7938–7949 *Nucleic Acids Research, 2017, Vol. 45, No. 13* *Published online 18 May 2017*
doi: 10.1093/nar/gkx437

The SmAP1/2 proteins of the crenarchaeon *Sulfolobus solfataricus* interact with the exosome and stimulate A-rich tailing of transcripts

Birgit Märtens[1,*], Linlin Hou[2], Fabian Amman[3], Michael T. Wolfinger[3,4], Elena Evguenieva-Hackenberg[2] and Udo Bläsi[1,*]

[1]Department of Microbiology, Immunobiology and Genetics, Max F. Perutz Laboratories, Center of Molecular Biology, University of Vienna, Vienna Biocenter, Dr. Bohrgasse 9, 1030 Vienna, Austria, [2]Institute of Microbiology and Molecular Biology, Justus Liebig University Gießen, Heinrich-Buff-Ring 26–32, 35392 Gießen, Germany, [3]Institute for Theoretical Chemistry, University of Vienna, Währingerstraße 17/3, 1090 Vienna, Austria. and [4]Center for Anatomy and Cell Biology, Medical University of Vienna, Währingerstraße 13, 1090 Vienna, Austria

Received June 27, 2016; Revised April 27, 2017; Editorial Decision April 29, 2017; Accepted May 03, 2017

ABSTRACT

The conserved Sm and Sm-like proteins are involved in different aspects of RNA metabolism. Here, we explored the interactome of SmAP1 and SmAP2 of the crenarchaeon *Sulfolobus solfataricus* (Sso) to shed light on their physiological function(s). Both, SmAP1 and SmAP2 co-purified with several proteins involved in RNA-processing/modification, translation and protein turnover as well as with components of the exosome involved in 3′ to 5′ degradation of RNA. In follow-up studies a direct interaction with the poly(A) binding and accessory exosomal subunit DnaG was demonstrated. Moreover, elevated levels of both SmAPs resulted in increased abundance of the soluble exosome fraction, suggesting that they affect the subcellular localization of the exosome in the cell. The increased solubility of the exosome was accompanied by augmented levels of RNAs with A-rich tails that were further characterized using RNA_{Seq}. Hence, the observation that the Sso SmAPs impact on the activity of the exosome revealed a hitherto unrecognized function of SmAPs in archaea.

INTRODUCTION

The evolutionarily conserved Sm and Sm-like (Lsm) proteins play important roles in RNA metabolism (1–4). Sm/Lsm proteins share affinity for single-stranded 3′ uridine or adenosine tracts (4) that provide vulnerability to 3′ exonucleolytic attack and are thus important determinants of RNA stability.

In Eukaryotes, the hetero-heptameric Lsm complexes are either localized in the nucleus (Lsm2–8) or in the cytoplasm (Lsm1–7) (4–6). Lsm2–8 functions in various RNA maturation processes as well as in decay of nuclear RNAs (4–8). Lsm1–7 binds to the 3′UTR of deadenylated mRNAs, which can prevent nucleolytic attack by the exosome (9–11) and simultaneously stimulate de-capping, which precedes 5′ to 3′ directional decay (4,6,12).

In *Escherichia coli*, the Sm-like protein Hfq facilitates the interactions of small RNAs with target mRNAs, which modulates their translation output and stability (13). In addition, Hfq can also influence mRNA decay directly by association with the 3′ end of the transcript and by promoting polyadenylation, which in turn triggers 3′ to 5′ degradation by exoribonuclease(s) (14–17). At variance with mRNAs, polyadenylation of tRNAs controls their processing and thereby regulates functional tRNA levels (18). Moreover, Hfq can inhibit the 3′ to 5′ exonuclease activity of polynucleotide phosphorylase (PNPase) (15,19), and was found in association with poly(A) polymerase (PAP) (20) and PNPase (20,21).

While bacterial and eukaryotic Lsm family members have been studied in more detail (2,4,5,14), the function of Sm-like archaeal proteins (SmAPs) remains poorly understood. Like other Sm/Lsm proteins, the SmAPs are composed of an N-terminal α-helix and five β-strands forming a continuous Sm-fold (22–26). However, the homo-heptameric SmAPs display characteristic differences from bacterial and eukaryotic Lsm proteins, e.g. they lack an extended C-terminal domain that is characteristic for some bacterial Hfq- and eukaryotic Sm proteins (2). Several studies

*To whom correspondence should be addressed. Tel: +43 1 4277 54609; Fax: +43 1 4277 9546; Email: Udo.Blaesi@univie.ac.at
Correspondence may also be addressed to Birgit Märtens. Tel: +43 1 4277 54609; Fax: +43 1 4277 9546; Email: Birgit.Maertens@univie.ac.at
Present address: Linlin Hou, Genome Regulation Laboratory, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China.

showed that SmAPs from different Archaea bind to oligo-U stretches of different lengths with K_d's ranging from 70 nM to 10 μM depending on their origin and the length of oligo-U (24,25,27–31). Co-immunoprecipitation (Co-IP) experiments revealed that the two SmAPs of the euryarchaeon *Archaeoglobus fulgidus* associate with RNaseP RNA *in vivo* (24), which may suggest a role in tRNA processing. Furthermore, a Co-IP approach with the sole SmAP of *Haloferax volcanii* (Hv) revealed potential interacting proteins that are involved in translation (aEF-2; aEF-1α), stress response (heat shock proteins; thermosome), nucleic acid metabolism (nucleases; mRNA 3′ end processing) and the cell cycle (28). In addition, the Hv SmAP was shown to co-purify with several uncharacterized non-coding RNAs, tRNAs and C/D box snoRNAs (28). A deletion of the Sm1 motif in the Hv SmAP encoding gene showed a gain of function in swarming, which agreed with the up-regulation of transcripts encoding proteins required for motility (32).

In the clade of crenarchaeaota 2–3 SmAPs are present, whereas in Euryarchaeota only 1–2 SmAPs are found (2). One of the best characterized crenarchaeota is *Sulfolobus solfataricus* (Sso), which can grow chemo-organotrophically at 80°C and at a pH of 2–4. Sso encodes three SmAP proteins (http://www-archbac.u-psud.fr/projects/sulfolobus), Sso 6454 (SmAP1), Sso 5410 (SmAP2) and Sso 0276 (SmAP3). SmAP1 and SmAP2 show 50% similarity, whereas they share only 30% similarity with SmAP3. In Sso different classes of non-coding RNAs and mRNAs were identified that interact either with SmAP1 or SmAP2 or with both proteins (26). The large number of associated intron-containing tRNAs and rRNA modifying RNAs suggested as well a role of these SmAPs in tRNA/rRNA processing (26).

In Eukaryotes and Archaea, the exosome can be regarded as a central 3′ to 5′ RNA processing and degradation machinery. The archaeal exosome is structurally similar to the nine-subunit core of the essential eukaryotic exosome and to bacterial PNPase (33,34). In contrast to the eukaryotic exosome, PNPase and the archaeal exosome exhibit metal ion-dependent phosphorolytic activities, and in addition to their exoribonucleolytic activity, synthesize heteropolymeric RNA tails (33). The Sso exosome consists of four orthologs of the eukaryotic exosomal subunits: the RNase PH-domain-containing subunits Rrp41 and Rrp42 form a hexameric ring with three active sites, whereas the S1-domain-containing subunits Rrp4 and Csl4 form an RNA-binding trimeric cap on the top of the ring (35). In Sso, the subunits Rrp4 and Csl4 confer different substrate specificities to the exosome (36). Rrp4 displays poly(A) specificity (36), whereas the Csl4-exosome degrades with high efficiency RNAs with an A-poor 3′ end (36). DnaG, which binds to the Csl4-exosome, functions as an additional RNA-binding subunit with poly(A) specificity (36,37).

In Eukaryotes, a spatial organization of RNA processing and degradation is ensured not only via compartmentalization, but also by sub-localization of RNases within specialized cytoplasmic foci (P-bodies) (38). A spatial organization of the degradosome has also been described in Bacteria (39–42). Here, the bacterial Sm protein Hfq co-localizes with the degradosome at the cytoplasmic membrane (43–45) and is also found in the nucleoid (46). In Sso, the exosome can likewise localize to the membrane, which has been suggested to be mediated by the DnaG subunit (47). The partitioning between the membrane and the cytoplasm might be important for regulation of the exosome activity, i.e. 3′ to 5′ decay and tailing, as suggested for the bacterial degradosome (41).

Here, using affinity purification in combination with mass spectrometry we identified proteins that interact with Sso SmAP1 and Sso SmAP2. Among others, the study disclosed DnaG as a putative interacting partner of both SmAPs. Follow-up studies corroborated a physical interaction of both SmAPs with DnaG. In addition, elevated levels of the SmAPs increased the abundance of the soluble exosome and that of RNAs with A-rich tails.

MATERIALS AND METHODS

Purification of His-tagged SmAPs and DnaG from Sso and identification of co-purifying proteins

The Sso strains PH1-16(pMJ05-SmAP1-His), PH1-16(pMJ05-SmAP2-His) (26) and PH1-16(pMJ05-DnaG-His) were generated as described (48) (Supplementary Data). For expression of the plasmid borne genes, the strains PH1-16(pMJ05), harboring the backbone vector (mock control), PH1-16(pMJ05-SmAP1-His), PH1-16(pMJ05-SmAP2-His) and PH1-16(pMJ05-DnaG-His) were grown at 75°C in arabinose-containing Brock's medium. The affinity purification was performed as described in Supplementary Data. The dialysed eluates (500 μl) were concentrated 10-fold to 50 μl using 3K Amicon® Ultra-0.5 centrifugal filter devices (Millipore), and then analyzed for co-purifying proteins by mass spectrometry described in detail in the Supplementary Data. Each experiment was performed in duplicate using two biological replicates. As input for the purifications, 20 μl (1/1000) of the cell lysates from the respective strains were loaded. Five microliter (for western-blot analysis) and 15 μl (for coomassie-blue staining) of the respective eluates from the affinity purifications were used. For further details see Supplementary Data.

