
Article:
Phansopa, C., Kozak, R.P., Liew, L.P. et al. (2015) Characterization of a sialate-O-acetylesterase (NanS) from the oral pathogen Tannerella forsythia that enhances sialic acid release by NanH, its cognate sialidase. Biochemical Journal, 472 (2). 157 - 167. ISSN: 0264-6021

https://doi.org/10.1042/BJ20150388

Characterization of a Sialate O-acetyl esterase (NanS) from the oral pathogen *Tannerella forsythia* that enhances sialic acid release by NanH, its cognate sialidase

Chatchawal Phansopa[1], Radoslaw P Kozak[2], Li Phing Liew[2], Andrew M Frey[1], Thomas Farmilo[1], Jennifer L Parker[1], David J Kelly[3], Robert J Emery[2], Rebecca I Thomson[4], Louise Royle[2], Richard A Gardner[2], Daniel I R Spencer[2] & Graham P. Stafford[1]

[1] Integrated BioSciences, School of Clinical Dentistry, University of Sheffield, Sheffield, S10 2TA, UK

[2] Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, UK

[3] Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield S10 2TN, UK

[4] Reading School of Pharmacy, University of Reading, Whiteknights, Reading RG6 6AP, UK

Keywords:
Bacteroidetes, glycans, sialic acid, acetlyesterase, Carbohydrate active enzyme, oral cavity

Short title: sialate-O-acetylesterase of *Tannerella*

Corresponding author
Graham P Stafford
Integrated BioSciences, School of Clinical Dentistry, University of Sheffield, Sheffield, S10 2TN, UK
g.stafford@sheffield.ac.uk

40-word summary
We characterize a novel bacterial sialate-O-acetylesterase potentially important for the nutrition of oral pathogens causing periodontal disease by enhancing their ability to harvest sialic acid sugar. Its high activity and stability indicate it can also be used in glycan pharmacoanalytics.

Abstract

Tannerella forsythia, a Gram-negative member of the *Bacteroidetes* has evolved to harvest and utilize sialic acid.. The most common sialic acid in humans is a mono-*N*-acetylated version termed Neu5Ac. Many bacteria are known to access sialic acid using sialidase enzymes. However, in humans a high proportion of sialic acid contains a second acetyl group attached via an O- group i.e. chiefly O-acetylated Neu5,9Ac_2 or Neu5,4Ac_2. This diacetylated sialic acid is not cleaved efficiently by many sialidases and in order to access diacetylated sialic acid, some organisms produce sialate-O-acetylesterases that catalyse removal of the second acetyl group. In this study we performed bioinformatic and biochemical characterization of a putative sialate-O-acetylesterase from *T. forsythia* (NanS), which contains two putative SGNH-hydrolase domains related to sialate-O-acetylesterases from a range of organisms. Purification of recombinant NanS revealed an esterase that has activity against Neu5,9Ac_2 and its glycolyl form Neu5Gc,9Ac. Importantly the enzyme did not remove acetyl groups positioned at the 4-O position (Neu5,4Ac_2). In addition NanS can act upon complex N-glycans released from a glycoprotein (EPO), Bovine submaxillary mucin and oral epithelial cell-bound glycans. When incubated with its cognate sialidase, NanS increased sialic acid release from mucin and oral epithelial cell surfaces, implying that this esterase improves sialic acid harvesting for this pathogen and potentially other members of the oral microbiome. In summary, we have characterized a novel sialate-O-acetylesterase that contributes to the siallobiology of this important human pathogen and has potential applications in analysis of sialic acid diacetylation of biologics in the pharmaceutical industry.

Introduction

The oral dwelling pathogen *Tannerella forsythia* is a member of a small group of bacteria that is associated with the prevalent oral disease of periodontitis (1). It is a member of a wider family of Gram-negative anaerobic human colonizing bacteria of the phylum *Bacteroidetes.* The *Bacteroidetes*, with *Tannerella* being no exception, are well adapted to life in the human body with a wide array of metabolic adaptations to their gut and oral environments, respectively (2, 3). In particular, our recent work and that of others on the physiology of *Tannerella* revealed that it has an extensive array of glycan harvesting capability encoded in its genome (3) with its ability to harvest and utilize sialic acid from glycoproteins present on human epithelial surfaces and secretory fluids most relevant, alongside potential dietary sources. This ability to utilize sialic acid is possibly the most pertinent *in vivo* and certainly the most well studied growth substrate for this enigmatic organism(4).

This ability to harvest and utilize sialic acid is encoded by a genetic locus in *Tannerella* that contains not only putative sialic acid catabolic genes (*nanA,E*) but also a unique sialic acid uptake system that is dependent on the periplasm spanning TonB protein and can thus be considered a sialic acid specific Polysaccharide Utilisation system (PUL) (5, 6). However, the substrate for this transport system, unlike many other *Bacteroidetes,* is the monomeric form of the sugar, rather than an oligosaccharide that is then acted upon by periplasmic sialidases and other enzymes before catabolism commences (7).

The sialic acid specific PUL of *Tannerella* secretes a sialidase enzyme of the GH33 family (8) with the ability to cleave sialic acids from fetuin and less efficiently from human and bovine mucin sources (9). One possible reason for this lower efficiency of cleavage from mucins has been hypothesized to be the high level of di-acetylation of sialic acid present that inhibits functional access of sialidases and slows cleavage rates for many sialidases (10, 11). This high level of di-acetylation is widespread in humans (12) and would represent a wasted nutritional source for oral bacteria such as *T. forsythia*. The diacetylated sialic acids are also important in viral infections with several viruses using the main form of di-acetylated sialic acid, Neu5,9,Ac_2, as a receptor, e.g. Influenza C and several coronaviruses (13,14). In addition both viruses and humans possess sialate-O-acetylesterase enzymes which in the case of the viruses act as receptor destroying enzymes involved in the infection cycle and in the case of humans are used for the modulation of di-acetylation of sialic acids on both internal and external proteins with biological roles in development and autoimmunity amongst others (15, 16), including certain leukaemias and cancers (17, 18). These enzymes are members of a large family of SGNH serine-esterases whose activity acts to remove O-acteyl groups from sialic acids, generally having specificity for acetyl groups at the 9- or 4- positions. In addition, recent evidence suggests that bacteria are also capable of producing sialate-O-acetylesterases that may be involved in survival in mucin rich environments (12), and the sialic acid catabolism operon of *Tannerella* appears to contain such an enzyme (19). However, information on this class of enzymes lags behind that of the sialidases and other classes of carbohydrate active enzymes.

In addition, this class of enzymes has potential value in the study and design of recombinant biopharmaceuticals. An example of this is the biological drug human erythropoietin (EPO) which is highly glycosylated with most glycans terminating in sialic acid residues which can contain two or more extra O-acetyl groups (20) which could consequently have an impact on the efficacy of the drug. As the specific structures of the glycans on a biopharmaceutical can be critical to the efficacy of the drug, it is essential to ensure that glycosylation is optimised and consistent from batch to batch. Thus, the availability of enzymes specific for additional O-acetylation on sialic acids could aid both the identification and quality control of glycosylation on such drugs.

In this study we set out to characterize the putative sialate-O-acetylesterase from the oral pathogen *Tannerella forsythia* and investigate its activity and ability to act on experimental and physiologically relevant diacetylated sialic acid substrates. The data presented here show that NanS from *Tannerella forsythia* is a novel, highly active, neutral sialate-O-acetylesterase that acts to remove 9-O-acetyl groups (and potentially those at 7-O and 8-O positions but not 4-O-acetylated sialic acid) from diacetylated sialic acids in mucinous and epithelial cell glycoproteins and on released highly branched N-glycans from

erythropoietin. We also provide evidence that it can work in concert with its cognate sialidase in release of sialic acid from mucin and oral epithelial cell surfaces, suggesting that *in vivo* it might act to increase sialic acid release by sialidases for *T. forsythia* and other bacteria in its environmental niche.

Materials and Methods

Bacterial cell culture — *Escherichia coli* BL21 (DE3) strains were cultured in Luria-Bertani (LB) broth at 37 °C, or on solid LB medium containing 1.5% bacteriological agar (Oxoid). Selective antibiotics were added to appropriate concentrations (i.e. 50 μg mL^{-1} ampicillin).

Cloning of nanS for expression— The predicted mature *TF0037 (nanS*) open reading frame (i.e. lacking the putative secretion sequence of *TF0037*) from *T. forsythia* genomic DNA was PCR amplified using primers (NanS-GST-Bam-For AAA GGA TCC CAG AAG ACT GTG AAA GTG GC and NanS-GST-Eco-Rev AAA GAA TTC TTA TTC AAT CAC AGC CTG AAA CG) using Phusion Polymerase according to manufacturers instructions. The gene was cloned into pGEX-4T-3 using *Bam*HI and *Eco*R1 to produce pGEX-NanS, and its sequence verified by DNA sequencing (Core Genomic Facility, University of Sheffield)- N.B this revealed 4 differences to the published genome sequence, but none in active site regions which is now known to be from strain *T. forsythia* 92A.2 (ATCC-BAA-2717) (personal communication Floyd Dewhirst- see supplementary Fig. S1 for alignment).

Production of recombinant NanS — *E. coli* BL21 (DE3) Origami B was transformed with the pGEX-NanS plasmid, grown in LB broth and induced at mid-exponential growth phase (A_{600} = 0.6) with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside; Sigma-Aldrich) for five hours at 37 °C with agitation. Cells were harvested, resuspended in purification buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2.5 mM $CaCl_2$), disrupted using a French pressure cell (1050 psi x 3) (Thermo Scientific), and soluble fractions clarified by further centrifugation (40,000 × *g*, 30 minutes, 4 °C). The *N*-terminally glutathione-S-transferase (GST)-tagged protein was purified by applying the cell-free extract on a 1-mL GSTrap column (GE Healthcare) at 1 mL min^{-1}, following which, unbound proteins were removed by washing with 20 column volumes of purification buffer. The GST tag of bound NanS was proteolytically removed by incubating the column in 1 U μL^{-1} thrombin (Sigma-Aldrich) in the purification buffer, sealed and left to stand at room temperature for 5 hours. Cleaved NanS was eluted with two bed volumes of purification buffer, then mixed gently with 50 μL *p*-aminobenzamidine-agarose (Sigma-Aldrich) that had been pre-equilibrated in the same buffer to remove thrombin. The purified protein was extensively dialysed against a dialysis buffer (50 mM sodium phosphate, pH 7.4, 150 mM NaCl) and concentration determined using the Pierce BCA protein assay kit (Thermo Scientific).

