Extremophiles (2012) 16:437–446
DOI 10.1007/s00792-012-0443-2

ORIGINAL PAPER

Nmag_2608, an extracellular ubiquitin-like domain-containing protein from the haloalkaliphilic archaeon *Natrialba magadii*

María Victoria Ordóñez · Débora Nercessian · Rubén Danilo Conde

Received: 4 January 2012 / Accepted: 22 March 2012 / Published online: 10 April 2012

Abstract Ubiquitin-like proteins (Ubls) and ubiquitin-like domain-containing proteins (Ulds) found in both eukaryotes and prokaryotes display an ubiquitin fold. We previously characterized a 124-amino acid polypeptide (P400) from the haloalkaliphilic archaeon *Natrialba magadii* having structural homology with ubiquitin family proteins. The reported *N. magadii*'s genome allowed the identification of the *Nmag_2608* gene for the protein containing P400, which belongs to specific orthologs of halophilic organisms. It was found that Nmag_2608 has an N-terminal signal peptide with a lipobox motif characteristic of bacterial lipoproteins. Also, it presents partial identity with the ubiquitin-like domain-containing proteins, soluble ligand binding β-grasp proteins. Western blots and heterologous expression tests in *E. coli* evidenced that Nmag_2608 is processed and secreted outside the cell, where it could perform its function. The analysis of *Nmag_2608* expression in *N. magadii*'s cells suggests a co-transcription with the adjoining *Nmag_2609* gene encoding a protein of the cyclase family. Also, the transcript level decreased in cells grown in low salinity and starved. To conclude, this work reports for the first time an extracellular archaeal protein with an ubiquitin-like domain.

Communicated by A. Oren.

Electronic supplementary material The online version of this article (doi:10.1007/s00792-012-0443-2) contains supplementary material, which is available to authorized users.

M. V. Ordóñez (✉) · D. Nercessian · R. D. Conde
Degradación de Proteínas, Instituto de Investigaciones Biológicas, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, CONICET, Funes 3250 CC 1245, 7600 Mar del Plata, Argentina
e-mail: mvordone@mdp.edu.ar

Keywords Halophilic archaea · *Natrialba magadii* · Ubiquitin-like domains · Extracellular proteins · Protein–substrate interaction

Introduction

Ubiquitin, a small protein highly conserved in eukaryotes, acts as post-translational modifier of proteins in many cellular events (Ciechanover 1994; Ciechanover and Iwai 2004; Hochstrasser 2009). For this, ubiquitin binds to proteins in a process catalyzed by three enzymes known as E1, E2, and E3 (Ciechanover 1994; Pickart 2001). This process known as ubiquitination controls the roles of target proteins by altering their association with other macromolecules (Kerscher et al. 2006). Ubiquitin displays a stable β-grasp fold composed of 5 β-sheets arranged in an antiparallel way, an α-helix between the second and third β-sheet, and a conserved di-glycine motif in its C-terminus (Burroughs et al. 2007a, b). Because the ubiquitin protein conjugation is crucial for cellular homeostasis, the question has arisen whether prokaryotes have an equivalent mechanism.

Modern DNA sequencing techniques made possible to know the genomes of several archaeal organisms in short periods of time (Soppa 2006). This allows to reveal the sequence of many archaeal proteins but not to define their physiological role, which is often not related to any bacterial or eukaryal counterpart. Thus, assignment of roles to many new identified proteins becomes a difficult task (Makarova and Koonin 2003). Fortunately, gene context and bioinformatic sequence analysis, including structural prediction programs, became strong tools in revealing the roles of genes from a sequenced archaeal genome. These approaches have been used by several groups in the search of ubiquitin-like proteins within prokaryotes.

Several eukaryotic and prokaryotic proteins having ubiquitin fold but low amino acid sequence identity with ubiquitin have been described (Hochstrasser 2009). They are classified as ubiquitin-like (Ubls) and ubiquitin-like domain-containing proteins (Ulds). As ubiquitin, eukaryotic Ubls act in conjugation pathways to modify proteins (Hochstrasser 2000; Dye and Schulman 2007). In prokaryotes, the best studied Ubls, ThiS and MoaD, display only 14 and 17 % sequence identity with ubiquitin, respectively. They are sulfur carriers involved in the biosynthesis of thiamine and molybdopterin cofactors, respectively (Marquet 2001; Rudolph et al. 2001; Wang et al. 2001). As eukaryotic Ubls, they require an activation step by an E1-like enzyme. Although this step will lead to the attachment of a sulfur atom to the C-terminus of a protein instead of binding with a protein, it resembles the ubiquitin conjugation pathways (Kerscher et al. 2006).

Archaeal genome survey has led to identify several genes encoding proteins with a β-grasp fold as well as E1- and E2-like proteins in these organisms (Iyer et al. 2006). So, looking for the presence in archaea of an ubiquitination-like system, two ubiquitin-like small archaeal modifier proteins (SAMP1 and 2) have been recently described (Humbard et al. 2010). The authors propose that SAMP1 and 2 are conjugated to target proteins and that this process is related to the proteasome 20S from the halophilic archaeon *Haloferax volcanii*. Also, these proteins display a similar structure as Ubls as well as the di-glycine motif in the C-terminus (Jeong et al. 2011; Ranjan et al. 2011).

The ubiquitin fold has also been found in larger multidomain proteins (Hartmann-Petersen and Gordon 2004). These Ulds have been found performing several roles in eukaryotes (Kiel and Serrano 2006; Grabbe and Dikic 2009). An Uld is defined as a 45- to 80-amino acid region that resembles the folding and, occasionally, the amino acid sequence of ubiquitin. It is found in one or more repeats inside either the N-terminal or the C-terminal region of proteins, usually lacking the di-glycine motif typical of ubiquitin and Ubls (Burroughs et al. 2007a; Grabbe and Dikic 2009). As a result, they do not bind covalently with other proteins. However, they can form non-covalent interactions with proteins containing either ubiquitin-associated or ubiquitin-like binding domains (Kiel and Serrano 2006).

We have previously characterized a 124-amino acid polypeptide (P400) from the haloalkaliphilic archaeon *Natrialba magadii* (Nercessian et al. 2009). Despite not displaying significant sequence identity with any known protein, including ubiquitin, P400 shows a folding similar to that of Ubls according with in silico secondary structure prediction. Also, we obtained a 3D model of P400 similar to that of human ubiquitin 1 (pdb code: 1UBQ). These results also fit with FTIR (Fourier transform infrared spectroscopy) studies showing that the recombinant P400 is mainly built of β-sheets, similar to Ubls (Ordóñez et al. 2011).

