RESEARCH ARTICLE

INTERNATIONAL MICROBIOLOGY (2009) 12:167-173
DOI: 10.2436/20.1501.01.95 ISSN: 1139-6709 **www.im.microbios.org**

INTERNATIONAL MICROBIOLOGY

Presence of structural homologs of ubiquitin in haloalkaliphilic *Archaea*

Débora Nercessian,[1]* Cristina Marino Buslje,[2] María V. Ordóñez,[1] Rosana E. De Castro,[1] Rubén D. Conde[1]

[1]Institute of Biological Research, Faculty of Exact and Natural Sciences, National University of Mar del Plata and CONICET, Mar del Plata, Argentina. [2]Institute of Chemistry and Physicochemical Biology, Faculty of Pharmacy and Biochemistry, National University of Buenos Aires and CONICET, Buenos Aires, Argentina

Received 30 January 2009 · Accepted 20 August 2009

Summary. Ubiquitin, a protein widely conserved in eukaryotes, is involved in many cellular processes, including proteolysis. While sequences encoding ubiquitin-like proteins have not been identified in prokaryotic genomes sequenced so far, they have revealed the presence of structural and functional homologs of ubiquitin in *Bacteria* and *Archaea*. This work describes the amplification and proteomic analysis of a 400-bp DNA fragment from the haloalkaliphilic archaeon *Natrialba magadii*. The encoded polypeptide, P400, displayed structural homology to ubiquitin-like proteins such as those of the ThiS family and Urm1. Expression of the P400 DNA sequence in *Escherichia coli* cells yielded a recombinant polypeptide that reacted with anti-ubiquitin antibodies. In addition, a putative open reading frame encoding P400 was identified in the recently sequenced genome of *N. magadii*. Together, these results evidence the presence in *Archaea* of structural homologs of ubiquitin-related proteins. [**Int Microbiol** 2009; 12(3):167-173]

Keywords: *Natrialba magadii* · halophilic archaea · ubiquitin-like proteins · structural homology

Introduction

The roles of most genes from *Archaea* are unknown and are often predicted only on the basis of sequence alignments. Many archaeal genes and functions are shared with *Bacteria*, as shown by recent studies [11,15]. However, sequence similarity between an archaeal protein and its functionally characterized homolog is usually too low to be detected by computational methods [28,40]. Thus, the roles of many proteins from these microorganisms remain unknown. Recently, transcriptomic and proteomic studies of the haloarchaea *Halobacterium salinarum*, *Haloferax volcanii*, and *Natronomonas pharaonis* [4,23,44] have provided relevant advances in understanding gene and protein control mechanisms, but not in identifying orthologs for most archaeal proteins. For this purpose, folding assessment of proteins encoded by the genomes of diverse organisms is useful [29,46]. Also, efforts in the structural genomics field, including novel bioinformatics tools, protein crystallization, and structure determinations, have yielded significant insights into protein function [21,43].

Ubiquitin has important functions in several processes within eukaryotic cells, the most relevant being the targeting of protein substrates destined for degradation into the proteasome [17]. Although widely conserved in eukaryotes, ubiquitin has not been identified in the prokaryotic genomes so far sequenced, even though archaea have a functional protea-

***Corresponding author:** D. Nercessian
Instituto de Investigaciones Biológicas
Facultad de Ciencias Exactas y Naturales
Universidad Nacional de Mar del Plata-CONICET
7600 Mar del Plata, Argentina
Fax +54-2234751392
E-mail: dnercess@mdp.edu.ar

Authors D.N. and C.M.B. contributed equally to the paper.

some for protein degradation. However, several proteins from prokaryotes and eukaryotes share the ubiquitin fold, despite their low amino acid sequence identities [20,48]. These proteins, known as ubiquitin-like proteins (Ubls), have a β-grasp fold and are often endowed with similar roles. For instance, although not related to proteolysis, some of them bind covalently to other proteins through a mechanism mediated by activating enzymes. These proteins are considered as putative ancestors of ubiquitin, providing further evidence that a simple and stable fold was used by proteins involved in a wide variety of biological processes [13,18,20,48].

Among the eukaryotic proteins that share folding and partial primary sequence identity with ubiquitin are: SUMO (small ubiquitin modifier), RUB1 (related to ubiquitin 1, named NEDD8 in mammals), and ISG15 (interferon-stimulated gene) [18]. In addition, Apg12 and Urm1 (ubiquitin-related modifier 1 of yeast) show structural homology to ubiquitin [13]. It is noteworthy that these proteins belong to different protein conjugation systems in eukaryotes and are considered functional homologs of ubiquitin [16,18].

