SCIENTIFIC REPORTS

Ionic strength-dependent conformations of a ubiquitin-like small archaeal modifier protein (SAMP2) from *Haloferax volcanii*

Shanhui Liao[1]*, Wen Zhang[1]*, Kai Fan[1,4]*, Kaiqin Ye[1], Xuecheng Zhang[2], Jiahai Zhang[1], Chao Xu[3] & Xiaoming Tu[1]

[1]Hefei National Laboratory for Physical Sciences at Microscale, School of Life Science, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China, [2]School of Life Sciences, Anhui University, Hefei, Anhui 230039, P. R. China, [3]Structural Genomics Consortium, University of Toronto, 101 College St., Toronto, Ontario, M5G 1L7, Canada, [4]Reproductive Medical Center, Nanjing General Hospital of Nanjing Military Command, Nanjing 21002, P.R. China.

Ubiquitin-like proteins play important roles in diverse biological processes. In this study, we present an unexpected finding that a ubiquitin-like small archaeal modifier protein (SAMP2) from *Haloferax volcanii* adopts two distinct states under low ionic condition. One of these is similar to the β-grasp structure conserved in ubiquitin-like proteins from eukaryotes; the other is disordered, like prokaryotic ubiquitin-like protein, Pup. Furthermore, our study reveals that the conformation of SAMP2 is dependent on ionic strength. With the increase of ion concentration, SAMP2 undergoes a conformational conversion from disorder to order, indicating that the ordered conformation is the functional form of SAMP2 under the physiological condition of *H. volcanii*.

Since the discovery of ubiquitin and ubiquitin-like modifiers (Ubls) in eukaryotes, the roles of post-translational modifications by Ubls have been revealed in various biological processes ranging from proteolysis to DNA repair[1,2]. The covalent conjugation of Ubls to substrates is catalyzed by a system composed of activating enzyme (E1), conjugating enzyme (E2) and ligase (E3), which is conserved for various Ubls[2,3]. In addition to the conjugation pathway, eukaryotic Ubls share a β-grasp fold and a C-terminal diglycine motif, although the sequence similarity between them is low[4,5]. The similarities suggest that all Ubls might have a common ancestor. Furthermore, small proteins with the β-grasp fold and C-terminal diglycine motif are widespread among prokaryotes and archaea[6,7], providing additional evidence for the existence of a common ancestor of Ubls.

Eukaryotic ubiquitin-like proteins have been studied intensively. However, identification of prokaryotic ubiquitin-like proteins was a long process, due to their low sequence similarity. Recently, prokaryotic ubiquitin-like protein (Pup), involved in protein degradation in *Mycobacterium tuberculosis* (Mtb), has been identified[8,9]. Similar to Ubls involved in ubiquitylation, Pup is covalently conjugated to substrates by a ligase (PafA) and detached by a deamidase (Dop)[10,11]. Previously, our group and others revealed that Pup is an intrinsically disordered protein, which is an interesting property distinguished from ubiquitin and Ubls[12,13].

Archaeal almost thrive in the extreme environments such as high salt concentration, extremely temperature and so on. And the molecular adaptation to extreme environment is different from the prokaryotic and eukaryotic organisms. Recently, two ubiquitin-like small archaeal modifiers (SAMP1 and SAMP2) from *Haloferax volcanii* (*H. volcanii*) which belongs to the moderate halophiles and grows well in the salt concentration between 0.5 M and 2.5 M, have been reported[14]. SAMP1 and SAMP2 share low sequence identity with ubiquitin-like proteins from eukaryotes and prokaryotes, but preserve a C-terminal diglycine motif. Mass spectrometry analyses further revealed covalent conjugation between SAMPs and substrates in a manner similar to that of ubiquitin[14]. In addition, the C-terminal diglycine motif of SAMPs is essential for SAMPylation, as the role of it in ubiquitin and Ubls[14]. Furthermore, a functional study on SAMP2 indicated that SAMP2 is required for the thiolation of tRNA in *H. volcanii*[15].

Recent studies on prokaryotic ubiquitin-like protein, Pup, revealed that Pup adopts an intrinsically disordered structure which is totally different from the conserved ordered β-grasp structure of eukaryotic ubiquitin-like proteins. Therefore, it is interesting to know what the structure of archaea ubiquitin-like protein is. Besides, It is

also an interesting question that how salt affects the structure of ubiquitin-like proteins from halophilic archea. In our study, we analyzed the solution structures under different ionic conditions. We present an unexpected finding that SAMP2 adopts two distinct conformations under low ionic condition. Intriguingly, the state of SAMP2 is dependent on ionic strength. The disordered conformation of SAMP2 would transform to the ordered one with the ionic concentration increasing, indicating that the ordered conformation should be the functional form under the physiological condition of *H. volcanii*.

Results

SAMP2 adopts two distinct conformations under low ionic condition. Primary sequence analysis showed SAMP2 shares low sequence identity with eukaryotic and prokaryotic ubiquitin-like proteins. However, it possesses a conserved C-terminal diglycine motif, a characteristic of ubiquitin-like family proteins (Fig. 1a). In order to investigate the structure of SAMP2, recombinant SAMP2 was expressed and purified, which has 66 residues in length (Fig. 1a). The purified SAMP2 was dissolved in buffer containing 25 mM NaH_2PO_4, 0.1 M NaCl at pH 6.7 for NMR experiments. To our surprise, the ^{1}H-^{15}N HSQC spectrum of SAMP2 under the low ionic condition displayed more than 130 resonances (Fig. 1b),

implying SAMP2 might adopt two distinct conformations. This was also implied by ion exchange chromatography result. SAMP2 was eluted in two separated peaks from ion exchange column (Supplemental Fig. S1a), indicating the presence of two conformations. The fractions collected for the two separated peaks were examined by mass spectrometry. The results confirmed the same component, SAMP2, in the fractions (Supplemental Fig. S1b). Each fraction of the two separated peaks was further applied to another round of ion exchange chromatography. Two separated eluted peaks again appeared in each run (data not shown). Furthermore, the ^{1}H-^{15}N HSQC spectrum of SAMP2 from either of the two separated peaks was the same as that of initial SAMP2 before separation (Supplemental Fig. S1c). These results indicated that the two distinct conformations were intrinsic features of SAMP2 and interconverted in a dynamic balance.