Based on the recent sequenciation of the *Natrialba magadii*'s genome, this work intended to settle the gene encoding P400 and analyze its expression in cells at both different growth phases and culture conditions. Also, an exploration on its biological role was performed.

Materials and methods

Bioinformatic analysis

A search for the complete gene to which *p400* DNA fragment correspond was done using sequence identity searches blastP against *Natrialba magadii*'s, ATCC 43099, Shotgun complete Genome at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). Protein sequence homology searches were performed with psi-blast. Protein sequences alignment of the proteins retrieved from psi-blast were done using ClustalW (Thompson et al. 1994). Manual adjustment of the alignment was performed accordingly with psi-blast results. Secondary structure predictions were calculated with Jpred and Psipred (Cole et al. 2008; Jones 1999). Gene context analysis was done using Integrated Microbial Genomes (IMG) Data Management System (Mavromatis et al. 2009).

Biological material and growth conditions

The haloalkaliphilic archaeon, *Natrialba magadii,* ATCC 43099, was grown in culture medium as described by Tindall et al. (1984), except that casamino acids were replaced by yeast extract (5 g/l). Also *N. magadii*'s cell cultures under starvation (yeast extract was removed) or low salt (15 % w/v of NaCl) conditions were developed. *Escherichia coli* cells were grown in LB medium supplied with antibiotics when needed. Cells were grown at 37 °C under aerobic conditions and constant agitation (170 r.p.m.). Growth was measured by optical density at 600 nm (OD_{600}).

RNA purification

Total RNA was isolated from *N. magadii*'s cells from exponential (OD_{600} ~0.7), early stationary (OD_{600} ~1.2) and late stationary (OD_{600} ~2) growth phases of control and treated cultures using TriZOL reagent (Invitrogen S.A.) as specified by the manufacturer. RNA was resuspended in DEPC-treated H_2O and its concentration was determined using an Ultraspect 1100 spectrophotometer.

Northern blot analysis

RNA from different growth phases was denatured by incubation with loading buffer containing 0.2 μg/ml BrEt and 6 M glyoxal at 55 °C for 1 h. RNA samples were loaded onto a 1.5 % (w/v) agarose gel and electrophoresis was performed at 100 V for 1 h. RNA was transferred from the gel to nitrocellulose Hybond N+ membrane (Amersham S.A.) by downward capillarity action for 16 h using 10× SSC (Ming et al. 1994). RNA was cross-linked to the membrane using a UV Stratalinker 1800 (Stratagene S.A.) and then hybridized to gene-specific probes. *Nmag_2608* probe, corresponding to the middle region of *Nmag_2608* gene, was obtained by PCR from *N. magadii*'s genomic DNA using specific primers: F1 sense 5′-GCGGATCCA GATCTTCGTCAAA-3′, R1 antisense 5′- TAAGCTTGAT CTTCGTGAAGA-3′. Also, a probe for the gene contiguous *Nmag_2609* was obtained by PCR from *N. magadii*'s genomic DNA using specific primers: F2 sense 5′-ATGT ACGTCGATCTTACCCA-3′, R2 antisense 5′-TTAGT CAGCGGTCGATGCGG-3′ and used to determine the possible co-expression of the gene *Nmag_2608* together with *Nmag_2609* gene. *7S*-specific probe for the constitutively expressed ribosomal 7S rRNA gene was obtained by PCR from *N. magadii*'s genomic DNA using the specific primers: F3 sense 5′-AGTTGCTGATGCCGGCGTGTC-3′ and R3 antisense 5′-GGTGGTCCGCTGCTCACTTC-3′ borrowed from Dr. De Castro and used as loading control. All probes were radioactively labeled using Random Primers DNA Labeling System (Invitrogen S.A.) and ^{32}P-α-dCTP as specified by manufacturer. For the hybridization step, membranes with cross-linked RNA samples were incubated with modified Church solution (0.5 M phosphate buffer pH 7.5, 7 % SDS and 10 mM EDTA pH 8) for 30 min at 65 °C followed by incubation with 12.5 ng of the denatured ^{32}P-labeled probe in the same solution at 65 °C for 16 h. Membranes were washed twice with 2× SSC-0.1 % SDS for 20 min followed by one wash with 0.5× SSC-0.1 % SDS for 20 min. For detection of the hybridized RNA, membranes were exposed in an amplification film Fuji Imagining Plate Type BAS III and analyzed by the bioimaging analyser Storm 840 Amersham Biosciences. Densitometry analysis was performed using ImageQuant 5.2 software.

Heterologous expression of *Nmag_2608* gene

The *Nmag_2608* gene was amplified by PCR from *N. magadii*'s genomic DNA. Two versions of the gene with (F4) and without the first twenty residues of the protein N-terminal region (F4′) were obtained using two sets of sense primers: F4 5′-CATATGCGACAGCTTGCCGCCC T-3′ and F4′ 5′-CATATGAGTCGTTCGACTCATCG CCA-3′ containing restriction sites for *Nde*I enzyme combined with the antisense primer R4 5′-TTAAG CTTGTCGTCAGCCGACTG-3′ containing restriction site for *Hin*dIII enzyme. The products were digested with *Nde*I and *Hin*dIII and cloned into the expression vector pET24b(+) (Novagen), digested with the same enzymes. The resulting constructs, pET24b-*nmag2608**C*** and pET24b-*nmag2608**T***, were transformed into *E. coli* Rosetta (DE3) competent cells, plated onto LB plates containing 50 μg/ml kanamycin and 25 μg/ml chloramphenicol, and incubated at 37 °C overnight. Positive clones were verified by colony-PCR. *E. coli* Rosetta (DE3) cells harboring each construction were grown in selective LB medium to an OD_{600nm} of 0.5 and expression was induced by adding 0.5 mM IPTG for 3 h at 37 °C. The cell pellets and clarified medium were separated by centrifugation.

Protein preparation

Cells were collected from different growth phase cultures of *N. magadii* or *E. coli* Rosetta strain expressing *Nmag_2608* and resuspended in buffer 100 mM Tris–HCl pH 8, 500 mM NaCl and 1 mM PMSF. The resuspended cells were sonicated and centrifuged at 17,000×*g* for 30 min and the supernatant was stored as total protein extracted. Extracellular protein samples were prepared by precipitation with 50 % acetone and 1 mM PMSF of the clarified culture medium, at 4 °C for 16 h. The protein pellets were resuspended in buffer 100 mM Tris–HCl pH 8 with 500 mM NaCl or SDS loading buffer. All proteins were conserved at −20 °C. Protein concentration in the samples was determined with bicinchoninic acid using bovine serum albumin (BSA) as standard (Smith et al. 1987). Extracellular proteins from *N. magadii* cultures at OD_{600nm} between 1 and 1.5 were also prepared. Cell pellets were harvested at 2–3 h intervals during the stationary phase of growth and extracellular proteins were obtained as described above.