In prokaryotes, ThiS and MoaD display structural and functional characteristics resembling those of ubiquitin, but they share low sequence identity with the protein (14 and 7%, respectively) [5,18,36,48]. ThiS is a sulfide carrier protein engaged in thiamine biosynthesis, and its activation by the enzyme ThiF is similar to the activation of ubiquitin by E1 [48]. In addition, the primary sequence of ThiF resembles that of E1. Therefore, ThiS and ThiF have been proposed as bacterial ancestors of the ubiquitin pathway. MoaD, the small subunit of the molybdopterin synthase that is also present in archaea, shares the ubiquitin fold [36]. The yeast Urm1 protein is similar to MoaD and ThiS, but not to ubiquitin. However, Urm1 is considered to be a functional homolog of ubiquitin based on the thioester linkage between its C-terminal Gly and the E1-like protein Uba4 [48]. Therefore, Urm1 has been proposed as a molecular fossil, providing the evolutionary bridge between the protein conjugation ThiS-ThiF of bacteria and the protein ubiquitination of eukaryotes [48]. According to these findings, a bioinformatic work has suggested the existence of a bacterial counterpart of Urm1 [Scheel H et al. (2005) FEBS Journal 272 (s1) B2 024P]. Recently, prokaryotic sequences that might be functional homologs to ubiquitin have been found also by computational methods [5,20]. The bacterial protein YukD from *Bacillus subtilis* was identified by van den Ent and Löwe [45]. YukD adopts a ubiquitin-like fold but has a shorter C-terminal tail that ends before the conserved double-glycine motif of proteins belonging to the ubiquitin superfamily. Experimental data showed that YukD cannot form covalent bonds with other proteins.

We previously detected proteins that react with anti-ubiquitin antibodies in several haloalkaliphilic archaea [31]. In the present work, we amplified a 400-bp DNA fragment from the haloalkaliphilic archaeon *Natrialba magadii*. This fragment encodes a polypeptide, P400, that shares structural homology to ubiquitin and the ubiquitin-like proteins ThiS, MoaD, and Urm1. Protein sequence alignments showed the conservation of several structurally and functionally relevant amino acid residues among these proteins. In addition, the recombinant polypeptide synthesized in *Escherichia coli* cells cross-reacted with anti-ubiquitin antibodies, demonstrating the existence of structural homologs to ubiquitin in haloarchaea.

Materials and methods

Biological material. The haloalkaliphilic archaeon *Natrialba magadii*, ATCC 43099, was grown in the culture medium described by Tindall et al. [42], except that casamino acids were replaced by yeast extract (20 g/l). The cells were grown at 37°C under oxic conditions and constant agitation at 150 rpm. Growth was measured by the optical density at 600 nm (OD_{600}).

Amplification by PCR. *Natrialba magadii* genomic DNA was extracted according to the method of Ng et al. [32]. The oligonucleotides used as primers were designed based on the alignment of conserved ubiquitin sequences from the eukaryotes *Saccharomyces cereviciae*, *Homo sapiens*, *Tetrahymena pyriformis*, *Trypanosoma cruzi*, *Dictyostelium discoideum*, and *Homarus americanus* [38,49], and that described for the cyanobacterium *Anabaena variabilis* [10]. The sense primer (I) was 5´-CAGATCTTCGT CAAGACCCTCAC-3´, and the antisense primer (II) 5´-GACTCCTTCTG GATGTTGTAGTC-3´. The third nucleotide from the codon corresponding to each amino acid was selected taking into account the codon usage table for halophilic archaea [47].

Sunflower or potato leaf genomic DNA was used as the positive control and sterile water as the negative control of the amplification reactions. The reaction mixture contained *N. magadii* genomic DNA, 100–200 ng; 1× PCR buffer; 200 μM dNTPs; 1.5 mM $MgCl_2$; 1 μM of each primer; 1.5 U *Taq* DNA polymerase, in a final volume of 25 μl. The PCR program was the following: 94ºC, 2 min, 1 cycle; 94ºC, 1 min; 47ºC, 1 min; 72ºC, 2 min, for 30 cycles. The amplification products were separated by agarose gel electrophoresis (1.8% w/v), purified using the QIA Quick gel extraction kit (Qiagen, Valencia, CA, USA), and cloned into pGEM-T Easy Vector System I (Promega, Madison, WI, USA), according to the manufacturer's specifications. *E. coli* XL-Blue competent cells were transformed with the ligation products and plated onto LB medium containing 100 μg ampicillin/ml, 2% (v/v) X-gal, and 0.1 M IPTG. White colonies were chosen at random to verify the presence of inserts with colony-PCR using universal M13 forward and reverse primers. Plasmid DNA was isolated using Qiagen-tip100 columns (Qiagen) and sequenced.

