BIOPHYSICS AND MEDICAL PHYSICS

Original article
DOI: https://doi.org/10.18721/JPM.17109

AN ANALYSIS OF THE ASGARDARCHAEA Cas1_3 PROTEIN: EXPERIMENTAL CHARACTERIZATION OF A POTENTIAL INTERMEDIATE IN THE EVOLUTION OF CRISPR-Cas SYSTEMS

M. V. Abramova[1], *A. S. Malykh*[1], *I. M. Gatieva*[1],
M. A. Kazalov[1], *P. A. Selkova*[1], *A. P. Yakimov*[1],
A. A. Vasileva[1, 2], *A. N. Arseniev*[1, 2], *M. A. Khodorkovskii*[1] ✉

[1] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia;
[2] Complex of NBICS Technologies, National Research Center "Kurchatov Institute", Moscow, Russia

✉ khodorkovskii@gmail.com

Abstract. The evolution of CRISPR-Cas systems and their possible origin from mobile genetic elements of transposons are currently being actively studied. Recently, unique systems have been discovered in a new group of Asgardarchaea that presumably function as transposons and contain Cas1-like proteins. In this study, the genetic and biochemical technologies were used, along with electrophoresis, affinity chromatography, and high-resolution mass spectrometry, to obtain and partially characterize a recombinant version of one of these proteins, Cas1_3. In particular, it was shown to have the ATPase activity, the quantitative value of the latter being determined by the spectrophotometric method. The results obtained may be useful in understanding the mechanisms of functioning the potential ancestor of CRISPR-Cas systems.

Keywords: characterization, CRISPR-Cas, evolution intermediate form, Asgardarchaea, transposon, recombinant protein

Funding: The reported study was funded by Russian Science Foundation (Grant No. 21-14-00122), and by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-15-2021-1062).

Citation: Abramova M. V., Malykh A. S., Gatieva I. M., Kazalov M. A., Selkova P. A., Yakimov A. P., Vasileva A. A., Arseniev A. N., Khodorkovskii M. A., An analysis of the Asgardarchaea Cas1_3 protein: Experimental characterization of a potential intermediate in the evolution of CRISPR-Cas systems, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 81–92. DOI: https://doi.org/10.18721/JPM.17109

Научная статья
УДК 577.322
DOI: https://doi.org/10.18721/JPM.17109

АНАЛИЗ БЕЛКА Cas1_3 АСГАРДАРХЕЙ: ЭКСПЕРИМЕНТАЛЬНАЯ ХАРАКТЕРИСТИКА ПОТЕНЦИАЛЬНОГО ПРОМЕЖУТОЧНОГО ЗВЕНА В ЭВОЛЮЦИИ СИСТЕМ CRISPR-Cas

М. В. Абрамова[1], *А. С. Малых*[1], *Я. М. Гатиева*[1],
М. А. Казалов[1], *П. А. Селькова*[1], *А. П. Якимов*[1],
А. А. Васильева[1, 2], *А. Н. Арсениев*[1, 2], *М. А. Ходорковский*[1] ✉

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;
[2] Комплекс НБИКС-природоподобных технологий, Национальный исследовательский центр «Курчатовский институт», Москва, Россия
✉ khodorkovskii@gmail.com

Аннотация. В настоящее время актуальны исследования эволюции систем CRISPR-Cas и их возможного происхождения от мобильных генетических элементов транспозонов. Недавно у новой группы асгардархей были обнаружены уникальные системы, которые предположительно функционируют как транспозоны и содержат Cas1-подобные белки. В настоящей работе для одного из таких белков, Cas1_3, впервые с использованием генетических и биохимических технологий, а также методов электрофореза, аффинной хроматографии и высокоразрешающей масс-спектрометрии получена и частично охарактеризована его рекомбинантная версия. В частности, показано наличие у него АТФ-азной активности, количественное значение которой было определено с помощью метода спектрофотометрии. Полученные результаты могут быть полезны в понимании механизмов функционирования потенциального предшественника систем CRISPR-Cas.

Ключевые слова: характеризация, CRISPR-Cas, промежуточное звено эволюции, асгардархеи, транспозон, рекомбинантный белок

Финансирование: Исследование выполнено при финансовой поддержке Российского научного фонда (грант № 21-14-00122), а также Министерства науки и высшего образования Российской Федерации (соглашение № 075-15-2021-1062).

Для цитирования: Абрамова М. В., Малых А. С., Гатиева Я. М., Казалов М. А., Селькова П. А., Якимов А. П., Васильева А. А., Арсениев А. Н., Ходорковский М. А. Анализ белка Cas1_3 асгардархей: экспериментальная характеристика потенциального промежуточного звена в эволюции систем CRISPR-Cas // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 1. С. 81–92. DOI: https://doi.org/10.18721/JPM.17109

Introduction

CRISPR-Cas are specific systems of bacteria and archaea protecting them against the invasion of harmful genetic material (these can be viruses, mobile genetic elements, etc.). Combined with ribonucleic acids (RNA), proteins of such a system are targeted recognize foreign elements and insert spacer sequences into them, leading to subsequent degradation. The CRISPR-Cas system includes a CRISPR cassette and genes encoding Cas proteins [1]. The CRISPR cassette is a locus in the genome transcribed as short palindromic repeats separated by small unique spacer sequences. The latter are fragments of foreign genetic material stored in the genome of the

microorganism from previous infection. Spacers appear due to the adaptation process, i.e., the insertion of foreign genetic material into the CRISPR cassette. The Cas1 and Cas2 proteins are responsible for the adaptation stage in most CRISPR-Cas systems [2].

