**EndoS**₂ is a unique and conserved enzyme of serotype M49 group A *Streptococcus* that hydrolyses N-linked glycans on IgG and α₁-acid glycoprotein

Jonathan SJÖGREN*1,2, Weston B. STRUWE†3, Eoin F. J. COSGRAVE†4, Pauline M. RUDD†, Martin STERVANDER‡, Maria ALLHORN*2, Andrew HOLLOANS§5, Victor NIZET*6 and Mattias COLLIN*2

*Department of Clinical Sciences, Division of Infection Medicine, Lund University, Biomedical Center B14, SE-22184 Lund, Sweden, †The National Institute for Bioprocess Research and Training, NIBRT, University College Dublin, Fosters Avenue, Mount Merrion, Blackrock, Co. Dublin, Ireland, §Molecular Ecology and Evolution Laboratory, Department of Biology, Lund University, Ecology Building, SE-223 62 Lund, Sweden, ‡School of Biological Sciences, University of Wollongong, Wollongong, NSW 2522, Australia, ¶Department of Pediatrics, University of California San Diego, 9500 Gilman Drive, Mail Code 0687, La Jolla, CA 92093, U.S.A., and †Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, Mail Code 0687, La Jolla, CA 92093, U.S.A.

Many bacteria have evolved ways to interact with glycosylation functions of the immune system of their hosts. *Streptococcus pyogenes* [GAS (group A *Streptococcus*)] secretes the enzyme EndoS that cleaves glycans on human IgG and impairs the effector functions of the antibody. The *ndoS* gene, encoding EndoS, has, until now, been thought to be conserved throughout the serotypes. However, in the present study, we identify EndoS₂, an endoglycosidase in serotype M49 GAS strains. We characterized EndoS₂ and the corresponding *ndoS2* gene using sequencing, bioinformatics, phylogenetic analysis, recombinant expression and LC–MS analysis of glycodic activity. This revealed that EndoS₂ is present exclusively, and highly conserved, in serotype M49 of GAS and is only 37% identical with EndoS. EndoS₂ showed endo-β-N-acetylglucosaminidase activity on all N-linked glycans of IgG and on biantennary and sialylated glycans of AGP (α₁-acid glycoprotein). The enzyme was found to act only on native IgG and AGP and to be specific for free biantennary glycans with or without terminal sialylation. GAS M49 expression of EndoS₂ was monitored in relation to carbohydrates present in the culture medium and was linked to the presence of sucrose. We conclude that EndoS₂ is a unique endoglycosidase in serotype M49 and differs from EndoS of other GAS strains by targeting both IgG and AGP. EndoS₂ expands the repertoire of GAS effectors that modify key glycosylated molecules of host defence.

Key words: α₁-acid glycoprotein, endo-β-N-acetylglucosaminidase, host–pathogen interaction, IgG glycosylation, *Streptococcus pyogenes*.

**INTRODUCTION**

Glycosylation is a common post-translational modification, and almost all key molecules in the immune system are glycosylated [1]. IgG is the most abundant antibody in serum with the capacity to bind and neutralize antigens, facilitate antibody-dependent cytotoxicity, opsonize antigens and initiate phagocytosis. IgG is composed of two light and two heavy chains, of which the latter are glycosylated with complex N-linked glycans at Asn297. The glycan is present in a pocket of the two heavy chains of the IgG molecule, where it has been shown to be flexible and dynamic allowing it to influence the glycan–protein interaction with FcγRs [5]. IgA, IgD, IgE and IgM each carry several occupied N- and O-linked glycosylation sites, and the study of the glycan’s impact on the effector functions of these immunoglobulins has only begun [6].

*Streptococcus pyogenes* [GAS (group A *Streptococcus*)] is a leading Gram-positive bacterial pathogen exhibiting a wide array of immune evasion mechanisms, including interference with host glycosylation [7]. Every year, this bacterium causes over 500 000 deaths due to severe infections and post-infectious immunological disorders: invasive infections, rheumatic fever, glomerulonephritis and hundreds of millions of cases of milder and self-limiting infections, such as pharyngitis and impetigo [8]. GAS is subdivided into serotypes on the basis of the antigenic M-protein on the bacterial surface and there are currently over 100 serotypes described [9].

An endoglycosidase from *S. pyogenes*, EndoS, was discovered in serotype M1 of GAS and found to hydrolyse the N-linked glycan on the heavy chain of native human IgG and in this way modulate the binding of IgG to FcγR [10–12]. EndoS (EC

---

**Abbreviations used:** 2-AB, 2-aminobenzamide; ABS, *Arthrobacter ureafaciens* sialidase; AGP, α₁-acid glycoprotein; AMF, almond meal α-fucosidase; BEH, bridged ethane–silicon hybrid; BKF, bovine kidney α-fucosidase; BTG, bovine testes β-galactosidase; CM, C-medium; CcpA, catabolite control protein A; FcγR, Fcγ receptor; FLD, fluorescence detection; GAS, group A *Streptococcus*; GH18, family 18 of glycoside hydrolases; HILIC, hydrophilic interaction liquid chromatography; HRP, horseradish peroxidase; LCA, *Lentilus culinaris* agglutinin; 4MU-GlcNAc, 4-methylumbelliferyl N-acetyl-β-D-glucosaminide; MWCO, molecular mass cut-off; NAN1, neuraminidase/sialidase 1; PNGase F, peptide N-glycosidase F; r, recombinant; UHPLC, ultra-HPLC.

1 To whom correspondence should be addressed (email jonathan.sjogren@med.lu.se).
2 Patents for the use of EndoS₂ have been applied for by Genovis AB. Jonathan Sjögren, Andrew Hollands, Victor Nizet, Maria Allhorn and Mattias Collin are listed as inventors on the application that is pending.
3 Present address: Waters Corporation, 34 Maple Street, Milford, MA 01757, U.S.A.
4 Present address: Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3TA, U.K.
5 The sequences of *Streptococcus pyogenes* *ndoS2* will appear in the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under accession numbers KC155346 (strain 3487-05), KC155348 (strain AP49), KC155347 (strain ACN49), KC155349 (strain AW1) and KC155350 (strain AW2).
Expression of EndoS was studied using growth of NZ131 in 50% CM (C-medium) [0.5% Proteose Peptone, 1.5% (w/v) yeast extract, 10 mM K$_2$PO$_4$, 0.4 mM MgSO$_4$, and 17 mM NaCl (pH 7.5)].

**Sequencing of ndoS2**

Five GAS serotype M49 strains were selected for sequencing of the ndoS2 gene; 3487-05, AP49, ACN49, AW1 and AW2. Sequencing was carried out using primers ndoS2-out-R, seq38-R, seq42-R, seq54-R, seq15-F, seq17-F, seq24-F and seq28-F and the Lightrun sequencing service of GATC Biotech (Konstanz, Germany). All primers used for sequencing are summarized in Supplementary Table S2. The sequences have been deposited in GenBank with accession numbers as follows: KC155346 (strain 3487-05), KC155348 (strain AP49), KC155347 (strain ACN49), KC155349 (strain AW1), KC155350 (strain AW2) (Supplementary Table S2).

**Recombinant expression of EndoS$_2$**

Recombinant expression of EndoS$_2$ in *E. coli* was established by PCR amplification of the ndoS2 gene from GAS NZ131 with the primers ndoS2-F-BamHI, 5′-CTGTAGCATCCAGAGACCTG-3′, and ndoS2-R-Xhol, 5′-GGAGCGCTGCATGTTATATATGGAGCTT-3′. The ndoS2 fragment was digested with restriction enzymes BamHI and Xhol (restriction sequences are underlined) and ligated into the expression vector pGEX-5X-3 (GE Healthcare) using DNA ligase T4 (Thermo Fisher Scientific) creating the plasmid pGEX-ndoS2. The expression vector was transformed into *E. coli* Top10 chemically competent cells and screened with PCR using primers ndoS2-F-BamHI and ndoS2-R-Xhol. Positive clones were isolated and the pGEX-ndoS2 plasmid was purified and transformed into the *E. coli* expression strain BL21 pLysS. One recombinant clone was grown overnight at 37°C with antibiotics, diluted 1:20 in lysogeny broth medium with antibiotics and grown for 3 h to mid-exponential phase. The expression of the protein GST–EndoS$_2$ was induced with 0.1 mM IPTG for 3 h. The cells were harvested and lysed with BugBuster Protein Extraction Reagent (Novagen/Merck). Recombinant GST–EndoS$_2$ was purified on a column with glutathione–Sepharose 4B (GE Healthcare) and eluted with reduced glutathione. The GST tag was cleaved off using Factor Xa (New England BioLabs). Site-directed mutagenesis was performed on pGEX-ndoS2 using a QuikChange II Site-Directed Mutagenesis Kit (Agilent) with primers ndoS2(E-L)-F and ndoS2(E-L)-R, to exchange the glutamate residue (E) in the active site to leucine (L). Recombinant EndoS$_2$(E186L) was cloned and expressed in a similar way to EndoS$_2$.