Bioinformatic analysis. The search for protein sequence homology was done with Psi-BLAST [http://www.ncbi.nlm.nih.gov/BLAST/]. Secondary structure prediction was calculated with Psipred [6], Prof [35] and JPred 3 [7]. An alignment of Swiss-Prot true-positive hits of ubiquitin 2 formed by 365 sequences (prosite, PS50053 [19]) was used to align our sequence to that of the profile obtained with ClustalX [24]. Final adjustments were manually done. Several fold-recognition methods were tested through the BioInfoBank metaserver [14]. A model was generated with

Modeller [37] and its stereochemical quality checked with Procheck [25] and verify3D [27], and by superposition with its own template structure. The model was superposed with human ubiquitin (protein data bank (pdb) code: 1ubq) using the combinatorial extension method [39]. Databases SCOP [1], CATH [34], Pfam [2], PDB [3], and PDBSum [26] were also used.

Heterologous expression, purification, and renaturalization of P400. The P400 DNA fragment was PCR-amplified from *N. magadii* genomic DNA, digested with *Bam*HI and *Hin*dIII, cloned into the expression vector pET24b(+) (Novagen, Madison, WI, USA), and digested with the same enzymes. The resulting construct, pET24b-*p400*, was transformed into *E. coli* Rosetta (DE3) competent cells, plated onto LB plates containing 50 μg kanamycin/ml and 25 μg chloramphenicol/ml, and incubated at 37ºC overnight. Positive clones were verified for the presence of the 400-bp fragment by colony-PCR using the sense primer (III) 5´-GCG GATCCAGATCTTCGTCAAA-3´ and the antisense primer (IV) 5´-T TAAGCTTGATCTTCGTGAAGA-3´. *E. coli* Rosetta (DE3) cells harboring pET24b-*p400* were grown in LB medium containing the above-mentioned antibiotics to an OD_{600} of 0.5. P400 expression was induced by the addition of 0.5 mM IPTG for 3 h at 37ºC. The cells were harvested by centrifugation and cell extracts were obtained by sonication followed by centrifugation at 3000 ×*g* for 15 min. As the recombinant peptide was found in the inclusion bodies, it was resuspended in 0.1 M Tris-HCl (pH 7.5) containing 6 M guanidine-HCl and 0.5 M NaCl. The recombinant P400 protein, containing a His6 tail in its C-terminus, was purified by affinity chromatography using Ni^{2+} columns (Ni-NTA Purification System; Invitrogen, Carlsbad, CA, USA) under denaturing conditions as recommended by the manufacturer. Refolding of the purified P400-His6 was carried out by dialysis in 0.1 M Tris-HCl (pH 7.5) containing 2 M KCl at 4ºC for 16 h. The protein was then concentrated by Centricon YM 10 (Millipore, Billerica, MA, USA) and stored at –20ºC.

SDS-PAGE and Western blot assays. Protein concentration was determined by the bicinchoninic acid method [41] using bovine serum albumin as standard. Semi-denaturing SDS-PAGE was carried out under non-reducing conditions and without boiling the samples. Purified and renatured recombinant P400 (4 μg) was loaded on a semi-denaturing polyacrylamide gel (12%). After electrophoresis, the separated proteins were transferred to PVDF membranes and subjected to Western blotting by standard protocols using 1:100 rabbit anti-ubiquitin primary antibody (Sigma, St. Louis, MO, USA), as described before [31]. The immunoreactive band was visualized after incubation with alkaline-phosphatase-conjugated secondary antibody (1:10,000), using NBT (nitroblue tetrasolium) and BCIP (5-bromo-4-chloro-3-indolyl phosphate disodium salt). Additionally, dot-blot assays with the refolded purified protein were done under native conditions. A similar protocol as that described for the Western blot assays was used to detect the reactive spots.

Results and Discussion

Although computational searches among the archaeal genomes sequences have not shown the existence of ubiquitin homologs, our previous results provided evidence of proteins that react with anti-ubiquitin antibodies in haloalkaliphilic archaea [31]. These findings led us to further search for the genes encoding these proteins by PCR. Oligonucleotide primers were designed by aligning several ubiquitin sequences from eukaryotes and the sequence described for the cyanobacterium *Anabaena variabilis* [10,38,49]. Amplification of genomic DNA from *N. magadii* yielded 200-, 300-, and 400-bp products. The presence of the 200- and 400-bp fragments was encouraging since ubiquitin is encoded as repeating units (polyubiquitin). The 400-bp fragment contained an open reading frame (ORF) with no stop codon that was translated and used in a homolog search with the BLAST suite of programs. However, this search gave no hits with any known protein at either the DNA or the amino acid level.

As mentioned above, proteins having no significant amino acid sequence similarity can nonetheless share the same fold and, sometimes, the same function. Several proteins with structural homology but no sequence identity to ubiquitin have been described [20,36,45,48]. Based on these findings, and considering the structural nature of antigen-antibody interactions, we explored the possibility of shape homology between the P400 polypeptide and ubiquitin. Therefore, the secondary and tertiary structures of P400 were analyzed using the programs described in Materials and methods. Computational results strongly supported the hypothesis that P400 has an α/β fold with all the elements of secondary structure that define a β-grasp (ubiquitin-like) fold. The ThiS family of prokaryotic proteins (*E. coli* ThiS and molybdopterin converting factor subunit 1 and *Methanothermobacter thermoautotrophicus* hypothetical protein Mth1743), the *Bacillus subtilis* protein YukD, yeast Urm1, and other proteins belonging to eukaryotic conjugation systems share this same fold [13,45]. Furthermore, the ubiquitin of *Homo sapiens*, the ubiquitin-like protein NEDD8 of *Mus musculus*, and RUB1 of *Arabidopsis thaliana* also include a β-grasp fold [18].

