Expression, purification and structural analysis of the Pyrococcus abyssi RNA binding protein PAB1135

Juliana S Luz1, João ARG Barbosa2,3, Celso RR Ramos1,4 and Carla C Oliveira*1

Abstract

Background: The gene coding for the uncharacterized protein PAB1135 in the archaeon Pyrococcus abyssi is in the same operon as the ribonuclease P (RNase P) subunit Rpp30.

Findings: Here we report the expression, purification and structural analysis of PAB1135. We analyzed the interaction of PAB1135 with RNA and show that it binds efficiently double-stranded RNAs in a non-sequence specific manner. We also performed molecular modeling of the PAB1135 structure using the crystal structure of the protein Af2318 from Archaeoglobus fulgidus (2OGK) as the template.

Conclusions: Comparison of this model has lead to the identification of a region in PAB1135 that could be involved in recognizing double-stranded RNA.

Background

Despite the recent progress in various genome analysis projects, about a quarter of the archaeal genomes encode functionally uncharacterized proteins, which are almost all only common to other archaeal species [1-4]. Pyrococcus abyssi PAB1135 protein function has not yet been characterized, and it is classified in the family domain of unknown function 54 (DUF54) and in the uncharacterized protein family 0201 (UPF0201). This group’s members have been annotated as conserved hypothetical proteins in 46 archaeal species. Some of these proteins were annotated as possible exosome subunits (TK1451 - GI: 57641386, Thermococcus kodakarensis; MK0388 - GI: 20093826, Methanopyrus kandleri; Msp1244 - GI: 84490032, Methanosphaera stadtmanae) [4-6], but analysis of completely sequenced Pyrococcus abyssi genome revealed the presence ofPAB1135 gene in the same operon as Pa1136, the ribonuclease P (RNase P) subunit Rpp30 [7].

RNase P is an endoribonuclease responsible for maturation of tRNAs in all domains of life. RNase P is a ribonucleoprotein (RNP) complex, formed by one RNA molecule and a variable number of protein subunits, depending on the organism. Bacterial RNase P contains one protein, whereas the archaeal relative contains at least four proteins, and in humans, it contains at least 10 protein subunits [8,9]. Pyrococcus horikoshii RNase P has been shown to be formed by one catalytic RNA and the proteins Ph1481, Ph1601, Ph1771, and Ph1877, which show homology to the human RNase P subunits hPop5, Rpp21, Rpp29, and Rpp30, respectively [10]. The structures of these P. horikoshii proteins have been solved and a possible arrangement of the protein complex has been proposed [11].

Although the function of Pyrococcus abyssi PAB1135 has not been characterized, nor its association with RNase P complex, the genome location of the gene suggests that PAB1135 is involved in RNA metabolism. Here we show that PAB1135 binds RNA in vitro, showing higher affinity for double-stranded RNAs. In addition, structural analysis of PAB1135 by molecular modeling indicates a possible region for protein-RNA interaction.

Methods

Cloning of PAB1135 sequence

The Escherichia coli strains used in this study were DH5α and BL21-CodonPlus (DE3)-RIL (Stratagene). Plasmid DNA was extracted using Qiagen plasmid purification
Expression and purification of the recombinant protein

The pET28a-PAB1135 was transformed into the E. coli BL21-CodonPlus (DE3)-RIL strain. The transformed cells were grown at 37°C in 2xTY medium supplemented with 20 mg/L kanamycin and 17 mg/L chloramphenicol. The expression of His-PAB1135 was induced for two hours with 0.5 mM IPTG. Cells were harvested by centrifugation at 20,000 × g for 15 min, the supernatant was fractionated by affinity chromatography in Ni-NTA-agarose (Qiagen). The purified fractions were analyzed by SDS-PAGE.

Thrombin and Trypsin digestion and analysis

To remove the His-tag from PAB1135, tagged protein was incubated with thrombin for 8 h at room temperature, as recommended by the manufacturer (GE Healthcare). Trypsin treatment was performed by incubating His-PAB1135 with 0.1% trypsin solution (0.1 mg/ml) for 2 hours at room temperature, followed by the addition of 1 mM PMSF. His-PAB1135 and its tryptic cleavage products were subjected to SDS-PAGE and transferred to PVDF membranes (BioRad), which were incubated with an anti-poly-histidine antibody (GE Healthcare). The immunoblots were analyzed using the ECL system (GE Healthcare).

