Combining LC-MS/MS, PMF and N-terminal amino acid sequencing for multiplexed characterization of a bacterial surfactant glycoprotein biosynthesized by Acinetobacter radioresistens S13

Marta Riva Violetta;Roberto Mazzoli;Cristina Barello;Paolo Fattori;Maria G. Giuffrida;Enrica Pessione. Combining LC-MS/MS, PMF and N-terminal amino acid sequencing for multiplexed characterization of a bacterial surfactant glycoprotein biosynthesized by Acinetobacter radioresistens S13. RSC ADVANCES. 4 pp: 10918-10927.
DOI: 10.1039/c4ra00692e

Combining LC-MS/MS, PMF and N-terminal amino acid sequencing for multiplexed characterization of a bacterial surfactant glycoprotein biosynthesized by *Acinetobacter radioresistens* S13

Marta Riva Violetta[a,b], Roberto Mazzoli[a], Cristina Barello[b], Paolo Fattori[a], Maria G. Giuffrida[b] and Enrica Pessione[a,*]

[a]Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123, Turin, Italy

[b]CNR-ISPA, c/o Bioindustry Park S. Fumero, Via Ribes 5, 10010 Colleretto Giacosa, Turin, Italy

***Corresponding Author:** Enrica Pessione

Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123, Turin, Italy

e-mail: enrica.pessione@unito.it

Phone: +39 011 670 4644

Fax: +39 011 670 4508

Table of contents

The present study has highlighted the mechanisms of bacterial protein glycosylation. Experimental results underline that the consensus sequon can be different from the one found in Eukaria.

Abstract

Surfactants of biological origin are highly requested both in the food industry and for cosmetic applications. In the present paper the glycosylated nature of a surfactant protein, previously identified in the *Acinetobacter radioresistens* S13 proteome, was demonstrated. A multiplexed analysis was performed to establish its amino acid sequence homology degree with AlnA, a previously characterized surfactant from *A. radioresistens* KA53. The amino acid sequence coverage study (N-terminal amino acid sequencing, peptide mass fingerprinting and LC-MS/MS experiments) revealed a 99.97% identity, that is the substitution of one amino acid only. Gly_{25} of *A. radioresistens* KA53 AlnA is replaced by Thr in the protein identified in *A. radioresistens* S13. Such change gives rise to Asn-Asp-Thr N-glycosylation consensus sequon, which is absent in *A. radioresistens* KA53 AlnA. Actually, Asn_{23} of the *A. radioresistens* S13 protein could be identified after N-glycosyl hydrolase treatment only. Asn-Asp-Thr is a peculiar glycosylation consensus sequence since Asp in the central position was shown to decrease the protein glycosylation efficiency in eukaryotes. Our findings provide additional support that glycosylation mechanisms in bacteria differ from those observed in eukaryotic cells. The emulsifying activity of the extracellular protein extracts of *A. radioresistens* S13 (containing the AlnA-like protein) and a commercial solubilizer widely used in cosmetic preparations, *i.e.* ethoxylated hydrogenated castor oil (EHCO), were compared. *A. radioresistens* S13 extracellular protein extracts showed greater emulsifying activity on every tested molecules, especially on peppermint essential oil, with respect to EHCO.

Key words: 2DE, Pro-Q Emerald, LC-MS/MS, MALDI-TOF/TOF, Glycosylation site, N-linked sugars, Emulsifying activity.

Introduction

Acinetobacter spp. are gram-negative bacteria that are often employed in environmental applications such as phosphate removal from waste waters,[1] hydrocarbon degradation,[2,3] lipase[4] and bio-emulsifier production[5,6]. Surfactants of biological origin are very requested[7] for environmental applications,[8] as immuno-modulators and signalling molecules,[9] as anti-adhesive and antiviral compounds[10] and as antimicrobial agents both for therapeutic use[11] and in the cosmetic industry. As far as cosmetic preparations are concerned, bio-surfactants can also be used as gentle emulsifiers.[12] Their mechanism of surface and interfacial tension reduction is exactly the same as the chemical surfactants, but biosurfactans can be produced by waste biomass fermentation processes, with significant economic advantages.[13]

Acinetobacter spp. can produce several structurally similar emulsifiers such as *A. calcoaceticus* A2 biodispersan (an anionic heteropolysaccharide)[14] and *A. calcoaceticus* BD4 and *A. calcoaceticus* RAG1 emulsans (polyanionic lipoheteropolysaccharides)[15]. Sometimes these surfactants are bound to the bacterial surface as either pseudo minicapsules,[16] or true capsules, as described in *A. venetianus*.[17]

A. radioresistens, in particular, has been extensively employed for its surfactant properties so as to enhance substrate bioavailability for efficient oil spill bioremediation.[18]

Several studies demonstrated that *A. radioresistens* KA53 can increase the solubility and stabilize emulsions of hydrophobic molecules, chiefly by an extracellular protein complex induced by ethanol.[19] This multimeric protein, named Alasan, consists of three subunits of 45, 31 and 16 kDa, respectively.[20] One of them, *i.e.* AlnA (apparent molecular weight of 45 kDa), is glycosylated and displays the highest emulsifying activity. The sugar component seems to improve both secretion and protection from proteolysis[19] and also contributes to emulsion stability, although it does not affect the emulsifying activity.[20]

In our previous investigations on *A. radioresistens* S13 proteome, which aimed to study aromatic compound catabolic pathways, a protein ascribable to *A. radioresistens* KA53 AlnA,

was detected in both the soluble (late stationary),[21] and the membrane-enriched (late exponential) extracts.[22] AlnA-like protein was overproduced by phenol- and benzoate-grown bacteria, while it was present in lower amount in acetate-grown cultures.

Toren and co-workers[23] reported that *A. radioresistens* KA53 AlnA shares several protein domains (85% N-terminal amino acid identity and 27% DNA deduced overall amino acid sequence identity) with *E. coli* OmpA, that is an outer membrane protein commonly found in gram-negative bacteria.[24,25] Actually, a later study demonstrated that AlnA and OmpA were the products of a single gene.[26] The Authors hypothesized that neo-synthesized Alasan molecules are stored in the outer membrane and secreted into the extracellular environment in the stationary growth phase only. The recombinant *A. radioresistens* KA53 OmpA/AlnA, recovered from *E. coli* cellular extracts, showed an apparent molecular mass of 36 kDa,[27,28] that is lower than the one (45 kDa) of the native protein.[20]

As far as we know, some information about the sugar composition of AlnA has been reported,[19] but no data are available on its glycosylation site(s). On the other hand, glycosylation mechanisms in bacteria are far to be fully elucidated.

