Structure-function relationships of bacterial sialate *O*-acetyltransferases

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades einer

Doktorin der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation
von

Dipl.-Biochem. Anne Katrin Bergfeld

geboren am 16. Dezember 1980 in Bergisch Gladbach

2009

Referentin: Prof. Dr. Rita Gerardy-Schahn

Korreferent: Prof. Dr. Walter Müller

Tag der Promotion: 24. Juni 2009

Schlagworte: Polysialinsäure, *O*-Acetyltransferasen, bakterielle Kapselpolysaccharide

keywords: polysialic acid, *O*-acetyltransferases, bacterial capsular polysaccharides

Erklärung zur Dissertation

Hierdurch erkläre ich, dass die Dissertation „Structure-function relationships of bacterial sialate *O*-acetyltransferases“ 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.

Table of Contents

Zusammenfassung

Die Sepsis- und Meningitis-Erreger *Escherichia coli* K1 und *Neisseria meningitidis* Serogruppe C sind von einer dicken Polysaccharid-Kapsel umgeben. Die Kapselpolysaccharide bestehen aus α2,8- bzw. α2,9-verknüpfter Polysialinsäure (PolySia) und stellen für beide Pathogene einen wichtigen Virulenzfaktor dar. In fast 90% aller Serogruppe C Stämme ist die Kapsel zusätzlich durch *O*-Acetylierung an Position C7 oder C8 der Sialinsäure modifiziert. Das Gen, welches die entsprechende *O*-Acetyltransferase OatC kodiert, ist Teil des Kapselgenkomplexes und weist keinerlei Sequenzhomologie zu anderen Proteinen auf. In *Escherichia coli* K1 wird eine phasenvariable *O*-Acetylierung des Kapselpolysaccharids an den Positionen C7 und C9 der Sialinsäuren beobachtet, die durch das Prophagen-kodierte Enzym NeuO katalysiert wird. Die Phasenvariation wird durch eine variable Anzahl an Heptanukleotid-Sequenzwiederholungen im 5'-Bereich des *neuO*-Gens vermittelt, wobei ausschließlich Wiederholungen, die ein Vielfaches von drei sind, eine vollständige Translation des Enzyms erlauben. Um die *O*-Acetylierung von PolySia-Kapseln biochemisch zu untersuchen, wurden NeuO und OatC rekombinant exprimiert, gereinigt und anschließend enzymatisch und strukturell charakterisiert. Im Falle von NeuO erfolgte dies vergleichend für vier Enzymvarianten, die durch Allele mit 0, 12, 24 bzw. 36 Heptanukleotid-Sequenzwiederholungen kodiert werden. Alle Varianten bildeten Hexamere, waren enzymatisch aktiv und wiesen eine hohe Substratspezifität für PolySia mit >14 Resten auf. Für die katalytische Effizienz wurde eine lineare Korrelation mit der Zahl der Sequenzwiederholungen beobachtet und somit ein neuartiger Mechanismus zur Modulation der NeuO-Aktivität aufgedeckt. Mit Hilfe eines Strukturmodells und der Identifizierung von zwei konservierten Aminosäuren, die kritisch für die enzymatische Aktivität sind (His-119 und Trp-143), konnten deutliche Übereinstimmungen zur Hexapeptidrepeat-Proteinfamilie aufgezeigt werden. Im Gegensatz zu NeuO wurde OatC als Homodimer gefunden, wobei die ersten 34 Aminosäuren eine effiziente Oligomerisierungsdomäne ausbilden, deren Funktion auch in einem anderen Proteinkontext erhalten blieb. Die Suche nach Sequenzmotiven führte zur Identifikation eines *nucleophile elbow*-Motivs (*GXS*^{286}*XGG*), welches ein Charakteristikum der α/β-Hydrolase *fold*-Enzyme ist. Mittels umfangreicher Mutagenese-Experimente konnte eine katalytische Triade identifiziert werden, die aus Ser-286, Asp-376 und His-399 besteht. In Übereinstimmung mit einem für α/β-Hydrolase *fold*-Enzyme typischen Ping-Pong-Mechanismus, wurde die Bildung eines kovalenten Acetyl-Enzym-Intermediats nachgewiesen. Zusammen mit Sekundärstrukturanalysen, die eine α/β-Hydrolase *fold*-Topologie vorhersagen, liefern die gezeigten Daten klare Hinweise, dass OatC zur α/β-Hydrolase *fold*-Familie gehört. Dies zeigt, dass OatC evolutionär unabhängig von allen anderen bakteriellen Sialinsäure-spezifischen *O*-Acetyltransferasen entstanden ist. Diesen gemeinsam ist die Zugehörigkeit zur Hexapeptidrepeat-Proteinfamilie, einer Gruppe von Acyltransferasen, die durch eine linksgängige β-Helix und die Ausbildung katalytischer Trimere gekennzeichnet sind. In ersten Kristallisations-Experimenten zur Aufklärung der 3D-Strukturen von PolySia-spezifischen *O*-Acetyltransferasen wurden OatC-Kristalle erhalten, die die Streuung von Röntgenstrahlen bis zu einer Auflösung von etwa 3,8 Å ermöglichten.

Abstract

Escherichia coli K1 and *Neisseria meningitidis* serogroup C are protected by thick polysaccharide capsules composed of α2,8- and α2,9-linked polysialic acid (polySia), respectively. Both pathogens cause bacterial sepsis and meningitis and their capsules represent major virulence factors. In the majority of the serogroup C strains, the capsular polysaccharide is further modified by *O*-acetylation at C7 or C8 of the sialic acids. The gene encoding the corresponding *O*-acetyltransferase OatC is part of the capsule gene complex and shares no sequence similarities with other proteins. In *Escherichia coli* K1, capsule *O*-acetylation at position C7 and C9 of the sialic acids is a phase-variable modification that is catalyzed by the prophage-encoded *O*-acetyltransferase NeuO. Phase-variation is mediated by changes in the number of heptanucleotide repeats within the 5'-region of *neuO*, and full-length translation is restricted to repeat numbers that are a multiple of three. To understand the biochemical basis of polySia capsule *O*-acetylation, NeuO and OatC were recombinantly expressed and purified to homogeneity, allowing the first enzymatic and structural characterization of polySia-specific *O*-acetyltransferases. For NeuO, enzyme variants encoded by alleles containing 0, 12, 24 and 36 heptanucleotide repeats were comparatively analyzed. All variants assembled into hexamers and were enzymatically active with a high substrate specificity towards polySia with >14 residues. The catalytic efficiency increased linearly with increasing numbers of repeats, revealing a new mechanism for modulating NeuO activity. Homology modeling and identification of two conserved amino acids critical for enzymatic activity (His-119 and Trp-143) highlighted a close relationship to the hexapeptide repeat family. The meningococcal sialate *O*-acetyltransferase OatC was found as a homodimer, with the first 34 amino acids forming an efficient oligomerization domain that worked even in a different protein context. Motif-scanning revealed a nucleophile elbow motif (*GXS*^{286}*XGG*), which is a hallmark of α/β-hydrolase fold enzymes. In a comprehensive site-directed mutagenesis approach, a catalytic triad composed of Ser-286, Asp-376, and His-399 was identified. Consistent with a double-displacement catalytic mechanism common to α/β-hydrolase fold enzymes, a covalent acetyl-enzyme intermediate was found. Together with secondary structure prediction showing an α/β-hydrolase fold topology, these data provide strong evidence that OatC belongs to the α/β-hydrolase fold family. This demonstrates that OatC evolved apart from other bacterial sialate *O*-acetyltransferases which all belong to the hexapeptide repeat family, a class of acyltransferases that adopt a left-handed β-helix fold and assemble into catalytic trimers. Initial crystallization approaches to solve the 3D-structures of polySia-specific *O*-acetyltransferases resulted in OatC crystals that diffracted X-rays already down to a resolution of approximately 3.8 Å.

Chapter 1 - General Introduction

1.1 - Bacterial capsular polysaccharides

Polysaccharides at the surface of bacteria have important functions in maintaining surface charge and structural integrity of the bacterial cell, in providing serum resistance and physical protection, and in mediating resistance to desiccation (Guerry and Szymanski, 2008; Curtis *et al.*, 2005). Due to their exposed location at the interface of cell and surrounding medium, cell-surface glycoconjugates often play a crucial role in host-pathogen interactions (Whitfield 2006). In gram-negative bacteria, the majority of these glycoconjugates are found as lipo-polysaccharides (LPS) or capsular polysaccharides (K-antigens) that are linked to the outer membrane (Vimr and Steenbergen, 2009; see Figure 1-1).

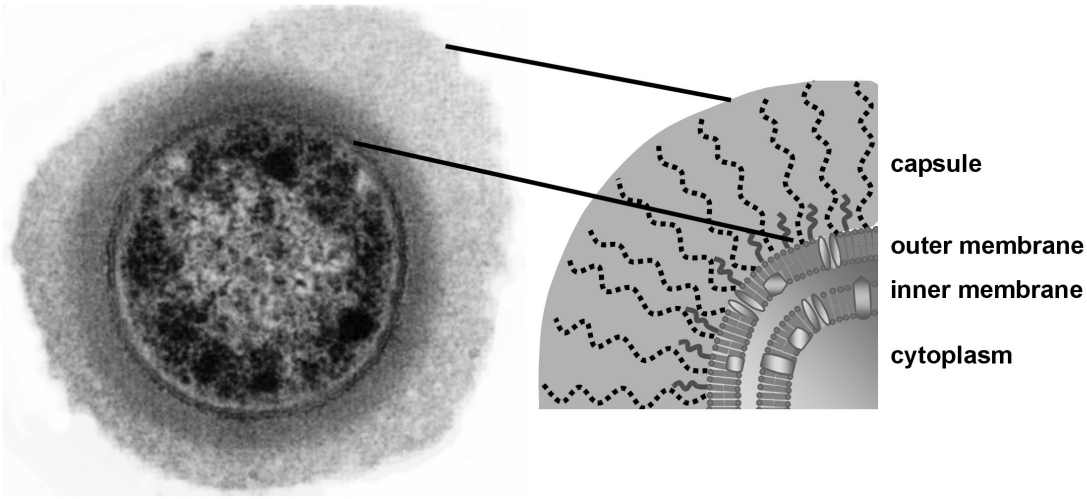


Figure 1-1: Cross-section of an encapsulated bacterium. Electron micrograph of *Escherichia coli* K1, encapsulated by a capsular polysaccharide composed of polysialic acid (left) (Roth *et al.*, 1993) and schematic representation of a gram-negative bacterium surrounded by a polysaccharide capsule (right).

Bacterial capsular polysaccharides have long been accepted as important virulence factors. During invasive infections, the interaction between the capsule and the host's immune system may decide about the outcome of the infection (Moxon & Kroll 1990). One of the most fascinating features of bacterial capsules is their structural diversity. For example, *Escherichia coli* (*E. coli*) as well as *Streptococcus pneumoniae* are capable to express over 80 chemically and serologically distinct capsules, and a single strain of *Bacteroides fragilis* is able to express multiple capsule types (Krinos *et al.* 2001). The capsular polysaccharides are usually hydrophilic and negatively charged. This allows a high water binding capacity to prevent desiccation of the organism and also protects the bacterium against complement-mediated lysis and phagocytosis (Rick and Silver, 1996; Jann and Jann, 1997; Roberts, 2000; Curtis *et al.*, 2005).

Biosynthesis and assembly of bacterial capsules is a complex process; Activated mono- or diphospho-sugars are synthesized in the cytoplasm and incorporated into the nascent polysaccharide chain. The polymer is subsequently dislocated through the periplasm and attached to components of the outer membrane to shape the cell surface. Despite the diversity in capsular polysaccharides, bacteria use only a limited repertoire of biosynthesis and assembly strategies (Whitfield *et al.*, 2006). The genes required for bacterial capsule biosynthesis and assembly are clustered in capsule gene complexes (designated *kps*). Almost three decades ago, the first *kps* was cloned from *Escherichia coli* K1 (Silver *et al.*, 1981) and meanwhile capsule gene complexes have been described for various gram-negative bacteria (for review see Whitfield, 2006). The clustering of the capsule genes at a single chromosomal locus allows coordinated regulation of all genes involved in biosynthesis and export of capsular polysaccharides (Roberts, 1996). Most gram-negative pathogenic bacteria express group 2 capsular polysaccharides, which include several carbohydrate structures that resemble vertebrate glycoconjugates such as the K-antigens K1 (α2,8-linked polysialic acid), K4 (a substituted chondroitin backbone), and K5 (an *N*-acetylheparosan backbone) of *E. coli* (Jann and Jann, 1997). Group 2 *kps* have a common organization composed of three regions (Whitfield and Roberts, 1999).

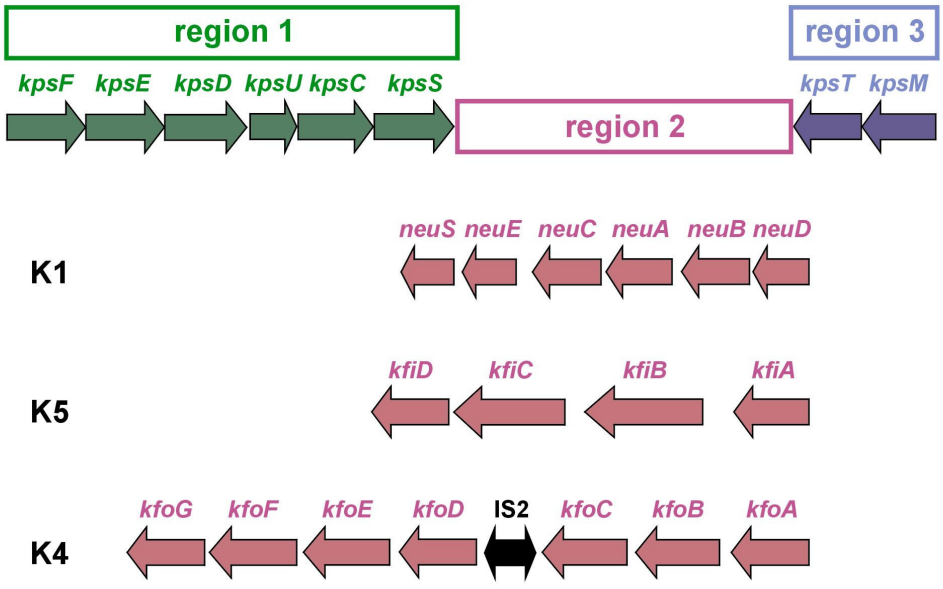
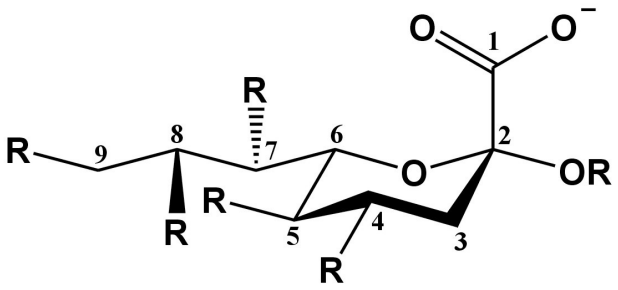


Figure 1-2: Organization of the genes required for expression of *E. coli* K1, K4 and K5 capsules. The *kps* locus comprises a serotype-specific region 2 flanked by two regions that are conserved among all bacteria possessing a group 2 *kps*. Region 1 and 3 gene products are involved in polymer export and translocation, whereas region 2 genes encode the enzymes required for biosynthesis of the respective polysaccharide. The *neu* genes encode proteins required for the synthesis of the K1-antigen composed of (→8)-α-5-*N*-acetylneuraminic acid-(2→) (McGuire and Binkley, 1964; Silver *et al.*, 1981). The K5 capsular polysaccharide is a heteropolymer composed of the disaccharide repeating unit (→4)-β-glucuronyl-(1→4)-α-*N*-acetylglucosaminyl (1→) and is synthesized by the enzymes encoded by the *kfi* genes (Vann *et al.*, 1981; Petit *et al.*, 1995). The K4 capsular polysaccharide consists of a backbone with the structure (→3)-β-D-glucuronyl-(1→4)-β-D-*N*-acetylgalactosaminyl(1→) to which β-fructofuranose is linked at the C-3 position of glucuronic acid (Rodriguez *et al.*, 1988) and is built up by the enzymes encoded by the *kfo* genes (Ninomiya *et al.*, 2002). Figure modified from Whitfield, 2006.

A central region encompasses all genes required for biosynthesis of the serotype-specific polysaccharide whereas two flanking regions contain genes essential for transport of the polymer across the membranes and attachment to the bacterial surface (Roberts *et al.*, 1996; Barrett *et al.*, 2002; Whitfield, 2006; see Figure 1-2). Whereas the gene content of the central region varies according to the structure of the formed polymer, the *kps* gene products that are involved in export across the membranes and assembly of the capsule on the cell surface are highly conserved among serogroups and species (Whitfield 2006; Vimr and Steenbergen, 2009).

1.2 - Polysialic acid capsules

Sialic acids (Sias) are nine-carbon sugars that are derived from neuraminic acid (see Figure 1-3). They all carry a carboxylic function at the C1 position and are typically found as terminal sugars of *N*-glycans, *O*-glycans and glycosphingolipids. Currently, more than 50 naturally occurring sialic acids have been described in various kingdoms of life. The diversity of Sias is based on a variety of substitutions at positions 4, 5, 7, 8, and 9 of the neuraminic acid (Schauer, 2000; Angata and Varki, 2002; Varki and Varki, 2007; see Figure 1-3). The most common modification is an *N*-acetyl group at position C5 which results in the formation of *N*-acetylneuraminic acid (Neu5Ac), the most prevalent Sia in humans. Beside *N*-5-acetylation, the most frequently observed substitution is *O*-acetylation, which can occur on hydroxyl groups at position C4, C7, C8, and C9 .



R2 = H in free Sia; alpha linkage to Gal(3/4/6), GalNAc(6), GlcNAc(4/6) or Sia(8/9)
R4 = H or O-acetyl
R5 = Amino, N-acetyl, N-glycolyl or Hydroxyl
R7 = H, O-acetyl
R8 = H, O-acetyl, O-methyl, O-sulfate or Sia
R9 = H, O-acetyl, O-lactyl, O-phosphate, O-sulfate or Sia

Figure 1-3: Structural diversity of Sialic acids. All Sias are nine carbon sugars with a carboxylic function at the C1 position and are linked to the underlying glycoconjugate via the C2 position. Various substituents at positions 4,5,7,8 and 9 combined with different *O*-glycosidic linkages generate a large set of naturally occurring sialoglycoconjugates (modified from Varki and Varki, 2007).

Sias are also found as components of homo- and heteropolymers called polysialic acid (polySia). These unusual carbohydrate polymers represent the capsular polysaccharides of

several pathogenic bacteria and occur in the animal kingdom as posttranslational modification of a limited number of glycoproteins. In vertebrates, polySia joined by α2,8-linkages is mainly found as dynamically regulated posttranslational modification of the neural cell adhesion molecule (NCAM) (Finne *et al.*, 1983; Mühlenhoff *et al.*, 1998). The presence of this polyanionic glycan structure with its enormous hydration volume controls NCAM-mediated interactions, modulates cell-cell interactions, and plays an important role in neural plasticity of the adult brain (Weinhold *et al.*, 2005; Hildebrandt *et al.*, 2007; Gascon *et al.*, 2007; Rutishauser 2008). Although NCAM represents the major polySia acceptor in vertebrates, additional polysialylated proteins have been identified such as the scavenger receptor CD36 in human milk, the α-subunit of the voltage-sensitive sodium channel in rat brain, and neuropilin-2 of maturating human dendritic cells (Yabe *et al.*, 2003; Zuber *et al.*, 1992; Curreli *et al.*, 2007).

In bacteria, polySia is found as linear homo- and heteropolymers of different linkages which form the capsule of several pathogens (see Figure 1-4). The capsular polysaccharide of *Neisseria meningitidis (N. meningitidis* or meningococcus*)* of the serogroups W-135 and Y are Neu5Ac-containing heteropolymers composed of the disaccharide repeating units (→6)-α-D-Gal*p*-(1→4)-α-Neu5Ac-(→2) and (→6)-α-D-Glc*p*-(1→4)-α-Neu5Ac-(→2), respectively (Bhattacharjee *et al.*, 1976). The capsule of *N. meningitidis* serogroup C is a homopolymer of Neu5Ac residues joined by α2,9-linkages, whereas *E. coli* K1 and serogroup B meningococci comprise structurally identical capsular polysaccharides consisting of α2,8-linked Neu5Ac of up to 200 residues (Bhattacharjee *et al.*, 1975; Rohr and Troy, 1980; Pelkonen *et al.*, 1988). Hence, the polySia found in *E. coli* K1 and serogroup B meningococci is chemically identical to polySia expressed in the host organism (e.g. on NCAM) (Mühlenhoff *et al.*, 1998). Due to this antigenic mimicry, these capsules are poorly immunogenic, and no effective polysaccharide-based vaccines are available.

In several K1 strains, modification of the capsule by *O*-acetylation of the Neu5Ac residues in positions O7 and O9 was observed (Ørskov *et al.*, 1979). With the single exception of *N. meningitidis* serogroup B, *O*-acetylation of the polySia capsules have also been found in serogroup C, W-135, and Y meningococci (see Figure 1-4).

In the capsular polysaccharide of serogroup C meningococci, the *O*-acetyl groups are distributed between position C7 and C8 of Neu5Ac residues, whereas in serogroups W-135 and Y, *O*-acetylation is found at positions C7 and C9 of the capsular sialic acids (Bhattacharjee *et al.*, 1976; Jennings *et al.*, 1977; Lemercinier and Jones, 1996). Studies from the United Kingdom demonstrated capsule *O*-acetylation for 88% and 79% of the serogroup C and Y strains, respectively, whereas only 8% of the W-135 strains displayed *O*-acetylated capsules (Borrow *et al.*, 2000; Longworth *et al.*, 2002).

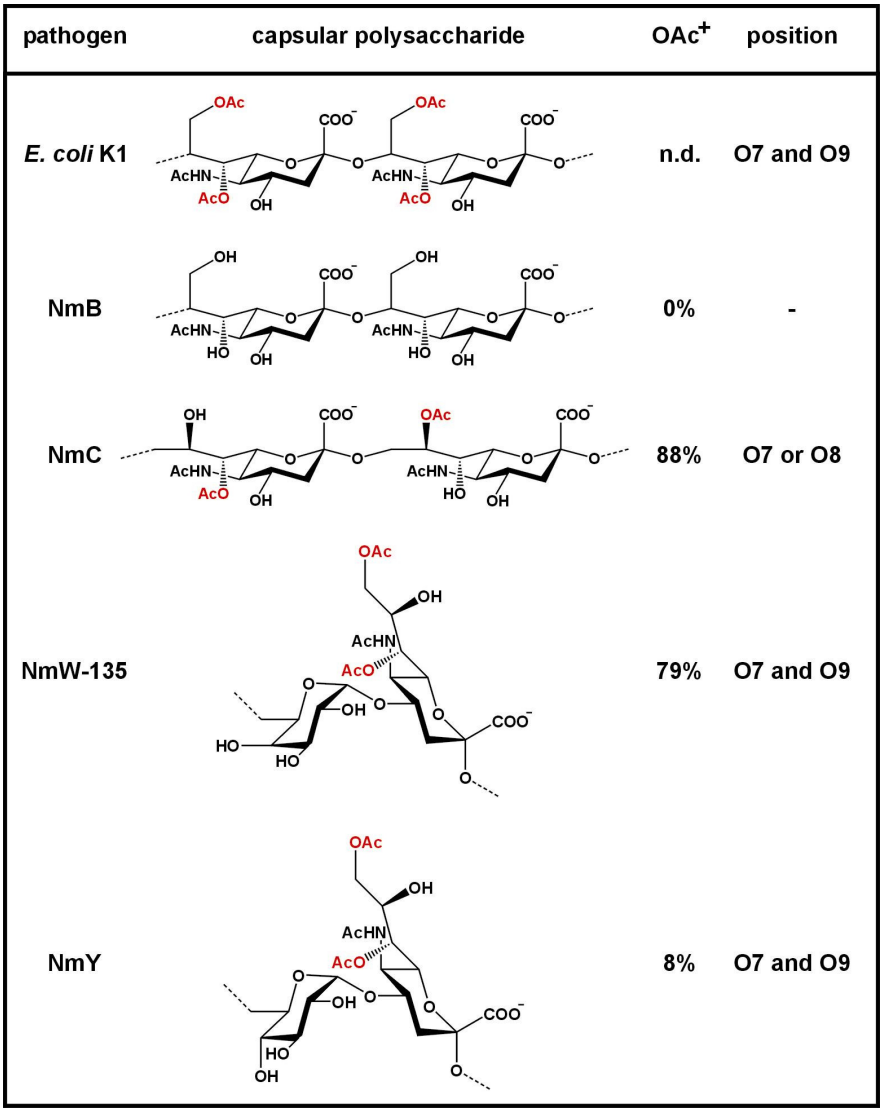


Figure 1-4: Polysaccharide structures of *E. coli* K1 and *N. meningitidis* serogroup B, C, W-135, and Y. Positions, where *O*-acetyl groups were found are shown in red (Bhattacharjee *et al.*, 1975; Bhattacharjee *et al.*, 1976; Ørskov *et al.*, 1979; Lemercinier and Jones, 1996). The percentage of strains that comprised *O*-acetylated capsular polysaccharides (OAc^+) in a study from the UK are indicated (Borrow *et al.*, 2000; Longworth *et al.*, 2002).

Although *O*-acetylation influences size and net charge of the capsular polysaccharides, the biological relevance of this modification with regard to pathogenicity is still largely unclear. In the case of *E. coli* K1, capsule *O*-acetylation is associated with increased immunogenicity, but nevertheless, a correlation between *O*-acetylation and increased virulence was observed in patients with bacteraemia (Frasa *et al.*, 1993). Because *O*-acetylated Sias are more resistant to hydrolysis by sialidases, capsule *O*-acetylation may favor colonization of *E. coli* K1 in the intestinal tract (Ørskov *et al.*, 1979). For *N. meningitidis* serogroups C and Y, an inverse correlation between immunogenicity and the level of capsule *O*-acetylation was observed, whereas no impact on immunogenicity was found for the *O*-acetylated capsular polysaccharide of serogroup W-135 meningococci (Michon *et al.*, 2000; Fusco *et al.*, 2007; Gudlavalleti *et al.*, 2007). These findings suggest that for serogroup C and Y meningococci,

capsule *O*-acetylation may play a general role in masking immunogenic epitopes and thereby enabling escape from immune surveillance. However, as shown recently, capsule *O*-acetylation does not impair recognition of serogroup C, W-135, and Y meningococci by human dendritic cells, indicating that modulatory effects of capsule *O*-acetylation on immune recognition are restricted to the humoral immune response (Villwock *et al.*, 2008).

1.3 - Genetic basis of capsule *O*-acetylation in *Escherichia coli* K1

E. coli is a versatile enterobacterium which can be subdivided into (i) nonpathogenic commensal strains, (ii) pathogenic intestinal strains and (iii) pathogenic extraintestinal strains (Russo and Johnson, 2000). This classification is mainly based on the presence or absence of particular DNA regions that are associated with certain pathotypes. This genetic information belongs to the flexible *E. coli* genome, which includes genetic elements like plasmids, bacteriophages and genomic islands, that are frequently acquired by horizontal gene transfer (Dobrindt, 2005). Extraintestinal *E. coli* (ExPEC) strains possess virulence traits that allow invasion, colonization and induction of diseases outside of the gastrointestinal tract. ExPEC are responsible for several human diseases including urinary tract infections, sepsis, pneumonia, surgical site infections and neonatal meningitis (Smith *et al.*, 2007). Concerning the mortality and morbidity of ExPEC infections, these pathogens have a great impact on the public health systems, with an economic cost of several billion dollars annually (Russo and Johnson, 2003). In the United States, more than 50% of all cases of neonatal sepsis and meningitis that are caused by gram-negative enteric bacteria were due to ExPEC strains, and approximately 80% of those strains were carrying the K1 capsular antigen (Pong and Bradley, 1999; Bonacorsi *et al.*, 2003; Johnson *et al.*, 2002). The K1 capsule has long been accepted as a major virulence factor of ExPEC (Johnson and Russo, 2005; Freitag *et al.*, 2005). The K1 capsular polysaccharide consists of α2,8-linked Neu5Ac with chain lengths of 160-230 residues (Pelkonen *et al.*, 1988). Neonatal immune responses to polySia-carrying bacteria such as *E. coli* K1 are most likely restricted due to the structural identity of the capsule to polySia found in vertebrates (Rohr and Troy, 1980; Mühlenhoff *et al.*, 1998; Stein *et al.*, 2006). Development of an *E. coli* K1-caused meningitis is a multistage process, initiated by colonization of the mucosa of the newborns gastrointestinal tract and followed by translocation into the blood stream after passage of the mucosal membrane (Kim *et al.*, 2001; Kim *et al.*, 2003; Kim *et al.*, 2005). In the circulation, the K1-capsule is essential for survival and multiplication of the pathogen as it protects the bacterium from humoral and cellular components of the immune system (Cross *et al.*, 1986; Mushtaq *et al.*, 2004; Bortolussi *et al.*, 1979; Mushtaq *et al.*, 2005). After reaching a specific threshold level of bacteraemia, the pathogen adheres to brain microvascular endothelial cells which form the blood-brain-barrier (BBB). Subsequently, *E. coli* K1 crosses this barrier by

transcytosis through ligand-receptor interactions, host cytoskeletal rearrangements and the activation of different signaling pathways, without affecting the integrity of the BBB (Kim, 2003; Kim *et al.*, 2005; Galanakis *et al.*, 2006; Kim *et al.*, 2008). The pathogen is able to transverse the BBB irrespective of the presence or absence of the K1 capsule, but only encapsulated bacteria survive the transcytosis (Kim *et al.*, 1992; Hoffman *et al.*, 1999). After reaching the subarachnoid space, *E. coli* K1 multiplies and elicits host inflammatory responses in the meninges which ultimately results in meningitis (Kim *et al.*, 2003; Xie *et al.*, 2004; Kim *et al.*, 2005). After invasion of the central nervous system, the bacterium ceases the expression of the protective capsular polysaccharide, indicating, that the capsule is dispensable after the neuropathogen colonized a protected niche (Zelmer *et al.*, 2008; Guerry and Szymanski, 2008).

Interestingly, several *E. coli* K1 strains are able to further modify their protective polySia capsule by *O*-acetylation (Ørskov *et al.*, 1979; Higa and Varki, 1988; Deszo *et al.*, 2005; see Figure 1-2). This surface modification was first observed in 1979 by Ørskov and coworkers and it was shown that certain K1 strains were capable to switch between an *O*-acetylated (OAc^+) and a non *O*-acetylated (OAc^-) capsule form with a constantly high frequency of 1:50 to 1:20 (Ørskov *et al.*, 1979). Later on, enzymatic activity of a polySia-specific *O*-acetyltransferase was proved but the genomic basis of capsule *O*-acetylation remained an enigma (Higa and Varki, 1988). All attempts to purify the endogenous enzyme required detergent extraction, and after partial purification on DEAE-Sepharose, further isolation steps abolished the enzymatic activity (Higa and Varki, 1988). In 2005, analysis of the partially sequenced genome of the *E. coli* K1 strain RS218 (http://www.genome.wisc.edu) revealed the presence of a 40 kb prophage (termed CUS-3) which is integrated into the *argW* gene of the bacterial genome (Deszo *et al.*, 2005). Based on sequence similarities and genomic organization, CUS-3 was characterized as a P22-like phage (Stummeyer *et al.*, 2006). All members of this family of temperent phages comprise one or more lysogenic conversion genes located between tailspike and integrase gene. The corresponding translation products modify surface structures of the host, a common strategy to prevent superinfection by homologous phages (Stummeyer *et al.*, 2006; Allison and Verma, 2000). In the equivalent locus, CUS-3 harbors the *neuO* gene, which was shown to mediate the phase-variable *O*-acetylation of the polySia capsule of *E. coli* K1 (Deszo *et al.*, 2005; see Figure 1-5).

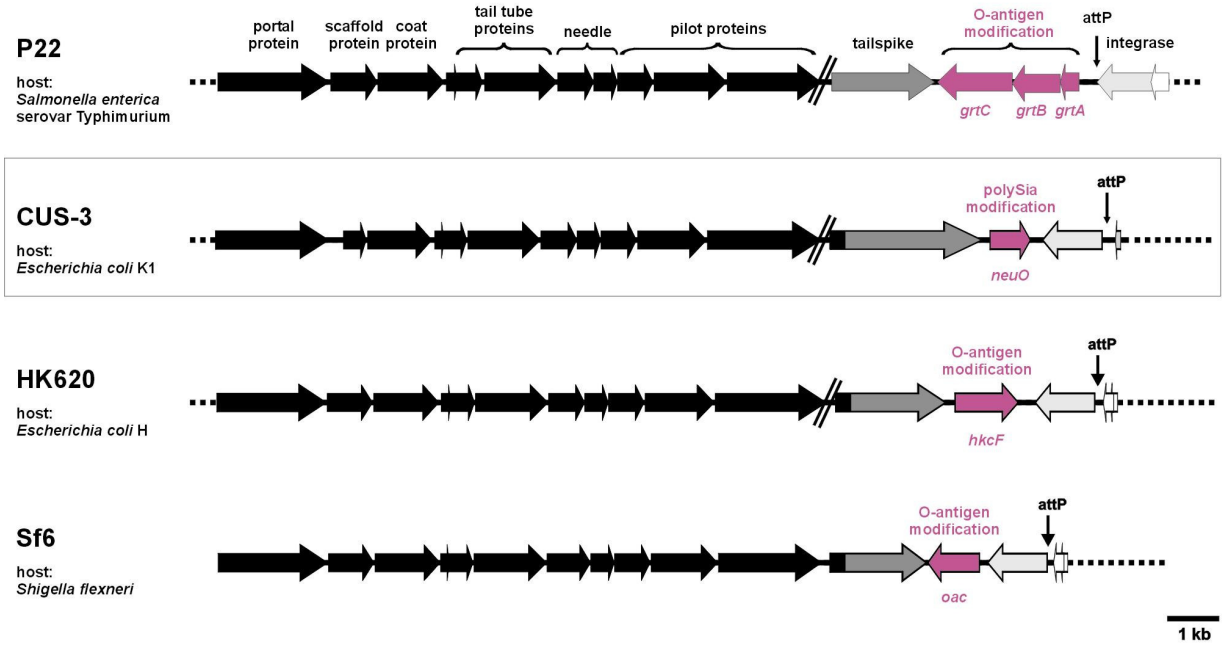


Figure 1-5: Genomic localization of the *neuO* gene. The *neuO* gene is part of the 40kb CUS-3 genome which is integrated into the *argW* tRNA gene of *E. coli* K1 strain RS218 (Deszo *et al*., 2005). Part of the second half of the CUS-3 genome encompassing virion, lysogenic conversion and integrase genes is shown in comparison with the corresponding genome parts of Salmonella phage P22, Coliphage HK620 and Shigella phage Sf6. The lysogenic conversion gene(s) (shown in pink) are flanked by a tailspike (dark grey) and an integrase (light grey) gene. Location of phage attachment sites (attP) is indicated by an arrow. Figure modified from Stummeyer *et al*., 2006.

Consequently, *O*-acetylation of the K1-capsule is restricted to *E. coli K1* strains that harbor the CUS-3 prophage. Analysis of 111 clinical *E. coli* K1 isolates revealed that over 60% of the strains comprised the *neuO* gene (King *et al*., 2007).

Sequence analysis of the *neuO* gene threw light on the puzzling observation that certain K1 strains were able to switch between OAc^+ and OAc^- capsule forms. This phase variation is mediated by a variable number of heptanucleotide repeats (VNTRs) of the sequence

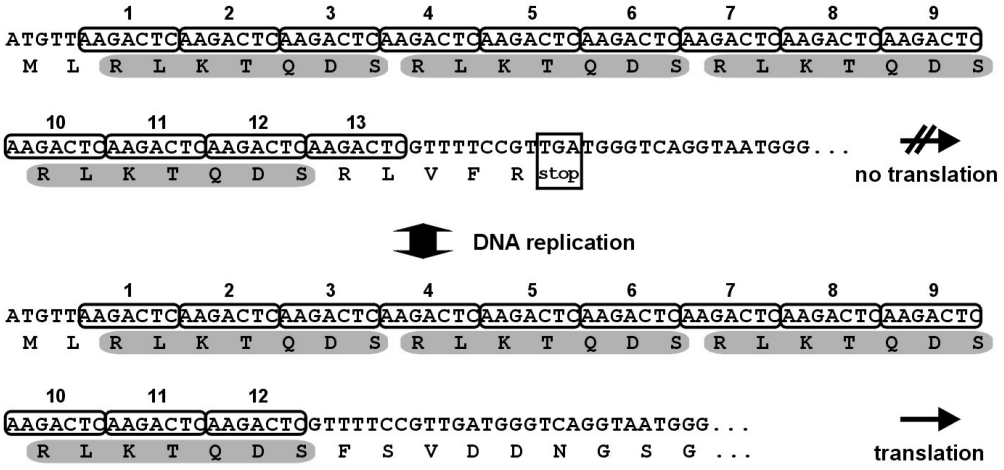


Figure 1-6: Sequence elements mediating phase-variable expression of *neuO*. Schematic representation of the VNTR region of *neuO* showing alleles with 13 (top) and 12 (bottom) 5'-AAGACTC-3' repeats (boxed). Three heptanucleotide repeats encode one protein heptad (gray). During DNA replication, changes in the number of nucleotide repeats can arise through slipped-strand mispairing.

5'-AAGACTC-3' which are located two base pairs downstream of the starting ATG (Deszo *et al.*, 2005; see Figure 1-6). VNTR loci are among the most variable regions of many bacterial genomes and arise through slippage and mispairing during DNA replication due to occasional DNA polymerase dissociation (van Belkum *et al.*, 1998; Henderson *et al.*, 1999; Martin *et al.*, 2005; Lukacova *et al.*, 2007). In the case of *neuO*, gain or loss of complete VNTRs due to slipped-strand mispairing results in an 'all or non' control mechanism of enzyme expression. Translation of functional NeuO is restricted to *neuO* alleles that harbor a repeat number divisible by three (phase ON), whereas all other repeat numbers move the start codon out of frame resulting in truncated and thereby inactive translation products (phase OFF) (Deszo *et al.*, 2005).

1.4 - Genetic basis of polySia *O*-acetylation in *Neisseria meningitidis*

N. meningitidis is a major cause of septicaemia and meningitis worldwide with incidence rates varying from 1 to 1000 cases per 100,000 (van Deuren *et al.*, 2000 ; Stephens, 2007). Meningococci are a transitory colonizer of the human nasopharynx, which represents the sole reservoir (Yazdankhah and Caugant, 2004). With up to 20% healthy carriers at any one time, meningococci are common commensals that only occasionally are responsible for life-threatening diseases (Greenfield *et al.*, 1971; Caugant *et al.*, 1994; Stephens, 1999; Janda & Knapp, 2003). Only a few hyper-invasive genetic clones of the species are pathogenic and may rapidly result in countrywide outbreaks of meningitis (Raymond *et al.*, 1997; Swartley *et al.*, 1997; Maiden *et al.*, 1998; Vogel *et al.*, 2000; Yazdankhah *et al.*, 2004; Caugant, 2008). In contrast to most bacteria, meningococci spontaneously produce a vast number of genetic variants, also in the absence of environmental changes or stimuli (Davidsen and Tonjum, 2006). The frequent genomic alterations (e.g. exchange of genetic islands by horizontal gene transfer, slipped-strand mispairing of repetitive nucleotides, intergenic recombination events or insertion sequence element movement) have drastic consequences on host-pathogen interactions, impacting structural and antigenic changes in surface components that are crucial for adherence and invasion (Davidsen and Tonjum, 2006; Caugant, 2008; Hotopp *et al.*, 2006; Kahler *et al.*, 2001). The pathogenesis of meningococcal meningitis starts with colonization of the human nasopharynx, followed by adherence to and invasion of epithelial cells by endocytosis enabling passage across the epithelial barrier into the bloodstream (Stephens and Farley, 1991; Yazdankhah and Caugant, 2004; Stephens, 2007). Meningococci that are able to survive and multiply in the circulation belong to the limited number of pathogens capable of crossing the blood-brain barrier to gain access to the cerebrospinal fluid (Cartwright and Ala'Aldeen, 1997; Doran *et al.*, 2005). Despite appropriate treatment, meningitis caused by invasive meningococci remains a prevalent disease with a

case-fatality rate of 10% and high morbidity among survivors (Connolly and Noah, 1999; Tzeng and Stephens, 2000; Brandtzaeg and van Deuren, 2002; Goldrace *et al.*, 2003).
Based on the structure of the capsular polysaccharide, meningococci are classified into thirteen distinct serogroups (Gotschlich *et al.*, 1969). Serogroups A, B, C, W-135, X, and Y are currently associated with significant pathogenic potential and are responsible for over 90% of the invasive diseases worldwide (Rosenstein *et al.*, 2001; Djibo *et al.*, 2003; Boisier *et al.*, 2007; Stephens, 2007; see Figure 1-7).

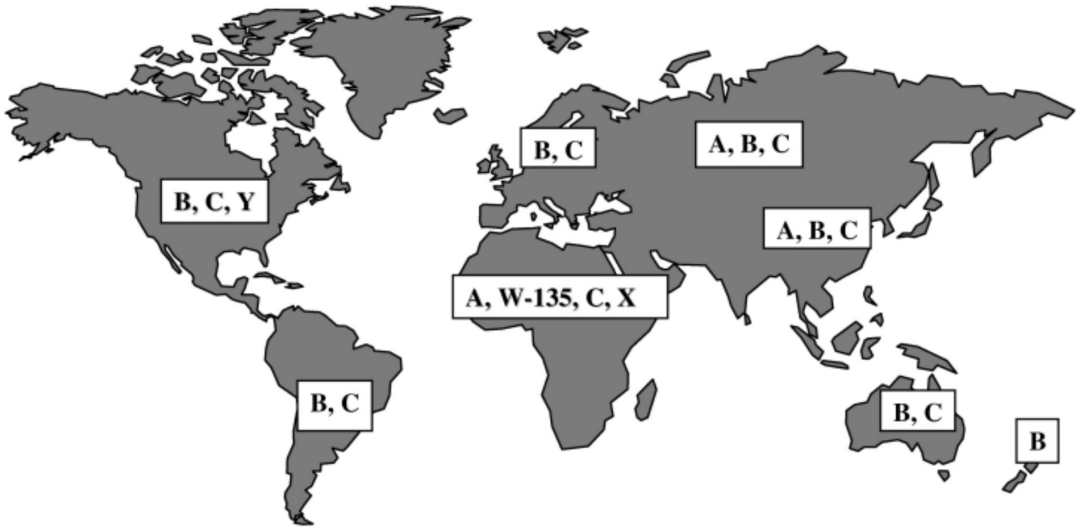


Figure 1-7: Global serogroup distribution of invasive meningococcal disease (modified from Stephens, 2007).

The polysaccharide capsules play contrasting roles in meningococcal pathogenesis. On the one hand side, the disease-associated capsules protect the bacteria from phagocytosis and complement-mediated lysis and are therefore essential for survival and multiplication of the pathogen in the bloodstream (Jarvis and Vedros, 1987; Vogel *et al.*, 1996; Spinosa *et al.*, 2007). On the other hand, the presence of a thick polysaccharide capsule masks the meningococcal adhesins and invasins and thereby inhibits colonization and invasion of the nasopharynx (Hammerschmidt *et al.*, 1996b). Expression of the capsule is dependent on both the possession and the expression of the capsular genes located in the *kps* (Frosch *et al.*, 1989; see Figure 1-8). The capsular genes are located within a genomic island of horizontally transferred DNA, explaining acquisition and loss of the *kps* among strains (Parkhill *et al.*, 2000; Tettelin *et al.*, 2000; Schön *et al.*, 2008). Moreover, expression of genes involved in capsule biosynthesis is regulated by phase variation via slipped-strand mispairing or reversible insertion of mobile elements (Hammerschmidt *et al.*, 1996a; Hammerschmidt *et al.*, 1996b). There is evidence that meningococci display an adaptive response upon contact to epithelial cells involving transcriptional regulators to down-regulate capsule expression during the early stages of the infection to facilitate adhesion to and invasion of the host cells (Deghmane *et al.*, 2002; Deghmane *et al.*, 2003).

Among the six clinically important meningococcal serogroups, four serogroups, i.e. B, C, W-135 und Y, express sialic acid containing capsular polysaccharides (Rosenstein *et al.*, 2001; Jódar *et al.*, 2002; Boisier *et al.*, 2007; see part 1.2). The group B capsular polysaccharide represents a homopolymer of α2,8-linked Sia which is structurally identical to polySia found in the human host, whereas the group C capsule is a homopolymer of α2,9-linked Sia (Bhattacharjee *et al.*, 1975; see Figure 1-2). Serogroups W-135 und Y express heteropolymeric capsules composed of the disaccharide repeating units (→6)-α-D-Gal*p*-(1→4)-α-Neu5Ac-(→2) and (→6)-α-D-Glc*p*-(1→4)-α-Neu5Ac-(→2), respectively (Bhattacharjee *et al.*, 1976; see Figure 1-2). Effective meningococcal glycoconjugate vaccines targeting the serogroup A, C, Y and W135 polysaccharides are currently available (Pace and Pollard, 2007).

Among the polySia containing capsules, polysaccharides of serogroup C, W-135, and Y can be found *O*-acetylated as described in part 1.2 (Bhattacharjee *et al.*, 1975; Bhattacharjee *et al.*, 1976; Lemercinier and Jones, 1996; see Figure 1-2). Although the biological role of polySia *O*-acetylation with regard to pathogenicity is still unclear, recent findings suggest a general impact of capsule *O*-acetylation in masking immunogenic epitopes to circumvent the hosts immune system (Michon *et al.*, 2000; Fusco *et al.*, 2007; Gudlavalleti *et al.*, 2007; Villwock *et al.*, 2008). Even though *O*-acetylation of meningococcal capsular sialic acids is known for decades, there is almost no information on the enzymes catalyzing this modification. Only one report on the capsule specific *O*-acetyltransferase activity in serogroup C meningococci is available (Vann *et al.*, 1978). In this study, spheroplast membranes of serogroup C meningococcal were used as an enzyme source, and Vann and coworkers demonstrated that the endogenous sialate *O*-acetyltransferase transfers acetyl groups from acetyl-CoA to endogenous membrane bound as well as to exogenous soluble serogroup C capsular polysaccharide (Vann *et al.*, 1978). However, reports on the isolation and further characterization of this enzyme are missing and no information is available on endogenous sialate *O*-acetyltransferase activity in serogroup W-135 and Y meningococci. The lack of information on these enzymes is also due to the fact that the genetic basis of polySia capsule *O*-acetylation was discovered only recently (Claus *et al.*, 2004). In 2004, the genes encoding the respective *O*-acetyltransferases of serogroup C, W-135 and Y meningococci were identified (Claus *et al.*, 2004; Vogel *et al.*, 2004), which was at that time the first identification of genes that encode sialic acid-specific *O*-acetyltransferases. The *O*-acetyltransferase genes of meningococcal serogroups C, W-135 and Y are located downstream of the polySia biosynthesis genes as part of the serotype specific region A within the capsule gene complexes (see Figure 1-8).

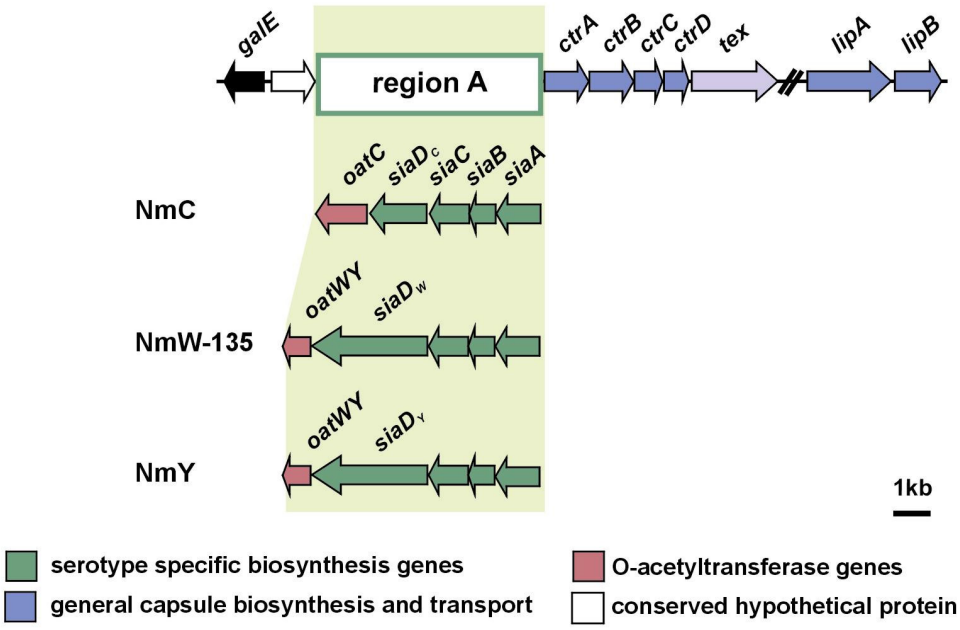


Figure 1-8: Genetic organization of the capsule biosynthesis and transport loci of *N. meningitidis* serogroups C, W-135, and Y. The capsule transport operon (*ctrA-D*) as well as the genes *lipA/lipB* are involved in transport of the polymer across the membranes (shown in blue) and are conserved among *Neisseria* species. Region A comprises all genes required for biosynthesis of the serotype-specific capsular polysaccharide and is highlighted in green. The enzymes encoded by *siaA–C* synthesize CMP-activated Sia and are conserved among the Sia-producing serogroups. Polysialyltransferases, which synthesize the serogroup-specific polysaccharide, are encoded by *siaD*, and vary from serogroup to serogroup. The *oat* genes responsible for capsule *O*-acetylation are located downstream of *siaD* (shown in pink). The capsule gene locus is located between the genes *tex* (shown in grey) and *galE* (shown in black) and is regarded as a genetic island taken up by horizontally transferred DNA (Tettelin *et al.*, 2000; Parkhill *et al.*, 2000). Region A is shown from strains Fam18 (NmC; Bentley *et al.*, 2007), 2232 (NmW-135; Claus *et al.*, 2004) and 2227 (NmY; Claus *et al.*, 2004).

Notably, the serogroups W-135 and Y comprise an identical *O*-acetyltransferase gene, which was therefore designated *oatWY*. The previously described widespread occurrence of OAc^- W-135 and Y capsule forms was found to correlate with the presence of non-functional oatWY genes (e.g. due to insertion sequence elements) (Claus *et al.*, 2004). The gene encoding the polySia specific *O*-acetyltransferase of serogroup C meningococci was termed *oatC*. Surprisingly, no sequence similarities between *oatC* and *oatWY* were detectable, although the two meningococcal genes are located at similar loci and have closely related functions. In line with the fact that *N. meningitidis* contains more genes that undergo phase variation than any pathogen studied to date and the observation that the *kps* is among the most variable regions, it is no surprise that also expression of *oatC* can be switched on and off. The presence of two homopolymeric tracts with seven nucleotides each allows phase variation with a frequency of approx. 1:20,000 that is most likely due to slipped-strand mispairing (Claus *et al.*, 2004).

Identification of the genetic basis of capsule *O*-acetylation in *E. coli* K1 and serogroup C, W-135, and Y meningococci now enables the production of recombinant enzymes, thereby facilitating the biochemical and structural characterization of polySia-specific *O*-acetyltransferases.

1.5 - Objectives

The aim of the present study was the molecular characterization of bacterial polySia-specific *O*-acetyltransferases. For this purpose, two genetically unrelated enzymes were chosen as model proteins: the prophage-encoded *O*-acetyltransferase NeuO, responsible for polySia capsule modification in *E. coli* K1 and OatC, the *O*-acetyltransferase that catalyzes modification of capsular sialic acids in serogroup C meningococci. For both enzymes, protocols for the recombinant expression, purification, and determination of enzymatic activity were established. Combined with the generation of site-specific mutants, this allowed the first biochemical characterization of polySia-specific *O*-acetyltransferases and the analysis of structure-function relationships. Points that were addressed included (i) determination of donor- and acceptor-substrate specificities, (ii) identification of structural and kinetic features that highlight phylogenetic relationships to known protein families, (iii) kinetic analysis of naturally occurring NeuO variants to investigate the biological impact of variations in the primary sequence and in the number of N-terminal RLKTQDS heptads, (iv) and crystallization of OatC to reveal first insight into the tertiary structure of a sialate *O*-acetyltransferase.

Chapter 2

Biochemical characterization of the polysialic acid-specific *O*-acetyltransferase NeuO of *Escherichia coli* K1

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

'Biochemical characterization of the polysialic acid-specific *O*-acetyltransferase NeuO of *Escherichia coli* K1.'

Anne K. Bergfeld[‡], Heike Claus[§], Ulrich Vogel[§] and Martina Mühlenhoff[‡1]

From the [‡]Department of Cellular Chemistry, Medical School Hannover, Carl-Neuberg-Strasse 1, Hannover 30625 and the Institute for Hygiene and Microbiology, [§]University of Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg, 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. 30, pp. 22217–22227, July 27, 2007

Received for publication, April 11, 2007, and in revised form, May 14, 2007 Published, JBC Papers in Press, May 22, 2007
DOI 10.1074/jbc.M703044200

Preface - About the Manuscript

Escherichia coli K1 is one of the main organisms causing bacterial sepsis and meningitis in neonates. The pathogen is protected by a thick capsular polysaccharide composed of α2,8-linked polysialic acid, which represents the major virulence factor. In several K1 strains, phase-variable *O*-acetylation of the sialic acid residues was observed, a modification that is catalyzed by the prophage-encoded *O*-acetyltransferase NeuO. The reversible switch between an on and off phase is mediated by changes in the number of heptanucleotide repeats in the 5'-coding region of the *neuO* gene since only repeat numbers that are a multiple of three allow full-length translation of *neuO*. During the on phase, the nucleotide repeats are translated into tandem copies of an RLKTQDS heptad. Previous attempts to isolate the endogenous enzyme failed and almost no information on the biochemical basis of K1 capsule *O*-acetylation was available. Therefore, my work was aiming at developing expression and purification strategies to isolate active recombinant NeuO. In parallel, I established two independent *in vitro* assays and determined donor and acceptor substrate specificities of NeuO as well as the minimal acceptor length. To investigate the role of N-terminal NeuO extensions by variable numbers of RLKTQDS heptads, I purified NeuO variants containing 0, 4, 8, and 12 heptads and performed a comparative kinetic analysis for these variants. Homology modeling predicted a left-handed parallel β-helix (LβH) as the central fold of NeuO, indicating that the enzyme is a member of the LβH-family of acyltransferases. To confirm this model and to gain insight into structure-function-relationships, I generated site-specific mutants and compared enzymatic activity and quaternary structure of purified wild-type and mutant forms of NeuO.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

Biochemical Characterization of The polysialic Acid-specific *O*-Acetyltransferase NeuO of *Escherichia coli* K1*[S]

Received for publication, April 11, 2007, and in revised form, May 14, 2007 Published, JBC Papers in Press, May 22, 2007, DOI 10.1074/jbc.M70344200

Anne K. Bergfeld[‡], Heike Claus[§], Ulrich Vogel[§], and Martina Mühlenhoff[‡1]

From the [‡]Department of Cellular Chemistry, Medical School Hannover, Carl-Neuberg-Strasse 1, Hannover 30625 and the Institute for Hygiene and Microbiology, [§]University of Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg

***Escherichia coli* K1 is a leading pathogen in neonatal sepsis and meningitis. The K1 capsule, composed of α2,8-linked polysialic acid, represents the major virulence factor. In some K1 strains, phase-variable *O*-acetylation of the capsular polysaccharide is observed, a modification that is catalyzed by the prophage-encoded *O*-acetyltransferase NeuO. Phase variation is mediated by changes in the number of heptanucleotide repeats within the 5'-coding region of *neuO*, and full-length translation is restricted to repeat numbers that are a multiple of three. To understand the biochemical basis of K1 capsule *O*-acetylation, NeuO encoded by alleles containing 0, 12, 24, and 36 repeats was expressed and purified to homogeneity via a C-terminal hexahistidine tag. All NeuO variants assembled into hexamers and were enzymatically active with a high substrate specificity toward polysialic acid with >14 residues. Remarkably, the catalytic efficiency (k_{cat}/K_m^{donor}) increased linearly with increasing numbers of repeats, revealing a new mechanism for modulating NeuO activity. Using homology modeling, we predicted a three-dimensional structure primarily composed of a left-handed parallel β-helix with one protruding loop. Two amino acids critical for catalytic activity were identified and corresponding alanine substitutions, H119A and W143A, resulted in a complete loss of activity without affecting the oligomerization state. Our results indicate that in NeuO typical features of an acetyltransferase of the left-handed β-helix family are combined with a unique regulatory mechanism based on variable N-terminal protein extensions formed by tandem copies of an RLKTQDS heptad.**

Escherichia coli K1 is one of the main organisms causing bacterial sepsis and meningitis during the neonatal period (1–3). The pathogenesis of *E. coli* meningitis is a complex process involving colonization of the gastrointestinal tract, intestinal translocation, bacteremia, passage of the blood-brain barrier, and invasion of the arachnoidal space (4). Cell culture experiments indicated that transfer across the blood-brain barrier is mediated by transcytosis through brain endothelial cells (5, 6) and that, in contrast to other bacterial agents causing meningitis, *E. coli* K1 invades meningioma cells directly, leading to rapid cell death before an inflammatory response can be induced (7). Despite advances in diagnostics and therapeutics of neonatal infections, *E. coli* neonatal meningitis is still characterized by high rates of mortality and a high incidence of permanent neurological sequelae in those that survive (8, 9). *E. coli* K1 is protected by a thick layer of capsular polysaccharide, which is the major virulence factor important for serum resistance and vital passage of the blood-brain barrier (5, 10, 11). The K1 capsule consists of polysialic acid (polySia) formed by the most prevalent sialic acid of humans, 5-*N*-acetyl neuraminic acid (Neu5Ac). Up to 200 residues are joined in α2,8-glycosidic linkages, resulting in a linear homopolymer that is structurally identical to polySia of the host organism (12, 13). Due to this antigenic mimicry, the K1 capsule is poorly immunogenic, and no effective polysaccharide-based vaccines are available.

In several K1 strains, modification of the capsule by *O*-acetylation of the Neu5Ac residues in positions *O*-7 and *O*-9 was observed (14–16). This modification is phase-variable, and individual strains can switch between *O*-acetylated (OAc^+) and non-*O*-acetylated (OAc^-) capsule variants (14). Although associated with increased immunogenicity, modification of the K1 capsule by *O*-acetylation correlates with increased virulence in patients with bacteremia (17). Because *O*-acetylated polySia resists hydrolysis by neuraminidases, capsule *O*-acetylation may favor colonization of the intestinal tract (14).

O-Acetylation of the K1 capsule is catalyzed by an acetyl-CoA-dependent *O*-acetyltransferase with preferential acceptor specificity toward polySia with >14 residues (15). The corresponding gene (*neuO*) is part of a 40-kb prophage of the P22 family, and phase variation is mediated by changes in the overall length of variable number of tandem repeats (VNTRs) within the 5'-coding region (16, 18). VNTR loci are among the most variable regions of many bacterial genomes and arise through slippage and mispairing during DNA replication due to occasional DNA polymerase dissociation (19, 20). In the *neuO*-positive K1 strains reported so far, 14–39 copies of the heptanucleotide unit 5'-AAGACTC-3' were inserted two nucleotides downstream of the start codon (16). Only

* This work was supported by grants of the Deutsche Forschungsgemeinschaft (Grant VO 718/4-1 to U.V. and Grant MU 1774/2-1 to M.M.). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) AJ783705.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1-S3.

