SUPPLEMENTARY DATA

RNA processing machineries in Archaea: the 5'-3' exoribonuclease aRNase J of the β-CASP family is engaged specifically with the helicase ASH-Ski2 and the 3'-5' exoribonucleolytic RNA exosome machinery

Duy Khanh Phung¹, Clarisse Etienne¹, Manon Batista¹, Petra Langendijk-Genevaux¹, Yann Moalic², Sébastien Laurent², Sophie Liu³, Violette Morales¹, Mohamed Jebbar², Gwennaele Fichant¹, Marie Bouvier¹, Didier Flament² and Béatrice Clouet-d'Orval*¹∗

¹ Laboratoire de Microbiologie et de Génétique Moléculaires, UMR5100, Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, Université Paul Sabatier, F-31062 Toulouse, France
² Ifremer, Univ Brest, CNRS, Laboratoire de Microbiologie des Environnements Extrêmes, F-29280 Plouzané, France
³ Micalis Institute, PAPPSO, INRA, AgroParisTech, Université Paris-Saclay, 78350, Jouy-en-Josas, France

§These authors have participated equally to this work

*To whom correspondence should be addressed. Tel: (33)561335875 Email: Beatrice.Clouet-dOrval@ibcg.biotoul.fr

Present Address [Duy Khanh Phung] University College London, Gower Street, Darwin Building, London WC1E 6BT United Kingdom

CONTENT:
Supplementary Figure S1-S5 and legends for Figure S1-S5
Supplementary Table S1-S4 and legends for Table S1-S4
**FIGURE S1** - The aRNase J and ASH-Ski2 phylogenetic trees are congruent. Tips from the same genome are linked by a red line. Colour code for taxonomic order as in Figure 4.
**Figure S2** - *In vitro* co-purification assays. (A) Affinity purification controls. Untagged aRNase J (top) and Pab-ASH-Ski2 (bottom) are not intrinsically retained on nickel column matrices. (B) Sedimentation profiles of Pab-aRNase J-(His)$_6$ (2µM) and Pab-(His)$_6$-ASH-Ski2 (1µM) on linear PBS1X 5-25% sucrose gradients (200 000g during 20h). Upper panel: purified protein alone. Lower panel: mixture of both proteins. Proteins were separated on a 4-15% SDS-polyacrylamide gel and monitored by Western blot using specific antibodies. (C) Co-purification assays in which *P. abyssi* aRNase J-(prey) was challenged with Pab-(His)$_6$-Rrp4 and Pab-DNaG-(His)$_6$ (baits) respectively. (D) Co-purification assays in which *P. abyssi* ASH-Ski2 (prey) was challenged with Pab-(His)$_6$-Rrp4, Pab-DNaG-(His)$_6$, Pab-(His)$_6$-Csl4, Pab-(His)$_6$-Rrp41 and Pab-DNaG-(His)$_6$ (baits) respectively. Legend as in Figure 6.
**FIGURE S3-** Csl4 structural domains. (A) Weblogo sequences derived from multiple alignments of Csl4 sequences co-distributed or not with aRNase J (referred to distribution of aRNase J & Csl4 displayed on Fig. 4). Several residues are specifically conserved in Csl4 sequences co-distributed with aRNase J. (B) The structural model for *P. abyssi* Csl4 is built using Phyre2 software. The solvent-exposed phenylalanine residues F121 and F129 are indicated by stars. S1 is for S1 domain and Zn for Zinc Finger motif containing domain.
FIGURE S4 - Sedimentation profiles of Pab-aRNase J, Pab-ASH-Ski2 and Pab-Rrp41 from WT strain P. abyssi cell extract performed in TK in (A) and TK-EDTA in (B) buffers, as for Figure 7. In here, the endogenous Pab-aCPSF1 β-CASP endo-RNase is also monitored as control.
FIGURE S5 - Sedimentation profiles of Tba-aRNase J, Tba-ASH-Ski2 and Tba-Rrp41 from the Tba-WT in (A), Tba-ΔASH-Ski2 in (B), and Tba-ΔaRNase J in (C), strain cell extracts prepared in TK-EDTA buffer. Legend as for Figure 7.
| GENES  | PROTEINS                          | N-His | C-His | pET11b |
|-------|-----------------------------------|-------|-------|--------|
| pab1751 | aRNase J                         | ✓     | ✓     | ✓      |
| pab2313 | ASH-Ski2                         | ✓     | ✓     |        |
|         | ΔC-ASH-Ski2                      | ✓     |       |        |
|         | ΔN-ASH-Ski2                      | ✓     |       |        |
|         | ASH-Ski2-C124A                   | ✓     |       |        |
|         | DomN                             | ✓     |       |        |
|         | DomN C124A                       | ✓     |       |        |
| pab0592 | Hel308                           | ✓     |       |        |
| pab2314 | Csl4                             | ✓     | ✓     |        |
|         | ΔC-Csl4                          | ✓     |       |        |
|         | ΔN-Csl4                          | ✓     |       |        |
|         | Csl4-F121A                       | ✓     |       |        |
|         | Csl4-F129A                       | ✓     |       |        |
| pab0316 | DnaG                             |       | ✓     |        |
| pab0419 | Rrp4                             | ✓     |       |        |
| pab0420 | Rrp41                            | ✓     |       |        |