SmAP antibodies

Antibodies directed against the Sso SmAPs were raised in rabbits (Pineda). The anti-SmAP1 serum was specific for SmAP1. The antibody raised against SmAP2 recognized both SmAPs, although a slightly reduced avidity for SmAP2 was noticed. Throughout the manuscript the designation anti-SmAP1/2 is used for the latter antibodies to note its cross-reactivity.

Co-Immunoprecipitation

SmAP1-His, SmAP2-His and DnaG-His were purified from *E. coli*. The SmAP1 coding gene was cloned as described in Supplementary Data. DnaG-His, SmAP1-His and SmAP2-His were purified as previously described (26,49). For Co-IP with purified components, 50 pmol of DnaG-His was either incubated alone (mock control) or SmAP1-His and SmAP2-His were incubated together with 50 pmol DnaG-His, respectively, in 200 μl Co-IP buffer (50 mM Tris–HCl pH 6.0, 100 mM KCl; 5% glycerol, 1 mM $MgCl_2$ and 0.1%

Triton X-100) for 10 min at 65°C. Then, 15 μl of anti-SmAP1/2 were added and the samples were incubated for 1h on ice. The Dynabeads® Protein G beads (Invitrogen) were equilibrated in Co-IP buffer and then added to the samples and incubated for 1 h at 4°C. The beads were captured by a magnetic device and washed three times with 1 ml Co-IP buffer. The elution was achieved by incubating the beads for 10 min at 98°C in 50 μl of sodium dodecyl sulphate (SDS)-loading buffer. The bound proteins were analyzed by western-blotting using the anti-SmAP1/2 antibodies or DnaG-specific antibodies as described in Supplementary Data.

For Co-IP of endogenous untagged SmAP1 and DnaG, cell lysates from 400 ml culture of the wild-type strain P2 were prepared. The cells were pelleted and lysed in 20 ml Co-IP buffer by sonication. The cell debris was removed and an aliquot of 100 μl of the P2-lysate was used as input material (In). Twenty microliter anti-SmAP1-antibodies or 20 μl of the anti-DnaG-antibodies were added to 10 ml of P2-lysate and incubated for 1 h on ice. After incubation, 20 μl of the Dynabeads® Protein G beads (Invitrogen) were equilibrated in Co-IP buffer, and then added to the samples and incubated on a rolling wheel at 4°C. The beads were captured with a magnetic device and the supernatant was used as flow through/unbound fraction (Ft). The beads were washed three times with 1 ml of Co-IP buffer and the last wash fraction (W) was precipitated with trichloroacetic acid and resuspended in 10 μl SDS-Laemmli-buffer. The beads were eluted (E) with 60 μl 100 mM Glycine pH 2.0 for 15 min at 30°C and then 20 μl 50 mM Tris pH 8.9 was added. Ten microliter of the Input (1/1000), 10 μl of the flow through/unbound fraction (1/1000) and the last wash fraction together with 20 μl of the eluted proteins/bound fraction (one-fourth) were analyzed by western-blotting using SmAP1-specific or DnaG-specific antibodies.

Exosome pelleting assay

The strains PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His) were inoculated and the synthesis of the SmAPs was induced with arabinose (Supplementary Data). To demonstrate increased synthesis of SmAP1 and SmAP2, 1 ml samples from non-induced cultures (+sucrose) and induced cultures (+arabinose) were withdrawn at an OD_{600} of 0.8. The cells were pelleted and then resuspended in SDS-Laemmli-buffer. Overproduction of the SmAPs (Figures 1A and 4A) was tested by western-blotting using the anti-SmAP1/2 antibodies. For the exosome pelleting assay (see Figure 4B), 50 ml each of the non-induced cultures (+sucrose) and induced cultures (+arabinose) was pelleted and resuspended in 1 ml 2-(N-morpholino)ethanesulfonic acid (MES) low salt buffer (20 mM MES, 2 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonylfluorid (PMSF) and 0.5 mM ethylenediaminetetraaceticacid (EDTA) pH 6.5). For the sucrose gradients 100 ml of the induced cultures (+arabinose) were pelleted and resuspended in 1 ml MES low salt buffer. The cells were lysed by sonication 3 × 20 s and the cell debris was removed by centrifugation at 5000 *g* at 4°C for 30 min. A total of 750 μl of the supernatant (S5) was then loaded onto a linear 10–30% sucrose gradient and analyzed as described below or centrifuged at 130 000 *g* at 4°C for 2 h. The supernatant was removed (S130) and the pellet was resuspended in 750 μl MES low salt buffer (P130). An equal volume of S130 and P130 was loaded on 12.5% SDS-polyacrylamide gels. The proteins were transferred to a nitrocellulose membrane and immunodetection was carried out with anti-DnaG, anti-Rrp41, anti-aIF2γ and anti-aIF2β antibodies as described in Supplementary Data.

Isolation of RNA and RNA_{Seq}

A total of 100 ml culture of non-induced or induced PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His) cells, respectively, were pelleted and total RNA was isolated using Trizol (Thermo Fisher Scientific). A total of 250 μg of total RNA of each sample (two biological replicates for each condition/strain) was used to isolate adenylated RNAs employing the poly(A) RNA purification kit Oligotex™ (Qiagen) following the manufacturer's instructions. Two microliter of the eluted A-rich RNAs were analyzed with the Agilent 2100 Bioanalyzer (Agilent Technologies) and the RNA 6000 Nano Kit (Agilent Technologies) following the manufacturer's instructions. The cDNAs libraries were constructed using the NEBNext® Ultra™ Directional RNA Library Prep Kit. A total of 100 bp single end sequence reads were generated by the next generation sequencing facility at the Vienna Biocenter Core Facilities GmbH (VBCF), member of Vienna Biocenter (VBC), using the Illumina HiSeq 2000 platform. Raw sequencing reads are available at the European Nucleotide Archive (ENA) under the accession number PRJEB20182. Adaptor sequences of the reads were removed using cutadapt (50) and mapped against the Sso P2 reference genome (NC_002754) using bwa (algorithm mem; clipping penalty -L 4) arXiv:1303.3997 [q-bio.GN]. Therefore, reads which map only partially to the reference genome with tails being not encoded in the reference genome are reported as softclipped in the read alignment. After removal of polymerase chain reaction duplicated and reads aligning to multiple genomic loci, the softclipped sequences and their corresponding anchor site where extracted using custom perl scripts (available at https://github.com/fabou-uobaf/Helferlein/blob/master/getMappingOverhang.pl). Only 3′ tails with at least 15 nt in length and with at least 5 reads were considered (parameters used: -l 15 -c 5 -cc 3 -s -1 -e 3). The tail consensus sequence was determined by majority vote on the natively aligned individual sequence tails and inspected with WebLogo (51). Replicas were merged by keeping only tail anchor points, which were detected in both replicas, thereby allowing for five bases inaccuracy. To quantify the expression levels of individual genes, reads per gene were counted using BEDtools (52) and normalized to transcripts per million using ViennaNGS (53). All tail attachment sites (i) positioned up to 100 nt downstream of an annotated gene, (ii) within a gene or in (iii) intergenic regions were included. The genes with a detected tail (within the coding region and/or within up to 100 nt in the 3′UTR) were tested for functional enrichment utilizing the functional annotation from the SulfoSys project (54). Thereby, annotated genes without any assigned function were added