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

[2] The abbreviations used are: polySia, polysialic acid; DP, degree of polymerization; GAT, galactoside acetyltransferase; IMAC, immobilized metal affinity chromatography; LβH, left-handed β-helix; mAb, monoclonal antibody; MAT, maltose O-acetyltransferase; Neu5Ac, *N*-acetylneuraminic acid; OAc, *O*-acetylation; SAT, serine O-acetyltransferase; VNTR, variable number of tandem repeats; Strep, Strep tag II.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

repeat numbers that are a multiple of three allow full-length translation of *neuO* (phase on), whereas other numbers move the start codon out of frame resulting in truncated and thereby inactive translation products (phase off). In the case of full-length translation, the presence of VNTRs will lead to significant N-terminal protein extensions with every three tandem heptanucleotides encoding an RLKTQDS heptad. Comparison of 10 different OAc^+ K1 strains indicated that VNTR-encoded protein extensions may affect NeuO activity (16). However, direct analysis of the enzymatic activity in bacterial lysates is hampered by the high frequency of phase variation, and thus far, isolation of endogenous NeuO failed (15). In the present study, we succeeded in purification of active recombinant enzyme, forming the basis for a detailed biochemical characterization of NeuO. Functional and structural analyses of variants encoded by *neuO* alleles with 0, 12, 24, and 36 heptanucleotide repeats revealed that NeuO shares the typical features of an acyltransferase of the left-handed-β-helix (LβH) superfamily and assembles into hexamers that might be formed by dimerization of two trimers. In contrast to other members of the LβH family, NeuO activity is regulated by changes in the number of N-terminal RLKTQDS heptads and the catalytic efficiency gradually increases with the length of this VNTR-encoded protein extension.

EXPERIMENTAL PROCEDURES

Materials - *E. coli* K1 strains A160 and C375 were kindly provided by M. Achtman and F. Ørskov, respectively. pET expression vectors and *E. coli* BL21(DE3) were obtained from Novagen. CMP-Neu5Ac, Neu5Ac, colominic acid, 5,5'-dithiobis(2-nitrobenzoic acid) (Ellman's reagent), acetyl-, propionyl-, and butyryl-CoA were purchased from Sigma. Isolated sialic acid dimers, trimers, tetramers, pentamers, and hexamers were obtained from Nacalai-Tesque (Tokyo, Japan).

Cloning of the Capsule-specific O-Acetyltransferase of E. coli K1 - In parallel to Deszo *et al.* (16), we identified the capsule-specific *O*-acetyltransferase of *E. coli* K1 in the unfinished genome of strain RS218 (serotype O18ac:H7:K1, University of Wisconsin) by screening for homologies to the amino acid sequence of the sialic acid *O*-acetyltransferase of serogroup W-135 and Y meningococci (*oatWY*, accession number Y13969). A corresponding DNA-sequence of the identified gene was amplified by PCR from genomic DNA of the *O*-acetylation positive K1 strain C375. The gene was termed *oat*K1, and the sequence was submitted to the DDBJ/EMBL/GenBank™ databases under the accession number AJ783705. DNA sequencing revealed the presence of 18 heptanucleotide repeats and that the coding region is identical to *neuO* published by Deszo *et al.* (16).

Homology Modeling - The structural fold of NeuO was predicted by the program 3D-PSSM (21), a profile-based method that relies on multiple sequence and multiple structural alignments. Highest scores were obtained for the structural model based on homology to the *E. coli* maltose *O*-acetyltransferase(PDB accession number 1ocx), and all structural data were visualized by the PyMOL Molecular Graphics System.

Generation of NeuO Expression Plasmids - Constructs for the expression of wild-type NeuO were constructed in pET22b-Strep. In this modified pET22b vector, the original pelB leader sequence was exchanged by the sequence encoding an N-terminal Strep tag II (WSHPQFEK). *NeuO* without heptanucleotide repeats was amplified by PCR using genomic *E. coli* K1 DNA as a template and the primer pair MM275/MM282 (5'-CGGGATCCATGTCGTTTTCCGTTGATG-3', 5'-GTCCGCTCGAGTTGCGTGAGCTTCGCATG-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 pET22b-Strep, resulting in constructs with an N-terminal Strep tag II and a C-terminal hexahistidine tag. *NeuO* variants with heptanucleotide repeats were amplified by PCR using the primer pair MM301 (5'-CGGGATCCATGTTAAGACTCAAGACTC-3' and MM282 and genomic DNA of an *E. coli* K1 strain harboring an *neuO* gene with 36 heptanucleotide repeats as a template. The forward primer MM301 covers two heptanucleotide repeats and, therefore, can bind randomly within the repeat stretch resulting in a wide variety of PCR products containing 2–36 heptanucleotide repeats. PCR products of different length were subcloned by BamHI/XhoI sites into pET22b-Strep as described above. Constructs containing 12, 24, and 36 repeats were selected, and the number of heptanucleotide repeats was confirmed by sequencing.

Site-directed Mutagenesis - Site-directed mutagenesis was performed by PCR using the QuikChange site-directed mutagenesis kit (Stratagene) following the manufacturer's guidelines and the following primer pairs: MM295/MM296 to introduce the amino acid exchange H119A (5'-GCGTGCATCAGATGGCGCGCCTATATTTGATATTC-3' and 5'-GAATATCAAATATAGGCGCGCCATCTGATGCAC-GC-3') and MM278/MM279 for the mutation W143A (5'-CATTATATCTAGTTACGTAGCGGTAGGGAGAAATGTCTC-3' and 5'-GAGACATTTCTCCCTACCGCTACGTAACTAGAT-ATAATG-3'). BamHI/XhoI fragments of the corresponding PCR products were subcloned in pET22b-Strep for recombinant expression in *E. coli* BL21(DE3). The identity of all constructs was confirmed by sequencing.

Expression and Purification of Recombinant NeuO - Freshly transformed *E. coli* BL21(DE3) were cultivated at 37 °C in Luria-Bertani (LB)-medium containing 200 µg/ml carbenicillin. At an optical density (A_{600}) of 0.6 expression was induced by adding 0.1 mM isopropyl 1-thio-β-*D*-galactopyranoside and bacteria were harvested 2 h after induction. Purification was performed by immobilized Ni^{2+}-affinity chromatography using HisTrap-chelating HP columns (GE Healthcare). Bacteria were resuspended in binding buffer (20 mM Tris-HCl, pH 8.0/500 mM NaCl/40 mM imidazole) and lysed by sonication, and the soluble fraction was applied to a 1-ml column according to the manufacturer's instructions. After washing with 25 ml of binding buffer, proteins were eluted with a linear imidazole gradient of 40–500 mM imidazole in binding buffer. Imidazole was removed by gel filtration using a HiPrep 26/10 desalting column (GE Healthcare) equilibrated with 100 mM Tris-HCl, pH 8.0/150 mM NaCl.

In Vitro Assay for the Determination of NeuO Activity - NeuO activity was determined in a spectrophotometric

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

assay according to Alpers *et al.* (22). In a total volume of 100 µl, the standard reaction mixture contained 20 mM Tris-HCl, pH7.5, 25 mM EDTA, 4 mM 5,5'-dithiobis(2-nitrobenzoic acid), 2 mg/ml colominic acid, and 1 mM acetyl-CoA, and the reaction was initiated by adding 24 pmol of purified enzyme. The reaction was performed at 25 °C and monitored continuously at 405 nm in half-area 96-well plates (Greiner) using a PowerWave 340 microplate spectrophotometer (BioTek). For comparison of NeuO activity toward colominic acid, CMP-Neu5Ac, Neu5Ac, and Neu5Ac oligomers of different length, the concentrations of all acceptor substrates were normalized to equal sialic acid content (6.85 mM) and thereby to an equal number of acceptor sites. For quantification of the sialic acid content, oligomers were hydrolyzed to free sialic acid by treatment with 100 mM trifluoroacetic acid for 4 h at 80 °C followed by the thiobarbituric acid assay according to Skoza and Mohos (23) using Neu5Ac as standard.

Kinetic Analysis - The acceptor and donor kinetics were measured on purified NeuO variants using the spectrophotometric assay described above. The steady-state parameters, K_m and k_{cat}, were determined by holding one substrate at saturating concentration while the concentration of the second substrate was varied. Twelve different concentrations of donor (0.025–2 mM acetyl-CoA) and acceptor substrate (2–215 µM colominic acid) were used. Kinetic parameters were obtained by fitting the initial rate data to the Michaelis-Menten equation using nonlinear regression analysis with Prism 4.0 software (GraphPad Software, Inc.).

Analysis of NeuO Activity in Vivo - *NeuO* lacking the VNTR region was amplified by PCR using the primers HC412 (5'-GCGCGCGGATCCCGTTTTCCGTTGATGAT-AATGGG-3') and HC406 (5'-GCGCGCCTGCAGTATTT-ATTGCGTGAGCTTCGC-3') containing BamHI and PstI sites (underlined), respectively. The above described pET-based plasmids containing *neuO* encoding either wild-type or mutated variants with the amino acid exchanges H119A and W143A were used as a template. PCR products were ligated into BamHI/PstI sites of pQE32(Qiagen), and the identity of all constructs was confirmed by DNA sequencing. The resulting plasmids were transformed into *E. coli* K1 strain A160, which lacks endogenous *neuO*. Bacteria were grown on LB-agar containing 100 µg/ml ampicillin and 20 µl of a bacterial suspension (optical density at 600 nm = 0.1) were added to each well of a microtiter plate (Greiner) coated with poly-D-lysine (Sigma). After drying, bacteria were fixed with 0.05% glutaraldehyde in phosphate-buffer saline. Capsule expression and modification by *O*-acetylation was monitored in an enzyme-linked immunosorbent assay as described previously (24) using the following monoclonal antibodies (mAbs): mAb735 (25), which is specific for α2,8-linked polySia (OAc^+ and OAc^- forms), and mAb 58-5 (Monosan), which binds exclusively to the *O*-acetylated form of the K1 antigen.

Determination of the Minimal Acceptor Substrate Length - In a final volume of 50 µl, 200 µg of colominic acid was incubated with 1.3 µg of purified NeuO in the presence of [^{14}C]acetyl-CoA (GE Healthcare) for 1 h at 37°C. The reaction mixture was separated by high percentage PAGE, and the degree of polymerization of the smallest radioactively labeled acceptor molecule was determined by autoradiography. For generation of a radioactively labeled oligosialic acid marker ladder, the solid-phase fixed polysialyltransferase ST8SiaIV/PST-1 was autopolysialylated *in vitro* in the presence of CMP-[^{14}C]Neu5Ac (GE Healthcare) as described (26). After removal of unbound substrate, radioactively labeled sialic acid oligomers were released by partial endosialidase cleavage.

Analysis of Polysialic Acid by High Percentage PAGE - Sialic acid oligomers and polymers were separated on 25% polyacrylamide gels as described (27). Electrophoresis was performed for 4 h at 4 °C and 400 V, using 14-cm gels (0.8-mm thick). Immediately after electrophoresis, acrylamide gels were vacuumdried and exposed to Hyperfilm-MP (GE Healthcare).

SDS-PAGE, Silver Staining, and Immunoblotting – SDS-PAGE was performed under reducing conditions using 2.5% (v/v) β-mercaptoethanol. Silver staining and Western blot analysis were performed as described previously (28). For detection of the hexahistidine tag, penta-His antibody (Qiagen) was used at a concentration of 1 µg/ml. Strep-tagged proteins were detected with StrepTactin-AP conjugate (IBA) at a 1:4000 dilution.

Size-exclusion Chromatography - The quaternary structure of wild-type and mutant NeuO was determined on a *Superdex* 200 HR 10/30 column (GE Healthcare). The column was equilibrated with 100 mM Tris-HCl, pH 8.0/150 mM NaCl and calibrated with molecular mass standards (Sigma) thyroglobulin (669 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome C (12.4 kDa). 1 mg of affinity-purified wild-type or mutant NeuO was applied to the column, and the eluted protein was monitored by absorbance at 280 nm.

RESULTS

Purification of NeuO Variants with Variable N-terminal Protein Extensions - In all OAc^+ *E. coli* K1 strains investigated so far, *neuO* contains VNTRs located two nucleotides downstream of the start codon (16). As shown in Fig. 1*A*, full-length translation and thereby expression of active NeuO depend on repeat numbers that are divisible by three. In the case of full-length translation, every three tandem heptanucleotides encode an RLKTQDS heptad.

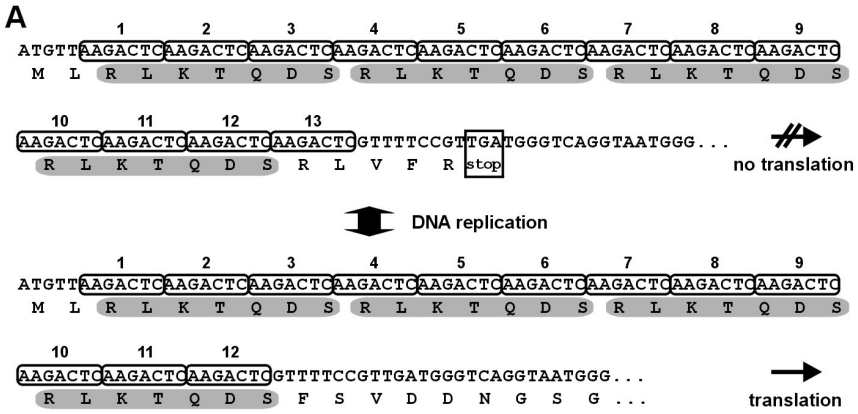
To investigate the role of N-terminal heptads on protein expression and activity, a set of *neuO* alleles with 0, 12, 24, and 36 heptanucleotide repeats was generated, and the corresponding translation products were termed NeuO+0, NeuO+12, NeuO+24, and NeuO+36, respectively. Each variant was expressed in *E. coli* BL21(DE3) with an N-terminal Strep- and a C-terminal hexahistidine tag (see Fig. 1*B* for schematic representation). Recombinant proteins were isolated by immobilized metal affinity chromatography (IMAC) using the C-terminal hexahistidine tag. Notably, high imidazole concentrations (≥450 mM) were required for the elution of all variants, indicating tight adsorption to the IMAC matrix. Because NeuO was unstable under these conditions, fractions eluted from the IMAC column were loaded immediately onto a Sephadex G-25 column to remove

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

imidazole. Using this purification protocol, all NeuO variants were purified to homogeneity (Fig. 1, *C* and *D*). Silver staining revealed single bands with an apparent molecular mass consistent with the calculated mass of the respective variant (Fig. 1, *B* and *D*). Western blot analysis using StrepTactin and anti-penta-His antibody (Fig. 1*D*, *middle* and *right panel*) revealed that *neuO* alleles with 12, 24, and 36 heptanucleotide units were translated in full-length proteins with the expected N-terminal extensions of 28, 56, and 84 amino acids, respectively, and no proteolytic cleavage or degradation of the heptad region was observed. From 1-liter cultures, 2.5 mg of purified NeuO+0 was obtained (Fig. 1*B*). Expression levels and protein yields increased with increasing numbers of VNTRs, suggesting that heptads contribute to protein stability. Because the insertion of 12, 24, and 36 nucleotide repeats increases the molecular mass of the translation products by 13, 25, and 37%, respectively, protein yields were compared on the molar level revealing a 2-fold higher protein yield for NeuO+36 than for NeuO+0 (Fig. 1*B*).

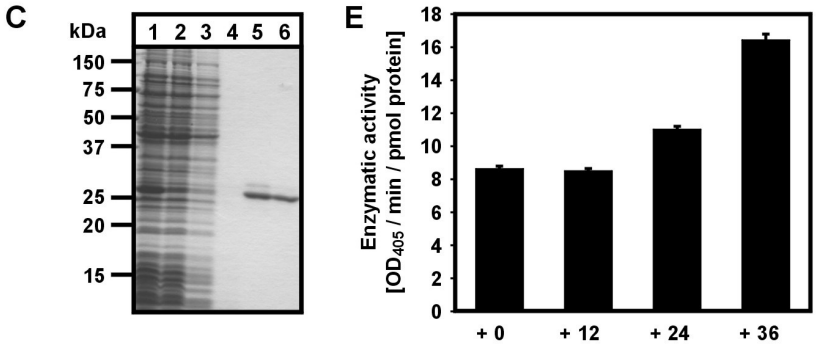
It is worth noting that none of the NeuO variants bound significantly to a StrepTactin column, although the N-terminal Strep tag was detectable by Western blotting (Fig. 1*D*, *middle panel*). This observation indicates that in native NeuO the N terminus is not accessible. Remarkably, this finding was irrespective of the presence or absence of heptad repeats (data not shown). For the unequivocal detection of full-length proteins, we kept the Strep tag in all used constructs. Direct comparison of NeuO with and without the Strep tag revealed that this tag did not affect NeuO activity (data not shown).

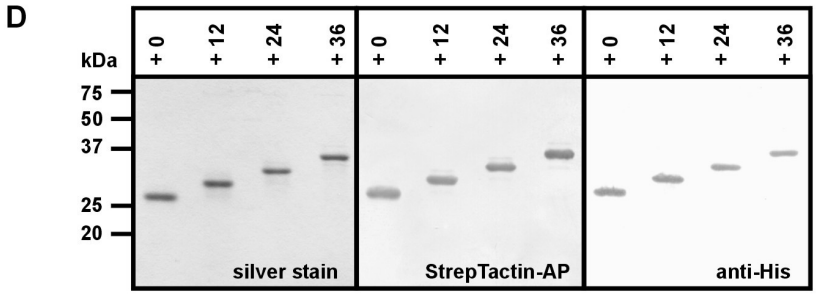
N-terminal Heptads Are Dispensable for Enzymatic Activity of Isolated NeuO - After demonstrating that heptanucleotide repeats are translated in stable protein extensions, we asked whether RLKTQDS heptads are essential for enzymatic activity of NeuO. To investigate this point, the *O*-acetyltransferase activity of NeuO variants with and without heptad repeats was determined in a modified Alpers assay (22). In this spectrophotometric assay, the transfer of acetyl groups from acetyl-CoA to polySia is followed by measuring the appearance of the free sulfhydryl group of CoA-SH with Ellman's reagent. Reactions were initiated by adding equimolar amounts of purified NeuO+0, NeuO+12, NeuO+24, or NeuO+36, and the optical density was monitored at 405 nm. Under the applied conditions of high donor and acceptor substrate concentrations, NeuO+0 and NeuO+12 showed comparable enzymatic activities (Fig. 1*E*), demonstrating



B

VNTRs	Translation product	RLKTQDS heptads	Additional amino acids	Molecular mass [kDa]	Yield [mg/L]	Relative molar yield
0	N–NeuO+0–C	0	0	26.8	2.5 ± 0.1	1.0
12	N–NeuO+12–C	4	28	30.2	3.7 ± 0.6	1.3
24	N–NeuO+24–C	8	56	33.5	5.0 ± 0.4	1.6
36	N–NeuO+36–C	12	84	36.8	6.8 ± 0.5	2.0





Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

TABLE 1
Kinetic parameters of NeuO variants with different numbers of heptad repeats
NeuO activity was measured in the spectrophotometric assay at pH 7.5 and 25 °C. Donor kinetics were determined using 0.025-2mM acetyl-CoA with colominic acid at 2 mg/ml. Acceptor kinetics were determined using 0.015–2 mg/ml colominic acid (corresponding to 0.05-6.85 mM sialic acid residues) with acetyl-CoA at 2 mM. The steady-state parameters, K_m and k_{cat}, were determined from initial velocity measurements from at least three independent experiments measured in triplicate and values are given as mean ± S.D.

NeuO variant	Donor substrate	Donor			Acceptor[a]		
		K_m	k_{cat}	k_{cat}/K_m	K_m	k_{cat}	k_{cat}/K_m
		mM	s^{-1}	$s^{-1}\ mM^{-1}$	*mM*	s^{-1}	$s^{-1}\ mM^{-1}$
+ 0	Acetyl-CoA	0.40 ± 0.01	0.51 ± 0.05	1.28	1.17 ± 0.14	1.99 ± 0.08	1.70
+ 12	Acetyl-CoA	0.22 ± 0.04	0.56 ± 0.07	2.54	1.10 ± 0.14	1.85 ± 0.14	1.68
+ 24	Acetyl-CoA	0.27 ± 0.01	1.07 ± 0.11	3.95	1.00 ± 0.04	2.57 ± 0.11	2.57
+ 36	Acetyl-CoA	0.22 ± 0.05	1.16 ± 0.16	5.26	1.06 ± 0.12	3.84 ± 0.10	3.62
+ 36	Propionyl-CoA	0.22 ± 0.02	0.09 ± 0.01	0.40			

[a] For acceptor kinetics, acceptor substrate concentrations were given as mM sialic acid residues.

that N-terminal RLKTQDS heptads are dispensable for enzymatic activity. Interestingly, increased activities were observed for NeuO+24 and NeuO+36, indicating that VNTR-encoded protein extensions influence the enzymatic properties of NeuO.

RLKTQDS Heptad Repeats Increase the Catalytic Efficiency of NeuO - To further investigate the effect of heptad repeats on the activity of NeuO, substrate kinetics of purified NeuO variants with and without RLKTQDS repeats were compared. In a first step, steady-state kinetic constants of NeuO+0, NeuO+12, NeuO+24, and NeuO+36 were determined for the donor substrate using 12 different concentrations of acetyl-CoA. All kinetic analyses were performed at 25 °C and yielded linear Lineweaver-Burk plots (see supplemental Figs. S1 and S2). The donor K_m value of NeuO+0 was determined to be 0.4 mM (Table 1), whereas all K_m values of heptad-containing variants were decreased by almost 50%. V_{max} and thereby k_{cat} values slightly increased with the number of heptads leading to a 2.3-fold higher k_{cat} for NeuO+36 compared with NeuO+0.

A lower K_m value in combination with an increased k_{cat} resulted in a 4-fold higher catalytic efficiency (k_{cat}/K_m^{donor}) of NeuO+36 compared with NeuO+0 (Table 1). Remarkably, k_{cat}/K_m^{donor} values increased linearly with the number of heptad repeats (Fig. 2*A*). These results suggest that the presence of heptad repeats result in gradual changes in enzyme conformation allowing an increase in enzyme-donor substrate affinity and facilitation of the reaction rate.

In a second step, kinetic parameters for the acceptor were determined using 12 different polySia concentrations. As an acceptor substrate, we used colominic acid, a commercially available mixture of oligomeric α2,8-linked sialic acid. Analysis of the batch used in the present study revealed oligomer lengths ranging from 2 to over 60 residues and an average degree of polymerization (DP) of 32. To acknowledge the fact that colominic acid is heterogeneous in polymer length, and that longer polymers contain more acceptor sites than shorter chains, acceptor concentrations were expressed as concentration of total sialic acids. Because acceptor affinities may vary with polySia chain length and also between internal and terminal sialic acids residues,

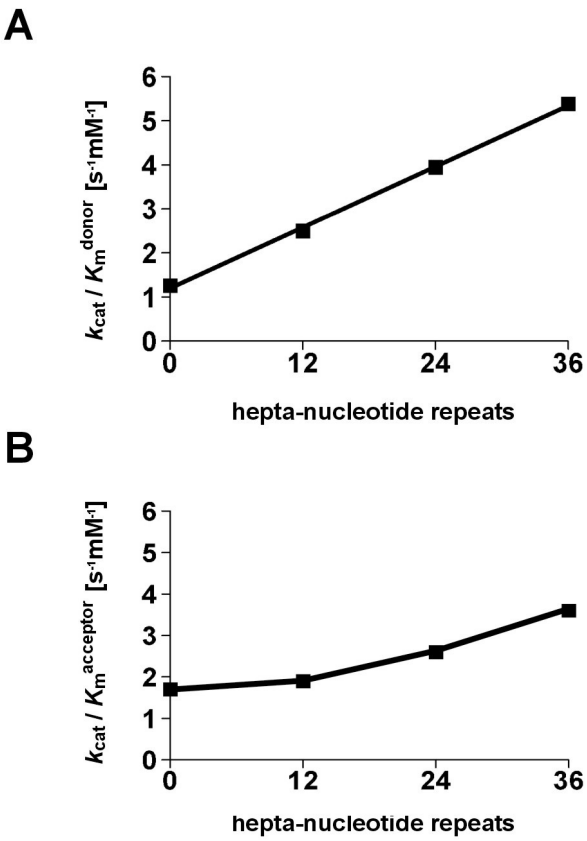


FIGURE 2. **Effect of heptad repeats on the catalytic efficiency of NeuO.** Catalytic efficiencies for the donor (*A*) and the acceptor substrate (*B*) were determined for NeuO encoded by alleles containing 0, 12, 24, and 36 heptanucleotide repeats (see Table 1), and the obtained *k*cat/*Km* values were plotted *versus* the number of heptanucleotide repeats.

FIGURE 1. **Purification of NeuO variants containing VNTR-encoded heptad repeats.** *A*, schematic representation of the VNTR region of *neuO* showing alleles with 13 (*top*) and 12 (*bottom*) 5'-AAGACTC-3' repeats. RLKTQDS heptads encoded by 3 heptanucleotide repeats are highlighted by *gray boxes*. *B*, schematic representation of NeuO variants encoded by alleles containing 0, 12, 24, and 36 repeats. The VNTR-encoded heptad region is shown in *black*, and N-terminal Strep- and C-terminal hexahistidine tag are shown in *light* and *dark gray*, respectively. All variants were expressed in *E. coli* BL21(DE3), and proteins were purified by IMAC on an Ni^{2+}-chelating column. The amount of purified protein obtained from a 1-liter culture is given as mean ± S.D. from three independent experiments. *C*, purification of NeuO lacking the VNTR region. After expression in *E. coli* BL21(DE3), NeuO+0 was isolated by IMAC. Purification steps were monitored by 14% SDS-PAGE and Coomassie staining, including aliquots of bacterial lysate (*lane 1*), flow through of the IMAC column (*lane 2*), wash fractions (*lanes 3* and *4*), NeuO eluted from IMAC column (*lane 5*), and purified NeuO after subsequent imidazole removal by gel filtration (*lane 6*). *D*, analysis of affinity-purified NeuO variants encoded by alleles containing 0, 12, 24, and 36 heptanucleotide repeats. Equal protein amounts of each variant were separated by 14% SDS-PAGE and analyzed by silver staining (*left panel*) and Western blotting using StrepTactin-AP for detection of the N-terminal Strep tag (*middle panel*) or anti-penta-His antibody for detection of the C-terminal hexahistidine tag (*right panel*). *E*, enzymatic activities of NeuO variants encoded by alleles containing the indicated number of heptanucleotide repeats. NeuO activity was determined spectrophotometrically in a modified Alpers assay using 24 pmol of purified protein.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

TABLE 2
Acceptor specificity of NeuO
Relative enzymatic activities of purified NeuO+0 and NeuO+36 were determined in the presence of different acceptor substrates. All acceptor substrates were used at a concentration of 6.85 mM of total sialic acids, and the activity obtained for colominic acid was set to 100%. Values are mean ± S.D. from two independent experiments measured in triplicate.

Acceptor	Relative enzymatic activity	
	NeuO+0	NeuO+36
	%	
Colominic acid	100.0 ± 1.3	100.0 ± 1.7
Neu5Ac	0.4 ± 0.2	0.9 ± 0.9
(-8-Neu5Ac-α2-)$_2$	0.9 ± 0.3	0.6 ± 0.4
(-8-Neu5Ac-α2-)$_3$	1.0 ± 0.7	0.1 ± 1.0
(-8-Neu5Ac-α2-)$_4$	0.2 ± 0.5	0.3 ± 0.8
(-8-Neu5Ac-α2-)$_5$	0.6 ± 1.0	0.2 ± 1.2
(-8-Neu5Ac-α2-)$_6$	1.1 ± 1.3	0.3 ± 0.7
CMP-Neu5Ac	0.6 ± 0.7	2.1 ± 1.3

acceptor K_m values will represent average values. For NeuO+0, the acceptor K_m value was determined to be 1.17 mM (Table 1). In contrast to the donor K_m values, no significant differences in $K_m^{acceptor}$ were observed for NeuO variants with and without heptad repeats. Therefore, the $k_{cat}/K_m^{acceptor}$ value was only 2-fold higher for NeuO+36 compared with NeuO+0 (Table 1), and the catalytic efficiency also increased with the number of repeats (Fig. 2*B*).

Donor Specificity of NeuO - In addition to acetyl-CoA, the alternative acyl donors propionyl-CoA and butyryl-CoA were assayed as substrates for NeuO. Butyryl-CoA turned out to be a very poor substrate, and even at a concentration of 1 mM, no transfer activity was observed. Propionyl-CoA, however, was used as a donor substrate. Similar to the results obtained with acetyl-CoA, increased enzymatic activities were monitored for NeuO+36 compared with NeuO+0 (data not shown). Kinetic parameters were evaluated only for NeuO+36, and results are summarized in Table 1 (bottom row).

Interestingly, identical K_m values were obtained for acetyl- and propionyl-CoA, indicating that the binding pocket for the donor substrate can accommodate the larger acyl-chain of propionyl-CoA. However, a 13-fold decrease in k_{cat} was observed for propionyl-CoA, indicating that the free energy for the transition state of the rate-limiting step is significantly higher for propionyl-CoA compared with acetyl-CoA.

Acceptor Specificity of NeuO - To investigate the acceptor substrate specificity of NeuO, the enzymatic activity of purified enzyme was monitored in the presence of sialic acid monomers (Neu5Ac), activated sialic acid (CMP-Neu5Ac), and a set of short α2,8-linked sialic acid oligomers with defined DPs in the range of dimer to hexamer. Because different kinetic properties were observed for NeuO variants with and without heptad repeats, acceptor substrate specificities were analyzed in parallel for NeuO+0 and NeuO+36. In contrast to colominic acid, none of the assayed compounds was used as an acceptor (Table 2). This result demonstrates that NeuO is highly specific for polySia and that even an α2,8-linked sialic acid hexamer is too short for NeuO-catalyzed *O*-acetylation.

To determine the minimal acceptor length of NeuO, we performed a radioactive incorporation assay. Purified NeuO+0 and NeuO+36 were incubated in the presence of ^{14}C-labeled acetyl-CoA and colominic acid as acceptor substrate. After separation by high percentage PAGE, reaction products were visualized by autoradiography. Because only those oligomers modified by *O*-acetylation will become radioactively labeled, the length of the shortest visible oligomer will represent the minimal acceptor length required. As a size marker, ^{14}C-labeled polySia was partially digested with endosialidase, resulting in an oligomer ladder starting with DP 3 (26). For the two enzyme variants, NeuO+0 and NeuO+36, reaction products of similar length were obtained and results are shown for NeuO+36 as an example (Fig. 3). Faint radioactive signals became visible for oligomers with DP 14 –16, and strong signals were obtained for oligomers >16 residues. Together, these results demonstrate high acceptor substrate specificity of NeuO for sialic acid oligomers >14 residues.

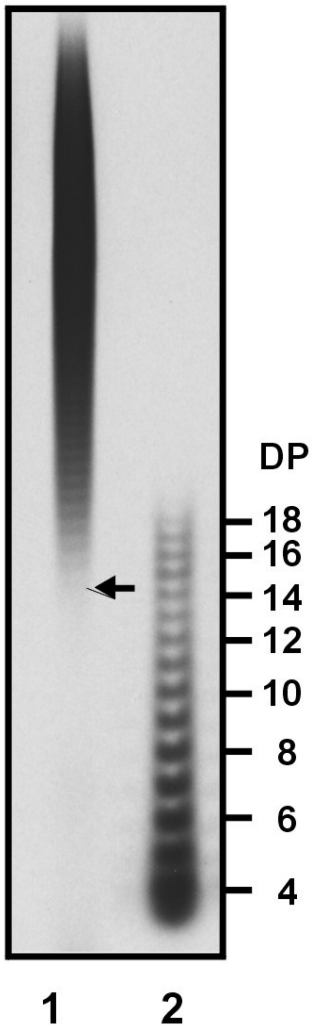


FIGURE 3. **Determination of the minimal acceptor length.** Purified NeuO+36 was incubated with colominic acid and [^{14}C]acetyl-CoA. The reaction mixture was separated by 25% PAGE, and radioactively labeled products were visualized by autoradiography (*lane 1*). The shortest detectable sialic acid oligomer is marked by an *arrow*. For size determination, radioactively labeled sialic acid oligomers with DP 3–18 were used as a marker (*lane 2*).

Homology Modeling of the Catalytic Part of NeuO - Primary sequence analysis of NeuO revealed similarity to the hexapeptide repeat family of acyltransferases (16, 29). Proteins of this family are characterized by tandem repeats of a hexapeptide with the consensus motif [LIV]-

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

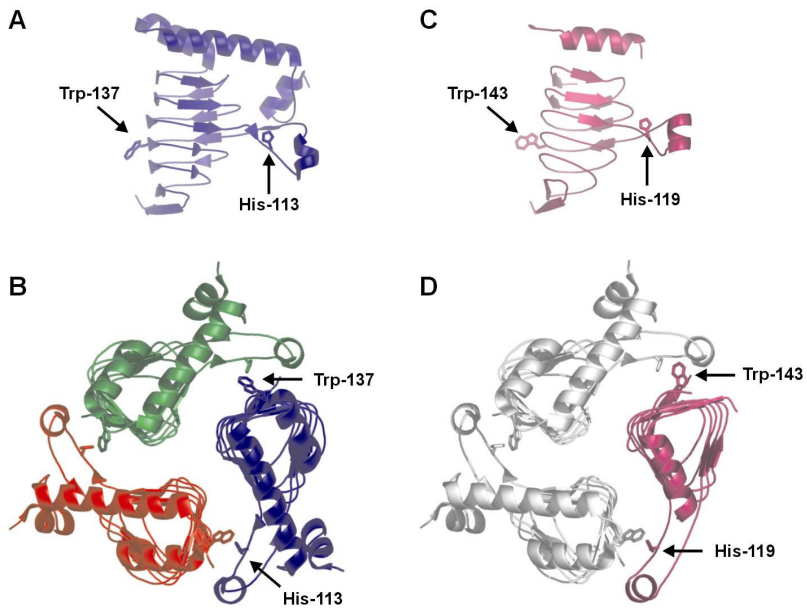


FIGURE 4. **Homology model of NeuO in comparison with the crystal structure of MAT.** *A*, ribbon representation of MAT (PDB accession number 1ocx; side view of the monomer). *B*, trimeric organization of MAT (*top view*), with the three subunits shown in *green*, *red*, and *blue*. The location of His-113 and Trp-137, which are part of the active site, are indicated by *arrows*. *C*, three-dimensional model of NeuO lacking VNTR-encoded heptads (side view of the monomer). The model was obtained by using the program 3D-PSSM (21) with the structure of MAT as a template. *D*, the predicted structure of the NeuO monomer (*purple*) was modeled into the trimer of MAT (*gray*). The locations of His-119 and Trp-143 of NeuO are indicated by *arrows*. All structural data were visualized by PyMOL.

[GAED]-X_2-[STAV]-*X*, which forms alternating short β-strands and tight turns resulting in a triangular left-handed β-helix (LβH) (30). All members of the hexapeptide acyltransferase superfamily assemble into catalytic trimers with three symmetrically arranged catalytic sites at the interface between the monomers.

To gain first insight into the structure of a polySia modifying *O*-acetyltransferase, homology models of NeuO lacking heptad repeats were generated using the program 3D-PSSM (21). The best structural match with an *E*-value of 2.2 x 10^{-4} (≥95% certainty) was obtained with the crystal structure of the maltose *O*-acetyltransferase (MAT) of *E. coli* (Protein Data Bank number 1ocx) (31) (see supplemental Fig. S3). The obtained model consists primarily of an LβH fold with five and one-third triangular coils (Fig. 4). The LβH domain covers the central and C-terminal part of NeuO (amino acids 61–216) and is interrupted by a single loop formed by residues 116–133 (Fig. 4*C*). For the first 60 amino acids of NeuO, however, only poor predictions were obtained, indicating structural differences between the N-terminal parts of NeuO and MAT. The LβH domain of MAT is N-terminally capped by three α-helices (residues 1–55), whereas only one α-helix was predicted for NeuO.

Notably, two amino acids that are part of the active site of MAT were highly conserved in NeuO. His-113, which corresponds to His-119 in NeuO, is located in the loop that protrudes between β-sheets four and five. This residue is proposed to abstract the proton from the hydroxyl group of the acceptor prior to acetyl transfer (31, 32). Trp-137, which corresponds to Trp-143 in NeuO, is located in a β-sheet of the LβH domain and is involved in binding of acetyl-CoA (31, 32). In the homo-trimer, the tryptophan of one subunit is located opposite to the catalytic histidine of the next subunit (Fig. 4*B*). In the predicted NeuO model, His-119 and Trp-143 are located at equivalent positions, suggesting similar functions in enzymatic catalysis.

Identification of Catalytic Residues - To prove whether His-119 and Trp-143 are crucial for enzymatic activity of NeuO, single alanine substitutions were performed. For both mutations, H119A and W143A, variants with and without 36 heptanucleotide repeats were generated, and the resulting proteins were expressed in *E. coli* BL21(DE3) with N-terminal Strep- and C-terminal hexahistidine tags. For either variant, similar expression levels than for the corresponding wild-type proteins were observed (data not shown), indicating that the introduced amino acid exchanges did not affect the stability of NeuO. The four mutant variants were purified to homogeneity (Fig. 5*A*) and assayed for enzymatic activity. However, in contrast to the wild-type forms of NeuO+0 and NeuO+36, no enzymatic activity was detected for the corresponding mutant forms (Fig. 5*B*). In addition to the *in vitro* analysis, the activity of NeuO variants was studied *in vivo*. To circumvent phase variation, the *in vivo* approach was restricted to variants lacking the VNTR region. Corresponding wild-type and mutant forms were cloned into the expression vector pQE32, and the resulting plasmids were transformed into *E. coli* K1 strain A160. As shown by PCR analysis, this strain lacks endogenous *neuO* (data not shown). Consequently, the expressed K1 capsule is OAc-negative as confirmed by an enzyme-linked immunosorbent assay using two capsule-specific antibodies: (i) mAb 735 (25), which binds to α2,8-linked polySia irrespective of the presence or absence of *O*-acetylation, and (ii) mAb 58-5, which recognizes exclusively the *O*-acetylated form of the K1 antigen (Fig. 5*C*). *O*-Acetylation of the K1 capsule of strain A160 was only observed after expression of wild-type NeuO. By contrast, no capsule *O*-acetylation was detected after expression of NeuO-H119A and NeuO-W143A, although K1 expression *per se* was not affected (Fig. 5*C*). Together, these results demonstrate that His-119 and Trp-143 are essential for NeuO activity *in vitro* and *in vivo*.

NeuO Assembles into Hexamers - The quaternary structure of all known hexapeptide acyltransferases is trimeric with the exception of the serine acetyltransferase (SAT) of *E. coli* and *Haemophilus influenzae*, which adopts a hexameric structure that is formed by dimers of trimers (33–35). Because trimerization is prerequisite for activity of hexapeptide acyltransferases, we asked whether the introduced mutations H119A and W143A induced conformational changes that prevent

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

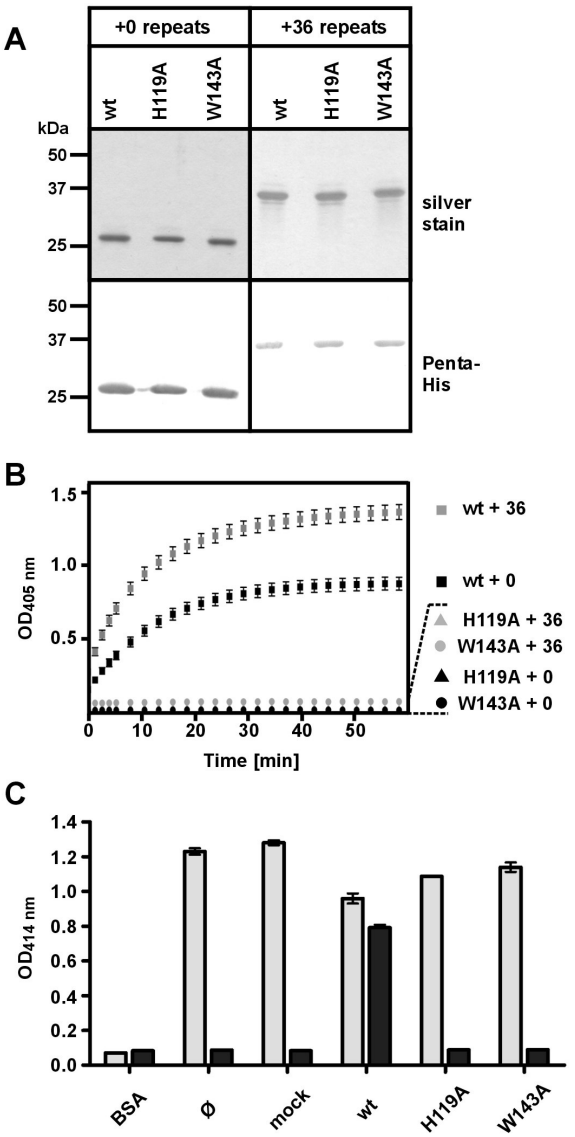


FIGURE 5. **Identification of amino acid residues critical for NeuO activity.** *A*, the single amino acid exchanges H119A and W143A were introduced into NeuO+0 (*left panel*) and NeuO+36 (*right panel*). After expression in *E. coli* BL21(DE3) and affinity purification on Ni^{2+}-chelating columns, wild-type (*wt*) and mutant forms were analyzed by 14% SDS-PAGE and silver staining (*upper panel*) or Western blotting using anti-penta-His antibody (*lower panel*). *B*, the enzymatic activities of NeuO+0 (*wt*), NeuO+36 (*wt+36*), and the corresponding mutants H119A and W143A were determined spectrophotometrically in the modified Alpers assay using equimolar concentrations of each purified protein. Data represent mean values ± S.D. of three independent experiments measured in triplicates. *C*, *in vivo* analysis of NeuO mutants. *E. coli* K1 strain A160 (*Ø*), which lacks endogenous *neuO*, was transformed with the empty vector pQE32 (*mock*), pQE32-based plasmids containing wild-type *neuO* lacking the VNTR region (*wt*), and the corresponding mutant variants H119A and W143A. Bacteria were analyzed for *O*-acetylated capsular polysaccharide by an enzyme-linked immunosorbent assay using mAb 58-5 (*black bars*). K1 capsule expression was controlled using mAb 735 (*gray bars*), which recognizes both OAc^+ and OAc^- forms of the K1 antigen. Negative controls were performed using bovine serum albumin instead of primary antibody.

oligomerization and thereby abolished enzymatic activity of NeuO. Therefore, the quaternary structure of wild- type and mutant NeuO was determined by size-exclusion chromatography, and results are shown for variants based on neuO alleles with 36 heptanucleotide repeats (Fig. 6). For wild-type NeuO+36, a molecular mass of 217 kDa was determined corresponding to an oligomeric state

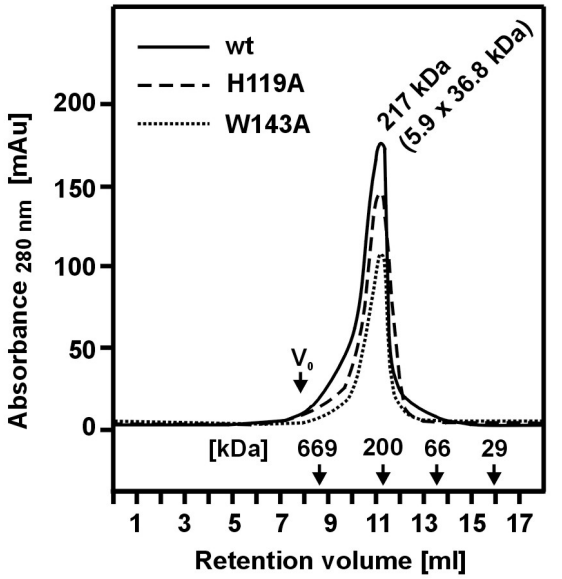


FIGURE 6. **Quaternary structure of NeuO.** The oligomeric state of purified NeuO+36 and the corresponding mutants H119A and W143A was analyzed by size exclusion chromatography on a Superdex 200 column. Void volume (*Vo*) and elution positions of standard proteins are indicated by *arrows*.

of 5.9. Although the shape of NeuO is not yet known, this result indicates that NeuO is another example of a hexapeptide acyltransferase that assembles into hexamers. Similar to the structure of SAT, NeuO hexamers might be formed by dimers of trimers. Notably, both mutants, H119A and W143A, showed an identical oligomeric state, demonstrating that the introduced amino acid exchanges did not interfere with complex assembly. In summary, our data indicate that alanine substitution of His-119 and Trp-143 did not result in substantial alterations in the overall structure of NeuO. However, exchange of either residue completely abolished activity, suggesting that both amino acids are part of the active site.

DISCUSSION

Surface sialic acids are a common component and virulence factor of a variety of pathogenic bacteria (36). In many cases, modification by O-acetylation has been observed and might play an important role in immunogenicity and survival in the host organism (14, 37–42). In the present study, we succeeded for the first time in purification and characterization of a polySia-specific O-acetyltransferase. NeuO, which is essential for O-acetylation of the capsular polysaccharide of *E. coli* K1, was purified to homogeneity using a C-terminal hexahistidine tag. Previous attempts to purify endogenous NeuO required detergent extraction, and, after partial purification on DEAE-Sepharose, further isolation steps abolished activity (15). By contrast, recombinant NeuO was obtained from the soluble protein fraction, and the isolated enzyme was enzymatically active as demonstrated in a spectrophotometric assay and by incorporation of radioactively labeled acetyl groups into polySia. A possible explanation for the different outcome during purification of endogenous and recombinant NeuO is that, in *E. coli* K1, efficient *O*-acetylation of the capsular polysaccharide might depend

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

on tight association of NeuO with the polySia biosynthesis and transport machinery. In this case, detergent might be required to dissociate NeuO from the membrane-bound protein complex, a step that can be omitted if NeuO is expressed in an *E. coli* strain lacking the K1 capsular polysaccharide gene complex.

The idea that *in vivo* NeuO-dependent *O*-acetylation occurs exclusively during or after polymerization of the polysaccharide but not on polySia building blocks is supported by our finding that *in vitro* neither Neu5Ac nor CMP-Neu5Ac served as acceptor substrates for NeuO. In line with the initial observation made by Higa and Varki (15), purified NeuO was highly specific for sialic acid oligomers with more than 14 residues, and efficient *O*-acetylation seemed to depend on oligomers with DP >16. However, to understand the structural basis of the acceptor specificity, further studies, including detailed product analyses and characterization of the protein-carbohydrate interactions are required.

Using homology modeling followed by mutational analysis, we characterized NeuO as a typical member of the superfamily of hexapeptide acyltransferases. Enzymes of this family use phosphopantothenyl-based cofactors to transfer acetyl, succinyl, or long chain fatty acyl groups to free hydroxyl or amino groups of a variety of acceptor substrates. All family members contain tandem-repeated imperfect copies of a hexapeptide repeat sequence with the consensus motif [LIV]-[GAED]-X_2-[STAV]-*X* (29, 30). Several members of the family have been crystallized, and the solved structures revealed that this characteristic hexapeptide repeat sequence encodes folding of a lefthanded parallel β-helix domain (31, 43–50). In all known cases, hexapeptide acyltransferases assemble into catalytic trimers with three symmetrical active sites, which are located at the interfaces of two subunits formed by a loop protruding out of one LβH domain embracing the LβH domain of the neighboring monomer. Based on the structural characteristics, the hexapeptide repeat family is also called LβH family and includes not only acyltransferases but also a zinc-dependent γ-carbonic anhydrase (51).

The homology modeling performed in the present study indicated close structural similarity of NeuO to MAT (31), which is closely related to the *lacA*-encoded galactoside *O*-acetyltransferase (GAT) of *E. coli* (45). Similar to MAT and GAT, the predicted model of NeuO contains a short LβH domain with five and one-third triangular coils and one protruding loop. Two amino acid residues, which are critical for activity of MAT and GAT, were found in equivalent positions in the NeuO model: His-119 located in the protruding loop and Trp-143 located within a β-sheet of the LβH domain. Subsequent single alanine substitutions completely abolished NeuO activity *in vitro* and *in vivo*, indicating important roles in catalysis. All acyltransferases of the LβH family crystallized so far posses a catalytic histidine located in a loop embracing the adjacent subunit. The predicted role of this residue is to abstract a proton from the acceptor hydroxyl group, facilitating the attack of the resulting carbonyl by the acyl-donor (32, 52). The second critical amino acid identified in the model of NeuO, Trp-143, is conserved only in a subset of acetyltransferases, including MAT, GAT, xenobiotic acetyltransferases, and the *Rhizobium leguminosarum* nodulation factor NodL (31, 32, 44, 53). In GAT, replacement of the corresponding Trp-139 by phenylalanine abolished the intrinsic fluorescence quench observed on acetyl-CoA binding (32), and the crystal structure of the binary complex of GAT with acetyl-CoA revealed a direct contact between the indole side chain of Trp-139 and the phosphopantothenyl arm of the cofactor (45). In addition to donor binding, Trp-139 may serve to position the catalytic histidine relative to the hydroxyl group of the acceptor and alter its p*Ka* (45). In line with this dual role in donor binding and positioning of the catalytic histidine, a complete loss of NeuO activity was observed when Trp-143 was substituted by alanine. Together, these results provided convincing evidence that NeuO consists primarily of an LβH fold with His-119 and Trp-143 as part of the active site.

Using size-exclusion chromatography, we demonstrated that NeuO assembles into hexamers. So far, the only example of an LβH-acyltransferase with a quaternary structure other than trimeric is SAT (54). Similar to NeuO, SAT of *E. coli* and *H. influenzae* adopts a hexameric structure, and the solved crystal structures revealed that two SAT trimers assemble into hexamers with the dimer of trimers interface formed by the N-terminal α-helical domain (34, 35). The observation that a Strep tag placed at the N terminus of NeuO was not accessible for purification might be an indication that, also in NeuO, the N-terminal domain is involved in dimerization of NeuO trimers.

Recently, several genes encoding sialic acid specific *O*-acetyltransferases have been cloned by us and others: (i) *oatC* responsible for *O*-acetylation of the α2,9-linked polySia capsule of *N. meningitidis* of serogroup C (55); (ii) *oatWY*, involved in *O*-acetylation of sialic acids within the galactose and glucose containing heteropolymeric polySia capsules of serogroup W-135 and Y meningococci, respectively (55); (iii) *neuO* of *E. coli* K1 (16); (iv) *neuD* of group B streptococci required for *O*-acetylation of terminal α2,3-linked sialic acids capping the group B streptococci capsule (56); and (v) *neuD* of *Campylobacter jejuni*, which was shown to *O*-acetylate terminal α2,8- linked sialic acids of the bacterial lipo-oligosaccharide (42). With the notable exception of OatC, which shares no sequence similarity to any known protein in the data base, all other sialic acid-specific *O*-acetyltransferases were found to contain hexapeptide repeat sequences, indicating that these proteins are all members of the LβH family. However, the NeuD proteins might form a separate branch within the family, because amino acid sequence alignments revealed the lack of a potential catalytic histidine, which might be substituted by a lysine residue (56).

Although NeuO shares the typical features of the LβH-family, the present study revealed a unique regulatory mechanism based on changes in VNTR-encoded N-terminal protein extensions. The *neuO* gene is characterized by the presence of a variable number of heptanucleotide repeats within the coding region (16). Because only repeat numbers that are a multiple of three allow full-length translation and thereby capsule *O*-acetylation, changes in the overall number provide a reversible switch between an on and off phase. Oscillation between OAc^+ and OAc^- variants might permit

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

adaptation to changes in environmental conditions. However, our results demonstrate that this is not a simple "all-or-none" mechanism. By expressing *neuO* alleles containing 0, 12, 24, and 36 heptanucleotide repeats, we found that VNTRs are translated in stable N-terminal protein extensions that affect the actual enzymatic activity of the translation product. Every three heptanucleotide repeats encode an RLKTQDS heptad, and, in the presence of 36 heptanucleotide repeats, the translation product is N-terminally elongated by 12 heptads, which account for 37% of the total molecular mass. The presence of RLKTQDS heptads was not essential for NeuO activity but decreased the K_m for acetyl-CoA by ~50% without affecting the K_m for the acceptor substrate. Remarkably, a linear correlation between the catalytic efficiency (k_{cat}/K_m^{donor}) and the number of repeats was observed. Thus, stochastic genetic variation generated by VNTRs during the on phase is translated in protein variation that provides a gradual modification of NeuO activity.

k_{cat}/K_m may be considered as a measure of the apparent first-order rate constant for enzyme-substrate interaction and reflects the affinity between enzyme and substrate to form a complex. Thus, the results are consistent with the interpretation that the presence of heptad repeats leads to a change in enzyme conformation with the resultant increase in enzyme donor substrate affinity and facilitation of the reaction rate.

As suggested by Deszo *et al.* (16), heptad repeats may assemble into a triple coiled-coil. In the hexameric NeuO complex, variation in the length of such a coil might gradually change the subunit arrangement and, thereby, the conformation of the catalytic sites formed at the subunit interfaces of each trimer. However, the complete arrangement of all six subunits and the structure of the heptad repeats will only unequivocally be solved by analyzing the crystal structure of the NeuO complex.

In vivo, direct analysis of NeuO activity is hampered by frequent oscillation between on and off phases. However, comparison of 10 different OAc^+ K1 strains indicated that VNTR-encoded protein extensions affect NeuO activity also *in vivo*, and the highest activities were found in strains expressing *neuO* alleles with >21 heptanucleotide repeats (16). The analysis of capsular polysaccharides from four individual OAc^+ K1 strains revealed different degrees of *O*-acetylation ranging from 5 to 85% (14, 15). Remarkably, the lowest *O*-acetyl content and lowest activity were found in a strain without detectable phase variation (14, 15), indicating that, *in vivo*, modulation of NeuO activity has an impact on the degree of capsule *O*-acetylation.