**Table S1** - pET vector constructions used to express recombinant proteins in *E.coli*. 
| Oligonucleotides | Sequences (5’-3’) |
|-----------------|-------------------|
| Reverse PCR     |                   |
| For pET15b      | CATATGGCTGCGCCGCGGC |
| Rev pET15b / pET11b | GGATCCGCCTGCTAAGAAAGCC |
| For pET11b      | CATGTTATATCTCTCTTTGATTT |
| For aRNase J pET11b | GAGATATACATATGTGGAGAGATAAACATGATCA |
| Rev aRNase J pET11b | TTAGAGCCGGATCTCTACGTCCTAAATGACG |
| For ASH-Ski2 pET15b | CGGCGAGCCATATGCTATTTTCGATGAGCTACAAGC |
| Rev ASH-Ski2 pET15b | AGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For ASH-Ski2 pET11b | GAGATATACATATGTGCTATTTTCGATGAGCTACAAGC |
| Rev ASH-Ski2 pET11b | TTAGAGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For ASH-Ski2ΔN-ter (194-855) pET15b | CCGCCGGAGCCATATGAGCAGATGGAGATGATC |
| Rev ASH-Ski2ΔN-ter (1-573) pET15b | AGCCGGATCTCTGAGTTAACAGCTATCAGCATGAG |
| Rev DomN (1-193) ASH-Ski2 pET15b | TTAGAGCCCGATCTTTAATAGGTTACATCTGAG |
| For Hel308 pET15b | CGGCGAGCCATATGCTATTTCGATGAGCTACAAGC |
| Rev Hel308 pET15b | TTAGAGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For Csl4 pET15b / pET11b | CGGCGAGCCATATGCTATTTTCGATGAGCTACAAGC |
| Rev Csl4 pET15b | TTAGAGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For Csl4 pET11b | GAGATATACATATGATGCTATTTTCGATGAGCTACAAGC |
| Rev Csl4 pET11b | TTAGAGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For Csl4ΔN-ter pET15b | AGGATCTTTTCGTCATCAGTGAATC |
| Rev Csl4ΔN-ter pET15b | CCTTCAATACCTAAGAAAAGGGG |
| For For DnaG pET21b | GAGATATACATATGATGAAAGATTTTTGTTCAAAAT |
| Rev DnaG pET21b | TTAGAGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For Rrp4 pET15b | CGGCGAGCCATATGCTATTTTCGATGAGCTACAAGC |
| Rev Rrp4 pET15b | TTAGAGCCGGATCTCTGTTATGTTTTCTTCTCCTAATGTC |
| For Rrp41 pET15b | AGGATCTTTTCGTCATCAGTGAATC |
| Rev Rrp41 pET15b | CCTTCAATACCTAAGAAAAGGGG |
| For aRNase J Δ1 pET15b | GATTAGAACCCTGTCGAGCCTATTTC |
| Rev aRNase J Δ1 pET15b | GACCTTAAAGAAGCTCGGAATAAC |
| For ASH-Ski2 C124A | GAGAGTACATAGCCGAGAGATGTGA |
| For ASH-Ski2 C124A | CACATCTCTCGCTATGTACCTC |
| For Csl4 F121A | GTTAAGGAGCGCGCCGTAGGAGATTAAGAAG |
| Rev Csl4 F121A | TAAATCCTACAGCGCCGCTCTTAACTTGAG |