In the present paper we demonstrated that the AlnA-like protein produced by benzoate-grown *A. radioresistens* S13 can be both membrane-bound and secreted in the extracellular environment and that it is glycosylated, as it was previously observed for ethanol-induced AlnA from KA53. Furthermore, *A. radioresistens* S13 AlnA-like glycosylation site was established, by combining MS techniques and N-terminal amino acid sequencing. The emulsifying activity of AlnA-like containing extracellular extracts of *A. radioresistens* S13 and ethoxylated hydrogenated castor oil (that is a widely used surfactant for cosmetic applications) were compared. The obtained results open new perspectives for biotechnological applications of AlnA-like protein from *A. radioresistens* S13.

Results

Structural studies

Glycoprotein detection in 2DE gels.

2DE analyses were performed on the membrane protein extracts of benzoate-grown *A. radioresistens* S13 harvested during the exponential growth phase, as shown in Fig. 1. Gels were stained with either Sypro Ruby (Fig. 1A), i.e. for total protein detection, or ProQ Emerald (Fig. 1B), for glycoprotein detection. A spot train of about 40 kDa retained both dyes (boxed in Fig. 1A and 1B). PMF analysis (data not shown) confirmed that ProQ Emerald stained spots were different isoforms of a AlnA-like protein found in previously reported studies (spot 19,[21] ; spots 1, 21,[22]).

Sequence coverage of *A. radioresistens* S13 AlnA-like protein and glycosylation site localization.

AlnA from *Acinetobacter radioresistens* KA53 (UniProt ID: Q8VPR9) was used as the reference sequence for amino acid sequence coverage studies. *In-gel* multiple enzyme digestions followed by MALDI-TOF/TOF and nano LC−ESI MS/MS analysis was performed (supplementary data, Table 1S). MS techniques allowed to achieve 95% of sequence coverage (Fig. 2). The Edman degradation technique was used to investigate the N-terminal residues from 1st to 28th, and improve protein coverage to 99.7 %. The covered sequence of the *A. radioresistens* S13 protein was perfectly identical to *A. radioresistens* KA53 AlnA, except for the 25th residue, where the Gly present in the KA53 AlnA was replaced by a Thr in the S13 derived protein (Fig. 2). Furthermore, amino acid residue 23rd gave an empty cycle during N-terminal sequence analysis, suggesting it was the glycosylation-bearing amino acid.

Sugar removal

Since a Thr was present in position 25th of *A. radioresistens* S13 AlnA-like protein, we hypothesized that it could be part of a Asn23-Asp24-Thr25 N-glycosylation sequon. To verify this hypothesis, the *A. radioresistens* S13 AlnA-like protein was incubated with the endoglycosidase PNGase F, which can hydrolyze N-linked sugars. MALDI-TOF zoomed spectra of Asp-N generated peptides prior and after enzymatic treatment are shown in Fig. 3A and 3B, respectively. In the PNGase F treated sample, a signal at 957.457 m/z, which is consistent with the peptide ^{15}DTQHNNGGD^{23}, was detected while this signal was absent in the native *A. radioresistens* S13 AlnA-like protein. Endoglycosidase digestion, therefore, allowed amino acid at position 23rd to be identified, proving that it was hidden by a saccharidic moiety in the native protein. It also proved that the sequence from 15th to 22nd is consistent with the one obtained by Edman degradation. The identified Asp residue in position 23rd actually is the result of the well known catalytic mechanism of PNGase F hydrolysis (the N-group remaining on the sugar side) that allowed us to deduce that position 23rd was occupied by an Asn.

Extracellular protein isolation and identification

SDS-PAGE of the extracellular raw extract of *A. radioresistens* S13 cells, collected during the exponential growth phase, is shown in Fig. 4. PMF analysis was performed on 6 protein bands, which are indicated in Fig. 4. Identified proteins are listed in Table 1.

Functional studies

Emulsifying activity

A preliminary study determined the minimal concentration of the raw extracellular protein extract (REPE) able to solubilize a fixed amount of substrate I (1:1 v/v hexadecane and 2-methylnaftalene). Maximum specific emulsifying activity (S.E.A. = 23.90 ± 0.11 U/mg) was obtained with 100 μg of REPE, while for higher amounts of REPE S.E.A. progressively

decreased (data not shown). This amount of REPE was therefore used for the following emulsifying tests. S.E.A. of REPE on the four substrates of interest (hexadecane and 2-methylnaftalene, peppermint essential oil, Vitamine E, methyl dihydrojasmonate) was compared to values obtained by using EHCO, a surfactant very common in the cosmetic industry (Table 2). REPE showed specific emulsifying activity higher than EHCO by over two orders of magnitude on every tested substrate.

Discussion

AlnA is a glycoprotein from *A. radioresistens* KA53 and it is the key component of the bioemulsifier Alasan.[20,28] The present study aimed to characterize the AlnA-like protein from *A. radioresistens* S13, with respect to both its glycosylation and emulsifying features.

The use of glycan specific staining (*i.e.* Pro-Q Emerald) indicated that AlnA-like from the membrane extract of *A. radioresistens* S13 contained a saccharidic moiety. In order to determine the glycan component binding site, the full amino acid sequence determination of the AlnA-like protein was performed by combining LC-MS/MS, PMF and amino acid sequencing techniques. *A. radioresistens* S13 protein sequence was identical to the DNA-deduced primary sequence of *A. radioresistens* KA53 AlnA, with the exception of position 25th, where a Thr replaced a Gly. Residue 23rd of *A. radioresistens* KA53 AlnA is Asn, while this position could not be identified in *A. radioresistens* S13 protein by either MS or amino acid sequencing. This finding is consistent with a covalent modification of this amino acid. We hypothesized that Thr25 of *A. radioresistens* S13 AlnA-like protein, could be part of an Asn23-Asp24-Thr25 N-glycosylation sequon. Indeed, N-glycosidase treatment confirmed that AlnA-like amino acid 23rd is an Asn which bears a glycan in the native protein.

