



HAL
open science

Identification of amino acid residues at the active site of endosialidase that dissociate the polysialic acid binding and cleaving activities in *Escherichia coli* K1 bacteriophages

Elina Jakobsson, Anne Jokilammi, Juha Aalto, Pauli Ollikka, Jukka V Lehtonen, Harri Hirvonen, Jukka Finne

► To cite this version:

Elina Jakobsson, Anne Jokilammi, Juha Aalto, Pauli Ollikka, Jukka V Lehtonen, et al.. Identification of amino acid residues at the active site of endosialidase that dissociate the polysialic acid binding and cleaving activities in *Escherichia coli* K1 bacteriophages. *Biochemical Journal*, 2007, 405 (3), pp.465-472. 10.1042/BJ20070177 . hal-00478750

HAL Id: hal-00478750

<https://hal.science/hal-00478750>

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identification of amino acid residues at the active site of endosialidase that dissociate the polysialic acid binding and cleaving activities in *Escherichia coli* K1 bacteriophages

Elina Jakobsson^{*,1}, Anne Jokilammi^{*,1}, Juha Aalto^{*}, Pauli Ollikka^{*}, Jukka V. Lehtonen[†], Harri Hirvonen^{*,‡} and Jukka Finne^{*,2}

^{*}Department of Medical Biochemistry and Molecular Biology, University of Turku, FI-20520 Turku, Finland

[†]FBD Ltd, Lemminkäisenkatu 14-18 A U, FI-20520 Turku, Finland

[‡]Department of Neurology, University Central Hospital of Turku, FI-20520 Turku, Finland

Keywords: endosialidase; endo-N-acetylneuraminidase; bacteriophage; polysialic acid; host range mutants; catalytic activity

Running title: Mutant phage endosialidase

⁰Abbreviations used: GFP, green fluorescent protein; NCAM, neural cell adhesion molecule; polySia, polysialic acid

¹These authors have contributed equally to this work.

²Corresponding author. Mailing address: Department of Medical Biochemistry and Molecular Biology, University of Turku, Kiinamylynkatu 10, FI-20520 Turku, Finland. Phone: +358-2-333 7240. Fax: +358-2-333 7229. Email address: jukka.finne@utu.fi.

Synopsis

Endosialidase is a tailspike enzyme of bacteriophages specific for human pathogenic *Escherichia coli* K1, which specifically recognizes and degrades polysialic acid. Polysialic acid is also a polysaccharide of the capsules of other meningitis- and sepsis-causing bacteria, and a post-translational modification of the neural cell adhesion molecule NCAM. We have cloned and sequenced three spontaneously mutated endosialidases of the PK1A phage and one of the PK1E phage which display lost or residual enzyme activity but retain the binding activity to polysialic acid. Single to triple amino acid substitutions were identified, and back-mutation constructs indicated that single substitutions accounted for only partial reduction of enzymic activity. A homology-based structural model of endosialidase revealed that all substituted amino acid residues localize to the active site of the enzyme. The results reveal the importance of non-catalytic amino acid residues for the enzymatic activity. The results reveal the molecular background for the dissociation of the polysialic acid binding and cleaving activities of endosialidase and for the evolution of host range mutants of *E. coli* K1 bacteriophages.

Introduction

Polysialic acid (polySia) is a homopolymer of α -(2,8)-linked N-acetylneuraminic acid and constitutes the capsule of the pathogenic bacteria *Escherichia coli* K1, *Neisseria meningitidis* group B, *Mannheimia (Pasteurella) hemolytica* A2, and *Moraxella nonliquefaciens* [1–3]. Polysialic acid, the K1 antigen, is a promising target of phage therapy [4, 5]. PolySia is also a post-translational modification of the neural cell adhesion molecule (NCAM) and promotes complex neurological processes, such as plasticity [6, 7]. Polysialic acid synthesis and its role to direct tumor cell growth has been studied extensively [8–10].

E. coli K1-specific bacteriophages encoding K1 depolymerase activity have been isolated [11–13]. The coding genes of the polySia degrading enzyme endosialidase have been cloned from some of the phages [13–15]. We have previously isolated host range mutants of the PK1A and PK1E phages using *E. coli* mutants with a sparse polySia capsule [16, 17]. The phage mutants have lost the catalytic activity of their endosialidase partly or totally, but still retain their polySia binding activity [18]. The structural basis for the dissociation of the catalytic and binding activities has remained unresolved.

In the present work we have cloned and sequenced the PK1A endosialidase gene of the wild type phage and its three mutants, as well as a mutant of the PK1E phage, and identified the amino acid substitutions in their sequence. A homology-based structural model of the endosialidase predicts the spatial arrangement of the mutant amino acids close to each other at the active site of the enzyme, which thus indicates the molecular basis of dissociation of the binding and catalytic activity, and the altered host range of the mutant bacteriophages.

Materials and Methods

Phages and bacterial strains

Bacteriophage PK1A was kindly supplied by Dr. B. Rowe, Central Public Health Laboratory, Colindale [11]. The mutant phages PK1A2, PK1A5, PK1A8 and PK1E3 have been described by Pelkonen *et al.* [16]. The following *E. coli* strains were used: IH3088 [19] and K1-defective PK1A-resistant mutant strains, EH954 and EH1008, derived from rough *E. coli*

IH3088 [17, 19]. IH3088 is the host of the native phage PK1A and PK1E. EH954 is infected by PK1A2. PK1A5, PK1A8 and PK1E3 have EH1008 host specificity. L-broth, L-agar plates, and soft agar [20] were used for the cultivation of the organisms. The phages were propagated by standard methods [21].

Purification of the genomic phage DNA

For the DNA purification 100-200 μg of phage was used (measured by the Lowry-method). Proteinase K (Promega) and sodium dodecyl sulfate (SDS) were added to a final concentration of 200 $\mu\text{g}/\text{ml}$ and 0.5%, respectively in a total volume of 450 μl . After incubating for 2.5 hours at 37 °C the DNA was extracted with phenol-chloroform and precipitated with 0.1 vol of 3M sodium acetate pH 5.2 and dissolved in water.

Cloning, sequencing and gene analysis

Part of the endosialidase gene, the Bgl II–Bcl I-fragment, was cloned into pBluescript vector. DNA was sequenced on both strands using the universal primers T3 and T7. The sequence was completed by primer walking using synthetic oligonucleotide primers (Table 1) and the total phage genome as template. The oligonucleotides were synthesized by Cyber Gene (Novum, SE-141 57 Huddinge, Sweden). The same oligonucleotides were used to determine the mutant sequences. The sequences were determined on an ABI PRISM 377 Sequencer; Department of Medical Genetics, University of Turku. The assembly and sequence analysis were performed with DNASTAR (Lasergene package), BLAST [22] and EMBOSS software [23]. EMBOSS Needle program was used to compare translated amino acid sequences.

Construction of fusion proteins

The catalytically active and inactive fusion proteins have been described earlier [24]. Constructions of point mutated fusion proteins (417- and 489-mutation) were made from catalytically inactive fusion protein by PCR, with the primers listed (Table 2). N- and C-terminal deletions of inactive EndoA2GFP fusion protein were made by PCR using the primers listed (Table 2). Constructs were monitored by sequencing. The products were expressed as described earlier [24].