RNA binding assays

RNA binding assays were carried out with 1 pmol 32P 5’-labeled oligoribonucleotides. The oligos used were: U6G5C8 (5′UUUUUUUUCCCCCAAAAAAAA3′), C8U5G6 (5′CCCCCCCCUUUUUGGGGGGGG3′), and UUA/C (5′UUUAUUACAUCAUUAAUCA3′). The assays were performed as described previously [14,15], in Tris-HCl pH 8.0, 20 mM KCl, 5 mM MnCl2, 1 mM DTT, 100 ug/ml BSA, 0.8 U Rnasin. Different amounts of trypsin-treated PAB1135 were incubated with the RNA oligos in 20 μl at 37°C for 30 minutes. The samples were resolved on 8% native polyacrylamide gels and visualized on a Phosphorimager (MolecularDynamics).

Molecular modeling and structural analysis

MODELLER ([16], version 9v1) was used to produce a homology molecular model of PAB1135. The structure of the conserved hypothetical protein Af2318 from Archaeoglobus fulgidus (PDB code 2OGK, [17]) was used as a template. This protein shares 40% identity with the PAB1135 sequence, being the top hit in a search through the PDB [18] using BLAST [19]. The parameters used during the modeling exercise were the default of the programs. The alignment used in MODELLER was produced with CLUSTALX [20]. The alignments of the N- and C-terminal regions were cut at the corresponding ends of the crystallographic model.

The homology models were validated with the VERIFY-3D [21] and PROCHECK [22] softwares. The analysis was made through visualization of the superimposed structures using PYMOL [23] and various alignments produced with CLUSTALX. COOT [24] was used for the superposition [25] of the atomic coordinates of the models and PDB files: 2OGK[17], J12[26], and 1MLI[27]. DALI [28] was also used for analysis of correlated structures.
Results

Purification of PAB1135

The localization of the Pyrococcus abyssi PAB1135 gene in the same operon as the RNase P subunit Rpp30 [7] suggested the involvement of the uncharacterized PAB1135 protein in RNA metabolism. To analyze PAB1135 structure and its association with RNA, PAB1135 gene was cloned and the recombinant protein His-PAB1135 was purified by affinity chromatography. The SDS-PAGE showed protein bands with the expected molecular weight of 21 kDa (Figure 1A,B). The His-tag was released from PAB1135 after cleavage with thrombin (Figure 1C). Limited proteolysis of His-PAB1135 with trypsin resulted in the detection of a very stable protein, corresponding to PAB1135, with a molecular weight close to that expected for the native protein 17.9 kDa (Figure 1D). To ascertain the identity of the peptide released after trypsin cleavage, immunoblot was performed with anti-His antibody and the results show that anti-His is only able to detect the protein before trypsin treatment (Figure 1E,F).

Circular dichroism analysis of PAB1135 show a spectrum with double minimum at 208 and 222 nm, and positive peak near 190 nm, as expected for a protein with α/β structure and its association with RNA, PAB1135 gene was cloned and the recombinant protein His-PAB1135 was purified by affinity chromatography. The SDS-PAGE showed protein bands with the expected molecular weight of 21 kDa (Figure 1A,B). The His-tag was released from PAB1135 after cleavage with thrombin (Figure 1C). Limited proteolysis of His-PAB1135 with trypsin resulted in the detection of a very stable protein, corresponding to PAB1135, with a molecular weight close to that expected for the native protein 17.9 kDa (Figure 1D). To ascertain the identity of the peptide released after trypsin cleavage, immunoblot was performed with anti-His antibody and the results show that anti-His is only able to detect the protein before trypsin treatment (Figure 1E,F).

Circular dichroism analysis of PAB1135 show a spectrum with double minimum at 208 and 222 nm, and positive peak near 190 nm, as expected for a protein with α/β content, indicating that PAB 1135 has a well defined structure (Figure 2A). The deconvolution of CD spectra using K2d algorithm (Figure 2A; blue line in box) indicates a >27% and >25% content for alpha and beta structures, respectively. A CD curve that better fits the experimental data corresponds to a protein containing 30% and 25% of alpha and beta structures, respectively (Figure 2A; red line in box). These estimations are in accordance with the obtained molecular model for Pa1135 protein.