Acknowledgments - We thank Rita Gerardy-Schahn for helpful discussions and continuous support and Katharina Stummeyer for her assistance with PyMOL

REFERENCES

1. 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
2. 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
3. McCracken, G. H., Sarff, L. D., Glode, M. P., Mize, S. G., Schiffer, M. S., Robbins, J. B., Gotschlich, E. C., Ørskov, I., and Ørskov, F. (1974) *Lancet* **2**, 246–250
4. Bonacorsi, S., and Bingen, E. (2005) *Int. J. Med. Microbiol.* **295**, 373–381
5. Kim, K. S. (2003) *Nat. Rev. Neurosci.* **4**, 376–385
6. Xie, Y., Kim, K. J., and Kim, K. S. (2004) *FEMS Immunol. Med. Microbiol.* **42**, 271–279
7. Fowler, M. I., Weller, R. O., Heckels, J. E., and Christodoulides, M. (2004) *Cell Microbiol.* **6**, 555–567
8. Harvey, D., Holt, D. E., and Bedford, H. (1999) *Semin. Perinatol.* **23**, 218–225
9. Saez-Llorens, X., and McCracken, G. H., Jr. (2003) *Lancet* **361**, 2139–2148
10. Pluschke, G., Mayden, J., Achtman, M., and Levine, R. P. (1983) *Infect. Immun.* **42**, 907–913
11. Leying, H., Suerbaum, S., Kroll, H. P., Stahl, D., and Opferkuch, W. (1990) *Infect. Immun.* 58, 222–227
12. Rohr, T. E., and Troy, F. A. (1980) *J. Biol. Chem.* **255**, 2332–2342
13. Mühlenhoff, M., Eckhardt, M., and Gerardy-Schahn, R. (1998) *Curr. Opin. Struct. Biol.* **8**, 558–564
14. Ørskov, F., Ørskov, I., Sutton, A., Schneerson, R., Lin, W., Egan, W., Hoff, G. E., and Robbins, J. B. (1979) *J. Exp. Med.* **149**, 669–685
15. Higa, H. H., and Varki, A. (1988) *J. Biol. Chem.* **263**, 8872–8878
16. Deszo, E. L., Steenbergen, S. M., Freedberg, D. I., and Vimr, E. R. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5564–5569
17. Frasa, H., Procee, J., Torensma, R., Verbruggen, A., Algra, A., Rozenberg-Arska, M., Kraaijeveld, K., and Verhoef, J. (1993) *J. Clin. Microbiol.* **31**, 3174–3178
18. Stummeyer, K., Schwarzer, D., Claus, H., Vogel, U., Gerardy-Schahn, R., and Mühlenhoff, M. (2006) *Mol. Microbiol.* **60**, 1123–1135
19. van Belkum, A., Scherer, S., van Alphen, L., and Verbrugh, H. (1998) *Microbiol. Mol. Biol. Rev.* **62**, 275–293
20. Viguera, E., Canceill, D., and Ehrlich, S. D. (2001) *EMBO J.* **20**, 2587–2595
21. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) *J. Mol. Biol.* **299**, 499–520
22. Alpers, D. H., Appel, S. H., and Tomkins, G. M. (1965) *J. Biol. Chem.* **240**, 10–13
23. Skoza, L., and Mohos, S. (1976) *Biochem. J.* **159**, 457–462
24. Vogel, U., Morelli, G., Zurth, K., Claus, H., Kriener, E., Achtman, M., and Frosch, M. (1998) *J. Clin. Microbiol.* **36**, 2465–2470
25. Frosch, M., Görgen, I., Boulnois, G. J., Timmis, K. N., and Bitter-Suermann, D. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 1194–1198
26. Mühlenhoff, M., Eckhardt, M., Bethe, A., Frosch, M., and Gerardy-Schahn, R. (1996) *EMBO J.* **15**, 6943–6950
27. Pelkonen, S., Häyrinen, J., and Finne, J. (1988) *J. Bacteriol.* **170**, 2646–2653
28. Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S., Lottspeich, F., and Frosch, M. (1995) *Mol. Microbiol.* **16**, 441–450
29. Vaara, M. (1992) *FEMS Microbiol. Lett.* 76, 249–254
30. Jenkins, J., and Pickersgill, R. (2001) *Prog. Biophys. Mol. Biol.* **77**, 111–175
31. Lo Leggio, L., Dal Degan, F., Poulsen, P., Andersen, S. M., and Larsen, S. (2003) *Biochemistry* **42**, 5225–5235
32. Lewendon, A., Ellis, J., and Shaw, W. V. (1995) *J. Biol. Chem.* **270**, 26326–26331
33. Hindson, V. J., Moody, P. C., Rowe, A. J., and Shaw, W. V. (2000) *J. Biol. Chem.* **275**, 461–466
34. Pye, V. E., Tingey, A. P., Robson, R. L., and Moody, P. C. (2004) *J. Biol. Chem.* **279**, 40729–40736
35. Olsen, L. R., Huang, B., Vetting, M. W., and Roderick, S. L. (2004) *Biochemistry* **43**, 6013–6019
36. Vimr, E., and Lichtensteiger, C. (2002) *Trends Microbiol.* **10**, 254–257
37. Lewis, A. L., Nizet, V., and Varki, A. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101**, 11123–11128
38. Bhattacharjee, A. K., Jennings, H. J., Kenny, C. P., Martin, A., and Smith, I. C. (1976) *Can. J. Biochem.* **54**, 1–8
39. Lemercinier, X., and Jones, C. (1996) *Carbohydr. Res.* **296**, 83–96
40. Gamian, A., Romanowska, E., Ulrich, J., and Defaye, J. (1992) *Carbohydr. Res.* **236**, 195–208
41. Feng, L., Senchenkova, S. N., Tao, J., Shashkov, A. S., Liu, B., Shevelev, S. D., Reeves, P. R., Xu, J., Knirel, Y. A., and Wang, L. (2005) *J. Bacteriol.* **187**, 758–764
42. Houliston, R. S., Endtz, H. P., Yuki, N., Li, J., Jarrell, H. C., Koga, M., van Belkum, A., Karwaski, M. F., Wakarchuk, W. W., and Gilbert, M. (2006) *J. Biol. Chem.* **281**, 11480–11486
43. Raetz, C. R., and Roderick, S. L. (1995) *Science* 270, 997–1000
44. Beaman, T. W., Sugantino, M., and Roderick, S. L. (1998) *Biochemistry* **37**, 6689–6696
45. Wang, X. G., Olsen, L. R., and Roderick, S. L. (2002) *Structure* **10**, 581–588
46. Olsen, L. R., and Roderick, S. L. (2001) *Biochemistry* **40**, 1913–1921

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

47. Beaman, T. W., Binder, D. A., Blanchard, J. S., and Roderick, S. L. (1997) *Biochemistry* **36**, 489–494
48. Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C., and Bourne, Y. (1999) *EMBO J.* **18**, 4096–4107
49. Sugantino, M., and Roderick, S. L. (2002) *Biochemistry* **41**, 2209–2216
50. Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F., and Bourne, Y. (2001) *J. Biol. Chem.* **276**, 11844–11851
51. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. C. (1996) *EMBO J.* **15**, 2323–2330
52. Wyckoff, T. J., and Raetz, C. R. (1999) *J. Biol. Chem.* **274**, 27047–27055
53. Dunn, S. M., Moody, P. C., Downie, J. A., and Shaw, W. V. (1996) *Protein Sci.* **5**, 538–541
54. Johnson, C. M., Roderick, S. L., and Cook, P. F. (2005) *Arch. Biochem. Biophys.* **433**, 85–95
55. Claus, H., Borrow, R., Achtman, M., Morelli, G., Kantelberg, C., Longworth, E., Frosch, M., and Vogel, U. (2004) *Mol. Microbiol.* **51**, 227–239
56. Lewis, A. L., Hensler, M. E., Varki, A., and Nizet, V. (2006) *J. Biol. Chem.* **281**, 11186–11192

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

SUPPLEMENTAL FIGURES

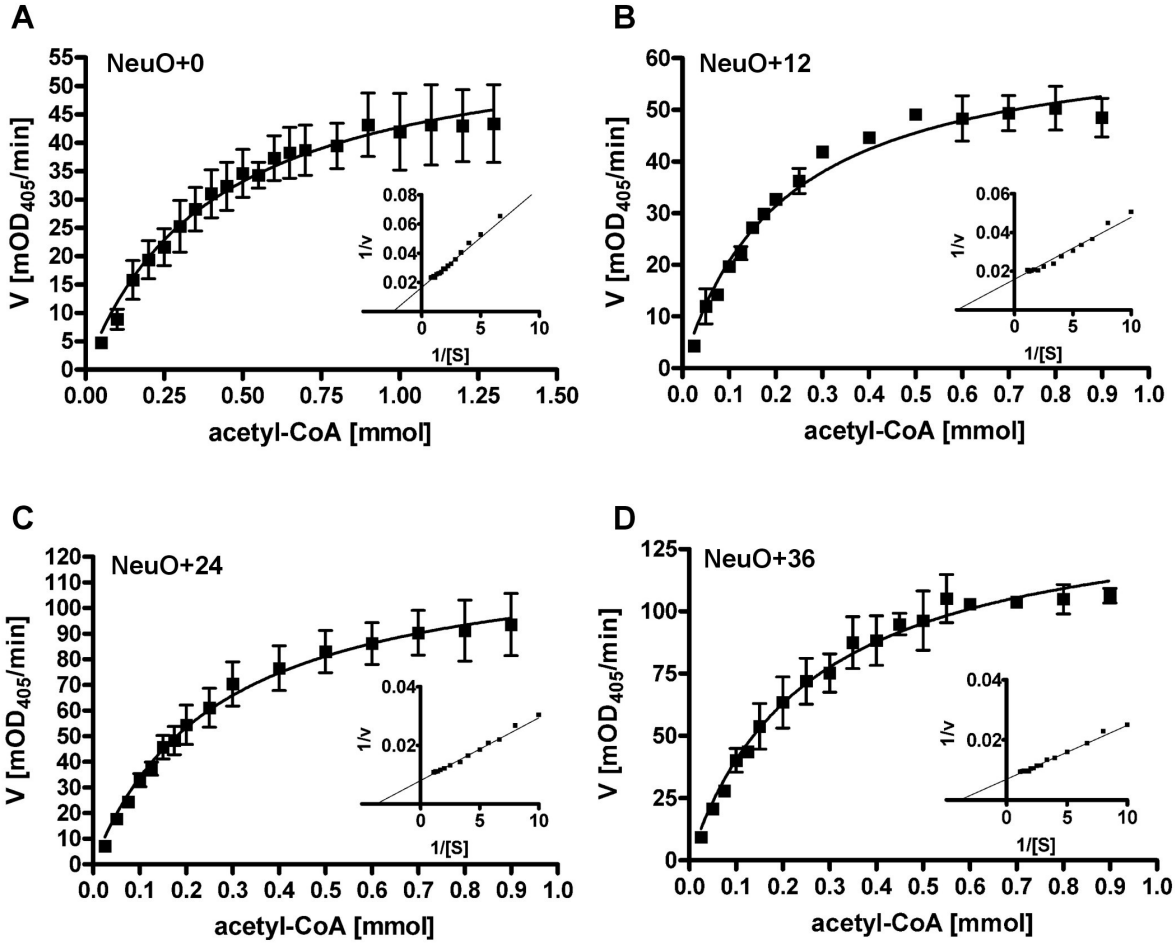


Fig. S1: Donor substrate kinetics of NeuO variants with different numbers of heptad repeats. NeuO activity was measured in the spectrophotometric assay at pH 7.5 and 25°C using equimolar amounts of purified enzyme variants. Donor kinetics were determined using 0.025-2 mM acetyl-CoA with colominic acid at 2 mg/ml. The initial reaction velocity was plotted versus the donor concentration. Data were acquired from at least three independent experiments measured in triplicate and values are given as mean ± S.D.. The corresponding Lineweaver-Burk plots are shown as inserts. Results obtained for NeuO variants encoded by alleles containing 0, 12, 24, and 36 heptanucleotide repeats are shown in (A), (B), (C), and (D), respectively.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

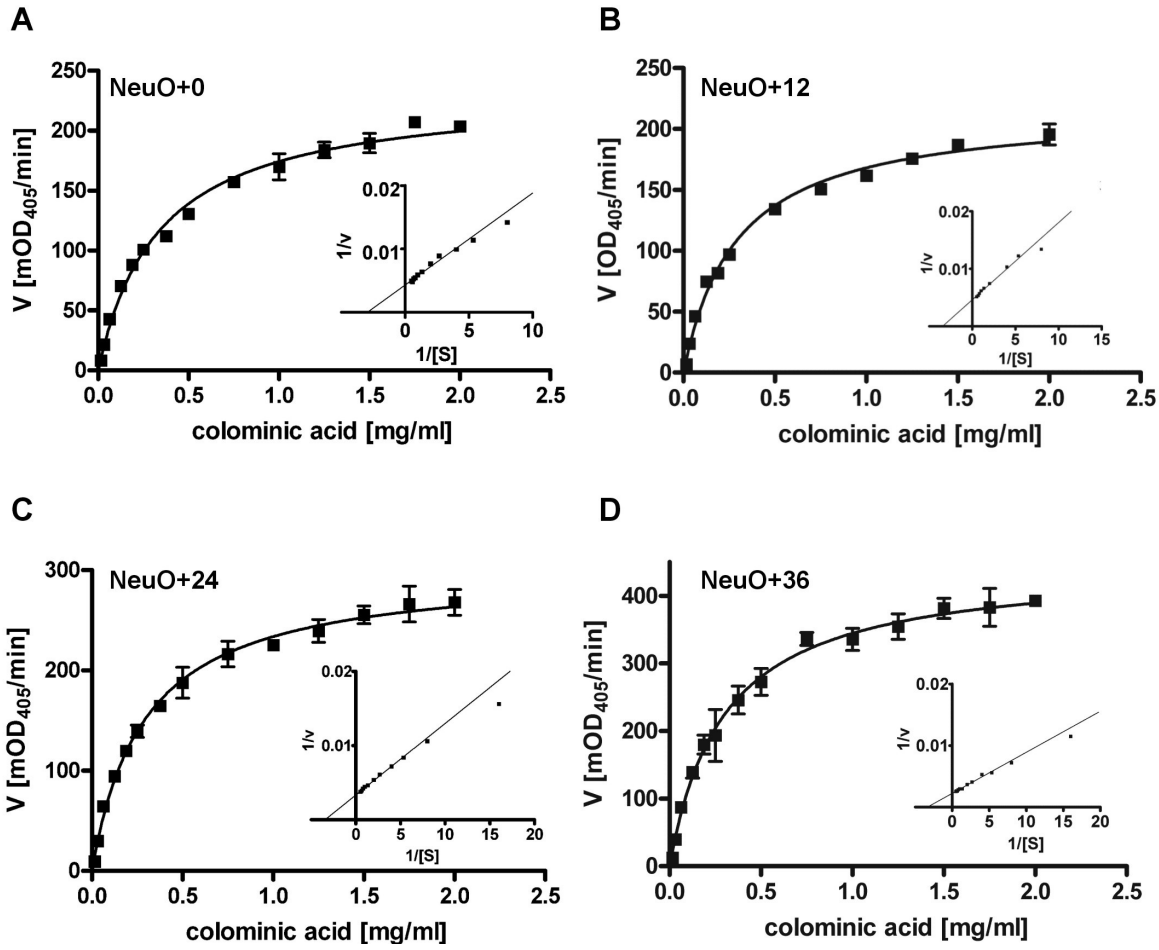


Fig. S2: Acceptor substrate kinetics of NeuO variants with different numbers of heptad repeats. NeuO activity was measured in the spectrophotometric assay at pH 7.5 and 25°C using equimolar amounts of purified enzyme variants. Acceptor kinetics were determined using 0.015-2 mg/ml colominic acid (corresponding to 0.05-8.85 mM sialic acid) with acetyl-CoA at 2 mM. The initial reaction velocity was plotted versus the acceptor concentration. Data were acquired from at least three independent experiments measured in triplicate and values are given as mean ± S.D.. The corresponding Lineweaver-Burk plots are shown as inserts. Results obtained for NeuO variants encoded by alleles containing 0, 12, 24, and 36 heptanucleotide repeats are shown in (A), (B), (C), and (D), respectively.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

A

Protein	PDB	PSSM e-value	% certainty
maltose-O-acetyltransferase (MAT)	1ocx	0.000221	95
galactoside acetyltransferase (GAT)	1kqa	0.0027	95
xenobiotic acetyltransferase (PAXAT)	1xat	0.0598	90
streptogramin A acetyltransferase (VatD)	1kk6	0.16	80

B

```
                10        20          30        40        50        60
           ---------+---------+-   --------+---------+---------+---------+--  -
NeuO       MSFSVDDNGSGNVFVCGDLVN---SKENKVQFNGNNNKLIIEDDVECRWLTVIFRGDNNYVRIHK--N
MAT        MSTEKEKMIAGELYRSADETLSRDRLRARQLIHRYNHSLAEEHTLRQQILADLF---------GQVTE

                70        80        90        100       110       120       130
           ------+---------+---------+---------+---------+---------+---------+-
NeuO       SKIKGDIVATKGSKVIIGRRTTIGAGFEVVTDKCNVTIGHDCMIARDVILRASDGHPIFDIHSKKRIN
MAT        AYIEPTFRCDYGYNIFLGNNFFANFDC-VMLDVCPIRIGDNCMLAPGVHIYTAT-HPIDPVARNSGAE

                140       150       160       170       180       190
           --------+---------+---------+---------+---------+---------+---------
NeuO       WAKDIIISSYVWVGRNVSIMKGVSVGSGSVIGYGSIVTKDVPSMCAAAGNPAKIIKRNIIWARTDKAE
MAT        LGKPVTIGNNVWIGGRAVINPGVTIGDNVVVASGAVVTKDVPDNVVVGGNPARIIKKL----------

           200       210
           +---------+------
NeuO       LISDDKRCSSYHAKLTQ
MAT        -----------------
```

C

```
                10          20        30        40        50        60
           ---------+--------   -+---------+---------+---------+---------+-----
NeuO       MSFSVDDNGSGNVFVCGD---LVNSKENKVQFNGNNNKLIIEDDVECRWLTVIFRGDNNYVRIHKNSK
GAT        NMPMTERIRAGKLFTDMCEGLPEKRLRGKTLMYEFNHSHPSEVEKRESLIKEMFA------TVGENAW

               70        80        90        100       110       120       130
           ----+---------+---------+---------+---------+---------+---------+---
NeuO       IKGDIVATKGSKVIIGRRTTIGAGFEVVTDKCNVTIGHDCMIARDVILRASDGHPIFDIHSKKRINWA
GAT        VEPPVYFSYGSNIHIGRNFYANFNLTIV-DDYTVTIGDNVLIAPNVTL-SVTGHPVHHELRKNGEMYS

                140       150       160       170       180       190       200
           ------+---------+---------+---------+---------+---------+---------+-
NeuO       KDIIISSYVWVGRNVSIMKGVSVGSGSVIGYGSIVTKDVPSMCAAAGNPAKIIKRNIIWARTDKAELI
GAT        FPITIGNNVWIGSHVVINPGVTIGDNSVIGAGSIVTKDIPPNVVAAGVPCRVIREINDRDKHYYFKDY

                200
           --------+------
NeuO       SDDKRCSSYHAKLTQ
GAT        KVES------------
```

Fig. S3: Scores and sequence alignments obtained for NeuO homology building using the program 3D-PSSM. (A) List of proteins with lowest E-values reflecting best structural matches with NeuO. Please note that all four proteins belong to the left-handed beta-helix family of acetyltransferases. (B) Sequence alignment between NeuO and MAT. (C) Sequence alignment between NeuO and GAT (aa 2-201). Amino acid numbering for NeuO is given on top. Highly conserved amino acid residues that are essential for activity are shown in grey boxes.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY;* VOL. **282**, NO. 30, pp. 22217-22227, July 27, 2007

Chapter 3

O-acetyltransferase gene *neuO* is segregated according to phylogenetic background and contributes to environmental desiccation resistance in *Escherichia coli* K1

This manuscript was first submitted on October, 29th 2008. A revised form is currently prepared for resubmission to *Environmental Microbiology*:

'*O*-acetyltransferase gene *neuO* is segregated according to phylogenetic background and contributes to environmental desiccation resistance in *Escherichia coli* K1'

Ines L. Mordhorst[1], Heike Claus[1], Christoph Schoen[1], Johannes Elias[1], Julia Batzilla[1], Ulrich Dobrindt[2], Christa Ewers[3], Lothar H. Wieler[3], **Anne K. Bergfeld**[4], Martina Mühlenhoff[4], and Ulrich Vogel[1+]

[1]University of Würzburg, Institute for Hygiene and Microbiology; [2]University of Würzburg, Institute for Molecular Infection Biology; [3]Free University Berlin, Institute of Microbiology and Epizootics; [4]Hannover Medical School, Department of Cellular Chemistry

[+]To whom correspondence should be addressed: Ulrich Vogel; Institute for Hygiene and Microbiology; University of Würzburg; Josef-Schneider-Str. 2/ E1; 97080 Würzburg, Germany; Tel.: 49-931-201 46802; Fax: 49-931-201 46445; E-mail: uvogel@hygiene.uni-wuerzburg.de

Manuscript-ID: EMI-2008-0851

Keywords: bacteria, microbial genetics, pathogen ecology, population biology/ clonal structure, evolution/ evolutionary processes/ gene transfer/ mutation

Preface - About the Manuscript

In this study, a population genetics approach was performed to investigate the distribution of the *O*-acetyltransferase gene *neuO* among *E. coli* K1 isolates from avian and human sources. Therefore, a total number of 184 *E. coli* K1 strains was analyzed by multilocus sequence typing at the *Institute for Hygiene and Microbiology* of the *University of Würzburg*. Evolutionary aspects concerning the presence of the CUS-3 prophage which harbors *neuO* were determined and the contribution of *O*-acetylation to desiccation resistance of *E. coli* K1 in the inanimate environment was investigated. Sequence analysis of the *neuO* gene in 103 K1 strains revealed the presence of five different allelic variants. Three out of the seven observed polymorphic sites were non-synonymous, leading to three distinct enzyme variants which differ at primary sequence level. My contribution to this study was to investigate whether these NeuO variants show different kinetic properties. After recombinant expression and purification of all variants, I performed a comparative kinetic analysis by determining the steady-state parameters K_m and k_{cat} for each variant.

O-acetyltransferase gene *neuO* is segregated according to phylogenetic background and contributes to environmental desiccation resistance in *Escherichia coli* K1

Running Title: Population biology of *neuO* in *E. coli* K1

Ines L. Mordhorst[1], Heike Claus[1], Christoph Schoen[1], Johannes Elias[1], Julia Batzilla[1§], Ulrich Dobrindt[2], Christa Ewers[3], Lothar H. Wieler[3], **Anne K. Bergfeld**[4], Martina Mühlenhoff[4], and Ulrich Vogel[1*]

[1] University of Würzburg, Institute for Hygiene and Microbiology;
[2] University of Würzburg, Institute for Molecular Infection Biology;
[3] Free University Berlin, Institute of Microbiology and Epizootics;
[4] Hannover Medical School, Department of Cellular Chemistry
[§] Present address: Max-von-Pettenkofer-Institute, Munich, Germany

[*] Correspondent footnote
Ulrich Vogel
Institute for Hygiene and Microbiology
University of Würzburg
Josef-Schneider-Str. 2 / E1
97080 Würzburg
Germany
Tel.: ++49-(0)931-201 46802
Fax: ++49-(0)931-201 46445

SUMMARY

Escherichia coli K1 causes disease in humans and birds. Its polysialic acid capsule can be *O*-acetylated via the acetyltransferase NeuO encoded by prophage CUS-3. The role of capsule *O*-acetylation in ecological adaptation or pathogenic invasion of *E. coli* K1 is largely unclear. A population genetics approach was performed to study the distribution of *neuO* among *E. coli* K1 isolates from human and avian sources. Multilocus sequence typing revealed 40 different sequence types (ST) among 184 *E. coli* K1 strains. The proportion of ST95 complex (STC95) was 44%. *NeuO* was found in 98% of the STC95 strains, but only in 24% of the strains with other STs. Grouping of multilocus sequence types and prophage genotypes revealed a segregation of prophage types according to MLST types, suggesting co-evolution of CUS-3 and housekeeping genes. Within the STC95, which is known to harbor both human and avian pathogenic isolates, CUS-3 genotypes were shared irrespective of host species. Strains, which expressed functional NeuO constitutively better withstood desiccation *in vitro* than strains expressing a non-functional NeuO:H119A mutant, suggesting an environmental role of *neuO*-mediated K1 capsule *O*-acetylation.

INTRODUCTION

Escherichia coli K1 strains are intestinal commensals of mammalians and birds, but may cause urinary tract infections, newborn meningitis or bacteremia in humans as well as colisepticemia in birds. The capsular polysaccharide of *E. coli* K1, a major pathogenicity factor, is composed of the poorly immunogenic α-2,8-linked N-acetyl neuraminic acid (also termed polysialic acid) (Barry and Goebel, 1957; Dewitt and Rowe, 1961; Kundig *et al.*, 1971; Mühlenhoff *et al.*, 1998). In contrast to the structurally identical capsular polysaccharide of serogroup B meningococci (Liu *et al.*, 1971), the *E. coli* K1 polysaccharide is phase-variably *O*-acetylated at a high frequency of 1:50 to 1:20 (Ørskov *et al.*, 1979). The *O*-acetyltransferase responsible for this capsule modification is encoded by the gene *neuO*. Primary sequence analysis of the full length translation product revealed that NeuO belongs to the super family of NodL-LacA-CysE hexapeptide acyltransferases (Downie, 1989; Jenkins and Pickersgill, 2001; Vaara, 1992). Phase variation is mediated by a poly(ψ) motif prone to DNA slippage and mispairing. This poly(ψ) motif consists of a variable number of heptanucleotide tandem repeats (VNTR) (5´-AAGACTC-3´) and is located two nucleotides downstream of the start codon (Deszo *et al.*, 2005). Bergfeld *et al.* showed that the enzymatic activity of purified recombinant enzyme increases with repeat numbers, thus supporting large numbers of repeats in the population (Bergfeld *et al.*, 2007).

The biological function of K1 capsule *O*-acetylation is not fully understood until now. In general, modification of the host cell surface by receptor modifying enzymes is a strategy of lysogenic phages to prevent superinfection by homologous phages (Allison and Verma, 2000). However, the high frequency of phase variation observed for NeuO-dependent capsule *O*-acetylation may also point to an impact of the capsule modification in the rapid adaptation of *E. coli* K1 to environmental changes (King *et al.*, 2007a). In patients with bacteremia, *O*-acetylated *E. coli* K1 strains tend to be more virulent (Frasa *et al.*, 1993). Furthermore, in the neonatal mouse model, *E. coli* K1 escapes from killing by anti-K1-antibodies via capsule *O*-acetylation (Colino and Outschoorn, 1999). On the other hand, after oral administration of *neuO*-positive strains only de-*O*-acetylated bacteria were recovered from the blood of rat pups (Zelmer *et al.*, 2008).

In contrast to *oatC* and *oatWY*, the genes encoding polysialic acid specific *O*-acetyltransferases in *Neisseria meningitidis* serogroups C and W135/Y, respectively (Claus *et al.*, 2004), *neuO* is not associated with the capsular polysaccharide synthesis cluster, but is part of a circa 40.2 kb lambdoid prophage termed CUS-3 (Deszo *et al.*, 2005; Stummeyer *et al.*, 2006)**.** Mitomycin treatment induced phage release, and subsequent infection by phage particles depended on K1 polysaccharide expression, which is the receptor for the phage (King *et al.*, 2007b). *In vitro* studies suggest that horizontal transfer of CUS-3 occurs naturally, although phage transduction was strain specific. In a selection of 111 clinical *E. coli* K1 strains characterized by serotyping, more than 60% of the strains harbored *neuO* (King *et al.*, 2007b). Nevertheless, a phylogenetic analysis of the CUS-3 distribution is still lacking, and there is no information, whether CUS-3 is randomly distributed among *E. coli* K1 isolates.

DNA sequencing has greatly enhanced the understanding of bacterial population structures. Multilocus sequence typing (MLST) has resulted in the unique opportunity to construct large internet-databases and virtual strain collections (Maiden *et al.*, 1998; Maiden, 2006; Urwin and Maiden, 2003). The data, i.e. partial DNA sequences of seven or more housekeeping genes, are phylogenetically informative and allow the unequivocal assignment of sequence types (ST) and ST complexes (STC) comprising closely related STs. MLST data have been used to study the population biology of heterogeneous and homogenous species of diverse pathovars, the extent of horizontal gene transfer, the biological age of species, and the geographic spread of bacteria (Falush *et al.*, 2001; Falush *et al.*, 2003; Linz *et al.*, 2007; Lacher *et al.*, 2007; Linz *et al.*,

2007; Reid *et al*., 2000; Sheppard *et al*., 2008). Several MLST schemes have been introduced for the study of *E. coli* (Adiri *et al*., 2003; Beutin *et al*., 2005; Noller *et al*., 2003). The MLST scheme used in this study has been developed by Wirth *et al*. (Wirth *et al*., 2006). Since its publication, the number of strains and STs deposited in the database has risen to 2114 and 1047, respectively (http://web.mpiib-berlin.mpg.de/mlst/dbs/Ecoli; as of October 23th, 2008). The K1 capsule was found in several *E. coli* lineages. K1 strains contributed a major share of strains of the phylogenetic group B2, which comprises the STC95 as defined by MLST (Wirth *et al*., 2006). This complex also harbors avian pathogenic *E. coli* (APEC) strains (Johnson *et al*., 2008).

A population genetics approach was performed to study the distribution of *neuO* among *E. coli* K1 isolates from human and avian sources. By grouping STs and phage genotypes, the phylogenetic relationship of *E. coli* K1 and CUS-3 was investigated. Finally, the impact of capsule *O*-acetylation on environmental survival of *E. coli* K1 was studied.

RESULTS

E. coli *K1 strain collection and multilocus sequence typing*

The *E. coli* K1 strain collection analyzed in the present study comprised 144 human, one rat, and 39 avian isolates, which were collected between 1943 and 2006 in 11 different countries (Tab. 1, Tab. S1).

Host	Clinical Condition	Country*	Number (n)
Bird	Colisepticemia	DE	24
Bird	Colisepticemia	FR	4
Bird	Colisepticemia	IL	4
Bird	Colisepticemia	NL	3
Bird	Colisepticemia	CA	1
Bird	Colisepticemia	JO	1
Bird	Colisepticemia	unknown	1
Bird	Colisepticemia	USA	1
Human	Appendicitis	SE	1
Human #	Cervical swab	DE	1
Human #	Feces	DE	52
Human	Feces	UK	3
Human	Feces	USA	2
Human	Feces	BE	1
Human	Feces	unknown	1
Human #	Invasive disease	DE	6
Human	Invasive disease	unknown	6
Human	Invasive disease	USA	5
Human	Invasive disease	FI	5
Human #	UTI	DE	58
Human	UTI	FI	1
Human	UTI	SE	1
Human	UTI	unknown	1
Rat	Feces	DE	1

Table 1: Summary of the *E. coli* K1 strain collection.

NOTE: The strain collection comprises strains collected during the course of this study (#), as well as isolates provided by L.W., U.D., Mark Achtman (formerly Berlin, Germany), and Lothar Beutin (Berlin, Germany).

* BE, Belgium; CA, Canada DE, Germany; FI, Finland; FR France; IL, Israel; JO Jordany; NL, Netherlands; SE Sweden; UK, United Kingdom USA, United States o America

All 184 *E. coli* K1 strains were typed by MLST, and a total of 40 different sequence types were identified (Tab. 2). The most prevalent STs were ST95 (n=61), ST141 (n=23), ST59 (n=13), and ST357 (n=12). According to the *E. coli* MLST database, five sequence type complexes were present in the collection, i.e. STC10 (n=3), STC59 (n=16), STC95 (n=80), STC399 (n=1), and STC568 (n=10) (Tab. 2).

STC95 strains were recovered from patients with UTI (n=19), human invasive disease (n=14), and healthy intestinal carriage (n=17). Furthermore, 29 of 39 APEC K1 strains belonged to the STC95, however; it should be noted that APEC strains with ST-95 were actively searched for at some point of the study, so that there was an intended sampling bias with regard to APEC strains. The STC59 was mostly associated with human intestinal carriage, and ST141 with human UTI (Tab.2).

STC (n)	ST (n)	Source (n)	*neuO*-positive (n)	CUS-3 genotype (n)
10 (3)	10 (3)	hu ca (2), hu UTI (1)	1	NA
59 (16)	59 (13), 379 (2), 415 (1)	hu ca (13), hu UTI (3)	0	
NA	62 (8)	hu ca (3), hu id (5)	7	GT 8 (1)
95 (80)	95 (61), 140 (5), 142 (1), 368 (1), 370 (1), 390 (7), 416 (1), 417 (1), 418 (1), 421 (1)	hu ca (17), hu UTI (19), hu id (14), hu ap (1), av cs (29)	78	GT1 (1), GT2 (38), GT3 (1), GT4 (1), GT5 (3), GT8 (5), GT10 (10), GT11 (1), GT16 (1)
NA	141 (23)	hu ca (4), hu UTI (16), av cs (3)	0	
NA	144 (5)	hu ca (1), hu UTI (3), id (1)	3	GT17 (1), GT18 (2)
NA	357 (12)	hu ca (6), hu UTI (4), hu id (1), av cs (1)	7	GT12 (1), GT14 (1)
NA	358 (3)	hu ca (1), hu UTI (1), av cs (1)	2	GT13 (1), GT15 (1)
NA	363 (2)	hu ca (1), av cs (1)	0	
399 (1)	399 (1)	hu ca (1)	0	
NA	420 (3)	hu ca (1), hu UTI (1), hu id (1)	0	
NA	428 (3)	hu ca (1), hu UTI (2)	2	GT6 (1), GT9 (1)
NA	567 (2)	hu UTI (2)	0	
568 (10)	80 (8), 143 (1), 568 (1)	hu ca (2), hu UTI (6), hu cvs (1), rat ca (1)	1	NA
NA	115, 126, 131, 139, 364, 419, 427, 429, 430, 569, 573, 577, 579	hu ca (6), hu UTI (3), av cs (4)	2	GT2 (1), GT7 (1)
Total	**184**		**103**	**73**

Table 2: STCs and STs of *E. coli* K1 and their characteristics.
Abbreviations: STC, sequence type complex; ST, sequence type; hu, human; av, avian; ca, intestinal asymptomatic carriage; UTI, urinary tract infection; id, invasive disease (bacteremia and newborn meningitis); cvs, cervical swab; ap, appendicitis; cs, colisepticemia; NA, not assigned.

The strain collection represented a large fraction of the expected genetic diversity of *E. coli* K1 as evidenced by exploring the MLST database (data not shown). A large number of strains originated from Germany and were isolated in 2005 and 2006 (n=106). These strains were compared with the rest of the collection (n=78) with regard to species richness and diversity to test whether these parameters of the collection were influenced by inclusion of these strains. Species richness, i.e. the number of different STs, was 28 and 22 for Germany 2005/2006 and the rest of the collection, respectively. The Simpson Diversity Index (1-D) was 0.88 and 0.77, respectively, implying that the Germany 2005/2006 group was even more diverse.

Networks of concatenated MLST sequences were established using the Neighbor-Net approach. Significant splits allowed the definition of five MLST groups designated 1 through 5 with 82, 10, 13, 41, and 33 isolates, respectively; only three STs could not be assigned to a group (Fig. 1). The MLST groups were later used to compare the *E. coli* K1 population with CUS-3 genotypes (see below).

Distribution of neuO among E. coli K1 lineages

The strain collection was analyzed for the presence of *neuO* encoding the K1 capsule *O*-acetyltransferase. Out of 184 *E. coli* K1 isolates tested, 103 harbored the *neuO* gene (Tab. 2), as evidenced by PCR and DNA-DNA hybridization. 91% of human invasive strains, 77% of avian colisepticemia isolates, 43.3% of human fecal isolates and 42.6% of human UTI strains were *neuO*-positive. The vast majority of all *neuO*-positive strains belonged to the STC95 (n=78), and only two of 80 STC95 strains lacked the gene. Among the 104 non-STC95 strains, only 25 *neuO*-positive K1 isolates were identified (Tab. 2).

Sequencing of *neuO* revealed five different allelic variants of the gene (Tab. S2). Three of seven polymorphic sites were non-synonymous, but none of the amino acid exchanges affected the function of the enzyme, as proven by enzyme kinetic analysis (Tab. S2).

Sequence analysis of CUS-3

CUS-3 is the prophage harboring *neuO*. Partial sequencing of CUS-3, i.e. ORF22 (438 bp), fragments of the endosialidase gene *endo*NK1 (933 bp in total), and *neuO* (648 bp) (Tab. S3) was conducted to investigate whether CUS-3 genotypes are grouped along the MLST groups of *E. coli* K1, or whether MLST groups and

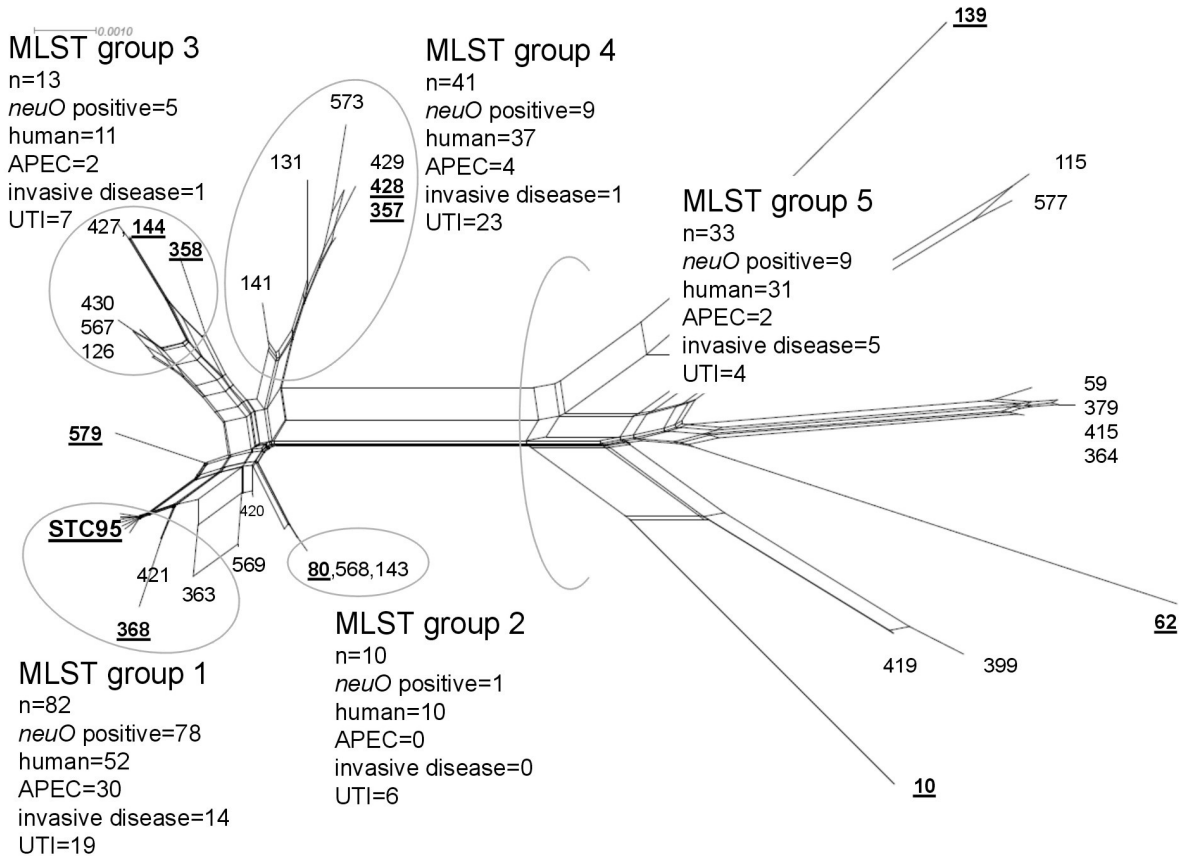


Figure 1: Neighbor-Net reconstruction based on concatenated multilocus sequence data of 184 *E. coli* K1 strains. The network identifiers represent STs. Underlined STs harbored at least one *neuO*-positive strain. Five MLST groups were defined according to splits with bootstrap values > 70% in 1000 iterations. Major characteristics of the groups are provided in inserted tables.

CUS-3 genotypes are randomly assorted. *NeuO* and ORF22 have been identified only in CUS-3, but not in other phages. The endosialidase of CUS-3 has distant relatives only in other *E. coli* K1 phages (i.e. the lytic phages K1A, K1E, K1F, K1-5). It was expected that this gene selection strategy would avoid the phylogenetic signals to be blurred by recombination between various phages. The concatenated partial CUS-3 sequence (2,019 bp) was analyzed using the Neighbor-Net approach.

Among the 103 *neuO*-positive isolates, 73 could be assigned to one of 18 CUS-3 genotypes (CUS3_GT) (Tab. 2 and Tab. S3); the remaining 30 strains were either devoid of ORF22, or lacked parts of *endo*NK1, or harbored an IS*629*-like sequence in *endo*NK1. Analysis of the aligned phage sequences using Splitstree and Neighbor-Net enabled us to identify four different CUS-3 groups, harboring 43, 18, 3, and 5 isolates (Fig. 2). Two CUS-3 genotypes could not be assigned unambiguously to one of the groups. CUS-3 group assignments inferred from Neighbor-Net analysis were further confirmed by principal component analysis (data not shown). 53 of 78 *neuO*-positive STC95 strains harbored either CUS3_GT2 (n=38), CUS3_GT10 (n=10) or CUS3_GT8 (n=5) (Tab. 2).

Subsequently, CUS-3 groups and MLST groups were compared. MLST group 1, which comprised the STC95, was shown to harbor CUS-3 groups 1 (n=41), 2 (n=16), and CUS-3 group 4 (n=1; Tab. 3). 57 of a total of 61 isolates of CUS-3 groups 1 and 2 belonged to MLST group 1. On the other hand, only 1 of 8 isolates of CUS-3 groups 3 and 4 was affiliated to MLST group 1, suggesting segregation of CUS-3 genotypes corresponding to phylogenetic groupings obtained by MLST analysis. Fisher´s exact test confirmed that CUS-3 groups 1 and 2 are associated to MLST group 1 (Tab. 3; p-value = 9.5×10^{-7}).

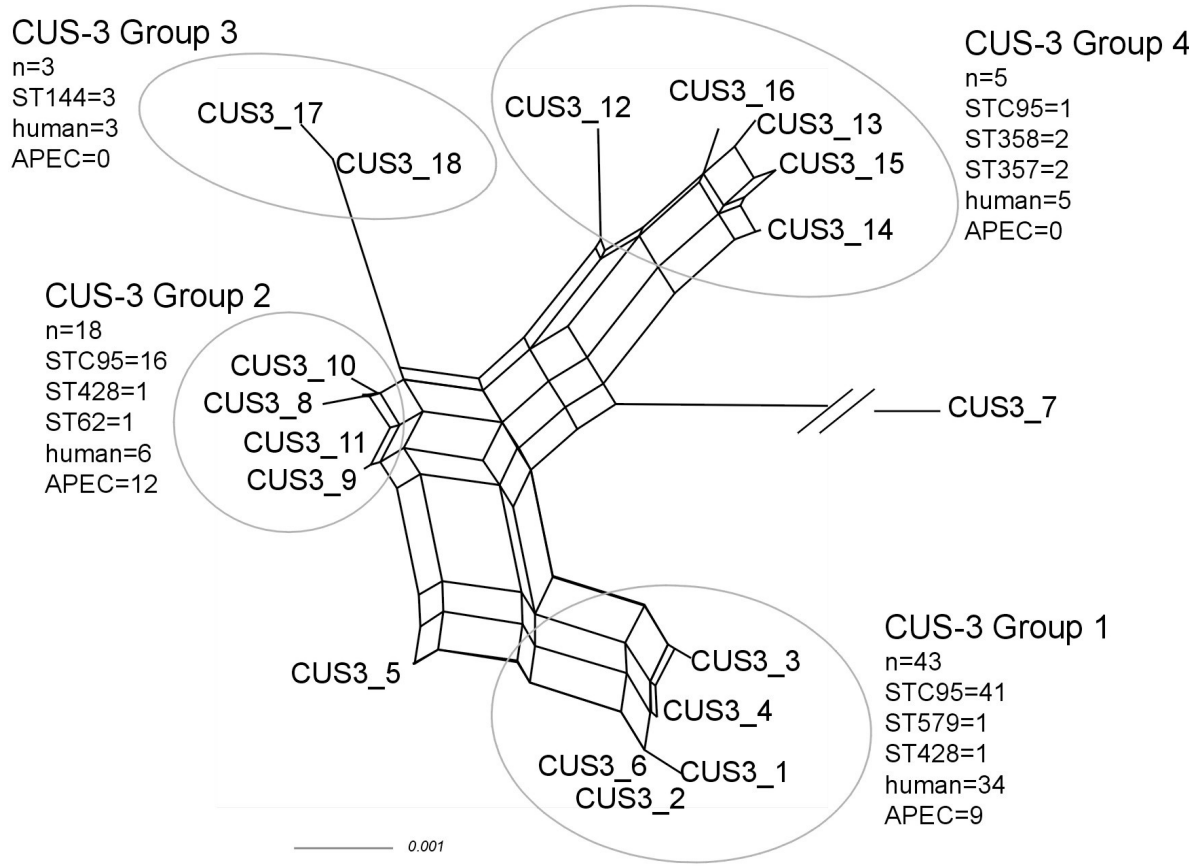


Figure 2: Neighbor-Net reconstruction based on CUS-3 genotypes of 73 CUS-3 positive strains. The network identifiers represent CUS-3_GTs. Four CUS-3 groups were defined according to splits with bootstrap values > 70% in 1000 iterations. Major characteristics of the groups are provided in inserted tables.

		CUS-3 group			
		1	**2**	**3**	**4**
MLST group	**1**	41	16	0	1
	3	0	0	3	2
	4	1	1	0	2
	5	0	1	0	0

Table 3: Association of *E. coli* K1 MLST groups and CUS-3 groups (as defined by Neighbor-Net reconstruction).

Of note, some human and avian strains were not distinguishable by sequence analysis. For example, six APEC and 21 human strains shared identical ST and CUS-3_GT, i.e. ST95 and CUS-3_GT2. To investigate whether these closely related isolates could be distinguished by gene content analyses, the presence or absence of ExPEC plasmid-associated traits, i.e. *etsA* and *hlyF*, as well as of ExPEC chromosome-associated traits, i.e. *ibeA*, *hlyD*, *malX*, and *fyuA* (Johnson *et al*., 2006; Johnson *et al*., 2007; Rodriguez-Siek *et al*., 2005) were investigated by DNA-DNA hybridization. Unweighted pair group mean average (UPGMA) cluster analysis revealed that the gene content of avian isolates corresponded to that of human strains (data not shown).

Functional relevance of NeuO for E. coli K1

The functional relevance of NeuO-mediated capsule *O*-acetylation for the interaction of *E. coli* K1 with its host or the inanimate environment is largely unknown. The uropathogenic strain A160 lacking *neuO*, and the *neuO* knock-out mutant of the neonatal meningitis strain RS218 were complemented *in trans* with a plasmid harboring either a functional NeuO (p4A) or a non-functional NeuO (p4A:H119A) (Bergfeld *et al*., 2007).

Adherence to or invasion of human brain microvascular endothelial cells (HBMEC) was monitored after 240 min of infection. Neither strain A160 nor strain RS218 depended on NeuO for adherence or invasion of HBMEC (data not shown).

Next, we investigated the role of NeuO-mediated capsule *O*-acetylation for environmental desiccation resistance. Dried chicken feces were used to mimic natural shedding of *E. coli* K1. Suspensions of chicken feces were inoculated with derivatives of the avian strain IMT5155, i.e. IMT5155Δ*neuO*/p4A (*O*-acetylated capsule) and IMT5155Δ*neuO*/p4A:H119A (non-*O*-acetylated capsule). The viability of the bacteria during desiccation was analyzed over a time period of eight weeks. During the first five weeks, similar die-off rates were observed for both variants. However, after six weeks, a significant difference was detected suggesting a more pronounced desiccation resistance in the variant expressing functional NeuO (Fig. 3A).

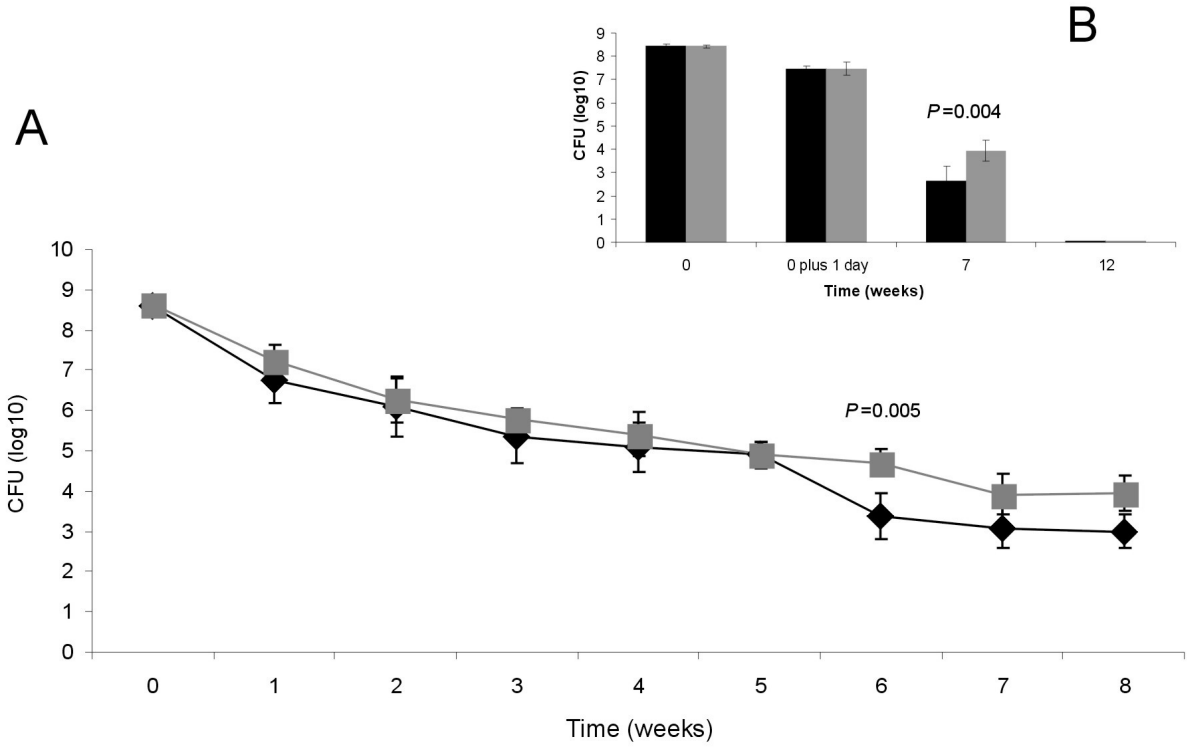


Figure 3: (A) Survival of the *O*-acetylated *E. coli* K1 mutant IMT5155Δ*neuO*/p4A (grey) and the non-*O*-acetylated *E. coli* K1 mutant IMT5155Δ*neuO*/p4A:H119A (black) in dried chicken feces over a period of eight weeks. Bacterial survival is given in means of colony forming units (CFU) ± standard deviation of five or six independent experiments performed in duplicates. (B) Confirmation of the results by an additional set of six independent experiments. Bacterial survival was determined after eight and twelve weeks.
P-values determined by 2-tailed paired student´s t-test are provided.

The results were confirmed by a second set of experiments targeting week eight of desiccation (Fig. 3B). Taken together, NeuO-mediated capsule *O*-acetylation had a moderate, but reproducible effect on the desiccation resistance of *E. coli* K1.

DISCUSSION

Differential distribution of virulence factors among pathovars of *E. coli* has been shown in various studies (reviewed in Dobrindt, 2005). A differential distribution of the K1 phenotype among different STs corresponding to the ancestral groups B2 and D has been shown recently (Wirth *et al.*, 2006). In this study, the distribution of *neuO* and the prophage CUS-3 was investigated among *E. coli* K1 isolates. Despite the fact that CUS-3 phage particles have been shown to be inducible and infectious *in vitro* (King *et al.*, 2007b),

CUS-3 and *neuO* were tightly associated with the STC95, and were not scattered in the *E. coli* K1 population. Furthermore, CUS-3 and housekeeping genes displayed signs of co-evolution. These findings implied that early after acquisition of the K1 capsule gene cluster, CUS-3 was imported into several K1 lineages, but that the consecutive impact of horizontal phage transfer was only moderate. As it is unlikely that CUS-3 phage transduction does not occur *in vivo* (King *et al*., 2007b), other mechanisms might contribute to the non-random distribution of CUS-3, such as recent clonal expansion of CUS-3 positive STC95 strains, or, alternatively, selection for this phage in STC95. The putative factor besides the K1 capsule that selects for CUS-3 is unknown. For another lambdoid phage encoded gene, i.e. *cif*, which encodes a protein conferring a cytopathic effect, the differential distribution in *E. coli* is clearly associated with the presence of the locus of enterocyte effacement (LEE), which provides the type III secretion system transporting Cif (Loukiadis *et al*., 2008). Positive selection for CUS-3 in STC95 might be driven by the environmental conditions in which the phage interacts with its host, and which might differ between *E. coli* K1 lineages. If *neuO* provides an advantage to its host bacterium, the presence of CUS-3 as a temperate prophage will be assured. The broad host range of *E. coli* K1, the dichotomy of carriage and infection displayed by this pathogen, and the various routes of transmission (fecal-oral, aerosols, uro-genital, transmission via the inanimate environment) provide a plethora of possibilities how *neuO* expression and subsequent capsule *O*-acetylation affects the ecology and distribution of *E. coli* K1. Recent evidence from neonatal rat infection experiments suggests that the K1 capsule was not *O*-acetylated after systemic invasion of blood and meninges (Zelmer *et al*., 2008). There are currently no data on the impact of the *O*-acetylated polysaccharide on the establishment of the bacteria in complex ecosystems such as the intestinal flora of animals and humans. In the present study, we investigated the extraintestinal survival of *O*-acetylated and non-*O*-acetylated *E. coli* K1 in dried chicken feces. Strains constitutively expressing NeuO were used to exclude phase variation. Irrespective of *O*-acetylation, *E. coli* K1 survived for more than two months, but after six weeks of desiccation, there was a moderate and reproducible difference in the survival rates in favor of the *O*-acetylated strain. It remains to be determined which impact the observed >10-fold benefit has on the survival in natural environments. Of note, the capsular polysaccharide colanic acid has been implicated to be involved in desiccation tolerance of *E. coli* (Obadia *et al*., 2007; Ophir and Gutnick, 1994). The molecule is *O*-acetylated (Garegg *et al*., 1971). It is tempting to speculate that the *O*-acetylation of extracellular polymeric substances leads to an increased protection against desiccation.

In this study, six APEC and 21 human isolates were identified with identical sequence type (ST95) and CUS-3 genotype (CUS3_GT2). The APEC isolates were indistinguishable from human isolates with regard to plasmid and chromosomal markers. These results underline the strong relationship between some human and avian K1 isolates, which has been proposed previously (Ott *et al*., 1991; Johnson,T.J. *et al*., 2007; Johnson, T.J. *et al.* 2008), and which suggests interspecies transmission. The recent publication of an avian *E. coli* K1 genome sequence demonstrated the genomic relatedness of strains from both sources (Johnson *et al*., 2007). Furthermore, a couple of chromosomal and plasmid markers were shown to be shared by APEC and human *E. coli* strains, especially those from human newborn meningitis (Ewers *et al*., 2007). These findings together with the data presented herein suggest that STC95 complex strains comprising APEC and human invasive isolates may be considered zoonotic.

In summary, bacteriophage CUS-3 entered *E. coli* K1 several times and displayed signs of divergent evolution. STC95 may represent a pool of zoonotic agents. NeuO-mediated K1 capsule *O*-acetylation contributes to desiccation resistance in the inanimate environment, but the role of *neuO* for *E. coli* K1 ecology remains only partially understood.

EXPERIMENTAL PROCEDURES

Bacterial strains, growth conditions and capsule typing

633 *E. coli* strains were anonymously isolated in 2005 and 2006 from routine diagnostic specimen processed at the Institute for Hygiene and Microbiology, Würzburg, Germany, and from fecal samples of healthy volunteers (staff of the Institute for Hygiene and Microbiology, Würzburg, and of the Department of Cellular Chemistry, Hannover Medical School; students of various University courses in medical microbiology, Würzburg). A single colony per sample was analyzed for expression of the K1 capsule by colony-immunoblotting and slide agglutination with the monoclonal antibody (mAb) 735 (Frosch *et al.*, 1985). 78 *E. coli* K1 strains were obtained from the strain collections of U.D., L.W., Mark Achtman (formerly Max Planck-Institute for Infection Biology, Berlin, Germany), and Lothar Beutin (formerly Robert Koch-Institute, Berlin, Germany). K1 capsule *O*-acetylation was detected by slide agglutination using the antibody 58-5 specific to this capsule modification (Torensma *et al.*, 1991) (Monosan, The Netherlands).

PCR-analysis, oligonucleotides, and DNA-sequencing

The gene *neuO* was detected in *E. coli* K1 by PCR using the oligonucleotides HC399 and HC400 (for oligonucleotide sequences see Tab. S4). The CUS-3 ORF22 and parts of the endosialidase gene *endo*NK1 were amplified with IM25/IM26 and IM30/HC404CR, respectively. PCR-products were sequenced with the respective oligonucleotides, and additionally with HC402 in the case of *neuO*, and with IM32, IM33 and HC407 in the case of *endo*NK1 parts. Sequencing was done using the BigDye® terminator v1.1 Cycle Sequencing Kit and an Applied Biosystems 377 sequencer (Applied Biosystems, USA). DNA sequences were analyzed with the programs EditSeq, MegAlign (DNAStar Inc., USA) and TraceEditPro (Ridom GmbH, Germany). Sequences of the *neuO* alleles determined in this study were submitted to the EMBL Nucleotide Sequence Database (accession numbers: FM863923 through FM863927). Concatenated CUS-3 sequences were numbered consecutively starting with 1. The sequences are provided in supplementary material S3.

DNA-DNA hybridization

The presence of *neuO* was confirmed by dot blot and Southern blot assays using a probe generated by PCR with the oligonucleotides HC402/HC403 (Tab. S4). Oligonucleotides for generating probes specific to the extraintestinal pathogenic *E. coli* (ExPEC) virulence associated traits *etsA*, *hlyF*, *ibeA*, *hlyD*, *malX*, and *fyuA* were described recently (Johnson *et al.*, 2006; Johnson *et al.*, 2007; Rodriguez-Siek *et al.*, 2005).

Multilocus sequence typing

MLST was performed essentially as described recently (Wirth *et al.*, 2006), however, an alternative forward primer for sequencing of the *mdh* fragment was used (Tartof *et al.*, 2005). Assembly of sequences was performed using TraceEditPro. STs and STCs were assigned using the *E. coli* MLST database hosted at the Max Planck-Institute for Infection Biology, Berlin, Germany (http://web.mpiib-berlin.mpg.de/mlst/dbs/Ecoli).

Sequence analysis

Sequences were concatenated using the program KonkatMeister written by J.E. (available upon request) and aligned by applying ClustalW. Aligned sequences were further analyzed using SplitsTree (Huson and Bryant, 2006). Networks were calculated based on uncorrected p-distances using the Neighbor-Net method (Bryant and Moulton, 2004). Groups of related STs and CUS-3 genotypes were defined on the basis of splits with bootstrap values of more than 70% in 1000 iterations. After alignment using MUSCLE (MUltiple Sequence Comparison by Log-Expectation) (Edgar, 2004), CUS-3 groups were confirmed by principal component analysis using Jalview (Clamp *et al.*, 2004) available at http://www.ebi.ac.uk/Tools/muscle/index.html.

Knock-out and complementation of neuO

The flanking regions of *neuO* were amplified by PCR using the oligonucleotides IM14/15 and IM16/17, respectively, and cloned into vector pUC19 (Fermentas GmbH, Germany) using the restriction sites indicated in Tab. S4. The CAT gene of vector pTnMax5 (Kahrs *et al*., 1995) was cloned between IM14/15 (5´-flanking region of *neuO*) and IM16/17 (3´-flanking region of *neuO*). The *neuO*-flanking regions together with the CAT gene were cloned into the suicide vector pMHH1, resulting in plasmid pIM1, and used for gene inactivation as described by Gunzer *et al* (Gunzer *et al*., 1998). *NeuO*-knock out mutants were verified by PCR, DNA sequencing, Southern blot analysis, and agglutination with the antibody 58-5. The Δ*neuO* strains were complemented with either plasmid p4A harboring the wild-type *neuO* gene, or p4A:H119A harboring a non-functional *neuO* variant with a single amino acid exchange (Bergfeld *et al*., 2007).

Expression, purification, and kinetic characterization of NeuO variants

Based on a pET22b-derived plasmid harboring *neuO*-1 with 36 heptanucleotide repeats (Bergfeld *et al*., 2007), constructs for expression of *neuO*-3 and *neuO*-4 were generated by introducing the required nucleotide exchanges to obtain *neuO* sequences that encode the respective protein variants NeuO-3 and NeuO-4. Nucleotide exchanges were introduced by PCR using the QuikChange site-directed mutagenesis kit (Stratagene, Germany) with the following primer pairs: AKB116_L2S/AKB117_L2S for introduction of the amino acid exchange Leu2Ser, AKB118_V5I/AKB119_V5I to introduce the exchange Val5Ile, and AKB120_V13I/AKB121_V13I for introduction of the exchange Val13Ile (Tab. S2 and S4). The identity of all constructs was confirmed by sequencing. NeuO variants encoded by *neuO*-1, *neuO*-3, and *neuO*-4 with 36 heptanucleotide repeats were expressed in *E. coli* BL21(DE3) with an N-terminal Strep-tag II and a C-terminal hexahistidine tag. Each variant was purified to homogeneity as described previously for NeuO-1 and donor kinetics were determined in a spectrophotometric assay as described (Bergfeld et al., 2007).

Cell culture experiments

Human Brain Microvascular Endothelial Cells (HBMECs) were kindly provided to the Institute for Hygiene and Microbiology by K.S. Kim (Baltimore, USA) (Stins *et al*., 1997). For strains A160 and RS218, the multiplicity of infection (MOI) was 600 and 40, respectively. The infection time was adjusted to 240 min. To determine bacterial invasion, cells were treated for 1 h with 200 µg/ml gentamicin, and bacteria were released by incubation with 0.1% saponine for 10 min. Gentamicin treatment was omitted for determination of both adherent and intracellular bacteria.

Desiccation experiments

Dried chicken feces (a commercially available organic fertilizer) (Anton Knoll GmbH, Germany) were sieved, sterilized, resuspended in sterile deionized water (14% w/v), and homogenised by ultrasonification for 10 min using a Branson Sonifier 450A via a cup horn (Branson Ultrasonic SA, Switzerland) with the following settings: constant duty cycle and an output of 8. The prepared chicken feces was mixed with either *O*-acetylated or non-*O*-acetylated bacteria to achieve a bacterial density of 2×10^{10} CFU/ml. 50 µl of the suspension were transferred onto glass disks with a diameter of 15 mm, placed in a 24-well-plate. Suspensions were dried for 4 h in a safety cabinet, and subsequently stored at 15-20° C and 40-50% humidity in the dark. For determination of the number of surviving bacteria, dried suspensions were resuspended in 1 ml PBS and incubated for 15 min at room temperature to resolve the bacteria.

Statistics

Two-tailed paired student´s t-test was employed to evaluate differences of CFU at various time points of desiccation. Fisher´s exact test was computed at http://www.langsrud.com/fisher.htm. Species richness and the Simpson Diversity Index were computed at http://www.changbioscience.com/genetics/shannon.html.

ACKNOWLEDGEMENT

This work was supported by grants VO 718/4-1 to UV and MU 1774/2-1 to MM, respectively, from the German Research Foundation. We are grateful to Mark Achtman and his colleagues for providing the *E. coli* MLST database. Mark Achtman and Lothar Beutin are kindly acknowledged for the donation of strains. The authors thank Matthias Frosch and Rita Gerardy-Schahn for constant support and helpful discussions. Provision of stool samples by anonymous healthy volunteers in Würzburg and Hannover is gratefully acknowledged. The Anton Knoll GmbH (Geeste-Bramhar, Germany) is thanked for providing dried chicken feces. K.S. Kim (Baltimore, USA) is thanked for providing HBMECs to the Institute for Hygiene and Microbiology.

REFERENCES

Adiri,R.S., Gophna,U., and Ron,E.Z. (2003) Multilocus sequence typing (MLST) of *Escherichia coli* O78 strains. *FEMS Microbiol Lett* **222:** 199-203.