SAMP2 presents an ordered conformation which is similar to the β-grasp fold conserved in ubiquitin-like proteins from eukaryotes. To better explore the two different conformations of SAMP2, a series of 3D NMR experiments were performed. Two sets of resonances, representing two different conformations of SAMP2, were assigned (Fig. 1b). Chemical shift index (CSI) was carried out for the two sets of resonances to evaluate the secondary structure characteristics of
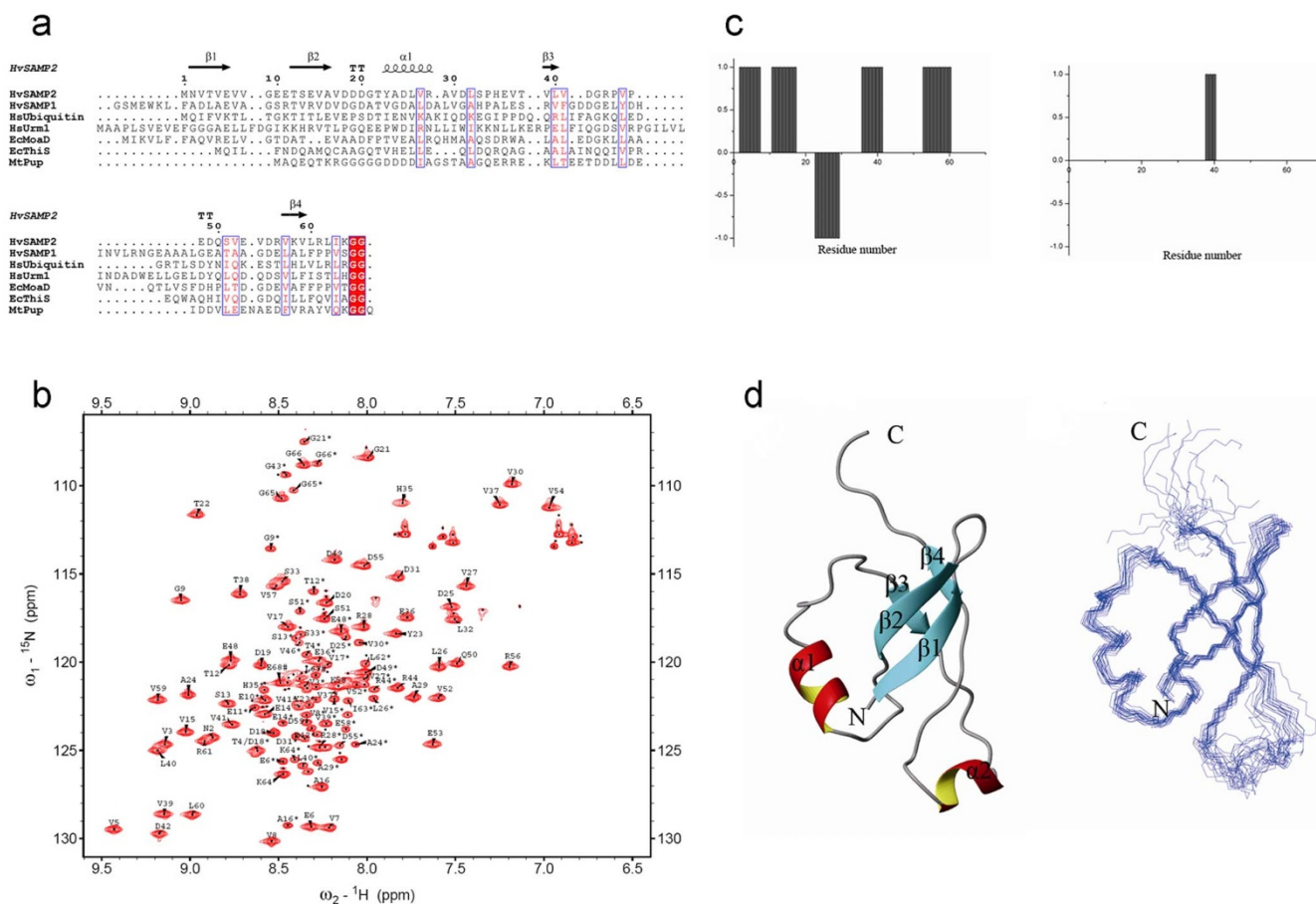
Figure 1 | Sequence alignment, chemical shift index and the solution structure of SAMP2-o under low ionic condition. (a) Multiple sequence alignment of SAMP2, SAMP1, Pup, ubiquitin and other β-grasp proteins. Alignment was performed with ClustalW2[29] and BOXSHADE version 3.2 (http://www.ch.embnet.org/software/BOX_form.html). Identical and similar amino acids are shaded in black and grey, respectively. HvSAMP2, *Haloferax volcanii* SAMP2; HvSAMP1, *Haloferax volcanii* SAMP1; HsUbiquitin, *Homo sapiens* Ubiquitin; HsUrm1, *Homo sapiens* Urm1; EcMoaD, *Escherichia coli* MoaD; EcThiS, *Escherichia coli* ThiS; MtPup, *Mycobacterium tuberculosis* Pup. (b) ^{1}H-^{15}N HSQC spectrum of SAMP2 under low ionic condition (0.1 M NaCl), peaks of SAMP2-d are labeled with '*'. (c) Chemical shift index (CSI) plots were tabulated based on chemical shifts of Hα, Cα, Cβ, CO of SAMP2 under low ionic condition (left for SAMP2-o, right for SAMP2-d). (d) Solution structure of SAMP2-o under low ionic condition (0.1 M NaCl). Left: Ribbon structure of a representative conformer of SAMP2-o. Right: 20 lowest-energy conformers calculated for SAMP2-o. This Figure was produced with MOLMOL[26].

