Published in final edited form as:

J Proteome Res. 2009 April 3; 8(4): 1972–1982. doi:10.1021/pr800923e.

S-layer Surface-Accessible and Concanavalin A Binding Proteins of *Methanosarcina acetivorans* and *Methanosarcina mazei*

Deborah R. Francoleon¹, Pinmanee Boontheung¹, Yanan Yang^{1,†}, Unmi Kim², A. Jimmy Ytterberg¹, Patricia A. Denny⁴, Paul C. Denny⁴, Joseph A. Loo^{1,3}, Robert P. Gunsalus^{2,*}, and Rachel R. Ogorzalek Loo^{3,*}

¹Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095

²Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095

³Department of Biological Chemistry, University of California, Los Angeles, CA 90095

⁴University of Southern California School of Dentistry, Los Angeles, CA 90089

Abstract

The outermost cell envelope structure of many archaea and bacteria contains a proteinaceous lattice termed the surface layer or S-layer. It is typically composed of only one or two abundant, often post-translationally modified proteins that self-assemble to form the highly organized arrays. Surprisingly, over a hundred proteins were annotated to be S-layer components in the archaeal species *Methanosarcina acetivorans* C2A and *Methanosarcina mazei* Gö1, reflecting limitations of current predictions. An *in vivo* biotinylation methodology was devised to affinity tag surface-exposed proteins while overcoming unique challenges in working with these fragile organisms. Cells were adapted to growth under N₂ fixing conditions, thus minimizing free amines reactive to the NHS-label, and high pH media compatible with the acylation chemistry was used. A 3-phase separation procedure was employed to isolate intact, labeled cells from lysed-cell derived proteins. Streptavidin affinity enrichment followed by stringent wash conditions removed non-specifically bound proteins. This methodology revealed S-layer proteins in *M. acetivorans* C2A and *M. mazei* Gö1 to be MA0829 and MM1976, respectively. Each was demonstrated to exist as multiple glycosylated forms using SDS-PAGE coupled with glycoprotein-specific staining, and by interaction with the lectin, Concanavalin A. A number of additional surface-exposed proteins and glycoproteins were identified and included all three subunits of the thermosome: the latter suggests that the chaperonin complex is both surface- and cytoplasmically-localized. This approach provides an alternative strategy to study surface proteins in the archaea.

rlloo@mednet.ucla.edu, robg@microbio.ucla.edu.

[†]Current Address: Agilent Technologies, 5301 Stevens Creek Blvd., Santa Clara, CA 95051-7201

Supporting Information Available: This material is available free at <http://pubs.acs.org>.

Table S-1. Surface-labeled proteins identified by at least one peptide of MOWSE score ≥ 30 ($p \leq 0.01$) from *M. acetivorans*.

Table S-2. Surface-labeled proteins identified by at least one peptide of MOWSE score ≥ 30 ($p \leq 0.01$) from *M. Mazei*.

Table S-3. Con A-eluted *M. acetivorans* proteins identified by at least one peptide of MOWSE score ≥ 30 ($p \leq 0.01$).

Table S-4. Con A-eluted *M. mazei* proteins identified by at least one peptide of MOWSE score ≥ 30 ($p \leq 0.01$).

Figure S1. Sypro Ruby-stained SDS-PAGE gel and near-western blot images from unlabeled *M. acetivorans* and *M. mazei* cells.

Data in Mascot Generic Format. HPLC-ESI-MS/MS data from a set of trypsin-digested SDS-PAGE bands of surface biotinylated *Methanosarcina acetivorans* and of *Methanosarcina mazei* have been provided in mgf format, as have analogous datasets for SDS-PAGE bands of Concanavalin-bound *M. acetivorans* and *M. mazei*. The *.txt file extension can be substituted with *.mgf.

Search Results. Mascot search results corresponding to the data above.

Keywords

S-layer; surface proteins; *Methanosarcina acetivorans*; *Methanosarcina mazei*; biotinylation; mass spectrometry; glycoproteins; Concanavalin A

Introduction

The *Methanosarcina*, including *M. acetivorans* C2A and *M. mazei* Gö1, are nutritionally versatile methanogens, capable of producing methane from the known range of substrates.¹⁻⁵ Their metabolic versatility is reflected in part by their large genomes of 5.75 Mb for *M. acetivorans* and 4.1 Mb for *M. mazei*, representing the largest and the fourth-largest sequenced archaeal genomes to date. The *Methanosarcinaceae* are distributed widely throughout the environment and inhabit both marine and fresh water ecosystems. They are present in ruminants, soils, sediments, and sewage sludge. As methane producers, they contribute greenhouse gases to the atmosphere and play a pivotal role in processing decaying organic matter within the global carbon cycle. They also have applications producing alternative fuels from low cost feed-stocks and biowaste. The genus is unique in its ability to propagate either as multicellular packets or as individual cells, depending on the cell environment. In freshwater (low osmolarity), they form large multi-cellular aggregates with each cell surrounded by a surface layer protein sheath (S-layer) and a methanochondroitin outer layer. At higher osmolarity (e.g., in marine medium), they propagate as individual cells surrounded by the S-layer only.⁶⁻¹⁰ Despite the abundance of these microbes and their environmental importance, relatively little is known about the *Methanosarcinaceae* with respect to the protein composition of this outermost cell envelope.

Present in archaea and in many bacteria, S-layers are highly-organized proteinaceous, two-dimensional crystalline arrays of one or more proteins that self-assemble to envelope the entire cell surface. S-layer proteins are sometimes glycosylated with molecular weights ranging from 40-200 kDa. The surface boundary layer, possibly the most ancient biological membrane,^{9, 11,12} is thought to play a critical role in the organism's interaction with the external environment, in nutrient uptake, cell excretion, signaling and surface interactions.^{13,14} As in most archaea that lack pseudomurein, the *M. acetivorans* and *M. mazei* S-layer is a protective coat surrounding the lipid membrane and contributes to cell size and shape.¹⁵⁻¹⁸ These proteins are also among the most abundant in the cell throughout the entire growth cycle.^{12, 19-21} Previously, *Methanosarcina mazei* strain S-6 DNA sequences were retrieved from antibody-reactive expression products in plasmid libraries. Those S-6 sequences correspond to orthologs MM1588 (*slgB*) and MM2440 in genome-sequenced *M. mazei* Gö1. Experimental data is needed to test and extend previous *in silico* predictions of the S-layer component proteins.

Traditionally S-layer proteins have been identified from cell-wall or membrane extracts or recovered from the medium in which cells had been grown. Here, we employ a newly devised method to retrieve exposed proteins directly from cell surfaces that involves *in vivo* biotin-tagging, affinity purification, SDS-PAGE and mass spectrometry (MS).²²⁻²⁸ These archaeal cells, as anaerobic, fragile cultivars, are uniquely challenging substrates for *in-vivo* tagging methods. From these efforts, the S-layer proteins of *M. acetivorans* and *M. mazei* are identified and information regarding post-translational modifications is revealed. Additional surface-exposed and Concanavalin A binding proteins are also identified.

Experimental Procedures

Cell Cultivation

M. acetivorans C2A and *M. maei* Göl were grown as single cells (disaggregating) as previously described with the following modifications.¹⁰ To minimize the presence of primary amines that could react with the biotinylation reagent, cells were grown in a medium lacking NH₄Cl or added organic nitrogen. The basal medium (marine medium) was prepared by the Hungate technique²⁹ using methanol at a final concentration of 0.05 M as the sole source of carbon and energy. Cells were cultivated in 10 mL medium in sealed anaerobe tubes (Difco, Sparks, MD) with a N₂-CO₂ (4:1) atmosphere headspace at 37°C. For *M. acetivorans*, the final pH was adjusted to 7.4 to favor subsequent cell labeling, while the *M. maei* medium was pH 7.0, because a higher pH did not support sufficient cell growth. These conditions differ from the optimal growth conditions for single cells in minimal medium.^{7,10} Cultures were harvested at an OD₆₀₀ of ~0.45, equivalent to approximately 4.5×10^8 cells/mL.

Biotin Labeling

Culture tubes were unsealed and their contents were transferred to 15-mL centrifuge tubes. EZ-Link Sulfo-NHS-LC-LC-biotin (Sulfosuccinimidyl-6'-(biotinamido)-6-hexanamido hexanoate, MW 669.75 g/mol, Pierce) was added (5 mg) and incubated aerobically at room temperature for 30 minutes.

After labeling, density centrifugation separated intact cells from the reaction mixture, employed as follows: the reaction mixture was transferred to a 15 mL plastic centrifuge tube and 0.5 mL of silicone oil (GE Versilube F-50, density = 1.038 g/mL) was added. Archaeal cells were sedimented by centrifugation at $1,125 \times g$ for 10 min. through the silicone oil underlay.³⁰ After centrifugation, three phases were apparent, the heaviest being the intact cell pellet, followed by the oil layer, and the aqueous top layer. The middle, oil phase served to isolate the cell pellet from soluble material in the reaction mixture, including excess label, secreted material, and proteins derived from lysed cells. Excess labeling reagent in the top layer was quenched with 100 mM glycine (1 mL) for 5 min. at room temperature. After quenching, the layered contents were frozen by dipping the tube in liquid nitrogen and stored at -80°C.

Immediately before use, the centrifuge tube tip containing the visible cell pellet was cut away and transferred to a 1.5 mL microcentrifuge tube containing 500 µL of lysis buffer (2% (w/v) CHAPS (3-(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate) in pH 7.0 phosphate buffered saline (PBS) supplemented with protease inhibitors (Sigma P8465, 2.25 µL/10 mL culture). The cell pellets were further disrupted by multiple freeze/thaw cycles and vortexing. Released insoluble material was removed by centrifugation at $16,000 \times g$ for 15 min. at 4°C. The soluble extract was passed through a Zeba™ desalt spin column (Pierce) by centrifugation at $1500 \times g$ for 2 min., as per the manufacturer's protocol.

Streptavidin Affinity Purification

Streptavidin-coated magnetic beads (200 µL, binding capacity ~ 0.2 µg/µL biotinylated protein) were washed four times with 200 µL of PBS (pH 7.0). The beads (Dynabeads®-MyOne-Streptavidin T1, Invitrogen) were then mixed with 200 µL of extract and incubated at room temperature for 30 min. The protein-bound beads were then pelleted and the unbound material (supernate) was discarded. The beads were washed six times with 200 µL of TPBS/SDS (PBS/0.1% Tween-20 (v/v)/2% SDS (w/v)) to remove non-specifically bound proteins. Biotinylated proteins were eluted by heating the beads to 95°C in NuPAGE® LDS sample buffer (Invitrogen). After pelleting the beads, the recovered eluate was analyzed by SDS-PAGE and near-western blotting.

Enrichment of α -D-Mannose- or α -D-Glucose-Linked Glycoproteins

Concanavalin A (Con A) coupled agarose beads (500 μ L, binding capacity $\sim$ 4 mg/mL, Vector Laboratories) were transferred to a Handee™ Mini Spin Column (Pierce). The beads were washed four times with 500 μ L of 50 mM Tris, 0.15 M NaCl (pH 7.5), followed by four washes with binding buffer (BB, 50 mM Tris, 0.15 M NaCl, 1 mM CaCl_2 , 1 mM $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ (pH 7.5) to equilibrate. Biotinylated cell lysate (250 μ L, $\sim$ 100 μ g labeled protein) together with 250 μ L of BB were applied to the column and incubated at room temperature for 30 min., after which the flow through was collected by centrifugation. The protein-bound beads were washed six times with 500 μ L of BB supplemented with 0.1 % of Tween-20, with the column spun and the supernate collected after each wash. The bound glycosylated proteins were eluted from the Con A beads in four 500 μ L additions of pH 7.5 elution buffer (50 mM Tris, 0.15 M NaCl, 1 mM CaCl_2 , 1 mM $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 0.2 M methyl- α -D-mannopyranoside and 0.2 M of methyl- α -D-glucopyranoside). The combined elutions were concentrated to 500 μ L by ultrafiltration through a cellulose membrane (Amicon 30 kDa cutoff, Millipore), after which they were analyzed by SDS-PAGE and near-western blotting.

To determine accurate masses for Con A-interacting forms of MM1976, cell lysates were prepared as described above from 8 pelleted 10 mL cultures ($\text{OD} \cong 1.48$) grown in marine medium containing NH_4Cl and 0.2 M $\text{NaCl}/\text{CH}_3\text{OH}$, pH 6.8, selected to yield higher cell densities. The scaled-up procedure was similar to that described above, but employed multiple centrifuge columns containing 2 mL of Con A agarose. Similar buffers were employed throughout, except that elution employed 50 mM $(\text{NH}_4)\text{HCO}_3/0.2$ M methyl- α -D-mannopyranoside/0.2 M methyl- α -D-glucopyranoside. Elutions were combined and concentrated to 1.6 μ g/ μ L as measured by BCA assay (Pierce).

For matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), 0.5 μ L of concentrated eluate was spotted onto a sample stage (previously layered with lanolin) followed by 0.5 μ L of matrix (saturated 2,4,6-trihydroxy acetophenone (THAP) in 7:3 $\text{CH}_3\text{CN}:\text{H}_2\text{O}$ (v/v) acidified to 0.1% (v/v) with trifluoroacetic acid). The Voyager DE-STR time-of-flight mass spectrometer (Applied BioSystems) was operated in linear mode with 337 nm irradiation and positive ion detection.

Enzymatic Deglycosylation on Bead

After binding and 6×200 μ L washes with TPBS/SDS, streptavidin-conjugated beads were washed three more times with PBS, and resuspended in 200 μ L of PBS for division into two aliquots. To one aliquot was added 3 μ L (2.5 U/mL) of PNGase F (N-glycosidase F from *Chryseobacterium meningosepticum*, Prozyme), while the second aliquot remained untreated. Both aliquots were incubated for 24 hr. at 37°C with mild rotation, after which bound and unbound proteins were recovered in NuPAGE® LDS sample buffer for SDS-PAGE and near-western blotting.

SDS-Polyacrylamide Gel Electrophoresis, Blotting, and Protein Visualization

Proteins were resolved on NuPAGE 4-12% Bis-Tris gels using MES Running Buffer (Invitrogen). Total proteins were fluorescence stained with SYPRO Ruby (Bio-Rad), and gel images were captured using a Molecular Imager FX scanner and PDQuest Image Analysis Software (Bio-Rad). Glycosylated proteins were revealed by Pro-Q Emerald 300 Glycoprotein Stain following the manufacturer's protocol (Molecular Probes). The glyco-stained gels were imaged using an Alpha Innotech Imager (280 nm excitation, 530 nm emission) with AlphaImager software (Alpha Innotech). To detect biotinylated proteins by near-western blotting, SDS-PAGE gels were electroblotted to Immobilon-P PVDF (Polyvinylidene difluoride) membrane (0.45 μ m, Millipore) using an XCell II blot module and NuPAGE® Transfer Buffer (Invitrogen). The membranes were blocked with 2% bovine serum albumin

(BSA, Sigma) in TPBS (pH 7.0) for 1 hr. at room temperature, after which they were incubated with 1:20,000 streptavidin-HRP (streptavidin conjugated to horseradish peroxidase) (GE Healthcare) in 2 % BSA for 1 hr. at room temperature. They were then washed four times with TPBS for 10 min. each, and the labeled bands were visualized by chemiluminescence (ECL reagent *PLUS* (PerkinElmer) and film (X-Omat Blue XB-1, Kodak Scientific Imaging)). Visualized blots were compared to SYPRO Ruby- and Pro-Q Emerald-stained gels. Protein molecular weight standards were obtained from Molecular Probes (Candy Cane™ glycoprotein standards) and Bio-Rad (biotinylated broad range protein standards).

Trypsin Proteolysis and nano-HPLC-MS/MS

Bands for protein identification were excised from the SYPRO Ruby-stained gels by a spot-excision robot (Proteome Works, Bio-Rad). The gel-embedded proteins were reduced, iodoacetamide-alkylated and trypsin-digested³¹ (Promega, sequencing grade modified trypsin). Product peptides were extracted in 50 % acetonitrile/0.1 % trifluoroacetic acid in water with the resulting extracts dried by vacuum centrifugation. Peptides were dissolved in 10 μ L of 0.1% formic acid (FA) solution and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) with electrospray ionization (ESI) on an Applied BioSystems QSTAR® Pulsar XL (QqTOF) mass spectrometer equipped with a nanoelectrospray interface (Protana), a Proxeon (Odense) nano-bore stainless steel emitter (30 μ m i.d.), and an LC Packings nano-LC system. The nano-LC was equipped with a homemade precolumn (150 μ m $\times$ 5 mm) and analytical column (75 μ m $\times$ 150 mm) packed with Jupiter Proteo C12 4- μ m resin (Phenomenex). Typically 6 μ L of sample was loaded onto the precolumn, washed with loading solvent (0.1% FA) for 4 min. and injected onto the LC column. To the column at solvent flow 200 nL/min., was applied the gradient: 3% B to 6% B in 6 sec, 6% B to 24% B in 18 min, 24% B to 36% B in 6 min, 36% B to 80% B in 2 min, and maintained at 80% B for 7.9 min. The column was finally equilibrated with 3% B for 15 min prior to the next run. Eluents used were 0.1% FA (aq) (solvent A) and 95% CH₃CN containing 0.1% FA (solvent B).

Peptide product ion spectra were recorded automatically during LC-MS/MS by information-dependent analysis (IDA) software on the mass spectrometer. Collision energies for maximum fragmentation efficiencies were calculated by the software using empirical parameters based on the charge and mass-to-charge ratio of the peptide precursor ion and argon was employed as collision gas. Proteins were identified by database searches utilizing the Mascot database search engine (Matrix Science). Searches were performed against *Methanosarcina acetivorans* protein¹ or *Methanosarcina mazei* protein³ databases supplemented with keratin and trypsin sequences. Protein sequence searches employed one missed cleavage and a mass tolerance of 0.3 Da for both precursor and product ions. Proteins hits were accepted based on ≥ 2 ascribed peptides, at least one of which possessed a MOWSE score ≥ 26 ($p \leq 0.03$ in *M. acetivorans*, $p \leq 0.02$ in *M. mazei*). Correspondences between MS/MS spectra and ascribed sequences were also verified manually.

The Sosui algorithm (<http://bp.nuap.nagoya-u.ac.jp/sosui/>) was employed to predict transmembrane helices and signal peptides.

Results

The *in-vivo* labeling of anaerobic, osmotically fragile archaea cells such as *M. acetivorans* and *M. mazei* presents several unique challenges, because (i) resuspending cells in typical labeling-optimized buffers or significantly altered media formulations can induce lysis, (ii) primary amines present in cell culture medium interfere with cell labeling chemistry, and (iii) intact cells need to be separated from lysis debris, post-labeling, to reduce background. Therefore, an experimental approach was devised to affinity tag, enrich, and analyze cell surface proteins from *Methanosarcina* spp. (Experimental Procedures, and illustrated in Fig. 1). Cells were

cultivated under nitrogen-fixing conditions to circumvent inclusion of ammonium salts or organic nitrogen in the media. Next, protein affinity tagging, binding, washing, and elution from streptavidin-conjugated beads were empirically optimized by comparing visualized total proteins in SDS-PAGE gels to the labeled proteins visualized from near-western blots (Fig. 2). The biotinylated proteins revealed in the whole cell lysate lanes (Fig. 2, right, lanes 1-2) demonstrate the efficiency of biotin labeling for both *Methanosarcina* spp. Biotinylated proteins from the indicated streptavidin-eluate lanes were excised and trypsin-digested. A total of thirteen *M. acetivorans* and ten *M. mazei* peptide mixtures analyzed by tandem mass spectrometry (Experimental Procedures) resulted in the identification of twelve *M. acetivorans* proteins (Table 1) and seventeen *M. mazei* proteins (Table 2). Expanded protein lists that include identifications based on a single peptide (score ≥ 30 , $p \leq 0.01$) are presented as supplemental material in Tables S-1 and S-2 for *M. acetivorans* and *M. mazei*, respectively. Unstained regions were also examined by mass spectrometry, but yielded no additional protein identifications.

For *M. acetivorans*, three gel-resolved species with apparent sizes of ~ 134 , ~ 119 , and ~ 114 kDa (Fig. 2A, bands 3-5) were identified by the tandem mass spectrometry procedure to contain MA0829 (Fig. 3). Annotated as a hypothetical protein,³ the primary sequence of MA0829 predicts a molecular weight (MW) of 74.3 kDa. Because the three SDS-PAGE bands migrated to positions 40 to 60 kDa higher in apparent mass (Fig. 2A), their anomalous migration suggests that each of the three MA0829 forms is modified post-translationally.

Within the list of seventeen *M. mazei* retrieved proteins (Table 2), ORF MM1976 was identified as an ortholog of MA0829. Also appearing at multiple positions on SDS-PAGE gels, *M. mazei* MM1976 exhibited apparent molecular weights of ~ 131 kDa, ~ 119 kDa, and ~ 101 kDa. Migrating 30-60 kDa above its sequence-calculated 74.0 kDa size, it also appears to be post-translationally modified. However, in contrast to the two highly abundant MA0829 species, only the ~ 101 kDa MM1976 form appeared to be highly abundant and biotinylated (Fig. 2B left, band 3 and Fig. 2B right). Although the other two MM1976 protein forms were purified through the streptavidin magnetic beads, no biotinylation was detected in the near-western blot (Fig. 2 right, lane 5). The lack of streptavidin-binding of the larger species (*i.e.*, slower migrating forms) in the western blot might reflect very strong non-covalent interactions between different forms of MM1976, such that they are withdrawn together by the streptavidin-conjugated beads. If so, this observation suggests that only the lower apparent molecular weight form is surface-exposed. Alternative explanations are that, under these pH conditions, biotin might modify some glycans or tyrosines in this protein, resulting in somewhat labile modifications, or that select modifications reduce labeling efficiency. For example, sulfated or negatively-charged sugars could electrostatically repel sulfo-NHS ester reagents, reducing biotin-tagging.

Investigating the Post-Translational Modifications of MA0829 and MM1976

The surface-labeled *M. acetivorans* MA0829 and *M. mazei* MM1976 proteins are each predicted to contain a signal peptide sequence plus one C-terminal transmembrane (TM) domain (discussed below). Tandem MS confirmed the predicted processed N-termini: VDVIEIR (MA0829) and ADVIEIR (MM1976). Interestingly, the MM1976 and MA0829 N-termini were found in unmodified and acetylated forms, with the former species most abundant (Fig. 3).

Glycosylated S-layer proteins have been observed in archaea,^{9,32-34} suggesting the modification as a potential cause for the anomalous migration of MA0829 and MM1976 by SDS-PAGE. Duplicate SDS-PAGE gels loaded with the streptavidin-fractionated proteins from biotinylated cells were incubated in glycosylation-specific, Pro-Q Emerald stain (Experimental Procedures). The fluorescence-visualized glyco-stained gel (Fig. 4) indicated

glycosylation of the two largest forms of *M. acetivorans* MA0829 (~134, and ~119 kDa) and in the smaller form of *M. mazei* MM1976 (~101 kDa).

To further examine the nature of the MA0829 and MM1976 modifications, protein deglycosylation using PNGase F (N-glycanase) was attempted. This enzyme was selected based on the suggested prevalence of archaeal N-linked glycosylation.³³ It specifically releases intact N-linked oligosaccharides from glycopeptides and less efficiently, from glycoproteins. N-linked glycans are covalently linked to proteins via asparagine side chains in Asn-Xaa-Ser/Thr motifs, where Xaa can be any amino acid, except proline. SDS-PAGE and streptavidin-visualized western blots of PNGase F-incubated proteins showed reduced intensities for all *M. acetivorans* MA0829 bands (see Fig. 5), suggesting excised N-linked glycans and corresponding altered electrophoretic migration. In contrast, no change was apparent after treatment of MM1976 with PNGase (figure not shown). Glycosidase-treated *M. acetivorans* proteins were excised from the gel in segments that were trypsin-digested and analyzed by nano-HPLC-MS/MS. In the indicated lane, MA0829 was identified at a distinct position of ~89.2 kDa (upper arrow in Fig. 5) in support of the proposed N-glycosylation.

Exploring the Glycan Composition of MA0829 and MM1976

Lectin affinity purification was employed to further elucidate glycosylation of MA0829 and MM1976. By binding to specific saccharide residues and combinations, lectins provide a useful probe of composition.³⁵⁻³⁷ Concanavalin A reacts with the ring forms of non-reducing α -D-mannose and α -D-glucose, and facilitates many glycoprotein purifications. Among N-linked sugar motifs, it binds especially strongly to high-mannose, less strongly to the hybrid- and biantennary complex, and poorly to highly branched complex glycans.

Biotinylated *Methanosarcinaceae* cell lysate fractions eluted from Con A-linked agarose were analyzed by SDS-PAGE, near-western blotting, and mass spectrometry (Fig. 6). Six *M. acetivorans* and nine *M. mazei* proteins bound to Con A, putatively possessing α -D-mannose or α -D-glucose in their glycans (Tables 3 and 4). Because glycosylation has been observed in both S-layer and membrane proteins of archaea,³² the proteins listed in Tables 3 and 4 were queried to predict signal peptides and transmembrane helices using the Sosui algorithm (Experimental Procedures). Three out of six *M. acetivorans* and six of nine *M. mazei* proteins were predicted to have signal peptides and at least one transmembrane helix.

The ~114 kDa MA0829 glycoform was recovered among the Con A-bound proteins. This form, previously observed bound to streptavidin with two more slowly migrating forms, was the least abundant of the three species as assessed by total protein staining (Fig. 6A). Nevertheless, it was readily detected in Con A-eluate by near-western blotting. Because the two other MA0829 glycoforms (~134 and ~119 kDa) were not observed in the Con A eluate, different glycan compositions are suggested.

For *M. mazei*, all three MM1976 forms (~131, ~119, and ~101 kDa) were retrieved by the Con A affinity procedure (Fig. 6B, left), as well as three variants, identified at ~72, ~65, and ~59 kDa in SDS-PAGE gels. The near-western blot revealed biotinylation of all six bands (Fig. 6B, right), implying either that Con A was more efficient than streptavidin in their retrieval, or that MM1976 was degraded during the Con A purification. Con A eluate was also spotted directly onto a sample stage for analysis by MALDI-MS. Singly and doubly charged ions indicative of proteins 73180, 74450, and 75890 Da +/- 150 Da were observed, which we attribute to the three MM1976 glycoforms.

Other identified affinity tagged proteins

The three subunits comprising the *M. mazei* thermosome (MM1379, MM0072, and MM1096), were recovered by streptavidin purification (Band 5 of Fig. 2B). This eukaryotic-type chaperonin complex is sometimes called the “rosettasome.” Also retrieved by streptavidin affinity tagging was the 63 kDa cytoplasmic pyruvate carboxylase subunit B (oxaloacetate decarboxylase) in both *Methanosarcina* species (*i.e.*, MA0674 and MM1827). This endogenously biotinylated subunit is homologous to the soluble biotin-containing subunit PYCB of *Methanosarcina barkeri* that is part of the $\alpha_4\beta_4$ -type acetyl CoA-independent pyruvate carboxylase (PYC) complex. This archaeal enzyme is involved in the CO₂ fixation needed for cell anabolic reactions.³⁸ The PYC B-subunit was also observed in unlabeled-cell control experiments performed to reveal endogenously biotinylated and non-specifically bound proteins (Fig. S1).

Discussion

The major affinity tagged proteins recovered from surface-labeled *M. acetivorans* and *M. mazei* cells were MA0829 and MM1976, respectively. As revealed by the tandem MS experiments, MA0829 and MM1976 proteins contain signal peptide sequences of 24 amino acids that are removed to yield the processed N-termini VDVIEIR (MA0829) and ADVIEIR (MM1976). The MM1976 and MA0829 N-termini were also found in both unmodified and acetylated forms.

Protein sequence analysis³⁹ of the two *Methanosarcina* proteins revealed tandem duplicated DUF1608 domains, related to motifs found within the *Methanococcus voltae* S-layer protein.⁴⁰ A consequence of these duplicated sequence domains is that peptide mapping of MM1976 yields seven tryptic peptides that occur twice, while MA0829 has four duplicated tryptic peptides. Our data demonstrate that *Methanosarcinaceae* surface exposure is clearly correlated to DUF1608 duplication domains. Pandit phylogenies⁴¹ based on these S-layer related duplication domains retrieved detected proteins MA0829, MA0068, MA3556, MM0467, MM1364, and MM1976, as well as proteins MA0884, MA3639, and MM1816, not detected in our streptavidin and Con A affinity experiments. DUF1608 domains occur in pairs within these proteins, except for MM1816, which has only one domain. Similarly, Syssters⁴² protein sequence clusters for the P137954 protein family include the above 6 detected proteins, along with undetected MA3598. MA3598 also lacks the second DUF1608 domain. Among the observed proteins, MM0467, eluted from Concanavalin A, contains a predicted signal peptide, as does MA0068, routinely observed from our *in vivo* biotinylated preparations, but less consistently bound to Concanavalin A. Similarly, MA0829, MM1976, and MA3556 are predicted by both the SignalP⁴³ and Sosui algorithms^{44,45} to possess signal peptides. A signal peptide was predicted for MM1364 by only SignalP.

The *M. acetivorans* MA0829 and *M. mazei* MM1976 proteins each share low similarity to S-Layer proteins identified in other organisms, but do possess several common traits.³⁴ First, each is abundant, consistent with S-Layer proteins' known cellular abundance, comprising 10-15% of an organism's protein content.^{12,19-21,46} Second, MA0829 and MM1976 are present as multiple bands by SDS-PAGE and migrate anomalously, as though 30-60 kDa larger than predicted by their primary sequences. Third, the glycosylation-specific protein staining and lectin binding experiments indicate that both proteins are glycosylated. MA0829 and MM1976 contain four and five potential N-glycosylation sites, respectively. All forms of MM1976 bound Concanavalin A strongly, indicative of α -D-linked mannose or α -D-glucose, while only one MA0829 band interacted with the lectin.

Comparison to Previous Predictions and Measurements

Macario and de Macario⁴⁷⁻⁵⁰ attributed genomic library expression products to S-layers from *M. mazei* strain S-6, guided by the products' reactivity to *M. mazei* cell surface antibodies and by homology to *Methanothermobacter fervidus* and *Methanothermobacter sociabilis* S-layer proteins.⁵¹ *In vivo* expression of two proteins was established indirectly, by demonstrating that similarly-sized components of *M. mazei* S-6 whole cell lysate, membrane, and S-layer fractions reacted with the surface antibodies.⁴⁷⁻⁵⁰ These *M. mazei* S-6 sequences encode proteins SlpB and "ORFs 492/378/783," and correspond to the *M. mazei* Gö1 MM1588 and MM2470 proteins. Interestingly, we did not detect MM1588 or MM2470 nor did we detect either of the two *M. acetivorans* homologues, MA0336 and MA1904. This discrepancy may reflect fundamental differences between the *M. mazei* S6, *M. mazei* Gö1, and *M. acetivorans* C2A proteomes, differences in cell cultivation procedures, or overlapping antibody specificities. Further studies are needed to resolve this point.

Annotation of *Methanosarcina* genomes reveal many proteins with predicted cell envelope functions. For example, Maeder, *et al.*² predicted 277 and 235 cell envelope genes in *M. acetivorans* and in *M. mazei*, respectively. The large number was rationalized as enabling *Methanosarcina* spp. to alter their cell surface compositions and/or methanochondroitin layers to aid cell survival and/or adapt to environmental changes. Many of the annotations were based only on sequence homology predictions to distantly related organisms rather than experimental data derived from *Methanosarcina* species. The *M. mazei* genome annotation³ listed 160 cell envelope-related genes (5.2% of the genome) with potential roles in cell surface protein or heteropolysaccharide layer synthesis. Of these, fourteen gene annotations contained "surface" in their description. However, none of these were observed in the *in vivo* cell labeling experiments. Similarly, eighty five *M. acetivorans* genes contained annotations with the keyword "surface" (<http://www-genome.wi.mit.edu/>). Sixty-two genes were annotated as surface-resident¹ based on homology to metazoan surface antigens and to the *M. mazei* strain S-6 proteins discussed above.^{49,52} Of three explicitly annotated as genes for "S-layer proteins", MA1286, MA1961, and MA2457, none were detected in our experiments. Of the predicted "surface-resident" proteins, only one (MA0336) was detected, albeit only when lower stringency (without SDS) washes was employed. In a subsequent study, Adindla, *et al.* predicted 23 *M. acetivorans* and 16 *M. mazei* proteins as surface-resident based on their novel sequence repeats.^{53,54} None of those proteins were observed in our *in vivo* labeling experiments. The above proteins previously predicted as "surface-exposed" may be present but at low levels. The major goal of these experiments was to identify the S-layer protein component(s) and abundant surface-exposed proteins by affinity labeling, and thus the very stringent washing conditions employed to reduce non-specific binding could have reduced *bona fide*, but low-abundance tagged proteins to levels below detection.

Glycoproteins MM1976, MA0829, and *Methanococcus voltae* S-layer protein Q50833⁴⁰ contain tandem duplicated domains of ~250 amino acids which, in the *Methanosarcinaceae*, are referenced variously as PF007752, IPR006457 or DUF1608. However, BlastP comparisons find no significant similarity between the *Methanococcus* and *Methanosarcinaceae* proteins. An ortholog to these *Methanosarcinaceae* proteins (MBUR_1690, Q12VE2_METBU, ZP_00562202.1)⁵⁵ was also previously identified in cell-free supernatant from *Methanococcoides burtonii*, a psychrotolerant organism in the *Methanosarcina* family. However, the *M. acetivorans* and *mazei* homologues closest to the most abundant secreted *Methanococcoides burtonii* protein (MBUR_1349) were not observed in our study, nor were the homologues closest to the 5 other described proteins (MBUR_1027, MBUR_1109, MBUR_1118, MBUR_2003, and MBUR_2064).⁵⁵ Of the 18 *M. acetivorans* and *mazei* proteins at least 40% identical to *M. burtonii* secreted proteins; *i.e.*, MM0467, MM1364, MM1976, MA0829, MA3556, MA0310, MA0315, MA0564, MA0876, MA0884, MA0957,

MA2981, MA3598, MA3639, MM1511, MM1575, MM1723, and MM1816, only the first 5 were observed by us.

Other Proteins Identified

As noted above, a total of twelve *M. acetivorans* proteins (Table 1) and seventeen *M. mazei* proteins (Table 2) were identified by the affinity tagging approach. For *M. acetivorans*, eight of twelve proteins identified were predicted to have signal peptides and seven were also predicted to possess transmembrane helices. For *M. mazei*, six of seventeen proteins identified had predicted signal peptides and four had transmembrane helices. These predictions are consistent with at least some of the identified proteins being secreted and/or membrane-anchored proteins with surface access.

Methanosarcina genomes contain both group I (bacterial-type) and group II (eukaryotic-type) chaperonins, both of which are expressed under standard growth conditions.

Methanosarcinaceae are unique in this regard; other archaeal species possess only group II chaperonins. All three subunits comprising the *M. mazei* group II chaperonin complex, known as the thermosome, were recovered from our streptavidin affinity preparations (Band 5 of Fig. 2B). The three subunits (MM1379 (α), MM0072 (β), and MM1096 (γ), respectively) are known to assemble preferentially in a molar ratio of 2:1:1 α : β : γ .^{56,57} The thermosome α -subunit was also recovered reliably from streptavidin-enriched *M. acetivorans* proteins, while the γ subunit was detected when less stringent washes were employed. Although the thermosome's presence within the cytosol is well-established, these experiments now confirm the proposal⁵⁸ that it is also an archaeal membrane element.

Consistent with these thermosome observations, Group I and II chaperonins (*i.e.*, 60 kDa heat shock proteins and GroEL), are known to be present on eukaryotic and bacterial cell surfaces, in addition to their primary cytosolic localization.^{22,24,59-62} Increased amounts of Hsp60 are also observed on the surfaces of stressed cells.⁶³⁻⁶⁸ However, the streptavidin elutions described here provided little evidence of surface-exposed GroEL: MA0631 was not detected, while *M. mazei* MM1798 was *only* observed from a run in which SDS-washing was not employed. Interestingly, hsp70 analogue MA1478 was reliably detected from Con A elutions, and frequently from streptavidin, while analogue MM2505 was detected occasionally.

Large subunits of membrane-bound F₄₂₀-nonreducing hydrogenases⁶⁹⁻⁷¹ were also retrieved from cells, consistent with previous work attributing *M. mazei* *vhoA*, to the outer membrane surface.⁷² MA1142, and MM2170/MM2313 were observed, but the high similarity between the *M. mazei* *vhtA* and *vhoA* genes prevented an MS/MS determination of whether one, or both, proteins were recovered. Glutamine synthetases, observed here from both organisms (MM3188/MA4216), have been linked to cell surfaces, based on reactivity to sera directed against group B streptococci.⁷³

InterPro⁷⁴ annotations for other observed proteins are consistent with surface exposure; *e.g.*, MM2893, recovered from streptavidin affinity chromatography, contains a DUF11 domain. This domain of about 53 amino acids has been observed in phylogenetically distant prokaryotes, including *Methanobacterium thermoautotrophicum* and *Chlamydia trachomatis*, where it may be involved in porin formation. Con A-binding protein MM1999 is annotated as containing a WD40/YVTN-like repeat domain, consistent with structures forming 7-bladed propellers in some archaeal surface proteins. MM3291, MM0001, and MM0002 contain ABC transporter domains, while MM1329 is predicted to be membrane-localized.

Considerations in Biotin Labeling Surface Proteins

Obtaining surface proteins by *in vivo* biotinylation provides an orthogonal alternative to protein extraction⁷⁵ or to recovering proteins shed into cultivation medium.⁵⁵ However, labeling these fragile organisms presented unique challenges. Cells were susceptible to lysis if transferred to buffers differing significantly in ionic strength or composition from that in which they were cultivated. *M. mazei* was more susceptible to lysis than *M. acetivorans*. This challenge was overcome by adapting cells to growth fixing N₂ (minimizing free amines reactive to the NHS-label) in a higher pH medium compatible with the acylation.

Reducing intracellular contaminants is the goal of all strategies interrogating surface proteins, and fractionating cell membranes, post-labeling, has been effective for that purpose.²⁷ Alternatively, a 3-phase separation (pelleted cells/dense oil/water) was devised and applied here to isolate labeled, intact cells from lysed-cell derived labeled proteins. The approach reduced intracellular background and should be applicable to future studies labeling other cell types.

Reactivity and size were considered when selecting the tagging reagent. Prevalent lysine residues are a prime target for tag attachment, and a fortuitous choice, given that both MA0829 and MM1976 lack cysteine, as do most S-layer proteins.¹⁹ Previous concerns that the smaller sulfo-NHS-LC-biotin (sulfosuccinimidyl-6-biotinamido)-hexanoate) tag may have entered gram-negative bacterial cells through outer-membrane porin channels led us to select the larger sulfo-NHS-LC-LC resin for our studies.²² Also, a small amount of the vitamin biotin (0.02 µg/mL) was supplied to cells in the culture medium.

The non-covalent biotin-avidin and biotin-streptavidin interactions support many useful tools to capture target species from complex mixtures. Although these strong interactions permit extensive washing to detach non-specifically bound species, they are less ideal later, when quantitative release of the specifically-bound species is sought. Typically biotin is eluted from avidin with 8M GuHCl, pH 1.5 or it is boiled in SDS-PAGE buffer.^{76,77} Monomeric avidin has been employed by several investigators to recover biotinylated surface proteins.^{24,27,78-82} Its reduced affinity to biotin is suggested to permit specific elution of tagged species in, *e.g.*, 2 mM D-biotin/PBS buffer, although investigators frequently resort to harsher elution conditions, potentially abrogating the unique specificity. Nunomora *et al.*²⁷ eluted biotin-tagged *peptides* from monomeric avidin with 30% acetonitrile/0.4% aqueous trifluoroacetic acid (TFA), carefully differentiating background- from labeled-peptides by requiring explicit evidence of biotin (*i.e.*, MS/MS product ions defining the biotinylated residue or the biotin-tag). Their *peptide-level* isolation was quite specific; only 10% of assigned peptides lacked biotin tags. However, in our *protein-level* isolations performed with monomeric avidin, recoveries from specific, 2 mM biotin elution were low, while the specificities obtained from denaturing elutions were also disappointing, (yielding significant quantities of non-biotinylated proteins), likely reflecting the tighter non-specific interactions of avidin to intact proteins, as compared to peptides. Alternative reduced-affinity biotin analogues could be employed, such as strongly-pH dependent avidin-binder 2-iminobiotin⁸³⁻⁸⁵ or reductively-cleaved sulfo-NHS-SS-biotin.^{59,86} However, the sulfo-NHS ester of 2-iminobiotin is not commercially available, while the disulfide linkage of sulfo-NHS-SS-biotin does not withstand the reducing conditions present in *Methanosarcina* cultivation media, nor will the cells withstand transfer to disulfide compatible labeling buffers.

Biotin's high affinity to avidin and streptavidin ($K_d \sim 10^{-14} - 10^{-15}$ M) permits stringent wash solutions that denature most proteins.^{76,77,87} By visualizing the total and biotinylated proteins eluted from streptavidin *via* incubation in 95°C loading buffer, we found that significant amounts of non-biotinylated proteins remained following multiple washes with PBS or tris-buffered saline, (recommended by some manufacturer protocols), and even after harsher 0.1%

SDS washes. Ultimately, we were driven to washing with room temperature 2% SDS solution to eliminate non-specifically bound proteins, trading recovery for specificity.

Protein glycosylation in archaea is suggested to occur on the outer cell surface, topologically analogous to eukaryal glycosylation, in which translocation across a membrane precedes protein modification.⁸⁸ Thus, affinity to Concanavalin A provides an alternative approach to retrieving surface proteins, although it is specific for appropriately glycosylated species, biased towards avidly binding high mannose and hybrid-type N-glycans, but against highly branched oligosaccharides. Previously, Yao, *et al.*⁴⁷ demonstrated that the major Concanavalin A binding species in *Methanosarcina mazei* S-6 was a single broad protein band migrating well beyond 100 kDa, observed for packets, single cells, and lamina. Similarly, we have found Con A affinity purification to be especially useful for rapidly enriching S-layer protein MM1976 from *M. mazei* whole cell lysates, but less useful for the *M. acetivorans* ortholog, of which only some glycoforms bound Con A. In considering the proteins recovered from Con A (Tables 2 and 4) it should be noted that the gentle binding and elution conditions may preserve non-covalent complexes, such that non-glycosylated interactors may also be recovered.

Conclusions

A number of surface proteins and glycoproteins have been identified in *M. acetivorans* and *M. mazei* by direct tagging with a cell-impermeable reagent or by binding to the lectin Concanavalin A. The S-layer proteins have been identified as MA0829 and MM1976, in contrast to predictions and to a previous experiment. The protein-level enrichments permitted protein modifications to be probed, including glycosylation and signal peptide cleavage sites. The *in vivo* tagging conditions and cultivation conditions were optimized for application to fragile organisms.^{48,49}

Having identified MM1976 and MA0829 as *Methanosarcina* S-layer proteins under the unique cultivation conditions required to biotin-tag surface proteins, we are investigating the question of whether the envelope-protein might actually change depending on cultivation condition. Antibody fluorescence cell imaging would provide an elegant means to query cells. Less specific, but also less costly, is fluorescently-tagged concanavalin A, which is allowing us to visualize con A interacting glycans on the surface of *M. mazei* and *M. acetivorans*. Finally, the abundance required for any protein to successfully envelope the cell provides us with means to conveniently track the S-layer by following the abundance of MM1976 or MA0829 under different cultivation conditions. Radical reductions in their abundance would signal a change in the identity of the S-layer protein, or at least the need to devise tagging experiments compatible with the altered conditions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors acknowledge support from the U.S. Department of Energy through the UCLA-DOE Laboratory for Genomics and Proteomics (DE-FC03-87ER60615) to J.A.L. and R.P.G. and through DE-FG03-86ER13498 to R.P.G., from the U.S. National Institute of Health (****) to J.A.L. and R.R.O.L., from the National Science Foundation NSF 0762 to R.P.G., and from the Ruth L. Kirschstein NRSA fellowship (NIH 5 F31AI061886-02) to D.R.F. The UCLA Mass Spectrometry and Proteomics Technology Center was established and equipped by a generous gift from the W. M. Keck Foundation. We thank Chris Bird for his efforts on preliminary experiments.

References

1. Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, DeArellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, de Macario EC, Ferry JG, Jarrell KF, Jing H, Macario AJL, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WW, Birren B. *Genome Research* 2002;12:532–542. [PubMed: 11932238]
2. Maeder DL, Anderson I, Brettin TS, Bruce DC, Gilna P, Han CS, Lapidus A, Metcalf WW, Saunders E, Tapia R, Sowers KR. *J Bacteriol* 2006;188:7922–7931. [PubMed: 16980466]
3. Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, Henne A, Wiezer A, Baumer S, Jacobi C, Bruggemann H, Lienard T, Christmann A, Bomeke M, Steckel S, Bhattacharyya A, Lykidis A, Overbeek R, Klenk HP, Gunsalus RP, Fritz HJ, Gottschalk G. *J Mol Microbiol Biotechnol* 2002;4:453–461. [PubMed: 12125824]
4. Li Q, Li L, Rejtar T, Karger BL, Ferry JG. *J Proteome Res* 2005;4:129–135. [PubMed: 15707367]
5. Li Q, Li L, Rejtar T, Karger L, Ferry JG. *J Proteome Res* 2005;4:112–128. [PubMed: 15707366]
6. Kandler O, Koenig H. *Archives of Microbiology* 1978;118:141–152. [PubMed: 697504]
7. Sowers KR, Baron SF, Ferry JG. *Appl Environ Microbiol* 1984;47:971–977. [PubMed: 16346552]
8. Kreisl P, Kandler O. *Systematic and Applied Microbiology* 1986;7:293–299.
9. Baumeister W, Lembecke G. *J Bioenergetics Biomembranes* 1992;24:567–575.
10. Sowers KR, Boone JE, Gunsalus RP. *Appl Environ Microbiol* 1993;59:3832–3839. [PubMed: 16349092]
11. Beveridge TJ, Graham LL. *Microbiol Rev* 1991;55:684–705. [PubMed: 1723487]
12. Messner P, Sleytr UB. *Advances in Microbial Physiology* 1992;33:213–275. [PubMed: 1636510]
13. Beveridge TJ, Pouwels PH, Sára M, Kotiranta A, Lounatmaa K, Kari K, Kerosuo E, Haapasalo M, Egelseer EM, Schocher I, Sleytr UB, Morelli L, Callegari ML, Nomellini JF, Bingle WH, Smit J, Leibovitz E, Lemaire M, Miras I, Salamitou S, Beëguin P, Ohayon H, Gounon P, Matuschek M, Sahm K, Bahl H, Grogono-Thomas R, Dworkin J, Blaser MJ, Woodland RM, Newell DG, Kessel M, Koval SF. *FEMS Microbiol Rev* 1997;20:99–149. [PubMed: 9276929]
14. Sara M, Sleytr UB. *J Bacteriol* 2000;182:859–868. [PubMed: 10648507]
15. Mescher MF, Strominger JL. *Proc Natl Acad Sci U S A* 1976;73:2687–2691. [PubMed: 1066681]
16. Wildhaber I, Baumeister W. *EMBO Journal* 1987;6:1475–1480. [PubMed: 16453767]
17. Phipps BM, Huber R, Baumeister W. *Mol Microbiol* 1991;5:253–265. [PubMed: 1904123]
18. Engelhardt H. *J Struct Biol* 2007;160:190–199. [PubMed: 17888677]
19. Boot HJ, Pouwels PH. *Mol Microbiol* 1996;21:1117–1123. [PubMed: 8898381]
20. Novotny R, Pfoestl A, Messner P, Schaffer C. *Glycoconj J* 2003;20:435–447. [PubMed: 15316277]
21. Sleytr UB, Egelseer EM, Ilk N, Pum D, Schuster B. *FEBS J* 2007;274:323–334. [PubMed: 17181542]
22. Sabarth N, Lamer S, Zimny-Arndt U, Jungblut PR, Meyer TF, Bumann D. *J Biol Chem* 2002;277:27896–27902. [PubMed: 12023975]
23. Zhang W, Ge Zhou G, Zhao Y, White MA, Zhao Y. *Electrophoresis* 2003;24:2855–2863. [PubMed: 12929181]
24. Shin BK, Wang H, Yim AM, Le Naour F, Brichory F, Jang JH, Zhao R, Puravs E, Tra J, Michael CW, Misek DE, Hanash SM. *J Biol Chem* 2003;278:7607–7616. [PubMed: 12493773]
25. Chen WNU, Yu LR, Strittmatter EF, Thrall BD, Camp DG II, Smith RD. *Proteomics* 2003;3:1647–1651. [PubMed: 12923789]
26. Rybak JN, Ettorre A, Kaissling B, Giavazzi R, Neri D, Giuliano Elia G. *Nature Methods* 2005;291–298. [PubMed: 15782212]
27. Nunomura K, Nagano K, Itagaki C, Taoka M, Okamura N, Yamauchi Y, Sugano S, Takahashi N, Izumi T, Isobe T. *Mol Cell Proteomics* 2005;4:1968–1976. [PubMed: 16176923]
28. Cordwell SJ. *Current Opinion in Microbiology* 2006;9:1–10.

29. Sowers KR, Gunsalus RP. *J Bacteriol* 1988;170:998–1002. [PubMed: 3338976]
30. Proctor LM, Lai R, Gunsalus RP. *Appl Environ Microbiol* 1997;63:2252–2257. [PubMed: 9172344]
31. Shevchenko A, Wilm M, Vorm O, Mann M. *Anal Chem* 1996;68:850–858. [PubMed: 8779443]
32. Schaffer C, Graninger M, Messner P. *Proteomics* 2001;1:248–261. [PubMed: 11680871]
33. Schaffer C, Messner P. *Biochimie* 2001;83:591–599. [PubMed: 11522387]
34. Claus H, Akca E, Debaerdemaeker T, Evrard C, Declercq JP, Harris JR, Schlott B, Koenig H. *Can J Microbiology* 2005;51:731–743.
35. Kornfeld R, Ferris C. *J Biol Chem* 1975;250:2614–2619. [PubMed: 1123324]
36. Baenziger JU, Fiete D. *J Biol Chem* 1979;254:2400–2407. [PubMed: 85625]
37. Debray H, Decout D, Strecker G, Spik G, Montreuil J. *Eur J Biochem* 1981;117:41–55. [PubMed: 7262089]
38. Mukhopadhyay B, Purwantini E, Kreder CL, Wolfe RS. *Journal of Bacteriology* 2001;183:3804–3810. [PubMed: 11371547]
39. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer ELL, Studholme DJ, Yeats C, Eddy SR. *Nucleic Acids Res* 2004;32:D138–D141. [PubMed: 14681378]
40. Konisky J, Lynn D, Hoppert M, Mayer F, Haney P. *J Bacteriol* 1994;176:1790–1792. [PubMed: 8132478]
41. Whelan S, de Bakker PIW, Quevillon E, Rodriguez N, Goldman N. *Nucleic Acids Research* 2006;34:D327–D331. [PubMed: 16381879]
42. Krause A, Stoye J, Vingron M. *BMC Bioinformatics* 2005;6:15. [PubMed: 15663796]
43. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. *J Mol Biol* 2004;340:783–795. [PubMed: 15223320]
44. Hirokawa T, Boon-Chieng S, Mitaku S. *Bioinformatics* 1998;14:378–379. [PubMed: 9632836]
45. Gomi M, Sonoyama M, Mitaku S. *Chem-Bio Info J* 2004;4:142–147.
46. Mesnage S, Fontaine T, Mignot T, Delepierre M, Mock M, Fouet A. *EMBO Journal* 2000;19:4473–4484. [PubMed: 10970841]
47. Yao R, Macario AJL, de Macario EC. *J Bacteriol* 1992;174:4683–4688. [PubMed: 1624456]
48. Yao R, Macario AJL, Conway de Macario E. *Biochim Biophys Acta* 1994;1219:697–700. [PubMed: 7948028]
49. Mayerhofer LE, Conway de Macario E, Macario AJL. *Gene* 1995;165:87–91. [PubMed: 7489922]
50. Mayerhofer LE, Conway de Macario E, Yao R, Macario AJL. *Arch Microbiol* 1998;169:339–345. [PubMed: 9531635]
51. Broeckl G, Behr M, Fabry S, Hensel R, Kaudewitz H, Biendl E, Koenig H. *Eur J Biochem* 1991;199:147–152. [PubMed: 1712296]
52. Jing H, Takagi J, Liu JH, Lindgren S, Zhang RG, Joachimiak A, Wang JH, Springer TA. *Structure* 2002;10:1453–1464. [PubMed: 12377130]
53. Adindla S, Inampudi KK, Guruprasad K, Guruprasad L. *Comparative and Functional Genomics* 2004;5:2–16. [PubMed: 18629042]
54. Adindla S, Inampudi KK, Guruprasad L. *Int J Biological Macromolecules* 2007;41:454–468.
55. Saunders NFW, Ng C, Raftery MJ, Guilhaus M, Goodchild A, Cavicchioli R. *J Prot Res* 2006;5:2457–2464.
56. Klunker D, Haas B, Hirtreiter A, Figueiredo L, Naylor DJ, Pfeifer G, Mueller V, Deppenmeier U, Gottschalk G, Hartl FU, Hayer-Hartl M. *J Biol Chem* 2003;278:33256–33267. [PubMed: 12796498]
57. Hirtreiter, AM. *Chemie und Pharmazie. Ludwig-Maximilians-Universität; München*: 2006. p. 191
58. Trent JD, Kagawa HK, Paavola CD, McMillan RA, Howard J, Jahnke L, Lavin C, Embaye T, Henze CE. *Proc Natl Acad Sci USA* 2003;100:15589–15594. [PubMed: 14673104]
59. Cicconi R, Delpino A, Piselli P, Castelli M, Vismara D. *Molecular and Cellular Biochemistry* 2004;259:1–7. [PubMed: 15124901]
60. Korbelik M, Sun J, Cecic I. *Cancer Res* 2005;65:1018–1025. [PubMed: 15705903]
61. Mayrhofer C, Krieger S, Allmaier G, K D. *Proteomics* 2006;6:579–585. [PubMed: 16372259]

62. Harding SV, Sarkar-Tyson M, Smither SJ, Atkins TP, Oyston PC, B KA, Liu Y, Wait R, Titball RW. Vaccine Early Edition 2007;25:2664–2672.published 26-Dec-2006
63. Xu Q, Schett G, Seitz CS, Hu Y, Gupta RS, Wick G. Circ Res 1994;75:1078–1085. [PubMed: 7525102]
64. Amini HR, Ascencio F, Cruz-Villacorta A, Ruiz-Bustos E, Wadstrom T. FEMS Immunol Med Microbiol 1996;16:163–172. [PubMed: 9116633]
65. Soltys BJ, Gupta RS. Exp Cell Res 1996;222:16–27. [PubMed: 8549659]
66. Garduno RA, Garduno E, Hoffman PS. Infect Immun 1998;66:4602–4610. [PubMed: 9746556]
67. Sapozhnikov AM, Ponomarev ED, Tarasenko TN, Telford WG. Cell Prolif 1999;32:363–378. [PubMed: 10646688]
68. Bergonzelli GE, Granato D, Pridmore RD, Marvin-Guy LF, Donnicola D, Cortesey-Theulaz IE. Infect Immun 2006;74:425–434. [PubMed: 16368998]
69. Deppenmeier U. J Bioenerg Biomembr 2004;36:55–64. [PubMed: 15168610]
70. Brodersen J, Baumer S, Abken HJ, Gottschalk G, Deppenmeier U. Eur J Biochem 1999;259:218–224. [PubMed: 9914496]
71. Fontecilla-Camps JC, Frey M, Garcin E, Hatchikian C, Montet Y, Piras C, Vernede X, Volbeda A. Biochimie 1997;79:661–666. [PubMed: 9479448]
72. Wu LF, Chanal A, Rodrigue A. Arch Microbiol 2000;173:319–324. [PubMed: 10896209]
73. Suvorov AN, Flores AE, Ferrieri P. Infect Immun 1997;65:191–196. [PubMed: 8975911]
74. Buillard V, Cerutti L, Copley R, Courcelle E, Das U, Daugherty L, Dibley M, Finn R, Fleischmann W, Gough J, Haft D, Hulo N, Hunter S, Kahn D, Kanapin A, Kejariwal A, Labarga A, Langendijk-Genevaux PS, Lonsdale D, Lopez R, Letunic I, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Nikolskaya AN, Orchard S, Orengo C, Petryszak R, Selengut JD, Sigrist CJA, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C. Nucleic Acids Research 2007;35:D224–D228. [PubMed: 17202162]
75. Koenig, H. Archaea A Laboratory Manual. Robb, FT.; Place, AR.; Sowers, KR.; Schreier, HJ.; DasSarma, S.; Fleischmann, EM., editors. Cold Spring Harbor Laboratory Press; Cold Spring Harbor: 1995. p. 315–328.
76. Holmberg A, Blomstergren A, Nord O, Lukacs M, Lundeberg J, Uhlén M. Electrophoresis 2005;26:501–510. [PubMed: 15690449]
77. Lee BS, Krisnanchettiar S, Salman Lateef S, Gupta S. Rapid Comm Mass Spectrom 2005;19:886–892.
78. Bernstein EM, Quick MW. J Biol Chem 1999;274:889–895. [PubMed: 9873028]
79. Zhao Y, Zhang W, Kho Y, Zhao Y. Analytical Chemistry 2004;76:1817–1823. [PubMed: 15053638]
80. Zhao Y, Zhang W, White MA, Zhao Y. Anal Chem 2003;75:3751–3757. [PubMed: 14572040]
81. Khare T, Giometti CS. BioTechniques 2006;40:584–588. [PubMed: 16708758]
82. Sostaric E, Georgiou AS, Wong CH, Watson PJ, Holt WV, Fazeli A. J Prot Res 2006;5:3029–3037.
83. Hofmann K, Wood SW, Brinton CC, Montibeller JA, Finn FM. Proc Natl Acad Sci USA 1980;77:4666–4668. [PubMed: 6933515]
84. Orr GA. J Biol Chem 1981;256:761–766. [PubMed: 6161128]
85. Gauthier DJ, Gibbs BF, Rabah N, Lazure C. Proteomics 2004;4:3783–3790. [PubMed: 15529337]
86. Scherurer SB, Rybak JN, Roesli C, Brunisholz RA, Potthast F, Schlapbach R, Neri D, Elia G. Proteomics 2005;5:2718–2728. [PubMed: 15986331]
87. Bayer EA, Ehrligh-Rogozinski S, Wilchek M. Electrophoresis 1996;17:1319–1324. [PubMed: 8874057]
88. Eichler J. Microbiology 2003;149:3347–3351. [PubMed: 14663068]

Figure 1. Experimental approach used to identify surface-exposed proteins in *M. acetivorans* and *M. mazei*. (1) Intact cells. (2) Surface amines tagged with biotin. (3) Cells separated from secreted proteins and lysis products. (4) Tagged proteins bound to streptavidin-conjugated beads. (5) SDS-PAGE separation and visualization. (6) HPLC-MS/MS analysis.

Figure 2. Fractionation of affinity purified proteins from *M. acetivorans* and *M. maei*. SYPRO Ruby-stained SDS-PAGE gels (Left), and streptavidin-HRP blots (Right) from (A) *M. acetivorans*, and (B) *M. maei*. Lanes as follows: (M) Biotinylated protein markers; (1) Whole cell lysate; (2) Whole cell lysate after desalting; (3) Unbound fraction; (4) Wash 1; (5) Final wash; (6) Eluted proteins. Gel bands analyzed by LC-MS/MS are numbered.

Figure 3. S-layer protein NH₂-termini from *M. acetivorans* MA0829 and *M. mazei* MM1976 are observed in acetylated and non-acetylated forms. MS/MS of 2+ ion from N-terminal tryptic peptide of (A) MA0829 at m/z 422, (B) MA0829 at m/z 443, (C) MM1976 at m/z 408, and (D) MM1976 at m/z 429.

