The Protein Journal, Vol. 23, No. 1, January 2004 (© 2004)

Complete Amino Acid Sequence and Location of Omp-28, an Important Immunogenic Protein from *Salmonella enterica* serovar typhi

Ana G. C. Neves-Ferreira,[1] Carlos M. de Andrade,[2] Marcos A. Vannier-Santos,[3] Jonas Perales,[1] Hilton J. Nascimento,[4] and José G. da Silva Junior[4,5,6]

Received August 21, 2003

Omp-28 isolated from *Salmonella enterica* serovar typhi presented a subunit molecular mass of 9,632 Da by MALDI-TOF MS. It was denatured, *S*-alkylated, and 1) directly submitted to Edman sequencing, 2) cleaved with CNBr, and 3) hydrolyzed either with endoproteinase Glu-C or Asp-N. The major CNBr peptide containing the *C*-terminal portion of Omp-28 was isolated by tricine-SDS-PAGE and electroblotted whereas Omp-28 enzymatic peptides were isolated by C18-RP-HPLC. All peptides were sequenced. This approach allowed the elucidation of the complete primary structure of Omp-28. Its amino acid sequence is identical to that deduced from part of the DNA of the "putative periplasmic transport protein" of either *S. enterica* serovar typhimurium and a multiple drug resistant *S. enterica* serovar typhi. Omp-28 homologous protein sequences were also deduced from *Escherichia coli* and *Yersinia pestis* genomic DNA. All proteins had their secondary structures predicted. Immunogold cytochemistry indicated that Omp-28 is found on the bacterium outer membrane.

KEY WORDS: Amino acid sequence; Omp-28; outer membrane protein; *Salmonella enterica*; serovar typhi.

1. INTRODUCTION

Salmonella enterica serovar typhi (former *Salmonella typhi*) is an enteric Gram-negative bacteria responsible for typhoid fever and other related clinical diseases that are distributed worldwide mainly in regions with limited sanitary conditions.

Gram-negative bacteria have an outer membrane formed by lipopolysaccharide and proteins. Despite the apparent low protein complexity of this bacterial coat, these molecules play central roles in the transport across membrane, infection, and immunity. Among these proteins, the porins Omp A, Omp C, Omp F, Lam B, and PhoE (Jeannin *et al.*, 2002; Jeanteur *et al.*, 1994; Schulz, 1996), the adhesion proteins (Krogfelt, 1991), the phospholipases A (Merck *et al.*, 1997), and the lipoproteins TraT and PAL (Achtman *et al.*, 1977; Mizuno, 1981) are well characterized.

[1] Departamento de Fisiologia e Farmacodinâmica, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil.
[2] Departamento de Ciências Biológicas, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
[3] Laboratório de Microscopia Eletrônica, Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz, Bahia, Brazil
[4] Departamento de Bioquímica, Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[5] Departamento de Química, Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, Brazil
[6] To whom correspondence should be addressed. E-mail: godinho@iq.ufrj.br

[7] Abbreviations: BLAST, basic local alignment search tool; CNBr, cyanogen bromide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; IEF-PAGE, isoelectric focusing in polyacrylamide gel electrophoresis; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; PBS, phosphate buffered saline; PSI-PRED, position specific iterated prediction; PVDF, polyvinylidene difluoride; RCA-Omp-28, reduced and carboxyamidated Omp-28; RPE-Omp-28, reduced and *S*-pyridylethylated Omp-28; RP-HPLC, reverse-phase high-performance liquid chromatography; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

1572-3887/04/0100-0071/0

We have previously isolated and partially characterized a *S. enterica* serovar typhi outer membrane protein named Omp-28 not reported in Gram-negative bacteria thus far (Andrade *et al.*, 1998). Omp-28 is formed by three identical subunits not linked by disulfide bonds, presents a molecular mass of 28 kDa, and an isoelectric point of 4.6. Twenty-four *N*-terminal amino acid residues of native Omp-28 were determined. Significant levels of antibodies against Omp-28 were detected in sera from typhoid fever convalescent patients. Mice antibody raised against Omp-28 showed anti-*Salmonella* activity *in vitro*. These results indicate the immunogenic importance of Omp-28 isolated from *Salmonella* and reinforce the relevance of further investigation on this protein, which is potentially useful in the protection against the disease caused by this bacteria.

This paper reports the complete Omp-28 amino acid sequence from *S. enterica* serovar typhi through a peptide approach as well as its cellular location. It confirms the deduced amino acid sequence from part of the bacterial genomic DNA related to a "putative periplasmic transport protein" (Parkhill *et al.*, 2001) and unambiguously demonstrates the precursor protein enzymatic cleavage site. The secondary structures of Omp-28 and its main homologous proteins from *Escherichia coli* and *Yersinia pestis* are predicted and compared.

2. MATERIALS AND METHODS

2.1. Isolation of Omp-28

The bacterium *Salmonella enterica* serovar typhi (S2154/IOC) was grown as previously described (Andrade *et al.*, 1998). The outer membrane protein extract was desalted and submitted to size exclusion in Sepharose CL-6B and to fast anion exchange chromatography in a Mono Q column (Amersham Biosciences, Uppsala, Sweden) (Andrade *et al.*, 1998). Omp-28 (0.5 mg) was repurified on a Vydac C4 column (The Separations Group, Hesperia, CA, USA) (0.46 cm $\times$ 25 cm) using a flow rate of 1 ml/min before structural studies. The solvent system used was 0.1% TFA in H_2O (solvent A) and 0.08% TFA in 80% acetonitrile (solvent B). The elution started with 100% solvent A for 10 min, followed by a linear gradient up to 90% B over 45 min.

2.2. Molecular Mass

Omp-28 subunit molecular mass was determined by MALDI-TOF MS[7] on a Voyager DE-PRO instrument (Perseptive Biosystems, Foster City, CA, USA). The matrix used was 3,5-dimethoxy-4-hydroxycinnamic acid. The molecular mass of Omp-28 was also determined by SDS-PAGE in a high-density pre-cast gel using the PhastSystem apparatus (Pharmacia Biotech, Uppsala, Sweden) as recommended by the manufacturer. Molecular weight protein markers (MW 14,400–67,000 and 2,500–17,000) (Sigma-Aldrich Co., St. Louis, MO, USA) were also used.

2.3. Reduction and S-Alkylation

2.3.1. Vinylpyridine

Omp-28 (35 μg) was dissolved in 0.5 *M* Tris-HCl buffer, pH 8.3, containing 6 *M* guanidine HCl and 0.001 *M* EDTA, under nitrogen, and incubated for 2 hr at 50°C. A 50-fold molar excess of DTT (assuming two cysteine residues in Omp-28) (Andrade *et al.*, 1998) was added to the reaction mixture. Following 30 min of incubation at 50°C, 4-vinylpyridine was added, and the alkylation mixture was allowed to incubate overnight at room temperature in the dark. Excess reagents were either washed with 0.1% TFA through the ProSorb cartridge (Perkin Elmer, Wellesley, MA, USA) for direct *N*-terminal sequencing or removed by size exclusion chromatography using a HiTrap desalting column (Amersham Biosciences) before the treatment with CNBr or endoproteinase Glu-C.

2.3.2. Iodoacetamide

Omp-28 (50 μg) was denatured with guanidine HCl and reduced with DTT as already described. Alkylation of free thiol groups was performed after incubation with iodoacetamide for 2 hr at room temperature in the dark. Excess reagents were removed by RP-HPLC on a Vydac C18 (The Separations Group) (0.21 cm $\times$ 15 cm) column using a flow rate of 0.2 ml/min. The absorbance was monitored at 215 nm. The solvent system used was 0.1% TFA in H_2O (solvent A) and 0.08% TFA in acetonitrile (solvent B). The gradient was 5% solvent B for 20 min, 80% B at 40 min, and 80% B at 60 min. Protein peaks were collected and speed vac dried for further digestion with Asp-N endoproteinase.

2.4. CNBr Cleavage

Pyridylethylated Omp-28 (5 μg) was cleaved with CNBr (60 μg) in 70% aqueous formic acid, overnight,

at 20°C (Allen, 1989). The mixture was diluted with 10 volumes of water and extensively freeze-dried. This last procedure was repeated twice. The CNBr peptides were isolated by tricine-SDS-PAGE (Schägger and von Jagow, 1987), transferred to a polyvinylidene difluoride (PVDF) membrane, and directly *N*-terminally sequenced.

2.5. Endoproteinase Glu-C Digestion

Pyridylethylated Omp-28 (30 μg) was dissolved in 0.025 *M* ammonium carbonate, pH 7.8, and digested with endoproteinase Glu-C at an enzyme to substrate ratio of 1:33 (w:w) at 30°C for 8 hr. Endoproteinase Glu-C peptides were isolated by RP-HPLC on a Vydac C18 (0.46 cm × 25 cm) column using a flow rate of 1 ml/min. The absorbance was monitored at 215 nm. The linear gradient was performed by the addition of 80% acetonitrile in 0.08% TFA (solvent B) to 0.1% TFA (solvent A). Protein peaks were collected, speed vac concentrated to 20 μl and Edman sequenced as described below.

2.6. Endoproteinase Asp-N Digestion

Omp-28 alkylated with iodoacetamide (50 μg) was dissolved in 0.05 *M* sodium phosphate pH 8.0 and digested with endoproteinase Asp-N at an enzyme to substrate ratio of 1:70 (w/w) at 35°C for 18 hr. The peptides obtained by Asp-N digestion were isolated by RP-HPLC on a Vydac C18 column (0.21 cm × 15 cm) as described in Sec. 2.3.2. The gradient program began with 5% solvent B for 5 min and was then ramped to 45% solvent B at 65 min and to 70% solvent B at 75 min. Protein peaks were collected, speed vac concentrated to 20 μl, and Edman sequenced.

2.7. Primary and Secondary Structures Analyses

N-terminal Edman sequencing of pyridylethylated Omp-28 and its peptides isolated by RP-HPLC or by tricine-SDS-PAGE was performed on a Shimadzu PSQ-23A protein sequencer (Shimadzu, Kyoto, Japan) according to the manufacturer's instructions. The complete Omp-28 amino acid sequence was used to scan the GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) and the SwissProt (http://bo.expasy.org/sprot/) databases for homologous sequences with the BLAST program (Altschul *et al.*, 1997). The PSIPRED method (Jones, 1999; McGuffin *et al.*, 2000) was used to predict Omp-28 and its putative homologous proteins (HdeB from *E. coli* and *Y. pestis*) secondary structures.

2.8. Immunogold Procedure

It was performed after Bendayan *et al.* (1987), modified by Vannier-Santos *et al.* (1996), as follows: after growth, bacterial cells were harvested, centrifuged, washed in sterile 0.15 *M* NaCl solution, and then fixed in 0.1 *M* sodium cacodilate buffer, pH 7.4, containing 4% formaldehyde and 1% glutaraldehyde, overnight at 4°C. After washing in the same buffer, free aldehyde groups were blocked by a solution of 0.1 *M* glycine in PBS, pH 7.2, for 60 min at room temperature. Mice polyclonal antiserum raised against an Omp-28 fraction was used as primary antibody. The serum was diluted in PBS with 5% BSA and 0.01% Tween 20, pH 8.0, and the reaction was performed in suspension for 1 hr with mild stirring at room temperature. As secondary antibody, 10 nm particles colloidal gold-labeled goat anti-mouse, diluted in the same buffer, was used for 2 hr. After this step, the cells were washed as described above. The cells were dehydrated in an acetone series and embedded in Polybed (Polyscience Inc., Warrington, PA). Polymerization was carried out for 48 hr at 60°C. After polymerization, ultrathin sections were contrasted with uranyl acetate, lead citrate, and observed in a CEM/Zeiss 900 electron microscope (Zeiss, Oberkochen, Germany).

3. RESULTS AND DISCUSSION

The use of the Vydac C4 RP-HPLC for the definitive polishing of Omp-28 was necessary for the elimination of slight protein contaminants present in the fraction obtained after rechromatography in Mono Q column (Silva Junior *et al.*, 1997). These contaminants were observed either directly by reverse phase fractionation [Fig. 1(A)] or indirectly through Pico-Tag amino acid analysis. In the last case, 0.7 mol of arginine/mol of Mono Q-Omp-28 preparation was estimated (results not shown) although, according to the sequential analysis, this protein is devoid of arginine.

Omp-28, a protein isolated from *Salmonella enterica* serovar typhi (Bhatnagar et al., 1982; Andrade *et al.*, 1998) had its complete amino acid sequence elucidated through a peptide approach. The peptides

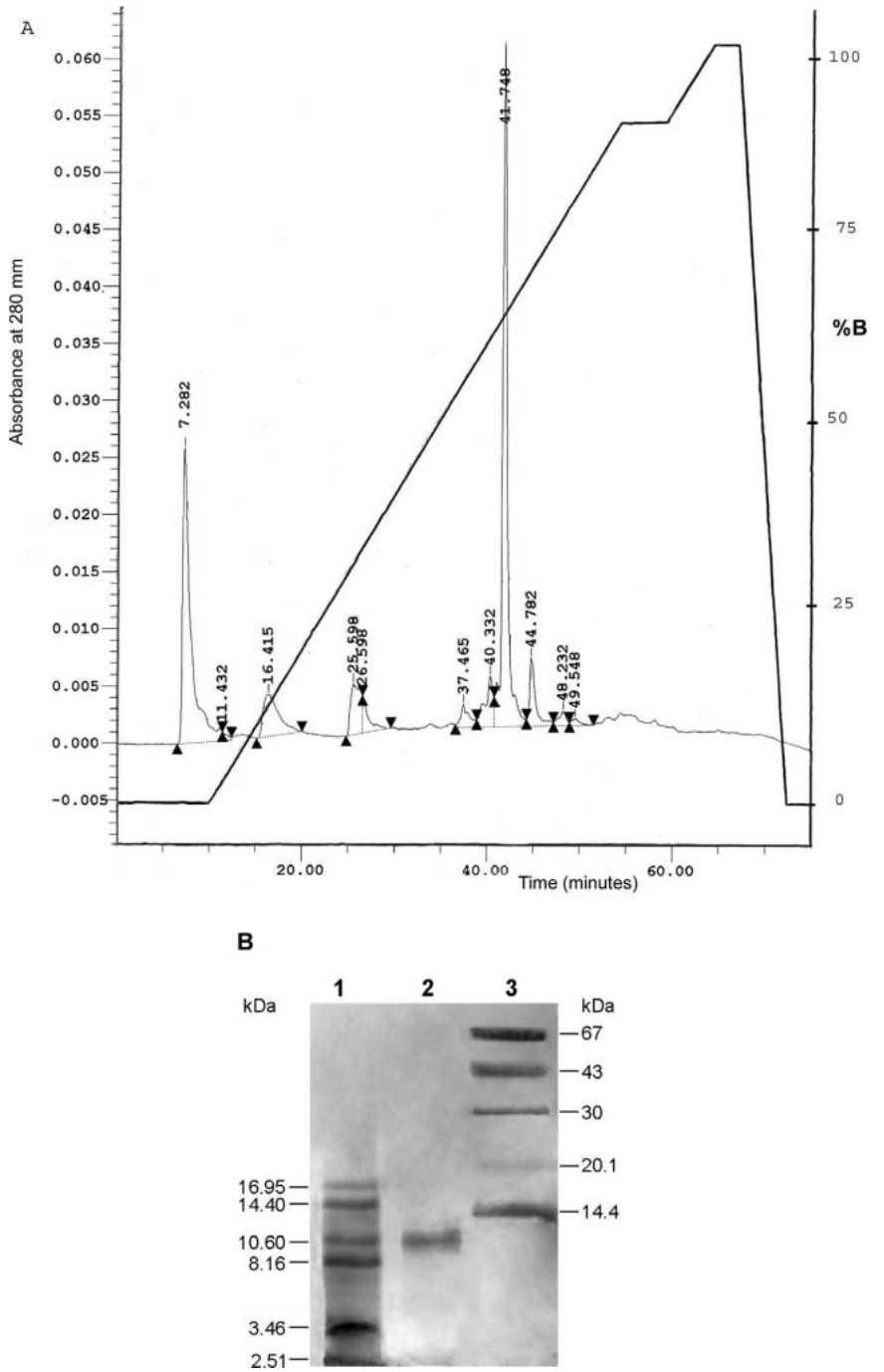

Fig. 1. (A) Polishing of Omp-28 for structural studies. Omp-28 isolated by Mono Q column was dialyzed against milli-Q water and dried in a speed vac. One-half milligram of a Mono Q rechromatographed fraction was dissolved in TFA 0.1% and eluted in a Vydac C-4 column (0.46 cm × 25 cm) using a linear acetonitrile gradient in TFA solution. The acetonitrile gradient was done by the addition of 80% acetonitrile in 0.08% TFA (solvent B) to 0.1% TFA (solvent A). The chromatographic profile was determined by absorbance readings at 220 nm (not shown) and 280 nm. The main fraction (retention time 41.748 min) was isolated and pooled. (B) High-density SDS-PAGE of Omp-28 obtained after C4 RP-HPLC (2 μg). Lane 1, low-range molecular mass markers; Lane 2, Omp-28; Lane 3, high-range molecular mass markers. Phast System pre-cast gel was silver stained as recommended by the manufacturers.

produced by chemical (cyanogen bromide) or enzymatic (Glu-C- and Asp-N- endoproteinases) cleavages of the denatured protein were purified by tricine-SDS-PAGE (*C*-terminal CNBr peptide) or by reverse-phase chromatographies (enzymatic peptides). All peptides had their amino acid sequences partially or totally determined by Edman's chemistry by the use of an automatic sequenator. Associating these data to that obtained in the partial *N*-terminal amino acid sequencing of the either native (Andrade *et al.*, 1998) or reduced and *S*-alkylated protein, it was possible to elucidate the complete primary structure of Omp-28 (Fig. 2). The molecular weights of Omp-28 determined by high density SDS-PAGE [Fig. 1(B)] and by MALDI-TOF MS were 9,905 and 9,632, respectively, which are in accordance with the presence of 85 residues in the protein Omp-28. These experimentally determined values are in close agreement with the theoretical molecular weight of 9,513 calculated from the amino acid sequence.

For the Omp-28 primary structure elucidation, it was necessary initially to associate the sequencing results obtained from RPE-Omp-28 (41 *N*-terminal residues with two gaps) with those determined for the CNBr major peptide (A^{28}. . .T^{52}, no gaps). The latter result filled the two gaps (W^{33} and E^{39}) observed up to Omp-28 52nd amino acid residue. Two RPE-Omp-28 endoproteinase Glu-C peptides comprised between the 63rd and 78th amino acid residues had their sequences totally elucidated. The link between the *N*- and *C*-terminal portions of Omp-28 as well as the overlapping

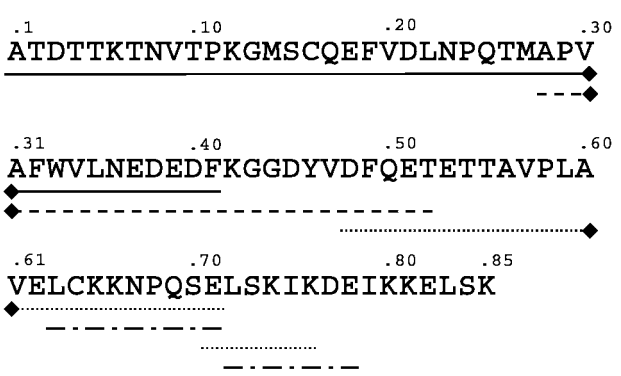

Fig. 2. Summary of the complete amino acid sequence of Omp-28. A minimum set of overlapping peptides was used to determine the sequence. Direct *N*-terminal sequence of RPE-Omp-28 gave the first 41 residues (—). The peptides obtained from different digestions are as follows: CNBr fragment (- - - -), endoproteinase Asp-N peptides (.....), endoproteinase Glu-C peptides (-------). Black lozenges indicate that the sequence continues in the next line.

among the Glu-C peptides present in the *C*-terminal portion was done through the RCA-Omp-28 endoproteinase Asp-N peptides. Fortunately, only one endoproteinase Asp-N peptide did not show an aspartic acid as its *N*-terminal residue, being undoubtedly found in this position one glutamic residue (^{71}E-L-S-K-I-K^{76}). This result was fundamental to complete Omp-28 *C*-terminal sequence. A summary of the Omp-28 complete amino acid sequence is shown in Fig. 2, along with all peptides and overlaps necessary for the proof of the sequence.

The complete Omp-28 amino acid sequence here described was compared with other protein sequences using protein and DNA sequence data banks. Total homology was observed with two hypothetical proteins present in the *Salmonella enterica* serovar typhi drug resistant strain CT18 (Parkhill *et al.*, 2001) and in the *Salmonella enterica* serovar typhimurium LT2 (McClelland *et al.*, 2001). According to the sequence alignment, Omp-28 also presented a considerable homology with two other Gram-negative bacterial proteins from *Y. pestis* (HdeB; 50% identity in 84 residues) and from *E. coli* (43.5% identity in 85 residues) (Gouet *et al.*, 1999)(Fig. 3). In these cases, it is possible to observe regions of the protein structures such as, for example, ^{38}M(S/T)C(Q/K)EF(V/I/L)DLNP^{48}, ^{53}PV (A/V) (F/W/Y)W(V/M)L^{59}, and ^{84}(A/V)(V/I)E(L/Y/V) CKK^{90}, where the amino acid sequences are considerably preserved (identity and positivity) (Fig. 3, clear boxes).

Using the PSIPRED method for Omp-28 secondary structure prediction (McGuffin *et al.*, 2000), which is based on the PSIBLAST program, it is noteworthy the high content of random coil structure (48.24%) followed by alpha helix (38.82%) and extend strand (12.94%), not only observed for the Omp-28 protein as well as for its more homologous proteins from *E. coli* (49.41; 37.65 and 12.94%, respectively) and *Y. pestis* (45.88; 41.18 and 12.94%, respectively). These three proteins share very similar secondary structures according to the prediction method used. The elucidation of the amino acid sequence of Omp-28 and its secondary structure allowed us to identify peptide epitopes very well conserved, as demonstrated by comparative analysis with the homologous *E. coli* and *Y. pestis* putative periplasmic proteins (Fig. 3, clear boxes). The possibility of these epitopes to be present in other species of Enterobacteriaceae will be investigated. These epitope peptides are also being evaluated as target molecules better than the entire Omp-28 in immunization purposes.

Fig. 3. Sequential homology of Omp-28. The amino acid sequence of Omp-28 was aligned with other significant homologous proteins using the computer software Multialin (http://prodes.toulouse.inra.fr/multalin/multalin.html). Identical residues in all sequences are gray boxed. Regions of considerable homology among Omp-28 and the HdeB proteins from *E. coli* and *Y. pestis* are clear boxed.

The location of Omp-28 is doubtful. The recent method used for Omp-28 isolation (Andrade *et al.*, 1998) began with the extraction of *S. typhi* surface proteins (Bhatnagar *et al.*, 1982). However, other proteins located at the periplasm may also be extracted using this protocol. Based on the perisplasmic location of *E. coli* HdeB (Yoshida *et al.*, 1993), it has been proposed the same location for its very homologous Gram-negative bacterial proteins (*S. typhi* Omp-28 and *Y. pestis* HdeB). Our evidence is due to the observation of recent immunogold microscopy where it was possible to observe a few number of black points related to positive reaction in the outside of the outer membrane of *Salmonella* cells (Fig. 4). Therefore, the result of the immunogold assay suggests that Omp-28 is located at the outer envelope of *Salmonella enterica* serovar typhi. However, the presence of Omp-28 in the periplasmic site cannot be ruled out. In this case, it could be transported to the bacterial surface.