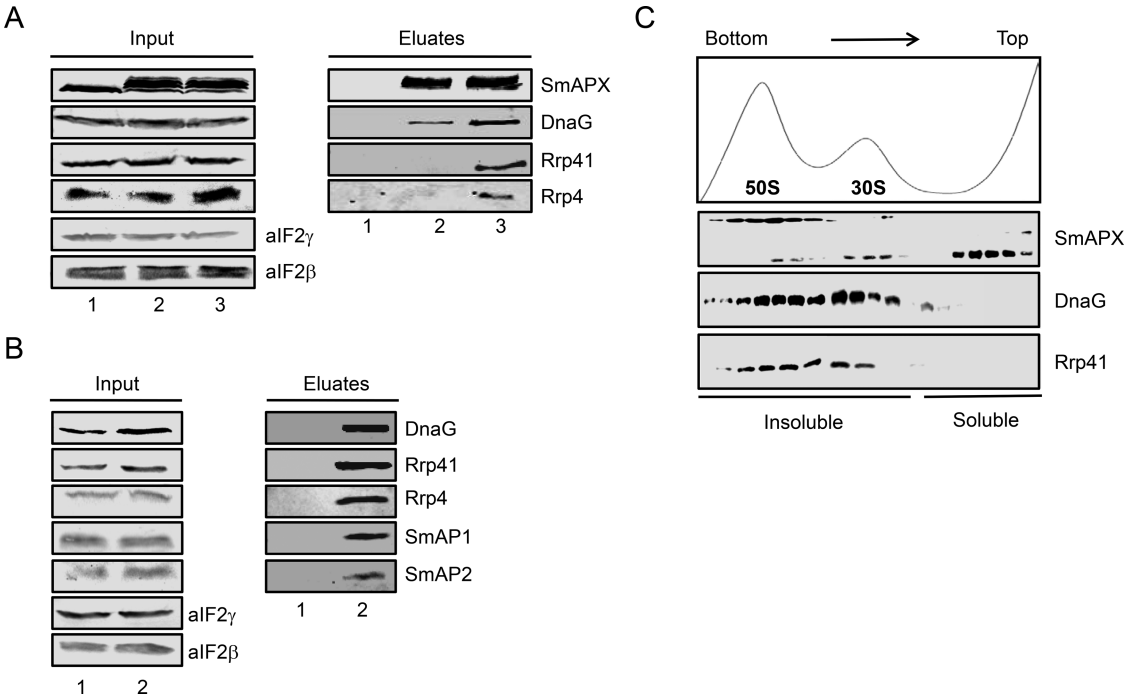

Figure 1. SmAP1 and SmAP2 associate with the exosome. (**A**) Exosomal subunits co-purifying with His-tagged SmAP1 and SmAP2. A total of 20 μl of the cell lysates comprising 1/1000 of the input materials (left panel) and 5 μl of the eluates from the affinity-purifications (right panel) obtained from strains PH1-16(pMJ05) (mock experiment; right panel, lane 1), PH1-16(pMJ05-SmAP1-His) (right panel, lane 2) and strain PH1-16(pMJ05-SmAP2-His) (right panel, lane 3) were subjected to western blot analysis. The presence of the bait proteins SmAP1 and SmAP2 were confirmed by using anti-SmAP1/2 antibodies and co-purifying exosomal subunits were detected using anti-DnaG-, anti-Rrp41- and anti-Rrp4-specific antibodies. The blot was also probed with antibodies directed against the aIF2β and aIF2γ subunits of translation initiation factor aIF2 (loading control). (**B**) SmAP1, SmAP2, Rrp4 and Rrp41 co-purify with His-tagged DnaG. A total of 20 μl of the cell lysates comprising 1/1000 of the input materials (left panel) and 5 μl of the eluates from the affinity-purifications (right panel) obtained from strains PH1-16(pMJ05) (mock experiment; right panel, lane 1), PH1-16(pMJ05-DnaG-His) (right panel, lane 2) were subjected to western-blot analysis. The presence of the bait protein DnaG was confirmed by using anti-DnaG and the co-purifying exosomal proteins were detected using anti-Rrp41- and anti-Rrp4- antibodies. SmAPX was detected using anti-SmAP1/2 antibodies. The blot was also probed with antibodies directed against the aIF2β and aIF2γ subunits of translation initiation factor aIF2 (loading control). (**C**) Co-migration of SmAPs with DnaG. A cell lysate from the Sso wild-type strain P2 was layered on top of a linear 10–30% sucrose gradient. 500 μl fractions were collected and the OD_{260} was measured to determine the positions of 30S and 50S ribosomal subunits (upper panel). The fractions were TCA precipitated and subjected to western-blot analysis. The heptameric (upper bands; SmAPX) and monomeric forms (lower bands) of the SmAPs, the DnaG- and the Rrp41-specific bands were detected with anti-SmAP1/2-, anti-DnaG- and anti-Rrp41-specific antibodies, respectively.

to the group 'Function unknown'. Enrichment of functions within the tail associated gene set versus the genome annotation background was tested with a two-sided Fisher's exact test. The obtained *P*-values were corrected for multiple testing and considered significantly enriched with an enrichment factor >1.5 and a corrected *P*-value < 0.05.

Sucrose gradients

One gram (wet weight) of wild-type cells grown at 75°C in Brock's medium was lysed in buffer containing 20 mM Tris–HCl pH 7.4, 10 mM MgAc, 40 mM NH_4Cl and 1 mM DTT. The cell lysate was centrifuged and 500 μg of total protein were loaded onto a linear 10–30% sucrose gradient. 100 ml of PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His) grown in the presence of arabinose were lysed as described for the exosome pelleting assay. A total of 750 μl of the S5 lysate was loaded onto a linear 10–30% sucrose gradient. After centrifugation at 100 000 *g* for 17 h at 4°C, the samples were collected by continuously measuring the OD_{260}. Then, 500 μl samples were TCA precipitated. The samples were loaded on 15% SDS-polyacrylamide gels and the proteins were blotted onto a nitrocellulose membrane. Immunodetection was carried out as described in Supplementary Data.

RESULTS AND DISCUSSION

Proteins co-purifying with SmAP1 and SmAP2 at a glance

To shed light on the physiological role(s) of Sso SmAP1/2 we started out to assess their interactome through the identification of co-purifying proteins. The genes encoding the SmAP1 and SmAP2 proteins were abutted at the 3′end with a 10×His-tag encoding sequence. Transcription of the genes from the virus-derived pMJ05-vector was controlled by an arabinose-inducible promoter in Sso PH1-16 (26). When compared with endogenous protein levels, this expression system permits a 3- to 10-fold overproduction of the respective proteins (48,55). Correspondingly, we observed ∼3 to 4-fold increase of the SmAP1 and SmAP2 levels after induction (Figure 1A). SmAP1-His and SmAP2-His were isolated from lysates of PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His), respectively, using stringent washing conditions. Unspecific binding to the affin-

ity matrix was controlled by a mock purification using cell lysates of strain PH1-16(pMJ05) (Supplementary Figure S1, lane 1). As Sm proteins display RNA-binding activity, the lysates were treated with DNase I, micrococcal nuclease and RNase A to minimize nucleic acid-mediated association of proteins with the SmAPs. After Ni-affinity purification, the co-purifying proteins (Supplementary Figure S1) were identified by mass spectrometry from two biological replicates (Table 1 and Supplementary Table S1). Protein identifications were accepted with a probability <99% and with a minimum of two unique peptides (Supplementary Data).

The majority of co-purifying proteins are involved in rRNA and tRNA modification and processing, RNA decay and translation (Table 1). The number of protein (Table 1) and RNA-interaction partners (26) suggests a multifunctional role of the SmAPs in RNA metabolism of crenarchaeota. It is worth noting that the identified proteins belong to similar or the same functional classes as the putative interaction partners identified for bacterial Hfq (56), for eukaryotic Lsm proteins (57–59) and for the SmAP of Hv (28). Additionally identified proteins with diverse or unknown function(s) as well as ribosomal proteins are listed in Supplementary Table S1.

Co-purification of SmAP1 and SmAP2

SmAP1 co-purified with SmAP2 and *vice versa* (Table 1). Thus, it is possible that the two SmAPs form either hetero-oligomeric assemblies or that homo-oligomeric proteins interact with each other. We favor the first possibility as both SmAPs displayed a similar pattern of co-purifying proteins (Supplementary Figures S1 and 2; Table 1 and Supplementary Table S1). In addition, the heptameric SmAP complexes $(SmAP1\text{-}His)_7$ and $(SmAP2\text{-}His)_7$ (Supplementary Figure S1A) isolated from polyacrylamide gels contained both SmAPs in a similar ratio as identified by mass spectrometry (Supplementary Figure S1B). Henceforth, we therefore use the designation SmAPX for the heptameric form of the SmAPs. Hetero-oligomeric SmAP assemblies have so far not been described in Archaea, and it remains to be seen whether a differential composition impacts their function(s) as shown for the hetero-oligomeric Lsm assemblies found in Eukaryotes (3). The third SmAP present in Sso (SmAP3/Sso*0276*) co-purified as well with SmAP1 and SmAP2 (Table 1). All three Sso SmAPs share a Lsm motif, whereas SmAP3 is larger in size (17 kDa) and additionally contains three more motifs, Lsm14, Sm-ATX and Lsm_C, that are often found in other Sm/Lsm protein family members (see Supplementary Figure S3).

Co-purification of proteins involved in RNA modification, turnover and translation

Proteins involved in methylation of rRNA e.g. Fibrillarin, Nop56 and Nep1 methyltransferase were co-purifying with the SmAPs (Table 1), which is in agreement with the recent observation that snoRNAs, which are components of the C/D box snoRNP, were found in complex with both SmAPs (26). The methyltransferase fibrillarin is involved in the first steps of pre-ribosomal processing and required for ribosome stability (60). Like fibrillarin, the Nop56 ortholog is also part of the nucleolar snoRNP. Nep1 has been described as a rRNA small subunit methyltransferase (61). These findings and the association of SmAPs with rRNA processing complexes (Table. 1) and rRNA (26) hints toward a function in pre-rRNA processing, which has been suggested before (2). This notion is supported by the co-migration of SmAPX with 50S and the monomeric form(s) of the SmAPs with 30S ribosomal subunits (Figure 1C) as well as by the co-purification with small and large ribosomal proteins (Supplementary Table S1).

The dimeric tRNA-splicing endonuclease (α- and β-subunit), which removes introns from archaeal pre-tRNAs (62) co-purified with the SmAPs (Table 1). Moreover, the two ALBA proteins (Sso*10b1*/Alba1; Sso*10b2*/Alba2) were enriched in both eluates (Table 1). ALBA proteins, which have been described as chromatin binding proteins (63), have also been shown to co-purify with SmAPs in halophilic Archaea (28). In addition, they share a high similarity with the Rpp20 subunit of eukaryotic nuclear, RNA-based RNase P involved in tRNA processing (64). Taken together with the findings that tRNAs co-purified with the Sso SmAPs (26) and that eukaryotic Lsm proteins are involved in tRNA processing (7), the co-purification of the tRNA-splicing endonuclease and the ALBA proteins with the SmAPs may hint to a role in tRNA processing.

Furthermore, we identified a putative 8.5-kDa ssh7a (Sso*10610*) endoribonuclease (65) and a putative metal-dependent phosphohydrolase co-purifying with both SmAPs (Table 1). The latter enzyme belongs to the HD superfamily, which constitutes a common domain for tRNA nucleotidyltransferases, poly(A)-polymerases and (p)ppGpp synthetase I (66).