These experimental results demonstrated that: (i) AlnA-like has only one site of glycosylation, (ii) AlnA-like is N-glycosylated and (iii) this N-glycosylation sequon is not present in AlnA from *A. radioresistens* KA53.

The ability of bacteria to glycosylate proteins was relatively recently discovered.[29,31] Our results are consistent with previous observations which indicated that proteins are preferentially glycosylated near their N-terminus.[32] In *A. radioresistens* KA53 AlnA, amino acid 23rd is located just upstream a beta-barrel domain: while the latter is embedded in the lipid bilayer, this amino acid is likely available for sugar binding.[23,28]

The present investigation proved that PNGase F is also able to remove the glycan component of bacterial glycoproteins, differently from previous observations by Balonova and co-workers.[33] With respect to eukaryotic consensus motifs for protein N-glycosylation (*i.e.* Asn-X-Ser/Thr), the consensus sequence that was identified in the present study (*i.e.* Asn-Asp-Thr) is quite unusual. Studies performed on N-glycosyltransferases from *Eukarya*[34] indicated that the glycosylation efficiency is decreased when an aspartate or a glutamate is present in the 2nd position of the sequon. It was hypothesized that negatively charged amino acids in this site hamper the oligosaccharyltransferase catalytic mechanism. Actually, these studies mainly concerned the Asn-X-Ser sequons, that, in general, are less efficiently glycosylated in respect to Asn-X-Thr. Furthermore, bacteria significantly differ from eukaryotes with respect to both the glycosyltransferase catalytic mechanism and the cell district where glycosylation occurs.[35] As far as glycosyltransferase catalytic mechanism is concerned, the regions surrounding the sequon triplet in target proteins contribute to catalytic activity modulation in both prokaryotic and eukaryotic cells. However, the presence of acidic amino acids (Asp and Glu) two positions left to Asn while it inhibits glycosylation in eukaryotes,[36] it is essential to promote the activity of bacterial glycosyltransferase.[37] Furthermore, protein glycosylation in bacteria generally occurs at the membrane level. In the present study, a glycosyltransferase, probably belonging to the membrane, was isolated from *A. radioresistens* S13 raw extracellular extracts (Table 1, band 3). This finding can be partly explained considering that weakly-linked membrane proteins can be extracellularly released, either during cell division[38] or due to pH changes.[39]

All these considerations indicate that further studies are needed for in-depth understanding of bacterial glycosylation.

The presence of glycoproteins on the bacterial surface of strains belonging to the *Acinetobacter* genus is not new. Surface glycoproteins and polysaccharides were referred to be involved in the protection of *Acinetobacter* cell envelope phospholipids from stressful direct contact with alkanes.[47] Similarly, benzoate induces a strong envelope stress response,[40] including the synthesis of peculiar LPS, in *A. radioresistens* S13.[41-42]

The AlnA-like surface glycoprotein that was characterized in the present investigation could therefore function as both an additional defense against hydrocarbon stress and a hydrocarbon emulsifying protein. Indeed, some studies indicate that in *Acinetobacter* strains surface glycoproteins act as bioemulsifying agents enhancing bioavailability of hydrophobic substrates. Apart from the extensively studied *A. radioresistens* KA53 AlnA,[19,20,23,26-28] a Concanavalin A-positive glycoprotein was detected in the outer membrane of a hydrocarbon-tolerant *A. venetianus*.[17] The Authors hypothesized that this glycoprotein was responsible for bioemulsifying activity on diesel nanodroplets at the cell envelope level. The glycoprotein characterized in the present study, although present (in low amounts) in acetate cultures, is over-produced in benzoate-grown cells. Since benzoate is quite hydrophobic because of its aromatic ring structure, AlnA-like protein from *A. radioresistens* S12 could play a similar role as the *A. venetianus* glycoprotein. On the other hand, aromatic compounds are not the only hydrophobic substrates that can be degraded by the strain in study. Preliminary unpublished data obtained by our group indicate that *A. radioresistens* S13 also harbors a gene (*estA*) encoding an esterase (as detected by PCR). This strain exhibits esterase C4 and esterase-lipase C8 activity as well (data not shown) and proved to be able to hydrolyze both tributyrine (tributylglycerol) and MU-butyrate (4-methylumbelliferyl-butyrate), in agar plate tests. Tributyrine esterase activity was also clearly detected in late exponential phase extracellular extracts. All these results underline the need for surfactant production to allow lipid

emulsification in this strain. Actually, *A. radioresistens* S13 cannot metabolize sugars, while several hydrophobic or amphipathic molecules such as aromatic compounds, caprate, adipate and short chain esters, can be used as growth substrates. The evolution of efficient emulsifying systems for improved substrate availability probably contributed to this metabolic phenotype.

For optimal emulsifying activity, bio-surfactants should be released extracellularly, although they can exert part of their action also when they are surface-bound. In the present investigation, the two most abundant extracellular proteins (Fig. 4, bands 4 and 5,) were identified as outer membrane proteins (Table 1). The PMF analysis confirmed that band 4 is consistent with the AlnA-like protein that was previously characterized in the membrane fraction. This protein therefore displays a similar behaviour to *A. radioresistens* KA53 AlnA with respect to its cellular location.[26] Band 5 was identified as a member of outer membrane proteins. Further minor proteins were also identified in the *A. radioresistens* S13 extracellular extracts. Band 1 corresponds to a 71 kDa protein conferring resistance to copper. Resistance to copper and other metals is a common feature in both clinical[43] and environmental *Acinetobacter* spp.[44] On the other hand, hydrocarbon contaminated environments often display high heavy metal content. Band 2 contains an isocitrate lyase, that is an enzyme of the glyoxylate pathway. Its extracellular location was previously demonstrated in *Acinetobacter* spp. by Barbaro and co-workers[45] although a specific function was for such moonlighting behaviour has not been elucidated yet. Band 3 is consistent with a glycosyl transferase, probably involved in the glycosylation process of AlnA like protein and other cell wall or outer membrane proteins. Band 6 contains a protein that is consistent with the *A. radioresistens* SK82 YiaD. YiaD was firstly annotated as an inner membrane protein. Nonetheless, more in depth investigations on its amino acid sequence rather suggested that it is an outer membrane lipoprotein possessing an OmpA domain.[46] Inside this domain, Asp153 seems to play a crucial role in the formation of hydrogen bond with the mesodiaminopimelate residues of peptidoglycan. According to more recent data, YiaD is a cell wall associated

protein anchored to the outer membrane through an N-terminal lipid-attachment and also interacting with peptidoglycan through non-covalent linkages.[47] In *E. coli*, its function is the transport and assembly of outer membrane beta-barrel proteins.