Interestingly, results from gel filtration assays suggest that PAB1135 forms a homodimer in solution, with an apparent molecular weight of about 30 kDa (Figure 2B), whereas a PAB1135 monomer runs as an approximately 20 kDa protein on SDS-PAGE. This is very similar to the calculated 35.8 kDa molecular weight of a homodimer. In the conditions tested, PAB1135 appears monodisperse in solution since 100% of the mass was accounted for by a single peak at 2.7 nm. The temperature variation between 25°C and 50°C did not affect the results.

We had previously analyzed the RNA binding ability of PAB1135 and observed that it does not bind single-stranded RNA oligos, but binds efficiently RNAs that can form low stability hydrogen bonds (J.S. Luz and C.C. Oliveira, unpublished results). Here, we extended these analyses by electrophoresis mobility shift assays. Different amounts of purified PAB1135 were incubated with 32P-labeled RNA oligos. Samples were separated by electrophoresis on native polyacrylamide gels and visualized by phosphorimaging (Figure 3). Very weak shifted RNA bands can be visualized when PAB1135 is incubated with the RNA oligos U8C5A8 (low stability secondary structure; ΔG = -1.9 kcal/mol) and UUA/C (single strand) (Figure 3; lanes 1-5 and 11-15, respectively). When incubated with an RNA that forms higher stability secondary structure (oligo C4U5G4, ΔG = -15.9 kcal/mol), however, PAB1135 binds it much more efficiently, and stronger RNA shifted bands can be visualized in the presence of the protein (Figure 3; lanes 6-10). These results indicate that PAB1135 binds RNAs in a non-sequence specific manner, with higher affinity for double-stranded RNAs.

A BLAST search with the PAB1135 sequence against the PDB shows the Archaeoglobus fulgidus 2OGK[16] as the closest sequence to PAB1135, with 40% identity and 63% similarity, followed by Sulfolobus solfataricus SSO0741 [16] with 27% identity and 48% similarity (Figure 4A). Interestingly, and similar to Pyrococcus abyssi, the gene encoding the Archaeoglobus fulgidus AF2318 protein is present in the same operon as RNase P subunit Rpp30 [7]. Sequence analysis has led to the clustering of these proteins in the UPF0201 family [17]. The structures of members of this protein family show a single α/β domain characterized by a twisted five-stranded anti parallel β-sheet with five α-helices on one side and an unpro-
Figure 2 Structure analysis of PAB1135. (A) Circular dichroism analysis of 2 mg/ml PAB1135. A spectrum with double negative peaks at 208 and 222 nm, and positive peak near 190 nm is detected. The deconvolution of CD spectra (blue line) indicates a >27% and >25% content of alpha and beta structures, respectively. (B) Size-exclusion chromatography of PAB1135 after cleavage of His-PAB1135 with trypsin for release of the His-tag. Protein standards used as controls were: 1, albumin, 67 kDa; 2, ovalbumin, 43 kDa; 3, chymotrypsinogen, 25 kDa; 4, ribonuclease A, 13.7 kDa.
A homology model of the PAB1135 based on 2OGK presents good stereochemistry as judged by PROCHECK and VERIFY-3D. The model was used in a DALI search for structural neighbors, yielding a list that was topped by four UPF0201 Archaea structures followed by 1IQ4 and 1JJ2, all with Z score above 7. After these structures the Z score dropped significantly to 3.5 and sequence identity was below 19%. The first three hits were expected and did not yield new information, but the latter two structures are from the ribosomal protein L5 from Bacillus stearothermophilus and the large ribosomal subunit from Halobaculamariismortui, respectively. In the case of the whole ribosomal subunit, the similarity was also with the L5 protein (Figure 4B).

In light of the functional data showing higher affinity of PAB1135 for double-stranded RNA oligonucleotides, we superposed the structures of the PAB1135, 2OGK, 1JJ2 (only Cα atoms) and 1MII. The latter structure, from the L5 ribosomal protein of Thermus thermophilus, has the complete protein bound to RNA. The superposition of the PAB1135 model onto the 1JJ2 structure is seen in Figure 5. The concave surface of the β-sheet of 1JJ2 is responsible for binding to a double-stranded RNA (dsRNA) region of the 23S rRNA. The size and shape are similar to the surface of PAB1135 although the nature of the residues does not have a clear equivalence (Figure 5).