CRISPR-Cas systems are actively studied from an evolutionary perspective. This not only expands the fundamental understanding of the biology of such systems, but also offers potential for developing new technologies and improveming existing techniques for gene editing. The search for possible evolutionary intermediates can lead to the discovery of proteins useful for subsequent practical applications.

Earlier, bioinformatics methods traced the evolution of different CRISPR-Cas systems. It has been hypothesized that part of the adaptation modules has an evolutionary origin from transposon-encoded integrases, i.e., DNA sequences capable of moving from one location in the genome to another [3, 4]. In 2020, unique systems were discovered in nucleotide sequences of a new group of Asgard archaea (Asgardarchaeota), presumably functioning as transposons. All these systems are united by the presence of large Cas1-like proteins, which are homologues of the Cas1 protein from the CRISPR-Cas system. One of such systems, aCas1_3, can be regarded as a model for one of the ancestors of the CRISPR-Cas system [5].

In this study, we obtained for the first time a recombinant version of the Cas1_3 protein, a giant Cas1-like protein from the presumably transposable aCas1_3 system in Asgard archaea. In addition to obtaining a recombinant version of the protein, we carried out its primary functional characterization.

Materials and methods

Bioinformatic analysis of the Cas1_3 protein sequence. The MOTIF library (https://www.genome.jp/tools/motif/)was used to identify the motifs in the amino acid sequence of the protein. The search was carried out over the PROSITE [6], NCBI-CDD [7] and Pfam libraries [8].

Cloning of the Cas1_3 protein gene for expression in producing cells. DNA sequences of Asgard archaea carrying proteins of the aCas1_3 system exist only as metagenomic data. This paper considered the aCas1_3 protein system, found in the *Candidatus Thorarchaeota archaeon* isolate Yap500.bin9.44 sequence from the GenBank database (ID: JAEOSL000000000.1).

The codon-optimized sequence of the Cas1_3 protein gene was synthesized at Cloning Facility (Moscow, Russia) and cloned into the pEXPR_001 vector. A PCR fragment carrying the Cas1_3 protein sequence was amplified from the resulting pEXPR_001_Cas1_3 plasmid (a small circular DNA molecule) by polymerase chain reaction (PCR). A PCR fragment carrying the sequence of the maltose binding protein (MBP) gene was also amplified from the standard pMJ806 vector (Addgene, USA).

The PCR fragments were amplified using a Q5 High-Fidelity 2X Master Mix (NEB, UK) in accordance with the manufacturer's protocol. The sequences of the primers used are given in Table.

The obtained PCR fragments carrying sequences of Cas1_3 (4986 bp) and MBP (1101 bp) proteins were cloned into the rET21a vector by the Gibson method [9] using Gibson Assembly Master Mix (NEB, UK) (recall that the length of sequenced DNA regions (genes, sites, chromosomes) or the entire genome is measured in base pairs of nucleotides (bp)). NdeI and NotI restriction endonuclease sites were selected for cloning.

Expression of the Cas1_3 protein gene in producing cells. The pET21a_Cas1_3_MBP plasmid was transformed into competent Escherichia coli (*E. coli*) BL21 AI cells using standard methods. The resulting cells were grown in lysogeny broth (LB) with ampicillin (100 mcg/ml) in a volume of 500 ml. Cell cultures were incubated at a temperature of 37 °C and uniform aeration on an orbital rocker until an optical density of 0.6 at a wavelength of 600 nm was reached. Transcription of the target Cas1_3 protein gene was induced by adding arabinose and isopropyl-β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1% and 1 mM, respectively. Cell culture cultivation was continued at a temperature of 18 °C for 16 hours. The cells were collected by centrifugation at a centrifugal acceleration of 3500 g for 30 minutes. The resulting pellet was stored at −20 °C.

The induction of Cas1_3 protein transcription in the obtained pellet was checked by the protocol described earlier [10].