**Phylogenetic analysis**

We searched non-redundant protein databases at NCBI with the BLASTP algorithm, submitting the EndoS$_2$ sequence of GAS strain ACN49 (M49). We retrieved similar protein sequences, setting a cut-off of the expect value at $<10^{-10}$. All but two sequences belonged to GH18, except for two sequences of other hypothetical proteins. All sequences were included in phylogenetic analyses of the proteins, and the non-GH18 proteins served as outgroup. A total of 101 protein sequences were aligned in Geneious version 6.0.3 (Biomatters Ltd, available from http://www.geneious.com/) using the ClustalW algorithm (Supplementary Table S3 at
EndoS<sub>2</sub> hydrolyses N-linked glycans on IgG and α<sub>1</sub>-acid glycoprotein

http://www.biochemj.org/bj/455/bj4550107add.htm). From the alignment of 1817 amino acids, a region of generally high coverage comprising 1099 amino acids was extracted and analysed in BEAST version 1.7.4 [7,26]. We used the Blosum62 substitution model [8,27] with no site heterogeneity specification, set a strict molecular clock, selected the Yule process [9,28] for tree prior and ran Bayesian MCMC sampling every 1000 generations for 10 million generations. The output was examined with Tracer version 1.5 (A. Rambaut and A.J. Drummond, available from http://beast.bio.ed.ac.uk/Tracer) in order to ensure that likelihood scores were stationary and that effective sample sizes were adequate (>500), setting burnin to 25%. A maximum clade credibility tree was calculated with TreeAnnotator version 1.7.4 [1,26]. 16S rRNA (RNA or cDNA) sequences were retrieved from the Ribosomal Project Database (http://rdp.cme.msu.edu) and GenBank<sup>®</sup> for taxonomic analyses of the corresponding taxa/strains in the EndoS dataset (Supplementary Table S3). In case a specific strain was lacking for taxa more distantly related in the EndoS protein tree dataset, another strain was chosen if available. This resulted in a total of 51 representative sequences, which were aligned according to the procedures described for the EndoS protein dataset above. From the resulting alignment of 2172 nt, a high coverage region of 51 representative sequences, which were aligned according to the procedures described for the EndoS protein dataset above.

**EndoS<sub>2</sub> expression analysis**

Overnight cultures of NZ131 in CM were diluted 1:50 and grown for 16 h at 37°C with 5% CO<sub>2</sub> and the cysteine protease inhibitor E-64 at 20 μM. Glucose, galactose, sucrose, acetylglucosamine and mannose were added at 0.01% concentration when indicated. Supernatants were concentrated using precipitation on 51 representative sequences, which were aligned according to the procedures described for the EndoS protein dataset above. From the resulting alignment of 2172 nt, a high coverage region of 51 representative sequences, which were aligned according to the procedures described for the EndoS protein dataset above.

**Chitinase assay**

4MU-GlcNAc (4-methylumbelliferyl N-acetyl-β-D-glucosaminide) (Sigma–Aldrich) was incubated at 0.2 mM with 0.3 unit of chitinase from *Streptomyces griseus* (Sigma–Aldrich) or 2 μg of rEndoS<sub>2</sub> (where r denotes recombinant) or 2 μg of rEndoS<sub>2</sub> or PBS in 100 μl of PBS. The reactions were incubated at 37°C for 1 h. Then, 100 μl of 0.1 M glycine (pH 10) was added to stop the reaction. Absorbance at 355/445 nm was measured in a black 96-well plate using a spectrophotometer. The experiments were carried out using five replicates and results are shown as means±S.D. The response in absorbance was analysed statistically by an unpaired Student’s t test, where differences were considered significant if *P* < 0.05. ****P < 0.001.

**Glycoprotein denaturing**

A 4 μg amount of IgG or AGP was incubated in 10 μl of PBS at 37°C, 40°C, 50°C, 60°C, 70°C or 80°C for 30 min. After the incubation, the samples were kept at 37°C. Then, 2 μg of rEndoS<sub>2</sub> was added to each reaction mixture and incubated further at 37°C for 2 h. The samples were analysed on a SDS/PAGE gel and for IgG with LCA lectin blotting as described above. For analysis of EndoS<sub>2</sub> specificity, 4 μg of α<sub>2</sub>-macroglobulin, ovalbumin, human lactoferrin, RNase B and fetuin (all Sigma–Aldrich) were incubated with 2 μg of EndoS<sub>2</sub> at 37°C overnight and subsequently analysed on SDS/PAGE gel as described.

**LC–FLD (fluorescence detection)–MS**

Online coupled LC–MS with FLD was performed using a Waters Xevo G2 QToF with Acuity UPLC and BEH (bridged ethane–silicon hybrid) glycan column (1.0 mm × 150 mm, 1.7 μm particle size). MS data was acquired in negative mode with the following conditions: 2500 V capillary voltage, 50 V cone voltage, 280°C desolvation temperature, 600 l/h<sup>−1</sup> desolvation gas and 100°C source temperature. The analyser was set to sensitivity mode. The fluorescence data rate was 1 point·s<sup>−1</sup> and a PMT gain of 10 with excitation and emission wavelengths set at 320 nm and 420 nm respectively. Samples were in 80% acetonitrile with an injection volume of 10 μl. The flow rate was 0.150 μl·min<sup>−1</sup>. Solvent A was 50 mM ammonium formate (pH 4.4) and solvent
which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Licence (CC-BY) (http://creativecommons.org/licenses/by/3.0/)

© 2013 The Author(s)

column (2.1 mm of 330 nm and 420 nm respectively. Eluting glycans were detected detection was achieved using excitation and emission wavelengths · sialic acid in all linkages was removed with 1 m-unit end monosaccharides were removed as follows: terminal (ultra-HPLC) with fluorescence detection using a BEH glycan labelled glycans were then prepared for separation on UHPLC MWCO centrifugal filters. Digested 2-AB (2-aminobenzamide)-digested glycan pool was then relabelled with 2-AB. Each FLD–MS using a Waters ACQUITY UPLC μ the presence of 80 % formate (pH 4.4) against acetonitrile with ammonium formate increasing from 30 % to 47 % over a 32 min period. Fluorescence detection was achieved using excitation and emission wavelengths of 330 nm and 420 nm respectively. Eluting glycans were detected in positive mode with the following settings: cone voltage of 80 V, capillary voltage of 3.0 kV, source temperature of 120 °C, desolvation temperature of 300 °C, and desolvation gas flow of 800 l h⁻¹. Mass data were acquired using sensitivity mode with a mass range of 750 m/z to 2000 m/z with a 1.0 s scan time. Both LC–FLD and LC–MS data were acquired and processed using Waters UNIFI version 1.6.

Figure 1 Genetic context analysis of ndoS and ndoS2
The genetic context of ndoS2 (from NZ131/M49) and ndoS (from MGAS5005/M1) was analysed by aligning and comparing the identity of the sequences in MacVector.