Figure 4. Streptavidin-affinity purification fractions visualized with Pro-Q Emerald glycoprotein stain from A) *M. acetivorans* and B) *M. maei* cultures. (M) Biotinylated protein markers; (G) Candy Cane™ glycoprotein markers; (1) Whole cell lysate; (2) Unbound lysate; (3) Wash 1; (4) Final wash; (5) Eluted proteins. *White Box* indicates location of MA0829 (A) and MM1976 (B) bands.

Figure 5. Streptavidin-bound *M. acetivorans* proteins treated with PNGase F, eluted, and resolved by SDS-PAGE. A) SYPRO-Ruby staining reveals total proteins, and B) Near-western blot probed with streptavidin-HRP reveals biotin-tagged proteins. (M) Biotinylated protein markers; (1) Protein eluate from beads untreated by PNGase F (negative control); (2) Protein eluate from beads treated with PNGase F. *Black Box*: MA0829 bands pre-treatment. *Upper Black Arrow*: MA0829 post-treatment. *Lower Black Arrow*: PNGase F.

Figure 6.

The presence of α -linked mannose and/or glucose in the oligosaccharides of *M. acetivorans* MA0829 and *M. mazei* MM1976. *M. acetivorans* and *M. mazei* whole cell lysates with biotinylated surface-exposed proteins were subjected to Con A affinity purification. Fractions resolved by SDS-PAGE and visualized by staining and near-western blotting for A) *M. acetivorans* and B) *M. mazei*. (G) Molecular weight standards; (M) Biotinylated markers; (1) whole cell lysate; (2) flow through; (3) first wash; (4) final wash; (5) concentrated eluate (monosaccharide released); (6) eluate released by boiling Con A beads; (7) Ultrafiltrate from 30 kDa membrane. Arrows: MA0829 and MM1976 glycoforms. Dots: Con A monomer (26 kDa) and its hydrolytic cleavage products. Numbered bands were analyzed by tandem MS (Tables 3 and 4).

Proteins Identified by LC-MS/MS from Gel Bands of Purified Biotinylated Cell Surface Proteins of *M. acetivorans*^a

Table 1

Bands	MA no.	Protein ID	Prot. Score	MW calc. (kDa)	Peptide no.	Signal Peptide	TMH no.
Band 1	MA4643	hypothetical protein	84	159.0	3	yes	1
	MA0068	hypothetical protein	31	126.9	3	yes	1
Band 2	MA4643	hypothetical protein	93	159.0	4	yes	1
Band 3	MA0829	hypothetical protein	166	74.3	7	yes	1
Band 4	MA0829	hypothetical protein	207	74.3	6	yes	1
Band 5	MA0829	hypothetical protein	208	74.3	6	yes	1
Band 6	MA0840	diptide ABC transporter, solute-binding protein	65	62.9	2	no*	1
Band 7	MA0674	pyruvate carboxylase, subunit B	75	63.2	4	no	0
Band 8	MA0674	pyruvate carboxylase, subunit B	182	63.2	11	no	0
	MA0086	thermosome, subunit alpha	120	58.9	5	yes	0
	MA3898	nitrogenase, subunit alpha	78	59.6	7	no	0
	MA1142	F420-nonreducing hydrogenase	56	64.6	3	no	0
Band 9	MA0455	Methanol-5-hydroxybenzimidazolylcobamide co-methyltransferase (methanol:mtaC methyltransferase)	84	50.3	5	no	0
	MA4216	glutamine synthetase	35	50.7	2	no	0
Band 10	No ID with multiple peptides						
Band 11	MA3895	nitrogenase (iron protein)	134	29.2	8	yes	0
	MA0456	methanol-5-hydroxybenzimidazolylcobamideco-methyltransferase (methanol-specific corrinoid-binding protein)	55	27.7	2	no	0
Band 12	MA3895	nitrogenase (iron protein)	53	29.2	3	no	0
Band 13	None	streptavidin background only					

^aNCBI protein accession numbers, the protein names, protein scores, theoretical masses, number of peptides identified, predicted signal peptide, and number of predicted transmembrane helices (TMH) are listed.

* Signal P algorithm predicted as present.

Table 2
Proteins Identified by LC-MS/MS from Gel Bands of Purified Biotinylated Cell Surface Proteins of *M. mazel*^a

Bands	MA no.	Protein ID	Prot. Score	MW calc. (kDa)	Peptide no.	Signal Peptide	TMH no.
Band 1	MM1976	hypothetical protein	249	74.0	9	yes	1
	MM1770	pyruvate phosphate dikinase	157	97.0	7	no	0
Band 2	MM1976	hypothetical protein	156	74.0	4	yes	1
Band 3	MM1976	hypothetical protein	195	74.0	7	yes	1
Band 4	MM1770	pyruvate phosphate dikinase	38	97.0	3	no	0
	MM1827	pyruvate carboxylase, subunit B	136	63.3	4	no	0
Band 5	MM0001	dipeptide ABC transporter, binding protein	56	58.1	2	no	0
	MM1096	thermosome, gamma subunit	221	58.2	8	no	0
	MM1827	pyruvate carboxylase, subunit B	156	63.3	5	no	0
	MM1379	thermosome, alpha subunit	135	58.8	4	yes	0
Band 6	MM2170	F420-nonreducing hydrogenase II, large subunit	74	64.1	2	no	0
	MM0780	A ₁ A ₀ H ⁺ ATPase subunit A	55	63.8	2	no	0
	MM0072	thermosome, beta subunit	40	60.8	3	yes	0
	MM3291	exinuclease ABC, subunit A	33	110.8	2	no	0
	MM3188	glutamine synthetase	203	56.9	5	no	0
	MM0686	acetyl-CoA decarboxylase/synthase complex subunit beta	114	52.2	3	no	0
Band 7	MM2893	hypothetical protein	75	49.4	2	yes	1
	MM1647	methanol:corrinoid methyltransferase MtaB	158	50.3	7	no	0
Band 8	MM1074	methanol:corrinoid methyltransferase	74	50.3	4	no	0
	MM1013	50S ribosomal protein L10	30	37.3	2	no	0
Band 9	MM2367	hexulose-6-phosphate isomerase	66	24.1	2	no	0
Band 10	None						

^a Genome accession numbers, the protein names, protein scores, theoretical masses, number of peptides identified, predicted signal peptide, and number of predicted transmembrane helices (TMH) are listed.

Table 3
Concanavalin A Binding Proteins Identified in *M. acetivorans*^a

Bands	MA no.	Protein ID	Prot. Score	MW calc. (kDa)	Peptide no.	Signal Peptide	TMH no.
Band 1	MA3471	hypothetical protein (multi-domain)	46	442.3	2	no	0
Band 2	MA3556	predicted protein	106	94.6	3	yes	1
	MA0829	hypothetical protein	102	74.3	5	yes	1
Band 3	MA1478	heat shock protein 70	122	66.2	8	no	0
Band 4	No ID with multiple peptides						
Band 5	MA3899	nitrogenase, beta subunit <i>nifK</i>	49	49.4	3	yes	0
Band 6	MA4547	methyl coenzyme M reductase, subunit gamma	121	27.6	8	no	0

^a Genome accession numbers, the protein names, protein scores, theoretical masses, number of peptides identified, predicted signal peptide, and number of predicted transmembrane helices (TMH) are listed.

Table 4
Concanavalin A Binding Proteins Identified in *M. macleod*^a

Bands	MA no.	Protein ID	Prot. Score	MW calc. (kDa)	Peptide no.	Signal Peptide	TMH no.
Band 1	MM1364	hypothetical protein	356	126.1	10	no [*]	2
	MM1976	hypothetical protein	76	74.0	2	yes	1
Band 2	MM1976	hypothetical protein	204	74.0	6	yes	1
	MM0467	hypothetical protein	142	94.6	4	yes	0
Band 3	MM1976	hypothetical protein	910	74.0	44	yes	1
	MM1329	methyl-accepting chemotaxis protein	74	76.6	2	no [*]	2
Band 4	MM0002	dipeptide ABC transporter, binding protein	214	58.7	6	no	0
	MM1976	hypothetical protein	171	74.0	6	yes	1
Band 5	MM1999	hypothetical protein	76	63.2	2	yes	0
	MM2567	ABC transporter, periplasmic binding protein	85	58.4	5	yes	0
Band 6	MM1976	hypothetical protein	81	74.0	2	yes	1
	MM1976	hypothetical protein	71	74.0	2	yes	1
Band 7	MM2567	ABC transporter, periplasmic binding protein	54	58.4	2	yes	0
	MM1647	methanol:corrinoid methyltransferase MtaB	85	50.3	2	no	0

^a Genome accession numbers, the protein names, protein scores, theoretical masses, number of peptides identified, predicted signal peptide, and number of predicted transmembrane helices (TMH) are listed.

^{*} Signal P algorithm predicted as present.