Allison,G.E. and Verma,N.K. (2000) Serotype-converting bacteriophages and *O*-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8:** 17-23.

Barry,G.T. and Goebel,W.F. (1957) Colominic acid, a substance of bacterial origin related to sialic acid. *Nature* **179:** 206.

Bergfeld,A.K., Claus,H., Vogel,U., and Mühlenhoff,M. (2007) Biochemical characterization of the polysialic acid specific *O*-acetyltransferase NeuO of *Escherichia coli* K1. *J Biol Chem* **282**:22217-22227

Beutin,L., Kaulfuss,S., Herold,S., Oswald,E., and Schmidt,H. (2005) Genetic analysis of enteropathogenic and enterohemorrhagic *Escherichia coli* serogroup O103 strains by molecular typing of virulence and housekeeping genes and pulsed-field gel electrophoresis. *J Clin Microbiol* **43:** 1552-1563.

Bryant,D. and Moulton,V. (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* **21:** 255-265.

Clamp,M., Cuff,J., Searle,S.M., and Barton,G.J. (2004) The Jalview Java alignment editor. *Bioinformatics* **20:** 426-427.

Claus,H., Borrow,R., Achtman,M., Morelli,G., Kantelberg,C., Longworth,E. *et al.* (2004) Genetics of capsule *O*-acetylation in serogroup C, W-135 and Y meningococci. *Mol Microbiol* **51:** 227-239.

Colino,J. and Outschoorn,I. (1999) The form variation of the capsular polysaccharide K1 is not a critical virulence factor of *Escherichia coli* in a neonatal mouse model of infection. *Microb Pathog* **27:** 187-196.

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.

Dewitt,C.W. and Rowe,J.A. (1961) Sialic acids (N,7-*O*-diacetylneuraminic acid and N-acetylneuraminic acid) in *Escherichia coli*. I. Isolation and identification. *J Bacteriol* **82:** 838-848.

Dobrindt,U. (2005) (Patho-)Genomics of *Escherichia coli*. *Int J Med Microbiol* **295:** 357-371.

Downie,J.A. (1989) The *nodL* gene from *Rhizobium leguminosarum* is homologous to the acetyl transferases encoded by *lacA* and *cysE*. *Mol Microbiol* **3:** 1649-1651.

Edgar,R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Ewers,C., Li,G., Wilking,H., Kiebetaling,S., Alt,K., Antao,E.M. *et al.* (2007) Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: How closely related are they? *Int J Med Microbiol* **297:** 163-176.

Falush,D., Kraft,C., Taylor,N.S., Correa,P., Fox,J.G., Achtman,M., and Suerbaum,S. (2001) Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: estimates of clock rates, recombination size, and minimal age. *Proc Natl Acad Sci U S A* **98:** 15056-15061.

Falush,D., Wirth,T., Linz,B., Pritchard,J.K., Stephens,M., Kidd,M. *et al.* (2003) Traces of human migrations in *Helicobacter pylori* populations. *Science* **299:** 1582-1585.

Frasa,H., Procee,J., Torensma,R., Verbruggen,A., Algra,A., Rozenberg-Arska,M. *et al.* (1993) *Escherichia coli* in bacteremia: *O*-acetylated K1 strains appear to be more virulent than non-*O*-acetylated K1 strains. *J Clin Microbiol* **31:** 3174-3178.

Frosch,M., Gorgen,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.

Garegg,P., Lindberg,B., Onn,T., and Sutherland,L.W. (1971) Comparative studies on the M-Antigen from Salmonella typhimurium, Escherichia coli and Aerobacter cloacae. *Acta Chem Scand* **25**: 2103-2108.

Gunzer,F., Bohn,U., Fuchs,S., Muhldorfer,I., Hacker,J., Tzipori,S., and Donohue-Rolfe,A. (1998) Construction and characterization of an isogenic slt-ii deletion mutant of enterohemorrhagic *Escherichia coli*. *Infect Immun* **66:** 2337-2341.

Huson,D.H., and D. Bryant, (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Mol Biol Evol* **23**:254-267.
Jenkins,J. and Pickersgill,R. (2001) The architecture of parallel beta-helices and related folds. *Prog Biophys Mol Biol* **77:** 111-175.

Johnson,T.J., Kariyawasam,S., Wannemuehler,Y., Mangiamele,P., Johnson,S.J., Doetkott,C. *et al.* (2007) Genome Sequence of Avian Pathogenic *Escherichia coli* Strain O1:K1:H7 Shares Strong Similarities with Human ExPEC Genomes. *J Bacteriol* **189:** 3228-3236.

Johnson,T.J., Siek,K.E., Johnson,S.J., and Nolan,L.K. (2006) DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. *J Bacteriol* **188:** 745-758.

Johnson,T.J., Wannemuehler,Y., Johnson,S.J., Stell,A.L., Doetkott,C., Johnson,J.R. *et al.* (2008) Comparison of Extraintestinal Pathogenic *Escherichia coli* from Human and Avian Sources Reveals a Mixed Subset Representing Potential Zoonotic Pathogens. *Appl Environ Microbiol*. published online ahead of print; doi:10.1128/AEM.01395-08

Kahrs,A.F., Odenbreit,S., Schmitt,W., Heuermann,D., Meyer,T.F., and Haas,R. (1995) An improved TnMax mini-transposon system suitable for sequencing, shuttle mutagenesis and gene fusions. *Gene* **167:** 53-57.

King,M.R., Steenbergen,S.M., and Vimr,E.R. (2007a) Going for baroque at the *Escherichia coli* K1 cell surface. *Trends Microbiol* **15:** 196-202.

King,M.R., Vimr,R.P., Steenbergen,S.M., Spanjaard,L., Plunkett,G., III, Blattner,F.R., and Vimr,E.R. (2007b) *Escherichia coli* K1-specific bacteriophage CUS-3 distribution and function in phase-variable capsular polysialic acid O acetylation. *J Bacteriol* **189:** 6447-6456.

Kundig,J.D., Aminoff,D., and Roseman,S. (1971) The sialic acids. XII. Synthesis of colominic acid by a sialyltransferase from *Escherichia coli* K-235. *J Biol Chem* **246:** 2543-2550.

Lacher,D.W., Steinsland,H., Blank,T.E., Donnenberg,M.S., and Whittam,T.S. (2007) Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis by multilocus sequence typing and virulence gene allelic profiling. *J Bacteriol* **189:** 342-350.

Linz,B., Balloux,F., Moodley,Y., Manica,A., Liu,H., Roumagnac,P. *et al.* (2007) An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* **445:** 915-918.

Liu,T.Y., Gotschlich,E.C., Dunne,F.T., and Jonssen,E.K. (1971) Studies on the meningococcal polysaccharides. II. Composition and chemical properties of the group B and group C polysaccharide. *J Biol Chem* **246:** 4703-4712.

Loukiadis,E., Nobe,R., Herold,S., Tramuta,C., Ogura,Y., Ooka,T. *et al.* (2008) Distribution, functional expression, and genetic organization of Cif, a phage-encoded type III-secreted effector from enteropathogenic and enterohemorrhagic *Escherichia coli*. *J Bacteriol* **190:** 275-285.

Maiden,M.C. (2006) Multilocus sequence typing of bacteria. *Annu Rev Microbiol* **60:** 561-588.

Maiden,M.C., Bygraves,J.A., Feil,E., Morelli,G., Russell,J.E., Urwin,R. *et al.* Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci (U.S.A.)* **95**: 3140-3145.

Mühlenhoff,M., Eckhardt,M., and Gerardy Schahn,R. (1998) Polysialic acid: three-dimensional structure, biosynthesis and function. *Curr Opin Struct Biol* **8:** 558-564.

Noller,A.C., McEllistrem,M.C., Stine,O.C., Morris,J.G., Jr., Boxrud,D.J., Dixon,B., and Harrison,L.H. (2003) Multilocus sequence typing reveals a lack of diversity among *Escherichia coli* O157:H7 isolates that are distinct by pulsed-field gel electrophoresis. *J Clin Microbiol* **41:** 675-679.

Obadia,B., Lacour,S., Doublet,P., Baubichon-Cortay,H., Cozzone,A.J., and Grangeasse,C. (2007) Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *J Mol Biol* **367:** 42-53.

Ophir,T. and Gutnick,D.L. (1994) A Role for Exopolysaccharides in the Protection of Microorganisms from Desiccation. *Appl Environ Microbiol* **60:** 740-745.

Ørskov,F., Ørskov,I., Sutton,A., Schneerson,R., Lin,W., Egan,W. *et al.* (1979) Form variation in *Escherichia coli* K1: determined by *O*-acetylation of the capsular polysaccharide. *J Exp Med* **149:** 669-685.

Ott,M., Bender,L., Blum,G., Schmittroth,M., Achtman,M., Tschape,H., and Hacker,J. (1991) Virulence patterns and long-range genetic mapping of extraintestinal *Escherichia coli* K1, K5, and K100 isolates: use of pulsed-field gel electrophoresis. *Infect Immun* **59:** 2664-2672.

Reid,S.D., Herbelin,C.J., Bumbaugh,A.C., Selander,R.K., and Whittam,T.S. (2000) Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* **406:** 64-67.

Rodriguez-Siek,K.E., Giddings,C.W., Doetkott,C., Johnson,T.J., and Nolan,L.K. (2005) Characterizing the APEC pathotype. *Vet Res* **36:** 241-256.

Sheppard,S.K., McCarthy,N.D., Falush,D., and Maiden,M.C. (2008) Convergence of *Campylobacter* species: implications for bacterial evolution. *Science* **320:** 237-239.

Stins,M.F., Gilles,F., and Kim,K.S. (1997) Selective expression of adhesion molecules on human brain microvascular endothelial cells. *J Neuroimmunol* **76:** 81-90.

Stummeyer,K., Schwarzer,D., Claus,H., Vogel,U., Gerardy-Schahn,R., and Mühlenhoff,M. (2006) Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1-specific phages. *Mol Microbiol* **60:** 1123-1135.

Tartof,S.Y., Solberg,O.D., Manges,A.R., and Riley,L.W. (2005) Analysis of a uropathogenic *Escherichia coli* clonal group by multilocus sequence typing. *J Clin Microbiol* **43:** 5860-5864.

Torensma,R., van Wijk,A., Visser,M.J., Bouter,A., Rozenberg-Arska,M., and Verhoef,J. (1991) Monoclonal antibodies specific for the phase-variant *O*-acetylated K1 capsule of *Escherichia coli*. *J Clin Microbiol* **29:** 1356-1358.

Urwin,R. and Maiden,M.C. (2003) Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol* **11:** 479-487.

Vaara,M. (1992) Eight bacterial proteins, including UDP-N-acetylglucosamine acyltransferase (LpxA) and three other transferases of *Escherichia coli*, consist of a six-residue periodicity theme. *FEMS Microbiol Lett* **76:** 249-254.

Wirth,T., Falush,D., Lan,R., Colles,F., Mensa,P., Wieler,L.H. *et al.* (2006) Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* **60:** 1136-1151.

Zelmer,A., Bowen,M., Jokilammi,A., Finne,J., Luzio,J.P., and Taylor,P.W. (2008) Differential expression of the polysialyl capsule during blood-to-brain transit of neuropathogenic *Escherichia coli K1*. *Microbiology* **154:** 2522-2532.

SUPPLEMENTAL FIGURES

Table S1:
Detailed characteristics of all strains investigated in this study.

ORIGINAL ID	WUE-ID	ST	STC	CUS3_GT	Serotype	Provided by	Host	Clinical source	Country*	Year	*neuO*	poly(ψ) motif	*neuO*-allele
A247	3223	95	95		O18:H7:K1	M. Achtman	Human	Appendicitis	SE	1943	+	10	*neuO*-1
A23	3225	59	59		O1:H-:K1	M. Achtman	Human	Feces	BE	1948			
RS179	3984	62			K1	L. Wieler	Human	Invasive Disease	USA	1974	+	23	*neuO*-5
RS167	3972	144			K1	L. Wieler	Human	Invasive Disease	USA	1974			
RS176	3799	62			O2:K1:H7	U. Dobrindt	Human	Invasive Disease	USA	1974	+	23	*neuO*-5
IHE3033 (A53)	3788	95	95	2	K1	U. Dobrindt	Human	Invasive Disease	FI	1976	+	17	*neuO*-1
IHE3034 (A54)	3789	95	95	2	K1	U. Dobrindt	Human	Invasive Disease	FI	1976	+	13	*neuO*-1
A21 (C94)	3804	62			K1	U. Dobrindt	Human	Invasive Disease	USA	1976	+	15	*neuO*-5
IHE3036 (A56)	3791	390	95	2	K1	U. Dobrindt	Human	Invasive Disease	FI	1977	+	22	*neuO*-1
IHE3035 (A148)	3790	95	95	2	K1	U. Dobrindt	Human	Invasive Disease	FI	1977	+	13	*neuO*-1
A50	3983	62			K1	L. Wieler	Human	Invasive Disease	USA	1977	+	34	*neuO*-5
A152	3229	142	95	2	O18:H7:K1	M. Achtman	Human	Feces	UK	1977	+	33	*neuO*-1
RS226 (78015)	3800	95	95	2	K1	U. Dobrindt	Human	Feces	USA	1978	+	8	*neuO*-1
A64	3230	143	568		O7:H7:K1	M. Achtman	Rat	Feces	DE	1979			
A198	3224	139		7	O1:K1	M. Achtman	Human	Feces	UK	1979	+	14	*neuO*-2
IMT9698	3835	95	95	5	O2:K1	L. Wieler	Chicken	Colisepticemia	DE	1980	+	34	*neuO*-1
IMT4517	3823	95	95		O2:K1:H7	L. Wieler	Turkey hen	Colisepticemia	DE	1990	+	46	*neuO*-1
MT78	3832	95	95	2	O2:K1:H+	L. Wieler	Chicken	Colisepticemia	France	1993	+	20	*neuO*-1
TK3	3834	95	95		O1:K1:H7	L. Wieler	Turkey hen	Colisepticemia	CA	1993			
IMT2121	3973	357			K1	L. Wieler	Turkey hen	Colisepticemia	DE	1994	+	38	*neuO*-3
IMT2097	3812	95	95	8	O1:K1:H7	L. Wieler	Laying hen	Colisepticemia	DE	1997	+	31	*neuO*-3
IMT2108	3814	95	95	2	Osp.:H7:K1	L. Wieler	Laying hen	Colisepticemia	DE	1997	+	8	*neuO*-1
IMT2099	3813	95	95	2	O18:H7:K1	L. Wieler	Laying hen	Colisepticemia	DE	1997	+	31	*neuO*-1
IMT2288	3817	140	95	10	O2:K1:H5	L. Wieler	Laying hen	Colisepticemia	DE	1999	+	17	*neuO*-3
IMT2297	3818	95	95	10	O2:K1:H5	L. Wieler	Chicken	Colisepticemia	DE	1999	+	15	*neuO*-3
IMT2251	3815	95	95	2	On.t.:K1:H7	L. Wieler	Turkey hen	Colisepticemia	DE	1999	+	44	*neuO*-1
IMT2254	3816	126			On.t.:H-:K1	L. Wieler	Duck	Colisepticemia	unknown	1999			
IMT2294	3975	115			O2:K1	L. Wieler	Chicken	Colisepticemia	DE	1999			
IMT2248	3974	358			K1	L. Wieler	Laying hen	Colisepticemia	DE	1999			
IMT2470	3820	95	95	11	O2:K1:H5	L. Wieler	Laying hen	Colisepticemia	DE	2000	+	5	*neuO*-3
IMT4534	3826	140	95	10	O2:K1:Hsp	L. Wieler	Laying hen	Colisepticemia	DE	2000	+	46	*neuO*-3
IMT2491	3821	95	95		O2:H6:K1	L. Wieler	Chicken	Colisepticemia	JO	2000	+	21	*neuO*-1
IMT2469	3819	140	95	10	O2:K1:H5	L. Wieler	Chicken	Colisepticemia	DE	2000	+	14	*neuO*-3
IMT2532	3822	95	95	4	O18:K1:H7	L. Wieler	Laying hen	Colisepticemia	DE	2000	+	19	*neuO*-1
IMT2477	3976	141			K1	L. Wieler	Laying hen	Colisepticemia	DE	2000			
IMT4525	3825	95	95	8	O1:K1:H7	L. Wieler	Laying hen	Colisepticemia	DE	2000	+	26	*neuO*-3
IMT4518	3824	95	95	8	O1:K1:H7	L. Wieler	Laying hen	Colisepticemia	DE	2000	+	27	*neuO*-3
IMT5155	3827	140	95	10	O2:K1	L. Wieler	Laying hen	Colisepticemia	DE	2001	+	18	*neuO*-3
IMT5119	3977	363			K1	L. Wieler	Laying hen	Colisepticemia	DE	2001			
IMT5144	3978	141			K1	L. Wieler	Laying hen	Colisepticemia	DE	2001			

ORIGINAL ID	WUE-ID	ST	STC	CUS3_GT	Serotype	Provided by	Host	Clinical source	Country*	Year	*neuO*	poly(ψ) motif	*neuO*-allele
IMT8894	3828	368	95	1	O2:K1:H7	L. Wieler	Flamingo	Colisepticemia	NL	2003	+	25	*neuO*-1
IMT8899	3829	95	95		O1:K1	L. Wieler	Ostrich	Colisepticemia	NL	2003	+	10 and 62	*neuO*-1
IMT8897	3979	141			K1	L. Wieler	Bird	Colisepticemia	NL	2003			
IMT9241	3831	95	95	2	O2:K1:H4	L. Wieler	Laying hen	Colisepticemia	DE	2004	+	16	*neuO*-1
IMT9232	3830	95	95		O1:K1:H7	L. Wieler	Chicken	Colisepticemia	DE	2004	+	14	*neuO*-1
	Delta 58	95	95	8	K1	this study	Human	Feces	DE	2005	+	35	*neuO*-3
	Delta 59	357			K1	this study	Human	Feces	DE	2005	+	18	*neuO*-3
	Delta 60	420			K1	this study	Human	Feces	DE	2005			
	Delta 61	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 62	95	95		K1	this study	Human	Feces	DE	2005	+	8	*neuO*-1
	Delta 63	10	10		K1	this study	Human	Feces	DE	2005			
	Delta 64	95	95	2	K1	this study	Human	Feces	DE	2005	+	16	*neuO*-1
	Delta 65	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 16	80	568		K1	this study	Human	Cervical swab	DE	2005			
	Delta 1	95	95	2	K1	this study	Human	Feces	DE	2005	+	16	*neuO*-1
	Delta 2	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 3	141			K1	this study	Human	Feces	DE	2005			
	Delta 4	419			K1	this study	Human	Feces	DE	2005			
	Delta 5	357			K1	this study	Human	Feces	DE	2005			
	Delta 6	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 7	357			K1	this study	Human	Feces	DE	2005			
	Delta 8	416	95	2	K1	this study	Human	Feces	DE	2005	+	2	*neuO*-1
	Delta 9	95	95	2	K1	this study	Human	Feces	DE	2005	+	29	*neuO*-1
	Delta 13	141			K1	this study	Human	Feces	DE	2005			
	Delta 17	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 18	95	95	2	K1	this study	Human	Feces	DE	2005	+	40	*neuO*-1
	Delta 33	62			K1	this study	Human	Feces	DE	2005	+	41	*neuO*-5
	Delta 34	95	95		K1	this study	Human	Feces	DE	2005	+	32	*neuO*-1
	Delta 43	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 44	357			K1	this study	Human	Feces	DE	2005			
	Delta 45	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 46	427			K1	this study	Human	Feces	DE	2005			
	Delta 47	358		13	K1	this study	Human	Feces	DE	2005	+	19	*neuO*-3
	Delta 48	399	399		K1	this study	Human	Feces	DE	2005			
	Delta 49	144		18	K1	this study	Human	Feces	DE	2005	+	10	*neuO*-4
	Delta 50	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 51	131			K1	this study	Human	Feces	DE	2005			
	Delta 52	59	59		K1	this study	Human	Feces	DE	2005			
	Delta 55	429			K1	this study	Human	Feces	DE	2005			
	Delta 23	420			K1	this study	Human	Invasive Disease	DE	2005			
	Delta 27	95	95	2	K1	this study	Human	Invasive Disease	DE	2005	+	49	*neuO*-1
	Delta 29	95	95		K1	this study	Human	Invasive Disease	DE	2005	+	24	*neuO*-1
	Delta 39	95	95		K1	this study	Human	Invasive Disease	DE	2005	+	31	*neuO*-1
	Delta 10	141			K1	this study	Human	UTI	DE	2005			
	Delta 11	95	95		K1	this study	Human	UTI	DE	2005	+	4	*neuO*-1

ORIGINAL ID	WUE-ID	ST	STC	CUS3_GT	Serotype	Provided by	Host	Clinical source	Country*	Year	*neuO*	poly(ψ) motif	*neuO*-allele
	Delta 12	357			K1	this study	Human	UTI	DE	2005	+	36	*neuO*-2
	Delta 14	141			K1	this study	Human	UTI	DE	2005			
	Delta 15	421	95		K1	this study	Human	UTI	DE	2005			
	Delta 20	417	95	2	K1	this study	Human	UTI	DE	2005	+	23	*neuO*-1
	Delta 21	418	95	2	K1	this study	Human	UTI	DE	2005	+	41	*neuO*-1
	Delta 22	390	95	2	K1	this study	Human	UTI	DE	2005	+	16	*neuO*-1
	Delta 24	141			K1	this study	Human	UTI	DE	2005			
	Delta 25	141			K1	this study	Human	UTI	DE	2005			
	Delta 26	420			K1	this study	Human	UTI	DE	2005			
	Delta 28	144			K1	this study	Human	UTI	DE	2005			
	Delta 30	95	95		K1	this study	Human	UTI	DE	2005	+	24	*neuO*-1
	Delta 31	415	59		K1	this study	Human	UTI	DE	2005			
	Delta 32	390	95	2	K1	this study	Human	UTI	DE	2005	+	3	*neuO*-1
	Delta 35	141			K1	this study	Human	UTI	DE	2005			
	Delta 36	141			K1	this study	Human	UTI	DE	2005			
	Delta 37	95	95		K1	this study	Human	UTI	DE	2005	+	42	*neuO*-1
	Delta 38	80	568		K1	this study	Human	UTI	DE	2005			
	Delta 40	80	568		K1	this study	Human	UTI	DE	2005	+	51	*neuO*-3
	Delta 41	59	59		K1	this study	Human	UTI	DE	2005			
	Delta 42	430			K1	this study	Human	UTI	DE	2005			
	Delta 53	357			K1	this study	Human	UTI	DE	2005	+	8	*neuO*-3
	Delta 54	95	95	2	K1	this study	Human	UTI	DE	2005	+	5	*neuO*-1
	Delta 56	428		6	K1	this study	Human	UTI	DE	2005	+	22	*neuO*-1
	Delta 57	428			K1	this study	Human	UTI	DE	2005			
	Delta 66	357			K1	this study	Human	UTI	DE	2005			
	Delta 67	95	95	2	K1	this study	Human	UTI	DE	2005	+	22	*neuO*-1
	Delta 68	95	95	2	K1	this study	Human	UTI	DE	2005	+	10	*neuO*-1
	Delta 84	95	95	8	K1	this study	Human	Feces	DE	2006	+	13	*neuO*-3
	Delta 98	141			K1	this study	Human	Feces	DE	2006			
	Delta 99	379	59		K1	this study	Human	Feces	DE	2006			
	Delta 100	141			K1	this study	Human	Feces	DE	2006			
	Delta 101	80	568		K1	this study	Human	Feces	DE	2006			
	Delta 102	80	568		K1	this study	Human	Feces	DE	2006			
	Delta 103	357			K1	this study	Human	Feces	DE	2006			
	Delta 104	357		12	K1	this study	Human	Feces	DE	2006	+	25	*neuO*-3
	Delta 105	428		9	K1	this study	Human	Feces	DE	2006	+	9	*neuO*-3
	Delta 70	357			K1	this study	Human	Invasive Disease	DE	2006	+	19	*neuO*-3
	Delta 69	568	568		K1	this study	Human	UTI	DE	2006			
	Delta 71	569			K1	this study	Human	UTI	DE	2006			
	Delta 72	80	568		K1	this study	Human	UTI	DE	2006			
	Delta 73	80	568		K1	this study	Human	UTI	DE	2006			
	Delta 74	144		18	K1	this study	Human	UTI	DE	2006	+	30	*neuO*-4
	Delta 75	95	95	2	K1	this study	Human	UTI	DE	2006	+	4	*neuO*-1
	Delta 76	357		14	K1	this study	Human	UTI	DE	2006	+	49	*neuO*-3
	Delta 77	573			K1	this study	Human	UTI	DE	2006			

ORIGINAL ID	WUE-ID	ST	STC	CUS3_GT	Serotype	Provided by	Host	Clinical source	Country*	Year	*neuO*	poly(ψ) motif	*neuO*-allele
	Delta 78	141			K1	this study	Human	UTI	DE	2006			
	Delta 79	141			K1	this study	Human	UTI	DE	2006			
	Delta 80	141			K1	this study	Human	UTI	DE	2006			
	Delta 81	10	10		K1	this study	Human	UTI	DE	2006			
	Delta 82	80	568		K1	this study	Human	UTI	DE	2006			
	Delta 83	141			K1	this study	Human	UTI	DE	2006	Deletion		
	Delta 85	141			K1	this study	Human	UTI	DE	2006			
	Delta 86	95	95		K1	this study	Human	UTI	DE	2006	+	32	*neuO*-1
	Delta 87	567			K1	this study	Human	UTI	DE	2006			
	Delta 88	95	95		K1	this study	Human	UTI	DE	2006	+	32	*neuO*-1
	Delta 89	95	95	2	K1	this study	Human	UTI	DE	2006	+	93	*neuO*-1
	Delta 90	95	95		K1	this study	Human	UTI	DE	2006	+	30	*neuO*-1
	Delta 91	95	95	3	K1	this study	Human	UTI	DE	2006	+	20	*neuO*-1
	Delta 92	59	59		K1	this study	Human	UTI	DE	2006			
	Delta 93	141			K1	this study	Human	UTI	DE	2006			
	Delta 94	358		15	K1	this study	Human	UTI	DE	2006	+	18	*neuO*-3
	Delta 95	141			K1	this study	Human	UTI	DE	2006			
	Delta 96	567			K1	this study	Human	UTI	DE	2006			
	Delta 97	95	95	2	K1	this study	Human	UTI	DE	2006	+	34	*neuO*-1
	Delta 106	95	95	16	K1	this study	Human	UTI	DE	2006	+	25	*neuO*-3
	Delta 107	141			K1	this study	Human	UTI	DE	2006			
IHE3080	3792	390	95	2	K1	U. Dobrindt	Human	Invasive Disease	FI	unknown	+	16	*neuO*-1
A160	3227	141			O2:H6:K1	M. Achtman	Human	UTI	FI	unknown			
BEN79	3793	95	95	2	O18:K1:H7	U. Dobrindt	Poultry	Colisepticemia	FR	unknown	+	30	*neuO*-1
BEN374	3794	95	95	2	O18:K1:H7	U. Dobrindt	Poultry	Colisepticemia	FR	unknown	+	2	*neuO*-1
BEN2908	3795	95	95	5	O2:K1:H7	U. Dobrindt	Poultry	Colisepticemia	FR	unknown	+	20	*neuO*-1
W1827	3803	95	95		O2:K1:H-	U. Dobrindt	Poultry	Colisepticemia	DE	unknown	+	27	*neuO*-1
4405/1	3808	95	95	2	O12:K1:H7	U. Dobrindt	Human	Invasive Disease	DE	unknown	+	19	*neuO*-1
1772	3807	577			O2:K1	U. Dobrindt	Bird	Colisepticemia	DE	unknown			
RB11	3796	95	95	10	O2:K1	U. Dobrindt	Bird	Colisepticemia	IL	unknown	+	15	*neuO*-3
RB12	3797	95	95	10	O2:K1	U. Dobrindt	Bird	Colisepticemia	IL	unknown	+	47	*neuO*-3
RB14	3798	579		2	O2:K1	U. Dobrindt	Bird	Colisepticemia	IL	unknown	+	28	*neuO*-1
F18	3801	95	95	10	K1	U. Dobrindt	Human	Feces	un-known	unknown	+	22	*neuO*-3
VE239/94	3805	62			K1	U. Dobrindt	Human	Invasive Disease	un-known	unknown	+	6	*neuO*-5
VE1140/94	3806	95	95	2	K1	U. Dobrindt	Human	Invasive Disease	unknown	unknown	+	11	*neuO*-1
VE8793/95	3809	390	95	2	K1	U. Dobrindt	Human	Invasive Disease	unknown	unknown	+	25	*neuO*-1
B13155	3811	390	95	2	K1	U. Dobrindt	Human	Invasive Disease	unknown	unknown	+	19	*neuO*-1
B616/96	3802	390	95	2	K1	U. Dobrindt	Human	Invasive Disease	unknown	unknown	+	18	*neuO*-1
B10363	3810	95	95		K1	U. Dobrindt	Human	Invasive Disease	unknown	unknown	+	33	*neuO*-1
A285	3231	144		17	O16:H6:K1	M. Achtman	Human	UTI	SE	unknown	+	20	*neuO*-4
A211	3226	140	95	10	O2:H5:K1	M. Achtman	Human	Feces	UK	unknown	+	43	*neuO*-3
248	3833	370	95	10	O35:K1	L. Wieler	Chicken	Colisepticemia	USA	unknown	+	25	*neuO*-3
A173	3228	62		8	O7:K1	M. Achtman	Human	Feces	USA	unknown	+	31	*neuO*-3
IMT9090	3980	141			K1	L. Wieler	Human	UTI	unknown	unknown			
KK3/2	3430	62			O7:K1:NM	L. Beutin	Human	Feces	DE	unknown			

ORIGINAL ID	WUE-ID	ST	STC	CUS3_GT	Serotype	Provided by	Host	Clinical source	Country*	Year	*neuO*	poly(ψ) motif	*neuO*-allele
KK4/2	3431	95	95	2	O2:K1:H7	L. Beutin	Human	Feces	DE	un-known	+	32	*neuO*-1
KK5/14	3432	10	10		O2:K1:NM	L. Beutin	Human	Feces	DE	un-known	+	16	*neuO*-1
KK7/1	3433	95	95	2	O2:K1:H4	L. Beutin	Human	Feces	DE	un-known	+	31	*neuO*-1
KK9/10	3434	363			O46:K1:H31	L. Beutin	Human	Feces	DE	un-known			
KK11/6	3435	364			O1:K1:NM	L. Beutin	Human	Feces	DE	un-known			
KK15/1	3436	59	59		O1:K1:NM	L. Beutin	Human	Feces	DE	un-known			
KK18/10	3437	95	95	2	O2:K1:H7	L. Beutin	Human	Feces	DE	un-known	+	28	*neuO*-1
KK20/2	3438	379	59		O1:K1:NM	L. Beutin	Human	Feces	DE	un-known			
KK27/1	3439	95	95	2	Orough:K1:H4	L. Beutin	Human	Feces	DE	un-known	+	31	*neuO*-1

NOTE: WUE-ID, Würzburg strain collection; ST, sequence type; STC, sequence type complex; CUS3_GT, CUS-3 genotype; Serotype, as available at the *E. coli* MLST website; poly(ψ) motif, number of heptanucleotide repeats; * BE, Belgium; CA, Canada; DE, Germany; FI, Finland; FR, France; IL, Israel; JO, Jordany; NL, Netherlands; SE, Sweden; UK, United Kingdom; USA, United States of America.

Table S2: Nucleotide polymorphisms resulting in five allelic variants of *neuO*.
Nucleotides that differ between *neuO* alleles are given in bold. Positions are arranged in vertical direction and the numbering disregards the poly(ψ) motif at the 5´-end of the coding sequence. *NeuO*-1 is the reference allele of *neuO* found also in the genome sequence of strain RS218 (Accession no. CP000711). Kinetic parameters of different NeuO variants were analyzed by measuring enzymatic activity in a spectrophotometric assay at pH 7.5 and 25° C. Donor substrate kinetics were determined using 0.025-1 mM acetyl-CoA with colominic acid at 2 mg/ml. Reactions were initiated by adding 885 ng of purified enzyme variant. The steady-state parameters K_m and k_{cat} were determined from initial velocity measurements from three independent experiments measured in triplicates and values are given as mean ± standard deviation. Kinetic parameters of NeuO-2 and NeuO-5 were not determined (n.d.), since these variants have the same amino acid sequence as NeuO-3 and NeuO-4, respectively.

Allele	Nucleotide position							Protein variant	Amino acid position			K_m [mM]	k_{cat} [s^{-1}]
	5	**13**	**27**	**37**	**159**	**459**	**597**		**2**	**5**	**13**		
												0.24 ± 0.02	1.86 ± 0.02
neuO-1	T	G	G	G	T	A	G	NeuO-1	Leu	Val	Val		
												n.d.	n.d.
neuO-2	C	G	G	A	T	G	G	NeuO-2	**Ser**	Val	**Ile**		
neuO-3	**C**	G	A	**A**	T	**G**	G	NeuO-3	**Ser**	Val	**Ile**	0.25 ± 0.02	1.88 ± 0.03
												0.27 ± 0.03	1.90 ± 0.06
neuO-4	**C**	A	G	**A**	G	**G**	G	NeuO-4	**Ser**	**Ile**	**Ile**		
neuO-5	**C**	**A**	G	**A**	**G**	**G**	A	NeuO-5	**Ser**	**Ile**	**Ile**	n.d.	n.d.
	Leu ▼ Ser	**Val ▼ Ile**		**Val ▼ Ile**									

Table S3:
Partial concatenated CUS-3 DNA sequence (2,019 bp) comprising ORF22 (438 bp), two fragments of the prophage *endo*NK1 gene (492 bp and 441 bp), and *neuO* (648 bp without poly(ψ) motif).

>CUS3_GT1
ATGTTGTTTTCCGTTGATGATAATGGGTCAGGTAATGTTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGAGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCTC
TTGAGAAGTATTACCCGCCTGATACAACAGCTGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATCCAGATTGAAACATCACCGATGAGTACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTTTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCAGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT2
ATGTTGTTTTCCGTTGATGATAATGGGTCAGGTAATGTTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGAGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCTC
TTGAGAAGTATTACCCGCCTGATACAACAGCTGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATCCAGATTGAAACATCACCGATGAGTACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCAGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT3
ATGTTGTTTTCCGTTGATGATAATGGGTCAGGTAATGTTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGAGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCTC
TTGAGAAGTATTACCCGCCTGATACAACAGCTGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATCCAGATTGAAACATCACCGATGAGTACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT4
ATGTTGTTTTCCGTTGATGATAATGGGTCAGGTAATGTTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGAGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCTC
TTGAGAAGTATTACCCGCCTGATACAACAGCTGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATCCAGATTGAAACATCACCGATGAGTACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT

>CUS3_GT5
ATGTTGTTTTCCGTTGATGATAATGGGTCAGGTAATGTTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGAGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCAGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT6
ATGTTGTTTTCCGTTGATGATAATGGGTCAGGTAATGTTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGTTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGAGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCAGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT7
ATGTCGTTTTCCGTTGATGATAATGGGTCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACGAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTATACAGACGCCGAACTCGCGGACTTCTTTGAAGTTAGCGAGGCAACAATCAACAATTGGAAATTGGAATATCCGGAGTTTTTAGAGTCCATAAAAAA
GGGTAAGGCAGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCTC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATCCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGGTTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAATTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGATTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCAGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCTCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCTGCTATGCCTTTCTTTGCTAT
>CUS3_GT8
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCGTTTCTTTGCTAT
>CUS3_GT9
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT

>CUS3_GT10
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT11
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAATTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGCGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT12
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGATATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAATTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGCGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT13
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCACGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAACTTGCTGATTTCTTTGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGATATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGCGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
AGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT14
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCACGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTTGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGATATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGATAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAACATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT

>CUS3_GT15
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCACGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAACTTGCTGATTTCTTTGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGATATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT16
ATGTCGTTTTCCGTTGATGATAATGGATCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGTGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCACGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGATATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACCGCTGCCATTTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTATTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGATTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAACTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGACTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCGTTTCTTTGCTAT
>CUS3_GT17
ATGTCGTTTTCCATTGATGATAATGGGTCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGGGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCGCGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGATCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGATATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGGTTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAATTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGATTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT
>CUS3_GT18
ATGTCGTTTTCCATTGATGATAATGGGTCAGGTAATATTTTTGTATGTGGAGATCTTGTAAATAGCAAAGAGAATAAAGTTCAGTTCAATGGAAACAATAACAAACTTATTATAGA
AGATGATGTTGAGTGTCGATGGCTTACCGTAATATTTAGGGGGGATAATAATTACGTAAGAATACATAAAAACAGTAAGATTAAAGGTGATATTGTCGCAACAAAAGGTTCAAAAG
TTATTATCGGTAGAAGAACGACAATAGGTGCAGGTTTTGAAGTCGTCACTGATAAGTGCAATGTTACAATTGGCCATGACTGCATGATAGCAAGAGATGTTATTTTGCGTGCATCA
GATGGGCATCCTATATTTGATATTCATAGCAAAAAAAGGATTAATTGGGCAAAAGATATCATTATATCTAGTTACGTATGGGTAGGGAGAAATGTCTCTATAATGAAAGGGGTATC
TGTTGGAAGCGGATCTGTCATTGGGTATGGGAGTATTGTAACTAAAGATGTGCCATCTATGTGTGCAGCAGCCGGTAATCCAGCAAAAATAATAAAAAGAAATATAATATGGGCAA
GAACGGATAAAGCGGAGCTAATTAGTGATGACAAGAGATGCTCCAGCTATCATGCGAAGCTCACGCAAATGGCACGCCCAACAAAGTATCAAGAGGCGTATGCCGAACAGGCACGC
AAACTGTGCTTGCTGGGCTACACCGATGCAGAGCTTGCTGATTTCTTCGAAGTCAGTGAGTCAACTATTAACAAGTGGAAGCTTGATTATCCTGAGTTTTCGGAGTCCATAAAAAA
GGGTAAGGCCGTCGCTGATGCAGAAGTTAGTGACCGTCTTTATCAACGCGCTATGGGCTTCGTGGCTCCAGACATCGATATTCGTGTTATTGAAAACAGAATTGTCGAAACTCCGC
TTGAGAAGTATTACCCGCCTGATACAACTGCCGCCATCTTCTGGCTTAAGAACCGACAGAAGGATAAATGGCGCGACAAGGTTGATCACGAGCTAACAGGCAAAGACGGCGGCGCA
ATTCAGATTGAAACATCACCGATGAGCACTCTGTTCGGAAAATCTGAGCCATGTGTAAAATATTATGATGGCATTCTGTATCTGACGACCAGGGGGACATTAAGTACTCAACCCGG
TAGTTCATTGCACAGAAGCTCTGATTTAGGTACATCATGGAATTCTCTTCGCTTCCCAAATAATGTTCATCACTCAAACCTTCCTTTTGCCAAAGTTGGCGATGAGCTGGTTATTT
TTGGCAGTGAGCGCGCATTTGGTGAGTGGGAAGGAGGAGAACCTGATAACCGTTATGCAGGAAATTATCCAAGAACATTTATGACCAGAGTTAACGTCAATGAGTGGAGTCTGGAT
AATGTAGAGTGGGTTAATGTTACTGATCAGATTTATCAGGGCGGAATAGTTAACTCTGCGGTTGGTGTTGGTTCAGTTTGTATCAAAGACAACTGGCTGTACTACATTTTCGGTGG
GGAAGATTTTCTAAACCCATGGAGCATAGGGGATAACAACAGAAAATATCCTTATGTTCACGATGGTCACAGCGGCGGTGAAACAACGTCTTCAGCAGATGGTGCAATGATAACGT
TGCATGGTTCCGGAAGCAGTACTCCACGTCGCGCGGTATATAACGCACTCGAACATCTTTTTGAGAACGGAGATGTTAAACCTTATCTTGATAATGTAAATGCTCTTGGTGGTCCG
GGAAACAGGTTCTCGACAGTTTATCTTGGCTCCAATCCTGTGGTTACCAGTGACGGAACATTAAAGACAGAGCCGGTCTCTCCTGACGAAGCATTGCTGGATGCCTGGGGTGACGT
CAGGTATATCGCTTATAAATGGCTGAACGCTGTCGCTATAAAGGGGGAAGAAGGGGCGAGGATACATCATGGTGTAATCGCGCAGCAACTTCGTGATGTTCTTATTTCTCACGGAC
TCATGGAAGAAGAAAGCACAACATGCCGCTATGCCTTTCTTTGCTAT

Table S4:
Oligonucleotides used in this study.

Oligonucleotide	Sequence (5´-3´)	Target	Position as referred to accession no.	Accession no.
HC399	CGCCAGCAGGATCTAGATG	CUS-3 *neuO*	2,647-2,629	CP000711
HC400	ATTGGCTTGATGAGAAGGTAG	CUS-3 *neuO*	1,324-1,344	CP000711
HC402	TGAGCTTCGCATGATAGCTG	CUS-3 *neuO*	1,597-1,616	CP000711
HC403	CCGTTGATGATAATGGGTCAG	CUS-3 *neuO*	2,229-2,209	CP000711
IM25	TCGAAGCGGTGCTGTAAATC	CUS-3 ORF22	22,763-22,745	CP000711
IM26	CTTTGGCGACTTTGTAGCGA	CUS-3 ORF22	22,205-22,224	CP000711
IM30	CGCAGACGCATACCTGCAG	CUS-3 *endo*NK1	4,049-4,031	CP000711
HC404CR	CGTTGACGTCAATGTTTTTTGC	CUS-3 *endo*NK1	2,449-2,470	CP000711
IM32	CTTCCTGTTTAATTTTCACCC	CUS-3 *endo*NK1	3,479-3,499	CP000711
IM33	CGTCATCATATACTGCGGG	CUS-3 *endo*NK1	2,762-2,780	CP000711
HC407	GATGATGTTACCGCTCAGAG	CUS-3 *endo*NK1	3,260-3,241	CP000711
IM14	gcgcgcGAGCTCGGGAAACAGGTTCTCGATAG	CUS-3 *endo*NK1	3,069-3,050	CP000711
IM15	gcgcgcCCCGGGATTAGTATCCTACGTTATTTTACC	CUS-3 *endo*NK1	2,373-2,396	CP000711
IM16	gcgcgcGGATCCAAAACATCATGCAGACATATTAC	CUS-3 integrase	1,584-1,562	CP000711
IM17	gcgcgcGCATGCATCGTATCGCTCATTGCGAC	CUS-3 integrase	846-865	CP000711
AKB116_L2S	GGATCCGATACTATAATGT**C**TAGACTCAAGACTCAAGAC	CUS-3 *neuO*	2381-2413	CP000711
AKB117_L2S	GTCTTGAGTCTTGAGTCTA**G**ACATTATAGTATCGGATCC	CUS-3 *neuO*	2413-2381	CP000711
AKB118_V5I	CTCAAGACTCGTTTTCC**A**TCGATGATAATGGGTCAGG	CUS-3 *neuO*	2244-2208	CP000711
AKB119_V5I	CCTGACCCATTATCATCGA**T**GGAAAACGAGTCTTGAG	CUS-3 *neuO*	2208-2244	CP000711
AKB120_V13I	GATGATAATGGGTCAGGTAAT**A**TTTTTGTATGTGGAGATCTTG	CUS-3 *neuO*	2224-2182	CP000711
AKB121_V13I	CAAGATCTCCACATACAAAAA**T**ATTACCTGACCCATTATCATC	CUS-3 *neuO*	2182-2224	CP000711
HC448	CGCAAAGCAGAAGCACGTC	*E. coli* K1 *argW*	2,569,840-2,569,858	CP000468

Restriction sites used for subsequent cloning of the PCR products are underlined.
Nucleotide exchanges in *neuO* leading to amino acid exchanges are indicated in bold letters.

Chapter 4

The polysialic acid specific *O*-acetyltransferase OatC from *Neisseria meningitidis* serogroup C evolved apart from other bacterial sialate *O*-acetyltransferases

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

'The Polysialic Acid-specific *O*-Acetyltransferase OatC from *Neisseria meningitidis* Serogroup C Evolved Apart from Other Bacterial Sialate *O*-Acetyltransferases.' ♦

Anne K. Bergfeld[‡], Heike Claus[§], Nina K. Lorenzen[‡], Fabian Spielmann[§], Ulrich Vogel[§] and Martina Mühlenhoff[‡1]

From the [‡]Department of Cellular Chemistry, Medical School Hannover, 30623 Hannover and the [§]Institute for Hygiene and Microbiology, University of Würzburg, 97080 Würzburg, Germany

♦ This article was selected as a Paper of the Week.

[1]To whom correspondence should be addressed: Abteilung Zelluläre Chemie, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, 30623 Hannover, Germany. Tel.: 49-511-532-9807; Fax: 49-511-532-3956 E-mail: muehlenhoff.martina@mh-hannover.de.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

Received for publication, September 29, 2008 Published, JBC Papers in Press, November 5, 2008
DOI 10.1074/jbc.M807518200

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

Preface - About the Manuscript

The human pathogen *Neisseria meningitidis* serogroup C is surrounded by a protective capsule composed of α2,9-linked polysialic acid. In almost 90% of all strains, the capsular polysaccharide is further modified by *O*-acetylation of the sialic acid residues at position C7 or C8. This modification is catalyzed by the capsule specific *O*-acetyltransferase OatC, which shares no sequence similarity to any other protein in the databases. My first aim in this study was to establish a successful expression and purification procedure to isolate recombinant OatC. After setting up *in vitro* test systems to monitor OatC activity, I determined the donor and acceptor substrate specificity of the enzyme. Using size-exclusion chromatography, I showed that OatC assembles into homo-dimers. By generating N-terminally truncated OatC variants and OatC hybrid proteins, I showed that the first 34 amino acids of OatC form an efficient oligomerization domain that works even in a different protein context. By combining bioinformatic methods (motif scanning and secondary structure prediction), a comprehensive site-directed mutagenesis approach and *in vitro* activity assays, I identified a catalytic triad composed of Ser-286, Asp-376, and His-399, with the catalytic serine located in a nucleophile elbow motif. I confirmed the crucial role of these residues by generating site-specific mutants and showed that OatC-mediated capsule *O*-acetylation proceeds *via* a covalent acetyl-enzyme intermediate that strictly depended on the presence of Ser-286. These data highlighted that OatC shares the typical features of an α/β hydrolase fold enzyme, indicating that OatC evolved apart from all other bacterial sialate *O*-acetyltransferases known so far.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

The Polysialic Acid-specific *O*-Acetyltransferase OatC from *Neisseria meningitidis* Serogroup C Evolved Apart from Other Bacterial Sialate *O*-Acetyltransferases*[S♦]

Received for publication, September 29, 2009 Published, JBC Papers in Press, November 5, 2009, DOI 10.1074/jbc.M807518200

Anne K. Bergfeld[‡], Heike Claus[§], Nina K. Lorenzen[‡], Fabian Spielmann[§], Ulrich Vogel[§], and Martina Mühlenhoff[‡1]

From the [‡]Department of Cellular Chemistry, Medical School Hannover, 30623 Hannover, Germany and the [§]Institute for Hygiene and Microbiology, University of Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg, Germany

Neisseria meningitidis **serogroup C is a major cause of bacterial meningitis and septicaemia. This human pathogen is protected by a capsule composed of α2,9-linked polysialic acid that represents an important virulence factor. In the majority of strains, the capsular polysaccharide is modified by *O*-acetylation at C-7 or C-8 of the sialic acid residues. The gene encoding the capsule modifying *O*-acetyltransferase is part of the capsule gene complex and shares no sequence similarities with other proteins. Here, we describe the purification and biochemical characterization of recombinant OatC. The enzyme was found as a homodimer, with the first 34 amino acids forming an efficient oligomerization domain that worked even in a different protein context. Using acetyl-CoA as donor substrate, OatC transferred acetyl groups exclusively onto polysialic acid joined by α2,9-linkages and did not act on free or CMP-activated sialic acid. Motif scanning revealed a nucleophile elbow motif (G*X*S[286]*X*GG), which is a hallmark of α/β-hydrolase fold enzymes. In a comprehensive site-directed mutagenesis study, we identified a catalytic triad composed of Ser-286, Asp-376, and His-399. Consistent with a double-displacement mechanism common to α/β-hydrolase fold enzymes, a covalent acetyl-enzyme intermediate was found. Together with secondary structure prediction highlighting an α/β-hydrolase fold topology, our data provide strong evidence that OatC belongs to the α/β-hydrolase fold family. This clearly distinguishes OatC from all other bacterial sialate *O*-acetyltransferases known so far because these are members of the hexapeptide repeat family, a class of acyltransferases that adopt a left-handed β-helix fold and assemble into catalytic trimers.**

Neisseria meningitidis (meningococcus) is a worldwide and devastating cause of epidemic meningitis and sepsis (1, 2). Despite progress in diagnostics, vaccination, and therapy, invasive meningococcal infections remain a prevalent disease with a mortality of 10% and high morbidity among survivors (3). Meningococci are usually found as commensal bacteria of the human upper respiratory tract, and only a minority of isolates cause systemic disease (4). Common to these invasive strains is the presence of a polysaccharide capsule that serves as an important virulence factor (5–7). Based on the capsular polysaccharide (CPS) structure, meningococci were classified into 13 serogroups. However, only five serogroups (A, B, C, W-135, and Y) account for nearly all reported invasive diseases worldwide (1, 8). With the exception of serogroup A, these clinically important serogroups are decorated with polysialic acid (polySia) capsules (1, 9). Sialic acid refers to a family of nine-carbon sugars that are derived from *N*-acetylneuraminic acid (Neu5Ac). They are widely found as terminal sugars of glycoproteins and glycolipids on vertebrate cells but only in a limited number of pathogenic microorganisms, where they might have evolved to mimic host cell surfaces and to avoid immune attack (10, 11). The CPS of *N. meningitidis* serogroup B (NmB-CPS) is composed of α2,8-linked Neu5Ac, a homopolymer that is structurally identical to polySia found as a post-translational modification of the neural cell adhesion molecule NCAM (12, 13). In the case of serogroup C CPS (NmC-CPS), the sialic acid residues are joined by α2,9-linkages, whereas the CPS of serogroups W-135 (NmW-CPS) and Y (NmY-CPS) are Neu5Ac-containing heteropolymers composed of the disaccharide repeating units (→6)-α-*D*-Gal*p*-(1→4)-α-Neu5Ac-(2→) and (→6)-α-*D*-Glc*p*-(1→4)-α-Neu5Ac-(2→), respectively (14, 15). Interestingly, in serogroups C, W-135, and Y, but not in serogroup B, the Neu5Ac residues can be further modified by *O*-acetylation. In NmC-CPS, the acetyl groups are distributed exclusively between the hydroxyl groups located at C-7 and C-8 of Neu5Ac, whereas in NmW-CPS and NmY-CPS, *O*-acetylation is found at C-7 and C-9 (15–17). Studies from the United Kingdom showed capsule *O*-acetylation for 88 and 79% of the serogroup C and Y strains, respectively, whereas only 8% of the W-135 strains displayed *O*-acetylated capsules (18, 19). However, the biological impact of this modification with regard to pathogenicity is still unclear. For NmC-CPS and NmY-CPS, an inverse correlation between immunogenicity and the level of *O*-acetylation was observed, whereas no impact on immunogenicity was

* This work was supported by Deutsche Forschungsgemeinschaft Grants VO 718/4-1 (to U.V.) and MU 1774/2-1 (to M.M.). 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.

♦ This article was selected as a Paper of the Week.

S The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1 and S2.

1 To whom correspondence should be addressed: Abteilung Zelluläre Chemie, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, 30623 Hannover, Germany. Tel.: 49-511-532-9807; Fax: 49-511-532-3956; E-mail: muehlenhoff.martina@mh-hannover.de.

2 The abbreviations used are: CPS, capsular polysaccharide(s); polySia, polysialic acid; Neu5Ac, N-acetylneuraminic acid; NmC, *N. meningitidis* serogroup C; NmW, *N. meningitidis* serogroup W-135; NmY, *N. meningitidis* serogroup Y; mAb, monoclonal antibody; MBP, maltose-binding protein; IMAC, immobilized metal affinity chromatography; OAc, *O*-acetylation.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

found for *O*-acetylation of NmW-CPS (20, 21). These findings suggest that for serogroup C and Y meningococci, capsule *O*-acetylation may play a general role in masking immunogenic epitopes and thereby enabling escape from immune surveillance. The sialate *O*-acetyltransferases that catalyze the capsule *O*-acetylation might therefore constitute a virulence or fitness factor.

O-Acetyltransferases are very common in bacteria and modify a diverse set of substrates, including carbohydrates, amino acids, and xenobiotics. They are found as cytosolic, inner membrane-associated or integral membrane proteins and include members of different protein families. In 2004, we reported the cloning of the genes encoding the capsule-specific *O*-acetyltransferases of serogroup C and serogroups W-135 and Y (*oatC* and *oatWY*, respectively), which was at that time the first identification of genes that encode sialic acid-specific *O*-acetyltransferases (22). In serogroups W-135 and Y, an identical *O*-acetyltransferase gene was found, which might reflect the structural similarity between NmW-CPS and NmY-CPS, two heteropolymers that differ only in the orientation of the C-4 hydroxyl group of the hexose moiety. Both *oatC* and *oatWY* were located downstream of the polySia synthesis genes *siaA–D* within the capsule gene complex of serogroup C and W-135/Y meningococci, respectively. Whereas no homology to any known protein in the data bases was found for OatC, analysis of the primary sequence of OatWY revealed the presence of imperfect tandem repeats of the consensus sequence (LIV)(GAED)X_2(STAV)X, the hallmark of enzymes belonging to the hexapeptide repeat family of acyltransferases. Members of this family adopt a unique left-handed β-helix fold, and three identical subunits assemble into a catalytic trimer (23, 24).

Since 2004, several genes encoding sialate *O*-acetyltransferases have been identified in bacteria that contain sialic acid as a component of their CPS or lipo-oligosaccharide, such as *Escherichia coli* K1, *Campylobacter jejuni*, and type III group B streptococci (25–28). Notably, all these enzymes belong to the hexapeptide repeat family, suggesting that the specificity for sialic acid evolved exclusively within this protein family.

To investigate whether OatC is just an estranged member of the hexapeptide repeat family that lacks sequence homology or whether it is the first example of a sialate *O*-acetyltransferase belonging to a different protein family, we started a detailed biochemical characterization of this enzyme. Using serogroup C meningococcal spheroplast membranes as an enzyme source, Vann *et al.* (29) demonstrated that endogenous OatC transfers acetyl groups from acetyl-CoA to endogenous membrane-bound as well as to exogenous soluble NmC-CPS. However, reports on the isolation and further characterization of the enzyme are missing. In this study, we succeeded in purifying active recombinant OatC, which was found as a homodimer. A comprehensive site-directed mutagenesis study revealed the presence of a Ser-Asp-His catalytic triad with the serine located in a nucleophile elbow motif, a characteristic feature of α/β-hydrolase fold enzymes. The identification of a covalent acetyl-enzyme intermediate gave the first insight into the catalytic mechanism of OatC. Together, our data obtained by biochemical analysis of wild-type and mutant OatC provide strong evidence that OatC is a polySia-specific *O*-acetyltransferase that belongs to the α/β-hydrolase fold superfamily. This result demonstrates that OatC evolved apart from all other bacterial sialate *O*-acetyltransferases known so far.

EXPERIMENTAL PROCEDURES

Materials - Purified de-*O*-acetylated CPS from *N. meningitidis* serogroup C was kindly provided by Baxter. Purified CPS from *N. meningitidis* serogroups W-135 and Y were kindly donated by Pasteur Mérieux Connaught. CPS from *E. coli* K1 was purified as described (30) and kindly provided by T. Scheper (Institut für Technische Chemie, Universität Hannover, Hannover, Germany). Monoclonal antibody (mAb) P1.2 was kindly provided by J. Suker and I. Feavers (National Institute for Biological Standards and Control, Hertfordshire, United Kingdom). pET expression vectors and *E. coli* BL21(DE3) were obtained from Novagen. CMP-Neu5Ac, Neu5Ac, colominic acid, 5,5'-dithiobis(2-nitrobenzoic acid), acetyl-CoA, propionyl-CoA, and butyryl-CoA were purchased from Sigma.

Generation of OatC Expression Plasmids - All constructs for the expression of OatC variants were generated in the prokaryotic expression vector pET22b-Δ*Strep*, resulting in OatC variants with a C-terminal hexahistidine tag. pET22b-Δ*Strep* was generated by adapter ligation of oligonucleotides AKB42 and AKB43 into the BamHI/XbaI sites of pET22b-*Strep* (26), resulting in a vector that lacks the pelB leader sequence of pET22b and the Strep tag sequence of pET22b-*Strep.* Wild-type *oatC* was amplified by PCR using the primer pair MM269/MM271 and genomic DNA of *N. meningitidis* serogroup C as template. The PCR product was subcloned by the BamHI/XhoI sites into pET22b-Δ*Strep*. Constructs for the expression of the N-terminally truncated forms ΔN34-OatC and ΔN103-OatC were generated by PCR using the primer pairs AKB67/MM271 and AKB69/MM271, respectively. The PCR products obtained were subcloned into pET22b-Δ*Strep* as described above.

For expression of a hybrid protein composed of the first 34 amino acids of OatC fused to the maltose-binding protein (MBP) from *E. coli* (N34-MBP), a pET22b-based plasmid was generated in a two-step cloning approach. (i) The sequence encoding the first 34 amino acids of OatC was amplified by PCR using the primer pair MM269/AKB81 and wild-type *oatC* as template. The PCR product was subcloned into pET22b-Δ*Strep* by BamHI/XhoI sites, resulting in the plasmid $pOatC_{fusion}$. (ii) The MBP gene was amplified by PCR using the primer pair AKB72/AKB128 and the vector pMAL-c (New England Biolabs) as template. The resulting PCR product was subcloned into the HindIII/XhoI sites of $pOatC_{fusion}$.

For expression of MBP alone, the corresponding coding sequence was amplified by PCR using the primer pair AKB129/AKB128 and subcloned into the BamHI/XhoI sites of pET22b-Δ*Strep*. The identity of all generated constructs was confirmed by sequencing. The sequences of the indicated oligonucleotides are given in supplemental Fig. S1.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

Site-directed Mutagenesis - Single amino acid substitutions were generated by PCR using the QuikChange site-directed mutagenesis kit (Stratagene) following the manufacturer's guidelines. Wild-type *oatC* subcloned in pET22b-Δ*Strep* was used as template together with the primer pairs listed in supplemental Fig. S1. BamHI/XhoI fragments of the PCR products obtained were subcloned in pET22b-Δ*Strep*, and the identity of the inserted PCR fragments was confirmed by sequencing.

Expression and Purification of Recombinant OatC - Freshly transformed *E. coli* BL21(DE3) cells were cultivated at 37 °C in 500 ml of Power Broth (AthenaES) containing 200 µg/ml carbenicillin. At A_{600} = 1.4, expression was induced by adding 0.1 mM isopropyl β-*D*-thiogalactopyranoside. Thereafter, bacteria were grown at 15°C for 20 h, harvested, and lysed by sonication. Recombinant proteins were isolated by immobilized metal affinity chromatography (IMAC) using 1 ml-HisTrap HP columns (GE Healthcare) equilibrated with binding buffer (50mM Tris-HCl (pH 7.5), 300 mM NaCl, and 20 mM imidazole). After washing with 15 ml of binding buffer, proteins were eluted with a linear imidazole gradient (20–500 mM imidazole in binding buffer). Enzyme-containing fractions were pooled, concentrated by ultrafiltration, and loaded onto a Superdex 200 HR 10/30 column (GE Healthcare) equilibrated with 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl.