Exoglycosidase digestion arrays
Analysis of glycan sequence, composition and linkage specificities was facilitated by the use of exoglycosidase digestion arrays. All digestion reactions were performed with enzymes from Prozyme. Fluorescently labelled glycans were digested in 50 mM sodium acetate (pH 5.5) at 37 °C overnight using a panel of enzymes with each digestion reaction brought to a final volume of 10 μl using double-distilled water. Digested glycans were then separated from the enzyme mixtures using 10 kDa MWCO centrifugal filters. Digested 2-AB (2-aminoazanilide)-labelled glycans were then prepared for separation on UHPLC (ultra-HPLC) with fluorescence detection using a BEH glycan column as described previously [5,29]. Specific non-reducing end monosaccharides were removed as follows: terminal sialic acid in all linkages was removed with 1 m-unit-¹ ABS (Arthrobacter ureafaciens sialidase); terminal galactose monosaccharides were removed using 0.5 m-unit-¹ BTG (bovine testes β-galactosidase), which releases both β(1,3)- and β(1,4)-linked galactose; terminal GlcNAc monosaccharides were released with 40 m-unit-¹ GUH (Streptococcus pneumoniae hexosaminidase), capable of cleaving β-linked GlcNAc moieties; core α(1,6)-fucose was selectively removed using 1 m-unit-¹ BTF (bovine kidney α-fucosidase), (2,3)-linked sialic acid was removed using 10 m-unit-¹ recombinant Streptococcus pneumoniae N1 (neuraminidase/sialidase 1), and AMF (almond meal α-fucosidase) at 6 m-unit-¹ was used to release (1,3)- and (1,4)-linked non-reducing terminal fucose residues.

Activity of EndoS2 on free N-glycans
N-glycans present on 80 μg of bovine fetuin (Sigma–Aldrich) were released using 2500 units of PNGase F (New England BioLabs), and labelled with 2-AB (Ludger). The labelled fetuin 2-AB glycan pool was then incubated at 37 °C for 16 h in the presence of 80 μg ml⁻¹ EndoS2 in PBS to determine the activity of EndoS2 on free N-glycans. The resulting EndoS2-digested glycan pool was then relabelled with 2-AB. Each glycan preparation was separated using a 1.7 μm BEH glycan column (2.1 mm×150 mm, Waters) and analysed by UHPLC–FLD–MS using a Waters ACQUITY UPLC® H-Class Bio with fluorescence detection coupled to a Waters Xevo G2-S Q-ToF mass spectrometer. The column temperature was 40 °C with a flow rate of 0.4 ml min⁻¹ using a linear gradient of 50 mM ammonium formate (pH 4.4) against acetonitrile with ammonium formate increasing from 30 % to 47 % over a 32 min period. Fluorescence detection was achieved using excitation and emission wavelengths of 330 nm and 420 nm respectively. Eluting glycans were detected in positive mode with the following settings: cone voltage of 80 V, capillary voltage of 3.0 kV, source temperature of 120 °C,

RESULTS
Identification of EndoS2 from GAS serotype M49
In the sequenced genome of GAS strain NZ131 (serotype M49), we identified ndoS2, a gene harbouring a GH18 domain [6,24]. ndoS2 from GAS serotype M49 was found in the same genetic context as ndoS from GAS serotype M1, but showed only 53 % nucleotide identity with ndoS (Figure 1). The surrounding genes, i.e. scrb, scra, scrk and pmi, showed a high degree of nucleotide identity when comparing the chromosomal context between strain NZ131 and serotype M1 strain MGAS5005 (Figure 1). One genome of serotype M49 is available to the public (NZ131, GenBank® accession number NC_011375) and therefore ndoS2 was sequenced in five M49 strains of different origin and isolation year (3487-05, ACN49, AP49, AW1 and AW2). The comparison revealed 100 % identity of ndoS2 in the five selected strains compared with ndoS2 found in NZ131. The ndoS2 sequences have been submitted to GenBank® (Supplementary Table S1). The deduced amino acid sequence of EndoS2 and EndoS revealed 37 % identity when aligned using ClustalW (Figure 2). The signal peptide was conserved, but three major sections of the EndoS amino acid sequence were lacking in EndoS2; at positions 45–83, 535–561 and 933–986, gaps can be seen in the alignment. A comparison of the active site of EndoS2 and EndoS revealed the GH18 motif (DXDXXDXE) with glutamate at position 186 as the catalytic amino acid to be conserved (Figure 2). Specific tryptophan residues have previously been shown to be important for the enzymatic activity of EndoS2, and when EndoS2 was aligned and compared with EndoS, tryptophan residues at positions 121, 164, 332, 361, 391, 809, 828 and 907 were found to be conserved [10–13].

In order to evaluate the evolutionary history of EndoS2, we reconstructed a protein specific phylogenetic tree, using BEAST version 1.7.4, on 101 protein sequences selected with the BLASTP algorithm on EndoS2. EndoS2 (depicted in blue) was found to be unique to GAS serotype M49 and relatively different from EndoS found in other serotypes of GAS as well as EndoS-like proteins in other Streptococcus species (Figure 3). This can be contrasted with the taxonomic phylogeny of the 16S rRNA sequences (Supplementary Figure S1 at http://www.biochemj.org/bj/455/bj4550107add.htm).

EndoS2 hydrolyses the N-linked glycan on the heavy chain of IgG
Previous work has concluded that EndoS hydrolyses the N-linked glycan on IgG [10,13]. Although the enzymes are different, we tested whether IgG is a substrate for EndoS2. A comparison of
EndoS2 hydrolysates N-linked glycans on IgG and α1-acid glycoprotein

Figure 2 ClustalW alignment of EndoS2 and EndoS

EndoS2 from GAS strain NZ131 and EndoS from GAS strain MGAS5005 was aligned using ClustalW. Depicted in blue is the GH18 active site (DXXDXDXE) and in green are conserved tryptophan residues.

The hydrolysis of the N-linked glycan on the heavy chain of IgG was carried out using EndoS2, EndoS and PNGase F as positive control [14,30]. PNGase F from Elizabethkingia meningoseptica cleaves between the GlcNAc and the asparagine residue of N-linked glycans, whereas EndoS cleaves between the two GlcNAc moieties in the chitobiose core of N-linked glycans [10,11]. EndoS2 was mutated in the active site through site-directed mutagenesis where the catalytically active glutamate residue was mutated to leucine, creating the enzyme EndoS2(E186L). EndoS2, EndoS(E186L), EndoS and PNGase F were incubated with human IgG in PBS at 37 °C overnight, and analysed by SDS/PAGE and a subsequent LCA (recognizing α-linked mannose) lectin blot (Figure 4A). The gel shows a ∼4 kDa shift of the heavy chain of IgG and a corresponding lack of LCA lectin signal when incubated with EndoS2, EndoS or PNGase F, but not with EndoS2(E186L) or PBS (Figure 4A). This result indicates that EndoS2 hydrolysates the N-linked glycan on the heavy chain of IgG and confirms the glutamate residue at position 186 of EndoS2 to be the catalytically active site.
active amino acid. To evaluate enzymatic activity of EndoS2 on the subclasses of IgG, recombinant EndoS2 and EndoS2(E186L) were incubated with human IgG subclasses 1–4 and showed activity on all four human subclasses as analysed by SDS/PAGE and LCA lectin blot (Figure 4B). The glycan-hydrolysing activity of EndoS2 on animal IgG was found for the following species: mouse, rat, monkey, sheep, goat, cow and horse. To investigate glycan specificity of EndoS2, the composition of the released glycans from pooled human serum IgG was analysed by HILIC (hydrophilic interaction liquid chromatography)–UHPLC–FLD–MS and compared with the glycan profile of IgG generated by PNGase F (Figure 5). The HILIC–UHPLC–FLD–MS revealed EndoS2 to cleave between the two GlcNAc residues in the chitobiose core of the N-linked glycan and thus leaving a single GlcNAc residue with or without α(1,6)-linked fucose attached to the protein backbone. All peaks present in the PNGase F chromatogram could be found in the glycan profile of IgG released by EndoS2 with the difference of one GlcNAc with or without α(1,6)-linked fucose.