Another abundant protein in the eluate was translation elongation factor 1-alpha (aEF1α). AEF1α is bound to the ribosome stalk (67) and acts as a carrier GTPase for tRNAs as well as for the tRNA mimicking proteins, archaeal release factor 1 and for aPelota. This suggests multiple roles for aEF1α in translational elongation and termination as well as in mRNA surveillance pathways (68), in which the SmAPs might be also involved. Analogously, in halophilic Archaea aEF1α co-purified with the SmAP (28). In Sso aEF1α was co-immunoprecipitated with DnaG and Rrp41 using DnaG-and Rrp41-specific antibodies, indicating that aEF1α is an interaction partner of the exosome (69). It is also worth noting that in Eukaryotes the Lsm proteins (Lsm1–4) have likewise been shown to co-purify with EF1α (70).

We further identified translation initiation factor aIF5A as putative interaction partner of Sm proteins (Table 1). The factor aIF5A is an ortholog of the eukaryotic eIF5A/bacterial EF-P proteins that promote translation of polyproline stretches in both kingdoms (71,72). In Haloarchaea it has been characterised as a ribonuclease (73), whereas the function in Sso remains elusive.

The confirmation of the interactions between the SmAPs and the proteins mentioned above as well as their physiological implications will have to be addressed in future studies. As the exosomal subunits DnaG, the cap-protein Rrp4 and the exosomal core subunits Rrp41 and Rrp42 co-purified with either both SmAPs or solely with SmAP2 (Table 1),

Table 1. Identification of proteins co-purifying with SmAP1 and SmAP2 by mass spectrometry

Sm proteins	ORF	MW	SmAP1	SmAP2
SmAP1	Sso*6454*	9 kDa	43	41
SmAP2	Sso*5410*	10 kDa	38	51
SmAP3	Sso*0276*	17 kDa	3	6
rRNA/tRNA modification and processing, RNA turnover, translation				
Fibrillarin	Sso*0940*	26 kDa	5	9
C/D box methylation guide RNP subunit aNOP56	Sso*0939*	47 kDa	7	27
Ribosomal RNA small subunit methyltransferase Nep1	Sso*2226*	26 kDa	25	13
tRNA-splicing endonuclease subunit α	Sso*0439*	21 kDa	5	3
tRNA-splicing endonuclease subunit β	Sso*0281*	20 kDa	4	4
Sso10b1 (ALBA) RNA-DNA binding	Sso*0962*	11 kDa	8	41
Sso10b2 (ALBA) RNA-DNA binding (Rpp20 homolog)	Sso*6877*	10 kDa	3	3
Endoribonuclease ssh7a	Sso*10610*	8 kDa	7	5
metal-dependent phosphohydrolase HD superfamily	Sso*0095*	47 kDa	7	3
Elongation factor 1-α	Sso*0216*	48 kDa	33	23
Translation initiation factor aIF5A	Sso*0970*	14 kDa	3	3
Exosome				
Bacterial-like DNA primase (DnaG)	Sso*0079*	45 kDa	3	10
Exosome complex subunit Rrp4	Sso*0736*	28 kDa	4	7
Exosome complex exonuclease Rrp41	Sso*0735*	27 kDa	0	4
Exosome complex subunit Rrp42	Sso*0732*	30 kDa	0	6

Proteins are only listed if they were detected in both biological replicates, i.e. when they were co-captured by either SmAP1 or SmAP2 in two independent experiments. Raw spectra were interpreted by Mascot 2.2.04 (Matrix Science). The spectral data were searched against the Archaea subset of the non-redundant protein database (NCBI). Results were further processed in Scaffold 3.0.2 (Proteome Software). Peptide identifications were accepted with a probability <95% as calculated by the Protein Prophet algorithm. Protein identifications were accepted with a probability <99% and with a minimum of two unique peptides. The corresponding open reading frame (ORF) and the molecular weight (MW) of the proteins are listed in the respective columns. The numbers in the SmAP1 and SmAP2 columns correspond to the number of identified peptides assignable to the identified proteins.

we focused in the present study on a possible link between the SmAPs and the exosome.

The SmAPs physically interact with the exosome via DnaG

For further analyzes, antibodies directed against both SmAPs were raised in rabbits and tested with the recombinant proteins for cross-reactivity (Supplementary Figure S4). The anti-SmAP1 serum was specific for SmAP1. In cell extracts (CE) both, the monomeric, SmAP1 and the SmAPX complexes were detected (Supplementary Figure S4A). The antibody raised against SmAP2, referred to as anti-SmAP1/2 antibody, recognized both SmAPs and showed a slightly reduced avidity for SmAP2 (Supplementary Figure S4B). Based on their size difference, the anti-SmAP1/2 antibody permitted the visualization of the monomeric forms of SmAP1 and SmAP2 (Supplementary Figure S4B), but the heptameric complexes could not be distinguished.

As mentioned above, when compared with the control strain (Figure 1A, left panel, lane 1), the input material used for the affinity purification contained ∼3 to 4-fold elevated levels of SmAP1 (Figure 1A, left panel, lane 2) and SmAP2 (Figure 1A, left panel, lane 3), respectively. The increased synthesis of the SmAPs had no influence on the levels of exosome components, i.e. DnaG, Rrp41 and Rrp4 (Figure 1A, left panel, lanes 1–3). The presence of the exosomal subunit DnaG in the eluates of affinity-purified SmAP1 and SmAP2, respectively, and the presence of Rrp41 and Rrp4 in the eluate of affinity-purified SmAP2 (Table 1) could be verified by immunodetection using antibodies against DnaG, Rrp41 and Rrp4 (Figure 1A, right panel lanes 2 and 3). The reason why higher quantities of DnaG co-purified with SmAP2 than with SmAP1 (Figure 1A, right panel, lanes 2 and 3) remains to be clarified. However, purely based on a qualitative scale, the latter result concurred with the mass spectrometry data as more DnaG-specific peptides were detected in the SmAP2-eluate (Table 1). As DnaG is tightly associated with the core-exosomal subunits (49), the lower abundance of DnaG in the SmAP1 eluate might explain why Rrp41 and Rrp4 could not be detected among the proteins co-purifying with SmAP1 (Figure 1A, right panel, lane 2).

In addition, DnaG-His was used as bait protein and the reverse experiment was performed with strain PH1-16(pMJ05-DnaG-His). The induction of the *dnaG-His* variant resulted in an increase of the intracellular DnaG levels (∼3-fold) (Figure 1B, left panel, lane 2). Optimization of the resolution conditions revealed that only DnaG-His was increased 3-fold, whereas the endogenous DnaG level was unchanged (Supplementary Figure S5). The increase in the DnaG-levels were also accompanied by a minor increase in the Rrp41 levels, whereas the Rrp4, SmAP1 and SmAP2 levels remained unchanged (Figure 1B, left panel, lane 2). The His-tagged DnaG was immobilized on a Ni-affinity column. The reverse experiment mirrored the results shown in Figure 1A in that Rrp41, Rrp4, SmAP1 and SmAP2 co-purified with DnaG (Figure 1B, right panel, lane 2).

The Sso exosome was shown to co-migrate with the ribosomal fraction (47,69). As shown in Figure 1C, like the exosomal subunits DnaG and Rrp41, SmAPX predominantly co-migrated together with 50S ribosomes. In addition, substantial amounts of monomeric SmAPs were detected in the top/soluble fraction where it is apparently not associated with the exosome (Figure 1C).

Clearly, these initial co-purification and co-migration experiments did not exclude the possibility of an indirect association of DnaG with the SmAPs. To demonstrate a direct

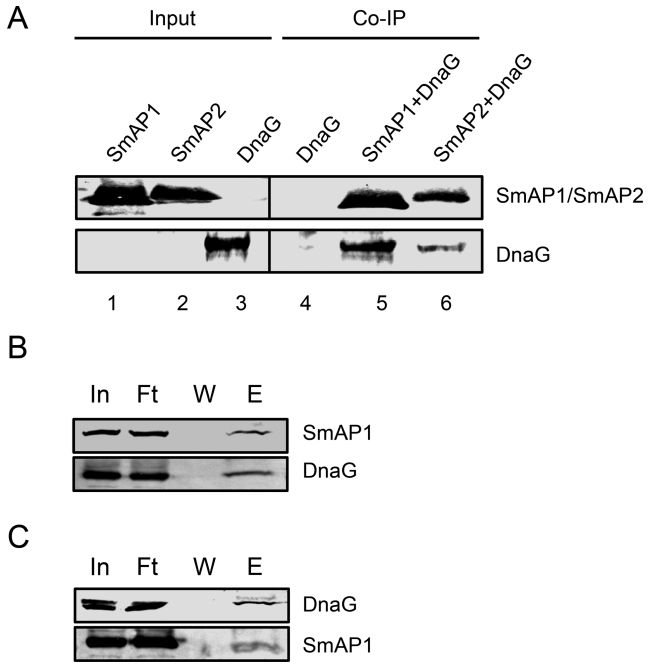

Figure 2. The SmAPs physically interact with DnaG. (**A**) Co-immunoprecipitation (Co-IP) with anti-SmAP1/2 antibodies. SmAP1, SmAP2 and DnaG were purified from *Eandscherichia coli*. Lanes 1–3, input of SmAP1 (lane 1), SmAP2 (lane 2) and DnaG (lane 3) used for the Co-IP assays. Lanes 4–6, Co-IP assays with DnaG alone (mock-control, lane 4), SmAP1 in the presence of DnaG (lane 5) or SmAP2 in the presence of DnaG (lane 6). The proteins were incubated for 10 min at 65°C and the anti-SmAP1/2 antibodies were used in the Co-IP assay. ProteinG-Dynabeads were used to capture the immunocomplexes. After washing, the proteins were eluted with sodium dodecyl sulphate (SDS)-loading buffer and analyzed by western-blotting using anti-SmAP1/2- and anti-DnaG-specific antibodies. (**B**) Co-IP with anti-SmAP1-antibodies. A cell extract (CE) derived from the Sso wild-type strain P2 was incubated with anti-SmAP1 antibodies to capture endogenous SmAP1 protein. The immunocomplexes were then immobilized on ProteinG-Dynabeads, washed and then eluted with 0.1 M glycine, pH 2.0. Ten microliter of the cell lysates (In) and the flowthrough/unbound fraction (Ft) was loaded together with the TCA-precipitated wash fraction (W) and 1/4 of the eluate (E). The fractions were then analyzed by western blotting using anti-SmAP1- or anti-DnaG-specific antibodies. (**C**) Co-IP with anti-DnaG-antibodies. A cell extract derived from the Sso wild-type strain P2 was incubated anti-DnaG antibodies to capture endogenous DnaG protein. Co-IP and western-blotting were performed as described in (B).

physical interaction His-tagged SmAP1, SmAP2 and DnaG were produced in *E. coli* and purified to homogeneity using Ni-affinity and size-exclusion chromatography (26,37). The recombinant His-tagged proteins (Figure 2A, lanes 1–3 (Input)) were used for the *in vitro* Co-IP experiments together with the anti-SmAP1/2 antibodies. First, a mock experiment was performed with DnaG and anti-SmAP1/2 antibodies. As shown in Figure 2A, lane 4, this experiment showed on the one hand that the anti-SmAP1/2 antibodies did not cross-react with DnaG and on the other hand eliminated the possibility that DnaG binds non-specifically to the ProteinG beads. To demonstrate a physical interaction with DnaG, either SmAP1-His or SmAP2-His were incubated together with DnaG-His and then the Co-IP was performed with anti-SmAP1/2 antibodies (Figure 2A, lanes 5 and 6). The precipitates were analyzed by western-blotting using anti-SmAP1/2- (Figure 2A, upper panel, lanes 5 and 6) and anti-DnaG antibodies (Figure 2A, lower panel, lanes 5 and 6). DnaG was co-captured with Sso-SmAP1 and to a lower extend with Sso-SmAP2 (Figure 2A, lane 5 and 6). We hypothesize that this due to the lower avidity of the anti-SmAP1/2 antibody for SmAP2 than for SmAP1 (see Supplementary Figure S4B), which likely results in a less efficient pull-down of SmAP2-His/DnaG-His complexes. The Co-IP experiment with isolated components clearly indicated a direct interaction between the SmAPs and DnaG *in vitro* although we cannot completely exclude that their binding properties are influenced by the addition of the His-tag.

To further demonstrate the interaction of endogenous SmAPs and DnaG in Sso lysates we employed the anti-SmAP1 antibody, specifically recognizing SmAP1 (Supplementary Figure S4A) and the anti-DnaG antibody to test whether DnaG co-precipitates with SmAP1 and SmAP1 co-precipitates with DnaG, respectively. In either experiment, we could confirm an interaction of SmAP1 and DnaG (Figure 2B and C), which again mirrored the co-migration studies (Figure 1C).

Elevated levels of the SmAPs increase the amounts of soluble exosome and of A-rich tails on RNAs

Since DnaG might be required for membrane localization of the Sso exosome (47), we next assessed whether an increase in the intracellular concentration of both SmAPs affects the cellular partitioning (soluble versus insoluble) of the exosome. We utilized the Sso strains PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His) to achieve elevated levels of the SmAPs. When compared to the control samples (non-induced), induction of the plasmid borne SmAP genes with arabinose resulted in ∼3 to 4-fold increase of SmAP1 and SmAP2, respectively (Figure 3A). The cell lysates were prepared from non-induced (SmAP1−/SmAP2−) and induced cultures (SmAP1+/SmAP2+) as described before (69) with the exception that the lysates were centrifuged at 130 000 *g* for 2 h. This led to a complete removal of the β- and γ-subunits of archaeal translation initiation factor aIF2 from the supernatant (S130) (Figure 3B), which most likely can be ascribed to its association with ribosomes present in the pellet (P130). Likewise, in non-induced cells the exosomal subunits Rpr41 and DnaG were only found in the P130 fraction (Figure 3B). In contrast, the increase of the intracellular SmAP levels (Figure 3A) concurred with the immunodetection of DnaG and Rpr41 in the S130 fraction (Figure 3B), indicating that elevated SmAP levels increased the abundance of the soluble exosome.

Next, we asked whether increased levels of either SmAP1 or SmAP2 can affect the intracellular partitioning of the exosome, i.e. the fractions of soluble/insoluble exosome. In contrast to the results shown in Figure 1C and as observed for the exosome pelleting assay (Figure 3B), a fraction of DnaG and Rrp41 shifted in 10–30% sucrose gradients toward the soluble fractions upon over-production of SmAP1 and SmAP2 (Figure 3C), respectively, whereas the main part of the insoluble portion remained unaltered. This effect was more pronounced in the SmAP2 over-producing strain (Figure 3B and C), which might be attributed to the

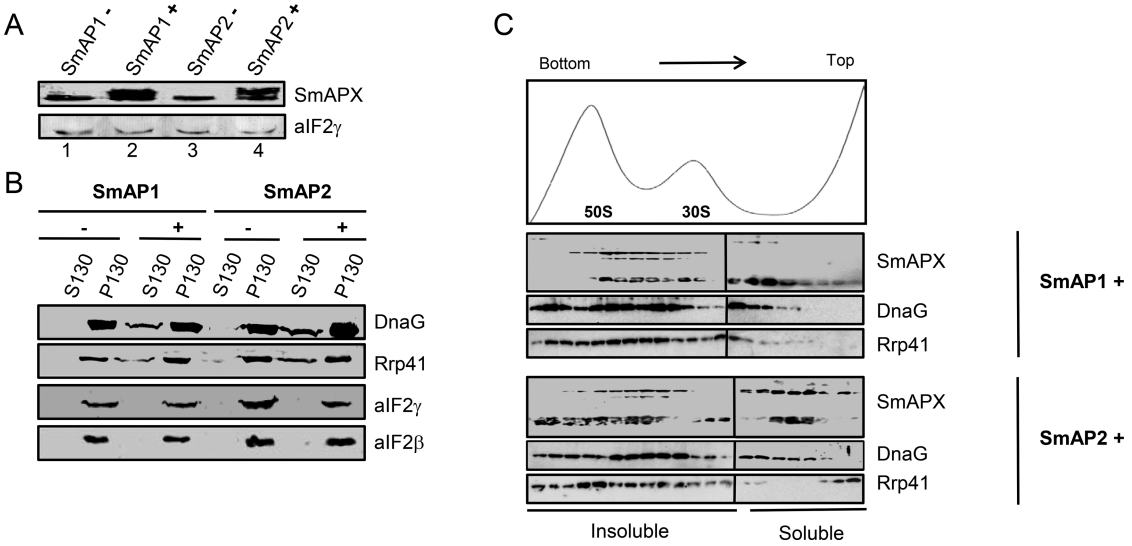

Figure 3. Increasing SmAP levels concur with an increase of the soluble exosome. (**A**) The Sso strains PH1-16(pMJO5-SmAP1-His) and PH1-16(pMJO5-SmAP2-His) were grown at 75°C in the presence of sucrose (−, non-induced) or arabinose (+, induced). The respective cells were lysed and the levels of SmAP1 (lanes 1–2) and SmAP2 (lanes 3–4) were assessed by quantitative western-blotting with anti-SmAP1/2 antibodies. To ensure equal loading the membranes were also probed with anti-aIF2γ antibodies. (**B**) Exosome pelleting assay. A total of 750 μl of the cell lysates derived from the Sso strains PH1-16(pMJO5-SmAP1-His) and PH1-16(pMJO5-SmAP2-His) grown at 75°C in the presence of sucrose (−, non-induced) or arabinose (+, induced) were centrifuged for 2 h at 130 000 *g*. The supernatant (S130) was removed and the pellet (P130) was dissolved in an equivalent volume. Equal volumes of S130 and P130 were analyzed by western-blotting for the presence of DnaG and Rrp41 using the respective antibodies. To ensure equal loading the membranes were also probed with anti-aIF2β and anti-aIF2γ antibodies. (**C**) Sucrose gradients. A total of 750 μl of the cell lysates obtained from strains PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His) grown in the presence of arabinose were layered on top of a linear 10–30% sucrose gradient. A total of 500 μl samples were collected and the OD_{260} was measured to determine the positions of 50S and 30S ribosomal subunits (upper panel). The proteins were TCA precipitated, separated on a 15% SDS polyacrylamide gel and then transferred to a nitrocellulose membrane. The heptameric (upper bands; SmAPX) and monomeric forms (lower bands), the DnaG-specific band and the Rrp41-specific band were detected with anti-SmAP1/2-, anti-DnaG- and anti-Rrp41- specific antibodies, respectively.

strength of the SmAP2-DnaG interaction (Figure 1A). In addition, when compared with the wild-type cells, where only the monomeric SmAPs were detected in the soluble fraction (Figure 1C), augmented levels of the SmAPs resulted in an increase of the SmAPX complexes in the soluble fraction (Figure 3C). This was again more pronounced in the strain with elevated SmAP2 levels.

DnaG is required for efficient polyadenylation of rRNA substrates (37). Considering the results shown in Figure 3B and C, we next tested whether the increase of the soluble/active exosome (33,69) upon over-production of the SmAPs impacts on the adenylation state of cellular RNAs. Total RNA was isolated from the Sso strains PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJ05-SmAP2-His) grown in the presence of sucrose (non-induced/SmAP-) and arabinose (induced/SmAP+). When equal amounts of these RNA preparations were passed over an oligo-T column, by far more A-rich RNAs were retained from the cultures that contained elevated levels of the SmAPs (Figure 4A). Most of the eluted RNAs isolated from the SmAP1 and SmAP2 over-producing strains had a size between 750 and 3000 nt.

Next, we used RNA_{Seq} to identify the most abundant RNAs present in the eluates of the oligo-T affinity purification and to determine the location of A-rich stretches longer than 15 nt within a given gene, at its 3′ end as well as in intergenic regions as outlined in 'Materials and Methods' section. These analyses (Supplementary Tables S2 and 3) revealed 324 and 327 transcripts (241 transcripts were in common) with A-rich stretches upon over-production of SmAP1 and SmAP2, respectively. Among these, 83 and 86 transcripts were specifically detected upon over-production of SmAP1 and SmAP2 (Supplementary Figure S6), respectively. The A-rich stretches were predominantly located in the coding region of the transcripts (80.9% for SmAP1 and 76.9% for SmAP2), whereas ∼10% were localized in the 3′ UTRs (10.53% for SmAP1 and 10.6% for SmAP2). In addition, some A-rich stretches were also detected in intergenic regions (8.6% for SmAP1 and 12.5% for SmAP2). Moreover, the two previously identified adenylated RNAs in Sso (74), viz 16S rRNA and the *nuoH* mRNA, were found among the tailed transcripts. The sequence composition of the tails were analyzed with WebLogo (51). They were predominantly A-rich with a noticeable G content (Figure 4B). The categorization of the transcripts revealed several functions that were significantly (q-value < 0.05) enriched. These included genes encoding functions involved in transcription, ribosomal structure and biogenesis, translation, post-translational modification, protein folding and turnover, and energy production and conversion (Figure 4C).

The increase in adenylated RNAs led to the question whether the SmAPs directly stimulate the tailing activity of the exosome or whether it is caused indirectly by increasing the level of the soluble/active (33,69) exosome (Figure 3). To address this, an *in vitro* polyadenylation assay with the native 3′ end of the Sso 16S rRNA was performed (37,74). The 163 nt long 3′ end of the Sso 16S rRNA was incubated

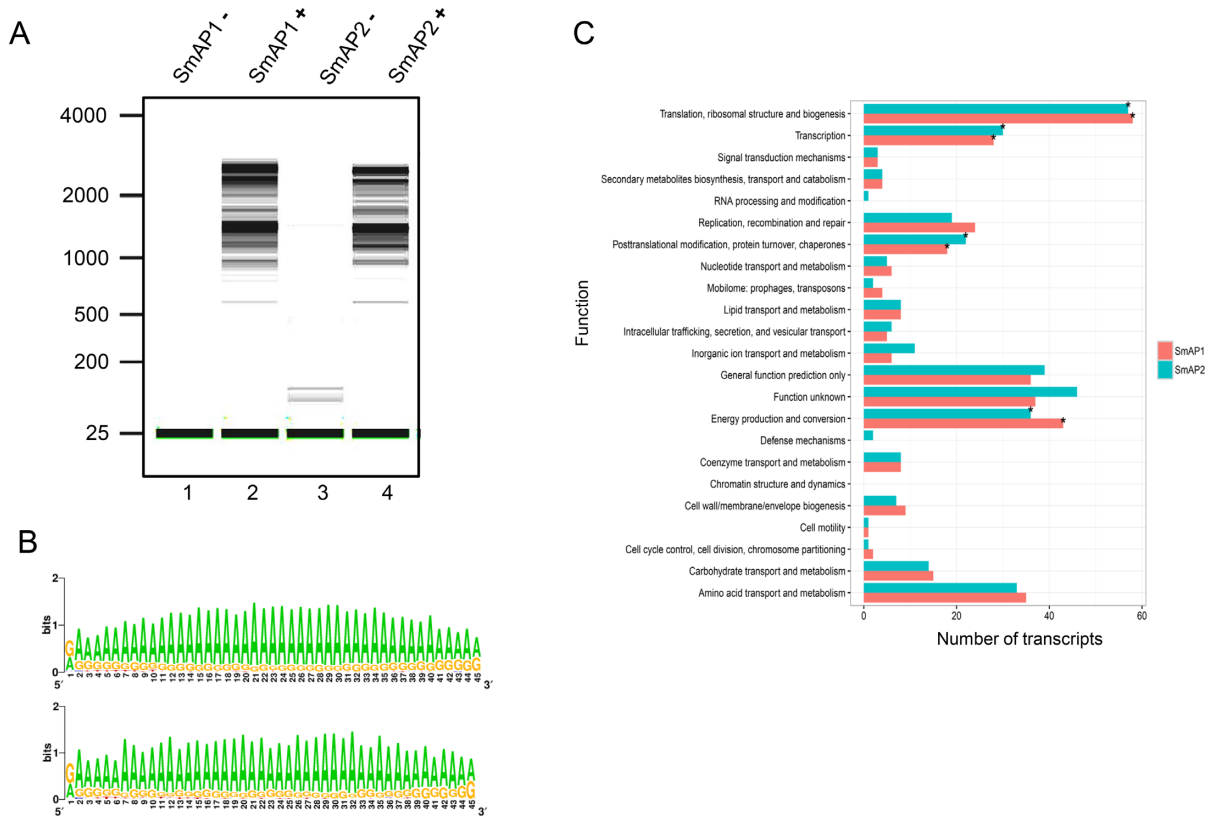

Figure 4. Elevated levels of the SmAPs increase the abundance of RNAs with A-rich tails. (**A**) Total RNA was isolated from strains PH1-16(pMJ05-SmAP1-His) and PH1-16(pMJO5-SmAP2-His) grown either in the presence of sucrose (−, non-induced) (lanes 1 and 3) or arabinose (+, induced) (lanes 2 and 4). Equal amounts of total RNA were used to isolate adenylated RNA with the Oligotex™ kit. Two microliter of each eluate was analyzed using with the Agilent 2100 Bioanalyzer (Agilent Technologies) and the RNA 6000 Nano Kit (Agilent Technologies). (**B**) The sequence composition of the A-rich tails obtained after over-production of SmAP1 (top) and SmAP2 (bottom) was determined using WebLogo (51). Only 3′ ends of the adaptor clipped reads, which do not map to the reference genome and which showed an overhang of longer than 15 nt were used for the analyzes. Only sites present in both replicas that are supported by at least five independent reads were analyzed. (**C**) Functional categorization of tailed RNAs. Functions, which are significant enriched (Fisher's exact test; $\alpha = 0.05$) are marked with an asterisk. Genes are annotated according to (54).

with the Csl4-Rpr41-Rpr42 exosome and the DnaG-Csl4-Rpr41-Rpr42 exosome in the absence and presence of the SmAP1 or SmAP2. The Csl4-Rpr41-Rpr42 exosome has been reported to be devoid of tailing activity on the 16S rRNA substrate (37). The presence of the SmAPs did not affect this trait (Supplementary Figure S7, lanes 1–7). In contrast, the DnaG-Csl4-Rpr41-Rpr42 exosome enabled tailing of the 16S rRNA substrate (Supplementary Figure S7, lanes 8–10). However, the tailing activity did not increase after addition of the SmAPs proteins to the adenylation assay (Supplementary Figure S7, lanes 11–12). Although we cannot exclude that this preliminary result is inherent to the substrate used, we favor the idea that the increase of 'A-rich' RNAs observed after *in vivo* over-production of the SmAPs results from elevated levels of the soluble/active exosome fraction (Figure 3B and C).

The more prominent 'A-rich tailing' in coding regions resembles that observed for bacterial mRNAs, where it may serve as a toehold for and accelerate 3′ to 5′ directional decay of the transcript by exoribonucleases (16,17). However, in Sso the A-rich tails seem to be longer as a minimum of 15 nt was used as a threshold for the A-rich stretches. The predominant presence of the A-rich stretches in coding regions might imply that tailing occurs on decay intermediates generated by endonucleolytic cleavage. Such a mechanism has been proposed for RNase E in *E. coli*, where the enzyme appears to affect poly(A) tailing indirectly through the generation of new 3′ termini that serve as substrates for poly(A) polymerase (17). We have recently identified in Sso and in *Sulfolobus acidocaldarius* aCPSF2 exoribonucleases with 5′ to 3′ directionality belonging to the group of β-CASP proteins (75). However, in contrast to other members of the β-CASP ribonucleases (76), no endonucleolytic activity has as yet been demonstrated for Sso-aCPSF2. Apart from CRISPR Cas6 (77) and a tRNA splicing endoribonuclease (62) the only other riboendonuclease described in Sso is Sso7d (65). Interestingly Sso7a, a paralog of Sso7d, as well as two subunits of the t-RNA splicing endonuclease co-purified with both SmAPs (Table 1). Whether these observations have any meaning with regard to 'A-rich tailing' within coding regions remains to be seen.

Alternatively, tailing within coding regions might result from 3′ to 5′ curtailing by the exosome followed by 'A-rich tailing'. This mode of action has been described for *E. coli* polynucleotide phosphorylase, which works exonucleolytically and biosynthetically at high and low intracellular levels of inorganic phosphate, respectively (16). Similary, the Sso exosome displays a degradative and a biosynthetic ac-

tivity in the presence of increased levels of inorganic phosphate and ADP, respectively (33). Clearly, the physiological consequence of the 'A-rich tailing' remains to be elucidated in Sso. Nevertheless, there are some indications that, like in Bacteria, tailed RNAs are degraded faster by the Sso exosome *in vitro* (37).