Spectrophotometric Activity Assay - The enzymatic activity of purified OatC variants was determined in a spectrophotometric assay. The reaction was performed at 25 °C in a total volume of 100 µl containing 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 20% (w/v) glycerol, 2 mM 5,5'-dithiobis(2-nitrobenzoic acid), 200 µg/ml purified NmC-CPS, and 1 mM acetyl-CoA. The reaction was initiated by adding 1.7 pmol of purified enzyme and monitored continuously at 405 nm in half-area 96-well plates (Greiner) using a PowerWave 340 microtiter plate spectrophotometer (BioTek).

Radioactive Incorporation Assay - For determination of OatC activity in crude bacterial lysates, a radioactive incorporation assay was performed. OatC variants were expressed in *E. coli* BL21(DE3) as described above. Bacteria from 1 ml of culture medium with A_{600} = 1 were resuspended in 25 µl of lysis buffer (50 mM Tris-HCl (pH 8.0), 2 mM EDTA, and 100 µg/ml lysozyme), incubated for 15 min at 4 °C, and lysed by sonication. The enzyme reaction was performed at 25 °C in a total volume of 50 µl containing 20 µl of the soluble lysate fraction, 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 20 µg of NmC-CPS, and 9.25 kBq of [^{14}C]acetyl-CoA (GE Healthcare). The reaction was stopped after 15 min by spotting 10 µl of the reaction mixture on Whatman No. 3MM paper. Free radioactivity was removed by descending paper chromatography with running buffer (300 mM ammonium acetate buffer (pH 7.5) and 70% ethanol), and radioactivity incorporated into the chromatographically immobile NmC-CPS was quantified by scintillation counting. To determine the enzymatic activity of purified wild-type OatC, the reaction was performed in a total volume of 200 µl containing the same reaction mixture as described for the spectrophotometric assay supplemented with 37 kBq of [^{14}C]acetyl-CoA. At different time points, aliquots of the reaction mixture were analyzed as described above.

Detection of Acetyl-enzyme Intermediates - Purified enzyme (3.2 µg) was incubated in a final volume of 50 µl at 25 °C in the presence of 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, and 9.25 kBq of [^{14}C]acetyl-CoA. After 5 min, one-half of the reaction mixture was removed. The remaining reaction mixture was supplemented with 20 µg NmC-CPS and incubated for 25 min. Enzyme reactions were stopped by adding an equal volume of loading buffer and separated by SDS-PAGE. Gels were analyzed by either phosphorimaging or combined Alcian blue/silver staining (31).

SDS-PAGE, Staining, and Immunoblotting - SDS-PAGE was performed under reducing conditions using 2.5% (v/v) β-mercaptoethanol. Silver staining, Coomassie staining, and Western blot analysis with anti-penta-His antibody (Qiagen) were performed as described previously (26).

Size-exclusion Chromatography - The quaternary structure of OatC variants was determined on a Superdex 200 HR 10/30 column. The column was equilibrated with 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl and calibrated with the following molecular mass standards (Sigma): thyroglobulin (669 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome *c* (12.4 kDa). Eluted proteins were monitored by absorbance at 280 nm.

Analysis of OatC Activity in Vivo – Isogenic NmC strains carrying different *oatC* variants were generated by homologous recombination. The downstream region of *oatC* that ought to provide the right anchor for homologous recombination was amplified using the primer pair FS1/FS2. The PCR product was subcloned into the ApaI sites of pBluescript II SK(+) (Stratagene), resulting in the plasmid pFS1. Subsequently, the SalI restricted kanamycin resistance cassette of pUC4K (GE Healthcare) was inserted into the XhoI site of pFS1, resulting in pFS2. The kanamycin cassette was inserted to later select for transformants. Finally, wild-type and mutant *oatC* variants with an additional sequence stretch encoding a C-terminal hexahistidine tag were excised from the pET-based expression plasmids described above by sequential restriction. After restriction with HpaII and subsequent blunt ending with T4 DNA polymerase, the 1.5-kb fragment obtained was restricted with BamHI and inserted into the BamHI and EcoRV sites of pFS2. The resulting plasmids were transformed into *N. meningitidis* serogroup C wild-type strain 2120 (32) to allow homologous recombination. Correct insertion into the *oatC* locus was confirmed by PCR, Southern blotting, and sequence analysis. Meningococci were grown on chocolate agar, and 20 µl of a bacterial suspension (A_{600} = 0.15) was added to each well of a microtiter plate (Greiner) that had been coated with poly-*D*-lysine (Sigma). After drying, bacteria were fixed with 0.05% glutaraldehyde in phosphate-buffered saline. Capsule expression and modification by *O*-acetylation were analyzed in an enzyme-linked immunosorbent assay as described previously (32) using the following mAbs: mAb924, which is specific for NmC-CPS (OAc^+ and OAc^- forms); mAb 1125, which is directed exclusively against the *O*-acetylated form of NmC-CPS (OAc^+); and mAb P1.2, which is directed against the surface protein PorA

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

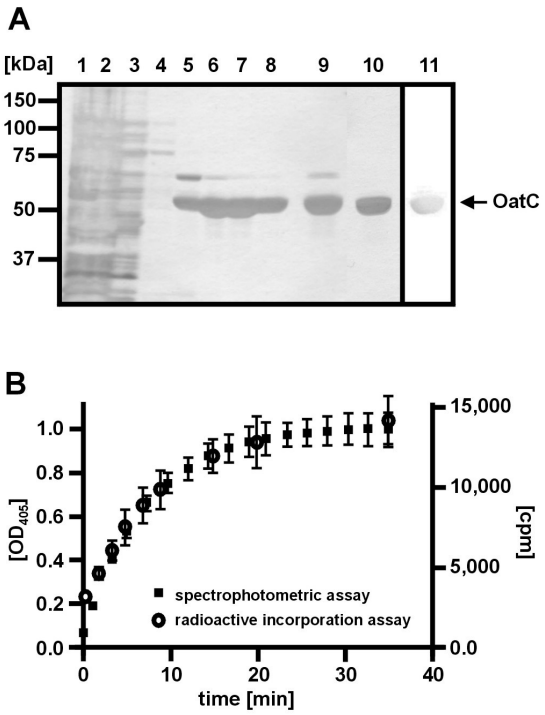


FIGURE 1. **Purification of OatC.** *A*, after expression in *E. coli* BL21(DE3), OatC carrying a C-terminal hexahistidine tag was isolated by IMAC on a Ni^{2+} chelating column followed by size-exclusion chromatography. Purification steps were monitored by 10% SDS-PAGE and silver staining, including aliquots of bacterial lysate (*lane 1*), flow-through of the IMAC column (*lane 2*), wash fractions (*lanes 3* and *4*), OatC-containing fractions eluted from the IMAC column (*lanes 5–8*), a pool of fractions 5-8 from the IMAC column (*lane 9*), and purified OatC after size-exclusion chromatography (*lane 10*). Purified OatC after size-exclusion chromatography was also analyzed on a Western blot stained with anti-penta-His antibody (*lane 11*). *B*, the enzymatic activity of purified OatC was determined in a spectrophotometric assay (data given as means ± S.D. of three independent experiments measured in triplicates) and a radioactive incorporation assay (data given as means ± S.D. of two independent experiments). A_{405} measured in the spectrophotometric assay is displayed on the *left y axis*, whereas the incorporated radioactivity is given as counts/min on the *right y axis*.

(33), as a loading control. *N. meningitidis* serogroup C strain 2948 lacking a functional *oatC* gene (22) was used as a negative control.

RESULTS

Purification of Active OatC - OatC was expressed in *E. coli* BL21(DE) with a C-terminal hexahistidine tag. The recombinant protein was isolated by IMAC and size-exclusion chromatography as described under "Experimental Procedures." Using this two-step purification protocol, OatC was purified to homogeneity (Fig. 1*A*), yielding 24 mg of enzyme from 500 ml of bacterial culture. Silver staining revealed a single band consistent with a calculated molecular mass of 54.3 kDa, and the identity of this band was confirmed by Western blot analysis using anti-penta-His antibody (Fig. 1*A*).

To determine acetyltransferase activity, isolated OatC was incubated with [^{14}C]acetyl-CoA in the presence of purified NmC-CPS. After removal of free acetyl-CoA, incorporation of radiolabeled acetyl groups was measured by scintillation counting (Fig. 1*B*). In parallel, a spectrophotometric assay was established that monitored the transfer of acetyl groups from acetyl-CoA onto NmC-CPS by measuring the appearance of the free sulfhydryl group on released CoA with Ellman's reagent at 405 nm (Fig. 1*B*). In both assays, enzymatic activity was detected, and similar time courses were observed for the incorporation of acetyl groups and the release of CoA (Fig. 1*B*).

TABLE 1

Substrate specificity of OatC

Relative enzymatic activities of purified OatC were determined in the presence of different acceptor and donor substrates. The concentration of all acceptor substrates was normalized to equal sialic acid content (6.47mM sialic acid residues) and thereby to an equal number of acceptor sites. The activity obtained for NmC-CPS and acetyl-CoA was set to 100%. Values are means ± S.D. from three independent experiments.

	Relative enzymatic activity
	%
Acceptor substrate	
NmC-CPS	100.0 ± 2.7
Neu5Ac	0.2 ± 0.6
CMP-Neu5Ac	0.3 ± 0.2
colominic acid	0.4 ± 0.7
K1-CPS	0.3 ± 0.6
NmW-CPS	0.0 ± 0.4
NmY-CPS	0.3 ± 0.1
Donor substrate	
Acetyl-CoA	100.0 ± 2.4
Propionyl-CoA	21.4 ± 1.2
Buturyl-CoA	0.1 ± 0.1

Donor and Acceptor Substrate Specificity of OatC - Using the spectrophotometric assay, we analyzed the acceptor substrate specificity of OatC. The enzymatic activity of purified OatC was monitored in the presence of monomeric Neu5Ac, activated Neu5Ac (CMP-Neu5Ac), the Neu5Ac-containing heteropolymers NmW-CPS and NmY-CPS, and polySia homopolymers composed of either α2,9-linked Neu5Ac residues (NmC-CPS) or α2,8-linked Neu5Ac residues (colominic acid and K1-CPS, the partially hydrolyzed and full-length CPS of *E. coli* K1, respectively). As summarized in Table 1, no enzymatic activity was detected for free Neu5Ac and CMP-Neu5Ac, indicating that OatC is specific for oligo- or polySia. However, of the tested polymeric substrates, only NmC-CPS served as an acceptor, demonstrating that OatC is highly specific for α2,9-linked polySia.

The donor substrate specificity of OatC was determined by measuring the enzymatic activity for NmC-CPS in the presence of acetyl-CoA and the alternative acyl donors propionyl-CoA and butyryl-CoA. Compared with acetyl-CoA, ~20% residual enzymatic activity was detected in the presence of propionyl-CoA, whereas no activity was observed with butyryl-CoA (Table 1).

OatC Assembles into Dimers Mediated by the First 34 Amino Acids – To study the oligomerization state of OatC, the purified enzyme was analyzed by size-exclusion chromatography on a Superdex 200 column. A single peak was obtained that corresponded to a molecular mass of 105 ± 1.6 kDa (Table 2), demonstrating that OatC is a homodimer with an apparent molecular mass that is in agreement with the calculated dimeric mass of 108.6 kDa.

Using the program HTH (34) and searching the Pfam Database of protein families (35), we identified a putative helix-turn-helix motif within the first 34 amino acids and a tetratricopeptide repeat spanning amino acids 71–104 of OatC. Because both motifs have been associated with protein-protein interactions (36–38), we asked whether

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

TABLE 2

Quaternary structures of OatC variants

The quaternary structures of purified OatC variants, the hybrid protein N34-MBP, and MBP were determined by size-exclusion chromatography. The molecular masses of the monomeric species were calculated from the respective amino acid sequences. The oligomeric state was calculated from the values obtained by gel filtration divided by the mass calculated for the monomer. Relative enzymatic activities of OatC variants were analyzed in the spectrophotometric assay, and the activity obtained for wild-type (WT) OatC was set to 100%. Values are means ± S.D. from three independent experiments. ND, not determined.

Protein variant	Molecular mass		Oligomeric state	Relative enzymatic activity
	Monomer (calculated)	Gel filtration		
	kDa			%
WT OatC	54.3	105.0 ± 1.6	2.0	100.0 ± 2.7
ΔN34-OatC	50.2	42.8 ± 0.8	0.9	24.9 ± 1.2
ΔN103-OatC	42.0	46.0 ± 0.9	1.1	0.4 ± 0.3
N34-MBP	46.0	114.4 ± 6.8	2.5	ND
MBP	41.9	29.1 ± 1.1	0.7	ND
OatC-S286A	54.3	100.1 ± 1.2	1.9	0.1 ± 0.3
OatC-D376A	54.3	92.4 ± 1.3	1.7	0.2 ± 0.1
OatC-H399A	54.3	92.7 ± 2.3	1.7	0.1 ± 0.4

the N-terminal part of OatC is involved in dimer formation. To address this question, the ΔN34-OatC and ΔN103-OatC variants, lacking the first 34 and 103 amino acids, respectively, were generated. After expression in *E. coli* BL21(DE3), both variants were purified to homogeneity by IMAC and size-exclusion chromatography (data not shown). Interestingly, both variants were found as monomers (Table 2), indicating that dimerization of OatC depends on the first 34 amino acids. Whereas no enzymatic activity was found for ΔN103-OatC, ΔN34-OatC retained 25% of the wild-type activity (Table 2). The latter finding demonstrates that dimerization is not an absolute requirement for enzymatic activity. However, deletion of the first N-terminal amino acids seems to affect protein folding and/or stability, leading to the observed drop in activity. To further evaluate the role of the N-terminal part in providing a dimerization interface, we fused the first 34 amino acids of OatC to MBP of *E. coli*, resulting in the hybrid protein N34-MBP. MBP is a monomeric protein, and in line with previous reports (39, 40), an oligomeric state of 0.7 was determined by gel filtration. By contrast, a quaternary state of 2.5 was found for N34-MBP, providing strong evidence that the first 34 amino acids of OatC are sufficient to mediate protein dimerization even in a different protein context.

Identification of a Catalytically Important Histidine Residue - For OatC, the identification of functionally important residues by visualizing highly conserved residues was impeded by a complete lack of sequence homology between OatC and any other protein in the data bases (22). However, because acetyltransferases of a variety of different protein families comprise catalytically important histidine residues (24, 41–44), we asked whether one of the six histidines found in the primary sequence of OatC (Fig. 2*A*) is involved in enzymatic activity. To address this question, single alanine substitutions were introduced, resulting in OatC variants H31A, H111A, H183A, H267A, H399A, and H456A. All proteins were expressed in *E. coli* BL21(DE3) with a C-terminal hexahistidine tag and were found at comparable levels in the soluble fraction. After isolation by IMAC and size-exclusion chromatography (Fig. 2*B*), equimolar amounts of wild-type and mutant enzymes were analyzed

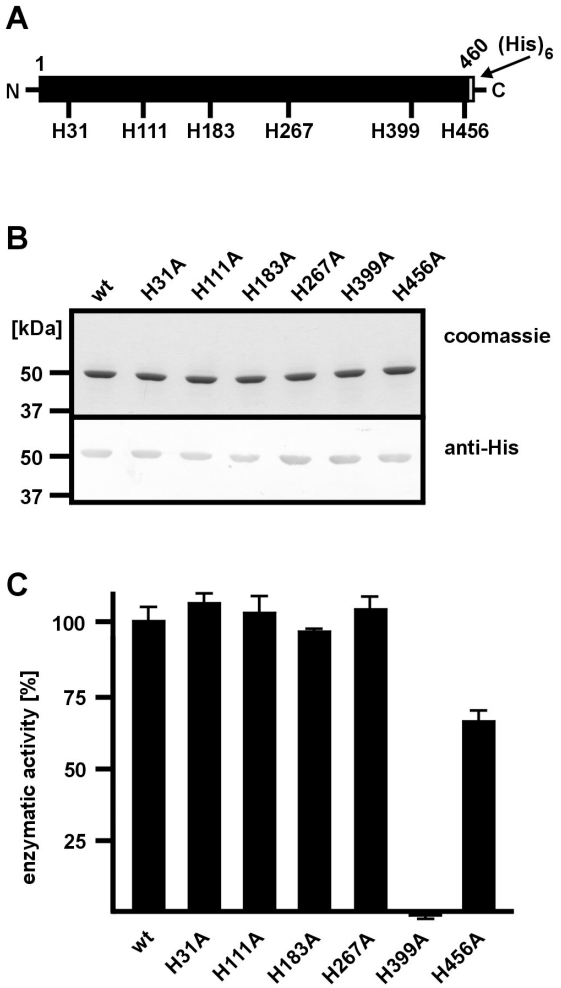


FIGURE 2. **Identification of a catalytic histidine.** *A*, shown is a schematic representation of OatC with a C-terminal hexahistidine tag. The positions of all six histidine residues present in OatC are highlighted. *B*, the single alanine substitutions of the indicated amino acids were introduced into OatC. After expression in *E. coli* BL21(DE3) and affinity purification on Ni^{2+} chelating columns, wild-type (*wt*) and mutant forms were analyzed by 10% SDS-PAGE and Coomassie staining (*upper panel*) and Western blotting using anti-penta-His antibody for detection (*lower panel*). *C*, the enzymatic activities of wild-type and mutant forms were determined in a spectrophotometric assay using equimolar concentrations of each purified enzyme variant. Data represent means ± S.D. of three independent experiments measured in triplicates, and the value obtained for wild-type OatC was set to 100%.

in the spectrophotometric assay. With the exception of OatC-H399A and OatC-H456A, all investigated variants showed enzymatic activity similar to that of the wild-type protein (Fig. 2*C*). Alanine substitution of the most C-terminal histidine residue (H456A) resulted in reduced activity. However, this mutant retained 70% of the wild-type level, excluding an essential role of His-456 in catalysis. By contrast, a complete loss of activity was observed for OatC-H399A, demonstrating that His-399 is critical for enzymatic activity.

OatC Contains a Nucleophile Elbow Motif - Motif scanning using the MyHits web server (45) revealed a putative nucleophile elbow motif (G*X*S^{286}*X*GG) in the primary sequence of OatC (Fig. 3*A*). With the consensus sequence Sm-*X*-Nu-*X*-Sm-Sm (where Sm is a small residue, Nu is a nucleophilic residue, and *X* is any residue), the nucleophile elbow represents a structural

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

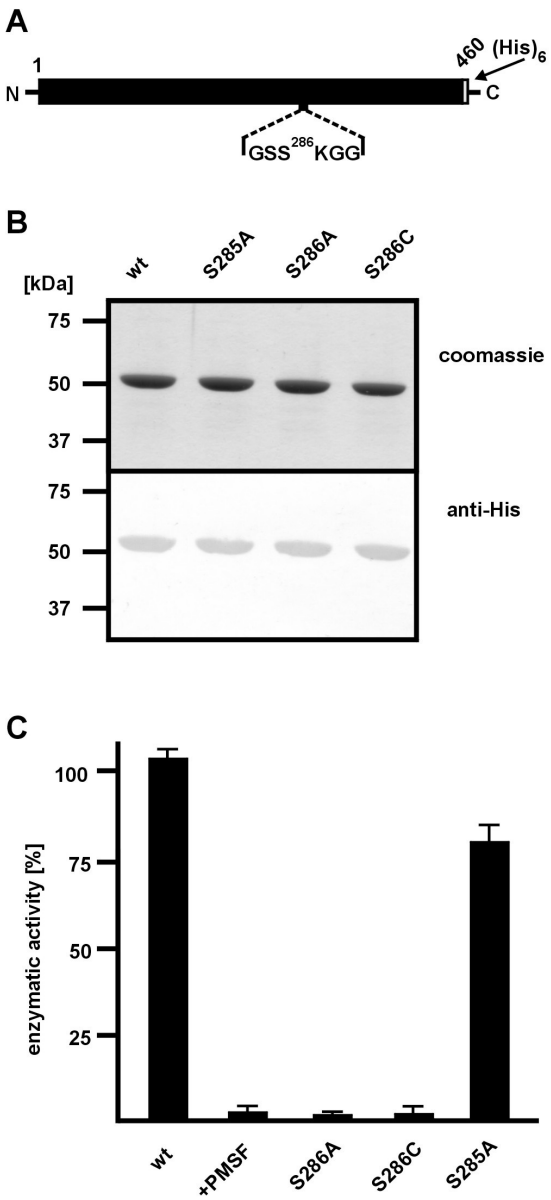


FIGURE 3. **Identification of a nucleophile elbow motif.** *A*, shown is a schematic representation of OatC. The relative positions of the identified nucleophile elbow motif and Ser-286 are indicated. *B*, the single amino acid exchanges S286A, S286C, and S285A were introduced into OatC. After expression in *E. coli* BL21(DE3) and affinity purification on Ni^{2+} chelating columns, wild-type (*wt*) and mutant forms were analyzed by 10% SDS-PAGE and Coomassie staining (*upper panel*) and Western blotting using anti-penta-His antibody for detection (*lower panel*). *C*, the enzymatic activities of wild-type and mutant forms were determined in a spectrophotometric assay using equimolar concentrations of each purified enzyme variant. In parallel, the enzymatic activity of purified wild-type OatC was analyzed in the presence of 10 mM phenylmethylsulfonyl fluoride (+*PMSF*). Data represent means ± S.D. of three independent experiments measured in triplicates, and the value obtained for wild-type OatC was set to 100%.

motif that places a nucleophile at the tip of a sharp turn to allow efficient presentation at the site of attack on the substrate. It is a typical feature of α/β-hydrolase fold enzymes, where the nucleophile (Ser, Cys, or Asp) within the motif is part of a catalytic triad composed of Nu-(Asp/Glu)-His (42, 46–50). Members of the α/β-hydrolase fold family that use serine as a nucleophile can be irreversibly inhibited by the serine-specific reagent phenylmethylsulfonyl fluoride, resulting in sulfonylation of the catalytic serine. Preincubation of purified OatC with phenylmethylsulfonyl fluoride resulted in an almost complete loss of OatC activity (Fig. 3*C*), indicating the presence of a catalytic serine. To determine whether Ser-286 located in the nucleophile elbow is indeed part of the catalytic site of OatC, this residue was exchanged with alanine and in parallel by cysteine to introduce an alternative nucleophile. Because the nucleophile elbow motif identified in OatC contains a second serine (Ser-285), this residue was also exchanged with alanine. The resulting variants, OatC-S285A, OatC-S286A, and OatCS286C, were expressed in *E. coli* BL21(DE3) and purified to homogeneity via the C-terminal hexahistidine tag by IMAC and size-exclusion chromatography (Fig. 3*B*). Enzymatic activity was investigated using the spectrophotometric activity assay described above. Whereas OatC-S285A retained 80% of the wild-type activity, alanine substitution of the adjacent Ser-286 resulted in a complete loss of activity (Fig. 3*C*). Replacement by a cysteine did not rescue activity, indicating that OatC activity depends on a serine residue that has to be properly positioned within the nucleophile elbow.

Identification of a Ser-Asp-His Catalytic Triad - To investigate whether His-399 and Ser-286 are part of a catalytic triad, we searched for an acidic residue that would complete the triad. The common order in which the catalytic residues appear in the primary sequence of α/β-hydrolase fold enzymes is Ser- (Asp/Glu)-His (42, 46). Therefore, aspartate and glutamate residues located between Ser-286 and His-399 of OatC (Fig. 4*A*) were likely candidates for the missing catalytic acid. Consequently, we generated OatC variants with the single alanine substitutions E336A, E345A, D359A, D371A, D376A, and E379A that were expressed in *E. coli* BL21(DE3) with a C-terminal hexahistidine tag. The analysis of bacterial lysates by SDS-PAGE combined with either Coomassie staining or Western blot analysis revealed similar expression levels for all variants (Fig. 4*B*). To omit purification of each variant, bacterial lysates were analyzed directly for enzymatic activity. Because the spectrophotometric assay is not compatible with a high background of protein-bound sulfhydryl groups as present in crude lysates, the radioactive incorporation assay was applied. With the single exception of OatC-D376A, all variants showed enzymatic activity similar to that of the wild-type enzyme. To confirm the observation that alanine substitution of Asp-376 results in a complete loss of activity, the respective OatC variant was purified to homogeneity (Fig. 4*D*) and was analyzed in the spectrophotometric assay. In this case, identical amounts of purified enzyme were used for wild-type OatC and OatC-D376A. However, no enzymatic activity was observed for OatC-D376A, demonstrating that Asp-376 is essential for catalysis (Fig. 4*E*) and completes the catalytic triad as the critical acid. It is noteworthy that all three OatC variants in which the catalytic triad was destroyed by the single alanine substitution S286A, D376A, or H399A were found as dimers (see Table 2), suggesting that the introduced mutations did not affect the overall folding of these proteins.

Capsule O-Acetylation in Serogroup C Meningococci Carrying Mutant Forms of oatC - To verify the critical role of Ser-286, Asp-376, and His-399 for OatC activity *in*

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

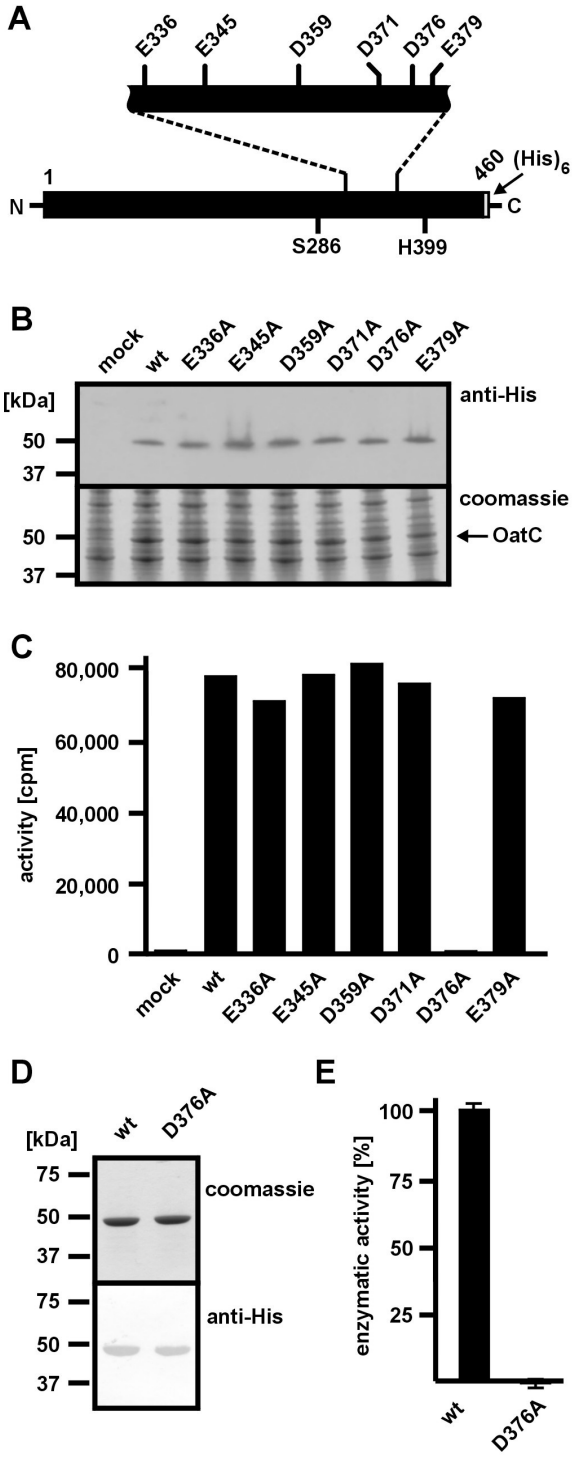


FIGURE 4. **Catalytic acid of OatC.** *A*, shown is a schematic representation of OatC, highlighting the positions of the six aspartate and glutamate residues located between Ser-286 and His-399. *B*, the single amino acid substitutions E336A, E345A, D359A, D371A, D376A, and E379A were introduced into OatC. After expression in *E. coli* BL21(DE3), bacterial lysates of all variants were separated by 10% SDS-PAGE and analyzed by Western blotting using anti-penta-His antibody for detection (*upper panel*) and Coomassie staining (*lower panel*). *C*, the OatC activity of wild-type (*wt*) and mutant forms was determined in bacterial lysates using a radioactive incorporation assay. *D*, wild-type OatC and the variant OatC-D376A were purified by IMAC and size-exclusion chromatography. Isolated proteins were analyzed by 10% SDS-PAGE followed by Coomassie staining (*upper panel*) and Western blot analysis with anti-penta-His antibody (*lower panel*). *E*, the enzymatic activities of the wild-type enzyme and OatC-D376A were determined in a spectrophotometric assay using equimolar concentrations of each purified enzyme variant. Data represent means ± S.D. of three independent experiments measured in triplicates, and the value obtained for wild-type OatC was set to 100%.

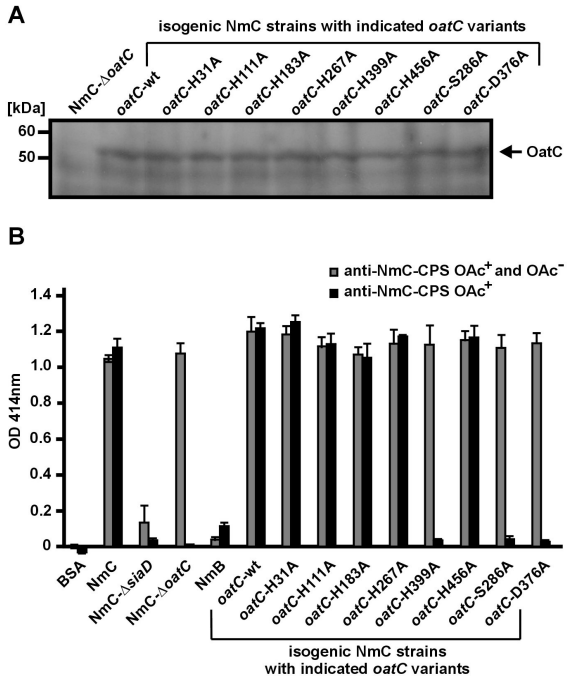


FIGURE 5. **Capsule *O*-acetylation in isogenic NmC strains with different *oatC* variants.** A set of isogenic NmC strains carrying the indicated *oatC* variants was generated by homologous recombination. In all cases, genes extended by a sequence encoding a C-terminal hexahistidine tag were used. *A*, shown are the results from Western blot analysis of bacterial lysates stained with anti-penta-His antibody. As a negative control, an NmC strain lacking *oatC* was used (NmC-Δ*oatC*). *B*, the *O*-acetylation status of the indicated strains was investigated in a whole-cell enzyme-linked immunosorbent assay with mAb1125 (*black bars*), which is specific for the *O*-acetylated (OAc^+) form of NmC-CPS. Capsule expression was controlled with mAb 924 (*gray bars*), which recognizes NmC-CPS irrespective of the *O*-acetylation status (OAc^+ and OAc^-). As a negative control, bovine serum albumin (*BSA*) was used instead of bacteria. The specificity of the primary antibodies was controlled by using the following strains: a wild-type (*wt*) NmC strain carrying a functional *oatC* gene (NmC), an NmC deletion mutant lacking the polysialyltransferase gene (*siaD*) essential for capsule synthesis (NmC-Δ*siaD*), an NmC deletion mutant lacking *oatC* (NmC-Δ*oatC*), and a serogroup B strain (NmB).

vivo, we generated a set of isogenic NmC strains carrying different *oatC* variants. Wild-type *oatC* and mutant forms, extended by a sequence encoding a C-terminal hexahistidine tag, were inserted by homologous recombination into the *oatC* locus of NmC strain 2120 (32), thereby replacing the endogenous wild-type gene. OatC expression was monitored in bacterial lysates by Western blot analysis with anti-penta-His antibody. As shown in Fig. 5*A*, all isogenic strains expressed comparable amounts of OatC, whereas no enzyme was detected in the deletion mutant NmC-Δ*oatC* (22), used as a negative control. The capsule *O*-acetylation competence of all strains was monitored in a whole-cell enzyme-linked immunosorbent assay using two capsule-specific antibodies: mAb 924, which binds to α2,9-linked polySia irrespective of the presence or absence of *O*-acetylation, and mAb 1125, which recognizes exclusively the *O*-acetylated form of NmC-CPS (32). As shown in Fig. 5*B*, both antibodies gave strong signals on *oatC*-positive NmC wild-type strain 2120. By contrast, complete loss of the OAc-specific signal was observed for the

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

respective null mutant NmC-Δ*oatC*, although capsule expression *per se* was not affected as demonstrated by a strong signal obtained with mAb 924. The specificity of the antibodies used was confirmed by the absence of antibody binding on NmC-Δ*siaD*, which lacks the polysialyltransferase essential for capsule synthesis (51), and on NmB, which expresses a polySia capsule composed of α2,8-linked Neu5Ac (Fig. 5*B*). Analysis of the newly generated isogenic NmC strains revealed capsule *O*-acetylation in those strains that express an enzyme variant that was proved active *in vitro* (here, variants H31A, H111A, H183A, H267A, and H456A are shown as an example). By contrast, no signal with OAc-specific antibody mAb 1125 was obtained for mutant strains comprising amino acid substitution H399A, S286A, or D376A. The lack of mAb 1125 binding was not due to reduced capsule expression because in all cases strong signals were obtained with mAb 924. Thus, in line with the *in vitro* findings, no enzymatic activity was found for OatC-H399A, OatC-S286A, and OatCD376A *in vivo*. Together, these results confirm the presence of a catalytic triad composed of Ser-286, Asp-376, and His-399.

Secondary Fold Prediction - Members of the α/β-hydrolase fold family share a conserved eight-stranded mostly parallel β-sheet, flanked on both sides by helices. The presence of this fold might become evident from the arrangement of the secondary structure elements. As shown in supplemental Fig. S2, we propose a topology diagram for OatC that is based on secondary structure prediction (52). Whereas the first helix (αA) and the first three β-strands (β1–3) could not be unequivocally assigned, secondary structure elements predicted for amino acids 259–422 show the typical topology of a "canonical" α/β-fold. The positions of the catalytic residues Ser-286, Asp-376, and His-399 after strands β5, β7, and β8, respectively, also matched the commonly found location of the catalytic triad (46, 53, 54).

Catalytic Mechanism of OatC - Most α/β-hydrolase fold enzymes act via a double-displacement (ping-pong) mechanism (42, 55–58). In the case of acetyltransferases belonging to the α/β-hydrolase fold family, this involves the formation of a covalent acetyl-enzyme intermediate that occurs by transfer of the acetyl group from acetyl-CoA onto the serine residue located in the nucleophile elbow (47, 58, 59). In the second half of the reaction, the acetyl group is further transferred onto the acceptor substrate.

To gain the first insight into the catalytic mechanism of OatC, we monitored the enzyme reaction for the occurrence of a stable acetyl-enzyme intermediate. Therefore, purified OatC was incubated with [^{14}C]acetyl-CoA, and the reaction mixture was analyzed before and after addition of the acceptor substrate by SDS-PAGE followed by autoradiography (Fig. 6*A*) or combined Alcian blue/silver staining as a loading control (Fig. 6*B*). For wild-type OatC, a radioactively labeled protein band migrating with the expected molecular mass of 54.3 kDa was observed, indicating the appearance of an acetyl-enzyme intermediate. Upon addition of NmC-CPS, this intermediate collapsed, and radiolabeled NmC-CPS appeared (Fig. 6*A*), which is consistent with the transfer of the radioactive acetyl group from the enzyme onto the acceptor substrate. By contrast, no acetyl-enzyme intermediate was detected for OatC-S286A, suggesting that the hydroxyl group of Ser-286 acts as acceptor site for the acetyl group. In the case of the site-specific variant D376A, the appearance of an acetyl-enzyme intermediate could be detected, although the amount of radiolabeled enzyme was lower than for the wild-type enzyme. However, addition of the acceptor substrate did not result in a collapse of the intermediate, and accordingly, no transfer onto NmC-CPS was detected (Fig. 6*A*). This result indicates that Asp-376 is dispensable for the first half of the reaction but essential for the second half. Similar to the S286A exchange, alanine substitution of His-399 abolished the formation of an acetyl-enzyme intermediate, suggesting an important role of His-399 during transfer of the acetyl group onto the enzyme. Thus, analysis of wild-type and mutant OatC demonstrated that *O*-acetylation of NmC-CPS proceeds via a covalent acetyl-enzyme intermediate that involves Ser-286.

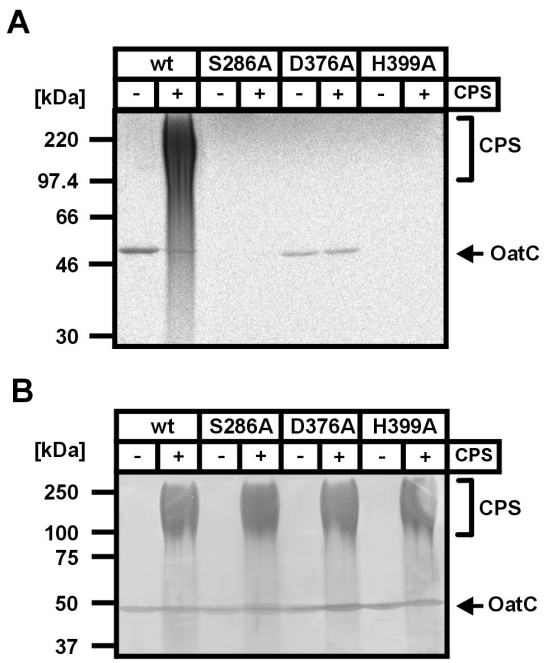


FIGURE 6. **Identification of an acetyl-enzyme intermediate.** Purified wild-type OatC and variants containing the single amino acid exchanges S286A, D376A, and H399A were incubated with [^{14}C]acetyl-CoA. Before (-) and after (+) addition of the acceptor substrate NmC-CPS, aliquots of the reaction mixture were analyzed by 10% SDS-PAGE followed by phosphorimaging to visualize radioactively labeled enzyme and acceptor substrate (*A*) or combined Alcian blue/silver staining to control identical loading of enzyme and NmC-CPS (*B*). Bands representing OatC and NmC-CPS are indicated on the *right*.

DISCUSSION

In this study, we performed the first purification and biochemical characterization of OatC, the capsule *O*-acetyltransferase of serogroup C meningococci. The enzyme proved to be highly specific for polySia joined by α2,9-linkages and did not act on α2,8-linked polySia or on sialic acid-containing heteropolymers such as CPS of serogroups W-135 and Y. Moreover, free Neu5Ac and CMP-Neu5Ac were not modified by OatC, indicating that *in vivo O*-acetylation of NmC-CPS occurs exclusively at

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

the polymer level. Similar observations have been made for NeuO, a prophage-encoded sialate *O*-acetyltransferase found in *E. coli* K1 strains that harbor the CUS-3 phage (25). NeuO catalyzes the *O*-acetylation of K1-CPS, and the enzyme acts exclusively on α2,8-linked polySia with at least 14 sialic acid residues (26, 60).

Despite the functional similarity between NeuO and OatC, our data presented in this study point out that the two enzymes evolved independently and belong to structurally diverse protein families. As shown recently by homology modeling and site-directed mutagenesis studies, NeuO belongs to the hexapeptide repeat family of acyltransferases (26). Members of this protein family contain a unique triangular left-handed β-helix fold with a protruding loop that harbors the catalytic histidine residue. They assemble into catalytic trimers in which the protruding loop of one subunit embraces the left-handed β-helix domain of the adjacent subunit, resulting in three identical active sites at the interfaces between the subunits (24, 61–63). In the case of NeuO, hexamers were found in solution, which might be dimers of trimers (26).

In contrast to NeuO, the primary sequence of OatC is devoid of hexapeptide repeats. Moreover, BLAST search analyses did not reveal significant sequence similarities between OatC and any other protein (22), impeding the identification of phylogenetic relationships. However, by combining bioinformatic approaches (motif scanning and secondary structure prediction), site-directed mutagenesis, and biochemical analyses, we have now identified several features that characterize OatC as a new member of the α/β-hydrolase fold family of proteins. The α/β-hydrolase fold is a versatile and widespread protein architecture that is found in enzymes with diverse functions such as proteases, lipases, esterases, dehalogenases, haloperoxidases, lyases, and epoxide hydrolases (42, 50, 64–70). The various members of this large superfamily usually lack significant amino acid sequence similarities but share a common fold that gave the family the name. The canonical α/β-hydrolase fold is composed of an eight-stranded mainly parallel β-structure that arranges in a twisted β-sheet, flanked on both sides by helices (42, 46). The fold provides a stable scaffold for a catalytic triad consisting of a nucleophile (serine, cysteine, or aspartate), an acid (aspartate or glutamate), and a histidine, with the nucleophile located at the tip of a sharp turn, the so-called “nucleophile elbow”. However, many variations from the canonical fold have been observed mainly due to the absence of particular strands or helices and/or insertions that can range from a few amino acids to entire domains. Based on the large number of solved crystal structures, the essential features to identify the fold are the presence of at least five parallel β-strands, the order of the catalytic triad (nucleophile-acid-histidine), and a nucleophile elbow at the top of canonical strand β5 (54).

Although a three-dimensional structure is required to unequivocally confirm the presence of a particular fold, our data provide strong evidence that OatC adopts an α/β-hydrolase fold. A series of single amino acid exchanges led to the identification of a catalytic triad composed of Ser-286, Asp-376, and His-399. Ser-286 was found to be part of a typical nucleophile elbow motif (G*X*S*X*GG), indicating that similar to other α/β-hydrolase

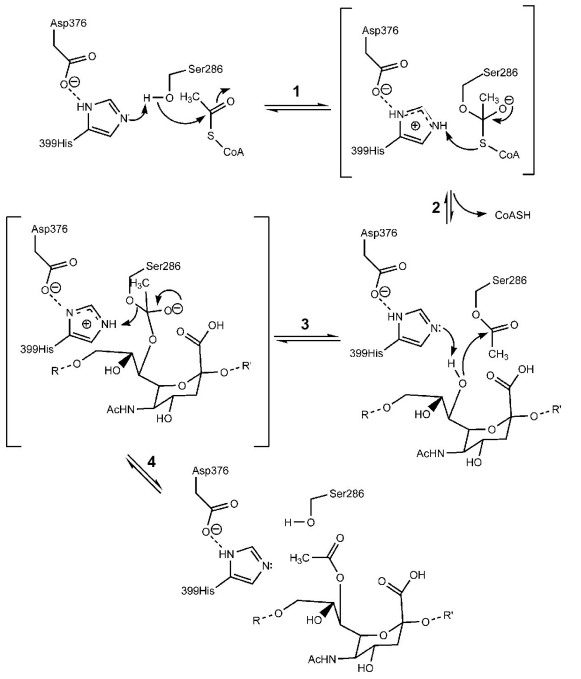


FIGURE 7. **Proposed reaction mechanism of the OatC-catalyzed *O*-acetylation of NmC-CPS.** The reaction proceeds through the following steps: *step 1*, nucleophilic attack of the carbonyl group of acetyl-CoA by Ser-286 and formation of the first tetrahedral intermediate; *step 2*, decomposition of the tetrahedral intermediate, release of free CoA, and formation of an acetyl-enzyme intermediate; *step 3*, nucleophilic attack by the C-7 (or C-8) hydroxyl group of the acceptor polysaccharide and formation of a second tetrahedral intermediate; and *step 4*, release of the second product (*O*-acetylated NmC-CPS) and regeneration of free enzyme.

fold proteins, the nucleophile is located in a tight turn to expose the hydroxyl group of Ser-286 to nucleophilic attack of the incoming substrate. Moreover, secondary structure prediction suggested for OatC a pattern of β-strands and α-helices that matches the topology of the canonical α/β-hydrolase fold (42, 46, 53, 54). In the proposed topology model, the catalytic residues are located after strands β5, β7, and β8 for Ser-286, Asp-376, and His-399, respectively, which is consistent with the common positioning of the catalytic triad in α/β-hydrolase fold enzymes.

Members of the α/β-hydrolase fold family catalyze a variety of transformations. However, for almost all enzymes, a double-displacement (ping-pong) kinetic mechanism is described that involves a covalent substrate-enzyme intermediate with the substrate linked to the nucleophile of the catalytic triad (42, 58, 66, 71–76). In line with this, a stable acetyl-enzyme intermediate was observed for OatC that crucially depended on the presence of Ser-286. In addition, a CoA burst was observed in the absence of polySia (data not shown). This indicates that the first substrate is released before binding of the second substrate, which confirms a ping-pong mechanism. On the basis of our data obtained for wild-type OatC and mutant forms, we propose the following catalytic mechanism (Fig. 7). The carboxyl group of Asp-376 forms a salt bridge with the N-δ1 of His-399, allowing the N-ε2 of the imidazole ring to abstract a proton from Ser-286. Thereby, the nucleophilicity of the hydroxyl group of Ser-286 is increased, facilitating the

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

attack of the carbonyl carbon of acetyl-CoA. The resulting tetrahedral transition state subsequently collapses to an acetyl-enzyme intermediate with the acetyl moiety covalently attached to Ser-286. During the acetylation step, the imidazole group of His-399 transfers the proton of the serine hydroxyl to the sulfhydryl leaving group, and free CoA is released. The second half-reaction is also assisted by His-399 through acid/base catalysis and starts with the abstraction of a proton from the C-7 (or C-8) hydroxyl group of the acceptor polysaccharide. This enables the nucleophilic attack of the carbonyl carbon of the acetyl group in the acetyl-enzyme intermediate, resulting in a second tetrahedral transition state. Subsequent decomposition of the tetrahedral intermediate leads to release of the second reaction product, *i.e. O*-acetylated NmC-CPS. In this last step, Ser-286 is returned to its protonated state, which reinitializes OatC for a new catalytic cycle. Further studies are required to investigate whether OatC transfers acetyl groups to both the C-7 and C-8 hydroxyl groups. Using ^{13}C NMR spectroscopy, Bhattacharjee *et al.* (14) detected di-*O*-acetylated residues (Neu5,7,8Ac$_3$) in NmC- CPS, whereas no evidence for di-*O*-acetylated sialic acids was found by Lemercinier and Jones (17). For NmC-CPS containing only mono-*O*-acetylated residues, non-enzymatic migration of acetyl groups was observed. In freshly prepared CPS, *O*-acetylation was found predominantly at O-8, whereas after storage at room temperature, the majority of acetyl groups were shifted to O-7 (17, 77, 78). Therefore, OatC might be selective for O-8, and acetylation of O-7 might occur exclusively by non-enzymatic migration.

In agreement with the proposed critical roles of the catalytic residues of OatC, single alanine substitutions of Ser-286, Asp-376, and His-399 completely abolished OatC activity *in vitro* and *in vivo*. Replacement of either His-399 or Ser-286 already abrogated the ability to form a covalent acetyl-enzyme intermediate, which is consistent with the proposed crucial function of His-399 as a general base catalyst and the central role of Ser-286 as an attachment site for the acetyl group. In contrast to OatC-H399A and OatC-S286A, OatC-D376A was still able to form a limited amount of acetyl-enzyme intermediate, indicating that Asp-376 is dispensable during the first half-reaction but essentially required for catalysis of the second half-reaction. This finding is in accordance with the less pronounced role of the catalytic acid observed in many α/β-hydrolase fold enzymes. In several cases, alanine substitution of the catalytic acid was even compatible with enzyme catalysis, and considerably high levels of activity were observed (66, 76, 79).

We conclude that OatC adopts an α/β-hydrolase fold structure with a Ser-Asp-His catalytic triad that forms the molecular basis for *O*-acetylation of NmC-CPS by a ping-pong mechanism. This clearly distinguishes OatC from all other bacterial sialate *O*-acetyltransferases known so far, which all belong to the hexapeptide repeat family: (i) NeuO, responsible for *O*-acetylation of the *E. coli* K1 capsule; (ii) OatWY, involved in *O*-acetylation of sialic acids within the galactose- and glucose-containing heteropolymeric polySia capsules of serogroup W-135 and Y meningococci, respectively (22); (iii) NeuD of group B streptococci, required for *O*-acetylation of terminal α2,3-linked sialic acids capping CPS (28); and (iv) NeuD of *C. jejuni*, catalyzing the *O*-acetylation of terminal α2,8-linked sialic acids of the lipo-oligosaccharide (27). Compared with α/β-hydrolase fold enzymes, members of the hexapeptide repeat family not only adopt an entirely different structural fold, the left-handed β-helix fold, but also act by a different kinetic mechanism. Although they possess a strictly conserved catalytic histidine, this residue is not part of a catalytic triad, and kinetic data obtained for this enzyme family revealed a sequential kinetic mechanism (24, 62, 80, 81).

The observed striking differences between OatC and other bacterial sialic acid-modifying *O*-acetyltransferases indicate that the enzymatic function of sialate *O*-acetylation evolved independently on two distinct structural frameworks by convergent evolution. Whereas the hexapeptide repeat family of acyltransferases is restricted to microorganisms, α/β-hydrolase fold enzymes are present in all kingdoms of life. Interestingly, the α/β-hydrolase fold is also used as the structural basis for viral sialic acid-specific acetylesterases, the receptor-destroying enzymes of influenza C and coronaviruses that use *O*-acetylated sialic acids as host receptor structures (82, 83). Thus, this fold appears to be a likely candidate for the structural framework of eukaryotic sialate *O*-acetyltransferases, for which the genetic and structural bases are still unknown. Attempts to clone the respective genes, *e.g.* by expression cloning, failed, and purification and characterization of the enzymatic activities proved to be challenging (84–88). The new insight in the evolution of sialate *O*-acetyltransferases obtained in bacteria may help to identify similar enzymes in eukaryotes by searching for genes encoding potential α/β-hydrolase fold proteins with unknown function.

Acknowledgments—We thank Rita Gerardy-Schahn for helpful discussions and continuous support and Astrid Oberbeck for excellent technical assistance.

REFERENCES

1. Rosenstein, N. E., Perkins, B. A., Stephens, D. S., Popovic, T., and Hughes, J. M. (2001) *N. Engl. J. Med.* **344**, 1378–1388
2. Stephens, D. S. (2007) *FEMS Microbiol. Rev.* **31**, 3–14
3. Connolly, M., and Noah, N. (1999) *Epidemiol. Infect.* **122**, 41–49
4. Yazdankhah, S. P., Kriz, P., Tzanakaki, G., Kremastinou, J., Kalmusova, J., Musilek, M., Alvestad, T., Jolley, K. A., Wilson, D. J., McCarthy, N. D., Caugant, D. A., and Maiden, M. C. (2004) *J. Clin. Microbiol.* **42**, 5146–5153
5. Vogel, U., Hammerschmidt, S., and Frosch, M. (1996) *Med. Microbiol. Immunol.* **185**, 81–87
6. Spinosa, M. R., Progida, C., Tala, A., Cogli, L., Alifano, P., and Bucci, C. (2007) *Infect. Immun.* **75**, 3594–3603
7. Villwock, A., Schmitt, C., Frosch, M., and Kurzai, O. (2008) *Int. J. Med. Microbiol.* **298**, 591–597
8. Yazdankhah, S. P., and Caugant, D. A. (2004) *J. Med. Microbiol.* **53**, 821–832
9. Jodar, L., Feavers, I. M., Salisbury, D., and Granoff, D. M. (2002) *Lancet* **359**, 1499–1508
10. Vimr, E., and Lichtensteiger, C. (2002) *Trends Microbiol.* **10**, 254–257
11. Vimr, E. R., Kalivoda, K. A., Deszo, E. L., and Steenbergen, S. M. (2004) *Microbiol. Mol. Biol. Rev.* **68**, 132–153
12. Hildebrandt, H., Mühlenhoff, M., and Gerardy-Schahn, R. (2008) *Neurochem. Res.* 10.1007/s11064-008-9724-7
13. Finne, J., Finne, U., Deagostini-Bazin, H., and Goridis, C. (1983) *Biochem. Biophys. Res. Commun.* **112**, 482–487
14. Bhattacharjee, A. K., Jennings, H. J., Kenny, C. P., Martin, A., and Smith, I. C. (1975) *J. Biol. Chem.* **250**, 1926–1932
15. Bhattacharjee, A. K., Jennings, H. J., Kenny, C. P., Martin, A., and Smith, I. C. (1976) *Can. J. Biochem.* **54**, 1–8

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

16. Jennings, H. J., Bhattacharjee, A. K., Bundle, D. R., Kenny, C. P., Martin, A., and Smith, I. C. (1977) *J. Infect. Dis.* **136**, (suppl.) S78–S83
17. Lemercinier, X., and Jones, C. (1996) *Carbohydr. Res.* **296**, 83–96
18. Longworth, E., Fernsten, P., Mininni, T. L., Vogel, U., Claus, H., Gray, S., Kaczmarski, E., and Borrow, R. (2002) *FEMS Immunol. Med. Microbiol.* **32**, 119–123
19. Borrow, R., Longworth, E., Gray, S. J., and Kaczmarski, E. B. (2000) *FEMS Immunol. Med. Microbiol.* **28**, 189–191
20. Michon, F., Huang, C. H., Farley, E. K., Hronowski, L., Di, J., and Fusco, P. C. (2000) *Dev. Biol.* **103**, 151–160
21. Fusco, P. C., Farley, E. K., Huang, C. H., Moore, S., and Michon, F. (2007) *Clin. Vaccine Immunol.* **14**, 577–584
22. Claus, H., Borrow, R., Achtman, M., Morelli, G., Kantelberg, C., Longworth, E., Frosch, M., and Vogel, U. (2004) *Mol. Microbiol.* **51**, 227–239
23. Vaara, M. (1992) *FEMS Microbiol. Lett.* **76**, 249–254
24. Johnson, C. M., Roderick, S. L., and Cook, P. F. (2005) *Arch. Biochem. Biophys.* **433**, 85–95
25. Deszo, E. L., Steenbergen, S. M., Freedberg, D. I., and Vimr, E. R. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5564–5569
26. Bergfeld, A. K., Claus, H., Vogel, U., and Mühlenhoff, M. (2007) *J. Biol. Chem.* **282**, 22217–22227
27. Houliston, R. S., Endtz, H. P., Yuki, N., Li, J., Jarrell, H. C., Koga, M., van Belkum, A., Karwaski, M. F., Wakarchuk, W. W., and Gilbert, M. (2006) *J. Biol. Chem.* **281**, 11480–11486
28. Lewis, A. L., Hensler, M. E., Varki, A., and Nizet, V. (2006) *J. Biol. Chem.* **281**, 11186–11192
29. Vann, W. F., Liu, T. Y., and Robbins, J. B. (1978) *J. Bacteriol.* **133**, 1300–1306
30. Rode, B., Endres, C., Ran, C., Stahl, F., Beutel, S., Kasper, C., Galuska, S., Geyer, R., Mühlenhoff, M., Gerardy-Schahn, R., and Scheper, T. (2008) *J. Biotechnol.* **135**, 202–209
31. Tsai, C. M., and Frasch, C. E. (1982) *Anal. Biochem.* **119**, 115–119
32. Vogel, U., Morelli, G., Zurth, K., Claus, H., Kriener, E., Achtman, M., and Frosch, M. (1998) *J. Clin. Microbiol.* **36**, 2465–2470
33. Poolman, J. T., Kriz-Kuzemenska, P., Ashton, F., Bibb, W., Dankert, J., Demina, A., Frøholm, L. O., Hassan-King, M., Jones, D. M., Lind, I., Prakash, K., and Xujing, H. (1995) *Clin. Diagn. Lab. Immunol.* **2**, 69–72
34. Dodd, I. B., and Egan, J. B. (1990) *Nucleic Acids Res.* **18**, 5019–5026
35. Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2008) *Nucleic Acids Res.* **36**, D281–D288
36. Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1997) *Nat. Struct. Biol.* **4**, 567–577
37. Hitomi, K., Oyama, T., Han, S., Arvai, A. S., and Getzoff, E. D. (2005) *J. Biol. Chem.* **280**, 19127–19135
38. D'Andrea, L. D., and Regan, L. (2003) *Trends Biochem. Sci.* **28**, 655–662
39. Blondel, A., and Bedouelle, H. (1990) *Eur. J. Biochem.* **193**, 325–330
40. Blondel, A., and Bedouelle, H. (1991) *Protein Eng.* **4**, 457–461
41. Westwood, I. M., Kawamura, A., Fullam, E., Russell, A. J., Davies, S. G., and Sim, E. (2006) *Curr. Top. Med. Chem.* **6**, 1641–1654
42. Holmquist, M. (2000) *Curr. Protein Pept. Sci.* **1**, 209–235
43. Jogl, G., Hsiao, Y. S., and Tong, L. (2004) *Ann. N. Y. Acad. Sci.* **1033**, 17–29
44. Hodawadekar, S. C., and Marmorstein, R. (2007) *Oncogene* **26**, 5528–5540
45. Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C. V., and Falquet, L. (2004) *Nucleic Acids Res.* **32**, W332–W335
46. Nardini, M., and Dijkstra, B. W. (1999) *Curr. Opin. Struct. Biol.* **9**, 732–737
47. Wang, M., Liu, L., Wang, Y., Wei, Z., Zhang, P., Li, Y., Jiang, X., Xu, H., and Gong, W. (2007) *Biochem. Biophys. Res. Commun.* **363**, 1050–1056
48. Milkowski, C., and Strack, D. (2004) *Phytochemistry* **65**, 517–524
49. Hendle, J., Mattevi, A., Westphal, A. H., Spee, J., de Kok, A., Teplyakov, A., and Hol, W. G. (1995) *Biochemistry* **34**, 4287–4298
50. Mirza, I. A., Nazi, I., Korczynska, M., Wright, G. D., and Berghuis, A. M. (2005) *Biochemistry* **44**, 15768–15773
51. Claus, H., Vogel, U., Mühlenhoff, M., Gerardy-Schahn, R., and Frosch, M. (1997) *Mol. Gen. Genet.* **257**, 28–34
52. Combet, C., Blanchet, C., Geourjon, C., and Deleage, G. (2000) *Trends Biochem. Sci.* **25**, 147–150
53. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G., and Goldman, A. (1992) *Protein Eng.* **5**, 197–211
54. Heikinheimo, P., Goldman, A., Jeffries, C., and Ollis, D. L. (1999) *Structure (Lond.)* **7**, R141–R146
55. Blow, D. M., Birktoft, J. J., and Hartley, B. S. (1969) *Nature* **221**, 337–340
56. Kraut, J. (1977) *Annu. Rev. Biochem.* **46**, 331–358
57. Ishida, T., and Kato, S. (2004) *J. Am. Chem. Soc.* **126**, 7111–7118
58. Nazi, I., and Wright, G. D. (2005) *Biochemistry* **44**, 13560–13566
59. Born, T. L., Franklin, M., and Blanchard, J. S. (2000) *Biochemistry* **39**, 8556–8564
60. Higa, H. H., and Varki, A. (1988) *J. Biol. Chem.* **263**, 8872–8878
61. Pye, V. E., Tingey, A. P., Robson, R. L., and Moody, P. C. (2004) *J. Biol. Chem.* **279**, 40729–40736
62. Olsen, L. R., Huang, B., Vetting, M. W., and Roderick, S. L. (2004) *Biochemistry* **43**, 6013–6019
63. Olsen, L. R., and Roderick, S. L. (2001) *Biochemistry* **40**, 1913–1921
64. Fujino, S., Akiyama, D., Akaboshi, S., Fujita, T., Watanabe, Y., and Tamai, Y. (2006) *Biosci. Biotechnol. Biochem.* **70**, 377–386
65. Kolkenbrock, S., Parschat, K., Beermann, B., Hinz, H. J., and Fetzner, S. (2006) *J. Bacteriol.* **188**, 8430–8440
66. Lee, L. C., Lee, Y. L., Leu, R. J., and Shaw, J. F. (2006) *Biochem. J.* **397**, 69–76
67. Nandhagopal, N., Yamada, A., Hatta, T., Masai, E., Fukuda, M., Mitsui, Y., and Senda, T. (2001) *J. Mol. Biol.* **309**, 1139–1151
68. Polgar, L. (2005) *CMLS Cell. Mol. Life Sci.* **62**, 2161–2172
69. Srivatsan, M. (2006) *Bioinformation* **1**, 281–284
70. Argiriadi, M. A., Morisseau, C., Hammock, B. D., and Christianson, D. W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10637–10642
71. Campobasso, N., Patel, M., Wilding, I. E., Kallender, H., Rosenberg, M., and Gwynn, M. N. (2004) *J. Biol. Chem.* **279**, 44883–44888
72. Dodson, G., and Wlodawer, A. (1998) *Trends Biochem. Sci.* **23**, 347–352
73. Ishida, T., and Kato, S. (2003) *J. Am. Chem. Soc.* **125**, 12035–12048
74. Ishida, T. (2006) *Biochemistry* **45**, 5413–5420
75. Wahab, H. A., Ahmad Khairudin, N. B., Samian, M. R., and Najimudin, N. (2006) *BMC Struct. Biol.* **6**, 23
76. Weadge, J. T., and Clarke, A. J. (2007) *Biochemistry* **46**, 4932–4941
77. Ho, M. M., Lemercinier, X., Bolgiano, B., Crane, D., and Corbel, M. J. (2001) *Biotechnol. Appl. Biochem.* **33**, 91–98
78. Ravenscroft, N., Averani, G., Bartoloni, A., Berti, S., Bigio, M., Carinci, V., Costantino, P., D'Ascenzi, S., Giannozzi, A., Norelli, F., Pennatini, C., Proietti, D., Ceccarini, C., and Cescutti, P. (1999) *Vaccine* **17**, 2802–2816
79. Wajant, H., and Pfizenmaier, K. (1996) *J. Biol. Chem.* **271**, 25830–25834
80. Wyckoff, T. J., and Raetz, C. R. (1999) *J. Biol. Chem.* **274**, 27047–27055
81. Hindson, V. J., and Shaw, W. V. (2003) *Biochemistry* **42**, 3113–3119
82. Rosenthal, P. B., Zhang, X., Formanowski, F., Fitz, W., Wong, C. H., Meier-Ewert, H., Skehel, J. J., and Wiley, D. C. (1998) *Nature* **396**, 92–96
83. Zeng, Q., Langereis, M. A., van Vliet, A. L., Huizinga, E. G., and de Groot, R. J. (2008) *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9065–9069
84. Lrhorfi, L. A., Srinivasan, G. V., and Schauer, R. (2007) *Biol. Chem.* **388**, 297–306
85. Butor, C., Diaz, S., and Varki, A. (1993) *J. Biol. Chem.* **268**, 10197–10206
86. Shen, Y., Tiralongo, J., Iwersen, M., Sipos, B., Kalthoff, H., and Schauer, R. (2002) *Biol. Chem.* **383**, 307–317
87. Shen, Y., Tiralongo, J., Kohla, G., and Schauer, R. (2004) *Biol. Chem.* **385**, 145–152
88. Iwersen, M., Dora, H., Kohla, G., Gasa, S., and Schauer, R. (2003) *Biol. Chem.* **384**, 1035–1047