SDS-PAGE and western blot assays

Total cellular and extracellular protein preparations from *N. magadii* and *E. coli* cultures expressing different versions of Nmag_2608 protein were loaded onto 12 % polyacrylamide gels. Following electrophoresis, proteins were transferred to a nitrocellulose membrane. Immunodetection assays were performed by incubation of membranes with anti-P400 primary polyclonal antibody (1:5,000) or anti-ubiquitin polyclonal primary antibody (1:100) (Sigma S.A.) for 16 h. The immunoreactive band was visualized after incubation with alkaline phosphatase-conjugated secondary antibody (1:10,000), using NBT (nitroblue tetrazolium) and BCIP (5-bromo-4-chloro-3-indolyl phosphate disodium salt).

Results

Identification of the gene containing P400 region: Nmag_2608

An in silico survey of the *N. magadii*'s genome using P400 sequence as query identified the ORF Nmag_2608 of 262 residues (blastP from NCBI). This ORF is annotated as conserved hypothetical protein without sequence similarity with any protein of known role. The protein encoded by Nmag_2608 is 90 amino acids longer in its C-terminus than P400 and has a signal peptide in the N-terminus (SignalP; Bendtsen et al. 2004) (Fig. 1). This N-terminal region belongs to signal peptides that contain a lipobox motif L-[I/G/A]-[A/G/S]-C (Hayashi and Wu 1990) present in bacterial lipoproteins for secretion (Prosite program; Hulo et al. 2007). In Nmag_2608, this motif contains Ser instead of Leu.

For further characterization, psi-blast sequence similarity searches using Nmag_2608 as query against the NCBI nr database with a threshold *e* value of 0.05 were performed. Proteins with no assigned role from different halophilic archaea were retrieved (Fig. 2). They show good sequence identity and similarity with Nmag_2608 (*e* value between e-42 and e-11 after 3 iterations). Also, all of them have a similar length to Nmag_2608, and several contain a lipobox-like signal motif. Nmag_2608 probably belongs to a halophilic archaeal family of secreted proteins having a putative lipoprotein signal peptide. Then, a group of proteins containing domains related to solute binding was retrieved by iteration 4 of this psi-blast search. Among them periplasmic solute-binding proteins (mainly phosphate binding proteins) with a lysR domain (pfam code PF03466); proteins with immunoglobulin fold domains (protein OM2255_1450 glycosyl hydrolase-like protein and fibronectin type III domain-containing proteins); and V12B01_23145 SLBB-like proteins from *Vibrio splendidus* (see supplementary material).

Analysis of the secondary structure of Nmag_2608 by prediction programs showed that, beside P400, residues in C-terminal region display a secondary structure mainly composed of β-sheets with an ending α-helix (Fig. 2).

Fig. 2 Multiple sequence alignment of Nmag_2608 and related halophilic proteins. Sequences were obtained by psi-blast sequence similarity search after 3 iterations (*e* value threshold 0.05) using NCBI tools (http://www.ncbi.nlm.nih.gov/BLAST/). Alignments were done using ClustalW and manually corrected according to psi-blast results. Protein names are indicated at *left* while their lengths at *right*. The secondary structure predicted by psiPred program is indicated in the *last lane* of each alignment, where β-sheet is signaled by **e** and α-helix by **h**. Nmag, *N. magadii*; Huta, *Halorhabdus utahensis*; Htur, *Haloterrigena turkmenica*; Hbor, *Halogeometricum borinquense*; rrnAC, *Haloarcula marismortui*; Hmuk, *Halomicrobium mukohataei*; HQ, *Haloquadratum walsbyi*; Hlac, *Halorubrum lacusprofundi*; HVO, *Haloferax volcanii*; VNG, *Halobacterium* sp. ▶

Since similarity searches did not retrieve any protein sequence with known function (threshold below 0.005 and identity higher than 15 %), a gene context analysis of Nmag_2608 was performed. It allows to find out functional relationships of annotated "conserved hypothetical proteins". This term refers to proteins without sequence similarity with those of known function (Mavromatis et al. 2009). Using these IMG tools, a group of nine halophilic archaeal ortholog genes to which *Nmag_2608* belongs was retrieved, which showed a divergent distribution around the gene of interest. Only *Hmuk_2777* and *Huta_2659* from *Halomicrobium mukohataei* and *Halorhabdus utahensis*, respectively, show a conserved gene neighborhood. Finally, *Nmag_2608* gene is found in the same chromosomal cassette with methyl-tRNA synthetase *Nmag_2601*, beta-lactamase *Nmag_2602*, pyruvate kinase *Nmag_2605*, cell surface adhesin *Nmag_2606*, cyclase *Nmag_2609*, ATP binding protein *Nmag_2610*, and several hypothetical proteins including the upstream gene *Nmag_2607*.

Expression of *Nmag_2608* gene during early stationary phase of growth of *N. magadii*'s cells

The existence of *Nmag_2608* gene encoding the P400-containing protein led to corroborate its expression in *N. magadii*'s cells. Northern blot tests using P400 DNA as probe to evaluate *Nmag_2608* expression during growth