the two conformations. The results indicated that one of the states (signified by SAMP2-o) SAMP2 possessed substantial secondary structures and the other (signified by SAMP2-d) sparsely populated structured regions except for one region with higher propensity for β-sheet structure. (Fig. 1c). We further determined the solution structure of SAMP2-o. The minimum-energy solution structure of SAMP2-o and the assembly of twenty lowest-energy structures were shown in Fig. 1d. The structural statistics of the NMR structure were shown in Table 1. The structure of SAMP2-o contains two α-helices (residues 22–29 for α1 and residues 47–50 for α2) and four β-stands i.e. strand 1 (residues 2–8), strand 2 (residues 11–17), strand 3 (residues 38–40) and strand 4 (residues 57–60). β1 and β2 of SAMP2-o are perpendicular to β3 and β4, distinct from the case in SAMP1 from *Haloferax volcanii* and other eukaryotic ubiquitin-like proteins wherein most β-stands are almost parallel to each other (Fig. 2). Eukaryotes ubiquitin-like proteins and archaea SAMP1[16] all have conserved β-grasp structures, which are folded compactly and maintained by the hydrophobic core formed by the β-sheet and one of the α-helices. SAMP2-o has secondary structure elements similar to those ubiquitin-like proteins (Fig. 2). However, compared with those of other ubiquitin-like proteins, the global fold of SAMP2-o is looser, due to much more flexible loops and a longer distance between the α-helices and the β-sheet (Fig. 2).

SAMP2 adopts a disordered state as well. In contrast to the folded state, the other state has only one region with helical character while the rest of the protein is largely disordered, indicated by CSI (Fig. 1c). NMR relaxation parameters can provide valuable insights into the internal molecular dynamics of disordered states. T_1, T_2 and heteronuclear $\{^1H\}$-^{15}N NOE were measured for SAMP2-d and SAMP2-o. The average values of the parameters are 579.7 ms, 134.97 ms, and 0.19 for SAMP2-d and 568.14 ms, 95.80 ms and 0.61 for SAMP2-o, respectively (Fig. 3a, 3b and 3c). Expectably, SAMP2-d exhibited much lower T_2 and $\{^1H\}$-^{15}N NOE, which indicated it was relatively flexible. We further evaluated residual or transient structure of SAMP2-d via NOE constraints. The results showed that there were only short-range NOE constraints in SAMP2-d (Fig. 3d), lacking medium-range and long-range NOEs. All these data unambiguously confirm that SAMP2-d was disordered state.

**SAMP2 undergoes a conformational conversion from disorder to order with the increase of ionic concentration.** Our results indicated that SAMP2 adopts two distinct conformations under low salt condition. However, *Haloferax volcanii* lives in a high salt environment, which implies that all proteins and protein complexes perform their biological functions under high ionic condition *in vivo*. In order to better understand the state of SAMP2 under its physiological condition, a series of ^{15}N-HSQC spectra of SAMP2 with different ionic concentrations were recorded. To our surprise, a number of resonances of SAMP2 were weakened severely and even disappeared when ion concentration of NaCl increased up to 1 M (Fig. 4a). Meanwhile, the intensities of the resonances from SAMP2-o increased. These observations imply a conversion from the disordered to ordered conformation of SAMP2. The HSQC spectra of SAMP2 in 2 M and 3 M NaCl were similar to that in 1 M NaCl (Supplemental Fig. S2), indicating SAMP2 adopts similar states under different conditions with ionic concentrations ranging from 1 M to 3 M. Meanwhile, gel filtration experiment showed that SAMP2 was eluted as two partially separated peaks under the low ionic condition and had only one peak under the high ionic condition (Supplemental Fig. S3), further supporting that SAMP2 underwent a conformational conversion from disordered to order with the increase of ionic concentration.

Table 1 | NMR and structural statistics of SAMP2

NMR restraints in the structure calculation	L*	H*
Intraresidue	220	176
Sequential ($\lvert i - j \rvert = 1$)	269	217
Medium-range ($\lvert i - j \rvert < 5$)	86	77
Long-range ($\lvert i - j \rvert >/= 5$)	87	85
Hydrogen bonds	0	0
Total distance restraints	662	555
Dihedral angle restraints	89	89
Lennard-jones potential energy, kcal.mol^{-1}	-135.23 ± 7.41	-149.71 ± 7.68
Rmsd from idealized covalent geometry		
Bonds, Å	0.0030 ± 0.0002	0.0030 ± 0.0020
Angles, °	0.3740 ± 0.0113	0.3604 ± 0.0700
Impropers, °	0.2254 ± 0.0135	0.1882 ± 0.0124
Rmsd from experimental restraints		
Distance, Å	0.0102 ± 0.0009	0.0078 ± 0.0012
Constrained dihedral, °	0.4384 ± 0.0769	0.3748 ± 0.0483
Chemical shift assignment %		
Backbone	90	95
Total	85	90
Coordinate rmsd for residues 2–60, Å		
All backbone atoms	0.92	1.13
All heavy atoms	1.62	1.78
Coordinate rmsd for secondary structure, Å		
All backbone atoms	0.86	1.08
All heavy atoms	1.64	1.71
Ramachandran plot, % residues		
Most favored regions	85.8	85.0
Additional allowed regions	10.3	13.1
Generously allowed region	2.6	0.4
Disallowed regions	1.3	1.5

L*: NMR and structural statistics of SAMP2-o under low ionic condition.
H*: NMR and structural statistics of SAMP2 under high ionic condition.
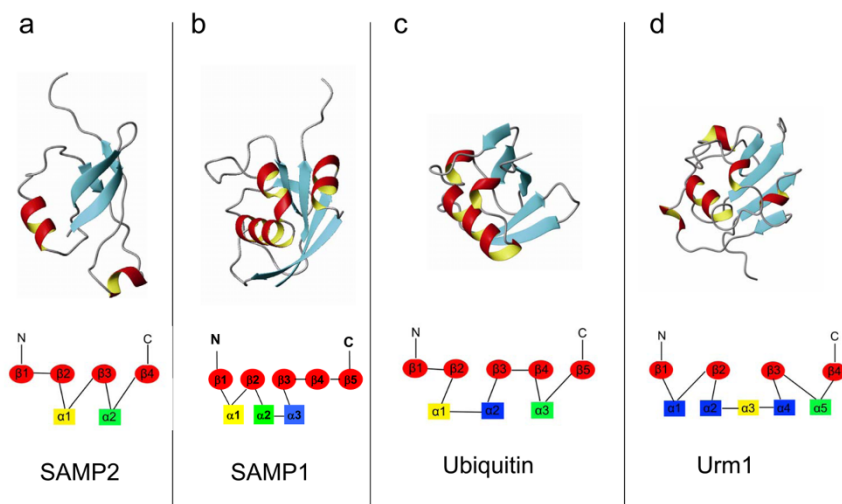
Figure 2 | Structural comparison of SAMP2-o, SAMP1, ubiquitin and Urm1. Structural comparison of (a) SAMP2-o from *Haloferax volcanii*, PDB ID code 2L32; (b) SAMP1 from *Haloferax volcanii*, PDB ID code 3PO0; (c) Ubiquitin from human, PDB ID code 1UBI; and (d) Urm1 from *T. brucei*, PDB ID code 2K9X. The corresponding SSEs in spatial structures are indicated with diagrams.

The solution structure of SAMP2 under high ionic condition (1 M NaCl) was further determined by NMR. The minimum-energy solution structure of SAMP2 and assembly of twenty lowest-energy structures were shown in Figure 4b. The structural statistics of the NMR structure were shown in Table 1. The structure was compared with the ordered conformation of SAMP2 under the low ionic condition (Fig. 4c). The RMSD between them is 3.2 Å, which demonstrates SAMP2 adopts similar folds for the ordered conformations under different ionic conditions.

The disordered state is an intrinsic feature of SAMP2 under low ionic condition. We elucidated above that SAMP2 may adopt both disordered and ordered states under low ionic conditions. However, one question is whether the disordered state of SAMP2 under low
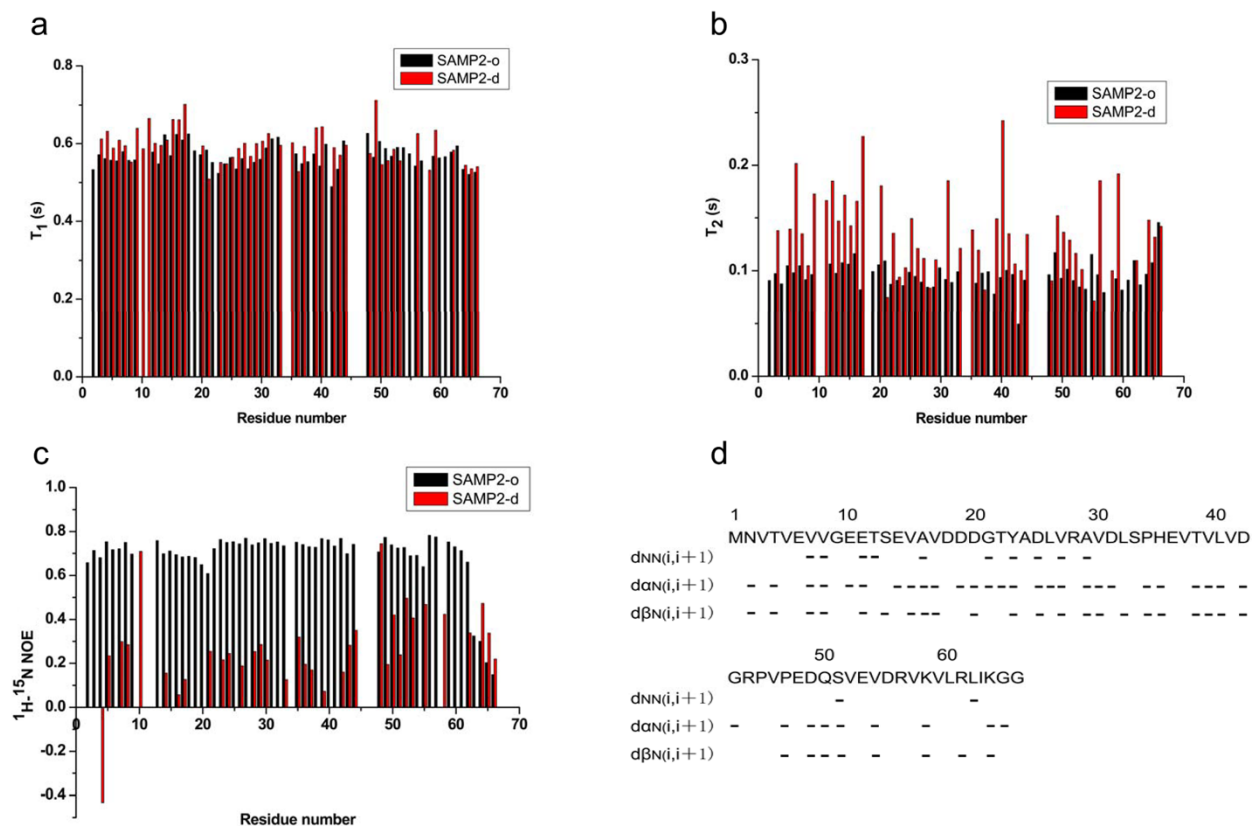
Figure 3 | NMR dynamic parameters of SAMP2 under low ionic condition. (a) ^{15}N longitudinal relaxation times (T_1) of SAMP2-o and SAMP2-d. (b) ^{15}N transverse relaxation times (T_2) of SAMP2-o and SAMP2-d. (c) The heteronuclear {^1H}-^{15}N NOEs of SAMP2-o and SAMP2-d. (d) Summary of inter-residual ^1H-^1H NOEs observed in SAMP2-d.
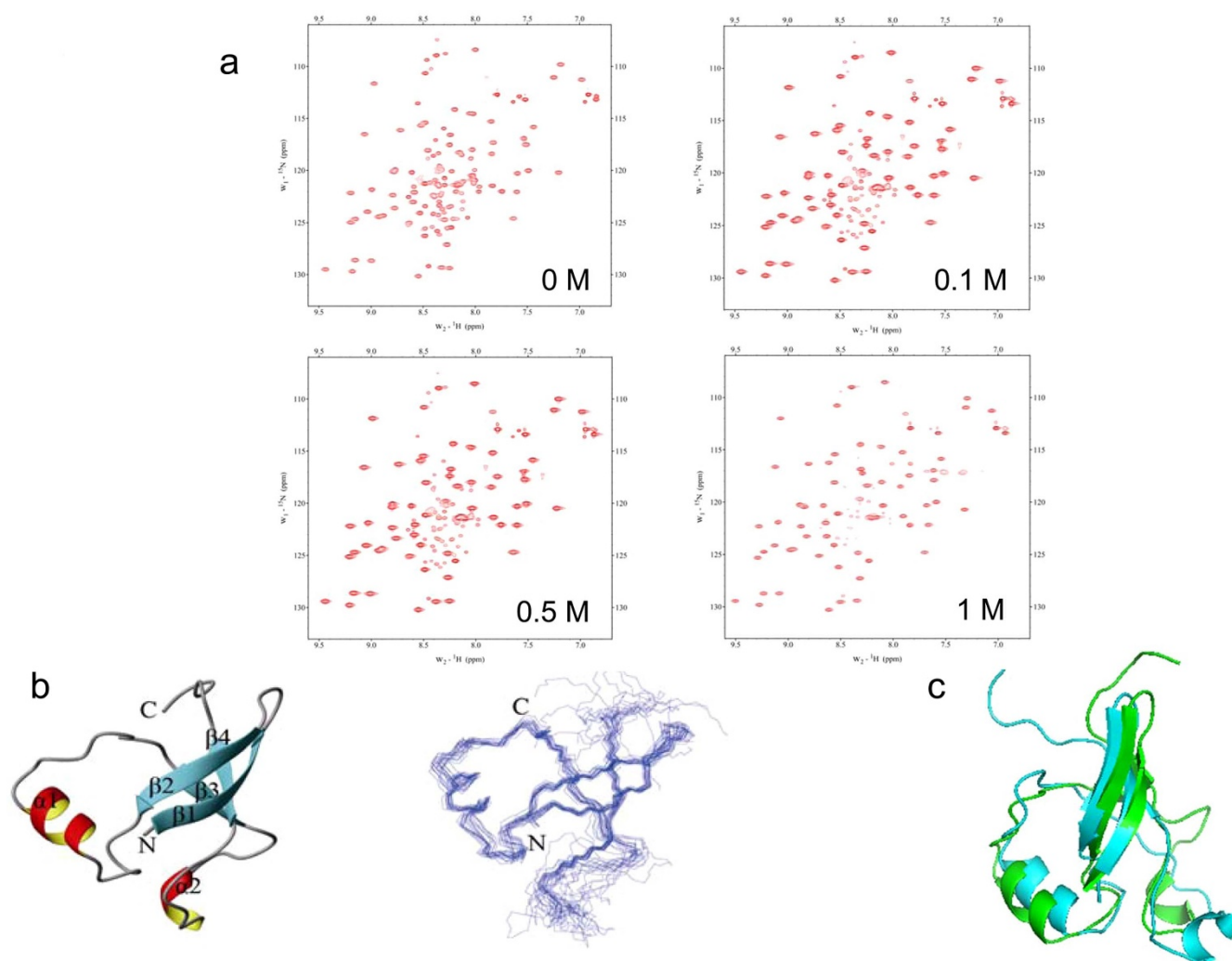
Figure 4 | ¹⁵N-HSQC spectra and solution structure of SAMP2 in different concentrations. (a) The ¹H-¹⁵N-HSQC spectra of SAMP2 in different concentrations of NaCl. The salt concentrations are 0 M, 0.1 M, 0.5 M and 1 M of NaCl, respectively. Note that all the weakened and disappeared resonances of SAMP2 belong to the disordered state and are distributed between 8.0–8.7 ppm of HN. (b) Solution structure of SAMP2 under high ionic condition (1 M NaCl). Left: Ribbon structure of a representative conformer of SAMP2 with the secondary structure elements highlighted. Right: Backbone superposition of 20 selected conformers with the lowest energy from the final CNS calculation. (c) Structural comparison of SAMP2 under low and high ionic conditions. Blue: structure in low salt. Green: structure in high salt.

ionic condition is merely an artifact or contamination produced during the expression or purification process. To answer it, we recorded ¹H-¹⁵N HSQC spectrum for SAMP2 in 1 M NaCl, then diluted it to 0.1 M NaCl by dialysis and recorded another spectrum. As expected, the spectrum in 1 M NaCl displayed only the resonances of the ordered conformation (Supplemental Fig. S4a). While, in the spectrum in diluted 0.1 M NaCl, additional resonances, which were identified from the disordered state, appeared besides of those of the ordered conformation (Supplemental Fig. S4b). This confirmed that the disordered form is an intrinsic feature of SAMP2 under low ionic condition.

Phylogenetic analysis. Phylogenetic analysis was performed on 36 ubiquitin-like proteins from eukaryotes, prokaryotes and archaea to obtain evolutionary information for SAMP2 (Supplemental Table S1). A neighbor-joining (NJ) tree demonstrates that SAMP2 is closer to MoaD, ThiS and Urm1 than Pup, Sumo and other eukaryotic Ubls, in evolution (Fig. 5). This is consistent with the recent report that SAMP2 was indicated to be required for the thiolation of tRNA, similar to MoaD, ThiS and Urm1[15]. As MoaD, ThiS and Urm1 are thought to be ancestors of the ubiquitin-like

superfamily[5,17], SAMP2 might represent the prototype of ubiquitin-like family proteins.

Discussion

Ancient archaea lived in the extreme environments, such as Dead Sea, which contains abundant resources of salt, and developed a large number of extremophile proteins. Generally accepted view is that *Haloferax volcanii* belongs to the moderate halophiles which grows well in medium containing 0.5–2.5 M salt and adopts "salt-in" strategy to maintain its internal environment which is accumulated salt from the external environment to the cell[18]. It is an interesting issue that what the mechanism of the adaptation of the proteins in the extreme halophiles is. In our study, we present a novel finding that SAMP2 from *Haloferax volcanii* may adopt two distinct states under low ionic condition. One of the two states is similar to the β-grasp structure conserved in ubiquitin-like proteins from eukaryotes; the other is disordered, like prokaryotic ubiquitin-like protein, Pup. Moreover, there is a dynamic equilibrium between the disordered form and the ordered form. With the small increase of ionic concentration, the resonances corresponding to the disordered state were weakened gradually and disappeared till the salt concentration
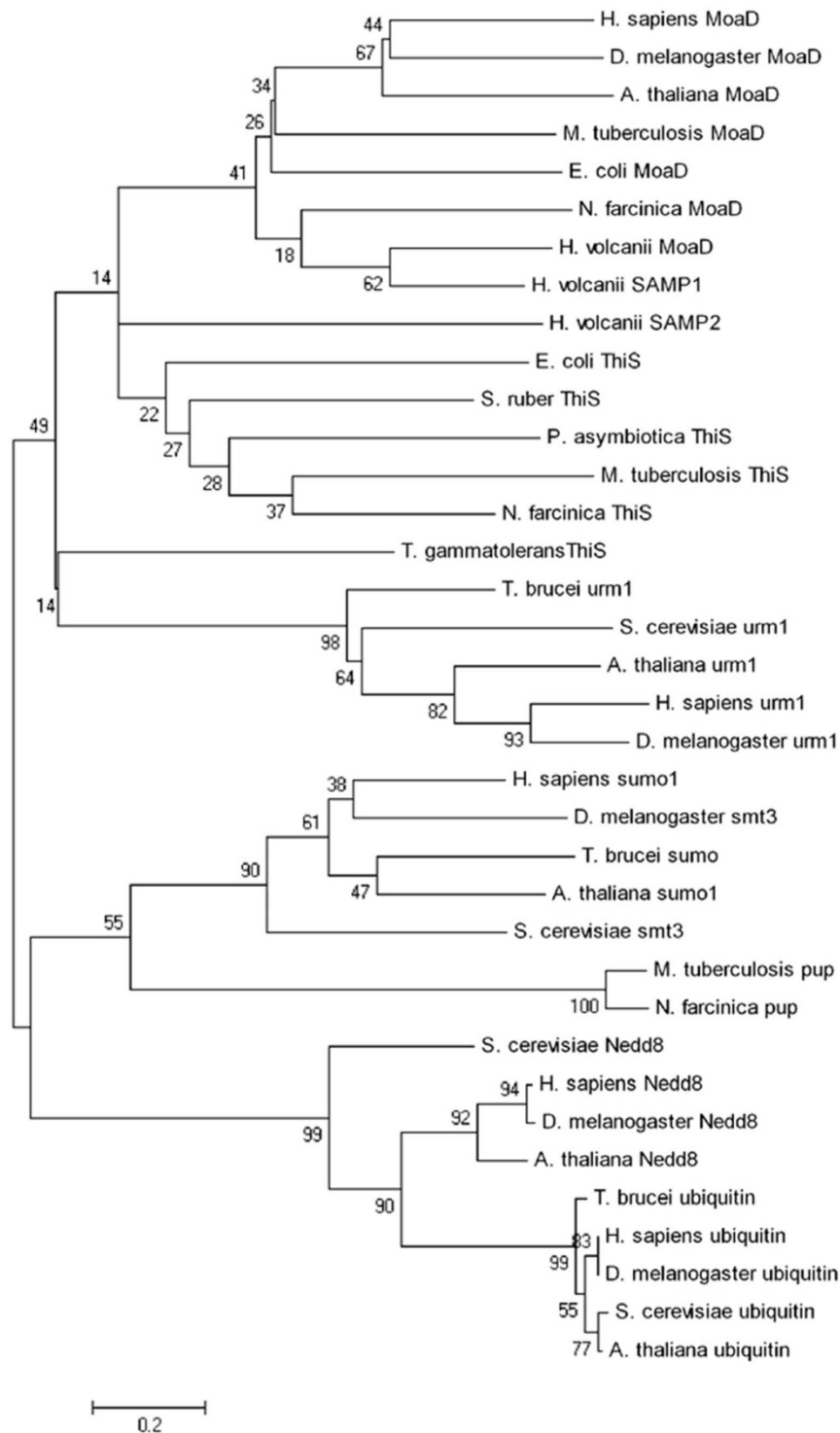
Figure 5 | Phylogenetic analysis of SAMP2 with other β-grasp proteins. Sequences were aligned with ClustalW and the phylogenetic tree was constructed with the Neighbour-Joining method using MEGA 4 program. The nonparametric bootstrap test was performed for 1,000 replicates. Sequences used in construction of phylogenetic tree were listed in Supplemental Table S1.

reaching to 1 M NaCl. It indicated that the disordered form of SAMP2 undergoes a conformational conversion to change to the ordered one. Furthermore, as a halophilic protein, the higher ionic strength is an advantage for *Hv* SAMP2 to form the ordered conformation to perform its biological function. In low salt concentration, however, SAMP2 is not inclined to fold an ordered conformation totally. Some molecules tend to adopt the disordered state which might be the inactive form of this protein.

Compared to eukaryotic ubiquitin-like proteins, sequence and structure analyses of SAMP2 showed that it has an increased abundance of acidic residues on protein surface, such as Asp and Glu, and decreased abundance of hydrophobic residues with long side chain in secondary structure elements. According to the theory proposed by Xavier Tadeo and coworkers[19], the increasing abundance of acidic residues on well-folded SAMP2 protein surface contribute to increase the protein solubility, while shorting side chain of

residues within secondary structure elements may improve the protein stability without influence on tendency of protein structure formation in high saline environments. Assumably, both the electrostatic and small side chain residues may play an important role in the adaptation of extremely high ionic strength environment. All these observations suggest SAMP2 plays its functional role via its ordered conformation under the physiological condition of *H. volcanii*.

It is well known that ubiquitin-like proteins in eukaryotes exhibit conserved ordered β-grasp structure. However, prokaryotic ubiquitin-like protein, Pup, is structurally distinguished from its eukaryotic counterparts in that it is an intrinsically disordered protein. The significant structural distinction between eukaryotic and prokaryotic ubiquitin-like proteins raises the question that what the evolutionary link between the order and the disorder is. Archaea including *Haloferax volcanii* is considered as, from evolutionary view, an ancient kingdom which comprises some characteristics of both prokaryotes and eukaryotes. Consistent with this, archaean ubiquitin-like protein SAMP2 displays structural characteristics which are combination of prokaryotic Pup and eukaryotic Ubls by presenting both disordered and ordered conformations under low salt conditions, which correspond to the environments modern prokaryotes and eukaryotes live. Phylogenetic tree analysis was performed on 36 representative ubiquitin-like modifier proteins from prokaryotic or eukaryotic organisms. The result indicated that SAMP2 is closed to the MoaD, an ancestor of the ubiquitin-like superfamily, in the evolution. So, SAMP2 might represent the most ancient species of ubiquitin-like family proteins.

Methods

Protein preparation. The gene encoding full-length SAMP2 was amplified and cloned into NdeI/XhoI site of pET-22 b (+) (Novagen). The recombinant plasmid with a 6-histidine tag fused to the C-terminus of the protein was transformed into *Escherichia coli* strain BL21 (DE3). Cells containing the recombinant plasmid were grown in Luria-Bertani medium containing 100 μg/mL of ampicillin at 37°C to an OD_{600} of 1.2 and induced with 0.5 mM isopropyl-β-D-thiogalacto-pyranoside (IPTG) for 5 hours. The induced cells were harvested and suspended in 30 mL ice-cold lysis buffer (20 mM Tris, 500 mM NaCl, pH 7.6), then lysed by sonication. The lysate was centrifuged at 14000 g for 20 min at 4°C. The supernatant was loaded on Ni-NTA resin filled column (QIAGEN). The column was washed with 20 mL lysis buffer and 25 mL lysis buffer containing 50 mM imidazole. Recombinant proteins were eluted with lysis buffer containing 500 mM imidazole. The fractions containing recombinant protein were collected and dialyzed against 500 mL dialysis buffer (25 mM NaH_2PO_4, 0.1 M NaCl, pH 6.7) three times. Dialyzed SAMP2 was then concentrated to about 4 mg/mL by using a centrifugal filter device (Millipore). For ion exchange chromatography, purified SAMP2 was loaded on an ion exchange column (Mono Q^{TM}, Pharmacia) using buffer A (25 mM NaH_2PO_4, pH 6.7) and buffer B (25 mM NaH_2PO_4, 1.5 M NaCl, pH 6.7) and the sample was separated by the gradient elution buffer with an increasing concentration of NaCl from 0 to 1.5 M. For gel filtration experiment, purified SAMP2 was loaded on a gel filtration column (Sephadex G-75) using elution buffer of 25 mM NaH_2PO_4, 0.1 M NaCl, pH 6.7 or 25 mM NaH_2PO_4, 1 M NaCl, pH 6.7. ^{15}N, ^{13}C-labeled SAMP2 was prepared in the same way except that LB was replaced by M9 medium containing 0.5 g/L 99% ^{15}N-labeled ammonium chloride and 2.5 g/L ^{13}C-labeled glucose as the sole nitrogen and carbon source, respectively.

Mass spectrometry. Fractions corresponding to peak1 and peak2 were loaded on SDS-PAGE gels. Gels containing proteins were excised and cut into ~1 mm³ pieces. Samples were washed with NH_4HCO_3 (50 mM). Samples were reduced by addition of 100 μL NH_4HCO_3 (50 mM) containing 10 mM dithiothreitol for 15 min at 50°C. Samples were alkylated by addition of 100 μL of freshly made NH_4HCO_3 (50 mM) containing 30 mM iodoacetamide for 15 min at room temperature in the dark and then washed with NH_4HCO_3 (50 mM). Commassie stain is removed by the addition of 200 μL of NH_4HCO_3 (50 mM) containing 30% (v/v) acetonitrile (ACN). Samples were washed with 200 μL of NH_4HCO_3 (50 mM) containing 50% (v/v) ACN, then with 200 μL 100% ACN. Drying of the samples was carried out by vacuum centrifugation. Digestion of the proteins in gels was performed by addition of 50 μL of NH_4HCO_3 (50 mM) containing 0.1 μg trypsin at 37°C overnight. In gel digested proteins were extracted by successive washes of gel slices with ACN. Extracted digested proteins were dried under vacuum centrifugation. The precipitated pellets were suspended in 25 μL 0.1% (v/v water) formic acid. LC-MS/MS analysis was performed on a ProteomeX-LTQ mass spectrometer (Thermo Fisher), and the data was analyzed by Bioworks Browser (Thermo Fisher).

NMR experiments and structure calculation. NMR spectra for structure calculation were recorded at 293 K on a Bruker DMX500 spectrometer. A set of standard triple-resonance spectra was recorded for backbone and side chain assignments. NOE distance restraints were obtained from 3D ^{15}N- and ^{13}C-edited NOESY spectra acquired with a mixing time of 130 ms. After these experiments, the sample was lyophilized and redissolved in 99.96% ^{2}H$_2$O. A series of ^{15}N-HSQC experiments was performed to monitor the disappearance of NH signals to obtain the hydrogen bond information. NMR data was processed with NMRPipe and analyzed with Sparky 3 software[20,21]. The interproton restraints were graded according to relative NOE intensity: strong NOE, 1.8–2.5 Å; intermediate NOE, 2.5–3.5 Å; weak NOE 3.5–5 Å. Based on analysis of Hα, Cα, Cβ and CO chemical shifts using the chemical shift index, the information of the φ and ψ backbone dihedral angles was obtained by TALOS program[22]. Hydrogen bond restraints were obtained by assignment of slow-exchange amide protons located in regular secondary structural elements (SSEs). The CNS program[23–25] was used to calculate 3-D structure of SAMP2 by distance restraints using the ARIA setup and protocols. Short-range NOEs and long-range NOEs were first used to determine SSEs of SAMP2. φ and ψ backbone dihedral angles and hydrogen bond restraints were added in consecutive steps to constraint the 3-D structure of SAMP2. 20 structures with the lowest energy from 200 calculated structures were selected for structural statistic calculation and analyzed with MOLMOL[26]. Ramachandran plot was analyzed with PROCHECK[27].

^{15}N longitudinal (T_1) and transverse (T_2) relaxation times and the heteronuclear ^{1}H-^{15}N NOE were recorded at 293 K on a Bruker DMX600 spectrometer. For the T_1 measurements, 8 time points were collected with delays of 11.15, 61.30, 141.54, 241.84, 362.20, 522.68, 753.37 and 1144.54 ms; For T_2, 7 time points with delays of 0, 17.6, 35.2, 52.8, 70.4, 105.6 and 140.8 ms were recorded. The heteronuclear ^{1}H-^{15}N NOEs were calculated from duplicate pairs of ^{1}H-^{15}N spectra recorded with and without amide proton saturation.

Phylogenetic analysis. A total of 36 sequences from ubiquitin-like modifier families, as well as the ThiS and MoaD families, were obtained from NCBI reference protein database and swissprot protein sequences database (accession numbers in Supplemental Table S1). Amino acid sequences of above proteins were aligned using MEGA 4[28]. Because of the low sequence similarity across these sequences, the gap separation distance was adjusted to 2.0 for a better alignment. The phylogenetic tree was constructed with Neighbor-Joining method using MEGA 4 program. The nonparametric bootstrap test was performed for 1000 replicates.

Accessions code. Protein Data Bank: The structures of SAMP2-o under low ionic condition and SAMP2 under high ionic condition have been deposited in Protein Data Bank with ID codes 2L32 and 2LJI, respectively.

1. Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* **458**, 422–429 (2009).
2. Morris, J. R. More modifiers move on DNA damage. *Cancer Res* **70**, 3861 (2010).
3. Haas, A. L. & Siepmann, T. J. Pathways of ubiquitin conjugation. *FASEB J* **11**, 1257–1268 (1997).
4. Tanaka, K., Suzuki, T. & Chiba, T. The ligation systems for ubiquitin and ubiquitin-like proteins. *Mol Cells* **8**, 503 (1998).
5. Hochstrasser, M. Evolution and function of ubiquitin-like protein-conjugation systems. *Nat Cell Biol* **2**, 153–157 (2000).
6. Burroughs, A. M., Balaji, S., Iyer, L. M. & Aravind, L. A novel superfamily containing the beta-grasp fold involved in binding diverse soluble ligands. *Biol Direct* **2** (2007).
7. Iyer, L. M., Burroughs, A. M. & Aravind, L. The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like β-grasp domains. *Genome Biol* **7**, R60 (2006).
8. Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P. & Darwin, K. H. Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science's STKE* **322**, 1104 (2008).
9. Burns, K. E., Liu, W. T., Boshoff, H. I. M., Dorrestein, P. C. & Barry, C. E. Proteasomal protein degradation in *Mycobacteria* is dependent upon a prokaryotic ubiquitin-like protein. *J Biol Chem* **284**, 3069–3075 (2009).
10. Striebel, F., Hunkeler, M., Summer, H. & Weber-Ban, E. The mycobacterial Mpa–proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. *EMBO J* **29**, 1262–1271 (2010).
11. Striebel, F. *et al.* Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat Struct Mol Biol* **16**, 647–651 (2009).
12. Liao, S. *et al.* Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein. *Biochem. J* **422**, 207–215 (2009).
13. Chen, X. *et al.* Prokaryotic ubiquitin-like protein pup is intrinsically disordered. *J Mol Biol* **392**, 208–217 (2009).
14. Humbard, M. A. *et al.* Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature* **463**, 54–60 (2010).
15. Miranda, H. V. *et al.* E1- and ubiquitin-like proteins provide a direct link between protein conjugation and sulfur transfer in archaea. *Proc Natl Acad Sci U S A* **108**, 4417 (2011).
16. Jeong, Y. J., Jeong, B. C. & Song, H. K. Crystal structure of ubiquitin-like small archaeal modifier protein 1 (SAMP1) from *Haloferax volcanii*. *Biochim Biophys Res Commun* **405**, 112–117 (2011).

17. Xu, J. *et al.* Solution structure of Urm1 and its implications for the origin of protein modifiers. *Proc Natl Acad Sci U S A* **103**, 11625–11630 (2006).
18. Oren, A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Systems* **4**, 13 (2008).
19. Tadeo, X. *et al.* Structural basis for the aminoacid composition of proteins from halophilic archea. *PLoS Biol* **7**, e1000257 (2009).
20. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293 (1995).
21. Goddard, T. & Kneller, D. SPARKY 3-NMR Assignment and Integration Software. *University of California, San Francisco* (2006).
22. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**, 289–302 (1999).
23. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D: Biological Crystallography* **54**, 905–921 (1998).
24. Linge, J. & Nilges, M. Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. *J Biomol NMR* **13**, 51–59 (1999).
25. Nilges, M. & O'Donoghue, S. I. Ambiguous NOEs and automated NOE assignment. *Prog Nucl Mag Res Sp* **32**, 107–139 (1998).
26. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14**, 51 (1996).
27. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477–486 (1996).
28. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599 (2007).
29. Larkin, M. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

Author contributions

S.L., W.Z., K.F. and K.Y. carried out the experiments, J.Z. collected the data. X.Z., C.X. and X.T. designed the experiments. All authors analyzed the data. All authors contributed in writing paper.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Liao, S. *et al.* Ionic strength-dependent conformations of ubiquitin-like small archaeal modifier protein from *Haloferax volcanii. Sci. Rep.* **3**, 2136; DOI:10.1038/srep02136 (2013).