The present study finding of a significant number of outer membrane proteins in the extracellular compartment suggests that these proteins are weakly linked to the bacterial surface: apart from Alasan like proteins, the biological functions of both copper-resistance protein and YiaD fully justify their double extracellular location (surface-attached or released).

AlnA-like protein is among the most abundant proteins of the raw extracellular protein extract (REPE) from benzoate-grown *A. radioresistens* S13. For this reason whole REPE were tested for their emulsifying properties. Comparative experiments demonstrated that REPE was a much more effective surfactant than ethoxylated hydrogenated castor oil (EHCO), *i.e.* a commercial emulsifier, active on a number of hydrophobic substrates used in cosmetic preparations. Purification of AlnA-like protein from *A. radioresistens* S13 REPE will be necessary to confirm its emulsifying properties and clearly establish its biotechnological potential. However, the direct use of REPE, by-passing complex and expensive purification processes, would be of great interest from an economical industrial point of view.

Experimental

Bacterial Strain.

Acinetobacter radioresistens S13 was isolated as previously described[2,48] and was selected because of its ability to grow on aromatic compounds as the sole carbon and energy source.

The strain was maintained in Luria-Bertani (LB) broth supplemented with benzoate (400 mg/L) and stored at –20°C after 20% v/v glycerol supplementation.

Culture conditions

Bacteria were grown in the Sokol and Howell minimal medium[48] (0.5 L in 2 L baffled Erlenmeyer flasks) supplemented with benzoate (400 mg/L). Flasks were incubated at 30°C in continuous agitation (210 rpm) using a Gallenkamp oscillating shaker. Culture optical density at 600 nm was monitored by using a Ultrospec 2000 (Pharmacia Biotech) spectrophotometer. Biomass was harvested by centrifugation (10,000 x *g*, 15 min at 4°C) in late exponential phase.

Protein separation

The biomass (25 mg dry weight) was washed twice with 50 mM Tris-HCl pH 7.3 (3000 x g, 15', 4°C) and re-suspended in 3 mL of the same buffer containing a nuclease mix (GE Healthcare). The re-suspended biomass was disrupted via ultrasonic treatment in a Microsonix sonicator ultrasonic liquid processor XL2020 by 10'' pulses at 20 KHz separated by 20'' pauses, for a total of 30 min pulses, keeping the cells on ice. The sonicated solution was centrifuged at 3000 x *g* for 20 min at 4°C (ALC Multispeed Refrigerated Centrifuge PK 131 R) to separate the unbroken cells, and the supernatant was collected. The pellet was re-suspended and sonicated again as described above. Both supernatants were combined, diluted with ice-cold 0.1 M sodium carbonate (pH 11) and slowly stirred on ice overnight.[22] Carbonate treated samples were centrifuged at 10,0000 x *g* for 1 h at 4°C using a Beckman L8-60M Ultracentrifuge. The pellet (containing the membrane fraction) was washed twice with 50 mM Tris-HCl pH 8.6, ultracentrifuged (100000 x *g*, 20', 4°C) and re-suspended in the "Molloy" Buffer.[49,50]

For the determination of the emulsifying activity, cell-free broths were obtained by centrifugation and were mixed with the Complete protease inhibitors (Roche), filtered using 0.45 µm pore-size membrane filters, mixed with ammonium sulphate to 80% saturation and then incubated overnight at 4°C with continuous agitation. The suspension was

ultracentrifuged at 100,000 x *g* for 90 min at 4°C (Beckman L8-60M). Pellets were re-suspended in MilliQ water and dialyzed in three volumes of water.

Protein quantification was performed using the 2-DE quant kit (Amersham Biosciences).

Electrophoresis

One-dimension electrophoresis (1-DE) was performed under reducing conditions on sodium dodecyl sulphate polyacrylamide gels (SDS-PAGE) following the Laemmli's method.[51] Five µg of total proteins were loaded in each lane. Molecular mass markers were the LMW from GE Healthcare.

For two-dimensional electrophoresis (2-DE), membrane proteins (175 µg) were loaded on 7 cm pI 4-7 immobilized pH Gradient (IPG) strips, (GE Healthcare) by the in-gel rehydration method[52] and were separated by using an IPGphor (GE Healthcare) up to 20,000 Vhr at 20°C. After IEF, proteins were reduced and alkylated by incubation with the "Sample Reducing Agent" (Invitrogen) and 125 mM of "Alkylating Solution" (Invitrogen), respectively. The equilibrated strips were embedded in 0.5 % (w/v) agarose prior to SDS–PAGE. The second dimension was performed by using 12% NuPage Novex Bis-Tris [Bis(2-hydroxyethyl)imino-tris(hydroxymethyl) methane-HCl] Pre-Cast Gels (Invitrogen). The separation was performed in 600 ml of MOPS [3-(N-morpholino) propane sulfonic acid] buffer by using a Xcell SureLock ™ Novex Mini-cell (Invitrogen). Separation was perfomed by applying 200V for 60 min. Molecular mass markers were from Invitrogen (Mark12 Unstained Standards).

Gels were automatically stained using the Processor Plus (GE Healthcare) with freshly prepared Colloidal Coomassie Blue stain.[53] Gel images were acquired with a GS-800 Densitometer (Bio-Rad).

Glycoprotein staining

Detection of glycoproteins in 2DE gels was performed by the Pro-Q Emerald 488 glycoprotein stain kit (Invitrogen). The 2DE gels were soaked in the fixing solution (50% methanol and 5%

acetic acid v/v in MilliQ water) for 1 h, then incubated overnight in fresh fixing solution to completely remove residual SDS. Gels were washed twice in 3% acetic acid (w/v in MilliQ water) for 10 min, and glycans were oxidized by incubation in 1% periodic acid, 3% acetic acid (w/v/v in MilliQ water) for 20 min. In order to remove the residual periodate, gels were washed three times for 10 min in 3% acetic acid and then incubated in Pro-Q Emerald 488 dye solution for two hours, following the producer's instruction. Afterwards, gels were washed again in 3% acetic acid, three times for 10 min and then briefly rinsed in water. In the negative control gel, the periodate oxidation step was omitted.

