Glycoconj J (2009) 26:675–689
DOI 10.1007/s10719-008-9223-8

Detection and characterization of a sialoglycosylated bacterial ABC-type phosphate transporter protein from patients with visceral leishmaniasis

Angana Ghoshal · Sumi Mukhopadhyay · Rodion Demine · Michael Forgber · Saulius Jarmalavicius · Bibhuti Saha · Shyam Sundar · Peter Walden · Chhabinath Mandal · Chitra Mandal

Received: 5 April 2008 / Revised: 23 September 2008 / Accepted: 15 December 2008 / Published online: 29 January 2009

Abstract We report the discovery and characterization of a glycosylated bacterial ABC-type phosphate transporter isolated from the peripheral blood mononuclear cell (PBMC) fraction of patients with visceral leishmaniasis (VL). Three disease-associated 9-*O*-acetylated sialoglycoproteins (9-*O*-AcSGPs) of 19, 56 and 65 kDa, respectively, had been identified and their purity, apparent mass and pI established by SDS-PAGE and isoelectric focusing. Western blot analyses showed that the 9-*O*-acetylated sialic acid is linked via α2→6 linkage to a subterminal *N*-acetylgalactosamine. For the 56 kDa protein, *N*- as well as *O*-glycosylations were demonstrated by specific glycosidase treatment and found to account for more than 9 kDa of the protein mass. The presence of sialic acids was further confirmed through thin layer chromatography, fluorimetric HPLC and electrospray ionization-mass spectrometry. The protein was identified by mass spectrometry and *de novo* sequencing of five tryptic fragments as a periplasmic ABC-type phosphate transporter of *Pseudomonas aeruginosa*. The amino acid sequences of the assigned peptides had 83–100% identity with the NCBI entry for a *Pseudomonas* transporter protein. Based on the recently reported X-ray structure of a human phosphate-binding protein, we predicted a 3D structural model for the 56 kDa protein using homology and threading methods. The most probable *N*- and *O*-glycosylation sites were identified by combinations of sequence motif-searching bioinformatics tools, solvent accessibility calculations, structural environment analyses and mass spectrometric data. This is the first reported glycosylation as well as sialylation of the periplasmic component of an ABC-type phosphate transporter protein and of one of few identified bacterial glycoproteins.

Electronic supplementary material The online version of this article (doi:10.1007/s10719-008-9223-8) contains supplementary material, which is available to authorized users.

A. Ghoshal · S. Mukhopadhyay · C. Mandal
Department of Infectious Disease and Immunology,
Indian Institute of Chemical Biology,
4 Raja S.C. Mullick Road,
Kolkata 700032, India

C. Mandal
Structural Biology and Bioinformatics Division,
Indian Institute of Chemical Biology,
4 Raja S.C. Mullick Road,
Kolkata 700032, India

R. Demine · M. Forgber · S. Jarmalavicius · P. Walden
Department of Dermatology, Charité-Universitätsmedizin Berlin,
Humboldt University,
10098 Berlin, Germany

B. Saha
Department of Tropical Medicine, School of Tropical Medicine,
Kolkata, India

S. Sundar
Department of Medicine, Institute of Medical Sciences,
Banaras Hindu University,
Varanasi, India

C. Mandal (✉)
Infectious Disease and Immunology Division,
Indian Institute of Chemical Biology,
4, Raja S. C. Mullick Road,
Kolkata 700032, India
e-mail: cmandal@iicb.res.in

Keywords *Pseudomonas aeruginosa* · Sialylated ABC-type phosphate transporter protein · *De novo* peptide sequencing · Molecular modeling · Visceral leishmaniasis

Abbreviations

9-*O*-AcSA	9-*O*-acetylated sialic acid
9-*O*-AcSGP	9-*O*-acetylated sialoglycoprotein
ABC	ATP-binding cassette
BSA	bovine serum albumin
BSM	bovine submandibular mucin
Da	dalton
DIG	digoxigenin
DSA	*Datura stramonium* agglutinin
DMB	1,2-diamino-4,5-methylenedioxybenzene
ddH_2O	double distilled water
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assays
ESI-MS	electrospray ionization mass spectrometry
FITC	fluorescein isothiocyanate
GalNAc	*N*-acetylgalactosamine
GNA	*Galanthus nivalis* agglutinin
GP-56	56 kDa *O*-acetylated sialoglycoprotein
GT	glycosyltransferase
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
HPBP	human phosphate-binding protein
IEF	isoelectric focusing
kDa	kilodalton
LD	*Leishmania donovani*
MAA	*Maackia amurensis* agglutinin
MALDI-TOF	matrix-assisted laser desorption/ionization
MS	time-of-flight mass spectrometry
Mab	monoclonal antibody
MS	mass spectrometry
GlcNAc	*N*-acetylglucosamine
Neu5Ac	*N*-acetylneuraminic acid
*O*AcSGP	*O*-acetylated sialoglycoproteins
PA	*Pseudomonas aeruginosa*
PB	potassium phosphate buffer
PBMC	peripheral blood mononuclear cells
PBS	phosphate-buffered saline
PDB	protein data base
PE	phycoerythrin
PMF	peptide mass fingerprint
PMSF	phenylmethylsulfonyl fluoride
PNA	peanut agglutinin
PSD	post-source decay
RIA	radioimmunoassay
RMS	root mean square
RMSD	root mean square deviation
SA	sialic acid
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNA	*Sambucus nigra* agglutinin
TBS	Tris-buffered saline
TLC	thin layer chromatography
TM	transmembrane
VL	visceral leishmaniasis

Introduction

Sialic acids (SA) constitute a family of *N*- and *O*-substituted 9-carbon-carboxylated monosaccharides, the major member being *N*-acetylneuraminic acid (Neu5Ac) [1]. Most frequent are *O*-acetylations at positions C-7/8/9 that generate a family of *O*-acetylated sialoglycoconjugates (*O*-AcSGs) [2]. Molecular modifications of sialic acids by *O*-acetylation and alterations of their patterns have been found to be of great importance in different pathological conditions and hence have become the focus of intense research in our and other laboratories around the world [3, 4]. There is growing evidence that a number of physiological and pathological processes involving cell–cell adhesion, signaling, differentiation and metastasis may be associated with *O*-acetylated sialoglycotopes [3]. The levels of SA may signal acuteness of disease and, thereby, indicate the need to combat the causative agents or processes [2, 3]. We used the specificity of the snail lectin Achatinin-H for glycoproteins with terminal 9-*O*-acetylated sialic acid (9-*O*-AcSA) derivatives linked in α2-6 linkage to subterminal *N*-acetylgalactosamine (GalNAc) to identify and analyze these specific sialoglycotopes [5]. In previous studies, we have detected 9-*O*-AcSGPs on erythrocytes of patients with visceral leishmaniasis (VL), not on normal erythrocytes and these sialylations correlate directly with the degree of complement-mediated hemolysis [6]. Likewise, VL-associated 9-*O*-AcSGPs corresponding to molecular masses of 19, 56 and 65 kDa were found on peripheral blood mononuclear cells (PBMC) of patients and not on PBMC of normal individuals [7].

VL is caused by the intracellular parasitic kinetoplastid protozoa *Leishmania donovani* (LD) [8]. An estimated 12 million humans are infected, with an incidence of 0.5 million cases of VL per year [9]. Approximately 50% of the world's VL cases occur on the Indian subcontinent, 90% of them in Bihar alone [9]. Because of the association of 9-*O*-AcSGPs with VL pathology, we investigated the over-all density of these sialoglycotopes on PBMC and report the purification, identification and molecular characterization of the PBMC-associated 56 kDa glycoprotein, here named GP-56. Both *N*- and *O*-linked glycosylations of this protein were demonstrated. Tryptic fragments of the protein were sequenced by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) with post source

decay (PSD) mass spectrometry that led to its identification as the periplasmic component of an ABC-type phosphate transporter of *Pseudomonas aeruginosa* (PA), which has homology with a recently discovered human plasma phosphate-binding protein (HPBP) [10]. We used the published structure of HPBP to develop a 3-D structure for GP-56 by molecular modeling and the resulting model to predict the glycosylation sites.

Materials and methods

Clinical samples

The study involved clinically confirmed VL patients (20 males, 10 females, median age, 30 years, range, 10–50 years) of the Kala-Azar Medical Research Center, Muzaffarpur and the School of Tropical Medicine, Kolkata. Diagnosis of VL was based on microscopic demonstration of *Leishmania* amastigotes in splenic aspirates according to WHO recommendations [11]. The median white blood cell count of the VL patients was 1.2×10^6/mL (range $1.0–1.4\times10^6$/mL). PBMC of patients with high parasitemia and spleen sizes of 5 to 8 cm were prepared from venous blood (5–6 mL) by density centrifugation (Ficoll-Hypaque, Amersham Pharmacia, Uppsala, Sweden), washed thoroughly at 4°C and processed within 2–12 h of collection. The Institutional Human Ethical Committee had approved the study and samples were taken with the consent of the donors, patients, or their parents or guardians.

Detection reagents and flow cytometry

All chemicals were purchased from Sigma (St. Louis, USA) unless stated otherwise. The lectin, Achatinin-H, was affinity-purified from the hemolymph of African giant land snail *Achatina fulica* using bovine submandibular mucin (BSM), known to contain a high percentage of 9-*O*-AcSAs, as affinity matrix [5]. It was conjugated with fluorescein isothiocyanate (FITC) for flow cytometry and coupled to Sepharose 4B [12] for purification of 9-*O*-AcSGPs. The monoclonal antibodies (Mabs) including anti CD3, CD13, CD16 and CD19 were purchased from Pharmingen (San Diego, CA). For flow cytometry, PBMC (1×10^6/100 μL) were suspended in RPMI-1640 supplemented with 2 mM glutamine, gentamicin and 10% heat-inactivated human AB serum and stained on ice for 1 h with FITC-Achatinin-H and phycoerythrin (PE)-anti-CD Mabs along with appropriate isotype controls. The cells were then washed, fixed in paraformaldehyde (1%) and analyzed with a FACS Calibur flow cytometer (Becton Dickinson, Mountain View, CA). The data were analyzed with the Cell Quest software (Becton Dickinson). The presence of 9-*O*-AcSGPs on the PBMC was confirmed by flow cytometry with FITC-Achatinin-H after 3 h at 37°C pretreatment or no pretreatment with 9-*O*-acetyl hemagglutinin esterase of influenza C virus known to specifically cleave off *O*-AcSAs [13].

