The Exosome Complex

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION

Received for publication, November 22, 2005, and in revised form, December 19, 2005 Published, JBC Papers in Press, December 28, 2005, DOI 10.1074/jbc.MS124952001

Celso Raul Romero Ramos 1,4, Cristiano L. P. Oliveira 2,5, Iris L. Torriani 4, and Carla Columbano Oliveira 1,3

From the 1Department of Biochemistry, Chemistry Institute, University of Sào Paulo, 05508-900 Sào Paulo, SP, Brazil, 2Brazilian Synchrotron Light Laboratory, 13084-971 Campinas, SP, Brazil, and 3Physics Institute, State University of Campinas, 13084-971 Campinas, SP, Brazil

The exosome is a conserved eukaryotic enzymatic complex that plays an essential role in many pathways of RNA processing and degradation. Here, we describe the structural characterization of the predicted archaeal exosome in solution using small angle x-ray scattering. The structure model calculated from the small angle x-ray scattering pattern provides an indication of the existence of a disk-shaped structure, corresponding to the “RNases PH ring” complex formed by the proteins aRrp41 and aRrp42. The RNases PH ring complex corresponds to the core of the exosome, binds RNA, and has phosphorylolytic and polymerolytic activities. Three additional molecules of the RNA-binding protein aRrp4 are attached to the core as extended and flexible arms that may direct the substrates to the active sites of the exosome. In the presence of aRrp4, the activity of the core complex is enhanced, suggesting a regulatory role for this protein. The results shown here also indicate the participation of the exosome in RNA metabolism in Archaea, as was established in Eukarya.

In Eukarya, the complex of 3′→5′ exoribonucleases named exosome participates in the maturation of the small nuclear RNAs, small nuclear RNAs, and rRNAs (1) as well as in the degradation of deadenylated mRNAs (2, 3) and aberrant pre-rRNAs (4). Consequently, the exosome plays a central role in the regulation of gene expression in these organisms.

The exosome was originally described in yeast as a complex consisting of nine core components, six of which have homology to bacterial RNase PH (Rrp41p, Rrp42p, Rrp43p, Rrp45p, Rrp46p, and Mtr3p), and three contain a putative S1 RNA binding domain (Rrp4p, Rrp40p, and Cls4p) (5). All these proteins were also identified in human (6) and Drosophila melanogaster (7). Some exosome components were also found in Trypanosoma brucei (5). All these proteins were also identified in human (6) and three contain a putative S1 RNA binding domain (Rrp4p, Rrp40p, and Rrp41p, Rrp42p, Rrp43p, Rrp45p, Rrp46p, and Mtr3p), and have phosphorylolytic and polymerolytic activities. Three additional molecules of the RNA-binding protein aRrp4 are attached to the core as extended and flexible arms that may direct the substrates to the active sites of the exosome. In the presence of aRrp4, the activity of the core complex is enhanced, suggesting a regulatory role for this protein. The results shown here also indicate the participation of the exosome in RNA metabolism in Archaea, as was established in Eukarya.

In Eukarya, the complex of 3′→5′ exoribonucleases named exosome participates in the maturation of the small nuclear RNAs, small nuclear RNAs, and rRNAs (1) as well as in the degradation of deadenylated mRNAs (2, 3) and aberrant pre-rRNAs (4). Consequently, the exosome plays a central role in the regulation of gene expression in these organisms.

The exosome was originally described in yeast as a complex consisting of nine core components, six of which have homology to bacterial RNase PH (Rrp41p, Rrp42p, Rrp43p, Rrp45p, Rrp46p, and Mtr3p), and three contain a putative S1 RNA binding domain (Rrp4p, Rrp40p, and Cls4p) (5). All these proteins were also identified in human (6) and Trypanosoma brucei (7). Some exosome components were also found in Drosophila melanogaster (8) and Arabidopsis thaliana (9). These findings indicate a conservative function for the exosome among eukaryotic organisms.

Currently, no high resolution structural data are available for the eukaryotic exosome, but electron microscopy and bioinformatics analyses led to a model that shows a structure similar to that of the polynucleotide phosphorylase (PNPase)4 of Streptomyces antibioticus (10, 11). According to that model, the core components of the eukaryotic exosome assemble into a doughnut-shaped structure, with the six RNase PH-type proteins forming a hexameric ring, and three proteins, which contain S1 RNA binding domain, are located on top of the RNases PH ring. This model is supported by two-hybrid interactions between components of the human (6), T. brucei (7), and yeast5 exosomes.

The existence of an archaeal counterpart of the eukaryotic exosome was previously proposed by comparing the gene clustering in archaeal genomes (12). In these organisms the orthologues of the exosome core subunits aRrp4p, aRrp41p, and aRrp42p are encoded by adjacent ORFs and are part of the so-called exosomal superoperon. Each of these proteins corresponds to one domain of bacterial PNPase and to three proteins of eukaryotic exosome (Fig. 1A). Recently, a complex containing these proteins was isolated from Sulfolobus solfataricus (13). The co-precipitation of 16 S rRNA and ribosomal proteins with this complex suggests its participation in rRNA processing in Archaea. In the exosome of S. solfataricus, the proteins aRrp41, aRrp42, and aRrp4 are in apparent equimolar amounts forming a protein complex of ~250 kDa, as determined by glycerol gradient centrifugation (13).