Enzyme activity assays

Thiobarbituric acid assay was used to measure endosialidase activity. The reagent volumes used in this study were half of the volumes of the modified Horgan procedure [25], except that the chromophore was extracted with 0.75 ml of acidified n-butanol. The total reaction volume was 50 μ l adjusted with 100 mM sodium phosphate buffer pH 6.6. 15 μ l cultured phage stock or 4 μ g protein were used as enzyme samples. Colominic acid (poly-2,8-N-acetylneuraminic acid sodium salt, Sigma) substrate 1 μ g/ μ l was added in 30 μ l 100 mM sodium phosphate buffer pH 6.6 for phage stocks or 50 mM sodium phosphate buffer pH 7.5, 300 mM NaCl for protein samples. Absorbance was measured at 549 nm against phosphate buffer and corrected with a substrate blank containing no enzyme. N-acetylneuraminic acid (Fluka) was used as reference. Results represent mean values from three independent assays. The protein concentration measurements of the phage stocks were taken into account in calculating the activity results.

Fluorescence microscopy of cells and paraffin sections

Human neuroblastoma SHSY-5Y cells (ECACC:94030304) and baby hamster kidney cells, BHK-21[C-13](ATCC:CCL-10), were cultured on 13-mm glass slides in 24-well plates. The cells were fixed in 600 μ l of 3 % paraformaldehyde in PBS (10 mM potassium phosphate buffer pH 7.5, 0.15 M NaCl) at room temperature for 15 minutes, washed three times with PBS, and blocked with normal horse serum (Vector Laboratories) half an hour at room temperature. 250 μ l of 10 μ g/ml each fusion protein in the buffer was added to the cells. Alternatively, a dilution series of the deletion fusion protein derivatives (10, 5, 3 and 0.5 μ g/ml) in PBS were added to the cells. The cells were incubated for one hour at room temperature, washed six times with the PBS, mounted on object glasses with Immu-mount and examined by fluorescence microscopy. Controls included staining on the presence of colominic acid (100 μ g/ml, 30 minutes) or preincubation on the presence of active endosialidase enzyme (10 μ g/ml) in PBS at room temperature for 30 min.

Paraffin embedded temporal section of human fetal (14 w) cortex cerebrum were sectioned (4 μ m), deparafinized and rehydrated in descending ethanol series. Staining was done as described for the cells. The tissue sample were obtained with the appropriate approval of the Joint Ethical

Committee of the University Hospital of Turku and the University of Turku, Finland.

Protein homology modeling

The sequence of the PK1A endosialidase was aligned to the sequence of chain D of the PK1F structure (PDB ID: 1V0E [26]) using the MALIGN [27] algorithm in the computer program Bodil [28]. The default structural families based scoring matrix [27] was used in the alignment, although – considering the similarity of the compared sequences – selection of a different selection matrix would not change the result significantly. The PK1A and PK1F endosialidases were assumed to contain a conserved active site structure due to the high sequence similarity. The residues corresponding to the point mutations of PK1A2, PK1A5, and PK1A8 endosialidase mutants were located from the PK1F structure via sequence alignment.

Sequence deposition

The PK1A endosialidase sequence data will appear in the GenBank Nucleotide Sequence Data Library under the accession number EF507428.

Results

Characteristics of PK1A endosialidase sequence

Phage PK1A is a helical bacterial virus with double stranded DNA. Endo-alpha-sialidase (EC 3.2.1.129), a glycosyl hydrolase, is a component of the phage's short tail structure. The length of the native PK1A endosialidase sequence is 2436 bp and encodes a polypeptide of 811 amino acid residues. According to the amino acid sequence (Figure 1) the molecular weight of the polypeptide would be 89 kDa, but due to post-translational modification the molecular weight of the mature protein monomer is 76 kDa (Figure 2). PK1E endosialidase is similar in size to the PK1A enzyme. The enzymes are of exactly the same length as the PK1-5 endosialidase [13], while the PK1F polypeptide is longer and consists of 1064 amino acid residues [29].

The bacteriophage K1A endosialidase gene contains two copies of the sialidase (neuraminidase) motif Ser-Xaa-Asp-Xaa-Gly-Xaa-Thr-Trp (Figure 1).

This Asp-motif, typical of glycosyl hydrolases, probably maintains the correct structural fold and affects the solubility properties [30–33].

Based on sequence alignment an eight-element fingerprint provides a signature for the family 58 hydrolases. The PK1A endosialidase contains the signature at the following amino acid sequence positions: 104–129, 169–198, 288–312, 358–385, 436–463, 492–517, 598–625, and 630–658 (<http://www.ebi.ac.uk/interpro/> and <http://afmb.cnrs-mrs.fr/CAZY/>). At the amino acid positions 38–45 the PK1A endosialidase matches the (Ala or Gly)-Xaa-Xaa-Xaa-Xaa-Gly-Lys-(Ser or Thr) pattern (Figure 1). This is the best-conserved motif of ATP/GTP-binding proteins named “A” consensus sequence (or P-loop) [34, 35].

The K1-5 promoter sequence has similarity with the consensus promoter of SP6. The putative PK1A endosialidase promoter (Figure 1) matches exactly to the K1-5 consensus promoter-like sequence predicted by Scholl *et al.* [36].

The PK1A nucleotide sequence upstream from the endosialidase coding region and Shine-Dalgarno (S-D) site has another putative S-D site (Figure 1). According to the standard genetic code translation there is one putative but rare start codon, GUG, which initiates to code a short peptide (18 amino acids). Downstream from the PK1A endosialidase coding sequence a putative S-D site and a start codon are found but the putative peptide (51 amino acids) has no significant similarity to EMBL database sequences. The other published coliphage endosialidase sequences and the K5 lyase (coliphage K5) have the highest similarity to the 3840bp PK1A sequence reported here. The translated native endosialidase PK1A amino acid sequence is very similar to the PK1E and K1-5 translations, 90 and 89% respectively.

Identification of point mutations in the mutant endosialidase

The three PK1A mutant endosialidases and the PK1E mutant endosialidase are equal in size as compared to the wild type of the enzyme. Each mutant has displayed one to three amino acid substitutions at different locations (Table 3). However, the substitution R614G in PK1A5 and PK1A8 is a conserved normal alleles present in the wild-type phages PK1E, PK1-5 and PK1F displaying full endosialidase activity.

Effects of point mutations on the catalytic and binding activities

Quantitative determination of the enzymatic activities revealed that PK1A2 and PK1A5 mutant phage endosialidases are catalytically inactive, whereas the PK1A8 and PK1E3 mutants exhibit greatly reduced activity (Figure 3). The results correlate well with the previously published qualitative estimation of the enzymic activities and binding properties [16].