Conclusions and perspectives

This study disclosed a novel function of SmAPs in crenarchaeal RNA metabolism in that the SmAPs bind to the archaeal exosome subunit DnaG and thereby seem to impact indirectly on the adenylation status of RNAs. Although one explanation for this finding could be an increase in the soluble/active form of the exosome (33,69), it remains puzzling why over-production of either SmAP resulted not only in tailing of a common set of RNA substrates but also of distinct ones.

In Bacteria and Eukarya, A-rich/poly-(A) tailing is linked with a short (20) and increased (12) longevity of a mRNA, respectively. Having identified distinct mRNAs that are tailed within the coding region and at their 3′ extremities opens up the possibility to address, for the first time, the consequences of A-rich tailing in an archaeon. Furthermore, the interaction of the SmAPs and DnaG could influence the degradation activity of the exosome either directly or by changing its binding affinity for poly(A). How the activity of DnaG is regulated by the SmAPs requires further experimentation.

Although it appears safe to say that the archaeal SmAP/Lsm proteins are involved in RNA metabolism there could be subtle functional differences in different archaeal phyla. For instance in halophilic Archaea the exosome is absent and RNA tailing does not occur (74), whereas DnaG is present (37). It seems therefore worthwhile to study the function of Sm proteins in different phyla to uncover unprecedented functions of archaeal SmAP/Lsm proteins.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are indebted to Drs S.V. Albers, B. Mayer and C. Witharana for materials and helpful discussions. The technical support of Diego Oxilia Speratti and the MFPL Mass Spectrometry Facility is acknowledged.

FUNDING

Austrian Science Fund (FWF) [P2288-B20 to U.B.]; Deutsche Forschungsgemeinschaft [Kl563/27 to E.E.-H.]. Funding for open access charge: Austrian Science Fund (FWF) [P2288-B20].
Conflict of interest statement. None declared.

REFERENCES

1. Salgado-Garrido,J., Bragado-Nilsson,E., Kandels-Lewis,S. and Seraphin,B. (1999) Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J.*, **18**, 3451–3462.
2. Mura,C., Randolph,P.S., Patterson,J. and Cozen,A.E. (2013) Archaeal and eukaryotic homologs of Hfq: a structural and evolutionary perspective on Sm function. *RNA Biol.*, **10**, 636–651.
3. Wilusz,C.J. and Wilusz,J. (2005) Eukaryotic Lsm proteins: lessons from bacteria. *Nat. Struct. Mol. Biol.*, **12**, 1031–1036.
4. Wilusz,C.J. and Wilusz,J. (2013) Lsm proteins and Hfq: life at the 3′ end. *RNA Biol.*, **10**, 592–601.
5. Beggs,J.D. (2005) Lsm proteins and RNA processing. *Biochem. Soc. Trans.*, **33**, 433–438.
6. Tharun,S. (2009) Roles of eukaryotic Lsm proteins in the regulation of mRNA function. *Int. Rev. Cell Mol. Biol.*, **272**, 149–189.
7. Kufel,J., Allmang,C., Verdone,L., Beggs,J.D. and Tollervey,D. (2002) Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. *Mol. Cell. Biol.*, **22**, 5248–5256.
8. Kufel,J., Bousquet-Antonelli,C., Beggs,J.D. and Tollervey,D. (2004) Nuclear pre-mRNA decapping and 5′ degradation in yeast require the Lsm2-8p complex. *Mol. Cell. Biol.*, **24**, 9646–9657.
9. He,W. and Parker,R. (2001) The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3′ termini from partial degradation. *Genetics*, **158**, 1445–1455.
10. Tharun,S., Muhlrad,D., Chowdhury,A. and Parker,R. (2005) Mutations in the *Saccharomyces cerevisiae* LSM1 gene that affect mRNA decapping and 3′ end protection. *Genetics*, **170**, 33–46.
11. Tharun,S. (2009) Lsm1-7-Pat1 complex: a link between 3′ and 5′-ends in mRNA decay? *RNA Biol.*, **6**, 228–232.
12. Tharun,S., He,W., Mayes,A.E., Lennertz,P., Beggs,J.D. and Parker,R. (2000) Yeast Sm-like proteins function in mRNA decapping and decay. *Nature*, **404**, 515–518.
13. Wagner,E.G. (2013) Cycling of RNAs on Hfq. *RNA Biol.*, **10**, 619–626.
14. Vogel,J. and Luisi,B.F. (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.*, **9**, 578–589.
15. Regnier,P. and Hajnsdorf,E. (2013) The interplay of Hfq, poly(A) polymerase I and exoribonucleases at the 3′ ends of RNAs resulting from Rho-independent termination: a tentative model. *RNA Biol.*, **10**, 602–609.
16. Mohanty,B.K. and Kushner,S.R. (2000) Polynucleotide phosphorylase functions both as a 3′ to 5′ exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11966–11971.
17. Mohanty,B.K. and Kushner,S.R. (2000) Polynucleotide phosphorylase, RNase II and RNase E play different roles in the *in vivo* modulation of polyadenylation in *Escherichia coli*. *Mol. Microbiol.*, **36**, 982–994.
18. Mohanty,B.K., Maples,V.F. and Kushner,S.R. (2012) Polyadenylation helps regulate functional tRNA levels in Escherichia coli. *Nucleic Acids Res.*, **40**, 4589–4603.
19. Silva,I.J., Saramago,M., Dressaire,C., Domingues,S., Viegas,S.C. and Arraiano,C.M. (2011) Importance and key events of prokaryotic RNA decay: the ultimate fate of an RNA molecule. *Wiley Interdiscip. Rev. RNA*, **2**, 818–836.
20. Mohanty,B.K., Maples,V.F. and Kushner,S.R. (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. *Mol. Microbiol.*, **54**, 905–920.
21. Obregon,K.A., Hoch,C.T. and Sukhodolets,M.V. (2015) Sm-like protein Hfq: composition of the native complex, modifications, and interactions. *Biochim. Biophys. acta*, **1854**, 950–966.
22. Collins,B.M., Harrop,S.J., Kornfeld,G.D., Dawes,I.W., Curmi,P.M. and Mabbutt,B.C. (2001) Crystal structure of a heptameric Sm-like protein complex from archaea: implications for the structure and evolution of snRNPs. *J. Mol. Biol.*, **309**, 915–923.
23. Mura,C., Cascio,D., Sawaya,M.R. and Eisenberg,D.S. (2001) The crystal structure of a heptameric archaeal Sm protein: implications for the eukaryotic snRNP core. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 5532–5537.
24. Törö,I., Thore,S., Mayer,C., Basquin,J., Seraphin,B. and Suck,D. (2001) RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J.*, **20**, 2293–2303.
25. Törö,I., Basquin,J., Teo-Dreher,H. and Suck,D. (2002) Archaeal Sm proteins form heptameric and hexameric complexes: crystal structures of the Sm1 and Sm2 proteins from the hyperthermophile *Archaeoglobus fulgidus*. *J. Mol. Biol.*, **320**, 129–142.

26. Märtens,B., Bezerra,G.A., Kreuter,M.J., Grishkovskaya,I., Manica,A., Arkhipova,V., Djinovic-Carugo,K. and Bläsi,U. (2015) The Heptameric SmAP1 and SmAP2 proteins of the Crenarchaeon *Sulfolobus Solfataricus* bind to common and distinct RNA targets. *Life (Basel)*, **5**, 1264–1281.
27. Achsel,T., Stark,H. and Luhrmann,R. (2001) The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 3685–3689.
28. Fischer,S., Benz,J., Spath,B., Maier,L.K., Straub,J., Granzow,M., Raabe,M., Urlaub,H., Hoffmann,J., Brutschy,B. *et al.* (2010) The archaeal Lsm protein binds to small RNAs. *J. Biol. Chem.*, **285**, 34429–34438.
29. Nikulin,A., Mikhailina,A., Lekontseva,N., Balobanov,V., Nikonova,E. and Tishchenko,S. (2016) Characterization of RNA-binding properties of the archaeal Hfq-like protein from *Methanococcus jannaschii*. *J. Biomol. Struct. Dyn.*, **35**, 1615–1628.
30. Mura,C., Kozhukhovsky,A., Gingery,M., Phillips,M. and Eisenberg,D. (2003) The oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs). *Protein Sci.*, **12**, 832–847.
31. Thore,S., Mayer,C., Sauter,C., Weeks,S. and Suck,D. (2003) Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya. *J. Biol. Chem.*, **278**, 1239–1247.
32. Maier,L.K., Benz,J., Fischer,S., Alstetter,M., Jaschinski,K., Hilker,R., Becker,A., Allers,T., Soppa,J. and Marchfelder,A. (2015) Deletion of the Sm1 encoding motif in the lsm gene results in distinct changes in the transcriptome and enhanced swarming activity of *Haloferax* cells. *Biochimie*, **117**, 129–137.
33. Evguenieva-Hackenberg,E. (2010) The archaeal exosome. *Adv. Exp. Med. Biol.*, **702**, 29–38.
34. Lorentzen,E., Walter,P., Fribourg,S., Evguenieva-Hackenberg,E., Klug,G. and Conti,E. (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nat. Struct. Mol. Biol.*, **12**, 575–581.
35. Büttner,K., Wenig,K. and Hopfner,K.P. (2005) Structural framework for the mechanism of archaeal exosomes in RNA processing. *Mol. cell*, **20**, 461–471.
36. Roppelt,V., Klug,G. and Evguenieva-Hackenberg,E. (2010) The evolutionarily conserved subunits Rrp4 and Csl4 confer different substrate specificities to the archaeal exosome. *FEBS Lett.*, **584**, 2931–2936.
37. Hou,L., Klug,G. and Evguenieva-Hackenberg,E. (2014) Archaeal DnaG contains a conserved N-terminal RNA-binding domain and enables tailing of rRNA by the exosome. *Nucleic Acids Res.*, **42**, 12691–12706.
38. Parker,R. (2012) RNA degradation in *Saccharomyces cerevisae*. *Genetics*, **191**, 671–702.
39. Liou,G.G., Jane,W.N., Cohen,S.N., Lin,N.S. and Lin-Chao,S. (2001) RNA degradosomes exist in vivo in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 63–68.
40. Montero Llopis,P., Jackson,A.F., Sliusarenko,O., Surovtsev,I., Heinritz,J., Emonet,T. and Jacobs-Wagner,C. (2010) Spatial organization of the flow of genetic information in bacteria. *Nature*, **466**, 77–81.
41. Deutscher,M.P. (2015) How bacterial cells keep ribonucleases under control. *FEMS Microbiol. Rev.*, **39**, 350–361.
42. Strahl,H., Turlan,C., Khalid,S., Bond,P.J., Kebalo,J.M., Peyron,P., Poljak,L., Bouvier,M., Hamoen,L., Luisi,B.F. *et al.* (2015) Membrane recognition and dynamics of the RNA degradosome. *PLoS Genet.*, **11**, e1004961.
43. Bandyra,K.J., Bouvier,M., Carpousis,A.J. and Luisi,B.F. (2013) The social fabric of the RNA degradosome. *Biochim. Biophys. Acta*, **1829**, 514–522.
44. Fortas,E., Piccirilli,F., Malabirade,A., Militello,V., Trepout,S., Marco,S., Taghbalout,A. and Arluison,V. (2015) New insight into the structure and function of Hfq C-terminus. *Biosci. Rep.*, **35**, doi:10.1042/BSR20140128.
45. Taghbalout,A., Yang,Q. and Arluison,V. (2014) The Escherichia coli RNA processing and degradation machinery is compartmentalized within an organized cellular network. *Biochem. J.*, **458**, 11–22.
46. Azam,T.A. and Ishihama,A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.*, **274**, 33105–33113.
47. Roppelt,V., Hobel,C.F., Albers,S.V., Lassek,C., Schwarz,H., Klug,G. and Evguenieva-Hackenberg,E. (2010) The archaeal exosome localizes to the membrane. *FEBS Lett.*, **584**, 2791–2795.
48. Albers,S.V., Jonuscheit,M., Dinkelaker,S., Urich,T., Kletzin,A., Tampe,R., Driessen,A.J. and Schleper,C. (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol.*, **72**, 102–111.
49. Hou,L., Klug,G. and Evguenieva-Hackenberg,E. (2013) The archaeal DnaG protein needs Csl4 for binding to the exosome and enhances its interaction with adenine-rich RNAs. *RNA Biol.*, **10**, 415–424.
50. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet*, **17**, 10–12.
51. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
52. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
53. Wolfinger,M.T., Fallmann,J., Eggenhofer,F. and Amman,F. (2015) ViennaNGS: a toolbox for building efficient next- generation sequencing analysis pipelines. *F1000Research*, **4**, 50.
54. Esser,D., Kouril,T., Zaparty,M., Sierocinski,P., Chan,P.P., Lowe,T., Van der Oost,J., Albers,S.V., Schomburg,D., Makarova,K.S. *et al.* (2011) Functional curation of the *Sulfolobus solfataricus* P2 and S. acidocaldarius 98-3 complete genome sequences. *Extremophiles*, **15**, 711–712.
55. Hasenöhrl,D., Lombo,T., Kaberdin,V., Londei,P. and Bläsi,U. (2008) Translation initiation factor a/eIF2(-gamma) counteracts 5′ to 3′ mRNA decay in the archaeon *Sulfolobus solfataricus*. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 2146–2150.
56. Butland,G., Peregrin-Alvarez,J.M., Li,J., Yang,W., Yang,X., Canadien,V., Starostine,A., Richards,D., Beattie,B., Krogan,N. *et al.* (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*, **433**, 531–537.
57. Uetz,P., Giot,L., Cagney,G., Mansfield,T.A., Judson,R.S., Knight,J.R., Lockshon,D., Narayan,V., Srinivasan,M., Pochart,P. *et al.* (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623–627.
58. Ho,Y., Gruhler,A., Heilbut,A., Bader,G.D., Moore,L., Adams,S.L., Millar,A., Taylor,P., Bennett,K., Boutilier,K. *et al.* (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, **415**, 180–183.
59. Krogan,N.J., Peng,W.T., Cagney,G., Robinson,M.D., Haw,R., Zhong,G., Guo,X., Zhang,X., Canadien,V., Richards,D.P. *et al.* (2004) High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell*, **13**, 225–239.
60. Rodriguez-Corona,U., Sobol,M., Rodriguez-Zapata,L.C., Hozak,P. and Castano,E. (2015) Fibrillarin from Archaea to human. *Biol. Cell*, **107**, 159–174.
61. Taylor,A.B., Meyer,B., Leal,B.Z., Kotter,P., Schirf,V., Demeler,B., Hart,P.J., Entian,K.D. and Wohnert,J. (2008) The crystal structure of Nep1 reveals an extended SPOUT-class methyltransferase fold and a pre-organized SAM-binding site. *Nucleic Acids Res.*, **36**, 1542–1554.
62. Kleman-Leyer,K., Armbruster,D.W. and Daniels,C.J. (1997) Properties of H. volcanii tRNA intron endonuclease reveal a relationship between the archaeal and eucaryal tRNA intron processing systems. *Cell*, **89**, 839–847.
63. Bell,S.D., Botting,C.H., Wardleworth,B.N., Jackson,S.P. and White,M.F. (2002) The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science*, **296**, 148–151.
64. Aravind,L., Iyer,L.M. and Anantharaman,V. (2003) The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. *Genome Biol.*, **4**, R64.
65. Shehi,E., Serina,S., Fumagalli,G., Vanoni,M., Consonni,R., Zetta,L., Deho,G., Tortora,P. and Fusi,P. (2001) The Sso7d DNA-binding protein from *Sulfolobus solfataricus* has ribonuclease activity. *FEBS Lett.*, **497**, 131–136.
66. Aravind,L. and Koonin,E.V. (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.*, **23**, 469–472.

67. Ito,K., Honda,T., Suzuki,T., Miyoshi,T., Murakami,R., Yao,M. and Uchiumi,T. (2014) Molecular insights into the interaction of the ribosomal stalk protein with elongation factor 1alpha. *Nucleic Acids Res.*, **42**, 14042–14052.
68. Saito,K., Kobayashi,K., Wada,M., Kikuno,I., Takusagawa,A., Mochizuki,M., Uchiumi,T., Ishitani,R., Nureki,O. and Ito,K. (2010) Omnipotent role of archaeal elongation factor 1 alpha (EF1alpha in translational elongation and termination, and quality control of protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 19242–19247.
69. Witharana,C., Roppelt,V., Lochnit,G., Klug,G. and Evguenieva-Hackenberg,E. (2012) Heterogeneous complexes of the RNA exosome in *Sulfolobus solfataricus*. *Biochimie*, **94**, 1578–1587.
70. Sobti,M., Cubeddu,L., Haynes,P.A. and Mabbutt,B.C. (2010) Engineered rings of mixed yeast Lsm proteins show differential interactions with translation factors and U-rich RNA. *Biochemistry*, **49**, 2335–2345.
71. Gutierrez,E., Shin,B.S., Woolstenhulme,C.J., Kim,J.R., Saini,P., Buskirk,A.R. and Dever,T.E. (2013) eIF5A promotes translation of polyproline motifs. *Mol. Cell*, **51**, 35–45.
72. Ude,S., Lassak,J., Starosta,A.L., Kraxenberger,T., Wilson,D.N. and Jung,K. (2012) Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science*, **339**, 82–85.
73. Wagner,S. and Klug,G. (2007) An archaeal protein with homology to the eukaryotic translation initiation factor 5A shows ribonucleolytic activity. *J. Biol. Chem.*, **282**, 13966–13976.
74. Portnoy,V., Evguenieva-Hackenberg,E., Klein,F., Walter,P., Lorentzen,E., Klug,G. and Schuster,G. (2005) RNA polyadenylation in Archaea: not observed in Haloferax while the exosome polynucleotidylates RNA in Sulfolobus. *EMBO Rep.*, **6**, 1188–1193.
75. Märtens,B., Amman,F., Manoharadas,S., Zeichen,L., Orell,A., Albers,S.V., Hofacker,I. and Bläsi,U. (2013) Alterations of the transcriptome of *Sulfolobus acidocaldarius* by exoribonuclease aCPSF2. *PLoS One*, **8**, e76569.
76. Zbigniew,D., Carpousis,A.J. and Clouet-d´Orval,B. (2013) Emergence of the β-CASP ribonucleases: highly conserved and ubiquitous metallo-enzymes involved in messenger RNA maturation and degradation. *Biochim. Biophys. Acta*, **1829**, 532–551.
77. Shao,Y. and Li,H. (2013) Recognition and cleavage of a non-structrured CRISPR RNA by its processing endoribonuclease Cas6. *Structure*, **21**, 385–393.