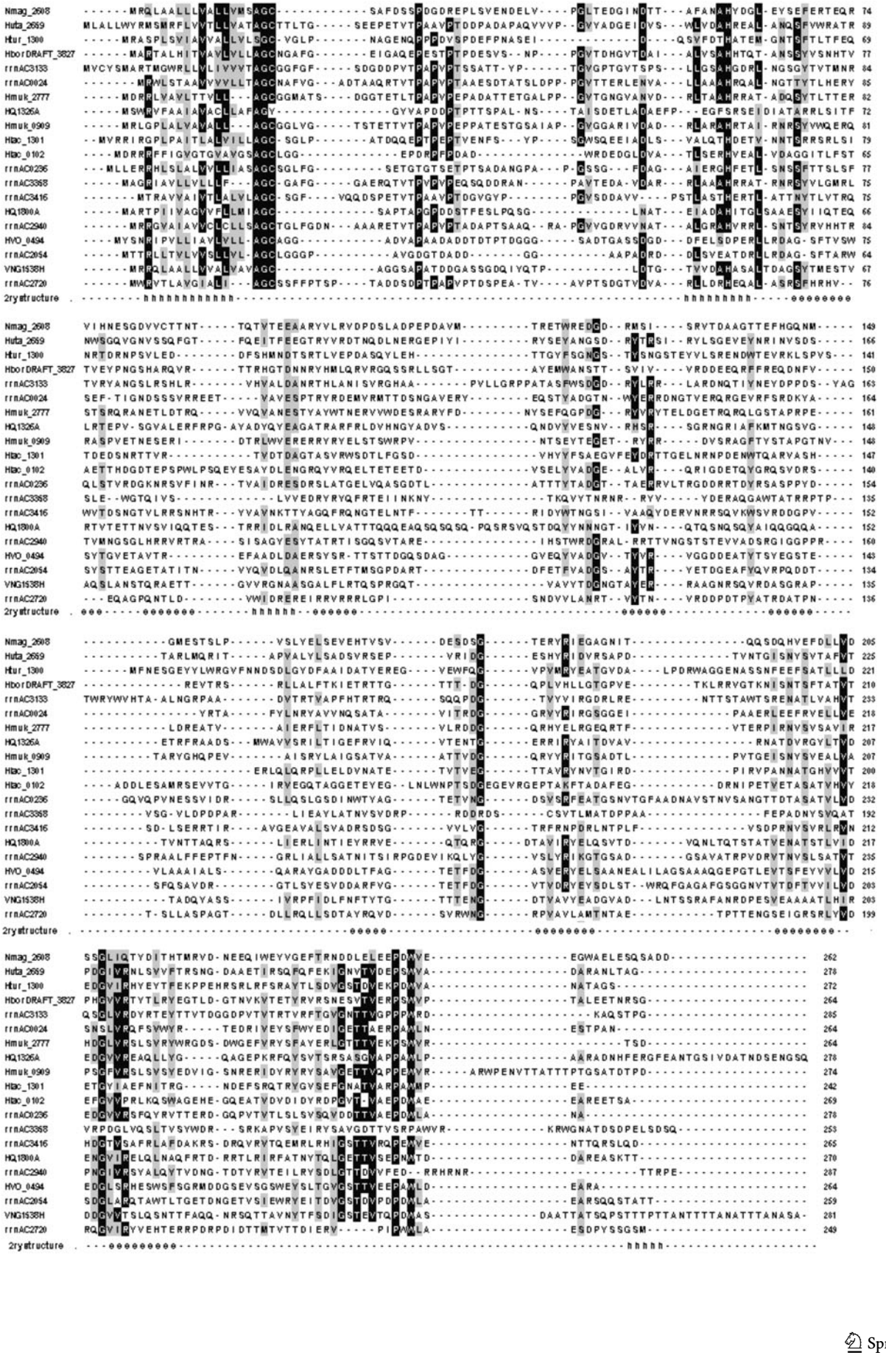
```
P400       ------------------------------------------------------------NAHYDGLEYSEFERTEQRVIHNES  24
Nmag_2608  MRQLAALLLVALLVMSAGCSAFDSSPDGDREPLSVENDELVPGLTEDGINDTTAFANAHYDGLEYSEFERTEQRVIHNES  80

P400       GDVVCTTNTTQTVTEEAARYVLRVDPDSLADPEPDAVMTRETWREDGDRMSISRVTDAAGTTEFHGQNMGMESTSLPVSL  104
Nmag_2608  GDVVCTTNTTQTVTEEAARYVLRVDPDSLADPEPDAVMTRETWREDGDRMSISRVTDAAGTTEFHGQNMGMESTSLPVSL  160

P400       YELSEVEHTVSV------------------------------------------------------------------------  116
Nmag_2608  YELSEVEHTVSVDESDSGTERYRIEGAGNITQQSDQHVEFDLLVDSSGLIQTYDITHTMRVDNEEQIWEYVGEFTRNDDL  240

P400       ----------------------- 116
Nmag_2608  ELEEPDWVEEGWAELESQSADD 262
```

Fig. 1 Alignment of the P400 polypeptide sequence with hypothetical protein Nmag_2608 from *Natrialba magadii*'s genome. A blastp was performed using the amino acidic sequence of P400 against the genome of *Natrialba magadii* ATCC 43099. Nmag_2608 correspond to the conserved hypothetical protein Nmag_2608 sequence. *Highlighted in grey* in Nmag_2608 sequence is a signal peptide predicted by SignalP 3.0 HMM and Prosite (signal peptide probability 1.000)



were performed. Figure 3a shows that a 1.4 kbp transcript appears at early stationary growth phase. Because this size was higher than the expected 786 bp for *Nmag_2608*, its genomic context was analyzed. Since it produces a 336 bp transcript in antisense mode, co-transcription of upstream *Nmag_2607* with *Nmag_2608* is unlikely. Conversely, *Nmag_2609,* noted as a cyclase, is located downstream from *Nmag_2608* and transcribed in the same direction. Since *Nmag_2609* has a 666 bp mRNA, its co-transcription with *Nmag_2608* could explain the transcript size (1.4 kbp) observed. Agreeing with the existence of a polycistronic mRNA harboring both genes, the same band was revealed with *Nmag_2609* and *Nmag_2608* probes (Fig. 3a). In other words, both genes are expressed in the same transcriptional cluster. Also, a manual search at the intergenic region between both genes for consensus archaeal promoter sequence BRE and tata box showed that these sequences are absent in the upstream region from the ATG of *Nmag_2609* (data not shown).

To get information on Nmag_2608 physiological role, northern blot tests with total RNA of cultures grown in low salt concentration (15 % NaCl) were performed. The *Nmag_2608* transcript amount decreased 39 % in cells grown in low salt (Fig. 3b). As expected for polycistronic genes, a similar result was obtained with *Nmag_2609* probe (Fig. 3b). Also, both *Nmag_2608* and *2609* decreased in starved cells, even though these results were less precise (data not shown).