Solubility prediction for Cas1_3 protein. A 1 ml cell culture was selected after incubation with arabinose and IPTG (see the previous section). The cells were collected by centrifugation

Table

Primer sequences used in the study

Sequence, 5′ → 3′ [primer designation]
AACCTGTACTTCCAATCCAATATTGGAATGACACGATCAAGAGGAAGGCCCG [Cas1_3 forward primer]
TGGTGCTCGAGTGCGGCCGCAAGCTTTACAACTGCTGCTCCTCAGACTCCGTCGGTTCGC [Cas1_3 reverse primer]
CTTTAAGAAGGAGATATACATATGCACCACCACCACCACCACATGAAAATCGAAGAAGGTAAAC [MBP forward primer]
ATATTGGATTGGAAGTACAGGTTTTCCTCGATAGTCTGCGCGTCTTTCAGGGCTTC [MBP reverse primer]

Notations: forward and reverse primers are given.

at 12,000 g for 10 minutes. The pellet was resuspended in 700 μl of lysing buffer comprising 50 mM Tris, 5% glycerol, NaCl solution of different concentrations (150, 500 or 1000 mM); the acidity (pH) of the buffer was 6.5 or 8.0. Cell lysis was performed by ultrasonic disruption in an ice bath. Ultrasonic disruption was performed with a LABSONIC device (Sartorius AG, Germany) at 30% amplitude for 0.2 s. The duration of each treatment cycle was 2 minutes, each subsequent cooling cycle lasted 2 minutes; the total treatment time was 10 minutes.

The sample was centrifuged at a centrifugal acceleration of 4000 g and a temperature of 4 °C for 10 minutes, then the supernatant was separated from the pellet. 700 μl of lysing buffer was added to the pellet, resuspended, and then 15 μl were taken from the pellet and from the supernatant for composition analysis. The analysis was performed by electrophoresis in denaturing 10% polyacrylamide gel (PAAG) using the Laemmli method [11].

Preparation of recombinant Cas1_3 protein. The pellet obtained from 500 ml of liquid cell culture containing the induced Cas1_3 protein was resuspended in 15 ml of lysing buffer with lysozyme and incubated on ice for 20 minutes.

The composition of the lysing buffer (pH = 8.0): 50 mM Tris ($C_4H_{11}NO_3$), 5% glycerin, NaCl solution with a concentration of 150 mM, 1 mg/ml lysozyme.

Cell lysis was performed by ultrasonic disruption in an ice bath by the protocol described in the previous section, with an extended treatment time (up to 30 minutes). The soluble part of the lysate and the pellet were separated by centrifugation at 16,000 g and 4 °C for 60 minutes. The supernatant was used for further purification.

Affinity chromatography was performed with MBPTrap HP Columns (GE Healthcare, USA). The samples with protein were concentrated using Amicon Ultra-4 Centrifugal Filter Unit (Merck Millipore, USA) centrifugal concentrators with a 100 kDa filter.

For gel filtration, the resulting sample was applied to a Superose 6 Increase 10/300 GL column (GE Healthcare, USA) in different buffers (see Results and Discussion below). Fractions containing the target protein were collected, combined and concentrated using centrifugal concentrators with a 100 kDa filter.

Glycerin was added to the protein concentrates to a final concentration of 10%. The purified protein was stored at a temperature of −80 °C.

Quantitative analysis of ATPase activity of Cas1_3 protein. The method described in [12] was used to quantify the ATPase activity (recall that ATP is adenosine triphosphate). A buffer consisting of 25 mM Tris (pH = 7.5), 10 mM $MgCl_2$, 2 mM ATP, 2−4 mM phosphoenolpyruvate (PEP), 50 units/ml lactate dehydrogenase, 50 units/ml pyruvate kinase and 4−5 mM reduced nicotinamide adenine dinucleotide (NADH) was used for this purpose.

First, we measured the background characterizing the hydrolytic activity of the buffer, then the activity of the protein in the buffer; at the last stage, the activity of the protein in the buffer was measured in the presence of oligonucleotides consisting of deoxythymines (oligo-dT).

Absorption curves for NADH were obtained at a wavelength of 340 nm using a Cary 5000 spectrophotometer (Varian, USA). To ensure the reliability of the data obtained, the experiment was repeated three times.

Results and discussion

Cas1_3 is a representative of a unique group of Cas1-like proteins of the aCas1_3 Asgard archae systems. This group of proteins has not been studied sufficiently and has not been characterized experimentally before.

Cas1_3, like all other representatives of this group, is a very large protein compared with the well-studied Cas1 proteins that are part of CRISPR-Cas systems. It contains 1,368 amino acid residues, while the Cas1 protein, for example, of the CRISPR-Cas system of the *Streptococcus pyogenes* SF370 (M1 GAS) type II-A bacterium, contains only 289 amino acid residues. Such a difference in size suggests the likely presence of additional domains in the structure of the Cas1_3 protein.

The first step in this study was to analyze which additional domains may be part of the Cas1_3 protein, allowing to draw immediate conclusions about their potential function. Since the spatial model of Cas1_3 has not been solved, such analysis was possible only based on available data on the amino acid sequence of the protein. We searched for the so-called motifs, i.e., the short characteristic sequences of amino acids associated with certain functions.

We were able to detect motifs similar to Walker motifs in the amino acid sequence of the Cas1_3 protein. In particular, the Walker A motif is a characteristic sequence GXXXXGKT/S of amino acid residues, where X is any amino acid residue [13]. The given motif can be represented by amino acid residues in the sequence of the Cas1_3 protein (G143-···-G150-K151-T152). Furthermore, a site similar to the Walker B motif can be found in the amino acid sequence of the Cas1_3 protein (GXXXLhhhD, where X is any amino acid residue and h is a hydrophobic amino acid residue). This site is formed by amino acid residues G167-···-L172-A173-H174-P175-D176.

It is known that the location of the Walker B motif downstream of the Walker A motif is quite characteristic for proteins with ATPase activity. The sequences of these motifs bind to the ATP and are involved in transferring the energy released during hydrolysis of the ATP molecule to the structural changes of the protein. This allows proteins containing Walker motifs to participate in functions involved in active transport, cellular mobility, and other biological processes. This refers to the processes associated with movement and transport within the cell, ensuring efficient use of the energy released during ATP hydrolysis [14]. As noted above, the Cas1_3 protein is part of the system that most likely acts as a transposon (a mobile genetic element). It is reasonable to assume that the additional energy obtained during ATP hydrolysis by the Cas1_3 protein may help the transposon to incorporate into the genome, thus triggering its mobility.

To test the ability of the Cas1_3 protein to hydrolyze ATP, we obtained its recombinant form. For this purpose, its gene was cloned into a construct for expression in producing cells. Additionally, we combined the gene of the target protein with the sequence of the MBP protein gene. The MBP is used to increase the solubility of the target protein in the lysing buffer for its further purification by affinity chromatography [15]. The resulting pET21a_Cas1_3_MBP plasmid is shown in Fig. 1,*a*.

The insertion of target genes into the vector was verified by restriction analysis: the resulting pET21a_Cas1_3_MBP plasmid was digested with restriction endonucleases NdeI, EcoRI, BamHI. Such digestion of the plasmid should produce fragments with a length of approximately 1334 bp, 1841 bp, 2643 bp and 5710 bp (see Fig. 1,*b*). In the case of such treatment of the original pET21a vector without insertion, fragments with the lengths of 5399, 38 and 6 bp should be formed . The sizes of the fragments obtained by restriction correspond to the expected ones (see Fig. 1,*b*), which confirms the insertion of target genes into the pET21a vector. For additional verification, the cloned gene was sequenced (determination of its primary nucleotide sequence) using the Sanger method.

To subsequently isolate the recombinant Cas1_3_MBP protein, we needed to test whether its synthesis is induced in producing cells and whether this protein is soluble.

The verification results for transcription induction of the Cas1_3_MBP protein gene are shown in Fig. 2,*a*. A product corresponding to the expected molecular weight of the target protein (232 kDa) was recorded in the 'after induction' sample, which indicated successful synthesis of Cas1_3_MBP in producing cells.

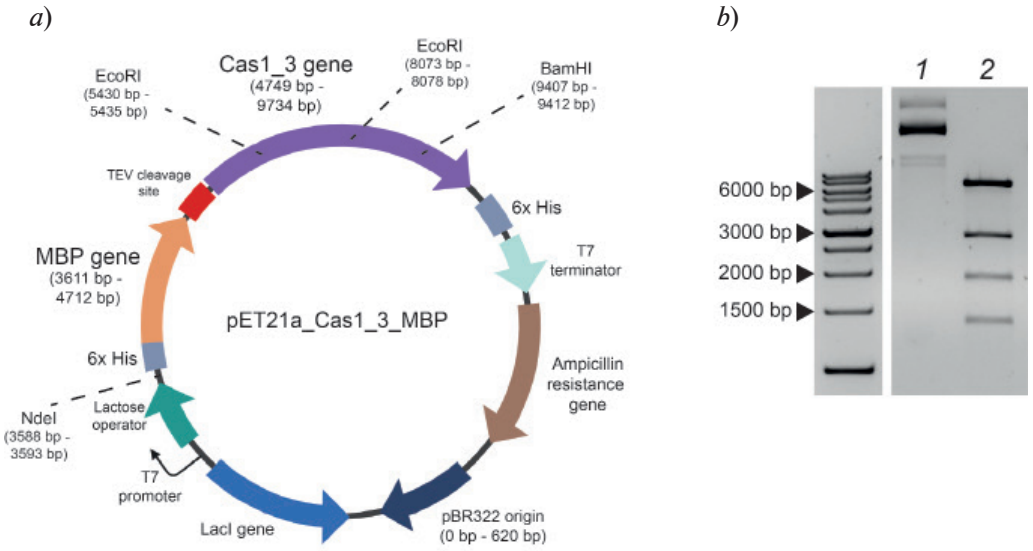

Fig. 1. Result of cloning the genes of Cas1_3 and MBP proteins into the pET21a vector: scheme of the pET21a_Cas1_3_MBP plasmid (carries the genes of the Cas1_3 and MBP proteins) (*a*); results of restriction analysis of this plasmid (*b*).

Fig. 1,*b*: analysis with control without restrictases (*1*) and restricted plasmid with NdeI, EcoRI, BamHI sites (*2*); bp is the base pair

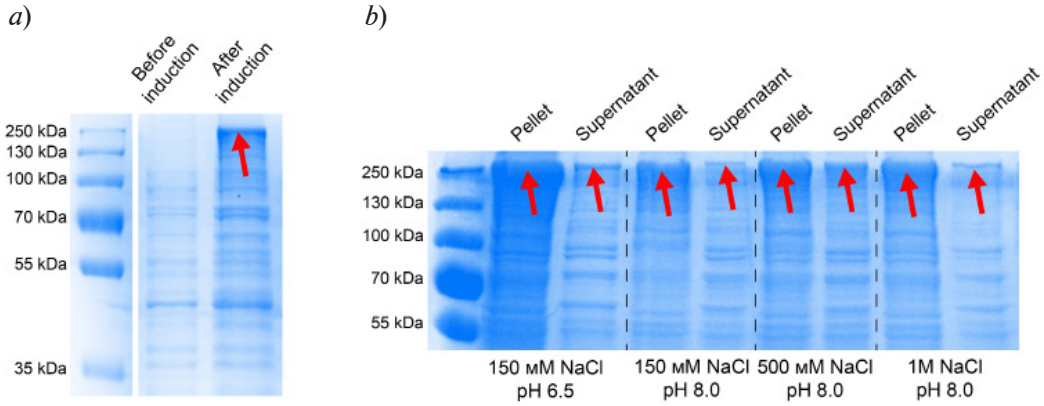

Fig. 2. Results of electrophoretic analyses for samples obtained during two tests: induction of Cas1_3_MBP protein synthesis (a) and solubility (b).
The expected molecular weight of the protein was 232 kDa.
Red arrows indicate the positions of the target protein

The solubility of the protein had to be verified chromatographic purification. Therefore, the conditions under which the Cas1_3_MBP protein would have the greatest solubility were selected at the next stage. To select the best conditions, lysing buffers with different concentrations of sodium chloride NaCl (150, 500 and 1000 mM) and pH (6.5 and 8.0) were used. Products corresponding to the molecular weight of the Cas1_3_MBP protein were found in all samples: both in the pellet and in the supernatant (Fig. 2,*b*), which indicated the solubility of the protein in all tested buffer solutions. A buffer with the lowest amount of NaCl (150 mM) and pH = 8.0 was used for further purification of the protein, since higher NaCl concentrations may further interfere with *in vitro* experiments with this protein.

To isolate the functionally active Cas1_3_MBP protein, two-stage purification was performed using affinity chromatography and gel filtration methods.

Gel filtration was carried out in several configurations. Buffers with different NaCl contents were used for this purpose. In addition, we carried out purification in a buffer with 2M NaCl with the protein solution treated with benzonase and TEV protease before gel filtration. Benzonase

treatment was performed to remove nucleic acids from the protein solution, TEV-protease treatment was performed to cleave the MBP peptide from Cas1_3. It turned out that a peak consisting of Cas1_3_MBP appears on the chromatogram in the region of 12–14 ml when the protein is purified in a buffer containing 2M NaCl (Fig. 3). However, as previously noted, buffers with a high NaCl concentration may interfere with further experiments. For this reason, we decided to use protein fractions obtained after the affinity chromatography stage to verify the presence of ATPase activity.

The ability of the recombinant Cas1_3_MBP protein to hydrolyze ATP was tested *in vitro*. The activity of the RecA protein was measured simultaneously as a positive control. Fig. 4 shows the results of the experiment as a dependence of the quantity of hydrolyzed ATP on time. Evidently, the purified Cas1_3_MBP protein has ATPase activity with a constant hydrolysis rate, as does the RecA protein.

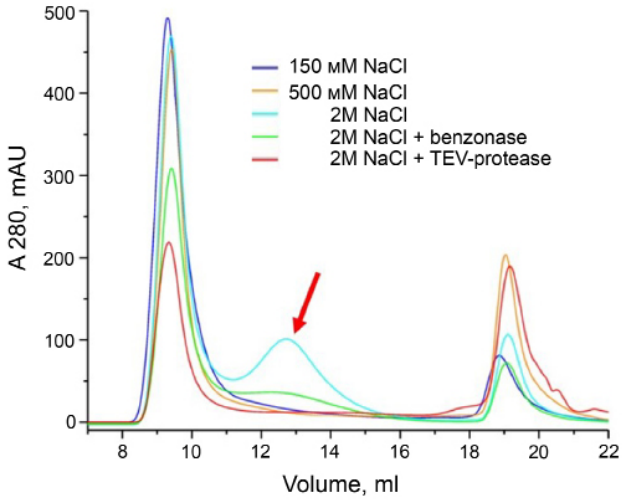

Fig. 3. Chromatographic profiles of gel-filtration purification of Cas1_3_MBP protein using various buffer conditions (see the legend). A peak related to this protein was detected at a volume of 13 ml (marked with a red arrow)upon isolation of this protein in 2M NaCl solution

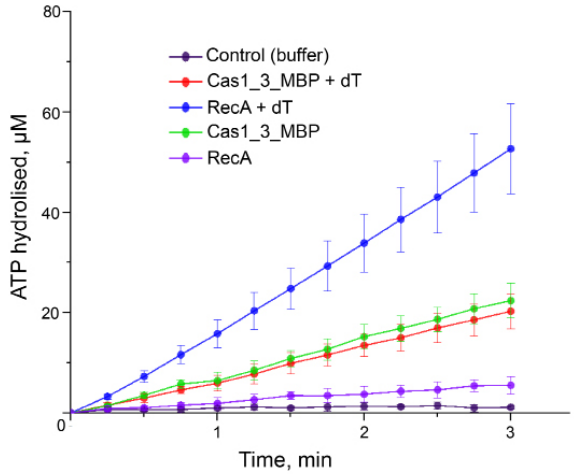

Fig. 4. Time dependences of ATP hydrolyzed by Cas1_3_MBP and RecA proteins in the presence and absence of oligonucleotides consisting of deoxythymines (oligo-dT), in solution