Reaction kinetics and pH optimum of purified NanS — The pH optimum of NanS was determined using the chromogenic esterase substrate 4-nitrophenyl acetate (pNP-Ac; Sigma-Aldrich). pNP-Ac by incubating 2.5 nM purified NanS with 0.1 mM pNP-Ac in 20 mM sodium citrate (pH 3.0 – 6.0), 20 mM sodium phosphate (pH 6.4 – 8.8) or 20 mM sodium carbonate (pH 9.2)buffers, and the presence of free pNP measured and expressed as the mean change in absorbance per minute, ± S.D. pNP-Ac hydrolysis was monitored every 15 s by removal of an aliquot of enzyme reaction, followed by heat inactivation (98 °C, 90 s), after which, an equal volume of quenching buffer (2 M sodium carbonate, pH 10.5) was added and the release of free pNP recorded by measuring its absorbance (A_{405}) in a spectrophotometer (Tecan M200).

For K_m determination, purified NanS was pre-incubated at 25 °C for 15 min in 50 mM sodium-potassium phosphate, pH 7.4, 150 mM NaCl, at a final concentration of 2.5 nM. pNP-Ac was then added at a range of final concentrations (2.5 to 1000 μM), and pNP-Ac hydrolysis was monitored as above. Time course data from each concentration of the pNP-Ac substrate were examined for linearity by plotting the data, followed by the calculation of the slope of the tangent. Known concentrations of pNP in the same buffer mixed with an equal volume of quenching buffer was measured as above to obtain a standard curve. Kinetic parameters K_m and V_{max} were calculated by fitting data from three biological replicates in Prism 6 (GraphPad Software) to equation 1,

$$V_0 = V_{max} \cdot Z / (K_m + Z) \qquad (1)$$

in which V_0 is the steady-state rate of pNP-Ac substrate turnover by NanS, V_{max} is the maximum rate ($\mu moles^{-1}min^{-1}mg$ NanS protein^{-1}) Z is the concentration of pNP-Ac (μM) in the reaction mixture and K_m is the Michaelis–Menten constant (μM) at which the reaction rate is half of V_{max}. We also used the equation V_{max}/[E] to determine K_{cat}.

Quantification of sialic acid release from mucin

A modified version of the thiobarbituric acid (TBA) assay (20, 21) was used to assess hydrolysis of sialic acid from mucin. 3μM mucin was incubated with 0.1μM NanH, NanS, or NanH+NanS in PBS pH 7.4 (Sigma Aldrich, UK) for 30 minutes at 37°C, in a total reaction volume of 50μl before addition of 25μl 25mM sodium periodate (Sigma Aldrich, UK) in 500mM H_2SO_4. This was incubated for 30 minutes at 37°C before oxidation of free sialic acid was halted by addition of 2% (w/v) Sodium Arsenite (Sigma Aldrich, UK) in 500mM HCl. Formation of the red chromophore was completed by the addition of 200μl 100mM Thiobarbituric Acid (Sigma Aldrich, UK) pH 9 and incubated for 7.5 minutes at 95°C. The chromophore was extracted by addition of 500μl acidified butanol (butan-1-ol with 5% (v/v) 600mM HCl), before vortexing and centrifugation at 10000xg for 3 minutes. 100μl of the butanol phase was then added to a clear 96 well plate (Greiner, UK), and the absorbance at 549nm was measured using a Tecan infinite M200 plate reader. Free sialic acid (carbosynth) was used to generate a standard curve.

Lectin staining of oral epithelial cells

The oral squamous carcinoma cell line H357 was cultured in T175cm^2 tissue culture flasks (Greiner), in Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich) supplemented with 10% v/v Foetal Bovine Serum (FBS, Sigma-Aldrich), and 2mM L-glutamine (Sigma-Aldrich, UK), 100 U/ml penicillin and 100μg/ml streptomycin (Sigma-Aldrich, UK). Cells were seeded at 100,000 cells/ml into 24 well tissue culture plates (Greiner) containing glass coverslips (VWR, UK), with 1ml of cell suspension/well, and incubated at 37°C in 5% CO_2 for 14-24 hours. To stain for cellular Neu5,9Ac_2, cells were incubated with either PBS or 50nM NanS in PBS for 2 hours at 37°C,washed twice with 500μl PBS and fixed using 2% paraformaldehyde for 15 minutes at 37°C., fixed cells were washed twice with 500μl PBS, incubated with 2μg/ml biotinylated *Cancer antennarius* (CCA) or Sambucus Nigra (SNA) lectin (EY labs) for 30 minutes at 37°C, washed twice with 500μl PBS, then stained with 2μg/ml streptavidin-FITC (Vector Labs, UK) for 30 minutes at 37°C. To stain for cellular α2-6 linked Neu5Ac, cells were incubated with either PBS, 50nM NanS, 50nM NanH or 50nM NanH and 50mM NanS, all in PBS, for 30 minutes at 37°C, washed twice with 500μl PBS and fixed using 2% paraformaldehyde for 15 minutes at 37°C. Coverslips with stained cells were mounted on glass slides using Prolong Gold Antifade mountant with DAPI (Life Technologies, UK). Cells were visualized using an Axiovert 200M fluorescence microscope (Zeiss, Germany) and Axiovision software. Fiji/ImageJ software was used to process images with three fields of view from two biological replicates assessed (23). Neu5,9Ac2 staining was expressed as the mean pixel brightness(FITC staining only)/number of cells=Mean fluorescence Intensity (MFI) per cell. FITC image acquisition was carried out using the same exposure time and % weight for each condition, and parameters during image analysis were kept constant between conditions.

Treatment of diacetylated sialic acids with NanS

A Neu5,9Ac_2 (Ludger Ltd, Oxford, UK), Neu5,4Ac_2 (Carbosynth UK) and sialic acid reference panel (SRP) (derived from bovine submaxillary mucin, Ludger Ltd Oxford UK) were incubated with 1μL of NanS enzyme (0.7mg/ml) in a final volume of 10μL (1× PBS, pH 7.2) for 16 hours at 25°C. Sialic acid standards (1μg of each sialic acid) were labelled with 1,2-diamino-4,5-methylenedioxybenzene.2HCl (DMB) using a Ludger DMB sialic acid release and labelling kit according to product guide. Briefly, dried standards were incubated with 20μL of labelling reagent in the dark for 3 hours at 50°C. The reaction was terminated upon the addition of 480μL of water. The DMB labelled samples were analysed by reverse phase HPLC using a LudgerSep-R1 column (4.6 x 150 mm, Ludger Ltd) at 30°C on a Waters 2795 HPLC with a 2475 fluorescence detector (λex = 373 nm, λem = 448 nm), controlled by empower software (Waters, Manchester, UK). Isocratic run conditions were at a flow rate of 0.5 mL/min with solvent acetonitrile: methanol: water (9:7:84 by volume). 10μL sample were injected undiluted onto the HPLC. The SRP standard was used as a system suitability standard as well as an external calibration standard to allocate sialic acid structures.

Treatment of N-glycans from rh-EPO with NanS

N-glycans were released from 10μg of recombinant human EPO expressed from Chinese hamster ovary (CHO) cells (gift from Antonio Vallin, Center for Molecular Immunology, Cuba) using PNGase F (E-PNG01; Ludger Ltd, UK). The EPO (in 17.5μL) was denatured following addition of 6.25 μL of 2% SDS, 1M beta-mercaptoethanol by heating at 100°C for 5 minutes, before incubation at 37°C for 16 hours with 1μL PNGaseF and 1.25μL 15% w/v Triton X-100. Released N-glycans were fluorescently labelled with 2-aminobenzamide (2-AB) according to Bigge *et al.* (24) using a Ludger 2-AB Glycan Labelling Kit. The released glycans were incubated with labelling reagents for 3 hours at 65°C. The 2-AB labelled glycans were cleaned up using LudgerClean S Cartridges (Ludger Ltd). NanS enzyme digestions were performed according to Royle *et al.* (25). The 2-AB labelled N-glycans were incubated with 1μL of NanS enzyme (0.7mg/mL) in a final volume of 10μL (50 mM sodium acetate buffer, pH 5.5) for 16 hours at 37°C. NanS was removed by binding onto a protein-binding plate LC-PBM-96 (Ludger Ltd) before analysis by HILIC-UPLC. 2-AB labelled samples were analysed by HILIC-ULPC using an ACQUITY UPLC® BEH-Glycan 1.7μm, 2.1 x 150 mm column at 40°C on an ACQUITY UPLC H Class instrument with a fluorescence detector (excitation 250 nm and emission 428nm), controlled by Empower software version 2, build 2154 (Waters, Manchester, UK). Solvent A was 50mM ammonium formate pH 4.4, Solvent B was acetonitrile. Samples were injected in in 28% aqueous/72% acetonitrile; injection volume 25 μL. Gradient conditions were 0 to 54 min, 28 to 48 % A; 54 to 57 min, 48 to 100% A; 57 to 58 min, 100% A; 58 to 59 min, 100 to 28% A; 59 to 60 min, 28% A; at a flow rate of 0.4mL/min. Waters GPC software with a cubic spline fit was used to allocate GU values to peaks. 2-AB labelled glucose homopolymer (CAB-GHP-30, Ludger Ltd) was used as a system suitability standard as well as an external calibration standard for GU allocation (24).

Basic bioinformatic analysis

The *nanS* gene sequence and its product NanS was analysed for the presence of a signal sequence using PSORTb and Cello. Alignments were generated using Multalin (26). Accession numbers for genes used are: *T. forsythia* 92A.2 (Accession WP_014225512.1), *Bacteroides fragilis* (B. fragilis- WP_005808991.1), *Prevotella denticola* (P. dent- WP_013672182.1), *Parabacteroides distastonis* (P. dista- WP_011967074.1), *Capnocytophaga sputigena* (C. sputi- WP_002678445.1) (Accession), Human SIAE1 enzyme (Hs SIAE1) (Accession), *Escherichia coli* O157:H7 (E. coli- WP_000991449), Influenza C (P07975).

RESULTS

NanS is a predicted sialate-O-acetylesterase containing putative SGHN type I and II domains

In our initial work on the sialic acid utilization operon of *T. forsythia* we noted the presence of a potential sialate-O-acetylesterase based on BLAST search data (6). The predicted size of the full NanS protein is 78.6 kDa, however it has a strong putative N-terminal secretion signal with a potential cleavage site after Alanine 18 (P-SORT) in the primary amino acid sequence meaning the size of the mature protein would be 76.3 kDa. Therefore this protein at the very least enters the periplasm, but we assume, given its potential role in sialic acid harvesting that it is secreted from the cell, similar to the NanH sialidase of this organism (27).

A more detailed bioinformatics analysis of the NanS amino acid sequence of *T. forsythia* indicates the presence of two distinct domains, both of which are related to different classes of SGNH serine-hydolase family of enzymes (Fig.1A)(28). The SGNH-hydrolases are a group of enzymes responsible for hydrolysis of a range of substrate types that includes both carbohydrate esterases and lipases where the catalytic action in both cases involves hydrolysis of ester bonds via use of a catalytic serine residue in the active site (28). Within the primary amino acid sequence of the *T. forsythia* NanS, residues 1-180 of the mature protein have homology with members of the SGNH-hydrolase superfamily with all the conserved sequence features present (Fig. 1B)(29). It is worth noting that, at least in the signature sequence blocks, that this domain can be aligned with viral sialate-O-acetylesterase active site residues such as those identified for the Haemagglutinin Esterase from Influenza C virus but that homology outside of this region is very low (30, 31).