**Table S2** - List of oligonucleotides used in this study
| ID     | ORF    | Name     | Description                                                                 | Ref. Spectra | Spec. Index | Nuclease Effect (%) |
|--------|--------|----------|------------------------------------------------------------------------------|--------------|-------------|---------------------|
| Q9V119 | PAB0420| Rrp41    | RNA Exosome core subunit Rrp41                                                | 7.64         | 0           | 76.78               |
| G8ZHS0 | PAB2163| RPA41    | RPA41 subunit                                                                | 6.69         | 0           | 8.94                |
| Q9V2F9 | PAB0071| AubA     | RNA-binding protein AU-1                                                      | 6.03         | 0           | 100.00              |
| Q9V118 | PAB0421| Rrp42    | RNA Exosome core subunit Rrp42                                                | 5.37         | 0           | 59.81               |
| Q9UYB6 | PAB1306| eIF-2B   | Translation initiation factor eIF-2B subunit 2-like                          | 4.67         | 0           | 10.53               |
| Q9V120 | PAB0419| Rrp4     | RNA Exosome cap subunit Rrp4                                                  | 4.65         | 0           | 55.20               |
| Q9V2M1 | PAB2313| ASH-Ski2  | ATP-dependant RNA Helicase                                                    | 4.61         | 0           | 57.80               |
| Q9V113 | PAB0425| RpoA2    | DNA-directed RNA polymerase subunit A''                                      | 4.28         | 0           | 85.48               |
| Q9V114 | PAB0424| RpoA1    | DNA-directed RNA polymerase A'                                                | 3.97         | 0           | 58.80               |
| Q9UY85 | PAB1284| RecJ-like| RecJ-like phosphoesterase                                                      | 3.68         | 0           | 0.00                |
| Q9UY07 | PAB1115| Grad-2   | Glucose-1-phosphate thymidylyltransferase                                    | 3.67         | 0           | 0.00                |
| Q9UYD9 | PAB1036| Unch,    | Uncharacterized protein                                                       | 3.33         | 0           | 30.22               |
| Q9UYJ6 | PAB1366| Unch,    | DHH family phosphohydrolase                                                   | 2.34         | 0           | 43.59               |
| Q9V1Z1 | PAB2165| RPA32    | RPA32 subunit                                                                | 2.34         | 0           | 0.00                |
| Q9V1F2 | PAB0316| DnaG     | DNA exosome cap subunit DnaG                                                   | 30.88        | 0.02456     | 8.14                |
| Q9V115 | PAB0423| RpoB     | DNA-directed RNA polymerase subunit beta                                     | 10.32        | 0.03366     | 46.07               |
| Q9V1T4 | PAB2119| Unch,    | Uncharacterized RNA methyl transferase-like                                  | 6.65         | 0.05610     | 1.43                |
