# Crystal structure of ubiquitin-like small archaeal modifier protein 1 (SAMP1) from *Haloferax volcanii* ☆

Young Jee Jeong [1], Byung-Cheon Jeong [1], Hyun Kyu Song *

*School of Life Sciences and Biotechnology, Korea University, Seoul 136-701, South Korea*

ABSTRACT

The ubiquitin-like (Ubl) system has been shown to be ubiquitous in all three kingdoms of life following the very recent characterization of ubiquitin-like small archaeal modifier proteins (SAMP1 and 2) from *Haloferax volcanii*. The ubiquitin (Ub) and Ubl molecules in eukaryotes have been studied extensively and their cellular functions are well established. Biochemical and structural data pertaining to prokaryotic Ubl protein (Pup) continue to be reported. In contrast to eukaryotes and prokaryotes, no structural information on the archaeal Ubl molecule is available. Here we determined the crystal structure of SAMP1 at 1.55 Å resolution and generated a model of SAMP2. These were then compared with other Ubl molecules from eukaryotes as well as prokaryotes. The structure of SAMP1 shows a β-grasp fold of Ub, suggesting that the archaeal Ubl molecule is more closely related to eukaryotic Ub and Ubls than to its prokaryotic counterpart. The current structure identifies the location of critical elements such a single lysine residue (Lys4), C-terminal di-glycine motif, hydrophobic patches near leucine 60, and uniquely inserted α-helical segments (α1 and α3) in SAMP1. Based on the structure of SAMP1, several Ub-like features of SAMPs such as poly-SAMPylation and non-covalent interactions have been proposed, which should provide the basis for further investigations concerning the molecular function of archaeal Ubls and the large super-family of β-grasp fold proteins in the archaeal kingdom.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

Post-translational modification by small protein ubiquitin (Ub) and ubiquitin-like (Ubl) proteins is involved in a wide variety of cellular processes [1,2]. Ub and Ubls are small proteins that are well conserved in terms of primary and tertiary structure, and are usually covalently attached to target proteins using their C-terminal glycine residue [1,3]. Interestingly, it was long believed that the Ub system was only present in eukaryotes, although Ubl-protein conjugation is thought to have evolved from prokaryotic sulfur transferase systems [4,5], and other components involved in the proteasomal degradation pathway including proteasome homologs and similar degradation signals such as N-degron have

also been identified in prokaryotes [6–9]. A prokaryotic Ubl protein (Pup) from *Mycobacterium tuberculosis* was eventually identified and shown to act as a degradation signal for the proteasome [10,11]. Ubiquitin-like small archaeal modifier proteins 1 and 2 (SAMP1 and SAMP2) from the archaeon *Haloferax volcanii* have only recently been biochemically characterized [12], although the existence of Ubl proteins in archaea had been suggested by previous studies [13,14]. Therefore, the Ub and Ubl modification systems are now truly ubiquitous in all three kingdoms of life and the aforementioned reports finally fill the gap in our knowledge concerning the evolutionary pathway of Ub and Ubl proteins.

Intriguingly, although Pup has a name implying a prokaryotic version of ubiquitin, it does not possess a β-grasp fold of Ub, but has an intrinsically disordered structure [10,11]. This unstructured state forms an ordered α-helix upon complex formation with the tentacle-like coiled coil domain in *Mycobacterium* proteasomal ATPase [15]. Furthermore, there is a di-glycine motif at the C-terminal region of Pup, although covalent modification is not directly mediated by a glycine residue from this motif [10,11]. Instead, "PUPylation" is mediated by a glutamic acid residue at the C-terminus. Therefore, Pup differs from ubiquitin both structurally and biochemically [11]. However, recent analyses of archaea have suggested that there are a large number of proteins sharing a β-grasp fold with a di-glycine motif at the C-terminus [12]. Although no

experimental data on "SAMPylation" enzymes, which are equivalent to ubiquitin E1-activing, E2-conjugating and E3-ligating enzymes, have been reported, a similar enzymatic cascade has been predicted [16]. In addition to ubiquitin-equivalent SAMP1 and 2, a homolog of eukaryotic E1-activating enzyme, MoeB, has been suggested as a potential E1 candidate for SAMPylation, though homologs of eukaryotic E2 and E3 enzymes have not been identified so far.