The secondary structure of the P400 polypeptide was predicted using Psi-Pred, Prof and JPred (see Materials and methods). These programs showed an almost perfect coincidence in their prediction of the secondary structure elements of P400 and the topology of ubiquitin. Minor differences between the predictors were not surprising as all the programs use a single sequence for the calculation instead of a profile derived from a multiple sequence analysis (MSA); this procedure notably lowers the program's accuracy. Although the fold-recognition methods failed to assign a fold to P400 sequence, the secondary structure prediction from JPred and Prof (and from PsiPred, setting aside the three extra helices predicted) was compatible with a ubiquitin-like or β-grasp fold. To further test this possibility, a homology model was constructed with Modeller [37], using the human ubiquitin high-resolution structure 1UBQ.pdb as template. 1UBQ.pdb shares a 13.7% sequence identity with P400 (Fig. 1); however, despite the low sequence identity between the superfamily members, an alignment of Swiss-Prot true-positive hits of ubiquitin 2 formed by 365 sequences (prosite,

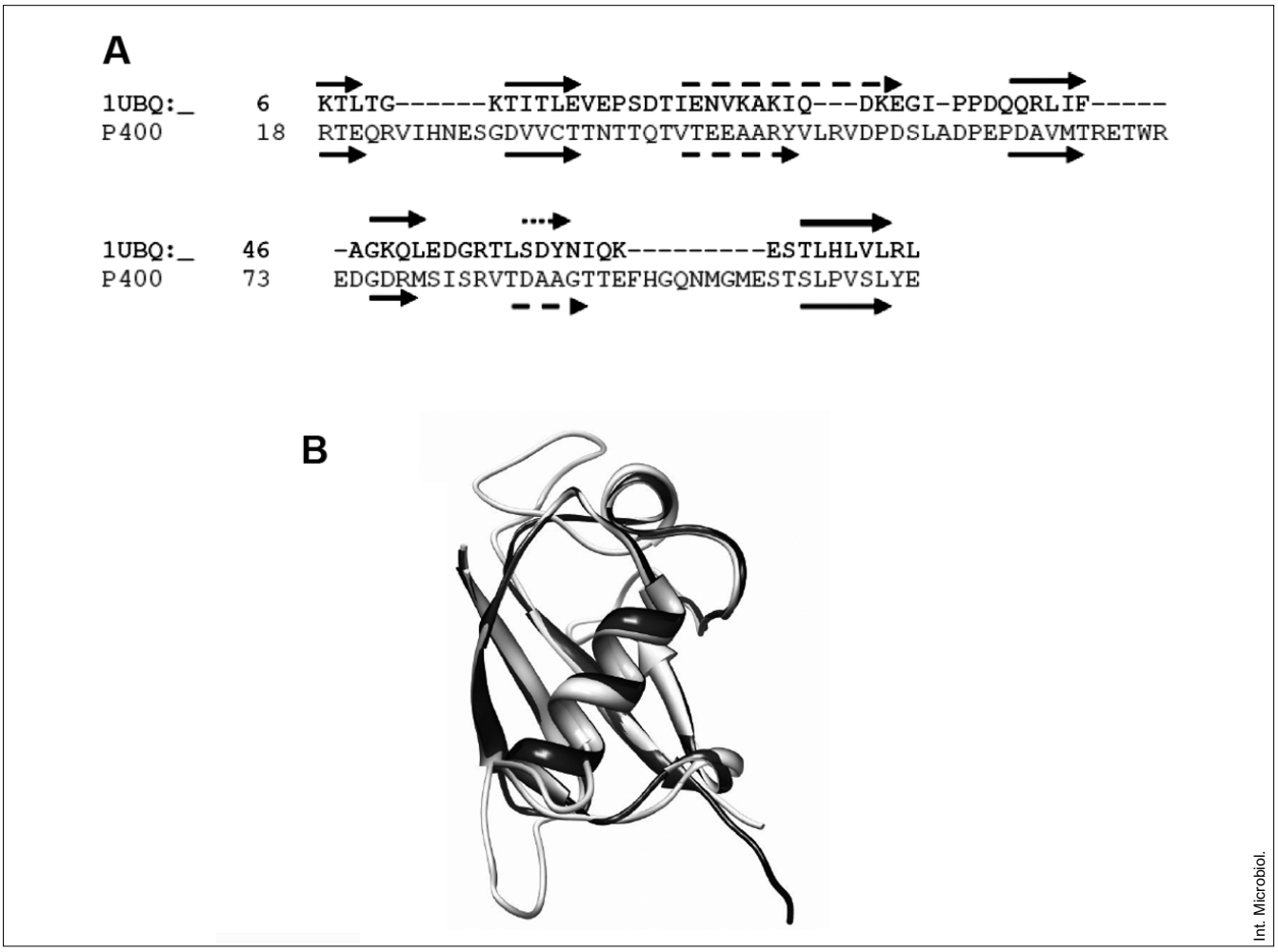

Fig. 1. (**A**) Structural alignment between 1UBQ and the modeled P400. Secondary structure elements are indicated as lines above and below each sequence: Full line: β chain; dashed line: α helix; dotted line: 3_{10} Helix. (**B**) Ribbon representation of the superimposed structures 1UBQ and P400; black 1UBQ. pdb; gray P400 (13 N-terminal and 16 C-terminal unstructured residues of P400 were removed as they were not modeled).

PS50053) was the profile to which our sequence was aligned using ClustalX (see Materials and methods).

Although there are many structures to be chosen as templates in the β-grasp fold (SCOP) classification, the one from the nearest taxonomic system/organism example is the MoaD/ThiS superfamily, which includes crystallized proteins from the archaea representatives *Pyrococcus furiosus* and *Methanobacterium thermoalcaliphilum*. However, we used a typical, well-characterized representative of the β-grasp fold (1UBQ) as the modeling template because there was no significant improvement either in sequence identity or in structure correctness when other, more poorly defined or less characterized proteins were used. The model obtained assigned 89.3% of the residues to the preferably region, 10.7% to the allow region, 0.0% to the generously allow region, and 0.0% in the disallowed regions in the Ramachandran plot. In addition, no negative values were found using the program verify3D. Also, a deviation of 1.2 Å rmsd (root-mean-square distance) between the model and the template structure was obtained in the superposed zones (aligned/gap positions = 73/25) (see Fig. 1B). The structure in our proposed model has both an N-terminal and a C-terminal tail, with no secondary structure elements of 13 and 16 amino acid residues, respectively. This is a relevant fact when the ubiquitin-like structures available in the PDB database are examined. Moreover, this fold admits enormous variation in secondary structure elements as well as tail and loop boundaries, while preserving the characteristic topology. For example, the N-terminal ubiquitin-like domain of mouse ubiquitin-specific protease 14 (USP14) (pdb code 1WGG) has a 21-residue C-terminal tail; the small ubiquitin-related modifier SUMO-12AWT (pdb code 1A5R) and the ubiqui-

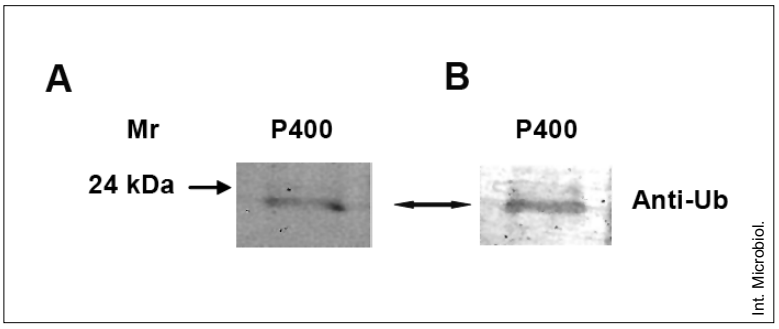

Fig. 2. Renatured recombinant peptide P400 immunodetected by anti-ubiquitin antibody. (**A**) Semi-denaturing SDS-PAGE. (**B**) Western blot assay of P400 using anti-ubiquitin antibody.

tin-like domain in mouse ubiquitin-like protein SB132 (pdb code 1X1M) have N-terminal tails that are 20 and 16 residues long, respectively. Structures with very different loop regions are also common. The novel identified ubiquitin-like domain in the human hypothetical protein FLJ35834 (pdb code 2DAF) is made up of 116 residues and has a C- and N-terminal unstructured region of 16 and 29 residues, respectively (Zhao C, et al. Solution structure of the novel identified ubiquitin-like domain in the human hypothetical protein FLJ35834; to be published [http://www.pdb.org/pdb]).

It should also be noted that many halophilic proteins are longer than their non-halophilic homologs [12]. Thus, ferredoxin 2Fe-2S of *Haloarcula marismortui* has homology with plant ferredoxins, but contains an extension of 30 amino acids in its amino terminus. This extension forms two α-helices with a negative charge that confers more solvation to the protein than is the case for the ferredoxins of non-halophilic organisms. Another halophilic protein with an amino acid extension is the serine protease halolysin, from the archaea *Natrialba asiatica* strain 172P1 and *N. magadii*. This enzyme has a long C-terminal extension of nearly 120 amino acids that is not present in other serine proteases from non-halophilic organisms [8,22]. Together, these considerations led us to conclude that a ubiquitin-like fold can be assigned to P400.

To further confirm the potential relationship between P400 and the ubiquitin superfamily, the P400 DNA fragment was expressed in *E. coli* cells. The recombinant peptide was refolded by several protocols [9] and detected with anti-ubiquitin antibodies. Following dilution of the peptide with either 3 M KCl or 3 M NaCl at 4 and 18ºC, it reacted to a low extent with the anti-ubiquitin antibody, suggesting that its refolding was not complete (not shown). Since the cytoplasm of haloarchaea contains a high concentration of KCl, the peptide was dialyzed against 2 M KCl. As expected, P400 was subsequently detected by anti-ubiquitin antibodies in Western blots (Fig. 2) and in dot blots carried out under native conditions. As seen following SDS-PAGE, P400 displayed a relative mass (24 kDa) higher than its predicted molecular mass (14 kDa), which is a common feature of certain ubiquitin-like proteins and many haloarchaeal proteins [33]. Therefore, this result supports the data gained by bioinformatic analysis suggesting that P400 adopts a ubiquitin-like fold.

The alignment between P400 and representative proteins with a known 3D structure of the ThiS family showed three contiguous identical amino acid residues, Asp53-Gly54-Asp55 (ThiS numbering) (Fig. 3). The yeast protein Urm1, proposed as a molecular fossil of the ubiquitin system, was included in the comparison since it is a homolog in sequence and structure to ThiS [13]. The alignment of P400 with ThiS and Urm1 showed that they shared these three contiguous identical residues. In ThiS and Urm1, these residues are found in a loop connecting the α-helix with the E β-sheet, the two last elements of the proteins' secondary structure. This accessible position, besides its invariability in different life domains, suggests that it plays a major functional role, albeit one that has not yet been described. Moreover, the sequence DGD (Asp-Gly-Asp) inside the BM-motif of the murine phosphoadenosine phosphosulfate (PAPS) synthetase might be necessary to preserve the kinase activity of the enzyme [40].

The analysis reported by Bienkowska et al. [5] included four prokaryotic sequences that have conserved the functional features of ubiquitin (TVN0569, AFO737, MTH1743, and Aq_025a). Thus, the authors proposed that the four proteins are distant homologs to ubiquitin. The alignment between P400 and these proteins showed that they have several common amino acids, and that the sequence Asp-Gly-Asp, shared between P400 and the ThiS family and Urm1, is conserved although with replacement of the first Asp by Glu, a conservative modification (data not shown). Alignments between different ubiquitin-like proteins reported in the literature also identified these shared residues [5,45].

Ubiquitin and proteins belonging to the ubiquitin family, such as ThiS family proteins, have two Gly residues in the carboxyl terminus that are necessary for the activation of these proteins by specific enzymes. These terminal Gly residues are not present in P400 nor have they been found in

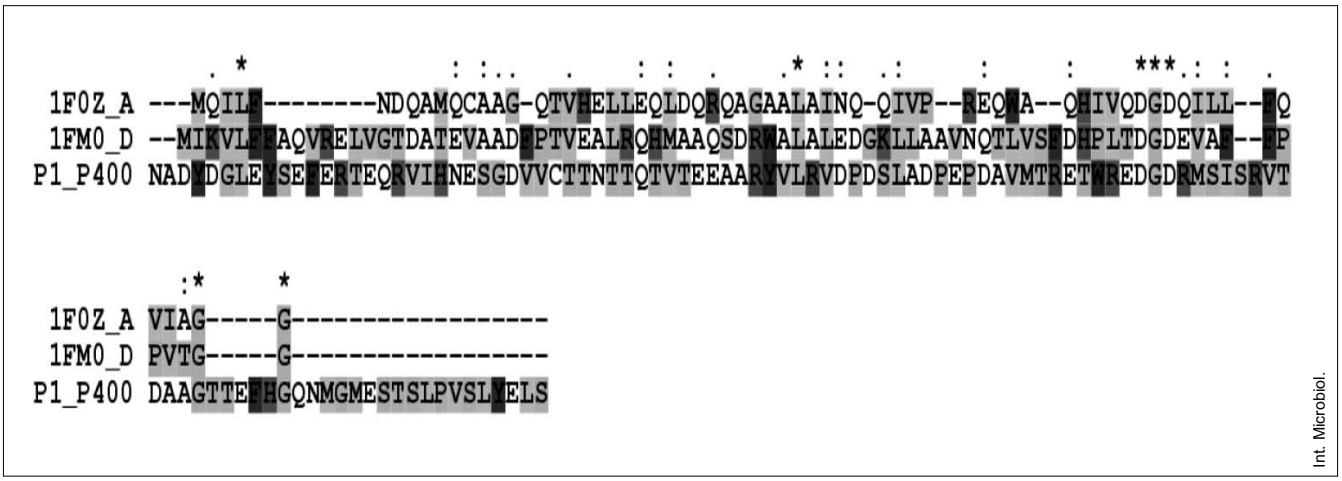

Fig. 3. ClustalW sequences alignment of ThiS, MoaD, and P400. Designations: 1f0z_A, the ThiS protein of *E. coli*, chain A; 1fm0_d, molybdopterin converting factor, subunit 1 of *E. coli*, chain D.

other several ubiquitin homologs, including YukD [45]. (For details, see the Pfam [2] and Homstrad [30] alignments, PF00240 and UBQ families, respectively.)