| Q9V1U8 | PAB2397| Rps4e    | 30S ribosomal protein S4e                                                     | 2.51         | 0.06917     | 14.00               |
| P61881 | PAB3089| HarA     | Archaeal histone A                                                           | 4.31         | 0.07685     | 0.00                |
| Q9V1F5 | PAB3117| HarB     | Archaeal histone B                                                           | 4.25         | 0.09445     | 0.00                |
| Q9UWR8 | PAB1166| Rpl1     | 50S ribosomal protein L1                                                       | 3.36         | 0.09646     | 91.01               |
| G8ZJV0 | PAB1300| ApeH     | ApoE acylamino-acid-releasing enzyme                                          | 1.83         | 0.12459     | 0.00                |
| Q9UX0  | PAB1686| Unch,    | Uncharacterized Cas7/DevR-like protein                                        | 2.29         | 0.15261     | 57.73               |
| Q9V2L5 | PAB2306| FlpA     | Fibrillarin-like rRNA/IRNA 2'-O-methyltransferase                            | 13.21        | 0.16112     | 47.16               |
| Q9UZ78 | PAB2428| TmcA     | RNA(Met) cytidine acetyltransferase                                          | 18.53        | 0.24122     | 0.00                |
| P61992 | PAB0361| Rps4     | 30S ribosomal protein S4                                                       | 1.69         | 0.38540     | 78.78               |
| Q9V1T5 | PAB2120| Rpl3     | 50S ribosomal protein L3                                                       | 2.57         | 0.40940     | 0.00                |
| Q9V174 | PAB1994| Unch,    | Uncharacterized DUF536 containing protein                                    | 1.58         | 0.48199     | 100.00              |
| Q9V2L4 | PAB2305| Nop5p    | Nop5 C/D snoRNP                                                               | 4.21         | 0.48716     | 0.00                |
| Q9UZ66 | PAB1633| PINA     | PIN domain ATPase                                                            | 2.89         | 0.49010     | 0.00                |
| Q9V0G8 | PAB1813| Rps19e   | 30S ribosomal protein S19e                                                     | 1.98         | 0.52322     | 100.00              |
| Q9UX5S | PAB1167| Rpl10    | 50S ribosomal protein L10                                                      | 2.31         | 0.57916     | 26.95               |
| Q9V018 | PAB1721| Unch,    | Uncharacterized hydroase, metallo-beta-lactamase family protein              | 1.65         | 0.61968     | 24.91               |
| Q9V1N3 | PAB7094| AlbA     | DNA/RNA-binding protein                                                        | 3.98         | 0.76378     | 100.00              |
| Q9UY49 | PAB1444| Rpl10e   | 50S ribosomal protein L10e                                                     | 1.50         | 0.77055     | 0.00                |
| Q9V135 | PAB0407| Top6B    | Type 2 DNA topoisomerase 6 subunit B                                         | 4.23         | 0.87969     | 0.00                |
| P62008 | PAB0460| Rpl7Ae   | 50S ribosomal protein L7Ae                                                     | 1.41         | 0.89414     | 24.30               |