The rate of ATP hydrolysis by RecA protein ranged from 17±3 mmol/min with the addition of oligonucleotides consisting of deoxythymines (oligo-dT), which coincides with the literature data [16]. The hydrolytic activity of RecA depends on oligo-dT, therefore, ATP hydrolysis was not observed in their absence.

The rate of ATP hydrolysis by the Cas1_3_MBP protein was 6.0±1.5 mmol/min both in the presence and in the absence of oligo-dT. The lack of dependence of hydrolytic activity of Cas1_3_MBP on oligo-dT may be due to the fact that we were unable to completely eliminate the DNA from the cell culture during protein purification.

We verified that the observed ATP hydrolysis is due to the action of the Cas1_3_MBP protein by detecting the presence of protein impurities that could in any way affect the result obtained.

For this purpose, the method of high-resolution mass spectrometry was used (see, for example, [17]), which made it possible to verify that proteins released together with Cas1_3_MBP (Fig. 5) cannot make a significant contribution to the measured values of the ATP hydrolysis rate.

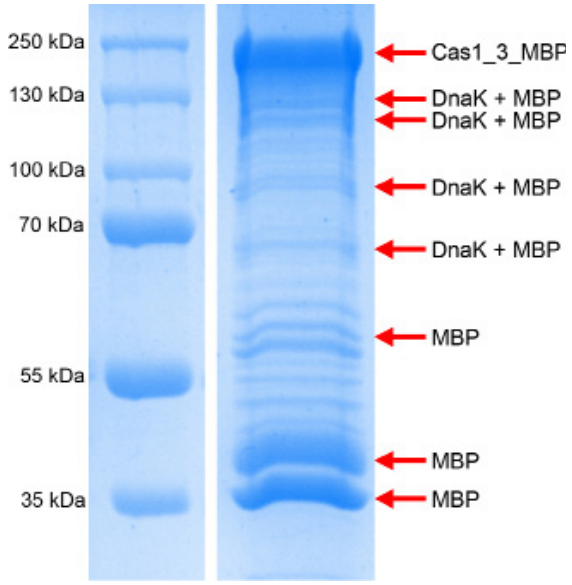

Fig. 5. Result of electrophoretic analysis of the sample obtained by purification of Cas1_3_MBP protein using affinity chromatography.
The red arrows indicate samples of protein impurities analyzed by mass spectrometry

Conclusion

We determined the conditions for obtaining a recombinant version of the Cas1_3 protein of the aCas1_3 Asgard archaea system, not characterized in any previous studies. This system most likely performs the functions of a transposon and is a potential evolutionary ancestor to CRISPR-Cas systems.

Our initial results demonstrate the ability of the Cas1_3 protein to hydrolyze ATP *in vitro*. The detection of ATPase activity allows to assess the possible role of ATP hydrolysis in the life cycle of the transposon as a mobile genetic element and ancestor of CRISPR-Cas systems.

New variants of the Cas1_3 protein should be created to further explain the mechanism of the detected ATPase activity, with mutated motifs potentially responsible for this activity.

REFERENCES

1. **Jansen R., Embden J. D., Gaastra W., Schouls L. M.,** Identification of genes that are associated with DNA repeats in prokaryotes, Mol. Microbiol. 43 (6) (2002) 1565–1575.

2. **Koonin E. V., Makarova K. S.,** Origins and evolution of CRISPR-Cas systems, Philos. Trans. R. Soc. B: Biol. Sci. 374 (1772) (2019) 20180087.

3. **Krupovic M., Shmakov S., Makarova K. S., et al.,** Recent mobility of casposons, self-synthesizing transposons at the origin of the CRISPR-Cas immunity, Genome Biol. Evol. 8 (2) (2016) 375–386.

4. **Faure G., Shmakov S. A., Yan W. X., et al.,** CRISPR-Cas in mobile genetic elements: counter-defence and beyond, Nature Reviews Microbiology. 17 (8) (2019) 513–525.

5. **Makarova K. S., Wolf Y. I., Shmakov S. A., et al.,** Unprecedented diversity of unique CRISPR-Cas-related systems and Cas1 homologs in Asgard archaea, CRISPR J. 3 (3) (2020) 156–163.

6. **Sigrist C. J., de Castro E., Cerutti L., et al.,** New and continuing developments at PROSITE, Nucl. Acids Res. 41 (D1, 1 Jan) (2013) D344–D347.

7. **Marchler-Bauer A., Zheng C., Chitsaz F., et al.,** CDD: conserved domains and protein three-dimensional structure, Nucl. Acids Res. 41 (D1, 1 Jan) (2013) D348–352.