In order to study the role of the two point mutations in the PK1A2 endosialidase, the gene was back-mutated at the individual locations and the mutant proteins harboring single amino acid substitutions were expressed as GFP-fusion proteins. Both back-mutants were able to cleave polySia, but with less activity than the native endosialidase (Figure 4). Thus, both of the two amino acid were needed for complete inactivation of the enzyme.

The influence of the single amino acid substitutions on the polySia binding activity of the enzyme were tested by binding to polySia expressing neuroblastoma cells (binding defined as sustained binding). Of the GFP-fusion protein derivatives only the catalytically inactive PK1A2 endosialidase remained bound to the cells (Figure 5).

Effects of N- and C-terminal deletions on the catalytic and binding activities of endosialidase

In order to investigate whether deletions in terminal areas have an effect on the binding or catalysis, we made N- and C-terminal deletions to the wild type and the mutant PK1A endosialidases (Table 4). The catalytic activity of wild type PK1A endosialidase was very sensitive to deletions: only the first N-terminally deletion derivate (1–47) was catalytically active, whereas all C-terminally truncated derivatives were inactive. In case of the PK1A mutant endosialidase only the two first N-terminally deleted derivatives, Δ 1–47 and 1–107, bound polySia in a sustained manner, but also in these cases, the expression levels were low. Interestingly, truncation of the residues 1–107 of the wild-type PK1A enzyme resulted in a derivate which was catalytically inactive but which retained its binding activity to polysialic acid (Figure 6). As a difference to the PK1A2 mutant endosialidase, the wild type truncated derivative with the same catalytic and binding properties was, however, expressed in reduced quantities.

Endosialidase derivatives with longer truncations were all insoluble. Retrospectively, based on comparison with the PK1F structure [26], the deletions that produced insoluble enzyme obviously removed essential parts of the central (catalytic) domain or the C-terminal portion, which is important for trimerisation.

Structural model of PK1A endosialidase

The crystal structure of PK1F endosialidase complexed with oligomeric sialic acid has been reported [26]. This endosialidase assembles into a homotrimer, and each monomer has three distinct domains: β -propeller, β -barrel, and tailspike. The PK1F sequence contains an amino-terminal extension, which is not found in PK1A, PK1E, and K1-5 endosialidases. The central regions of the enzymes are highly conserved including two “Asp-box” motifs and all four enzymes are posttranslationally processed from their C-terminus.

The sequence alignment of the PK1A and PK1F endosialidases indicates homology of the proteins: out of 666 residues in PK1F 663 were aligned with PK1A residues and 555 of those pairs are identical. Residues 1–37 of PK1A did not align with any residue of PK1F. The PK1F endosialidase structure lacks its amino-terminal residues and thus no correspondence can be established. However, the aligned section of the PK1A and PK1F sequences have 83% identity. Thus, the PK1F endosialidase structure can be considered a viable template for building of a homology-based structural model of residues 38–706 of PK1A endosialidase.

The high identity of the homologs warranted us to initially forfeit the actual building of the homology model and study the residues of the PK1F endosialidase structure that correspond to the point mutations we observed in PK1A endosialidase. The distribution of conserved residues within the alignment supports this approach: the β -propeller domain, which contains the active site of the enzyme, is well conserved, the sequences are 93% identical. Furthermore, the β -barrel domain is 100% identical in sequence. Thus, most differences in the sequences are located in the tailspike domain and in the amino-terminal fragment, which forms a “cap” for the β -propeller domain.

Four spontaneous point mutations implicated in the reduced activity of the PK1A mutants are located on or near the putative active site, where cleavage of polySia has been proposed to occur (Figure 7). Also the amino

acid substitution in the PK1E3 endosialidase localizes to this site. Interestingly, the amino acid substitutions do not correspond to any of the three residues (Glu371, Arg386, and Arg437 in PK1A) that in PK1F were identified as active site residues. Furthermore, the four residues are all conserved between the PK1A and PK1F endosialidases. A common feature for all the residues involved in PK1A and PK1E activity is that single amino acid substitution did not inactivate the catalytic activity totally. For complete inactivation a combination of two substitutions, H417Y and N489D in PK1A2, H332N and N489D in PK1A5, or a combination of two of the three active site residues in PK1F, are required.

Tryptophan 118 forms one side of the hydrophobic pocket of the active cleft. A cysteine in the same position would clearly change the hydrophobic interactions that may occur between the substrate. Thus, this mutation is likely to change specificity and affinity.

Histidine 332 is located next to glutamic acid 315, which in turn is next to the glutamic acid 371 identified as an active site residue in PK1F endosialidase. A histidine to asparagine mutation (as in PK1A5) is likely to change the coordination of Glu371 and thus affect its interaction with polySia.

Histidine 417 is buried and therefore not in a position to directly bind to the substrate. Thus, the mutation to tyrosine (PK1A2) may have an indirect effect, since the surrounding residues have to compensate for the change of size, shape and hydrogen bonding tendencies.

Asn489 is on the edge of the putative “cup-like” active site. Asn489 forms a part of a ridge that is between the putative active site and polySia binding site formed by an other chain of the trimer [26]. Thus, it is possible that when the polySia chain is bound into the active site for cleavage, the chain forms interactions with asparagine at position 489. However, when the asparagine has mutated into aspartic acid, the charge of the side chain has changed so that the carbohydrate chain cannot form all required interactions.

The two endosialidases from PK1E and PK1E3 differ only by one amino acid residue. The PK1E3 endosialidase has a glutamic acid in position 370, while the PK1E, PK1A, and all its spontaneous mutations have alanine in that position. The effect of this single amino acid substitution on the enzyme activity is large (Figure 4). The magnitude of the effect can be explained by the rather radical change from a hydrophobic residue to an acidic one, the difference of the side-chain size, and in particular by the position. The

residue 370 precedes glutamic acid 371, which in PK1F has been identified to have a major influence on activity [26] (Figure 7). The PK1F structure contains a serine residue at the equivalent position [26] and is in fact the only residue in the active site area that differs between the native endosialidase PK1A and that of PK1F.

Discussion

Phages have a strong tendency to produce point mutations to the host identifying tail structure, in order to extend their host range [37]. Head proteins are usually more stable [38]. The spontaneous point mutations in endosialidase reflect nature's susceptibility to create new wider host specificities. By keeping the binding property while decreasing the catalytic efficiency the mutants acquire the property to infect *E. coli* with a sparse polysialic acid capsule. Sustained binding apparently provides the time required for the infection to occur in an environment of decreased number of substrate chains.

The spontaneous point mutations detected in this work significantly reduce the ability of the endosialidase enzyme to cleave polySia. The point mutations reside around the active site residues and thus likely affect the shape of the active cleft, coordination of the putative catalytic residues, and the interactions with the polySia substrate. In the mutant bacteriophage endosialidases the substrate binding property of the mutants is left, but the binding of polySia does not promote the transfer to the transition state as in the native enzyme, and thus the mutant enzymes cannot act as effective catalysts.