Localization of Nmag_2608 protein in the extracellular protein fraction

Nmag_2608 protein is predicted to contain a lipoprotein signal peptide. To corroborate its membrane localization, western blot tests of different cellular fractions of *N. magadii*'s cultures with an antibody raised against P400 region were performed. Total cellular protein, cytosolic, and membrane fractions from *N. magadii*'s cultures obtained at exponential, early and late stationary phase of growth did not show reactive bands (data not shown). Because northern blot analysis showed that *Nmag_2608* transcript appears only in early stationary growth phase (Fig. 3), western blot tests of extracellular protein fractions from close time intervals within this growth phase were made. As viewed in Fig. 4b a unique 17 kDa band, smaller than the expected theoretical 29 kDa mass of Nmag_2608 complete protein, was revealed in all the samples tested. Similarly, this band was also detected by anti-ubiquitin antibody (Fig. 4c) with an intensity lower than that observed with anti-P400 antibody. Mass spectrometry analysis of the reactive band identified one peptide mass corresponding to the region between 121 and 129 of Nmag_2608. This result confirms that the reactive secreted protein corresponds to a fragment of Nmag_2608 protein.

Since in silico analysis of Nmag_2608 predicted that the signal for translocation of the gene product is coded by the 20 first residues of the protein, we decided to confirm this

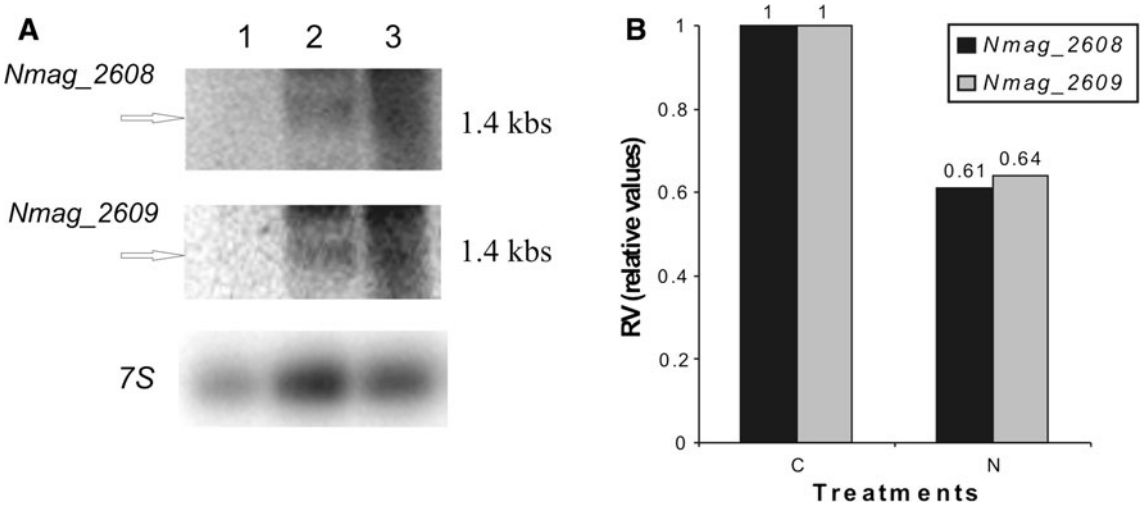


Fig. 3 Northern blot analysis of *Natrialba magadii*'s *Nmag_2608* transcript. Membranes containing total mRNA from *N. magadii* were hybridized with gene-specific probes. **a** Northern blot of mRNA from different growth phases of cultures grown under normal conditions. *Lane 1* exponential, *2* early stationary, *3* late stationary phase of growth of *N. magadii* culture. *Nmag_2608 panels*: *Nmag_2608* gene internal probe labeled with ^{32}P-α-dCTP. *Nmag_2609 panels*: *Nmag_2609* gene probe labeled with ^{32}P-α-dCTP**.** *Arrows* indicate the detected transcript with both probes which have a relative size of 1.4 kbp. *7S panels*: hybridization of the membrane with a *7S rRNA* ^{32}P-radioactive probe. **b** Graphical representation of relative amounts of *Nmag_2608* and *Nmag_2609* transcript in early stationary phase of growth for normal (*C*) and treated cultures, 15 % NaCl (*N*) of *N. magadii* obtained by densitometry analysis. The densitometry was corrected to the amounts of 7S rRNA loaded and for the treatment cultures the values were normalized to the control culture values. *Inset* indicates color reference for the *bars* in the graph. The figure is a representation of four independent experiments

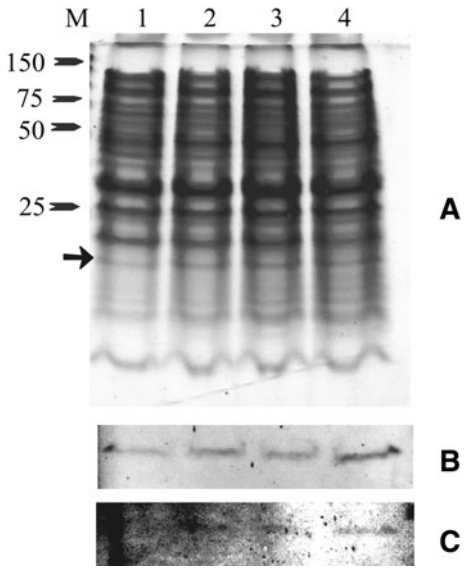


Fig. 4 Localization of Nmag_2608 protein by western blot analysis of extracellular protein samples of *N. magadii*'s cultures. Extracellular protein preparations were obtained from clarified medium cultures from stationary growth phase grown to OD_{600} of 1.02 (*lane 1*), 1.20 (*lane 2*), 1.27 (*lane 3*) and 1.37 (*lane 4*). Proteins were separated by SDS-PAGE (**a**) and subjected to western blot analysis against Nmag_2608-specific anti-P400 primary antibody (**b**) or anti-ubiquitin primary antibody (**c**). The *arrow* indicates the corresponding reactive band in the gel. Molecular weight standard (*M*) positions are indicated to the *right* of the image

by heterologous expression tests. Thus, two constructs with or without the predicted N-terminus signal peptide were expressed in *E. coli* Rosetta cells. Western blot assays against anti-P400 were performed to intracellular and extracellular protein fractions of the IPTG-induced cultures of *E. coli* containing either pET24b-*nmag2608**C*** (complete) or pET24b-*nmag2608**T*** (truncated) construct. The complete version of Nmag_2608 expressing cultures displayed two bands of ~48 and 38 kDa by SDS-PAGE in the extracellular fraction (Fig. 5a). Western blot tests showed that both bands reacted with the anti-P400 primary antibody. Also, both bands showed identity with Nmag_2608 protein sequence when analyzed by mass spectrometry (data not shown). On the other hand, the truncated version of the protein (Nmag2608Tr) expressed in *E. coli* was ~48 kDa by SDS-PAGE, which was not secreted but remained associated with the cells (Fig. 5b).