After glycoprotein detection by Pro-Q Emerald and image acquisition, gels were post-stained with SYPRO Ruby (Invitrogen) by overnight incubation without fixing step. Before the SYPRO Ruby image acquisition, gels were washed twice in 10 % methanol and 7 % acetic acid (v/v) and rinsed three times with ultrapure water. Gel images were acquired using a ProXPRESS 2D Proteomic Imaging System (PerkinElmer Life and Analytical Sciences) equipped with suitable excitation and emission filters (480/30 nm- 530/30 nm for Pro-Q Emerald 488 and 460/80 nm-650/150 nm for SYPRO Ruby).

Enzymatic digestion

Multiple enzymatic digestion on selected 2DE spots by using Trypsin, Endoproteinase Asp-N, Elastase and Proteinase K was performed according to Speicher et al.[54] method, with slight modifications. Briefly, protein spots were excised from the gels and destained for two hours in 40% ethanol containing 50 mM NH_4HCO_3. Spots were dehydrated three times for 10 min in acetonitrile and dried in a vacuum concentrator for 30 min. Rehydration was performed for 5 min at room temperature with 5 µL of trypsin solution (75 ng/µl trypsin in 25 mM NH_4HCO_3) (Promega Corp.), or 5 µl of 75 ng/µl Asp-N solution (Sigma), or 3 µl of 200 ng/µl elastase solution (Sigma) or 5 µl of 75 ng/µl proteinase K solution (Sigma). For trypsin digestions only, an aliquot of the peptide solution was collected after 5 min and analysed by MS. Then,

each enzyme-treated sample (including trypsin) was supplemented with 10 µL of 25mM NH_4HCO_3 and were kept on stirring for one hour at 37°C. An aliquot of the supernatant, containing proteinase K generated peptides, was collected and directly submitted to LC-MS analysis. Trypsin, Asp-N and elastase digestions were performed overnight at 37°C after addition of 15 µl of 10% ACN in 25mM NH_4HCO_3. The day after, each supernatant was harvested and submitted to MS analysis. To obtain more amino acid sequence information, tryptic peptide solution was digested again with AspN endoproteinase in the same conditions described above.

The SDS-PAGE bands of interest were "in gel" digested with trypsin as described above.

Sugar removal from glycoproteins

The endoglycosidase PNGaseF from *Elizabethkingia meningoseptica* (Sigma – Aldrich) was used for hydrolytic removal of saccharides from proteins. 2DE spots were excised from gels and destained in 40% ethanol containing 50 mM NH_4HCO_3 for two hours. Spots were dehydrated in acetonitrile three times for 10 min and dried in a vacuum concentrator for 30 min. Rehydration was performed in 10 µL of 500 U.I./ml PNGase F (Sigma-Aldrich), at 37°C for 30 min with constant stirring. Then 20 µl of water were added and samples were kept on stirring at 37°C for additional 2 hours. After centrifugation, supernatant was collected and concentrated to 3 µl in a vacuum concentrator, while gels were rehydrated in 200 µl of water and sonicated at room temperature for 30 min for 3 times. After centrifugation, the supernatant was collected, concentrated and gathered with the other supernatants. Both deglycosylated and non-deglycosylated spots were treated by multiple enzymatic digestion as described above.

Mass Spectrometry Analysis

For MALDI-TOF mass spectrometry analysis, 0.5 µl of each peptide mixture were applied to a target disk and allowed to air-dry. Subsequently, 0.5 µl of matrix solution (1%w/v α-cyano-

4hydroxycinnamic acid in 30% acetonitrile, 0.1% TFA) were applied to the dried sample and allowed to dry under vacuum. Spectra were collected on a Ultraflex II MALDI-TOF/TOF (Bruker Daltonik GmbH, Germany). The reported mass spectra are the result of signals of approximately 800 laser shots. The spectra were externally calibrated using the standard calibration mixture from Bruker Daltonik (GmbH, Germany). Manual/visual evaluation of the mass spectra was performed by using Flex Analysis software (Bruker Daltonik GmbH, Germany). The tandem mass spectra were acquired on the same mass spectrometer, run LID experiments using LIFT TOF/TOF acquisition. The MS/MS spectra were automatically analysed by Flex Analysis software (Bruker Daltonik GmbH, Germany).

For LC-MS/MS experiments an Agilent's 1100 Series liquid chromatograph and LC/MSD XCT series ion trap mass spectrometer equipped with nano-ESI source were used. The peptide mixtures were separated by RP C18 column (Zorbax 300SB-C18 3,5 µm 150x0,075mm) (CPS), which was equilibrated with 0.1% formic acid in MilliQ water. The peptides were eluted over 55 min with a linear gradient of 5-70% (v/v) of 0.5% water/0.1% formic acid in acetonitrile. The flow rate was 300 nL/min. The mass spectrometer was operated in positive ion mode. Data were processed using DataAnalysis software (Agilent Technologies, Santa Clara, U.S.).

Spectra interpretation

For Peptide Mass Fingerprinting (PMF) identification method the MS-Fit software package in ProteinProspector v5.10.13 (http://prospector.ucsf.edu/prospector/) was used for spectra interpretation.[55] For the search the UniProtKB database (http://www.uniprot.org/) was used. Two missed cleavages were allowed for enzymatic digestions, cysteins were considered completely modified by iodoacetamide and methionine partially oxidized. Mass accuracy was 20 ppm.

To study the MS spectra, the amino acid sequence of the OmpA-like protein from *Acinetobacter radioresistens* KA53 (UniProt ID: Q8VPR9) was used as template. By using the MS-digest (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest) and PeptideMass (http://expasy.org/tools/peptide-mass.html) software, the theoretical trypsin and endoproteinase AspN peptides were listed and their values were used to identify the peaks in the MALDI-TOF spectra.

The LC-MS/MS and MALDI-TOF/TOF spectra were analysed using MASCOT MS/MS Ions Search (http://www.matrixscience.com/). Parameters used were peptide mass tolerance (0.8 Da), MS/MS ion mass tolerance (0.6 Da), allowance of two missed cleavage sites, differential modification (methionine as oxidation), and default selection of charge states ions (+2, +3). Peptides showing a score greater than 50 were considered significant ($p<0.05$).