EndoS2 releases biantennary and sialylated glycans on AGP

AGP, also known as orosomucoid, is a 41–45 kDa human plasma glycoprotein, a major positive acute-phase protein, up-regulated severalfold during inflammation and a member of the lipocalin family [11,31]. The immunomodulatory effects of AGP is linked to the carbohydrate composition of the five N-linked
Glycans (Asn^{33}, Asn^{56}, Asn^{72}, Asn^{93} and Asn^{103}) that make up 45% of the molecular mass [15,31]. When incubating AGP with recombinant EndoS₂, subsequent SDS/PAGE revealed a new band at ~38 kDa and a decrease in the intensity of the band at 45 kDa (Figure 4C). No activity was detected with EndoS or PBS in the same assay (Figure 4C). To elucidate the enzymatic activity of EndoS₂ on AGP in detail, we analysed the glycans released from AGP by EndoS₂ using HILIC–UHPLC–MS and exoglycosidase arrays in UHPLC (Figure 6). The sequence, composition and linkage specificities of all glycoforms of AGP released by PNGase F were determined in the same assay. To test whether the enzyme is specific for core α(1,3) non-reducing terminal fucose linked to galactose residues and not core α(1,6)-fucose, BKF treatment, which is specific for core α(1,6)-linked fucose residues, did not result in glycan digest products.

EndoS₂ is specific for IgG and AGP and not a general chitinase

Previous work on EndoS has shown that the enzyme is specific for the native form of IgG [14]. To test whether this is valid for EndoS₂, IgG and AGP were incubated at temperatures ranging from 37 to 80°C for 30 min before the addition of EndoS₂ or PBS and a 2 h incubation at 37°C. SDS/PAGE analysis revealed a shift of IgG incubated at 37–50°C and loss of signal was seen in a corresponding LCA lectin blot, whereas only partial shift could be seen at 60°C and no shift and intact LCA signal at temperatures 70°C and 80°C (Figure 7A). Glycans from AGP were hydrolysed at 37°C, but not at 40–70°C (Figure 7B). The activity of EndoS₂ was tested further on a range of glycoproteins, i.e. α₂-macroglobulin, ovalbumin, lactoferrin, RNase B and fetuin, but no activity could be detected (Figure 7C). To study whether EndoS₂ shows general chitinase activity, we employed the substrate 4MU-GlcNAc, which fluoresces when cleaved, to compare the enzymatic activity of EndoS₂ and EndoS with that of a chitinase from S. griseus. The results indicate that neither EndoS₂ nor EndoS has a general chitinase activity compared with the positive control (Figure 7D).

EndoS₂ hydrolyses free biantennary glycans

Following the findings of the specificity of EndoS₂, we asked the question whether the enzyme is substrate-specific and/or has glycoform selectivity. To test this, we analysed the activity of EndoS₂ on free glycans. All glycoforms from bovine fetuin were released using PNGase F and were 2-AB-labelled; in a secondary enzymes were BTG, BKF and AMF (Supplementary Figure S2 at http://www.biochemj.org/bj/455/bj4550107add.htm). AMF digestion removed α(1,3) non-reducing terminal fucose linked to galactose residues and not core α(1,6)-fucose. BKF treatment, which is specific for core α(1,6)-linked fucose residues, did not result in glycan digest products.
Figure 6  Glycan fluorescent profiles from human AGP released by PNGase F and EndoS<sub>2</sub>

HILIC–FLD–MS of 2-AB-labelled glycans released from human AGP by PNGase F (left) and EndoS<sub>2</sub> (right) respectively. 2-AB-labelled glycans were digested further with NAN1 and ABS, and subsequent bi-, tri- and tetra-antennary structures are indicated using the Oxford glycan nomenclature [45]. Ions were detected as [M – 2H]<sup>−</sup> (*) and [M – H]<sup>−</sup> (**) species.

Figure 7  Activity of EndoS<sub>2</sub> on native and denatured IgG and AGP, other glycoproteins and chitinase assay

(A) IgG was incubated at temperatures ranging from 37 to 80°C for 30 min followed by incubation with EndoS<sub>2</sub> at 37°C for 2 h and analysis by SDS/PAGE and LCA lectin blot. (B) AGP was incubated at 37–70°C followed by incubation with EndoS<sub>2</sub> at 37°C for 2 h and analysis by SDS/PAGE. (C) EndoS<sub>2</sub> was incubated with α<sub>2</sub>-macroglobulin, ovalbumin, human lactoferrin, RNase B and fetuin at 37°C overnight and analysed by SDS/PAGE. (D) EndoS<sub>2</sub>, EndoS and a chitinase from S. griseus was incubated with the fluorescent substrate 4MU-GlcNAc for 1 h and fluorescence was measured at 355/445 nm. The experiments were carried out using five replicates and results are means±S.D. The response in absorbance was analysed statistically by an unpaired Student’s t test, where differences was considered significant if P < 0.05. ****P < 0.001.
reaction, the free glycans were incubated with EndoS\textsubscript{2}, relabelled with 2-AB and analysed using HILIC–UHPLC–FLD–MS. Three structures in the chromatogram (labelled 1, 2 and 3) were modified by EndoS\textsubscript{2} compared with the PNGase F glycan pool (Figure 8A and 8B). The \text{m/z} of the \([M+H]^+\) ions of these structures were identified and revealed structures for A\textsubscript{2}G\textsubscript{2} (1558.5513), A\textsubscript{2}G\textsubscript{2}S\textsubscript{1} (925.3294) and A\textsubscript{2}G\textsubscript{2}S\textsubscript{2} (1070.8700) less one GlcNAc residue (Figures 8C–8E). The results indicate that EndoS\textsubscript{2} specifically hydrolyses free biantennary glycoforms with or without terminal sialylation.

**Expression of EndoS\textsubscript{2} is linked to carbohydrate utilization**

To confirm the findings with recombinant EndoS\textsubscript{2}, the expression levels and enzymatic activity of EndoS\textsubscript{2} were analysed in GAS supernatants. The expression of EndoS\textsubscript{2} in serotype M1 of GAS is maximized in the nutrient-poor CM [10,16–21]. Therefore expression of EndoS\textsubscript{2} was analysed by Western blotting of 16 h culture medium and the expression of EndoS\textsubscript{2} was studied. Adding glucose, galactose, GlcNAc or mannose to 50% CM inhibited EndoS\textsubscript{2} expression, whereas additional sucrose increased the amount of EndoS\textsubscript{2} in the supernatant (Figure 9A).

**DISCUSSION**

The study of bacterial glycosidases has emerged as a field at the intersection of microbial pathogenesis and glyobiology. By studying the mechanisms by which bacteria interfere with host glycosylation, new insight can be gained into both bacterial pathogenesis and the impact of glycosylation on the immune system. Interfering with the glycosylation of the host defence is widespread among pathogenic bacteria for modulation of the functions of the immune system or as a way of utilizing the glycans of glycoproteins as nutrients [24,33].

For example, *Enterococcus faecalis*, a Gram-positive gut bacterium and opportunistic, secretes EndoE, an endoglycosidase with activity on the Fc-glycan on IgG and on the glycoprotein RNase B that promotes bacterial growth when nutrients are scarce [24,34]. The endoglycosidases EndoF\textsubscript{1}–3 from *E. meningoseptica* and EndoH from *Streptomyces plicatus* has been shown to be glycan-specific: high-mannose and hybrid oligosaccharides are cleaved by EndoF\textsubscript{1} and EndoH, whereas complex biantennary glycans are released by EndoF\textsubscript{2} and EndoF\textsubscript{3} respectively [24,25,35–38]. An N-glycan deglycosylation complex in *Capnocytophaga canimorsus* has been found to cleave off N-linked glycans from IgG and to transport the glycans across the cell membrane for glycocalyx degradation [24,25,39]. *S. pneumoniae* has three surface-anchored exoglycosidases that work in concert to remove sialic acid, galactose and GlcNAc on human glycoproteins [40]. GAS EndoS was thought to be conserved throughout the GAS serotypes, and only minor variations are found when comparing ndoS among the sequenced GAS strains. It was therefore surprising to find that GAS strain

---

**Figure 8** EndoS\textsubscript{2} hydrolyses biantennary free glycans

Bovine fetuin N-glycans were released with PNGase F, labelled with 2-AB, and analysed by HILIC–UHPLC–FLD–MS. Released N-glycans were digested further with EndoS\textsubscript{2} to determine enzymatic activity on free glycans. Comparison of the fluorescent chromatograms of glycans after PNGase F (B) and subsequent EndoS\textsubscript{2} digestion (A) identified three unique peaks (labelled 1, 2 and 3). These peaks correspond to three isomeric structures (A\textsubscript{2}G\textsubscript{2}, A\textsubscript{2}G\textsubscript{2}S\textsubscript{1} and A\textsubscript{2}G\textsubscript{2}S\textsubscript{2}) and were detected primarily as \text{m/z} 1558.55 \([M+H]^+\), 925.33 [M + 2H]\textsubscript{2}\textsuperscript{+} and 1070.87 \([M + 2H]\textsubscript{2}\textsuperscript{+}\) ions respectively (C–E). Extracted ion chromatograms of A\textsubscript{2}G\textsubscript{2}S\textsubscript{1} and A\textsubscript{2}G\textsubscript{2}S\textsubscript{2} precursor ions identified structural isomers, presumably from variation in sialic acid linkages.
NZ131 harboured ndoS2, with 53% identity with ndoS. The sequenced ndoS2 in five different M49 strains revealed high identity, arguing that this gene is conserved throughout the serotype.