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

SUPPLEMENTAL FIGURES

construct	oligo		sequence 5'-3'
pET22b-Δ*Strep*	AKB 42	s	CTAGAAATAATTTTTCGCGACTTTAAGAAGGAGACG
	AKB 43	as	GATCCGTCTCCTTCTTAAAGTCGCGAAAAATTATTT
OatC	MM 269	s	CGGGATCCATGTCAATCAATACGTTTG
	MM 271	as	GTCCGCTCGAGACATAAGAATTTATGTAATC
OatC-ΔN34	AKB 67	s	CGCGGATCCATGTTGGAAACCATCCGAATG
	MM 271	as	GTCCGCTCGAGACATAAGAATTTATGTAATC
OatC-ΔN103	AKB 69	s	CGCGGATCCATGATCTCAGCAAGATTTAAAC
	MM 271	as	GTCCGCTCGAGACATAAGAATTTATGTAATC
MBP	AKB 129	s	CGCGGATCCATGAAAACTGAAGAAGGTAAACTG
	AKB 128	as	GGCTGCCTCGAGCCGGCCGGGTACCGAGC
MBP-OatC	AKB 72	s	GGCTGCAAGCTTATGAAAACTGAAGAAGGTAAACTGG
	AKB 128	as	GGCTGCCTCGAGCCGGCCGGGTACCGAGC
$OatC_{fusion}$	MM 269	s	CGGGATCCATGTCAATCAATACGTTTG
	AKB 81	as	GGCTGCCTCGAGAAGCTTTAGTTTTGGATTTAATAATAATGCTTTC
OatC-H31A	AKB 53	s	CGCATTACGGGATTTAGTAACGCGAAAGCGCTAAACTTGGAAACC
	AKB 54	as	GGTTTCCAAGTTTAGCGCTTTCGCGTTACTAAATCCCGTAATGCG
OatC-H111A	AKB 55	s	CTCAGCAAGATTTAAAGCGATGTGTCTCAAATATAGGC
	AKB 56	as	GCCTATATTTGAGACACATCGCTTTAAATCTTGCTGAG
OatC-H183A	AKB 57	s	CCCGATTATTTATCAATAATAGCGGCTATTGAAAATTTAGAG
	AKB 58	as	CTCTAAATTTTCAATAGCCGCTATTATTGATAAATAATCGGG
OatC-H267A	AKB 59	s	GCTTATGAATTGATCGCAGAGGCGATTATTAGCTTAG
	AKB 60	as	CTAAGCTAATAATCGCCTCTGCGATCAATTCATAAGC
OatC-H399A	AKB 61	s	CCTTTAAACTATTCAGGAGCCGGCTCTCTTATCCCATTAAC
	AKB 62	as	GTTAATGGGATAAGAGAGCCGGCTCCTGAATAGTTTAAAGG
OatC-H456A	AKB 65	s	CCAATCCTGATATGCGATTAGCGAAATTCTTATGTCTCGAGCACC
	AKB 66	as	GGTGCTCGAGACATAAGAATTTCGCTAATCGCATATCAGGATTGG
OatC-S286A	AKB 84	s	CTATTATTGGCTCAGCAAAAGGAGGAACTGCC
	AKB 85	as	GGCAGTTCCTCCTTTTGCTGAGCCAATAATAG
OatC-S286C	AKB 112	s	CTATTATTGGCTCATGCAAAGGAGGAACTGCC
	AKB 113	as	GGCAGTTCCTCCTTTGCATGAGCCAATAATAG
OatC-S285A	AKB 82	s	CTATTATTGGCGCATCAAAAGGAGGAACTGCC
	AKB 83	as	GGCAGTTCCTCCTTTTGATGCGCCAATAATAG
OatC-E336A	AKB 122	s	CAAAAATTTGCAGCATATTTCTCTTATAAC
	AKB 123	as	GTTATAAGAGAAATATGCTGCAAATTTTTG
OatC-E345A	AKB 124	s	AACTCAATTCTAGCAAGCAAATGTGCTC
	AKB 125	as	GAGCACATTTGCTTGCTAGAATTGAGTT
OatC-D359A	AKB 86	s	GCTCAAAAGTTAATAAATTTTGCTATTCTCTATAGAAATAAACTAAC
	AKB 87	as	GTTAGTTTATTTCTATAGAGAATAGCAAAATTTATTAACTTTTGAGC
OatC-D371A	AKB 88	s	CTAACTATTTTTTATGGCGCCAAATTTAAAATGGACGCCC
	AKB 89	as	GGGCGTCCATTTTAAATTTGGCGCCATAAAAAATAGTTAG
OatC-D376A	AKB 90	s	GGTGATAAATTTAAAATGGCCGCCCAAGAAGTATCTAC
	AKB 91	as	GTAGATACTTCTTGGGCGGCCATTTTAAATTTATCACC
OatC-E379A	AKB 126	s	ATGGACGCCCAAGCAGTATCTACAATTAG
	AKB 127	as	CTAATTGTAGATACTGCTTGGGCGTCCAT
pFS1	FS 1	s	GCGCGCGGGCCCATCCAATCAAAAAAGCGTGAAG
	FS 2	as	GCGCGCGGGCCCGCATGATTTTAGAATTGCAGAG

Fig. S1. **Oligonucleotides used in the present study.** Sequences of sense (s) and antisense (as) primers used in this study are listed in 5'-3'-direction.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

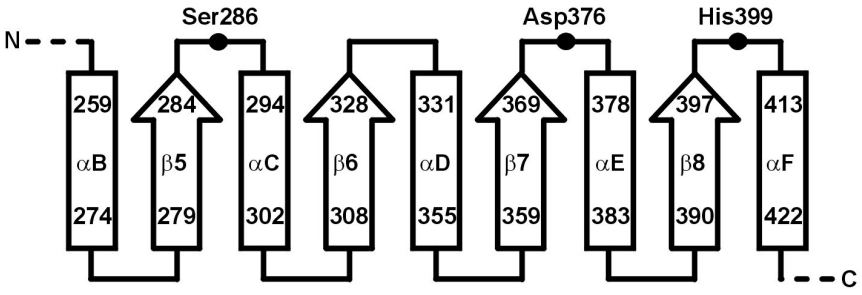


Fig. S2. **Topology diagram of OatC based on secondary structure prediction.** The primary sequence of OatC was analysed by secondary structure prediction. For amino acids 259-422 the typical topology of a 'canonical' fold could be assigned. The position of the catalytic residues Ser-286, Asp-376 and His-399 after β-strands 5, 7 and 8, respectively, also matched the commonly found localisation of the catalytic triad.

Originally published in *THE JOURNAL OF BIOLOGICAL CHEMISTRY*; VOL. **284**, NO. 1, pp. 6-16, January 2, 2009

Chapter 5 - Crystallization of the polysialic acid specific *O*-acetyltransferase OatC from *Neisseria meningitidis* serogroup C

5.1 Introduction

Neisseria meningitidis (*N. meningitidis* or meningococcus) is a common commensal that colonizes the human nasopharynx in up to 30% of healthy carriers at any one time (Claus *et al.*, 2005). Nonetheless, meningococci occasionally cause countrywide outbreaks of life-threatening systemic disease, making the pathogen a leading cause of sepsis and meningitis worldwide (van Deuren *et al.*, 2000 ; Stephens, 2007; Caugant, 2008). Although the pathogenic potential of the bacterium is associated with only a few hyper-invasive genetic clones of the species, the reasons how meningococci become invasive are not yet fully understood (Raymond *et al.*, 1997; Swartley *et al.*, 1997; Maiden *et al.*, 1998; Vogel *et al.*, 2000; Yazdankhah and Caugant, 2004; Caugant, 2008). Meningococci are differentiated into thirteen distinct serogroups based on the polysaccharide composition of their capsules (Gotschlich *et al.*, 1969). Among the six serogroups that account for nearly all reported invasive diseases worldwide, the four serogroups B, C, W-135 and Y are decorated with polysialic acid (polySia) capsules (Rosenstein *et al.*, 2001; Jodar *et al.*, 2002; Jazdankhah and Caugant, 2004; Boisier *et al.*, 2007; Stephens, 2007). The homopolymeric capsular polysaccharide (CPS) of *N. meningitidis* serogroup C is composed of α2,9-linked polySia and the capsular Sia residues have been found to be further modified by *O*-acetylation *in vivo* (Bhattacharjee *et al.*, 1975). In 2004, the gene encoding the respective *O*-acetyltransferase OatC was identified and shown to undergo phase variation, which explains the high variability in this phenotypic trait (Claus *et al.*, 2004). *OatC* is located downstream of the polySia synthesis genes *siaA–D* within the capsule gene complex of serogroup C meningococci. With the notable exception of OatC, primary sequence analyses of all known bacterial sialate *O*-acetyltransferases revealed significant similarities towards the hexapeptide repeat family of acyltransferases (Claus *et al.*, 2004; Deszo *et al.*, 2005; Lewis *et al.*, 2006; Houliston *et al.*, 2006; Bergfeld *et al.*, 2007). The hallmark of these enzymes is the presence of imperfect tandem repeats with the consensus sequence (LIV)(GAED)X_2(STAV)X which adopt into a unique left-handed β-helix fold that assembles into catalytic trimers (Vaara, 1992; Johnson *et al.*, 2005). OatC lacks these tandem repeats and biochemical analyses showed that OatC contains several typical features of an α/β-hydrolase fold enzyme: a Ser-Asp-His catalytic triad, a nucleophile elbow motif containing the catalytic serine, a secondary structure topology matching the ”canonical” α/β-hydrolase fold, and a double-displacement kinetic mechanism (Bergfeld *et al.*, 2009). However, so far, no

structural data are available for any sialic acid-specific *O*-acetyltransferase and nothing is known about the molecular basis of substrate selectivity for sialic and polysialic acid. The first successful isolation of polySia-specific *O*-acetyltransferases was recently described for the capsule modifying *O*-acetyltransferase NeuO of *Escherichia coli* K1 and for OatC (Bergfeld *et al*., 2007; Bergfeld *et al*., 2009). The availability of the recombinant enzymes in high yields and purity is now forming the basis for comprehensive structural analyses. A currently ongoing project is therefore the determination of the 3D-crystal structures of NeuO and OatC. This chapter summarizes the recent progress made in crystallizing OatC. First crystallization trials were accomplished with an OatC variant carrying a C-terminal hexa-histidine tag and initial crystallization conditions were identified. Refinement of the initial hit conditions resulted in OatC crystals that diffracted X-rays to almost 3.8 Å. To get better diffracting crystals, several optimization strategies, including crystallization of OatC lacking a hexa-histidine tag, were implemented and first results are presented.

5.2 Materials and methods

5.2.1 Materials

Purified de-*O*-acetylated CPS from *N. meningitidis* serogroup C was kindly provided by Baxter. pET expression vectors and *E. coli* BL21(DE3) were obtained from Novagen. 5,5'-dithiobis(2-nitrobenzoic acid), dithiothreitol, D-(+)-glucose and acetyl-CoA, were purchased from Sigma. Polyethylene glycols (PEGs), PEG-MME 550, 2-methyl-2,4-pentanediol, bicine and bistris were obtained from Fluka and HEPES as well as glycerol were purchased from AppliChem. Tris(hydroxymethyl)aminomethane (Tris-HCl), sodium acetate, ammonium sulfate and calcium chloride dihydrate were bought from Merck and 2-(*N*-morpholino)-ethanesulfonic acid (MES) was obtained from Calbiochem. Sodium chloride, 2-propanol and ethanol were purchased from J.T.Baker and Jeffamine ED-2001® as well as pentaerythritol ethoxylate (15/4 EO/OH) were obtained from Hampton Research.

5.2.2 Generation of OatC expression plasmids

All constructs for the expression of OatC variants were generated in the prokaryotic expression vector pET22b-Δ*Strep*, allowing the expression of OatC variants with a C-terminal hexahistidine tag. Generation of the respective vector pET22b-Δ*Strep* as well as the plasmids for expression of OatC, ΔN34-OatC and OatC-H456A was described previously (Bergfeld *et al.*, 2009). Constructs for the expression of the N-terminally or C-terminally truncated forms ΔN17-OatC and OatC-ΔC2 were generated by PCR using the primer pairs AKB76/MM271 and MM269/AKB70, respectively. The obtained PCR products were subcloned into BamHI/XhoI sites of pET22b-Δ*Strep*. For isolation of untagged OatC, a CRV-protease cleavage site was introduced between the sequence of OatC and the C-terminal hexahistidine tag, allowing to cleave off the affinity tag after purification. Therefore adapter ligation of the oligonucleotides AKB135/AKB136 into the XhoI site of the pET22b-Δ*Strep* expression plasmid harboring the wild-type *oatC* gene (Bergfeld *et al.*, 2009) was performed. The primers used in the present study are listed in Table 5-1 and the identity of all generated constructs was confirmed by sequencing.

construct	oligo		sequence 5'-3'
ΔN17-OatC	AKB 76	s	CGGGATCCATGAATTTCGCAGAAATTGCTCGCATTAC
	MM 271	as	GTCCGCTCGAGACATAAGAATTTATGTAATC
OatC-ΔC2	MM 269	s	CGGGATCCATGTCAATCAATACGTTTG
	AKB 70	as	GTCCGCTCGAGGAATTTATGTAATCGCATATCAGGATTG
OatC-CRV-His	AKB 135	s	TCGAGGTACTGTTCCAGGGCCCGC
	AKB 136	as	TCGAGCGGGCCCTGGAACAGTACC

Table 5-1: Oligonucleotides used in the present study. Sequences of sense (s) and antisense (as) primers are listed in 5'-3' direction.

5.2.3 Large-scale purification of $(His)_6$-tagged OatC

Freshly transformed *E. coli* BL21(DE3) cells were cultivated at 37°C in 500 ml of Power Broth (AthenaES) containing 200 µg/ml carbenicillin. At OD_{600} = 1.4, expression was induced by adding 0.1 mM isopropyl-β-D-thiogalactopyranoside. Thereafter, bacteria were grown at 15°C for 20 h, harvested, and lysed by sonication. Recombinant OatC was isolated by immobilized metal affinity chromatography (IMAC) using a 5 ml HisTrapHP column (GE Healthcare) equilibrated with binding buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, and 20 mM imidazole). After washing with 20 ml of binding buffer, proteins were eluted with a 25 ml linear imidazole gradient (20–500 mM imidazole in binding buffer). Enzyme-containing fractions were pooled and desalted on a HiPrep^{TM}26/10 Desalting column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 7.5 and 100 mM NaCl to remove imidazole. Thereafter protein containing fractions were pooled and loaded onto a HiLoad^{TM} 16/60 Superdex^{TM} 200 preparative size-exclusion chromatography column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 7.5 and 100 mM NaCl at 4°C. OatC-containing fractions were pooled, concentrated to 8.4 mg/ml by ultra filtration (Amicon Ultra-15 ultracel-30k, Millipore), aliquoted and subsequently flash-frozen in liquid nitrogen.

Expression and purification of truncated OatC-variants as well as single alanine substitutions were performed as described previously (Bergfeld *et al.*, 2009).

5.2.4 Purification of untagged OatC

Transformed *E. coli* BL21 gold (DE3) cells were cultivated at 37°C in 500 ml of Power Broth (AthenaES) containing 200 µg/ml carbenicillin. At OD_{600} = 1.5, expression was induced by adding 0.1 mM isopropyl-β-D-thiogalactopyranoside. Thereafter, bacteria were grown at 15°C for 20 h, harvested, and lysed by sonication. Recombinant OatC-CRV-His was isolated by immobilized metal affinity chromatography (IMAC) using a 5 ml HisTrapHP column (GE Healthcare) equilibrated with binding buffer (50 mM TrisHCl pH 7.5, 300 mM NaCl, and 20 mM imidazole). After washing with 20 ml of binding buffer, proteins were eluted with a 30 ml linear imidazole gradient (20–500 mM imidazole in binding buffer). Enzyme-containing fractions were pooled and desalted on a HiPrep^{TM}26/10 Desalting column (GE Healthcare) equilibrated with 10 mM Tris-HCl pH 7.5 and 150 mM NaCl to remove imidazole. Thereafter protein containing fractions were pooled and supplemented with 10 U of $(His)_6$-tagged HRV-3C-protease (Novagen), whereupon the concentration of the TrisHCl buffer was increased to 50 mM and the concentration of NaCl was kept at 150 mM. The protease cleavage was performed over night at 4°C. Afterwards, the reaction mixture was run through a 1 ml HisTrapHP column (GE Healthcare) equilibrated with binding buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, and 20 mM imidazole) to remove the $(His)_6$-tagged protease and residual uncleaved OatC-CRV-His. The flow-through which contained the cleaved untagged OatC

was subsequently loaded onto a HiLoad™ 16/60 Superdex™ 200 size-exclusion chromatography column (GE Healthcare) equilibrated with 10 mM TrisHCl pH 7.5 and 150 mM NaCl at 4°C. OatC-containing fractions were pooled, concentrated to 9.5 mg/ml by ultra filtration (Amicon Ultra-15 ultracel-30k, Millipore), supplemented with 4 mM DTT, aliquoted and subsequently flash-frozen in liquid nitrogen.

5.2.5 SDS-PAGE, staining, and immunoblotting

SDS-PAGE was performed under reducing conditions using 2.5% (v/v) β-mercaptoethanol. Coomassie staining, and Western blot analysis with anti-penta-His antibody (Qiagen) were performed as described previously (Bergfeld *et al.*, 2007).

5.2.6 Analytical size-exclusion chromatography

The quaternary structure of OatC variants was determined on a Superdex 200 HR 10/30 column. The column was equilibrated with 50 mM TrisHCl pH 7.5 and 100 mM NaCl and calibrated with the following molecular mass standards (Sigma): thyroglobulin (669 kDa), *β*-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome *c* (12.4 kDa). Eluted proteins were monitored by absorbance at 280 nm.

5.2.7 Spectrophotometric activity assay

Enzymatic activity of isolated OatC variants was determined in a spectrophotometric activity assay as described previously (Bergfeld *et al.*, 2009). The reaction was performed at 25 °C in a total volume of 100 µl containing 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 20% (w/v) glycerol, 2 mM 5,5'-dithiobis(2-nitrobenzoic acid), 200 ng/ml purified NmC-CPS, and 1 mM acetyl-CoA. The reaction was initiated by adding 1.7 pmol of purified enzyme and monitored continuously at 405 nm in half-area 96-well plates (Greiner) using a PowerWave 340 microtiter plate spectrophotometer (BioTek).

5.2.8 Initial crystal screening

Initial screenings for OatC-crystals were performed at Europe's largest high-throughput crystallization facility at the EMBL Hamburg Outstation (Mueller-Dieckmann, 2006) using the *Cryo I* & *II sparse matrix* and *Wizard I* & *II sparse matrix* crystallization screens (Emerald BioSystems) as well as *Crystal Screen™ & Crystal Screen 2™, Grid Screen™ PEG 6000* & *Grid Screen™ Ammonium Sulfate, Index™,* and *SaltRX HT™* (Hampton Research Corp.). In addition the crystallization screens *The JCSG+ Suite* (Qiagen) as well as *JBScreen Classic 1-4* & *JBScreen Classic 5-8* (Jena Bioscience) were used. The crystallization conditions were assessed as 400 nL sitting drops (200 nL enzyme sample + 200 nL stock solution) in *Greiner*

Low profile crystallization plates. The plates were set up and stored at 19°C and automatic imaging of each condition was performed over a period of 6 month. The resulting 16 picture sets of each condition taken after 0, 0.5, 1, 3, 6, 10, 13, 17, 24, 38, 52, 59, 81, 109, 137 and 165 days, respectively, were accessible for downloading via the internet (http://icarus.embl-hamburg.de:2000).

5.2.9 Refinement of initial hits

All crystallization experiments which were performed in house to reproduce and optimize the initial hit conditions identified by the initial crystal screening described above, were carried out in 24-well Cryschem plates (Hampton Research) using the sitting drop vapor-diffusion technique. A reservoir volume of 500 µl was provided before 1 µl OatC-containing sample (8.4 mg/ml of $(His)_6$-tagged OatC and 9.5 mg/ml of untagged OatC, respectively) was mixed with 1 µl of the respective buffer reservoir. Screens stored at 20°C were set up at room temperature whereas screens stored at 4°C were prepared on ice using pre-cooled buffer solutions.

5.2.10 X-ray diffraction analyses

X-ray diffraction analyses were performed at the Department of Molecular Structural Biology, headed by Prof. Dr. R. Ficner, at the Georg-August-Universität Göttingen. Crystals suitable in size for X-ray diffraction measurements were flash-frozen in nylon-fiber loops at 100 K after cryo-protection. Initial X-ray diffraction analyses were carried out at the Department of Molecular Structural Biology. Promising crystals were kept in liquid nitrogen and reanalyzed using the beamline 23.2 (microfocus beamline) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France).

5.3 Results

5.3.1 Purification of OatC

The purification of active recombinant OatC with a C-terminal hexahistidine tag has been described previously (Bergfeld *et al.*, 2009; see chapter 4). However, since crystallization trials require large amounts of chemically and structurally homogenous protein, the original purification procedure was scaled-up to obtain higher yields. OatC with a C-terminal $(His)_6$-tag was expressed in *E. coli* BL21(DE3) and isolated by immobilized metal affinity chromatography (IMAC) using a Ni^{2+}-chelating column. The OatC-containing fractions were pooled and further purified by preparative size-exclusion chromatography. Fractions of the major peak were pooled (Fig. 5-1 A) and concentrated by ultra-filtration to a final concentration of 8.4 mg/ml. This 2-step purification protocol resulted in a total yield of 82.4 mg OatC from 4 L of expression culture. Analysis of the isolated enzyme by SDS-PAGE and Coomassie staining revealed a major band consistent with the calculated molecular mass of 54.3 kDa and the identity of this band was confirmed by Western blot analysis using an anti-His antibody (Fig. 5-1 B). Additional bands that were still visible in the protein fraction obtained by IMAC were almost completely removed by subsequent size-exclusion chromatography (compare lanes 1 and 3 in Fig. 5-1 B), leading to an OatC preparation that contained only trace amounts of impurities.

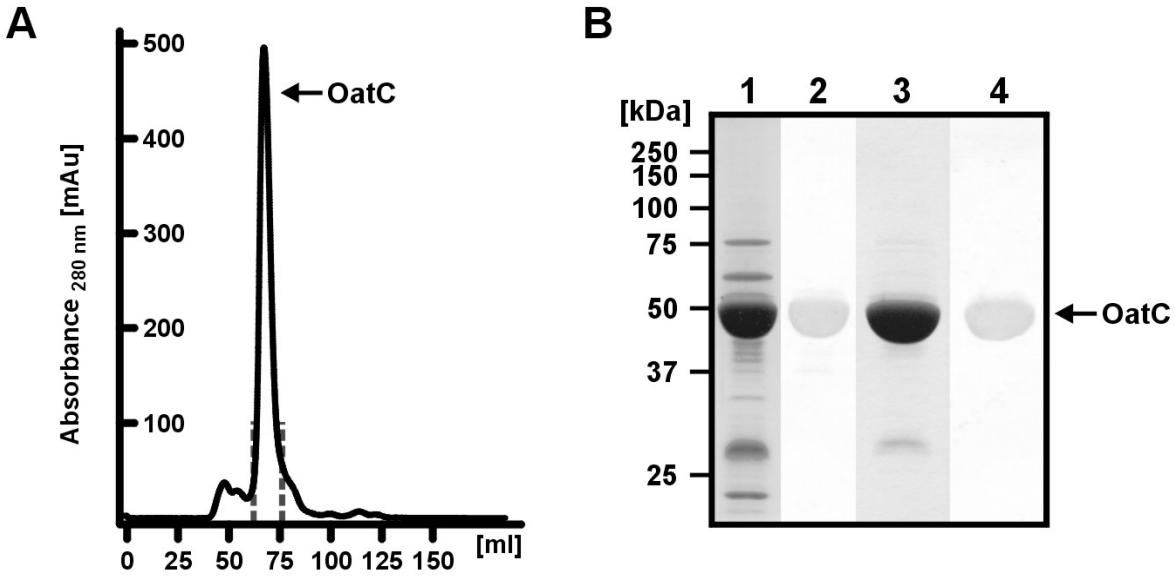


Figure 5-1: Purification of OatC for crystallization. (A), Elution profile of OatC separated by size-exclusion chromatography. OatC carrying a C-terminal $(His)_6$-tag was expressed in *E. coli* BL21(DE3). After a first isolation step by IMAC, OatC-containing fractions were pooled and subsequently loaded onto a HiLoad™ 16/60 Superdex™ 200 column. Enzyme containing peak fractions (indicated by dotted lines) were pooled and concentrated by ultra filtration to 8.4 mg/ml. (B), Isolated OatC was analyzed by 10% SDS-PAGE followed by Coomassie staining (lanes 1 and 3) or Western blot analysis with anti-His antibody (lanes 2 and 4). Shown are the pooled OatC-containing fractions after IMAC (lanes 1 and 2) and the pooled peak fractions of the size-exclusion chromatography after concentration to 8.4 mg/ml (lanes 3 and 4).

5.3.2 Crystallization of $(His)_6$-tagged OatC

Crystallization is an empirical process in which hundreds of solutions are scrutinized for their ability to convert soluble biological macromolecules into well ordered crystals that diffract X-rays to high resolution. The initial screening for suitable conditions that allow crystallization of OatC was done in an automated process at Europe's largest high-throughput crystallization facility at the EMBL Hamburg Outstation (Mueller-Dieckmann, 2006). The crystallization robot requires only 200 nL of sample volume per condition, thereby facilitating the testing of a multitude of conditions with relatively low amounts of protein. For initial crystallization trials of recombinant $(His)_6$-tagged OatC, 864 conditions were tested at 19°C. Over a period of 6 month, 16 sets of pictures were automatically taken from all conditions, monitoring the development of each condition over time. Most of the conditions tested for OatC led to the formation of precipitate or pre-crystalline compounds, and a total of 21 conditions resulted in needle-shaped crystals. Growth of these crystals started within 6 days and no additional positive conditions were identified within the observed timeframe of 6 month. The majority of crystals were obtained in the presence of polyethylene glycol (PEG) of different molecular masses (Fig. 5-2 A). Five initial hit conditions were based on ethanol (Fig. 5-2 B), one contained 2-propanol (Fig. 5-2 C), and one was based on the macromolecular reagent Jeffamine ED-2001® (O,O-Bis(2-aminopropyl)-polypropylene-glycol-block polyethylene-glycol-block poly-propylene-glycol 1,900) (Fig. 5-2 D).

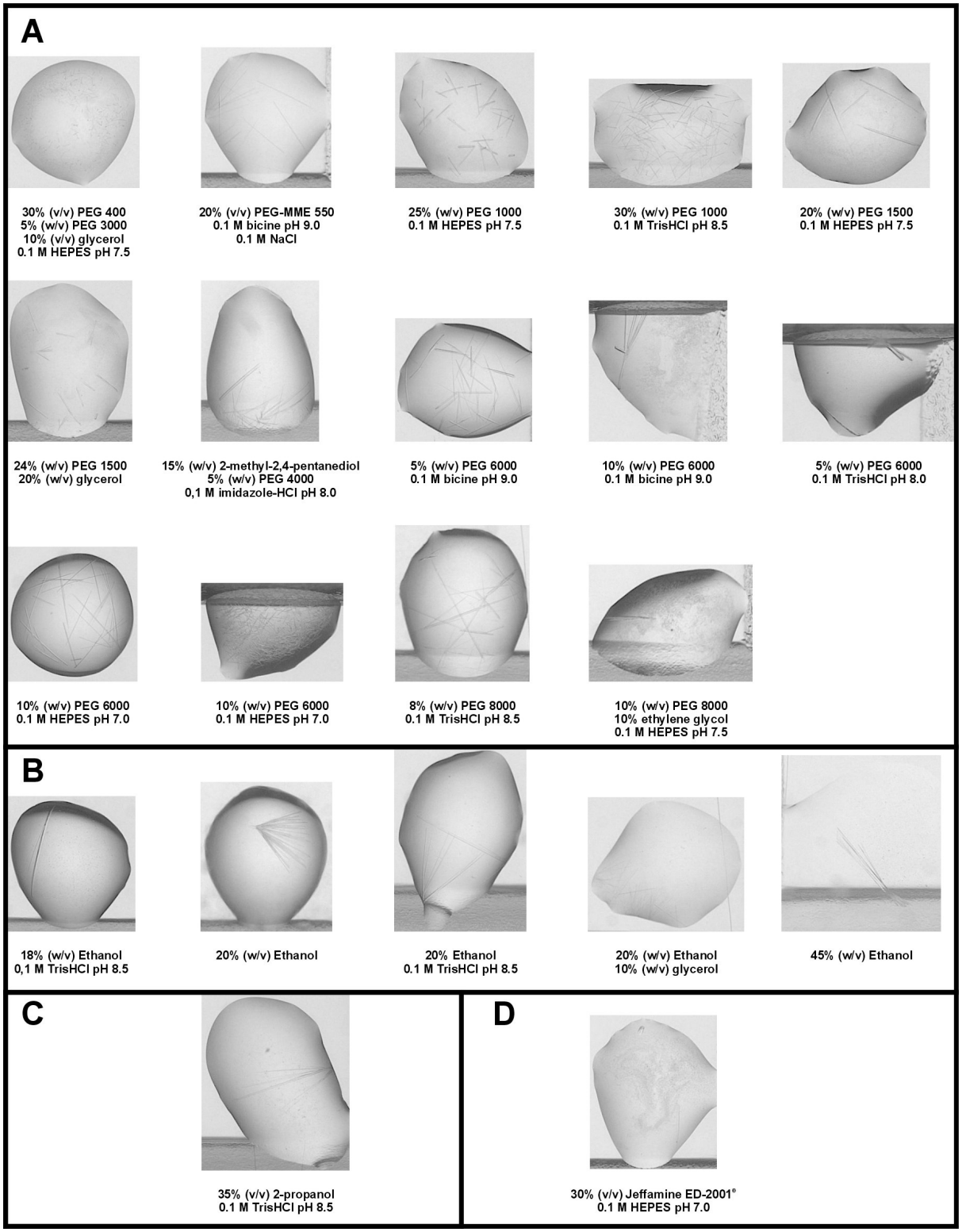


Figure 5-2: Initial hit conditions for the crystallization of $(His)_6$-tagged OatC. Purified OatC with a concentration of 8.4 mg/ml was used for initial crystallization trials performed at 19°C. Crystal conditions were assessed at the high-throughput crystallization facility at the EMBL Hamburg Outstation as 400 nL sitting drops (200 nL OatC + 200 nL stock solution) in Greiner Low profile crystallization plates. A total of 864 conditions was screened for crystal growth and pictures taken at day 6 are shown for the 21 conditions that contained clearly visible crystals. Compositions of the used stock solutions are displayed below each photo. (A) Crystals obtained with conditions based on different polyethylene glycols as precipitants; (B) Crystals obtained with ethanol as crystallant; (C) Crystals grown in the presence of 2-propanol, and (D) Crystals based on Jeffamine ED-2001® as precipitant.

5.3.3 Refinement of initial hit conditions

To achieve crystal sizes sufficient for X-ray diffraction measurements, crystallization of OatC-$(His)_6$ was repeated in house using the sitting drop vapor-diffusion technique and a drop size of 2 µl (1 µl OatC and 1 µl stock solution). With the single exception of the condition based on 2-propanol as precipitant (Fig. 5-2 C), all initial hits from the HT-screen could be reproduced or at least closely related conditions were found that allowed crystallization of OatC. To optimize the morphology of the crystals, conditions were refined by step-wise variation of the concentration of the precipitants, iterative testing of different buffers and pH values, and by performing additive screens. However, despite intensive screening, all crystals obtained in ethanol-containing conditions appeared as very long and extremely thin needles which in most cases were found as clusters (Fig. 5-3).

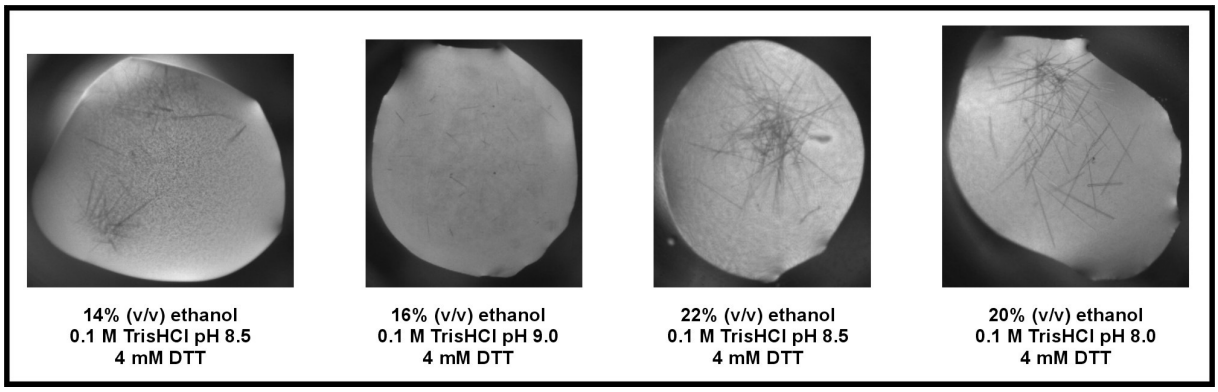


Figure 5-3: Optimization of OatC crystals based on ethanol as crystallant. Purified OatC with a concentration of 8.4 mg/ml was used for crystallization trials performed at 20°C. Crystal conditions were assessed in house as 2 µL sitting drops (1 µL OatC + 1 µL stock solution) in 24-well Cryschem plates. Pictures were taken 48 hours after the setup. Compositions of the used stock solutions are displayed below each photo.

For crystals obtained under conditions with Jeffamine ED-2001® as precipitant, the morphology was improved and crystal sizes sufficient for X-ray diffraction were obtained (Fig. 5-4).

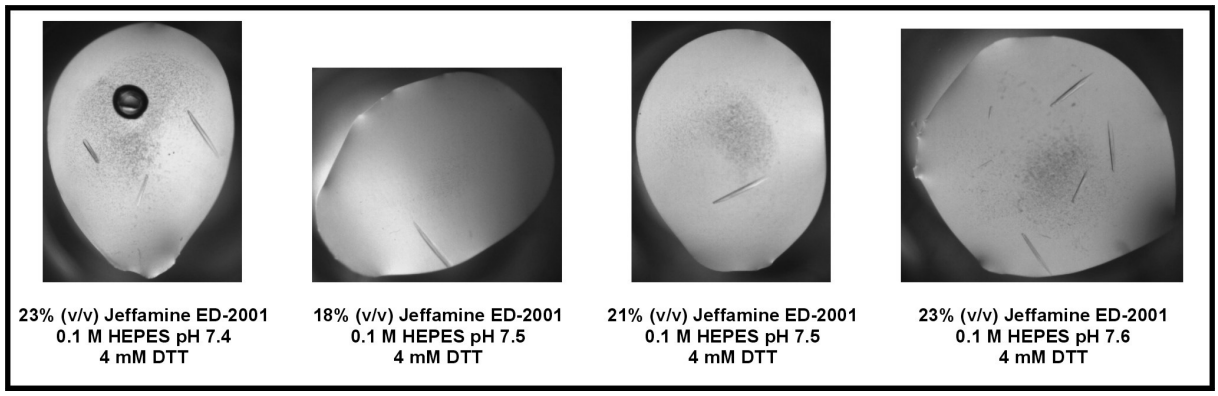


Figure 5-4: Refinement of OatC crystals based on Jeffamine ED-2001® as crystallant. Purified OatC with a concentration of 8.4 mg/ml was used for crystallization trials performed at 20°C. Crystal conditions were assessed in house as 2 µL sitting drops (1 µL OatC + 1 µL stock solution) in 24-well Cryschem plates. Pictures were taken 48 hours after the setup. Compositions of the used stock solutions are displayed below each photo.

Unfortunately, the subset of crystals grown with Jeffamine ED-2001® did not diffract X-rays. Although the crystals appeared uniform when viewed under the microscope, they are most likely not well ordered.

Among the various PEG conditions tested, PEG 1,000 turned out to be most suitable to get single crystals of sufficient size for X-ray diffraction analyses. A subset of crystals obtained with PEG 1,000 as precipitant is shown in Figure 5-5.

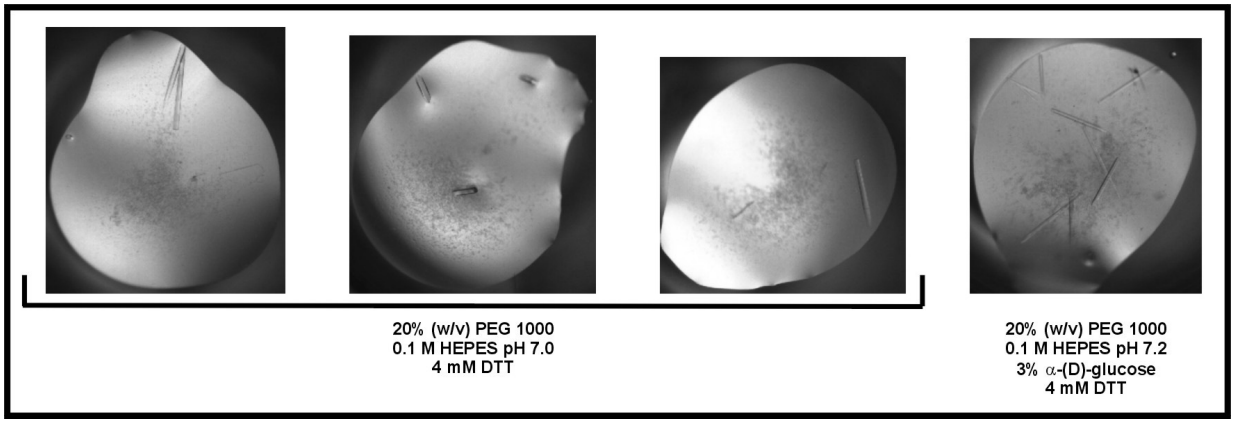


Figure 5-5: Optimization of OatC crystals based on PEG 1,000 as crystallant. Purified OatC with a concentration of 8.4 mg/ml was used for crystallization trials performed at 20°C. Crystal conditions were assessed in house as 2 µL sitting drops (1 µL OatC + 1 µL stock solution) in 24-well Cryschem plates. Pictures were taken 7 days after the setup. Compositions of the used stock solutions are displayed below each photo.

Various crystals grown in the presence of PEG 1,000 were analyzed and shown to diffract X-rays. Despite intensive screening to optimize crystallization conditions of $(His)_6$-tagged OatC with PEG 1,000 as precipitant, the obtained crystals diffracted X-rays only to a poor resolution of > 3.8 Å with 20% glycerol as cryoprotectant.

5.3.4 Crystallization trials with $(His)_6$-tagged OatC at 4°C

One possible explanation for the formation of poorly ordered crystals might be that OatC is not stable over time under the used crystallization conditions at 20°C. In this case, increasing incorporation of deranged enzyme molecules into the crystals will prevent regular crystal packing (Cudney *et al.*, 1994). To circumvent this problem, the crystallization temperature was shifted to 4°C and results are shown in Figure 5-6. At room temperature, OatC crystals appeared within 24 h, whereas growth rates were slightly decreased at 4°C and the first crystals were visible earliest at day 2.

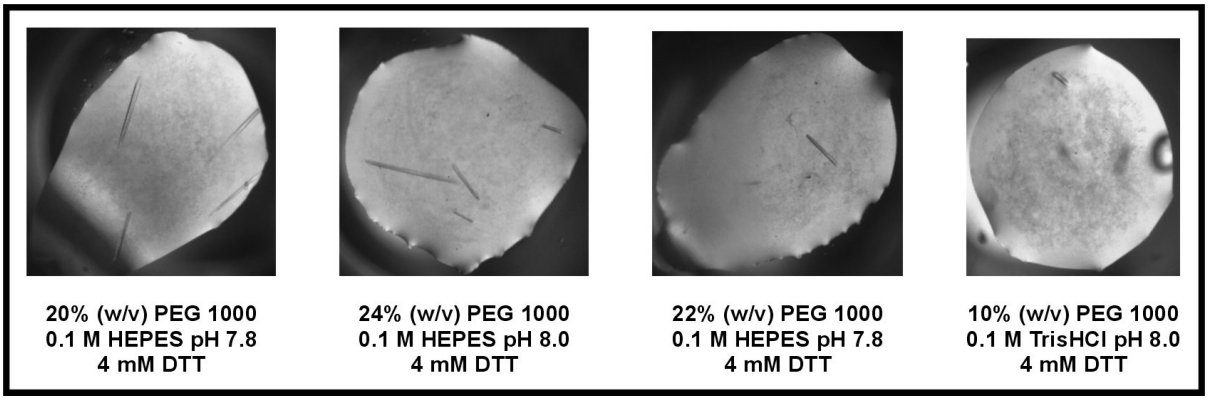


Figure 5-6: Crystals of $(His)_6$-tagged OatC grown at 4°C. Purified OatC with a concentration of 8.4 mg/ml was used for crystallization trials. Crystal conditions were assessed in house as 2 µL sitting drops (1 µL OatC + 1 µL stock solution) in 24-well Cryschem plates. Screens were set up on ice with pre-cooled reservoir solutions and subsequently stored at 4°C. Pictures were taken 9 days after the setup. Compositions of the used stock solutions are displayed below each photo.

So far, only a few initial crystals obtained at 4°C were analyzed and for these crystals the resolution of the diffraction pattern reached only ~ 4.5 Å.

5.3.5 Crystallization trials with $(His)_6$-tagged OatC in the presence of the donor substrate acetyl-coenzyme A

In the case of enzymes, the formation of co-crystals with substrates or substrate analogs is an alternative strategy to improve crystal packing (e.g. by changing the space group, stabilizing the enzyme, or inducing a particular conformation). Therefore, additional crystallization trials were set up for an OatC sample containing 8.4 mg/ml of enzyme and a 100-fold molar excess of the donor substrate acetyl-CoA. Because crystallization conditions of enzyme-substrate complexes frequently differ from those found for the corresponding apo-enzyme, a comprehensive initial screening was performed at the high-throughput crystallization facility at the EMBL Hamburg Outstation as described above for the apo-enzyme. Among the 864 conditions tested, seven resulted in crystal growth (Fig. 5-7). Two of them resembled crystallization conditions that were already found for the apo-enzyme (Fig 5-7 A). In addition, novel conditions were identified, such as small PEGs in combination with high pH-values of 9.5 (Fig. 5-7 B). Moreover, new conditions based either on PEG or propylene glycol as precipitants yielded crystals (Fig. 5-7 C) as well as a condition containing 3 M ammonium sulfate (Fig. 5-7 D). The presence of substrate during crystallization does not necessarily results in the formation of co-crystals. However, the identification of crystallization conditions that differ considerably from those found for the apo-enzyme, suggests that at least some of these conditions resulted in crystallization of an enzyme-substrate complex.

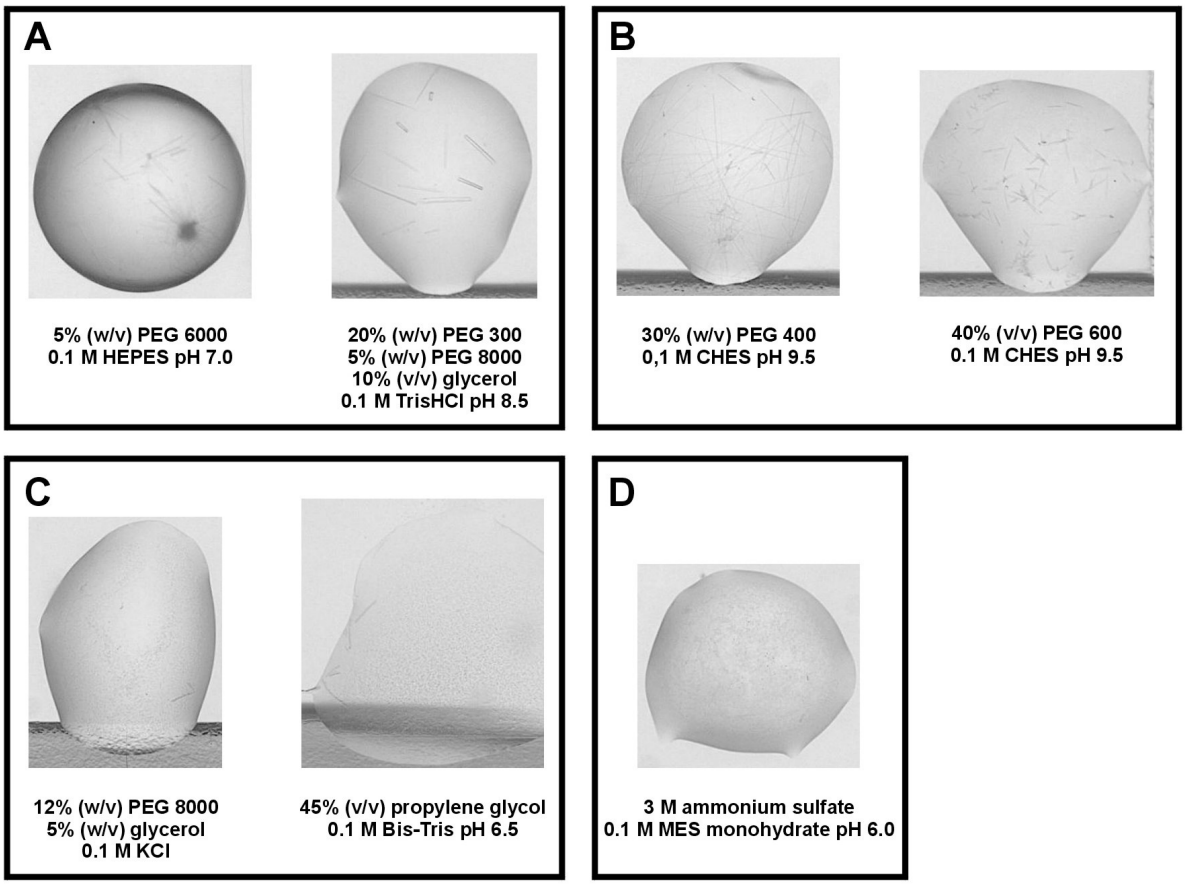
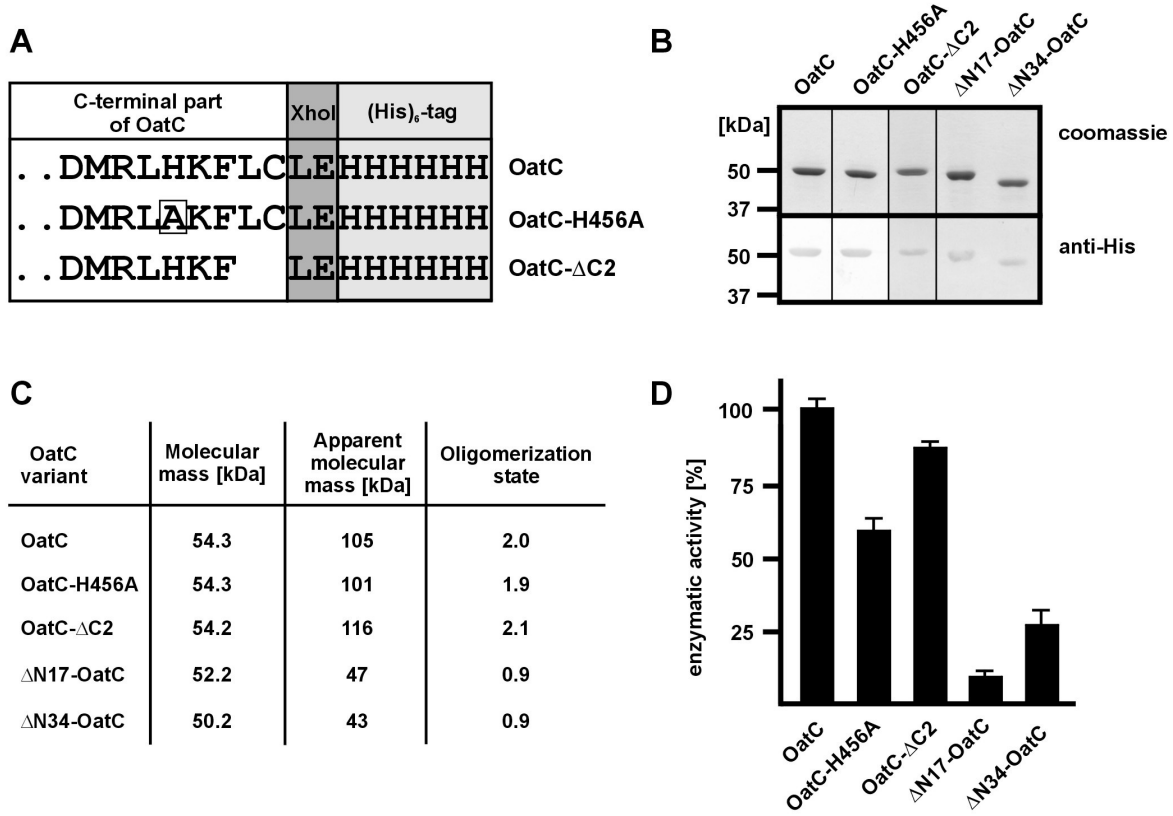


Figure 5-7: Initial hits of EMBL HT-screenings with $(His)_6$-tagged OatC in the presence of a 100-fold molar excess of the donor substrate acetyl-CoA. Purified OatC with a concentration of 8.4 mg/ml was used for initial crystallization trials performed at 19°C. Crystal conditions were assessed at the high-throughput crystallization facility at the EMBL Hamburg Outstation as 400 nL sitting drops (200 nL OatC supplemented with a 100-fold molar excess of acetyl-CoA + 200 nL stock solution) in Greiner Low profile crystallization plates. A total of 864 conditions was screened for crystal growth and pictures taken at day 6 are shown for the 7 conditions that contained clearly visible crystals. Compositions of the used stock solutions are displayed below each photo. (A), initial hits obtained with stock solutions that resemble the conditions that were identified for growth of OatC crystals in the absence of acetyl-CoA; (B) crystals based on small PEGs as crystallants were found in conditions containing CHES-buffer at pH 9.5; (C), crystals based on PEG 8,000 and propylene glycol as precipitants; and (D) crystals obtained in the presence of 3 M ammonium sulfate.

5.3.6 Influences of N-terminal and C-terminal truncations on the structural integrity of OatC

Poor crystal packing can also be due to flexible protein parts that contribute to micro-heterogeneity. Thus, deleting or removing the flexible regions at the termini or within a protein may help to get diffraction-quality crystals (Dale *et al.*, 2003). To analyze whether truncated OatC-variants are better suitable for crystallization, a set of N- and C-terminally truncated enzyme variants was generated and in a first step analyzed with regard to their oligomeric state and enzymatic activity. It has been shown previously, that the enzymatic activity of OatC containing the single amino acid substitution His456Ala was reduced to approximately 70% of the wild-type activity (Bergfeld *et al.*, 2009). As histidine-456 is located within the very C-terminal part of OatC (Fig. 5-8 A), this finding indicated that the C-terminus might be critical for correct protein folding. To further investigate the influence of the C-

terminal part in maintaining the overall structure of OatC, the enzyme was C-terminally truncated by two amino acids resulting in the variant OatC-ΔC2 (Fig. 5-8 A). The truncated OatC-variant was purified to homogeneity as demonstrated by SDS-PAGE (Fig. 5-8 B) and the oligomerization state was determined by size-exclusion chromatography (Fig. 5-8 C). In accordance to wild-type OatC, also OatC-ΔC2 was found as a dimer. Compared to wild-type OatC, the variant truncated by only two amino acids possessed slightly reduced enzymatic activity when analyzed in the spectrophotometric assay system (Fig. 5-8 D). OatC-variants that were C-terminally truncated by six or more amino acids were found inactive when analyzed *in vivo* (MD-thesis F. Spielmann, University of Würzburg, in preparation).



OatC variant	Molecular mass [kDa]	Apparent molecular mass [kDa]	Oligomerization state
OatC	54.3	105	2.0
OatC-H456A	54.3	101	1.9
OatC-ΔC2	54.2	116	2.1
ΔN17-OatC	52.2	47	0.9
ΔN34-OatC	50.2	43	0.9

Figure 5-8: Analyses of truncated OatC-variants. (A), Overview on the different amino acid sequences at the C-termini of wild-type OatC, the single alanine substitution OatC-H456A as well the C-terminally truncated enzyme variant OatC-ΔC2. All variants comprise the amino acids leucine [L] and glutamine [E] as consequence of the translated XhoI cloning site (dark gray) as well the hexahistidine tag (light gray). (B) After expression in *E. coli* BL21(DE3) and affinity purification on Ni^{2+}-chelating columns, OatC variants were analyzed by 10% SDS-PAGE followed by Coomassie staining (upper panel) and Western blotting using anti-His antibody for detection (lower panel). (C), The quaternary structure of purified OatC variants was analyzed by size-exclusion chromatography on a Superdex 200 column, equilibrated with molecular weight standards. The apparent molecular mass and the oligomeric state of each protein variant was calculated from the retention volume. (D), The enzymatic activities of wild-type and mutant forms were determined in a spectrophotometric assay using equimolar concentrations of each purified enzyme variant. Data represent means ± S.D. of three independent experiments, and the value obtained for wild-type OatC was set to 100%.

Taken together, these findings strongly suggest an important role of the C-terminal part in proper folding of OatC and rule out any benefits from using C-terminally truncated OatC variants for crystallization. Concerning the N-terminal part of OatC, strong evidence has been provided, that the first 34 amino acids mediate dimerization of OatC (Bergfeld *et al.*, 2009). In

addition to ΔN34-OatC described previously (see Chapter 4), another N-terminally truncated OatC-variant was generated which lacks only the first 17 amino acids (ΔN17-OatC). Both N-terminally truncated variants were purified to homogeneity as demonstrated by SDS-PAGE (Fig. 5-8 B) and were found as monomers when analyzed by size-exclusion chromatography (Fig. 5-8 C). Comparative analyses of the enzymatic activity revealed only 6% and 25% of the wild-type activity for ΔN17-OatC and ΔN34-OatC, respectively (Fig. 5-8 D). These results demonstrated that the N-terminus of OatC is crucial for both dimerization and enzymatic activity, thus indicating an important role for the structural integrity of OatC. In summary, neither N-terminal nor C-terminal truncations of OatC seem to be applicable to generate alternative enzyme variants for crystallization attempts.

As both protein termini of OatC were found essential for the maintenance of the overall structure, the addition of artificial affinity tags to either end might be critical as well. OatC-variants harboring an N-terminal Strep-tag II and a C-terminal hexahistidine tag exhibited an intermediate oligomerization state of 1.6 (data not shown), whereas OatC containing only a C-terminal hexahistidine tag was found as a dimer (2.0). The OatC-variant harboring the additional N-terminal Strep-tag II also possessed slightly reduced enzymatic activity and the expression level of soluble protein was lower compared to $(His)_6$-tagged OatC (data not shown). Nevertheless, also the C-terminal hexahistidine tag might influence the structural integrity of OatC and moreover, artificial affinity-tags were frequently found to disturb crystallization (Chant *et al.*, 2005; Bucher *et al.*, 2002). Since purification of native OatC lacking convenient affinity-tags requires the establishment and optimization of a new purification procedure, an OatC-variant was generated which harbors a CRV-3C protease cleavage site between OatC and the C-terminal affinity tag (OatC-CRV-His). This allows removal of the hexa-histidine tag after affinity purification although six artificial amino acids, representing remnants of the protease cleavage site, will remain at the C-terminus (Fig. 5-9 A).

5.3.7 Purification of OatC after proteolytic cleavage of the C-terminal $(His)_6$-tag

The enzyme variant OatC-CRV-His was expressed in *E. coli* BL21 gold (DE3) and isolated by immobilized metal affinity chromatography (IMAC) using a Ni^{2+}-chelating column. The OatC-containing fractions were pooled, desalted, and incubated with $(His)_6$-tagged CRV-3C protease to cleave off the C-terminal affinity tag of OatC. After removal of protease and residual uncleaved OatC-CRV-His by IMAC, the obtained untagged OatC was further purified by size-exclusion chromatography and concentrated by ultra-filtration to a final concentration of 9.5 mg/ml. Using this purification scheme, a total amount of 28.5 mg protein was isolated from 2 L of expression culture. Analysis of OatC-CRV-His after IMAC by SDS-PAGE and Coomassie staining revealed a major band consistent with the calculated molecular mass of

54.0 kDa and the identity of this band was confirmed by Western blot analysis using anti-His antibody (Fig. 5-9 B). After protease cleavage, removal of protease, and size-exclusion chromatography, only minor impurities were detectable in the untagged OatC sample.