The dissociation constant, K_d , determined for the PK1A2 mutant fusion endosialidase and polySia earlier is high [24]. The value $19 \times 10^{-9} \text{M}$ is comparable and even higher than typical values, 10^{-6} to 10^{-8} , for carbohydrate antibodies [39]. Anyway, the binding in the case of mutant endosialidase PK1A2 is sustained and multiple conformational changes characteristic for the enzymatic activity are hindered. To determine the dissociation constant for catalytically active endosialidase and polySia is not possible.

The effects of the relatively conservative spontaneous mutations resemble the effects of engineered alanine-substitutions in endosialidases [26]. In PK1F the substitutions E581A, R596A, or R647A reduce catalytic activity by 95% [26], and in PK1E G138A by 80% [31]. In the present study the activities of

the mutants N489D (Back a) and H417Y (Back b) have decreased by 70% and 50%, respectively. In all cases one substitution is alone not enough to abolish catalytic activity completely. Double mutants are needed for total inactivation.

The host range mutant bacteriophages studied were isolated using *E. coli* mutants with a highly reduced amount of polysialic acid capsule. The low or absent activity of the mutant endosialidases seems to be the essential requirement of the mutant bacteriophages to bind the sparse capsule of their host bacteria. Identification of individual amino acid substitutions in the active site of endosialidases characterized the molecular background of the changes that contributed to the development of the host range bacteriophage mutants of *E. coli* K1. Identification of the molecular changes that lead to the evolution of host range bacteriophages infecting bacteria resistant to the parent phages contributes to our understanding of the molecular background of phage evolution and may also contribute to the development of practical applications, such as phage therapies to control pathogenic bacterial infections [40].

Acknowledgements

The research was supported by grants from the Finnish Academy and the Finnish Cultural Foundation.

We thank Dr. M. Venojärvi for help with confocal microscopy, Dr. O. Pentikäinen for insights to molecular modeling, and Mrs. T. Jompero, Mr. J. Karhu, and Ms. A. Mikkola for technical assistance.

References

1. Jennings, H. J., Katzenellenbogen, E., Lugowski, C., Michon, F., Roy, R. and Kasper, D. L. (1984) Structure, conformation and immunology of sialic acid-containing polysaccharides of human pathogenic bacteria. *Pure Appl. Chem.* **56**, 893–905
2. Devi, S. J. N., Schneerson, R., Egan, W., Vann, W. F., Robbins, J. B. and Shiloach, J. (1991) Identity between polysaccharide antigens of *Moraxella nonliquefaciens*, group B *Neisseria meningitidis*, and *Escherichia coli* K1 (non-O acetylated). *Infect. Immun.* **59**, 732–736

3. Adlam, C., Knights, J. M., Mugridge, A., Williams, J. M. and Lindon, J. C. (1987) Production of colominic acid by *Pasteurella haemolytica* serotype A2 organism. FEMS Microbiol. Lett. **42**, 23–26
4. 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
5. 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
6. Rutishauser, U. (1998) Polysialic acid at the cell surface: biophysics in service of cell interactions and tissue plasticity. J. Cell. Biochem. **70**, 304–312
7. Finne, J. (1982) Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. J. Biol. Chem. **257**, 11966–11970
8. Mühlenhoff, M., Manegold, A., Windfuhr, M., Gotza, B. and Gerardy-Schahn, R. (2001) The impact of N-glycosylation on the functions of polysialyltransferases. J. Biol. Chem. **276**, 34066–34073
9. Horstkorte, R., Mühlenhoff, M., Reutter, W., Nohring, S., Zimmermann-Kordmann, M. and Gerardy-Schahn, R. (2004) Selective inhibition of polysialyltransferase ST8SiaII by unnatural sialic acids. Exp. Cell Res. **298**, 268–274
10. Seidenfaden, R., Krauter, A., Schertzinger, F., Gerardy-Schahn, R. and Hildebrandt, H. (2003) Polysialic acid directs tumor cell growth by controlling heterophilic neural cell adhesion molecule interactions. Mol. Cell Biol. **23**, 5908–5918
11. Gross, R. J., Cheasty, T. and Rowe, B. (1977) Isolation of bacteriophages specific for the K1 polysaccharide antigen of *Escherichia coli*. J. Clin. Microbiol. **6**, 548–550

12. Machida, Y., Miyake, K., Hattori, K., Yamamoto, S., Kawase, M. and Iijima, S. (2000) Structure and function of a novel coliphage-associated sialidase. *FEMS Microbiol. Lett.* **182**, 333–337
13. Scholl, D., Rogers, S., Adhya, S. and Merrill, C. R. (2001) Bacteriophage K1-5 encodes two different tail fiber proteins, allowing it to infect and replicate on both K1 and K5 strains of *Escherichia coli*. *J. Virol.* **75**, 2509–2515
14. Long, G. S., Bryant, J. M., Taylor, P. W. and Luzio, J. P. (1995) Complete nucleotide sequence of the gene encoding bacteriophage E endosialidase: implication for K1E endosialidase structure and function. *Biochem. J.* **309**, 543–550
15. Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S., Lottspeich, F. and Frosch, M. (1995) Molecular cloning and functional expression of bacteriophage PK1E-encoded endoneuraminidase Endo NE. *Mol. Microbiol.* **16**, 441–450
16. Pelkonen, S., Aalto, J. and Finne, J. (1992) Differential activities of bacteriophage depolymerase on bacterial polysaccharide: Binding is essential but degradation is inhibitory in phage infection of K1-defective *Escherichia coli*. *J. Bacteriol.* **174**, 7757–7761
17. Pelkonen, S. (1990) Capsular sialyl chains of *Escherichia coli* K1 mutants resistant to K1 phage. *Curr. Microbiol.* **21**, 23–28
18. Aalto, J., Pelkonen, S., Kalimo, H. and Finne, J. (2001) Mutant bacteriophage with non-catalytic endosialidase binds to both bacterial and eukaryotic polysialic acid and can be used as probe for its detection. *Glycoconj. J.* **18**, 751–758
19. Pelkonen, S. and Finne, J. (1987) A rapid turbidimetric assay for the study of serum sensitivity of *Escherichia coli*. *FEMS Microbiol. Lett.* **42**, 53–57
20. Maniatis, T., Fritsch, E. F. and Sambrook, J., eds. (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

21. Mäkelä, P. H. (1985) The use of bacteriophages in the study of bacterial cell surface structures. In *Enterobacterial surface antigens: methods for molecular characterization* (Korhonen, T. K., Dawes, E. A. and Mäkelä, P. H., eds.), pp. 155–178. Elsevier Science Publishers, Amsterdam
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410
23. Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M. D., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N. J., Oinn, T. M., Pagni, M., Servant, F., Sigrist, C. J. and Zdobnov, E. M. (2000) InterPro – an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* **16**, 1145–1150
24. Jokilampi, A., Ollikka, P., Korja, M., Jakobsson, E., Loimaranta, V., Haataja, S., Hirvonen, H. and Finne, J. (2004) Construction of antibody mimics from a noncatalytic enzyme-detection of polysialic acid. *J. Immunol. Methods* **295**, 149–160
25. Horgan, I. E. (1981) A modified spectrophotometric method for determination nanogram quantities of sialic acid. *Clin. Chim. Acta* **116**, 409–415
26. Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R. and Ficner, R. (2005) Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nat. Struct. Mol. Biol.* **12**, 90–96
27. Johnson, M. S. and Overington, J. P. (1993) A structural basis for sequence comparisons. An evaluation of scoring methodologies. *J. Mol. Biol.* **233**, 716–738
28. Lehtonen, J. V., Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., Pentikäinen, O., Nyrönen, T., Salminen, T., Gyllenberg, M. and Johnson, M. S. (2004) BODIL: a molecular modeling environment for structure-function analysis and drug design. *J. Comput. Aided Mol. Des.* **18**, 401–419

29. Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M. and Gerardy-Schahn, R. (2003) Proteolytic processing and oligomerization of bacteriophage-derived endosialidases. *J. Biol. Chem.* **278**, 12634–12644
30. Chien, C. H., Shann, Y. J. and Sheu, S. Y. (1996) Site-directed mutations of the catalytic and conserved amino acids of the neuraminidase gene, nanH, of *Clostridium perfringens* ATCC 10543. *Enzyme Microb. Technol.* **19**, 267–276
31. Leggate, D. R., Bryant, J. M., Redpath, M. B., Head, D., Taylor, P. W. and Luzio, J. P. (2002) Expression, mutagenesis and kinetic analysis of recombinant K1E endosialidase to define the site of proteolytic processing and requirements for catalysis. *Mol. Microbiol.* **44**, 749–760
32. Crennell, S., Garman, E., Laver, G., Vimr, E. and Taylor, G. (1994) Crystal structure of *Vibrio cholerae* neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. *Structure* **2**, 535–544
33. Gaskell, A., Crennell, S. J. and Taylor, G. L. (1995) The tree domains of a bacterial sialidase: a beta-propeller an immunoglobulin module and a galactose-binding jelly-roll. *Structure* **3**, 1197–1205
34. Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951
35. Saraste, M., Sibbald, P. R. and Wittinghofer, A. (1990) The P-loop – a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430–434
36. Scholl, D., Kieleczawa, J., Kemp, P., Rush, J., C., R. C., Merrill, C., Adhya, S. and Molineux, I. J. (2004) Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. *J. Mol. Biol.* **335**, 1151–1171
37. Haggård-Ljungquist, E., Halling, C. and Calendar, R. (1992) DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J. Bacteriol.* **174**, 1462–1477

38. Chibani-Chennoufi, S., Canchaya, C., Bruttin, A. and Brüssow, H. (2004) Comparative genomics of the T4-Like *Escherichia coli* phage JS98: implications for the evolution of T4 phages. *J. Bacteriol.* **186**, 8276–8286
39. Pellequer, J. L. and Van Regenmortel, M. H. (1993) Affinity of monoclonal antibodies to large multivalent antigens: influence of steric hindrance on antibody affinity constants calculated from Scatchard plots. *Mol. Immunol.* **30**, 955–958
40. Qimron, U., Marintcheva, B., Tabor, S. and Richardson, C. C. (2006) Genomewide screens for *Escherichia coli* genes affecting growth of T7 bacteriophage. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19039–19044
41. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242

Table 1: The oligonucleotides used to determine the endosialidase sequence of native phage PK1A and its three host range mutants and the oligonucleotides for sequencing of the PK1E3 mutant.

Number of oligonucleotide		5'-3' Sequence
PK1A	1	(=T3) 5'-AATTAACCCTCACTAAAGGG
	2	(=T7) 5'-TAATACGACTCACTATAGGG
	3	5'-CTTCAGAGGCTTCACCTTTG
	4	5'-TGAGATAGAAAGTACCTGCA
	5	5'-GCTGGAAAGAATTGGCACAT
	6	5'-ATGTTTGACCATCGTCACCA
	7	5'-TACCATCGTGACATCCCTC
	8	5'-TTACAAAGTATCATCTTTAC
	9	5'-AAAATGAATGGGAAGCAGGT
	10	5'-TAACCTCCTTTTCTGAAAGG
	11	5'-AAGTACAAGTGCTAATATTC
	12	5'-ACGTTCTGAACCAAACATAA
	13	5'-TGTGCACCACACTCACTTTG
	14	5'-CACAAAGTGAGGCGAATCAT
	15	5'-TTTAGTACACTCTCATCAAT
	16	5'-AAGGTAGGTATAGCCTGTCC
	17	5'-TTCCAGCGTTATTCAAATCT
PK1E3	Left 1	5'-CGTGGGCATCACTTGTATCA
	2	5'-CGGTACTGGTGCTACTTACA
	3	5'-GTGGTATCACTAAGGCTGCA
	4	5'-ACCTTATCACTCGTGGCACT
	5	5'-GCAGATTGCAAATGACAGTC
	6	5'-GATGCTGAACACAATTACGG
	Right 1	5'-AGCTGACACACGCCTATAC
	2	5'-TTAGTGGCATAACATCTCCTC
	3	5'-AGCAGGCATGATTGGTGCAC
	4	5'-GTTGTCACCATAAGTCATTG
	5	5'-ACTCCAGATGGTATCTGACG
	6	5'-ATGCACAATTGGTTAGTGCG

Table 2: The oligonucleotides used in the backmutations from PK1A2 endosialidase.

	Name	5'-3' Sequence
Back-mutations	Y417H	5'-CCACATAATGTGCATCATACTAC
	Y417H	5'-AAATCTTAGTGACTCCCAAGTC
	D489N	5'-GACATTGTGAACTCTAGTGTAGG
	D489N	5'-ACCCTGATAGATTTGGTCTGTG
N-terminal deletions	PK1A-A	5'-TCAGGTAAGAGAATTGACGG
	PK1A-B	5'-GCTTGGCCTCAAGACAAAGC
	PK1A-C	5'-CTGCATCCAGATTACCCTAC
	PK1A-D	5'-GGTGGTATCACTAAGGCTGC
	STALKEND	5'-CGTGACAGCTTATCTAACC
C-terminal deletions	PK1A-1	5'-TTCGGATGCAAAAAGCAGCTAC
	PK1A-2	5'-CTTATGAGTGGCATCAGACG
	PK1A-3	5'-AGAAGTACTAGAACCTGCTCC
	PK1A-4	5'-TGAGCGAATATTAGCACTTGTAC
	STOP-pri	5'-AGCTGAGCTTGGACTCCTG

Table 3: Amino acid substitutions in mutant endosialidases

Phage	Position in sequence					
	118	332	370	417	489	614
PK1A*	W	H	A	H	N	R
PK1A2*	W	H	A	Y	D	R
PK1A5*	W	N	A	H	D	G
PK1A8*	C	H	A	H	D	G
PK1E [†]	W	H	A	H	N	G
PK1E3*	W	H	E	H	N	G
K1-5 [‡]	W	H	A	H	N	G
PK1F ^{§,}	W	H	S	H	N	G

*This study

[†]Long *et al.*, [14] Gerardy-Schahn *et al.* [15]

[‡]Scholl *et al.* [13]

[§]Mühlenhoff *et al.* [29]

^{||}The residue numbers in PK1F are 328, 542, 580, 627, 699 and 819, respectively.

Table 4: Catalytic and binding activities of deletion derivatives of wild type PK1A and mutant PK1A2 endosialidases. The derivatives were expressed and purified as their GFP fusion proteins, and assayed for catalytic and binding activities as described in Figures 3 and 5. Binding is defined as sustained binding, distinct from the transient binding during catalysis. Endosialidases with the deletions 1–158, 1–211, 616–811 and 565–811 were insoluble.

	Deleted sequence	Catalytic activity	Observed binding
PK1A (wild type)	Unmodified	+	–
	Δ 1–47	+	–
	Δ 1–107	–	+
	Δ 719–811	–	–
	Δ 668–811	–	–
PK1A2 (mutant)	Unmodified	–	+
	Δ 1–47	–	+
	Δ 1–107	–	+
	Δ 719–811	–	–
	Δ 668–811	–	–

Figure legends

Figure 1 Primary nucleotide sequence and derived amino acid sequence of the endosialidase PK1A gene. A K1-5-like putative promoter region, the putative Shine-Dalgarno sites, the two sialidase motif sequences and ATP/GTP-binding site motif A (P-loop) are shown as boxed. An asterisk (*) marks the translation stop codon. Nucleotide and amino acid positions are given on both sites. This sequence data will appear in the GenBank Nucleotide Sequence Data Library under the accession number EF507428.

Figure 2 Molecular size of the native and mutant endosialidases. Sodium dodecyl sulfate polyacrylamide gel electrophoresis of the purified wild-type PK1A (1) and mutant PK1A2 (2) endosialidases. LMW, low molecular weight marker proteins.

Figure 3 Endosialidase activity of the wild-type phages PK1A and PK1E and the mutants PK1A2, PK1A5, PK1A8 and PK1E3. The results are represented in nanomoles of released N-acetylneuraminic acid per minute per milligram of phage protein.

Figure 4 Endosialidase activities of PK1A2 single amino acid mutants. Endosialidase derivatives back mutated for either of the two amino acid substitutions in PK1A2 endosialidase were constructed (Back a and Back b). The endosialidase activities of the GFP-fusion proteins were determined and expressed as relation to that of the native PK1A enzyme.

Figure 5 Sustained binding of endosialidase derivatives to polySia-expressing cells. Fluorescence microscopy of SHSY-5Y cells with endosialidase and endosialidase derivatives fused to GFP. Paraformaldehyde-fixed SHSY-5Y cells stained with catalytically active native endosialidase PK1A, catalytically inactive double mutant endosialidase PK1A2 or with the catalytically partially active single back mutants Back a and Back b. Bars 20 μm .

Figure 6 Binding of N-terminally truncated wild type PK1A endosialidase to polysialic acid. Paraffin-embedded temporal sections of fetal (14w) cortex cerebrum were stained using wild type PK1A endosialidase with the terminal truncation $\Delta 1-107$. (a) Staining with GFP-fusion protein of the endosialidase derivative (b) as in a but pretreated with active endosialidase before staining. Bars 100 μm .

Figure 7 A structural model of PK1A endosialidase active site based on the PK1F endosialidase structure (PDB [41] ID: 1V0E [26]). The molecular surface is shown with cyan, except that the three residues corresponding to those identified in endoNF as active site residues are drawn with gray and the mutated residues with transparent red color and as space-filling models. Alanine 370, representing the mutation in PK1E3 (Table 3) is shown with a green surface. The residue numbers refer to the PK1A endosialidase sequence.

Figure 1:

```

-699 GATCACTTCAGGGTGTGACAGCGCGTTACCAGAGAATCCAAAGGTGAAGCCTCTGAAGCCGACGGCTACGGCAGTCTACGAGGTGGACGATAAAGTCTGTTTCAGTAAGCCTGACGCTAC -580
-579 AAAATTAGGTCGTGATATTCTCTCGCTAGAACCGCGGTATAAATTAACATAATTTTATGTATCAGTGCTTACAGATTACTGGACACTATAGAAGAGAGATAAGATAGGCCGCTCTCTTTGA -460
                                     Promoter
-459 GCGGCCTATTACTAGCCAATCTTCATAGGAGGGTTGGGACGTAATAGGAGAGTATATGGTAAAGTAACTAAACCCCAAGACAACGGGCTTACTACATAGAGATACTGACTAGTACCT -340
-339 TATTAGATAATTTACTATCTAAAAGCGGTGTACATTTGAAGGTAGTCCAAGTAAATACTAAAATTGAGATAGAAGTACCTGCAACAGGAGATTGATGTTCTCGTGGGCATCA -220
                                     Shine-Dalgarno
-219 CTGTATCATTAAAGTACACCTACACTTTGTACAGCTTGGATTACTAAAGTGTGACATACTAGATTAGAAGTACATGTGTTCCATACAGCACAAGTTGAAATAGACATAGATGTAGTATGTT -100
-99 TACATTCTAGGTAAACACATTTGTCAGTCCGTAAGCACTGCTTTTCGCGCAACTTTCTTAAAGGTTATCATGATGGTACGCTTTTCAGAAAAGGAGTTACATGATTCAAGACTAAGTTC 20
1                                     Shine-Dalgarno
21 TTCATTAGTTAAATTCAAGAGTAAAATAGCAGGTGCAATCTGGCGTAACTTGGATGACAAGCTCACCAGGTTGTATCGCTTAAAGATTGGAGCCAAAGGTGATGGTAAAGACAACGA 140
8 S L V K F K S K I A G A I W R N L D D K L T E V V S L K D F G A K G D G K T N D 47
                                     P-loop
141 CCAAGATGCAGTAAATGCAGCGATGGCTTCAGGTAAGAGAATGACGGTGTGGTCTACTTACAAGTATCATCTTACCTGATATGGAGCGATTCTATAACACCCGCTCTGATGGGA 260
48 Q D A V N A A M A S G K R I D G A G A T Y K V S S L P D M E R F Y I N T R F V W E 87
261 ACGTTTACAGGTCAACCTCTTACTATGTGAGTAAAGGTTTTATCAATGGTGAACCTATATAAAATCAGGATAACCCCTTATTACAATGCTTGGCCTCAAGACAAGCGTTTGTATATGA 380
88 R L S G Q P L Y Y V S K G F I N G E L Y K I T D N P Y Y N A W P Q D K A F V Y E 127
381 GAACGTGATATATGCACCTTACATGGGTAGCGACCGTCAATGGTGTAGTGTCTGCATGTATCATGGGTTAAGTCTGGTACCGTCAACATGCTACACAGAGTGGTAACTGA 500
128 N V I Y A P Y M G S D R H G V S R L H V S W V K S G D D G Q T W S T P E W L T D 167
                                     Sialidase motif
501 TCTGCATCCAGATTACCCTACAGTAACTATCTTGTATGAGCATGGGTATGTGCAACCGTCTGTTGCCATGATTGAAACAGCTACTTTAGCCAAAGACGCACTAACCAATTGTGC 620
168 L H P D Y P T V N Y H C M S M G V C R N R L F A M I E T R T L A K N A L T N C A 207
621 ATTTGGGATCGCCCTATGCTCTGATGCTGATCTTACGGGTGATCACTAAGGCTGCAATCAGGATATGCAACAATACATGTACCAGATCAGGACTATCTGTTGGGCGATTGTTG 740
202 L W D R P M S R S L H L T G G I T K A A N Q R Y A T I H V P D H G L Q I P S E Y E P 367
741 TAACCTCTCTAATCTGCGGTAAACAGGTGTATCAGGTGATATGACTGTTGCAACGGTAATAGATAAAGACAACCTTCCAGGTTCTTACAACCTAACCCAGCAGACTCAGATTGATAA 860
248 N F S N S A V T G V S G D M T V A T V I D K D N F T V L T P N Q Q T S D L N N A 287
861 TGGAAAGAATGGCACATGGGTACTTCTTCCATAAGTCTCCGCGGTAAAGACAGATCTGGTCTAATCCCTCATGTCCAGAGGTGATAGCTTTGCTACTATTGATAACAATGGCT 980
288 G K N W H M G T S F H K S P W R K T D L G L I P H V T E V H S F A T I D N N G F 327
981 TGTTATGGGCTATCATCAAGGTGATGTAGCTCCACGAGAAGTTGGGCTTTTCTACTCCCTGATGCTTCAATAGCCCTCTAATATGTTGCTGTCAGATACCATCTGATGATGAACC 1100
328 V M G Y H Q G D V A P R E V G L F Y F P D A F N S P S N Y V R R Q I P S E Y E P 367
1101 AGATCGCGCAGAGCCATGCATCAAGTACTATGACGGTGTATTATACCTTATCACTCGTGTACTGCGGCGACCGACTAGGAAGCTCTCTGCATCGTGTAGAGATATAGTCAAGCTTG 1220
368 D A A E P C I K Y Y D G V L Y L I T R G T R G D R L G S S L H R S R D I G Q T W 407
                                     Sialidase motif
1221 GGAGTCACTAAGATTTCCACATAATGTGCATATACTACTTTACCGTTTGCTAAGTAGGAGATGACCTTATTATGTTGGTTCAGAACGTCGACAAAATGAATGGGAAGCAGGTGCACC 1340
408 E S L R F P H N V H H T T L P F A K V G D D L I M F G S E R A E N E W E A G A P 447
1341 AGATGATCGTTACAAGGCTCTTATCCTGCTACCTTCTATGCAAGTAAATGGAATGCAAGCTGATATGGAATGGGTTAAACATCAGACAAAATCTATCAGGDTGACAT 1460
448 D S A S Y P R T F Y A R L N V N W N A A D I E W V N I T D P Q I G G T I 487
1461 TGTGAACCTAGTGTAGGTAGTGTAGTGTAAAGACAGCTACATTTACTATATCTTGGTGGGAAACCAATTCACCCCAATGACTTATGGTGAACAAGAGATAAAGACCC 1580
488 V N S S V G V G S V V V K D S Y I Y Y I F G G E N H F N P M T Y G D N R D K D P 527
1581 ATTTAAAGTCAAGGACACCCCTACTGATATACTACTATAAGATGACAGTGAATGCAAAATGACAGCTGTGATCTCGTAAAGTTACATATGGTCAACTCCCGGACAGGCTATACCTACCT 1700
528 F K G H G H P T D I Y C Y K M Q I A N D S R V S R K F T Y G A T P G Q A I P T F 567
1701 CATGGTACTGATGGTATCTGCTACTTCTGCTCCACTTCACTTATCAGGAGAGTGTGGCCGAGATGACTGTGCAACATCTTACCCTGAAGGCAAGTCAAGTGTCAATATTCC 1820
568 M G T D G I R T I P A P L H L S G E V V A A D M T V E H L T L K A S T S A N I R 607
1821 CTCAGAGATGCTGATGGAACGTAATATGGGTTCAATGTAAGACTATACCAAAGGATAACCCCACTGCCAACGTCATCATCTCTGGTGGAGGGGTACAGCAGCAGATACCGGTGC 1940
608 S E M L M E R E Y G F I G K T I P K D N P T A Q R V I I S G G G T A A D T G A 647
1941 GCAAATTAACCTACATGGAGCAGTCTGACTACTCTAGACCGCAGTATATAATGCAATGAGCACTTGTCCAAAGCGGGCTATCATGCCCTACAATGATAACGTGATCCTGCTGG 2060
648 Q I T L H G A G S S T S R R A V Y N A N E H L F Q S G A I M P Y N D N V Y P A G 687
2061 CGGCCCTCGTAATAGATTCACTACTATCCACTTACAAGTGAACCTATTATACGCTGATGCCACTCATAAGTATGGATCACAAGGATTGATGAGAGTGTACTAAAGGATGGGGTAA 2180
688 G P R N R F T T I H L T S D P I I T S D A T H K Y G S Q G I D E S V L K A W G K 727
2181 GGTAGCTTCAAGCAGTACAAGTTGATTTGGTGTAGTGTCTGCTGGTGTGACCACACTCACTTTGGTGTATTGGCACAAGACATTGTAGCTGCTTTGTCATCGAAGGGTTAGATGCAAT 2300
728 V S F K Q Y K L I G E M S R G V H H T H F G V L A Q D I V A A F A S E G L D A I 767
2301 CGACTTCGGCATTGTCTCCTTCSAAGAAGTCAAGTTGGTGTAGATATAGTAAATCTTAATCTGGAGGCTGCTTATCTCGTTATCGTTAGATAAGTTAGAGAGATGTATGCCAC 2420
768 D F G I V S F E E G Q F G V R Y S E I L I L E A A Y T R Y R L D K L E E M Y A T 807
2421 TAATAAAATCAGTTAAGCAAGTCTGCTACTCCAGAAATCAGAGAAGCACTTATTCAATCAGGACGTGACCCCTAAGCAGGCATACGCCATTGCCAAGGATGACAACGTAGAGCATTGAGTA 2540
808 N K I S * 811
2541 AGCCCTCTGATCGTCTGCGAAGCAGTAACTAGTCTACACAAGGCTGGTTTAGTTGGCCCTTAGTGCAACCCAAAGGAGTTGATGATTGGATTCACTACTATGGAGTCGGCC 2660
                                     Shine-Dalgarno
2661 CATTCAAACCTAAGTACTCAGGAAGATACATGACACTGAAGAGTCTGCACTGGCTAGCCAGAAAGTCTTGGTGAAGTGTACAGGCTTACCTGAAGCAACCCAAAGGATTCTACTTAG 2780
2781 TGTAATCAAATGGAGATAACATGTTAGTAGGCTTAATGATTTTATCTTAGGTGATGCGTAGGTTTCTGTATTATGCCTGAGCAATAGCTTGTGCACTTAAAGGCTGAGGATAGGTA 2900
2901 CCTTAGTCTGTACAAGGATCAATCAAAGTCAAATTTGATATAGCGGTGTGTCAAGTCTCTCGCCCTGCGCCCTGCGCGACTTCCCCATATGGGGTCCGCTGGGTTGGCTGGGA 3020
3021 ATTTGGCTAGGCTGGGCGCTCTTCAACCTGCTGCCGAGGAAGCTGACGGGCTGGTGGAGTTGGCTTGTACTCGCTTAGTGTACACAAGGAAAGCCCTAGGAAGCCCTAGGCGAC 3140