8. **Finn R. D., Bateman A., Clements J., et al.,** Pfam: the protein families database. Nucl. Acids Res. 42 (D1, 1 Jan) (2014) D222–D230.

9. **Gibson D. G., Young L., Chuang R. Y., et al.,** Enzymatic assembly of DNA molecules up to several hundred kilobases, Nat. Methods. 6 (5) (2009) 343–345.

10. **Vasileva A. A., Aliukas S. A., Selkova P. A., et al.,** Type II CRISPR-Cas system nucleases: a pipeline for prediction and in vitro characterization. Molecular Biology (Moscow). 57 (3) (2023) 546–560 (in Russian).

11. **Laemmli U. K.,** Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature. 227 (5259) (1970) 680–685.

12. **Kreuzer K. N., Jongeneel C. V.,** *Escherichia coli* phage T4 topoisomerase. Methods Enzymol. 100 (1983) 144–160.

13. **Walker J. E., Saraste M., Runswick M. J., Gay N. J.,** Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold, EMBO J. 1 (8) (1982) 945–951.

14. **Ramakrishnan C., Dani V. S., Ramasarma T.,** A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins, Protein Eng. Des. Sel. 15 (10) (2002) 783–798.

15. **Sun P., Tropea J. E., Waugh D. S.,** Enhancing the solubility of recombinant proteins in *Escherichia coli* by using hexahistidine-tagged maltose-binding protein as a fusion partner, In book: Evans Jr., T., Xu M.Q. (Eds.) Heterologous Gene Expression in *E. coli*, Book series: Methods in Molecular Biology, Vol. 705. Humana Press, Totowa, New Jersey, USA (2011) 259–274.

16. **Kowalczykowski S. C., Clow J., Krupp R. A.,** Properties of the duplex DNA-dependent ATPase activity of *Escherichia coli* RecA protein and its role in branch migration, Proc. Natl. Acad. Sci. 84 (10) (1987) 3127–3131.

17. **Yakunina M., Artamonova T., Borukhov S., et al.,** A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage, Nucl. Acids Res. 43 (21) (2015) 10411–10420.

СПИСОК ЛИТЕРАТУРЫ

1. **Jansen R., Embden J. D., Gaastra W., Schouls L. M.** Identification of genes that are associated with DNA repeats in prokaryotes // Molecular Microbiology. 2002. Vol. 43. No. 6. Pp. 1565–1575.

2. **Koonin E. V., Makarova K. S.** Origins and evolution of CRISPR-Cas systems // Philosophical Transactions of the Royal Society B: Biological Sciences. 2019. Vol. 374. No. 1772. P. 20180087.

3. **Krupovic M., Shmakov S., Makarova K. S., Forterre P., Koonin E. V.** Recent mobility of casposons, self-synthesizing transposons at the origin of the CRISPR-Cas immunity // Genome Biology and Evolution. 2016. Vol. 8. No. 2. Pp. 375–386.

4. **Faure G., Shmakov S. A., Yan W. X., Cheng D. R., Scott D. A., Peters J. E., Makarova K. S., Koonin E. V.** CRISPR-Cas in mobile genetic elements: Counter-defence and beyond // Nature Reviews Microbiology. 2019. Vol. 17. No. 8. Pp. 513–525.

5. **Makarova K. S., Wolf Y. I., Shmakov S. A., Liu Y., Li M., Koonin E. V.** Unprecedented diversity of unique CRISPR-Cas-related systems and Cas1 homologs in Asgard archaea // The CRISPR Journal. 2020. Vol. 3. No. 3. Pp. 156–163.

6. **Sigrist C. J., de Castro E., Cerutti L., Cuche B. A., Hulo N., Bridge A., Bougueleret L., Xenarios I.** New and continuing developments at PROSITE // Nucleic Acids Research. 2013. Vol. 41. No. D1, 1 January. Pp. D344–D347.

7. **Marchler-Bauer A., Zheng C., Chitsaz F., et al.** CDD: conserved domains and protein three-dimensional structure // Nucleic Acids Research. 2013. Vol. 41. No. D1, 1 January. Pp. D348–D352.

8. **Finn R. D., Bateman A., Clements J., et al.** Pfam: the protein families database // Nucleic Acids Research. 2014. Vol. 42. No. D1, 1 January. Pp. D222–D230.

9. **Gibson D. G., Young L., Chuang R.Y., Venter J. C., Hutchison C. A. 3rd, Smith H. O.** Enzymatic assembly of DNA molecules up to several hundred kilobases // Nature Methods. 2009. Vol. 6. No. 5. Pp. 343–345.

10. **Васильева А. А., Алюкас С. А., Селькова П. А., Арсениев А. Н., Чернова В. Е., Мушарова А. С., Климук Е. И., Ходорковский М. А., Северинов К. В.** Нуклеазы CRISPR-Cas II типа: алгоритм поиска и *in vitro* характеристика // Молекулярная биология. 2023. Т. 57. № 3. С. 546–560.

11. **Laemmli U. K.** Cleavage of structural proteins during the assembly of the head of bacteriophage T4 // Nature. 1970. Vol. 227. No. 5259. Pp. 680–685.