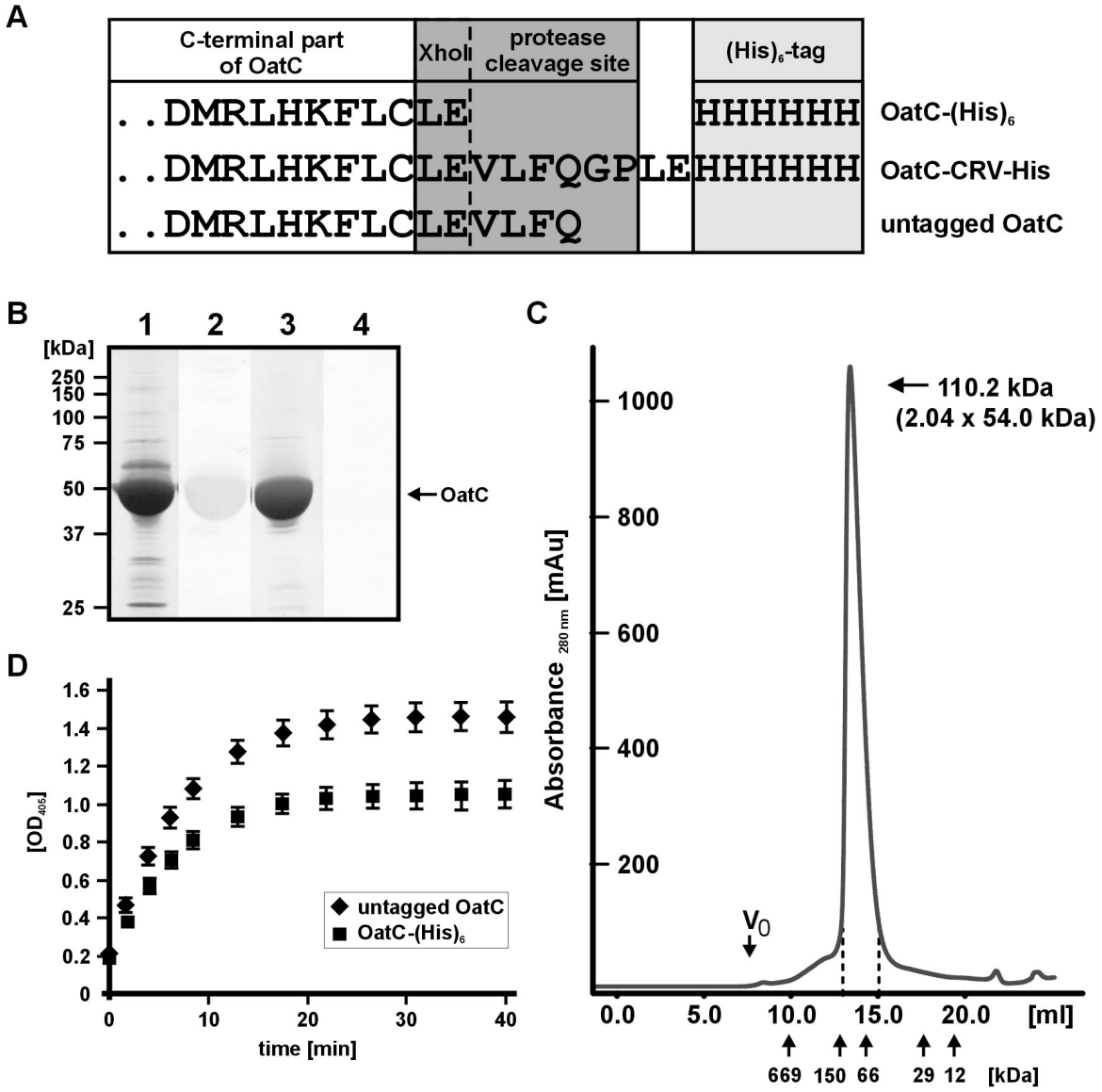


Figure 5-9: Isolation of active untagged OatC for crystallization. (A), Overview of the C-terminal primary sequences of $(His)_6$-tagged OatC, OatC-CRV-His and untagged OatC. A CRV-3C protease cleavage site is present in the variant OatC-CRV-His (dark gray). The first two artificially introduced amino acids (LE) are present in all constructs due to translation of an XhoI site used for cloning and are coincidentally part of the protease cleavage site. The hexahistidine tag is shown in light gray. Proteolytic cleavage of OatC-CRV-His results in the release of untagged OatC. Compared to native OatC, this untagged variant still comprises six additional amino acids at the C-terminus which represent remnants of the CRV-3C protease cleavage site. (B), OatC-CRV-His was expressed in *E. coli* gold BL21(DE3). The enzyme was isolated by IMAC on a Ni^{2+}-chelating column followed by preparative size-exclusion chromatography on a HiLoad™ 16/60 Superdex™ 200 column. Isolation of pure enzyme was confirmed by analysis on 10% SDS-PAGE followed by coomassie staining and the identity of the purified protein as well as the completeness of protease cleavage was verified by Western blotting with anti-His antibody. Shown are pooled OatC-containing fractions eluted form the IMAC column (lane 1 coomassie; lane 2 Western blot) and pooled fractions after protease cleavage followed by size-exclusion chromatography and concentration to 9.5 mg/ml (lane 3 coomassie; lane 4 Western blot). (C), Elution profile of untagged OatC from a Superdex 200 analytical size-exclusion chromatography column. The oligomeric state of purified untagged OatC was analyzed to confirm the dimeric nature of the enzyme which was previously observed for $(His)_6$-tagged OatC. Void-volume (V_0) and elution positions of standard proteins are indicated by arrows. (D), enzymatic activity of purified untagged OatC in comparison to $(His)_6$-tagged OatC was determined in a spectrophotometric assay using equimolar enzyme concentrations (data given as means ± S.D. of three independent experiments).

In the corresponding Western blot analysis with anti-His antibody, neither the OatC band nor any other bands were visible, demonstrating that the $(His)_6$-tag was completely removed from OatC and that the untagged enzyme fraction does not contain detectable amounts of $(His)_6$-tagged CRV-3C protease (Fig. 5-9 B). The oligomerization state of untagged OatC was determined by size-exclusion chromatography and the enzyme was found as a dimer with an apparent molecular mass of 110.2 kDa (Fig. 5-9 C), which is consistent with the molecular mass of 108.0 kDa calculated for an OatC dimer. Enzymatic activity of untagged OatC versus $(His)_6$-tagged OatC was determined in the spectrophotometric assay using equimolar amounts of enzymes. Untagged OatC revealed a significantly higher enzymatic activity compared to the $(His)_6$-tagged variant (Fig. 5-9 D). This finding indicates that the C-terminal affinity tag used so far, affects OatC activity and might therefore also impact the overall folding of the enzyme. Consequently, untagged OatC is most likely more suitable for further crystallization trials.

5.3.8 Crystallization trials with untagged OatC

Initially, a broad crystal screening of isolated untagged OatC was accomplished that covered all positive crystallization conditions which were previously identified for $(His)_6$-tagged OatC. Nonetheless, not a single crystal was obtained, indicating that untagged OatC requires different crystallization conditions. Therefore, untagged OatC was again screened at the high-throughput crystallization facility at the EMBL Hamburg Outstation, as described above for the $(His)_6$-tagged enzyme. Among the 864 conditions tested, most of the conditions tested for untagged OatC led to the formation of precipitate or pre-crystalline compounds, and a total of 11 conditions resulted in crystals. Growth of these crystals started within 6 days and no additional positive conditions were identified within the observed timeframe of 6 month. Remarkably, comparable to the $(His)_6$-tagged enzyme, most crystals of untagged OatC were also based on PEGs as crystallants, whereupon the pH-value was found significantly lower in starter conditions of untagged OatC (Fig. 5-2 A and 5-10 A). Crystals of untagged OatC were also obtained in two conditions containing ethanol as precipitant, but in contrast to $(His)_6$-tagged OatC, these conditions also comprised sodium acetate (Fig. 5-2 B and 5-10 B). Furthermore, one condition containing crystals of untagged OatC was identified which was based on the macromolecular reagent pentaerythritol ethoxylate (15/4 EO/OH) and ammonium sulfate as crystallants (Fig. 5-10 C).

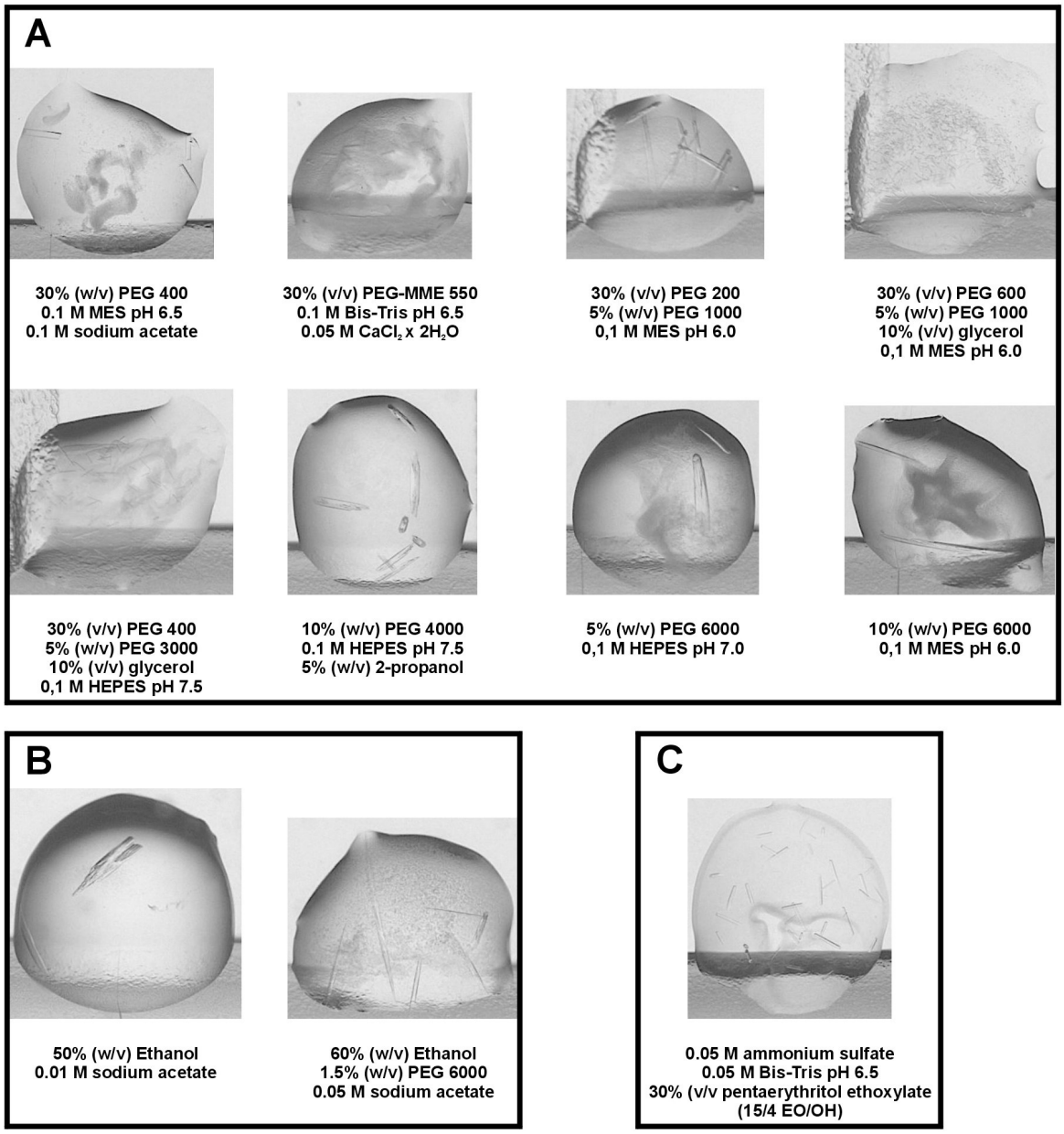


Figure 5-10: Initial hits of HT-screenings with untagged OatC. Purified untagged OatC with a concentration of 9.5 mg/ml was used for initial crystallization trials performed at 19°C. Crystal conditions were assessed at the high-throughput crystallization facility at the EMBL Hamburg Outstation as 400 nL sitting drops (200 nL untagged OatC + 200 nL stock solution) in Greiner Low profile crystallization plates. A total of 864 conditions was screened for crystal growth and pictures taken at day 6 are shown for the 11 conditions that contained clearly visible crystals. Compositions of the used stock solutions are displayed below each photo. (A) 8 positive conditions based on polyethylene glycols as precipitants; (B) 2 positive conditions obtained with ethanol as crystallant and (C) crystals based on ammonium sulfate and pentaerythritol ethoxylate (15/4 EO/OH) as precipitants.

5.3.9 Refinement of initial hit conditions of untagged OatC

Based on the results from the HT-screening, broader screenings were performed in house to analyze the reproducibility of the initial crystallization conditions. Whereas crystal growth obtained in the presence of ethanol (Fig 5-10 B) was not reproducible in first trials, the presence of different PEGs and the combination of pentaerythritol ethoxylate (15/4 EO/OH) and ammonium sulfate as crystallants yielded crystals that were in several cases already of sufficient size for initial X-ray diffraction measurements (Fig. 5-11).

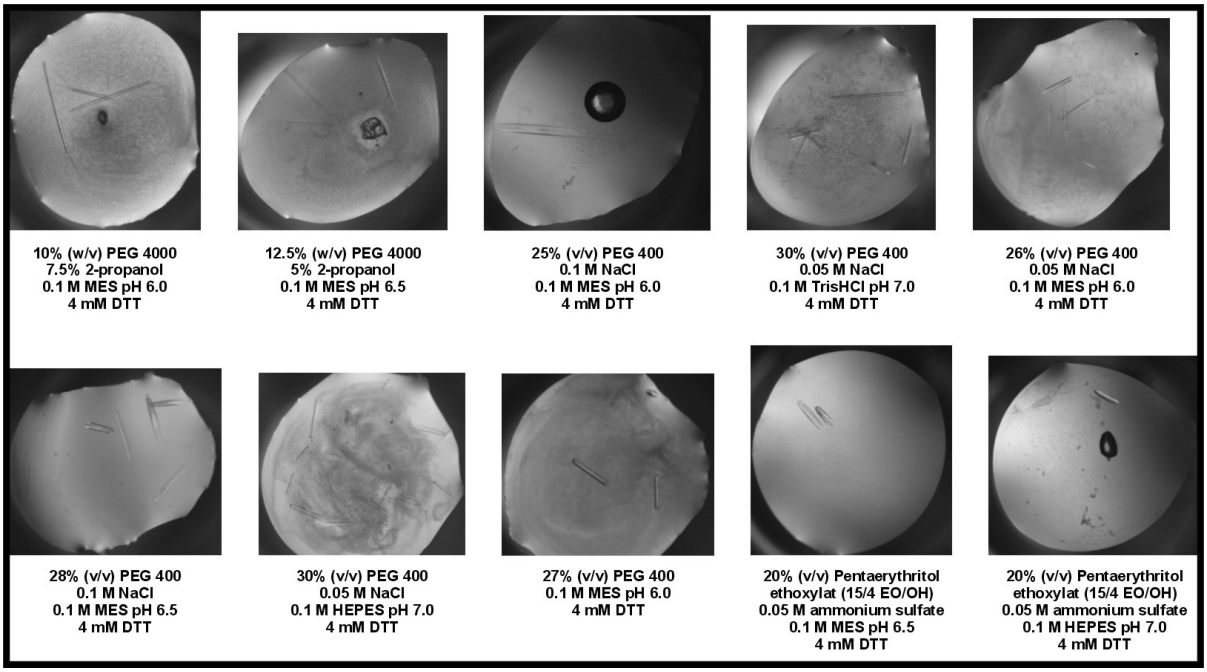


Figure 5-11: Initial refinement of crystallization conditions for untagged OatC. Purified untagged OatC with a concentration of 9.5 mg/ml was used for crystallization trials performed at 20°C. Crystal conditions were assessed in house as 2 µL sitting drops (1 µL OatC + 1 µL stock solution) in 24-well Cryschem plates. Pictures were taken 2 days after the setup. Compositions of the used stock solutions are displayed below each photo.

Although the conditions were not yet refined (e.g. by step-wise varying the concentration of the precipitants, iteratively testing of different buffers at various pH values, by performing additive screens and by refinement of the cryoprotectant condition), the few crystals analyzed so far diffracted already down to a resolution of approximately 4 Å. Additional optimization steps are currently in progress to further improve the quality of crystals of untagged OatC.

5.4 - Discussion

In the present study, protein crystals of the polysialic acid-specific *O*-acetyltransferase OatC were obtained. For the apo-enzyme with a C-terminal hexa-histidine tag, PEG 1,000 based conditions yielded crystals that diffracted X-rays to a resolution of approximately 3.8 Å. Several approaches to achieve well diffracting OatC crystals were investigated and initial results and further steps towards structure determination of OatC will be discussed.

Since diffracting OatC crystals have been obtained, further attempts to improve the diffraction-quality of these crystals will include the screening for optimal cryo-protection conditions as only a small subset of conditions was tested so far (data not shown). Moreover, an extensive screen for additives that enhance the formation of well-ordered crystals might yield crystals of higher quality. Furthermore, the setup of crystallization experiments with freshly purified OatC could be advantageous as the enzyme variants used in the present study might have been slightly distorted as a result from the flash-freezing procedure followed by storage at -80°C and thawing. Although no significant difference between freshly purified OatC and the once frozen enzyme was detectable in the spectrophotometric activity assay or by SDS-PAGE analysis (data not shown). An alternative strategy to improve the quality of OatC crystals might be achieved by seeding procedures as some OatC crystals remained small or formed clusters (e.g. crystals based on ethanol as crystallant). The formation of clusters results from additional nucleation processes that occur on the crystal surface. Seeding is a powerful tool for the separation of nucleation and growth. In this technique, previously nucleated crystals are used as seeds and introduced into new drops equilibrated at lower levels of supersaturation, whereupon this crystal might grow to a size amenable for data collection (Caylor *et al.*, 1999; Bergfors, 2003; Messick and Marmorstein, 2004; Bergfors, 2007).

An important finding of the present study was that the C-terminal hexa-histidine tag used for affinity-isolation of OatC impacts enzyme activity and therefore most likely affects the overall structure of the enzyme. Removal of the affinity-tag from an OatC variant that contained a proteolytic cleavage site in front of the $(His)_6$-tag significantly increased the enzymatic activity compared to the $(His)_6$-tagged enzyme. Moreover, $(His)_6$-tagged and untagged OatC did not crystallize under the same conditions, indicating that the C-terminal $(His)_6$-tag influenced crystal packing. These results implicate that untagged OatC is more suitable for further crystallization trials. Nonetheless, the C-terminus of untagged OatC used in the present study still contains remnants of the proteolytic cleavage site and it remains unknown whether C-terminal addition of this LEVLFQ hexapeptide affects the structural integrity of OatC. Therefore, crystallization of native OatC might provide an alternative strategy to yield high-quality crystals of OatC but so far, no purification protocols have been established for the isolation of OatC lacking an affinity tag.

For decades, crystallographers have made use of the fact that fragments or domains of proteins either crystallize better, or form more well-diffracting crystals, compared to the intact protein (Dong *et al.*, 2007). This most likely originates from the loss of flexible protein regions which might interfere with the assembly of regular crystals (McPherson *et al.*, 1995; Chayen and Saridakis, 2008). Protein fragments lacking flexible parts can be obtained either by protease treatment, by identification of domain boundaries on the basis of sequence alignments followed by recombinant expression of the corresponding protein fragment or by screening a large set of recombinant proteins differing slightly in their N- and C-terminal ends. In the present study, N- and C-terminally truncated OatC variants were generated and the purified enzymes were analyzed with regard to enzymatic activity and oligomerization state. However, already short truncations of either end of the protein impacted the maintenance of the overall structure as indicated by a severe drop in enzymatic activity and/or shift in the oligomeric state from dimer to monomer. Consequently, none of the truncated OatC variants were suitable for crystallization trials.

Recently, *in situ* proteolysis was successfully applied for crystallization of several proteins that otherwise yielded either no or only poorly diffracting crystals (Dong *et al.*, 2001; Mandel *et al.*, 2006; Dong *et al.*, 2007; Xiang *et al.*, 2007). In this method, trace amounts of proteases (such as trypsin or chymotrypsin) are added to the crystallization condition, which results in the removal of flexible loops and/or unstructured N- or C-terminal extensions. A comprehensive approach to evaluate *in situ* proteolysis resulted in crystal growth for 9 out of 20 tested proteins that never crystallized before and the quality of crystals from 6 out of 35 proteins was sufficiently enhanced to solve their 3D-structures (Dong *et al.*, 2007). If flexible internal loops hamper the growth of high quality OatC crystals, *in situ* proteolysis might be the method of choice to discard these loops without affecting the overall protein structure and to improve crystal packing.

For many enzymes, binding of substrates induces conformational changes which can involve the fixation of otherwise flexible loop regions. Thus, the formation of co-crystals of OatC and substrates or substrate analogs provides an alternative strategy to get well-ordered crystals. In the present study, initial co-crystallization trials of OatC with the donor substrate acetyl-CoA were performed. Notably, several of the identified crystallization conditions differed considerably from those found for the apo-enzyme, increasing the possibility that crystals from these conditions are of better quality. Several acetyltransferases of different protein families have already been successfully co-crystallized with acetyl-CoA (Modis and Wierenga, 2000; Peneff *et al.*, 2001; Yan *et al.*, 2002; Wang *et al.*, 2002; Sugantino and Roderick, 2002; Kursula *et al.*, 2002; Kehoe *et al.*, 2003; Han *et al.*, 2006; Hung *et al.*, 2006; Schuetz *et al.*, 2007; Vetting *et al.*, 2008). In the case of OatC, addition of acetyl-CoA results in the formation of a covalent acetyl-enzyme intermediate that involves the active site serine

residue and thus indicates a double-displacement mechanism (Bergfeld *et al.*, 2009). The kinetics as well as the stability of this acetyl-enzyme intermediate are unknown and co-crystallization conditions may lead to mixture of acetylated and non-acetylated forms of OatC and thereby to unfavorable protein heterogeneity. Nevertheless, structures of acetyl-enzyme intermediates of related α/β-hydrolase fold enzymes have already been described (Modis and Wierenga, 2000; Kursula *et al.*, 2002). To increase the homogeneity of the OatC sample, further experiments on growing co-crystals with acetyl-CoA are also planned with the enzyme variant OatC-S286A. As described previously, the single alanine substitution OatC-S286A lacks the acceptor hydroxyl group of the active site serine and was therefore incapable to form the acetyl-enzyme intermediate (Bergfeld *et al.*, 2009). Moreover, the reaction product CoA and the corresponding derivative desulpho-CoA are considered as candidates for additional co-crystallization approaches using wild-type OatC.

As shown in Chapter 4, OatC catalyzes exclusively *O*-acetylation of α2,9-linked polysialic acid and does not act on α2,8-linked polysialic acid, sialic acid containing heteropolymers or monomeric sialic acid. To address the question how OatC gains its high acceptor substrate specificity towards α2,9-linked polysialic acid and to investigate the underlying protein-carbohydrate interactions, co-crystallization of OatC with the acceptor polysaccharide is one of the future goals. However, the capsular polysaccharide of *N. meningitidis* serogroup C (NmC-CPS) is a natural product with a high degree of polymerization with up to 200 sialic acid residues and is therefore comparatively heterogeneous with respect to polymer size. So far, the endosialidase of bacteriophage K1F, which hydrolyses α2,8-linked polysialic acid, is the only polysialic acid binding protein that has been crystallized in the presence of acceptor substrate (Stummeyer *et al.* 2005). In this case, crystals were soaked with pentameric α2,8-linked sialic acid. For α2,9-linked sialic acid, short oligomers of defined length are not commercially available and protocols for partial hydrolysis and purification have to be established. Thermal degradation of NmC-CPS in the presence of 100 mM trifluoric acid followed by separation on a MonoQ anion exchange column resulted in separation of oligomers with a degree of polymerization of 1 to 15 (data not shown). This material will be sufficient to determine the minimal acceptor size required by OatC. However, to obtain high amounts of oligomers of homogenous size as required for co-crystallization trials, the procedure needs further optimization and up-scaling.

Determining the 3D-structures of proteins by X-ray crystallography involves a series of steps: cloning, recombinant expression, isolation of pure enzyme in high yields, crystallization, collection of diffraction data, and determination of atomic positions. During this process, the production of high-quality single crystals remains a major bottleneck (Chayen and Saridakis, 2008). In the present study the first crystallization of a sialate *O*-acetyltransferase was achieved. Intensive refinement of initial crystallization conditions led to reproducible OatC

crystals that diffracted X-rays already down to approximately 3.8 Å. Several strategies were pointed out which are currently accomplished to further improve the quality of OatC crystals and preliminary results were presented. The promising outcome of the actual study is now forming the basis for further optimization trials that will hopefully result in the determination of the first 3D crystal structure of a sialate *O*-acetyltransferase.

Chapter 6 - General Discussion

Surface sialic acids are a common component and virulence factor of a variety of pathogenic bacteria (Vimr and Lichtensteiger, 2002). In many cases, modification by *O*-acetylation has been observed and might play an important role in immunogenicity and survival in the host organism (Bhattacharjee *et al.*, 1976; Ørskov *et al.*, 1979; Gamian *et al.*, 1992; Lemercinier and Jones, 1996; Lewis *et al.*, 2004; Feng *et al.*, 2005; Houliston *et al.*, 2006). The present study describes the first purification and biochemical characterization of NeuO and OatC, the polySia-specific capsule *O*-acetyltransferases of *E. coli* K1 and *N. meningitidis* serogroup C (NmC), respectively. Both enzymes proved highly specific for their endogenous acceptor substrates, α2,8- and α2,9-linked sialic acid, respectively. Free Neu5Ac or CMP-Neu5Ac was not modified and a minimal substrate length of >14 residues was found for NeuO. This indicates that *in vivo*, *O*-acetylation of the polySia capsules of NmC and *E. coli* K1 occurs exclusively at the polymer level. Despite the functional similarity between NeuO and OatC, the data presented in this study point out that the two enzymes evolved independently and belong to structurally diverse protein families (see Figure 6-1).

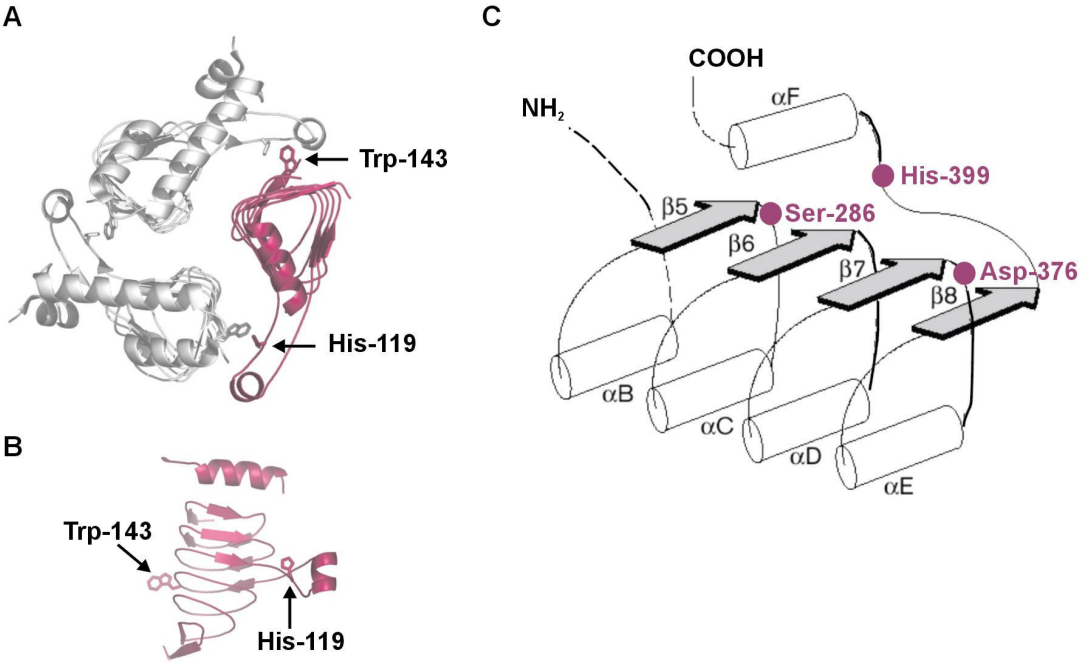


Figure 6-1: The sialate *O*-acetyltransferases NeuO and OatC belong to structurally diverse protein families. *A*, the predicted three-dimensional structure of NeuO was obtained by using the program 3D-PSSM (Kelley *et al.*, 2000) with the structure of the maltose *O*-Acetyltransferase (MAT) of *E. coli* (PDB accession number 1ocx) as a template. The NeuO monomer (*purple*) was modeled into the trimer of MAT (*gray*). *B*, side view on the three-dimensional model of a NeuO monomer. The model highlights the presence of a left-handed β-helix fold for NeuO with one protruding loop carrying a catalytic histidine and a conserved tryptophan is located at the opposite site of the β-helix. The position of the two active site residues His-119 and Trp-143 of NeuO are indicated by *arrows*. All structural data were visualized by PyMOL. *C*, Predicted topology of OatC based on secondary structure prediction (Combet *et al.*, 2000). For amino acids 259-422 the typical topology of a 'canonical' α/β-hydrolase fold could be assigned. The position of the catalytic residues Ser-286, Asp-376 and His-399 (*purple*) after β-strands 5, 7 and 8, respectively, also matched the commonly found localization of a catalytic triad which is the hallmark of α/β-hydrolase fold enzymes.

6.1. The sialate *O*-acetyltransferase NeuO is a member of the left-handed β-helix family of acyltransferases

In the first part of this work, the prophage encoded *O*-acetyltransferase NeuO, which is essential for *O*-acetylation of the capsular polysaccharide of *E. coli* K1 was purified to homogeneity using a C-terminal hexahistidine tag. Previous attempts to isolate endogenous NeuO required detergent extraction and resulted in only partially purified enzyme as all further isolation steps abolished activity (Higa and Varki, 1988). By contrast, recombinant NeuO was obtained from the soluble protein fraction, and the isolated enzyme was enzymatically active as demonstrated in a spectrophotometric assay and by transfer of radioactively labeled acetyl groups from acetyl-CoA to α2,8-linked polySia. A possible explanation for the different outcome during purification of endogenous and recombinant NeuO is that, *in vivo*, the *O*-acetyltransferase might be associated with the polySia biosynthesis and export machinery (Vimr and Steenbergen, 2006). In this case, detergent might be required to dissociate NeuO from the membrane-bound protein complex, a step that can be omitted if NeuO is expressed in an *E. coli* strain lacking the K1 capsule gene complex.

Using homology modeling followed by mutational analysis, NeuO was characterized as a typical member of the superfamily of hexapeptide acyltransferases. Enzymes of this family use phosphopantothenyl-based cofactors to transfer acetyl, succinyl, or long chain fatty acyl groups to free hydroxyl or amino groups of a variety of acceptor substrates. The unique feature of this protein family is the presence of tandem-repeated imperfect copies of a hexapeptide repeat sequence with the consensus motif [LIV]-[GAED]-X_2-[STAV]-X (Vaara, 1992; Jenkins and Pickersgill, 2001). Several members of the family have been crystallized, and the solved crystal structures revealed that this characteristic hexapeptide repeat sequence encodes folding of a left-handed parallel β-helix domain (Raetz and Roderick, 1995; Beaman *et al.*, 1997; Beaman *et al.*, 1998; Brown *et al.*, 1999; Olsen and Roderick, 2001; Sulzenbacher *et al.*, 2001; Sugantino and Roderick, 2002; Wang *et al.*, 2002; Lo *et al.*, 2003). In all known cases, hexapeptide acyltransferases assemble into catalytic trimers with three symmetrical active sites, which are located at the interfaces of two subunits formed by a loop protruding out of one LβH-domain embracing the LβH-domain of the adjacent monomer. Based on the structural characteristics, the hexapeptide repeat family is also called LβH family and includes not only acyltransferases but also a zinc-dependent γ-carbonic anhydrase (Kisker *et al.*, 1996).

The homology modeling performed in the present study indicated close structural similarity of NeuO to the maltose and galactoside *O*-acetyltransferases (GAT and MAT, respectively) of *E. coli* (Lo *et al.*, 2003; Wang *et al.*, 2002; see Figure 6-1). Similar to MAT and GAT, the

predicted 3D-model of NeuO was found to contain a short LβH domain with five and one-third triangular coils and one protruding loop. To amino acids, which are critical for activity of MAT and GAT, were found in equivalent positions in the NeuO model: His-119 located in the protruding loop and Trp-143 located within a β-sheet of the LβH domain (see Figure 6-1). Subsequent alanine substitutions completely abolished NeuO activity *in vitro* and *in vivo*, indicating important roles in catalysis. All crystallized acyltransferases of the LβH family possess a catalytic histidine located in a loop embracing the neighboring subunit. The predicted role of this residue is to abstract a proton from the acceptor hydroxyl group, facilitating the attack of the resulting carbonyl by the acyl-donor (Lewendon *et al.*, 1995; Dunn *et al.*, 1996). The second critical amino acid identified in the model of NeuO, Trp-143, is conserved only in a subset of acetyltransferases, including MAT, GAT, xenobiotic acetyltransferases, and the *Rhizobium leguminosarum* nodulation factor NodL (Lewendon *et al.*, 1995; Dunn *et al.*, 1996; Beaman *et al.*, 1998; Lo *et al.*, 2003). In GAT, replacement of the corresponding Trp-139 by phenylalanine abolished the intrinsic fluorescence quench observed on acetyl-CoA binding (Lewendon *et al.*, 1995), and the crystal structure of the binary complex of GAT with acetyl-CoA revealed a direct contact between the indole side chain of Trp-139 and the phosphopantothenyl arm of the cofactor (Wang *et al.*, 2002). In addition to donor binding, Trp-139 may serve to position the catalytic histidine relative to the hydroxyl group of the acceptor and alter its p*Ka* (Wang *et al.*, 2002). In line with this dual role in donor binding and positioning of the catalytic histidine, a complete loss of NeuO activity was observed *in vitro* and *in vivo* when Trp-143 was substituted by alanine. Taken together, these results provide convincing evidence that NeuO consists primarily of an LβH fold with His-119 and Trp-143 as part of the active site.

Size-exclusion chromatography of recombinant NeuO demonstrated that the enzyme assembles into hexamers. So far, the only example of an LβH-acyltransferase with a quaternary structure other than trimeric is the serine *O*-acetyltransferase (SAT) (Johnson *et al.*, 2005). Similar to NeuO, SAT of *E. coli* and *H. influenzae* adopts a hexameric structure, and the solved crystal structures revealed that two SAT trimers assemble into hexamers with the dimer of trimers interface formed by N-terminal α-helices (Pye *et al.*, 2004, Olsen *et al.*, 2004). The observation that a Strep tag II placed at the N-terminus of NeuO was not accessible for purification might indicate that, also in NeuO, the N-terminal domain is involved in dimerization of NeuO trimers.

In addition to *neuO*, four sialate *O*-acetyltransferase genes have been cloned so far: (i) *oatC* responsible for *O*-acetylation of the α2,9-linked polySia capsule of *N. meningitidis* serogroup C (Claus *et al.*, 2004); (ii) *oatWY*, involved in *O*-acetylation of sialic acids within the galactose and glucose containing heteropolymeric polySia capsules of serogroup W-135 and Y

meningococci, respectively (Claus *et al.*, 2004); (iii) *neuD* of group B streptococci required for *O*-acetylation of terminal α2,3-linked sialic acids capping the group B streptococci capsule (Lewis *et al.*, 2006); and (iv) *neuD* of *Campylobacter jejuni*, which was shown to *O*-acetylate terminal α2,8-linked sialic acids of the bacterial lipo-oligosaccharide (Houliston *et al.*, 2006). With the single exception of OatC, all other sialate *O*-acetyltransferases were found to contain hexapeptide repeat sequences, indicating that they are all members of the LβH family. However, NeuD proteins might form a separate branch within the family, because amino acid sequence alignments revealed the lack of a potential catalytic histidine, which might be substituted by a lysine residue (Claus *et al.*, 2004).

Although NeuO shares the typical features of the LβH family, the present study revealed a unique regulatory mechanism based on changes in VNTR-encoded N-terminal protein extensions. The *neuO* gene is characterized by the presence of a variable number of heptanucleotide repeats (VNTRs of the sequence 5'-AAGACTC-3') within the 5'-end of the coding region (Deszo *et al.*, 2005). Because only repeat numbers that are divisible by three allow full-length translation and thereby capsule *O*-acetylation, changes in the overall number of repeats provide a reversible switch between an *O*-acetylated on phase and a non *O*-acetylated off phase. Oscillation between OAc^+ and OAc^- variants might permit adaptation to changes in environmental conditions. However, findings of the present study demonstrated that this is not a simple "all-or-none" mechanism. The expression of *neuO* alleles containing 0, 12, 24, and 36 heptanucleotide repeats showed that VNTRs are translated into stable N-terminal protein extensions that affect enzymatic activity of the translation product. Every three heptanucleotide repeats encode an RLKTQDS heptad. Remarkably, the presence of RLKTQDS heptads was not essential for NeuO activity but the catalytic efficiency (k_{cat}/K_m^{donor}) increased linearly with the number of heptads. These results suggest that the presence of heptad repeats results in gradual changes in enzyme conformation allowing an increase in enzyme-donor substrate affinity and facilitation of the reaction rate.

The N-terminal heptad repeats of NeuO were previously suggested to assemble into a triple coiled-coil (Deszo *et al.*, 2005). In the hexameric NeuO complex, variation in the length of such a coil might gradually change the subunit arrangement and, thereby, the conformation of the active sites formed at the subunit interfaces of each trimer. However, a solved crystal structure is required to unravel the definite arrangement of all six subunits of the NeuO complex and the influence of the number of N-terminal repeats on the overall structure of the enzyme.

The influence of N-terminal protein extension on NeuO activity was also analyzed *in vivo*, although it has to be taken into account that direct analysis of NeuO activity is hampered due to the occurrence of phase variation. However, comparison of 10 different OAc^+ K1 strains

indicated that the highest enzymatic activities were found in strains expressing *neuO* alleles with >21 heptanucleotide repeats (Deszo *et al.*, 2005). Moreover, analysis of the capsular polysaccharide from four OAc^+ K1 strains revealed that the degree of *O*-acetylation varied between 5 and 85 %, whereupon the lowest *O*-acetyl content and lowest activity were found in a strain without detectable phase variation (Ørskov *et al.*, 1979; Higa and Varki, 1988). These studies indicate that modulation of NeuO activity impacts the degree of capsule *O*-acetylation *in vivo*. In addition to allelic variation caused by VNTRs, sequencing of the *neuO* gene in 102 strains of a diverse *E. coli* K1 strain collection revealed the occurrence of five distinct allelic variants due to seven polymorphic sites. Translation of these five allelic variants results in three NeuO enzyme variants with altered amino acid sequences. However, kinetic characterization of the respective isolated NeuO variants performed in the present study, demonstrated that none of the amino acid exchanges significantly affected NeuO activity.

6.2 The capsule *O*-acetyltransferase OatC belongs to the α/β hydrolase fold family

In the second part of this thesis, structure-function relationships of the sialate *O*-acetyltransferase OatC from *Neisseria meningitis* serogroup C (NmC) were investigated. Endogenous OatC was previously shown to transfer acetyl groups from acetyl-CoA to endogenous membrane-bound as well as to exogenous soluble capsular polysaccharide (CPS) of NmC by using serogroup C meningococcal spheroplast membranes as an enzyme source (Vann *et al.*, 1978). However, reports on the isolation and further characterization of the enzyme were missing. For the first time, a successful expression and purification procedure for recombinant OatC was established that allowed a detailed biochemical characterization of the enzyme.

In contrast to all other bacterial sialate *O*-acetyltransferases known so far, the primary sequence of OatC is devoid of hexapeptide repeats, indicating that the enzymes does not belong to the LβH family of acyltransferases. Moreover, BLAST search analyses did not reveal significant sequence similarities between OatC and any other protein (Claus *et al.*, 2004), impeding the identification of phylogenetic relationships. However, by combining bioinformatic approaches (motif scanning and secondary structure prediction), site-directed mutagenesis, and biochemical analyses, several features were identified in the present study that characterize OatC as a new member of the α/β-hydrolase fold family of proteins. The α/β-hydrolase fold is a versatile and widespread protein architecture that is not restricted to acyltransferases but is found in enzymes with diverse functions such as proteases, lipases, esterases, dehalogenases, haloperoxidases, lyases, and epoxide hydrolases (Argiriadi *et al.*,

1999; Holmquist, 2000; Nandhagopal *et al.*, 2001; Mirza *et al.*, 2005; Polgar, 2005; Fujino *et al.*, 2006; Kolkenbrock *et al.*, 2006; Lee *et al.*, 2006; Srivatsan, 2006). The members of this large superfamily usually lack significant sequence similarities but share a common fold that gave the family its name. Various solved crystal structures allowed a detailed insight into the architecture and common features of α/β-hydrolase fold enzymes (Brady *et al.*, 1990; Pathak and Ollis, 1990; Sussman *et al.*, 1991; Verschueren *et al.*, 1993; Hecht *et al.*, 1994; Lawson *et al.*, 1994; Wagner *et al.*, 1996; Silton *et al.*, 1997; Medrano *et al.*, 1998; Argiriadi *et al.*, 1999; Nandhagopal *et al.*, 2001; Kashiwagi *et al.*, 2002; Lo *et al.*, 2003; Campobasso *et al.*, 2004; Mirza *et al.*, 2005; Larsen *et al.*, 2006 ; Byun *et al.*, 2007; Wang *et al.*, 2007). The canonical α/β-hydrolase fold is composed of an eight-stranded mainly parallel β-sheet, flanked on both sides by α-helices. The fold provides a stable scaffold for a highly conserved catalytic triad consisting of a nucleophile (serine, cysteine or aspartate) which is located at the tip of a sharp turn called the "nucleophile elbow", an acidic residue (aspartate or glutamate) and a histidine (Nardini and Dijkstra, 1999; Heikinheimo *et al.*, 1999; Holmquist 2000). However, many variations from the 'canonical' fold have been observed mainly due to the absence of particular strands or helices and/or insertions that can range from a few amino acids to entire domains.

For OatC, a comprehensive site-directed mutagenesis approach led to the identification of a catalytic triad, composed of Ser-286, Asp-376, and His-399, consistent with the conserved order of the catalytic triad residues (nucleophile-acid-histidine) found in α/β-hydrolase fold enzymes (see Figure 6-1). Moreover, Ser-286 was found to be located in a typical nucleophile elbow motif ($GXS^{286}XGG$), indicating that, similar to other α/β-hydrolase fold proteins, the nucleophile is positioned in a tight turn to expose the hydroxyl group of Ser-286 to nucleophilic attack of the incoming substrate. Furthermore, secondary structure prediction (Combet *et al.*, 2000) suggested a pattern of β-strands and α-helices for OatC which matches that of the canonical α/β-hydrolase fold (Heikinheimo *et al.*, 1999; Nardini and Dijkstra 1998; Ollis *et al.*, 1992; Holmquist 2000; see Figure 6-1). In the proposed topology model, the identified catalytic triad residues are located after strands β5, β7, and β8 for Ser-286, Asp-376, and His-399, respectively, which is consistent with the common positioning of the catalytic triad in α/β-hydrolase fold enzymes (see Figure 6-1). Moreover, OatC was found to assemble into dimers, which is in line with the oligomeric state found for a wide variety of α/β-hydrolase fold enzyme (Argiriadi *et al.*, 1999; Born *et al.*, 2000; Campobasso *et al.*, 2004; Mirza *et al.*, 2005; Wahab *et al.*, 2006; Byun *et al.*, 2007; Wang *et al.*, 2007). Nonetheless, a three-dimensional structure is required to unequivocally confirm the presence of a particular fold. In the present study, first successful crystallization attempts were accomplished with OatC and crystals were obtained that diffracted X-rays already down to approximately 3.8 Å.

Almost all enzymes of the α/β-hydrolase fold family follow a double-displacement (ping-pong) kinetic mechanism involving a covalent substrate-enzyme intermediate (Dodson and Wlodawer, 1998; Holmquist 2000; Ishida and Kato, 2003; Campobasso *et al.*, 2004; Nazi and Wright, 2005; Ishida, 2006; Lee *et al.*, 2006; Wahab *et al.*, 2006). In line with this, a stable acetyl-enzyme intermediate was observed for OatC that crucially depended on the presence of Ser-286. In addition, a CoA burst was observed in the absence of polySia (data not shown). This indicates that the first product, CoA, is released before binding of the second substrate, α2,9-linked polySia, occurs which confirms the presence of a ping-pong kinetic mechanism. Based on the biochemical data for wild-type OatC and mutant forms, the following catalytic mechanism was proposed (Chapter 4, Fig. 7): The formation of a salt bridge between the carboxyl group of Asp-376 and the imidazolium group of His-399 allows the N-ε2 of the imidazole ring to abstract a proton from Ser-286. Therewith, the nucleophilicity of the hydroxyl group of Ser-286 increases, facilitating the attack of the carbonyl carbon of the acetyl group of acetyl-CoA. The resulting first tetrahedral transition state subsequently collapses to a covalent acetyl-enzyme intermediate while free CoA is released. In the second half-reaction the proposed salt bridge between the carboxyl group of Asp-376 and the imidazolium group of His-399 allows the imidazole ring to abstract a proton from the C7 (or C8) hydroxyl group of a sialic acid residue within the acceptor NmC-CPS. This enables the hydroxyl group of the sialic acid residue to attack the carbonyl carbon of the acetyl-group in the acetyl-enzyme intermediate resulting in a second tetrahedral transition state. Subsequent decomposition of the transition state leads to release of the second reaction product, i.e. *O*-acetylated NmC-CPS, and thereby returns Ser-286 back to its protonated state which reinitializes OatC for a new catalytic cycle.

It remains to be investigated whether or not OatC transfers acetyl groups on both, the C7 and the C8 hydroxyl group. The presence of di-*O*-acetylated sialic acids (Neu5,7,8Ac_3) in NmC-CPS was demonstrated by ^{13}C NMR spectroscopy (Bhattacharjee *et al.*, 1975), whereas no evidence for di-*O*-acetylated sialic acids was found in another study (Lemercinier and Jones, 1996). For NmC-CPS containing only mono-*O*-acetylated residues, non-enzymatic migration of acetyl groups was observed. In freshly prepared CPS, *O*-acetylation was found predominantly at O8, whereas after storage at room temperature, the majority of acetyl groups were shifted to O7 (Lemercinier and Jones, 1996; Ho *et al.*, 2001; Ravenscroft *et al.*, 1999). Therefore, OatC might be selective for O8, and acetylation of O7 might occur exclusively by non-enzymatic migration.

In line with the proposed critical roles of the catalytic triad residues of OatC, single alanine substitutions of Ser-286, Asp- 376, and His-399 completely abrogated OatC activity *in vitro* and *in vivo*. Replacement of either Ser-286 or His-399 by alanine abolished the ability of OatC to form a covalent acetyl-enzyme intermediate, which is consistent with the proposed

function of His-399 as the crucial general base catalyst and the role of Ser-286 as the acceptor site for the acetyl group. By contrast, single alanine substitution of Asp-376 resulted in an enzyme variant that was still able to produce limited amounts of acetyl-enzyme intermediate, indicating that Asp-376 is dispensable during the first half-reaction but essentially required for catalysis of the second half-reaction. In several α/β-hydrolase fold enzymes, alanine substitution of the catalytic acid was even compatible with considerably high levels of residual activity, indicating a less pronounced role of this catalytic triad residue (Wajant and Pfitzenmaier, 1996; Lee *et al.*, 2006; Weadge and Clarke, 2007). Interestingly, replacement of the catalytic Ser-286 by the alternative nucleophile cysteine did not rescue enzymatic activity of OatC, demonstrating that the enzyme crucially depends on a properly positioned serine residue in the nucleophile elbow. So far, all characterized *O*-acetyltransferases within the α/β-hydrolase fold family comprise a serine residue in the nucleophile elbow motif (Born *et al.*, 2000; Mirza *et al.*, 2005; Nazi and Wright, 2005; Wang *et al.*, 2007), whereas several other α/β-hydrolase fold enzymes catalyzing different reactions accommodate a cysteine residue (Pathak and Ollis, 1990; Sfakianos *et al.*, 2002; Amara and Rehm, 2003; Campobasso *et al.*, 2004; Sandy *et al.*, 2005; Wang *et al.*, 2005). However, corresponding substitutions of the nucleophile (serine by cysteine and *vice versa*) was analyzed for various α/β-hydrolase fold enzymes resulting in a reduction or complete loss of enzymatic activity (Higaki *et al.*, 1989; Leuveling *et al.*, 1994; Witkowski *et al.*, 1994; Li *et al.*, 1996; Sfakianos *et al.*, 2002; Amara and Rehm, 2003; Nazi and Wright, 2005).

The identification of a double-displacement (ping-pong) kinetic mechanism clearly distinguishes OatC from all other bacterial sialate *O*-acetyltransferases known so far which all belong to the hexapeptide repeat family. Compared with α/β-hydrolase fold enzymes, members of the hexapeptide repeat family act via sequential kinetic mechanisms (Wyckoff and Raetz, 1999; Hindson and Shaw, 2003; Olsen *et al.*, 2004; Johnson *et al.*, 2005). Although a highly conserved catalytic histidine residue is required in both protein families as a general base catalyst, this residue is not part of a catalytic triad in the hexapeptide repeat family.

The detailed biochemical analysis of OatC performed in the present study revealed striking differences between OatC and all other known bacterial sialate *O*-acetyltransferases. Together these findings strongly indicate that the enzymatic function of sialate *O*-acetylation evolved independently on two distinct structural frameworks by convergent evolution. The specificity of NeuO and OatC for polySia might have emerged from two distinct ancestral enzymes belonging to the LβH and α/β-hydrolase fold family, respectively, that already used acetyl-CoA as donor substrate but transferred the acetyl group to different polysaccharide acceptors. In meningococci, *O*-acetylation of the capsular polysaccharide is not restricted to serogroups carrying sialic acid containing capsules. Among the clinically important

serogroups, also *N. meningitidis* serogroup A (NmA) is able to modify the capsular polysaccharide composed of repeating units of α1,6-linked *N*-acetyl-*D*-mannosamine-1-phosphate by *O*-acetylation of the *N*-acetylmannosamine at position O3 and O4 (Liu *et al.*, 1971; Gudlavalleti *et al.*, 2004). The *mynC* gene encoding the respective capsule *O*-acetyltransferase is, like *oatC*, located downstream of the polysaccharide polymerase within the capsule gene complex (Swartley *et al.*, 1998; Gudlavalleti *et al.*, 2004). Although the two *O*-acetyltransferases MynC and OatC share no significant similarities on the primary sequence level, both enzymes posses a nucleophile elbow motif with the sequence GSSKGG and might share a common α/β-hydrolase fold.

6.3 Outlook

The identification of a sialate *O*-acetyltransferase that belongs to the α/β-hydrolase fold family of enzymes might have direct implications on the identification of related enzymes in eukaryotes. Whereas the hexapeptide repeat family of acyltransferases is restricted to microorganisms, α/β-hydrolase fold enzymes are present in all kingdoms of life. Interestingly, the α/β-hydrolase fold is also used as the structural basis for viral sialic acid-specific acetylesterases, the receptor-destroying enzymes of influenza C and coronaviruses that use *O*-acetylated sialic acids as host receptor structures (Rosenthal *et al.*, 1998; de Groot, 2006; Zeng *et al.*, 2008). Thus, the α/β-hydrolase fold appears to be a likely candidate for the structural framework of eukaryotic sialate *O*-acetyltransferases, for which the genetic and structural bases are still unknown. In eukaryotes, *O*-acetylation of sialic acids is found on the outer most part of the carbohydrate moiety of membrane-bound and secreted glycoconjugates (Klein and Roussel, 1998). *O*-acetylation of hydroxyl groups at the C9 position of sialic acids is a tissue-specific, tightly regulated modification, which is involved in a variety of physiological and pathological processes, such as cell-cell and host-pathogen interactions, immunology, as well as development (Varki, 1992; Schauer and Kamerling, 1997; Varki, 1997; Klein and Roussel, 1998; Varki and Varki, 2007). However, attempts to clone the respective sialate *O*-acetyltransferase genes, e.g. by expression cloning, failed, and purification and characterization of the enzymatic activities proved to be challenging (Butor *et al.*, 1993; Shen *et al.*, 2002; Iwersen *et al.*, 2003; Shen *et al.*, 2004; Lrhorfi *et al.*, 2007).

By searching for genes encoding potential α/β-hydrolase fold proteins with unknown function, *CASD1* was identified as a candidate gene encoding a putative human sialate *O*-acetyltransferase. *CASD1* represents a homologue of the *cas1* gene of the fungus *Cryptococcus neoformans*. The Cas1 enzyme was previously reported to be an *O*-acetyltransferase involved in *O*-acetylation of mannose residues in the cryptococcal capsule composed of glucuronoxylomannan (Janbon *et al.*, 2001; Kozel *et al.*, 2003). The human *CASD1* gene encodes a protein of 797 amino acids which appears to be composed of two distinct parts: an N-terminal catalytic domain for which homology modeling predicted an α/β-hydrolase fold with a Ser-Asp-His catalytic triad and a C-terminal part with 12 predicted transmembrane domains (data not shown). CASD1 is currently investigated with regard to sialate *O*-acetyltransferase activity.

Chapter 7 - References

Allison,G.E., Verma,N.K. (2000) Serotype-converting bacteriophages and *O*-antigen modification in *Shigella flexneri Trends Microbiol.* **8**: 17-23.

Amara,A.A., Rehm,B.H. (2003) Replacement of the catalytic nucleophile cysteine-296 by serine in class II polyhydroxyalkanoate synthase from *Pseudomonas aeruginosa*-mediated synthesis of a new polyester: identification of catalytic residues *Biochem.J.* **374**: 413-421.

Angata,T., Varki,A. (2002) Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective *Chem.Rev.* **102**: 439-469.

Argiriadi,M.A., Morisseau,C., Hammock,B.D., and Christianson,D.W. (1999) Detoxification of environmental mutagens and carcinogens: structure, mechanism, and evolution of liver epoxide hydrolase *Proc.Natl.Acad.Sci.U.S.A* **96**: 10637-10642.

Barrett,B., Ebah,L., and Roberts,I.S. (2002) Genomic structure of capsular determinants *Curr.Top.Microbiol.Immunol.* **264**: 137-155.

Beaman,T.W., Binder,D.A., Blanchard,J.S., and Roderick,S.L. (1997) Three-dimensional structure of tetrahydrodipicolinate N-succinyltransferase *Biochemistry* **36**: 489-494.

Beaman,T.W., Sugantino,M., and Roderick,S.L. (1998) Structure of the hexapeptide xenobiotic acetyltransferase from *Pseudomonas aeruginosa Biochemistry* **37**: 6689-6696.

Bentley,S.D., Vernikos,G.S., Snyder,L.A., Churcher,C., Arrowsmith,C., Chillingworth,T., Cronin,A., Davis,P.H., Holroyd,N.E., Jagels,K., Maddison,M., Moule,S., Rabbinowitsch,E., Sharp,S., Unwin,L., Whitehead,S., Quail,M.A., Achtman,M., Barrell,B., Saunders,N.J., and Parkhill,J. (2007) Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18 *PLoS.Genet.* **3**: e23.

Bergfeld,A.K., Claus,H., Lorenzen,N.K., Spielmann,F., Vogel,U., and Mühlenhoff,M. (2009) The polysialic acid-specific *O*-acetyltransferase OatC from *Neisseria meningitidis* serogroup C evolved apart from other bacterial sialate *O*-acetyltransferases *J.Biol.Chem.* **284**: 6-16.

Bergfeld,A.K., Claus,H., Vogel,U., and Mühlenhoff,M. (2007) Biochemical characterization of the polysialic acid-specific *O*-acetyltransferase NeuO of *Escherichia coli* K1 *J.Biol.Chem.* **282**: 22217-22227.

Bergfors,T. (2003) Seeds to crystals *J.Struct.Biol.* **142**: 66-76.

Bergfors,T. (2007) Screening and optimization methods for nonautomated crystallization laboratories *Methods Mol.Biol.* **363**: 131-151.

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 *J.Biol.Chem.* **250**: 1926-1932.

Bhattacharjee,A.K., Jennings,H.J., Kenny,C.P., Martin,A., and Smith,I.C. (1976) Structural determination of the polysaccharide antigens of *Neisseria meningitidis* serogroups Y, W-135, and BO1 *Can.J.Biochem.* **54**: 1-8.

Bliss,J.M., Silver,R.P. (1996) Coating the surface: a model for expression of capsular polysialic acid in *Escherichia coli* K1 *Mol.Microbiol.* **21**: 221-231.

Boisier,P., Nicolas,P., Djibo,S., Taha,M.K., Jeanne,I., Mainassara,H.B., Tenebray,B., Kairo,K.K., Giorgini,D., and Chanteau,S. (2007) Meningococcal meningitis: unprecedented incidence of serogroup X-related cases in 2006 in Niger *Clin.Infect.Dis.* **44**: 657-663.

Bonacorsi,S., Clermont,O., Houdouin,V., Cordevant,C., Brahimi,N., Marecat,A., Tinsley,C., Nassif,X., Lange,M., and Bingen,E. (2003) Molecular analysis and experimental virulence of French and North American Escherichia coli neonatal meningitis isolates: identification of a new virulent clone *J.Infect.Dis.* **187**: 1895-1906.

Born,T.L., Franklin,M., and Blanchard,J.S. (2000) Enzyme-catalyzed acylation of homoserine: mechanistic characterization of the *Haemophilus influenzae* met2-encoded homoserine transacetylase *Biochemistry* **39**: 8556-8564.

Borrow,R., Longworth,E., Gray,S.J., and Kaczmarski,E.B. (2000) Prevalence of de-*O*-acetylated serogroup C meningococci before the introduction of meningococcal serogroup C conjugate vaccines in the United Kingdom *FEMS Immunol.Med.Microbiol.* **28**: 189-191.

Bortolussi,R., Ferrieri,P., Bjorksten,B., and Quie,P.G. (1979) Capsular K1 polysaccharide of *Escherichia coli*: relationship to virulence in newborn rats and resistance to phagocytosis *Infect.Immun.* **25**: 293-298.

Brady,L., Brzozowski,A.M., Derewenda,Z.S., Dodson,E., Dodson,G., Tolley,S., Turkenburg,J.P., Christiansen,L., Huge-Jensen,B., Norskov,L., and . (1990) A serine protease triad forms the catalytic centre of a triacylglycerol lipase *Nature* **343**: 767-770.

Brandtzaeg,P., Van Deuren,M. (2002) Current concepts in the role of the host response in *Neisseria meningitidis* septic shock *Curr.Opin.Infect.Dis.* **15**: 247-252.

Brown,K., Pompeo,F., Dixon,S., Mengin-Lecreulx,D., Cambillau,C., and Bourne,Y. (1999) Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily *EMBO J.* **18**: 4096-4107.