Direct evidence showing that P400 reacts with anti-ubiquitin antibodies, together with the data from the *in silico* analysis, suggests that P400 is structurally related to ubiquitin-like proteins. Note that the immunoreactive protein bands previously detected in several haloarchaea strains are induced under conditions of accelerated proteolysis, which is a common response of the ubiquitin system for protein conjugation [31]. This fact leads us to speculate that P400 is part of a conjugation system within a proteolytic route, such as that mediated by ubiquitin in eukaryotic organisms. The participation of a ubiquitin-like protein from *Mycobacterium tuberculosis* in the proteasome pathway is in agreement with this hypothesis [33]. Moreover, the *M. tuberculosis* protein not only does not show significant sequence identity with ubiquitin, but makes use of a chemistry distinct from that of ubiquitin to modify proteins.

The presence of a DNA sequence potentially encoding P400 (gi 224819784) was identified in *N. magadii*, whose genomic sequence has been recently determined. This finding encourages our future investigations into the occurrence of P400 in *N. magadii* cells. Indeed, this information will greatly contribute to future experimental and computational work, and thus to advances in identifying and characterizing the P400 protein.

Acknowledgements. This work was supported by grants from the National Scientific and Technical Research Council (CONICET) and the National University of Mar del Plata (UNMdP), Argentina. D.N., C.M.B., R.E.D.C. and R.D.C. are researchers from CONICET. M.V.O. performed the heterologous expression experiments as part of her PhD thesis; she is a doctoral fellow from CONICET. Authors are grateful to Dr. Julie Maupin-Furlow for providing *E. coli* (DE3) Rossetta cells and the expression vector pET24b(+).

References

1. Andreeva A, Howorth D, Brenner SE, Hubbard TJP, Chothia C, Murzin AG (2004) SCOP database in 2004: refinements integrate structure and sequence family data. Nucl Acid Res 32:D226-D229
2. Bateman A, Coin L, Durbin R, et al. (2004) The Pfam protein families database. Nucl Acid Res (database issue) 32:D138-D141
3. Berman HM, Westbrook J, Feng Z, Gilliland G, et al. (2000) The Protein Data Bank. Nucl Acid Res 28:235-242
4. Bidle KA, Kirkland PA, Nannen JL, Maupin-Furlow JA (2008) Proteomic analysis of *Haloferax volcanii* reveals salinity-mediated regulation of the stress response protein PspA. Microbiology 154:1436-1443
5. Bienkowska JR, Hartman H, Smith TF (2003) A search method for homologs of small proteins. Ubiquitin-like proteins in prokaryotic cells? Protein Eng 16:897-904
6. Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT (2005) Protein structure prediction servers at University College London. Nucl Acid Res 33 (Web Server issue):W36-38. DOI: 10.1093/nar/gki41
7. Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server. Nucl Acd Res 36 (Web Server issue):W197-201. DOI: 10.1093/nar/gkn238
8. De Castro RE, Ruiz DM, Giménez MI, Paggi RA, Silveyra MX, Maupin-Furlow JA (2008) Gene cloning and heterologous synthesis of ahaloalkaliphilic extracellular protease of *Natrialba magadii* (Nep). Extremophiles 12:677-687
9. Díaz S, Pérez-Pomares F, Pire C, Ferrer C, Bonete MJ (2006) Gene cloning, heterologous overexpression and optimized refolding of the NAD-glutamate dehydrogenase from *Haloferax mediterranei*. Extremophiles 10:105-115
10. Durner J, Böger P (1995) Ubiquitin in the prokaryote *Anabaena variabilis*. J Biol Chem 270:3720-3725