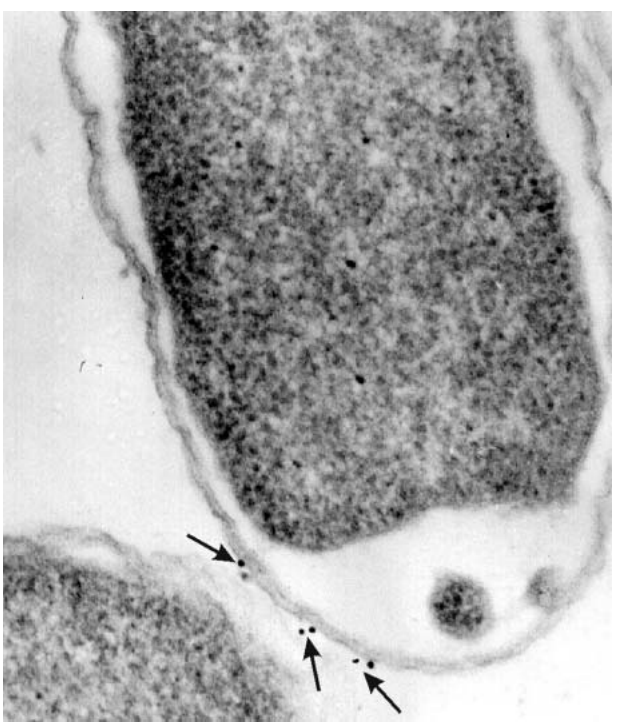

Fig. 4. Immunogold location of *Salmonella enterica* serovar typhi Omp-28 by electron microscopy. Bacterial cells were treated with a polyclonal antibody raised against an Omp-28 rich fraction that was revealed with a secondary antibody colloidal gold conjugate. Gold particles (arrows) represent positive staining for the protein Omp-28 at the bacterial outer envelope surface.

ACKNOWLEDGMENT

This study was supported by Brazilian grants from CNPq, FAPERJ, and PAPES-Fiocruz. We thank Dr. Alex Chapeaurouge from Departamento de Fisiologia e Farmacodinâmica, Instituto Oswaldo Cruz, Brazil, for the MALDI-TOF MS analysis.

REFERENCES

Achtman, M., Kennedy, N. and Skurray, R. (1977). *Proc. Natl. Acad. Sci. U .S. A.* **74:** 5104–5108.

Allen, G., (1989). Specific cleavage of the protein. In: Burdon, R. H. & van Knippenberg, P. H. (eds.), *Sequencing of Proteins and Peptides*, Elsevier Science Publishers B. V., Amsterdam, pp. 73–104.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., *et al.* (1997). *Nucleic Acids Res.* **25:** 3389–3402.

Andrade, C. M., Ferreira, A. G., da Silva, J. D., Nascimento, H. J., and Silva Junior, J. G. (1998). *Microbiol. Immunol.* **42:** 521–526.

Bendayan, M., Nanci, A. and Kan, F. W. (1987). *J. Histochem. Cytochem.* **35:** 983–996.

Bhatnagar, N., Müller, W., and Schlecht, S. (1982). *Zentralbl. Bakteriol. Mikrobiol. Hyg. [A]* **253:** 88–101.

Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999). *Bioinformatics* **15:** 305–308.

Jeannin, P., Magistrelli, G., Goetsch, L., Haeuw, J. F., Thieblemont, N., Bonnefoy, J. Y., and Delneste, Y. (2002). *Vaccine* **20**(Suppl 4): A23–27.

Jeanteur, D., Lakey, J. H., and Pattus, F., (1994).The porin superfamily: diversity and common features cell wall In: *Bacterial Cell Wall* (Ghuysen, J. M. & Hakenbeck, R., eds.), Elsevier, Amsterdam.

Jones, D. T. (1999). *J. Mol. Biol.* **292:** 195–202.

Krogfelt, K. A. (1991). *Rev. Infect. Dis.* **13:** 721–735.

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwilko, S., Ali, J. Daute, M. Du, F. (2001). *Nature* **413:** 852–856.

McGuffin, L. J., Bryson, K., and Jones, D. T. (2000). *Bioinformatics* **16:** 404–405.

Merck, K. B., de Cock, H., Verheij, H. M. and Tommassen, J. (1997). *J. Bacteriol.* **179:** 3443–3450.

Mizuno, T. (1981). *J. Biochem. (Tokyo)* **89:** 1039–1049.

Parkhill, J., Dougan, G., James, K. D., Thomson, N. R., Pickard, D., Wain, J., Churcher, C., Mungall, K. L., Bentley, S. D., Holden, M. T. (2001). *Nature* **413:** 848–852.

Schägger, H., and von Jagow, G., (1987). *Anal. Biochem.* **166:** 368–379.

Schulz, G. E. (1996). *Curr. Opin. Struct. Biol.* **6:** 485–490.

Silva Junior, J. G., Ferreira, A. G. P., Nascimento, H. J., Candido, M. S., and Andrade, C. M. (1997). *FASEB J.* **11:** A1149.

Vannier-Santos, M. A., de Souza, S. J., Brentani, R. R., and de Souza, W. (1996). *Biochem. Biophys. Res. Commun.* **221:** 837–842.

Yoshida, T., Ueguchi, C., and Mizuno, T. (1993). *J. Bacteriol.* **175:** 7747–7748.