**Table S3.** List of potential proteins partners identified by pulldown-MS/MS with **Pab-aRNase J-(His)_{6}** as bait protein. Controls were run 6 times whereas “no nuclease” and “nuclease” assays were run in triplicate.
| ID      | ORF    | Name               | Description                              | Ref, Spectra | Spec, Index | Nucl, Effect (%) |
|---------|--------|--------------------|------------------------------------------|--------------|-------------|------------------|
| G8ZI82  | PAB1751| aRNaseJ            | 5'-3' Exoribonuclease                     | 10.73        | 0           | 24.14            |
| Q9V120  | PAB0419| Rrp4               | RNA Exosome cap subunit Rrp4             | 9.95         | 0           | 100              |
| Q9V133  | PAB2412| Unch,             | Uncharacterized DUF530 containing protein| 8.37         | 0           | 0                |
| Q9V2F9  | PAB0071| AubA              | RNA-binding protein AU-1                 | 8.08         | 0           | 97.28            |
| Q9U1Y85 | PAB1284| RecJ-like         | RecJ-like phosphoesterase                | 8.00         | 0           | 36.25            |
| Q9V114  | PAB0424| RpoA1             | DNA-directed RNA polymerase A'           | 7.30         | 0           | 40.17            |
| Q9V118  | PAB0421| Rrp42             | RNA Exosome core subunit Rrp42           | 6.40         | 0           | 100              |
| Q9V119  | PAB0420| Rrp41             | RNA Exosome core subunit Rrp41           | 5.28         | 0           | 100              |
| Q9U1Y91 | PAB1064| Unch,             | Uncharacterized Cas6-like protein        | 4.26         | 0           | 100              |
| Q9V113  | PAB0425| RpoA2             | DNA-directed RNA polymerase subunit A''   | 4.00         | 0           | 39.13            |
| Q9UWR8  | PAB1166| Rpl1              | 50S ribosomal protein L1                  | 3.49         | 0           | 100              |
| Q9V236  | PAB0161| Unch,             | Uncharacterized translation initiation factor IF2/IF5-like | 3.49       | 0           | 54.41            |
| Q9UX7   | PAB1136| Rnp3              | Ribonuclease P protein component 3       | 3.47         | 0           | 21.67            |
| Q9UZK6  | PAB1613| Unch,             | CRISPR-associated endoribonuclease       | 2.91         | 0           | 89.50            |
| P61992  | PAB0361| Rps4              | 30S ribosomal protein S4                  | 2.75         | 0           | 0                |
| GZI80   | PAB2163| RPA41             | RPA41 subunit                            | 14.21        | 0.00957     | 5.48             |
| Q9V1V6  | PAB2137| Rpl30             | 50S ribosomal protein L30                | 4.93         | 0.01919     | 10.52            |
| Q9UZ3   | PAB1584| Unch,             | Uncharacterized DUF2184 containing protein| 4.34       | 0.02184     | 100              |
| Q9V0G8  | PAB1813| Rps19e            | 30S ribosomal protein S19e               | 2.86         | 0.04021     | 20.37            |
| Q9UX5   | PAB1167| RplP0             | Acidic ribosomal protein P0 homolog      | 2.39         | 0.05071     | 4.54             |
| Q9V089  | PAB2390| Sun protein       | Sun protein (Fmu protein)                | 20.47        | 0.05175     | 22.35            |
| P62008  | PAB0460| Rpl7Ae            | 50S ribosomal protein L7Ae               | 8.33         | 0.10606     | 70.18            |
| Q9V196  | PAB0365| Rpl13             | 50S ribosomal protein L13                | 3.07         | 0.13018     | 1.21             |
| Q9V1V5  | PAB2136| Rps5              | 30S ribosomal protein S5                  | 6.13         | 0.14079     | 10.91            |
| Q9V115  | PAB0423| RpoB              | DNA-directed RNA polymerase subunit beta  | 20.87        | 0.14375     | 39.93            |
| Q9UZ6   | PAB1633| PINA              | PIN domain ATPase                        | 2.74         | 0.18422     | 0                |
| Q9V1U6  | PAB2436| Rpl14             | 50S ribosomal protein L14                | 2.72         | 0.18868     | 42.52            |
| Q9VL4   | PAB2305| Nop5p             | Nop5 C/D snoRNP                          | 11.04        | 0.27236     | 2.64             |
| Q9V1F2  | PAB0316| DnaG              | RNA exosome cap subunit DnaG             | 11.42        | 0.30179     | 21.99            |
| Q9V1N3  | PAB7094| AlbA              | DNA/RNA-binding protein                  | 19.59        | 0.40021     | 95.47            |
| Q9V1V4  | PAB2135| Rpl18             | 50S ribosomal protein L18                | 2.16         | 0.47113     | 89.28            |
| Q9UZ78  | PAB2428| TmcA              | tRNA(Met) cytidine acetyltransferase     | 24.85        | 0.51726     | 0                |
| Q9VL5   | PAB2306| FlpA              | Fibrillar-like RNA/rRNA 2'-O- methyltransferase | 8.84       | 0.52605     | 22.83            |
| Q9VI15  | PAB2120| Rpl3              | 50S ribosomal protein L3P                 | 3.63         | 0.57003     | 29.48            |

**Table S4.** List of potential proteins partners identified by pulldown-MS/MS with **Pab-(His)6-ASH-Ski2** and **Pab-ASH-Ski2-(His)6** as bait proteins, respectively. Controls were run 8 times whereas "no nuclease" and "nuclease" assays were run in quadruplicate.