11. Empadinhas N, da Costa MS (2008) Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. Int Microbiol 11:151-161
12. Frolow F, Harel M, Sussman JL, Mevarech M, Shoham M (1996) Insights into protein adaptation to a saturated salt environment from the crystal structure of halophilic 2Fe-2S ferredoxin. Nature Struct Biol 3:452-458
13. Furukawa K, Mizushima N, Noda T, Ohsumi Y (2000) A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes. J Biol Chem 275:7462-7465
14. Ginalski K, Elofsson A, Fischer D, Rychlewski L (2003) 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics 19:1015-1018
15. Griffiths E, Gupta RS (2007) Identification of signature protein that are distinctive of the Deinococcus-Thermus phylum. Int Microbiol 10:201-207
16. Herrmann J, Lerman LO, Lerman A (2007) Ubiquitin and ubiquitin-like proteins in protein regulation. Circ Res 100:1276-1291
17. Hershko A, Ciechanover A (1992) The ubiquitin system for protein degradation. Annu Rev Biochem 61:761-807
18. Hochstrasser M (2000) Evolution and function of ubiquitin-like protein-conjugation systems. Nature Cell Biol 2:E153-E157
19. Hulo N, Bairoch A, Bulliard V, et al. (2008) The 20 years of PROSITE. Nucl Acid Res 36 (Database issue):D245-249
20. Iyer LM, Burroughs AM, Aravind L (2006) The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like β-grasp domains. Genome Biol 7:R60
21. Jenney Jr FE, Adams MWW (2008) The impact of extremophiles on structural genomics (and vice versa). Extremophiles 12:39-50
22. Kamekura M, Seno Y, Holmes ML, Dyall-Smith ML (1992) Molecular cloning and sequencing of the gene for a halophilic alkaline serine protease (halolysin) from an unidentified halophilic archaea strain (172P1) and expression of the gene in *Haloferax volcanii*. J Bacteriol 174:736-742
23. Konstantinidis K, Tebbe A, Klein C, et al. (2007) Genome-wide proteomics of *Natronomonas pharaonis*. J Proteome Res 6:185-193
24. Larkin MA, Blackshields G, Brown NP, et al. (2007) Clustal W and Clustal X version 2.0 Bioinformatics 23:2947-2948
25. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283-291
26. Laskowski RA, Chistyakov VV, Thornton JM (2005) PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. Nucl Acid Res (Database issue) 33:D266-D268
27. Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356:83-85
28. 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
29. Michnick SW, Shakhnovich E (1998) A strategy for detecting the conservation of folding-nucleus residues in protein superfamilies. Fold Des 3:239-251
30. Mizuguchi K, Deane CM, Blundell TL, Overington JP (1998) HOMSTRAD: a database of protein structure alignments for homologous families. Protein Sci 7:2469-2471
31. Nercessian D, De Castro RE, Conde RD (2002) Ubiquitin-like proteins in halobacteria. J Basic Microbiol 42:277-283
32. Ng WL, Yang CF, Halladay JT, Arora P, DasSarma S (1995) Isolation of genomic and plasmid DNAs from *Halobacterium halobium*. In: DasSarma S, Fleischmann EM (eds) Archaea: a laboratory manual. Halophiles. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 179-184
33. Pearce MJ, Mintseris J, Ferreyra J, Gygi SP, Darwin KH (2008) Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. Science 322:1104
34. Pearl FMG, Lee D, Bray JE, et al. (2000) Assigning genomic sequences to CATH. Nucl Acid Res 28:277-282
35. Rost B (1996) PHD: predicting one-dimensional protein structure by profile-based neural networks. Methods Enzymol 266:525-539
36. 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
37. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779-815
38. Shean BS, Mykles DL (1995) Polyubiquitin in crustacean striated muscle: increased expression and conjugation during molt-induced claw muscle atrophy. Biochim Biophys Acta 1264: 312-322
39. Shindyalov IN, Bourne PE (2001) A database and tools for 3-D protein structure comparison and alignment using the Combinatorial Extension (CE) algorithm. Nucl Acid Res 29:228-229
40. Singh B, Schwartz NB (2003) Identification and functional characterization of the novel BM-motif in the murine phosphoadenosine phosphosulfate (PAPS) synthetase. J Biol Chem 278:71-75
41. Smith PK, Krohn RI, Hermanson GT, et al. (1987) Measurment of protein using bicinchoninic acid. Anal Biochem 150:76-85
42. Tindall BJ, Ross HN, Grant WD (1984) *Natronobacterium* gen. nov. and *Natronococcus* gen. nov., two new genera of the haloalkaliphilic archaebacteria. System Appl Microbiol 5:41-57
43. Todd AE, Marsden RL, Thornton JM, Orengo CA (2005) Progress of structural genomics initiatives: an analysis of solved target structures. J Biol Mol 348:1235-1260
44. Van PT, Schmid AK, King NL, et al. (2008) *Halobacterium salinarum* NRC-1 PeptideAtlas: toward strategies for targeted proteomics and improved proteome coverage. J Proteome Res 7:3755-3764
45. van den Ent F, Löwe J (2005) Crystal structure of the ubiquitin-like protein YukD from *Bacillus subtilis*. FEBS Letters 579:3837-3841
46. Vriend G, Sander C (1991) Detection of common three-dimensional substructures in proteins. Proteins 11:52-58
47. Wada K, Wada Y, Ishibashi T, Gojobori T, Ikemura T (1992) Codon usage tabulates from the GenBank genetic sequence data. Nucl Acid Res 20:2111-2118
48. 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
49. Wolf S, Lottspeich F, Baumeister W (1993) Ubiquitin found in the archaebacterium *Thermoplasma acidophilum* FEBS Lett 326:42-44