N-terminal sequencing

Spots were passively eluted from gels as described elsewhere.[56] Membrane was analyzed by direct micro sequencing on a Procise 492 protein sequencer (Applied Biosystems, Foster City, CA, USA). All chemicals used were from Applied Biosystems.

Emulsifying activity determination

The emulsification activity was determined as described by Toren et al.,[20] with some modifications. The solubilization activity of the raw extracellular protein extract (REPE) from *A. radioresistens* S13 was compared to ethoxylated hydrogenated castor oil (EHCO) commercial solubilizer (Cremophor*RH40, BASF, Italy). Four substrates were used: (I) a solution containing 1:1 (v/v) hexadecane and 2-methylnaftalene (Sigma-Aldrich), (II) peppermint essential oil (Maraschi e Quirici S.p.a, Italy), (III) Vitamine E solution (Firmenich International S.A. Geneva, Switzerland), (IV) methyl dihydrojasmonate (Firmenich

International S.A. Geneva, Switzerland). Substrate I was the same as used by Toren et al.,[20] while the other substrates were chosen among the most used in cosmetic applications.

Fifty µl of either REPE or EHCO were mixed with 1.4 ml TM buffer (20 mM Tris-HCl buffer, pH 7.0, 10 mM Mg_2SO_4) and 20 µl of substrate in 2 ml test tubes. The tubes were vigorously mixed for 30 min at room temperature and the absorbance at 600 nm (A_{600}) was measured by using a spectrophotometer Ultrospec 2000 (Pharmacia Biotech). All tests were performed in triplicate and the average and the standard deviation (SD) were calculated. One unit of Generic Emulsifying Activity (G.E.A.) is defined as the amount of emulsifier that yields a 0.1 A_{600} increase. The Specific Emulsifying Activity (S.E.A.) is defined as the ratio between the G.E.A. and the amount in mg of surfactant that was used.

So as to determine the minimal amount of REPE able to solubilize a fixed amount of substrate, concentrations of REPE ranging from 1.0 to 10.0 mg/ml were mixed with 20 µl of substrate I.

The S.E.A. of REPE for the four substrates tested was then determined and compared to those of EHCO and Bovin Serum Albunin (BSA). The latter was used as the negative control.

Conclusions

Apart from roles in immunogenicity and adhesiveness of pathogenic strains, surface glycoproteins are important for bacteria to survive in harsh environments, such as hydrocarbon polluted soil and water, and to enhance hydrophobic substrate bioavailability.

The results obtained in the present study established that an *A. radioresistens* S13 protein, homologous to the AlnA component of the *A. radioresistens* KA53 Alasan, is a glycoprotein. Sequence coverage studies revealed that only one amino acid residue differs between the two proteins (Gly25 of *Acinetobacter radioresistens* KA53 AlnA is replaced by a threonine in *A. radioresistens* S13 protein). This difference results in a consensus sequon Asn-Asp-Thr for N-glycosylation of *Acinetobacter radioresistens* S13 AlnA-like protein. AlnA-like protein glycosylation was confirmed to occur on Asn23. The presence of Asp in the middle of an N-

glycosylation consensus sequence is peculiar since it is considered as unfavourable for glycosylation mechanisms, at least in eukaryotes. This finding therefore provides additional support that glycosylation mechanisms in bacteria significantly differ from eukaryotic cells. The puzzle of bacterial glycoproteins is still incomplete: in depth understanding of protein glycosylation in prokaryotes is far to be achieved. Research in this domain will likely bring useful information for biotechnological applications such as pollutant bioremediation, cellulosic biomass bioprocessings, food and cosmetic preparations and in the understanding of host-microbiota relationship at the gut level.[57-58]

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work has been supported by "Ricerca autofinanziata 2011" grant from the University of Torino. A particular thank to Microna s.r.l. for supplying substrates of cosmetical interest.

References

1 C. D. Boswell, R. E. Dick, H. Eccles and L. E. Macaskie, *J. Ind. Microbiol. Biot.*, 2005, **26**, 333-340.

2 E. Pessione and C. Giunta, *Microbios*, 1996, **88**, 213-221.

3 S. Y. Siehler, S. Dal, R. Fischer, P. Patz and U. Gerischer, *Appl. Environ. Microb.*, 2007, **73**, 232-242.

4 E. A. Snellman and R. R. Colwell, *J. Ind. Microbiol. Biot.*, 2004, **31**, 391-400.

5 E. Rosenberg and E. Z. Ron, *Curr. Opin. Biotechnol.*, 1997, **8** , 313-316.

6 E. Z. Ron and E. Rosenberg, *Curr. Opin. Biotechnol.*, 2002, **13**, 249-252.

7 J. D. Desai and I. M. Banat, *Microbiol. Mol. Biol. Rev*., 1997, **61**, 47-64.

8 C. N. Mulligan, M. Kamali and B. F. Gibbs, *J. Hazard. Mater.*, 2004, **110**, 77-84.

9 S. S. Cameotra and R. S. Makkar, *Curr. Opin. Microbiol.*, 2004, **7**, 262-266.

10 L. Rodrigues, I. M. Banat, J. Teixeira and R. Oliveira, *J. Antimicrob. Chemoth.*, 2006, **57**, 609-618.

11 P. Singh and S. S. Cameotra, *Trends Biotechnol*., 2004, **22**, 142-146.

12 A. Singh, J. D. Van Hamme and O. P. Ward, *Biotechnol. Adv.*, 2007, **25**, 99-121.

13 K. Muthusamy, S. Gopalakrishnan, T. K. Ravi and P. Sivachidambaram, *Curr. Sci.* 2008, **94**, 736-747.

14 E. Rosenberg, C. Rubinovitz, A. Gottlieb, S. Rosenhak and E. Z. Ron, *Appl. Environ. Microb.*, 1998, **54**, 317-322.

15 E. Rosenberg, A. Zuckerberg, A. Rubinovitz and D. L. Gutnick, *Appl. Environ. Microb.*, 1979, **37**, 402-408.

16 N. Kaplan and E. Rosenberg, *Appl. Environ. Microb.*, 1982, **44**, 1335-1341.

17 F. Baldi, M. Pepi, A. Capone, C. Della Giovampaola, C. Milanesi, R. Fani and R. Focarelli, *Res. Microbiol.*, 2003, **154**, 417-424.