Purification of 9-*O*-AcSGPs from PBMC

PBMC (1×10^7) of clinically confirmed VL patients (n=30) with 80–90% 9-*O*-AcSGP-positive cells were extensively washed with ice-cold phosphate buffered saline (PBS), suspended in cold homogenization buffer Tris–HCl (50 mM, pH 7.2; TB) containing 1 mM ethylenediaminetetraacetic acid (EDTA) with a protease inhibitor cocktail consisting of 1 mM phenylmethylsulfonyl fluoride (PMSF), 10 μg/mL leupeptin and 10 μg/mL aprotinin, and homogenized on ice with a Teflon pestle in glass tube. Cell debris was removed by centrifugation at 1,000×g for 5 min at 4°C and the supernatant centrifuged at 21,000×g for 30 min at 4°C. The membrane pellet was resuspended in solubilization buffer, *i.e.* TB with 150 mM NaCl (TBS), glycerol (10%), 10 mM EDTA, 1.5 mM $MgCl_2$ and protease inhibitor cocktail as above by sonication for 15 s intermittently on ice [12, 14]. The suspension was adjusted to 0.1% SDS and centrifuged. The supernatant containing 0.8 mg protein was dialyzed against TBS containing 0.03 M $CaCl_2$ and 0.02% sodium azide (TBS–Ca^{2+}) at 4°C and then passed through an Achatinin-H Sepharose 4B affinity column (1×2 cm; 1.0 mg Achatinin-H/mL gel) equilibrated with TBS–Ca^{2+} at 4°C. After washing off unbound material with the same buffer, bound 9-*O*-AcSGPs were eluted at 25°C with TBS containing 0.04 M sodium citrate, dialyzed at 4°C against TBS and stored at −70°C.

Electrophoresis

Affinity-purified 9-*O*-AcSGPs (30 μg/lane) were analyzed by SDS-PAGE with 10% acryl amide in a Minigel apparatus (BioRad, Hercules, CA) and stained with Coomassie brilliant blue R-250 [12]. For Western blot analysis, the proteins were electroblotted onto nitrocellulose at 100 V for 2 h. After blocking the nitrocellulose with 10% BSA in TBS, the membrane was probed with Achatinin-H (150 μg/mL) in presence of 0.03 M Ca^{2+}. After washing, the blots were incubated with rabbit anti-Achatinin-H antibodies (1:400) at 4°C and washed. The antigen–antibody complexes were then detected with horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (1:10,000; Cappel, St. Louis, MO). To obtain pure protein, Coomassie-stained bands were excised from the electrophoresis gels and the proteins eluted using an Electro-Eluter Model 422 (Bio Rad) according to the manufacturer's instruction. IEF was carried out with 9-*O*-AcSGPs and purified GP-56 (1.5 μg) in capillary tubes with ampholine

polyacrylamide gels (4%), pH range 3.5–10.0 using Mini-PROTEAN II tube cell apparatus (Bio-Rad). The samples were focused at a constant voltage of 400 V for 6 h. The gels were then fixed, washed and silver-stained [12]. The isoelectric points (pI) of the individual proteins were determined as a function of their migration from the cathode using standard pI markers for a pH range of 3.5 to 10 (Bio-Rad).

Analysis of carbohydrates

The presence of terminal sugars on 9-*O*-AcSGPs was analyzed by a Digoxigenin (DIG) enzyme assay using a commercially available DIG-Glycan differentiation kit (Roche Applied Science, Mannheim, Germany) with the plant lectins *Galanthus nivalis* agglutinin (GNA), *Sambucus nigra* agglutinin (SNA), *Maackia amurensis* agglutinin (MAA), *peanut agglutinin* (PNA) and *Datura stramonium* agglutinin (DSA) according to the manufacturer's protocol. Total SA contents were measured fluorimetrically using pure SA as standard. The SA of the glycoproteins was mildly oxidized with sodium metaperiodate and the resulting formaldehyde determined by reaction with acetylacetone and ammonium acetate, which produces a fluorogen, detected at 510 nm upon excitation at 410 nm with an F-4010 spectrofluorimeter (Hitachi, Tokyo, Japan). 8- or 9-*O*-acetylated sialic acids were not oxidized under these conditions and are quantified by measuring the amount of fluorogen before and after de-*O*-acetylation with 0.1 M NaOH [15]. To differentiate *N*- or *O*-linked glycosylation, glycoproteins were deglycosylated with specific glycosidases using the deglycosylation kit as per manufacturer's instructions (Roche Applied Science). Briefly, the GP-56 was incubated overnight with *Arthrobacter ureafaciens* neuraminidase (specific activity, 0.5 mU) in denaturation buffer (100 mM sodium phosphate, 10 mM EDTA 0.5% (*w*/*v*) Triton X-100, 0.05% (*w*/*v*) SDS, 1% (*v*/*v*) 2-mercaptoethanol, pH 8.6) at 37°C. The reaction mixture was heated at 100°C for 3 min and centrifuged. Then, reaction buffer at pH 7.2 containing CHAPS and *N*-glycosidase F was added and the sample incubated overnight at 37°C. For *O*-glycosidase (1–2 mU) or combination of *N*- and *O*-glycosidase treatment, neuraminidase treated samples were incubated overnight with the enzymes in potassium phosphate buffer (20 mM) containing EDTA (10 mM), Triton X-100 (0.5%) and SDS (0.05%) at pH 7.3 (PB) and 37°C. For IEF, the samples were desialylated overnight with neuraminidase (0.5–1 mU) in PB at 37°C.

Detection of sialic acids of GP-56 through TLC

Gel-eluted GP-56 (25 μg) was hydrolyzed with 4 M propionic acid for 4 h at 80°C to release sialic acid. It was subsequently passed onto Dowex 50WX8 (100–200 mesh) cation exchange and Dowex 2X8 (200–400 mesh) anion exchange columns. The eluted free sialic acids were collected, lyophilized and resuspended in double distilled water (ddH_2O, 40 μL), spotted (5 μg/10 μL) onto TLC plates (Merck KGaA, Germany) and run in 1-propanol/H_2O (7:3 *v*/*v*). Sialic acids released from BSM along with Neu5Ac (Sigma) were used as internal standards. Plates were developed by spraying with orcinol/HCl/$FeCl_3$ spray reagent and heated at 180°C for 20 min [16].

Fluorimetric HPLC for estimation of sialic acids of GP-56

Gel-eluted GP-56 was lyophilized and subjected to mild acid hydrolysis with 200 μL of 4 M propionic acid for 4 h at 80°C to release glycosidically bound Neu5Ac. Reaction mixture was cooled on ice for 15 min. An aliquot of this sample was derivatized with 1,2-diamino-4,5-methylenedioxybenzene (DMB) for fluorimetric reverse-phase HPLC analysis [12, 13, 16]. Analysis of the derivatized sialic acids was done on an RP-18 column (LichroCART 125-4 HPLC-cartridge, 5 μm; Merck, Germany) using isocratic elution with water/acetonitrile/methanol (84:9:7, *v*/*v*/*v*) at a flow rate 1 mL/min and compared with authentic sialic acid purified from BSM. Fluorescence was detected using an excitation and emission wavelength of 373 and 448 nm, respectively.

ESI-Q-TOF-MS of DMB derivatized sialic acids obtained from GP-56

The mass spectrometric data was acquired on a Q-Tof micromass spectrometer (Waters Corporation, Massachusetts, USA). The fraction corresponding to DMB derivatized Neu5Ac from GP-56 was collected after fluorimetric HPLC, dried, dissolved (water/acetonitrile/methanol; 84:9:7, *v*/*v*/*v*) and analyzed directly via ESI-MS at a flow rate of 10 μL/min. The ion spray voltage was operated at 5 kV. The tuning and calibration solution consisted of an equimolecular mixture of polypropylene glycol 425, 1,000, and 2,000 in 50/50/0.1 H_2O/methanol/formic acid (*v*/*v*/*v*), NH_4OAc (1.0 mM). Spectra were acquired by scanning from m/z=80 to m/z=800 [17].

Binding assays

To analyze the kinetics of sialoglycoproteins binding by Achatinin-H, the lectin was iodinated with ^{125}I (Amersham Pharmacia Biotech, Sweden) by the chloramine T method [18] yielding a specific activity of 1.5×10^6 cpm/μg. For measuring the binding to cells, PBMC (2×10^6/150 μL per tube) from five patients were suspended in 10 μL RPMI-1640 medium containing $CaCl_2$ (0.3 M) and bovine serum albumin (BSA, 0.2%), pH 7.2 (reaction buffer). The cells were incubated with increasing doses of ^{125}I-Achatinin-H for 90 min at 4°C and washed. The bound radioactivity in the cell pellets was measured with a gamma counter

(Electronic Corporation, India). For controlling the specificity of binding, a 100-fold excess of unlabeled Achatinin-H was added. The data were processed by Scatchard analyses and the number of binding sites for Achatinin-H per cell was calculated from the intersection of the curve with the *X*-axis and the dissociation constant, Kd, obtained by dividing the number of receptors per cell by the bound/free ratio derived from the *Y*-axis intercept of the Scatchard plot [18].

Protein/peptide mass spectrometry

The measurements to identify the glycoprotein were done with a Bruker-Daltonik reflector MALDI-TOF mass spectrometer Reflex IV with PSD capabilities (Bruker Daltonik, Bremen, Germany). The samples were prepared by dried-droplet procedure using 2,5-dihydroxybenzoic acid as matrix. Calibration was done externally with a mixture of Angiotensin I, Angiotensin II, Substance P, Bombesine, ACTH clip 1–17 and ACTH clip 18–39. The accuracy of peptide mass measurement was 0.1 Da and that of peptide fragment mass measurement 0.5 Da. 10–100 pmol of peptide samples were used for analysis. Tryptic fragments were generated by digestion of GP-56 from a SDS-PAGE gel band overnight in ammonium hydrocarbonate (5 mM) with 10 ng trypsin (Promega, Mannheim, Germany) per sample. For sequence analysis, selected tryptic fragments of the protein were subjected to PSD and fragment mass detection. MALDI-PSD data was processed using the Sequit! 4.0 *de novo* sequencing software with subsequent searches in the NCBI database version July 25, 2006 using BLAST version 2.2.13 [19]. MS/MS ion searches and peptide mass fingerprint (PMF) analyses were done with MASCOT using the same settings as for Sequit!.