Analysis of the concave surface in the model shows a patch of basic residues formed by Arg11, Arg13, Arg106, His127, Arg129 and Lys131 on the exposed surface of the β-sheet and Lys65 and Lys133 on nearby loops (Figure 5B). Based on the presence of these residues we raised the hypothesis that this region could complex with and stabilize other negatively charged molecules such as the outside of a double-stranded RNA. These residues are not all present in other members of the UPF0201 family which might not bind to RNA. In addition to binding the 23S rRNA, the Thermus thermophilus L5 ribosomal protein also interacts with the 5S rRNA through a portion of the protein that is not present in the PAB1135 model and therefore was not considered a good RNA binding site.

Conclusions

We show that PAB1135 is highly conserved and its structure can be inferred by molecular modeling based on the crystal structures of archaeal UPF0201 proteins. Furthermore, as shown here, PAB1135 binds RNA in vitro, with higher affinity for structured RNAs, in accordance with the model’s suggestion for the presence of an RNA binding scaffold in the protein. It is possible that PAB1135 binds structured RNAs in vivo, such as RNase P RNA and tRNAs.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

JSL purified the recombinant proteins and carried out the activity assays, protein interaction assays and drafted portions of the manuscript. CRRR, established the protein purification protocols, helped with the CD analysis, and drafted portions of the manuscript. JARGB coordinated parts of the work, performed the molecular modeling and structural analysis, participated in the interpretation of data and drafted portions of the manuscript. CCO designed,
Figure 4 Sequence alignment of PAB1135. (A) Alignment of PAB1135 from Pyrococcus abyssi with Af2318 from Archaeoglobus fulgidus and SSO0741 from Sulfolobus solfataricus in CLUSTALW. (B) Cartoon of the Pa1135 model from MODELLER showing the twisted five-stranded β-sheet with its concave face on one side and α-helices on the other. The side chains of the basic residues on the surface of the β-sheet are shown. (C) Cartoon of the model of PAB1135 interacting with RNA.
organized and coordinated the experiments, drafted the manuscript and edited the final text. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by a FAPESP grant (07/57096-9 to C.C.O.). J.S.L. was recipient of a FAPESP fellowship.

Author Details
1Department of Biochemistry, Institute of Chemistry, University of São Paulo, São Paulo, SP, Brazil, 2Center for Structural Molecular Biology, Brazilian Synchrotron Light Laboratory, LNLS, 13083-970, Campinas, SP, Brazil, 3Current address: Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, 70790-160, Brasília, DF, Brazil and 4Current address: Department of Helminthology, Fiocruz, 21045-900, Rio de Janeiro, RJ, Brazil

Received: 20 November 2009 Accepted: 9 April 2010
Published: 9 April 2010

References
1. Cohen GN, Barbé V, Fletchmann RD, Sutton GG, Blake JA, Fitzgerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CJ, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Goble A, Scott JL, Geoghegan NS, Venter JC: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. Science 1996, 273:1058-1073.

2. Bult CJ, White O, Olsen GI, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CJ, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Goble A, Scott JL, Geoghegan NS, Venter JC: Complete genome sequence of the hyperthermophic archaeon Pyrococcus abyssi. Mol Microbiol 2003, 47:1495-1512.

3. She Q, Singh RK, Contaloiyer F, Zivancovic Y, Allard G, Ayawyez MJ, Chan-Weher CC, Clausen IG, Curtis BA, De Moors A, Erausso G, Fletcher C, Gordon PM, Helkaamp-de Jong I, Jeffries AC, Koerez CJ, Medina N, Peng X, Th-Ngoc HP, Redder P, Schenk M, Theurant C, Tolstrup N, Charlebois RL, Doolittle WF, Duquet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Oost J Van der: The complete genome of the crenarchaeon Sulfolobus solfataricus P2. PNAS 1996, 93:7835-7840.

4. Fuku T, Atomi H, Kanal T, Matsumi R, Fujieda S, Imnakana T: Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus genomes. Genome Res 2005, 15:352-363.