Although knowledge about eukaryotic ubiquitin-mediated pathways and E1–E2–E3 enzymes has been firmly established by a vast amount of information [3,17,18], and the mechanism of PUPylation in bacteria has begun to be understood as evidenced by recent serial reports [10,15,19–21], the mechanism of SAMPylation in archaea remains largely unknown. In an effort to improve our understanding of archaeal SAMPylation mechanisms, structural information relating to each reaction component should furnish valuable insights. As a first step, we have determined a high resolution structure of SAMP1 and compared the structure with other ubiquitin-like molecules from eukaryotes and prokaryotes. SAMP1 shows structural similarity with eukaryotic Ub and Ubls, but differs radically from Pup. A single lysine residue and di-glycine motif at the C-terminus are spatially isolated and form a poly-SAMPylated chain. A hydrophobic surface patch near Leu60, which is structurally equivalent to that found near Ile44 of Ub, is also present, suggesting that several Ub-like features such as poly-SAMPylation and non-covalent interactions might be conserved in the archaeal kingdom.

## 2. Materials and methods

### 2.1. Sample preparation

The codon-optimized full-length *samp1* gene was synthesized (Mr. Gene GmbH) and amplified by the polymerase chain reaction (PCR). The PCR product was cloned into pRSF-GST vector using the BamHI and XhoI restriction enzyme sites. The resultant plasmid for expressing GST-tagged SAMP1 was transformed into BL21(DE3) RIL cells, and was confirmed by DNA sequencing. Transformed cells were cultured in LB medium containing 50 μm/ml kanamycin at 37 °C until reaching an $OD_{600\ nm}$ of 0.8, and then induced by the addition of 0.5 mM IPTG and incubated at 18 °C for 20 h. Cells were harvested by centrifugation, resuspended in phosphate-buffered saline containing 2 mM DTT, and subsequently disrupted by ultrasonication. SAMP1 was purified by GST affinity chromatography (GE Healthcare) and treated with protease TEV at a 1:40 (w/w) ratio at 22 °C overnight in order to cleave the GST-tag. Further purification was carried out by successive anion exchange (HiTrap Q Fast Flow, GE Healthcare) and size exclusion (Hiload 16/60 Superdex 75, GE Healthcare) chromatography. Eluents from columns were analyzed by SDS–PAGE. Final protein was concentrated to 50 mg/ml in 20 mM HEPES (pH 7.5), 150 mM NaCl and 2 mM DTT.

### 2.2. Crystallization

Protein solution was mixed with the same volume of reservoir solution containing 100 mM HEPES (pH 7.5), 50 mM cadmium sulfate hydrate and 1.0 M sodium acetate trihydrate. Thin and multiple needle-shaped crystals grew within a day using a hanging-drop vapor-diffusion method at 22 °C. The best crystals were grown in 100 mM HEPES (pH 7.3), 50 mM cadmium sulfate hydrate and 1.1 M sodium acetate trihydrate. Since there is no methionine in SAMP1, leucine 80 was mutated to methionine for multi-wavelength anomalous dispersion (MAD) phasing [22], and the mutated plasmid was transformed into B834(DE3) cells. Selenomethionyl L80M mutant SAMP1 was expressed in M9 media in the presence of 40 mg/ml selenomethionine (SeMet). Subsequent induction, purification and crystallization were performed in the same manner employed for wild-type SAMP1 protein. Crystals of both native and selenomethionyl protein were transferred into cryoprotectant solution containing reservoir solution with 20% (w/v) glycerol and then flash-frozen in a cold nitrogen stream at 100 K.

### 2.3. Structure determination

Native and SeMet diffraction data were collected on an ADSC quantum charge-coupled device detector at the 6C beamline of the Pohang Light Source, Korea and the AR-NE3 beamline of the Photon Factory, Tsukuba, Japan, respectively. Crystals of SAMP1 diffracted to an approximate resolution of 1.4 Å (Table 1). The diffraction data were processed and scaled using the HKL2000 software package [23] and statistics for the data collection are described in Table 1. The structure of SeMet SAMP1 was solved with a four-wavelength MAD data set using SOLVE as implemented in the PHENIX software suite [24]. Two selenium sites in the asymmetric unit were identified and the resultant phases had a figure-of-merit of 0.342. Subsequent density modification using RESOLVE improved the phases greatly to yield a figure-of-merit of 0.708 at 2.0 Å resolution. An initial model was built with ARP/wARP [25] and the model was rebuilt using the program COOT [26]. Solvent molecules were added using model-phased difference-Fourier maps [27]. The SAMP1 model was refined with CNS including the bulk solvent correction [27]. The assessment of model geometry and assigned secondary structural elements was achieved using the PROCHECK program [28]. Statistics for the refinement and model quality are shown in Table 1. The modeled structure of SAMP2 was obtained using the ModWeb server (http://modbase.compbio.ucsf.edu/Mod-Web20-html/modweb.html). For structure comparison, the DALI server was used (http://ekhidna.biocenter.helsinki.fi/dali_server). All figures were generated using PyMOL (http://www.pymol.org).