Database searches, prediction of glycosylation sites

For searches in sequence databases, standard BLAST with the NCBI database version as cited was used. Sequence alignments were done with CLUSTALW [20]. For prediction of *N*-glycosylation sites we employed PROSITE software [21]. In addition, NetNglyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc) was used for the prediction of *N*-glycosylation sites. Predictions of *O*-linked glycosylation sites were done with the neural network-based NetOGlyc 3.1.

3-D structural modeling and refinement

3-D structural modeling was done with homology modeling software SWISS-MODEL [22] and with the threading method-based LOOPP software package [23]. LOOPP uses sequence alignment, sequence profile, threading, secondary structure and exposed surface area predictions, and merges them into a single score, generating atomic coordinates based on alignment of few homologous template structures. Both approaches were based on structures available in the protein data bank (PDB). 3-D structure alignments were done with ABGEN [24] to select core structures as starting scaffolds of target proteins for subsequent refinement. The starting structures were refined using the software suite Insight II (2005) (Accelrys, San Diego, CA) equipped with DISCOVER for energy minimization and molecular dynamics simulation. Structural optimization involved molecular dynamics simulation followed by energy minimization (100 steps each of steepest descent and conjugate gradient methods) of the lowest energy conformation using cff91 force field. A typical dynamics run for structure regularization consisted of 10,000 steps of 1 fs after 1,000 steps of equilibration with a conformation sampling of one in 10 steps at 300 K. At the end of the molecular dynamics simulation, the conformation with the lowest potential energy was picked for energy minimization with the ANALYSIS module of Insight II followed by a new cycle of refinement. The energy-minimized structure was subjected to various validation tests to check the quality of the refined segments. The conformations of each identified segment were regularized separately by this combination of molecular dynamics and energy minimization while keeping the rest of the molecule fixed by applying positional constraints. This was continued until satisfactory conformational parameters were obtained. Structures, thus obtained, were energy-minimized as before keeping all the atoms of the protein free. The resulting model was taken as the initial crude model and used for side chain regeneration and backbone conformation refinement. Side chain regeneration was done with SCWRL 3.0 [25]. This is the most recent version of the SCWRL using graph theory to solve combinatorial problems encountered in side-chain predictions and allows sophisticated energy functions for incorporation of side-chain flexibility around rotameric positions. PROCHECK [26] was used for checking the Ramachandran's plots of regularized individual loops as well as the overall protein models for their stereochemical quality. Amino acids whose Φ and Ψ dihedral angle pair fell outside the core and allowed zones of the Ramachandran's plot were identified and grouped into segments. Backbone conformations of these segments were refined to drive the amino acids to adopt proper Φ–Ψ values by molecular dynamics and energy minimization while applying constraints on the dihedral angles of the amino acids with Φ–Ψ combinations mapping to the core and allowed zones of the plot. For molecular dynamics and energy minimization, the CHARM module of Insight II was used with CharmM27 force field. Conformations with correct Φ–Ψ combinations were picked from stored conformations of the molecular dynamics runs. This procedure of backbone conformation refinement was done for each identified segment separately until all segments were covered and the Ramachandran's plot checked with

PROCHECK. Overall structural validation of the refined structures was done with MOLPROBITY [27] which measures all-atom contacts in terms of clash scores (number of atoms having atom pair overlaps ≥0.4 Å per 1000 atoms) and calculates rotamer outliers of the side chains. Our in-house software MODELYN [28] was used for calculation of dihedral angles of protein backbones as well as standard deviations of bond lengths and bond angles.

Solvent accessibility and stereo-chemical environment of potential glycosylation sites

Differential solvent accessibilities of all atoms grouped into amino acids of the refined models were calculated with the ACCESS software [29] and the most probable *N*- and *O*-linked glycosylation sites were identified by the degree of surface exposure. X-ray crystallographic structures with known *N*-linked glycosylation taken from PDB were superimposed with the homology model of the target protein. Then the distances between the nitrogen (ND2) of Asn and oxygen (OG1) of the Ser/Thr residue at the third position of the consensus sequence Asn-Xaa-Ser/Thr were measured for assessing the stereo-chemical environment. Likewise, for *O*-linked glycosylation the X-ray structure of an *O*-glycosylated protein was superimposed and distances between two oxygen atoms (OG1) of Thr measured and compared with the model. The capacity of accessible sites of the modeled protein to be glycosylated was checked by fitting in a covalently linked glycotope.

Computational platforms

ABGEN, PROCHECK and SCWRL 3.0 were run on a FUEL workstation of Silicon Graphics, Inc. in the IRIX 6.5 environment. Insight II equipped with the bio-science modules BIOPLOYMER, DISCOBER, CHARMm, ANALYSIS, AFFINITY and others was run on a FUEL workstation as well as, or on an ALTRIX high-end server of Silicon Graphics, Inc. in the IRIX 6.5 environment. MODELYN and ANALYN were run on IBM-compatible PCs in Windows environment. SWISS-MODEL, PROSITE, LOOPP, CLUSTALW, MOLPROBITY, NetNGlyc and NetOGlyc were used at the web-servers at the internet sites http://swissmodel.expasy.org//SWISS-MODEL.html, http://www.expasy.org/prosite/, http://cbsuapps.tc.cornell.edu/loopp.aspx, http://www.ebi.ac.uk/clustalw/ and http://molprobity.biochem.duke.edu/, www.cbs.dtu.dk/services/netnGlyc, http://www.cbs.dtu.dk/services/NetOGlyc.

Statistical analysis

Results are expressed as mean±SD for individual sets of data. Each experiment was performed two to three times and the results shown are representatives of each set of experiments. One or two-tailed *t* tests for significance were performed where applicable. $P<0.05$ was considered significant.

Results

9-*O*-acetyl-sialoglycoproteins in VL patients

The PBMCs of VL patients were tested by flow cytometry for 9-*O*-AcSGPs using FITC-Achatinin-H. Consistently, 80–90% of the PBMC were positive for 9-*O*-AcSGPs. Two-color flow cytometry with FITC-Achatinin-H and PE-labeled leukocyte subpopulation-specific antibodies showed that 9-*O*-AcSGPs were present at the surfaces of T cells (CD3, 60.43±5.78%), NK cells (CD16, 11.76±3.53%), monocytes (CD13, 47.82±0.86%) and B cells (CD19, 19.54±0.26%). Increased 9-*O*-AcSGPs were also detected on erythrocytes of these patients as determined by flow cytometry, enzyme-linked immunosorbent assay (ELISA) and hemagglutination assay [15]. The sera of the patients had high-levels of anti-9-*O*-AcSGPs antibodies as detected by ELISA using BSM as coating antigen [30]. Moreover, the sera also had high levels of parasite-specific antibodies as determined by parasite-specific ELISA with crude parasite lysate as coating antigen [30]. The average numbers of 9-*O*-AcSGP molecules determined by binding measurements with ^{125}I-Achatinin-H and Scatchard analysis done with the PBMC of five VL patients were found to be 2.8×10^7/cell (Fig. 1a). For confirmation of the binding specificity, an excess of up to 100-fold unlabelled Achatinin-H was added and the apparent dissociation constant (Kd) calculated from the inhibition of binding of the radio labeled probe as 3.7±0.02 pM.

Purification and characterization of 9-*O*-AcSGPs

PBMC of clinically confirmed VL patients ($n=25$) with high parasitemia were selected for purification of 9-*O*-AcSGPs from the membrane fractions by affinity chromatography on Achatinin-H-Sepharose 4B. The yield of purified 9-*O*-AcSGP from 1.2×10^7cells was 0.79±0.12 mg, which corresponds to 42.7±2.15% of the total membrane proteins (1.85±0.25 mg). SDS-PAGE of the purified 9-*O*-AcSGPs (Fig. 1b, lane 2) revealed three VL-associated sialoglycoproteins with apparent masses of 19, 56 and 65 kDa, respectively. These three glycoproteins are not found on the PBMC of healthy individuals unlike the two other sialoglycoproteins of 36 and 144 kDa, as has been demonstrated by Western blot analysis earlier [7]. The specificity of these preparations was confirmed by Western blot analysis with Achatinin-H as reagent for 9-*O*-AcSGPs (Fig. 1b, lane 3). The first lane of Fig. 1b shows protein