In contrast, residues 260-564 seem to be related to the DUF303 pfam grouping (PF03629) whose founder member is the *E. coli* NanS sialate-O-acetylesterase(29, 32). All SGNH proteins contain several signature motifs with the presence of the invariant residues within conserved blocks spread throughout the amino acid sequence: Serine (block I), Glycine (block II), Asparagine (block III) and a catalytic Histidine (block V)(25). However, recent work by Rangarajan et al., (29) identified two types of SGNH hydrolases, based on differences in the signature sequences in blocks I,II,III and V between classic SGNH hydrolases and the DUF303 type, which they termed Type I and II. As indicated in Figure 1 the NanS from *T. forsythia* displays a domain architecture not reported in the SIAE family to date, i.e. contains both SGNH family I and II domains linked within one protein.

As mentioned above the c-terminal domain of the protein has strong homology the type II SGNH enzymes identified by Rangarajan *et al.,* (29) (Fig. 1A). However, while it is possible to align the NanS from *T. forsythia* and other *Bacteroidetes* based on the sequence blocks first characterized by Rangarajan *et al.* (29)(Fig. 1B) with high levels of conservation, the identity with the *E. coli* sequence that is the founder member of the group is low. When one examines the presence of the conserved sequence blocks it is clear that the *Bacteroidetes* NanS C-terminal domains contain the signature extended GQSN motif (block I), the conserved Glycine in block II, which actually seems to be part of a larger GGS motif and the QGES motif of Block III. Unlike *E. coli*, which lacks an obvious N residue within the SGNH schema, both the *Bacteroidetes*, *Capnocytophaga* and the Human sequence contain a putative catalytic N residue (N403 in *T. forsythia*), however its role in these enzymes is unclear given its absence from *E. coli* NanS, which is fully active. One absentee from the *Bacteroidetes, Capnocytophaga* and Human sequences however is an obvious catalytic histidine (His 301 within RSSH motif in *E. coli*), although we suggest that the DVH motif that is conserved across the *Bacteroidetes* group might represent a variant on the DXXH present in group I SGNH hydrolases, although this has not yet been confirmed experimentally by point mutagenesis or other means. It is also of note that at least in the conserved blocks these bacterial domains also align well with the human SIAE gene (Fig.1) (30), indicating that this GQSN containing SIAE domain may well represent a common catalytic domain for SIAE activity in biology.

Purified NanS enzyme is a highly active sialate-o-acetylesterase

To test our hypothesis that *nanS* encodes an esterase enzyme with specificity for acetyl groups contained within diacetylated sialic acids (i.e. Neu5,9Ac_2 or Neu5,4Ac_2) we first investigated whether purified NanS was active against the model and widely used esterase substrate p-Nitrophenol-acetate (pNP-Ac). To

achieve this we cloned NanS into a GST-fusion vector and expressed it in *E. coli* before purification via GST-affinity chromatography and release of the soluble untagged protein via thrombin cleavage of the NanS protein from its GST fusion partner (Supplementary Fig. 2).

When pNp-Ac was incubated with NanS protein we observed rapid cleavage of this substrate via measurement of release of free pNP at 405nm (not shown). In order to more fully characterize activity of NanS we then assessed its ability to cleave pNP-Ac over a range of pH values by incubating NanS (25nM) in the presence of 0.1mM pNP-Ac. As illustrated in Fig 2B NanS is active over a wide range of pH values (6-8.4) with an optimum in the neutral range of 7.2-7.6. Purified NanS was also highly stable, with only marginal loss of activity when the enzyme was incubated at ambient temperature for 72 hours (not shown). We then set out to establish the kinetic parameters of the enzyme with this pNP-Ac by assessing initial reaction velocity (V_0, $\mu moles^{-1}min^{-1}mg\ NanS^{-1}$) over a range of substrate concentrations as illustrated in Fig 2C. These data reveal that the enzyme seems to obey Michaelis-Menten type, one-step binding kinetics with the pNP-Ac substrate over the range of concentrations used with a K_m or 51.2 μM (±1.9 μM) and a V_{max} of 166.6 $\mu moles^{-1}min^{-1}mg\ NanS^{-1}$ (±0.004 μM). In addition we calculated K_{cat} at 111 s^{-1} and a specificity constant (K_{cat}/K_m) of 2.2 $x10^6\ M^{-1}s^{-1}$. Despite the presence of two potential catalytic domains we found no evidence, using pNP-acetate as substrate, that any co-operativity between these potential domains existed from a Hill plot of the data (i.e. log(v/vmax-v)) which revealed a Hill number 0.97 (data not shown).

In order to assess whether NanS in fact acted upon its hypothesized *in vivo* substrate, albeit in its monomeric unconjugated form, i.e. the diacetylated sialic acid of Neu5,9Ac we performed hydrolysis reactions by incubating NanS (1 μL, 0.7 mg mL^{-1}) with (DMB-) Neu5,9Ac standards at 25°C overnight, i.e. to completion. These sialic acids were then labelled with the fluorophore DMB before analysis by reverse-phase-UHPLC. The data revealed that the Neu5,9Ac was completely transformed into Neu5Ac by NanS (Fig. 3). It is also of note that the Neu5,9Ac_2 standard contained a small amount of another contaminating peak in the standard trace which are also degraded by NanS (indicated in middle panel of Fig. 3) which we propose are Neu5,8Ac_2 and Neu5.7Ac_2. In order to establish whether NanS acted upon the other prominent mammalian sialic acid Neu5,4Ac_2 we performed a similar experiment with a pure preparation of Neu5,4Ac and revealed that no degradation of the Neu5,4Ac peak at approximately 16.5mins occurred and indicating that NanS does not act upon this sugar.

To investigate the substrate specificity more thoroughly we incubated NanS with a Sialic acid reference panel sample that contains a range of di-acetylated NeuAc and NeuGc monosaccharides (Fig 4). These data revealed that NanS caused the disappearance of not only Neu5,9Ac_2 peaks but also Neu5Gc,9Ac alongside Neu5,7Ac_2 and Neu5,8Ac_2 peaks- although Neu5,7 and Neu5,8Ac_2 may not be acted upon directly but rather reflect transesterification occurring in the samples. At this point we have no knowledge of its action on triacetylated forms.

NanS potentiates activity of Tannerella sialidase activity on physiologically relevant substrates

Based on genomic data, our hypothesis was that the role of NanS *in vivo* is to aid in the efficient harvest of sialic acid from human glycoproteins in the oral cavity for uptake and utilization by *T. forsythia*. We therefore designed an experiment to test whether NanS enhanced the action of the *T. forsythia* NanH sialidase enzyme in removing sialic acids from the highly sialylated and diacetylated substrate, bovine submaxillary mucin where 25-80% of sialic acid is 9-O-acetylated depending on the batch (10, 33, 34) and which we had previously shown was a less efficient substrate for NanH than a non-diacetylated substrate such as fetuin (9).

As illustrated in Fig. 5A, incubation of mucin with NanS alone resulted in no release of sialic acid as measured by TBA assay (22). However, on the addition of 100 nM NanH only 90 pmoles min^{-1} was released per minute, while addition of equimolar amounts of NanS (100 nM) resulted in the amount of sialic acid released increasing over 4-fold, indicating that NanS allows more access to sialic acid for cleavage by NanH.

In addition these data reveal indirectly that NanS is able release sialic acid from an intact glycoprotein containing Neu5,9Ac and other variants.

NanS acts upon gingival cell associated diacetylated sialic acid

To further test the ability of NanS to de-acetylate sialic acids from protein substrates we also tested the ability of NanS to cleave acetyl groups from the surface of oral epithelial cells by incubating NanS with H357 oral squamous cell carcinoma (tongue) cells, followed by staining with the diacetylated sialic acid specific lectin CCA (from Cancer Antennarius, green channel). Fig. 4B shows that upon addition of NanS CCA specific Green-staining is lost from the surface of H357 cells. When this green fluorescence was measured in relation to the number of cells, this visual observation was corroborated (Fig. 5B and C), further illustrating the ability of NanS to access potential *in vivo* substrates on relevant human cell lines.

To further elucidate the potential role for NanS *in vivo* in interactions with sialic acids on the surface of human cells we next asked whether NanS would increase removal of sialic acid from human oral epithelial cells by its cognate sialidase NanH (much like the mucin experiment outlined above). In this case we performed staining of the cells with the alpha 2,6-linkage specific Neu5Ac lectin SNA and then treated cells with NanS or NanH alone or NanH plus NanS in concert, revealing an increase in sialic acid release from the surface of these cells as assayed by means of mean fluorescence per cell (Fig. 5D and E).

NanS activity against total N-glycans from a heavily glycosylated human protein

We next tested the ability of NanS to remove acetyl groups from diacetylated sialic acids from a well-characterised human glycoprotein, namely recombinant human Erythropoietin (rH-EPO). Following digestion of the 2-AB labelled N-glycans released from rH-EPO with NanS *in vitro* overnight at 37°C samples were analysed by HILIC-UPLC. We observed that a number of the peaks either disappear (peaks 10, 19 and 23a in Fig. 6; Supplementary Table 1) or reduce following treatment with NanS enzyme. When the glycan structure of these peaks was assessed from the undigested samples by Mass-Spectrometry they were revealed to contain 1 O-Acetyl group within the 3 or 4 terminal sialic acid residues of these tri- and complex tetrantennary glycan structures (see Figure 6 upper panel, Supplementary table 2). Since NanS is unable to act upon Neu5,4Ac_2 we can infer that these peaks contain Neu5,9 (7 or 8) and that NanS can act upon native complex human type N-glycans such as those known to be present in rH-EPO (20).

DISCUSSION

In this paper we report the initial characterization of a novel type of sialate-acetylesterase from the oral pathogen *Tannerella forsythia*. Activity of the purified enzyme, NanS, was high with the model esterase substrate pNP-Acetate where the enzyme displayed a neutral pH optimum, which one might expect given the physiological environment in which *Tannerella* finds itself *in vivo*, i.e. in the periodontal pocket, where pH is known to be neutral to alkaline (35). This pH optimum is typical of other sialate-O-acetylesterases from both viral and mammalian sources (36). It displays a similarly high affinity (μM) for pNP-Ac when compared to other sialate-O-acetylesterases from viral sources (37, 38), but differs in that it is able, based on evidence presented here, to target mucin conjugated diacetylated sialic acids, in-keeping with the potential role of *T. forsythia* NanS in enabling release of sialic acids from mucinous sources in their environment.

We also established that NanS is able to hydrolyse acetyl groups from a diacetylated Neu5,9Ac monosaccharide standard but not Neu5,4Ac_2, as assayed using DMB labelled derivatives (Fig. 3). Both the main Neu5,9Ac_2 peak and two additional peaks that are Neu5,8Ac_2 and Neu5,7Ac_2 (15) disappeared after incubation of monosaccharide with NanS. Although it is unclear whether the enzyme acts directly on Neu5,8Ac or Neu5,7Ac_2 because these exist in equilibrium with Neu5,9Ac_2 and can spontaneously convert to Neu5,9Ac_2 under the neutral conditions and long incubation times used here (39). Thus we propose that NanS is likely to be specific for Neu5,9Ac_2 although we cannot rule out Neu5,7 and 8Ac-specific activity. In this study we also show activity of NanS against both cell-bound glycoproteins, as evidenced by loss of CCA-lectin staining on H357 cells and mucinous secretions alongside released human N-glycans. It is also of note that NanS can act upon the Glycolyl containing Neu5Gc,9Ac, a form of sugar that is prevalent in

mammalian hosts but also likely to be present in secretions and cell surface glycans in the oral cavity due to cross-feeding of oral glycans from dietary sources.

In addition we show that NanS is active in de-acetylation of di-acetylated sialic acids on three highly branched triantennary and tetrantennary glycans (FA3G3S3Ac1, FA4G4S4Ac1 and FA4G4Lac1S4Ac1) from the important pharmaceutical protein erythropoietin (EPO), which is heavily N-glycosylated and contains a high amount of terminally sialylated glycans including a significant amount of diacetylated sialic acids (40, 41). Several brands of EPO are produced in Chinese hamster ovary (CHO) cells which have been identified as capable of acetylating sialic acids with up to 6 extra acetyl groups in this biopharamaceutical (20). The NanS enzyme could be a useful tool in identifying the structures present on recombinant EPO with respect to the presence of different variants of di-acetylated sialic acids, as changes in the glycan fingerprints for EPO need to be monitored for batch to batch consistency for biopharmaceutical productions and in identification of microheterogeneity in glycan fingerprints of EPO produced from different sources that are important in monitoring anti-doping activity(41, 42). Taken together, while we do not have kinetic data for these complex substrates it is clear from our evidence that NanS can act on physiologically relevant complex N- and O-glycans in free (monosaccharide) as protein-conjugated or free glycans.

Our hypothesis at the beginning of this study was that the sialate acetylesterase activity of NanS contributed to opening up the accessibility of Sialic acid for cleavage by the *T. forsythia* sialidase NanH. Thus it was reassuring to find that when the NanH sialidase is co-incubated with NanS its ability to catalyse the release of sialic acid from mucin was increased significantly (Fig. 5). This is not only of importance to *T. forsythia* where this ability to access diacetylated sialic acid means *T. forsythia* maximizes its ability to harvest and utilize sialic acid in its environmental niche in the oral cavity, but also to the oral microbial community. For example, while several cohabiting organisms possess sialidases enzymes, namely *Porphyromonas gingivalis* and *Treponema denticola*, they do not possess NanS homologues raising the possibility that *Tannerella* may contribute to overall community ability to access sialic acids, i.e. a "public-good" phenotype. Our data also hint that NanS may contribute to human cell interactions since they both access human glycan diacetylated sialic acid and potentiate sialidase access when incubated with human cells. This co-activity of a sialate esterase and sialidases is not unprecedented in biology, since there is evidence that the influenza C virus not only uses sialate-O-acetylesterase as a receptor destroying enzyme but also to enhance the activity of its neuraminidase but certainly this is the first example in the context of bacteria in oral cavity community (43). At a mechanistic level our previous work has suggested that the Nan operon of *T. forsythia* should be considered a sialic acid specific PUL that contains catabolic, transport and harvesting genes (3, 5). It is therefore tempting to speculate that NanS might interact with either or both the NanH sialidase and NanOU transport system either being localized on the cell surface or in a secreted form, as seems to be the case for NanH and NanU (3). These questions are currently under investigation in the laboratory.

Another biological consequence of NanS activity on top of its potential importance to the nutritional physiology of *T. forsythia* is that the ability to produce a sialate-O-acetylesterase might influence the immune response to this organism. Specifically there is evidence that conversion of Neu5,9Ac_2 to Neu5Ac on the surface of B-cells influences several autoimmune diseases (16), leading one to speculate that this might also be occurring in periodontal disease where alterations in immune responses are known to influence the local symptoms of the disease but also its systemic sequelae (44, 45). A role for the influence of diacetylation of sialic acids in cancer is also documented, although how this might be influenced by bacterial enzymes either in the oral cavity or elsewhere is not clear (18, 46).

As mentioned above, analysis of the predicted domain structure of NanS, suggests a novel domain architecture that differs from the only other characterized bacterial sialate-O-acetylesterase, namely the NanS from *E.coli* (29, 32). In short NanS from *T. forsythia* contains two putative sialate-o-acetylesterase domains, one of the SGNH Type I and one of the SGNH Type II, proposed recently in relation to NanS from *E.coli* (29). This domain architecture seems to be present in a range of bacteria of the phylum

Bacteroidetes, which includes *Prevetolla, Capnocytophaga, Bacteroides* and *Parabacteroides spp.* among others, but not outside of this group. The Type I SGNH sialate esterase domains contain the invariant SGNH residues with the *Bacteroidetes* sequences clustering closely but also sharing these with viral enzymes. In contrast the Type II SGNH sialate o-acetylesterases contain conserved sequence motifs that are shared across sialate esterases from *Bacteroidetes*, *Enterobacteriaciae* and humans, namely GQSN and QGES (blocks I and III). However, many questions remain, namely which is the catalytic histidine in the *Bacteroidetes* Type II domain, and more importantly what the role of the two putative acetylesterase domains is in interaction with and hydrolysis of di-acetylated sialic acids.

In summary we have begun to characterize a novel prototype member of a two-domain class of sialate-O-acetylesterases that may play a role in both the *in vivo* physiology of human pathogens and commensals in the oral cavity and potentially the gut. This ability to harvest diacetylated sialic acids might confer an advantage to *T. forsythia* and other oral organisms in harvesting of sialic acid for nutrition. In addition there is much evidence that sialic acid is key to the interaction of many human bacterial and viral pathogens with human cells in both epithelial layers (27) but also the immune system and one assumes that possessing sialate-O-acetylesterase activity contributes to this process. Whether this hypothesis is true or not it is likely that oral bacteria encounter diacetylated sialic acids in abundance with both secreted and membrane bound mucins containing significant amounts of diacetylated sialic acids but also the presence of diacetylated sialic acid epitopes present on oral epithelial cells as evidenced by CCA lectin staining and co-incubations with sialidases reducing SNA staining. One observation in this study is that the enzyme characterized here bears strong homology with other putative enzymes encoded in the chromosomes of several gut-dwelling *Bacteroidetes,* such as *B. fragilis,* indicating that this ability to harvest diacetylated sialic acids is also be important in the mucin rich environment that is the human gut (12) as well as the oral cavity.

Acknowledgements
Work in the GS lab was funded by a Dunhill medical trust grant (DMT-R185/0211) and a BBSRC CASE award (BB/K501098/1, AF). The authors thank Antonio Vallin (Center for Molecular Immunology, Cuba) for the erythropoietin (EPO) sample and K Honma and Ashu Sharma for the NanH sequence from *T.forsythia* 43037. We also declare that Ludger and Dr Stafford have a licensing agreement centred on use of NanS commercially.

Author contributions
Chatchawal Phansopa- Performed protein biochemistry and purifications,
Radoslaw P Kozak- Performed Glycan analysis of EPO with Louise Royle, Robert J Emery and Richard Gardner
Phing Liew – performed DMB analyses of NanS incubated sialic acid standards.
Daniel Spencer- designed Ludger's studies for this paper.
Rebecca Thomson- performed and designed analysis on EPO glycan identification
Andrew M Frey –performed/ developed NanH/S dual enzyme assays, cell culture and fluorescent microscopy
Jennifer L Parker- performed/ developed NanH/S dual enzyme assays
Thomas Farmilo- cloned NanS and established purification protocols and performed cellular CCA staining
David J Kelly- advised on biochemical assays
Graham P. Stafford- designed the study and wrote the paper
All authors contributed to the preparation of the manuscript.

1. Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C., and Kent, R. L. (1998) Microbial complexes in subgingival plaque. *J. Clin. Periodontol.* **25**, 134–44

2. Xu, J., Mahowald, M. A., Ley, R. E., Lozupone, C. A., Hamady, M., Martens, E. C., Henrissat, B., Coutinho, P. M., Minx, P., Latreille, P., Cordum, H., Van, B. A., Kim, K., Fulton, R. S., Fulton, L. A.,

Clifton, S. W., Wilson, R. K., Knight, R. D., and Gordon, J. I. (2007) Evolution of symbiotic bacteria in the distal human intestine. *PLoS.Biol.* **5**, e156

3. Douglas, C. W. I., Naylor, K., Phansopa, C., Frey, A. M., Farmilo, T., and Stafford, G. P. (2014) Physiological adaptations of key oral bacteria. *Adv. Microb. Physiol.* **65**, 257–335

4. Stafford, G., Roy, S., Honma, K., and Sharma, A. (2011) Sialic acid, periodontal pathogens and Tannerella forsythia: stick around and enjoy the feast! *Mol. Oral Microbiol.* **26**, 11–22

5. Phansopa, C., Roy, S., Rafferty, J. B., Douglas, C. W. I., Pandhal, J., Wright, P. C., Kelly, D. J., and Stafford, G. P. (2014) Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species. *Biochem. J.* **458**, 499–511

6. Roy, S., Douglas, C. W. I., and Stafford, G. P. (2010) A novel sialic acid utilization and uptake system in the periodontal pathogen Tannerella forsythia. *J. Bacteriol.* **192**, 2285–93

7. Martens, E. C., Koropatkin, N. M., Smith, T. J., and Gordon, J. I. (2009) Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *J. Biol. Chem.* **284**, 24673–7

8. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–5

9. Roy, S., Honma, K., Douglas, I., Sharma, A., Stafford, G. P., Douglas, C. W. I., Sharma, A., and Stafford, G. P. (2011) Role of sialidase in glycoprotein utilisation by Tannerella forsythia. *Microbiology* **157**, 3195–3202

10. Varki, A., and Diaz, S. (1983) A neuraminidase from Streptococcus sanguis that can release O-acetylated sialic acids. *J. Biol. Chem.* **258**, 12465–71

11. Corfield, T. (1992) Bacterial sialidases--roles in pathogenicity and nutrition. *Glycobiology* **2**, 509–21

12. Corfield, A. P., Wagner, S. A., Clamp, J. R., Kriaris, M. S., and Hoskins, L. C. (1992) Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. *Infect. Immun.* **60**, 3971–8

13. Rogers, G. N., Herrler, G., Paulson, J. C., and Klenk, H. D. (1986) Influenza C virus uses 9-O-acetyl-N-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. *J. Biol. Chem.* **261**, 5947–51

14. Shahwan, K., Hesse, M., Mork, A.-K., Herrler, G., and Winter, C. (2013) Sialic acid binding properties of soluble coronavirus spike (S1) proteins: differences between infectious bronchitis virus and transmissible gastroenteritis virus. *Viruses* **5**, 1924–33

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

16. Pillai, S., Cariappa, A., and Pirnie, S. P. (2009) Esterases and autoimmunity: the sialic acid acetylesterase pathway and the regulation of peripheral B cell tolerance. *Trends Immunol.* **30**, 488–93

17. Pal, S., Chatterjee, M., Bhattacharya, D. K., Bandhyopadhyay, S., and Mandal, C. (2000) Identification and purification of cytolytic antibodies directed against O-acetylated sialic acid in childhood acute lymphoblastic leukemia. *Glycobiology* **10**, 539–549

18. Ghosh, S., Bandyopadhyay, S., Mukherjee, K., Mallick, A., Pal, S., Mandal, C., Bhattacharya, D. K., and Mandal, C. (2007) O-acetylation of sialic acids is required for the survival of lymphoblasts in childhood acute lymphoblastic leukemia (ALL). *Glycoconj. J.* **24**, 17–24

19. Thompson, H., Homer, K. a, Rao, S., Booth, V., and Hosie, A. H. F. (2009) An orthologue of Bacteroides fragilis NanH is the principal sialidase in Tannerella forsythia. *J.Bacteriol.* **191**, 3623–3628

20. Shahrokh, Z., Royle, L., Saldova, R., Bones, J., Abrahams, J. L., Artemenko, N. V., Flatman, S., Davies, M., Baycroft, A., Sehgal, S., Heartlein, M. W., Harvey, D. J., and Rudd, P. M. (2011) Erythropoietin produced in a human cell line (Dynepo) has significant differences in glycosylation compared with erythropoietins produced in CHO cell lines. *Mol. Pharm.* **8**, 286–296

21. Aminoff, D. (1961) Methods for the quantitative estimation of N-acetylneuraminic acid and their application to hydrolysates of sialomucoids. *Biochem. J.* **81**, 384–92

22. Romero, E. L., Pardo, M. F., Porro, S., and Alonso, S. (1997) Sialic acid measurement by a modified Aminoff method: a time-saving reduction in 2-thiobarbituric acid concentration. *J. Biochem. Biophys. Methods* **35**, 129–34

23. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–82

24. Bigge, J. C., Patel, T. P., Bruce, J. A., Goulding, P. N., Charles, S. M., and Parekh, R. B. (1995) Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal. Biochem.* **230**, 229–38

25. Royle, L., Mattu, T. S., Hart, E., Langridge, J. I., Merry, A. H., Murphy, N., Harvey, D. J., Dwek, R. A., and Rudd, P. M. (2002) An analytical and structural database provides a strategy for sequencing O-glycans from microgram quantities of glycoproteins. *Anal. Biochem.* **304**, 70–90

26. Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890

27. Honma, K., Mishima, E., and Sharma, A. (2011) Role of Tannerella forsythia NanH sialidase in epithelial cell attachment. *Infect. Immun.* **79**, 393–401

28. Mølgaard, A., Kauppinen, S., and Larsen, S. (2000) Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. *Structure* **8**, 373–83

29. Rangarajan, E. S., Ruane, K. M., Proteau, A., Schrag, J. D., Valladares, R., Gonzalez, C. F., Gilbert, M., Yakunin, A. F., and Cygler, M. (2011) Structural and enzymatic characterization of NanS (YjhS), a 9-O-Acetyl N-acetylneuraminic acid esterase from Escherichia coli O157:H7. *Protein Sci.* **20**, 1208–19

30. Pleschka, S., Klenk, H. D., and Herrler, G. (1995) The catalytic triad of the influenza C virus glycoprotein HEF esterase: Characterization by site-directed mutagenesis and functional analysis. *J. Gen. Virol.* **76**, 2529–2537

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

32. Steenbergen, S. M., Jirik, J. L., and Vimr, E. R. (2009) YjhS (NanS) is required for Escherichia coli to grow on 9-O-acetylated N-acetylneuraminic acid. *J.Bacteriol.* **191**, 7134–7139

33. Chatterjee, M., Sharma, V., Mandal, C., Sundar, S., and Sen, S. (1998) Identification of antibodies directed against O-acetylated sialic acids in visceral leishmaniasis: Its diagnostic and prognostic role. *Glycoconj. J.* **15**, 1141–1147

34. Schauer, R. (1982) Chemistry, Metabolism, and Biological Functions of Sialic Acids. *Adv. Carbohydr. Chem. Biochem.* **40**, 131–234

35. Kobayashi, K., Soeda, W., and Watanabe, T. (1998) Gingival crevicular pH in experimental gingivitis and occlusal trauma in man. *J. Periodontol.* **69**, 1036–43

36. Garcia-Sastre, A., Villar, E., Manuguerra, J. C., Hannoun, C., and Cabezas, J. A. (1991) Activity of influenza C virus O-acetylesterase with O-acetyl-containing compounds. *Biochem. J.* **273(Pt 2)**, 435–41

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

38. Wurzer, W. J., Obojes, K., and Vlasak, R. (2002) The sialate-4-O-acetylesterases of coronaviruses related to mouse hepatitis virus: a proposal to reorganize group 2 Coronaviridae. *J. Gen. Virol.* **83**, 395–402

39. Higa, H. H., Manzi, A., and Varki, A. (1989) O-Acetylation and de-O-acetylation of sialic acids. Purification, characterization, and properties of a glycosylated rat liver esterase specific for 9-O-acetylated sialic acids. *J. Biol. Chem.* **264**, 19435–19442

40. Hokke, C. H., Bergwerff, A. A., Van Dedem, G. W., Kamerling, J. P., and Vliegenthart, J. F. (1995) Structural analysis of the sialylated N- and O-linked carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells. Sialylation patterns and branch location of dimeric N-acetyllactosamine units. *Eur. J. Biochem.* **228**, 981–1008

41. Goh, J. S. Y., Liu, Y., Liu, H., Chan, K. F., Wan, C., Teo, G., Zhou, X., Xie, F., Zhang, P., Zhang, Y., and Song, Z. (2014) Highly sialylated recombinant human erythropoietin production in large-scale perfusion bioreactor utilizing CHO-gmt4 (JW152) with restored GnT I function. *Biotechnol. J.* **9**, 100–9

42. Harazono, A., Hashii, N., Kuribayashi, R., Nakazawa, S., and Kawasaki, N. (2013) Mass spectrometric glycoform profiling of the innovator and biosimilar erythropoietin and darbepoetin by LC/ESI-MS. *J. Pharm. Biomed. Anal.* **83**, 65–74

43. Muñoz-Barroso, I., García-Sastre, a, Villar, E., Manuguerra, J. C., Hannoun, C., and Cabezas, J. a (1992) Increased influenza A virus sialidase activity with N-acetyl-9-O-acetylneuraminic acid-containing substrates resulting from influenza C virus O-acetylesterase action. *Virus Res.* **25**, 145–53

44. Koziel, J., Mydel, P., and Potempa, J. (2014) The link between periodontal disease and rheumatoid arthritis: an updated review. *Curr. Rheumatol. Rep.* **16**, 408

45. Settem, R. P., Honma, K., Stafford, G. P., and Sharma, A. (2013) Protein-linked glycans in periodontal bacteria: prevalence and role at the immune interface. *Front. Microbiol.* **4**, 310

46. Fukuda, M. (1991) Leukosialin, a major O-glycan-containing sialoglycoprotein defining leukocyte differentiation and malignancy. *Glycobiology* **1**, 347–356

Figure Legends:

Figure 1. Basic bioinformatic analysis of *T. forsythia* NanS
A. Domain architecture of NanS, illustrating predicted signal sequence (light grey), SGNH domain (type I, dark grey) and DUF303 domain (Black) (to scale).
B. Amino acid sequence alignments of the conserved sequence blocks of the SGNH Type I and DUF303 domains of confirmed and putative sialate-O-acteylesterases from a range of organisms: *T. forsythia* 92A.2 (Accession WP_014225512.1), *Bacteroides fragilis* NCTC 9343 (B. fragilis) (Accession), *Prevotella denticola* (P. dent- WP_013672182.1) (Accession), *Parabacteroides distastonis* (P. dista) (Accession), *Capnocytophaga sputigena* (C. sputi- WP_002678445.1) (Accession), Human SIAE1 enzyme (Hs SIAE1) (Accession), Escherichia coli O157:H7 (E. coli) (Accession), Influenza C (P07975.1). Residue numbers from *T. forsythia* 92A.2 shown.

Figure 2.
B. pH optimum of NanS with p-nitrophenol-acetate (pNP-Ac, 0.1mM) using 2.5nM NanS. Released pNP was measured at 405nM. A combination of buffers was used to assay across pH ranges with reactions quenched with 1M Na_2CO_3. Experiments were performed with three separate protein preparations each in triplicate, Mean and SD are shown.
C. Initial reaction velocities of NanS (2.5nM) with pNP-Ac substrate (concentrations indicated) in 50mM Phosphate, 150mM NaCl (pH7.4). Rates were established using a pNP standard curve that were used to derive values for K_m. Experiments were performed with three separate protein preparations with mean and SD shown.

Figure 3.
Fluorescently labelled (DMB) A. Neu5,9Ac2 and B. Neu5,4Ac2 was incubated overnight (25°C) with 5µl (0.7mg/ml) NanS (bottom) and compared to DMB-Neu5Ac standards (not shown) by U-HPLC. Elution times are shown on the x-axis and relative intensity as stacked images.

Figure 4. DMB labelled A. Neu5,9Ac2 and B. Neu5,4Ac2 was incubated overnight (25°C) with 1µL (0.7mg/ml) NanS (bottom) and compared to DMB-Neu5Ac standards (not shown) by reverse phase HPLC . Elution times are shown on the x-axis and relative intensity as stacked images.

Figure 5.
A. Co-incubation of 3µM Bovine submaxillary mucin with 100nM *T. forsythia* NanH and NanS, as indicated. Reactions were incubated for 30mins at 37°C before free sialic acid was assayed using a modified Thiobarbiturate assay (see methods). Free Neu5Ac was used to generate a standard curve. Assays are shown from triplicate enzyme reactions with SD displayed. One-way ANOVA was used to test statistical significance ($P<0.001$) (*).
B. Oral Squamous cell carcinoma H357 cells were incubated with NanS (50nM) for 1hour at 37°C, fixed and stained with FITC (Green) labelled Cancer Antennarius (CCA) lectin and DAPI (Blue). Exposure time and weighting for both DAPI and FITC fluorescence was kept consistent between conditions.
C. Images were processed in image J; background subtraction and mean fluorescence intensity (MFI) per cell calculations were carried out using the same parameters for each condition- MFI per cell is shown with Standard deviation for 3 fields of view (approx.. 60 cells), T-test was used to assess significance ($P=0.0276$).

D. Oral Squamous cell carcinoma H357 cells were incubated with NanS (50nM) or NanS and NanH (50nM each) for 30 minutes at at 37°C, stained with FITC (Green) labelled Sambucus Nigra (SNA) lectin and DAPI (Blue) before fixing. During fluorescent microscopy, exposure time and weighting for both DAPI and FITC fluorescence was kept consistent between conditions.
E. Images were processed in image J; background subtraction and mean fluorescence intensity (MFI) per cell calculations were carried out using the same parameters for each condition and MFI per cell shown s shown with Standrad deviation for 3 fields of view (approx.. 60 cells) , T-test was used to assess significance (P=0.0219).

Figure 6.
Plots of HILIC-UHLPC profiles of 2-AB labelled N-glycans released from Erythropoietin (EPO) using PNGase F after overnight incubation of EPO with (middle) or without (upper) 1μl NanS ($0.7 mgml^{-1}$). An overlaid plot is also shown (lower) and peaks of interest that are assumed to contain Neu5,9Ac2 are highlighted with a circle and their glycan structures (as derived by Mass Spectrometry of procainamide labelled N-glycans; Supplementary Table 2) are depicted according to the CFG notation; blue square, *N*-acetylglucosamine; green circle, mannose; yellow circle, galactose; red triangle, fucose; purple diamond, sialic acid. Acetyl group present on one of the sialic acids is also labelled as Ac1.

Supplementary Figure 1.

Primary amino acid sequence alignment between the published *T. forsythia* NanS (strain 92A.2) and the cloned version from strain 43037, sequenced by Valentin Friedrich, BOKU, Austria and deposited under accession NC_016610.1. The conserved SGNH residues are highlighted in bold with sequence blocks I,II,III and V underlined for the n- and c- terminal regions.

Supplementary Figure 2.

Purification of soluble NanS prepared from pGEX4T3-NanS by cleavage using thrombin; Lane1- NEB Protein ladder 10-250kDa , Lane 2-4- elution fractions 1-3, stained with Coomassie Blue.

Supplementary table 1

Comparison of relative abundance of N-glycans from rh-EPO that have been treated with NanS enzyme. Relative peak areas shown in Figure 5. Undig, undigested; nd, not detected.

Supplementary Table 2

Masses and compositions for selected (peaks 10, 19 and 23a) N-glycans released from rh-EPO. The N-glycans were procainamide labelled and analysed by by LC-ESI-MS as decribed in Kozak et al., Each sample was injected onto an ACQUITY UPLC® BEH-Glycan 1.7 μm, 2.1 x 150 mm column at 40°C on the Dionex Ultimate 3000 UHPLC instrument with a fluorescence detector (λex = 310nm, λem = 370nm), attached to a Bruker Amazon Speed ETD.

Kozak RP, Tortosa CB, Fernandes DL, Spencer DI. (2015) Comparison of procainamide and 2-aminobenzamide labeling for profiling and identification of glycans by liquid chromatography with fluorescence detection coupled to electrospray ionization-mass spectrometry. Anal Biochem. 2015 Oct 1;486:38-40

Figure 1

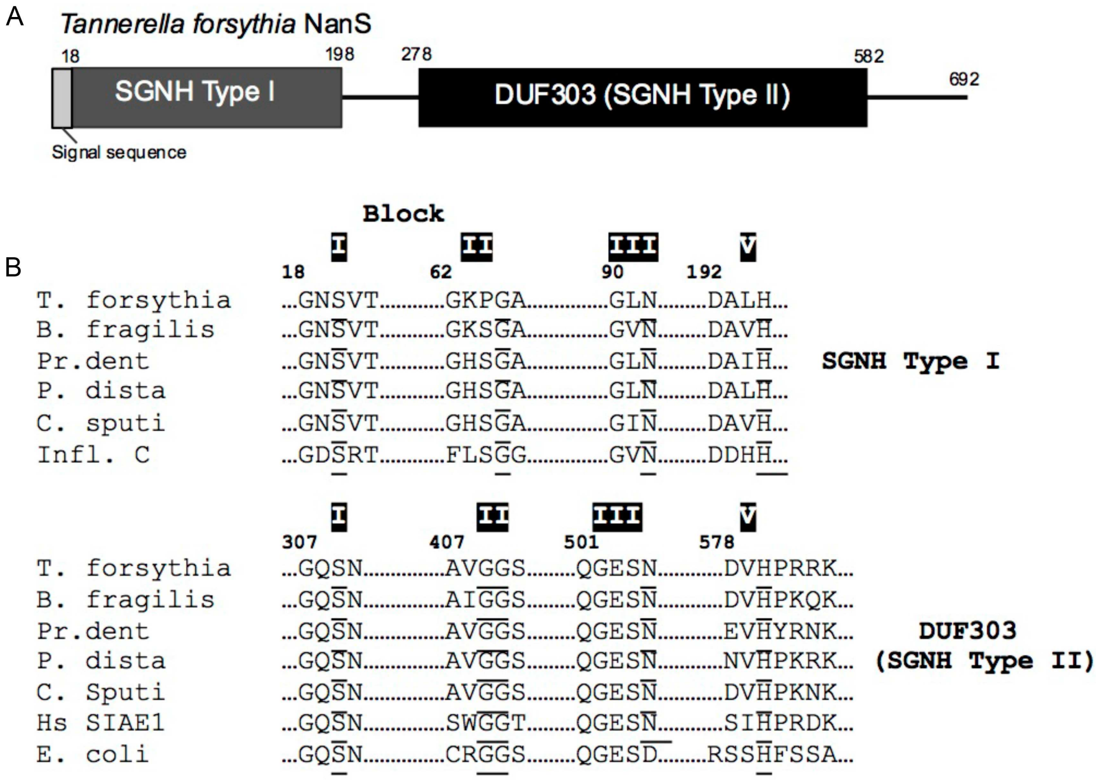

A
Tannerella forsythia NanS
18
198
278
582
692
SGNH Type I
DUF303 (SGNH Type II)
Signal sequence
B
Block
I
II
III
V
18
62
90
192
T. forsythia ...GNSVT...........GKPGA..............GLN.........DALH...
B. fragilis ...GNSVT...........GKSGA..............GVN.........DAVH...
Pr.dent ...GNSVT...........GHSGA..............GLN.........DAIH...
P. dista ...GNSVT...........GHSGA..............GLN.........DALH...
C. sputi ...GNSVT...........GHSGA..............GIN.........DAVH...
Infl. C ...GDSRT...........FLSGG..............GVN.........DDHH...
SGNH Type I
307
407
501
578
T. forsythia ...GQSN..............AVGGS.........QGESN............DVHPRRK...
B. fragilis ...GQSN..............AIGGS.........QGESN............DVHPKQK...
Pr.dent ...GQSN..............AVGGS.........QGESN............EVHYRNK...
P. dista ...GQSN..............AVGGS.........QGESN............NVHPKRK...
C. Sputi ...GQSN..............AVGGS.........QGESN............DVHPKNK...
Hs SIAE1 ...GQSN..............SWGGT.........QGESN............SIHPRDK...
E. coli ...GQSN..............CRGGS.........QGESD.........RSSHFSSA...
DUF303
(SGNH Type II)

Figure 2

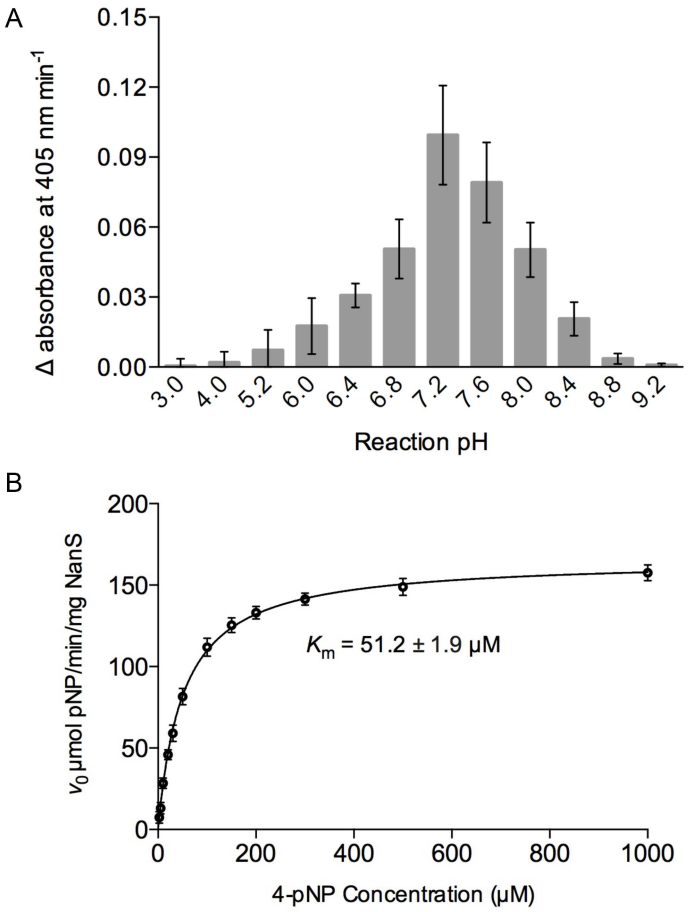

A
Δ absorbance at 405 nm min-1
0.15
0.12
0.09
0.06
0.03
0.00
3.0
4.0
5.2
6.0
6.4
6.8
7.2
7.6
8.0
8.4
8.8
9.2
Reaction pH
B
v0 µmol pNP/min/mg NanS
200
150
100
50
0
Km = 51.2 ± 1.9 µM
0
200
400
600
800
1000
4-pNP Concentration (µM)

Fig.3

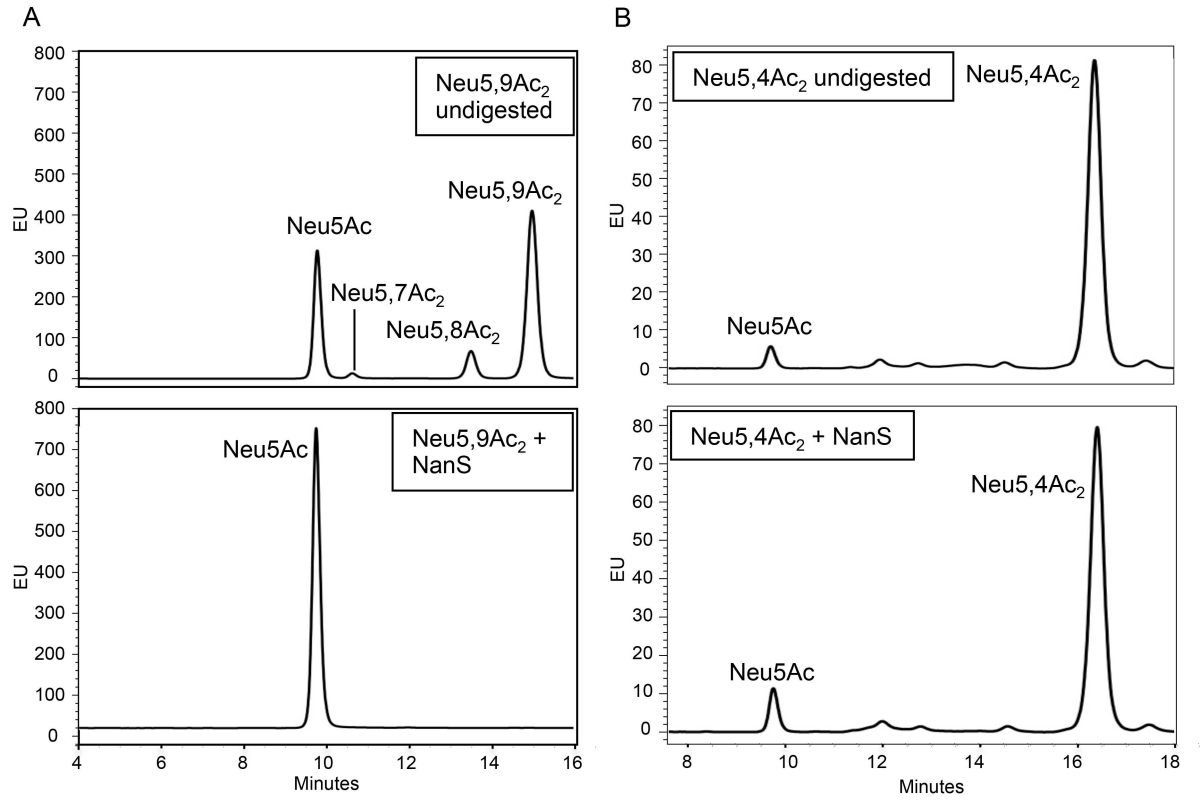

A
B
Neu5,9Ac2 undigested
Neu5Ac
Neu5,7Ac2
Neu5,8Ac2
Neu5,9Ac2
Neu5,9Ac2 + NanS
Neu5Ac
Neu5,4Ac2 undigested
Neu5Ac
Neu5,4Ac2
Neu5,4Ac2 + NanS
Neu5Ac
Neu5,4Ac2
EU
Minutes

Fig.4

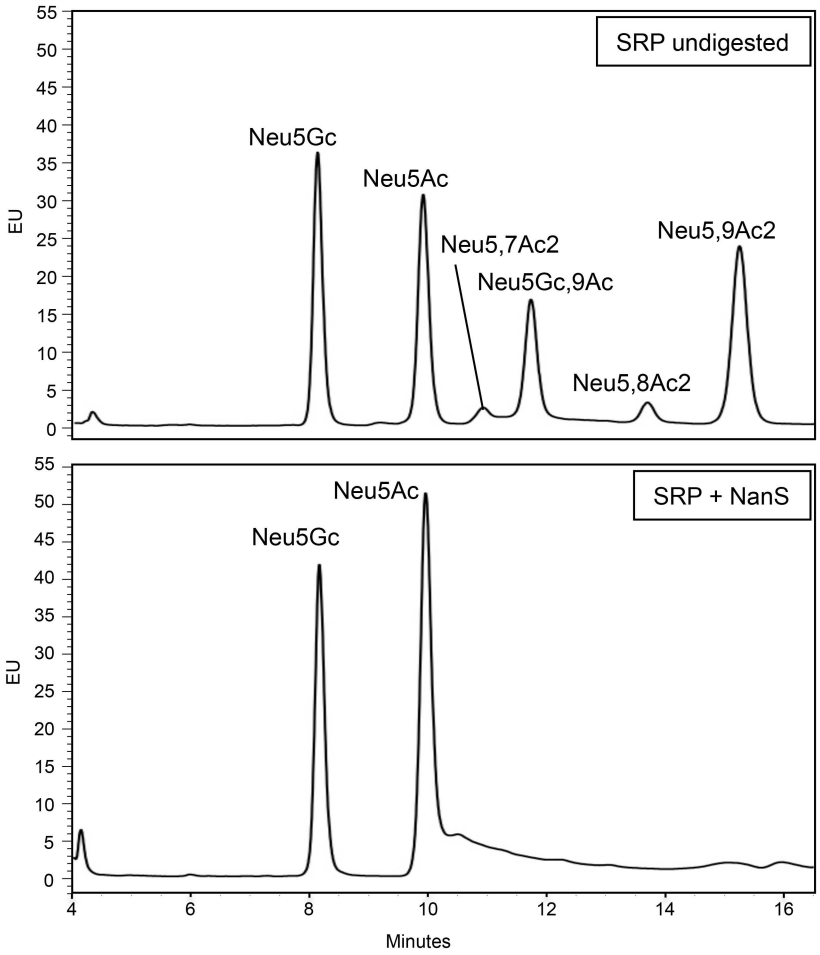

SRP undigested
Neu5Gc
Neu5Ac
Neu5,7Ac2
Neu5Gc,9Ac
Neu5,9Ac2
Neu5,8Ac2
SRP + NanS
Neu5Ac
Neu5Gc
EU
Minutes
0
5
10
15
20
25
30
35
40
45
50
55
4
6
8
10
12
14
16

Fig.5

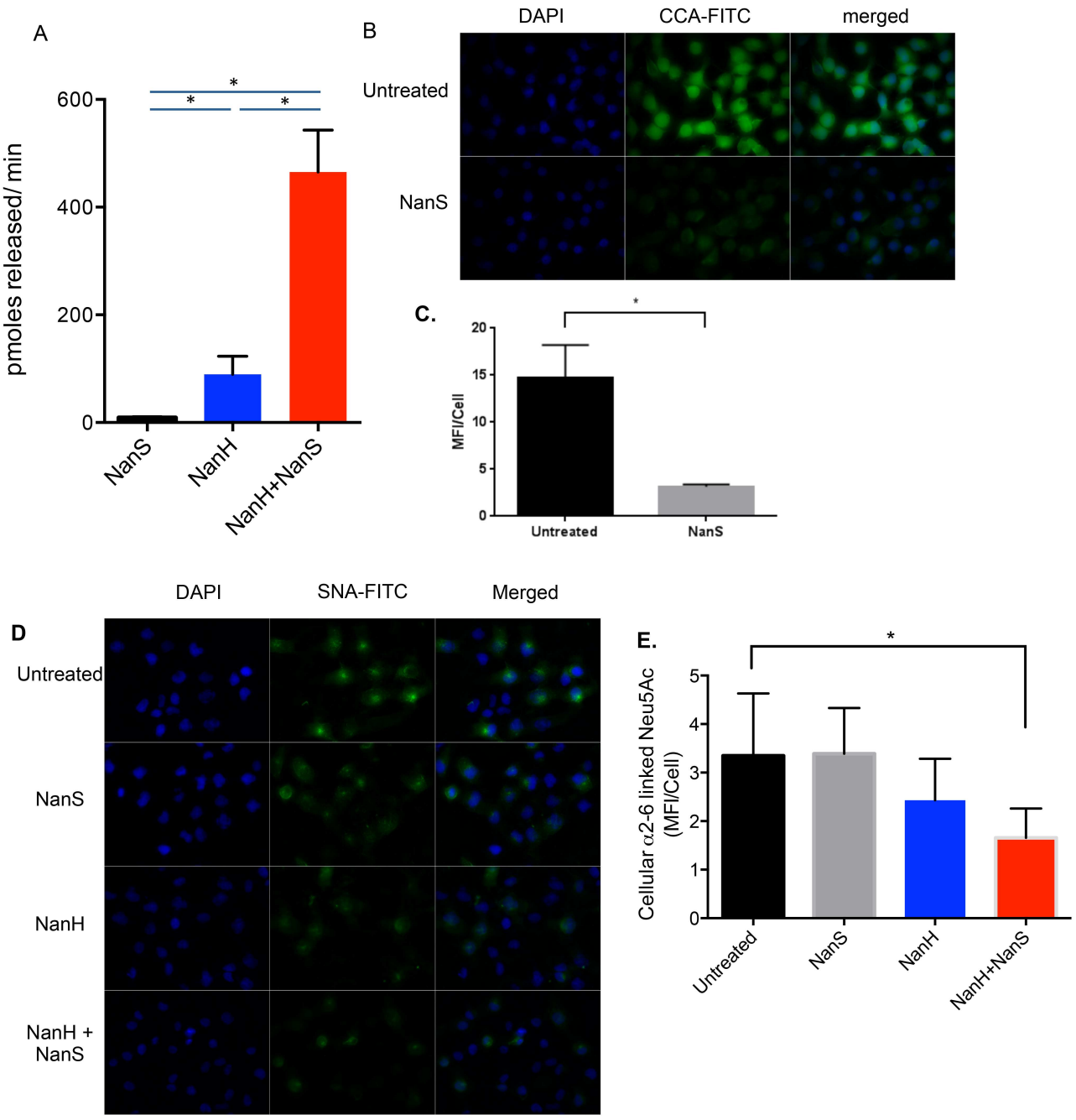
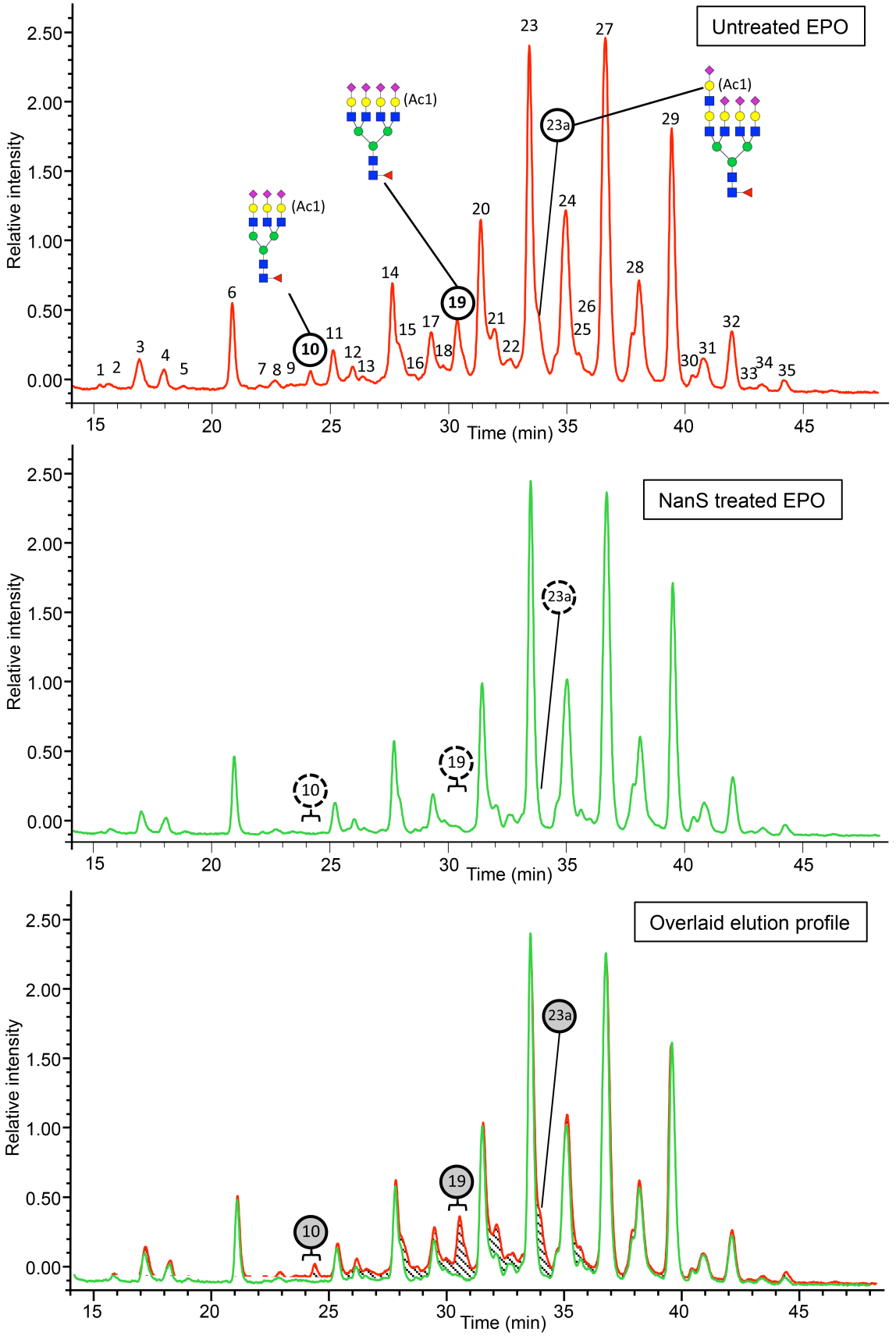