## 3. Results and discussion

SAMP1 contains no methionine residues except for the first N-terminal residue. Therefore, a hydrophobic residue, leucine 80, was mutated to methionine and then the structure of SAMP1 was determined by MAD phasing. The wild-type structure was also determined and basically no difference was observed between the wild-type and L80M mutant structures. The structure of wild-type SAMP1 refined at 1.55 Å resolution shows a globular compact domain comprising three α-helices, four β-strands and connecting loops, and a protruding C-terminal di-glycine motif (Fig. 1A). The entire 88 residues beginning from the two additional N-terminal residues (Gly-Ser) from the expression vector (the first methionine was removed during vector construction) to the C-terminal glycine residue (Gly87) were observed in the electron density map, although residues between 82 and 87 containing a critical di-glycine motif demonstrated relatively weak electron density, indicating that this C-terminal segment possesses high flexibility. It is reasonable to suppose that this flexible C-terminal di-glycine motif is extended far from the core of the molecule to allow for covalent attachment to a lysine residue at the surface of target proteins (Fig. 1A). The electrostatic potential surface of SAMP1 shows a high negative charge, except for some positively charged patches near Lys4 (Fig. 1B), and several hydrophobic residues, which might be important for interactions with molecular surfaces of other molecules (Fig. 1C).

A recent study predicted that SAMP1 belongs to the β-grasp fold Ub family [12], in contrast to prokaryotic Pup [29,30]. This needs to be confirmed since SAMP1 shares very limited sequence identity

**Table 1**
Data collection, phasing and refinement statistics.

| | Native | MAD (L80M mutant) | | | |
| --- | --- | --- | --- | --- | --- |
| | | Peak | Edge | RemoteH | RemoteL |
| *Data collection* | | | | | |
| Space group | $P2_12_12_1$ | | $P2_12_12_1$ | | |
| Cell dimension | | | | | |
| $a, b, c$ (Å) | | | 41.917, 42.298, 43.443 | | |
| X-ray sources[a] | PAL 6C | | PF AR-NE3 | | |
| Resolution (Å)[b] | 1.55 (1.61–1.55) | | 1.38 (1.43–1.38) | | |
| Wavelength (Å) | 1.23985 | 0.97951 | 0.97973 | 0.95000 | 0.98361 |
| Total reflections | 150,278 | 225,797 | 225,608 | 225,807 | 222,655 |
| Unique reflections | 11,712 | 16,027 | 16,044 | 16,117 | 16,025 |
| Completeness (%)[b] | 99.8 (98.8) | 97.2 (91.8) | 97.3 (93.2) | 97.4 (92.6) | 97.0 (90.0) |
| Overall ($I/\sigma I$)[b] | 30.0 (4.6) | 69.4 (9.3) | 67.7 (7.6) | 65.8 (6.2) | 65.0 (4.9) |
| $R_{sym}$ (%)[c,b] | 6.4 (39.1) | 6.7 (24.2) | 6.4 (26.4) | 6.3 (29.8) | 6.0 (33.6) |
| FOM[d] | | 0.342/0.708 (SOLVE/RESOLVE) | | | |
| *Refinement* | | | | | |
| Resolution range (Å) | 50.0–1.55 | | | | |
| Reflections used | 11,458 | | | | |
| $R_{work}/R_{free}$ (%)[e] | 19.9/24.3 | | | | |
| Number of atoms | | | | | |
| Proteins | 639 | | | | |
| Water | 90 | | | | |
| Ions | 2 ($Cd^{2+}$), 2 ($Mg^{2+}$), 4 (acetate) | | | | |
| RMS deviations | | | | | |
| Bond length (Å) | 0.006 | | | | |
| Bond angles (°) | 0.964 | | | | |
| Ramachandran outlier | 0.0% (0) | | | | |
| PDB ID | 3PO0 | | | | |

[a] PAL, Pohang Accelerator Laboratory, Korea; PF, Photon Factory, Japan.
[b] Values in parentheses are for reflections in the highest resolution bin.
[c] $R_{sym} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the *i*th measurement of *h* and $\langle I(h)\rangle$ is the corresponding average value for all *i* measurements.
[d] Figure of merit = $|\sum P(\alpha)e^{i\alpha}/\sum P(\alpha)|$, where $P(\alpha)$ is the phase probability distribution and $\alpha$ is the phase.
[e] $R_{work}$ and $R_{free} = \sum ||F_O| - |F_C||/\sum |F_O|$ for the working set and test set (10%) of reflections.