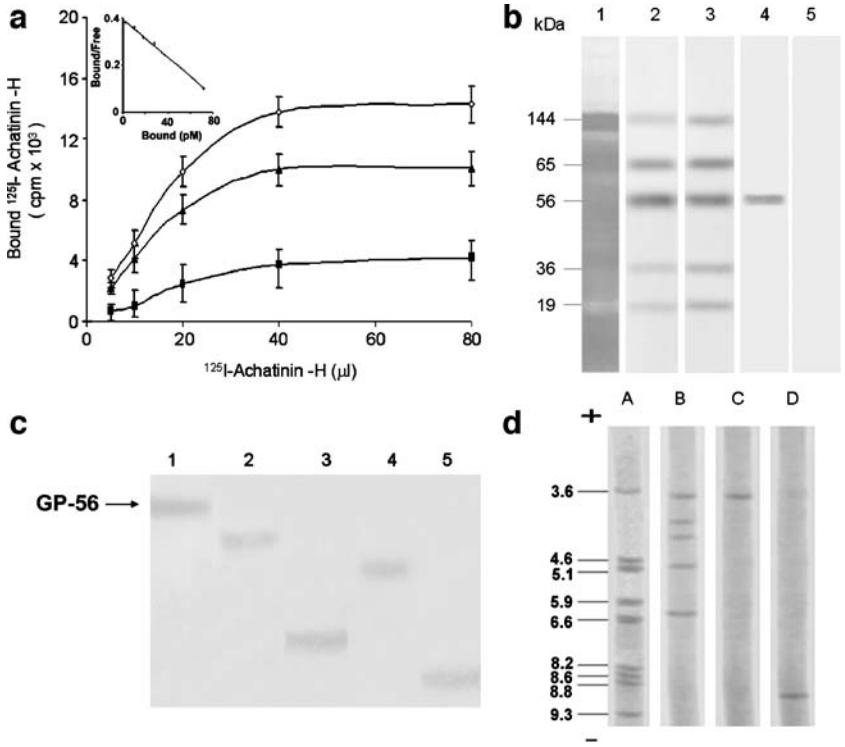

Fig. 1 Characterization of the 9-*O*-AcSGPs present on PBMC of VL patient. **a** Binding of ^{125}I-Achatinin-H. A fixed amount of PBMC (2×10^6/150 µL) was incubated with increasing amounts of ^{125}I-Achatinin-H. For evaluating the specific nature of binding, a 100-fold excess of unlabeled Achatinin-H was added. Bound and unbound Achatinin-H was separated at 4°C and the bound radioactivity determined. Specific binding (*filled triangles*) was calculated as difference between total binding (*open diamonds*) and binding in presence of a 100-fold excess of unlabeled Achatinin-H (*filled squares*). *Inset* Scatchard plot of the binding data. The results are expressed as mean±SD of data from triplicate experiments. **b** SDS PAGE (10%) and Western blot of 9-*O*-AcSGP purified from PBMC of VL patients. Crude membrane preparation of PBMC (10 µg, *lane 1*) from VL patients and 30 µg purified 9-*O*-AcSGPs (*lane 2*) were electrophoresed on SDS-PAGE. The gel was fixed and stained with Coomassie brilliant blue R-250. Affinity purified 9-*O*-AcSGPs (30 µg) from PBMC of VL patients after SDS-PAGE was subjected to Western blot analysis with Achatinin-H (*lane 3*). GP-56 was electro eluted from SDS-PAGE and a sample (5 µg) reanalyzed by SDS-PAGE (*lane 4*). Binding of Achatinin-H was completely abolished by de-*O*-acetylation of the blot by alkali treatment (*lane 5*). **c** SDS-PAGE of GP-56 before and after treatment with neuraminidase, *N*- and *O*-glycosidase. Affinity purified GP-56 from PBMC of VL patients was electro eluted from SDS-PAGE (*lane 1*) and was treated with neuraminidase from *Arthrobacter ureafaciens* (*lane 2*) followed by *N*-glycosidase F (*lane 3*), *O*-glycosidase (*lane 4*) or a combination of *N*- and *O*-glycosidase (*lane 5*) and further analyzed by SDS-PAGE. **d** Isoelectric focusing of 9-*O*-AcSGP and purified GP-56 from PBMC of VL patients. 9-*O*-AcSGPs (1.5 µg, *lane B*), GP-56 electro eluted from SDS-PAGE (*lane C*) and GP-56 after treatment with *Arthrobacter ureafaciens* neuraminidase (*lane D*) was analyzed by IEF with a pH gradient of 3.5 to 10. *Lane A* shows pI markers

staining of the crude membrane extract. For further characterization, GP-56 was eluted from the gel, quantified and re-analyzed by SDS-PAGE (Fig. 1b, lane 4). The amount of GP-56 was 66.36±1.22 µg which corresponded to 8.40±1.15% of the total affinity-purified 9-*O*-AcSGPs. Treatment of the 9-*O*-AcSGPs with alkali prior to SDS-PAGE completely abolished the binding by Achatinin-H further supporting the specificity of the detection of *O*-acetylated glycoproteins (Fig. 1b, lane 5). To further characterize the glycosylation of GP-56, the purified molecule (Fig. 1c, lane 1) was treated with neuraminidase from *Arthrobacter ureafaciens* followed by glycosidases. Treatment with neuraminidase resulted in a shift 1.03 kDa (Fig. 1c, lane 2) corresponding to a band of 54.97 kDa. Subsequent treatment with *N*-glycosidase F resulted in a shift of the band to an apparent molecular mass of 49.3 kDa, corresponding to a reduction by 6.7 kDa (Fig. 1c, lane 3). After *O*-glycosidase treatment, a shift by 2.13 kDa was detected resulting in a molecular mass of 53.87 kDa (Fig. 1c, lane 4). Desialylation and deglycosylation with *N*- and *O*-glycosidase yielded a protein band with an apparent mass of 47 kDa (Fig. 1c, lane 5) indicating that the glycosylations account for about 9 kDa of the total mass of GP-56. The 9-*O*-AcSGPs associated with PBMC of VL patients were further investigated by IEF (Fig. 1d) which produced five distinct bands within a pI range of 3.7 and 6.3, indicating five different sialoglycoproteins with a similar type of glycotopes (lane B). The distinct band confirms their homogeneity. The purified GP-56 showed a single band with an acidic pI of 3.7 (lane C). Neuraminidase treatment of this protein resulted in a single band with a pI value of about 8.9 (lane D) proving the presence of SA. Lane A shows the pI markers.

To establish the linkage-specific arrangement of the terminal sugars present on the purified 9-*O*-AcSGPs, we used a DIG glycan differentiation kit with various lectins of known specificities for different terminal sugar moieties (Fig. 2a). For the purified 9-*O*-AcSGPs the binding was 4.0 fold higher with SNA (specific for α2→6 linked Neu5Ac) than with MAA (specific for α2→3 linked Neu5Ac). The corresponding densitometric scores were 62,313±1,749 versus 15,222±1,041 ($P<0.0001$). Concordantly, the densitometric scores for GP-56 binding to SNA were 8.74 fold higher than to MAA (55,169±1,102 versus 6,308±137, $P<0.0001$). In contrast, both the 9-*O*-AcSGPs and the purified GP-56 showed low binding to GNA, which is specific for terminal Man (α1→3), (α1→6) and (α1→2) Man, binding being 20,211±1,055 and 10,212±1,023, respectively. Similarly low were the bindings to DSA, which is specific for Gal (β1→4) GlcNAc and PNA, which is specific for Gal (β1→3) GalNAc. Fluorimetric analysis of the purified 9-*O*-AcSGPs demonstrated 316±4.5 μg SA, which corresponds to 40.0% of the 9-*O*-AcSGPs purified by Achatinin-H affinity chromatography from 1.2×10^7 cells. 81±11.16% of this total SA was 9-*O*-acetylated. For GP-56, 29.53±1.82 μg SA was detected which corresponds to 44.0% of total GP-56, of which 84.0±9.2% was 9-*O*-acetylated (Fig. 2b).

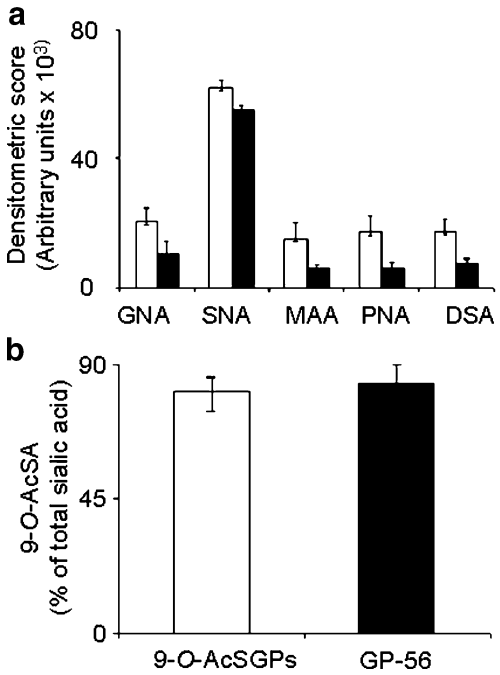

Fig. 2 Characterization of the terminal sugars on 9-*O*-AcSGPs and purified GP-56. **a** Differentiation of the terminal glycans by dot blot analyses. For determination of the terminal linkage-specific carbohydrates of 9-*O*-AcSGPs and GP-56 using different lectins in a DIG-linked enzyme assay, densitometric scoring of the scanned blots was performed. The intensities of the spots of 9-*O*-AcSGPs (*open squares*) and GP-56 electro eluted from SDS-PAGE (*filled squares*) were quantified densitometrically and represented in arbitrary units using the Master Totallab Software, version 1.11 (Amersham Pharmacia Biotech, Sweden). The results are shown as mean±SD. **b** Fluorimetric quantitation of 9-*O*-acetylated sialic acid on purified 9-*O*-AcSGPs and GP-56 from PBMC of VL patients. The *O*-acetyl groups of 9-*O*-AcSGPs (*open squares*) and GP-56 (*filled squares*) were oxidized with sodium metaperiodate before and after saponification and the resulting formaldehyde quantified by reaction with acetyl acetone and ammonium acetate by fluorimetric determination of the fluorescent product. The fractions (in percent) of 9-*O*-acetylated sialic acids (mean±SD) were calculated by subtracting fluorescence signal measured for the SA from the signals obtained after de-*O*-acetylation. Pure SA was used as a standard

The presence of sialic acid and *O*-acetylated sialic acid on GP-56 was further substantiated by several analytical techniques. When the glycosidically bound sialic acids of GP-56 were separated by acid hydrolysis and spotted onto the TLC plate, both sialic acid and *O*-acetylated sialic acid were detected by development with orcinol reagent (Fig. 3a). The R_F values corresponded to standard Neu5Ac and free sialic acids purified similarly from BSM.

To confirm the presence of SA in GP-56, freed SA was derivatized with DMB and analyzed by fluorimetric HPLC. Typically, the fluorimetric HPLC chromatogram of acid hydrolysates of GP-56 exhibited a well-resolved intense peak, which coincided with that of Neu5Ac (Fig. 3b ii). Also a peak comigrating with Neu5,9Ac$_2$ was obtained (Fig. 3b ii) comparable with the profile of sialic acids from BSM used as internal standard (Fig. 3b i).

For further confirmation of the presence of Neu5Ac (Fig. 3c) and Neu5,9Ac$_2$ (Fig. 3d) in GP-56, ESI-MS studies of DMB derivatized sialic acid was performed. The fraction corresponding to Neu5Ac was collected after fluorimetric HPLC, dried and analysed via ESI-MS. The ESI-MS spectra of Neu5Ac demonstrated the characteristic two groups of ions for Neu5Ac. The first group corresponds to the molecular ion (*m*/*z* 426, $[M+H]^+$), and its sodium adduct (448, $[M^+Na]^+$). The fragmentation also produced the characteristic second group of ions at *m*/*z* 408, $[M+H-18]^+$, corresponding to a first dehydration reaction, and ions (a), (b), (c), and (d) at *m*/*z* 313, 295, 283, and 229 in Fig. 3c. Similarly, the ESI-MS spectra of Neu5,9Ac$_2$ was identified by the characteristic ions at *m*/*z* at 450, 468, 490 corresponding to $[M+H-18]^+$, $[M+H]^+$, $[M+Na]^+$ respectively (Fig. 3d).

Identification of GP-56

For identification, the 56 kDa band was excised from a Coomassie-stained SDS-PAGE gel with Achatinin-H affinity-purified glycoproteins, destained and treated with trypsin. The fragments were extracted from the gels and analyzed by MALDI-TOF mass spectrometry given as a ESM Figure S1 (part A). Using the MASCOT software, the resulting PMF spectrum was compared to profiles of tryptic fragments generated *in silico* from all proteins in the NCBI sequence

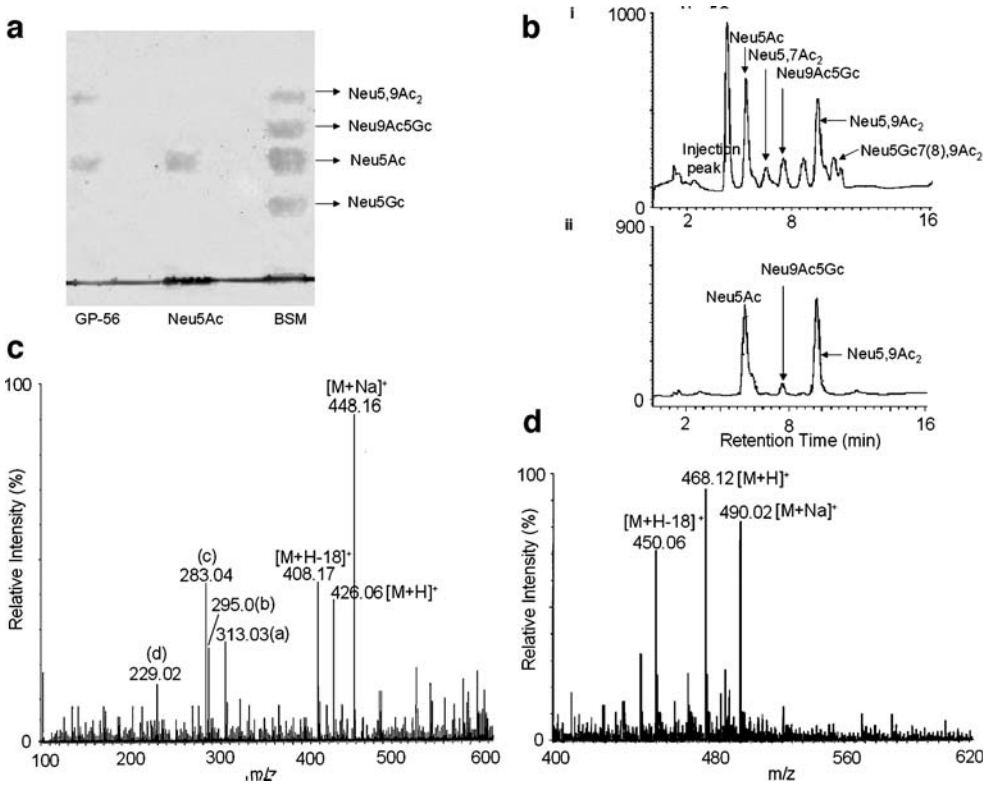

Fig. 3 Identification of sialic acids in GP-56 by analytical methods. **a** A representative TLC profile of liberated sialic acids of affinity purified, gel eluted GP-56. Glycosidically bound sialic acids of gel-eluted GP-56 (25 μg) were subjected to acid hydrolysis and purified by passing through Dowex cation and anion exchange columns. The eluted free sialic acids were lyophilized, dissolved and spotted onto a TLC plate. The TLC was developed with orcinol/HCl spray reagent as described in materials and methods. In parallel, free sialic acids released from BSM and commercially available Neu5Ac was similarly processed serving as internal standards were used for comparison. **b** Determination of sialic acids by fluorimetric HPLC. Glycosidically bound sialic acids from gel eluted GP 56 were subjected to acid hydrolysis, derivatized with DMB and analyzed using HPLC as described in materials and methods. (*i*) A representative profile of HPLC of fluorescent derivatives of free sialic acids from BSM. (*ii*) A representative chromatogram of fluorescent derivatives of free sialic acids from GP-56. **c** and **d** ESI-MS mass spectrum of Neu5Ac and $Neu5,9Ac_2$ of GP-56. The fractions corresponding to DMB derivatized Neu5Ac and $Neu5,9Ac_2$ from GP-56 were collected after fluorimetric HPLC dried and analysed via ESI-MS. The ion spray was operated at 5 kV. The spectra were acquired as described in Materials and Methods. A representative mass spectrum of Neu5Ac (**c**) and $Neu5,9Ac_2$ (**d**) are shown

databases without that a match was found. Five of the fragments with the masses 965, 1,088 (ESM Figure S1, part B), 1,304 (ESM Figure S1, part C), 1,314, and 1,698 Da (ESM Figure S1, part D) were subjected to MALDI-PSD mass spectrometry and the resulting fragment spectra analyzed with Sequit!, a database-independent *de novo* sequencing software, and MASCOT, a database-dependent peptide fragment fingerprint sequencing software. The complete amino acid sequences of all five fragments could be determined with the Sequit! software (ESM Figure S1, Table 1). When compared to the protein sequences of the NCBI sequence database, all five sequences independently matched the same and only one protein sequence, an ABC-type transporter of PA (GI: 116048595, Table 1). The identified protein was a phosphate transporter with conserved PstS domain or phosphate-specific transport system, a periplasmic component essentially responsible for transport and metabolism of inorganic phosphate ion. No other protein was indicated by the searches. For the two fragments with 965 and 1,314 Da these comparisons gave exact matches. Only these exactly matching sequences were identified by the MASCOT software. For the 1,088, 1,304 and 1,698 Da fragments, the sequences deviated by one or two amino acids from the database entry. In the case of the 1,088 Da fragment this was an Ala to Asp amino acid exchange at sequence position 360, for the 1,304 Da fragments the two amino acids exchanges Pro to Thr at position 153 and Ser to Ala at position 155 and for the 1,698 fragment the amino acid exchange Asn to Arg at position 123, which creates a new trypsin cleavage site. In addition to the amino acid difference within the sequenced part of the protein, a change of the amino acid to a positively charged amino acid Arg or Lys must be assumed for sequence position 107 immediately preceding the sequence of the 1,698 Da fragment, as otherwise the trypsin digest of the protein could not have produced the identified fragment. The 1,826 Da fragment is

Table 1 Identification of GP-56 by mass spectrometric *de novo* sequencing and peptide mass analyses

No.	Predicted mass[a] [Da]	Observed mass[b] [Da]	Sequence[c]	Sequence position[d]	GI: 116048595 sequence[e]	Method of identification[f]		
						De novo Sequit!	PFF MASCOT	Peptide mass
1	965.5414	965.5700	SGAITVVYR	158–166	SGAITVVYR	✓	✓	✓
2	1,044.5877	1,088.4400	FVPLPDSWK	355–363	FVPLPASWK	✓		
3	1,100.5118	1,100.5300	NVHWAGSDSK	78–87	NVHWAGSDSK			✓
4	1,314.6172	1,314.6200	SESSGTTELFTR	167–178	SESSGTTELFTR	✓	✓	✓
5	1,316.6594	1,303.9310	LTDWSQITGAGR	146–157	LTDWSQIPGSGR	✓		
6	[g]	1,698.0840	LIQVPSVATSVALPFR	108–123	LIQVPSVATSVALPFN	✓		
7	[g]	1,826.1640	LIQVPSVATSVALPFRK	108–124	LIQVPSVATSVALPFNK			✓
8	1,974.9919	1,975.0300	AAFLNNDYTKFVAGTTNK	60–77	AAFLNNDYTKFVAGTTNK			✓

S. Sequence of the *Pseudomona aeruginosa* ABC-type phosphate transporter, periplasmic domain, GI: 116048595 with the identified tryptic fragments of GP-56. Sites where amino acid exchanges have created new trypsin cleavage sites: exchange of P to either K or R in position 107, N to R exchange in position 123 identified by sequencing the 1698 Da fragment are denoted by an asterisk. All identified deviations from the GI: 116048595 sequence are indicated with bold face letter. The predicted *N*-glycosylation site NETN (92–95) and the two predicted *O*-glycosylation sites T273 and 306 are indicated with bold face and underlined letter

```
1
MYKRSLIAASLSVAALVSAQAMADINGGGATLPQQLYQEPGVLTAGFAAYIGVGSGNGKAAFLNNDYTKFVAGTTNKNVHWAGSDSKLSKTNETNPYLSA
                                                        AAFLNNDYTKFVAGTTNK
                                                                    NVHWAGSDSK
101   *               *
HGSAWGPLIQVPSVATSVALPFNKSGSNAVNFADVNTLCGVFSGRLTDWSQIPGSGRSGAITVVYRSESSGTTELFTRFLNASCSSTLEGGTFAITTSFG
       LIQVPSVATSVALPFR                    LTDWSQITGAGR         SESSGTTELFTR
       LIQVPSVATSVALPFRK                              SGAITVVYR
201
SSFSGGLPAGAVSAQGSQAVMNALNAAQGRITYMSPDFAAPTLAGLDDATKVAQVRGVSPAPANVSAAIGAVTPPTTAQRSDPNNWVPVFAATANPNDPS

301
VRPYPTSGYPILGFTNLIFSQCYANATQTQQVRDFFTRHYGATANNDTAITNHRFVPLPASWKLAVRQSFLTSTNNLYIGHSNVCNGIGRPL
                                                        FVPLPDSWK
```

[a] Masses of tryptic fragments predicted from the sequence of the *Pseudomonas aeruginosa* ABC-type phosphate transporter, periplasmic domain, GI: 116048595 (MASCOT)
[b] Mass of the tryptic fragments determined by MALDI-TOF mass spectrometry
[c] Sequence of the tryptic fragment of GP-56
[d] Sequence position of the tryptic fragment of GP-56 according to the GI: 116048595 sequence
[e] Sequence of the fragments of the *Pseudomona aeruginosa* ABC-type phosphate transporter corresponding to the identified tryptic fragment of GP-56
[f] Identification of the tryptic fragment of GP-56 by complete database-independent *de novo* sequencing using the Sequit! Software, peptide fragment fingerprint using MASCOT, peptide mass including the mutations identified with this work
[g] The indicated fragments could not be predicted from the database entry GI: 116048595 because of mutation framing or in this fragment that have created new trypsin cleavage sites

an extension of the 1,698 Da fragment by a C-terminal Lys. This fragment could now be identified by incorporating the new sequence of the 1,698 Da fragment into the database sequence for the protein. Once GP-56 was identified as the ABC-type phosphate transporter of PA recorded in the NCBI database, two more fragments could be assigned to the same protein by their masses of 1,100 Da and 1,975 Da. Thereby, eight masses of the tryptic digest of GP-56 could be assigned to the bacterial phosphate transporter. The sequence of the PA periplasmic ABC-like phosphate transporter plus the sequences of the identified fragments are shown in Table 1 with the variant amino acids in boldface letters. The sequence coverage of the five sequenced fragments plus the three additional fragments identified by their fragment masses was 27% of the full length of the protein sequence. However, the five completely sequenced fragments could be matched unequivocally by sequence comparison thus identifying GP-56 as a variant of the PA periplasmic ABC-like phosphate transporter listed in the NCBI database.

BLAST searches with the amino acid sequence of the PA ABC-like phosphate transporter as query sequence yielded a number of proteins with conserved PstS (ABC-type phosphate transport system, periplasmic component) domain. The sequence of the protein, we identified, showed a 92% identity with the hypothetical periplasmic binding protein PaerP_01000347 of PA7 and 73% identity with a phosphate ABC transporter, periplasmic phosphate binding protein of *Pseudomonas fluorescens*. The periplasmic

component of another ABC-type phosphate transport system of PA showed only 45% sequence identity. A close similarity of GP-56 was also found with the HPBP, which has 65% sequence identity. The other eukaryotic protein with PstS domains and similarity to GP-56 was p27SJ of *Hypericum perforatum* with 73% sequence identity. Using the sequences of GP-56 and HPBP, we searched the whole human and mouse genome databases for other homologues and found no other sequences with significant similarity besides the above which excluded the possibility that GP-56 is a human protein and by all likelihood, GP-56 is a protein of *Pseudomonas aeruginosa*.

3-D structural modeling of the GP-56

We used two approaches to predict 3-D structural models for GP-56. First, homology modeling based on the recently released (January 2007) X-ray structure of the human phosphate-binding apolipoprotein (HPBP), PDB code 2CAP. A BLAST search with the sequence of GP-56 showed 65% amino acid identity with HPBP, 74% positive score, 6% gaps and an expect value of 2×10^{-110} implying a high confidence in sequence match.

Second, we employed a threading method-based approach using LOOPP software, which compares secondary conformations of proteins of known structures for building an initial model for the target sequence. With the input of the GP-56 sequence, five models were predicted from the structures of the PDB entries 1QUJ, 1A55, 1QUK, 1IXH and 1A40. These five proteins are all of the class of periplasmic phosphate-binding and transport proteins of *E. coli* and contain a defined phosphate-binding domain. Multiple sequence alignment of the sequences of the five proteins and GP-56 with CLUSTALW showed that all five were of the same size and cover the same segments as GP-56. The model based on 1QUJ had the best score. The other four models were superimposed onto this model to identify structural segments common to all five. The RMS deviations for the identified core structures ranged between 0.47 and 0.66 Å and included 55.5% of the Cα atoms indicating close similarity of the core structures of the models but deviations in a number of loops. The loops were identified for the 1QUJ-based model and, in order to obtain the best loop conformation, subjected to regularization by molecular dynamic simulation and energy minimizations. The resulting initial threading-based model was used for subsequent analysis and refinement of the backbone and side chain conformations. Using the PROCHECK software [26], additional refinement, showed 73.7% of the backbone residues in the most favored core regions, 26.3% in allowed regions and none in unfavorable regions of Ramachandran's plot. Finally, the side chain conformations were regenerated using SCWRL3.0. The same analyses and refinements were done for the backbone conformations of the 2CAP-(HPBP)-based homology model. Initially, 87.9% and 11.4% of the Φ–Ψ pairs were in the favorable core and allowed regions, respectively, and only three amino acids (0.7%) in unfavorable regions of the Ramachandran's plot. After backbone refinement, 87.6% and 12.4% of the Φ–Ψ pairs were in the core and allowed regions, respectively. The side chain conformations were then regenerated on the refined backbone structure as before.

Clash scores and rotamer outliers of the refined threaded and homology models were computed using MOLPROBITY and compared with those of the X-ray structure of HPBP (2CAP). The clash scores of the 3D structures of the threading and homology models, with respective values of 3.92 and 3.83, were comparable to that of the X-ray structure 2CAP (7.10%) suggesting good structural compatibility with all-atom contacts. With 4.78% and 4.90% rotamer outliers for the threading and homology models, respectively, the side chain conformations were in good agreement with the 1.08% of the X-ray structure of HPBP. The RMS deviations from the respective standard values of all bond lengths and bond angles were 0.0185 Å and 3.3° for the threading model, 0.0161 Å and 2.6° for the homology model, and 0.0268 Å and 2.6° for the X-ray structure 2CAP indicating good general structural properties of the models.

The secondary structures and their arrangements in the homology model of GP-56 are displayed in Fig. 4a and compared to those of the X-ray structure of HPBP (not shown). Superimposition of the modeled structure onto the X-ray structure of HPBP with respect to the Cα atoms gave an RMSD of 0.101 Å for 342 (87%) of the 392 Cα atoms. As the homology model was derived from the X-ray structure of HPBP that has 65% amino acid identity with GP-56 such close resemblance of the two structures is expected. The overall folding pattern of the two structures is also very similar. Both have elongated structures with two similar globular domains connected by a hinge-like region. The two-disulphide bonds present in HPBP are conserved in the modeled structure of GP-56. Major differences are in loops, where the number of amino acids varies between the two structures. Of the 4 prominent loops of HPBP, loop 2 [10] is completely missing from the structure of GP-56 (arrow in Fig. 4a). Computation of the electrostatic potentials at the surface of the homology-modeled GP-56 with the MOLMOL software revealed a predominantly neutral stretch at the front face of the molecule and an intensely positively charged region at the reverse side (not shown).

Prediction of *N*- and *O*-linked glycosylation sites of GP-56

As we had established that GP-56 is *N*- and *O*-glycosylated, we opted for the identification of the glycosylation sites.

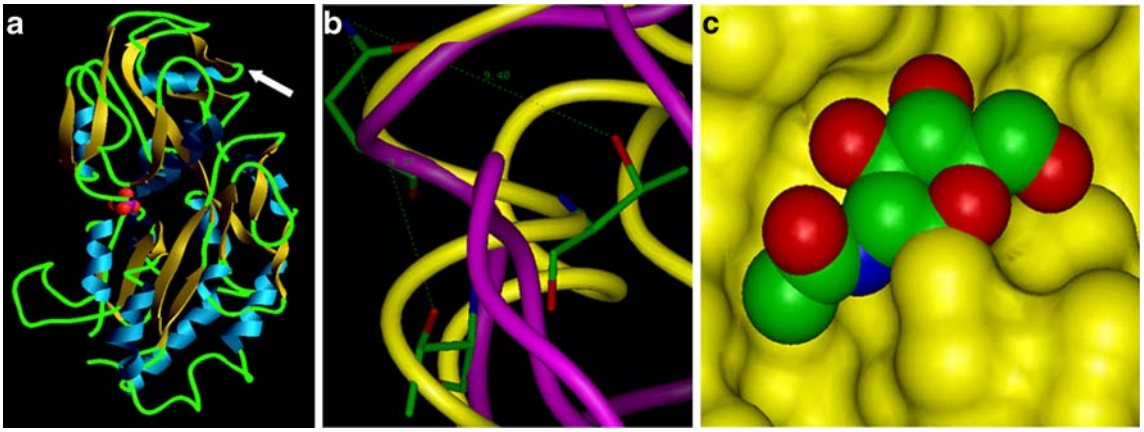

Fig. 4 3D structural modeling and suitability of the stereo-chemical environment of the predicted, accessible glycosylation sites. **a** 3D structural of the homology model of GP-56. The molecular model of GP-56 is shown as *ribbon* representation with the phosphate ion as space filling model colored by atom (*red* for oxygen, *pink* for phosphor atom). The secondary structures are shown with different color codes: *bronze* for β-sheets, *blue* for α-helices and *green* for loops. The *arrow* indicates a missing elongated loop that was observed in X-ray structure. **b** *N*-glycosylation site in the homology model of GP-56. Superposition of the relevant atoms at the *N*-glycosylation site Asn 92 of the homology model of GP-56 with the corresponding atoms of the X-ray structure 1NC4 with known *N*-glycosylation at Asn-88. The Asn residues (optimally superimposed) and the equivalent accessory Thr residues of both the structures are shown in *stick* model with the atoms colored *red* for oxygen, *blue* for nitrogen and *green* for carbon. The neighboring protein backbone is shown in *ribbon* representations for the homology model of GP-56 (*yellow*) and X-ray structures (*pink*). **c** Proper fit of a sugar moiety to the most probable predicted site for *N*-glycosylationin the homology model of GP-56. Space-filling representations of GlcNAc attached to Asn 92 residue, the most probable *N*-glycosylation site of GP-56 and the equivalent Asn 88 site in the X-ray structure of 1NC4. The atoms of the sugar moiety are colored *red* for oxygen, *blue* for nitrogen and *green* for carbon. The rest of the protein structures are shown as Connolly surface colored in *yellow* for the homology model of GP-56

Using the programs PROSITE and NetNglyc 1.0 we independently identified six potential *N*-glycosylation sites that all match Asn-containing consensus sequences required for such modifications (Table 2). With the NetOglyc 3.1 software, also six *O*-linked glycosylation sites were predicted, all containing Thr, with scores above the threshold value of the program of 0.5. As both *N*- and *O*-linked glycosylations are mediated by glycosyl transferases, the glycosylated residues need to be accessible at the surface of the proteins. We, therefore, calculated the solvent accessibilities of the predicted glycosylation sites using the software ACCESS. Only one of the predicted *N*-linked glycosylation sites, Asn-Glu-Thr-Asn (92–95), was highly accessible. The accessible surface of the Asn 92 was calculated to be 114.3 Å^2. For all other predicted *N*-glycosylation sites the accessible surfaces are less than half of this value (Table 2) identifying Asn-Glu-Thr-Asn (92–95) as the most probable *N*-glycosylation site. Of the six predicted *O*-linked glycosylation sites only two, Thr 273 and Thr 306 are sufficiently accessible at the surface of the protein. Their solvent-accessible surface areas are 141.1 and 122.0 Å^2, respectively (Table 2). For the predicted glycosylation sites Asn-Glu-Thr-Asn (92–95), Thr 273 and Thr 306 no corresponding tryptic fragments were identified (Table 1), which is consistent with the proposed modifications of these fragments.

The hydroxyl group Ser or Thr in the third position of the *N*-glycosylation consensus sequence Asn-Xaa-Ser/Thr is believed to mediate the glycosyl transfer by polarizing

Table 2 Solvent accessibility of *N*- and *O*-glycosylation sites in the homology model of GP-56

N-glycosylation sites[a]	Accessibile surface area [Å^2][b]	*O*-glycosylation sites[b]	Accessibile surface area [Å^2][b]
Asn-Glu-Thr-Asn [92–95]	114.3	Thr 250	34.3
Asn-Lys-Ser-Gly [123–126]	56.9	*Thr 273*	141.1
Asn-Ala-Ser-Cys [181–184]	70.4	Thr 276	82.8
Asn-Val-Ser-Ala [264–267]	38.3	Thr 277	107.9
Asn-Ala-Thr-Gln [325–328]	49.3	Thr 293	105.0
Asn-Asp-Thr-Ala [346–349]	1.9	*Thr 306*	122.0

[a] The predictions of potential glycosylations sites were based on the NCBI database sequence for the periplasmic component of the ABC-type phosphate transporter of PA, GI: 116048595

[b] In case of *N*-glycosylation the Asn residues and in case of *O*-glycosylation the Thr residues with more than 110 square Å solvent-accessible surface areas were taken as potential glycosylation sites and are shown in italicized letters

the glycosyl acceptor residue in the activation complex through hydrogen bonding. This function requires proper positioning of the hydroxyl group of the accessory Ser or Thr side chains with respect to the nitrogen of the Asn side chain (ND2). To establish whether these conditions are fulfilled in the predicted *N*-glycosylation site Asn-Glu-Thr-Asn (92–95), we calculated the distance between the side chain nitrogen atom of Asn in the first and oxygen atom of Thr in the third position of the motif. This distance was compared with the distances of the equivalent atoms in the X-ray structures 1NC4 and 1AYO for which *N*-glycosylations are established. The distance between ND2 and Thr-OH in the homology model of GP-56 was 9.40 Å compared to 9.37 and 9.44 Å in the X-ray structures 1NC4 and 1AYO, respectively. In order to confirm the suitability of the stereochemical environment of the glycosylation site, we superimposed the relevant atoms of 1NC4 and the homology model of GP-56 and found the environments to be very similar in both structures (Fig. 4b). Finally, to assess the fit of a sugar moiety to the predicted *N*-glycosylation site Asn-Glu-Thr-Asn (92–95), an *N*-acetylglucosamine (GlcNAc) was modeled into the structure of GP-56 using the structure of the GlcNAc residue at an *N*-glycosylation site of 1NC4. This GlcNAc structure was transferred to the *N*-linked glycosylation site of the GP-56 model and the structural parameters were optimized by molecular dynamics simulation and energy minimization. As shown in Fig. 4c, the sugar moiety fits well into the predicted site GP-56.

For *O*-glycosylation sites, no dependence on the neighboring amino acids in the sequence is known. To assess the suitability of the structural environment of the predicted *O*-glycosylation sites Thr 273 and Thr 306 of GP-56, we superimposed the local structures of these sites with the local structures of the two known *O*-glycosylation sites in the X-ray structure 1LSL, Thr 432 and Thr 489, and calculated the distances between the two sets of equivalent oxygen atoms. By applying torsion around the CB-OG single bond of the homology model, the distances between the corresponding oxygen atoms at the *O*-glycosylation sites could be brought to as close as 0.02 and 0.03 Å, respectively, without causing any conflicting contacts with other atoms (not shown).

Discussion

We report here the identification of a glycosylated GP-56 as a periplasmic component of an ABC-type phosphate transporter of PA from the PBMC of VL patients (Fig. 1b, Table 1, ESM Figure S1). Firstly, sequencing of five tryptic fragments of GP-56 independent of database information [19] evidenced a match with the same protein (ESM Figure S1). In addition, three more tryptic fragments were assigned to the same protein by their fragment masses. Although the sequence coverage by the eight assigned fragments is, with 27%, borderline, with five complete sequences of tryptic fragments all fitting the same protein the identification is unambiguous. The sequence of the isolated protein is, however, not exactly identical to the corresponding NCBI database entry GI: 116048595. Amino acid variations and side chain modifications account, in parts, for the low sequence coverage of the PMF, which is the typical method for protein identification by proteomics. Hence, the identification was only possible by *de novo* sequencing of tryptic fragments, a much more definite approach than PMF. The sequence differences in a few amino acids could indicate that GP-56 is a closely related protein or just that it is a variant of the database entry protein. Since the database entry is from complete genome sequencing, the second is the more likely option and it may be envisaged that GP-56 is an allele (maybe homologue) of the ABC phosphate transporter of PA listed in the genome database.

Secondly, the identity, homogeneity and glycosylation of GP-56 with regard to sialylation and 9-*O*-acetylation was reconfirmed by several biochemical and analytical techniques (Figs. 1, 2, and 3). It is a new protein that had not been purified, characterized or cloned for biochemical studies before with no prior report on its glycosylation status. It is, to the best of our knowledge, the first report of the presence of *O*-AcSAs on a bacterial protein and it is one of the few bacterial glycoproteins reported so far. In PA, another *O*-linked glycan was found in pilin [31]. However, PA also possesses a sialic acid-like sugar called pseudaminic acid that has an *O*-acetyl group at C-4 position [32].

The identification followed by subsequent confirmation of a sialylated glycoprotein of PA, leaves us with an open-ended question on the presence of sialylation machinery and status of SA in the bacteria. In fact, the genes for key enzymes for glycosylation are found in the genome of PA, like a cluster of pel genes encoding putative glycosyltransferases [33]. The possible involvement of sialidase, in colonization of PA in the respiratory epithelium has been reported [34]. Considering, the present state of the ever-expanding field of sialobiology, it is probable that PA can either synthesize sialic acids or may scavenge from the environment. However, till date there are no reports on the presence of sialylation machinery and is a matter of future research.

Using the canonical sequence motif for *N*-glycosylation, Asn-X-Thr/Ser-X, and, an AN*N*-based algorithm for *O*-glycosylation, six potential glycosylation sites of each class were identified (Table 2). One of the predicted *N*-glycosylation sites, *i.e.* Asn-Lys-Ser-Gly (123–126) could be excluded because, as found by MS, Asn (123) is altered to Arg in GP-56. In the absence of an X-ray structure, to substantiate or exclude the predicted *N*- and *O*-glycosylation sites, we modeled its structure by homology modeling

based on the X-ray structure of HPBP [10]. Secondly, a threading approach on the basis of known X-ray structures of six *E. coli* proteins was used. After iterative refinements, we obtained two comparable models of similar quality. We do, therefore, feel that the models are valid representations of the structure of GP-56 and suited for testing the predicted glycosylation sites. The homology model (ESM Figure S1, part A) was slightly superior and, therefore, used for the subsequent analyses.

Calculations of the solvent accessible surfaces of the amino acids as a measure of accessibility for glycosyltransferases narrowed down the predicted sites to three, Asn-Glu-Thr-Asn (92–95) for *N*-, and Thr 273 and Thr 306 for *O*-glycosylation. For *N*-glycosylations, the hydroxyl groups of Thr or Ser in the third position of the sequence motif need to be in defined proximity to the nitrogen of the Asn side chain. This condition was fulfilled for the predicted Thr 94 (ESM Figure S1, part B). Whereas, for *O*-glycosylation sites, we superimposed the atom arrangements around known *O*-glycosylations onto the predicted sites for the GP-56 structural model and compared the distances of functional groups in the vicinities of the hydroxyl groups of the threonines. For both predicted sites the structural arrangements were very similar to those in the X-ray structures used for comparison. Modeled sugar moieties into the three sites confirmed that they could accommodate glycans without steric conflicts (ESM Figure S1, part C). We feel that the consistency of all structural arrangements of glycosylation sites based on models and experimental data validates our analysis and suggest their probability of being glycosylated.

PA, the epitome of opportunistic bacteria is a major cause of co-infection in immune-compromised clinically confirmed VL patients [35, 36]. Hence, one of the intriguing questions remaining to be addressed is the presence of a protein of PA origin in the VL PBMC. Except for a few reports of host recognition via SA in *Campylobacter jejuni*, *Nesseria* sp., *Streptococcus* sp. [37–39], our study possibly opens a door for proposing a SA-mediated recognition of PA in the immunosuppressed host. It is conceivable that GP-56 or GP-56 like proteins are shed by PA and become associated with the VL-PBMC. It appears to have very distinct positively charged patches at its molecular surface which may mediate or assist in its binding to PBMC membranes. Alternatively, it could be envisaged that GP-56 may exhibit lectin-like activities [40]. Investigations in this direction are ongoing.

Acknowledgements The authors do not have a commercial or other association that might pose a conflict of interest. Ms. Angana Ghoshal and Dr. Sumi Mukhopadhyay are Senior Research Fellows of Council of Scientific and Industrial Research (CSIR), Govt. of India. This work received financial support from the CSIR, Indian Institute of Chemical Biology, Indian Council of Medical Research, New Delhi, Govt. of India, the Volkswagen Foundation, Germany, German Research Council (DFG) and the ProFit programme, EU/Berlin. We are grateful to Mr. Kalyan K. Sarkar for the ESI-MS study and Mr. Asish Mallick and Mr. Arthur O'Connor for their technical help.

References

1. Varki, A.: Multiple changes in sialic acid biology during human evolution. Glycoconj. J. (2008) (in press)
2. Schauer, R.: Achievements and challenges in sialic acid research. Glycoconj. J. **17**, 485–499 (2000). doi:10.1023/A:1011062223612
3. Ghosh, S., Bandyopadhyay, S., Mukherjee, K., Mallick, A., Pal, S., Mandal, C., Bhattacharya, D.K., Mandal, C.: *O*-acetylation of sialic acids is required for the survival of lymphoblasts in childhood acute lymphoblastic leukemia (ALL). Glycoconj. J. **24**, 17–24 (2007). doi:10.1007/s10719-006-9007-y
4. Mukhopadhyay, S., Mandal, C.: Targeting glycoproteins and glycolipids and their metabolic pathways for antiparasite therapy. In: Mazumdar, H. K. (ed.) Drug targets in kinetoplastid parasites, vol 625, pp. 87–99. Landes Biosciences, Austin (2008)
5. Sen, G., Mandal, C.: The specificity of the binding site of Achatinin-H, a sialic-acid binding lectin from *Achatina fulica*. Carbohydr. Res. **268**, 115–125 (1995). doi:10.1016/0008-6215(94)00311-3
6. Chava, A.K., Chatterjee, M., Sharma, V., Sundar, S., Mandal, C.: Variable degree of alternative complement pathway-mediated hemolysis in Indian visceral leishmaniasis induced by differential expression of 9-*O*-acetylated sialoglycans. J. Infect. Dis. **189**, 1257–1264 (2004). doi:10.1086/382752
7. Bandyopadhyay, S., Chatterjee, M., Sundar, S., Mandal, C.: Identification of 9-*O*-acetylated sialoglycans on peripheral blood mononuclear cells in Indian visceral leishmaniasis. Glycoconj. J. **20**, 531–536 (2004). doi:10.1023/B:GLYC.0000043289.86611.44
8. Mukhopadhyay, S., Mandal, C.: Glycobiology of *Leishmania donovani*. Indian J. Med. Res. **123**, 203–220 (2006)
9. Croft, S.L., Sundar, S., Fairlamb, A.H.: Drug resistance in leishmaniasis. Clin. Microbiol. Rev. **19**, 111–126 (2006). doi:10.1128/CMR.19.1.111-126.2006
10. Morales, R., Berna, A., Carpentier, P., Contreras-Martel, C., Renault, F., Nicodeme, M., Chesne-Seck, M.L., Bernier, F., Dupuy, J., Schaeffer, C., Diemer, H., Dorsselaer, A.V., Fontecilla-Camps, J.C., Masson, P., Rochu, D., Chabriere, E.: Serendipitous discovery and X-ray structure of a human phosphate binding apolipoprotein. Structure **14**, 601–609 (2006). doi:10.1016/j.str.2005.12.012
11. WHO: The leishmaniases. W. H. O. Tech. Rep. Ser. **793**, 154 (1990), (Geneva)
12. Pal, S., Ghosh, S., Mandal, C., Kohla, G., Brossmer, R., Isecke, R., Merling, A., Schauer, R., Schwartz-Albiez, R., Bhattacharya, D.K., Mandal, C.: Purification and characterization of 9-*O*-acetylated sialoglycoproteins from leukemic cells and their potential as immunological tool for monitoring childhood acute lymphoblastic leukemia. Glycobiology **14**, 859–870 (2004). doi:10.1093/glycob/cwh111
13. Chatterjee, M., Chava, A.K., Kohla, G., Pal, S., Merling, A., Hinderlich, S., Unger, U., Strasser, P., Gerwig, G.J., Kamerling, J.P., Vlasak, R., Crocker, P.R., Schauer, R., Schwartz-Albiez, R., Mandal, C.: Identification and characterization of adsorbed serum sialoglycans on *Leishmania donovani* promastigotes. Glycobiology **13**, 351–361 (2003). doi:10.1093/glycob/cwg027
14. Desai, R., Peretz, A., Idelson, H., Lazarovici, P., Attali, B.: Molecular cloning, biochemical and functional characterization. J. Biol. Chem. **275**, 39954–39963 (2000). doi:10.1074/jbc.M001562200

15. Sharma, V., Chatterjee, M., Mandal, C., Sen, S., Basu, D.: Rapid diagnosis of Indian visceral leishmaniasis using AchatininH, a 9-*O*-acetylated sialic acid binding lectin. Am. J. Trop. Med. Hyg. **58**, 115–125 (1998)
16. Reuter, G., Schauer, R.: Determination of sialic acids. Methods Enzymol. **230**, 168–199 (1994). doi:10.1016/0076-6879(94)30012-7
17. Klein, A., Diaz, S., Ferreira, I., Lamblin, G., Roussel, P., Manzi, A.E.: New sialic acids from biological sources identified by a comprehensive and sensitive approach: liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) of SIA quinoxalinones. Glycobiology **7**, 421–432 (1997). doi:10.1093/glycob/7.3.421
18. Bandyopadhyay, S., Chatterjee, M., Das, T., Bandyopadhyay, S., Sundar, S., Mandal, C.: Antibodies directed against *O*-acetylated sialoglycoconjugates accelerate complement activation in *Leishmania donovani* promastigotes. J. Infect. Dis. **190**, 2010–2019 (2004). doi:10.1086/425519
19. Demine, R., Walden, P.: Sequit: software for *de novo* peptide sequencing by matrix-assisted laser desorption/ionization post-source decay mass spectrometry. Rapid Commun. Mass Spectrom. **18**, 907–913 (2004). doi:10.1002/rcm.1420
20. Higgins, D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G., Gibson, T.J.: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680 (1994). doi:10.1093/nar/22.22.4673
21. de Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., Hulo, N.: ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. **34**, W362–W365 (2006). doi:10.1093/nar/gkl124
22. Schwede, T., Kopp, J., Guex, N., Peitsch, M.C.: SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. **31**, 3381–3385 (2003). doi:10.1093/nar/gkg520
23. Levefelt, C., Lundh, D.: A fold-recognition approach to loop modeling. J. Mol. Model. **12**, 125–139 (2006). doi:10.1007/s00894-005-0003-0
24. Mandal, C., Kingery, B.D., Anchin, J.M., Subramaniam, S., Linthicum, D.S.: ABGEN: a knowledge-based automated approach for antibody structure modeling. Nat. Biotechnol. **14**, 323–328 (1996). doi:10.1038/nbt0396-323
25. Canutescu, A.A., Shelenkov, A.A., Dunbrack Jr., R.L.: A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci. **12**, 2001–2014 (2003). doi:10.1110/ps.03154503
26. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., Thornton, J.M.: AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR **8**, 477–486 (1996). doi:10.1007/BF00228148
27. Arendall 3rd., W.B., Tempel, W., Richardson, J.S., Zhou, W., Wang, S., Davis, I.W., Liu, Z.J., Rose, J.P., Carson, W.M., Luo, M., Richardson, D.C., Wang, B.C.: A test of enhancing model accuracy in high-throughput crystallography. J. Struct. Funct. Genomics **6**, 1–11 (2005). doi:10.1007/s10969-005-3138-4
28. Mandal, C.: MODELYN—a molecular modelling program version PC-1.0, Indian Copyright No. 9/98 (1998)
29. Ahmad, S., Gromiha, M., Fawareh, H., Sarai, A.: ASAView: database and tool for solvent accessibility representation in proteins. BMC Bioinformatics **5**, 1471–2105 (2004). doi:10.1186/1471-2105-5-51
30. Chatterjee, M., Sharma, V., Mandal, C., Sundar, S., Sen, S.: Identification of antibodies directed against *O*-acetylated sialic acids in visceral leishmaniasis: its diagnostic and prognostic role. Glycoconj. J. **15**, 1141–1147 (1998). doi:10.1023/A:1006 963806318
31. DiGiandomenico, A., Matewish, M.J., Bisaillon, A., Stehle, J.R., Lam, J.S., Castric, P.: Glycosylation of *Pseudomonas aeruginosa* 1244 pilin:glycan substrate specificity. Mol. Microbiol. **2**, 519–530 (2002). doi:10.1046/j.1365-2958.2002.03171.x
32. Knirel, Y.A., Kocharova, N.A., Shashkov, A.S., Dmitriev, B.A., Kochetkov, N.K., Stanislavsky, E.S., Mashilova, G.M.: Somatic antigens of *Pseudomonas aeruginosa*. The structure of O-specific polysaccharide chains of the lipopolysaccharides from *P. aeruginosa* O5 (Lányi) and immunotype 6 (Fisher). Eur. J. Biochem. **163**, 639–652 (1987). doi:10.1111/j.1432-1033.1987.tb10913.x
33. Lee, D.G., Urbach, J.M., Wu, G., Liberati, N.T., Feinbaum, R.L., Miyata, S., Diggins, L.T., He, J., Saucier, M., Deziel, E., Friedman, L., Li, L., Grills, G., Montgomery, K., Kucherlapati, R., Rahme, L.G., Ausubel, F.M.: Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. Genome Biol. **7**, R90 (2006). doi:10.1186/gb-2006-7-10-r90
34. Cacalano, G., Kays, M., Saiman, L., Prince, A.: Production of the *Pseudomonas aeruginosa* neuraminidase is increased under hyperosmolar conditions and is regulated by genes involved in alginate expression. J. Clin. Invest. **89**, 1866–1874 (1992). doi:10.1172/JCI115791
35. Xia, B., Sachdev, G.P., Cummings, R.D.: *Pseudomonas aeruginosa* mucoid strain 8830 binds glycans containing the sialyl-Lewis x epitope. Glycoconj. J. **24**, 87–95 (2007). doi:10.1007/s10719-006-9015-y
36. Sipahi, T., Tavil, B., Oksal, A.: Visceral leishmaniasis and *Pseudomonas* septicemia associated with hemophagocytic syndrome and myelodysplasia in a Turkish child. Turk. J. Pediatr. **47**, 191–194 (2005)
37. Avril, T., Wagner, E.R., Willison, H.J., Crocker, P.R.: Sialic acid-binding immunoglobulin-like lectin 7 mediates selective recognition of sialylated glycans expressed on *Campylobacter jejuni* lipooligosaccharides. Infect. Immun. **74**, 4133–4141 (2006). doi:10.1128/IAI.02094-05
38. Jones, C., Virji, M., Crocker, P.R.: Recognition of sialylated meningococcal lipopolysaccharide by Siglecs expressed on myeloid cells leads to enhanced bacterial uptake. Mol. Microbiol. **49**, 1213–1225 (2003). doi:10.1046/j.1365-2958.2003.03634.x
39. Lewis, A.L., Hensler, M.E., Varki, A., Nizet, V.: The group B streptococcal sialic acid *O*-acetyl transferase is encoded by neuD, a conserved component of bacterial sialic acid biosynthetic gene clusters. J. Biol. Chem. **281**, 11186–11192 (2006). doi:10.1074/jbc.M513772200
40. Sonawane, A., Jyot, J., Ramphal, R.: *Pseudomonas aeruginosa* LecB is involved in pilus biogenesis and protease IV activity but not in adhesion to respiratory mucins. Infect. Immun. **74**, 7035–7039 (2006). doi:10.1128/IAI.00551-06