In the present study we report the in vitro assembly of a protein complex corresponding to the archaeal exosome using recombinant proteins of Pyrococcus abyssi (PAB) and P. horikoshii (PH); archaeal rRNA processing protein 4 (aRrp4, ORF PAB0419), aRrp42 (PH1548), and aRrp41 (PH1549). Analyses of these proteins separately or in complex by small angle x-ray scattering (SAXS) allowed us to propose a structural model for the archaeal exosome. In this model aRrp41 and aRrp42 form a disk-shaped structure that correlates with the PNPase hexameric ring formed by the RNase PH domains, as predicted by sequence analysis (Fig. 1A). The RNases PH ring complex showed characteristic activities typical of bacterial PNPase and RNase PH enzymes, including RNA binding, phosphorolysis, and polymerization. During the final preparation of this work a paper was published on the archaeal core exosome structure (14). In that work the crystal structure of the RNase PH ring from another Archaea, S. solfataricus, was obtained, revealing a hexameric ring formed by aRrp41-aRrp42 dimers, confirming our prediction about the homologous structure in Pyrococcus. Our data also complement those on Sulfolobus with the analysis of the Pyrococcus complex formed by aRrp4, aRrp41, and aRrp42, which constitute the archaeal exosome. In this complex three aRrp4 molecules are bound to the core complex as extended arms, which can drive the RNA substrates to the catalytic centers of the exosome. In the presence of aRrp4, 4 The abbreviations used are: PNPase, polynucleotide phosphorylase; ORF, open reading frame; SAXS, small angle x-ray scattering.

4 The abbreviations used are: PNPase, polynucleotide phosphorylase; ORF, open reading frame; SAXS, small angle x-ray scattering.

J. R. Tavares and C. C. Oliveira, unpublished results.

1 Recipient of Fundaçã˜o de Amparo a Pesquisa do Estado de Sào Paulo fellowship 02/00207.

2 Recipient of Fundaçã˜o de Amparo a Pesquisa do Estado de Sào Paulo fellowship 00/15087-4 and of a Conselho Nacional de Desenvolvimento Científico e Tecnológico fellowship (381600/2005-0).

3 To whom correspondence should be addressed: Dept. of Biochemistry, Chemistry Institute, University of Sào Paulo, Av. Prof. Lineu Prestes 748, CEP 05508-900 Sào Paulo, SP, Brazil. Tel.: 55-11-30913810 (ext. 208); Fax: 55-11-38155579; E-mail ccoliv@iq.usp.br.

5 J. R. Tavares and C. C. Oliveira, unpublished results.

© 2006 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.
Archaeal Exosome Complex Solution Structure

the activity of the core complex is enhanced, suggesting a regulatory role for this protein. The functional implications of these data and its relation with the eukaryotic exosome structure are discussed.

EXPERIMENTAL PROCEDURES

Microorganisms, Plasmids, Enzymes, and DNA Manipulation—The *Escherichia coli* strains used in this study were DH5α and BL21-Codon-Plus (DE3)-RIL (Stratagene). Plasmid DNA was extracted using a Qia-gen plasmid purification kit. Restriction enzymes and other DNA-modifying enzymes were used as recommended by their manufacturer (New England Biolabs). Genomic DNA of *Pyrococcus horikoshii* OT3 was kindly provided by Dr. F.J. Medrano (Laboratário Nacional de Luz Síncrotron, Campinas, Brazil), and Genomic DNA of *P. abyssi* GE5 was kindly provided by Dr. Patrick Forterre (Institut de Génétique et Microbiologie, Université Paris Sud, France). Genomic DNA of *P. horikoshii* was used as a template to amplify aRrp4 (PH1549) and aRrp42 (PH1548), whereas genomic DNA of *P. abyssi* was used for aRrp4 (PAB0419) amplification.

Gene Cloning and Construction of Expression Vectors—aRrp4 was amplified using primers aRrp4for (5′-GAGATCG-3′) and aRrp4rev (5′-TCACTTTC-3′); the restriction sites for NdeI and EcoRI are underlined. Ndel- and EcoRI-digested DNA of aRrp4 was inserted into the expression vector pET29a (+) (Novagen), aRrp42 (PH1548) was amplified using primers aRrp42for (5′-ATCTGCAGATTAGGTATAGA-3′) and aRrp4rev2 (5′-TCTGTACGATATGCTAACAATGTCGAATC-3′); the restriction sites for BamHI, Ndel, and PstI are underlined. Ndel- and PstI-digested DNA of aRrp42 was inserted into the expression vector pAE15 (aRrp4-4aRrp4 was amplified using primer aRrp4for (5′-AGGGATCCATATGACTGATAGATA-3′) and reverse primer for aRrp4rev). The PCR-amplified DNA fragment containing the ORFs aRrp4-aRrp42-digested with Ndel and HindIII (HindIII cuts just downstream of the aRrp4 stop codon), and the aRrp4 ORF was inserted into the pET29a (+) expression vector.

Expression and Purification of Proteins—All recombinant proteins were expressed in *E. coli* BL21-Codon-Plus (DE3)-RIL strain, transformed with the correspondent plasmids. After the addition of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside, the cells were harvested and resuspended in buffer A (30 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol) and lysed in a French press. The lysate was heated at 85 °C for 30 min and cooled on ice for 15 min. After centrifugation at 20,000 × g for 30 min, the supernatant was fractionated by chromatography.

Heat-treated supernatant containing aRrp4 was loaded onto an anion-exchange column (HiTrap Q, Amersham Biosciences) and eluted with a linear gradient of 0–500 mM NaCl in 30 mM Tris-HCl, pH 8.0, 0.1 mM EDTA. The proteins were dialyzed against 10 mM Tris-HCl, pH 8.0. Protein concentration was determined by optical density at 280 nm using the extinction coefficient calculated by ProtParam tools (www.expasy.ch).

**Small Angle X-ray Scattering—**SAXS experiments were performed at the D11A-SAXS beamline of the Brazilian Synchrotron Light Laboratory (16). The experimental setup included a temperature-controlled glass capillary (17) and Gabriel type linear position-sensitive detector (18). The RNAses PH ring and exosome complexes at 3.5–7 mg ml⁻¹ (in buffer containing 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM EDTA) were kept at 20 °C during the exposures. The data acquisition was performed taking several 600-s frames for each sample, which allowed the control of any possible radiation damage. Several sample-detector distances were used in the experiments, which enabled detection of scattering vector q values (q = 4πλ sin θ, where 2θ is the scattering vector, and λ is the wavelength) in several angular ranges for each sample. Data treatment of the scattering intensities was performed using the software package trati1d (19). As output, the data treatment program calculates the q values and the corrected experimental errors. Data at low and high concentrations as well as short and long distances were joined to obtain good scattering intensity curves in a wide angular range. From the Guinier Plot of the experimental data (ln(I) versus q²), we calculated the Guinier radius of gyration Rg, and the forward scattering intensity I(0) (using the extrapolation approximation \( I(q) \approx I(0)\exp(-q^2R_g^2/3) \), valid for very small angles (q < 1.3/Rg)). These parameters could also be obtained from the theoretical fit of the entire intensity curve using the program package Gnom (20), which also calculates the pair distance distribution function p(r). From this function the particle maximum dimension Dmax as well as the real space value of the radius of gyration Rg and forward scattering I(0) of the scattering particle were obtained.

The molecular mass of the studied proteins and protein complexes were evaluated by comparison of the forward scattering with that from a reference protein solution. The accuracy of this method was limited by the uncertainty in the measured protein concentration used for data calculation. The standards used were catalase (molecular mass 232 kDa), aldolase (molecular mass 158 kDa), and bovine serum albumin (molecular mass 67 kDa). Additionally, we calculated the so-called...
**Archaeal Exosome Complex Solution Structure**

**FIGURE 1.** A comparison of the organization of the PNPase domains with the archaeal locus containing the ORFs that encode the exosome core components. The homolog domains and proteins are indicated on the same gray scale. The yeast orthologues of each component (11) are also shown. For convenience, in the present work we use the nomenclature aRrp4, aRrp41, and aRrp42 for archaeal proteins, as was previously employed (13). H denotes the α-helical domain in PNPase. B, SDS-PAGE (12%) of isolated exosome subunits. M, molecular mass marker; 1, aRrp42; 2, aRrp41; 3, aRrp4.

Krakty plots ($Iq^2 \times q$), which give information on the compactness of the protein structure (21).

**SAXS Model Calculations**—Several modeling tools were applied for the data analysis, and the procedures are described under “Results.” The ab initio program DAMMIN (22) was applied to restore the protein shapes. The program DAMAVER was used to average independent calculated ab initio models in order to retrieve the most probable configuration (23). The Rigid Body Modeling program MASSHA (24) was used to search and calculate quaternary structures against the experimental data. The program MOLMOL (25) was used to generate the model pictures and superpose the calculated structures. The program CRY SOL (26) was used to fit the experimental intensity with the composed models and also with known crystallographic model structures. The program HYDROPRO (27) was used to calculate theoretical hydrodynamic properties for the generated models.

**Limited Proteolysis Experiments**—Protein suspension in 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM EDTA was incubated at 25 °C with trypsin at a 250:1 protein/enzyme ratio. At timed intervals, aliquots were withdrawn, and the reaction was stopped by the addition of phenylmethylsulfonyl fluoride (Sigma). The samples were analyzed by 15% SDS-PAGE.

**In Vitro Transcription of RNA for Activity Assays**—Uniformly labeled un Specific RNA was produced by in vitro transcription of XhoI-linearized pBS7 plasmid (28) in the presence of NTPs, 50 μCi of [α-32P]UTP (Amersham Biosciences), and SP6 RNA polymerase (Promega). This RNA probe is the antisense of the 3'-UTR of the Saccharomyces cerevisiae cytochrome C pre-mRNA (90 ribonucleotides).

**Electrophoresis Mobility Shift Assays**—The labeled RNA was mixed with different quantities of the proteins (0–8 pmol) in binding buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM KCl, 1 mM MgCl2) and incubated for 30 min at 37 °C. The RNA-protein complexes were subsequently fractionated by electrophoresis on 5% native polyacrylamide gels.

**RNase PH Phosphorolysis and Polymerase Assays**—RNA phosphorolysis was assayed by incubating the proteins with labeled RNA in 20 mM Tris-HCl, pH 8.0, 50 mM KCl, 10 mM MgCl2, and 10 mM NaH2PO4. RNA polymerization was accomplished under the same conditions but using ADP instead of NaH2PO4. The mixture was incubated for 30 min at 37 °C, and products were resolved on a 5% denaturing polyacrylamide gel.

**RESULTS**

**In Vitro Reconstitution of the Archaeal Exosome Complex**

To gain information on the structure and function of the exosome complex, we decided to study the predicted archaeal exosome with corresponding recombinant proteins from two closely related organisms: P. horikoshii and P. abyssi (Fig. 1A). The exosome subunits aRrp4, aRrp41, and aRrp42 from these Archaea show 95, 97, and 91% of identity and 97, 99, and 97% of similarity, respectively. This high similarity allows the combination of proteins from these organisms to reconstitute the exosome. The ORFs corresponding to aRrp4, aRrp41, and aRrp42 (PH1549) and aRrp42 (PH1548) were amplified by PCR, cloned in E. coli expression vectors, and purified as described under “Experimental Procedures” (Fig. 1B).

Exosome subunits that are homologous to the bacterial RNase PH (aRrp4 and aRrp42) are predicted to associate, forming a hexameric ring, as observed for PNPase RNase PH domains (10). This hypothesis was confirmed by coexpressing aRrp41 and aRrp42 in E. coli. The analysis of the single elution peak observed in the first step of purification confirmed the presence of both proteins (Supplemental Fig. 1). The intensities of the protein bands suggest a 1:1 stoichiometry. The complex purified by this procedure is referred to as RNases PH ring.

To reconstitute the archaeal exosome, purified aRrp4 was added to the RNases PH ring complex and subjected to co-purification through anion-exchange chromatography. This procedure resulted in the co-purification of the three subunits in a complex (Fig. 2A). These results indicate the specific association between the aRrp4, aRrp41, and aRrp42 proteins. The resulting complex, obtained by the combination of P. horikoshii and P. abyssi proteins, corresponds to the archaeal exosome.

**Ge Filtration Analysis**

The recombinant protein aRrp4 as well as the complexes were characterized by gel filtration on a Superose 6 HR 10/30 column. The data obtained are summarized in Table 1. aRrp4 is monomeric and eluted as a 34-kDa protein, a very close to its theoretical molecular mass. The observed molecular mass for the RNases PH ring was 198 kDa, whereas the theoretical molecular masses of aRrp41 and aRrp42 are 27.8 and 30 kDa, respectively. Considering the similar intensities of the corresponding bands in SDS-PAGE analysis (Fig. 2A and Supplemental Fig. 1), the molecular mass of 198 kDa suggests that the RNases PH ring complex is composed of three copies of each of these proteins, which is in accordance with the hexameric ring hypothesis.

The addition of the protein aRrp4 (30 kDa) resulted in an increase of ~90 kDa in the observed molecular mass of the RNases PH ring complex that corresponds to three molecules of aRrp4 for complex. Accordingly, the observed molecular mass for the exosome complex (287 kDa) is similar to its theoretical molecular mass (261 kDa). The gel filtration profiles of RNases PH ring and exosome complexes and the analysis of the corresponding peaks indicate that the size difference between the RNases PH ring and the exosome is due to the association of three aRrp4 subunits (Fig. 2). In summary, the gel filtration analysis is consistent with the structural similarity between the exosome and the bacterial PNPase.
Archaeal Exosome Complex Solution Structure

Table 1

| Structural parameters of archaeal exosome complex obtained by gel filtration and SAXS | aRrp4 | RNases PH ring | Exosome |
|---|---|---|---|
| Theoretical molecular mass (kDa) | 29.3 | 173.4* | 261.3* |
| Gel filtrationb | | | |
| Molecular mass (kDa) | 34 ± 1 | 198.5 ± 14.8 | 287.8 ± 3 |
| Stokes radius (Å) | 25.9 ± 0.4 | 53 ± 1.2 | 59.2 ± 0.18 |
| SAXS | | | |
| Molecular mass (kDa) | 36 ± 3 | 179 ± 5 | 257 ± 7 |
| Rs (Å) | 41.3 ± 0.8 | 38.5 ± 0.2 | 59.9 ± 0.3 |
| Maximum dimension (Å) | ~150 | ~112 | ~190 |
| SAXS modelsc | | | |
| Stokes radius (Å) | 39.01 | 49.82 | 86.60 |

* Theoretical values correspond to the hexameric ring hypothesis, i.e. the three folds of the sum of theoretical molecular mass of each component
b The values are means ± S.D. of three replica experiments.
c Values obtained using program HYDROPRO.

Small Angle X-ray Scattering Experiments

The results of the SAXS analyses for the isolated proteins and the complexes show that in all cases the values of Guinier radius of gyration RgG and real space Rg were close. This agreement in association with the calculated values of molecular weights indicates that in all cases we had satisfactory monodispersity of the solutions. The experimental intensity data as a function of the modulus of the scattering vector q are shown for aRrp4 protein (Fig. 3A), RNases PH ring (Fig. 3B), and exosome complex (Fig. 3C).

The analysis of the SAXS data indicates that the aRrp4 protein has a radius of gyration of 41.3 ± 0.8 Å (RgG = 42Å), and maximum dimension of ~150 Å. The estimated molecular mass for this protein using appropriate protein standards is 36 ± 3 kDa, which is close to the theoretical value for a monomeric protein (~30 kDa; Table 1). The function p(r) indicates that aRrp4 possesses an extremely elongated conformation in solution. However, gel filtration data yields a Stokes radius of 25.9 Å. As shown in the next section, the predicted Stokes radius for the ab initio model is 39 Å, and these results indicate that in addition to the elongated conformation (see Table 1), the protein aRrp4 is quite flexible (29). The RNases PH ring complex has a radius of gyration of 38.5 ± 0.2 Å (RgG = 40 Å) and maximum dimension of ~112 Å. The estimated molecular mass for this complex is 179 ± 5 kDa, which is in good agreement with the expected value for the RNases PH protein complex (173.4 kDa; see Table 1). The distance distribution function (p(r)) indicates that the RNases PH ring possesses a globular conformation. The exosome complex has a radius of gyration of 59.9 ± 0.3 Å (RgG = 62 Å) and maximum dimension of ~190 Å. The calculated molecular mass for this complex is 257 ± 7 kDa, close to the theoretical value for this complex (261.3 kDa; Table 1). The observed molecular mass difference between the exosome and RNases PH ring complex (87.9 kDa) indicates that three molecules of aRrp4 (~30 kDa) are attached to the RNases PH ring to form the exosome complex. These results corroborate the gel filtration data. The behavior for the p(r) function was similar for the RNases PH ring and exosome complexes, and the major differences occur for large r values. The differences in size and radius of gyration are due to the addition of aRrp4 protein to the RNase PH ring complex. Also, the Kratky plots suggest the flexibility of the aRrp4 protein and the compactness of the RNases PH ring complex (Supplemental Fig. 2, A and B, respectively). Also, this analysis indicates that the exosome complex possesses a globular compact domain. However, it is less structured than the RNase PH ring, indicating that the aRrp4 monomers have flexibility also when in the exosome complex (Supplemental Fig. 2C).

Model Calculations

aRrp4—Ab initio model calculations for this protein gave a very elongated conformation in solution as indicated in Fig. 4A. As shown in Fig. 3A, this model fits perfectly the experimental data. Interestingly, this structural information is not in agreement with previous predictions using the structure of RNA binding domains from S. antibioticus PNPase as in reference (10), which indicated that these domains should be more compact and linked to the RNases PH ring in a layered fashion. We do not have any bona fide solved structure of a homologous protein to compare with our aRrp4 model.

RNases PH Ring—Ab initio model calculations for the RNases PH ring complex leads to a disk-shaped rigid structure with a central protuberance (Fig. 4B), giving a very good fit of the experimental data (Fig. 3B). To compare the calculated models with homologous crystallographic structures recently published, we applied the program CRYSOI (26) using the structural information of the RNases PH ring from the exosome of S. solfataricus (2BR2), Bacillus subtilis RNase PH (1OYR), and RNases PH domains of S. antibioticus PNPase (1E3P). The best fit was obtained using the model of the RNases PH ring of PNPase (Fig. 3B). The difference in the data for values above q = 0.07 Å−1 arises mainly from differences in the monomers since they are homologous proteins. Interestingly, the similarity of P. horikoshii aRrp4 and aRrp42 proteins with B. subtilis RNase PH (46 and 40%, respectively) is higher than the similarity of these proteins with the corresponding RNase PH domain of S. antibioticus PNPase (Fig. 1). SAXS data indicates that despite sequence similarity, the structure of RNases PH ring of the Pyrococcus exosome, formed by aRrp4 and aRrp42, is more closely related to the PNPase ring, formed by two distinct domains, than to the homohexameric ring structure of bacterial RNases PH. Also, the structure of S. solfataricus exosomal core, which was solved by molecular substitution using the structure of the bacterial RNase PH (1OYR) as template (14), does not fit the SAXS data for the homologous P. horikoshii complex despite the similarity of the corresponding proteins (59 and 49% of

FIGURE 2. Gel filtration chromatography of the archaeal exosome complexes. A, SDS-PAGE (12%) analysis of the archaeal RNases PH ring (1) and exosome (2) complexes. B, the complexes were chromatographed on a Superose 6 HR 10/30 column. Proteins were separated isocratically and detected by their absorbance at 280 nm. The positions of protein standards are indicated. Numbers represent the molecular mass in kDa. ord. u., arbitrary units.
identity and 77 and 59% of similarity, respectively). The ab initio structure and the RNases PH domains of PNPase are superposed in Fig. 4C, showing their high similarity.

Exosome Complex—The building of the exosome complex had several steps. Initially we built an ab initio dummy atom model imposing the symmetry constraint P3, which gives the resulting model shown in Fig. 5A, with a very good fit of the experimental data (Fig. 3C). From this model we conclude that the exosome presents a rigid central structure, which can be correlated with the RNases PH ring complex, and three extended arms, which correspond to the aRrp4 protein. Using the ab initio models obtained for aRrp4 and RNases pH ring, we could reconstruct the quaternary structure of the exosome complex (Fig. 5C) by superposition of the dummy atom model of the exosome and the calculated models of its components. Although this is a simple approach, the fit of the experimental data is very good principally in the low q region of the scattering intensity (<0.07 Å⁻¹), indicating that the overall shape is correctly described by this model.

The structure of the region that corresponds to the PNPase RNA binding domains was not resolved in the crystal, but the electronic density indicates that these domains are on one side of the RNases PH ring (10). The KH and S1 RNA binding domains were modeled onto PNPase ring structure to represent this region (10). This model (data kindly provided by Dr. M. Symmons, University of Cambridge, Cambridge, UK) without the α-helical domain (Fig. 5B) was used for SAXS data fitting, giving a poor fit (Fig. 3C), with a radius of gyration lower than the experimental value. This result indicates that the RNA-binding protein aRrp4 is projected out from the RNases PH ring, as shown in Fig. 5A, and not retracted and sitting above it, as expected from the PNPase structure model (Fig. 5B). Consequently, the exosome solution structure should have a larger size due to the long and flexible aRrp4 arms. This conformation may provide an extended RNA binding surface, as has been observed for NusA transcription factor from Thermotoga maritima, that contains a succession of S1 and KH RNA binding domains (30).
Limited Proteolysis Experiments

Proteolytic enzymes are commonly used to probe overall protein flexibility in solution because the susceptibility of a peptide bond to proteolytic attack is not only influenced by the accessibility of that bond to the active site of the protease, but it is mainly dictated by the local conformation of the peptide segment in which proteolysis occurs (31). The conformational flexibility of the polypeptide backbone is proportional to the efficiency of proteolytic cleavage, and therefore, more rigid local environments make the polypeptide chain more resistant to proteolysis.

Analyzing the SAXS data, we inferred that the protein aRrp4 is as flexible in solution as when in the exosome complex. To check this hypothesis we performed limited proteolysis experiments with aRrp4 alone or as part of the exosome. The results of the digestions of aRrp4, RNases PH ring, and exosome with trypsin show that, whereas the RNases PH ring complex is stable for up to 2 h, aRrp4 is rapidly cleaved, originating two bands of ~10 and 20 kDa that may correspond to the KH and the S1 RNA binding domains (Fig. 6). After 90 min of reaction most of the aRrp4 protein was digested (data not shown). The cleavage of aRrp4 into two bands is independent of whether this protein is alone or it is in the exosome complex. Similar results were obtained using thermolysin (data not shown). These results indicate the existence of a highly susceptible site for protease attack in aRrp4 that is not protected when it is complexed in the exosome, in accordance with the SAXS model for the Rrp4 protein.

RNA Binding Analysis of Archaeal Exosome Subunits

To characterize the RNA binding properties of the protein subunits and complexes, we performed an RNA electrophoretic mobility shift assay. An essay with single subunits revealed that only aRrp4 was able to bind RNA (Fig. 7). This result was expected because aRrp4 contains S1 and KH RNA binding domains. Although aRrp41 and aRrp42 did not show RNA binding activity on their own, the RNases PH ring complex was able to bind RNA (Fig. 7). This binding activity was even higher when aRrp4 was added to the RNases PH ring for the formation of the exosome. Upon binding to RNA, the RNases PH ring and the exosome form several complexes, identified by multiple bands on the autoradiography (Fig. 7). In the case of the RNases PH ring complex, the shifted bands are still visible in the presence of 8 pmol of the complex, but in the case of the exosome, at that concentration of exosome (8 pmol) the RNA-protein complexes did not enter the gel. A similar pattern of mobility shift has been observed for E. coli PNPase and its RNA binding defective mutant (32). The results shown here indicate that the exosome binds RNA more tightly than the RNases PH ring, probably due to the presence of aRrp4.

RNA Phosphorolysis and Polymerization Activities

We also tested the RNA phosphorolysis and polymerization activities of all three archaeal exosome subunits. The single subunits aRrp4, aRrp41, and aRrp42 did not show either RNA phosphorolysis or polymerization (Fig. 8, A and C). However, when aRrp41 and aRrp42 are associated in the RNases PH ring complex, both activities are detectable under the same conditions (Fig. 8, A and C). A change in Mg²⁺ concentration from 5 to 10 mM increases RNA phosphorolysis (Fig. 8B). The polymerization is also observed in the presence of 10 mM ADP but with a complex pattern (Fig. 8D). Association of the RNA-binding protein aRrp4 to the RNases PH ring complex in order to obtain the exosome
slightly increased both catalytic activities, as observed by adding different amounts of the complexes to the reaction (Fig. 8, B and D). Phosphorylase of RNA was detected with starting concentrations of 1 pmol of the complexes RNases PH ring and exosome. RNA polymerization was detected in greater than 0.5-pmol concentrations of these complexes. This may be the first evidence of regulation of the exosome activity by an RNA-binding protein and would be similar to the RNA phosphorylase activity of full-length and RNA binding deletion mutant of bacterial PNPase (33). Detailed biochemical analyses are still necessary to understand the influence of RNA binding properties of aRrp4 in the RNase activity or substrate selection by the complex.

**DISCUSSION**

Several structural models have been proposed for the eukaryotic exosome (10, 11, 34, 35). However, no direct evidence is available to confirm those predictions. In the present study we obtained recombinant proteins corresponding to the archaeal exosome subunits which were used to reconstitute this complex.

The *Pyrococcus* exosome was obtained by expressing its subunits in *E. coli* and reconstituting the complex in *vitro*. Gel filtration and SAXS analyses confirmed an RNases PH hexameric ring as a central structural motif, formed by aRrp41 and aRrp42, and three molecules of aRrp4 bound to the ring. The structure of the archaeal exosome is similar to the *S. antibioticus* PNPase and to the *S. solfataricus* RNases PH ring (10, 14).

The structural data obtained from aRrp4 analyses indicates that this protein has an elongated conformation with intrinsic flexibility. The *ab initio* structural models obtained for the RNases PH ring and aRrp4 agree perfectly well with the model of the complete structure of the exosome (Fig. 5C). Also, the addition of aRrp4 to the RNases PH ring results in an increment of radius of gyration and flexibility of the complex. The analysis of the crystal structure of the bacterial PNPase showed that its RNA binding region is also flexible, as the structure corresponding to these regions was not determined in the crystal structure but is proposed to have a compact conformation above RNases PH ring (10) (see Fig. 5B) in contrast to our model for aRrp4 protein. The differences in this region may be due to different amino acid lengths and succession of RNA binding domains in the PNPase and aRrp4 (Fig. 1). On the other hand the electron microscopy analysis of the yeast exosome indicates that, although this complex is similar in size and shape with PNPase, it has great differences with respect to the S1 and KH domains possibly due to its flexibility or different organization (11).

In agreement with the model proposed here, aRrp4 has shown high susceptibility to cleavage in limited proteolysis experiments regardless of whether it was part of the exosome or not. This result is similar to that observed during the preparation of the T (primer-dependent) form of PNPase from the I (primer-independent) form by limited trypic digestion (36). During the preparation of this manuscript, Stickney et al. (33) clearly showed that limited proteolytic digestion removes the RNA binding domains from *E. coli* PNPase (33), confirming the susceptibility of this region to proteolysis. In that work it is also proposed that the RNA binding domains of PNPase are in an extended conformation (33).

Binding of the RNases PH ring and the exosome to RNA results in a complex pattern of shifted bands in electrophoretic mobility shift assays, similar to those observed for the *E. coli* PNPase (32). This effect may be due to the presence of more than one RNA binding site in these complexes. Additionally, the archaeal exosome has three copies of aRrp4, which has two RNA binding domains, although its monomeric form produced a single shifted band. These observations are consistent with the formerly proposed existence of two groups of RNA binding sites in bacterial PNPase; one group in the catalytic site binds the 3’-end of the RNA, and the second region (that corresponds to KH-S1 RNA binding motifs) binds to the 5’-end and/or internally to the RNA molecule (37). However, at this point we cannot ascertain the possible RNA binding sites formed in the complex. Further characterization of the RNA binding properties of the proteins and complexes studied in the present work by alternative methods such as atomic force or electron microscopy might shed some light on this matter.

aRrp41 and aRrp42 do not show any activity (RNA binding, phosphorylation, or polymerization) when tested alone but only when assembled into the RNases PH ring. Therefore, the catalytic sites of the RNases PH ring may only be formed when aRrp41 and aRrp42 bind to each other and assemble into the ring. The importance of the quaternary structure for the catalytic activity was previously shown for homologous complexes; mutations in *B. subtilis* and *P. aeruginosa* RNases PH that prevent the assembly of the hexameric ring result in the loss of their catalytic activity (38, 39). Also, data published during the final preparation of this manuscript on the core complex from *S. solfataricus* exosome (14) corroborate the importance of formation of the RNases PH ring by aRrp41 and aRrp42 for the RNA degradation and polymerization activities of the exosome. In addition to those data, in the present work we showed the RNA binding activity of the corresponding complex in *Pyrococcus* and the formation of the whole exosome complex by the addition of aRrp4 to the RNases PH complex. *Pyrococcus* and *S. solfataricus* are two distinct Archaea that belong to the *Euryarchaeota* and *Crenarchaeota* phylum, respectively. Our data, together with recent studies (12, 14) suggest that the exosome protein complex is formed in all Archaea containing the exosomal gene cluster, and this complex may perform functions similar to those of the eukaryotic exosome.

Based on the data shown here and the structure and function of the
bacterial PNPases and RNases PH, we suggest a model of the archaeal exosome (Fig. 9). Pyrococcus aRrp41 contains a phosphate binding site, as determined by molecular modeling (data not shown), and has higher similarity to bacterial RNase PH. On the other hand, the Rrp42 protein has higher variability among different organisms and does not have a phosphate binding site but participates in conjunction with aRrp41 in the formation of catalytic centers, which are located on one side of the RNases PH ring complex (Fig. 9). In our model, the 3'-end of RNA reaches the catalytic centers through the central channel that can accommodate a single-stranded RNA molecule (10, 33). As predicted by the crystal structure of *S. solfataricus* exosome core (14), the aRrp4 protein may be located at the hydrophobic patches in the opposite side of the ring. The aRrp4 function is to bring the RNA substrate to the catalytic centers in a way similar to the recently proposed model for bacterial PNPase (33) and then regulate the exosome activity. Although the SAXS models indicate an extended conformation of aRrp4 in solution, it is possible that this protein may adopt a more globular shape in some conditions (as was observed in gel filtration experiments). This aRrp4 conformational flexibility may be important to perform the RNA uptake function.

There is also evidence for the importance of the complex assembly for activity of the eukaryotic exosome. (a) In the yeast exosome all core subunits are essential for cell viability (40); (b) expression of Rrp43p mutants that no longer interact with Rrp46p causes defects in rRNA processing and mRNA degradation that are similar to those observed in single exosome subunit depletion (3); (c) in *T. brucei* depletion of some of the RNase PH core components results in exosome instability (7). These data combined with the results presented here suggest that the RNases PH ring of the exosome would be active only when all of its components are assembled together. In the case of eukaryotic Rrp41p, however, the isolated protein showed RNase activity (41, 42). Further biochemical characterization of the eukaryotic exosome is necessary to establish the relationship with its archaeal counterpart.

Acknowledgments—We thank Nilson I. T. Zanchin for suggestions and critical reading of this manuscript and Tereza C. Lima Silva and Zildene G. Correa for DNA sequencing.
REFERENCES

1. Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999) *EMBO J.* 18, 5399–5410
2. Anderson, J. S., and Parker, R. P. (1998) *EMBO J.* 17, 1497–1506
3. Oliveira, C. C., Gonzales, F. A., and Zanchin, N. I. (2002) *Nucleic Acids Res.* 30, 4186–4198
4. Allmang, C., Mitchell, P., Petfalski, E., and Tollervey, D. (2000) *Nucleic Acids Res.* 28, 1684–1691
5. Mitchell, P., and Tollervey, D. (2000) *Nat. Struct. Biol.* 7, 843–846
6. Raajmakers, R., Egberts, W. V., van Venrooij, W. J., and Pruijn, G. J. (2002) *Nucleic Acids Res.* 30, 695–700
7. Symmons, M. F., Jones, G. H., and Luisi, B. F. (2000) *Structure Fold. Des.* 8, 1215–1226
8. Aloy, P., Ciccarelli, F. D., Leutwein, C., Gavin, A. C., Superti-Furga, G., Bork, P., Bottcher, B., and Russell, R. B. (2002) *EMBO Rep.* 3, 628–635
9. Koonin, E. V., Wolf, Y. I., and Aravind, L. (2001) *Genome Res.* 11, 240–252
10. Symmons, M. F., Williams, M. G., Luisi, B. F., Jones, G. H., and Carpousis, A. J. (2002) *EMBO Rep.* 3, 880–883
11. Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G., and Conti, E. (2005) *Nat. Struct. Mol. Biol.* 12, 575–581
12. Worbs, M., Bourenkov, G. P., Bartunik, H. D., Huber, R., and Wahl, M. C. (2001) *Mol. Cell.* 7, 1177–1189
13. Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E., and Zambonin, M. (1997) *Fold. Des.* 2, 17–26
14. Regonesi, M. E., Briani, F., Ghetta, A., Zangrossi, S., Ghisotti, D., Tortora, P., and Deho, G. (2004) *Nucleic Acids Res.* 32, 1006–1017
15. Stickney, L. M., Hankins, J. S., Miao, X., and Mackie, G. A. (2005) *J. Bacteriol.* 187, 7214–7221
16. Symmons, M. F., Williams, M. G., Luisi, B. F., Jones, G. H., and Carpousis, A. J. (2002) *Trends Biochem. Sci.* 27, 11–18
17. Raijmakers, R., Schilders, G., and Pruijn, G. J. (2004) *Eur. J. Cell Biol.* 83, 175–183
18. Klee, C. B. (1969) *J. Biol. Chem.* 244, 2558–2566
19. Sulewski, M., Marchese-Ragona S. P., Johnson, K. A., and Benkovic S. J. (1989) *Biochemistry* 28, 5855–5864
20. Svergun, D. I. (1999) *Biophys. J.* 76, 2879–2886
21. Volkov, V. V., and Svergun, D. I. (2003) *J. Appl. Crystallogr.* 36, 860 –864
22. Konarev, P. V., Petoukhov, M. V., and Svergun, D. I. (2001) *J. Appl. Crystallogr.* 34, 527–532
23. Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E., and Zambonin, M. (1997) *Fold. Des.* 2, 17–26
24. Regonesi, M. E., Briani, F., Ghetta, A., Zangrossi, S., Ghisotti, D., Tortora, P., and Deho, G. (2004) *Nucleic Acids Res.* 32, 1006–1017
25. Stickney, L. M., Hankins, J. S., Miao, X., and Mackie, G. A. (2005) *J. Bacteriol.* 187, 7214–7221
26. Symmons, M. F., Williams, M. G., Luisi, B. F., Jones, G. H., and Carpousis, A. J. (2002) *Trends Biochem. Sci.* 27, 11–18
27. Raijmakers, R., Schilders, G., and Pruijn, G. J. (2004) *Eur. J. Cell Biol.* 83, 175–183
28. Klee, C. B. (1969) *J. Biol. Chem.* 244, 2558–2566
29. Sulewski, M., Marchese-Ragona S. P., Johnson, K. A., and Benkovic S. J. (1989) *Biochemistry* 28, 5855–5864
30. Harlow, L. S., Kadziola, A., Jensen, K. F., and Larsen, S. (2004) *Protein Sci.* 13, 668–677
31. Choi, J. M., Park, E. Y., Kim, J. H., Chang, S. K., and Cho, Y. (2004) *J. Biol. Chem.* 279, 755–764
32. Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., and Tollervey, D. (1999) *Genes Dev.* 13, 2148–2158
33. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997) *Cell* 91, 457–466
34. Chekanova, J. A., Shaw, R. J., Wills, M. A., and Belostotsky, D. A. (2000) *J. Biol. Chem.* 275, 33158–33166