12. **Kreuzer K. N., Jongeneel C. V.** *Escherichia coli* phage T4 topoisomerase // Methods in Enzymology. 1983. Vol. 100. Pp. 144–160.

13. **Walker J. E., Saraste M., Runswick M. J., Gay N. J.** Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold // The EMBO (European Molecular Biology Organization) Journal. 1982. Vol. 1. No. 8. Pp. 945–951.

14. **Ramakrishnan C., Dani V. S., Ramasarma T.** A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins // Protein Engineering, Design, and Selection. 2002. Vol. 15. No. 10. Pp. 783–798.

15. **Sun P., Tropea J. E., Waugh D. S.** Enhancing the solubility of recombinant proteins in *Escherichia coli* by using hexahistidine-tagged maltose-binding protein as a fusion partner // Evans Jr., T., Xu M.Q. (Eds.) Heterologous Gene Expression in *E. coli*. Book series: Methods in Molecular Biology. Vol. 705. Totowa, New Jersey, USA: Humana Press, 2011. Pp. 259–274.

16. **Kowalczykowski S. C., Clow J., Krupp R. A.** Properties of the duplex DNA-dependent ATPase activity of *Escherichia coli* RecA protein and its role in branch migration // Proceedings of the National Academy of Sciences. 1987. Vol. 84. No. 10. Pp. 3127–3131.

17. **Yakunina M., Artamonova T., Borukhov S., Makarova K. S., Severinov K., Minakhin L.** A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage // Nucleic Acids Research. 2015. Vol. 43. No. 21. Pp. 10411–10420.

THE AUTHORS

ABRAMOVA Marina V.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
abramova.mv07@gmail.com
ORCID: 0000-0003-4310-1189

MALYKH Anastasia S.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
anastasiyaa2s3d4f5g6@gmail.com

GATIEVA Iana M.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
yana.gatieva6@gmail.com

KAZALOV Maksim A.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
maksim.kazalov@gmail.com
ORCID: 0000-0001-5640-8142

SELKOVA Polina A.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
polina.nanobio@gmail.com
ORCID: 0000-0002-9562-1998

YAKIMOV Alexander P.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
yaleks@gmail.com
ORCID: 0000-0003-0462-1157

VASILEVA Aleksandra A.
Peter the Great St. Petersburg Polytechnic University
Complex of NBICS Technologies, National Research Center "Kurchatov Institute"(Moscow)
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
daucussativus7@gmail.com
ORCID: 0000-0002-0070-4898

ARSENIEV Anatolii N.
Peter the Great St. Petersburg Polytechnic University,
Complex of NBICS Technologies, National Research Center "Kurchatov Institute"(Moscow)
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
arsenievanatoly@gmail.com
ORCID: 0000-0003-0901-4188

KHODORKOVSKII Mikhail A.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
khodorkovskii@gmail.com
ORCID: 0000-0003-0562-0156

СВЕДЕНИЯ ОБ АВТОРАХ

АБРАМОВА Марина Викторовна – *инженер-исследователь лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
abramova.mv07@gmail.com
ORCID: 0000-0003-4310-1189

МАЛЫХ Анастасия Сергеевна – *лаборант-исследователь лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
malyh.as@edu.spbstu.ru

ГАТИЕВА Яна Маратовна – *лаборант-исследователь лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
yana.gatieva6@gmail.com

КАЗАЛОВ Максим Алексеевич – *инженер-исследователь лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
maksim.kazalov@gmail.com
ORCID: 0000-0001-5640-8142

СЕЛЬКОВА Полина Анатольевна – *инженер-исследователь лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
polina.nanobio@gmail.com
ORCID: 0000-0002-9562-1998

ЯКИМОВ Александр Павлович – *кандидат физико-математических наук, научный сотрудник лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
yaleks@gmail.com
ORCID: 0000-0003-0462-1157

ВАСИЛЬЕВА Александра Андреевна – *инженер-исследователь лаборатории молекулярной микробиологии Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург; инженер комплекса НБИКС-природоподобных технологий Национального исследовательского центра «Курчатовский институт», Москва, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
daucussativus7@gmail.com
ORCID: 0000-0002-0070-4898

АРСЕНИЕВ Анатолий Николаевич – *научный сотрудник Научно-исследовательского комплекса «Нанобиотехнологии» Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург; лаборант-исследователь комплекса НБИКС-природоподобных технологий Национального исследовательского центра «Курчатовский институт», Москва, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
arsenievanatoly@gmail.com
ORCID: 0000-0003-0901-4188

ХОДОРКОВСКИЙ Михаил Алексеевич – *кандидат физико-математических наук, директор Научно-исследовательского комплекса «Нанобиотехнологии» Института биомедицинских систем и биотехнологий Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
khodorkovskii@gmail.com
ORCID: 0000-0003-0562-0156

Received 08.12.2023. Approved after reviewing 26.02.2024. Accepted 26.02.2024.
Статья поступила в редакцию 08.12.2023. Одобрена после рецензирования 26.02.2024. Принята 26.02.2024.