5. Slesarev AI, Mezeheva KV, Makarova KS, Polushin NV, Shcherbina OA, Shakhtova VV, Belova GI, Avraind L, Natale DA, Rogozin IB, Tatusov RL, Wolf YI, Stetter KO, Malvich AG, Koonin EV, Kozyaevik SA: The complete genome of hyperthermophilic Methanopyrus kandleri AV19 and monophy of archaeal methanogens. Proc Natl Acad Sci 2009, 99:4644-4649.

6. Fricke WF, Seedorf H, Henne A, Kruer M, Liesegang H, Hedderich R, Gottschalk G, Thauker RK: The Genome Sequence of Methanosphaera stadtmaneae Reveals Why This Human Intestinal Archaeon is Restricted to Methanol and H2 for Methane Formation and ATP Synthesis. J Bacteriol 2006, 188:642-658.

7. Koonin EV, Wolf YI, Aravind L: Prediction of the archaeal exosome and its interactions with the proteasome and the translation and transcription machineries by a comparative-genomic approach. Genome Res 2001, 11:240-252.

8. Evans D, Marquez SM, Pace NR: RNase P: interface or the RNA and protein worlds. TIBS 2006, 31:333-341.

9. Altman S: A view of RNase P. Mol Bioly 2007, 5:604-607.

10. Kouzuma Y, Mizuguchi M, Takagi H, Fukuhara H, Tsukamoto M, Numata T, Kimura M: Reconstitution of archaeal ribonuclease P from RNA and four protein components. Bichem Bioph Res Commun 2003, 306:666-673.

11. Kawano S, Nakeshima T, Kakuta Y, Tanaka I, Kimura M: Crystal structure of protein P1481p in complex with protein P1877p of archaeal RNase P from Pyrococcus horikoshii OT3: implication of dimer formation of the holozyme. J Mol Biol 2006, 257:583-591.

12. Andrade MA, Chacón P, Merelo JJ, Morán F: Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. Protein Eng 1993, 6:383-390.

13. Chang CT, Wu CS, Yang JT: Circular dichroic analysis of protein conformation: inclusion of the beta-turns. Anal Biochem 1979, 91:13-31.

14. Luz JS, Tavares JR, Gonzales FA, Santos MCT, Oliveira CC: Analysis of the Saccharomyces cerevisiae exosome architecture and of the RNA binding activity of Rrp40p. Biochimie 2007, 89:688-691.

15. Ramos-CRR, Oliveira CLP, Torriani IL, Oliveira CC: The Pyrococcus exosome complex: Structural and functional characterization. J Mol Chem 2006, 281:6751-6759.

16. Salis A, Blundell TL: Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993, 234:779-815.

17. Rao KN, Burley SK, Swaminathan S: UPF201 archaeal specific family members reveal structural similarity to RNA-binding proteins but low likelihood for RNA-binding function. J Mol Biol 2006, 363:890.

18. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyanov IN, Bourne PE: The Protein Data Bank. Nucleic Acids Research 2000, 28:235-242.

19. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J Mol Biol 1990, 215:403-410.

20. Thompson JD, Gibson TJ, Paponok F, Jeannouan F, Higgins DG: The CLUSTAL windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 1997, 25:4876-4882.

21. Bowie JU, Luthy R, Eisenberg D: A method to identify protein sequences that fold into a known three-dimensional structure. Science 1991, 253:164-170.
22. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 1993, 26:283-291.
23. DeLano WL: The PyMOL Molecular Graphics System (2002) on World Wide Web. [http://www.pymol.org].
24. Emsley P, Cowtan K: Coot: model-building tools for molecular graphics. Acta Crystallogr 2004, 60:2126-2132.
25. Krissinel E, Henrick K: Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr 2004, 60:2256-2268.
26. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: The complete atomic structure of the large ribosomal subunit at 2.4 a resolution. Science 2000, 289:905-920.
27. Perederina A, Nevskaya N, Nikonov O, Nikulin A, Dumas P, Yao M, Tanaka I, Garber M, Gongadze G, Nikonov S: Detailed analysis of RNA-protein interactions within the bacterial ribosomal protein L5/S5 rRNA complex. RNA 2002, 8:1548-1557.
28. Holm L, Sander C: Searching protein structure databases has come of age. Proteins 1994, 19:165-173.

doi: 10.1186/1756-0500-3-97
Cite this article as: Luz et al., Expression, purification and structural analysis of the Pyrococcus abyssi RNA binding protein PAB1135 BMC Research Notes 2010, 3:97

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit