic than with viral exosialidases. It contains the catalytic site (*orange star*) located in a cleft in the centre of the propeller structure. Three amino acid residues were identified to be essential for enzymatic activity: Glu-581, Arg-596 and Arg-647 (*yellow sticks*). These residues are highly conserved in all known sialidases (endo- and exosialidases). Interestingly, the two catalytic residues essential for the proposed reaction mechanism of exosialidases (Chong *et al.*, 1992; Watts *et al.*, 2003; Amaya *et al.*, 2004) are missing in endoNF. This argues for significant differences in endo- and exosialidase reaction mechanisms.

By soaking experiments with pentameric sialic acid, a di-sialic acid binding site was identified to be located in the β -barrel domain of endoNF (*spheres* bound to the *green* subunit). The non-reducing end of the di-sialic acid points towards the active site cleft of the adjacent *blue* subunit which indicates that endoNF is a catalytic trimer with substrate binding and cleavage modules located in different subunits.

The stalk region of the mushroom shaped trimer is composed of the intertwining C-terminal tailspike domains of endoNF and folds into a left-handed triple β -helix that is interrupted by a small triple β -prism domain (Stummeyer *et al.*, 2005). Both folds are common for phage tailspike proteins and are considered to attribute to the unusual complex stability of these proteins (Chappell *et al.*, 1997; van Raaij *et al.*, 2001; Kanamaru *et al.*, 2002; Weigele *et al.*, 2003; Weigele *et al.*, 2005). The trimeric endoNF stalk domain is stabilised by an extensive network of intersubunit hydrogen bonds and hydrophobic interactions. More than 60% of the monomer’s solvent accessible surface is buried in the trimer (Stummeyer *et al.*, 2005).

However, the stalk domain appears to function not only in stabilising the catalytic trimer, but also in substrate binding, since a sialic acid binding site has been identified in the β -prism domain of the stalk (*spheres* bound to the *red* subunit). The amino acids Arg-837, Ser-848 and Gln-853 were shown to interact with a sialic acid as identified by soaking experiments. In principle, a polySia chain that is wrapped around the trimer of endoNF could simultaneously interact with all three subunits, for instance with the sialic acid binding site of the *red* subunit, the di-sialic acid binding site of the *green* subunit and the active site of the *blue* subunit (Stummeyer *et al.*, 2005). Yet, a more detailed picture of endoNF-

polySia interactions remained to be elucidated by further structure-function based analyses and was part of this present study.

3.5 – Biological and Biomedical Applications of Endosialidases

Phage-borne endosialidases are currently the only known enzymes that specifically degrade polySia. They catalyse a highly specific degradation of polySia that does not interfere with mono- or short oligosialylated structures. Therefore, these enzymes are widely applied in polySia research, including neuro- and oncobiology. Removal of polySia by treatment with endosialidase demonstrated intervention of polySia in dynamic cellular processes as different as migration of neuronal precursor cells, axonal outgrowth, synaptogenesis, physiological and morphological synaptic plasticity, and control of circadian rhythm (Vimr *et al.*, 1984; Rutishauser *et al.*, 1985; Becker *et al.*, 1996; Dityatev *et al.*, 2000; Aalto *et al.*, 2001; Durbec and Cremer, 2001; Jokilampi *et al.*, 2004; Dityatev *et al.*, 2004; Burgess *et al.*, 2007; Freiburger *et al.*, 2007; Oltmann-Norden *et al.*, 2007). Additionally, endosialidases may even be of interest as therapeutic agents. For instance, they have been described to efficiently improve the conditions in septicaemia and meningitis in *E. coli* K1 infected rat (Mushtaq *et al.*, 2005), and the application of the enzyme in a mouse model system led to a significant reduction of the metastatic potential of polySia positive tumor cells (Daniel *et al.*, 2001).

An additional novel aspect of endosialidase applications is given by a recent research initiative that evaluates the poorly immunogenic polysialic acid as a potential scaffold material for tissue engineering. These studies are carried out by an interdisciplinary research group in Hannover which integrates departments from the *Medizinische Hochschule Hannover*, the *Leibniz Universität Hannover* and the *Deutsche Institut für Kautschuktechnologie e.V.* One part of the project comprises the induced and controlled degradation of polySia-based scaffold materials. Since no endogenous polysialic acid degrading activity has been detected in mammals as yet, endosialidases should exhibit an excellent tool to induce the removal of artificial polysialic acid scaffolds at well-defined time points (Gerardy-Schahn *et al.*, 2004).

In summary, the versatile applications of endosialidases described above clearly illustrate the relevance and biotechnological potential of these enzymes. However, a more detailed understanding of the molecular mechanisms underlying the complex formation and maturation of endosialidases as well as their properties in binding to and cleavage of polySia is required to improve and extend their range of potential biological and biomedical applications.

3.6 – Objectives

Little is known about the evolution of K1-specific bacteriophages due to limited available genetic information. In line with the concept of modular evolution, the question arises, whether the different K1-phages evolved from recombination events that resulted in incorporation of endosialidase genes by different progenitor phages, or whether they emerged by divergent evolution from a single progenitor phage. Therefore, the first aim of the present study was to gain insight into the evolution of anti-K1 phages. To this end, a comparative genome analysis of the lytic coliphages K1E and K1F, and the temperate K1-phage CUS-3 was performed and evaluated.

The major part of this work was focussed on structure-function related studies of endoNF, the endosialidase of coliphage K1F. The enzyme contains a C-terminal part that is proteolytically cleaved off from the mature protein and is not found in the mature complex. This work aimed at revealing, whether the C-terminal part fulfils a chaperone-like function or acts as an oligomerisation platform, as previously suggested (Mühlenhoff *et al.*, 2003). Since the cleavage is no prerequisite for enzymatic activity, efforts were undertaken to elucidate the biological significance for the removal of this part from the mature enzyme. Furthermore, the impact of the C-terminal region on the complex stability or the functionality of endosialidases in terms of enzymatic activity or binding to polySia was investigated. Since the C-terminal region is common to endosialidases and several otherwise unrelated bacteriophage tailspike and fibre proteins, it was investigated whether the C-terminal part fulfils a more general function in these proteins.

This present study was carried out as part of the DFG Research Unit 548 that aimed at evaluating polySia as a new basis material for the use as a scaffold in nerve regeneration. Another objective of the present work was therefore to study whether endoN could be used to control epitope conservation in chemically and physicochemically modified polySia scaffolds. Materials tested in this part of the study, were synthesised at the *Institut für Organische Chemie* of the *Leibniz Universität Hannover*.

Chapter 4 – Evolution of Bacteriophages Infecting Encapsulated Bacteria: Lessons from *Escherichia Coli* K1 Specific Phages

This manuscript has originally been published in *Molecular Microbiology*.

'Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1 specific phages'

Katharina Stummeyer¹, **David Schwarzer**¹, Heike Claus², Ulrich Vogel², Rita Gerardy-Schahn¹, and Martina Mühlenhoff^{1*}

¹ *Abteilung Zelluläre Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Carl Neuberg-Str. 1, 30625 Hannover, Germany*

² *Institut für Hygiene und Mikrobiologie, Universität Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg, Germany*

* For correspondence. E-mail muehlenhoff.martina@mh-hannover.de; Tel.: (+49) 511 532 9807; Fax: (+49) 511 532 3956.

Molecular Microbiology (2006) **60**(5), 1123–1135.

© 2006 The Authors

Journal compilation © 2006 Blackwell Publishing Ltd

Accepted 22 March 2006. First published online 21 April 2006.

DOI:10.1111/j.1365-2958.2006.05173.x

Preface – About the Manuscript

The following study gives an insight into the evolution of bacteriophages infecting *Escherichia coli* K1. The lytic phages K1E and K1F as well as the temperate phage CUS-3 were investigated in a comparative genome analysis. In the course of this study, I contributed to the annotation of the Coliphage genomes. Therefore, it was searched for putative open reading frames within the genome sequences and an NCBI database analysis was performed to identify homologous genes and a putative function of the potential gene product. We revealed that the three K1-phages K1E, K1F and CUS-3 were determined to belong to distinct phage groups. Since CUS-3 phage was assigned to the group of the temperate P22-like phages HK620, Sf6 and P22, that together remarkably differ in certain gene clusters, I performed a phylogenetic analysis of the four complete genomes and a subset of 17 common genes to elucidate the basic relationship between these phages.

Summary

Bacterial capsules are important virulence factors but also provide attachment sites for bacteriophages that possess capsule degrading enzymes as tailspike proteins. To gain insight into the evolution of these specialized viruses, we studied a panel of tailed phages specific for *Escherichia coli* K1, a neuroinvasive pathogen with a polysialic acid capsule. Genome sequencing of two lytic K1-phages and comparative analyses including a K1-prophage revealed that K1-phages did not evolve from a common ancestor. By contrast, each phage is related to a different progenitor type, namely T7-, SP6-, and P22-like phages, and gained new host specificity by horizontal uptake of an endosialidase gene. The new tailspikes emerged by combining endosialidase domains with the capsid-binding module of the respective ancestor. For SP6-like phages, we identified a degenerated tailspike protein which now acts as versatile adaptor protein interconnecting tail and newly-acquired tailspikes and demonstrate that this adapter utilizes an N-terminal undecapeptide interface to bind otherwise unrelated tailspikes. Combining biochemical and sequence analyses with available structural data, we provide new molecular insight into basic mechanisms that allow changes in host specificity while a conserved head and tail architecture is maintained. Thereby, the present study contributes not only to an improved understanding of phage evolution and host-range extension but may also facilitate the on purpose design of therapeutic phages based on well characterized template phages.

Introduction

Capsular polysaccharides have long been identified as important virulence factors of many pathogenic bacteria causing severe invasive infections including septicaemia, meningitis, pneumonia, osteomyelitis, septic arthritis and pyelonephritis (Cross, 1990; Moxon and Kroll, 1990; Taylor and Roberts, 2005). As highly hydrated polymer gels they provide a thick (400 nm or more) layer protecting the bacterium from hostile environments and host immune defence. By masking underlying surface structures, capsules can confer resistance against complement-mediated lysis, which is a crucial step in the development of systemic infections. However, the capsule is also the Achilles heel of the bacterium since it provides an attachment site for specialized bacteriophages. These phages possess specific capsule depolymerases as tailspike proteins (Stirm and Freund-Mölbart, 1971), which enable the phage to penetrate the capsule and to gain access to the outer

membrane. Although evidence for phage-borne polysaccharide depolymerases has been recorded for almost 50 years (Adams and Park, 1956), little is known about the evolution of phages infecting encapsulated bacteria. To understand in more detail how these phages have evolved, we used a set of *Escherichia coli* K1 (*E. coli* K1) specific phages as model system. *E. coli* K1 is a leading pathogen in neonatal sepsis and meningitis (Robbins *et al.*, 1974; Sarff *et al.*, 1975) associated with high rates of mortality and severe neurologic sequelae (Kaper *et al.*, 2004; Saez-Llorens and McCracken, Jr., 2003). The K1 capsule, composed of α 2,8-linked polysialic acid (polySia) with up to 200 residues per polymer chain, is poorly immunogenic due to structural identity with host polySia present as posttranslational modification of the neural cell adhesion molecule NCAM (Mühlenhoff *et al.*, 1998). The capsular polysaccharide represents the major virulence factor of *E. coli* K1 and is essential for serum resistance (Leying *et al.*, 1990) and viable passage across the blood brain barrier (Kim, 2003).

Lytic bacteriophages specific for *E. coli* K1 have been isolated from sewage samples and belong to the class of linear double-stranded DNA viruses (Gross *et al.*, 1977; Kwiatkowski *et al.*, 1982; Kwiatkowski *et al.*, 1983; Miyake *et al.*, 1997; Scholl *et al.*, 2001; Smith and Huggins, 1982; Vimr *et al.*, 1984). Though different morphologies were found, *podoviridae* with short non-contractile tails and *myoviridae* with a long and contractile tail apparatus, a common feature of all K1-phages is the presence of an endosialidase which selectively degrades α 2,8-linked polySia (Finne and Mäkelä, 1985; Hallenbeck *et al.*, 1987; Pelkonen *et al.*, 1989). Endosialidase tailspike genes have been cloned from different phages (Gerardy-Schahn *et al.*, 1995; Long *et al.*, 1995; Mühlenhoff *et al.*, 2003; Petter and Vimr, 1993) and recently, we solved the first crystal structure of an endosialidase using the recombinantly expressed enzyme from phage K1F (Stummeyer *et al.*, 2005). The protein forms a homo-trimer with a unique modular architecture combining structural elements characteristic for sialidases (6-bladed β -propeller) and bacteriophage tailspike proteins (triple α -helix and triple β -prism).

So far, little is known on the evolutionary origin of K1-specific phages and genetic information is very limited because only one completely sequenced genome of a lytic K1-phage with dual host specificity for *E. coli* K1 and K5 (Scholl *et al.*, 2004) is available. Recently, a 40 kb prophage termed CUS-3 was identified in the partially sequenced genome of *E. coli* K1 strain RS218 (Deszo *et al.*, 2005), representing the first temperate phage with an endosialidase gene. By sequencing two new genomes of lytic K1-phages and

comparative genome analysis including the temperate phage CUS-3, we provide detailed insight into the evolution of K1-phages. We furthermore cloned the prophage endosialidase and showed by biochemical analysis including endosialidases from 3 different K1-phages that these specialized tailspikes emerged from extensive shuffling of catalytic- and capsid-binding domains. Together with the identification and biochemical characterization of a versatile tailspike adapter protein, our data provide insight into different molecular mechanisms by which K1-phages replaced the tailspike or tail fibre of their respective progenitor phages with capsule degrading endosialidases – the main incident towards K1-specificity.

Results and discussion

Modular architecture of endosialidase tailspike proteins

Genes encoding active endosialidases have been cloned from *E. coli* K1 specific phages CUS-3 (this study), K1F (Mühlenhoff *et al.*, 2003), and K1E (Gerardy-Schahn *et al.*, 1995; Long *et al.*, 1995) which is virtually identical to the one identified in the genome of phage K1-5 (Scholl *et al.*, 2001). As shown in Fig. 1A, all endosialidases share a highly conserved catalytic part followed by a C-terminal assembly domain which is essential for proper folding before it is released by proteolytic cleavage (Mühlenhoff *et al.*, 2003). The N-terminal parts, however, differ remarkably. The first 113 amino acids of the CUS-3 endosialidase, here termed endoNK1, share 75-95% identity with the corresponding parts of the tailspike proteins of Salmonella phage P22, Shigella phage Sf6 and Coliphage HK620 (Fig. 1B). The N-terminal part of the K1F derived enzyme (endoNF) shows about 40% identity to the first 134 amino acids of the tail fibre proteins (gp17) of Coliphages T7 and T3 and Yersiniophage ΦYeO3 (Pajunen *et al.*, 2001). By contrast, the endosialidases of phages K1E (endoNE) and K1-5 (endoN1-5) lack an extended N-terminal module and their first 11 amino acids are shared by gp47, one of two tailspike proteins of Salmonella phage SP6.

The N-terminal domains of P22 tailspikes and T7 fibres mediate attachment to the capsid (Steinbacher *et al.*, 1997; Steven *et al.*, 1988). To verify that in endosialidases the corresponding parts have a similar function and are therefore dispensable for enzymatic activity, we deleted the first 145, 245, and 38 amino acids of endoNK1, endoNF, and endoNE, respectively. All truncated variants were expressed as soluble proteins of the calculated molecular mass (see Fig. 1A,C) and enzymatic activities in the range of the wild-type proteins were detected (Fig. 1D). These results clearly

indicate that the N-terminal domains are neither required for proper folding nor for enzymatic activity, providing further evidence for an independent function in attaching the endosialidase to the phage tail.

Comparative genome analysis of the temperate K1-phage CUS-3

Analysis of the partially sequenced genome of *E. coli* K1 strain RS218 (<http://www.genome.wisc.edu>) revealed the presence of a prophage which harbours an endosialidase gene. In parallel to our work, this phage was also identified by Deszo *et al.* (2005) and termed CUS-3. The CUS-3 genome is inserted into the *argW* tRNA gene of *E. coli* K1 and encompasses 40,184 bp flanked by two 5'-aatgggtccc-3' repeats. 40% of the genome are 91%, 64%, and 59% identical with the genomes of Coliphage HK620, Shigella phage Sf6 and Salmonella phage P22, respectively. All four phages share a highly conserved organization of the virion genes with the tailspike gene located downstream of the capsid genes (Fig. 2). Since HK620, Sf6 and P22 infect O-antigen expressing strains, their tailspikes are endorhamnosidases specific for the respective host lipopolysaccharide (LPS). However, high sequence similarity is restricted to the common capsid-binding domain. In CUS-3, the endosialidase gene is positioned at exactly the same locus, neatly fused to a capsid-binding domain that is almost identical to the one shared by the tailspikes of the P22-like phages. Strikingly, all capsid proteins of CUS-3 are similar to the P22 cognate proteins, suggesting a highly conserved head and tail architecture. CryoEM data revealed that the P22 tail machine is formed by a central needle which is fixed to the tail by six copies of gp10 (Tang *et al.*, 2005). Attachment of the six trimeric tailspikes is mediated by a large contact area between their capsid-binding domains and the tail tube proteins gp4 and gp10 (Tang *et al.*, 2005). The latter protein might be the major interaction partner since gp10 is highly conserved (>92% amino acid identity) within the P22-like phages, whereas gp4 shares only 40% identity. Since the capsid-binding domain is the only contact site between tailspike and phage tail, exchange of the original endorhamnosidase part by an endosialidase is fully compatible with a highly conserved tail architecture.

CUS-3 also shares the common characteristic of P22-like phages that lysogenization results in modification of the host surface, a strategy to prevent superinfection by homologous phages (Allison and Verma, 2000). The lysogenic conversion genes of P22, HK620 and Sf6 are located downstream of the tailspike genes (Fig. 2) and their translation products catalyse glucosylation or acetylation of the bacterial O-antigen (Verma

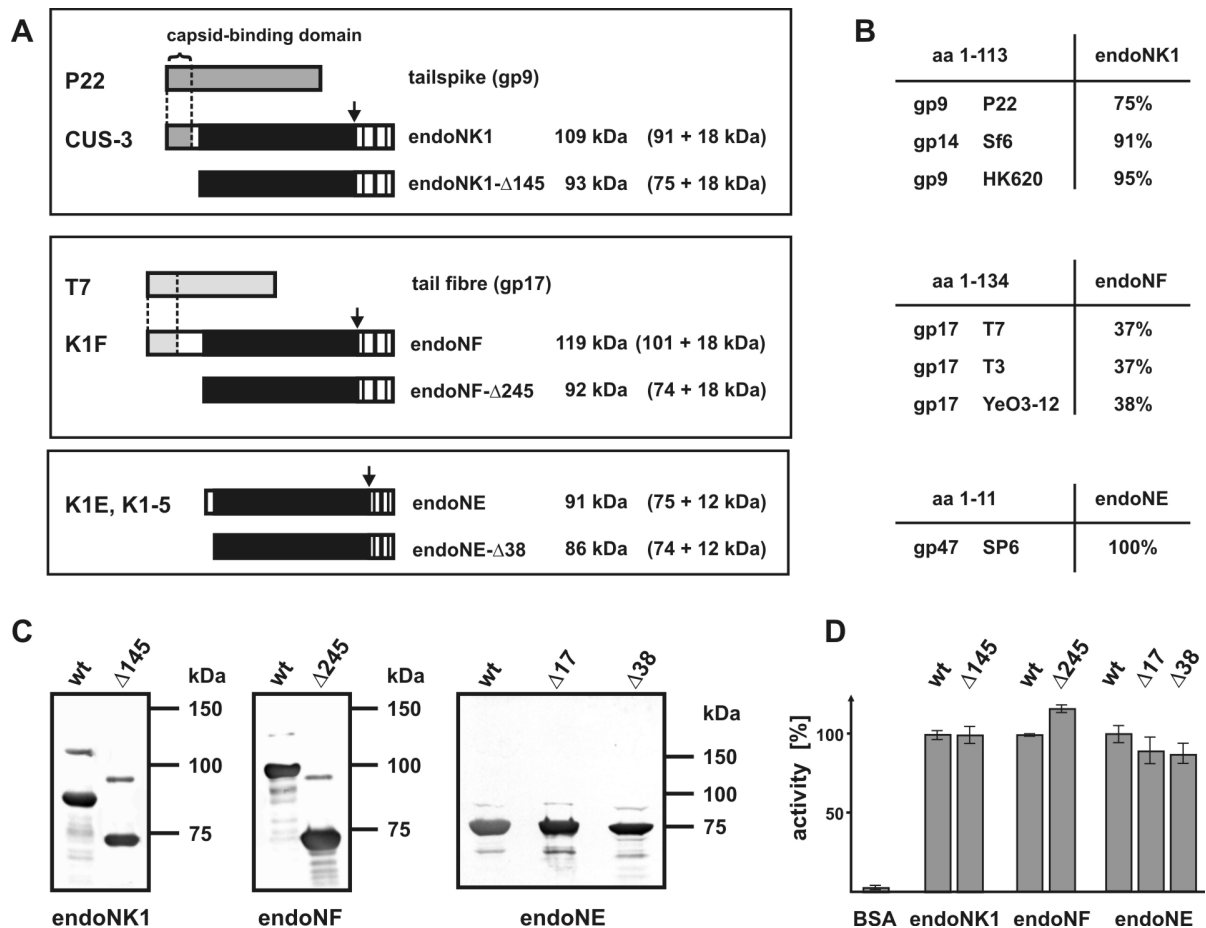


Fig. 1. Modular composition of endosialidases.

A. Schematic representation of the endosialidases of phages CUS-3 (endoNK1), K1F (endoNF), K1E (endoNE), and K1-5 (identical with endoNE). Amino acid stretches conserved in all endosialidases are shown in black and stretches with sequence similarities to the capsid-binding domain of tailspike protein gp9 of Salmonella phage P22 or tail fibre protein gp17 of Coliphage T7 are shown in dark and light grey, respectively. Endosialidase variants lacking the variable N-terminal part are indicated by the number of deleted amino acids ($\Delta 145$, $\Delta 245$ and $\Delta 38$ for endoNK1, endoNF, and endoNE, respectively). Molecular masses are given for the primary translation products and in brackets for the two fragments generated by proteolytic cleavage at a highly conserved serine residue as indicated by an arrow.

B. Sequence comparison between the N-terminal endosialidase modules with other phage proteins. Sequence identities between the indicated amino acids (aa) were determined by ClustalW using sequences with the following accession numbers: AM084332 (endoNK1), AJ505988 (endoNF), X78310 (endoNE), AAF75060 (Salmonella phage P22, gp9), AAQ12204 (Shigella phage Sf6, gp14), AAK28905 (Coliphage HK620, gp9), AAP33957 (Coliphage T7, gp17), CAC86305 (Coliphage T3, gp17), CAB63638 (Yersiniophage YeO3-12, gp17), and AAL86890 (Salmonella phage SP6, gp47).

C. Western blot analysis of wild-type and N-terminally truncated endosialidases. EndoNK1, endoNF and endoNE variants lacking the indicated number of amino acids were expressed in *E. coli* BL21(DE3). Protein lysates were separated by SDS-PAGE and the N-terminally epitope-tagged endosialidases were detected by Western blot analysis with anti-T7-tag antibody.

D. Relative enzymatic activity of wild-type and N-terminally truncated endosialidases. Endosialidase activity was determined in protein lysates with similar expression levels and activity of the respective wild-type enzyme was set to 100%.

et al., 1991; Weintraub *et al.*, 1992). In the same locus, CUS-3 harbours an O-acetyltransferase gene (*oat*) which mediates phase-variable O-acetylation of the polysialic acid capsule of *E. coli* K1 (Deszo *et al.*, 2005; Higa and Varki, 1988). However, in contrast to the integral acetyltransferase of Sf6 (Slauch *et al.*, 1996), the CUS-3 enzyme is a soluble protein that belongs to the left-handed β -helix family (M. Mühlenhoff, unpublished; Jenkins and Pickersgill, 2001).

In P22, tailspike and lysogenic conversion genes are separated from the virion genes by the *imml* regulatory gene which includes the superinfection exclusion gene *sieA*, the repressor genes *mnt* and *arc*, and the antirepressor gene *ant* (Vander and Kropinski, 2000). This region is completely deleted in Sf6 while in HK620

remnants of the P22 sequence are found together with two genes (*hkcB* and *hkcC*) that are not present in P22. In CUS-3, the *imml* region resembles a mixture of HK620 and P22 sequences, with *hkcB* of HK620 and the *ant* and *mnt* genes of P22 (Fig. 2). In addition, two open reading frames (ORFs) with sequence similarity to proteins of unknown function from *Photorhabdus luminescence* are inserted which have no counterpart in other phage genomes. Differences to P22 are also observed for the CUS-3 integrase which shows no significant similarity to its P22 counterpart but instead shares 98-99% identity with the Sf6 and HK620 cognates. The overall high nucleotide sequence identity between CUS-3 and HK620 indicates that CUS-3 evolved from an HK620-like phage which already showed

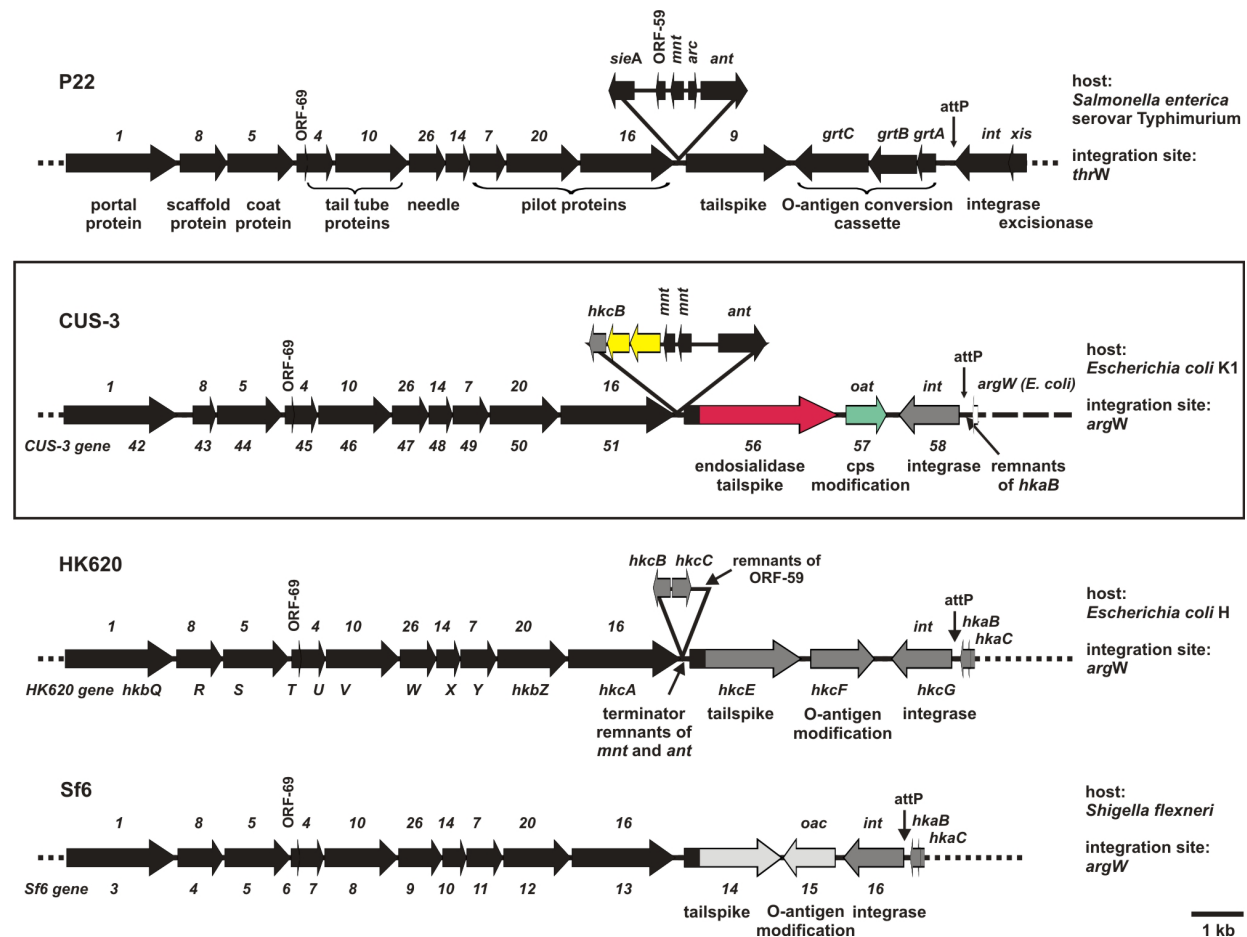


Fig. 2. Comparative genome analysis of *E. coli* K1 prophage CUS-3. The second half of the CUS-3 genome encompassing virion, lysogenic conversion and integrase genes is shown in comparison with the corresponding genome parts of Salmonella phage P22, Coliphage HK620 and Shigella phage Sf6 with the accession numbers NC_002371, NC_002730 and NC_005344, respectively. ORFs and amino acid stretches of CUS-3 that are homologous to P22 and HK620 are shown in black and dark grey, respectively, while ORFs restricted to Sf6 are depicted in light grey. P22 gene numbers are given for all genomes above the ORFs, while the individual numbering systems for CUS-3, HK620 and Sf6 are given below. CUS-3 genes with no counterpart in P22-like phages are highlighted in red (endosialidase gene), green (O-acetyltransferase gene), and yellow (two ORFs with sequence similarity to two proteins of unknown function from *Photothabdus luminescens*). Location of phage attachment sites (*attP*) are indicated by an arrow.

considerable variations to P22 in the region downstream of the tailspike gene (see also supplementary Fig. S1). Differences observed in the *imml* region suggest that the endosialidase was acquired by an ancestor of HK620 before all P22-like genes of the *imml* region were lost.

Comparative genome analysis of Φ K1F

The double stranded linear genome of the lytic K1-phage K1F consists of 39,699 bp of DNA flanked by terminal repeats of 178 bp. Sequence analysis revealed a close relationship to Coliphage T7 and the related Yersiniophages Φ A1122 and Φ YeO3-12 and Coliphage T3 (Garcia *et al.*, 2003; Pajunen *et al.*, 2001; Pajunen *et al.*, 2002). Of the 58 ORFs identified in K1F, 40 have a homologue in T7 (Dunn and Studier, 1983), including all essential proteins with known biological functions. On protein level, sequence identity ranges from 28 to 80% (see Table S2 in the supplementary material) with an average of 54%. In addition, 18

unique ORFs were identified which are mainly located in the first half of the genome, encompassing the genes expressed during early and mediate phase of infection. In contrast to T7, the K1F genome contains a putative self-splicing group I intron integrated into the coding sequence of the DNA-polymerase gene (gene 5, Table S2). The intron seems closely related to two recently identified introns in phages Φ I and W31 (Bonocora and Shub, 2004), since all three are inserted into the same position of gp5 and encode homing endonucleases (gene 5.3, Table S2) that are 78% identical on protein level. During submission of this manuscript, the genomic sequence of phage K1F was also published by Scholl and Merrill (2005). In comparison to their study, we annotated 16 additional ORFs, including three ORFs (1.2, 4.2, and 19.2) that show sequence similarity to cognate proteins in T7 (see Table S2). Seven of the ORFs identified in both studies (1.7, 2.5, 3.2, 4A, 5.5, 6.8, and 19.0) slightly vary in the number of amino acids either due to differences in the predicted start sites or due to minor sequence variations.

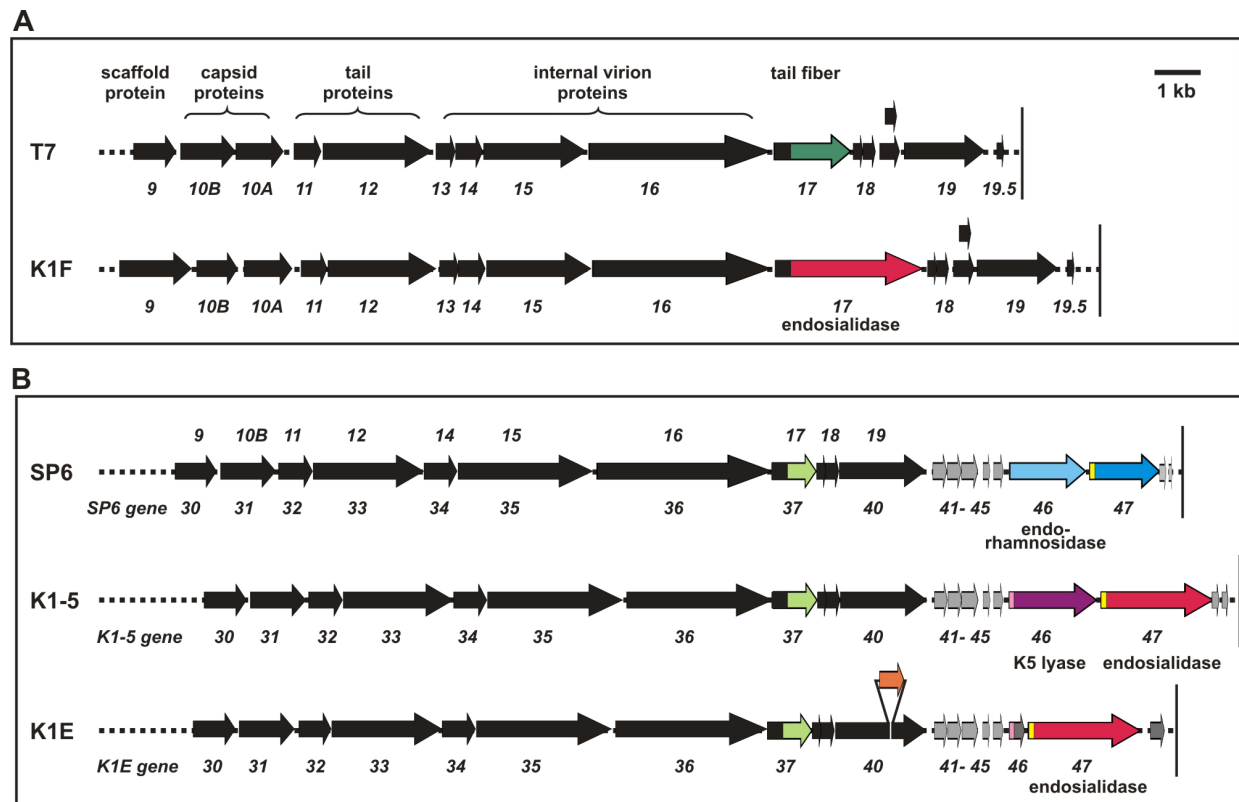


Fig. 3. Comparative genome analysis of the lytic *E. coli* K1 phages K1E and K1F. The second half of the genomes of K1E and K1F comprising mainly the virion genes is shown in comparison to the respective genome parts of closely related phages.

A. K1F (AM084414) in comparison with Coliphage T7 (NC_001604).

B. K1E (AM084415) in comparison with Salmonella phage SP6 (AY370673) and Coliphage K1-5 (AY370674).

ORFs and amino acid stretches that are homologous to T7 are shown in black, while ORFs depicted in light grey are merely conserved among the SP6-subfamily. Individual gene numbers are given below all ORFs, while for SP6 additionally the T7-numbering system is depicted. Unique ORFs exclusively found in K1E are shown in dark grey and an ORF encoding a homing endonuclease that is located in a putative group I intron is shown in orange. Tailspike and tail fibre genes are highlighted by different colours with endosialidase genes shown in red. The 5'-regions encoding identical amino acid stretches are shown in yellow for the sequence MIQRLGSSLVK and pink for MAKLTQPKTT. Dark and light green indicates the distant relationship between T7 gp17 and gp37 of SP6, K1-5 and K1E.

The genomes of K1F and T7 are largely co-linear and show an identical organization in the second half of the genome, encompassing the late genes with morphogenetic functions (Fig. 3A). Primary determinant of the host range in T7-like phages is gp17 which builds up the six kinked tail fibres that mediate adsorption to bacterial LPS (Steven *et al.*, 1988). For T7-mediated lysis, the polySia capsule of *E. coli* K1 provides the only barrier and can be overcome by the addition of exogenous endosialidase (Scholl *et al.*, 2005). Accordingly, K1F has acquired new host specificity by exchanging gene 17 with an endosialidase gene (Fig. 3A). However, the original gene was not completely displaced and the N-terminal capsid-binding part of gp17 was retained for proper connection of the new tailspike to the otherwise conserved capsid proteins.

Comparative genome analysis of Φ K1E

Sequence analysis of the lytic K1-phage K1E revealed that this phage also belongs to the T7 supergroup. Yet in contrast to K1F, K1E is much closer related to Coliphage K1-5 and Salmonella phage SP6 than to T7,

suggesting that K1E is the third phage to be grouped into the SP6 subgroup (Dobbins *et al.*, 2004; Scholl *et al.*, 2004). The linear genome of K1E encompasses 45,251 bp which are flanked by terminal repeats of 288 bp. Even on DNA level, 90% of the K1E genome is 92% identical to K1-5. This remarkably high sequence identity indicates a very close evolutionary relationship of the two phages, while SP6 is more distantly related with 78% nucleotide sequence identity over 60% of the genome. The genomic organization of K1E is essentially co-linear to those of K1-5 and SP6. Of the 62 ORFs identified in K1E, 48 have counterparts in K1-5 with 58–98% amino acid identity (see Table S3 in the supplementary material). Twelve ORFs were assigned that are unique for K1E and show no significant sequence similarity to any known protein. Most of these mainly smaller ORFs cluster in the first third of the genome (Table S3). Genome organization between SP6, K1-5 and K1E is particularly highly conserved in the second half of the genomes, encompassing the virion genes (Fig. 3B). However, a distinct feature of the K1E genome is a 644 bp fragment encoding a putative homing endonuclease (gp40.5, Table S3)

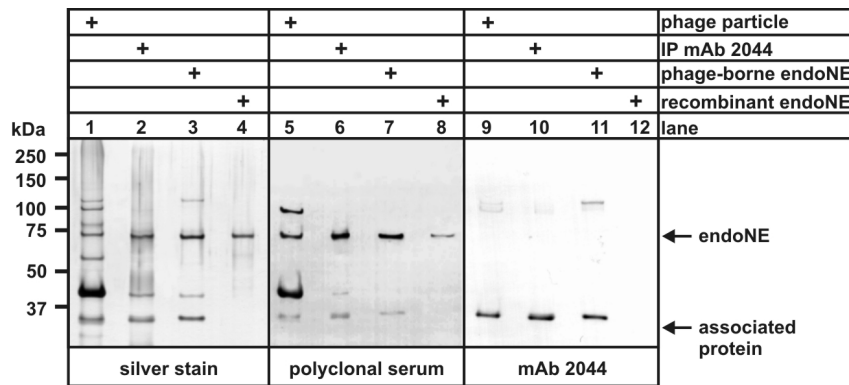


Fig. 4. Characterization of a 35 kDa protein tightly associated with phage-borne endoNE. Purified K1E phage particles, proteins isolated from K1E phage lysates by immunoprecipitation (IP) with mAb 2044, purified endoNE isolated from phage lysates (phage-borne endoNE), and purified recombinantly expressed endoNE were separated by SDS-PAGE. Proteins were displayed by silver stain (lanes 1–4) or Western blot analysis with a polyclonal antibody raised against whole phage particles (lanes 5–8) or with mAb 2044 (lanes 9–12). Bands corresponding to endoNE and the associated 35 kDa protein are indicated with an arrow.

which is inserted into gene 40, thereby splitting the gene into two open reading frames (gp40A and gp40B, Table S3). Initial analysis indicates that the insertion might be a self-splicing group I intron, an element that would still allow full-length expression of gp40, which is as putative large terminase subunit a crucial component of the DNA packaging machinery (Cerritelli *et al.*, 2003; Morita *et al.*, 1995).

Substantial diversity among the members of the SP6 subgroup is primarily seen within their tailspike genes which are located in a 3–4 kb region that is fused to the 3'-end of the genome. Interestingly, this region encompasses two tailspike loci, resulting in expanded host specificity due to the presence of two tailspike proteins, gp46 and gp47. In K1-5, the presence of an endosialidase (gp47) together with a lyase (gp46) specific for the capsular polysaccharide of *E. coli* K5 enables the phage to infect *E. coli* K1 and K5 (Scholl *et al.*, 2001). In SP6, gene 46 encodes a putative endorhamnosidase that is closely related to the P22 tailspike but lacks the corresponding capsid-binding domain (Dobbins *et al.*, 2004; Scholl *et al.*, 2004). The second tailspike of SP6 is not yet characterized but since phage SP6 is, compared to P22, not restricted to smooth *S. typhimurium* LT2 strains but can infect rough strains with equal efficiency (Scholl *et al.*, 2004), this might be due to the presence of gp47.

In the newly sequenced genome of K1E, the endosialidase gene 47 is found in the same locus as the virtually identical gene of K1-5 (Fig. 3B). Gene 46, however, encodes a small ORF of only 111 amino acids that might be a degenerated tailspike gene, explaining the single host specificity of K1E for *E. coli* K1. Interestingly, the first 10 amino acids of gp46 (MAKLTQPKTT) are shared by the N-terminal part of the gp46 lyase in K1-5 but not by the endorhamnosidase (gp46) of SP6, whereas the rest of the ORF shows no significant homology to any known protein. By contrast, all three gp47 proteins, the endosialidases of K1E and K1-5 as well as gp47 of SP6, have a common N-terminal MIQRLGSSLVK-peptide. Beside these

11 amino acids, no sequence similarity was found between endosialidases and the SP6 tailspike (gp47).

Identification of gp37 as a versatile tailspike adapter protein

As shown in Figure 1A, endoNE lacks a potential capsid-binding domain, raising the question how this tailspike is fixed to the phage. In contrast to endoNF, purification of endoNE from phage lysates always resulted in co-purification of an unknown protein which formed an SDS-resistant complex with endoNE (Gerardy-Schahn *et al.*, 1995; Tomlinson and Taylor, 1985). Using K1E phage particles for immunization, we obtained mAb 2044 which specifically recognizes a 35 kDa protein associated with K1E (Fig. 4, lanes 1 and 9). The protein was also detected in endoNE purified from phage lysates (lanes 3 and 11) but not in samples of the recombinantly expressed enzyme (lanes 4 and 12). Using mAb 2044 for immunoprecipitation, we co-precipitated endoNE from phage lysates (see lanes 2 and 10), confirming a tight association of the 35 kDa protein with endoNE. N-terminal peptide sequencing of the immunisolated material demonstrated, that the associated protein is encoded by gene 37 located 5.3 kb upstream of the endoNE gene (Fig. 3B). The 930 bp ORF encodes a 35 kDa protein of 320 amino acids and shares 73% and 65% sequence identity with gp37 of K1-5 and SP6, respectively. Subsequently, we cloned and recombinantly expressed the newly identified protein from K1E. In the absence of endoNE, gp37 assembled into a homo-trimer which was SDS-resistant at room temperature as indicated by a 105 kDa protein band observed in non-heated samples (Fig. 5A, lane 14). Co-expression of gp37 with endoNE resulted in assembly of an SDS-resistant high molecular weight complex which contained gp37, as shown by staining with mAb 2044, and endoNE, as shown by staining with an antibody specific for the T7-epitop-tag fused to the recombinantly expressed enzyme (Fig. 5A, lane 10). SDS-resistance of endoNE depends on association

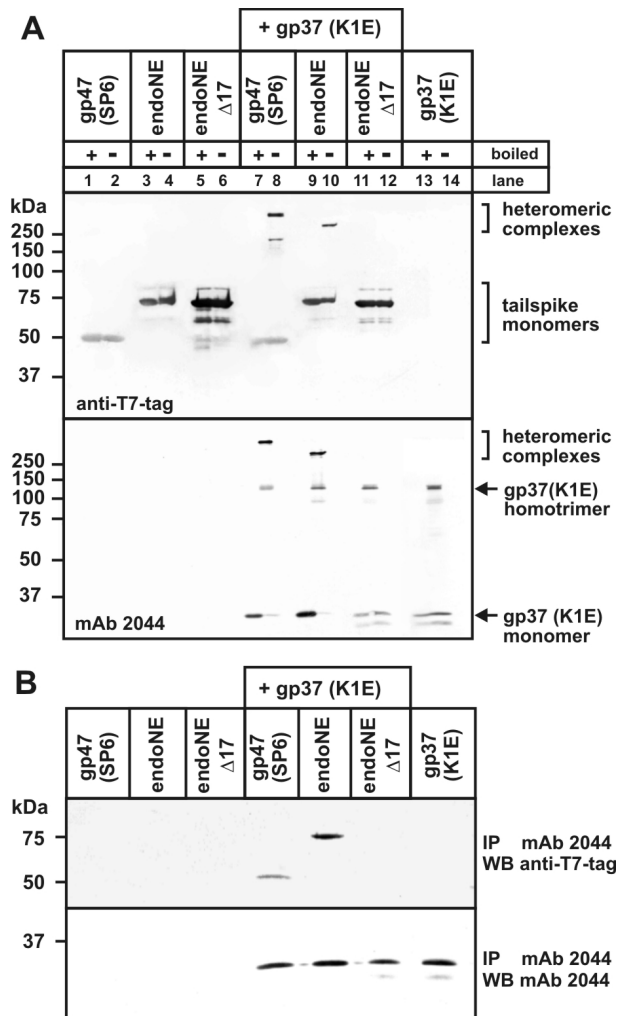


Fig. 5. Characterization of heteromeric complexes formed by gp37 of phage K1E.

A. The tailspike proteins gp47 of phage SP6, endoNE of phage K1E and the N-terminally truncated variant endoNE Δ 17 were either expressed separately (lanes 1-6) or in combination with gp37 of phage K1E (lanes 7-12). Individually expressed gp37 is shown in lanes 13 and 14. Soluble fractions of bacterial lysates were separated by SDS-PAGE. Western blot analysis with anti-T7-tag antibody for detection of the N-terminally epitope-tagged tailspike proteins is shown in the upper panel and staining with mAb 2044 recognizing gp37 is shown in the lower panel. Formation of SDS-resistant complexes was monitored by omitting the boiling step prior to electrophoresis.

B. Immunoprecipitation experiments with mAb 2044. Tailspike proteins and gp37 were either expressed individually or in combination as described above. Gp37 was immunoprecipitated from the soluble fraction of bacterial lysates. The precipitated material was analysed by SDS-PAGE and Western blotting (WB) using either anti-T7-tag antibody (upper panel) or mAb 2044 (lower panel).

with gp37 since no complexes were seen for the individually expressed enzyme (Fig. 5A, lane 4). Deletion of the first 17 amino acids of endoNE prevented assembly into the heteromeric complex (Fig. 5A, lane 12), indicating that the N-terminus of endoNE constitutes the binding site for gp37. As mentioned above, the first 11 amino acids of endoNE are shared by the second SP6 tailspike gp47 and after cloning and co-expression of gp47 with the K1E protein gp37 we could in fact demonstrate that both proteins assemble into a common complex (Fig. 5A, lane 8). The interaction of gp37

with both complex partners, endoNE and gp47 of SP6, was further confirmed by co-immunoprecipitation experiments using mAb 2044 (Fig. 5B). Together these experiments show that interactions of gp37 with the tailspike proteins endoNE of K1E and gp47 of SP6 depend on the N-terminal MIQRLGSSLVK-peptide which is the only common part in both tailspikes. However, further experiments are required to conversely dissect which part of gp37 is involved in binding of the tailspike proteins.

We next investigated whether gp37 of phage SP6 serves as an adapter protein for both of the SP6 tailspikes, gp46 and gp47. Individual expression of the SP6 gp37 resulted exclusively in insoluble protein (Fig. 6A). Also co-expression with gp46 proved to be difficult since only low amounts of the two proteins were expressed as soluble proteins. However, the level of soluble gp37 was markedly increased in the presence of gp47 (Fig. 6A). Furthermore, SDS-resistant complexes containing the SP6 adapter gp37 were only observed with proteins that exhibit the N-terminal MIQRLGSSLVK-peptide: SP6 gp47 and endoNE (Fig. 6B). By contrast, neither the SP6 nor the K1E gp37 formed a heteromeric complex with the P22-like tailspike gp46 (data not shown), confirming that interaction between adapter and tailspike protein depends on the MIQRLGSSLVK-peptide.

Comparative genome analysis revealed that gene 37 of SP6-like phages and the T7 tail fibre gene 17 are present in the same gene locus (see Fig. 3). Moreover, the corresponding translation products share sequence similarity, in particular within their N-terminal capsid-binding domains. Thus, it was already speculated that gp37 is capable of attaching the tailspikes of SP6-like phages to the virion (Dobbins *et al.*, 2004; Scholl *et al.*, 2004). In the present study we analysed the binding interactions biochemically and showed that gp37 mediates binding of the gp47 tailspikes by interaction with their common MIQRLGSSLVK-sequence. However, how the gp46 tailspikes are connected to the tail and if the N-terminal MAKLT KPKT sequence shared by gp46 of K1E and K1-5 is involved in capsid-binding is not yet known.

Evolution of K1-specific phages

All K1-phages identified so far possess endosialidase tailspike proteins. These are the key determinants of host specificity since they enable the phage to penetrate the thick polySia capsule of the host. In contrast to exosialidases which cleave single terminal sialic acid residues, endosialidases are highly specific for polySia and cleave within the polymer chain (Finne and Mäkelä, 1985; Hallenbeck *et al.*, 1987; Pelkonen *et al.*,

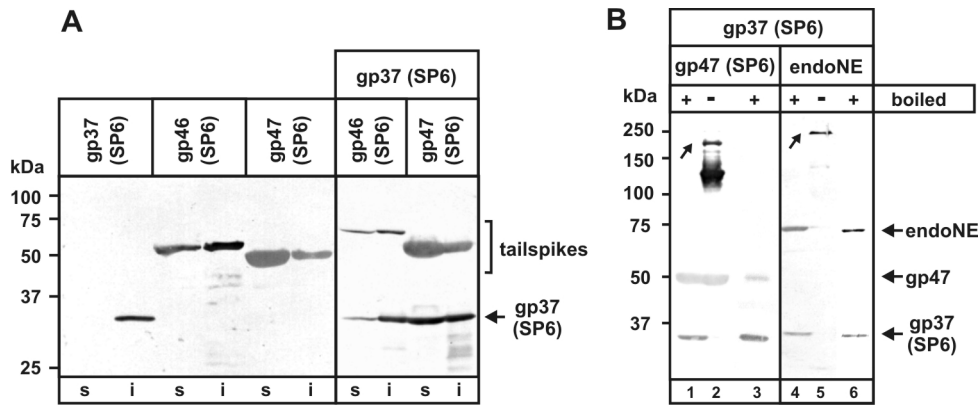


Fig. 6. Characterization of gp37 of phage SP6. (A) The adapter protein gp37 of phage SP6 and the SP6 tailspike proteins gp47 and gp46 were either expressed individually or in the indicated combinations. Soluble and insoluble fractions of bacterial lysates were analysed by SDS-PAGE and Western blotting using anti-T7-tag antibody. (B) The SP6 gp37 was co-expressed with either gp47 of phage SP6 or endoNE of phage K1E. Soluble bacterial lysates were analysed by SDS-PAGE and Western Blotting using anti-T7-tag antibody (lanes 1 and 4). The formation of SDS-resistant complexes was monitored by omitting the boiling step prior to electrophoresis and bands corresponding to heteromeric complexes are indicated with an arrow (lanes 2 and 5). To prove the presence of both, adapter and tailspike protein, indicated bands were excised and eluted proteins were reanalysed by SDS-PAGE after boiling (lanes 3 and 6). Bands corresponding to monomeric proteins are indicated on the right.

1989). Interestingly, endosialidases are exclusively found as tailspike proteins of K1-specific phages and have no counterpart in any other organism. Recently, we solved the crystal structure of the central catalytic part of endoNF which forms an SDS-resistant homotrimer (Stummeyer *et al.*, 2005). The structure revealed that a 6-bladed β -propeller, the basic structural element of all sialidases (Taylor, 1996), is fused to structural motifs characteristic for bacteriophage tail proteins: triple β -helix and triple β -prism (Kanamaru *et al.*, 2002; Steinbacher *et al.*, 1994; Stummeyer *et al.*, 2005; van Raaij *et al.*, 2001; Weigele *et al.*, 2003). This indicates that the endosialidase tailspike evolved from shuffling of a complete sialidase domain into the structural context of a phage tail protein with triple β -helix fold. Most probably, the sialidase is of bacterial origin since the β -propeller of endoNF shows higher structural similarity to bacterial than to viral sialidases (Stummeyer *et al.*,

2005) and contains two short Asp-boxes, structural motifs that were not found in viral sialidases. Sialidases were identified only in a restricted set of gram-positive and gram-negative bacteria, including *Arthrobacter ureafaciens*, *Micromonospora viridifaciens*, *Clostridium perfringens*, *Vibrio cholerae*, and *Salmonella typhimurium* (Taylor, 1996). Therefore, Salmonella phages like P22 or SP6 provide likely candidates for the first uptake and subsequent integration of a sialidase β -propeller encoding sequence into the existing tailspike gene. This potential evolutionary origin of endosialidases is supported by identification of P22- and SP6-related K1-phages (i. e. CUS-3 and K1E). Since the parallel β -helix fold of the P22 endorhamnosidase differs fundamentally from the canonical β -propeller of endoNF, it seems unlikely that endosialidases evolved from a P22-like tailspike by a series of point mutations or exchange of several small gene fragments as sug-

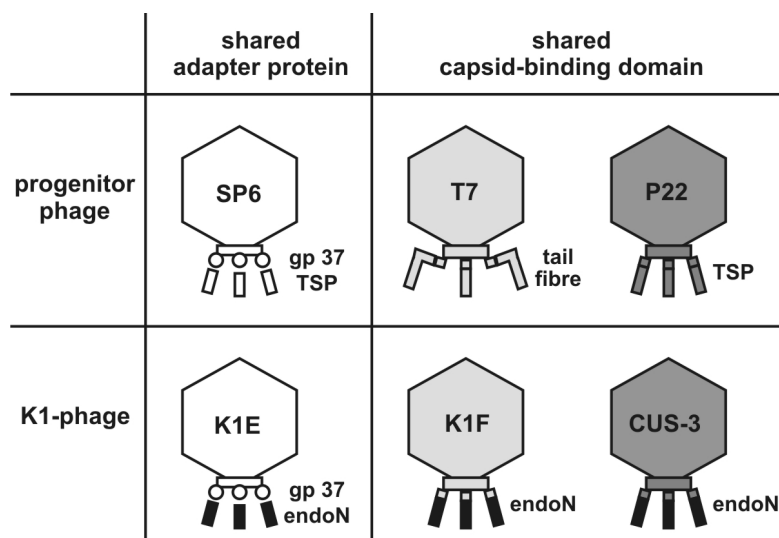


Fig. 7. Evolution of *E. coli* K1 specific phages. The three analysed K1-phages K1E, K1F and CUS-3 as well as prototypes of their respective progenitors are depicted schematically. Cognate phages are shown in identical grey scales (white, light grey, dark grey). The shared capsid-binding domains found in endosialidases (endoN) of K1F and CUS-3 as well as in the tailspikes of their related phages are coloured accordingly, while endosialidase domains are shown in black. The capsid-binding adapter (gp37) of phages K1E and SP6 is represented as open circle.

gested for the evolution of long tail fibre proteins (Haggard-Ljungquist *et al.*, 1992; Sandmeier *et al.*, 1992; Tetart *et al.*, 1998).

The high amino acid sequence identity between the catalytic parts of all known endosialidases shows that the endosialidase gene is well conserved which might be dictated by structural and functional constraints. Transmission of this gene to different phage genomes is mediated by horizontal gene transfer. The comparative genome analysis provided in the present study suggests that this step occurred by illegitimate recombination into the tailspike or tail fibre locus of a P22-, a T7-, and a SP6-like progenitor phage, resulting in the K1-phages CUS-3, K1F, and K1E, respectively. Even within the small subset of four K1-phages, members of three different phage groups were identified, indicating that the uptake of the endosialidase gene might have occurred with high frequency. Since simultaneous infection of the same bacterium by two lytic phages is a rare event (Brüssow and Hendrix, 2002; Lawrence *et al.*, 2002), the involvement of a temperate endosialidase carrying phage like CUS-3 may have played an important role in transmitting this tailspike to incoming lytic phages.

Notably, we identified two different strategies used by tailed phages to incorporate new tailspikes into their highly conserved capsid structures. As illustrated in Fig. 7, either the original capsid-binding domain of the progenitor phage is kept, resulting in fusion proteins combining capsid-binding and host range determining functions, or the newly acquired tailspikes are connected to the tail by a versatile adapter protein which requires only an N-terminal undecapeptide for binding. Interestingly, the interaction of the adapter protein gp37 is independent of the tailspike fused to the peptide, as shown for the two unrelated tailspike proteins gp47 of Salmonella phage SP6 and the endosialidase of K1E. This finding might be of particular relevance for the on purpose design of well characterized lytic phages with defined host specificity as required for phage therapy. With the increasing prevalence of antibiotic-resistant bacteria, phage therapy in general came to the fore again as alternative treatment strategy (Brüssow, 2005; Merril *et al.*, 2003; Thiel, 2004). However, due to the narrow host range of most phages, careful selection and sometimes isolation of a novel therapeutic phage or alternatively, the use of phage cocktails is required - strategies that are incompatible with Western regulatory agencies. In this context, the use of a well characterized phage as progenitor to engineer a large set of phages with different tailspikes would be a significant advantage. Detailed molecular insight into basic mechanisms underlying host range extension provided

in this study might help to design clearly defined phages for therapeutic use.

Experimental procedures

Materials

pET expression vectors and *E. coli* BL21(DE3) were obtained from Novagen. Purified phage DNA of Salmonella phage SP6 was kindly provided by Graham F. Hatfull (University of Pittsburgh, USA). The plasmid pMCL210 (Nakano *et al.*, 1995) was a gift from Yoshio Nakano (Kyushu University, Japan).

Propagation of bacteriophages and isolation of phage DNA

The *E. coli* K1 strain U9/41 (O2:K1:H4) was used for propagation of K1-specific bacteriophages (Gross *et al.*, 1977). Purification of phage particles and isolation of phage DNA was performed as described previously (Gerardy-Schahn *et al.*, 1995).

DNA sequencing, sequence assembly and analysis

Phage genomes of K1E and K1F were sequenced by a combination of shotgun sequencing and primer walking. Generation of the shotgun library and fluorescent dye dideoxy chain-terminating DNA sequencing was carried out at GATC Biotech (Konstanz, Germany) using an Applied Biosystems 3730 DNA analyzer. Collected sequence data were revised by removing poor-quality data and assembled using Seqman II (DNASTAR Inc.) resulting in a final 8-fold (K1F) and 5-fold (K1E) coverage. Open reading frames (ORFs) were analysed using GeneMark.hmm (Lukashin and Borodovsky, 1998) and EditSeq (DNASTAR Inc.) and scanned for sequence similarities to other proteins using BLASTP (Altschul *et al.*, 1997). Sequences of identified homologues were compared using ClustalW (Chenna *et al.*, 2003) or ALIGN (Pearson *et al.*, 1997). If not stated otherwise, amino acid identities were determined comparing full-length ORFs. Sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession numbers AM084414 (K1F), AM084415 (K1E) and AM084332 (endoNK1).

N-terminal amino acid analysis of K1E gp37

Using the monoclonal antibody (mAb) 2044 bound to protein G sepharose (Pharmacia), the endoNE/gp37 complex was immuno-isolated from freshly prepared phage lysates of K1E. The immunoprecipitated material was separated by 10% SDS-PAGE and proteins were blotted onto PVDF-membrane (Millipore). The Coomassie stained 35 kDa protein band was excised and N-terminal amino acid sequencing of the immobilized protein was performed at Giessen University (Biochemisches Institut am Klinikum) resulting in the following peptide sequence: ²SYTFTE(X)IANGTQVTY-PFSGAG(X)DKGYLRASDVIV³⁶.

Cloning of gene 37 of Coliphage K1E

Using purified K1E phage DNA as a template, the K1E gene 37 was amplified by PCR using the primers MM168A (5'-GGAATTCCATATGAGTTATACTTTCACAG-3') and MM169A (5'-CCGCTCGAGTTACTACTAGTGTAGCTTC-3') containing NdeI and XhoI sites (underlined), respectively. The PCR product was subcloned into the respective sites of pET23a and the sequence identity was confirmed by sequencing.

Cloning of genes 37, 46, and 47 of Salmonella phage SP6

SP6 genes were amplified by PCR using purified SP6 phage DNA and the following primer pairs: MM207/MM208 for gene 37 (5'-CGGGATCCATGAGT-TATACTTTCACAGAAC-3', 5'-GTCCGCTCGAGCTATAAAGTGACTACATCCG-3'), MM204/MM205 for gene 46 (5'-CGGGATCCATGTCATTAATAAACCACG-3', 5'-GTCCGCTCGAGTTAAGCAAGCTTCATTCGTG-3') and MM252/MM253 for gene 47 (5'-CGGGATCCATG-ATACAAAGATTAGGTTTC-3', 5'-GTCCGCTCGAGTCACTTGAATATGTAGCTG-3'). BamHI and XhoI sites (underlined) in forward and reverse primers, respectively, were used for subcloning of the PCR products into the BamHI/XhoI sites of pET23a, resulting in constructs with an N-terminal T7-epitope tag (MASMTGGQMQG). The sequence identity of all constructs was confirmed by sequencing.

Generation of wild-type and N-terminally truncated endosialidases

Constructs for the expression of wild-type endoNE and endoNF were constructed in pET23a as described (Mühlenhoff *et al.*, 2003). EndoNK1 was amplified by PCR using genomic *E. coli* K1 DNA as a template and the primer pair MM220/MM221 (5'-CGGGATCCATGACAGACATTACAGCC-3', 5'-GTCCGCTCGAGGTGCTTCCCTTCTATTAAG-3'). N-terminally truncated variants were generated by PCR using the respective wild-type constructs as a template and the following primer pairs: MM231/MM221 for endoNK1Δ145 (5'-CGGGATCCGCTATTGGTGATGGTGTTC-3', 5'-GTCCGCTCGAGGTGCTTCCCTTCTATTAAG-3'), MM147/MM98 for endoNFΔ245 (5'-CGGGATCCGCTAAAGGGGATGGTGTTC-3', 5'-GTCCGCTCGAGCTTCTGTTCGAAGCAGCAAAG-3'), MM257/MM89 for endoNEΔ17 (5'-CGGGATCCGGTGCAATCTGGCGTAAC-3', 5'-TCCGCTCGAGACTGATTTTATTAGTGGCATAAC-3'), and MM258/MM89 for endoNEΔ38 (5'-CGGGATCCGCCAAAGGTGATGGTAAG-3', 5'-TCCGCTCGAGACTGATTTTATTAGTGGCATAAC-3'). BamHI and XhoI sites (underlined) in forward and reverse primers, respectively, were used for subcloning of the respective PCR products into the BamHI/XhoI sites of pET23a, resulting in constructs with an N-terminal T7-epitope and a C-terminal hexahistidin tag. The identity of all constructs was confirmed by sequencing.

Constructs for co-expression experiments

For co-expression of tailspike proteins with gp37 we used the above described pET23a-based tailspike

constructs and pMCL210-based constructs with PA15 origin of replication and chloramphenicol resistance for the expression of gp37 of either K1E or SP6. To generate the latter constructs, we isolated the respective BglIII/XhoI fragments from the pET23a-based gp37 constructs described above and ligated the isolated fragments containing T7-promoter, ribosomal binding site, and the gp37 coding sequence into the BamHI and XhoI sites of pMCL210. The identity of both 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 or in the case of co-expression experiments containing 100 µg/ml carbenicillin and 34 µg/ml chloramphenicol. At an optical density (OD₆₀₀) of 0.6 expression was induced by adding 1 mM IPTG and bacteria were harvested 2 h after induction. Cells were resuspended in 50 mM Tris-HCl pH 8.0/ 2 mM EDTA and incubated for 15 min at 30°C in the presence of 100 µg/ml of lysozyme. Bacteria were lysed by sonication and soluble fractions were obtained after centrifugation (12,000xg, 15 min, 4°C).

Determination of endosialidase activity

Endosialidase activity was determined as described earlier (Stummeyer *et al.*, 2005). Briefly, microtiter plates (Becton Dickinson) were coated for 1 h with 20 µl/well purified capsular polysaccharide (1 µg/ml) of *E. coli* K1 (Weisgerber *et al.*, 1990). Non-specific binding sites were blocked with 1% BSA in PBS (10 mM sodium phosphate pH 7.4, 150 mM NaCl) and endosialidase containing samples were applied for 30 min at 37°C. After three washing steps in PBS the remaining uncleaved substrate was detected by immunostaining using the polySia specific mAb 735 (Frosch *et al.*, 1985). The enzymatic activity of different samples was compared by endpoint dilution, whereby loss of mAb 735 reactivity indicated endosialidase activity.

SDS-PAGE, silver staining, and immunoblotting

SDS-PAGE was performed under reducing conditions using 2.5% (v/v) β-mercaptoethanol. Gradient gels were prepared with ProSieve 50 gel solution (Bio Whittaker) and run in Tris/Tricine electrode buffer according to the manufacturer's instructions. Silver staining and Western blot analysis were performed as described previously (Gerardy-Schahn *et al.*, 1995). Proteins containing a T7-epitope tag were detected with 0.1 µg/ml anti-T7 antibody (Novagen), gp37 of phage K1E was detected with mAb 2044 using 5 µg/ml, and a polyclonal guinea pig serum raised against purified K1E phage particles (Gerardy-Schahn *et al.*, 1995) was used at a dilution of 1:1000. To prove the identity of heteromeric complexes containing two T7-tagged proteins, corresponding bands were excised, cut into small pieces, covered by PBS, and stored over night at 4°C. Eluted proteins were separated from gel slices by centrifugation through a 0.22 µm filter and filtrates were concentrated 3-fold by ultra-filtration using a 10 kDa cut-off membrane. After boiling in Laemmli sample buffer, eluted proteins were reanalysed by SDS-PAGE and Western blotting as described above.

Acknowledgements

We thank Dr. Dietmar Linder for protein sequencing and Melanie Grove, Astrid Oberbeck und Andrea Bethe for excellent technical assistance. This work was supported by grants from the Deutsche Forschungsgemeinschaft to R.G.-S., U.V., and M.M. and the Fonds der Chemischen Industrie.

References

- Adams, M.H., and Park, B.H. (1956) An enzyme produced by a phage-host cell system. II. The properties of the polysaccharide depolymerase. *Virology* **2**: 719-736.
- Allison, G.E., and Verma, N.K. (2000) Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8**: 17-23.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.
- Bonocora, R.P., and Shub, D.A. (2004) A self-splicing group I intron in DNA polymerase genes of T7-like bacteriophages. *J Bacteriol* **186**: 8153-8155.
- Brüssow, H. (2005) Phage therapy: the *Escherichia coli* experience. *Microbiology* **151**: 2133-2140.
- Brüssow, H., and Hendrix, R.W. (2002) Phage genomics: small is beautiful. *Cell* **108**: 13-16.
- Cerritelli, M.E., Conway, J.F., Cheng, N., Trus, B.L., and Steven, A.C. (2003) Molecular mechanisms in bacteriophage T7 procapsid assembly, maturation, and DNA containment. *Adv Protein Chem* **64**: 301-323.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**: 3497-3500.
- Cross, A.S. (1990) The biologic significance of bacterial encapsulation. *Curr Top Microbiol Immunol* **150**: 87-95.
- Deszo, E.L., Steenbergen, S.M., Freedberg, D.I., and Vimr, E.R. (2005) *Escherichia coli* K1 polysialic acid O-acetyltransferase gene, neuO, and the mechanism of capsule form variation involving a mobile contingency locus. *Proc Natl Acad Sci U S A* **102**: 5564-5569.
- Dobbins, A.T., George, M., Basham, D.A., Ford, M.E., Houtz, J.M., Pedulla, M.L. *et al.* (2004) Complete genomic sequence of the virulent *Salmonella* bacteriophage SP6. *J Bacteriol* **186**: 1933-1944.
- Dunn, J.J., and Studier, F.W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol* **166**: 477-535.
- Finne, J., and Mäkelä, P.H. (1985) Cleavage of the polysialosyl units of brain glycoproteins by a bacteriophage endosialidase. Involvement of a long oligosaccharide segment in molecular interactions of polysialic acid. *J Biol Chem* **260**: 1265-1270.
- Frosch, M., Görgen, I., Boulnois, G.J., Timmis, K.N., and Bitter-Suermann, D. (1985) NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of *Escherichia coli* K1 and group B meningococci. *Proc Natl Acad Sci U S A* **82**: 1194-1198.
- Garcia, E., Elliott, J.M., Ramanculov, E., Chain, P.S., Chu, M.C., and Molineux, I.J. (2003) The genome sequence of *Yersinia pestis* bacteriophage phiA1122 reveals an intimate history with the coliphage T3 and T7 genomes. *J Bacteriol* **185**: 5248-5262.
- Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S. *et al.* (1995) Molecular cloning and functional expression of bacteriophage PK1E- encoded endoneuraminidase Endo NE. *Mol Microbiol* **16**: 441-450.
- Gross, R.J., Cheasty, T., and Rowe, B. (1977) Isolation of bacteriophages specific for the K1 polysaccharide antigen of *Escherichia coli*. *J Clin Microbiol* **6**: 548-550.
- Haggard-Ljungquist, E., Halling, C., and Calendar, R. (1992) DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J Bacteriol* **174**: 1462-1477.
- Hallenbeck, P.C., Vimr, E.R., Yu, F., Bassler, B., and Troy, F.A. (1987) Purification and properties of a bacteriophage-induced endo-N-acetylneuraminidase specific for poly-alpha-2,8-sialosyl carbohydrate units. *J Biol Chem* **262**: 3553-3561.
- Higa, H.H., and Varki, A. (1988) Acetyl-coenzyme A:polysialic acid O-acetyltransferase from K1-positive *Escherichia coli*. The enzyme responsible for the O-acetyl plus phenotype and for O-acetyl form variation. *J Biol Chem* **263**: 8872-8878.
- Jenkins, J., and Pickersgill, R. (2001) The architecture of parallel beta-helices and related folds. *Prog Biophys Mol Biol* **77**: 111-175.
- Kanamaru, S., Leiman, P.G., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., Arisaka, F., and Rossmann, M.G. (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**: 553-557.
- Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**: 123-140.
- Kim, K.S. (2003) Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nat Rev Neurosci* **4**: 376-385.
- Kwiatkowski, B., Boschek, B., Thiele, H., and Stirm, S. (1982) Endo-N-acetylneuraminidase associated with bacteriophage particles. *J Virol* **43**: 697-704.
- Kwiatkowski, B., Boschek, B., Thiele, H., and Stirm, S. (1983) Substrate specificity of two bacteriophage-associated endo-N-acetylneuraminidases. *J Virol* **45**: 367-374.
- Lawrence, J.G., Hatfull, G.F., and Hendrix, R.W. (2002) Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. *J Bacteriol* **184**: 4891-4905.
- Leying, H., Suerbaum, S., Kroll, H.P., Stahl, D., and Opferkuch, W. (1990) The capsular polysaccharide is a major determinant of serum resistance in K-1-positive blood culture isolates of *Escherichia coli*. *Infect Immun* **58**: 222-227.
- Long, G.S., Bryant, J.M., Taylor, P.W., and Luzio, J.P. (1995) Complete nucleotide sequence of the gene encoding bacteriophage E endosialidase: implications for K1E endosialidase structure and function. *Biochem J* **309** (Pt 2): 543-550.
- Lukashin, A.V., and Borodovsky, M. (1998) GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* **26**: 1107-1115.
- Merril, C.R., Scholl, D., and Adhya, S.L. (2003) The prospect for bacteriophage therapy in Western medicine. *Nat Rev Drug Discov* **2**: 489-497.
- Miyake, K., Muraki, T., Hattori, K., Machida, Y., Watanabe, M., Kawase, M. *et al.* (1997) Screening of bacteriophages producing endo-N-acetylneuraminidase. *J Ferm Bioeng* **84**: 90-93.
- Morita, M., Tasaka, M., and Fujisawa, H. (1995) Structural and functional domains of the large subunit of the bacteriophage T3 DNA packaging enzyme: importance of the C-terminal region in prohead binding. *J Mol Biol* **245**: 635-644.
- Moxon, E.R., and Kroll, J.S. (1990) The role of bacterial polysaccharide capsules as virulence factors. *Curr Top Microbiol Immunol* **150**: 65-85.
- Mühlenhoff, M., Eckhardt, M., and Gerardy-Schahn, R. (1998) Polysialic acid: three-dimensional structure, biosynthesis and function. *Curr Opin Struct Biol* **8**: 558-564.
- Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M., and Gerardy-Schahn, R. (2003) Proteolytic processing

- and oligomerization of bacteriophage-derived endosialidases. *J Biol Chem* **278**: 12634-12644.
- Nakano, Y., Yoshida, Y., Yamashita, Y., and Koga, T. (1995) Construction of a series of pACYC-derived plasmid vectors. *Gene* **162**: 157-158.
- Pajunen, M.I., Elizondo, M.R., Skurnik, M., Kieleczawa, J., and Molineux, I.J. (2002) Complete nucleotide sequence and likely recombinatorial origin of bacteriophage T3. *J Mol Biol* **319**: 1115-1132.
- Pajunen, M.I., Kiljunen, S.J., Soderholm, M.E., and Skurnik, M. (2001) Complete genomic sequence of the lytic bacteriophage phiYeO3-12 of *Yersinia enterocolitica* serotype O:3. *J Bacteriol* **183**: 1928-1937.
- Pearson, W.R., Wood, T., Zhang, Z., and Miller, W. (1997) Comparison of DNA sequences with protein sequences. *Genomics* **46**: 24-36.
- Pelkonen, S., Pelkonen, J., and Finne, J. (1989) Common cleavage pattern of polysialic acid by bacteriophage endosialidases of different properties and origins. *J Virol* **63**: 4409-4416.
- Petter, J.G., and Vimr, E.R. (1993) Complete nucleotide sequence of the bacteriophage K1F tail gene encoding endo-N-acylneuraminidase (endo-N) and comparison to an endo-N homolog in bacteriophage PK1E. *J Bacteriol* **175**: 4354-4363.
- Robbins, J.B., McCracken, G.H., Jr., Gotschlich, E.C., Orskov, F., Orskov, I., and Hanson, L.A. (1974) *Escherichia coli* K1 capsular polysaccharide associated with neonatal meningitis. *N Engl J Med* **290**: 1216-1220.
- Saez-Llorens, X., and McCracken, G.H., Jr. (2003) Bacterial meningitis in children. *Lancet* **361**: 2139-2148.
- Sandmeier, H., Iida, S., and Arber, W. (1992) DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. *J Bacteriol* **174**: 3936-3944.
- Sarff, L.D., McCracken, G.H., Schiffer, M.S., Glode, M.P., Robbins, J.B., Orskov, I., and Orskov, F. (1975) Epidemiology of *Escherichia coli* K1 in healthy and diseased newborns. *Lancet* **1**: 1099-1104.
- Scholl, D., Adhya, S., and Merrill, C. (2005) *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. *Appl Environ Microbiol* **71**: 4872-4874.
- Scholl, D., Kieleczawa, J., Kemp, P., Rush, J., Richardson, C.C., Merrill, C. *et al.* (2004) Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. *J Mol Biol* **335**: 1151-1171.
- Scholl, D., and Merrill, C. (2005) The genome of bacteriophage K1F, a T7-like phage that has acquired the ability to replicate on K1 strains of *Escherichia coli*. *J Bacteriol* **187**: 8499-8503.
- Scholl, D., Rogers, S., Adhya, S., and Merrill, C.R. (2001) Bacteriophage K1-5 encodes two different tail fiber proteins, allowing it to infect and replicate on both K1 and K5 strains of *Escherichia coli*. *J Virol* **75**: 2509-2515.
- Slauch, J.M., Lee, A.A., Mahan, M.J., and Mekalanos, J.J. (1996) Molecular characterization of the *oafA* locus responsible for acetylation of *Salmonella typhimurium* O-antigen: *oafA* is a member of a family of integral membrane trans-acylases. *J Bacteriol* **178**: 5904-5909.
- Smith, H.W., and Huggins, M.B. (1982) Successful treatment of experimental *Escherichia coli* infections in mice using phage: its general superiority over antibiotics. *J Gen Microbiol* **128**: 307-318.
- Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R., and Huber, R. (1997) Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorham-
- nosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. *J Mol Biol* **267**: 865-880.
- Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science* **265**: 383-386.
- Steven, A.C., Trus, B.L., Maizel, J.V., Unser, M., Parry, D.A., Wall, J.S. *et al.* (1988) Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. *J Mol Biol* **200**: 351-365.
- Stirm, S., and Freund-Mölbert, E. (1971) *Escherichia coli* capsular bacteriophages. II. Morphology. *J Virol* **8**: 330-342.
- Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., and Ficner, R. (2005) Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol* **12**: 90-96.
- Tang, L., Marion, W.R., Cingolani, G., Prevelige, P.E., and Johnson, J.E. (2005) Three-dimensional structure of the bacteriophage P22 tail machine. *EMBO J* **24**: 2087-2095.
- Taylor, C.M., and Roberts, I.S. (2005) Capsular polysaccharides and their role in virulence. *Contrib Microbiol* **12**: 55-66.
- Taylor, G. (1996) Sialidases: Structures, biological significance and therapeutic potential. *Current Opinion in Structural Biology* **6**: 830-837.
- Tetart, F., Desplats, C., and Krisch, H.M. (1998) Genome plasticity in the distal tail fiber locus of the T-even bacteriophage: recombination between conserved motifs swaps adhesin specificity. *J Mol Biol* **282**: 543-556.
- Thiel, K. (2004) Old dogma, new tricks--21st Century phage therapy. *Nat Biotechnol* **22**: 31-36.
- Tomlinson, S., and Taylor, P.W. (1985) Neuraminidase associated with coliphage E that specifically depolymerizes the *Escherichia coli* K1 capsular polysaccharide. *J Virol* **55**: 374-378.
- van Raaij, M.J., Schoehn, G., Burda, M.R., and Miller, S. (2001) Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J Mol Biol* **314**: 1137-1146.
- Vander, B.C., and Kropinski, A.M. (2000) Sequence of the genome of *Salmonella* bacteriophage P22. *J Bacteriol* **182**: 6472-6481.
- Verma, N.K., Brandt, J.M., Verma, D.J., and Lindberg, A.A. (1991) Molecular characterization of the O-acetyl transferase gene of converting bacteriophage SF6 that adds group antigen 6 to *Shigella flexneri*. *Mol Microbiol* **5**: 71-75.
- Vimr, E.R., McCoy, R.D., Vollger, H.F., Wilkison, N.C., and Troy, F.A. (1984) Use of prokaryotic-derived probes to identify poly(sialic acid) in neonatal neuronal membranes. *Proc Natl Acad Sci U S A* **81**: 1971-1975.
- Weigele, P.R., Scanlon, E., and King, J. (2003) Homotrimeric, beta-stranded viral adhesins and tail proteins. *J Bacteriol* **185**: 4022-4030.
- Weintraub, A., Johnson, B.N., Stocker, B.A., and Lindberg, A.A. (1992) Structural and immunochemical studies of the lipopolysaccharides of *Salmonella* strains with both antigen O4 and antigen O9. *J Bacteriol* **174**: 1916-1922.
- Weisgerber, C., Husmann, M., Frosch, M., Rheinheimer, C., Peuckert, W., Görgen, I., and Bitter-Suermann, D. (1990) Embryonic neural cell adhesion molecule in cerebrospinal fluid of younger children: age-dependent decrease during the first year. *J Neurochem* **55**: 2063-2071.

4.1 – Supplemental Material

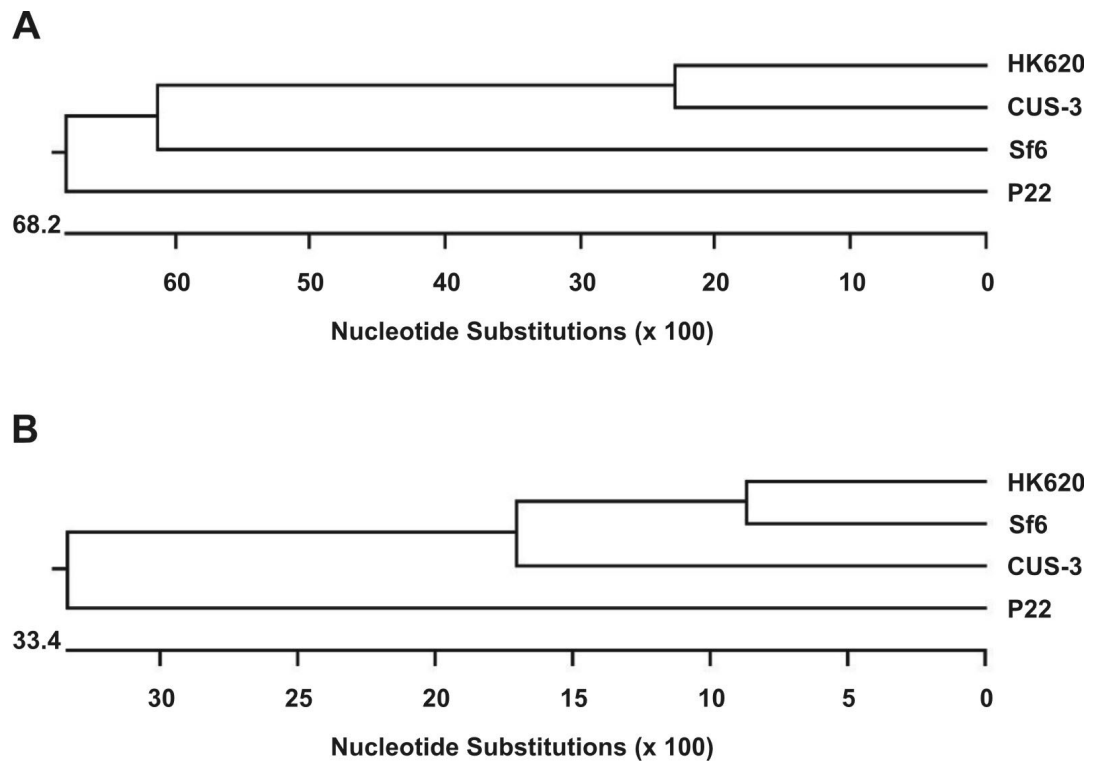


Figure S1: Phylogenetic comparison of the P22-like phages CUS-3, HK620, Sf6 and P22. (A) Comparison of phage genomes. (B) Comparison of 17 genes present in all four genomes. The selected genes encode for DNA polymerase, DNA helicase, holin, lysozyme, endopeptidase, large subunit of the DNA terminase, capsid proteins, and integrase (i. e. P22 genes 18, 12, 13, 19, 15, 2, 1, 8, 5, 4, 10, 26, 14, 7, 20, 16, and *int*, respectively, and the cognate genes of CUS-3, HK620 and Sf6). Phylogenetic trees were obtained using the ClustalW method of the MegAlign programme version 6.1 (DNASTAR Inc.).

Table S2. Bacteriophage K1F – predicted genes and gene products

ΦK1F ¹⁾ ORF	Position	Translation initiation region	Stop	Length (aa)	homologue in T7 ²⁾ ORF / identity (%)	comments/ putative function
0.1	352-567	GACTTAAAGGTCAGAGACATG	TAG	71	-	
0.2	785-1030	ACTTAAAAGGATGGACACTATG	TGA	81	-	
0.3	1114-1623	CTTTGAGAGGATATACCTATG	TGA	169	gp0.3 / 39	ocr protein
0.35	1623-1772	TGAGGAGGACTGCGAGTGATG	TGA	49	-	
0.36	1840-2007	ATTGATAGGAGTTAACTAATG	TAA	55	-	
0.37	2011-2208	GGACTGGCGGATAATCTATG	TAA	65	-	
0.38	2208-2468	TATTGGGAGCAAACCTGTAATG	TAA	86	-	
1	2566-5247	ATAACAAGGACTTTAAGTATG	TAA	893	gp1 / 62	RNA polymerase
1.1	5262-5477	TTAACAGGAGAGAGACTATG	TGA	71	-	
1.15	5648-5827	TACTGGAGATTTAACCGTATG	TAA	59	-	
1.2	5832-6095	TCGTAGGAGCATAACGCTATG	TAA	87	gp1.2 / 52	dGTP triphosphohydrolase inhibitor
1.3	6095-7117	CACATGGAGCAAATCATAATG	TAA	340	gp1.3 / 51	DNA ligase
1.6	7229-7486	CAACCTAAGGAGACTCTATG	TGA	85	gp1.6 / 38	
1.7	7483-7950	CGTAAAGGAGGTGTTCTTATG	TGA	155	gp1.7 / 80	
1.75	7940-8167	AACATAAGGATAAGTGCTATG	TAA	75	-	
1.8	8248-8346	GGGAGATAGAGGTGGAATATG	TGA	32	gp1.8 / 59	
2	8343-8501	ATGGAAGGAGCGCGTTTATG	TAA	52	gp2 / 47	host RNA polymerase inhibitor
2.5	8559-9257	AATAGGAGATTTACACCGATG	TAA	232	gp2.5 / 46	single-stranded DNA-binding protein
2.55	9100-9315	GACTGAAGGAGTGGAAAGGTG	TGA	71	-	
3	9257-9709	GGTGGCGATTACGACTTCTAATG	TGA	150	gp3.0 / 61	endonuclease
3B	9470-9709	ACCAAAAGGAATCTTCGCAGTG	TGA	79	gp3.0 / 64	
3.2	9706-10179	AAGAAAGGAGTTAAGAAATG	TAA	157	-	
3.5	10169-10627	GAATTGGTGGTGTACAATG	TAA	152	gp3.5 / 68	lysozyme/ RNA polymerase
3.5B	10274-10627	GGGGTAGAGACCATCCGATG	TAA	117	gp3.5 / 71	
3.7	10652-10867	CTAAGGGAGAACAACCAATG	TAA	71	-	
4A	10936-12636	GTGATTAAGGAGTGACCAATG	TAA	566	gp4A / 63	primase/helicase
4B	11086-12636	CATGGAAGAAAGGAGACATG	TAA	516	gp4B / 65	
4.2	12365-12688	AGAGAAAGGTAATTCGCATG	TAG	107	gp4.2 / 43	
4.3	12688-12852	ATACGGAGGACACCATTATG	TAG	54	-	
4.4	12743-12892	TGTACAAATGGTGGAGCGATG	TGA	49	-	
5	12917-14935 14936-15530 15531-15686	TAAAGCGGAGGATTAACCATG	TAA	723	gp5 / 55	DNA polymerase (aa 1-672) putative group I intron DNA polymerase (aa 673-723)
5.3	15015-15410	AAGCCATAGGGAGGATGTATG	TAG	131	-	homing endonuclease, homologues in phages Φ ³ and W31 ⁴⁾
5.4	15513-15686	AGTTAGCCGAAGTATTCAATG	TAA	57	-	
5.5	15686-15976	GGAAGGAGTGTCACTAATG	TGA	96	gp5.5 / 28	permits growth on Lambda-lysogens
5.7	15973-16182	GGGTGGAGTTCGAACGATG	TGA	69	gp5.7 / 76	
6	16179-17045	CACGGAGGCTGTCTATG	TAA	288	gp6 / 51	exonuclease
6.5	17256-17528	TAAAGAGGAGATTCAAGATG	TAA	90	gp6.5 / 36	
6.6	17308-17505	TTACAGGAAGAAAAGACTTG	TGA	65	-	
6.7	17543-17767	AACTGTAGGAGGACTATG	TAA	74	gp6.7 / 45	
6.8	17619-18176	CCTCTGAGTGAGGAAGTTG	TAA	185	-	
7.3	18169-18417	ACTAAGGAGATTAATCGTG	TAA	82	gp7.3 / 40	host range
8	18429-19997	TTTAAAGGAGGTGACTAATG	TAA	522	gp8 / 68	head-tail connector protein
9	20106-20990	AAAGGAGAATGACTTAATG	TAG	294	gp9 / 40	capsid assembly protein
10A	21131-22174	ATAGGAGAATTAGCATATG	TAA	347	gp10a / 70	major capsid protein
11	22368-22934	AAGTGAGAGGAGACTTATG	TAA	188	gp11 / 50	tail tubular protein
12	22934-25291	CATTGGGAGGGACTAATG	TAA	785	gp12 / 50	tail tubular protein B
13	25368-25838	GTTATAAGGAGACTTTATG	TAA	156	gp13 / 52	internal virion protein A
14	25823-26410	TCTCTAAGGAGGACGTATG	TAA	195	gp14 / 38	internal virion protein B
15	26422-28704	GCTAAGAGGAGGACTAATG	TAA	760	gp15 / 46	internal virion protein C
16	28709-32596	GGTAAGGAGTAACATAATG	TAA	1295	gp16 / 53	internal virion protein D
17	32661-35855	TCACAAAAGGAGGCATAATG	TAA	1064	-	endosialidase
17.5	35989-36183	TTAAACTGGAGGTTTTATG	TGA	64	gp17.5 / 72	lysis protein
18	36180-36443	AGGAGGTGGTAAAACATG	TAA	87	gp18 / 44	DNA maturase A
18.5	36546-36995	AACCGGGAGGGCAACTATG	TGA	149	gp18.5 / 50	endopeptidase
18.7	36646-36921	TGGGAGGCTAAGTCAATG	TGA	91	gp18.7 / 53	
19	36992-38752	AGAAGGAGGTGACGCATG	TAG	586	gp19 / 65	DNA maturation protein
19.2	37632-37799	CGTGAGCTGGAAGGTCGTG	TGA	55	gp19.2 / 36	
19.4	39048-39209	AGGAGGTATGTAGCATATG	TAA	53	-	

1) accession number [AM084414], complete genome of bacteriophage K1F (this study)

2) accession number [NC_001604], complete genome of bacteriophage T7 (Dunn and Studier, 1983)

3) accession number [AAV53690], intron-encoded homing endonuclease, bacteriophage ΦI (Bonocora and Shub, 2004)

4) accession number [AAV53693], intron-encoded homing endonuclease, bacteriophage W31 (Bonocora and Shub, 2004)

Table S3. Bacteriophage K1E – predicted genes and gene products

ΦK1E ¹⁾ ORF	Position	Translation initiation region	Stop	Length (aa)	homologues in K1-5 & SP6 ORF / identity (%)		comments/ putative function
					K1-5 ²⁾	SP6 ³⁾	
1	1057-1242	TTTGTGGAGGCTTTATCATG	TGA	61	gp1 / 87	gp1 / 69	
1.5	1226-1381	TTGATGAAGGAGTTAAGTAATG	TGA	51	-	-	
2	1368-1592	TAAAGGAGGTAAAGAGGATG	TAA	74	gp2 / 98	-	
2.5	1719-1931	CTACTTACTAAGGTGAACTATG	TAG	70	-	-	
2.8	2031-2195	CTAACAAAGGTGAACTATCATG	TAA	54	-	-	
3	2346-2672	CTAATGAGGATTAATCATG	TAA	108	gp3 / 94	gp3 / 54	ocr protein
4	2675-2788	AGACGAGGAAGAATAAGATG	TGA	37	gp4 / 97	gp4 / 75	
5	2865-3092	CTTACTAGGAGTTAAGACTATG	TAA	75	-	-	
6	3156-4040	ACTATATAAGGTGACTATTATG	TAA	294	gp6 / 94	gp6 / 65	
7	4104-6731	TTACTAGGAGTAAACAAGATG	TAA	875	gp7 / 98	gp7 / 84	RNA polymerase
8	7056-7184	TCACTTTGGAGATTAAACCATG	TAA	42	gp8 / 97	gp8 / 74	
8.5	7188-7406	GTGGGAGTTCTAAGTTATG	TAA	72	-	-	
9	7406-9394	GAACTATTAAGGTGAACTAATG	TAA	662	gp9 / 97	gp9 / 91	DNA-helicase
9.5	9394-9555	ACGATGAACAGGAGTTTTAATG	TAG	53	-	-	
10	9628-10311	AAATACTGGAGATTGATTATG	TAA	227	gp10 / 89	gp10 / 38	
10.5	10322-10540	AATATCTAAGGATTAATCATG	TAA	72	gp10.5 / 83	-	
11	10530-10739	GGCAGCTTTATGGTGAACATG	TAA	69	gp11 / 94	gp11 / 72	
11.5	10797-10904	TATTGATTGGAGGTGCATTATG	TAA	35	-	-	
12	10891-11109	ACGAAACAGAGGCAATGTATG	TGA	72	-	-	
13	11096-13642	ATACTTTGGAGCAATCTATG	TGA	848	gp13 / 98	gp13 / 89	DNA-polymerase
13.1	11985-12362	TTTAAACCGGAAGAAAAGAATG	TAA	125	-	-	
13.2	12363-12569	CCTTGGAGTGGGAAGATAAATG	TAG	68	-	-	
13.5	13642-13758	GTTGGAAGGAAACACTGATG	TGA	38	gp13.5 / 100	-	
14	13758-13856	AAGAGAGGGTTACGATTGATG	TGA	32	gp14 / 100	gp14 / 87	
15	14012-14389	TATAGGAGATATAAATATG	TAA	125	gp15 / 99	gp15 / 90	
15.5	14396-14695	CTTGACTAAATAAGGAGTTATG	TAA	99	gp15.5 / 93	-	
16	14865-15671	ACAATAGGAGACTAAAATAATG	TAA	268	gp16 / 97	gp16 / 81	
17	15681-15899	AATATTAAGGAGAACGGATG	TAA	72	gp17 / 98	gp17 / 68	
18	16005-16376	AACAAGGAGGAATTTAATATG	TGA	123	gp18 / 97	gp18 / 94	
19	16494-16766	AGTTATTGGAGAATGAATTATG	TAA	90	gp19 / 96	gp19 / 46	
20	16679-17704	TCTTGAAGGAGTGGCCTACATG	TAA	341	gp20 / 97	gp20 / 82	
21	17689-18102	TCTTGTGGGAGATGAAGGATG	TAA	137	gp21 / 62	gp21 / 67	
22	18095-19102	GAAGAAGGAGGTTAAAGTTG	TTG	335	gp22 / 98	gp22 / 94	
23	19170-19709	GTGAATGGAGGAATTAATATG	TAA	179	gp23 / 58	gp23 / 70	
24	19881-20828	TGTGATGAAGGAGGTGAAATG	TGA	315	gp24 / 88	gp24 / 83	DNA-ligase
25	20946-21143	ATTAAGGAGGTTAATTATG	TGA	65	gp25 / 95	gp25 / 83	
26	21106-21276	GAGAAGTTGAGGAACACTATG	TGA	56	gp26 / 100	gp26 / 59	
27	21273-21734	GCACAACGGTGGTCCAATG	TAA	153	gp27 / 83	gp27 / 75	hypothetical acetyltransferase
28	21744-21953	GAAGTTCTAAGGAGATCACATG	TAA	69	gp28 / 95	gp28 / 80	
29	21955-23505	TCTAGCTTGAAGGTGTAATATG	TAA	516	gp29 / 98	gp29 / 86	head-tail connector protein
30	23505-24425	GAACTTAAGGAGCGTAATG	TAA	306	gp30 / 88	gp30 / 64	scaffolding protein
30.5	23587-23796	ACCAAGGAGGTAACAACGATG	TAA	69	-	-	
31	24502-25704	TATAAGGAGGTTTATTAATG	TAA	400	gp31 / 84	gp31 / 93	major capsid protein
32	25759-26496	TTGTTAGGAGGATTCATG	TAA	245	gp32 / 98	gp32 / 87	tail tubular protein A
33	26496-28898	CCGTGGGAGGATCGTTAATG	TAA	800	gp33 / 96	gp33 / 82	tail tubular protein B
34	28898-29620	CCAAAGGAGGCTTAATG	TAA	240	gp34 / 98	gp34 / 70	internal virion protein
35	29620-32568	AACAAGAGGTATTTAATG	TAG	982	gp35 / 98	gp35 / 41	
36	32633-35941	CTATGGAGGAACTATG	TGA	1102	gp36 / 93	gp36 / 37	internal virion protein
37	35941-36903	TGAACGGAGGATTGATG	TAA	320	gp37 / 74	gp37 / 65	tail protein / adapter protein
38	36913-37107	GTGTAAGGAGGTGAGATG	TAA	64	gp38 / 89	gp38 / 78	holin
39	37091-37390	CTATGGAGGAAGAACCGTG	TGA	99	gp39 / 98	gp39 / 78	small terminase subunit
40A	37390-38607	TAAGGAGGATTGATG	TGA	405	gp40 / 99	gp40 / 81	large terminase subunit
40.5	38777-39310	CAATTGAAGGGAGTTAAATATG	TAA	177	-	-	homing endonuclease (Phage Aeh1, AY266303)
40B	39252-40034	TATGCGTAAGGTGGCACGTG	GTG	260	gp40 / 99	gp40 / 68	large terminase subunit
41	40199-40477	GGAGAGAGAGAGAGAGAATG	TAA	92	gp41 / 72	gp41 / 35	
42	40490-40777	TTAAGAGGAGAGTTACATG	TAA	95	gp42 / 77	gp42 / 52	
43	40798-41142	AAAGACGAGGTCAAATAGTG	TAA	114	gp43 / 84	gp43 / 74	
44	41251-41400	AATGGAAAAGGAGATGCAATG	TAA	49	gp44 / 93	gp44 / 85	
44.5	41411-41614	GAATAAGGTGCTAGAGGTG	TAA	67	-	gp48 / 68 ⁴⁾	
45	41478-41687	AGAGAAGGAGGCACAAAAGTG	TAA	69	gp45 / 85	gp45 / 76	
46	41819-42154	GGAAGTAATAGGAGAGTATATG	TAA	111	gp46 / *	-	* N-terminal 10 aa identical
47	42222-44657	TCAGAAAAGGAGGTTACATG	TAA	811	gp47 / 97	-	endosomalidase
48	44868-45083	AACATGCAAGAGGAGAATTG	TGA	71	-	-	

- 1) accession number [AM084415], complete genome of bacteriophage K1E (this study)
- 2) accession number [AY370674], complete genome of bacteriophage K1-5 (Scholl *et al.*, 2004)
- 3) accession number [AY370673], complete genome of bacteriophage SP6 (Scholl *et al.*, 2004)
- 4) accession number [NC_004831], complete genome of bacteriophage SP6 (Dobbins *et al.*, 2004)

Chapter 5 – Characterization of a Novel Intramolecular Chaperone Domain Conserved in Endosialidases and Other Bacteriophage Tailspike and Fibre Proteins

This manuscript has originally been published in *The Journal of Biological Chemistry*.

'Characterization of a Novel Intramolecular Chaperone Domain Conserved in Endosialidases and Other Bacteriophage Tail Spike and Fiber Proteins'

David Schwarzer, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff¹

Abteilung Zelluläre Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Carl Neuberg-Str. 1, 30625 Hannover, Germany

¹ To whom correspondence should be addressed: Tel.: 49-511-532-9807; Fax: 49-511-532-3956;
E-mail: muehlenhoff.martina@mh-hannover.de

THE JOURNAL OF BIOLOGICAL CHEMISTRY; VOL. **282**, NO. 5, pp. 2821–2831, February 2, 2007
© 2007 by 'The American Society for Biochemistry and Molecular Biology', Inc.

Received for publication, October 10, 2006, and in revised form, November 16, 2006 Published, JBC Papers in Press, December 7, 2006,

DOI 10.1074/jbc.M609543200

Preface – About the Manuscript

Endosialidases have been described to contain a highly conserved C-terminal part that is released from the mature protein complex after the formation of active oligomeric complexes. Interestingly, the C-terminal part is also found in bacteriophage tailspike and tail fibre proteins that are otherwise unrelated to endosialidases. This part of my study was aimed at (i) revealing the definite functions of the C-terminal part of these proteins and (ii) defining the molecular requirements for their proteolytic release.

First, I performed a database search using the primary sequence of the C-terminal part of endoNF to identify novel tailspike proteins harbouring the conserved C-terminal part. In total, 13 tailspike proteins were identified that could be grouped into four different protein families due to the phylogenetic analysis. One representative of each family was selected to perform comparative analyses addressing the questions of: (i) the role of highly conserved amino acids in this domain (ii) the potential function of this domain as an intramolecular chaperone and (iii) the mechanism of proteolytic maturation. Using truncated and point mutated proteins as well as assay systems to determine the thermostability of wild type and mutant proteins, the C-terminal domain could be confirmed as a functional chaperon, which is interchangeable between otherwise unrelated N-terminal domains. Moreover, the study demonstrates that the release of the C-terminal part from the folded proteins is essential to kinetically stabilise the functional N-terminal parts.

SUMMARY

Folding and assembly of endosialidases, the trimeric tail spike proteins of *Escherichia coli* K1 specific bacteriophages, crucially depend on their C-terminal domain (CTD). Homologous CTDs were identified in phage proteins belonging to three different protein families: neck appendage proteins of several *Bacillus* phages, L-shaped tail fibers of coliphage T5, and K5-lyases, the tail spike proteins of phages infecting *Escherichia coli* K5. By analyzing a representative of each family, we show that in all cases the CTD is cleaved off after a strictly conserved serine residue and alanine substitution prevented cleavage. Further structural and functional analyses revealed that: (i) CTDs are autonomous domains with a high α -helical content; (ii) proteolytically released CTDs assemble into hexamers which are most likely dimers of trimers; (iii) highly conserved amino acids within the CTD are indispensable for CTD-mediated folding and complex formation; (iv) CTDs can be exchanged between proteins of different families; and (v) proteolytic cleavage is essential to stabilize the native protein complex. Data obtained for full-length and proteolytically processed endosialidase variants suggest that release of the CTD increases the unfolding barrier, trapping the mature trimer in a kinetically stable conformation. In summary, we characterize the CTD as a novel C-terminal chaperone domain, which assists folding and assembly of unrelated phage proteins.

The largest group of bacteriophages, the order of tailed phages (*Caudovirales*), has evolved tail spike and fiber proteins for efficient virus-host-interactions (1). These specialized adhesins mediate recognition and attachment to the bacterial surface and are the key determinants for host specificity. Since a variety of pathogenic bacteria are protected by thick layers of lipo- or capsular polysaccharides, many tail spike proteins possess enzymatic activity. In particular phages infecting encapsulated bacteria crucially depend on polysaccharide depolymerases (endoglycosidases or lyases) to gain access to the bacterial membrane (2, 3).

In mature phage particles, tail spikes and fibers are exposed structures, which require high stability to maintain their functional conformation even under extreme conditions like high salt concentrations, the presence of extra-cellular proteases, and variations in pH and temperature. Interestingly, many spikes and fibers are build-up by homotrimers that contain stretches of intertwined subunits like coiled-coil, triple β -helix or triple β -spiral folds, leading to protein complexes which remain stable even in the presence of sodium dodecyl sulphate (SDS) (4).

One example are endosialidases, the tail spike proteins of *Escherichia coli* (*E. coli*) K1 specific phages. *E. coli* K1, a leading cause of sepsis and meningitis in newborns (5, 6), is encapsulated by α 2,8-linked polysialic acid (7). The K1 capsule protects the bacterium from the immune system (8) but simultaneously provides the attachment site for specialized phages that are equipped with endosialidase tail spikes (9–14). In line with the concept of modular evolution of bacteriophages (15, 16), K1-specific phages evolved by insertion of an endosialidase gene into the tail spike or fiber locus of diverse progenitor phages (17). Thereby, the N-terminal capsid-binding domain of the original spike or fiber protein is kept to ensure proper integration of the newly acquired tail spike into existing tail architectures.

Recently, we solved the crystal structure of the conserved catalytic part of endosialidase endoNF² (aa 246–911), the tail spike of phage K1F (18). The protein assembles into a catalytic homotrimer with an overall mushroom like outline. In the cap region, each subunit is folded into a 6-bladed β -propeller typical for sialidases and a lectin like β -barrel domain which is involved in binding of polysialic acid. The distal part of the protein forms a stalk region where all three subunits intertwine into a β -helix fold which is interrupted by a β -prism domain. The catalytic part is flanked by the N-terminal capsid binding domain (aa 1–245) and a short C-terminal domain (CTD; aa 912–1064). During maturation, the CTD is released by proteolytic cleavage after a highly conserved serine residue (19). Exchange of this residue by alanine (S911A in endoNF) prevents cleavage but not assembly into active trimers, demonstrating that proteolytic processing is not prerequisite for complex formation and activity. Although the CTD is not part of the native trimer, folding and trimerization of endosialidases crucially depend on the presence of an intact CTD in the nascent polypeptide. Truncation or even the exchange of a single histidine residue within the domain prevented trimerization and led to accumulation of inactive and predominantly insoluble protein (19, 20). Remarkably, the CTD is not restricted to endosialidases. Homologous domains were found as part of other tail spike and fiber proteins, suggesting a general role in folding and assembling of phage proteins (19).

In the present study we focused on the structural and functional characterization of the CTD as part of an endosialidase but also of other spike and fiber proteins derived from distinct phages. The CTD acts as an intramolecular chaperone, probably by lowering a high-energy barrier on the folding pathway. By analyzing a novel C-terminal chaperone domain, essential for folding and assembly of unrelated proteins, this work contributes not only to our understanding of bacteriophage evolution but might form the basis for engineering highly stable protein complexes.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of DFG Research Unit 548 (Ge801/7-1). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

² The abbreviations used are: aa, amino acids; CTD, C-terminal domain; endoNF, endo-N-acyl-neuraminidase or endosialidase of bacteriophage K1F (E.C. 3.2.1.129); gp12, gene product 12; LTF, L-shaped tail fiber protein.

EXPERIMENTAL PROCEDURES

Materials – Colominic acid was purchased from Sigma. The vector pBlueScript SK⁻ was obtained from Stratagene and pET expression vectors from Novagen.

Bacteria and Bacteriophages – *Bacillus* phage GA-1 (DSM 5548), *Bacillus pumilus* strain G1R (DSM 5549), and *E. coli* DSM 613 were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and coliphage T5 from the American Tissue Culture Collection (ATCC No. 11303-B5). The *E. coli* K5 strain used in this study was isolated from urine at the department of Medical Microbiology of the Medizinische Hochschule Hannover (Hannover, Germany). *E. coli* BL21(DE3) Gold was purchased from Stratagene.

Cloning of Tail Spike and Fiber Proteins – The coding sequences of the tail genes encoding endoNF, gp12, LTF and ElmA were amplified from purified phage DNA of coliphage K1F, *Bacillus* phage GA-1, coliphage T5, and from genomic DNA of *E. coli* K5, respectively, using the following primer pairs: endoNF: MM97 (5'-cg ggatcc atg tcc acg att aca caa ttc-3') and MM98 (5'-gtccg ctcgag ctt ctg ttc aag agc aga aag-3'); ΔN-endoNF lacking the first 245 aa: MM147 (5'-cgggatcc gct aaa ggg gat ggt gtc-3') and MM98 (5'-gtccg ctcgag ctt ctg ttc aag agc aga aag-3'); GA-1 gene 12: DS01 (5'-cg ggatcc atg cat aga ccg cca ttc-3') and DS02 (5'-ggc ctcgag act caa cct ttc aag ctt c-3'); LTF: DS03 (5'-cg ggatcc atg gct ata act aaa ata att c-3') and DS04 (5'-ggc ctcgag cat acc taa ttt atc ctc-3'); ElmA: DS07 (5'-cg ggatcc atg acc gtc tca acc gaa-3') and DS08, (5'-ggc ctcgag att ccc tgt taa ttg caa-3'). Sense and antisense primers are flanked by BamHI and XhoI sites (underlined), respectively. PCR products were ligated into BamHI/XhoI sites of the expression vector pET22b-Strep, a modified pET22b vector containing the sequence encoding an N-terminal Strep-tag II followed by a Thrombin cleavage site (WSHPQFEKGALVPRGS) and a C-terminal His₆ tag. The sequence identity of all PCR products was confirmed by sequencing. The amplified *lrf* gene showed five sequence deviations to the sequence published by Kaliman *et al.* (21) but was identical to the recently published sequence of Wang *et al.* (22).

Site-directed Mutagenesis – Mutagenesis was performed by PCR using the QuikChange site-directed mutagenesis kit (Stratagene) following the manufacturer's guidelines with the following primer pairs (mutated nucleotides are shown in bold): MM148 (5'-c cct att gtt act tct **gcc** ggg gag agg aaa aca g-3') and MM149 (5'-c tgt ttt cct ctc ccc **ggc** aga agt aac aat agg g-3') for endoNF-N912A; MM150 (5'-gct cgt att cac ttc gcg gtt att gct cag c-3') and MM151 (5'-g ctg agc aat aac **cg** gaa gtg aat acg agc-3') for endoNF-G956A; DS45 (5'-g ggt gag gag tgg ggt gtt **g**cg cct gac gga att ttc ttt gc-3') and DS46 (5'-gc aaa gaa aat tcc gtc agg **cg**e aac acc cca ctc ctc acc c-3') for endoNF-R1035A; DS25 (5'- cg ggt gct att aac aca **g**ct gac gag aga cat aaa acg g-3') and DS26 (5'-c cgt ttt atg tct ctc gtc **ag**e tgt gtt aat agc acc cg-3') for gp12-S620A; DS63 (5'-cg ggt gct att aac aca tct **g**ca gag aga cat aaa acg gac-3') and DS64 (5'- gtc cgt ttt atg tct ctc **g**ca aga tgt gtt aat agc acc cg-3') for gp12-

D621A; DS65 (5'-a ggc gaa gaa gct aga tat **g**ca ttt ggg gtt atc gct caa c-3') and DS66 (5'-a ttg agc gat aac ccc aaa **t**gc ata tct agc ttc ttc gcc t-3') for gp12-H663A; DS67 5'-gaa gaa gct aga tat cat ttt **g**cg gtt atc gct caa cag att g-3' and DS68 5'-c aat ctg ttg agc gat aac **cg**e aaa atg ata tct agc ttc ttc-3' for gp12-G665A; DS49 5'-ggc aac atc tac agc ata **g**ca cct acc gaa tgt caa tgg-3' and DS50 5'-cca ttg aca ttc ggt agg **t**gc tat gct gta gat gtt gcc-3' for gp12-R720A; DS27 5'-cc gta aat ggt aca att aac **a**ca **g**ct gat gct aga ttg aag aac gat gtt cg-3' and DS28 5'-cg aac atc gtt ctt caa tct agc atc **ag**e tgt gtt aat tgt acc att tac gg-3' for LTF-S1264A; DS41 5'-ca gcg ttt acg gtg **c**ta **g**cc gat gcg cgt ttc aag act gc-3' and DS42 5'-gc agt ctt gaa acg cgc atc **g**gc tag cac cgt aaa cgc tg-3' for elmA-S675A. Plasmids containing the corresponding wild-type sequence were used as a template. PCR products were subcloned into the BamHI and XhoI sites of pET22b-Strep resulting in a construct with an N-terminal Strep- and a C-terminal His₆ tag.

Generation of C-terminally Truncated Proteins – Constructs encoding C-terminally truncated proteins were generated by PCR with the following primer pairs: MM124, 5'-ggaattc catatg gag ata act cag-3' (NdeI) and MM126 5'-atccg ctcgag tta aga agt aac aat agg gtt-3' (XhoI) for endoNFΔC (aa 246–911), DS29 5'-ggttctac tctaga aga tgg tgc agg-3' (XbaI) and DS30 5'-atccg ctcgag taa gat gtg tta ata gca cc-3' (XhoI) for gp12ΔC (aa 621–740), DS15 5'-gcatg cccggg tgc att c-3' (SmaI) and DS31 5'-atccg ctcgag tta aga tgt gtt aat agc acc-3' (XhoI) for LTFΔC (aa 1265–1396), DS37 5'-cg gaa act gcc cca acg-3' and DS44 5'-atccg ctcgag tta gga cag cac cgt aaa cgc tg-3' (XhoI) for elmAΔC (aa 676–820). PCR products were subcloned into pET22b-Strep using the respective restriction sites given in brackets, resulting in constructs that encode for proteins with an N-terminal Strep-tag.

Generation of an Expression Plasmid Encoding the CTD of EndoNF – For separate expression of the endoNF CTD, DNA encoding aa 913–1064³ of endoNF was amplified by PCR with the primer pair MM183 5'-G GAA TTC CATATG GGG GAG AGG AAA ACA GAG C-3' (NdeI site) and MM98 5'-G TCC CTC GAG CTT CTG TTC AAG AGC AGA AAG-3' (XhoI site). The obtained PCR product was subcloned into the NdeI and XhoI restriction sites of pET22b resulting in a construct encoding the CTD of endoNF with a C-terminal His₆ tag.

Generation of Chimera – The chimera ΔN-endoNF-gp12 was generated by homologous recombination in *E. coli* YZ2000 cells (Gene Bridges) (23) using the primers DS69 5'-GC AAC CGT TTC ACC ACT GCA TAC CTC GGA AGC AAC CCT ATT GTT ACT TCT gag gag aga cat aaa acg gac ata gc-3' (upper case: nt 2684–2733 of the endoNF gene; lower case: nt 1861–1886 of gene 12) and DS70 5'-g taa tac gac tca cta tag ggc gaa ttg ggt acc ggg ccc ccc CTCGAG ACT CAA CCT TTC AAG CTT CCT TC-3', (lower case:

³ Cleavage of endoNF occurs at the highly conserved Ser-911, releasing a CTD which start with Asn-912. Asparagine residues, in particular in front of a glycine, are very labile and can undergo spontaneous degradation. Consequently, N-terminal sequencing of the released CTD revealed Gly-913 as the first amino acid (19) and the construct generated for separate expression of the endoNF CTD was designed accordingly.

nt 2198–2220 of gene 12; upper case: nt 615–673 of pBluescript-SK⁻; underlined: XhoI site). A plasmid containing the coding sequence of wild-type gp12 was used as a template and the resulting PCR product was recombined with the linearized plasmid pBluescript- Δ N-endoNF containing the coding sequence of endoNF lacking the N-terminal 245 amino acids. The product was subcloned into pET22b-Strep resulting in a construct encoding aa 246–911 of endoNF with an N-terminal Strep-tag fused to aa 621–740 of gp12 with a C-terminal His₆ tag.

Protein Expression and Purification – Proteins were expressed in *E. coli* BL21-Gold(DE3) in the presence of 100 μ g/ml Carbenicillin. Bacteria were cultivated in Luria-Bertani broth at 37°C (endoNF) or in PowerBroth (Athena Enzyme Systems) at 15°C (gp12, LTF, and ElmA). At an optical density (OD_{600nm}) of 0.6 expression of endoNF was induced by adding 1 mM isopropyl-1-thio- β -D-galactopyranoside (IPTG) and bacteria were harvested 2.5 h after induction. Expression of gp12, LTF, and ElmA was induced by adding 0.1 mM IPTG at an OD_{600nm} of 1.5 and bacteria were harvested 20 h after induction. Bacterial lysates were obtained by adding 2% (w/v) SDS and 0.1 M β -mercaptoethanol in 10 mM Tris-HCl, pH 7.4, 1 mM EDTA followed by incubation at 95°C for 5 minutes. Proteins were precipitated by adding 10 volumes of acetone and pellets were resuspended in sample buffer containing 1.5% (w/v) SDS. For the analysis of soluble and insoluble proteins, bacteria were lysed in BugBuster MasterMix (Novagen) and soluble and insoluble fractions were obtained after centrifugation (22,000 \times g, 20 min, 4°C). For protein purification bacteria were lysed by sonication in appropriate loading buffer containing protease inhibitors and soluble fractions were loaded on StrepTactin-Superflow (IBA) and HisTrap HP columns (Amersham) according to the manufacturer's guidelines. To isolate proteolytically matured protein, the minor fraction of non-cleaved protein was removed by passing samples over a HisTrap column prior to StrepTactin-chromatography. For isolation of CTDs, proteolytically processed and full-length proteins were removed by StrepTactin-chromatography before samples were loaded on a HisTrap column. Buffer exchange was performed on a HiPrep 26/10 Desalting column (Amersham).

Size Exclusion Chromatography – Size exclusion chromatography was performed on a Superdex 200 HR 10/30 column (Amersham) equilibrated with 10 mM sodium phosphate buffer, pH 7.4, 300 mM NaCl and 10 mM Tris-HCl, pH 8.0 for endoNF and gp12, respectively. The column was calibrated using the Gel Filtration Molecular Weight Markers (MW-GF-200) from Sigma.

SDS-PAGE and Immunoblotting – SDS-PAGE was performed under reducing conditions using 2.5% (v/v) β -mercaptoethanol and either 1.5% or 1% (w/v) sodium dodecyl sulphate (SDS) in the sample buffer. For Western blot analysis proteins were blotted onto nitrocellulose (Whatman). Proteins containing an N-terminal Strep-tag II were detected by StrepTactin-alkaline phosphatase-conjugate (StrepTactin-AP, IBA) according to the manufacturer's guidelines. His₆-tagged pro-

teins were detected with 1 μ g/ml Penta-His antibody (QIAGEN) followed by goat anti-mouse-IgG-AP (Dianova).

Determination of Endosialidase Activity – The enzymatic activity of endoNF variants was determined by means of the thiobarbituric acid assay as described earlier (19).

Circular Dichroism and Secondary Structure Estimation – Purified proteins were analyzed in 10 mM sodium phosphate buffer (pH 7.4) by far-UV circular dichroism (CD) on a J-715 spectropolarimeter (Jasco) in 1-mm- and 0.1-mm-path quartz cells. The 1 mm spectra (4-fold accumulation, 250 to 200 nm) were used for scaling the 0.1 mm spectra (16-fold accumulation, 250 to 182 nm) adjusting path length deviations of the 0.1 mm cell. The 0.1 mm spectra were analyzed for estimation of the secondary structure by use of the program CDpro suite (24, 25).

Protein Sequence Analysis – Multiple sequence alignment was performed using the ClustalV method of the MegAlign program version 6.1 (DNASTAR Inc.) with the following parameters: gap penalty 15.0 and gap length penalty 10.0. The phylogenetic tree was calculated from the aligned sequences using the maximum likelihood method with molecular clock (PromlK V3.6a3) in Phylip. The tree was printed with tree view (V1.6.6).

Thermostability Assay – Purified protein in 1.5% SDS sample buffer was incubated for 5 minutes at a variety of temperatures in the range of 4°C to 95°C. Samples were placed on ice until further analysis by SDS-PAGE.

RESULTS

Phage Proteins Containing an Endosialidase CTD – Using BLAST search analysis with the C-terminal sequence of endoNF (aa 908–1064) we identified a total of 13 phage proteins containing a similar domain. The identified proteins belong to diverse phages and can be grouped into four protein families (Table 1). Proteins of family I and II are capsular polysaccharide depolymerases specific for the respective host capsule, whereas members of family III and IV are neck appendage and tail fiber proteins, respectively, involved in phage adsorption. According to their different enzymatic or binding specificities, no sequence similarity is observed in the N-terminal part of proteins belonging to different families. The only observed relation between these proteins is found in their CTDs. By multiple sequence alignment six amino acid clusters with conserved amino acids were identified which are dispersed in regions of low sequence similarity (see Figure S1 in the Supplementary Material). The proteolytic cleavage site Ser ↓ Asx identified in endosialidases (19) is located in cluster I and is strictly conserved in all proteins. The CTDs are of variable length (105 to 163 aa) with overall amino acid sequence similarities ranging from 10.8% to 44.9% (see Table I). A phylogenetic tree based on the 13 CTD sequences revealed that similarities are not necessarily highest between members of one family (Fig. 1A) since CTDs of family I and III are split up into two different branches.

TABLE 1
Bacteriophage proteins containing an endosialidase CTD

Protein ^a	Origin	Function	Protein length	Cleavage site	Length of CTD	Similarity ^b
			aa		aa	%
Endosialidases						
endoNE	Coliphage K1E	Degradation of α 2,8-linked polysialic acid	811	Ser-706	105	23.9
endoNF	Coliphage K1F		1064	Ser-911	153	100.0
endoN63D	Coliphage 63D		984	Ser-852	132	44.9
endoNK1-5	Coliphage K1-5		811	Ser-706	105	22.9
K5 capsule depolymerases						
K5 lyase (KflA)	Coliphage K5	Degradation of poly-(β 1,4-GlcA- α 1,4-GlcNAc)	635	Ser-506	129	31.6
K5 lyase	Coliphage K1-5		632	Ser-505	127	31.3
K5 lyase (ElmA)	<i>E. coli</i> K5 ^c		820	Ser-675	145	26.8
Neck appendage proteins						
gp12	<i>Bacillus</i> phage GA-1	Binding of the host cell wall receptor	740	Ser-620	120	37.1
gp12	ϕ 29		854	Ser-691	163	11.5(22.3) ^d
gp12	PZA		854	Ser-691	163	10.8(21.4) ^d
gp12	B103		860	Ser-697	163	10.8(18.8) ^d
Neck appendage protein	<i>Zymomonas mobilis</i> ZM4		333	Ser-190	143	26.5(43.5) ^d
Tail fiber proteins						
L-shaped tail fiber protein (LTF)	Coliphage T5	Host cell receptor binding	1396	Ser-1264	132	21.3

^a Accession numbers of the used protein sequences are given in the legend to Fig. S1 in the supplemental material.

^b Primary sequence similarity of the respective CTD in comparison to the CTD of endoNF.

^c The *elmA* gene is found in some *E. coli* K5 strains and might be part of a not yet characterized prophage (28).

^d Primary sequence similarity of the respective CTD compared with the CTD of gp12 of *Bacillus* phage GA-1 as determined by the Jotun-Hein method of the MegAlign program.

Maturation of proteins with endosialidase CTD – To investigate whether all proteins containing the endosialidase CTD undergo proteolytic processing, representatives of each family shown in Table 1 were selected and compared to endoNF: the K5-lyase ElmA, the neck appendage protein gp12 of *Bacillus* phage GA-1, and the L-shaped tail fiber protein (LTF) of coliphage T5 (see Fig. 1B for schematic representation). Proteins were expressed in *E. coli* BL21 with an N-terminal Strep- and a C-terminal His₆ tag. Protein maturation was analyzed by Western blotting and results are shown in Fig. 2. Since the N-terminal capsid binding domain of endosialidases is not required for trimerization and activity (17, 18), an endoNF variant lacking the first 245 amino acids (Δ N-endoNF) was used in this study (see Fig. 1B for schematic representation). For each selected protein, two mutants were generated by (i) deletion of the respective CTD (Δ C) and (ii) alanine substitution of the highly conserved serine in the Ser \downarrow Asx motif.

Notably, all investigated proteins were proteolytically cleaved resulting in a large N-terminal (Fig. 2, A–D) and a short C-terminal fragment (Fig. 2E). Depending on the cleavage efficiency, variable amounts of full length proteins were observed (Fig. 2, A–D), which were detectable by the N- and C-terminal epitope-tags (see upper and lower panels in Fig. 2, A–D, respectively). Proteolytically processed proteins (marked with an asterisk) were only detectable by the N-terminal Strep-tag due to loss of the His₆-tagged CTD during maturation. In all cases, the apparent molecular mass of the mature protein was similar to the mass of the respective truncated variant lacking the CTD (Δ C), indicating that all proteins were cleaved at the predicted Ser \downarrow Asx cleavage site. Moreover, exchange of the

highly conserved serine residue to alanine prevented cleavage, resulting in the expression of full-length proteins (Fig. 2, A–D). In summary, these results demonstrate that cleavage at the Ser \downarrow Asx site is used for release of CTDs in proteins of all four families and that this cleavage is independent of the N-terminal protein context.

Functional role of the CTD in non-endosialidase proteins – To investigate the functional role of the CTD in a protein context other than endosialidases, the neck appendage protein gp12 was chosen as a model protein and compared to endoNF. The two molecules derive from completely unrelated phages, namely *Bacillus* phage GA-1 infecting the Gram-positive soil bacterium *Bacillus pumilus* and phage K1F infecting Gram-negative, pathogenic *E. coli* K1, respectively. EndoNF and gp12 were expressed in *E. coli* BL21 with an N-terminal Strep- and a C-terminal His₆ tag and the soluble fraction of the obtained bacterial lysates was analyzed as depicted in Fig. 3. Like endoNF, gp12 formed an oligomeric complex which is resistant to SDS at room temperature. By omitting the boiling step prior to SDS-PAGE, a high molecular weight band appeared for either protein, while in parallel the intensity of the band corresponding to monomeric, proteolytically processed protein (endoNF* and gp12*) decreased. SDS-resistant complexes were only detectable by the N-terminal Strep-tag, demonstrating that they are built up exclusively by mature protein lacking the CTD. In non-boiled samples, a larger proportion of gp12* than Δ N-endoNF* was found in the monomeric state (compare second lanes from the left in Fig. 3A and B), an indication that gp12 oligomers are less stable in the presence of SDS than endoNF trimers.

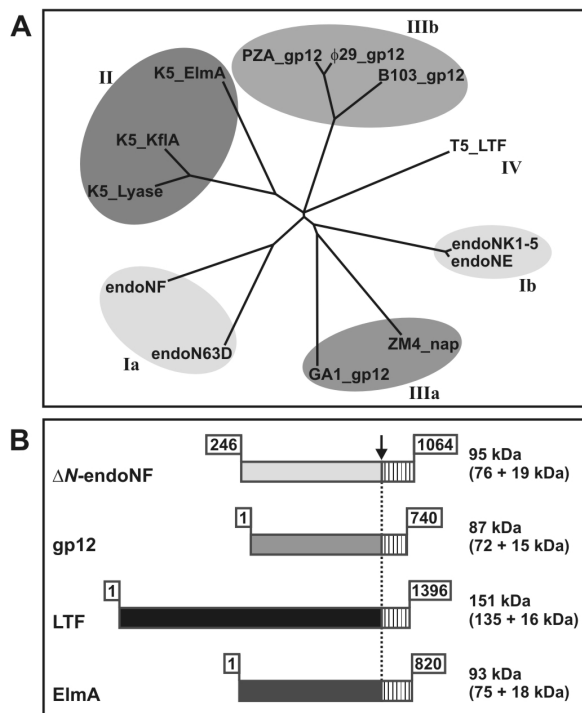


FIGURE 1. Phage proteins sharing a homologous CTD. **A**, phylogenetic tree based on the amino acid sequence alignment of 13 CTDs derived from the phage proteins listed in Table 1. CTDs derived from proteins with similar N-terminal parts are depicted in the same shade of gray and numbered according to the protein families given in Table 1: endosialidases (Ia and Ib), K5 capsule depolymerases (II), neck appendage proteins (IIIa and IIIb), and tail fiber proteins (IV). Accession numbers of the used protein sequences are given in the legend to Fig. S1 in the supplemental material. **B**, schematic representation of the four different phage proteins analyzed in the present study: the endosialidase of phage K1F, lacking the N-terminal capsid-binding domain (ΔN -endoNF), the neck appendage protein gp12 of *Bacillus* phage GA-1, the L-shaped tail fiber protein of coliphage T5 (LTF), and the K5 lyase ElmA of an *E. coli* K5 prophage. The nonrelated N-terminal parts are shown in different shades of gray, whereas homologous CTDs are shown in white with the amino acid numbers given above. Molecular masses are indicated for the primary translation products containing an N-terminal Strep-tag and a C-terminal His₆ tag, and the values in parentheses indicate the masses for the respective cleavage products generated by proteolytic processing at a highly conserved serine residue as indicated by an arrow.

Similar to the results obtained for endoNF, truncated gp12 which lacks the CTD already in the nascent protein (gp12 Δ C) was found exclusively as insoluble, SDS-sensitive species (lane *i*). This observation demonstrates that the presence of the CTD in the nascent protein is essential for proper folding of endoNF and gp12. However, in both cases, release of the CTD during maturation is not a prerequisite for oligomerization since SDS-resistant complexes were observed for the non-cleavable mutants endoNF-S911A and gp12-S620A.

Mutation of single conserved amino acids interferes with the chaperone function of the CTD – To further characterize the chaperone function of the CTD, we performed a comparative mutagenesis study with endoNF and gp12. Previously, we have shown that alanine substitution of a single histidine residue located in the center of the CTD (His-954 in endoNF) disturbed proper folding of endosialidases, leading to the accumulation of monomeric, inactive protein (19). In the present study, the identification of more proteins with a homologous CTD revealed that the targeted histidine is not conserved in all proteins but is substituted by isoleucine in a subset of neck appendage proteins (see

cluster III in Fig. S1). However, this position is followed by a glycine (Gly-956 in endoNF) that is strictly conserved in all proteins (see Fig. S1). To proof that highly conserved residues are essential for the chaperone function of the CTD in endoNF and gp12, the following amino acids were exchanged by alanine: the Asx residue that is part of the cleavage site (endoNF-N912A and gp12-D612A), the glycine residue in cluster III (endoNF-G956A and gp12-G665A) and an arginine residue in cluster V (endoNF-R1035A and gp12-R720A). In addition, the mutants endoNF-H954A and gp12-H663A were analyzed. Notably, all mutated proteins were expressed only partially as soluble proteins and endoNF variants were hardly detectable in the soluble fraction (Fig. 3A). For none of the mutants release of the CTD was observed and exclusively un-cleaved, full-length proteins were detected. However, in contrast to the non-cleavable mutants endoNF-S911A and gp12-S620A described in the previous paragraph, all amino acid exchanges within the CTD prevented the formation of SDS-resistant complexes. In addition, the endoNF mutants N912A, H954A, G956A, and R1035A completely lost enzymatic activity, highlighting the importance of a functional CTD for folding and assembly of the catalytic trimer. Together, the data presented in Fig. 3 demonstrate that the CTD plays a crucial role in folding and assembling not only of endosialidases but also of other, unrelated phage proteins.

CTD swapping between endoNF and gp12 is compatible with the assembly of active endoNF trimers – Assuming similar functions of the C-terminal part in all tail spike and fiber proteins listed in Table 1, our next questions was whether CTDs are interchangeable between proteins belonging to different families. To address this point, the chimera ΔN -endoNF-gp12 was constructed by replacing the CTD of endoNF by the respective fragment of gp12 (Fig. 4A). The resulting construct was expressed in *E. coli* BL21 with N-terminal Strep- and C-terminal His₆ tag. Released protein fragments were purified by affinity chromatography using a Strep-Tactin column for isolation of ΔN -endoNF* and a Ni²⁺-chelating column for isolation of CTDs. Proteins were analyzed by SDS-PAGE showing N-terminal fragments in Fig. 4B and CTDs in Fig. 4D. Notably, the chimera was proteolytically processed, releasing N- and C-terminal fragments of the expected molecular masses. Irrespective of the origin of the CTD (chimera or wild-type protein), processed ΔN -endoNF* fragments formed trimeric complexes as determined by gel-filtration (Table 2). However, different complex stabilities were observed in the presence of SDS. Trimers obtained after maturation of wild-type endoNF were stable in the presence of 1.5% SDS (see Fig. 4B, lane 2). By contrast, complexes derived from the chimera were SDS-sensitive under these conditions and exclusively monomeric species were observed (Fig. 4B, lane 4). Only in the presence of $\leq 1\%$ SDS, resistant oligomers were detectable as indicated by the appearance of high molecular weight bands in non-boiled samples (Fig. 4B, lane 8). Despite slight differences in SDS-resistance, similar molar activities were observed for either ΔN -endoNF* variant (Fig. 4C), demonstrating that the gp12 CTD is able to mediate folding and assembly of

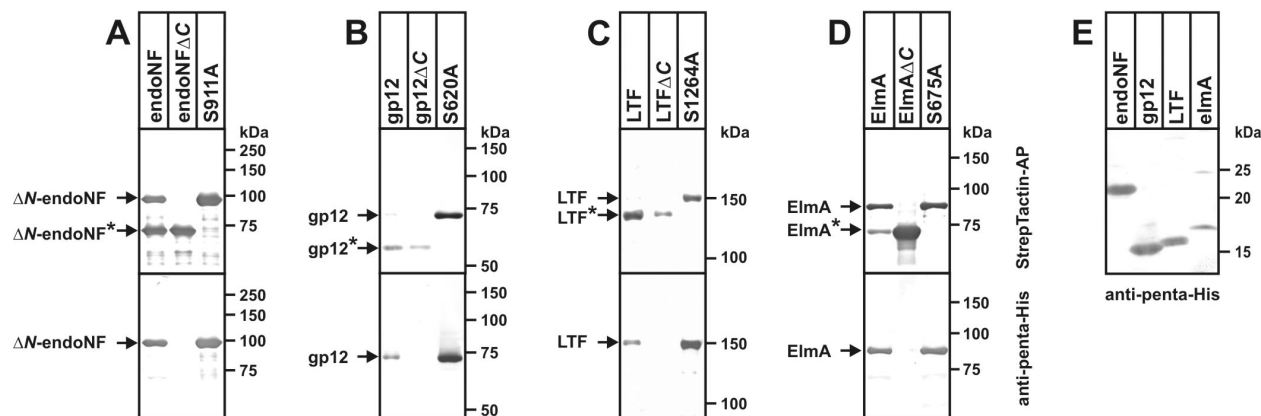


FIGURE 2. Proteolytic maturation of phage proteins containing an endosialidase CTD. A–D, Western blot analysis of ΔN -endoNF (A), gp12 (B), LTF (C), and ElmA (D). For each protein, the following variants were expressed in *E. coli* BL21: wild type, C-terminally truncated protein lacking the CTD (ΔC), and a mutant in which the highly conserved serine residue that is part of the cleavage site Ser ↓ Asx was exchanged by alanine. All three variants were expressed with an N-terminal Strep-tag, and wild-type and serine mutants were additionally epitope-tagged with a C-terminal His₆. Bacterial lysates were separated by SDS-PAGE followed by Western blot analysis using StrepTactin-AP (top) and anti-penta-His antibody (bottom) for N- or C-terminal detection, respectively. Bands corresponding to full-length and proteolytically processed protein (marked with an asterisk) are indicated by an arrow. E, Western blot analysis of released CTDs. Bacterial lysates of *E. coli* BL21 expressing ΔN -endoNF, gp12, LTF, or ElmA were separated by 14% SDS-PAGE, and the C-terminally His₆-tagged protein fragments were detected by Western blot analysis with anti-penta-His antibody.

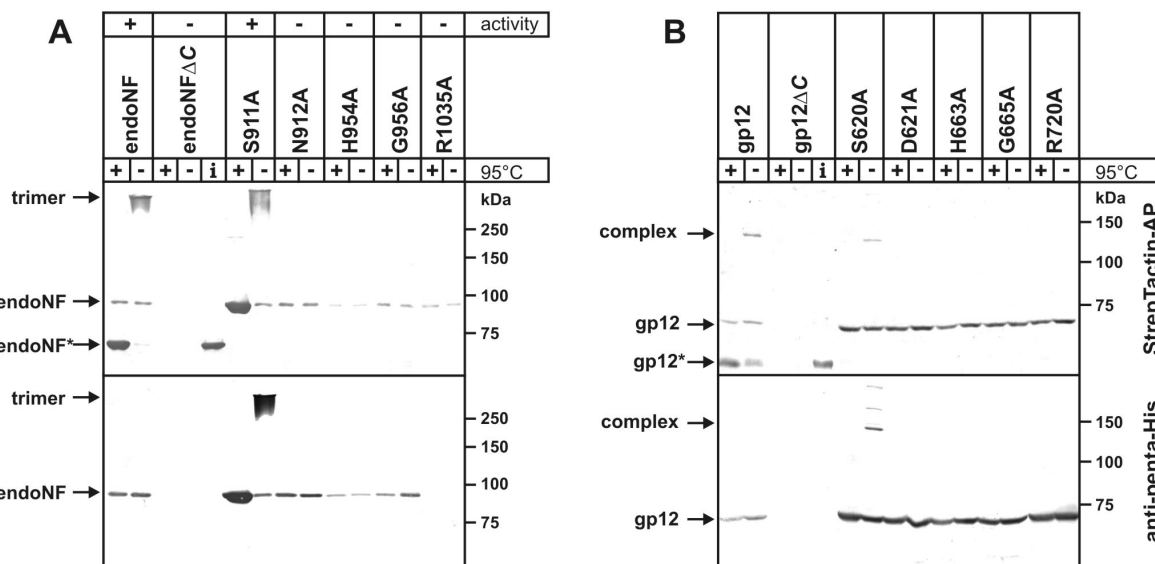


FIGURE 3. Complex formation of ΔN -endoNF and gp12 of *Bacillus* phage GA-1. For ΔN -endoNF and gp12, the following variants containing an N-terminal Strep-tag and C-terminal His₆ tag were expressed in *E. coli* BL21: wild-type protein, a variant lacking the CTD, including His₆ tag (ΔC), and a set of mutants with the indicated amino acid exchanges. The soluble protein fractions were analyzed by Western blotting using StrepTactin-AP (top) and anti-penta-His antibody (bottom) for N- and C-terminal detection, respectively. Prior to electrophoresis, samples were incubated for 5 min in the presence of SDS either at room temperature or at 95 °C as indicated. 1.5 and 1% SDS were used for the analysis of endoNF and gp12, respectively. In the case of the C-terminally truncated variants, also the insoluble protein fraction (indicated with \ddagger) was analyzed. Results obtained for ΔN -endoNF and gp12 are shown in A and B, respectively. The enzymatic activity of ΔN -endoNF variants was determined in the soluble fractions by the thiobarbituric acid assay, and results are given in the upper lane of A.

catalytically active endosialidase trimers. CTDs released from chimeric proteins had, in perfect agreement with the corresponding fragments derived from wild-type gp12, an apparent molecular weight of 15 kDa (Fig. 4D) and, notably, were found to attain the same oligomeric state (see Table 2).

Structural characterization of the CTD of endoNF – In contrast to the ΔN -endoNF* trimer for which the crystal structure has been solved (18), no structural information is available for an endosialidase CTD. To learn more about the structural characteristics of this intra-molecular chaperone domain, we analyzed the secondary structure of the CTD of endoNF purified from two different sources: (i) as cleavage product generated by proteolytic processing and (ii) as a separately expressed protein. For this purpose, a construct encoding

only the CTD of endoNF was generated and expressed in the absence of the N-terminal catalytic part. Both CTD variants were affinity purified on Ni²⁺-chelating columns by C-terminally added His₆ tags. As shown by size exclusion chromatography, either CTD variant formed hexameric complexes (Table 2). Most probably, these complexes are dimers of trimers and dimerization is presumably mediated by N-terminal regions of the CTD trimer which become accessible only after cleavage. Accordingly, the non-cleavable mutant ΔN -endoNF-S911A was found exclusively as a trimer and no dimerization via the trimeric CTDs was observed (Table 2). To determine if the CTD is an independent folding unit, we compared the secondary structure of CTDs obtained by separate expression and derived from proteolytically processed endoNF. Purified CTDs were

analyzed by circular dichroism (CD). As summarized in Fig. 5, almost identical CD spectra were obtained for both variants, demonstrating that the CTD represents an autonomous domain which folds independent of any N-terminal protein context. Interestingly, the CTD was found predominantly α -helical. This structural characteristic clearly separates the domain from the proteolytically processed Δ N-endoNF* trimer which is build-up almost exclusively by β -folds (18).

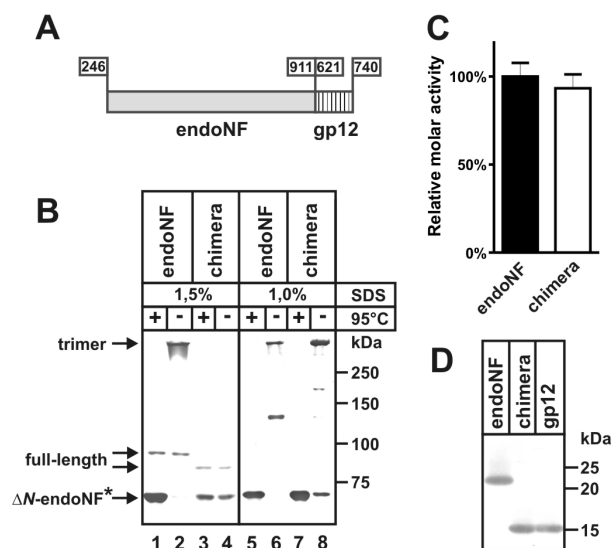


FIGURE 4. Analysis of an endoNF-gp12 chimera. A, schematic representation of the generated chimera containing the catalytic part of endoNF (aa 246–911) and the CTD of gp12 (aa 621–740). B, chimeric Δ N-endoNF-gp12 was expressed in *E. coli* BL21 as a variant containing an N-terminal Strep-tag and a C-terminal His₆ tag. Proteins were analyzed by Western blotting using StrepTactin-AP for detection. Soluble protein fractions incubated in the presence of 1.5% SDS are shown in lanes 1–4. Affinity-purified proteolytically processed protein (Δ N-endoNF*) derived either from wild-type Δ N-endoNF (lanes 5 and 6) or the chimera (lanes 7 and 8) was incubated in the presence of 1.0% SDS prior to electrophoresis. C, relative molar activity of affinity-purified Δ N-endoNF* derived from either Δ N-endoNF or chimeric Δ N-endoNF-gp12. Activities were determined by the thiobarbituric acid assay. Data are means \pm S.D. of two independent experiments performed in duplicates, and values obtained for Δ N-endoNF-derived fragments were set to 100%. D, CTDs released from Δ N-endoNF, the chimera, and gp12 were affinity-purified by Ni²⁺-chelating chromatography. Proteins were separated by SDS-PAGE and analyzed by Western blotting with anti-penta-His antibody.

TABLE 2

Quaternary structure of endoNF and gp12.

The oligomerization state of the indicated protein fragments was determined for affinity purified proteins by size exclusion chromatography. The molecular masses (M_R) of the monomeric species were calculated from respective amino acid sequences including epitope tags. The oligomeric state was calculated from the values obtained by gel filtration divided by the M_R calculated for the monomer.

Protein fragment	Molecular mass (monomer, calculated) kDa	Molecular mass (gel filtration) kDa	Oligomeric state
ΔN-endoNF*			
Obtained from wild-type Δ N-endoNF	76.1	195	2.6
Obtained from the chimera Δ N-endoNF-gp12	76.1	191	2.5
ΔN-endoNF-S911A	95.0	288	3.0
C-terminal domain of endoNF			
Obtained from Δ N-endoNF	18.9	102	5.4
Obtained by separate expression of aa 913-1064	18.9	100	5.3
gp12*	71.9	410	5.7
gp12-S620A	87.1	635	7.3
C-terminal domain of gp12			
Obtained from wild-type gp12	15.2	69	4.5
Obtained from the chimera Δ N-endoNF-gp12	15.2	65	4.3

Impact of N- and C-terminal domains on the thermostability of endoNF trimers – Analysis of truncated endosialidases demonstrated that the N-terminal capsid binding domain is not essential for trimer assembly and activity (17,18). However, so far nothing is known whether N-terminal capping or CTD-release stabilize the catalytic trimer. Aiming to understand in more detail the impact of N- and C-terminal domains on thermostability of endosialidases, a comparative analysis of the following endoNF variants was performed: (i) wild-type, (ii) Δ N-endoNF lacking the capsid binding domain, and (iii) Δ N-endoNF-S911A lacking the capsid binding domain but retaining the CTD (see Fig. 6 for schematic representation of the isolated fragments). Equal aliquots of the purified proteins were incubated for five minutes in the presence of 1.5% SDS at temperatures ranging from 4° to 95°C and trimer stability was monitored by SDS-PAGE. Dissociation of wild-type trimers (Fig. 6A) into monomers with an apparent molecular mass of 100 kDa started at 49°C and was completed at 60°C. An identical melting behavior was observed for Δ N-endoNF* (Fig. 6B), demonstrating that the N-terminal capsid binding domain has no significant influence on the stability of the functionally folded trimer. By contrast, the non-cleavable mutant Δ N-endoNF-S911A showed a different picture (Fig. 6C). At temperatures above 40°C, an additional band of low electrophoretic mobility became visible, paralleled by dissociation of the complex as indicated by appearance of monomeric species. Most likely, the additional high molecular weight band represents a partially unfolded and therefore more loosely packed intermediate, a topological variant that forms before disintegration of the trimeric complex. Interestingly, the non-cleavable variant Δ N-endoNF-S911A seems to be less stable and disintegration of trimers was observed already at temperatures 6°C lower than for complexes built-up by proteolytically processed Δ N-endoNF* fragments. This observation suggests that cleavage of the CTD increases complex stability by increasing the energy barrier for the unfolding process.

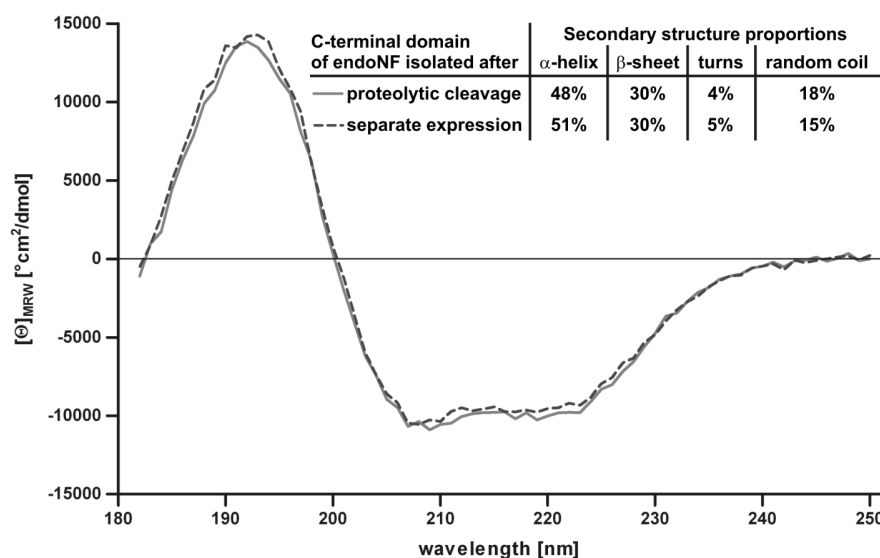


FIGURE 5. **Secondary structure analysis of the CTD of endoNF.** The CTD of endoNF was purified after release from wild-type endoNF and after separate expression of aa 913–1064. For both variants, the secondary structure content was analyzed by far-UV CD spectroscopy. The CD signals are given as mean residue weight ellipticity $[\Theta]_{MRW}$. Spectra obtained for CTDs derived from proteolytically processed endoNF and from separate expression are shown as *solid* and *dashed lines*, respectively. Secondary structure proportions were calculated using CDpro suite (24, 25).

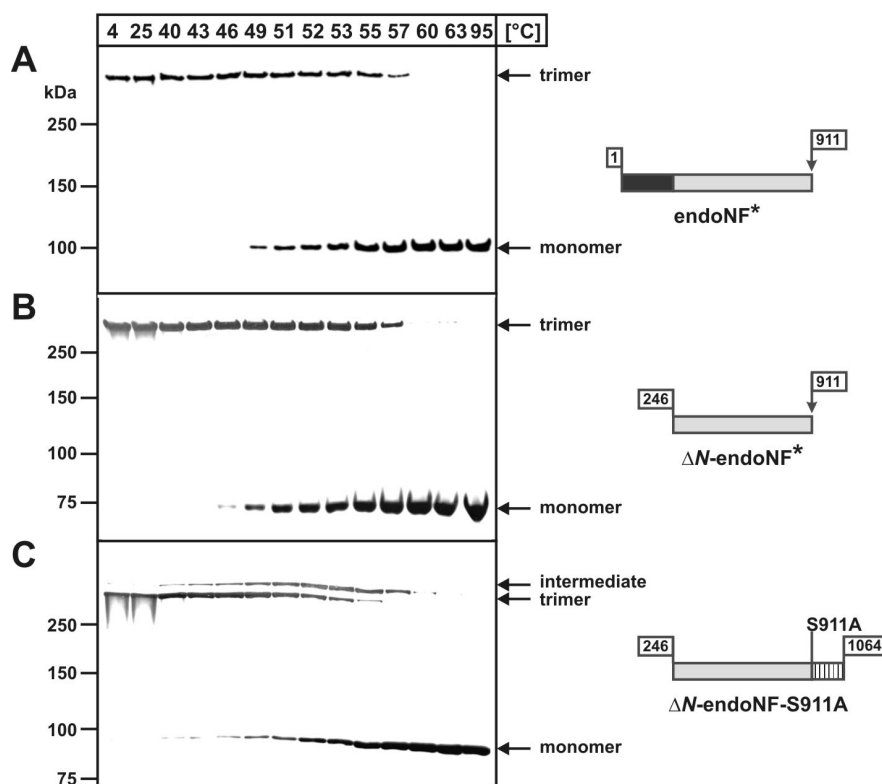


FIGURE 6. **Thermostability of endoNF variants.** Wild-type endoNF, ΔN -endoNF, and the mutant ΔN -endoNF-S911A were expressed in *E. coli* BL21. Proteolytically processed *N*-terminal parts of endoNF (endoNF*) and ΔN -endoNF (ΔN -endoNF*) and full-length ΔN -endoNF-S911A were affinity purified on StrepTactin columns. Schematic representations of the purified fragments are shown on the *right*. For each variant, aliquots of purified protein were incubated in the presence of 1.5% SDS for 5 min at the indicated temperatures. After separation by 6% SDS-PAGE, proteins were visualized by Western blotting with StrepTactin-AP. Results obtained for endoNF*, ΔN -endoNF*, and ΔN -endoNF-S911A are shown in A–C, respectively. Bands corresponding to monomeric and trimeric protein species are indicated with an *arrow*.

DISCUSSION

In this study we characterized a novel C-terminal intra-molecular chaperone domain, essential for assembling and folding of endosialidases and other tail spike and fiber proteins of diverse phages. Altogether 13 phage proteins belonging to 4 different protein families

were identified that contain this CTD: (i) endosialidases, the tail spike proteins of *E. coli* K1 specific phages which are required for depolymerization of the polysialic acid capsule of the host (26,27); (ii) K5-lyases, the capsule degrading tail spike proteins of phages infecting *E. coli* K5 (28, 29); (iii) the L-shaped tail fibers of coliphage T5, which accelerate phage adsorption by reversible binding to the O-antigen of the host (30); and (iv) the neck appendage protein gp12 of *Bacillus* phages GA-1, $\phi 29$, PZA, and B103, which forms appendages that are essential for phage attachment (31).

By analyzing a representative of each family (endoNF, ElmA, LTF, and gp12 of *Bacillus* phage GA-1), we showed that independent of the *N*-terminal sequence, the CTD is released during protein maturation. Proteolytic cleavage occurs at a highly conserved Ser ↓ Asx cleavage site. The essential role of the serine residue was confirmed by alanine substitution, which prevented cleavage but did not affect folding and complex assembly.

Within the CTD, we identified four critical amino acids (Asn-912, His-954, Gly-956, and Arg-1035 in endoNF), which are essential for proper function of the chaperone domain. As shown for endoNF and gp12, single amino acid exchanges prevented the formation of functional oligomers and only inactive monomeric species were observed. Notably, single alanine substitutions

of the strictly conserved glycine and arginine residues led to the expression of full-length proteins, although the cleavage site was not affected. Since oligomerization seems to be a prerequisite for proteolytic processing, cleavage at the Ser ↓ Asx-site might be mediated by an autocatalytic reaction. By losing the capability to

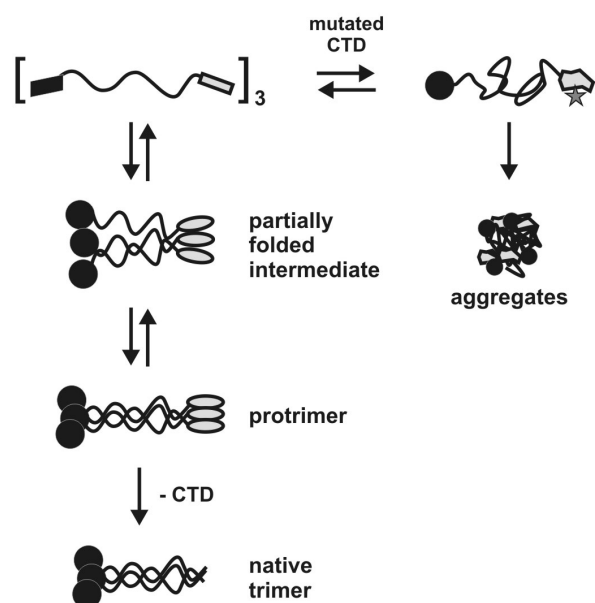


FIGURE 7. **Proposed folding pathway of endoNF.** In the schematic representation of endoNF, the CTD is shown in gray, and the folded β -propeller is shown as a black circle.

trimerize, the mutants H954A, G956A, and R1035A might have lost the spatial conditions required for inter- or intra-molecular self-cleavage.

As shown for endoNF and gp12, cleavage is not prerequisite for oligomerization and both non-cleavable mutants, endoNF-S911A and gp12-S620A, formed SDS-resistant complexes. Whether cleavage is required for assembly onto phage particles is not yet known. However, in *Bacillus* phage ϕ 29, which is closely related to GA-1, only processed gp12 was found as a building block of the neck appendages (32, 33). Using gel filtration, we determined the oligomeric state of mature GA-1 gp12 to be hexameric. This is in agreement with the stoichiometry found in ϕ 29 particles where 60 gp12 subunits form the 10 neck appendages (34, 35). However, because no crystal structure is available for gp12, it is unknown whether the hexameric neck appendages are arranged as dimers of trimers.

In contrast to gp12 of *Bacillus* phage GA-1, only very weak similarity (~11% amino acid identity) was observed between CTDs of the three neck-appendage proteins of *Bacillus* phages ϕ 29, PZA, and B103, and the CTD of endosialidases. In the phylogenetic tree based on all identified endosialidase-like CTDs they form an individual branch clearly separate from CTDs of other neck-appendage proteins (branch IIIa and b in Fig. 1). Nevertheless, the cleavage site and all amino acids shown in this study to be essential for CTD function are highly conserved. Moreover, the cleavage site seems to be functional since for mature gp12 of ϕ 29 a molecular mass of 73,888 Da was determined by mass spectrometry (34), which is in agreement with a molecular mass of 73,963 Da calculated for cleavage after Ser-691 within the Ser \downarrow Asx cleavage site.

Proteolytic maturation is a common theme among phage proteins involving phage-derived proteases or autocatalytic processes (36–40). In contrast to pro-enzymes like mammalian digestive serine proteases, the function of proteolytic maturation of phage proteins is

not restricted to enzyme activation. In the case of the major head proteins of coliphage T4, cleavage is required for head expansion and DNA packaging (41). Similarly, release of the endosialidase CTD is not prerequisite for an active enzyme (19) but might provide a mechanism to increase protein stability. Based on the results obtained in the present study, we propose a folding pathway as depicted exemplarily for endoNF in Fig. 7. Folding of the nascent polypeptide chain starts with formation of independent domains like the endosialidase β -propeller and the CTD. Upon release from the ribosome, monomers assemble into trimers. This step crucially depends on a functional CTD since alanine substitutions of critical amino acid residues within this domain prevented oligomerization and caused accumulation of SDS-sensitive species. CTD-assisted trimer assembly may involve several partially folded intermediates leading to an SDS-resistant protrimer that still contains CTDs. By generation of the non-cleavable variant endoNF-S911A we mimicked the protrimer state. Before dissociation into monomers, trimeric endoNF-S911A became partially unfolded and an SDS-resistant intermediate with lower electrophoretic mobility than the native trimer appeared. Compared to processed complexes, full-length trimers showed decreased thermostability, indicating that CTD-release stabilizes the native trimer. Thus, proteolytic cleavage may lead to a kinetically stable protein, a trimer which is trapped in a specific conformation due to an unusually high unfolding barrier. In this model, the role of the CTD would reside in overcoming the even higher energy barrier during the folding process. Consequently, mutants with a defective or missing CTD are unable to adopt the transition state conformation required for the formation of functional trimers.

A similar mechanism to increase protein stability was described for the α -lytic protease (α LP) of *Lyso-bacter enzymogenes*, a secreted serine protease involved in degradation of other soil bacteria (42). Folding of α LP essentially depends on the Pro domain which stabilizes the folding transition state and the native state of α LP. Release of the Pro domain traps the mature protein in a metastable conformation of high rigidity and low susceptibility to proteases (42). In contrast to the endosialidase CTD, the Pro region is located at the N-terminus. In addition, Pro can be expressed in *trans* as a separate polypeptide chain to assist folding of α LP, whereas the endosialidase CTD acts only in *cis* as part of the precursor protein (19). These differences may reflect the fact that Pro mediates folding of a monomeric protein, while the endosialidase CTD is involved in folding and assembling of an oligomeric complex.

Increasing longevity through kinetic stability seems to be a general mechanism observed for SDS-resistant proteins (43). Many of these are viral proteins and form oligomeric complexes with a high content of β -folds. Interestingly, some also possess an assembly module. In the tail spike protein of *Salmonella* phage P22, an endorhamnosidase with triple β -helix domain (44), the C-terminal domain is essential for folding and assembly of the SDS-resistant trimer (45, 46). However, unlike the endosialidase CTD, the caudal fin domain of P22 remains part of the native trimer. This is also the case

for the human adenovirus type 2 adhesin (Ad2) which forms trimeric fibers with a triple- β -spiral shaft (47). The globular C-terminal domain is not only involved in receptor binding but also required for complex assembly. Another example for C-terminally located folding domains is the “foldon”, the 27-aa CTD of fibritin, a segmented coiled-coil homotrimer which forms the fibrous whiskers of coliphage T4 (48). The foldon forms a β -propeller and is essential for correct fibritin assembly. Interestingly, function of the foldon is independent of the N-terminal protein part and the foldon can substitute the corresponding assembly domain of Ad2 (49). Moreover, the foldon was successfully used to stabilize HIV gp120, collagen, and T4 long tail fibers (50–52). However, in all cases the C-terminal assembly domain remains part of the trimeric protein complex and is composed of β -folds. In contrast, CD analysis performed in the present study revealed that the endosialidase CTD is predominantly α -helical, even though the catalytic trimer of endoNF is an all- β -protein (18). Both proteolytically released and separately expressed CTDs showed identical secondary and quaternary structures, demonstrating that the CTD represents an independent folding unit. Although this domain is trimeric in the non-cleaved mutant endoNF-S911A, hexamers were observed for processed CTDs, suggesting that dimerization of trimers occurs after release from the protimer. Interestingly, gp57, a phage-derived chaperone involved in assembly of the T4 long and short tail fibers (53, 54), is also largely α -helical and forms trimers which dimerize to hexamers (55). Future studies aiming at solving the 3D-structures will reveal whether both proteins adopt similar folds despite lack of sequence similarity.

Similar to the foldon of fibritin, which can substitute the CTD of Ad2, the C-terminal chaperone domain characterized in the present study promotes stable complex formation independent of the N-terminal protein context. In the endoNF-gp12-chimera, the CTD of gp12 mediated the formation of SDS-resistant endoNF trimers, which were proteolytically processed and enzymatically as active as wild-type proteins. However, compared to wild-type trimers, a slightly decreased SDS-resistance was observed for chimera-derived complexes, indicating minor structural differences. Sequence variations observed between CTDs of endoNF and gp12 might, therefore, reflect an adaptation to the respective N-terminal part.

In contrast to the P22 tail spike, Ad2, and fibritin, the C-terminal chaperone domain of endosialidases, K5-lyases, LTF and gp12 is removed after completing the job. Proteolytic processing may have evolved not only to increase the unfolding barrier but also to avoid structural constraints imposed by the presence of the CTD.

Acknowledgements – We thank Prof. Dr. Robert Seckler (Institut für Physikalische Biochemie, Universität Potsdam, Potsdam-Golm, Germany) for access to the circular dichroism facility and helpful discussions. Miriam Schiff and Jasmin Hotzy are acknowledged for expert technical assistance.

REFERENCES

- Ackermann, H. W. (1998) *Adv. Virus Res.* **51**, 135-201
- Stirm, S. and Freund-Mölbart, E. (1971) *J. Virol.* **8**, 330-342
- Lindberg, A. A. (1977) in *Surface Carbohydrates of the Prokaryotic Cell* (Sutherland, I. W., ed) pp. 289-356, Academic Press, Inc., London
- Weigle, P. R., Scanlon, E., and King, J. (2003) *J. Bacteriol.* **185**, 4022-4030
- Robbins, J. B., McCracken, G. H., Jr., Gotschlich, E. C., Ørskov, F., Ørskov, I., and Hanson, L. A. (1974) *N. Engl. J. Med.* **290**, 1216-1220
- Sarff, L. D., McCracken, G. H., Schiffer, M. S., Glode, M. P., Robbins, J. B., Ørskov, I., and Ørskov, F. (1975) *Lancet* **1**, 1099-1104
- Mühlenhoff, M., Eckhardt, M., and Gerardy-Schahn, R. (1998) *Curr. Opin. Struct. Biol.* **8**, 558-564
- Leying, H., Suerbaum, S., Kroll, H. P., Stahl, D., and Opferkuch, W. (1990) *Infect. Immun.* **58**, 222-227
- Kwiatkowski, B., Boschek, B., Thiele, H., and Stirm, S. (1982) *J. Virol.* **43**, 697-704
- Kwiatkowski, B., Boschek, B., Thiele, H., and Stirm, S. (1983) *J. Virol.* **45**, 367-374
- Gross, R. J., Cheasty, T., and Rowe, B. (1977) *J. Clin. Microbiol.* **6**, 548-550
- Vimr, E. R., McCoy, R. D., Vollger, H. F., Wilkison, N. C., and Troy, F. A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1971-1975
- Miyake, K., Muraki, T., Hattori, K., Machida, Y., Watanabe, M., Kawase, M., Yoshida, Y., and Iijima, S. (1997) *J. Ferm. Bioeng.* **84**, 90-93
- Smith, H. W. and Huggins, M. B. (1982) *J. Gen. Microbiol.* **128**, 307-318
- Botstein, D. (1980) *Ann. N.Y. Acad. Sci.* **354**, 484-490
- Casjens, S. R. (2005) *Curr. Opin. Microbiol.* **8**, 451-458
- Stummeyer, K., Schwarzer, D., Claus, H., Vogel, U., Gerardy-Schahn, R., and Mühlenhoff, M. (2006) *Mol. Microbiol.* **60**, 1123-1135
- Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., and Ficner, R. (2005) *Nat. Struct. Mol. Biol.* **12**, 90-96
- Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M., and Gerardy-Schahn, R. (2003) *J. Biol. Chem.* **278**, 12634-12644
- Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S., Lottspeich, F., and Frosch, M. (1995) *Mol. Microbiol.* **16**, 441-450
- Kaliman, A. V., Kulshin, V. E., Shlyapnikov, M. G., Ksenzenko, V. N., and Kryukov, V. M. (1995) *FEBS Lett.* **366**, 46-48
- Wang, J., Jiang, Y., Vincent, M., Sun, Y., Yu, H., Wang, J., Bao, Q., Kong, H., and Hu, S. (2005) *Virology* **332**, 45-65
- Zhang, Y., Muylers, J. P., Testa, G., and Stewart, A. F. (2000) *Nat. Biotechnol.* **18**, 1314-1317
- Sreerama, N. and Woody, R. W. (2004) *Methods Enzymol.* **383**, 318-351
- Sreerama, N. and Woody, R. W. (2000) *Anal. Biochem.* **287**, 252-260
- Pelkonen, S., Pelkonen, J., and Finne, J. (1989) *J. Virol.* **63**, 4409-4416
- Hallenbeck, P. C., Vimr, E. R., Yu, F., Bassler, B., and Troy, F. A. (1987) *J. Biol. Chem.* **262**, 3553-3561
- Clarke, B. R., Esumeh, F., and Roberts, I. S. (2000) *J. Bacteriol.* **182**, 3761-3766

29. Scholl, D., Rogers, S., Adhya, S., and Merrill, C. R. (2001) *J. Virol.* **75**, 2509-2515
30. Heller, K. and Braun, V. (1979) *J. Bacteriol.* **139**, 32-38
31. Meijer, W. J., Horcajadas, J. A., and Salas, M. (2001) *Microbiol. Mol. Biol. Rev.* **65**, 261-287
32. Tosi, M. E., Reilly, B. E., and Anderson, D. L. (1975) *J. Virol.* **16**, 1282-1295
33. Carrascosa, J. L., Camacho, A., Moreno, F., Jimenez, F., Mellado, R. P., Vinuela, E., and Salas, M. (1976) *Eur. J. Biochem.* **66**, 229-241
34. Peterson, C., Simon, M., Hodges, J., Mertens, P., Higgins, L., Egelman, E., and Anderson, D. (2001) *J. Struct. Biol.* **135**, 18-25
35. Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G., and Baker, T. S. (1998) *Cell* **95**, 431-437
36. Wang, S., Chandramouli, P., Butcher, S., and Dokland, T. (2003) *Virology* **314**, 1-8
37. Laemmli, U. K. (1970) *Nature* **227**, 680-685
38. Conway, J. F., Duda, R. L., Cheng, N., Hendrix, R. W., and Steven, A. C. (1995) *J. Mol. Biol.* **253**, 86-99
39. Marvik, O. J., Jacobsen, E., Dokland, T., and Lindqvist, B. H. (1994) *Virology* **205**, 51-65
40. 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
41. Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Ruger, W. (2003) *Microbiol. Mol. Biol. Rev.* **67**, 86-156, table
42. Cunningham, E. L., Jaswal, S. S., Sohl, J. L., and Agard, D. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11008-11014
43. Manning, M. and Colon, W. (2004) *Biochemistry* **43**, 11248-11254
44. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) *Science* **265**, 383-386
45. Gage, M. J. and Robinson, A. S. (2003) *Protein Sci.* **12**, 2732-2747
46. Kreisberg, J. F., Betts, S. D., Haase-Pettingell, C., and King, J. (2002) *Protein Sci.* **11**, 820-830
47. van Raaij, M. J., Schoehn, G., Burda, M. R., and Miller, S. (2001) *J. Mol. Biol.* **314**, 1137-1146
48. Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V., and Rossmann, M. G. (1997) *Structure* **5**, 789-798
49. Papanikolopoulou, K., Forge, V., Goeltz, P., and Mitraki, A. (2004) *J. Biol. Chem.* **279**, 8991-8998
50. Yang, X., Lee, J., Mahony, E. M., Kwong, P. D., Wyatt, R., and Sodroski, J. (2002) *J. Virol.* **76**, 4634-4642
51. Frank, S., Kammerer, R. A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bachinger, H. P., and Engel, J. (2001) *J. Mol. Biol.* **308**, 1081-1089
52. Miroshnikov, K. A., Marusich, E. I., Cerritelli, M. E., Cheng, N., Hyde, C. C., Steven, A. C., and Mesyanzhinov, V. V. (1998) *Protein Eng.* **11**, 329-332
53. Dickson, R. C. (1973) *J. Mol. Biol.* **79**, 633-647
54. Burda, M. R. and Miller, S. (1999) *Eur. J. Biochem.* **265**, 771-778
55. Ali, S. A., Iwabuchi, N., Matsui, T., Hirota, K., Kidokoro, S., Arai, M., Kuwajima, K., Schuck, P., and Arisaka, F. (2003) *Biophys. J.* **85**, 2606-26

5.1 – Supplemental Material

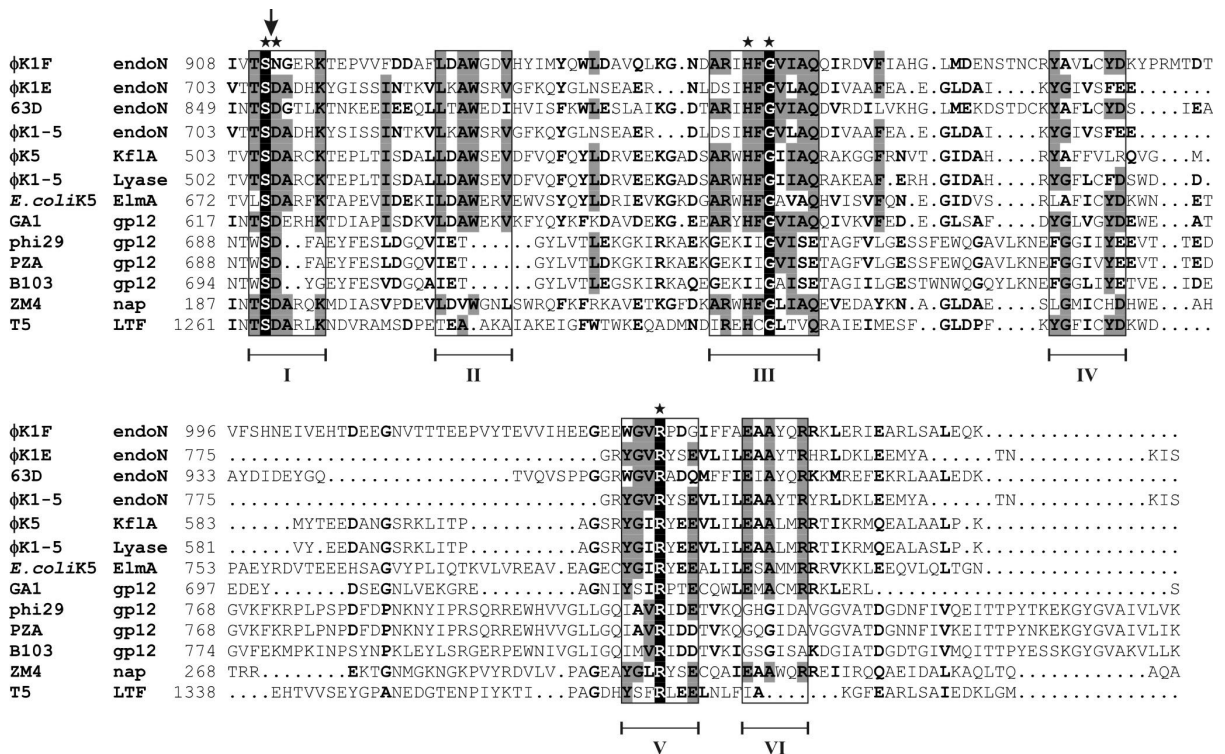


Figure S1. **Partial multiple sequence alignment of phage proteins sharing a homologous C-terminal domain.** Based on the C-terminal amino acid sequence of the indicated phage proteins a multiple sequence alignment was determined by ClustalV using sequences with the following accession numbers: CAD44528 (endoNF of coliphage K1F), CAA85449 (endoNE of coliphage K1E), BAA32990 (endoN63D of coliphage 63D), AAG59822 (endoNK1-5 of coliphage K1-5), CAA71133 (K5-lyase Kf1A of coliphage K5), AAG59821 (K5-lyase of coliphage K1-5), CAA65353 (K5-lyase ElmA of a not yet characterised *E. coli* K5 prophage), CAC21533 (neck appendage protein gp12 of *Bacillus* phage GA-1), AAA32285 (gp12 of *Bacillus* phage ϕ 29); NP_040730 (gp12 of *Bacillus* phage PZA), NP_690646 (gp12 of *Bacillus* phage B103), YP_162134 (potential neck appendage protein precursor identified in the genome of *Zymomonas mobilis* strain ZM4), and CAJ29339 (L-shaped tail fibre protein LTF of coliphage T5). Amino acid numbers are shown on the left. The location of the proteolytic cleavage site Ser \downarrow Asx is indicated by an arrow. Residues that are conserved in seven or more of the 13 sequences are shaded with grey, whereas residues that are identical in all sequences are shaded with black. Conservative residues shared by seven or more of the 13 sequences are shown in bold. Clusters of conserved amino acids are boxed and numbered. Amino acids that were exchanged to alanine in the present study are marked with asterisks.

Chapter 6 – Proteolytic Maturation of Endosialidase F Is Essential to Allow Efficient Binding to Polysialic Acid

This manuscript is to be submitted to *The Journal of Biological Chemistry* after review by the co-authors

'Proteolytic Maturation of Endosialidase F Is Essential to Allow Efficient Binding to Polysialic Acid'

David Schwarzer¹, Katharina Stummeyer¹, Thomas Haselhorst², Bastian Rode³, Thomas Scheper³, Mark von Itzstein², Martina Mühlhoff¹, and Rita Gerardy-Schahn^{1*}

¹ *Abteilung Zelluläre Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany*

² *Institute for Glycomics, Griffith University (Gold Coast Campus), PMB 50 Gold Coast Mail Centre, Queensland 9726, Australia*

³ *Institut für Technische Chemie, Leibniz Universität Hannover, Callinstrasse 3, 30167 Hannover, Germany*

* To whom correspondence should be addressed: Tel.: 49-511-532-9802; Fax: 49-511-532-8801;
E-mail: gerardy-schahn.rita@mh-hannover.de

Preface – About the Manuscript

In the preceding study it was suggested that the release of the C-terminal chaperone domain of endosialidases provides an essential step in the formation of kinetically stable complexes. In this work I analysed the impact of the C-terminal domain on the enzymatic activity or the binding to the substrate with the help of active and functionally inactive endoNF mutants. Therefore, the mutant S911A which is unable to release the C-terminal domain was comparatively tested together with maturation competent enzyme variants using ELISA-based binding assays and activity assays involving soluble as well as surface-bound polysialic acid. It could be demonstrated that the presence of the CTD decreased the binding to polySia. Since two binding sites had been identified in the crystal structure of endoNF, I investigated whether the effect observed for the S911A mutant can be resembled by rational mutagenesis of amino acid residues within one of the two sites. While depletion of this site significantly lowered polySia binding and enzymatic activity when tested with solid phase bound polySia, the enzymatic activity was increased when tested with the soluble polymer.

SUMMARY

Endosialidases, the tailspike proteins of *Escherichia coli* K1 bacteriophages, specifically bind and degrade the K1 capsule composed of α 2,8-linked polysialic acid (polySia). As identified in the crystal structure of the trimeric catalytic part of endoNF – the endosialidase of coliphage K1F, interaction with polySia is mediated by the active site and additionally by two binding sites per protein subunit. Folding and assembly of functional endosialidases crucially depend on their C-terminal chaperone domain that is released from the mature protein by cleavage at a highly conserved Serine residue. Exchanging this Serine into Alanine in endoNF (S911A) results in a non-cleavable mutant. Recently, we suggested that proteolytic cleavage is a prerequisite for gaining kinetically stable protein complexes.

Interestingly, endoNF-S911A containing the chaperone domain shows an increased molar activity, when active on soluble polySia compared to wild-type enzyme. In contrast, the mutant shows reduced enzymatic activity on surface bound polySia. However, the active site was not affected in the S911A mutant as determined with tetrameric sialic acid as minimal substrate revealing identical kinetic parameters for mutant and wild-type enzyme. Here we demonstrate that the presence of the C-terminal domain interferes with the sialic acid binding site in the stalk domain. Single and double amino acid substitutions to alanine changed the kinetic and binding properties towards the effects measured for endoNF-S911A. In total, gaining the sialic acid binding site in endosialidases, in combination with establishing kinetically stable complexes seem to be the evolutionary driving forces to maintain the proteolytic cleavage site.

Endosialidases (endo-N-acylneuraminidases, endoN) are found as specialized tailspike proteins of bacteriophages infecting the neuroinvasive bacterium *Escherichia coli* K1. This pathogen represents a major cause of meningitis and sepsis in neonates and is encapsulated in α 2,8-linked polysialic acid (polySia[#]) (1, 2). Endosialidases specifically bind to and degrade the dense polysaccharide capsule of *E. coli* K1, which was shown to function as efficient infection barrier for Coliphages lacking polySia degrading tailspikes (3). The enzymes are required to anchor the phage to the bacterial outer membrane for initiating the bacteriophage infection cycle (4-7).

The sugar polymer polySia is found not only as capsular structure of pathogenic bacteria but is also widely expressed during vertebrate development and remains an important modulator of neuronal plasticity in the adult brain (8). Since Phage-borne endosialidases are the only known enzymes that specifically degrade polySia, they are widely applied in neurosciences and polySia research. In contrast to exosialidases that cleave off terminal α -ketosidically linked sialic acid residues (exo-acting), endosialidases require α 2,8-glycosidically linked sialic acid oligo- or polymers with a minimal degree of polymerization between 5 and 8 residues for cleavage (9, 10). Enzymatically active endosialidases have been cloned from several bacteriophage sources (11-16). They all share a common modular architecture that can be subdivided in three major parts: (i) the N-terminal part was identified as attachment domain that anchors the endosialidase tailspikes to the phage particle. This domain can be of variable length and was shown to be dispensable for enzymatic activity (15, 17). (ii) The highly conserved central catalytic part comprises the polySia binding and cleavage activities while (iii) the small C-terminal domain (CTD) is proteolytically cleaved off during maturation of the enzymes but was found to be required for proper folding and assembly of endosialidases (15, 18).

Endosialidases form very stable trimeric complexes that show resistance against denaturation with SDS (15). The first crystal structure of an endosialidase comprising the central catalytic domain of the enzyme cloned from Φ K1F (endoNF) has been solved in our group (6). The homotrimer resembles the cap and stalk outline of a mushroom. Each endoNF subunit folds into three distinct domains and reveals a unique domain architecture in combining structural elements characteristic for exosialidases and bacteriophage tailspike proteins. The ‘mushroom cap’ harbors the three active sites of the trimer that are located in three six-bladed β -propeller structures common for exosialidases. Close to each active site, a di-sialic acid binding site was identified that is provided by a lectin-like β -barrel domain extending from the β -propeller (6). The ‘mushroom stalk’ is build up from the intertwining C-terminal parts of the three central domains and is composed of a triple β -helix that is interrupted by a short triple β -prism domain. Both folds are common for phage tailspike proteins and were suggested to mediate the unusual complex stability of these proteins (19). Interestingly, the stalk domain of endoNF additionally contains a

sialic acid binding site indicating that this domain may also be involved in substrate binding (6).

In a recent publication we furthermore analyzed the proteolytically cleaved off CTD of endosialidases in more detail and could identify this domain as a general intramolecular chaperone in phage tailspike and fiber proteins that are otherwise unrelated to endosialidases. Remarkably, the CTD was found to be proteolytically cleaved off the mature tailspike at a highly conserved serine residue in all analyzed proteins while an exchange of this serine to alanine lead to non-cleavable mutants. Interestingly, the non-cleaved endosialidase mutants were still found to be active thereby demonstrating that cleavage is not a prerequisite for enzymatic activity (15, 18). However, we could demonstrate that the presence of the CTD decreases the complex stability of the non-cleavable mutants and therefore suggested that proteolytic cleavage is necessary to trap the enzymes in a kinetically stable fold (18).

In this study, we describe the enzymatic properties of the non-cleavable endoNF-mutant – i.e. containing the CTD – in more detail and show a clear effect of this domain on substrate binding and enzymatic activity. By mutagenesis studies we could demonstrate that the presence of the CTD interferes with the substrate binding site located in the stalk domain of endoNF, which results in increased cleavage of soluble but decreased cleavage of surface bound polySia. Therefore, proteolytic cleavage may not only increase complex stability of the endosialidase tailspikes but also be beneficial for phage infection. Additionally we revealed insight in the interplay of binding and enzymatic activity of endosialidases that could serve as a model for other depolymerases.

EXPERIMENTAL PROCEDURES

Materials – pBluescript SK⁺ was purchased from Stratagene. pET expression vectors were obtained from Novagen. Polysialic acid (polySia) for ELISA based assays were purified according to Decher *et al.* from *Escherichia coli* K1 membrane fractions, while polySia lacking the lipid anchor for thiobarbituric acid assays (soluble polySia) was purified from the culture supernatant (20). Sialic acid oligomers from DP2 to DP4 were obtained from Nacalai Tesque.

Bacteria – The *E. coli* K1 strain used in this study is a clinical isolate obtained from the department of Medical Microbiology of the Medizinische Hochschule Hannover (Hannover, Germany). *E. coli* BL21-Gold(DE3) were purchased from Stratagene.

Site-directed mutagenesis – Mutagenesis of the sialic acid binding site in the stalk region was performed by PCR using the QuikChange site-directed mutagenesis

kit (Stratagene) following the manufacturer's guidelines with the following primer pairs (mutated nucleotides are shown in upper case): KS084 5'-c ccg gca ggg cag GCA atc ata ttt tgc ggg gg-3' and KS085 5'-cc ccc gca aaa tat gat TGC ctg ccc tgc cgg g-3' for ΔN-endoNF-R837A; KS080 5'-t gaa ggc acc agt Gca acg act ggc gc-3' and KS081 5'-gc gcc agt cgt tgC act ggt gcc ttc a-3' for ΔN-endoNF-S848A; KS082 5'-tea acg act ggc gca GCg att acg cta tat ggt gc-3' and KS083 5'-gc acc ata tag cgt aat cGC tgc gcc agt cgt tga-3' for ΔN-endoNF-Q853A. Plasmids containing the corresponding wild-type sequence were used as a template. Double and triple mutations in this binding site were subsequently generated in an appropriate order by using the respective second or third primer pair and a plasmid as template containing the respective single or double mutation. The sequence identity of all PCR products was confirmed by sequencing. PCR products were ligated into BamHI/XhoI sites of the expression vector pET22b-Strep, a modified pET22b vector containing the sequence encoding an N-terminal Strep-tag II followed by a Thrombin cleavage site (WSHPQFEKGALVPR'GS) and a C-terminal His₆ tag. Inactive endoNF variants were generated by exchanging the 3' 1009 bp endoNF NdeI/XhoI-fragment in pET22b-Strep containing the gene of inactive ΔN-endoNF-R596A/R647A (6) with the corresponding fragment from the respective endoNF binding site mutant.

Protein expression and purification – Proteins were expressed in *E. coli* BL21-Gold(DE3) in the presence of 100 μg/ml Carbenicillin. Bacteria were cultivated in PowerBroth (Athena Enzyme Systems) at 30°C. Expression of endoNF was induced by adding 0.1 mM IPTG at an optical density (A₆₀₀) of 1.5 and bacteria were harvested 6–7 h after induction. For the analysis of soluble and insoluble proteins, bacteria were lysed by sonication and soluble and insoluble fractions were obtained after centrifugation (22,000 × g, 20 min, 4°C). For protein purification soluble fractions of 500 ml cultures were loaded on StrepTactin-Superflow (IBA) and HisTrap HP columns (GE Healthcare Life Sciences) according to the manufacturer's guidelines. To isolate the proteolytically matured catalytic part (ΔN-endoNF*), the minor fraction of non-cleaved protein was removed by passing samples over a HisTrap HP column prior to StrepTactin-chromatography. Buffer exchange was performed on a HiPrep 26/10 Desalting column (Amersham Biosciences) equilibrated with 10 mM sodium phosphate buffer, pH 7.4.

Size exclusion chromatography – Size exclusion chromatography was carried out on a Superdex 200 HR 10/30 column (Amersham Biosciences) equilibrated with 10 mM sodium phosphate buffer, pH 7.4. The column was calibrated using the Gel Filtration Molecular Weight Markers (MW-GF-200) from Sigma.

SDS-PAGE – SDS-PAGE was performed using 1% (w/v) sodium dodecyl sulfate (SDS) in the sample buffer at 85 V and 20°C (Protein standard: Precision Plus (BioRad)). Proteins were incubated in the presence of 1% SDS for 5 minutes at 95°C prior to loading to the

gel. For the detection of SDS-resistant complexes, the incubation step was omitted. Proteins were stained using RotiBlue (Carl Roth GmbH) according to the manufacturer's guidelines. Scanning was performed in Odyssey Infrared imaging system (LI-COR Biosciences, version 2.1.12; scan parameters: 42 μ m resolution, quality: highest, channel: 800 nm, focus offset: 0.5 mm, intensity: 4.0).

Nuclear magnetic resonance – All NMR experiments were performed on a Bruker Avance 600 MHz spectrometer, equipped with a 5 mm TXI probe with triple axis gradients. Deuterated phosphate buffer (10 mM sodium phosphate buffer, pH 7.4, 150 mM NaCl) was used for all NMR experiments. The measurements were performed at 298 K without sample spinning. ^1H NMR spectra were acquired with 32 scans, a 2 s relaxation delay over a spectral width of 6,000 Hz. Solvent suppression of the residual HDO peak was achieved by continuous low power presaturation pulse during the relaxation delay. Data acquisition and processing were performed with XWINNMR software. Chemical shift assignment of polySia DP4 was achieved by COSY, TOCSY, NOESY and HSQC NMR experiments. Digestion of polySia DP4 was monitored by acquiring ^1H NMR spectra in 10 min time intervals over 16 h.

Determination of endosialidase activity using soluble polySia as substrate – The enzymatic activity of purified endoNF variants was determined by means of the thiobarbituric acid assay as described previously (15, 21). Kinetic parameters K_M and k_{cat} values were globally determined in Prism 4.03 (GraphPad Software Inc.).

Coating of micro plates for in-vitro binding and activity assay – Cellstar 96 well plates (Greiner BIO-ONE) were coated with 40 ng of polySia in PBS (10 mM sodium phosphate buffer, pH 7.4, 150 mM NaCl) per well for 1 h. Saturation was performed with 200 μ l/well 1% (w/v) BSA (bovine serum albumin; AppliChem: albumin fraction V) in PBS. Coating and blocking were each followed by three washing steps with PBS to remove unbound material.

Determination of endosialidase activity using surface bound polySia as substrate – Purified active endoNF variants were applied in 40 μ l per well to polySia-coated micro plates in 1:5 and 1:2 dilution steps in 1% BSA/PBS starting with appropriate concentrations, and incubated for 30 minutes at 37°C. 1% BSA/PBS was used as negative control. EndoNF and cleavage products of polySia were removed by three washing steps, each with 200 μ l PBS for 5 minutes at shaking (1,000 rpm). Residual bound polySia was detected incubating 200 ng/well monoclonal antibody 735 (mAb 735 (22); in 1% BSA/PBS) for 1 h at room temperature and shaking (1,000 rpm) followed by three washing steps with PBS, incubation with goat anti-mouse conjugated to horse radish peroxidase (1:2,000 diluted in 1% BSA/PBS), and another three washing steps with PBS. Chromogenic reaction was performed with ABTS (Roche) according to manufacturer's guidelines. Data were analyzed using Prism 4.03 (GraphPad Software Inc.). Curves were best fitted using the model

“Sigmoidal dose-response curve; top to bottom, variable slope” with equation 1:

$$A_{405} = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{1 + 10^{(\log(\text{EC}_{50}) - \log(c(\text{Protein}))) * \text{HillSlope}}}$$

The parameters are defined as A_{405} := Absorbance at 405 nm of the ABTS⁺ radical corresponding to residual bound polySia, *Bottom* := A_{405} value at the bottom plateau, and *Top* := A_{405} value at the top plateau. The calculated EC_{50} value corresponds to the respective endosialidase concentration of half-maximal polySia removal and represents the inflection point of the fitted curve. The *HillSlope* factor describes the steepness of the curve but is of no relevance for the results and therefore not shown.

Analysis of PolySia Cleavage Products in HPAEC-PAD – 80 μ g of polySia oligomers were digested using 1 pmol endosialidase for 30 minutes at 37°C. Enzymatic reaction was stopped adding one volume of 1 M sodium hydroxide to the reaction mixture. Aliquots were analysed by HPAEC-PAD (high performance anion-exchange chromatography with pulsed amperometric detection; DIONEX Bio-LC; software: Chromeleon 6.60) equipped with a 10 μ l injection loop, CarboPac PA-100 2 x 250 mm analytical column and CarboPac PA-100G, 2 x 50 mm guard column system (DIONEX). Elution was carried out using curved gradient 2 in 55 ml at 0.25 ml/min: 0.05 mM to 1 M NaAc in 100 mM NaOH. Charge was measured with the gold working electrode and silver/AgCl reference electrode in the electrochemical detector Bio-LC ED50A using waveform A, according to manufacturer's guidelines. The elution profiles were processed with MS Excel and Prism 4.03 (GraphPad Software Inc.).

In-vitro Binding Assay – Inactive endoNF variants were applied in 40 μ l per well to polySia coated microplates in 1:5 and 1:2 dilution steps in 1% BSA/PBS starting with appropriate concentrations, and incubated for 1 h at room temperature and shaking (1,000 rpm). 1% BSA/PBS was used as negative control. Unbound endoNF was removed in three washing steps each with 200 μ l PBS for 5 minutes at shaking (1,000 rpm). The N-terminal Strep-tag II of EndoNF was detected incubating Strep-Tactin conjugated to horse radish peroxidase (IBA, diluted 1:5,000 in 1% BSA/PBS) for 1 h at room temperature and shaking (1,000 rpm) followed by three PBS washing steps. Color development was performed with ABTS (Roche) according to manufacturer's guidelines. Data were analysed using Prism 4.03 (GraphPad Software Inc.). Curves were best fitted using the model “Sigmoidal dose-response curve; bottom to top, standard slope” with equation 2:

$$A_{405} = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{1 + 10^{\log(K_D) - \log(c(\text{Protein}))}}$$

The parameters are defined as A_{405} := Absorbance at 405 nm of the ABTS⁺ radical corresponding to bound endoNF, *Bottom* := A_{405} value at the bottom plateau, and *Top* := A_{405} value at the top plateau. The calculated K_D value corresponds to the respective endosialidase

concentration of half-maximal binding and represents the inflection point of the fitted curve.

RESULTS

Comparison of the enzymatic activity of mature and unprocessed endoNF – Previous studies revealed that proteolytic processing of endoNF is not a prerequisite to obtain enzymatically active trimers. However, no information is available whether the release of the C-terminal chaperone domain affects the kinetic properties of the enzyme. To address this point, we compared molar activities of mature endoNF lacking the CTD (endoNF*, control) and the non-cleavable mutant endoNF-S911A (c(S)), which mimics the precursor protein of endoNF*. Since the N-terminal capsid binding domain of endoNF is dispensable for trimerization, activity, and thermostability of the enzyme (6, 15, 18), variants lacking the first 245 amino acids (Δ N-endoNF) were used throughout this study (see Fig. 1A for schematic representation and Table 1 for notation of endoNF-variants). Proteins were expressed with an N-terminal Strep-tag II and a C-terminal His₆ tag in *E. coli* BL21 to allow efficient affinity purification (6). Enzymatic activities of purified proteins were analyzed by means of the thiobarbituric acid (TBA) assay (15, 21), which monitors the depolymerization of soluble polySia by measuring the increase of reducing ends of sialic acid. Interestingly, the non-processed c(S)-variant showed a 3-fold higher molar activity compared to the control (Fig. 1C), indicating that the presence of the CTD enhances depolymerization of polySia. However, completely opposing results were obtained when the activity towards surface bound polySia was determined in an ELISA. In this assay, polySia coated to microtiter plates was incubated with serial dilutions of purified endoNF. After removal of the enzyme, remaining substrate was detected with the polySia-specific mAb 735 which binds to α 2,8-linked polySia ≥ 8 residues (22-24). Under these conditions, the c(S)-variant showed a drastically reduced activity (Fig. 1D), and compared to the control, a 100-fold higher enzyme concentration was required to reach half-maximal depolymerization of the substrate (EC_{50} of 678.3 and 6.8 pM for c(S) and control, respectively). The unexpected outcome of the two experiments indicated that the presence of the CTD can either increase or decrease the enzymatic activity of endoNF, depending on how the polymeric substrate is presented.

Sialic acid tetramer is the minimal substrate size required for endoNF activity – The crystal

structure of endoNF revealed that in addition to the active site (site *a*), sialic acid binding sites are located in the lectin-like β -barrel domain (site *b1*) and in the β -prism domain of the stalk (*b2*) (adapted from (6), Fig. 1B). The geometry of the binding sites suggests that the polymeric substrate is wrapped around the trimeric endoNF complex, interacting simultaneously with all three binding sites (6). Accordingly, the CTD preserved in the unprocessed c(S) variant could either affect the conformation of the active site or interfere with polymer binding outside the catalytic center. To discriminate between these possibilities, we determined the kinetic parameters for cleavage of an oligomeric substrate that is too short to reach sites *b1* or *b2* when bound to the active site. Hallenbeck *et al.* reported that endoNF requires a minimum of five sialyl residues for activity (9). To re-investigate the minimal substrate length of endoNF, we monitored digestion of polySia oligomers with a defined degree of polymerization (DP) by ¹H NMR spectroscopy. Whereas no activity towards oligomers with DP3 was detected, cleavage of oligomers with DP4 was clearly detected and results are shown in Fig. 2A. In the ¹H NMR spectrum at *t*=0 min the distinct chemical shift values for the equatorial protons at position C3 (H3eq, see Fig. 2B for the definition of H3eq protons) are clearly visible. The H3eq proton of the reducing end sialic acid residue Sia1 (H3eq (β_e)) resonates at 2.22 ppm due to the predominance of the β -anomer. The ¹H NMR signal at 2.70 ppm corresponds to the H3eq protons of the two internal sialic acid moieties (Sia2 and Sia3 (H3eq (α_i))) whereas the H3eq proton of the non-reducing sialic acid residue (H3 of Sia4 (H3eq (α_e))) assigns a signal at 2.78 ppm. As expected for a tetramer, calculation of the relative integrals obtained at *t*=0 min revealed a ratio of 1:2:1 for H3eq (α_e) : H3eq (α_i) : H3eq (β_e) (Table 2). However, incubation with Δ N-endoNF* clearly shows a drastically changed ratio of the protons H3eq (α_e) : H3eq (α_i) : H3eq (β_e) from 1:2:1 towards 1:1:2 after 1 h leading to the assumption that Δ N-endoNF* has cleaved the sialic acid tetramer. Notably, the obtained ratio of 1:1:2 does not correspond to a cleavage of the polySia DP4 substrate in two identical di-sialic acid residues as a ratio of 1:0:1 for the protons H3eq (α_e) : H3eq (α_i) : H3eq (β_e) would have been expected. Interestingly, a ratio of 1:1:2 has been detected that is only in agreement with cleavage of the polySia DP4 substrate into oligoSia with DP3 and a sialic acid monomer. Since no change in ratio between the protons was observed between 1–16 h, the digest was completed after 1 h.

Processed and unprocessed endoNF show similar kinetic parameters for the cleavage of sialic acid tetramers – Substrate length requirements and cleavage pattern were further investigated for the mature enzyme and the unprocessed c(S)-variant. In a first step, cleavage of sialic acid oligomers with defined DP from 2 to 4 was analyzed in the TBA assay (Fig. 3A). Whereas no cleavage was observed for oligomers of DP2 and DP3, sialic acid tetramers were hydrolyzed by both, processed and unprocessed endoNF. As shown by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD), sialic acid tetramers were cleaved only once, producing trimers and monomers, and no further hydrolysis of the cleavage products was observed (Fig. 3B). This confirms the data obtained in the ¹H NMR studies. Thus, tetramers provide an ideal substrate to measure steady-state kinetics. Notably, both c(S)-variant and control exhibited identical kinetic parameters for the tetrameric substrate ($K_M=0.85$ mM, and $k_{cat}=4.37$ sec⁻¹; Fig. 3C), demonstrating that the conformation of the active site is not affected by the presence of the CTD. Therefore, effects of the CTD on the enzymatic activity of endoNF might be due to interference with polymer binding mediated by the additional binding sites b1 or b2.

Proteolytic maturation and integrity of the b2-site are essential for efficient polySia binding – In the unprocessed full length protein of endoNF the CTD is connected to the C-terminal end of the stalk domain. Due to its proximity to the stalk, it is likely that the CTD affects the sialic acid binding site b2 located within this domain. To investigate whether impaired polymer binding due to interference with binding site b2 is the reason for the observed changes in enzymatic properties of the c(S) mutant, the b2-site was selectively disrupted. For this purpose, we performed single, double, and triple alanine exchanges of the amino acids R837, S848, and Q853, which were identified to be involved in sialic acid binding of the b2-site (6). Altogether, 7 protein variants with mutated b2-site (see Table 1) were generated and expressed as described above. Whereas variants containing the double exchange R837A/Q853A (b2(RQ)) and the triple exchange R837A/S848A/Q853A (b2(RSQ)) were found exclusively in the insoluble fraction (data not shown), all other variants were expressed as soluble, proteolytically processed proteins. To prove whether the introduced exchanges result in decreased polySia binding, we also generated the respective inactive forms. As shown previously, double alanine-substitution of the active site

residues R596 and R647 results in complete inactivation of endoNF without affecting expression, complex formation or maturation (6). Therefore, the R596A/R647A double exchange (here termed a(RR), see Table 1) was introduced into the variants b2(R), b2(S), b2(Q), b2(RS), and b2(SQ). For both, active and inactive b2-variants, the mature enzymes that released the CTD were isolated by affinity chromatography and purified proteins were analyzed by size exclusion chromatography and SDS-PAGE. Since identical results were obtained for active and inactive variants, results are shown exclusively for the inactive proteins (Table 3 and Fig. 4). Notably, all variants were found as homo-trimers, clearly demonstrating that the introduced amino acid exchanges did not affect trimer formation. However, further studies revealed that almost all mutations affected trimer stability in the presence of SDS. In our previous studies, we showed that wild-type endoNF, Δ N-endoNF lacking the N-terminal capsid binding domain, and the unprocessed c(S)-variant form SDS-resistant complexes which can be visualized by SDS-PAGE, when the boiling step is omitted (6, 15). As shown in Fig. 4, the presence of SDS-resistant trimers is indicated by appearance of a high molecular mass band above 250 kDa and disappearance of the band that corresponds to monomeric endoNF. Interestingly, all variants with mutated b2-site lost their ability to form SDS-resistant protein complexes, with the single exception of b2(S). Thus, alanine substitution of either Arg-837 or Gln-853 did not affect trimer formation but abrogated SDS-resistance, indicating conformational changes in the stalk domain.

In a next step, all inactive a(RR)-forms of the generated b2-mutants as well as a respective a(RR)-form of the control and the c(S)-variant were analyzed for binding to surface bound polySia in an ELISA-based binding assay (Fig. 5). For the a(RR)-form of the control, a dissociation constant (K_D) of 1.9 nM was determined, demonstrating that endoNF binds with high affinity to its polymeric substrate. By contrast, drastically reduced binding with a 190-fold higher K_D value was observed for the unprocessed c(S)-variant, demonstrating that the CTD strongly interferes with binding to the polymer. The b2-mutants showed a clear tendency in decreased binding to polySia towards the c(S)-mutant indicating that this binding site is affected in the c(S)-mutant.

Interference with polySia binding by disrupting the b2-binding site increases endoNF activity towards soluble substrate – To investigate,

whether interference with the b2-site increases enzymatic activity towards soluble polySia as observed for the c(S)-variant, the relative molar activities of the b2-mutants were determined by the TBA assay (Fig. 6). The enzymatic activities of the SDS-resistant b2(S), as well as b2(Q), and b2(SQ) were found to be within the control range. In contrast, variants containing the amino acid exchange Arg837Ala (b2(R) and b2(RS)) showed a significant increase in molar activity, pointing out a trend towards the c(S)-mutant (cf. Fig. 1C) and this emphasizes that Arg-837 seems to be the most critical residue in the b2-site for binding to sialic acid.

In summary, our findings clearly corroborate the hypothesis that the presence of the CTD in the non-cleavable mutant c(S) interferes with the b2-site and that proteolytic maturation is a prerequisite to gain the sialic acid binding site.

DISCUSSION

Proteolytic maturation is a common feature among many phage proteins including endosialidases (15, 25-29). In the endosialidases a short C-terminal chaperone domain is released from the mature protein (15). This cleavage was recently described also for other tailspike and tail fiber proteins with a homologous CTD. Since the presence of the CTD decreases the thermal stability of the homotrimeric complex, as shown for the non-cleavable mutant c(S) (Δ N-endoNF-S911A), we recently suggested that proteolytic release of this general endogenous chaperone domain is required to establish kinetically stable complexes (18).

However, in this study, we clearly demonstrated that the presence of the CTD drastically reduces binding of endoNF to the substrate polySia, as revealed for the inactive and non-cleavable mutant a(RR)/c(S). Furthermore, the binding deficiency is reflected in a reduced enzymatic activity towards surface bound polySia, although the active site is not affected in c(S). Therefore, efficient binding of endoNF to polySia is crucial for proper degradation of the substrate. It has been suggested that the function of tailspike proteins of phages infecting encapsulated bacteria is not limited to degradation of the respective bacterial antigen but they also anchor the phage to the host surface. Despite structural differences, these tail spike proteins share common features like a specific depolymerase activity in combination with substrate binding subsites in a substrate-binding groove (30-33), or a set of binding sites as shown for the endosialidases (6). These additional sites fulfill the function of accelerating enzyme-

substrate binding, migration along the chain and finally attaching the phage to the bacterial membrane using remnant chain fragments of the polymer. Two sialic acid binding sites per subunit were identified in the endoNF crystal structure (cf. Fig. 1B, (6)). Since the presence of the CTD influences the binding properties of endoNF to polySia, we analyzed the binding site b2 in the stalk domain by site-directed mutagenesis due to the proximity to the C-terminus. The a(RR)/b2-mutants investigated in this study showed a clear tendency in binding deficiency towards the effect observed for c(S). Since binding was most influenced in a(RR)/b2-mutants containing the mutation R837A, Arg-837 seems to be the most important amino acid in this binding site. Taking into account that the a(RR)/b2-mutants did not completely resemble the binding deficiency obtained for the a(RR)/c(S)-mutant, we assume that there are more subsites involved in polySia binding within the stalk domain and solely the binding sites with higher affinities were identified as yet in the crystal structure. As a consequence, it would be possible that also oligomers are bound in the stalk domain. This hypothesis is supported by NMR-studies revealing that all five sialic acid moieties of DP5 interact with endoNF-R596A (34).

Depending on the accessibility of the substrate, the binding deficiency provokes an advantage or a disadvantage for endoNF-c(S) in digesting polySia. When surface bound polySia is used as substrate, the reduced binding of the c(s)-mutant is reflected in a decreased activity. Thus, association of the enzyme to the fixed substrate is an essential step for the depolymerase reaction. In contrast, using soluble polySia as substrate the enzymatic activity of the c(S)-mutant was found to be increased compared to the control. In this case, the dissociation of the polySia cleavage products from the c(S)-mutant should be facilitated due to the abolished b2-site. Since the substrate polySia is used in excess, the increased dissociation rate effect an advantage for the c(S)-mutant.

Interestingly, similar results of activities that depend on substrate-accessibility were reported for several cellulases and chitinases. These depolymerases also contain polymer binding subsites like capsule-degrading tail spike proteins of bacteriophages, arranged along a groove. Acting on the insoluble materials chitin and cellulose, these additional binding sites gain access to single polymer chains of the solid material for instantaneous degradation. Mutations of critical aromatic amino acid residues in binding subsites drastically decrease enzymatic activities on recalcitrant or crystalline substrate, while

increased or not affected activities were observed towards soluble substrates (35-38). Horn and coworkers recently elucidated that the differences in activities on differently accessible substrate of the chitinase B (ChiB) of *Serratia marcescens* are also reflected in a changed processivity of the enzyme (39). Furthermore, they underlined that loss of processivity is a disadvantage for the degradation of recalcitrant substrate but an advantage for the activity on soluble substrate. This could also turn out for endosialidases and other depolymerases that unify degradation and additional adhesion capabilities. Bacteriophages that infect encapsulated bacteria create a narrow path through the capsule towards the cell surface (40), indicating that depolymerization is processive. Recently, processivity of the coliphages K1E and K1-5 was suggested to be a result of the interplay of the six endosialidase molecules attached to the base plate (7). With the data presented in this study, the non-cleavable c(S)-mutant with the preserved CTD clearly reproduces the phenotype described for the chitinase and cellulase mutants, what could argue for endoNF exhibiting processivity itself that might be reduced in the c(S)-mutant. Inferentially, this could be a general feature for bacteriophage-derived depolymerases. Furthermore, it is possible that the trimeric endoNF can interact with three polysialic acid chains simultaneously and that processive migration along the chains is managed by independent and staggered association-cleavage-dissociation processes of the three subunits. Interfering with one binding site could influence these processes and thus probably decrease the processivity resulting in the effects observed for the c(S)-mutant.

In summary, we suggest that the preserved CTD could influence the binding site b2 and the subsites suggested in this study by two possible folding models as schematically depicted in Figure 7B and C. The correctly folded and processed control (Δ N-endoNF*) is depicted in Figure 7A. As shown in Fig. 7B for the c(S)-mutant, the presence of the CTD could lead to an improperly folded stalk domain. This would lead to incorrectly folded b2 and subsites unable to bind polySia in the stalk. Cleavage of the C-terminal domain could effort the folding enthalpy necessary to gain the proper folded stalk domain. Another model in Figure 7C suggests the CTD to be wrapped around the stalk domain and hide b2 and the subsites. We recently showed that highly conserved residues within the

CTD are critical for proper folding of endoNF (18). Assuming the model in Figure 7C, it could be possible that these critical amino acids in the CTD interact with certain amino acids in the stalk domain during the folding process. This interaction is terminated by cleavage and dissociation of the C-terminal domain from the mature tailspike protein. In the c(S)-mutant these interactions could be stable enough that the CTD hides the sialic acid binding sites in the stalk. In line with this model one could explain the abrogated SDS-resistance of the b2-mutants as depicted in Figure 4. The amino acids Arg-837 and Gln-853 might be interaction partners of certain amino acids of the CTD. Single mutations of these two residues result in SDS-sensitive protein complexes because the stalk domain might be poorly formed due to a loss of interactions with the CTD. Notably, the double mutation b2(RQ) and similarly the triple mutation b2(RSQ) in the sialic acid binding site resulted in exclusively insoluble proteins. Additionally, the CTD of tailspike proteins is crucially adapted to the respective N-terminal part by sequence and length. Exchange of the CTDs of endoNF (153 amino acids) and endoNE (105 amino acids) resulted in the processed chimera endoNE-NF with a slightly reduced activity. On the other hand, the processed but enzymatically inactive chimera endoNF-NE was obtained. The CTD of endoNE presumably does not achieve the required length for the N-terminal part of endoNF (15). The chimera Δ N-endoNF-gp12, in which the CTD of endoNF is substituted by the CTD of the otherwise unrelated tailspike protein gp12 of *Bacillus* phage GA-1, was processed into Δ N-endoNF* and found as active enzyme. However, compared with wild-type trimer, a slightly decreased SDS-resistance was observed for the chimera-derived complexes (18). Finally, the attempt to solve the crystal structure of the CTD of endoNF or of the non-cleavable c(S)-mutant was not yet accomplished. In the electron density obtained from the c(S)-mutant we were able to rebuild the complete N-terminal portion of Δ N-endoNF* that was revealed previously (6). However, the C-terminal domain could not be build into the density of the c(S)-mutant, but electron density located at the C-terminus and also surrounding the stalk domain of endoNF correspond to about one third of the C-terminal domain (data not shown). Taken together, these findings support the model shown in Figure 7C.

FOOTNOTES

The abbreviations used are: aa, amino acids; BSA, bovine serum albumin; CTD, C-terminal domain; DP, degree of polymerization; Δ N-endoNF, endo-N-acylneuraminidase or endosialidase (E.C. 3.2.1.129) of coliphage K1F lacking the N-terminal 245 amino acids of the capsid binding domain (6, 18); Δ N-endoNF*, catalytic part (aa 246-911) proteolytically released from Δ N-endoNF (18); polySia, polysialic acid; TBA, thiobarbituric acid. Further abbreviations are listed in Table 1.

Andrea Bethe and Melanie Grove are acknowledged for excellent technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of DFG Research Unit 548 (Ge801/7-1 and 7-2).

REFERENCES

1. Robbins, J. B., McCracken, G. H., Jr., Gotschlich, E. C., Orskov, F., Orskov, I., and Hanson, L. A. (1974) *New England Journal of Medicine* **290**, 1216-1220
2. Sarff, L. D., McCracken, G. H., Schiffer, M. S., Glode, M. P., Robbins, J. B., Orskov, I., and Orskov, F. (1975) *Lancet* **1**, 1099-1104
3. Scholl, D., Adhya, S., and Merrill, C. (2005) *Appl. Environ. Microbiol.* **71**, 4872-4874
4. Bayer, M. E., Thurow, H., and Bayer, M. H. (1979) *Virology* **94**, 95-118
5. Pelkonen, S., Aalto, J., and Finne, J. (1992) *J Bacteriol.* **174**, 7757-7761
6. Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., and Ficner, R. (2005) *Nature structural & molecular biology* **12**, 90-96
7. Leiman, P. G., Battisti, A. J., Bowman, V. D., Stummeyer, K., Mühlenhoff, M., Gerardy-Schahn, R., Scholl, D., and Molineux, I. J. (2007) *J. Mol. Biol.*
8. Kleene, R. and Schachner, M. (2004) *Nat. Rev. Neurosci.* **5**, 195-208
9. Hallenbeck, P. C., Vimr, E. R., Yu, F., Bassler, B., and Troy, F. A. (1987) *J Biol. Chem.* **262**, 3553-3561
10. Pelkonen, S., Pelkonen, J., and Finne, J. (1989) *J Virol.* **63**, 4409-4416
11. Long, G. S., Bryant, J. M., Taylor, P. W., and Luzio, J. P. (1995) *Biochem J* **309**, 543-550
12. Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S., Lottspeich, F., and Frosch, M. (1995) *Molecular Microbiology* **16**, 441-450
13. Machida, Y., Hattori, K., Miyake, K., Kawase, Y., Kawase, M., and Iijima, S. (2000) *Journal of bioscience and bioengineering* **90**, 62-68
14. Scholl, D., Rogers, S., Adhya, S., and Merrill, C. R. (2001) *Journal of Virology* **75**, 2509-2515
15. Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M., and Gerardy-Schahn, R. (2003) *J. Biol. Chem.* **278**, 12634-12644
16. Jakobsson, E., Jokilampi, A., Aalto, J., Ollikka, P., Lehtonen, J. V., Hirvonen, H., and Finne, J. (2007) *Biochem. J.* **405**, 465-472
17. Stummeyer, K., Schwarzer, D., Claus, H., Vogel, U., Gerardy-Schahn, R., and Mühlenhoff, M. (2006) *Molecular Microbiology* **60**, 1123-1135
18. Schwarzer, D., Stummeyer, K., Gerardy-Schahn, R., and Mühlenhoff, M. (2007) *J. Biol. Chem.* **282**, 2821-2831
19. Weigele, P. R., Scanlon, E., and King, J. (2003) *Journal of bacteriology* **185**, 4022-4030
20. Decher, G., Ringsdorf, H., Venzmer, J., Bitter-Suermann, D., and Weisgerber, C. (1990) *Biochim. Biophys. Acta* **1023**, 357-364
21. Skoza, L. and Mohos, S. (1976) *Biochem J* **159**, 457-462
22. Frosch, M., Gorgen, I., Boulnois, G. J., Timmis, K. N., and Bitter-Suermann, D. (1985) *Proc. Natl. Acad. Sci. U.S.A* **82**, 1194-1198
23. Frosch, M., Roberts, I., Gorgen, I., Metzger, S., Boulnois, G. J., and Bitter-Suermann, D. (1987) *Microb. Pathog.* **2**, 319-326
24. Häyrynen, J., Haseley, S., Talaga, P., Mühlenhoff, M., Finne, J., and Vliegthart, J. F. (2002) *Mol. Immunol.* **39**, 399-411
25. Marvik, O. J., Jacobsen, E., Dokland, T., and Lindqvist, B. H. (1994) *Virology* **205**, 51-65
26. Conway, J. F., Duda, R. L., Cheng, N., Hendrix, R. W., and Steven, A. C. (1995) *Journal of Molecular Biology* **253**, 86-99
27. 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
28. Wang, S., Chandramouli, P., Butcher, S., and Dokland, T. (2003) *Virology* **314**, 1-8
29. Laemmli, U. K. (1970) *Nature* **227**, 680-685
30. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) *Science* **265**, 383-386

31. Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., and Huber, R. (1996) *PNAS* **93**, 10584-10588
32. Freiberg, A., Morona, R., van den Bosch, L., Jung, C., Behlke, J., Carlin, N., Seckler, R., and Baxa, U. (2003) *J Biol. Chem.* **278**, 1542-1548
33. Smith, N. L., Taylor, E. J., Lindsay, A. M., Charnock, S. J., Turkenburg, J. P., Dodson, E. J., Davies, G. J., and Black, G. W. (2005) *Proc.Natl.Acad.Sci.U.S.A* **102**, 17652-17657
34. Haselhorst, T., Stummeyer, K., Mühlenhoff, M., Schaper, W., Gerardy-Schahn, R., and von Itzstein, M. (2006) *ChemBiochem.* **7**, 1875-1877
35. Koivula, A., Kinnari, T., Harjunpaa, V., Ruohonen, L., Teleman, A., Drakenberg, T., Rouvinen, J., Jones, T. A., and Teeri, T. T. (1998) *FEBS Lett.* **429**, 341-346
36. Zhang, S., Irwin, D. C., and Wilson, D. B. (2000) *Eur.J.Biochem.* **267**, 3101-3115
37. Watanabe, T., Ariga, Y., Sato, U., Toratani, T., Hashimoto, M., Nikaidou, N., Kezuka, Y., Nonaka, T., and Sugiyama, J. (2003) *Biochem.J.* **376**, 237-244
38. Katouno, F., Taguchi, M., Sakurai, K., Uchiyama, T., Nikaidou, N., Nonaka, T., Sugiyama, J., and Watanabe, T. (2004) *J.Biochem.(Tokyo)* **136**, 163-168
39. Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sorlie, M., Synstad, B., Vriend, G., Varum, K. M., and Eijsink, V. G. H. (2006) *PNAS* **103**, 18089-18094
40. Sutherland, I. W. (1977) Enzymes Acting on Bacterial Surface Carbohydrates. In Sutherland, I. W., editor. *Surface Carbohydrates of the Prokaryotic Cell*, Academic Press, Inc., New York, USA

FIGURES

Figure 1

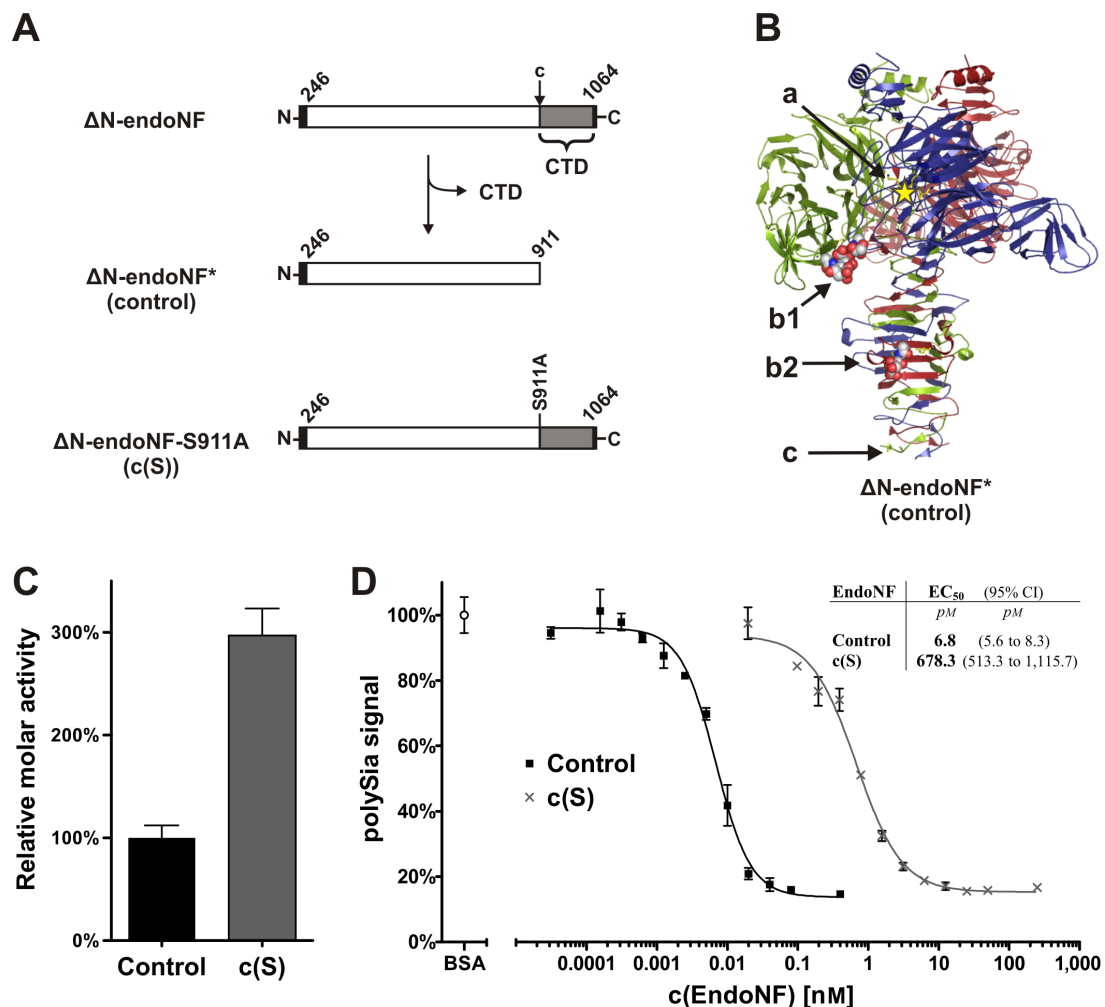


FIGURE 1. PolySia interaction sites and proteolytic cleavage site of endoNF, and comparison of endoNF activities in two different activity assays. A, schematic representation of endoNF variants lacking the N-terminal capsid binding domain (ΔN): the endosialidase F precursor (ΔN -endoNF), the processed catalytic part ΔN -endoNF* (control) and the non-cleavable mutant ΔN -endoNF-S911A (c(S)). The catalytic part is shown in white, the C-terminal chaperone domain (CTD), which is released from the mature protein, in grey. The N-terminal Strep-tag II and C-terminal His₆ tag are given in black. The first and the last amino acid number of ΔN -endoNF are depicted diagonally. The amino acid substitution S911A is written vertically. The proteolytic cleavage site Ser-911, as identified previously, is indicated as c.

B, ribbon diagram of ΔN -endoNF* (adapted from Ref. 6). The three monomers are red, green, and blue. The active site (a) of the blue subunit is schematically depicted with a yellow asterisk. Spheres represent di-sialic acid and sialic acid bound in the binding sites b1 in the beta-barrel domain of the green subunit and b2 in the beta-prism domain of the red subunit, respectively; amino acids involved in enzymatic activity or polySia binding as determined in Ref. 6 are depicted as yellow sticks. The proteolytic cleavage site is indicated as c.

C, relative molar activities of ΔN -endoNF* (Control) and ΔN -endoNF-S911A (c(S)) were determined by the thiobarbituric acid assay. Data are means \pm S.D. of two independent experiments performed in duplicates. Control was set to 100%.

D, relative amount of residual surface bound polySia for determining enzymatic activity of ΔN -endoNF* (control) and ΔN -endoNF-S911A (c(S)). BSA (mean of four values) was used as negative control to monitor maximal polySia bound to the surface and was set to 100%. The ELISA based activity assay was performed as described in the experimental procedures in two independent experiments performed in duplicates. Error bars are representing S.D. Curves were best fitted in Prism 4.03 using equation 1. The concentration (x-axis) is scaled logarithmically. Inset table: EC₅₀ values determined in the ELISA based assay of the respective endoNF variant are giving the endoNF concentration of half maximal polySia removal. Note that the 95% confidence interval (95% CI) is asymmetric due to the determination of $\log(\text{EC}_{50})$.

Figure 2

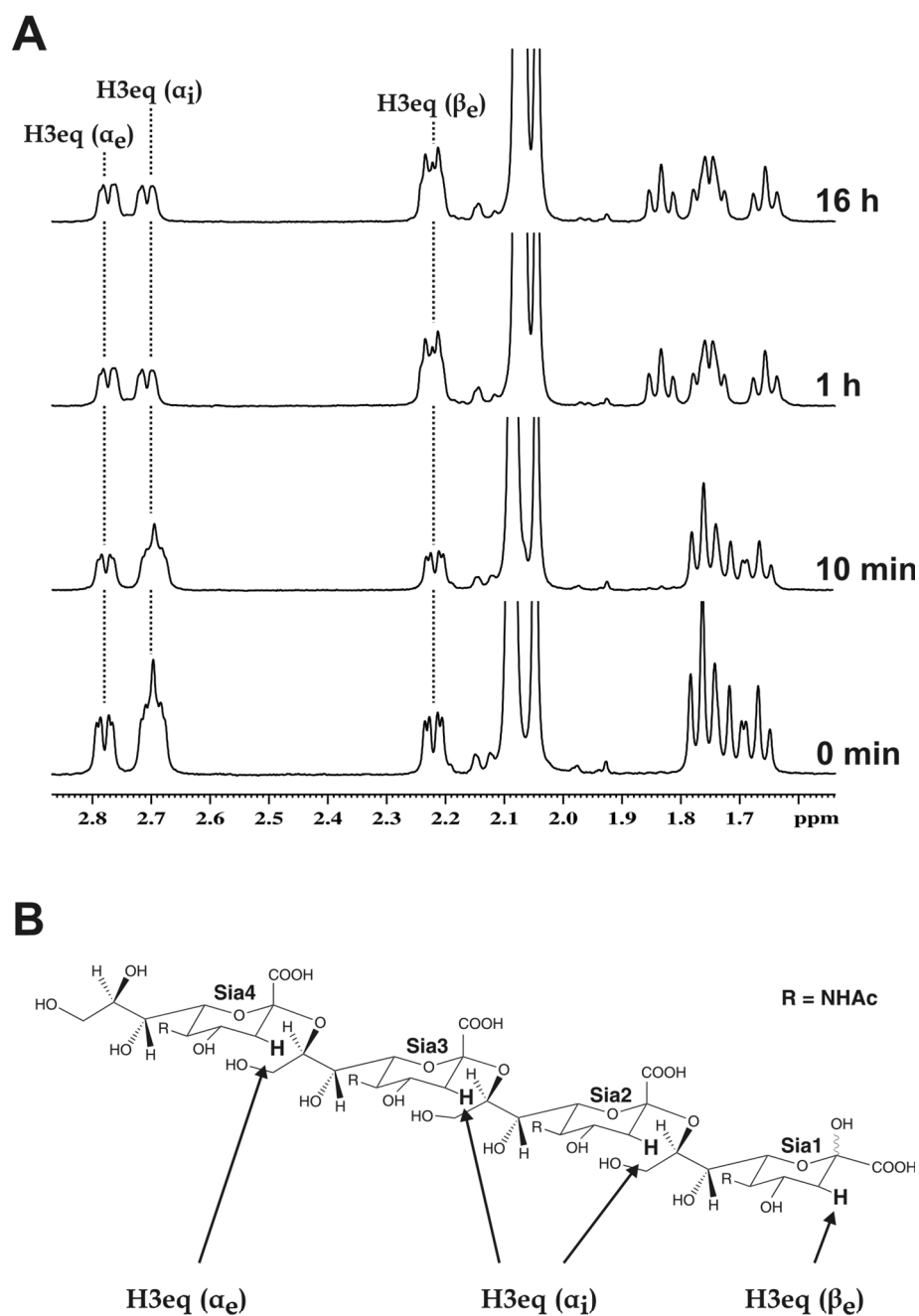


FIGURE 2. **Cleavage of sialic acid tetramer (DP4) by ΔN -endoNF* (Control) monitored by ^1H NMR.** A, ^1H NMR spectra of DP4 sialic acid incubated with ΔN -endoNF* recorded at different time intervals: from bottom to top: 0 min, 10 min, 1 h and 16 h. All NMR spectra were acquired in deuterated 10 mM phosphate buffer, 150 mM NaCl, pH 7.4 in D_2O at 298 K. A continuous low power presaturation pulse during the relaxation delay was used for water suppression. B, structure of DP4. The equatorial protons at position C3 of DP4 are additionally depicted in *bold* and indicated below.

Figure 3

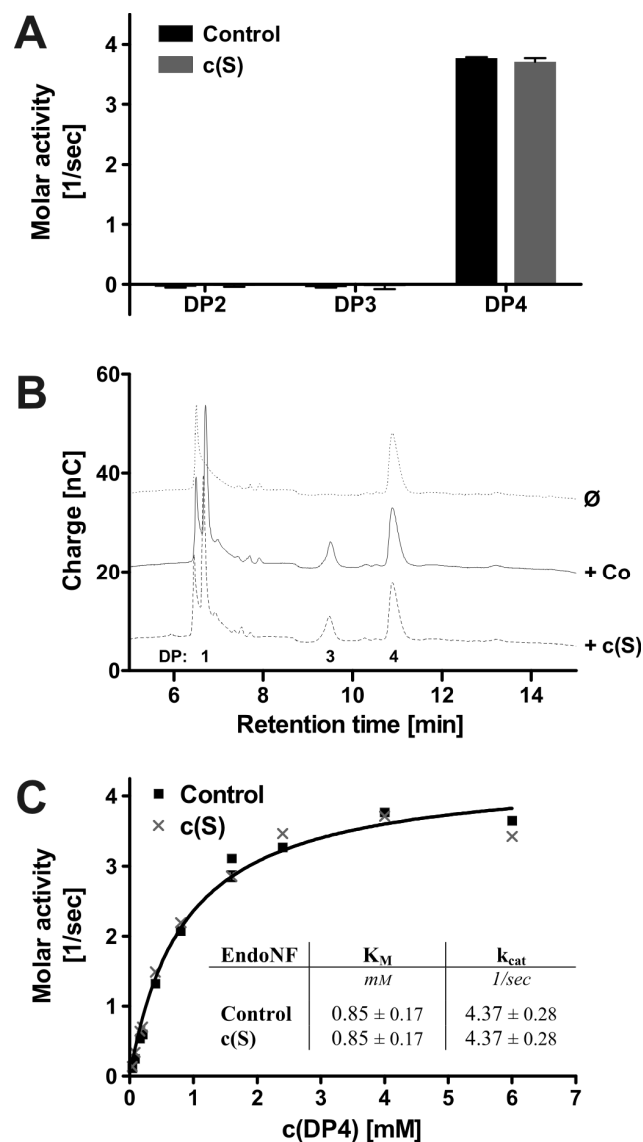


FIGURE 3. Cleavage of sialic acid oligomers by ΔN -endoNF*(Control) and ΔN -endoNF-S911A (c(S)). **A**, molar activity on sialic acid oligomers DP2–4 was determined by the thiobarbituric acid assay in two independent experiments. The molar activity is the amount of released reducing ends of sialic acid in mol per mol enzyme and sec (1/sec). **B**, detection of digested sialic acid DP4 in HPAEC-PAD analysis. Enzymatic digest and chromatography was performed with endoNF variants (+ Co, or + c(S), respectively) or without enzyme (\emptyset). The elution profiles are separated by a y-axis offset of 15 nC. Identified peaks are labeled below (DP1 to DP4). **C**, kinetic parameters of endoNF variants on the minimal substrate DP4 were determined in the thiobarbituric acid assay performed in two independent experiments and analysed in Prism. *Inset table* in C, kinetic parameters determined for both endoNF variants with the 95% confidence interval. The curves were best fitted globally with shared parameters.

Figure 4

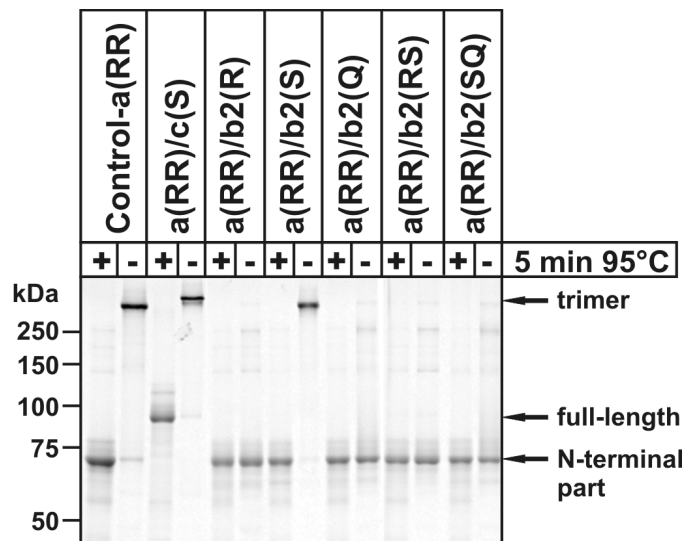


FIGURE 4. **Complex formation of enzymatically inactive Δ N-endoNF-a(RR) variants.** Affinity purified proteins were analysed in 7% SDS-PAGE. To determine SDS-resistance (*trimer*) of the respective variant, the initial boiling step (+) in the presence of 1% (w/v) SDS was omitted (-). *Full-length* protein represents the monomer of the non-cleavable mutant c(S), whereas all other mutants are proteolytically processed (monomeric *catalytic part*). Numbers on the left side represent protein standards.

Figure 5

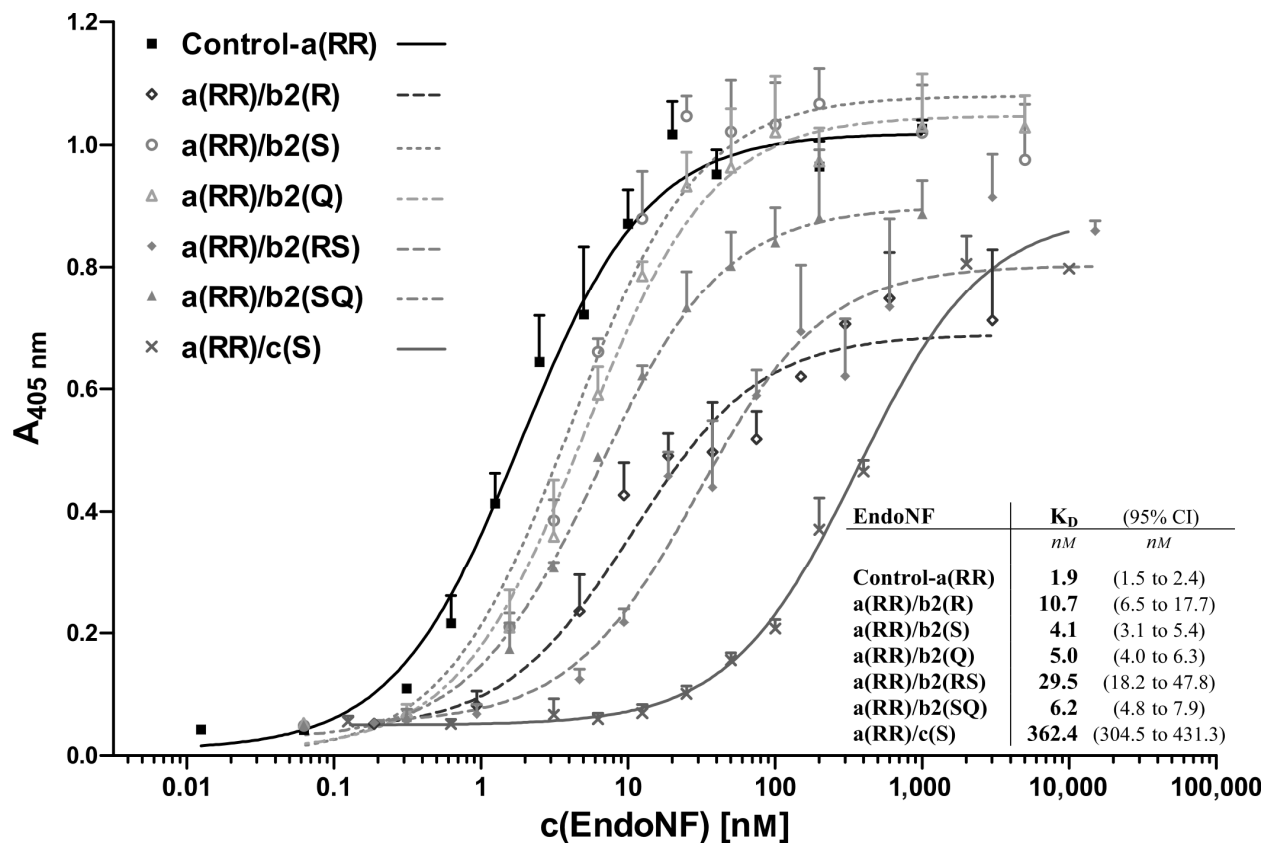


FIGURE 5. Binding of enzymatically inactive endoNF-a(RR) variants to surface bound polySia. The binding assay was performed as described in the experimental procedures in four independent experiments ($n=4$). Curves were best fitted in Prism 4.03 using equation 2. The concentration (x-axis) is scaled logarithmically. Error bars above the curves are representing S.D. *Inset table:* K_D values determined in the in-vitro binding assay of the respective inactive endoNF variant are giving the protein concentration of half maximal binding. Note that the 95% confidence interval (95% CI) is asymmetric due to the determination of $\log(K_D)$.

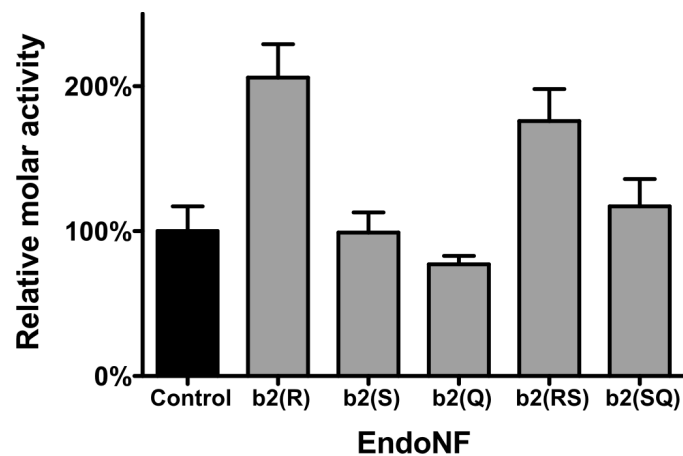
Figure 6

FIGURE 6. Relative molar activities of Δ N-endoNF variants with single and double amino acid exchanges in the sialic acid binding site b2 of the stalk region. Activities were determined by the thiobarbituric acid assay using soluble polySia as substrate. Data are means \pm S.D. of two independent experiments performed in duplicates. The molar activity of control was set to 100%.

Figure 7

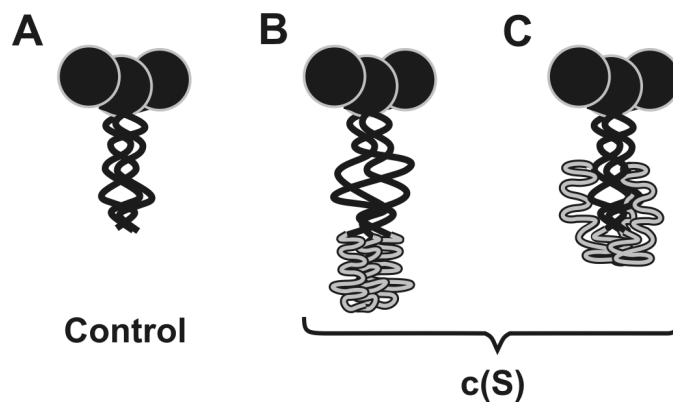


FIGURE 7. **Two proposed models of structural influence of the C-terminal domain on the b2-binding site.** Schematic representations of Δ N-endoNF* (Control, A), and Δ N-endoNF-S911A (c(S), B and C), respectively. The β -propeller is shown as *black circle*, the stalk domain as *black curve*, and the CTD is shown in *grey*. A, Δ N-endoNF* (*control*, adapted from (18)) with a proper folded stalk domain and an intact b2-binding site. B and C, two possible folding models of the C-terminal domain (*grey*) in the S911A mutant (c(S)): the preserved C-terminal domain induces a conformational change in the stalk domain, which abolishes the b2-binding site (B), or the C-terminal domain interacts with regions of the stalk domain and hides the b2-binding site (C).

TABLES

TABLE 1

Notation of EndoNF variants

The Δ N-endoNF variants listed below were analysed in this study. A convenient notation is introduced to facilitate reading (*left column*), with *a*, active site (*lower panel*), *b2*, binding site in the β -prism domain, and *c(S)* proteolytic cleavage site at Ser-911 (according to fig. 1B). The respective amino acids exchanged into alanine (see right column) are given in brackets. All variants contain an N-terminal Strep-tag II. The non-cleavable mutants c(S) and a(RR)/c(S) additionally contain a C-terminal His₆ tag.

Protein	amino acid substitutions
Control	-
b2(R)	R837A
b2(S)	S848A
b2(Q)	Q853A
b2(RS)	R837A/S848A
b2(RQ)	R837A/Q853A
b2(SQ)	S848A/Q853A
b2(RSQ)	R837A/S848A/Q853A
c(S)	S911A
Control-a(RR)	R596A/R647A
a(RR)/b2(R)	R596A/R647A/R837A
a(RR)/b2(S)	R596A/R647A/S848A
a(RR)/b2(Q)	R596A/R647A/Q853A
a(RR)/b2(RS)	R596A/R647A/R837A/S848A
a(RR)/b2(SQ)	R596A/R647A/S848A/Q853A
a(RR)/c(S)	R596A/R647A/S911A

TABLE 2

Relative integrals of the equatorial protons at position C3 of sialic acid moieties in different oligomers. Measured relative integrals of H3eq of DP4 before (A) and after (B) 16 h incubation with Δ N-endoNF*. Calculated relative integrals of H3eq that potentially occur by an endosialidase digest of DP4: symmetric cleavage into two DP2 moieties (C) or asymmetric cleavage into one DP3 and one DP1 moiety (D), respectively.

	Sialic acid oligomers	H3eq (α_e)	H3eq (α_f)	H3eq (β_e)
A	DP4	1	2	1
B	DP4 + Δ N-endoNF*	1	1	2
C	2 DP2	1	0	1
D	1 DP3 + 1 DP1	1	1	2

TABLE 3

Quaternary structure of endoNF. The oligomeric state of the indicated endoNF variants was determined for affinity-purified proteins by size-exclusion chromatography. 480 μ g in 300 μ l of each protein were loaded to the column. The oligomeric state was calculated from the values obtained by size-exclusion chromatography divided by the mass calculated for the monomer from the amino acid sequences including the N-terminal Strep-tag II (76.1 kDa).

EndoNF variant	Molecular mass (size-exclusion chromatography)	Oligomeric state
	<i>kDa</i>	
Control-a(RR)	225.7	3.0
a(RR)/b2(R)	256.8	3.4
a(RR)/b2(S)	252.6	3.3
a(RR)/b2(Q)	239.5	3.1
a(RR)/b2(RS)	245.3	3.2
a(RR)/b2(SQ)	231.2	3.0

Chapter 7 – Synthesis and Biological Evaluation of a New Polysialic Acid Hydrogel as Enzymatically Degradable Scaffold Material for Tissue Engineering

This manuscript was submitted on March, 31st 2008 to *Biomacromolecules* in an updated version including additional data not included in this thesis version:

'Synthesis and Biological Evaluation of a Polysialic Acid based Hydrogel as Enzymatically Degradable Scaffold Material for Tissue Engineering'

Silke Berski¹, Jeroen van Bergeijk², **David Schwarzer**³, Yvonne Stark⁴, Cornelia Kasper⁴, Thomas Scheper⁴, Claudia Grothe², Rita Gerardy-Schahn³, Andreas Kirschning¹, and Gerald Dräger^{1*}

¹ *Institut für Organische Chemie und Zentrum für Biomolekulare Wirkstoffchemie (BMWZ), Gottfried Wilhelm Leibniz Universität Hannover, Schneiderberg 1B, 30167 Hannover, Germany*

*To whom correspondence should be addressed: Tel: (+49)511 762 4619; Fax: (+49)511 762 3011;
E-mail: draeger@oci.uni-hannover.de

² *Institute of Neuroanatomy, Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany*

³ *Department of Cellular Chemistry, Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany*

⁴ *Institut für Technische Chemie, Gottfried Wilhelm Leibniz Universität Hannover, Callinstrasse 3, 30167 Hannover, Germany*

Manuscript ID: bm-2008-00327s

KEYWORDS: cross-linking; degradation; hydrogel; nerve tissue engineering; polysialic acid.

Preface – About the Manuscript

In this study, endoNF was used in application based assays. The aim of this study was to elucidate whether polySia can be used as a biodegradable scaffold material. Since polySia is a highly hydrated and water soluble polymer, novel insoluble hydrogels were generated at the *Institut für Organische Chemie* of the *Leibniz Universität Hannover* by chemical cross-linking of polysialic acid. My contribution to this study was to establish an *in vitro* testing system in which it was investigated whether the new materials can be degraded by endoNF in an induced and controlled manner.

ABSTRACT

Restorative medicine has a constant need for improved scaffold materials. Degradable biopolymers often suffer from uncontrolled chemical or enzymatic hydrolysis by the host. The need for a second surgery on the other hand is a major drawback for non-degradable scaffold materials. In this paper we report the design and synthesis a novel polysialic acid hydrogel with auspicious properties. Hydrogel synthesis was optimized and enzymatic degradation was studied using a phage-born endosialidase. After addition of endosialidase hydrogel readily degraded depending on the amount of initially used cross-linker within 2 to 11 days. Polysialic acid hydrogel is fully biocompatible, completely stable under physiological conditions and could be evaluated as growth support for PC12 cells. Here, additional coating with collagen I, PLL or matrigel is mandatory to improve the properties of the novel material.

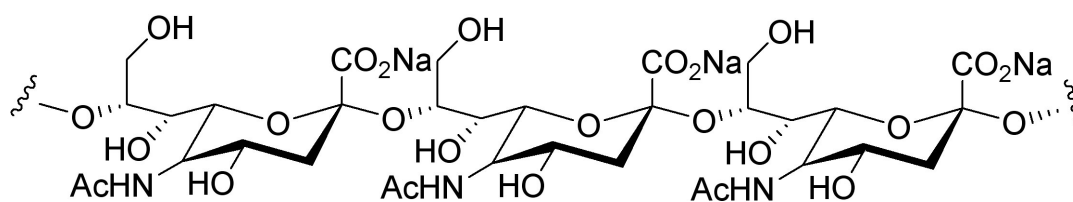
INTRODUCTION

Transplantation as a restorative therapy to replace severely injured, largely lost or dysfunctional organs or tissue is limited by the low availability of biological substitutes [1]. A great variety of polymer materials have been studied so far aiming the reconstruction of organs and tissue, e.g. skin, liver and nervous tissue [2]. After transection injury of a nerve and substantial tissue loss, for example, a guidance channel consisting of natural (e.g. collagen) or synthetic polymers (e.g. silicone [3]) is implanted to facilitate nerve regeneration. These grafts are either filled with components of the extracellular matrix [4] or additionally seeded with cells (e.g. physiological or genetically modified Schwann cells) prior to implantation which enhances growth of axons into the desired direction [5, 6]. It was demonstrated that the tube material and its pore size has a strong influence on the progression of nerve regeneration [7].

A critical point in every *de novo* creation of biological structures is the scaffold material itself, because it has to fulfil a variety of requirements. The main elements are high porosity of the artificial, three-dimensional matrix, biocompatibility, controlled biodegradability and importantly the encouragement of cell adhesion, proliferation and differentiation [8].

In the last decades natural polysaccharides have emerged as important biomaterials. The diverse number of functional groups (hydroxyl groups of monomers, amino groups as found in chitosan and carboxylate groups in hyaluronic acid and alginate) allows pursuing many routes for structural modifications. The opportunity to carry out chemical modifications is essential for optimizing solubility, porosity and mechanical properties of the material. Another advantage of polysaccharides as biomaterials is their biocompatibility because monomeric and oligomeric saccharides are part of cellular structures within the body and cause a markedly weak immunological response even after being modified. One of the best studied examples is hyaluronic acid. However, this polysaccharide cannot be used as a biomaterial in its native form because it is highly water soluble and is rapidly degraded within the body [9].

In this paper we report on the chemical modification and biological evaluation of polysialic acid (polySia, Scheme 1) as a new scaffold material for tissue engineering. PolySia is a negatively charged polysaccharide consisting of α 2,8-linked sialic acid units. It is a dynamically added posttranslational modification of the neural cell adhesion molecule (NCAM) and serves as a modulator of cell-cell and cell-substrate interactions in the processes of neuronal development [10]. Consequently, polySia is abundantly expressed in the embryonic nervous tissue [11] and ablation of its biosynthesis in mice generates a lethal phenotype [12]. Moreover and important in the context of this study, aimed at designing biocompatible nerve growth supports, upregulation of polySia in the adult brain has recently been demonstrated to promote regenerative processes [13, 14].



SCHEME 1. Structural formula of α 2,8-polysialic acid.

While hyaluronic acid is easily removed by the endogenous enzyme hyaluronidase, no polySia degradation mechanisms in mammals are known so far. But highly specific degradation of polySia can be induced by phage-born endosialidases (endo-N-acyl-neuraminidase, endoN) [15–18]. And most important for the use of polySia as biomaterial in regenerative therapy is the fact, that even degradation products of polySia are not toxic as recently evaluated by *in vitro* studies of glial and neuronal cells [19].

MATERIALS AND METHODS

Preparation of polysialic acid hydrogel

Colominic acid (obtained from Nacalai Tesque, Inc., Japan) was dialysed against distilled water at pH 9 using a membrane (Roth, Germany) with a molecular weight cut-off of 14,000 g/mol. For the hydrogel preparation colominic acid (0.2 g) was dissolved in NaOH (1 mL, 0.5 M) containing 10% sodium dodecylsulfate (SDS). di-epoxyoctane (96 μ L) was added and the solution was gently shaken for 3 d at room temperature. Afterwards the hydrogel was dialyzed for 7 d against sodium phosphate buffer (1 L, pH 7.0). The buffer was changed twice a day.

Enzymatic degradation of polysialic acid hydrogels

Swollen polySia hydrogel (39 mg, containing 1.4 mg dry material, cross-linked with 1 equivalent di-epoxyoctane; uncoated or coated with matrigel, collagen I, poly-L-lysine or a mixture containing PLL collagen I as described below) was added to phosphate buffered saline (400 μ L, pH 7.4) containing endoNF (4.26 μ g of the proteolytically processed endoNF lacking the N-terminal head-binding domain, i.e. Δ N-endoNF*; [20]). After incubation at 37 °C the hydrogel liquified in comparison to the control. The same procedure was employed for hydrogels with 0.6, 0.8, 2 and 3 equivalents di-epoxyoctane.

Gel electrophoresis and alcian blue silver staining of polysialic acid gels

Analysis of the degree of polymerization of colominic acid and the crude reaction mixtures was performed by using a polyacrylamide gel electrophoresis (PAGE) [21, 22]. Solutions for this method were prepared as follows: For a 50% acrylamide stock solution in H₂O 49.6 g acrylamide and 0.4 g bisacrylamide were solved in a volume of 100 mL. A 10-fold concentrated electrophoresis buffer contained 0.98 M tris base, 0.98 M boric acid and 0.02 M EDTA. 25% polyacrylamide gels were prepared by using 15 mL of 50% stock solution, 3 mL of 10-fold buffer and 12 mL H₂O. Polymerization was started by adding 125 µL of a 10% ammonium persulfate solution and 25 µL tetramethylethylenediamine. After complete polymerization the gel was preelectrophoresed at 10 V/cm for 1 h at 4 °C. The samples were mixed with 2 M sucrose in 10-fold electrophoresis buffer in a ratio of 1:1. Electrophoresis of tracking dyes (0.05% trypan blue, 0.02% xylene cyanol, 0.02% bromphenol blue, 0.02% bromcrescol purple and 0.02% phenol red) and of samples containing 2 and 5 µg colominic acid was performed at 6 V/cm for 14 h at 4 °C. Afterwards alcian blue silver staining of the polymeric compounds was performed. Therefore 0.5 g alcian blue was dissolved in 100 mL H₂O and filtrated. The gel was dyed for 30 min, rinsed with H₂O until the polysaccharide bands were visible and fixed in a solution of 5% acetic acid and 40% ethanol. After addition of oxidizer (0.7 g periodic acid, 5% acetic acid and 40% ethanol) for 5 min, the gel was rinsed three times with 300 mL H₂O for 15 min. Incubation in silver nitrate solution (0.6 g silver nitrate, 1.6 mL 25% ammonia, 1 mL 2 M sodium hydroxide filled up with H₂O to a volume of 100 mL) for 10 min was followed by washing with H₂O for three times for 10 min. The gel was incubated in developer (135 µL formaldehyde, 100 mL 240 µM citric acid) for 1 to 3 min and the reaction stopped by addition of 5% acetic acid in H₂O.

Mass spectrometric analysis

Hydrogels (containing 1 equivalent di-epoxyoctane) were hydrolyzed in 0.1 M trifluoroacetic acid at 80 °C for 2h, 4 h and 6 h and analyzed by ESI-MS using a Aquity UPLC (Waters) coupled Q-ToF premier (Waters) equipped with a LockSpray™ dual ion source (sample concentration 10 µg/mL in water, -2800 V capillary voltage, 120 °C source temperature, negative ESI mode). The sample prepared after 6 h hydrolysis showed the following main species: $m/z = 308.10$ (100%, sialic acid - H⁺); 468.21 (50%, sialic acid + linker - H⁺); 628.32 (47%, sialic acid + 2 linker - H⁺); 759.30 (20%, 2 sialic acid + linker - H⁺); 788.43 (25%, sialic acid + 3 linker - H⁺). To verify the identity of $m/z = 759.30$ LC-MS/MS studies were performed (collision energy ramped from 15 to 30 eV). MS/MS spectra and additional discussion are attached as supporting information.

Biological evaluation

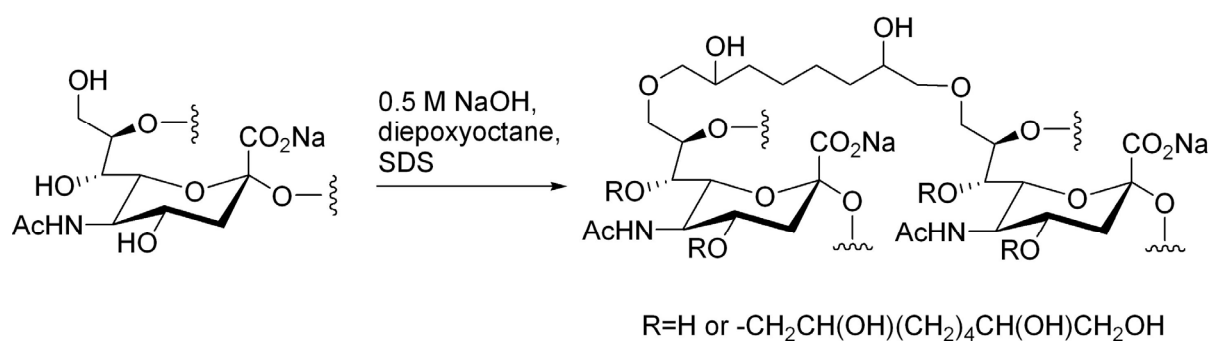
PC12 cells were cultivated under proliferation conditions in PC12 medium (Dulbecco's Modified Eagle's medium (DMEM) containing 10% (v/v) horse serum, 5% (v/v) fetal calf serum (FCS), 100 U/mL penicillin/0.1 mg/mL streptomycin, 1 mM sodium pyruvate and 6 mM L-glutamine). Pure, sterilized hydrogels and hydrogels preincubated with 0.2 mg/mL matrigel, 0.06 mg/mL collagen I, 0.5 mg/mL poly-L-lysine (PLL), and a mixture containing 0.5 mg/mL PLL and 0.06 mg/mL collagen I were placed on a 24-well culture plate. Afterwards, PC12 cells were cultured on the prepared hydrogels for 1 d under proliferation conditions and 3 to 6 d in differentiation medium (DMEM supplemented with 1% horse serum, 100 U/mL penicillin/0.1 mg/mL streptomycin, 1 mM sodium pyruvate, 6mM glutamine and 100 ng/mL nerve growth factor (NGF)). For fixation 4% (w/v) paraformaldehyde in PBS was added to the cells. PC12 cells were incubated for 40 min and washed three times for

overall 35 min with PBS supplemented with 0.3% triton® X 100 (polyethylene glycol tert-octylphenyl ether). Afterwards, DNA was stained with 4',6-diamidino-2-phenylindole (DAPI) and neurite outgrowth evaluated under phase contrast and fluorescence microscopy. In some cases additional immunocytochemistry was performed. Therefore, hydrogels were incubated in bovine serum albumin (BSA, 5% in 0.3% triton X 100/PBS) for 12 h to minimize unspecific interactions of the antibody with the hydrogel. The primary antibody anti-tubulin diluted in 1% (w/v) BSA/PBS was added for 180 min and the hydrogels were washed twice with PBS. The anti-mouse Alexa 555 antibody was subsequently added for 120 min and the wells washed with PBS for three times. The morphology of PC12 cells was studied under fluorescence microscopy.

RESULTS

Synthesis of PolySia hydrogels

For the synthesis of polySia based scaffold materials we chose di-epoxyoctane [23, 24] as a hydrophobic cross-linking agent to increase the molecular mass of the polymer and decrease its hydrophilicity (Scheme 2). First, we had to optimize the reaction conditions for obtaining three-dimensional hydrogels. Initially, we varied the concentration of sodium hydroxide and the total reaction volume and monitored the degree of polymerization (DP) of cross-linked polySia by means of polyacrylamide gel electrophoresis and subsequent alcian blue silver staining [21, 22]. Although all polymer chains exhibit the same ratio of charge to molecular weight, they are retained differently by the polyacrylamide gel network - small particles having a higher mobility than larger polySia.



SCHEME 2. Cross-linking of α 2,8-polysialic acid.

TABLE 1. Synthesis of polySia hydrogel. With V_R = reaction volume

Condition	c(NaOH)	V_R	di-epoxyoctane	Observation
	<i>M</i>	<i>mL</i>	<i>eq</i>	
A	0.25	0.5	6	liquid
B	0.5	0.5	12	liquid
C	0.5	0.25	6	liquid
D	0.5	0.1	3	hydrogel

We used the literature-known technique to qualitatively evaluate reaction conditions A, B and C (Table 1). PolySia shows a DP of over 52 and was used as a control (Figure 1). We observed no increase in molecular mass using 0.25 M NaOH (condition A). So we doubled the NaOH concentration (condition B) as well as halved the total volume of the reaction mixture (condition C). As shown in Figure 1 conditions B and C led to modified polySia derivatives with a higher molecular weight. We then increased the concentration of polySia to a maximum in 0.5 M NaOH (condition D) and finally observed formation of a clear insoluble hydrogel (Figure 2) and an increase of the polySia DP. The final protocol (20 mg scale) comprises the use of 20% (w/v) polySia in 0.5 M sodium hydroxide solution and three equivalents di-epoxyoctane.

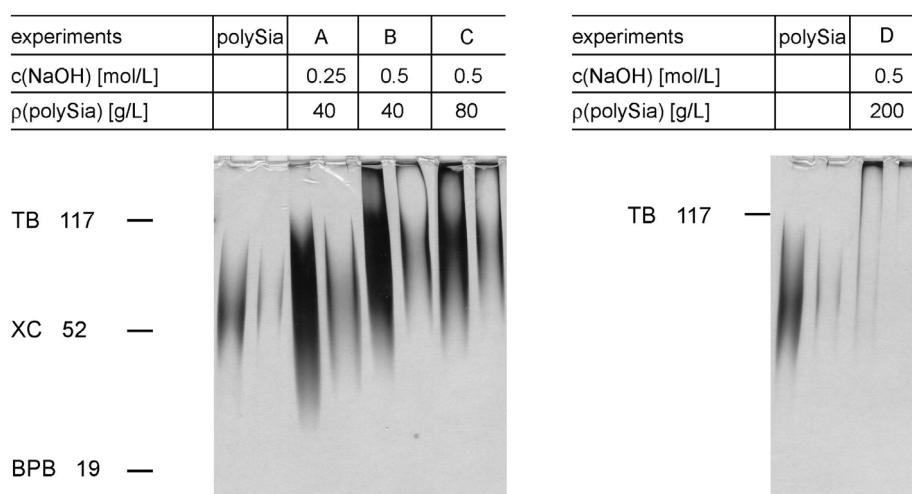


FIGURE 1. Polyacrylamide gels after alcin blue silver staining. The samples contained 2 μ g and 5 μ g of pure polySia as a control and the crude reaction mixtures of conditions A–D. The NaOH concentration c and polySia mass per volume ρ are given in the table above. Trypan blue (TB), xylene cyanol (XC) and bromphenol blue (BPB) were used as tracking dyes. Their mobility matched the numbers of sialyl residues: TB= 117; XC= 52 and BPB= 19.

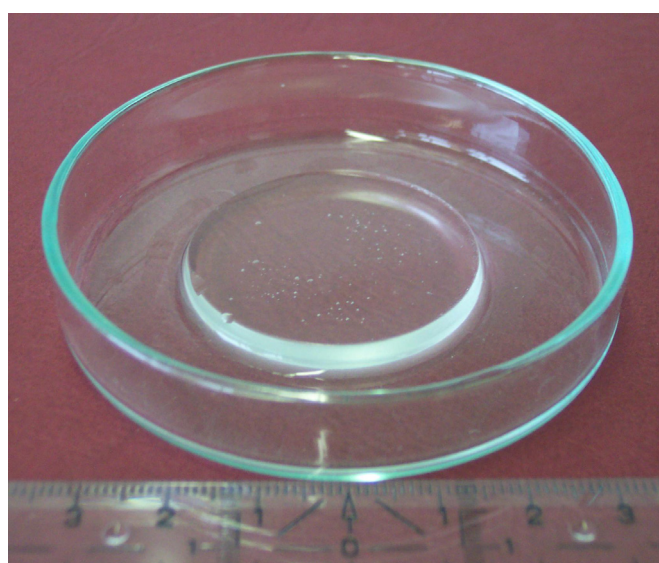


FIGURE 2. PolySia hydrogel (200 mg scale).

Incomplete dispersion of the cross-linker prohibited the formation of a hydrogel in a larger scale. This problem was solved by addition of sodium dodecyl sulphate (SDS) as a detergent which was removed after the reaction by extensive dialysis. Attempts to circumvent SDS by using water soluble ethylene glycol diglycidyl ether [25] failed but further experiments proofed that SDS was sufficiently removed. No negative side effects could be noticed in cellular studies [19, 26].

In order to judge whether an increase in hydrophobicity without additional cross-linking leads to the formation of insoluble hydrogels, the reaction was carried out with polySia and monoepoxyoctane under the same conditions. Although we used up to three equivalents of epoxide no formation of insoluble hydrogel was observed.

Enzymatic degradation experiments

First degradation experiments showed no activity of endoNF [20] towards hydrogels with three equivalents of di-epoxyoctane at physiological conditions of pH 7.4 and 37°C. In response to these biochemical results the amount of cross-linker was decreased to 0.6 equivalents, which turned out to be the minimum ratio for the formation of insoluble hydrogels. Enzymatic degradation experiments were carried out and liquidation of the polySia hydrogel was now detectable in comparison to the control sample. In contrast to the rapid degradation process of the natural substrate we noticed a drastically reduced turn-over of cross-linked hydrogels with 1, 0.6 and 0.8 equivalents of di-epoxyoctane (Table 2). The slow but controlled degradation of the cross-linked derivative of polySia is an advantage and a further essential requirement for application as a biomaterial for tissue engineering.

TABLE 2. Hydrogel degradation by endoNF. PolySia equivalents were calculated based on sialic acid units.

PolySia	Di-epoxyoctane	Degradation by endoNF
<i>eq</i>	<i>eq</i>	
1	3	no
1	2	no
1	1	completely after 11 d
1	0.8	completely after 2 d
1	0.6	completely after 2 d

Mass spectrometric analysis of polysialic acid hydrogel

The polymeric hydrogel composed of one equivalent of cross-linker was hydrolyzed in order to determine the degree of derivatization and visualize the product distribution. Thus, the α 2,8-glycosidic linkage of polySia was cleaved in 0.1 M trifluoroacetic acid at 80 °C and progression of hydrolysis was determined by mass spectrometric analysis after 2, 4 and 6 h. After 6 h hydrolysis was complete and no sialic acid dimers were detectable by ESI-MS. Based on the assumption that the monomeric products exhibit similar ionizabilities we found that the main fraction contained unmodified sialic acid and sialic

acid derivatized with one, two and three di-epoxyoctane molecules in a ratio of about 3.9 : 2 : 1.8 : 1. Additionally, dimers linked with one cross-linker were found in about 20% with regard to sialic acid (10% with regard to all sialic acid containing ions) and as was verified by LC-MS/MS fragmentation (see supplemental material). These dimers are important subunits as they should be responsible for hydrogel formation.

Cultivation of PC12 cells on polySia hydrogels

Primary biological evaluation of the synthesized hydrogels was performed by cultivation of PC12 cells on the hydrogel. PC12 cells were first obtained from a tumor of the adrenal medulla of a rat in 1976 [27]. During cultivation on collagen coated surfaces in the presence of nerve growth factor (NGF), PC12 cells differentiate and exhibit a neuronal phenotype. They display properties similar to sympathetic, peripheral neurons in primary cell culture and can be used as model system for neurons. In the present study, PC12 cells were used for biological evaluation of polySia based hydrogels as a potential matrix for nerve regeneration. Therefore, PC12 cells were cultivated on sterilized hydrogels under proliferation conditions for 12 h and incubated in serum reduced medium supplemented with NGF (differentiation medium) for 72 h. Afterwards they were fixed, stained with DAPI and analysed under phase contrast and fluorescence microscopy. Toxic effects could not be observed, but PC12 cells adhered poorly on the hydrogel surface. Morphologically, they exhibited a round shape and tended to grow in clumps both in proliferation and differentiation medium (Figure 3). Neurite outgrowth could not be observed under incubation with NGF.

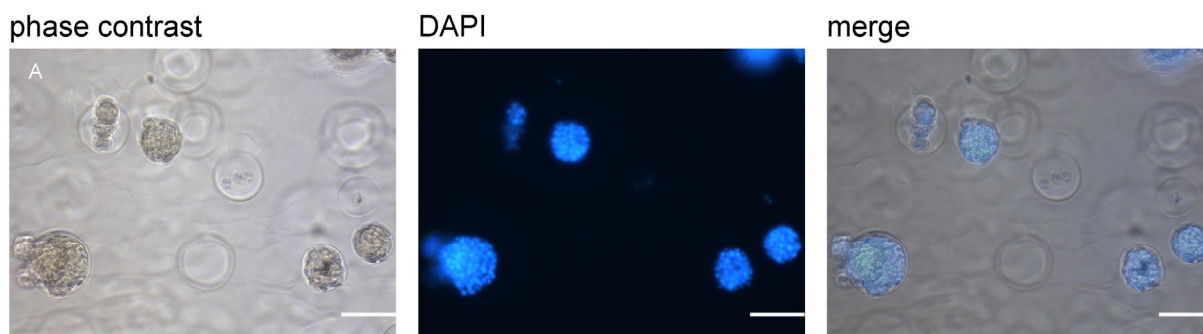


FIGURE 3. Morphology of PC 12 cells after cultivation on pure polySia hydrogel. PC12 cells were fixed and stained with DAPI after 24 d under proliferation and 72 d under differentiation conditions. Their morphology was examined by fluorescence microscopy. Only deepening arisen from di-epoxyoctane droplets during reaction were filled with PC12 cells (bar: 100 μ m).

In the following we masked the hydrogel by coating the surface with different substrates: poly-L-lysine (PLL), collagen I, mixtures of both of them and diluted matrigel[®]. We neither observed an improved cell adhesion nor neurite outgrowth using hydrogels preincubated with PLL after differentiation for 72 h. But the preincubation with collagen I, PLL/collagen I and diluted matrigel had a high impact on cell morphology and adhesion behaviour (Figure 4). The adhesion of PC12 cells seemed to be increased and neurite outgrowth was observed using collagen I (Figure 4A) and matrigel[®] (Figure 5A and B) as coating substrates.

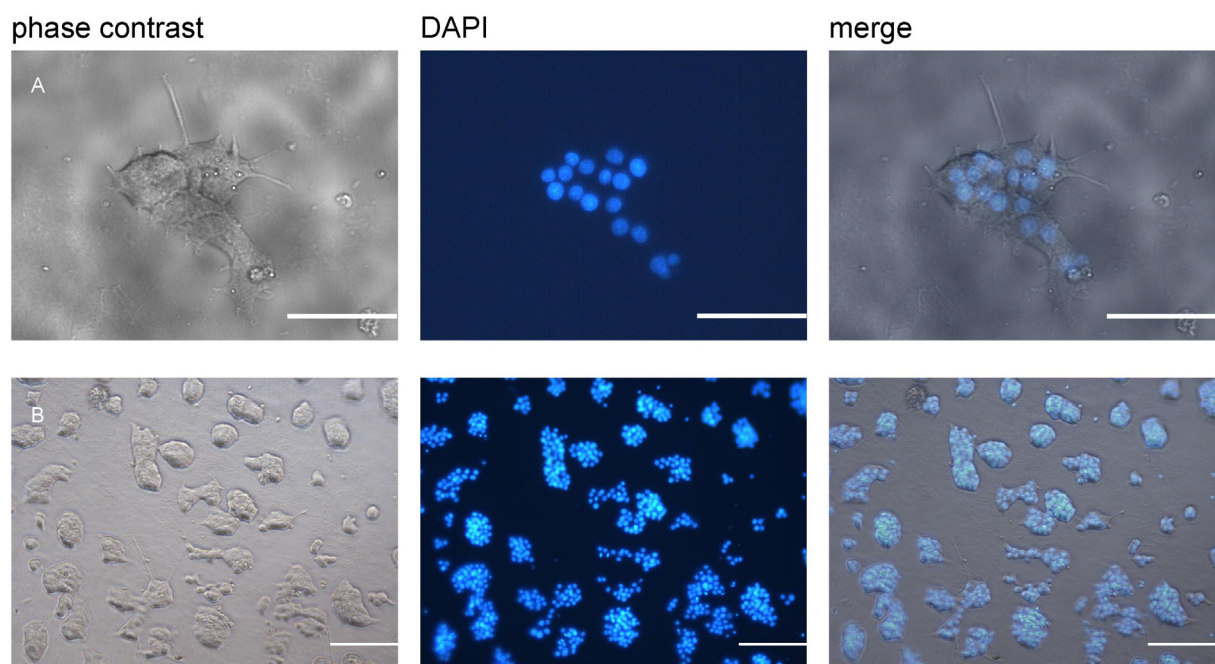


FIGURE 4. Morphology of PC12 cells after cultivation on polySia hydrogels. PC12 cells were fixed and stained with DAPI after 24 h under proliferation and 72 h differentiation conditions. Hydrogels were preincubated with collagen I (A; bar: 50 μm) and a mixture of collagen I and PLL (B; bar: 100 μm). PC12 cells tend to grow in small aggregates under optimal cell culture conditions [27]. This tendency is increased on collagen I coated hydrogels compared to coated cell culture dishes (not shown).

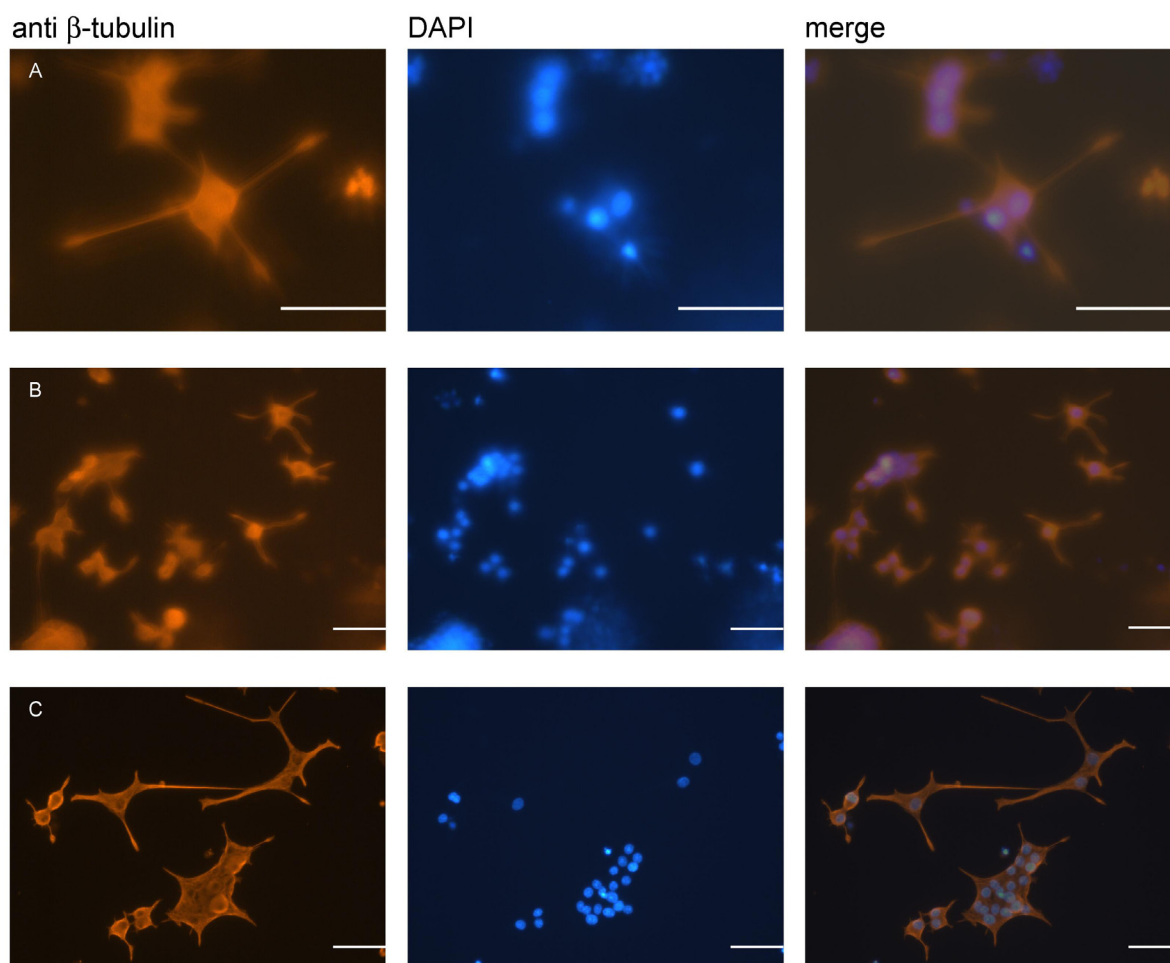


FIGURE 5. Morphology of PC12 cells on hydrogels preincubated with a diluted solution of matrigel. (A, bar: 40 μm and B, bar: 50 μm) after 6 days under differentiation conditions. As a control PC12 cells were cultivated on a plastic well coated with collagen I (C, bar: 50 μm). Cells were fixed and counterstained with an antibody against β -tubulin (secondary antibody anti-mouse Alexa 555) and DAPI. PC12 cells in picture A and B could not be focused due to an uneven hydrogel surface.

DISCUSSION

In this study we present the synthesis and biological evaluation of a new polysialic acid hydrogel as enzymatically degradable scaffold material for tissue engineering.

In contrast to hyaluronic acid polysialic acid has a very low molecular weight. The commercial material colominic acid (CA) has a degree of polymerization (DP) of approximately 50 (15.000 Da) according to PAGE analysis. Hydrogel formation using this material quality proved to be challenging since methods usually applied for hyaluronic acid initially failed and had to be optimized to give stable hydrogels and reproducible results. First experiments just resulted in slight enhancement of the molecular weight after cross-linking. Using a concentrated aqueous CA solution was crucial to further increase the reaction efficiency. To remove smaller polysialic acid oligomers which initially hampered the cross-linking reaction, we decided to purify commercial CA by extensive dialysis using $1.4 \cdot 10^4$ Da membranes. Using this material the reproducible synthesis of polysialic acid hydrogels was possible. Production of larger amounts of hydrogel was still problematic because di-epoxyoctane separated from the reaction

mixture. Since water miscible diepoxides gave no positive results, we decided applying additives to solubilize the cross-linking reagent. Although complete removal of SDS after cross-linking is challenging, this method allowed scaling the reaction to produce enough hydrogel for comparative studies without problems due to possible batch-to-batch variations. Still it is noteworthy that smaller batches (20 mg CA) can be easily obtained without SDS.

Although higher amounts of cross-linker were mandatory for gel formation, the resulting material was still degradable using the phage-born endosialidase (endo-N-acetylneuraminidase, endoN) [15–17]. EndoN can be applied *in vivo*. In mice no toxic effects were observed after injection of high doses of endoN [18] and sialic acid containing glycans were not affected [28]. The speed of degradation could be adjusted depending on the amount of cross-linker used. Under laboratory conditions degradation took place within 2 to 11 day which will be in good agreement with a stepwise tissue regeneration process. Moreover, even degradation products have no negative effect on cell cultures [26]. It is most important to point out that in contrast to hyaluronidase, which degrades hyaluronic acid based materials, there is no known human endosialidase. Polysialic acid is perfectly stable in humans. Degradation can be externally triggered applying endoN. This makes polysialic acid hydrogel an auspicious material because after implantation a second surgery, one major drawback a non-degradable scaffolds, will not be necessary.

Unfortunately PC12 cells cultivated on pure polyanionic hydrogel morphologically exhibited a round shape and tended to grow in clumps (Figure 3). Neurite outgrowth could not be observed under incubation with NGF. This effect can be explained with regard to the negative charge of polySia and its role in nervous tissue, wherein upregulation of polySia results in a higher motility of cells. Our results indicated that the strong negatively charged hydrogels impeded neurite outgrowth. To provide support for cell adhesion *in vitro* poly-L-lysine, collagen I, mixtures of both and diluted matrigel[®] were evaluated as coatings for the hydrogel. This approach is also used for other polyanionic materials as hyaluronic acid derived gels [25]. Preincubation with collagen I, PLL/collagen I and diluted matrigel had a high impact on cell morphology and adhesion behaviour (Figure 4). After coating all hydrogels can be easily degraded by endoN. Although future synthetic adaptations and surface modifications may be needed to optimize the quality for *in vivo* applications, the observed differentiation and neurite outgrowth of PC12 cells on the hydrogel surface as well as controlled degradability by the phage-born endoNF demonstrates the high potential of modified polySia hydrogels as innovative biomaterial for nerve regeneration.

CONCLUSIONS

In conclusion we reported a robust and scalable synthesis of a polySia hydrogel. The resulting material is stable under physiological conditions, enzymatically degradable by an externally applied endosialidase, fully biocompatible and could be evaluated as growth support for PC12 cells. Coating with collagen I, PLL or matrigel improves the properties of the novel material. The combination of the polySia hydrogel building up a three-dimensional matrix with a negative charge, and coating of the hydrogel with different components from the extracellular matrix leads to a new field of applications for tissue engineering.

ACKNOWLEDGEMENTS

We thank Dr. K. Stummeyer for her expertise and kind help with biochemical techniques. Y. Haile is gratefully acknowledged for helpful discussion. This work was supported by the Deutsche Forschungsgemeinschaft (grants Ki 397 / 8-1 and Ki 397 / 9-1) and the Fonds der Chemischen Industrie

REFERENCES

1. Tabata Y. Role of Gelatin in the Release Carrier of Growth Factor for Tissue Engineering. In Ma PX, Elisseeff JH, editors. *Scaffolding in Tissue Engineering*. Boca Raton: CRS Press, Taylor & Francis 2006. p. 45-60.
2. Langer R, Vacanti JP. Tissue engineering. *Science* 1993;260:920-926.
3. Timmer M, Robben S, Müller-Ostermeyer F, Nikkhah G, Grothe C. Axonal regeneration across long gaps in silicone chambers filled with Schwann cells overexpressing high molecular weight FGF-2. *Cell Transplant* 2003;12:265-277.
4. Madison R, da Silva CF, Dikkes P, Chiu TH, Sidman RL. Increased rate of peripheral nerve regeneration using bioresorbable nerve guides and a laminin-containing gel. *Exp Neurol* 1985;88:767-772.
5. Guénard V, Kleitman N, Morrissey TK, Bunge RP, Aebischer P. Syngeneic Schwann cells derived from adult nerves seeded in semipermeable guidance channels enhance peripheral nerve regeneration. *J Neurosci* 1992;12:3310-3320.
6. Haastert K, Lipokatic E, Fischer M, Timmer M, Grothe C. Differentially promoted peripheral nerve regeneration by grafted Schwann cells over-expressing different FGF-2 isoforms. *Neurobiol Dis* 2006;21:138-53.
7. Chamberlain LJ, Yannas IV, Hsu HP, Spector M. Connective tissue response to tubular implants for peripheral nerve regeneration: The role of myofibroblasts. *J Comp Neurol* 2000;417:415-430.
8. Jones AC, Milthorpe B, Averdunk H, Limaye A, Senden TJ, Sakellariou A, Sheppard AP, Sok RM, Knackstedt MA, Brandwood A, Rohner D, Hutmacher DW. Analysis of 3D bone ingrowth into polymer scaffolds via micro-computed tomography imaging. *Biomaterials* 2004;25:4947-4954.
9. Laurent TC, Fraser JR. Hyaluronan. *FASEB J* 1992;6:2397-2404.
10. Kiselyov VV, Soroka V, Berezin V, Bock E. Structural biology of NCAM homophilic binding and activation of FGFR. *J Neurochem* 2005;94:1169-1179.
11. Brusés JL, Rutishauser U. Roles, regulation, and mechanism of polysialic acid function during neural development. *Biochimie* 2001;83:635-643.
12. Weinhold B, Seidenfaden R, Röckle I, Mühlhoff M, Schertzinger F, Conzelmann S, Marth JD, Gerardy-Schahn R, Hildebrandt H. Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. *J Biol Chem* 2005;280:42971-42977.
13. El Maarouf A, Kolesnikov Y, Pasternak G, Rutishauser U. Polysialic acid-induced plasticity reduces neuropathic insult to the central nervous system. *Proc Natl Acad Sci USA* 2005;102:11516-11520.
14. El Maarouf A, Petridis AK, Rutishauser U. Use of polysialic acid in repair of the central nervous system. *Proc Natl Acad Sci USA* 2006;103:16989-16994.
15. Mühlhoff M, Stummeyer K, Grove M, Sauerborn M, Gerardy-Schahn R. Proteolytic processing and oligomerization of bacteriophage-derived endosialidases. *J Biol Chem* 2003;278:12634-12644.
16. Hallenbeck PC, Vimr ER, Yu F, Bassler B, Troy FA. Purification and properties of a bacteriophage-induced endo-N-acetylneuraminidase specific for poly-alpha-2,8-sialosyl carbohydrate units. *J Biol Chem* 1987;262:3553-3561.
17. Stummeyer K, Dickmanns A, Mühlhoff M, Gerardy-Schahn R, Ficner R. Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol* 2005;12:90-96.
18. Daniel L, Durbec P, Gautherot E, Rouvier E, Rougon G, Figarella-Branger D. A nude mice model of human rhabdomyosarcoma lung metastases for evaluating the role of polysialic acids in the metastatic process. *Oncogene* 2001;20:997-1004.

19. Haile Y, Haastert K, Cesnulecicius K, Stummeyer K, Timmer M, Berski S, Dräger G, Gerardy-Schahn R, Grothe C. Culturing of glial and neuronal cells on polysialic acid. *Biomaterials* 2007;28:1163-1173.
20. Schwarzer D, Stummeyer K, Gerardy-Schahn R, Mühlhoff M. Characterization of a novel intramolecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fiber proteins. *J Biol Chem* 2007;282:2821-2831.
21. Pelkonen S, Häyrinen J, Finne J. Polyacrylamide gel electrophoresis of the capsular polysaccharides of *Escherichia coli* K1 and other bacteria. *J Bacteriol* 1988;170:2646-2653.
22. Tikkanen K, Häyrinen J, Pelkonen J, Finne J. Immunoblot analysis of bacterial polysaccharides: Application to the type-specific polysaccharides of *Streptococcus suis* and *Streptococcus agalactiae*. *J Immunol Methods* 1995;187:233-244.
23. Šimkovic I, Hricovini M, Mendichi R, van Soest JJG. Cross-linking of starch with 1, 2, 3, 4- di-epoxybutane or 1, 2, 7, 8- di-epoxyoctane. *Carbohydr Polym* 2004;55:299-305.
24. Zhao XB, Fraser JE, Alexander C, Lockett C, White BJ. Synthesis and characterization of a novel double crosslinked hyaluronan hydrogel. *J Mater Sci: Mater Med* 2002;13:11-16.
25. Segura T, Anderson BC, Chung PH, Webber RE, Shull KR, Shea LD. Cross-linked hyaluronic acid hydrogels: a strategy to functionalize and pattern. *Biomaterials* 2005;26:359-371.
26. Haile Y, Berski S, Dräger G, Nobre A, Stummeyer K, Gerardy-Schahn R, Grothe C. Modified polySia hydrogel improved the adhesion and viability of primary neurons and glial cells. Manuscript in preparation.
27. Greene LA, Tischler AS. Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. *Proc Natl Acad Sci USA* 1976;73:2424-2428.
28. El Maarouf A, Rutishauser U. Removal of polysialic acid induces aberrant pathways, synaptic vesicle distribution, and terminal arborization of retinotectal axons. *J Comp Neurol* 2003;460:203-211.

7.1 – Supplemental Material

Hydrogels (containing 1 equivalent diepoxyoctane) were hydrolyzed in 0.1 M trifluoroacetic acid at 80°C for 4 h and 6 h and analyzed by LC-MS/MS. In particular, we concentrated on the sialic acid dimer with one linker unit (Scheme S1 and S2). Both molecules **S1** and **S2** have the same molecular mass ($m/z = 759.30$), but should give a different MS/MS pattern. After 4 h and 6 h hydrolysis the MS/MS of $m/z=759.30$ indeed display a different fragmentation pattern: The spectrum after 4 h hydrolysis exhibits a higher amount of the daughter ion $m/z = 581.1819$ consisting of a dehydrated sialic acid dimer, which can only originate from the mother ion of two glycosidically linked sialic acids and one linker unit (Scheme S1). As hydrolysis of glycosidic bonds proceeds from 4 h to 6 h, the spectrum completely changes: The intensity of the peaks of the ions with $m/z = 170.0451$, 290.0874 (dehydrated sialic acid) and 468.2091 (sialic acid with one linker unit) increases (Figure S2). Now the crosslinked dimer with a fragmentation shown in scheme S2 is the main dimeric compound.

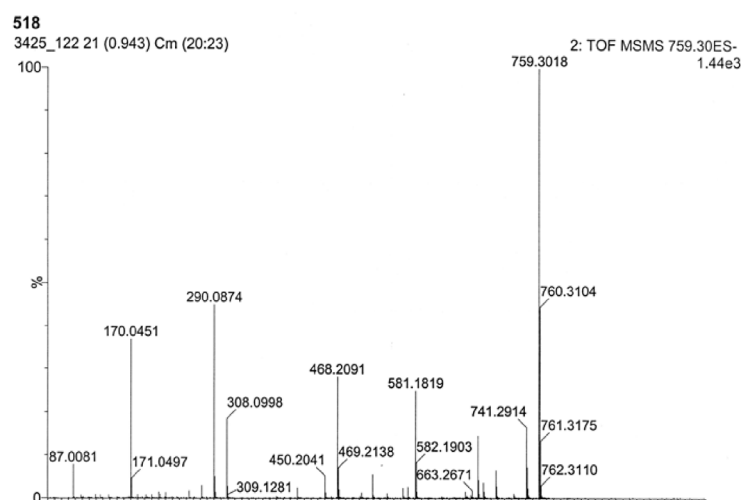


FIGURE S1. LC-MS/MS spectrum of two type of dimers after 4 h in 0.1 M trifluoroacetic acid at 80°C: two sialic acids crosslinked with diepoxyoctane (**S1**) and a glycosidically linked dimer with one linker unit (**S2**).

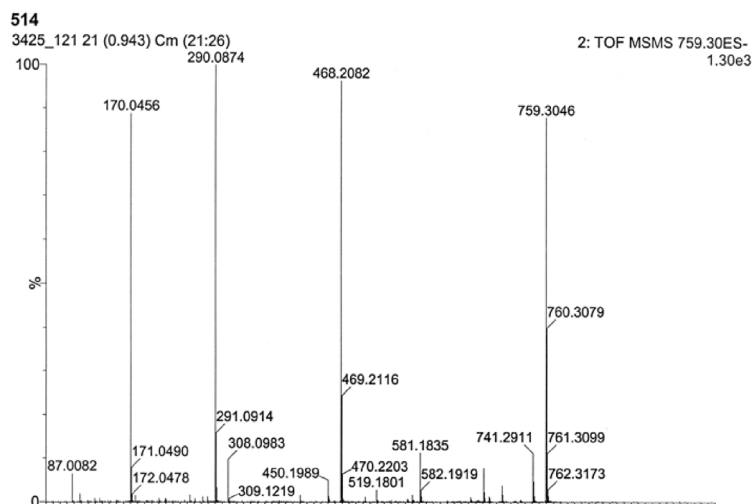
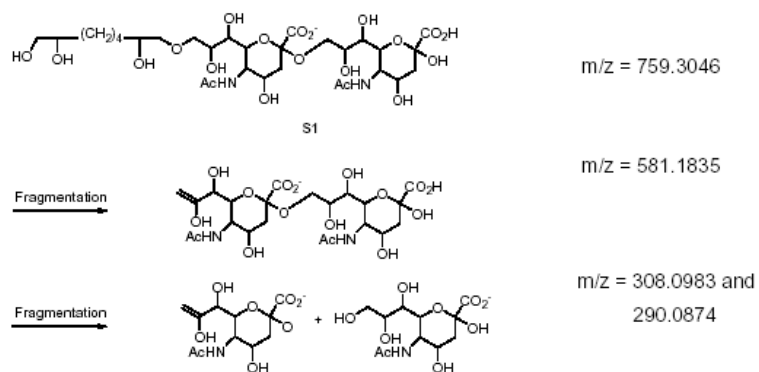
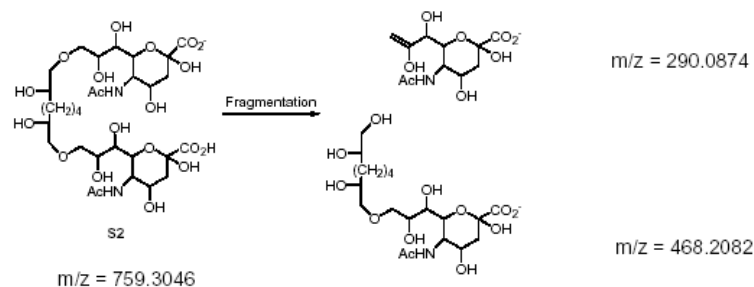


FIGURE S2. LC-MS/MS spectrum of sialic acid dimers crosslinked with one diepoxyoctane molecule after 6 h in 0.1 M trifluoroacetic acid at 80°C. The fragmentation pattern of a component consisting of two glycosidically linked sialic acids and one linker unit can be seen with low intensity.



SCHEME S1. Fragmentation of the glycosidically linked sialic acid dimer containing one linker unit.



Scheme S2. Fragmentation of the sialic acid dimer crosslinked with di-epoxyoctane.

Chapter 8 – General Discussion

8.1 – Evolution of *Escherichia coli* K1 specific Bacteriophages

In the first part of this thesis, a panel of *Escherichia coli* K1 specific phages was investigated as a model system to study the evolution of bacteriophages infecting encapsulated bacteria. The common feature among all K1-phages identified so far is the endosialidase tailspike protein. This polysialic acid degrading enzyme provides the key determinant of host specificity as it enables the phage to bind to its host and penetrate the thick polySia capsule. By comparative genome analysis, it was shown that K1-specific phages did not evolve from a single common ancestor. In contrast, the genomic organisation of the investigated K1-phages indicated a close relationship to different progenitor phages and strongly suggested that lateral acquisition of the endosialidase tailspike gene was the main evolutionary event in gaining host specificity for *E. coli* K1. Even within the small subset of four K1-phages investigated, members of three different phage groups, namely P22-, T7- and SP6-like phages, were identified, indicating that the uptake of the endosialidase gene might have occurred with high frequency.

Interestingly, endosialidases are exclusively found as tailspike proteins of K1-specific phages and do not have a counterpart in any other organism. Since there are no sequence similarities between endosialidases and other known proteins, structural information is essential to gain insight into how these specialised tailspike proteins have evolved. The crystal structure of the central catalytic part of endoNF, the tailspike protein of Coliphage K1F, revealed the presence of a 6-bladed β -propeller (Stummeyer *et al.*, 2005), a structural element found in all sialidases crystallised up to now (Taylor, 1996). In endoNF, the β -propeller is fused to two structural motifs characteristic for bacteriophage tail proteins: triple β -helix and triple β -prism (Stummeyer *et al.*, 2005). Triple-beta folds are associated with trimeric proteins of unusually high stability and have been identified in several bacteriophage tail proteins (Steinbacher *et al.*, 1994; van Raaij *et al.*, 2001; Kanamaru *et al.*, 2002). The structural data indicate that endosialidase tailspikes evolved from shuffling of a complete sialidase domain into the structural context of a phage tail protein. Moreover, the high amino acid sequence identity between the catalytic parts of all known endosialidases suggests that an endosialidase gene evolved only once and was then transmitted to different phage genomes by horizontal gene transfer. Since simultaneous infection of the same bacterium by two lytic phages is a rare event (Brüssow and Hendrix, 2002; Lawrence *et al.*, 2002), the involvement of a temperate endosialidase carrying phage like CUS-3 may have played an important role in transmitting this tailspike to incoming lytic phages. The

comparative genome analysis provided in the present study suggests that the uptake of the endosialidase gene occurred by illegitimate recombination into the tailspike or tail fibre locus of a P22-, a T7-, and a SP6-like progenitor phage, resulting in the K1-phages CUS-3, K1F, and K1E, respectively. Notably, the generated endosialidases of CUS-3 and K1F are chimeric tailspikes that kept the capsid-binding domain of their progenitor phages to ensure proper connection of the newly acquired tailspike to the phage. In contrast, in SP6-like phages, including K1E, the tailspikes share an N-terminal undecapeptide (MIQRLGSSLVK) that mediates the connection to the phage via the same adapter protein. Thus, two different strategies were identified how tailed phages incorporate new tailspikes into their highly conserved capsid structures: (i) either the original head-binding domain of the progenitor phage is kept, resulting in fusion proteins combining head-binding and host range determining functions, or (ii) the newly acquired tailspikes are connected to the tail by a versatile adapter protein which requires only an N-terminal undecapeptide for binding.

8.2 – Characterisation of the C-terminal Chaperone Domain of Endosialidases

In addition to the highly conserved central catalytic part, all endosialidases share a short C-terminal domain (CTD), which is released during maturation by proteolytic cleavage at a conserved Ser ↓ Asx-site. In the second part of this study, the CTD was characterised as a novel C-terminal intra-molecular chaperone domain, essential for assembling and folding of endosialidases and other tailspike and fibre proteins of diverse phages. Altogether, 13 phage proteins containing this domain were identified, which can be grouped into 4 different protein families: (i) endosialidases; (ii) K5-lyases, the capsule degrading tailspike proteins of phages infecting *E. coli* K5 that are also found in the genome of an *E. coli* K5 strain (Clarke *et al.*, 2000; Scholl *et al.*, 2001); (iii) the L-shaped tail fibres (LTF) of coliphage T5, which accelerate phage adsorption by reversible binding to the O-antigen of the host (Heller and Braun, 1979); and (iv) the neck appendage protein gp12 of *Bacillus* phages GA-1, Φ29, PZA, and B103, which forms appendages that are essential for the phage attachment (Meijer *et al.*, 2001). By analyzing one representative of each protein family (endoNF, elmA, LTF, and gp12 of *Bacillus* phage GA-1), it could be shown that regardless of the N-terminal sequence, the CTD is released during protein maturation. Furthermore, the highly conserved proteolytic cleavage site Ser ↓ Asx that had been identified in previous studies on endoNE and endoNF (Leggate *et al.*, 2002; Mühlenhoff *et al.*, 2003), was found in all proteins investigated. The essential role of the serine residue (Ser-911 in endoNF) was confirmed by alanine substitution. This mutation prevented cleavage but did not affect folding and complex assembly. For all proteins analysed, an intact CTD was essential for proper folding

and oligomerisation. Proteins lacking the CTD due to C-terminal truncation were exclusively expressed as insoluble proteins. CTD swapping between endoNF and gp12 resulted in an active, trimeric endoNF-gp12-chimera, suggesting that the CTD has a general chaperone function that promotes folding and complex formation independent of the N-terminal protein context.

A first structural characterisation of the CTD of endoNF by circular dichroism analysis revealed that this domain is predominantly composed of α -helices. This is in clear contrast to the catalytic part of endoNF, which is build up almost exclusively by β -folds (Stummeyer *et al.*, 2005). Both, proteolytically released and separately expressed CTDs showed identical secondary and quaternary structures, demonstrating that the CTD represents an independent folding unit. Although this domain is trimeric in the non-cleaved mutant endoNF-S911A, hexamers were observed for processed CTDs, suggesting that dimerization of trimers occurs after release. Interestingly, gp57, a phage-derived chaperone involved in assembly of the long and short tail fibres of Coliphage T4 (Dickson, 1973; Burda and Miller, 1999), is also largely α -helical and forms trimers which dimerise to hexamers (Ali *et al.*, 2003). Future studies that aim at solving the 3D-structures will clarify whether both proteins adopt similar folds despite the lack of sequence similarity.

Proteolytic maturation is a common theme among phage proteins involving either phage-derived proteases or autocatalytic processes (Laemmli, 1970; Marvik *et al.*, 1994; Conway *et al.*, 1995; Kanamaru *et al.*, 2002; Wang *et al.*, 2003). In contrast to pro-enzymes like mammalian digestive serine proteases, the function of proteolytic maturation of phage proteins is not restricted to enzyme activation. In the case of the major head proteins of coliphage T4, cleavage is required for head expansion and DNA packaging (Miller *et al.*, 2003). Similarly, the release of the endosialidase CTD is not a prerequisite to gain enzymatic activity (Mühlenhoff *et al.*, 2003). Therefore, a major aim of this study was to investigate the function of proteolytic maturation of endoNF. A comprehensive comparison of processed endoNF and the non-processed variant endoNF-S911A demonstrated that release of the CTD has a twofold impact on maturation of the endosialidase complex: (i) by unmasking polySia-binding sites located outside the catalytic center and (ii) by increasing trimer stability.

The crystal structure of endoNF revealed that in addition to the active site (site *a*), sialic acid binding sites are located in the lectin-like β -barrel domain (site *b1*) and in the β -prism domain of the stalk (*b2*) (Stummeyer *et al.*, 2005). The geometry of the binding sites suggests that the polymeric substrate is wrapped around the trimeric endoNF complex, interacting simultaneously with all three binding sites. As summarised in Chapter 6, the

CTD preserved in the unprocessed variant endoNF-S911A interferes with efficient binding to polymeric substrate. As a consequence, this variant showed drastically reduced activity towards surface-bound substrate, whereas a 3-fold higher molar activity was observed towards soluble polySia. However, for both, mature and unprocessed endoNF, identical kinetic parameters ($K_M=0.85$ mM, and $k_{cat}=4.37$ sec⁻¹) were determined for sialic acid tetramers, a substrate which is too short to reach binding sites b1 or b2 when bound to the active site. This clearly demonstrates that the conformation of the active site is not affected by the presence of the CTD and that reduced polySia binding is due to interference of the CTD with substrate binding sites located outside the active site. This was confirmed by endoNF variants with single and double alanine substitutions of amino acids involved in sialic acid binding within the b2-site. Similar to the unprocessed endoNF-S911A variant, mutants with a mutated b2 site were characterised by reduced polySia binding combined with increased activity towards soluble substrate.

The observation that reduced polymer binding either increases or decreases the enzymatic activity, depending on the accessibility of the substrate, was previously reported for cellulases and chitinases that act on insoluble sugar polymers (Koivula *et al.*, 1998; Zhang *et al.*, 2000; Watanabe *et al.*, 2003; Katouno *et al.*, 2004). In these enzymes, the substrate-binding clefts are lined with aromatic residues that are thought to facilitate sliding of the polymer chain through the cleft during a processive mode of action. As shown for chitinase B from *Serratia marcescens*, single alanine substitution of aromatic residues in the binding cleft reduced processivity and thereby the ability to degrade crystalline chitin, but increased the degradation rate for soluble substrate (Horn *et al.*, 2006). In line with this, endoNF may also act as a processive enzyme and processivity might be a common feature of depolymerases that combine degradation and binding capabilities. The view that tailspike proteins with depolymerase activity like endosialidases are processive enzymes is supported by the observation that bacteriophages infecting encapsulated bacteria create a narrow tunnel through the polysaccharide capsule of the host (Sutherland, 1977). Furthermore, it is conceivable that the trimeric endoNF can interact with three polysialic acid chains simultaneously and that processive migration along the chains is mediated by independent and staggered association-cleavage-dissociation processes of the three subunits. Removal of polymer binding sites will affect the directionality of these processes, resulting in decreased processivity.

In Chapter 6 it is demonstrated that proteolytic maturation of endosialidases and removal of the C-terminal chaperone domain is required to gain effective binding of endoNF to polysialic acid. Hence, this requirement represents an evolutionary driving force to emerge

the Ser ↓ Asx cleavage site, and K1-phages with a preserved CTD in the endosialidase might exhibit a decreased infection rate when compared to wild type phages containing a processed endosialidase. However, since the data presented in Chapter 5 indicate that the proteolytic release of the CTD occurs also in other bacteriophage tailspike and fibre proteins unrelated to endosialidases, cleavage may have other functions beyond unmasking substrate binding sites. Comparison of the thermal stability of processed and unprocessed endoNF (Chapter 5) revealed that release of the CTD increases trimer stability. Based on these results, a folding pathway was proposed, which is exemplarily depicted for endoNF in *Figure 8-1*.

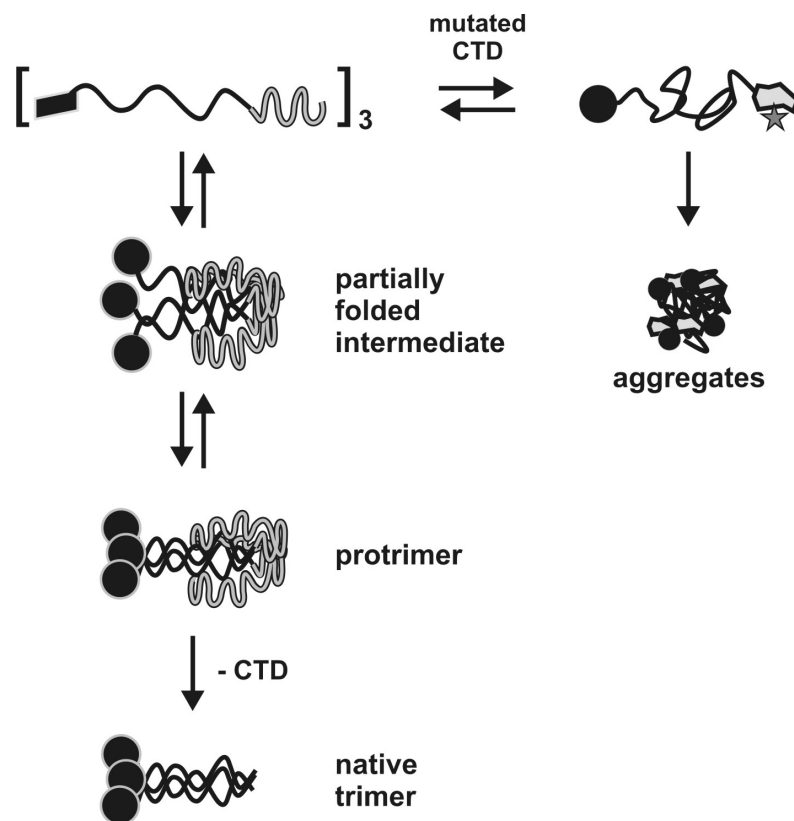


FIGURE 8-1: Proposed folding pathway of endoNF. In the schematic representation of endoNF, the CTD is shown in *grey*, and the folded β -propeller is shown as a *black circle*. Adapted from Chapter 5, *Figure 7* and Chapter 6, *Figure 7*.

The folding of the nascent polypeptide chain starts with the formation of independent N-terminal domains such as the endosialidase β -propeller and the β -barrel domain. Upon release from the ribosomes, the monomers assemble into trimers. A functional CTD is crucial in this step, since alanine substitutions of critical amino acid residues within this domain prevented oligomerisation and caused the formation of aggregates. To fulfil its chaperone function, it was suggested in Chapter 6 that the CTDs might interact with the stem region. CTD-assisted trimer assembly may involve several partially folded intermediates leading to an SDS-resistant protrimer that still contains CTDs and which can

be mimicked by the non-cleavable variant endoNF-S911A. Before dissociation into monomers, trimeric endoNF-S911A became partially unfolded and an SDS-resistant intermediate with a lower electrophoretic mobility than the native trimer was detected. Compared to processed complexes, full-length trimers showed decreased thermostability, indicating that CTD-release stabilises the native complex. Thus, proteolytic cleavage may lead to a kinetically stable protein, a trimer which is trapped in a specific conformation due to an unusually high unfolding barrier. In this model, the role of the CTD would reside in overcoming the high energy barrier during the folding process. Consequently, mutants with a defective or absent CTD are unable to adopt the transition state conformation required for the formation of functional trimers.

Increased protein stability due to proteolytic cleavage was also described for the α -lytic protease (α LP) of *Lysobacter enzymogenes*, a secreted serine protease involved in the degradation of other soil bacteria (Cunningham *et al.*, 1999). The folding of α LP essentially depends on the Pro domain which stabilises the folding transition state and the native state of α LP. A release of the Pro domain traps the mature protein in a metastable conformation of high rigidity and low susceptibility to proteases (Cunningham *et al.*, 1999). In contrast to the endosialidase CTD, the Pro region is located at the N-terminus. In addition, Pro can be expressed in *trans* as a separate polypeptide chain to assist folding of α LP, whereas the endosialidase CTD only acts in *cis* as part of the precursor protein (Mühlenhoff *et al.*, 2003). These differences may reflect the fact that Pro mediates folding of a monomeric protein, whereas the endosialidase CTD is involved in folding and assembly of an oligomeric complex.

Enhancing longevity achieved by kinetic stability seems to be a general mechanism observed for SDS-resistant proteins (Manning and Colón, 2004). Many of these are viral proteins which form oligomeric complexes with a high content of β -folds. Interestingly, some of these also possess an assembly module. In the tailspike protein of *Salmonella* phage P22, an endorhamnosidase with triple β -helix domain (Steinbacher *et al.*, 1994), the C-terminal domain is essential for folding and assembly of the SDS-resistant trimer (Kreisberg *et al.*, 2002; Gage and Robinson, 2003). However, unlike the endosialidase CTD, the caudal fin domain of P22 remains part of the native trimer. This is also true in case of the human adenovirus type 2 adhesin (Ad2) which forms trimeric fibres with a triple- β -spiral shaft (van Raaij *et al.*, 2001). The globular C-terminal domain is not only involved in receptor binding, but is also required for the complex assembly. Another example for C-terminally located folding domains is the “foldon”, the 27-aa CTD of fibrinin, a segmented coiled-coil homotrimer which forms the fibrous whiskers of coliphage T4 (Tao *et al.*, 1997).

The foldon forms a β -propeller and is essential for a correct fibrin assembly. Interestingly, the function of the foldon is independent of the N-terminal protein part, and the foldon can substitute the corresponding assembly domain of Ad2 (Papanikolopoulou *et al.*, 2004). Moreover, the foldon was successfully used to stabilise HIV gp120, collagen, and T4 long tail fibres (Miroshnikov *et al.*, 1998; Frank *et al.*, 2001; Yang *et al.*, 2002). However, in all cases, the C-terminal assembly domain remains part of the trimeric protein complex and is composed of β -folds. In contrast to the P22 tailspike, Ad2, and fibrin, the C-terminal chaperone domain of endosialidases is predominantly α -helical and is removed after completing the job. As shown for endoNF, proteolytic processing might have evolved not only to increase the unfolding barrier but also to avoid structural constraints leading to reduced substrate binding.

In total, the present study has shown that proteolytic maturation is a common feature amongst bacteriophage tailspike and tail fibre proteins sharing an endosialidase C-terminal chaperone domain. The idea that the proteolytic cleavage leads to the formation of kinetically stable complexes is complemented by the finding that cleavage exposes a sialic acid binding site in the stalk domain of endoNF. Since structural data of the other bacteriophage proteins characterised in this study are not available yet, there is no information whether these proteins also contain specific substrate binding sites that could be influenced by the presence of the CTD. However, one important role of the C-terminal domain is clearly represented by the chaperone function essential for folding and assembly of tailspike and fibre proteins with unrelated N-terminal protein parts. Considering the different functions of these tailspike and fibre proteins, it is likely that the CTD including the cleavage site was fused to different tail proteins by horizontal gene transfer. As suggested above, the proteolytic cleavage of the CTD might represent a general mechanism to gain kinetically stable protein complexes. This is in line with the fact that tailspikes and tail fibres are exposed structures, which require high stability to maintain their functional conformation even under extreme environmental conditions such as high salt concentrations, the presence of extra-cellular proteases, and wide variations in pH and temperature (Steinbacher *et al.*, 1994; Chappell *et al.*, 1997; van Raaij *et al.*, 2001; Freiberg *et al.*, 2003; Weigele *et al.*, 2003). Therefore, the major driving force to evolve and maintain the Ser \downarrow Asx proteolytic cleavage site was most likely to ensure that these proteins remain stable under harsh conditions.

8.3 – The Potential of Endosialidases in Medical Applications

Endosialidases are powerful tools for the specific depolymerization of polySia and gained major importance in different research fields, including neuro- and tumorbiology (cf. Chapter 3.5). As shown in Chapter 6 for enzymatically inactive endoNF, endosialidases bind to polySia in the nanomolar range. Therefore, inactive endosialidases can be used as antibody mimics for the specific detection of polySia (Jokilammi *et al.*, 2004) and might become valuable reagents to target polySia-positive tumor cells.

In the last part of this work, the application of endosialidases for selective degradation of polySia-based scaffold materials has been evaluated (Chapter 7). Since polySia is a poor immunogen (Jódar *et al.*, 2002) and a permissive substrate for nerve regeneration processes (Kleene and Schachner, 2004), this biomolecule provides a promising material for tissue engineering and in particular for neuro-reconstruction approaches. Since no endogenous endosialidases have been detected in mammals, polySia is a very stable polymer in the human body. In this study, it was demonstrated that endoNF can be used for the specific and time-controlled degradation of polySia-based scaffolds. Even though rather high concentrations of cross-linker were required to generate polySia-based hydrogels, the introduced chemical modifications did not abrogate specific degradation by endoNF. Notably, the released degradation products had no adverse effects on the used cell cultures, which is important with regard to therapeutic approaches. In summary, our data showed that endosialidases provide highly specific tools for the directed degradation of polySia-based scaffolds.

Chapter 9 – References

- Aalto, J., Pelkonen, S., Kalimo, H., and Finne, J. (2001) Mutant Bacteriophage With Non-Catalytic Endosialidase Binds to Both Bacterial and Eukaryotic Polysialic Acid and Can Be Used As Probe for Its Detection. *Glycoconj. J* **18**(10):751-758.
- Ackermann, Hans-Wolfgang, Department of Microbiology, Faculty of Medicine, Felix d'Herelle Reference Center for Bacterial Viruses, Laval University, Québec; Canada.
- Ackermann, H.-W. (1998) Tailed Bacteriophages: the Order Caudovirales. *Advances in Virus Research* **51**:135-201.
- Ackermann, H.-W. (2006) Classification of Bacteriophages (Chapter 2). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 8-16, Oxford University Press, Inc., New York, USA.
- Ackermann, H.-W. (2007) 5,500 Phages Examined in the Electron Microscope. *Arch. Virol.* **152**(2):227-243.
- Ali, S.A., Iwabuchi, N., Matsui, T., Hirota, K., Kidokoro, S., Arai, M., Kuwajima, K., Schuck, P., and Arisaka, F. (2003) Reversible and Fast Association Equilibria of a Molecular Chaperone, Gp57A, of Bacteriophage T4. *Biophysical Journal* **85**(4):2606-2618.
- Amaya, M.F., Watts, A.G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A., Paris, G., Frasch, A.C., Withers, S.G., and Alzari, P.M. (2004) Structural Insights into the Catalytic Mechanism of Trypanosoma Cruzi Trans-Sialidase. *Structure* **12**(5):775-784.
- Barry, G.T. and Goebel, W.F. (1957) Colominic Acid, a Substance of Bacterial Origin Related to Sialic Acid. *Nature* **179**(4552):206.
- Becker, C.G., Artola, A., Gerardy-Schahn, R., Becker, T., Welzl, H., and Schachner, M. (1996) The Polysialic Acid Modification of the Neural Cell Adhesion Molecule Is Involved in Spatial Learning and Hippocampal Long-Term Potentiation. *J Neurosci. Res.* **45**(2):143-152.
- Bhattacharjee, A.K., Jennings, H.J., Kenny, C.P., Martin, A., and Smith, I.C. (1975) Structural Determination of the Sialic Acid Polysaccharide Antigens of Neisseria Meningitidis Serogroups B and C With Carbon 13 Nuclear Magnetic Resonance. *Journal of Biological Chemistry* **250**(5):1926-1932.
- Botstein, D. (1980) A Theory of Modular Evolution for Bacteriophages. *Ann. N. Y. Acad. Sci* **354**:484-490.
- Bradley, D.E. (1967) Ultrastructure of Bacteriophage and Bacteriocins. *Bacteriol. Rev.* **31**(4):230-314.
- Brüssow, H. and Hendrix, R.W. (2002) Phage Genomics: Small Is Beautiful. *Cell* **108**(1):13-16.
- Brüssow, H., Canchaya, C., and Hardt, W.D. (2004) Phages and the Evolution of Bacterial Pathogens: From Genomic Rearrangements to Lysogenic Conversion. *Microbiol. Mol. Biol. Rev.* **68**(3):560-602.
- Brüssow, H. (2006) Prophage Genomics (Chapter 3). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 17-25, Oxford University Press, Inc., New York, USA.
- Brüssow, H. and Desiere, F. (2006) Evolution of Tailed Phages: Insights from Comparative Phage Genomics (Chapter 4). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 26-36, Oxford University Press, Inc., New York, USA.
- Burda, M.R. and Miller, S. (1999) Folding of Coliphage T4 Short Tail Fiber in Vitro. Analysing the Role of a Bacteriophage-Encoded Chaperone. *European Journal of Biochemistry / FEBS* **265**(2):771-778.
- Burgess, A., Weng, Y.Q., Ypsilanti, A.R., Cui, X., Caines, G., and Aubert, I. (2007) Polysialic Acid Limits Septal Neurite Outgrowth on Laminin. *Brain Research* **1144**:52-58.
- Cabezas, J.A. (1991) Some Questions and Suggestions on the Type References of the Official Nomenclature (IUB) for Sialidase(s) and Endosialidase. *Biochem J* **278**(Pt 1):311-312.
- Campbell, A. (2003) The Future of Bacteriophage Biology. *Nat. Rev. Genet.* **4**(6):471-477.
- Canchaya, C., Fournous, G., and Brussow, H. (2004) The Impact of Prophages on Bacterial Chromosomes. *Molecular Microbiology* **53**(1):9-18.
- Casjens, S.R. (2005) Comparative Genomics and Evolution of the Tailed-Bacteriophages. *Current Opinion in Microbiology* **8**(4):451-458.
- Cerritelli, M.E., Cheng, N., Rosenberg, A.H., McPherson, C.E., Booy, F.P., and Steven, A.C. (1997) Encapsidated Conformation of Bacteriophage T7 DNA. *Cell* **91**(2):271-280.

- Chappell, J.D., Gunn, V.L., Wetzel, J.D., Baer, G.S., and Dermody, T.S. (1997) Mutations in Type 3 Reovirus That Determine Binding to Sialic Acid Are Contained in the Fibrous Tail Domain of Viral Attachment Protein Sigma1. *Journal of Virology* **71**(3):1834-1841.
- Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.L., and Brüssow, H. (2004) Phage-Host Interaction: an Ecological Perspective. *Journal of Bacteriology* **186**(12):3677-3686.
- Chong, A.K., Pegg, M.S., Taylor, N.R., and von Itzstein, M. (1992) Evidence for a Sialosyl Cation Transition-State Complex in the Reaction of Sialidase From Influenza Virus. *Eur. J. Biochem.* **207**(1):335-343.
- Clarke, B.R., Esumeh, F., and Roberts, I.S. (2000) Cloning, Expression, and Purification of the K5 Capsular Polysaccharide Lyase (KflA) From Coliphage K5A: Evidence for Two Distinct K5 Lyase Enzymes. *Journal of Bacteriology* **182**(13):3761-3766.
- Conchonaud, F., Nicolas, S., Amoureux, M.C., Menager, C., Marguet, D., Lenne, P.F., Rougon, G., and Matarazzo, V. (2007) Polysialylation Increases Lateral Diffusion of Neural Cell Adhesion Molecule in the Cell Membrane. *Journal of Biological Chemistry* **282**(36):26266-26274.
- Conway, J.F., Duda, R.L., Cheng, N., Hendrix, R.W., and Steven, A.C. (1995) Proteolytic and Conformational Control of Virus Capsid Maturation: the Bacteriophage HK97 System. *Journal of Molecular Biology* **253**(1):86-99.
- Cross, A.S. (1990) The Biologic Significance of Bacterial Encapsulation. *Curr. Top. Microbiol. Immunol.* **150**:87-95.
- Cunningham, E.L., Jaswal, S.S., Sohl, J.L., and Agard, D.A. (1999) Kinetic Stability as a Mechanism for Protease Longevity. *Proceedings of the National Academy of Sciences of the United States of America* **96**(20):11008-11014.
- Curreli, S., Arany, Z., Gerardy-Schahn, R., Mann, D., and Stamatou, N.M. (2007) Polysialylated Neuropilin-2 Is Expressed on the Surface of Human Dendritic Cells and Modulates Dendritic Cell-T Lymphocyte Interactions. *Journal of Biological Chemistry* **282**(42):30346-30356.
- d'Herelle, F. (1917) Sur Un Microbe Invisible Antagoniste Des Bacilles Dysentériques. *Comptes Rendus De L'Académie Des Sciences (Paris)* **165**(10):373-375.
- Daniel, L., Durbec, P., Gautherot, E., Rouvier, E., Rougon, G., and Figarella-Branger, D. (2001) A Nude Mice Model of Human Rhabdomyosarcoma Lung Metastases for Evaluating the Role of Polysialic Acids in the Metastatic Process. *Oncogene* **20**(8):997-1004.
- Desiere, F., McShan, W.M., van Sinderen, D., Ferretti, J.J., and Brüssow, H. (2001) Comparative Genomics Reveals Close Genetic Relationships Between Phages From Dairy Bacteria and Pathogenic Streptococci: Evolutionary Implications for Prophage-Host Interactions. *Virology* **288**(2):325-341.
- Deszo, E.L., Steenbergen, S.M., Freedberg, D.I., and Vimr, E.R. (2005) Escherichia Coli K1 Polysialic Acid O-Acetyltransferase Gene, NeuO, and the Mechanism of Capsule Form Variation Involving a Mobile Contingency Locus. *Proc. Natl. Acad. Sci. U. S. A* **102**(15):5564-5569.
- Dickson, R.C. (1973) Assembly of Bacteriophage T4 Tail Fibers. IV. Subunit Composition of Tail Fibers and Fiber Precursors. *Journal of Molecular Biology* **79**(4):633-647.
- Die Fantastischen Vier (2004) TROY. Composers: Michael DJ Beck, Thomas Dürr, Andreas Rieke, Michael B.Schmidt, Joachim Witt, and Thomas Burchia.
- Dityatev, A., Dityateva, G., and Schachner, M. (2000) Synaptic Strength As a Function of Post- Versus Presynaptic Expression of the Neural Cell Adhesion Molecule NCAM. *Neuron* **26**(1):207-217.
- Dityatev, A., Dityateva, G., Sytnyk, V., Dellinger, M., Toni, N., Nikonenko, I., Müller, D., and Schachner, M. (2004) Polysialylated Neural Cell Adhesion Molecule Promotes Remodeling and Formation of Hippocampal Synapses. *J. Neurosci.* **24**(42):9372-9382.
- Duda, R.L., Hendrix, R.W., Huang, W.M., and Conway, J.F. (2006) Shared Architecture of Bacteriophage SPO1 and Herpesvirus Capsids. *Curr. Biol.* **16**(1):R11-R13.
- Durbec, P. and Cremer, H. (2001) Revisiting the Function of PSA-NCAM in the Nervous System. *Mol. Neurobiol.* **24**(1-3):53-64.
- Feucht, A., Schmid, A., Benz, R., Schwarz, H., and Heller, K.J. (1990) Pore Formation Associated With the Tail-Tip Protein Pb2 of Bacteriophage T5. *Journal of Biological Chemistry* **265**(30):18561-18567.
- Finne, J., Finne, U., Deagostini-Bazin, H., and Goridis, C. (1983) Occurrence of Alpha 2-8 Linked Polysialosyl Units in a Neural Cell Adhesion Molecule. *Biochem. Biophys. Res. Commun.* **112**(2):482-487.

- Frank, S., Kammerer, R.A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bachinger, H.P., and Engel, J. (2001) Stabilization of Short Collagen-Like Triple Helices by Protein Engineering. *Journal of Molecular Biology* **308**(5):1081-1089.
- Freiberg, A., Morona, R., van den Bosch, L., Jung, C., Behlke, J., Carlin, N., Seckler, R., and Baxa, U. (2003) The Tailspike Protein of Shigella Phage Sf6. A Structural Homolog of Salmonella Phage P22 Tailspike Protein Without Sequence Similarity in the Beta-Helix Domain. *J Biol. Chem.* **278**(3):1542-1548.
- Freiberger, F., Claus, H., Gunzel, A., Oltmann-Norden, I., Vionnet, J., Mühlenhoff, M., Vogel, U., Vann, W.F., Gerardy-Schahn, R., and Stummeyer, K. (2007) Biochemical Characterization of a Neisseria Meningitidis Polysialyltransferase Reveals Novel Functional Motifs in Bacterial Sialyltransferases. *Mol. Microbiol.* **65**(5):1258-1275.
- Frosch, M., Roberts, I., Gorgen, I., Metzger, S., Boulnois, G.J., and Bitter-Suermann, D. (1987) Serotyping and Genotyping of Encapsulated Escherichia Coli K1 Sepsis Isolates With a Monoclonal IgG Anti K1 Antibody and K1 Gene Probes. *Microb. Pathog.* **2**(5):319-326.
- Furowicz, A.J. and Orskov, F. (1972) Two New Escherichia Coli O Antigens, O150 and O157, and One New K Antigen, K92, in Strains Isolated From Veterinary Diseases. *Acta Pathol. Microbiol. Scand. [B] Microbiol. Immunol.* **80**(3):441-444.
- Gage, M.J. and Robinson, A.S. (2003) C-Terminal Hydrophobic Interactions Play a Critical Role in Oligomeric Assembly of the P22 Tailspike Trimer. *Protein Sci.* **12**(12):2732-2747.
- Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S., Lottspeich, F., and Frosch, M. (1995) Molecular Cloning and Functional Expression of Bacteriophage PK1E-Encoded Endoneuraminidase Endo NE. *Molecular Microbiology* **16**(3):441-450.
- Gerardy-Schahn, R., Behrens, P., Scheper, T., Grothe, C., Schuster, R.H., Kirschning, A., and Wilk, P. (2004) *Polysialic Acid: Towards the Evaluation of a New, Bio-Identical Scaffold Material*. Antrag zur Einrichtung einer Forschergruppe bei der Deutschen Forschungsgemeinschaft.
- Gross, R.J., Cheasty, T., and Rowe, B. (1977) Isolation of Bacteriophages Specific for the K1 Polysaccharide Antigen of Escherichia Coli. *J Clin. Microbiol.* **6**(6):548-550.
- Guihard, G., Boulanger, P., and Letellier, L. (1992) Involvement of Phage T5 Tail Proteins and Contact Sites Between the Outer and Inner Membrane of Escherichia Coli in Phage T5 DNA Injection. *J Biol. Chem.* **267**(5):3173-3178.
- Hallenbeck, P.C., Vimr, E.R., Yu, F., Bassler, B., and Troy, F.A. (1987) Purification and Properties of a Bacteriophage-Induced Endo-N-Acetylneuraminidase Specific for Poly-Alpha-2,8-Sialosyl Carbohydrate Units. *J Biol. Chem.* **262**(8):3553-3561.
- Hambly, E. and Suttle, C.A. (2005) The Viriosphere, Diversity, and Genetic Exchange Within Phage Communities. *Curr. Opin. Microbiol.* **8**(4):444-450.
- Heller, K. and Braun, V. (1979) Accelerated Adsorption of Bacteriophage T5 to Escherichia Coli F, Resulting From Reversible Tail Fiber-Lipopolysaccharide Binding. *Journal of Bacteriology* **139**(1):32-38.
- Heller, K.J. (1984) Identification of the Phage Gene for Host Receptor Specificity by Analyzing Hybrid Phages of T5 and BF23. *Virology* **139**(1):11-21.
- Heller, K.J. and Schwarz, H. (1985) Irreversible Binding to the Receptor of Bacteriophages T5 and BF23 Does Not Occur With the Tip of the Tail. *Journal of Bacteriology* **162**(2):621-625.
- Hendrix, R.W., Lawrence, J.G., Hatfull, G.F., and Casjens, S. (2000) The Origins and Ongoing Evolution of Viruses. *Trends Microbiol.* **8**(11):504-508.
- Hendrix, R.W. (2002) Bacteriophages: Evolution of the Majority. *Theoretical Population Biology* **61**(4):471-480.
- Horn, S.J., Sikorski, P., Cederkvist, J.B., Vaaje-Kolstad, G., Sorlie, M., Synstad, B., Vriend, G., Varum, K.M., and Eijsink, V.G.H. (2006) Costs and Benefits of Processivity in Enzymatic Degradation of Recalcitrant Polysaccharides. *Proceedings of the National Academy of Sciences* **103**(48):18089-18094.
- Hughes, K.A., Sutherland, I.W., and Jones, M.V. (1998) Biofilm Susceptibility to Bacteriophage Attack: the Role of Phage-Borne Polysaccharide Depolymerase. *Microbiology* **144** (Pt 11):3039-3047.
- Ibarra, B., Valpuesta, J.M., and Carrascosa, J.L. (2001) Purification and Functional Characterization of P16, the ATPase of the Bacteriophage Phi29 Packaging Machinery. *Nucleic Acids Res.* **29**(21):4264-4273.

- Jakobsson, E., Jokilammi, A., Aalto, J., Ollikka, P., Lehtonen, J.V., Hirvonen, H., and Finne, J. (2007) Identification of Amino Acid Residues at the Active Site of Endosialidase That Dissociate the Polysialic Acid Binding and Cleaving Activities in Escherichia Coli K1 Bacteriophages. *Biochem. J.* **405**(3):465-472.
- Jardine, P.J. and Anderson, D.L. (2006) DNA Packaging in Double-Stranded DNA Phages (Chapter 6). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 49-65, Oxford University Press, Inc., New York, USA.
- Jennings, H.J., Roy, R., and Michon, F. (1985) Determinant Specificities of the Groups B and C Polysaccharides of Neisseria Meningitidis. *J. Immunol.* **134**(4):2651-2657.
- Jódar, L., Feavers, I.M., Salisbury, D., and Granoff, D.M. (2002) Development of Vaccines Against Meningococcal Disease. *Lancet* **359**(9316):1499-1508.
- Jokilammi, A., Ollikka, P., Korja, M., Jakobsson, E., Loimaranta, V., Haataja, S., Hirvonen, H., and Finne, J. (2004) Construction of Antibody Mimics From a Noncatalytic Enzyme-Detection of Polysialic Acid. *J Immunol. Methods* **295**(1-2):149-160.
- Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F., and Hendrix, R.W. (2000) Genomic Sequences of Bacteriophages HK97 and HK022: Pervasive Genetic Mosaicism in the Lambdoid Bacteriophages. *Journal of Molecular Biology* **299**(1):27-51.
- Kanamaru, S., Leiman, P.G., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., Arisaka, F., and Rossmann, M.G. (2002) Structure of the Cell-Puncturing Device of Bacteriophage T4. *Nature* **415**(6871):553-557.
- Kataoka, Y., Miyake, K., and Iijima, S. (2006) Coliphage Derived Sialidase Preferentially Recognizes Nonreducing End of Polysialic Acid. *J Biosci. Bioeng.* **101**(2):198-201.
- Katouno, F., Taguchi, M., Sakurai, K., Uchiyama, T., Nikaidou, N., Nonaka, T., Sugiyama, J., and Watanabe, T. (2004) Importance of Exposed Aromatic Residues in Chitinase B From *Serratia Marcescens* 2170 for Crystalline Chitin Hydrolysis. *J. Biochem. (Tokyo)* **136**(2):163-168.
- Kleene, R. and Schachner, M. (2004) Glycans and Neural Cell Interactions. *Nat. Rev. Neurosci.* **5**(3):195-208.
- Koivula, A., Kinnari, T., Harjunpaa, V., Ruohonen, L., Teleman, A., Drakenberg, T., Rouvinen, J., Jones, T.A., and Teeri, T.T. (1998) Tryptophan 272: an Essential Determinant of Crystalline Cellulose Degradation by *Trichoderma Reesei* Cellobiohydrolase Cel6A. *FEBS Lett.* **429**(3):341-346.
- Kreisberg, J.F., Betts, S.D., Haase-Pettingell, C., and King, J. (2002) The Interdigitated Beta-Helix Domain of the P22 Tailspike Protein Acts As a Molecular Clamp in Trimer Stabilization. *Protein Sci.* **11**(4):820-830.
- Kwiatkowski, B., Boschek, B., Thiele, H., and Stirm, S. (1982) Endo-N-Acetylneuraminidase Associated With Bacteriophage Particles. *J Virol.* **43**(2):697-704.
- Kwiatkowski, B., Boschek, B., Thiele, H., and Stirm, S. (1983) Substrate Specificity of Two Bacteriophage-Associated Endo-N-Acetylneuraminidases. *J Virol.* **45**(1):367-374.
- Laemmli, U.K. (1970) Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4. *Nature* **227**(5259):680-685.
- Lawrence, J.G., Hatfull, G.F., and Hendrix, R.W. (2002) Imbroglios of Viral Taxonomy: Genetic Exchange and Failings of Phenetic Approaches. *J. Bacteriol.* **184**(17):4891-4905.
- Leggate, D.R., Bryant, J.M., Redpath, M.B., Head, D., Taylor, P.W., and Luzio, J.P. (2002) Expression, Mutagenesis and Kinetic Analysis of Recombinant K1E Endosialidase to Define the Site of Proteolytic Processing and Requirements for Catalysis. *Mol. Microbiol.* **44**(3):749-760.
- 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 Salome, M. (1996) N-Acetyl-Heparosan Lyase of Escherichia Coli K5: Gene Cloning and Expression. *Journal of Bacteriology* **178**(24):7260-7264.
- Leiman, Petr G., Department of Biological Sciences, Purdue University, West Lafayette, IN; USA.
- Leiman, P.G., Kanamaru, S., Mesyanzhinov, V.V., Arisaka, F., and Rossmann, M.G. (2003) Structure and Morphogenesis of Bacteriophage T4. *Cell Mol. Life Sci* **60**(11):2356-2370.
- Leiman, P.G., Battisti, A.J., Bowman, V.D., Stummeyer, K., Mühlenhoff, M., Gerardy-Schahn, R., Scholl, D., and Molineux, I.J. (2007) The Structures of Bacteriophages K1E and K1-5 Explain Processive Degradation of Polysaccharide Capsules and Evolution of New Host Specificities. *J. Mol. Biol.*

- Lindberg, A.A. (1977) Bacterial Surface Carbohydrates and Bacteriophage Adsorption (Sutherland, I.W., ed) pp. 289-356, Academia Press, London; United Kingdom.
- Long, G.S., Bryant, J.M., Taylor, P.W., and Luzio, J.P. (1995) Complete Nucleotide Sequence of the Gene Encoding Bacteriophage E Endosialidase: Implications for K1E Endosialidase Structure and Function. *Biochem J* **309**(Pt 2):543-550.
- Machida, Y., Hattori, K., Miyake, K., Kawase, Y., Kawase, M., and Iijima, S. (2000a) Molecular Cloning and Characterization of a Novel Bacteriophage-Associated Sialidase. *Journal of Bioscience and Bioengineering* **90**(1):62-68.
- Machida, Y., Miyake, K., Hattori, K., Yamamoto, S., Kawase, M., and Iijima, S. (2000b) Structure and Function of a Novel Coliphage-Associated Sialidase. *FEMS Microbiol. Lett.* **182**(2):333-337.
- Manning, M. and Colón, W. (2004) Structural Basis of Protein Kinetic Stability: Resistance to Sodium Dodecyl Sulfate Suggests a Central Role for Rigidity and a Bias Toward Beta-Sheet Structure. *Biochemistry* **43**(35):11248-11254.
- Marvik, O.J., Jacobsen, E., Dokland, T., and Lindqvist, B.H. (1994) Bacteriophage P2 and P4 Morphogenesis: Assembly Precedes Proteolytic Processing of the Capsid Proteins. *Virology* **205**(1):51-65.
- McGuire, E.J. and Binkley, S.B. (1964) The Structure and Chemistry of Colominic Acid. *Biochemistry* **3**:247-251.
- Meijer, W.J., Horcajadas, J.A., and Salas, M. (2001) Phi29 Family of Phages. *Microbiology and Molecular Biology Reviews : MMBR* **65**(2):261-287.
- Mesyanzhinov, V.V., Leiman, P.G., Kostyuchenko, V.A., Kurochkina, L.P., Miroshnikov, K.A., Sykilinda, N.N., and Shneider, M.M. (2004) Molecular Architecture of Bacteriophage T4. *Biochemistry (Mosc.)* **69**(11):1190-1202.
- Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Ruger, W. (2003) Bacteriophage T4 Genome. *Microbiol. Mol. Biol. Rev.* **67**(1):86-156, table.
- Miroshnikov, K.A., Marusich, E.I., Cerritelli, M.E., Cheng, N., Hyde, C.C., Steven, A.C., and Mesyanzhinov, V.V. (1998) Engineering Trimeric Fibrous Proteins Based on Bacteriophage T4 Adhesins. *Protein Engineering* **11**(4):329-332.
- Miyake, K., Muraki, T., Hattori, K., Machida, Y., Watanabe, M., Kawase, M., Yoshida, Y., and Iijima, S. (1997) Screening of Bacteriophages Producing Endo-N-Acetylneuraminidase. *Journal of Fermentation and Bioengineering* **84**(1):90-93.
- Moebius, J., Widera, D., Schmitz, J., Kaltschmidt, C., and Piechaczek, C. (2007) Impact of Polysialylated CD56 on Natural Killer Cell Cytotoxicity. *BMC Immunology* **8**(1):13.
- Molineux, I. (2006) The T7 Group (Chapter 20). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 277-301, Oxford University Press, Inc., New York, USA.
- Mosig, G. and Eiserling, F. (2006) T4 and Related Phages: Structure and Development (Chapter 18). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 225-267, Oxford University Press, Inc., New York, USA.
- Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M., and Gerardy-Schahn, R. (2003) Proteolytic Processing and Oligomerization of Bacteriophage-Derived Endosialidases. *Journal of Biological Chemistry* **278**(15):12634-12644.
- Mushtaq, N., Redpath, M.B., Luzio, J.P., and Taylor, P.W. (2005) Treatment of Experimental Escherichia Coli Infection With Recombinant Bacteriophage-Derived Capsule Depolymerase. *Journal of Antimicrobial Chemotherapy* **56**(1):160-165.
- Oltmann-Norden, I., Galuska, S.P., Hildebrandt, H., Geyer, R., Gerardy-Schahn, R., Geyer, H., and Mühlenhoff, M. (2007) Impact of the Polysialyltransferases ST8SiaII and ST8SiaIV on Polysialic Acid Synthesis During Postnatal Mouse Brain Development. *Journal of Biological Chemistry* :, in press.
- Pajunen, M., Kiljunen, S., and Skurnik, M. (2000) Bacteriophage PhiYeO3-12, Specific for Yersinia Enterocolitica Serotype O:3, Is Related to Coliphages T3 and T7. *Journal of Bacteriology* **182**(18):5114-5120.
- Papanikolopoulou, K., Forge, V., Goeltz, P., and Mitraki, A. (2004) Formation of Highly Stable Chimeric Trimers by Fusion of an Adenovirus Fiber Shaft Fragment With the Foldon Domain of Bacteriophage T4 Fibrin. *Journal of Biological Chemistry* **279**(10):8991-8998.

- Pelkonen, S., Pelkonen, J., and Finne, J. (1989) Common Cleavage Pattern of Polysialic Acid by Bacteriophage Endosialidases of Different Properties and Origins. *J Virol.* **63**(10):4409-4416.
- Petter, J.G. and Vimr, E.R. (1993) Complete Nucleotide Sequence of the Bacteriophage K1F Tail Gene Encoding Endo-N-Acylneuraminidase (Endo-N) and Comparison to an Endo-N Homolog in Bacteriophage PK1E. *J Bacteriol.* **175**(14):4354-4363.
- Prasadarao, N.V., Wass, C.A., Weiser, J.N., Stins, M.F., Huang, S.H., and Kim, K.S. (1996) Outer Membrane Protein A of Escherichia Coli Contributes to Invasion of Brain Microvascular Endothelial Cells. *Infect. Immun.* **64**(1):146-153.
- Rossmann, M.G., Mesyanzhinov, V.V., Arisaka, F., and Leiman, P.G. (2004) The Bacteriophage T4 DNA Injection Machine. *Curr. Opin. Struct. Biol.* **14**(2):171-180.
- Roth, J., Zuber, C., Wagner, P., Taatjes, D.J., Weisgerber, C., Heitz, P.U., Goridis, C., and Bitter-Suermann, D. (1988) Reexpression of Poly(Sialic Acid) Units of the Neural Cell Adhesion Molecule in Wilms Tumor. *Proc. Natl. Acad. Sci. U. S. A* **85**(9):2999-3003.
- Rutishauser, U., Watanabe, M., Silver, J., Troy, F.A., and Vimr, E.R. (1985) Specific Alteration of NCAM-Mediated Cell Adhesion by an Endoneuraminidase. *J Cell Biol.* **101**(5 Pt 1):1842-1849.
- Rutishauser, U. and Landmesser, L. (1996) Polysialic Acid in the Vertebrate Nervous System: a Promoter of Plasticity in Cell-Cell Interactions. *Trends Neurosci.* **19**(10):422-427.
- Sayers, J.R. (2006) Bacteriophage T5 (Chapter 19). In *The Bacteriophages*, 2nd edition (Calendar, R. and Abedon, S.T., ed) pp. 268-276, Oxford University Press, Inc., New York, USA.
- Scholl, D., Rogers, S., Adhya, S., and Merrill, C.R. (2001) Bacteriophage K1-5 Encodes Two Different Tail Fiber Proteins, Allowing It to Infect and Replicate on Both K1 and K5 Strains of Escherichia Coli. *Journal of Virology* **75**(6):2509-2515.
- Scholl, D., Adhya, S., and Merrill, C. (2005) Escherichia Coli K1's Capsule Is a Barrier to Bacteriophage T7. *Applied and Environmental Microbiology* **71**(8):4872-4874.
- Smith, H.W. and Huggins, M.B. (1982) Successful Treatment of Experimental Escherichia Coli Infections in Mice Using Phage: Its General Superiority Over Antibiotics. *J Gen. Microbiol.* **128**(2):307-318.
- Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Crystal Structure of P22 Tailspike Protein: Interdigitated Subunits in a Thermostable Trimer. *Science* **265**(5170):383-386.
- Steven, A.C., Trus, B.L., Maizel, J.V., Unser, M., Parry, D.A., Wall, J.S., Hainfeld, J.F., and Studier, F.W. (1988) Molecular Substructure of a Viral Receptor-Recognition Protein. The Gp17 Tail-Fiber of Bacteriophage T7. *J. Mol. Biol.* **200**(2):351-365.
- Stirm, S. and Freund-Mölbert, E. (1971) Escherichia Coli Capsule Bacteriophages. II. Morphology. *Journal of Virology* **8**(3):330-342.
- Stummeyer, K., Dickmanns, A., Mühlhoff, M., Gerardy-Schahn, R., and Ficner, R. (2005) Crystal Structure of the Polysialic Acid-Degrading Endosialidase of Bacteriophage K1F. *Nature Structural & Molecular Biology* **12**(1):90-96.
- Sutherland, I.W. (1977) Enzymes Acting on Bacterial Surface Carbohydrates. (In *Surface Carbohydrates of the Prokaryotic Cell*, Sutherland, I.W., ed) pp. 209-245, Academic Press, Inc., New York, USA.
- Tao, Y., Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M.G. (1997) Structure of Bacteriophage T4 Fibrin: a Segmented Coiled Coil and the Role of the C-Terminal Domain. *Structure* **5**(6):789-798.
- Tao, Y., Olson, N.H., Xu, W., Anderson, D.L., Rossmann, M.G., and Baker, T.S. (1998) Assembly of a Tailed Bacterial Virus and Its Genome Release Studied in Three Dimensions. *Cell* **95**(3):431-437.
- Taylor, G. (1996) Sialidases: Structures, Biological Significance and Therapeutic Potential. *Curr. Opin. Struct. Biol.* **6**(6):830-837.
- Taylor, C.M. and Roberts, I.S. (2005) Capsular Polysaccharides and Their Role in Virulence. *Contrib. Microbiol.* **12**:55-66.
- Tinsley, C.R., Bille, E., and Nassif, X. (2006) Bacteriophages and Pathogenicity: More than Just Providing a Toxin? *Microbes and Infection* **8**(5):1365-1371.
- Tomlinson, S. and Taylor, P.W. (1985) Neuraminidase Associated With Coliphage E That Specifically Depolymerizes the Escherichia Coli K1 Capsular Polysaccharide. *J Virol.* **55**(2):374-378.
- Twort, F.W. (1915) An Investigation on the Nature of Ultramicroscopic Viruses. *Lancet* **189**(2):1241-1243.

- van Raaij, M.J., Schoehn, G., Burda, M.R., and Miller, S. (2001) Crystal Structure of a Heat and Protease-Stable Part of the Bacteriophage T4 Short Tail Fibre. *J Mol. Biol.* **314**(5):1137-1146.
- Vimr, E.R., McCoy, R.D., Vollger, H.F., Wilkison, N.C., and Troy, F.A. (1984) Use of Prokaryotic-Derived Probes to Identify Poly(Sialic Acid) in Neonatal Neuronal Membranes. *Proc. Natl. Acad. Sci. U. S. A* **81**(7):1971-1975.
- Wang, S., Chandramouli, P., Butcher, S., and Dokland, T. (2003) Cleavage Leads to Expansion of Bacteriophage P4 Procapsids in Vitro. *Virology* **314**(1):1-8.
- Watanabe, T., Ariga, Y., Sato, U., Toratani, T., Hashimoto, M., Nikaidou, N., Kezuka, Y., Nonaka, T., and Sugiyama, J. (2003) Aromatic Residues Within the Substrate-Binding Cleft of *Bacillus Circulans* Chitinase A1 Are Essential for Hydrolysis of Crystalline Chitin. *Biochem. J.* **376**(Pt 1):237-244.
- Watts, A.G., Damager, I., Amaya, M.L., Buschiazzi, A., Alzari, P., Frasch, A.C., and Withers, S.G. (2003) Trypanosoma Cruzi Trans-Sialidase Operates Through a Covalent Sialyl-Enzyme Intermediate: Tyrosine Is the Catalytic Nucleophile. *J. Am. Chem. Soc.* **125**(25):7532-7533.
- Weigele, P.R., Scanlon, E., and King, J. (2003) Homotrimeric, Beta-Stranded Viral Adhesins and Tail Proteins. *Journal of Bacteriology* **185**(14):4022-4030.
- Weigele, P.R., Haase-Pettingell, C., Campbell, P.G., Gossard, D.C., and King, J. (2005) Stalled Folding Mutants in the Triple Beta-Helix Domain of the Phage P22 Tailspike Adhesin. *J Mol. Biol.* **354**(5):1103-1117.
- Weinhold, B., Seidenfaden, R., Rockle, I., Mühlhoff, M., Schertzinger, F., Conzelmann, S., Marth, J.D., Gerardy-Schahn, R., and Hildebrandt, H. (2005) Genetic Ablation of Polysialic Acid Causes Severe Neurodevelopmental Defects Rescued by Deletion of the Neural Cell Adhesion Molecule. *J Biol. Chem.* **280**(52):42971-42977.
- Yabe, U., Sato, C., Matsuda, T., and Kitajima, K. (2003) Polysialic Acid in Human Milk. CD36 Is a New Member of Mammalian Polysialic Acid-Containing Glycoprotein. *Journal of Biological Chemistry* **278**(16):13875-13880.
- Yang, X., Lee, J., Mahony, E.M., Kwong, P.D., Wyatt, R., and Sodroski, J. (2002) Highly Stable Trimers Formed by Human Immunodeficiency Virus Type 1 Envelope Glycoproteins Fused With the Trimeric Motif of T4 Bacteriophage Fibrin. *Journal of Virology* **76**(9):4634-4642.
- Zhang, S., Irwin, D.C., and Wilson, D.B. (2000) Site-Directed Mutation of Noncatalytic Residues of *Thermobifida Fusca* Exocellulase Cel6B. *Eur. J. Biochem.* **267**(11):3101-3115.
- Zuber, C., Lackie, P.M., Catterall, W.A., and Roth, J. (1992) Polysialic Acid Is Associated With Sodium Channels and the Neural Cell Adhesion Molecule N-CAM in Adult Rat Brain. *Journal of Biological Chemistry* **267**(14):9965-9971.

Appendix 1 – Abbreviations

aa	Amino acid
Ab	Antibody
<i>B. pumilus</i>	<i>Bacillus pumilus</i>
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
bp	Base pairs
BSA	Bovine serum albumin
CD	Circular dichroism
C-terminal	Carboxy-terminal
CTD	C-terminal domain
Da	Dalton; 1 Da = 1 g/mol = 1,6605402*10 ⁻²⁷ kg
ΔC	C-terminal truncation (lacking the C-terminal domain)
ddH ₂ O	Di-distilled water
DNA	Deoxyribonucleic acid
ΔN-EndoNF	Endosialidase of coliphage K1F lacking 245 N-terminal amino acids
ΔN-EndoNF*	Catalytically active N-terminal part of ΔN-EndoNF released by proteolytic cleavage
ds	Double stranded
<i>E. coli</i>	<i>Escherichia coli</i>
ElmA	Eliminase of an <i>Escherichia coli</i> strain
EndoNF	Endosialidase of coliphage K1F (precursor)
gp	Gene product
kb	Kilobase pairs
kDa	Kilodalton
Kdn	2-keto-3-deoxynonulosonic acid
LTF	L-shaped tail fibre protein
mAb	Monoclonal antibody
β-ME	β-Mercaptoethanol
NCAM	Neural cell adhesion molecule
NCBI	National Center for Biotechnology Information
n.d.	Not determined
Neu5Ac	5-N-acetyl-neuraminic acid
Neu5Gc	5-N-glycolyl-neuraminic acid
nt	Nucleotide
N-terminal	Amino-terminal
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
polySia	Polysialic acid
RNA	Ribonucleic acid
SDS	Sodium dodecyl-sulphate
U	Unit

Peer Reviewed Publications

- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff (2007) Characterization of a Novel Intra-molecular Chaperone Domain Conserved in Endosialidases and Other Bacteriophage Tail Spike and Fiber Proteins. *J. Biol. Chem.* **282**(5):2821-2831.
- Katharina Stummeyer, **David Schwarzer**, Heike Claus, Ulrich Vogel, Rita Gerardy-Schahn, and Martina Mühlenhoff (2006) Evolution of bacteriophages infecting encapsulated bacteria: Lessons from *Escherichia coli* K1 specific phages. *Mol. Micro.* **60**(5):1123-1135.

Submitted Manuscript

- Silke Berski, Jeroen van Bergeijk, **David Schwarzer**, Claudia Grothe, Rita Gerardy-Schahn, Andreas Kirschning, and Gerald Dräger: ‘Synthesis and Biological Evaluation of a new Polysialic Acid Hydrogel as Enzymatically Degradable Scaffold Material for Tissue Engineering’, submitted to *Biomacromolecules* on March 31st 2008

Manuscript in Preparation

- **David Schwarzer**, Katharina Stummeyer, Thomas Haselhorst, Bastian Rode, Thomas Scheper, Mark von Itzstein, Martina Mühlenhoff, and Rita Gerardy-Schahn: ‘Proteolytic Maturation of Endosialidase F Is Essential to Allow Efficient Binding to Polysialic Acid’

Oral Presentations

- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: „Characterisation of a novel intra-molecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fibre proteins”; Sunday, March 11th, 2007, at 3rd International meeting: Polysialic acid - Chemistry, Biology, Translational Aspects, March 10th – 13th, 2007, Bad Lauterberg, Germany
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: “The Proteolytic Cleavage of Endosialidases and Other Tailspike Proteins: A Key Step for Establishing Kinetically Stable Complexes?”; Monday, November 6th, 2007 at 17th Joint Meeting 2006 of the “Studiengruppe Glykobiologie der GBM”, the “Netherlands Society for Glycobiology” the “Groupe Lillois de Glycobiologie”, and the “Belgian Working Group for Glycosciences”, November 5th – 7th 2006, Brugge, Belgium
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: „Molecular characterization of the C-terminal domain of endosialidases“; Within the scope of the Basic Biological Questions (BBQ) seminar, September 30th, 2005, Medizinische Hochschule Hannover, Hannover, Germany
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: „Characterization of the C-terminal assembly domain of endosialidases“; Monday, September 5th, 2005, at XVIII. International Symposium on Glycoconjugates, September 4th – 9th, 2005, Florence, Italy

Poster Presentations

- **David Schwarzer**, Katharina Stummeyer, Martina Mühlenhoff, and Rita Gerardy-Schahn: “New Insights into PolySia Binding and Enzymatic Activity of Endosialidase F”, 18th Joint Meeting 2007 of the “Studiengruppe Glykobiologie der GBM”, the “Netherlands Society for Glycobiology” the “Groupe Lillois de Glycobiologie”, and the “Belgian Working Group for Glycosciences”, November 4th – 6th 2007, Lille, France
- Manuela Rollenhagen, **David Schwarzer**, Rita Gerardy-Schahn, and Martina Mühlenhoff: “Characterisation of the Novel K5-Lyase of *Escherichia coli* Phage K5B”, 18th Joint Meeting 2007 of the “Studiengruppe Glykobiologie der GBM”, the “Netherlands Society for Glycobiology” the “Groupe Lillois de Glycobiologie”, and the “Belgian Working Group for Glycosciences”, November 4th – 6th 2007, Lille, France
- **David Schwarzer**, Katharina Stummeyer, Martina Mühlenhoff, and Rita Gerardy-Schahn: “The Proteolytic Cleavage of Endosialidases and Other Tailspike Proteins: A Key Step for Establishing Kinetically Stable Complexes”, Glykostrukturfabrik - The 3rd Glycan Forum, February 22nd – 23rd, 2007, Berlin, Germany
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: “The C-terminal Assembly Module of Endosialidases and Other Tailspike Proteins: an Example for Divergent or Convergent Evolution?”, Glycobiology 2006; Annual Conference of the Society for Glycobiology, November 15th – 18th, 2006, Los Angeles, California, USA
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: „Characterisation of the C-terminal assembly domain of endosialidases and other tailspike proteins“, 16th Joint Meeting 2005 of the “Studiengruppe Glykobiologie der GBM”, the “Netherlands Society for Glycobiology” the “Groupe Lillois de Glycobiologie”, and the “Belgian Working Group for Glycosciences”, October 27th – 29th 2005, Hannover, Germany
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: “The C-terminal assembly domain of endosialidase NF is shared by several tailspike proteins of diverse bacteriophages“, Mosbacher Kolloquium 2005, March 31st – April 2nd, 2005 "Molecular Machines", Mosbach, Germany
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: “The endogenous assembly domain of endosialidases is also present in tailspike proteins of diverse bacteriophages”, Poster abstract for Glycobiology 2004 (US/Japan Glyco 2004), November 17th – 20th 2004, Honolulu, Hawaii, USA
- **David Schwarzer**, Katharina Stummeyer, Rita Gerardy-Schahn, and Martina Mühlenhoff: “Formation of active endosialidases depends on the presence of an endogenous chaperone domain”, Poster abstract for Sialobiology 2004, July 27th – 30th, 2004, St. Andrews, Scotland
- Martina Mühlenhoff, Katharina Stummeyer, **David Schwarzer**, and Rita Gerardy-Schahn: “Identification of the small subunit that stabilizes the endosialidase complex of bacteriophage K1E”, abstract for Sialobiology 2004, July 27th – 30th, 2004, St. Andrews, Scotland
- Martina Mühlenhoff, Katharina Stummeyer, Melanie Grove, **David Schwarzer**, and Rita Gerardy-Schahn: “Homo- and Heteromeric complex formation of endosialidases“, abstract for 14th Joint meeting on Glycobiology of the „Studiengruppe Glykobiologie der Gesellschaft für Biochemie und Molekularbiologie“, the „Netherlands Society for Glycobiology“ and the „Group Lillois de Glycobiology“, November 20th – 22nd, 2003, Lille, France

Appendix 3 – Danksagungen

Zunächst möchte ich mich ganz herzlich bei Prof. Dr. Rita Gerardy-Schahn für die Überlassung des interessanten Themas, die Übernahme des Referates, die erstklassigen Arbeitsvoraussetzungen und das Lesen der Manuskripte bedanken.

Ich danke Prof. Dr. Ralf Ficner für die Übernahme des Korreferates.

Mein besonderer Dank gilt Dr. Martina Mühlenhoff für die exzellente wissenschaftliche Begleitung während der letzten Jahre, ihre stete Hilfsbereitschaft, den Weitblick und das Lesen der Manuskripte.

Für die Aufklärung einer wundervollen Kristallstruktur, für viele produktive Anregungen und das Lesen der Manuskripte danke ich Dr. Katharina Stummeyer.

Melanie Grove, Andrea Bethe, Miriam Schiff und Jasmin Hotzy danke ich für Ihre professionelle Unterstützung im Labor.

Leider lässt es der Rahmen dieser Danksagungen nicht zu, das gesamte *Labteam* der Abteilung Zelluläre Chemie namentlich zu nennen, daher geht ein kollektiver Dank an das gesamte *Labteam* für die freundschaftliche und stete Hilfsbereitschaft, die unzähligen Kuchenschlachten, sowie für die Schaffung des positiven Arbeitsklimas.

Ebenso danke ich der Abteilung Zelluläre Chemie, dass mir die Teilnahme an internationalen Kongressen ermöglicht wurde, um meine Daten einem internationalen Fachpublikum präsentieren zu können.

Ich danke Dr. Thomas Haselhorst und Prof. Dr. Mark von Itzstein für die professionelle Zusammenarbeit.

Bei Prof. Dr. Thomas Scheper, Prof. Dr. Andreas Kirschning, Dr. Gerald Dräger, Silke Berski und Bastian Rode möchte ich mich ganz herzlich für die produktive Zusammenarbeit im Rahmen der Forschergruppe bedanken.

Für das Lesen des Manuskriptes der vorliegenden Arbeit danke ich Sebastian Seth, Dr. Birgit Weinhold, Tim Keys und Philipp Schmalhorst.

Prof. Dr. Hans-Wolfgang Ackermann und Dr. Petr Leiman danke ich für die Überlassung der elektronenmikroskopischen Aufnahmen.

Allen Freunden gilt mein Dank, die mich während der letzten Jahre stützend, aufbauend und mitreißend durch das Studium begleitet haben.

Tiefer Dank für das entgegengebrachte Vertrauen in mein Studium sowie für die Schaffung des dazu nötigen Fundaments gilt meinen Eltern, Großeltern und allen Verwandten.

Zum Schluss danke ich dem Glück und dem Zufall, alle diese Menschen zu kennen. Ohne jene Menschen wäre das vorliegende Werk nicht in dieser Form möglich gewesen.

Wir bleiben TROY (Die Fantastischen Vier, 2004)

Erklärung zur Dissertation

Hierdurch erkläre ich, dass die vorgelegte Dissertation „*Characterisation of Bacteriophage-derived Tailspike and Tail Fibre Proteins*“ selbstständig verfasst, und alle benutzten Hilfsmittel sowie evtl. zur Hilfeleistung herangezogene Institutionen vollständig angegeben wurden.

Die Dissertation wurde nicht schon als Diplom- oder ähnliche Prüfungsarbeit verwendet.

“Science has proof without any certainty.
Creationists have certainty without any proof.”

Ashley Montagu (1905-1999)