18 T. Barkay, S. Navon-Venezia, E. Z. Ron and E. Rosenberg, *Appl. Environ. Microbiol.*, 1999, **65**, 2697-2702.

19 S. Navon-Venezia, Z. Zosim, A. Gottlieb, R. Legmann, S. Carmeli, E. Z. Ron and E. Rosenberg, *Appl. Environ. Microb.* 1995, **61**, 3240-3244.

20 A. Toren, S. Navon-Venezia, E. Z. Ron and E. Rosenberg, *Appl. Environ. Microbiol.* 2001, **67**, 1102-6.

21 M. G. Giuffrida, E. Pessione, R. Mazzoli, G. Dellavalle, C. Barello, A. Conti and C. Giunta, *Electrophoresis*, 2001, **22**, 1705-1711.

22 E. Pessione, M. G. Giuffrida, L. Prunotto, C. Barello, R. Mazzoli, D. Fortunato, A. Conti and C. Giunta, *Proteomics*, 2003, **3**, 1070-1076.

23 A. Toren, E. Orr, Y. Paitan, E. Z. Ron and E. Rosenberg, *J. Bacteriol.*, 2002, **184**, 165-170.

24 A. Pautsch and G. E. Schulz GE, *Nat. Struct. Biol.*, 1998, **5**, 1013-1017.

25 A. Pautsch and G.E. Schulz, *J. Mol. Biol.*, 2000, **298**, 273-282.

26 G. Walzer, E. Rosenberg and E. Z. Ron, *Environ. Microbiol.*, 2006, **8**, 1026-1032.

27 A. Toren, E. Z. Ron, R. Bekerman and E. Rosenberg, *Appl. Microbiol. Biot.*, 2002, **59**, 580-584.

28 A. Toren, G. Segal, E. Z. Ron and E. Rosenberg, *Environ. Microbiol.*, 2002, **4**, 257-261.

29 S. Moens and J. Vanderleyden, *Arch. Microbiol.*, 1997, **168**, 169-175.

30 C. Schäffer and P. Messner, *Biochimie*, 2001, **83**, 591-599.

31 H. Nothaft and C. M. Szymanski, *Nat. Rev. Microbiol.*, 2010, **8**, 765-778.

32 M. Nita-Lazar, M. Wacker, B. Schegg, S. Amber and M. Aebi, *Glycobiology*, 2005, **15**, 361-367.

33 L. Balonova, L. Hernychova and Z. Bilkova, *Expert. Rev. Proteomics*, 2009, **6**, 75-85.

34 S. H. Shakin-Eshleman, S. L. Spitalnik and L. Kasturi, *J. Biol. Chem.*, 1996, 271, 6363-6366.

35 A. J. Petrescu, A .L. Milac, S. M. Petrescu, R. A. Dwek and R. Wormald, *Glycobiology*, 2004, **14**, 103-114.

36 M. Kowarik, N. M. Young, S. Numao, B. L. Schulz, I. Hug, N. Callewaert, D. C. Mills, D. C. Watson, M. Hernandez, J. F. Kelly, M. Wacker and M. Aebi, *EMBO J.*, 2006, **25**, 1957-1966.

37 B. Sánchez, P. Bressollier and M. C. Urdaci, *FEMS Immunol. Med. Microbiol.*, 2008, **54**, 1-17.

38 J. Antikainen, V. Kuparinen, K. Lahteenmaki and T. K. Korhonen, *J. Bacteriol.*, 2007, **189**, 4539-4543.

39 T. P. Pirog, T. A. Grinberg and Y. R. Malashenko, *Mikrobiologiya*, 1997, **66**, 335–340.

40 R. Mazzoli, P. Fattori, C. Lamberti, M. G. Giuffrida, M. Zapponi, C. Giunta and E. Pessione, *Mol. Biosyst.*, 2011, **7**, 598-607.

41 S. Leone, A. Molinaro, E. Pessione, R. Mazzoli, C. Giunta, L. Sturiale, D. Garozzo, R. Lanzetta and M. Parrilli, *Carbohydr. Res.*, 2006, **341**, 582-590.

42 S. Leone, L. Sturiale, E. Pessione, R. Mazzoli, C. Giunta, R. Lanzetta, D. Garozzo, A. Molinaro and M. Parrilli, *J. Lipid. Res.*, 2007, **48**, 1045-1051.

43 L. M. Deshpande, B. P. Kapadnis and B. A. Chopade, *Biometals*, 1993, **6**, 55-59.

44 R. Andreazza, B. C. Okeke, S. Pieniz and F. A. Camargo, *Biol. Trace Elem. Res.*, 2012, **146**, 107-15.

45 S. E. Barbaro, J. T. Trevors and W. E. Inniss, *Can. J. Microbiol.*, 2011, **47**, 194-205.

46 L. M. Parsons, F. Lin and J. Orban, *Biochemistry*, 2006, **45**, 2122-2128.

47 T. Tachikawa and J. Kato, *Biosci. Biotechnol. Biochem.*, 2011, 75, 162-164.

48 E. Pessione and C. Giunta, *Microbios*, 1997, **89**, 105-117.

49 M. P. Molloy, *Anal. Biochem.*, 2000, **280**, 1-10.

50 M. P. Molloy, B. R. Herbert, B. J. Walsh, M. I. Tyler, M. Traini, J. C. Sanchez, D. F. Hochstrasser, K. L. Williams and A. A. Gooley, *Electrophoresis*, 1998, **19**, 837-844.

51 U. K. Laemmli, *Nature*, 1970, **227**, 680-685.

52 J. C. Sanchez, V. Rouge, M. Pisteur, F. Ravier, L. Tonella, M. Moosmayer, M. R. Wilkins and D. F. Hochstrasser, *Electrophoresis*, 1997, **18**, 324-327.

53 V. Neuhoff, N. Arold, D. Taube and W. Ehrhardt, *Electrophoresis*, 1988, **9**, 255-256.

54 K. D. Speicher, O. Kolbas, S. Harper and D. W. Speicher, *J. Biomol. Tech.*, 2000, **11**, 74-86.

55 D. J. Pappin, P. Hojrup and A. J. Bleasby, *Curr. Biol.*, 1993, **3**, 327-332.

56 A. Reuter, D. Fortunato, L. P. Garoffo, L. Napolitano, S. Scheurer, M. G. Giuffrida, S. Vieths and A. Conti, *Proteomics*, 2005, **5**, 282-9.