Bucher,M.H., Evdokimov,A.G., and Waugh,D.S. (2002) Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein *Acta Crystallogr.D.Biol.Crystallogr.* **58**: 392-397.

Butor,C., Diaz,S., and Varki,A. (1993) High level *O*-acetylation of sialic acids on N-linked oligosaccharides of rat liver membranes. Differential subcellular distribution of 7- and 9-*O*-acetyl groups and of enzymes involved in their regulation *J.Biol.Chem.* **268**: 10197-10206.

Byun,J.S., Rhee,J.K., Kim,N.D., Yoon,J., Kim,D.U., Koh,E., Oh,J.W., and Cho,H.S. (2007) Crystal structure of hyperthermophilic esterase EstE1 and the relationship between its dimerization and thermostability properties *BMC.Struct.Biol.* **7**: 47.

Campobasso,N., Patel,M., Wilding,I.E., Kallender,H., Rosenberg,M., and Gwynn,M.N. (2004) *Staphylococcus aureus* 3-hydroxy-3-methylglutaryl-CoA synthase: crystal structure and mechanism *J.Biol.Chem.* **279**: 44883-44888.

Cartwright,K.A., Ala'Aldeen,D.A. (1997) *Neisseria meningitidis*: clinical aspects *J.Infect.* **34**: 15-19.

Caugant,D.A. (2008) Genetics and evolution of *Neisseria meningitidis*: importance for the epidemiology of meningococcal disease *Infect.Genet.Evol.* **8**: 558-565.

Caugant,D.A., Hoiby,E.A., Magnus,P., Scheel,O., Hoel,T., Bjune,G., Wedege,E., Eng,J., and Froholm,L.O. (1994) Asymptomatic carriage of *Neisseria meningitidis* in a randomly sampled population *J.Clin.Microbiol.* **32**: 323-330.

Caylor,C.L., Dobrianov,I., Lemay,S.G., Kimmer,C., Kriminski,S., Finkelstein,K.D., Zipfel,W., Webb,W.W., Thomas,B.R., Chernov,A.A., and Thorne,R.E. (1999) Macromolecular impurities and disorder in protein crystals *Proteins* **36**: 270-281.

Chant,A., Kraemer-Pecore,C.M., Watkin,R., and Kneale,G.G. (2005) Attachment of a histidine tag to the minimal zinc finger protein of the *Aspergillus nidulans* gene regulatory protein AreA causes a conformational change at the DNA-binding site *Protein Expr.Purif.* **39**: 152-159.

Chayen,N.E., Saridakis,E. (2008) Protein crystallization: from purified protein to diffraction-quality crystal *Nat.Methods* **5**: 147-153.

Claus,H., Borrow,R., Achtman,M., Morelli,G., Kantelberg,C., Longworth,E., Frosch,M., and Vogel,U. (2004) Genetics of capsule *O*-acetylation in serogroup C, W-135 and Y meningococci *Mol.Microbiol.* **51**: 227-239.

Claus,H., Maiden,M.C., Wilson,D.J., McCarthy,N.D., Jolley,K.A., Urwin,R., Hessler,F., Frosch,M., and Vogel,U. (2005) Genetic analysis of meningococci carried by children and young adults *J.Infect.Dis.* **191**: 1263-1271.

Combet,C., Blanchet,C., Geourjon,C., and Deleage,G. (2000) NPS@: network protein sequence analysis *Trends Biochem.Sci.* **25**: 147-150.

Connolly,M., Noah,N. (1999) Is group C meningococcal disease increasing in Europe? A report of surveillance of meningococcal infection in Europe 1993-6. European Meningitis Surveillance Group *Epidemiol.Infect.* **122**: 41-49.

Cross,A.S., Kim,K.S., Wright,D.C., Sadoff,J.C., and Gemski,P. (1986) Role of lipopolysaccharide and capsule in the serum resistance of bacteremic strains of *Escherichia coli J.Infect.Dis.* **154**: 497-503.

Cudney,R., Patel,S., Weisgraber,K., Newhouse,Y., and McPherson,A. (1994) Screening and optimization strategies for macromolecular crystal growth *Acta Crystallogr.D.Biol.Crystallogr.* **50**: 414-423.

Curreli,S., Arany,Z., Gerardy-Schahn,R., Mann,D., and Stamatos,N.M. (2007) Polysialylated neuropilin-2 is expressed on the surface of human dendritic cells and modulates dendritic cell-T lymphocyte interactions *J.Biol.Chem.* **282**: 30346-30356.

Curtis,M.A., Slaney,J.M., and Aduse-Opoku,J. (2005) Critical pathways in microbial virulence *J.Clin.Periodontol.* **32 Suppl 6**: 28-38.

Dale,G.E., Oefner,C., and D'Arcy,A. (2003) The protein as a variable in protein crystallization *J.Struct.Biol.* **142**: 88-97.

Davidsen,T., Tonjum,T. (2006) Meningococcal genome dynamics *Nat.Rev.Microbiol.* **4**: 11-22.

de Groot,R.J. (2006) Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses *Glycoconj.J.* **23**: 59-72.

Deghmane,A.E., Giorgini,D., Larribe,M., Alonso,J.M., and Taha,M.K. (2002) Down-regulation of pili and capsule of *Neisseria meningitidis* upon contact with epithelial cells is mediated by CrgA regulatory protein *Mol.Microbiol.* **43**: 1555-1564.

Deghmane,A.E., Larribe,M., Giorgini,D., Sabino,D., and Taha,M.K. (2003) Differential expression of genes that harbor a common regulatory element in *Neisseria meningitidis* upon contact with target cells *Infect.Immun.* **71**: 2897-2901.

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.

Djibo,S., Nicolas,P., Alonso,J.M., Djibo,A., Couret,D., Riou,J.Y., and Chippaux,J.P. (2003) Outbreaks of serogroup X meningococcal meningitis in Niger 1995-2000 *Trop.Med.Int.Health* **8**: 1118-1123.

Dobrindt,U. (2005) (Patho-)Genomics of *Escherichia coli Int.J.Med.Microbiol.* **295**: 357-371.

Dodson,G., Wlodawer,A. (1998) Catalytic triads and their relatives *Trends Biochem.Sci.* **23**: 347-352.

Dong,A., Xu,X., Edwards,A.M., Chang,C., Chruszcz,M., Cuff,M., Cymborowski,M., Di Leo,R., Egorova,O., Evdokimova,E., Filippova,E., Gu,J., Guthrie,J., Ignatchenko,A., Joachimiak,A., Klostermann,N., Kim,Y., Korniyenko,Y., Minor,W., Que,Q., Savchenko,A., Skarina,T., Tan,K., Yakunin,A., Yee,A., Yim,V., Zhang,R., Zheng,H., Akutsu,M., Arrowsmith,C., Avvakumov,G.V., Bochkarev,A., Dahlgren,L.G., Dhe-Paganon,S., Dimov,S., Dombrovski,L., Finerty,P., Jr., Flodin,S., Flores,A., Graslund,S., Hammerstrom,M., Herman,M.D., Hong,B.S., Hui,R., Johansson,I., Liu,Y., Nilsson,M., Nedyalkova,L., Nordlund,P., Nyman,T., Min,J., Ouyang,H., Park,H.W., Qi,C., Rabeh,W., Shen,L., Shen,Y., Sukumard,D., Tempel,W., Tong,Y., Tresagues,L., Vedadi,M., Walker,J.R., Weigelt,J., Welin,M., Wu,H., Xiao,T., Zeng,H., and Zhu,H. (2007) In situ proteolysis for protein crystallization and structure determination *Nat.Methods* **4**: 1019-1021.

Dong,A., Yoder,J.A., Zhang,X., Zhou,L., Bestor,T.H., and Cheng,X. (2001) Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA *Nucleic Acids Res.* **29**: 439-448.

Dunn,S.M., Moody,P.C., Downie,J.A., and Shaw,W.V. (1996) Crystallization and preliminary diffraction studies of NodL, a rhizobial *O*-acetyl-transferase involved in the host-specific nodulation of legume roots *Protein Sci.* **5**: 538-541.

Feng,L., Sechenkova,S.N., Tao,J., Shashkov,A.S., Liu,B., Shevelev,S.D., Reeves,P.R., Xu,J., Knirel,Y.A., and Wang.L. (2005) Structural and genetic characterization of enterohemorrhagic *Escherichia coli* O145 O antigen and development of an O145 serogroup-specific PCR assay *J. Bacteriol.* **187**: 758-764.

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**: 482-487.

Frasa,H., Procee,J., Torensma,R., Verbruggen,A., Algra,A., Rozenberg-Arska,M., Kraaijeveld,K., and Verhoef,J. (1993) *Escherichia coli* in bacteremia: *O*-acetylated K1 strains appear to be more virulent than non-*O*-acetylated K1 strains *J.Clin.Microbiol.* **31**: 3174-3178.

Freitag,T., Squires,R.A., Schmid,J., and Elliott,J. (2005) Feline uropathogenic *Escherichia coli* from Great Britain and New Zealand have dissimilar virulence factor genotypes *Vet.Microbiol.* **106**: 79-86.

Frosch,M., Weisgerber,C., and Meyer,T.F. (1989) Molecular characterization and expression in *Escherichia coli* of the gene complex encoding the polysaccharide capsule of *Neisseria meningitidis* group B *Proc.Natl.Acad.Sci.U.S.A* **86**: 1669-1673.
Fujino,S., Akiyama,D., Akaboshi,S., Fujita,T., Watanabe,Y., and Tamai,Y. (2006) Purification and characterization of phospholipase B from *Candida utilis Biosci.Biotechnol.Biochem.* **70**: 377-386.

Fusco,P.C., Farley,E.K., Huang,C.H., Moore,S., and Michon,F. (2007) Protective meningococcal capsular polysaccharide epitopes and the role of O acetylation *Clin.Vaccine Immunol.* **14**: 577-584.

Galanakis,E., Di Cello,F., Paul-Satyaseela,M., and Kim,K.S. (2006) *Escherichia coli* K1 induces IL-8 expression in human brain microvascular endothelial cells *Eur.Cytokine Netw.* **17**: 260-265.

Gamian A., Romanowska E., Ulrich J., and Defaye J. (1992) The structure of the sialic acid-containing *Escherichia coli* O104 O-specific polysaccharide and its linkage to the core region in lipopolysaccharide *Carbohydr Res.* **236**: 195-208.

Gascon,E., Vutskits,L., and Kiss,J.Z. (2007) Polysialic acid-neural cell adhesion molecule in brain plasticity: from synapses to integration of new neurons *Brain Res.Rev.* **56**: 101-118.

Gotschlich,E.C., Liu,T.Y., and Artenstein,M.S. (1969) Human immunity to the meningococcus. 3. Preparation and immunochemical properties of the group A, group B, and group C meningococcal polysaccharides *J.Exp.Med.* **129**: 1349-1365.

Greenfield,S., Sheehe,P.R., and Feldman,H.A. (1971) Meningococcal carriage in a population of "normal" families *J.Infect.Dis.* **123**: 67-73.

Gudlavalleti,S.K., Datta,A.K., Tzeng,Y.L., Noble,C., Carlson,R.W., and Stephens,D.S. (2004) The *Neisseria meningitidis* serogroup A capsular polysaccharide O-3 and O-4 acetyltransferase *J.Biol.Chem.* **279**: 42765-42773.

Gudlavalleti,S.K., Lee,C.H., Norris,S.E., Paul-Satyaseela,M., Vann,W.F., and Frasch,C.E. (2007) Comparison of *Neisseria meningitidis* serogroup W135 polysaccharide-tetanus toxoid conjugate vaccines made by periodate activation of *O*-acetylated, non-*O*-acetylated and chemically de-*O*-acetylated polysaccharide *Vaccine* **25**: 7972-7980.

Gudlavalleti,S.K., Szymanski,C.M., Jarrell,H.C., and Stephens,D.S. (2006) In vivo determination of *Neisseria meningitidis* serogroup A capsular polysaccharide by whole cell high-resolution magic angle spinning NMR spectroscopy *Carbohydr.Res.* **341**: 557-562.

Guerry,P., Szymanski,C.M. (2008) Campylobacter sugars sticking out *Trends Microbiol.* **16**: 428-435.

Hammerschmidt,S., Hilse,R., van Putten,J.P., Gerardy-Schahn,R., Unkmeir,A., and Frosch,M. (1996a) Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element *EMBO J.* **15**: 192-198.

Hammerschmidt,S., Müller,A., Sillmann,H., Mühlenhoff,M., Borrow,R., Fox,A., van Putten,J., Zollinger,W.D., Gerardy-Schahn,R., and Frosch,M. (1996b) Capsule phase variation in *Neisseria meningitidis* serogroup B by slipped-strand mispairing in the polysialyltransferase gene (siaD): correlation with bacterial invasion and the outbreak of meningococcal disease *Mol.Microbiol.* **20**: 1211-1220.

Han,B.W., Bingman,C.A., Wesenberg,G.E., and Phillips,G.N., Jr. (2006) Crystal structure of *Homo sapiens* thialysine Nepsilon-acetyltransferase (HsSSAT2) in complex with acetyl coenzyme A *Proteins* **64**: 288-293.

Hecht,H.J., Sobek,H., Haag,T., Pfeifer,O., and van Pee,K.H. (1994) The metal-ion-free oxidoreductase from *Streptomyces aureofaciens* has an alpha/beta hydrolase fold *Nat.Struct.Biol.* **1**: 532-537.

Heikinheimo,P., Goldman,A., Jeffries,C., and Ollis,D.L. (1999) Of barn owls and bankers: a lush variety of alpha/beta hydrolases *Structure.* **7**: R141-R146.

Henderson,I.R., Owen,P., and Nataro,J.P. (1999) Molecular switches--the ON and OFF of bacterial phase variation *Mol.Microbiol.* **33**: 919-932.

Higa,H.H., 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.

Higaki,J.N., Evnin,L.B., and Craik,C.S. (1989) Introduction of a cysteine protease active site into trypsin *Biochemistry* **28**: 9256-9263.

Hildebrandt,H., Mühlenhoff,M., and Gerardy-Schahn,R. (2008) Polysialylation of NCAM *Neurochem.Res.*

Hildebrandt,H., Mühlenhoff,M., Weinhold,B., and Gerardy-Schahn,R. (2007) Dissecting polysialic acid and NCAM functions in brain development *J.Neurochem.* **103 Suppl 1**: 56-64.

Hindson,V.J., Shaw,W.V. (2003) Random-order ternary complex reaction mechanism of serine acetyltransferase from *Escherichia coli Biochemistry* **42**: 3113-3119.
Ho,M.M., Lemercinier,X., Bolgiano,B., Crane,D., and Corbel,M.J. (2001) Solution stability studies of the subunit components of meningococcal C oligosaccharide-CRM197 conjugate vaccines *Biotechnol.Appl.Biochem.* **33**: 91-98.

Hoffman,J.A., Wass,C., Stins,M.F., and Kim,K.S. (1999) The capsule supports survival but not traversal of *Escherichia coli* K1 across the blood-brain barrier *Infect.Immun.* **67**: 3566-3570.

Holmquist,M. (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms *Curr.Protein Pept.Sci.* **1**: 209-235.

Hotopp,J.C., Grifantini,R., Kumar,N., Tzeng,Y.L., Fouts,D., Frigimelica,E., Draghi,M., Giuliani,M.M., Rappuoli,R., Stephens,D.S., Grandi,G., and Tettelin,H. (2006) Comparative genomics of *Neisseria meningitidis*: core genome, islands of horizontal transfer and pathogen-specific genes *Microbiology* **152**: 3733-3749.

Houliston,R.S., Endtz,H.P., Yuki,N., Li,J., Jarrell,H.C., Koga,M., van Belkum,A., Karwaski,M.F., Wakarchuk,W.W., and Gilbert,M. (2006) Identification of a sialate *O*-acetyltransferase from *Campylobacter jejuni*: demonstration of direct transfer to the C-9 position of terminalalpha-2, 8-linked sialic acid *J.Biol.Chem.* **281**: 11480-11486.

Hung,M.N., Rangarajan,E., Munger,C., Nadeau,G., Sulea,T., and Matte,A. (2006) Crystal structure of TDP-fucosamine acetyltransferase (WecD) from *Escherichia coli*, an enzyme required for enterobacterial common antigen synthesis *J.Bacteriol.* **188**: 5606-5617.

Ishida,T. (2006) Low-barrier hydrogen bond hypothesis in the catalytic triad residue of serine proteases: correlation between structural rearrangement and chemical shifts in the acylation process *Biochemistry* **45**: 5413-5420.

Ishida,T., Kato,S. (2003) Theoretical perspectives on the reaction mechanism of serine proteases: the reaction free energy profiles of the acylation process *J.Am.Chem.Soc.* **125**: 12035-12048.

Iwersen,M., Dora,H., Kohla,G., Gasa,S., and Schauer,R. (2003) Solubilisation and properties of the sialate-4-*O*-acetyltransferase from guinea pig liver *Biol.Chem.* **384**: 1035-1047.

Janbon,G., Himmelreich,U., Moyrand,F., Improvisi,L., and Dromer,F. (2001) Cas1p is a membrane protein necessary for the *O*-acetylation of the *Cryptococcus neoformans* capsular polysaccharide *Mol.Microbiol.* **42**: 453-467.

Jann,K., Jann,B. (1992) Capsules of *Escherichia coli*, expression and biological significance *Can.J.Microbiol.* **38**: 705-710.

Jarvis,G.A., Vedros,N.A. (1987) Sialic acid of group B *Neisseria meningitidis* regulates alternative complement pathway activation *Infect.Immun.* **55**: 174-180.

Jenkins,J., Pickersgill,R. (2001) The architecture of parallel beta-helices and related folds *Prog.Biophys.Mol.Biol.* **77**: 111-175.

Jennings,H.J., Bhattacharjee,A.K., Bundle,D.R., Kenny,C.P., Martin,A., and Smith,I.C. (1977) Strucutres of the capsular polysaccharides of *Neisseria meningitidis* as determined by 13C-nuclear magnetic resonance spectroscopy *J.Infect.Dis.* **136 Suppl**: S78-S83.

Jodar,L., Feavers,I.M., Salisbury,D., and Granoff,D.M. (2002) Development of vaccines against meningococcal disease *Lancet* **359**: 1499-1508.

Johnson,C.M., Roderick,S.L., and Cook,P.F. (2005) The serine acetyltransferase reaction: acetyl transfer from an acylpantothenyl donor to an alcohol *Arch.Biochem.Biophys.* **433**: 85-95.

Johnson,J.R., Russo,T.A. (2002) Extraintestinal pathogenic *Escherichia coli*: "the other bad *E. coli*" *J.Lab Clin.Med.* **139**: 155-162.

Johnson,J.R., Russo,T.A. (2005) Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli Int.J.Med.Microbiol.* **295**: 383-404.

Kahler,C.M., Blum,E., Miller,Y.K., Ryan,D., Popovic,T., and Stephens,D.S. (2001) exl, an exchangeable genetic island in *Neisseria meningitidis Infect.Immun.* **69**: 1687-1696.

Kehoe,L.E., Snidwongse,J., Courvalin,P., Rafferty,J.B., and Murray,I.A. (2003) Structural basis of Synercid (quinupristin-dalfopristin) resistance in Gram-positive bacterial pathogens *J.Biol.Chem.* **278**: 29963-29970.

Kelley,L.A., MacCallum,R.M., and Sternberg,M.J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.,* **299,** 499-520.

Kim,B.Y., Kang,J., and Kim,K.S. (2005) Invasion processes of pathogenic *Escherichia coli Int.J.Med.Microbiol.* **295**: 463-470.

Kim,K.S. (2001) *Escherichia coli* translocation at the blood-brain barrier *Infect.Immun.* **69**: 5217-5222.

Kim,K.S. (2002) Strategy of *Escherichia coli* for crossing the blood-brain barrier *J.Infect.Dis.* **186 Suppl 2**: S220-S224.

Kim,K.S. (2003) Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury *Nat.Rev.Neurosci.* **4**: 376-385.

Kim,K.S., Itabashi,H., Gemski,P., Sadoff,J., Warren,R.L., and Cross,A.S. (1992) The K1 capsule is the critical determinant in the development of *Escherichia coli* meningitis in the rat *J.Clin.Invest* **90**: 897-905.

Kim,Y.V., Pearce,D., and Kim,K.S. (2008) Ca(2+)/calmodulin-dependent invasion of microvascular endothelial cells of human brain by *Escherichia coli* K1 *Cell Tissue Res.* **332**: 427-433.

King,M.R., Steenbergen,S.M., and Vimr,E.R. (2007) Going for baroque at the *Escherichia coli* K1 cell surface *Trends Microbiol.* **15**: 196-202.

King,M.R., Vimr,R.P., Steenbergen,S.M., Spanjaard,L., Plunkett,G., III, Blattner,F.R., and Vimr,E.R. (2007) *Escherichia coli* K1-specific bacteriophage CUS-3 distribution and function in phase-variable capsular polysialic acid O acetylation *J.Bacteriol.* **189**: 6447-6456.

Klein,A., and Roussel,P. (1998) *O*-Acetylation of sialic acids Biochimie **80**: 49-57.

Kolkenbrock,S., Parschat,K., Beermann,B., Hinz,H.J., and Fetzner,S. (2006) N-acetylanthranilate amidase from *Arthrobacter nitroguajacolicus* Ru61a, an alpha/beta-hydrolase-fold protein active towards aryl-acylamides and -esters, and properties of its cysteine-deficient variant *J.Bacteriol.* **188**: 8430-8440.

Kozel,T.R., Levitz,S.M., Dromer,F., Gates,M.A., Thorkildson,P., and Janbon,G. (2003) Antigenic and biological characteristics of mutant strains of *Cryptococcus neoformans* lacking capsular O acetylation or xylosyl side chains *Infect.Immun.* **71**: 2868-2875.

Krinos,C.M., Coyne,M.J., Weinacht,K.G., Tzianabos,A.O., Kasper,D.L., and Comstock,L.E. (2001) Extensive surface diversity of a commensal microorganism by multiple DNA inversions *Nature* **414**: 555-558.

Kursula,P., Ojala,J., Lambeir,A.M., and Wierenga,R.K. (2002) The catalytic cycle of biosynthetic thiolase: a conformational journey of an acetyl group through four binding modes and two oxyanion holes *Biochemistry* **41**: 15543-15556.

Larsen,N.A., Lin,H., Wei,R., Fischbach,M.A., and Walsh,C.T. (2006) Structural characterization of enterobactin hydrolase IroE *Biochemistry* **45**: 10184-10190.

Lawson,D.M., Derewenda,U., Serre,L., Ferri,S., Szittner,R., Wei,Y., Meighen,E.A., and Derewenda,Z.S. (1994) Structure of a myristoyl-ACP-specific thioesterase from *Vibrio harveyi Biochemistry* **33**: 9382-9388.

Lee,L.C., Lee,Y.L., Leu,R.J., and Shaw,J.F. (2006) Functional role of catalytic triad and oxyanion hole-forming residues on enzyme activity of *Escherichia coli* thioesterase I/protease I/phospholipase L1 *Biochem.J.* **397**: 69-76.

Lemercinier,X., Jones,C. (1996) Full 1H NMR assignment and detailed *O*-acetylation patterns of capsular polysaccharides from *Neisseria meningitidis* used in vaccine production *Carbohydr.Res.* **296**: 83-96.

Lewendon,A., Ellis,J., and Shaw,W.V. (1995) Structural and mechanistic studies of galactoside acetyltransferase, the *Escherichia coli* LacA gene product *J.Biol.Chem.* **270**: 26326-26331.

Lewis,A.L., Hensler,M.E., Varki,A., and Nizet,V. (2006) The group B streptococcal sialic acid *O*-acetyltransferase is encoded by neuD, a conserved component of bacterial sialic acid biosynthetic gene clusters *J.Biol.Chem.* **281**: 11186-11192.

Lewis,A.L., Nizet,V., and Varki,A. (2004) Discovery and characterization of sialic acid *O*-acetylation in group B Streptococcus *Proc.Natl.Acad.Sci.U.S.A* **101**: 11123-11128.

Li,J., Szittner,R., Derewenda,Z.S., and Meighen,E.A. (1996) Conversion of serine-114 to cysteine-114 and the role of the active site nucleophile in acyl transfer by myristoyl-ACP thioesterase from *Vibrio harveyi Biochemistry* **35**: 9967-9973.

Liu,T.Y., Gotschlich,E.C., Jonssen,E.K., and Wysocki,J.R. (1971) Studies on the meningococcal polysaccharides. I. Composition and chemical properties of the group A polysaccharide *J.Biol.Chem.* **246**: 2849-2858.

Lo,L.L., Dal Degan,F., Poulsen,P., Andersen,S.M., and Larsen,S. (2003a) The structure and specificity of *Escherichia coli* maltose acetyltransferase give new insight into the LacA family of acyltransferases *Biochemistry* **42**: 5225-5235.

Lo,Y.C., Lin,S.C., Shaw,J.F., and Liaw,Y.C. (2003b) Crystal structure of *Escherichia coli* thioesterase I/protease I/lysophospholipase L1: consensus sequence blocks constitute the catalytic center of SGNH-hydrolases through a conserved hydrogen bond network *J.Mol.Biol.* **330**: 539-551.

Longworth,E., Fernsten,P., Mininni,T.L., Vogel,U., Claus,H., Gray,S., Kaczmarski,E., and Borrow,R. (2002) *O*-Acetylation status of the capsular polysaccharides of serogroup Y and W135 meningococci isolated in the UK *FEMS Immunol.Med.Microbiol.* **32**: 119-123.

Lrhorfi,L.A., Srinivasan,G.V., and Schauer,R. (2007) Properties and partial purification of sialate-*O*-acetyltransferase from bovine submandibular glands *Biol.Chem.* **388**: 297-306.

Lukacova,M., Barak,I., and Kazar,J. (2008) Role of structural variations of polysaccharide antigens in the pathogenicity of Gram-negative bacteria *Clin.Microbiol.Infect.* **14**: 200-206.

Maiden,M.C., Bygraves,J.A., Feil,E., Morelli,G., Russell,J.E., Urwin,R., Zhang,Q., Zhou,J., Zurth,K., Caugant,D.A., Feavers,I.M., Achtman,M., and Spratt,B.G. (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms *Proc.Natl.Acad.Sci.U.S.A* **95**: 3140-3145.

Mandel,C.R., Gebauer,D., Zhang,H., and Tong,L. (2006) A serendipitous discovery that in situ proteolysis is essential for the crystallization of yeast CPSF-100 (Ydh1p) *Acta Crystallogr.Sect.F.Struct.Biol.Cryst.Commun.* **62**: 1041-1045.

McGuire,E.J., Binkley,S.B. (1964) The structure and chemistry of colominic acid *Biochemistry* **3**: 247-251.

McPherson,A., Malkin,A.J., and Kuznetsov,Y.G. (1995) The science of macromolecular crystallization *Structure.* **3**: 759-768.

Medrano,F.J., Alonso,J., Garcia,J.L., Romero,A., Bode,W., and Gomis-Ruth,F.X. (1998) Structure of proline iminopeptidase from *Xanthomonas campestris* pv. citri: a prototype for the prolyl oligopeptidase family *EMBO J.* **17**: 1-9.

Messick,T., Marmorstein,R. (2004) Crystallization of macromolecules *Curr.Protoc.Protein Sci.* **Chapter 17**: Unit 17.4.

Michon,F., Huang,C.H., Farley,E.K., Hronowski,L., Di,J., and Fusco,P.C. (2000) Structure activity studies on group C meningococcal polysaccharide-protein conjugate vaccines: effect of *O*-acetylation on the nature of the protective epitope *Dev.Biol.(Basel)* **103**: 151-160.

Milkowski,C., Strack,D. (2004) Serine carboxypeptidase-like acyltransferases *Phytochemistry* **65**: 517-524.

Mirza,I.A., Nazi,I., Korczynska,M., Wright,G.D., and Berghuis,A.M. (2005) Crystal structure of homoserine transacetylase from *Haemophilus influenzae* reveals a new family of alpha/beta-hydrolases *Biochemistry* **44**: 15768-15773.

Modis,Y., Wierenga,R.K. (2000) Crystallographic analysis of the reaction pathway of *Zoogloea ramigera* biosynthetic thiolase *J.Mol.Biol.* **297**: 1171-1182.

Moxon,E.R., Kroll,J.S. (1990) The role of bacterial polysaccharide capsules as virulence factors *Curr.Top.Microbiol.Immunol.* **150**: 65-85.

Mueller-Dieckmann,J. (2006) The open-access high-throughput crystallization facility at EMBL Hamburg *Acta Crystallogr.D.Biol.Crystallogr.* **62**: 1446-1452.

Mühlenhoff,M., Eckhardt,M., and Gerardy-Schahn,R. (1998) Polysialic acid: three-dimensional structure, biosynthesis and function *Curr.Opin.Struct.Biol.* **8**: 558-564.

Mushtaq,N., Redpath,M.B., Luzio,J.P., and Taylor,P.W. (2004) Prevention and cure of systemic *Escherichia coli* K1 infection by modification of the bacterial phenotype *Antimicrob.Agents Chemother.* **48**: 1503-1508.

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 *J.Antimicrob.Chemother.* **56**: 160-165.

Nandhagopal,N., Yamada,A., Hatta,T., Masai,E., Fukuda,M., Mitsui,Y., and Senda,T. (2001) Crystal structure of 2-hydroxyl-6-oxo-6-phenylhexa-2,4-dienoic acid (HPDA) hydrolase (BphD enzyme) from the *Rhodococcus sp.* strain RHA1 of the PCB degradation pathway *J.Mol.Biol.* **309**: 1139-1151.

Nardini,M., Dijkstra,B.W. (1999) Alpha/beta hydrolase fold enzymes: the family keeps growing *Curr.Opin.Struct.Biol.* **9**: 732-737.

Nazi,I., Wright,G.D. (2005) Catalytic mechanism of fungal homoserine transacetylase *Biochemistry* **44**: 13560-13566.

Ninomiya,T., Sugiura,N., Tawada,A., Sugimoto,K., Watanabe,H., and Kimata,K. (2002) Molecular cloning and characterization of chondroitin polymerase from *Escherichia coli* strain K4 *J.Biol.Chem.* **277**: 21567-21575.

Ollis,D.L., Cheah,E., Cygler,M., Dijkstra,B., Frolow,F., Franken,S.M., Harel,M., Remington,S.J., Silman,I., Schrag,J., and . (1992) The alpha/beta hydrolase fold *Protein Eng* **5**: 197-211.

Olsen,L.R., Huang,B., Vetting,M.W., and Roderick,S.L. (2004) Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor *Biochemistry* **43**: 6013-6019.

Olsen,L.R., Roderick,S.L. (2001) Structure of the *Escherichia coli* GlmU pyrophosphorylase and acetyltransferase active sites *Biochemistry* **40**: 1913-1921.

Ørskov,F. (1978) Virulence factors of the bacterial cell surface *J.Infect.Dis.* **137**: 630-633.

Ørskov,F., Ørskov,I., Sutton,A., Schneerson,R., Lin,W., Egan,W., Hoff,G.E., and Robbins,J.B. (1979) Form variation in *Escherichia coli* K1: determined by *O*-acetylation of the capsular polysaccharide *J.Exp.Med.* **149**: 669-685.

Pace,D., Pollard,A.J. (2007) Meningococcal A, C, Y and W-135 polysaccharide-protein conjugate vaccines *Arch.Dis.Child* **92**: 909-915.

Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G., and Barrell,B.G. (2000) Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491 *Nature* **404**: 502-506.

Pathak,D., Ollis,D. (1990) Refined structure of dienelactone hydrolase at 1.8 A *J.Mol.Biol.* **214**: 497-525.

Pelkonen,S., Hayrinen,J., and Finne,J. (1988) Polyacrylamide gel electrophoresis of the capsular polysaccharides of *Escherichia coli* K1 and other bacteria *J.Bacteriol.* **170**: 2646-2653.

Peneff,C., Mengin-Lecreulx,D., and Bourne,Y. (2001) The crystal structures of Apo and complexed *Saccharomyces cerevisiae* GNA1 shed light on the catalytic mechanism of an amino-sugar N-acetyltransferase *J.Biol.Chem.* **276**: 16328-16334.

Petit,C., Rigg,G.P., Pazzani,C., Smith,A., Sieberth,V., Stevens,M., Boulnois,G., Jann,K., and Roberts,I.S. (1995) Region 2 of the *Escherichia coli* K5 capsule gene cluster encoding proteins for the biosynthesis of the K5 polysaccharide *Mol.Microbiol.* **17**: 611-620.

Polgar,L. (2005) The catalytic triad of serine peptidases *Cell Mol.Life Sci.* **62**: 2161-2172.

Pong,A., Bradley,J.S. (1999) Bacterial meningitis and the newborn infant *Infect.Dis.Clin.North Am.* **13**: 711-33, viii.

Pye,V.E., Tingey,A.P., Robson,R.L., and Moody,P.C. (2004) The structure and mechanism of serine acetyltransferase from *Escherichia coli J.Biol.Chem.* **279**: 40729-40736.

Raetz,C.R., Roderick,S.L. (1995) A left-handed parallel beta helix in the structure of UDP-N-acetylglucosamine acyltransferase *Science* **270**: 997-1000.

Ravenscroft,N., Averani,G., Bartoloni,A., Berti,S., Bigio,M., Carinci,V., Costantino,P., D'Ascenzi,S., Giannozzi,A., Norelli,F., Pennatini,C., Proietti,D., Ceccarini,C., and Cescutti,P. (1999) Size determination of bacterial capsular oligosaccharides used to prepare conjugate vaccines *Vaccine* **17**: 2802-2816.

Raymond,N.J., Reeves,M., Ajello,G., Baughman,W., Gheesling,L.L., Carlone,G.M., Wenger,J.D., and Stephens,D.S. (1997) Molecular epidemiology of sporadic (endemic) serogroup C meningococcal disease *J.Infect.Dis.* **176**: 1277-1284.

Roberts,I.S. (1996) The biochemistry and genetics of capsular polysaccharide production in bacteria *Annu.Rev.Microbiol.* **50**: 285-315.

Rodriguez,M.L., Jann,B., and Jann,K. (1988) Structure and serological characteristics of the capsular K4 antigen of *Escherichia coli* O5:K4:H4, a fructose-containing polysaccharide with a chondroitin backbone *Eur.J.Biochem.* **177**: 117-124.

Rohr,T.E., Troy,F.A. (1980) Structure and biosynthesis of surface polymers containing polysialic acid in *Escherichia coli J.Biol.Chem.* **255**: 2332-2342.

Rosenstein,N.E., Perkins,B.A., Stephens,D.S., Popovic,T., and Hughes,J.M. (2001) Meningococcal disease *N.Engl.J.Med.* **344**: 1378-1388.

Rosenthal,P.B., Zhang,X., Formanowski,F., Fitz,W., Wong,C.H., Meier-Ewert,H., Skehel,J.J., and Wiley,D.C. (1998) Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus *Nature* **396**: 92-96.

Russo,T.A., Johnson,J.R. (2000) Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC *J.Infect.Dis.* **181**: 1753-1754.

Russo,T.A., Johnson,J.R. (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem *Microbes.Infect.* **5**: 449-456.

Rutishauser,U. (2008) Polysialic acid in the plasticity of the developing and adult vertebrate nervous system *Nat.Rev.Neurosci.* **9**: 26-35.

Schauer,R., anf Kamerling,J.R. (1997) Chemistry, biochemistry and biology of sialic acids *Glycoproteins II*, Montreuil,J., Vliegenthart,J.F.G., and Schachter,H., eds. (Amsterdam, The Netherlands: Elsevier), pp. 243-402.

Schauer,R. (2000) Achievements and challenges of sialic acid research *Glycoconj.J.* **17**: 485-499.

Schoen,C., Joseph,B., Claus,H., Vogel,U., and Frosch,M. (2007) Living in a changing environment: insights into host adaptation in *Neisseria meningitidis* from comparative genomics *Int.J.Med.Microbiol.* **297**: 601-613.

Schuetz,A., Bernstein,G., Dong,A., Antoshenko,T., Wu,H., Loppnau,P., Bochkarev,A., and Plotnikov,A.N. (2007) Crystal structure of a binary complex between human GCN5 histone acetyltransferase domain and acetyl coenzyme A *Proteins* **68**: 403-407.

Sfakianos,M.K., Wilson,L., Sakalian,M., Falany,C.N., and Barnes,S. (2002) Conserved residues in the putative catalytic triad of human bile acid Coenzyme A:amino acid N-acyltransferase *J.Biol.Chem.* **277**: 47270-47275.

Shen,Y., Tiralongo,J., Iwersen,M., Sipos,B., Kalthoff,H., and Schauer,R. (2002) Characterization of the sialate-7(9)-*O*-acetyltransferase from the microsomes of human colonic mucosa *Biol.Chem.* **383**: 307-317.

Shen,Y., Tiralongo,J., Kohla,G., and Schauer,R. (2004) Regulation of sialic acid *O*-acetylation in human colon mucosa *Biol.Chem.* **385**: 145-152.

Silver,R.P., Finn,C.W., Vann,W.F., Aaronson,W., Schneerson,R., Kretschmer,P.J., and Garon,C.F. (1981) Molecular cloning of the K1 capsular polysaccharide genes of E. coli *Nature* **289**: 696-698.

Smith,J.L., Fratamico,P.M., and Gunther,N.W. (2007) Extraintestinal pathogenic *Escherichia coli* *Foodborne.Pathog.Dis.* **4**: 134-163.

Spinosa,M.R., Progida,C., Tala,A., Cogli,L., Alifano,P., and Bucci,C. (2007) The *Neisseria meningitidis* capsule is important for intracellular survival in human cells *Infect.Immun.* **75**: 3594-3603.

Srivatsan,M. (2006) An analysis of acetylcholinesterase sequence for predicting mechanisms of its non-catalytic actions *Bioinformation.* **1**: 281-284.

Stein,D.M., Robbins,J., Miller,M.A., Lin,F.Y., and Schneerson,R. (2006) Are antibodies to the capsular polysaccharide of *Neisseria meningitidis* group B and Escherichia coli K1 associated with immunopathology? *Vaccine* **24**: 221-228.

Stephens,D.S. (1999) Uncloaking the meningococcus: dynamics of carriage and disease *Lancet* **353**: 941-942.

Stephens,D.S. (2007) Conquering the meningococcus *FEMS Microbiol.Rev.* **31**: 3-14.

Stephens,D.S., Farley,M.M. (1991) Pathogenic events during infection of the human nasopharynx with *Neisseria meningitidis* and *Haemophilus influenzae Rev.Infect.Dis.* **13**: 22-33.

Stummeyer,K., Schwarzer,D., Claus,H., Vogel,U., Gerardy-Schahn,R., and Mühlenhoff,M. (2006) Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1-specific phages *Mol.Microbiol.* **60**: 1123-1135.

Sugantino,M., Roderick,S.L. (2002) Crystal structure of Vat(D): an acetyltransferase that inactivates streptogramin group A antibiotics *Biochemistry* **41**: 2209-2216.

Sulzenbacher,G., Gal,L., Peneff,C., Fassy,F., and Bourne,Y. (2001) Crystal structure of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridyltransferase bound to acetyl-coenzyme A reveals a novel active site architecture *J.Biol.Chem.* **276**: 11844-11851.

Sussman,J.L., Harel,M., Frolow,F., Oefner,C., Goldman,A., Toker,L., and Silman,I. (1991) Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein *Science* **253**: 872-879.

Swartley,J.S., Liu,L.J., Miller,Y.K., Martin,L.E., Edupuganti,S., and Stephens,D.S. (1998) Characterization of the gene cassette required for biosynthesis of the (alpha1-->6)-linked N-acetyl-D-mannosamine-1-phosphate capsule of serogroup A *Neisseria meningitidis J.Bacteriol.* **180**: 1533-1539.

Swartley,J.S., Marfin,A.A., Edupuganti,S., Liu,L.J., Cieslak,P., Perkins,B., Wenger,J.D., and Stephens,D.S. (1997) Capsule switching of *Neisseria meningitidis Proc.Natl.Acad.Sci.U.S.A* **94**: 271-276.

Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R., and Venter,J.C. (2000) Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58 *Science* **287**: 1809-1815.

Troy,F.A. (1992) Polysialylation: from bacteria to brains *Glycobiology* **2**: 5-23.

Tzeng,Y.L., Stephens,D.S. (2000) Epidemiology and pathogenesis of *Neisseria meningitidis Microbes.Infect.* **2**: 687-700.

Vaara,M. (1992) Eight bacterial proteins, including UDP-N-acetylglucosamine acyltransferase (LpxA) and three other transferases of *Escherichia coli*, consist of a six-residue periodicity theme *FEMS Microbiol.Lett.* **76**: 249-254.

van Belkum,A., Scherer,S., van Alphen,L., and Verbrugh,H. (1998) Short-sequence DNA repeats in prokaryotic genomes *Microbiol.Mol.Biol.Rev.* **62**: 275-293.

Van Deuren,M., Brandtzaeg,P., and van der Meer,J.W. (2000) Update on meningococcal disease with emphasis on pathogenesis and clinical management *Clin.Microbiol.Rev.* **13**: 144-66.

Vann,W.F., Liu,T.Y., and Robbins,J.B. (1978) Cell-free biosynthesis of the *O*-acetylated N-acetylneuraminic acid capsular polysaccharide of group C meningococci *J.Bacteriol.* **133**: 1300-1306.

Vann,W.F., Schmidt,M.A., Jann,B., and Jann,K. (1981) The structure of the capsular polysaccharide (K5 antigen) of urinary-tract-infective *Escherichia coli* 010:K5:H4. A polymer similar to desulfo-heparin *Eur.J.Biochem.* **116**: 359-364.

Varki,A. (1992) Diversity in the sialic acids *Glycobiology* **2**: 25-40.

Varki,A. (1997) Sialic acids as ligands in recognition phenomena *FASEB J.* **4**:248-55.

Varki,N.M., Varki,A. (2007) Diversity in cell surface sialic acid presentations: implications for biology and disease *Lab Invest* **87**: 851-857.

Verschueren,K.H., Franken,S.M., Rozeboom,H.J., Kalk,K.H., and Dijkstra,B.W. (1993) Refined X-ray structures of haloalkane dehalogenase at pH 6.2 and pH 8.2 and implications for the reaction mechanism *J.Mol.Biol.* **232**: 856-872.

Vetting,M.W., Bareich,D.C., Yu,M., and Blanchard,J.S. (2008) Crystal structure of RimI from *Salmonella typhimurium* LT2, the GNAT responsible for N(alpha)-acetylation of ribosomal protein S18 *Protein Sci.* **17**: 1781-1790.

Villwock,A., Schmitt,C., Frosch,M., and Kurzai,O. (2008) Capsule acetylation does not impair recognition of serogroup C, W-135 and Y meningococci by human dendritic cells *Int.J.Med.Microbiol.* **298**: 591-597.

Vimr,E.R., Steenbergen,S.M. (2006) Mobile contingency locus controlling *Escherichia coli* K1 polysialic acid capsule acetylation *Mol.Microbiol.* **60**: 828-837.

Vimr,E.R., Steenbergen,S.M. (2009) Early molecular-recognition events in the synthesis and export of group 2 capsular polysaccharides *Microbiology* **155**: 9-15.

Vogel,U., Claus,H., and Frosch,M. (2000) Rapid serogroup switching in *Neisseria meningitidis N.Engl.J.Med.* **342**: 219-220.

Vogel,U., Claus,H., and Frosch,M. (2004) Genetic lineages and their traits in *Neisseria meningitidis Int.J.Med.Microbiol.* **294**: 75-82.

Vogel,U., Hammerschmidt,S., and Frosch,M. (1996) Sialic acids of both the capsule and the sialylated lipooligosaccharide of *Neisseria meningitis* serogroup B are prerequisites for virulence of meningococci in the infant rat *Med.Microbiol.Immunol.* **185**: 81-87.

Wagner,U.G., Hasslacher,M., Griengl,H., Schwab,H., and Kratky,C. (1996) Mechanism of cyanogenesis: the crystal structure of hydroxynitrile lyase from *Hevea brasiliensis Structure.* **4**: 811-822.

Wahab,H.A., Ahmad Khairudin,N.B., Samian,M.R., and Najimudin,N. (2006) Sequence analysis and structure prediction of type II *Pseudomonas sp*. USM 4-55 PHA synthase and an insight into its catalytic mechanism *BMC.Struct.Biol.* **6**: 23.

Wajant,H., Pfizenmaier,K. (1996) Identification of potential active-site residues in the hydroxynitrile lyase from *Manihot esculenta* by site-directed mutagenesis *J.Biol.Chem.* **271**: 25830-25834.

Wang,H., Liu,L., Hanna,P.E., and Wagner,C.R. (2005) Catalytic mechanism of hamster arylamine N-acetyltransferase 2 *Biochemistry* **44**: 11295-11306.

Wang,M., Liu,L., Wang,Y., Wei,Z., Zhang,P., Li,Y., Jiang,X., Xu,H., and Gong,W. (2007) Crystal structure of homoserine *O*-acetyltransferase from *Leptospira interrogans Biochem.Biophys.Res.Commun.* **363**: 1050-1056.

Wang,X.G., Olsen,L.R., and Roderick,S.L. (2002) Structure of the lac operon galactoside acetyltransferase *Structure.* **10**: 581-588.

Weadge,J.T., Clarke,A.J. (2007) *Neisseria gonorrheae O*-acetylpeptidoglycan esterase, a serine esterase with a Ser-His-Asp catalytic triad *Biochemistry* **46**: 4932-4941.

Weinhold,B., Seidenfaden,R., Röckle,I., Mühlenhoff,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**: 42971-42977.

Whitfield,C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli Annu.Rev.Biochem.* **75**: 39-68.

Whitfield,C., Roberts,I.S. (1999) Structure, assembly and regulation of expression of capsules in *Escherichia coli Mol.Microbiol.* **31**: 1307-1319.

Witkowski,A., Witkowska,H.E., and Smith,S. (1994) Reengineering the specificity of a serine active-site enzyme. Two active-site mutations convert a hydrolase to a transferase *J.Biol.Chem.* **269**: 379-383.

Wyckoff,T.J., Raetz,C.R. (1999) The active site of *Escherichia coli* UDP-N-acetylglucosamine acyltransferase. Chemical modification and site-directed mutagenesis *J.Biol.Chem.* **274**: 27047-27055.

Xiang,S., Usunow,G., Lange,G., Busch,M., and Tong,L. (2007) Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis *J.Biol.Chem.* **282**: 2676-2682.

Xie,Y., Kim,K.J., and Kim,K.S. (2004) Current concepts on *Escherichia coli* K1 translocation of the blood-brain barrier *FEMS Immunol.Med.Microbiol.* **42**: 271-279.

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 *J.Biol.Chem.* **278**: 13875-13880.

Yan,Y., Harper,S., Speicher,D.W., and Marmorstein,R. (2002) The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate *Nat.Struct.Biol.* **9**: 862-869.

Yazdankhah,S.P., Caugant,D.A. (2004) *Neisseria meningitidis*: an overview of the carriage state *J.Med.Microbiol.* **53**: 821-832.
Yazdankhah,S.P., Kriz,P., Tzanakaki,G., Kremastinou,J., Kalmusova,J., Musilek,M., Alvestad,T., Jolley,K.A., Wilson,D.J., McCarthy,N.D., Caugant,D.A., and Maiden,M.C. (2004) Distribution of serogroups and genotypes among disease-associated and carried isolates of *Neisseria meningitidis* from the Czech Republic, Greece, and Norway *J.Clin.Microbiol.* **42**: 5146-5153.

Zelmer,A., Bowen,M., Jokilammi,A., Finne,J., Luzio,J.P., and Taylor,P.W. (2008) Differential expression of the polysialyl capsule during blood-to-brain transit of neuropathogenic *Escherichia coli* K1 *Microbiology* **154**: 2522-2532.

Zeng,Q., Langereis,M.A., van Vliet,A.L., Huizinga,E.G., and de Groot,R.J. (2008) Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution *Proc.Natl.Acad.Sci.U.S.A* **105**: 9065-9069.

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 *J.Biol.Chem.* **267**: 9965-9971.

Appendix 1 - Abbreviations

anti-His	monoclonal antibody directed against the hexahistidine tag
BBB	blood-brain-barrier
BSA	bovine serum albumin
CoA	coenzyme A
CPS	capsular polysaccharide
DP	degree of polymerization
E. coli	*Escherichia coli*
ExPEC	extraintestinal *E. coli*
$(His)_6$	hexahistidine tag
IMAC	immobilized metal affinity chromatography
kps	capsule gene complex
LPS	lipo-polysaccharide
mAb	monoclonal antibody
MBP	maltose-binding protein
MLST	multilocus sequence typing
MynC	capsule-specific *O*-acetyltransferase of *N. meningitidis* serogroup A
NCAM	neural cell adhesion molecule
n.d.	not determined
Neu5Ac	5-N-acetylneuraminic acid
NeuO	sialate *O*-acetyltransferase of *E. coli* K1
N. meningitidis	*Neisseria meningitidis*
NmB	*N. meningitidis* serogroup B
NmC	*N. meningitidis* serogroup C
NmW	*N. meningitidis* serogroup W-135
NmY	*N. meningitidis* serogroup Y
OAc	*O*-acetylation
OatC	sialate *O*-acetyltransferase of *N. meningitidis* serogroup C
OatWY	sialate *O*-acetyltransferase of *N. meningitidis* serogroup W-135 and Y
PMSF	phenylmethylsulfonyl fluoride
polySia	polysialic acid
Sia	sialic acid
Strep	Strep tag II
VNTR	variable number of tandem repeats
wt	wild-type

Appendix 2 – Curriculum Vitae and Publications

Personal data

Home address	Anne Katrin Bergfeld Sedanstr. 6 30161 Hannover Germany
E-mail address	Anne.Bergfeld@gmx.de
Date and place of birth	16. December 1980 in Bergisch Gladbach, Germany
Nationality	German

Scientific education

Doctoral thesis	December 2005 - June 2009 Topic: "Structure-function relationships of bacterial sialate *O*-acetyltransferases"; Supervisor: Prof. Dr. Rita Gerardy-Schahn, Institute for Cellular Chemistry, Hannover Medical School
Diploma	October 2005; Degree: Diplom-Biochemikerin (Dipl.-Biochem.)
Diploma thesis	April 2005 - October 2005 Topic: "Biochemical characterization of bacterial *O*-acetyltransferases"; Supervisor: Prof. Dr. Rita Gerardy-Schahn, Institute for Cellular Chemistry, Hannover Medical School
University	2000 - 2005 Studies in biochemistry (Diplom) at the *Gottfried Wilhelm Leibniz Universität Hannover*
Graduation	June 2000: Abitur (High School Diploma) at the *Landrat-Lucas-Schule* (Gymnasium) in Leverkusen as university entrance qualification

Peer Reviewed Publications

- Ines L. Mordhorst, Heike Claus, Christa Ewers, Martin Lappann , Christoph Schoen, Johannes Elias, Julia Batzilla, Ulrich Dobrindt, Lothar H. Wieler, **Anne K. Bergfeld**, Martina Mühlenhoff and Ulrich Vogel (2009) *O*-acetyltransferase gene *neuO* is segregated according to phylogenetic background and contributes to environmental desiccation resistance in *Escherichia coli* K1. *Environ. Microbiol.* in press, doi:10.1111/j.1462-2920.2009.02019.x

- **Anne K. Bergfeld**, Heike Claus, Nina K. Lorenzen, Fabian Spielmann, Ulrich Vogel, and Martina Mühlenhoff (2009) The polysialic acid-specific *O*-acetyltransferase OatC from *Neisseria meningitidis* serogroup C evolved apart from other bacterial sialate *O*-acetyltransferases. *J. Biol. Chem.* **284**(1):6-16

- **Anne K. Bergfeld**, Heike Claus, Ulrich Vogel, and Martina Mühlenhoff (2007) Biochemical characterization of the polysialic acid-specific *O*-acetyltransferase NeuO of *Escherichia coli* K1. *J. Biol. Chem.* **282**(30):22217-27

- Anne-Christin Lamerz, Thomas Haselhorst, **Anne K. Bergfeld**, Mark von Itzstein, Rita Gerardy-Schahn (2006) Molecular cloning of the *Leishmania major* UDP-glucose pyrophosphorylase, functional characterization, and ligand binding analyses using NMR spectroscopy. *J. Biol. Chem.* **281**(24):16314-22

Oral Presentations

* presenting author marked with an asterisk

- **Anne K. Bergfeld***, Heike Claus, Nina K. Lorenzen, Fabian Spielmann, Ulrich Vogel and Martina Mühlenhoff: “The polysialic acid specific *O*-acetyltransferase OatC from *Neisseria meningitidis* serogroup C evolved apart from other bacterial sialate:*O*-acetyltransferases”; December 2nd 2008; at the 19th Joint Meeting of the *Studiengruppe Glykobiologie der GBM*, the *Netherlands Society for Glycobiology*, the *Groupe Lillois de Glycobiologie*, and the *Belgian Working Group for Glycosciences*; November 30th - December 2nd 2008; Wageningen, The Netherlands

- **Anne K. Bergfeld***, Heike Claus, Nina K. Lorenzen, Fabian Spielmann, Ulrich Vogel and Martina Mühlenhoff: “The polysialic acid specific *O*-acetyltransferase of *Neisseria meningitidis* serogroup C - an alpha/beta hydrolase fold enzyme”; July 23rd 2008; at the International Conference on Biology and Chemistry of Sialic Acids, SialoGlyco2008; July 21st - 26th 2008; Moscow - St Petersburg, Russia

- Ulrich Vogel*, Heike Claus, **Anne K. Bergfeld**, Katharina Stummeyer, Ines Mordhorst, and Martina Mühlenhoff: “Genetics and population biology of the PolySia-Synthesis in the bacterial pathogens *Neisseria meningitidis* and *Escherichia coli*”; March 11th 2007; at the 3rd International meeting: Polysialic acid - Chemistry, Biology, Translational Aspects; March 10th - 13th 2007; Bad Lauterberg, Germany

- **Anne K. Bergfeld***, Heike Claus, Ulrich Vogel and Martina Mühlenhoff: “Purification and characterization of the polysialic acid specific *O*-acetyltransferase of *Escherichia coli* K1”; March 11th 2007; at the 3rd International meeting: Polysialic acid - Chemistry, Biology, Translational Aspects; March 10th - 13th 2007; Bad Lauterberg, Germany

- Martina Mühlenhoff*, Heike Claus, **Anne K. Bergfeld** and Ulrich Vogel: “Molecular characterization of the polysialic acid specific *O*-acetyltransferase of *Escherichia coli* K1”; September 7th 2005; at the XVIII. International Symposium on Glycoconjugates; September 4th - 9th 2005; Florence, Italy

Poster Presentations

* presenting author marked with an asterisk

- Myriam Elschami*, **Anne K. Bergfeld**, Nina K. Lorenzen, Heike Claus, Ulrich Vogel and Martina Mühlenhoff: "Characterization of the polysialic acid specific *O*-acetyltransferase of *Neisseria meningitidis* serogroup W-135 and Y"; 19th Joint Meeting of the *Studiengruppe Glykobiologie der GBM*, the *Netherlands Society for Glycobiology*, the *Groupe Lillois de Glycobiologie*, and the *Belgian Working Group for Glycosciences*; November 30th - December 2nd 2008; Wageningen, the Netherlands

- **Anne K. Bergfeld**, Heike Claus, Ulrich Vogel and Martina Mühlenhoff*: "Characterization of the capsular polysaccharide modifying *O*-acetyltransferase of *Escherichia coli* K1"; June 11th - 14th 2007; Benzon Symposium No. 54 Glycosylation, opportunities in drug development; Copenhagen, Denmark

- **Anne K. Bergfeld***, Heike Claus, Ulrich Vogel and Martina Mühlenhoff: „Biochemical characterization of the polysialic acid specific *O*-acetyltransferase of *Escherichia coli* K1"; November 5th – 7th 2006; 17th Joint Meeting of the *Studiengruppe Glykobiologie der GBM*, the *Nederlandse Vereniging voor Glycobiologie*, the *Groupe Lillois de Glycobiologie*, and the *Belgian Working Group for Glycosciences*; Brugge, Belgium

- **Anne K. Bergfeld***, Heike Claus, Ulrich Vogel and Martina Mühlenhoff: „Biochemical characterization of the capsular polysaccharide modifying *O*-acetyltransferase of *Escherichia coli* K1"; October 1st – 4th 2006; 58. Tagung der Deutschen Gesellschaft für Hygiene und Mikrobiologie e.V.; Würzburg, Germany

- **Anne K. Bergfeld***, Heike Claus, Ulrich Vogel and Martina Mühlenhoff: „Characterization of the polysialic acid specific *O*-acetyltransferase of *E. coli* K1"; March 29th 2006; Workshop Biomaterials - Synthesis, Processing and Biological Evaluation; Hannover, Germany

Appendix 3 - Danksagung

Zuerst möchte ich mich bei Dr. Martina Mühlenhoff für die Bereitstellung des sehr interessanten Themas und die tolle Betreuung während dieser Arbeit bedanken. Deine vielen guten Ideen und Denkanstöße, unsere Diskussionen und Deine freundliche und positive Art haben dazu geführt, dass ich immer voll motiviert und mit viel Spaß gearbeitet habe! Ganz herzlichen Dank für die produktive und schöne Zeit!

Bei Prof. Dr. Rita Gerardy-Schahn möchte ich mich ganz herzlich bedanken, dass ich im Institut für Zelluläre Chemie meine Doktorarbeit anfertigen durfte und im Verlaufe dieser Arbeit die Gelegenheit hatte, an einer Vielzahl von spannenden internationalen Tagungen teilzunehmen. Außerdem danke ich Prof. Gerardy-Schahn für die Übernahme des Referates und Prof. Dr. Walter Müller für die Übernahme des Korreferates.

Ich möchte mich auch bei unseren Kooperationspartnern Prof. Dr. Ulrich Vogel, Dr. Heike Claus und Ines Mordhorst von der Universität Würzburg für die sehr gute Zusammenarbeit auf diesem Projekt bedanken. Ein ganz herzliches Dankeschön gilt auch Dr. Katharina Stummeyer für die vielen nützlichen Tips und die gute Betreuung der Kristallisations-Experimente. Außerdem danke ich unseren Kooperationspartnern Prof. Dr. Ralf Ficner, Dr. Achim Dickmanns und Eike Schulz von der Universität Göttingen für die sehr gute Zusammenarbeit bei der Röntgenstrukturanalyse.

Ein ganz herzlicher Dank gilt Astrid Oberbeck und Maike Hartmann für die hervorragende technische Hilfestellung in Teilen dieser Arbeit. Desweiteren bedanke ich mich beim gesamten „Team aus Labor 1210“ für die tolle Arbeitsatmosphäre und insbesondere Melanie Grove und Andrea Bethe für die vielen praktischen Tips im Laboralltag.
Dem gesamten Labteam gilt ein besonders herzlicher Dank für die stete Hilfsbereitschaft, fachliche und einfach lustige Gespräche sowie natürlich auch für die vielen leckeren Kuchen. Es hat mir richtig viel Spaß gemacht in dieser positiven Arbeitsatmosphäre mit Euch zu arbeiten!

Ein besonderer Dank geht an Tanja Herrmann, Wiebke Schaper und Aike Schweda für ihre Freundschaft und dafür, dass sie immer ein offenes Ohr für mich haben.

Ein ganz großer Dank gilt an dieser Stelle meinen Eltern, Helga und Karl Wilhelm Bergfeld, für eure Liebe, euer Vertrauen, eure Unterstützung und dafür, dass ihr einfach immer für mich da seid!