with Ub and Ubl molecules (approximately 20% sequence identity allowing many sequence gaps). The Z-scores as determined by the DALI server for SAMP1 are 14.4, 12.3, 12.3, 12.2 and 8.2 for MoaD-related protein (PDB ID: 1v8c), molybdopterin converting factor (PDB ID: 1vjk), Urm1 (PDB ID: 2pko), cysteine synthase B (PDB ID: 3dwg) and Ub (PDB ID: 1ubi), respectively. When approximately 70–80 matching Cα atoms of SAMP1 were superposed with equivalent atoms of the above proteins, the root mean square (RMS) deviation was in the range of 1.5–2.5 Å. As shown in Fig. 2, SAMP1 and Ubl molecules are structurally conserved. It is intriguing that SAMP1 shows higher structural similarity with molecules involved in sulfur transfer reactions, which are regarded as ancient versions of ubiquitin [5,31]. Recent analyses suggested that there are a large number of proteins that possess a β-grasp fold with a di-glycine motif at the C-terminus in archaea [12], and these have been divided into two major groups comprising the ThiS and MoaD families [32]. The two ubiquitin-like molecules from *H. volcanii*, SAMP1 and SAMP2, which share limited sequence identity of around 20% with different chain lengths (87 *vs.* 66 amino acid residues; Fig. 3A), also belong to different families [32]. Our structural results showing that the molecule most structurally similar to SAMP1 is MoaD-related protein from *Thermus thermophilus* HB8 (1v8c) experimentally confirms recent bioinformatic analysis [32]. However, the bioinformatic study referred to classified SAMP2 in the ThiS family. Unfortunately, we were unable to obtain crystals of SAMP2 to initiate structural investigations. Instead, we performed a modeling experiment with the current structure and other homologous proteins present in the Protein Data Bank. The best template found for the final SAMP2 model was the uncharacterized PF1061 protein from *Pyrococcus furiosus* [33], and the SAMP2 model is shown in Fig. 3B (right). A structure-based alignment between SAMP1 and SAMP2 is also shown in Fig. 3A. In contrast to SAMP1, the structure of SAMP2 is more similar to that of Ub (Fig 3B, middle). The main structural difference between the two SAMPs relates to the presence of two additional α-helices (α1 and α3) in SAMP1 (Fig. 3B). Interestingly, these helices are also found in protein molecules with Z-scores of over 10, but not in Ub (Fig. 2). Therefore, it is tempting to speculate that additional helical segments in MoaD family proteins play a specialized role relating to their cellular function.