In the phylogenetic protein tree, the EndoS2 group is relatively different from EndoS in both *S. pyogenes* and in other *Streptococcus* species (but it groups with a hypothetical protein found in *Streptococcus iniae*). In general, the patterns are not uniform: whereas within-species or within-genus similarity of the EndoS-like proteins is high for some taxonomic groups (e.g. *Bifidobacterium longum*), there is also considerable within-genus variation in *Corynebacterium* (Figure 3). Strikingly, EndoS and EndoS2 from *Streptococcus* are more closely related to EndoS-like proteins of the fungi *Cordyceps militaris* and *Beauveria bassiana* than to EndoS-like proteins of bacteria such as *Melittosoccus*, *Corynebacterium* and *Lactobacillus* (Figure 3), in sharp contrast with the taxonomic relationships (Supplementary Figure S1). Notably, some taxa are paraphyletic in the EndoS-like protein phylogeny, to which could possibly be ascribed the inclusion of non-verified hypothetical proteins. However, the biologically verified EndoS-like proteins of *Enterococcus gallinarum* and *E. faecalis* do not form a monophyletic clade (Figure 3). In all, this picture indicates the occurrence of horizontal gene transfer of ndoS-like genes. Even though no known proteins were found to be closely related to EndoS2, the differentiation from *S. pyogenes* EndoS and the high degree of similarity between serotype M49 and other *S. pyogenes* strains combined with the conserved genetic context points towards horizontal gene transfer of ndoS2 into serotype M49. The strain NZ131 also has an unusually high frequency of transformation, and horizontal gene transfer has been described on several places in the genome [24]. The alternative interpretation, that a particularly strong directional selection on the ancestral ndoS gene in serotype M49 resulted in ndoS2, seems less plausible.

The active site and tryptophan residues important for activity in EndoS were found to be conserved in EndoS2 even though the proteins are only 37% identical [13]. Despite this substantial difference in amino acid sequences, EndoS2 hydrolysed the glycan on IgG in a similar fashion to EndoS. In the chromatograms comparing the glycan profile of IgG generated by EndoS2 and PNGase F (Figure 5), a shift was observed that could be explained by the site of action. PNGase F is an amidase that cleaves between the asparagine residue and the first GlcNAc residue of the glycan, whereas EndoS2 cleaves after the first GlcNAc and thus leaves one GlcNAc with or without fucose attached to the protein backbone. Owing to lack of one reducing end GlcNAc in the EndoS2 glycan profile, there is a loss in resolution, which explains why the A2G1 peak could not be separated in the EndoS2 chromatogram, but can be seen as two separate peaks in the PNGase F profile. It has been argued previously that EndoS does not cleave bisecting glycans [41,42]. From the LC–MS data of the present study, we argue that EndoS2 cleaves all glycoforms present of human serum IgG, including bisecting glycans, since all peaks present in the PNGase F glycan profile could be found in the EndoS2 profile (Figure 5).

A striking difference between EndoS2 and EndoS was found when incubated with the human acute-phase protein AGP. The observed activity of EndoS2 was confirmed with LC–MS and revealed that EndoS2 specifically releases biantennary and sialylated structures of AGP (Figure 6). Again, the peaks annotated in the EndoS2 profile could be found in the PNGase F release with the difference of one GlcNAc residue. It is clear that EndoS2 only releases a fraction of the glycans present on AGP. EndoS2 does not cleave tri- and tetra-antennary glycans, with or without outer arm fucosylation, although they are present in great numbers on AGP.

The activity on IgG and AGP raised several questions regarding the specificity of EndoS2. To answer these, we tested the activity of EndoS2 on heat-denatured IgG and AGP, on other glycoproteins, in a chitinase assay and on a pool of free N-glycans. EndoS2 was only active on native IgG and AGP and we draw the conclusion that EndoS2 requires a protein–protein interaction with its substrates for glycan hydrolysis to occur. The activity of EndoS2 on AGP may be the result of reduced protein recognition, since early studies indicate sequence homology between IgG and AGP [43]. On glycoproteins with a completely different fold, we detected no activity with similar assays to the activity on IgG and AGP detected. It was therefore not surprising to find that EndoS2 had no general chitinase activity compared with a chitinase from *S. griseus*. Taken together, these data indicate that EndoS2 specifically interacts with protein folds including IgG and AGP. Furthermore, we dissected the glycoform specificity of EndoS2 by incubating the enzyme with the N-glycan pool from fetuin released by PNGase F and showed that EndoS2 hydrolysed only free biantennary structures with or without terminal sialylation. No bisecting glycans are present on fetuin, which explains why such structures are not present in the chromatograms. On the basis of our findings, we believe that EndoS2 is both site- and glycoform-specific which is a unique property of an endoglycosidase.

The hydrolysis of the glycan of IgG has been shown to have major consequences on the effector functions of the antibody by modulating the binding to FcγR [12]. Since EndoS and EndoS2 have similar hydrolysing activity on the glycan of IgG, both enzymes are expected to affect the functionality of this antibody. The functional consequence for AGP when biantennary sialylated glycans are cleaved off is unknown and lies beyond the scope of the present study.

The expression of EndoS2 was found to depend on the availability of carbohydrates in the bacterial culture medium. C-medium is a poor medium for GAS and expression of EndoS, could only be detected when GAS was grown in 50% diluted C-medium. Incubating the supernatants with IgG confirmed the previous work carried out with recombinant EndoS and a clear correlation between expression of EndoS and hydrolysis of the Fc-glycan on IgG confirmed this. The genes *scrb*, *scra* and *scrk*, surrounding ndoS2, are part of a sucrose utilization operon and this could explain the increase of EndoS2 expression when sucrose was added to the culture medium. The presence of glucose, galactose, GlcNAc or mannose completely inhibited expression of EndoS2 indicating that this enzyme is tightly regulated by a mechanism.
sensitive to the presence of carbohydrates. Research has indicated that the virulence of GAS is linked to the utilization of available carbohydrates via CcpA (catabolite control protein A), but in the present study of EndoS2, we can only hypothesize that CcpA is involved in the regulation mechanism [32]. This indicates that in the infection scenario, EndoS2 is strictly regulated and that the enzyme is used in an environment where nutrition is scarce, e.g. the human skin. This indicates further that the virulence of GAS is linked to the utilization of complex carbohydrates [44].

The present study shows that the endoglycosidase EndoS2 is conserved and uniquely present in GAS serotype M49. We show that EndoS2 hydrolysed all glycoforms on human serum IgG and biantennary and sialylated glycans on AGP. EndoS2 is secreted by GAS during starvation and the expression is linked to the carbohydrate composition of the culture medium. The enzymatic activity on two key players of the immune system argues that EndoS2 has a role in immunomodulation of the host that could potentially be linked to the pathogenesis of GAS serotype M49 infections.

AUTHOR CONTRIBUTION
Jonathan Sjögren and Mattias Collin conceived the study. Jonathan Sjögren performed experiments and drafted the paper. Weston Struve, Eoin Cosgrave and Pauline Rudd performed glycan analysis and contributed to the paper. Martin Stenvander performed phylogenetic analyses and contributed to the paper. Victor Nizet, Andrew Hollands and Maria Althorn provided material and valuable input on the text. All authors read and approved the final paper.

FUNDING
This work was supported by the Swedish Research Council [grant number 2010-57X-20240], the Åke Wiberg Foundation, the Alfred Österlund Foundation, the Petrus and Augusta Hedlund Foundation, the Clas Groschinsky Foundation, the Torsten och Ragnar Söderberg Foundation, the Swedish Society for Medicine, the Royal Physiographic Society in Lund, King Gustaf Vs 80 Years Fund, the Medical Faculty at Lund University, Swedish governmental funding for clinical research (ALF), and Hansa Medical AB. The funders had no role in the preparation of the paper or the decision to publish.

REFERENCES
1 Rudd, P., Elliott, T., Cresswell, P., Wilson, I. and Dwek, R. (2001) Glycosylation and the immune system. Science 291, 2370–2376
2 Raju, T. S. (2008) Terminal sugars of Fc glycans influence antibody effector functions of IgG. Curr. Opin. Immunol. 20, 471–478
3 Lux, A. and Nimmerjahn, F. (2012) Impact of differential glycosylation on IgG activity. Adv. Exp. Med. Biol. 780, 113–124
4 Lux, A., Aschermann, S., Bitburger, M. and Nimmerjahn, F. (2009) The pro and anti-inflammatory activities of immunoglobulin G. Ann. Rheum. Dis. 69, 192–196
5 Barb, A. W. and Prestegard, J. H. (2011) NMR analysis demonstrates immunoglobulin G N-glycans are accessible and dynamic. Nat. Chem. Biol. 7, 147–153
6 Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. and Dwek, R. A. (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. Annu. Rev. Immunol. 25, 21–50
7 Nizet, V. (2007) Understanding how leading bacterial pathogens subvert innate immunity to reveal novel therapeutic targets. J. Allergy Clin. Immunol. 120, 13–22
8 Carapetis, J., Steer, A., Mulholland, E. and Weber, M. (2005) The global burden of group A streptococcal diseases. Lancet Infect. Dis. 5, 685–694
9 Cunningham, M. W. (2000) Pathogenesis of group A streptococcal infections. Clin. Microbiol. Rev. 13, 470–511
10 Collin, M. and Olsén, A. (2001) EndoS, a novel secreted protein from Streptococcus pyogenes with endoglycosidase activity on human IgG. EMBO J. 20, 3046–3055
11 Collin, M., Svensson, M., Sjöholm, A., Jonsenius, J., Sjöbring, U. and Olsén, A. (2002) EndoS and SpdB from Streptococcus pyogenes inhibit immunoglobulin-mediated opsonophagocytosis. Infect. Immun. 70, 6646
12 Althorn, M., Olin, A. I., Nimmerjahn, F. and Collin, M. (2008) Human IgG/FcγR interactions are modulated by streptococcal IgG glycan hydrolysis. PLoS ONE 3, e1413
13 Althorn, M., Olsén, A. and Collin, M. (2008) EndoS from Streptococcus pyogenes is hydrolyzed by the cysteine proteinase SpeB and requires glutamic acid 235 and tryptophans for IgG glycan-hydrolyzing activity. BMC Microbiol. 8, 3
14 Collin, M. and Olsén, A. (2001) Effect of SpeB and EndoS from Streptococcus pyogenes on human immunoglobulins. Infect. Immun. 69, 7187–7189
15 Sjögren, J., Okumura, C. Y., Collin, M., Nizet, V. and Hollands, A. (2011) Study of the IgG endoglycosidase EndoS in group A streptococcal phagocyte resistance and virulence. BMC Microbiol. 11, 120
16 Nandakumar, K. S., Collin, M., Olsén, A., Nimmerjahn, F., Blom, A. M., Ravetch, J. V. and Holmdahl, R. (2007) Endoglycosidase treatment abrogates IgG arthritogenicity: importance of IgG glycosylation in arthritis. Eur. J. Immunol. 37, 2973–2982
17 Althorn, M., Brizeno, J. G., Baudino, L., Loo, C., Orsson, M. L., Iuzzi, S. and Collin, M. (2010) The IgG-specific endoglycosidase EndoS inhibits both cellular and complement-mediated autoimmune hemolysis. Blood 115, 5080–5088
18 Albert, H., Collin, M., Dudzak, D., Ravetch, J. V. and Nimmerjahn, F. (2008) In vivo enzymatic modulation of IgG glycosylation inhibits autoimmune disease in an IgG subclass-dependent manner. Proc. Natl. Acad. Sci. U.S.A. 105, 15005–15009
19 van Timmeren, M. M., van der Veen, B. S., Stegeman, C. A., Petersen, A. H., Hellmark, T., Collin, M. and Heeringa, P. (2010) IgG glycan hydrolysis attenuates ANCA-mediated glomerulonephritis. J. Am. Soc. Nephrol. 21, 1103–1114
20 Collin, M., Shannon, O. and Björk, L. (2008) IgG glycan hydrolysis by a bacterial enzyme as a therapy against autoimmune conditions. Proc. Natl. Acad. Sci. U.S.A. 105, 4265–4270
21 Berenghoustra, M., Molnarit, N., Santiago-Raber, M.-L., Weber, M. S., Merkler, D., Collin, M. and Lalive, P. H. (2012) IgG glycan hydrolysis by EndoS inhibits experimental autoimmune encephalomyelitis. J. Neuroimmun. 9, 209
22 Huang, W., Giddens, J., Fan, S.-Q., Toonstra, C. and Wang, L.-X. (2012) Chemoregulatory glycoengineering of intact IgG antibodies for gain of functions. J. Am. Chem. Soc. 134, 12308–12318
23 Goodfellow, J. J., Banah, K., Yamamoto, K., Bonomelli, C., Krishna, B., Harvey, D. J., Crispin, M., Scanlan, C. N. and Davis, B. G. (2012) An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. J. Am. Chem. Soc. 134, 8030–8033
24 McShan, W. M., Ferretti, J. J., Kasara, T., Suzorov, A. N., Lin, S., Qin, B., Jia, H., Kenton, S., Najaf, F., Wu, H. et al. (2008) Genome sequence of a niphogenic and highly transformable M49 strain of Streptococcus pyogenes. J. Bacteriol. 190, 7773–7785
25 Chausse, M. S., Ajdic, D. and Ferretti, J. J. (1999) The rgg gene of Streptococcus pyogenes N2131 positively influences extracellular SpeB production. Infect. Immun. 67, 1715–1722
26 Drummond, A. J., Suchard, M. A., Xie, D. and Rambaut, A. (2012) Bayesian phylogenetic analysis with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973
27 Henikoff, S. and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. U.S.A. 89, 10915–10919
28 Gernhard, T. (2008) The conditioned reconstructed process. J. Theor. Biol. 253, 769–778
29 Ahn, J., Jones, B., Yu, Y. Q., Rudd, P. M. and Gilar, M. (2010) Separation of 2-amino-2-deoxy and 2-amino-2-deoxy-2-fluoro-glycans via CcpA (catabolite control protein A), but in the present study of EndoS2, we can only hypothesize that CcpA is involved in the regulation mechanism [32]. This indicates that in the infection scenario, EndoS2 is strictly regulated and that the enzyme is used in an environment where nutrition is scarce, e.g. the human skin. This indicates further that the virulence of GAS is linked to the utilization of complex carbohydrates [44].

The present study shows that the endoglycosidase EndoS2 is conserved and uniquely present in GAS serotype M49. We show that EndoS2 hydrolysed all glycoforms on human serum IgG and biantennary and sialylated glycans on AGP. EndoS2 is secreted by GAS during starvation and the expression is linked to the carbohydrate composition of the culture medium. The enzymatic activity on two key players of the immune system argues that EndoS2 has a role in immunomodulation of the host that could potentially be linked to the pathogenesis of GAS serotype M49 infections.
Tarentino, A. L., Quinones, G., Changchien, L. M. and Plummer, Jr, T. H. (1993) Multiple endoglycosidase F activities expressed by Flavobacterium meningosepticum endoglycosidases F2 and F3: molecular cloning, primary sequence, and enzyme expression. J. Biol. Chem. 268, 9702–9708

Renzi, F., Manfredi, P., Marly, M., Moes, S., Jenö, P. and Cornelis, G. (2011) The N-glycan glycoprotein deglycosylation complex (Gpd) from Capnocytophaga canimorsus deglycosylates human IgG. PLoS Pathog. 7, 17

King, S. J., Hippe, K. R. and Weiser, J. N. (2006) Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by Streptococcus pneumoniae. Mol. Microbiol. 59, 961–974

Anumula, K. R. (2012) Quantitative glycan profiling of normal human plasma derived immunoglobulin and its fragments Fab and Fc. J. Immunol. Methods 382, 167–176

Baruah, K., Bowden, T. A., Krishna, B. A., Dwek, R. A., Crispin, M. and Scanlan, C. N. (2012) Selective deactivation of serum IgG: a general strategy for the enhancement of monoclonal antibody receptor interactions. J. Mol. Biol. 420, 1–7

Schmid, K., Kaulthmann, H., Isemura, S., Bauer, F., Emura, J., Motoyama, T., Ishiguro, M. and Nanno, S. (1973) Structure of α1-acid glycoprotein: complete amino acid sequence, multiple amino acid substitutions, and homology with the immunoglobulins. Biochemistry 12, 2711–2724

Shelburne, S. A., Davenport, M. T., Keith, D. B. and Musser, J. M. (2008) The role of complex carbohydrate catabolism in the pathogenesis of invasive streptococci. Trends Microbiol. 16, 318–325

Harvey, D. J., Merry, A. H., Royle, L., P Campbell, M., Dwek, R. A. and Rudd, P. M. (2009) Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. Proteomics 9, 3796–3801
SUPPLEMENTARY ONLINE DATA

EndoS2 is a unique and conserved enzyme of serotype M49 group A Streptococcus that hydrolyses N-linked glycans on IgG and α1-acid glycoprotein

Jonathan SJÖGREN*1,2, Weston B. STRUWE†3, Eoin F. J. COSGRAVE‡4, Pauline M. RUDD†, Martin STERVANDER‡, Maria ALLHORN*2, Andrew HOLLANDS§∥2, Victor NIZET*∥¶2 and Mattias COLLIN*2

*Department of Clinical Sciences, Division of Infection Medicine, Lund University, Biomedical Center B14, SE-22184 Lund, Sweden, †The National Institute for Bioprocess Research and Training, NIBRT, University College Dublin, Fosters Avenue, Mount Merrion, Blackrock, Co. Dublin, Ireland, ‡Molecular Ecology and Evolution Laboratory, Department of Biology, Lund University, Ecology Building, SE-223 62 Lund, Sweden, §School of Biological Sciences, University of Wollongong, Wollongong, NSW 2522, Australia, ¶Department of Pediatrics, University of California San Diego, 9500 Gilman Drive, Mail Code 0687, La Jolla, CA 92093, U.S.A., and §§Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, Mail Code 0687, La Jolla, CA 92093, U.S.A.

Figure S1 Phylogenetic reconstruction of 16S rRNA sequences, for inference of taxonomic relationship between taxa of the EndoS-like protein phylogeny, based on 1576 nucleotides

All internal nodes were supported by a Bayesian posterior probability (PP) of 0.99–1.0 except for the grouping of the Listeriaceae family bacterium, highlighted with a black squares signifying PP < 0.80. The scale bar indicates genetic distance. Note that the branches connecting the ingroup and outgroup are truncated owing to the large taxonomic difference (genetic distances stated below branches). For accession numbers, see Table S3.

1 To whom correspondence should be addressed (email jonathan.sjogren@med.lu.se).
2 Patents for the use of EndoS2 have been applied for by Genovis AB. Jonathan Sjögren, Andrew Hollands, Victor Nizet, Maria Allhorn and Mattias Collin are listed as inventors on the application that is pending.
3 Present address: Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3TA, U.K.
4 Present address: Waters Corporation, 34 Maple Street, Milford, MA 01757, U.S.A.

The sequences of Streptococcus pyogenes ndoS2 will appear in the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under accession numbers KC155346 (strain 3487-05), KC155348 (strain AP49), KC155347 (strain ACN49), KC155349 (strain AW1) and KC155350 (strain AW2).
Figure S2  HILIC–FLD–MS and exoglycosidase sequencing of AGP N-glycans

The fucose linkage position was confirmed as α(1,3) to galactose residues on tri- and tetra-antennary glycans and were detected as [M – 2H]$^{2-}$ ions 972.9 and 1074.5.

Table S1  E. coli and group A streptococcal strains used in the present study

| Strain          | Source/accession number | ndoS2 accession number |
|-----------------|-------------------------|------------------------|
| E. coli         |                         |                        |
| Top10           | Invitrogen N/A          |                        |
| BL21 pLysS      | Invitrogen N/A          |                        |
| GAS (serotype)  |                         |                        |
| NZ131 (M49)     | ASM1812v1 ACI61688      |                        |
| 3487-05 (M49)   | S. pyogenes, serotype M49 KC155346 |                |
| ACN49 (M49)     | Strain 3274-98 from Center for Disease Control and Prevention, A gift from Actinova KC155347 |            |
| AP49 (M49)      | S. pyogenes strain collection, Institute of Hygiene and Epidemiology, Prague, Czech Republic KC155348 |          |
| AW1 (M49)       | Nephritis strain H 9449 B from the collection of L. Wannamaker KC155349 |             |
| AW2 (M49)       | Strain H 5424 from the collection of L. Wannamaker KC155350 |               |
Table S2  Plasmids and primers used in the present study

Plasmids and primers used for recombinant expression and sequencing.

(a) Plasmids

| Name               | Source            |
|--------------------|-------------------|
| pGEX-5X-3          | GE Healthcare     |
| pGEX-ndoS2         | The present study |
| pGEX-ndoS2(E-L)    | The present study |
| pGEX-ndoS          | [1]               |
| pCR2.1             | Invitrogen        |

(b) Primers

| Name               | Sequence (5′→3′) |
|--------------------|------------------|
| ndoS2-out-R        | GCGCCTTCTGAGTCTAAAC |
| Seq38-R            | TTTGGGACCTAGCGCTAG |
| Seq42-R            | GAATTCAAGCTTACGT |
| Seq54-R            | TTAGCCCTTTTGGACAG |
| Seq15-F            | CCTGAAGAAGATGCTG |
| Seq17-F            | TGAGGACCTGAAAGAACG |
| Seq24-F            | GAGACGAAGAAGATGCTG |
| Seq28-F            | TCACCTGAGCTAAGCTAAG |
| ndoS2-F-BamHI      | CTGTAAGGATCCAGGAAGACTG |
| ndoS2-R-XhoI       | GAAACCTCGAGTCTTGGCTAGGACCTT |
| ndoS2(E-L)-F-BamHI | CTAGATTTGAATTCACAGGAATTCAGAAC |
| ndoS2(E-L)-R-XhoI  | GTCTGAAATTCCTGAGAATATCAATATCTAG |
Table S3: Accession numbers for EndoS-like protein and 16S rRNA sequences

For the EndoS-like protein dataset, outgroup sequences are indicated with an asterisk. For the EndoS2 sequences of the present study, see Table S1.

| Taxa                             | Strain                | GenBank® accession number | EndoS-like protein | 16S rRNA |
|----------------------------------|-----------------------|---------------------------|--------------------|----------|
| Actinomyces neuii                | BVG02395              | ZP_11040533               | EJP67240           |          |
| Beauveria bassiana               | ARSEF 2660            | EJP67240                  |                    |          |
| Bifidobacterium longum subsp. infantis | 157F                | YP_002332900               | AP010889           |          |
| Bifidobacterium longum subsp. infantis | ATCC 15697           | YP_002332900               | AP010889           |          |
| Bifidobacterium longum subsp. longum | 48B                   | EIJ2312                   |                    |          |
| Bifidobacterium longum subsp. longum | 1-6B                  | EIJ2312                   |                    |          |
| Brachybacterium faecium          | DSM-4810              | YP_003135977; YP_003155571| X91032             |          |
| Carnobacterium maltaromaticum    | ATCC 3586             | ZP_10279968               | JF749289           |          |
| Carnobacterium maltaromaticum    | LMA28                 | YP_006993899              |                    |          |
| Chryseobacterium meningosepticum | ATCC 12354            | AJ704541                  |                    |          |
| Cordyceps militaris             | CM01                  | EGX89767                  | H058075            |          |
| Corynebacterium diphtheriae      | 241                   | YP_006214375              | CP003421           |          |
| Corynebacterium diphtheriae      | 31                    | YP_006214375              | CP003421           |          |
| Corynebacterium diphtheriae      | 258                   | YP_006353466              | CP003540           |          |
| Corynebacterium diphtheriae      | Cpl62                 | YP_006437976              | CP003652           |          |
| Corynebacterium diphtheriae      | FRC41                 | YP_003784295              | CP002097           |          |
| Corynebacterium diphtheriae      | PAT10                 | YP_005691331              | CP002924           |          |
| Corynebacterium pseudotuberculosis | X81907                | YP_005304514              | CP003077           |          |
| Corynebacterium pseudotuberculosis | 314                   | YP_005645446              | CP003077           |          |
| Corynebacterium ulcerans         | 809                   | YP_005711595              | CP002790           |          |
| Corynebacterium ulcerans         | BR-AD22               | YP_006430749              | CP002791           |          |
| Elizabethkingia anophelis        | Ag1                   | ZP_09415888               | EF426425           |          |
| Elizabethkingia meningoseptica   | EB4                   | FLAME                     |                    |          |
| Enterococcus faecalis            | ATCC 29200            | ZP_04493022               | ACHK01000081       |          |
| Enterococcus faecalis            | DO1RF                 | CP002841                  |                    |          |
| Enterococcus faecalis            | VS83                  | NP_813917                 |                    |          |
| Enterococcus faecalis            | HER1044               | AAR20477                  |                    |          |
| Enterococcus faecalis            | DS5                   | ZP_05561082               |                    |          |
| Enterococcus faecalis            | ERV85                 | EJ95658                   |                    |          |
| Enterococcus faecalis            | PC1.1                 | ZP_0674687                |                    |          |
| Enterococcus faecalis            | RS08                  | EJ959420                  |                    |          |
| Enterococcus faecalis            | RT12                  | ZP_06630726               |                    |          |
| Enterococcus faecalis            | T1                    | ZP_05422239               |                    |          |
| Enterococcus faecalis            | T2                    | ZP_05422504               |                    |          |
| Enterococcus faecalis            | T3                    | ZP_05502539               |                    |          |
| Enterococcus faecalis            | T8                    | ZP_05560447               |                    |          |
| Enterococcus faecalis            | TX0012                | EFT96360; EFT95516; EFT95520 |                    |          |
| Enterococcus faecalis            | TX0104                | ZP_09850001               |                    |          |
| Enterococcus faecalis            | TX0109                | ZP_07567140               |                    |          |
| Enterococcus faecalis            | TX0309B               | EFT96360; EFT95516        |                    |          |
| Enterococcus faecalis            | TX0470                | ZP_07750646               |                    |          |
| Enterococcus faecalis            | TX0585                | ZP_07595326               |                    |          |
| Enterococcus faecalis            | TX0800                | ZP_07593514               |                    |          |
| Enterococcus faecalis            | TX1322                | ZP_05440574               |                    |          |
| Enterococcus faecalis            | TX1341                | EFT12187                  |                    |          |
| Enterococcus faecalis            | TX1342                | EFU15396                  |                    |          |
| Enterococcus faecalis            | TX1346                | EFU16620                  |                    |          |
| Enterococcus faecalis            | TX1467                | EGS58918                  |                    |          |
| Enterococcus faecalis            | TX2137                | EFT39634                  |                    |          |
| Enterococcus faecalis            | TX244                 | EFT91060                  |                    |          |
| Enterococcus gallinarum           | EG2                   | ZP_05648160               |                    |          |
| Enterococcus saccharolyticus     | 30.1                  | ZP_09110633               |                    |          |
| Hectococcus kunzii               | ATCC 51366            | ZP_09737336; ZP_09738693  |                    |          |
| Hectococcus kunzii               | 22                    | NR_029237                 |                    |          |
| Hectococcus kunzii               | M112                  | JN61738                   |                    |          |
| Lactobacillus raffinolacticus     | KCTC 2804             | ZP_09465258               |                    |          |
| Lactobacillus garvieae           | ATCC 49156            | YP_004779109              | AP003332           |          |
| Listeria species                 | ITU M-001             | ZP_09890591               | JQ287762           |          |
| Melissococcus plutonius          | ATCC 35311            | YP_004455741              | AP012200           |          |
### Table S3  Continued

| Taxa                          | Strain | GenBank® accession number | 16S rRNA |
|-------------------------------|--------|--------------------------|---------|
| Melissococcus plutonius       | DAT561 | YP_005320150              | AP012282|
| Paenibacillus alvei           | DSM 29 | ZP_10866627               | AJ320491|
| Paenibacillus dendritiformis  | C454   | ZP_09678120               |        |
| Paenibacillus popilliae       | ATCC 14706 | GAC41484                | AB073198|
| Streptococcus canis           | FSL Z3-227 | ZP_10274531             |        |
| Streptococcus equi subsp. equi| 4047   | YP_002745815              | FM204883|
| Streptococcus equi subsp. zooepidemicus | MGCS10565 | YP_002122753           | CP001129|
| Streptococcus equi subsp. zooepidemicus | ATCC 35246 | AEJ24585               | CP002904|
| Streptococcus equi subsp. zooepidemicus | H70    | YP_002745127              |        |
| Streptococcus ictaluri        | 707-05 | ZP_09127116; ZP_09126970  |        |
| Streptococcus ictaluri        | 706-05 | DQ462420                 |        |
| Streptococcus iniae           | 9117   | ZP_11067943               |        |
| Streptococcus pyogenes        | Aab49  | YP_00650262               | CP003068|
| Streptococcus pyogenes        | M1 GAS | NP_269818                 | AE006615|
| Streptococcus pyogenes        | MGAS315| NP_665372                 | AE014074|
| Streptococcus pyogenes        | MGAS5005| YP_282903                | CP000017|
| Streptococcus pyogenes        | MGAS6180| YP_280992                | CP000056|
| Streptococcus pyogenes        | MGAS8232| NP_607886                | AE009554|
| Streptococcus pyogenes        | MGAS9429| YP_597275                | CP000259|
| Streptococcus pyogenes        | MGAS10270| YP_599216              | CP000260|
| Streptococcus pyogenes        | MGAS10394| YP_0060848              | CP000003|
| Streptococcus pyogenes        | MGAS10750| YP_603093               | CP000262|
| Streptococcus pyogenes        | MGAS15252| YP_005389405             | CP003116|
| Streptococcus pyogenes        | str. Manfredo | YP_001127900        | AM295007|
| Streptococcus pyogenes        | NZ131  | YP_002286383              | CP000829|
| Streptococcus pyogenes        | ATCC 10782| ZP_07460044             |        |
| Streptococcus pyogenes        | HKU QMH11M0907901 | EIK41346             |        |
| Streptococcus pyogenes        | M49 591| ZP_00365754               |        |
| Streptococcus pyogenes        | CS101  | U78969_1*                |        |
| Streptococcus sp. group C     | CT16   | AEC53484                 |        |
| Trichosporon asahii var. asahii| CBS 2479 | EJ746459               |        |

### Reference

1 Collin, M. and Ölsén, A. (2001) EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. EMBO J. **20**, 3046–3055

Received 23 January 2013/10 July 2013; accepted 19 July 2013 Published as BJ Immediate Publication 19 July 2013, doi:10.1042/BJ20130126

© 2013 The Author(s)

The author(s) has paid for this article to be freely available under the terms of the Creative Commons Attribution Licence (CC-BY) (http://creativecommons.org/licenses/by/3.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.