A
pmoles released/ min
600
400
200
0
*
*
*
NanS
NanH
NanH+NanS
B
DAPI
CCA-FITC
merged
Untreated
NanS
C.
*
MFI/Cell
20
15
10
5
0
Untreated
NanS
D
DAPI
SNA-FITC
Merged
Untreated
NanS
NanH
NanH +
NanS
E.
*
Cellular α2-6 linked Neu5Ac
(MFI/Cell)
5
4
3
2
1
0
Untreated
NanS
NanH
NanH+NanS

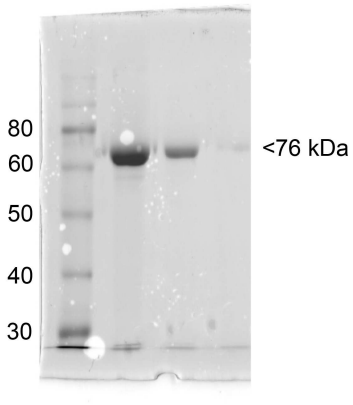
Fig.6


Untreated EPO
NanS treated EPO
Overlaid elution profile
Relative intensity
Time (min)
(Ac1)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 23a 24 25 26 27 28 29 30 31 32 33 34 35

Supplementary fig.1

```
T. forsythia 92A.2          QKTVKVACVGNSVTYGTGIKNRETDSYPAQLQRLLGKGYEVGSFGKPGATLLNKGFRPYT 60
T. forsythia 43037          QKTVKVACVGNSVTYGTGIKNRETDSYPAQLQRLLGKGYEVGSFGKPGATLLNKGFRPYT 60
                            ************************************************************

T. forsythia 92A.2          QQEEYRRAVEFAGDRVIIHLGLNDTDPKAWPNYRDEFLHDYLALIDTFRRVNPACRIWVC 120
T. forsythia 43037          QQEEYRRAVEFAGDRVIIHLGLNDTDPKAWPNYRDEFLHDYLALIDTFRRVNPACRIWVC 120
                            ************************************************************

T. forsythia 92A.2          RMTPIAHRHRRFKSGTRDWYWMIQEKIEQVARWAGVGLIDLQEPLYHRPDLLPDALHPDE 180
T. forsythia 43037          RMTPIAHRHRRFKSGTRDWYWMIQEKIEQVARWAGVGLIDLQEPLYHRPDLLPDALHPDE 180
                            ************************************************************

T. forsythia 92A.2          EGAGLLAKTVYQAVTGDFGGLKMPAVYSDNMVLQRDRALTIEGTADAGETVTVSIAKQTK 240
T. forsythia 43037          EGAGLLAKTVYQAVTGDFGGLRMPAIYSDNMVLQRDRALTIEGTADAGETVTVSIAKQTK 240
                            ********************:***:************************************

T. forsythia 92A.2          TARTPDNGRWTVTLDPMPAQTGLVLTVATPSRKLTYRQVAVGEVWLCSGQSNMAFRTSEA 300
T. forsythia 43037          TARTPDNGRWTVTLDPMPAQTGLVLTVATPSRKLTYRQVAVGEVWLCSGQSNMAFQTSEA 300
                            ************************************************:****

T. forsythia 92A.2          DAAERAKLLDFARKNPQIRFFDMKPRWHTGALAWSKTALDSLNHLQYYHDTRWEPCNEQT 360
T. forsythia 43037          DAAERAKLLDFARKNPQIRFFDMKPRWHTGASAWSKTALDSLNHLQYYHDTRWEPCNEQT 360
                            ******************************* ****************************

T. forsythia 92A.2          ADRFSAIAMAFGRMLSDSLGVPVGLISNAVGGSPAEAWIDRKTVEFEFPDLLYDWTKNDF 420
T. forsythia 43037          ADRFSAIAMAFGRMLSDSLGVPVGLISNAVGGSPAEAWIDRKTVEFEFPDLLYDWTKNDF 420
                            ************************************************************

T. forsythia 92A.2          IQDWVRERATLNLKEATIPLQRHPYEPCYLFEAGIAPLARYPLKGVLWYQGESNAHNLEA 480
T. forsythia 43037          IQDWVRERATLNLKEATIPLQRHPYEPCYLFEAGIAPLARYPLKGVLWYQGESNAHNLEA 480
                            ************************************************************

T. forsythia 92A.2          YSKLFPMLINSWRSYWRETLPFYFVQLSSLNRPTWAWFRDTQRRLAEEIPSCEMAVSSDR 540
T. forsythia 43037          YSKLFPMLVNSWRSYWRETLPFYFVQLSSLNRPTWAWFRDTQRRLAEEIPSCDMAVSSDR 540
                            ********:***************************************:*******

T. forsythia 92A.2          GDSLDVHPRRKTDVGERLARIALNRSYGQSSVIPSGPSFLSATFSEQAVWLSFRYAEGLT 600
T. forsythia 43037          GDSLDVHPKRKTDVGERLARIALNRSYGQSSVIPSGPSFLSATFSEQAVWLSFRYAEGLT 600
                            ********:***************************************************

T. forsythia 92A.2          TSDGKPLCTFELAGEDERFYPAEAEIVGHRLRLRSPQVSRPRKVRYGWQPFTRANLINGA 660
T. forsythia 43037          TSDGKPLCTFELAGEDERFYPAEAEIVGHRLRLRSPQVTRPRKVRYGWQPFTRANLINGA 660
                            ************************************:***********************

T. forsythia 92A.2          GLPASTFQAVIE 672
T. forsythia 43037          GLPASTFQAVIE 672
                            ************
```

Supplementary fig.2



Supp Table 1:

rh-EPO 2-AB labelled N-glycans			
Peak Id	GU	Undig	NanS
		% Area	
1	7.66	0.07	0.06
2	7.76	0.12	0.27
3	8.09	1.30	1.13
4	8.35	0.77	0.82
5	8.57	0.10	0.12
6	9.08	2.39	2.81
7	9.37	0.12	0.08
8	9.54	0.33	0.30
9	9.71	0.05	0.15
10	9.93	0.39	nd
11	10.18	1.33	1.34
12	10.40	0.82	0.73
13	10.51	0.50	0.28
14	10.86	5.71	4.68
15	11.12	nd	0.09
16	11.22	nd	0.12
17	11.32	2.88	1.71
18	11.47	0.85	0.28
19	11.64	3.49	nd
20	11.93	7.33	7.45
21	12.10	2.75	1.33
22	12.31	1.58	1.20
23	12.55	15.75	16.18
23a			nd
24	13.04	10.96	10.82
25	13.24	1.19	1.12
26	13.35	nd	0.55
27	13.60	16.29	18.81
28	14.09	7.26	8.34
29	14.58	9.70	11.83
30	14.90	0.47	0.70
31	15.07	2.00	2.37
32	15.53	2.50	3.10
33	15.83	0.14	0.19
34	16.03	0.37	0.47
35	16.40	0.47	0.57

Supplementary Table 2.

U(H)PLC Peak ID	Possible structure	LC-ESI-MS												
		Composition						$[M/Z]^{2+}$ calculated	$[M/Z]^{2+}$ observed	$[M/Z]^{3+}$ calculated	$[M/Z]^{3+}$ observed	$[M/Z]^{4+}$ calculated	$[M/Z]^{4+}$ observed	[M/Z] characteristic fragment ions (composition)
		Hex (H)	HexNAc (N)	Fuc (F)	Neu5Ac (S)									
					0 OAc	1 OAc	2 OAc							
10	FA3G3S3Ac1	6	5	1	2	1	0	1644.13	1644.12	1096.42	1096.36	822.57	822.52	1642.70 (H4N3F1-PROC) 1004.54 (S1H2N2) 699.26 (S1OAc1H1N1) 657.19 (S1H1N1)
19	FA4G4S4Ac1	7	6	1	3	1	0	-	-	1315.22	1315.09	986.63	986.55	1932.76 (S1H4N3F1-PROC) 1642.71 (H4N3F1-PROC) 1475.34 (S2H3N2) 699.22 (S1OAc1H1N1) 657.38 (S1H1N1)
23a	FA4G4S4LacNAc1Ac1	8	7	1	3	1	0	-	-	-	-	1077.91	1077.86	699.23 (S1OAc1H1N1) 657.23 (S1H1N1)