Interestingly, there is only one lysine residue in SAMPs (Lys4 in SAMP1 and Lys58 in SAMP2), whereas seven lysine residues (Lys6, 11, 27, 29, 33, 48 and 63) are present in Ub, which are involved in diverse cellular processes [34]. Previous results concerning the accumulation of SAMP1 and 2 in proteasomal subunit deletion mutants suggests that both SAMPs are involved in the proteasomal degradation pathway [12]. However, the lysine residue at the N-terminal region of SAMP1 is not located in a spatially equivalent position to Ub Lys48, but is closely located to Ub Lys6. Although covalently modifying SAMP2 Lys58 is also not equivalent to Ub Lys48 structurally, the direction of the lysyl side chain is rather similar to SAMP1 Lys4, and Ub Lys6 (Fig. 3B). In fact, the directions of Lys6 and Lys48 in the β-grasp fold of Ub differ, suggesting that the poly-SAMPylated chain with both SAMPs might display marked structurally differences compared with the poly-ubiquitin chain linked by Lys48. However, the poly-ubiquitin chain mediated by all lysine residues in Ub except for Lys63 may act as a degradation signal. Residue Lys6 of Ub is the most readily modified lysine residue in Ub and its modification is known to inhibit proteasomal degradation [35]. The poly-ubiquitylation of paired helical filament-Tau is mediated through Lys6 and Lys11 of Ub [36], and BRCA1–BARD1 ubiquitin ligase mediates Lys6-linked poly-Ub chains [37]. Similarly, poly-SAMPylation must occur easily with solvent accessible lysine residues (Lys4 and Lys58 of SAMP1 and 2) and the far extended conformation of the C-terminal di-glycine motif (Fig. 3B).
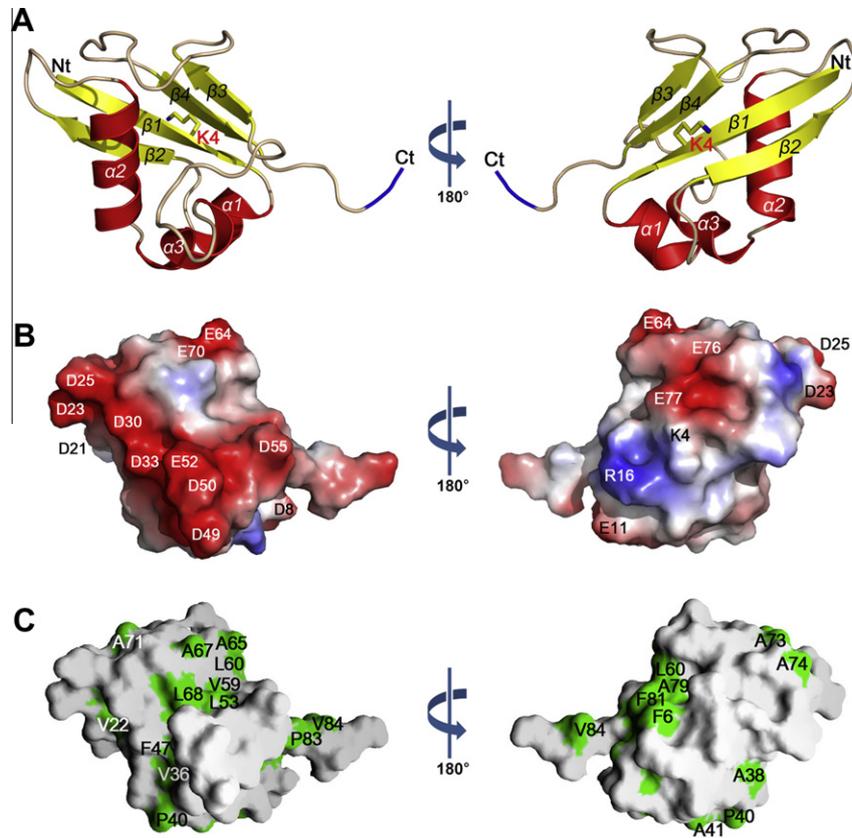
**Fig. 1.** Overall structure of SAMP1. (A) Ribbon diagram showing the overall structure of SAMP1 and on the right is a 180° rotation along the vertical axis as indicated. With the exception of the C-terminal di-glycine residues colored in blue, α-helices, β-sheets and connecting loops of SAMP1 are colored, red, yellow and pink, respectively. The secondary structural elements are sequentially labeled, and the side chain of the only lysine residue (Lys4) is shown and labeled. The N- and C-termini of SAMP1 are also indicated. (B) Electrostatic potential surface of SAMP1 viewed the same as in panel (A). Positive and negative electrostatic potentials are colored blue and red, respectively. (C) Hydrophobic surface of SAMP1 viewed in the same orientation. Residues forming hydrophobic surfaces are colored green and labeled. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
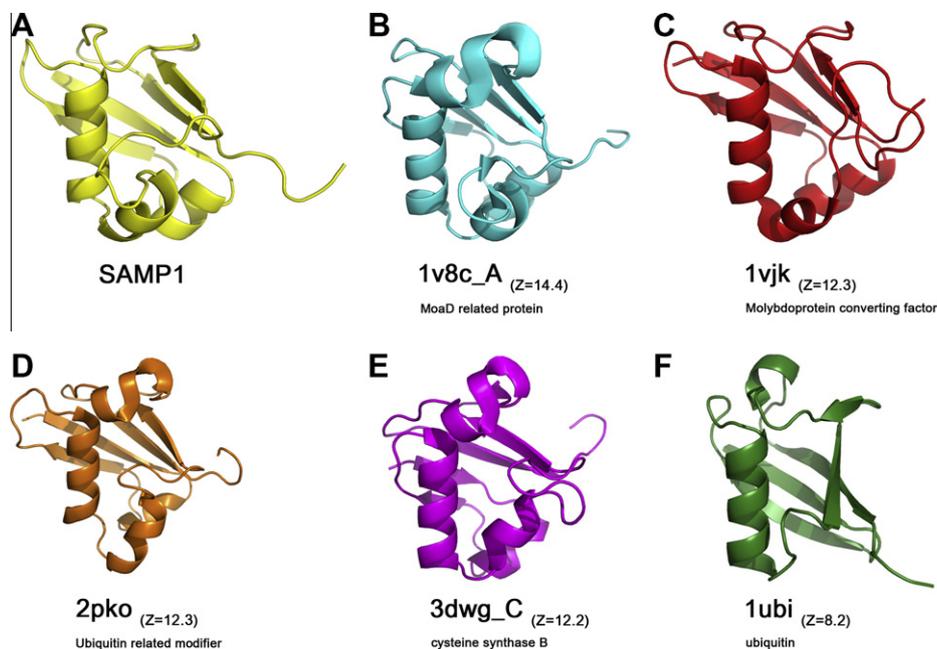


**Fig. 2.** Ribbon diagram comparing the overall structures of (A) SAMP1, (B) MoaD-related protein, (C) molybdopterin converting factor, (D) ubiquitin related modifier (Urm), (E) cysteine synthase B, and (F) ubiquitin. The view is the same as in Fig. 1A (right). PDB ID codes and Z-scores from DALI server for each structure are provided.
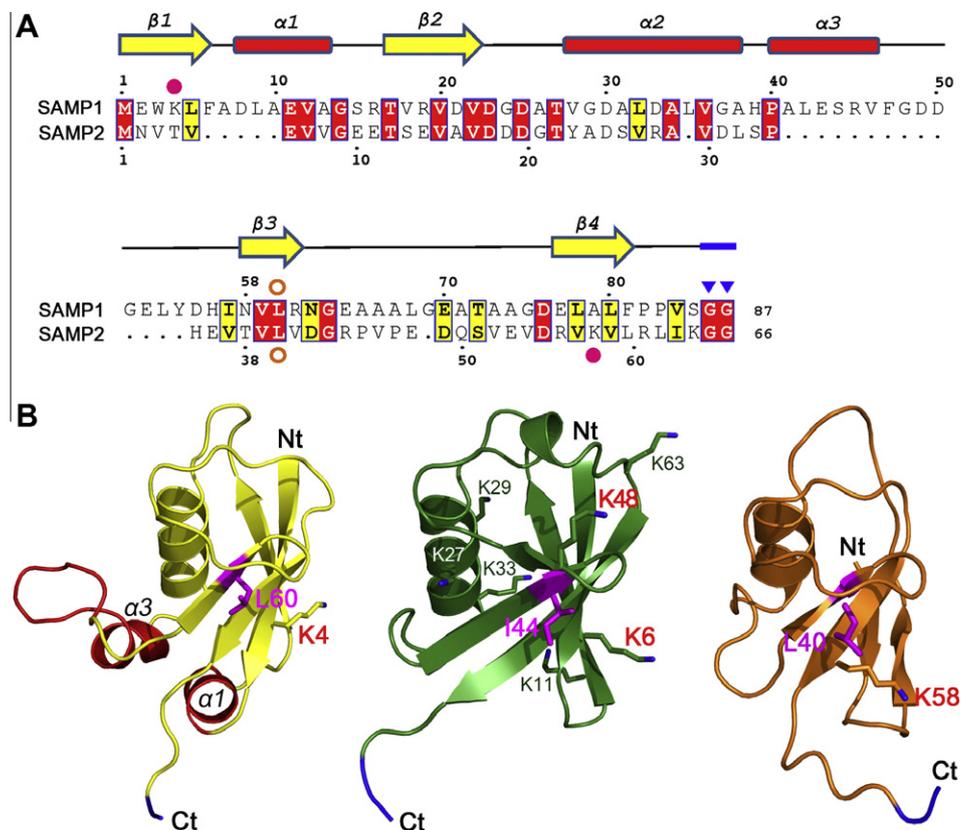
**Fig. 3.** (A) Structure-based sequence alignment between SAMP1 and SAMP2 from *H. volcanii*. The secondary structural elements at the top of each alignment correspond to those of SAMP1. Strictly conserved residues boxed in red and yellow indicate conservatively substituted residues. Residue Lys4 of SAMP1 and Lys58 of SAMP2 are marked by a filled red circle. Residue Leu60 of SAMP1 and Leu40 of SAMP2 are marked by an open orange circle. Di-glycine motif at C-terminus is marked by filled blue triangles. (B) Structural comparison between SAMP1 (left), ubiquitin (middle), and SAMP2 (right). The structures are oriented to show Ile44 of ubiquitin at the front center. The common C-terminal di-glycine motif is colored in blue and the unique α-helical addition in SAMP1 is highlighted in red and labeled. Lysine residues that might be involved in potential SAMPylation and ubiquitylation are shown as stick models (Lys4 in SAMP1; Lys6, 48, 63 in Ub; Lys58 in SAMP2). A critical ubiquitin residue (Ub Ile44) involved in non-covalent interactions with interacting proteins and its equivalent SAMPs' residues (SAMP1 Leu60; SAMP2 Leu40) are also shown as stick models and colored in pink. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

It is known that ubiquitin possesses a hydrophobic patch which includes Ile44, and is critical for protein-protein interaction [38]. Our structural study suggests that the SAMP1 and SAMP2 residues spatially equivalent to Ub Ile44 are Leu60 and Leu40, respectively. Amino acid residues located near Ub Ile44 are $_{43}$Leu-Ile-Phe-Ala-Gly$_{47}$, and those located near SAMP1 Leu60 and SAMP2 Leu40 are $_{59}$Val-Leu-Arg-Asn-Gly$_{63}$ and $_{39}$Val-Leu-Val-Asp-Gly$_{43}$, respectively (Fig. 3A). Thus, it can be supposed that the hydrophobic patch found on the molecular surface of SAMP1 (and SAMP2) which is similar to Ub might also be involved in non-covalent interactions with other binding proteins in *H. volcanii*. Indeed, there are several hydrophobic residues (Phe6, Ala79 and Phe81) near SAMP1 Leu60, and these form an exposed hydrophobic patch (Fig. 1C).

This study identified structurally unique characteristics of SAMP1, including its similarities to Ub and Ubl, and the manner by which it differs from SAMP2. Due to technical reasons, the poly-SAMPylation site of SAMP2 has only been identified, however, previous predictions and the current structural data suggest that poly-SAMPylation of SAMP1 might also be feasible due to the spatial orientation of Lys4 and the di-glycine motif. More importantly, it is tempting to speculate that non-covalent interactions between SAMPs and binding partners might play a significant role in archaea. Eukaryotic Ub contains seven lysine residues, which are involved in a variety of cellular functions and participate in different types of modifications such as mono-, multi- and poly-ubiquitylation [34]. Non-covalent interactions between Ub and Ub-binding domains also play a critical role in many cellular pathways [39].

Whether structurally similar SAMPs in archaea also participate in similar non-covalent interactions with binding partners and the functional consequences of such interaction remain unclear. The current structure of SAMP1 provides detailed information that should assist in delineating the molecular function of archaeal Ubls in addition to the large super-family of β-grasp fold proteins.

**Note added in proof**

The solution structure of SAMP1 from Methanosarcina acetivorans has been published very recently [40].

**References**

[1] M. Hochstrasser, Origin and function of ubiquitin-like proteins, Nature 458 (2009) 422–429.

[2] A. Varshavsky, Discovery of cellular regulation by protein degradation, J. Biol. Chem. 283 (2008) 34469–34489.

[3] J.M. Winget, T. Mayor, The diversity of ubiquitin recognition: hot spots and varied specificity, Mol. Cell 38 (2010) 627–635.

[4] M.J. Rudolph, M.M. Wuebbens, O. Turque, K.V. Rajagopalan, H. Schindelin, Structural studies of molybdopterin synthase provide insights into its catalytic mechanism, J. Biol. Chem. 278 (2003) 14514–14522.

[5] M.J. Rudolph, M.M. Wuebbens, K.V. Rajagopalan, H. Schindelin, Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation, Nat. Struct. Biol. 8 (2001) 42–46.

[6] M. Bochtler, L. Ditzel, M. Groll, C. Hartmann, R. Huber, The proteasome, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 295–317.

[7] M. Groll, M. Bochtler, H. Brandstetter, T. Clausen, R. Huber, Molecular machines for protein degradation, Chembiochem 6 (2005) 222–256.

[8] A. Varshavsky, The N-end rule: functions, mysteries, uses, Proc. Natl. Acad. Sci. USA 93 (1996) 12142–12149.

[9] A. Varshavsky, The N-end rule at atomic resolution, Nat. Struct. Mol. Biol. 15 (2008) 1238–1240.

[10] M.J. Pearce, J. Mintseris, J. Ferreyra, S.P. Gygi, K.H. Darwin, Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*, Science 322 (2008) 1104–1107.

[11] K.H. Darwin, Prokaryotic ubiquitin-like protein (Pup), proteasomes and pathogenesis, Nat. Rev. Microbiol. 7 (2009) 485–491.

[12] M.A. Humbard, H.V. Miranda, J.M. Lim, D.J. Krause, J.R. Pritz, G. Zhou, S. Chen, L. Wells, J.A. Maupin-Furlow, Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*, Nature 463 (2010) 54–60.

[13] S. Wolf, F. Lottspeich, W. Baumeister, Ubiquitin found in the archaebacterium *Thermoplasma acidophilum*, FEBS Lett. 326 (1993) 42–44.

[14] D. Nercessian, R.E. de Castro, R.D. Conde, Ubiquitin-like proteins in halobacteria, J. Basic Microbiol. 42 (2002) 277–283.

[15] T. Wang, K.H. Darwin, H. Li, Binding-induced folding of prokaryotic ubiquitin-like protein on the *Mycobacterium* proteasomal ATPase targets substrates for degradation, Nat. Struct. Mol. Biol. 17 (2010) 1352–1357.

[16] K.H. Darwin, K. Hofmann, SAMPyling proteins in archaea, Trends Biochem. Sci. 35 (2010) 348–351.

[17] R.J. Deshaies, C.A. Joazeiro, RING domain E3 ubiquitin ligases, Annu. Rev. Biochem. 78 (2009) 399–434.

[18] W.S. Choi, B.C. Jeong, Y.J. Joo, M.R. Lee, J. Kim, M.J. Eck, H.K. Song, Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases, Nat. Struct. Mol. Biol. 17 (2010) 1175–1181.

[19] K.E. Burns, F.A. Cerda-Maira, T. Wang, H. Li, W.R. Bishai, K.H. Darwin, "Depupylation" of prokaryotic ubiquitin-like protein from mycobacterial proteasome substrates, Mol. Cell 39 (2010) 821–827.

[20] F. Striebel, F. Imkamp, M. Sutter, M. Steiner, A. Mamedov, E. Weber-Ban, Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes, Nat. Struct. Mol. Biol. 16 (2009) 647–651.

[21] M. Sutter, F. Striebel, F.F. Damberger, F.H. Allain, E. Weber-Ban, A distinct structural region of the prokaryotic ubiquitin-like protein (Pup) is recognized by the N-terminal domain of the proteasomal ATPase Mpa, FEBS Lett. 583 (2009) 3151–3157.

[22] E.Y. Park, B.G. Lee, S.B. Hong, H.W. Kim, H. Jeon, H.K. Song, Structural basis of SspB-tail recognition by the zinc binding domain of ClpX, J. Mol. Biol. 367 (2007) 514–526.

[23] Z. Otwinoski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.

[24] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221.

[25] S.X. Cohen, M. Ben Jelloul, F. Long, A. Vagin, P. Knipscheer, J. Lebbink, T.K. Sixma, V.S. Lamzin, G.N. Murshudov, A. Perrakis, ARP/wARP and molecular replacement: the next generation, Acta Crystallogr. D Biol. Crystallogr. 64 (2008) 49–60.

[26] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132.

[27] A.T. Brunger, Version 1.2 of the crystallography and NMR system, Nat. Protoc. 2 (2007) 2728–2733.

[28] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck – a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.

[29] S. Liao, Q. Shang, X. Zhang, J. Zhang, C. Xu, X. Tu, Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein, Biochem. J. 422 (2009) 207–215.

[30] X. Chen, W.C. Solomon, Y. Kang, F. Cerda-Maira, K.H. Darwin, K.J. Walters, Prokaryotic ubiquitin-like protein pup is intrinsically disordered, J. Mol. Biol. 392 (2009) 208–217.

[31] C. Wang, J. Xi, T.P. Begley, L.K. Nicholson, Solution structure of ThiS and implications for the evolutionary roots of ubiquitin, Nat. Struct. Biol. 8 (2001) 47–51.

[32] K.S. Makarova, E.V. Koonin, Archaeal ubiquitin-like proteins: functional versatility and putative ancestral involvement in tRNA modification revealed by comparative genomic analysis, Archaea 2010.

[33] H. Valafar, K.L. Mayer, C.M. Bougault, P.D. LeBlond, F.E. Jenney Jr., P.S. Brereton, M.W. Adams, J.H. Prestegard, Backbone solution structures of proteins using residual dipolar couplings: application to a novel structural genomics target, J. Struct. Func. Genom. 5 (2004) 241–254.

[34] F. Ikeda, I. Dikic, Atypical ubiquitin chains: new molecular signals. 'Protein modifications: beyond the usual suspects' review series, EMBO Rep. 9 (2008) 536–542.

[35] F. Shang, G. Deng, Q. Liu, W. Guo, A.L. Haas, B. Crosas, D. Finley, A. Taylor, Lys6-modified ubiquitin inhibits ubiquitin-dependent protein degradation, J. Biol. Chem. 280 (2005) 20365–20374.

[36] D. Cripps, S.N. Thomas, Y. Jeng, F. Yang, P. Davies, A.J. Yang, Alzheimer disease-specific conformation of hyperphosphorylated paired helical filament-Tau is polyubiquitinated through Lys-48, Lys-11, and Lys-6 ubiquitin conjugation, J. Biol. Chem. 281 (2006) 10825–10838.

[37] H. Nishikawa, S. Ooka, K. Sato, K. Arima, J. Okamoto, R.E. Klevit, M. Fukuda, T. Ohta, Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1–BARD1 ubiquitin ligase, J. Biol. Chem. 279 (2004) 3916–3924.

[38] C. Raiborg, T. Slagsvold, H. Stenmark, A new side to ubiquitin, Trends Biochem. Sci. 31 (2006) 541–544.

[39] L. Hicke, H.L. Schubert, C.P. Hill, Ubiquitin-binding domains, Nat. Rev. Mol. Cell Biol. 6 (2005) 610–621.

[40] F.F. Ranjan, M. Damberger, F.H. Sutter, E. Allain, Weber-Ban, Solution structure and activation mechanism of ubiquitin-like small archaeal modifier proteins, J. Mol. Biol. 405 (2011) 1040–1055.