Discussion

We previously proved that P400, a polypeptide of *N. magadii*, folds similar to ubiquitin family proteins (Nercessian et al. 2009; Ordóñez et al. 2011). We now characterized Nmag_2608 protein holding P400 (Fig. 1) which belongs to a group of orthologs from different

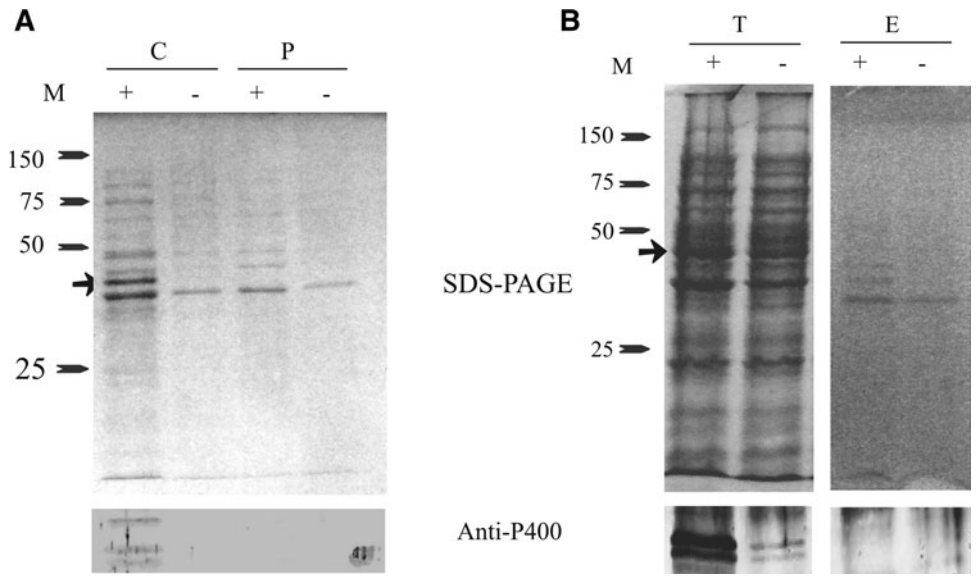


Fig. 5 Localization of complete and truncated recombinant versions of Nmag_2608 by western blot against anti-P400. *E. coli* Rosetta cell cultures containing the construction pET24b::*Nmag2608**C*** or pET24b::*Nmag2608**T*** were incubated with (+) or without (−) 0.5 mM IPTG, and protein from cellular and extracellular fractions was prepared as described (see "Materials and methods"). **a** SDS-PAGE (12 %) and western blot with anti-P400 primary antibody assay of extracellular protein fractions from cultures expressing the complete version Nmag2608rC (*lanes C*) and cultures containing the empty plasmid pET24b(+) (*lanes P*). **b** SDS-PAGE (12 %) and western blot with anti-P400 primary antibody assay of cellular (*lanes T*) and extracellular (*lanes E*) proteins of cultures expressing the truncated version Nmag2608rT. Molecular marker positions (*M*) are indicated in kDa at the *left* of each image. *Arrows* indicate the bands corresponding to the recombinant versions of Nmag_2608 protein

halophilic archaeal organisms without sequence identity with any known protein. Then Nmag_2608 and related proteins would be conserved among halophilic archaeal genomes. Bioinformatic data reveal that Nmag_2608 contains a signal peptide with the lipobox motif L-[I/G/A]-[A/G/S]-C found in bacterial lipoproteins and archaeal proteins (Hayashi and Wu 1990; Mattar et al. 1994; Giménez et al. 2007). Then, Nmag_2608 possibly links to the cell membrane by a thioester bond of the lipobox Cys (Hayashi and Wu 1990). Scarce information is available about membrane proteins with ubiquitin fold. Downes et al. (2006) described a group of membrane-anchored proteins, membrane-anchored ubiquitin fold (MUBs) found in several eukaryotic organisms. Nevertheless, these proteins would be linked to the membrane through prenylation of its c-terminal region, a different mechanism from the lipoproteins anchorage.

Unlike bacteria, archaea would not have a signal peptidase II to process lipoproteins signal peptides (Tokuda and Matsuyama 2004; Ng et al. 2007). However, several *Haloferax volcanii* lipoproteins have been reported to join the cell membrane through their lipobox Cys (Giménez et al. 2007). In addition, a signal peptide peptidase gene (*Nmag_2612*) is located within the *Nmag_2608* genomic context. The protein encoded by this gene could then take part in Nmag_2608 processing.

It has been shown that β-grasp fold is stable within proteins having neither sequence nor functional similarities. This fold takes part in many biochemical tasks by association with proteins and non-protein-soluble substrates (Burroughs et al. 2007a, b). For example, NqoI class proteins, involved in polysaccharide export and predicted for secretion and anchoring into the cell membrane display a N-terminal β-sheet rich domain and 1–8 soluble ligand binding β-grasp (SLBB) domains with a β-grasp fold (Burroughs et al. 2007b). After 4 iterations of a psi-blast search using as query the Nmag_2608 sequence a 326-amino acid V12B01_23145 protein was retrieved, which holds a SLBB-like domain from residues 137–326. Despite having 13 % identity with Nmag_2608, this domain aligns with the region holding P400 (Fig. 2 and supplementary material). Then, Nmag_2608 resembles an SLBB domain-containing protein with ubiquitin-like structural homology. Also, a group of proteins with β-grasp fold, extracellular localization and possible role in protein–protein interaction has been described in *Mycobacterium* species (Wang et al. 2007). Finally, like several prokaryotic and eukaryotic ubiquitin-like proteins, Nmag_2608 displays a DGD motif in its β-grasp domain (Nercessian et al. 2009) also present in SLBB domain (Burroughs et al. 2007b). All together, the interaction of Nmag_2608 with other molecules could be predicted.

Acquired bioinformatic data led to examine whether *N. magadii*'s cells express the *Nmag_2608* gene. Although larger than expected, a 1.4 kbp transcript was detected only at early stationary phase of growth (Fig. 3). Analysis of the genomic context of *Nmag_2608* suggests its co-transcription with the adjoining *Nmag_2609* gene encoding for a protein of the cyclase family. Cyclases take part in the transport and metabolism of amino acids (COG001878). Although distant from *Nmag_2608* stop codon, *Nmag_2609* plus *Nmag_2608* can produce a transcript of $\sim$1.4 kbp since their sizes are 666 and 786 bp, respectively. This was corroborated by northern blot assay (Fig. 3a). Nmag_2609 shows homology with metal-dependent hydrolases involved in the tryptophan metabolism by converting *N*-formyl kynurenine in kynurenine and formic acid (Pabarcus and Casida 2005). This pathway is mainly responsible for the synthesis of NAD/NADP in eukaryotes (Dobrovolskya et al. 2005). Conversely, it takes part in the synthesis of siderophores and antibiotics within some bacteria (Keller et al. 2010; Kurnasov et al. 2003; Lomovskaya et al. 1998; Matthijs et al. 2004; Narui et al. 2009). Then, co-expression may suggest that Nmag_2608 is involved in the metabolism of tryptophan. By the other way, the archaeon *Methylococcus capsulatus* uses the extracellular membrane-associated protein MopE for copper uptake (Helland et al. 2008). For this, a kynurenine residue formed by oxidation of tryptophan in the MopE copper-binding site is required. Then, co-expression of *Nmag_2608* with *Nmag_2609* could be necessary to change one of its three tryptophans to kynurenine for interaction with ligands of the processed gene product. This hypothesis will be considered in further research.

To get information on physiological function, Nmag_*2608* and *Nmag_2609* transcription in stressed cells was assessed. Both transcripts levels decreased nearly 37.5 % in cells grown under low salinity at stationary growth phase (Fig. 3b). These changes agree with those displayed by genes associated with the transport of metabolites across the membrane of *Halobacterium* sp. NRC-1 grown in low salinity (Coker et al. 2007). Then, cells keep proper intracellular ionic conditions by modulation of transporter genes. Agreeing with this, present data suggest that Nmag_2608 takes part in the uptake of solutes from the environment.

To check the Nmag_2608 localization in *N. magadii*'s cells, western blot tests using an anti-P400 antibody were performed. They revealed a reactive band in the extracellular fraction of *N. magadii* at early stationary growth phase (Fig. 4), which proves protein secretion, also confirmed by reaction of the same band with an anti-ubiquitin antibody (Fig. 4c). The size of the revealed protein result was minor than expected, suggesting a further processing besides signal peptide removal, after translocation across

the membrane. Evenmore, mass spectrometry analysis of the reactive band allowed the identification of a peptide within the P400 region of the protein. These results indicate that the ubiquitin-like domain of Nmag_2608 is kept during processing and released to the environment.

In accordance with the above, expression of *Nmag_2608* in *E. coli* also produced two extracellular proteins (Fig. 5a). In this case, the 48 kDa band would correspond to the leaderless protein, while possibly, the smaller band (38 kDa) corresponds to the mature processed version of Nmag_2608. Conversely, and revealing a signal peptide need for translocation, truncated *Nmag_2608* expression produced a protein associated with cells (Fig 5b). Thus, Nmag_2608 would be processed differently from other bacterial or archaeal lipoproteins described so far (Mattar et al. 1994; Giménez et al. 2007). On the other hand, the different sizes of the revealed protein bands compared to those expected for Nmag_2608 could be due to the an unspecific interaction of the recombinant protein in *E. coli* with a 20 kDa polypeptide of unknown identity. Another possibility is that this difference results from reduced migration in SDS-PAGE as reported for some halophilic proteins (Madern et al. 2000). Supporting this, the identity of these proteins as Nmag_2608 was also confirmed by mass spectrometry.

Most haloarchaeal predicted lipoproteins are secreted by the Tat translocation system (Storf et al. 2010). Since it does not contain a Tat motif, Nmag_2608 is most likely to be secreted by a different system (data not shown). The protein expressed in *E. coli* is also translocated outside the cell, suggesting that the same mechanism of secretion works in both bacteria and archaea. Finally, whether Nmag_2608 interacts with solutes of the extracellular medium must be explored. Nevertheless, present results would be in good agreement with the previously described roles of many ubiquitin-like domains in protein–protein and protein–soluble ligand interactions.

To sum up, this work reports the first description of an archaeal protein containing an ubiquitin-like domain, which is processed and secreted to the extracellular medium.

References

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795

Burroughs AM, Balaji S, Iyer LM, Aravind L (2007a) Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold. Biol Direct 2:18. doi:10.1186/1745-6150-2-18

Burroughs AM, Balaji S, Iyer LM, Aravind L (2007b) A novel superfamily containing the β-grasp fold involved in binding diverse soluble ligands. Biol Direct 2:4. doi:10.1186/1745-6150-2-4

Ciechanover A (1994) The ubiquitin-proteasome proteolytic pathway. Cell 79:13–21

Ciechanover A, Iwai K (2004) The ubiquitin system: from basic mechanisms to the patient bed. IUBMB Life 56:193–201

Coker JA, DasSarma P, Kumar J, Müller JA, DasSarma S (2007) Transcriptional profiling of the model Archaeon *Halobacterium* sp. NRC-1: responses to changes in salinity and temperature. Saline Systems 3:6. doi:10.1186/1746-1448-3-6

Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server. Nucl Acids Res 36(Web Server issue):W197–W201. doi:10.1093/nar/gkn238

Dobrovolskya VN, Bowyerb JF, Pabarcusc MK, Heflicha RH, Williamsd LD, Doerged DR, Arvidssone B, Bergquiste J, Casidac JE (2005) Effect of arylformamidase (kynurenine formamidase) gene inactivation in mice on enzymatic activity, kynurenine pathway metabolites and phenotype. Biochim Biophys Acta 1724:163–172

Downes BP, Saracco SA, Lee SS, Crowell DN, Vierstra RD (2006) MUBs, a family of ubiquitin-fold proteins that are plasma membrane-anchored by prenylation. J Biol Chem 281: 27145–27157

Dye BT, Schulman BA (2007) Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins. Annu Rev Biophys Biomol Struct 36:131–150

Giménez MI, Dilks K, Pohischröder M (2007) *Haloferax volcanii* twin-arginine translocation substates include secreted soluble, c-terminally anchored and lipoproteins. Mol Microbiol 66:1597–1606

Grabbe C, Dikic I (2009) Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins. Chem Rev 109:1481–1494

Hartmann-Petersen R, Gordon C (2004) Integral UBL domain proteins: a family of proteasome interacting proteins. Semin Cell Dev Biol 15:247–259

Hayashi S, Wu HC (1990) Lipoproteins in bacteria. J Bioenerg Biomembr 22:451–471

Helland R, Fjellbirkeland A, Karlsen OA, Ve T, Lillehaug JR, Jensen HB (2008) An oxidized tryptophan facilitates copper binding in *Methylococcus capsulatus*-secreted protein MopE. J Biol Chem 283:13897–13904

Hochstrasser M (2000) Evolution and function of ubiquitin-like protein-conjugation systems. Nat Cell Biol 2:53–157

Hochstrasser M (2009) Origin and function of ubiquitin-like proteins. Nature 458:422–429

Hulo N, Bairoch A, Bulliard V, Cerutti L, Cuche BA, de Castro E, Lachaize C, Langendijk-Genevaux PS, Sigrist CJA (2007) The 20 years of PROSITE. Nucl Acids Res 36:D245–D249

Humbard MA, Miranda HV, Lim J, Krause DJ, Pritz JR, Zhou G, Chen S, Wells L, Maupin-Furlow JA (2010) Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. Nature 463:54–60

Iyer LM, Burroughs AM, Aravind L (2006) The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like beta-grasp domains. Genome Biol 7:R60. doi:10.1186/gb-2006-7-7-r60

Jeong YJ, Jeong BC, Song HK (2011) Crystal structure of ubiquitin-like small archaeal modifier protein 1 (SAMP1) from *Haloferax volcanii*. Biochem Biophys Res Commun 405:112–117

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292:195–202

Keller U, Lang M, Crnovcic I, Pfennig F, Schauwecker F (2010) The actinomycin biosynthetic gene cluster of *Streptomyces chrysomallus*: a genetic hall of mirrors for synthesis of a molecule with mirror symmetry. J Bacteriol 192:2583–2595

Kerscher O, Felberbaum R, Hochstrasser M (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol 22:159–180

Kiel C, Serrano L (2006) The ubiquitin domain superfold: structure-based sequence alignments and characterization of binding epitopes. J Mol Biol 355:821–844

Kurnasov O, Goral V, Colabroy K, Gerdes S, Anantha S, Osterman A, Begley TP (2003) NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria. Chem Biol 10:1195–1204

Lomovskaya N, Doi-katayama Y, Filippins S, Nastro C, Fonstein L, Gallo M, Colombo AL, Hutchinson CR (1998) The *Streptomyces peucetius dpsY* and *dnrX* genes govern early and late steps of daunorubicin and doxorubicin biosynthesis. J Bacteriol 180:2379–2386

Madern D, Ebel C, Zaccai G (2000) Halophilic adaptation of enzymes. Extremophiles 4:91–98

Makarova KS, Koonin EV (2003) Comparative genomics of archaea: how much have we learned in six years, and what's next? Genome Biol 4:115

Marquet A (2001) Enzymology of carbon-sulfur bond formation. Curr Opin Chem Biol 5:541–549

Mattar S, Scharf B, Kent SBH, Rodewald K, Oesterhelt D, Engelhard M (1994) The primary structure of halocyanin, an archaeal blue copper protein, predicts a lipid anchor for membrane fixation. J Biol Chem 269:14939–14945

Matthijs S, Baysse C, Koedam N, Tehrani KA, Verheyden L, Budzikiewicz H, Schäfer M, Hoorelbeke B, Meyer JM, De Greve H, Cornelis P (2004) The *Pseudomonas* siderophore quinolobactin is synthesized from xanthurenic acid, an intermediate of the kynurenine pathway. Mol Microbiol 52:371–384

Mavromatis K, Chu K, Ivanova N, Hooper SD, Markowitz VM, Kyrpides NC (2009) Gene context analysis in the Integrated Microbial Genomes (IMG) data management system. PLoS One 4:e7979

Ming YZ, Di X, Gomez-Sanchez EP, Gomez-Sanchez CE (1994) Improved downward capillary transfer for blotting of DNA and RNA. Biotechniques 16:58–59

Narui K, Noguchi N, Saito A, Kakimi K, Motomura N, Kubo K, Takamoto S, Sasatsu M (2009) Anti-infectious activity of tryptophan metabolites in the L-tryptophan–L-kynurenine pathway. Biol Pharm Bull 32:41–44

Nercessian D, Marino Buslje C, Ordóñez MV, De Castro RE, Conde RD (2009) Presence of structural homologs of ubiquitin in haloalkaliphilic Archaea. Int Microbiol 12:167–173

Ng SYM, Chaban B, VanDyke DJ, Jarrell KF (2007) Archaeal signal peptidases. Microbiology 153:305–314

Ordóñez MV, Guillén J, Nercessian D, Villalain J, Conde RD (2011) Secondary structure determination by FTIR of an archaeal ubiquitin-like polypeptide from *Natrialba magadii*. Eur Biophys J 40:1101–1107

Pabarcus MK, Casida JE (2005) Cloning, expression, and catalytic triad of recombinant Arylformamidase. Protein Expr Purif 44:39–44

Pickart CM (2001) Mechanisms underlying ubiquitination. Ann Rev Biochem 70:503–533

Ranjan N, Damberger FF, Sutter M, Allain FHT, Weber-Ban E (2011) Solution structure and activation mechanism of ubiquitin-like small archaeal modifier proteins. J Mol Biol 405:1040–1055

Rudolph MJ, Wuebbens MM, Rajagopalan KV, Schindelin H (2001) Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. Nat Struct Biol 8:42–46

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC (1987) Measurement of protein using bicinchoninic acid. Anal Biochem 150:76–85

Soppa J (2006) From genomes to function: haloarchaea as model organisms. Microbiology 152:585–590

Storf S, Pfeiffer F, Dilks K, Chen ZQ, Imam S, Pohlschröder M (2010) Mutational and bioinformatic analysis of haloarchaeal lipobox-containing proteins. Archaea. doi:10.1155/2010/410975

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22:4673–4680

Tindall BJ, Ross HN, Grant WD (1984) *Natronobacterium* gen. nov. and *Natronococcus* gen. nov., two new genera of the haloalkaliphilic archaebacteria. Syst Appl Microbiol 5:41–57

Tokuda H, Matsuyama S (2004) Sorting of lipoproteins of the outer membrane in *E. coli*. Biochim Biophys Acta 1693:5–13

Wang C, Xi J, Begley TP, Nicholson LK (2001) Solution structure of ThiS and implications for the evolutionary roots of ubiquitin. Nat Struct Biol 8:47–51

Wang Z, Potter BM, Gray AM, Sacksteder KA, Geisbrecht BV, Laity JH (2007) The solution structure of antigen MPT64 from *Mycobacterium tuberculosis* defines a new family of beta-grasp proteins. J Mol Biol 366:375–381