57 E. Pessione, *Front. Cell. Infect. Microbiol.*, 2012, **2**, 86.

58 R. Mazzoli, C. Lamberti and A. Pessione, *Trends Biotechnol.*, 2012, **30**, 111-119.

Figure captions

Fig. 1. 2DE maps of membrane protein extracts (pI 4-7) from *Acinetobacter radioresistens* S13 grown on benzoate as the sole carbon source. (A) Sypro Ruby stained gel. (B) ProQ Emerald stained gel.

Fig. 2. Amino acid sequence coverage of the *A. radioresistens* S13 glycoprotein AlnA-like. Amino acids in italic were determined by N-terminal sequencing, underlined amino acids were determined by MS/MS, amino acids in bold were covered by PMF. X: amino acid residue not identified by N-terminal sequencing. For detailed information on peptides from the different enzymatic digestions see Table 1S.

Fig. 3. Zoomed MALDI-TOF spectra of Asp-N peptides prior (A) and after (B) PNGase removal of sugar chain (for the mass spectra interpretation see the text).

Fig. 4. SDS-PAGE of raw extracellular protein extract (REPE) from *A. radioresistens* S13 grown on benzoate as the sole carbon source and collected during the exponential phase. Lane 1, molecular weight markers; lane 2, REPE. Bands numbered 1-6 were analysed by PMF for protein identification.

MALDI identifications									
Band code	**Identified proteins**	**UniProt Entry**	**N°identified peptides**	**Coverage**	**T/Exp* MW (Da)**	**MOWSE score**	**Position**	**m/z - peptide sequence**	**Miss cleavage**
1	copper resistance protein A *(A.radioresistens* SK82*)*	C6RRE5	3	8.8%	70.9/71.0	2.94	223-238	1779.743 - (K)TFYDQVK(K)	0
							428-446	2005.914 - (R)TAFHFQPEK(N)	1
							256-275	2339.260 - (K)GWASLQSIPR(T)	2
2	isocitrate lyase *(A.radioresistens* SK82*)*	C6RMG1	5	9.9%	59.4/60.0	15.5	18-23	781.359 - (K)FGDTWR(D)	0
							144-150	906.471 - (K)ELNDLFR(A)	0
							318-326	1054.579 - (R)LASGLYQFR(E)	0
							124-138	1751.936 - (K)TSVPALIEEIYTFLR(Q)	0
							245-260	1812.907 - (R)YAFLEMoxGLEDGIIVAR(T)	0
3	glycosyltransferase *(A.radioresistens* SK82*)*	C6RKJ8	3	11.3%	48.2/48.0	9.01	39-397	834.427 - (K)AAWDMLK(S)	0
							93-110	2134.050 - (K)DRTQEIADRMAAEYPVIR(V)	2
							49-70	2587.413 - (R)WPMLSVLIPAYNEEVVIEDTLR(A)	0
4	Outer Membrane Protein *(A.radioresistens* SK82*)*	Q8VPR9	11	39.9%	37.9/41.2	5325	159-166	901.512 - (R)INDALSLR(T)	0
							330-338	951.530 - (R)VFATITGSR(T)	0

							129-137	1073.457 - (K)YEFDTDAGR(A)	0
							255-264	1133.596 - (K)LVEYPNATAR(I)	0
							265-275	1182.560 - (R)IDGHTDNTGPR(A)	0
							116-128	1430.800 - (K)IKPYALLGAGHYKI(Y)	0
							291-303	1493.707 - (K)SSLVNEYNIDPSR(L)	0
							238-250	1533.803 - (K)SNIKDQYKPEIAK(V)	1
							141-158	1862.896 - (R)GLDEEGTLGNAGLGAFWR(I)	0
							94-110	1875.905 - (K)GENIAGNFYVTSDLFTK(N)	0
							304-321	1962.952 - (R)LTAQGFAWDQPIADNSTK(E)	0
5	Outer Membrane Protein *(Acinetobacter sp.* ATCC 27244*)*	C0VM65	3	9.2%	27.9/30.0	151	60-71	1287.669 - (R)NAPLAEAAFLDR(A)	0
							229-239	1290.688 - (K)FINQQVSLEGR(V)	0
							228-239	1418.766 - (R)KFINQQVSLEGR(V)	1
6	Inner membrane lipoprotein YiaD *(A.radioresistens* SK82*)*	C6RPK3	8	48.8%	21.7/22.0	1518	43-53	1162.519 - (K)SNANSSAQNNR(A)	0
							200-209	1179.622 - (R)RVEISIYATQ(-)	1
							78-89	1305.603 - (R)QQMAGTGVEVNR(N)	0
							158-172	1514.786 - (R)AQAVANYLAGQGVPR(T)	0
							110-122	1526.793 - (R)LDSGQSWTITVNPGTTNAR(I)	0
							175-195	2011.989 - (R)INAQGLGASNPIASNATAEGRE)	0
							90-109	2090.017 -	0

		(R)NPDGSVGLIMPGNITFDTNK(S)	
	90-109	2106.014 - (R)NPDGSVGLIMoxPGNITFDTNK(S)	0

Table 1 Proteins identified by MALDI TOF/TOF analysis from the SDS-PAGE of raw extracellular protein extract (REPE) from benzoate-grown *A. radioresistens* S13.

	Specific emulsifying activity (S.E.A.) (U/mg)			
Emulsifier	Substrate I	Substrate II	Substrate III	Substrate IV
EHCO	0.0440 ± 0.0010	0.0329 ± 0.0004	0.0108 ± 0.001	0.0328 ± 0.0004
REPE	23.90 ± 0.11	18.43 ± 0.58	19.88 ± 0.57	21.31 ± 0.55

Table 2 EHCO and REPE specific emulsifying activity (S.E.A.) on different hydrophobic substrates. Substrate I, 1:1 (v/v) hexadecane:2-methylnaftalene; Substrate II, peppermint essential oil; Substrate III, vitamine E; substrate IV, methyl dihydrojasmonate.