```


Figure 2:

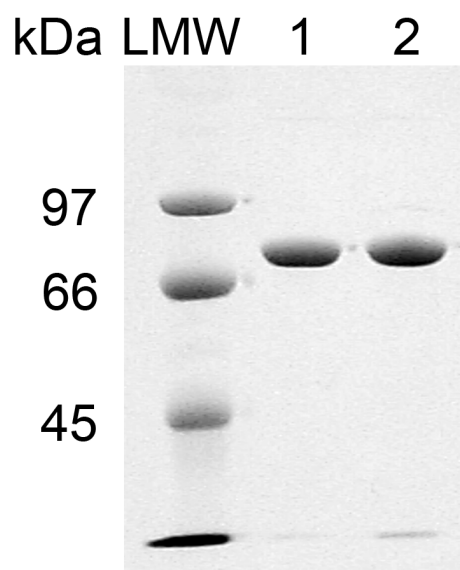


Figure 3:

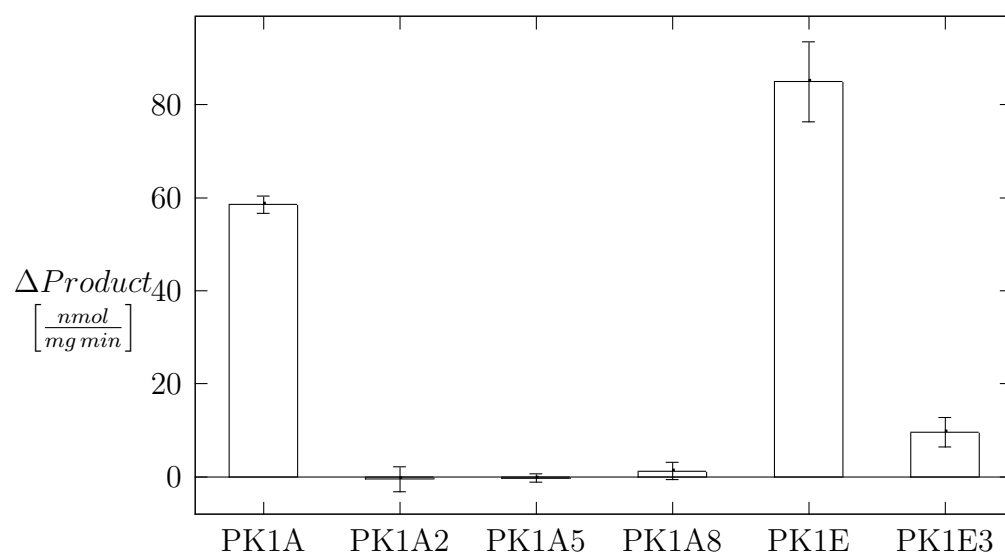


Figure 4:

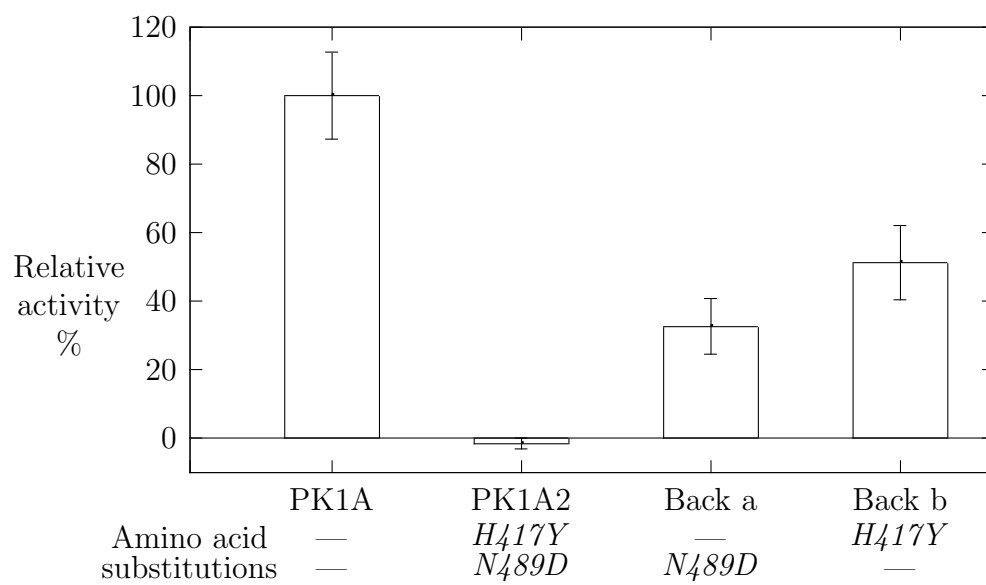


Figure 5:

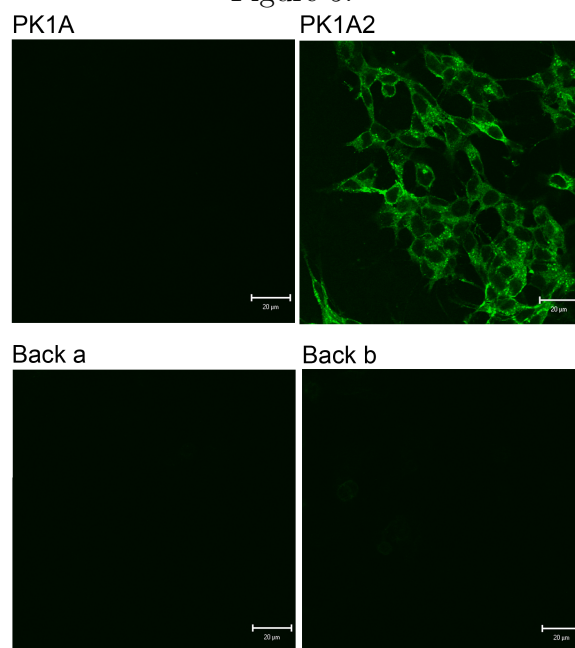


Figure 6:

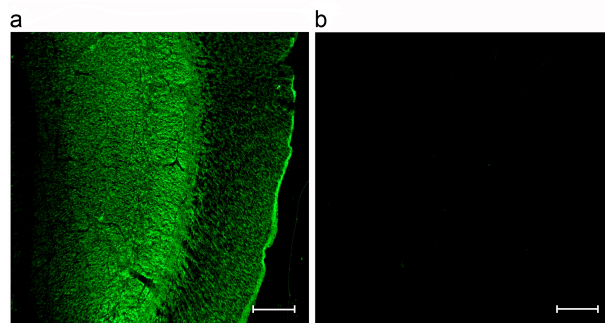


Figure 7:

