

Study on Purification and Crystallization of Hypothetical Protein AF1514 from *Archeoglobus fulgidus*

PAZILAT Bahti¹, LIU Zhi-Jie², ABDULLA Abbas^{1,*}

(1.College of Life Science and Technology, Xinjiang University, Urumqi 830046, China;

2.National Laboratory of Biomacromolecules, Institution of Biophysics, Chinese Academy of Sciences, Beijing 100101, China)

Abstract : *Archeoglobus fulgidus* DSM 4304 genome was transformed and overproduced in *Escherichia coli* (*E. coli* BL21) at 12 °C and the protein was purified to 95% purity by Ni²⁺-affinity column and Superdex-75 gel filtration. Crystals active and methylated proteins were obtained by the hanging-drop vapour-diffusion method at 16 °C under three conditions among the wide range of screened conditions (>400), but principally from the solution containing 0.1mol/L sodium acetate pH 5.0, 0.1mol/L sodium chloride, and 8%~14% (W/V) 2-methyl-2,4-dipentanol (MPD). The crystal is tetragonal in shape and diffraction data to 2.09Å⁰ are detected by X-ray detector.

Key words: *Archeoglobus fulgidus*; protein AF1514; crystallization; purification

古细菌 - 硫细菌蛋白 AF1514 的纯化及晶体生长研究

帕孜来提·拜合提¹, 刘志杰², 阿不都拉·阿巴斯^{1,*}

(1.新疆大学生命科学与技术学院, 新疆 乌鲁木齐 830046;

2.中国科学院生物物理研究所结构与分子生物学中心, 北京 100101)

摘 要: 本研究将 AF1514 蛋白的重组基因转化到大肠杆菌 (*E. coli* BL21) 并通过在 12℃ 条件下的诱导表达产生了大量的目的蛋白。用镍柱亲和层析和分子筛 Superdex-75 凝胶过滤层析的两步蛋白纯化方法进一步纯化蛋白 AF1514 后, 获得纯度较高的蛋白 (纯度 > 95%)。纯化的母体蛋白和甲基化处理后的蛋白衍生物分别用悬滴汽相扩散法在 16℃ 结晶。初步筛选所用的 400 多种结晶条件中, 适合晶体生长的条件只有 3 种, 其中最佳晶体生长条件溶液的成分包含 0.1mol/L 醋酸钠, pH 5.0, 0.1mol/L 氯化钠和 8%~14% (W/V) 2-甲基 2,4-戊二醇。蛋白晶体属于四方晶系, 初步 X-射线衍射分辨率为 2.09Å⁰。

关键词: 古细菌硫细菌; 蛋白 AF1514; 晶体生长; 纯化

中图分类号: TS201

文献标识码: A

文章编号: 1002-6630(2008)09-0338-06

Recent developments in automated techniques for DNA sequencing have led to an explosion over two-dozen micro organisms and two eukaryotes are available now, and soon the genomes information on the sequences of the genomes of several organisms. Complete genomic sequences of several dozen additional organisms will be completed [1-2]. But a large part of the gene products, proteins are repre-

sented by hypothetical proteins (HP). HPs, the predicted proteins, are predicted from nucleic acid sequences only and their function is still unknown, in other words these HPs have not shown to exist by chemical evidence of protein. These hypothetical proteins is of utmost importance to complete genomic and proteomic information. Structural studies of new HPs not only offers presentation of new structures but

收稿日期: 2007-11-30

作者简介: 帕孜来提·拜合提(1969-), 女, 副教授, 在职博士研究生, 主要从事资源生物学及细胞学研究。

E-mail: pbahti@yahoo.com

* 通讯作者: 阿不都拉·阿巴斯(1951-), 男, 教授, 主要从事资源植物学研究。E-mail: abdulla@xju.edu.cn

also new functions^[3-5].

Archaeoglobus fulgidus is the first sulphate-reducing microorganism, whose genome sequence has been determined, and it is a strictly anaerobic, hyperthermophilic, sulfate-reducing, marine archaeon. The organism thrives in an extreme and dynamic environment located at the interface between superheated anaerobic vent fluids and frigid aerobic seawater^[6-7]. *A. fulgidus* has been studied widely since its discovery ten years ago. Its completed genome sequence of the organism provides a wealth of new information about how this unusual organism exploits its environment. Its genome of 2178400 base pairs contains 2436 open reading frames (ORFs). Among the 2436 open reading frames (ORFs), over half of the *A. fulgidus* ORFs (1290) have no assigned biological role. Of these, 639 have no database match. The remaining 651 ORFs, designated 'onserved hypothetical proteins', encodes functionally uncharacterized yet conserved proteins^[8]. Structural study of these hypothetical proteins will probably add to our understanding of the genetic repertoire of the *Archaea*. Solution of more and more archaeal protein's structure will provide new information for the understanding of the function of these unknown genes, as well as better understanding of prokaryotic diversity^[9].

Open-reading frame ORF-1514 in the genome of the hyperthermophilic archaeon *Archaeoglobus fulgidus* DSM 4304 encodes a 91-residue protein (GenBank accession No. NP_070343) of unknown function with a molecular weight of 10.5 D. Sequence analysis (protein sequence MEIMDEIKVNLQKEVSLEEAERYAKNIASKYGDGILLSVHDSKTGYRAPEVYCCGEKPWEVYACNRGANLKISVNQFEFYFRIEVEGQAKY) by Wu Blast showed that AF 1514 protein has no alignment in the PDB (E value>1). So we have chosen it for structural studies aimed at characterizing its structure and identifying its possible function.

1 Material and Method

1.1 Protein production

The hypothetical protein AF1514 from *Archeoglobus fulgidus* was expressed in *Escherichia coli* strain BL21-DE3 (Invitrogen) with TEV (tobacco etch virus protease) cleaving N-terminal hexa-His-tag and MBP (maltose binding protein). The glycerol stocks of clone was inoculated to LB

medium containing 100 mg/ml ampicillin and grown at 37 °C temperature while shaking at 200 r/min for 2~4 h, when OD reached 0.6~0.8, the culture was cooled to 4 °C and then supplemented with 0.2 mmol/L isopropyl-D-thiogalactopyranoside (IPTG) for induction. After some time of fermentation (20h at 16 °C, 40 h at 12 °C), the cells were harvested by centrifugation at 4000 r/min for 30 min and suspended in 40 ml PBS (pH 7.0) sonication buffer. The cells were disrupted by three cycles of sonication each consisting of a 10 s pulse followed by 10 s stop. The cell debris was separated by centrifugation at 16 000 × g for 30 min and the soluble fraction containing target protein was collected for further purification.

1.2 Ni-affinity chromatography

Supernatant after centrifugation was applied onto a Ni²⁺-affinity column (Ni-NTA agarose, Qiagen) equilibrated with PBS buffer (pH 7.0). After 30~60 min Ni-binding at 4 °C, the column was washed with washing buffer (PBS containing 10 mmol/L imidazole), finally eluted with 10 ml 500 mmol/L imidazole (in PBS buffer pH 7.0) at room temperature (about 20 °C). The eluted protein in imidazole was treated to buffer exchange (PBS) by using concentrator Amicon Ultra-45000 MWCO (Millipore) and then digested with TEV at 30 °C for 3 h or at 4 °C overnight. After digestion, the supernatant passed through Ni²⁺-affinity column (Ni-NTA agarose, Qiagen) once again to remove N-terminal His-tag and MBP. The flow through containing target protein was concentrated to 1 ml by ultra-filtration (Amicon Ultra-45 kD Millipore) and further purified by gel filtration. 20~30 µl of sample was taken after each purification step and detected by 12%~15% SDS-AGE (load 3~5 µl sample).

1.3 Gel filtration chromatography

Concentrated protein sample was applied onto Superdex 75 10/30 gel-filtration column (Amersham Biosciences) equilibrated with crystallization buffer containing 200 mmol/L Tris, 100 mmol/L NaCl, and 1 mmol/L DTT for the further purification. The identity and purity of the prepared sample was judged by 15% SDS-AGE stained with Coomassie Brilliant Blue.

1.4 Chemical modification: methylation

The purified protein was chemically modified using a reductive methylation protocol as described previously^[10]. 20 µl 1 mol/L dimethylamine-borane complex (DMAB) and

40 $\mu\text{mol/L}$ formaldehyde were added to 1 ml 10 $\mu\text{g/ml}$ protein solution. The reaction mixture was incubated in the dark while shaking at 220 r/min. The chemical additions were repeated twice at 2 h intervals. Finally, 10 μl 1mol/L dimethylamine-borane complex (DMAB) was added and the reaction mixture was incubated overnight at $-4\text{ }^{\circ}\text{C}$. Excess chemicals were removed by size-exclusion chromatography using 200 mmol/L Tris, 100 mmol/L NaCl, and 1 mmol/L DTT buffer solution. The methylated protein was concentrated to 15~20 mg/ml and used for crystallization similar to the nonmethylated (native) protein.

1.5 Preliminary crystallization study

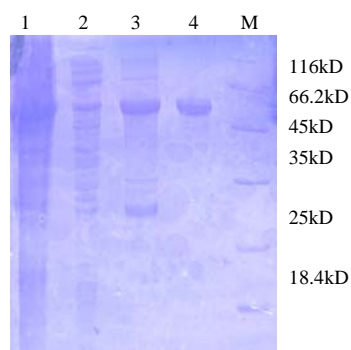
Protein obtained by gel filtration was crystallized by the hanging-drop vapour-diffusion method at $16\text{ }^{\circ}\text{C}$. 1 μl droplet of protein solution mixed with the same amount of reservoir solution was equilibrated against 300 ml reservoir solution.

The initial crystallization conditions were examined using commercially available screening kits from Hampton Research (Crystal Screen 1, Crystal Screen 2, Index and PEG/Ion Screen) and Emerald Biosystems (Wizard I and II).

Because of the importance of good crystal shape and better diffraction quality in protein crystallographic study, the crystallization conditions need further optimizing after preliminary crystallization study.

Good crystals appearing under different conditions were harvested in cryoloops and soaked for 1 min in paraffine solution. The crystal mounting in cryoloop was ash-cooled in nitrogen-gas stream at $-196\text{ }^{\circ}\text{C}$. Diffraction patterns of the crystals were detected using a X-ray RAXIS4 detector with Cr resource at the RIGAKU demo X-ray laboratory, Institution of Biophysics, CAS.

2 Results and Analysis



M.marke; 1.supernatant after centrifugation; 2. flow through of Ni-affinity; 3.washing buffe; 4. elution.

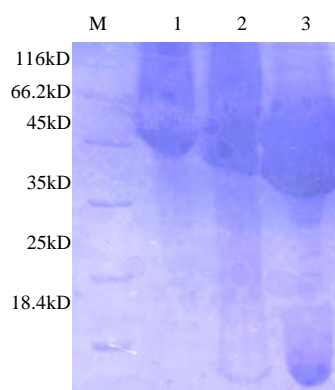
Fig.1 SDS-PAGE (12%) of protein after each purification steps (first Ni-affinity)

2.1 Protein production

Target protein with molecular weight of 54 (contained N-terminal hexa-his-tag & MBP) can be detected on SDS-page (Fig.1). Expression of the protein is very high. After centrifugation, most of the target protein exists in supernatant, and it is clear that the protein AF1514 is soluble. Results of the study also showed that induced protein production is much higher at lower temperature, and for longer fermentation time ($12\text{ }^{\circ}\text{C}$ for 40 h). The protein purification step was conducted at room temperature ($20\sim 25\text{ }^{\circ}\text{C}$) and AF1514 seems to be not very sensitive to temperature.

2.2 Ni-affinity chromatography

The almost pure target protein can be obtained after purification with Ni-binding affinity column. Although some of the target protein can be washed out with 10 mmol/L imidazole washing buffer, still desirable amount of the target protein is bind well to the column and can be eluted with 500

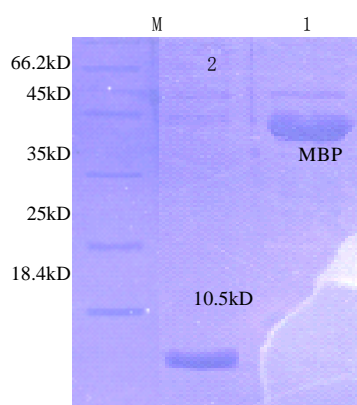


M. marker; 1.Before TEV; 2.After TEV treatment ($30\text{ }^{\circ}\text{C}$ 3 h); 3.After TEV treatment (overnight).

Fig.2 SDS-PAGE (15%) of TEV treatment

mmol/L imidazole (Fig.2).

For high throughput cloning and expression of target proteins, the cloned target protein included a polyhistidine affinity purification sequence (His-tag), the maltose binding protein (MBP), which improves the solubility of expressed proteins and a recognition sequence for the highly specific tobacco etch virus protease (TEV-site). Therefore TEV treatment is necessary for the further purification. The cleavage of MBP and hexa-His tag from target protein with TEV protease is better at $4\text{ }^{\circ}\text{C}$ overnight than at $30\text{ }^{\circ}\text{C}$ for 3 h (Fig.2). After TEV treatment, protein AF1514 with MW of 10.5 kD can be successfully separated from His-tag and MBP by



M.marker; 1.flow through after TEV treatment; 2.elution TEV treatment.

Fig.3 SDS-PAGE (15%) of after TEV treatment
(second Ni-affinity)

second Ni-affinity chromatography (Fig.3).

2.3 Gel filtration chromatography

Gel filtration is the simplest and mildest among all the chromatography methods used for protein purification. Target protein was purified by gel filtration using buffer containing 200 mmol/L Tris, 100 mmol/L NaCl and 1 mmol/L DTT at pH 7.0.

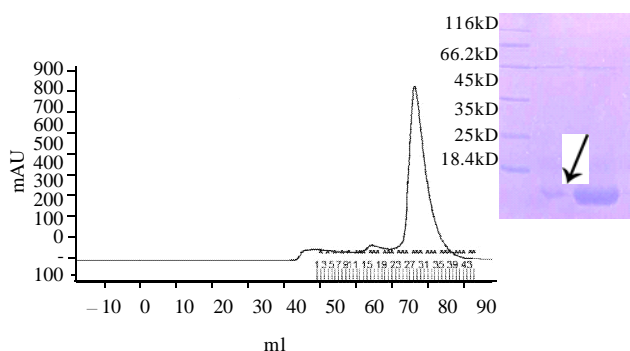


Fig.4 SDS-PAGE (15%) of protein after gel filtration

A single band corresponding to the hypothetical protein AF1514 (91 residues and molecular weight 10.5 kD) was observed after size exclusion (Fig.4). Because suitable protein concentration is important in crystal growth, the concentration of the final pure protein (45.8 mg/ml) was detected by spectrophotometer at 280 nm and diluted to 15~20 mg/ml with same buffer before setting up the crystals.

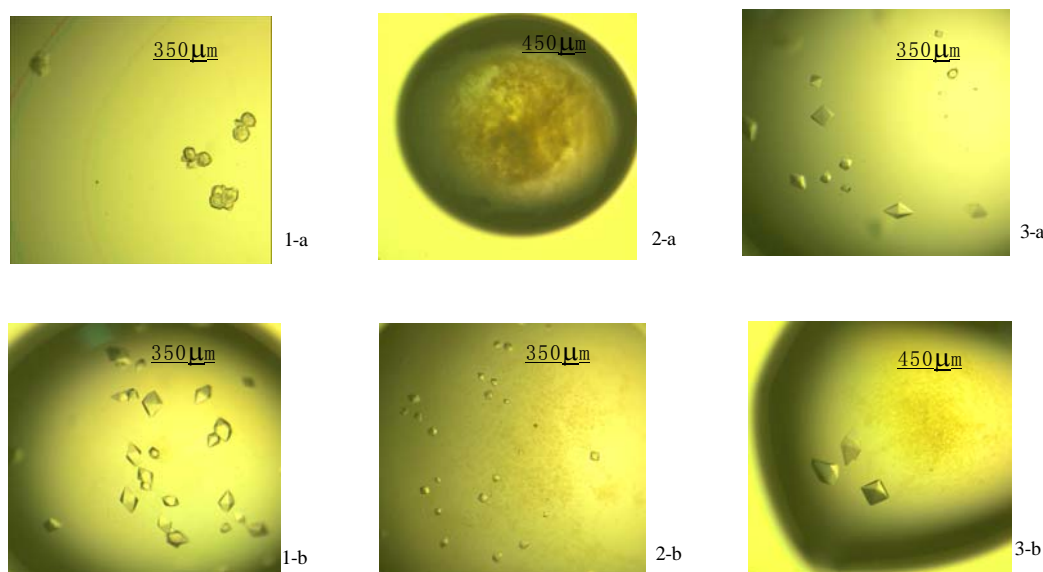
2.4 Crystallization

Crystals grew under some of the conditions of total 432 different conditions used for crystal screen (Table 1). For native protein, crystals appear within 2~3 d of incubation under several conditions such as Crystal Screen 1-condition 24, PEG/Ion-condition 48 and MembFac condition No.1. But the crystal growth of the native (nonmethylated) protein and methylated protein are different under these screened conditions. Crystal shape is improved in methylated protein under Crystal Screen 1-condition 24 and PEG/Ion-condition 48 in which native crystal growth isn't good in shape or not grows at all. The same shape of crystals (tetragonal) grows under MembFac condition 1 for both nonmethylated native protein and methylated protein (Fig.5).

One of the most promising crystallization condition, MembFac condition 1 was further optimized in order to obtain good diffraction patterns of the protein crystals. The condition was further refined by changing the pH (4.0, 4.5 and 5.0) of the buffer and the MPD concentration (10% ~ 20%). Finally the diffraction patterns of the crystals were tested in a couple of days after its growth using X-ray detector (RAXIS4, Rigaku) and the crystals which have better diffractions (resolution 2.09 Å) appear only under optimized conditions of MembFac condition No.1 (0.1 mol/L sodium acetate pH 5.0, 0.1 mol/L sodium chloride, 8% ~ 14% (W/V) 2-methyl-2,4-pentanediol (MPD)). (Data are not shown).

Table 1 Results of preliminary crystallization condition

Screening kit	Total condition	Result	Crystal growth condition	Chemical composition of crystal formed condition
Crystal screen 1	1~48	Yes	Condition No.24	Salt: 0.2 mol/L calcium chloride dihydrate Buffer: 1 mol/L sodium acetate trihydrate Precipitant: 20% (V/V) 2-Propanol
Crystal screen 2	1~48	No	—	—
Wizard I	1~48	No	—	—
Wizard II	1~48	No	—	—
PEG/ION	1~48	Yes	Condition No.48	Buffer: 0.2 mol/L Di-ammonium hydrogen citrate Precipitant: 20% (W/V) Polyethylene glycol 3350 pH 5.0
Index	1~96	No	—	—
Mem-fac	1~96	Yes	Condition No.1	Buffer: 0.1 mol/L sodium chloride Salt: 0.1 mol/L sodium acetate trihydrate pH 4.6 Precipitant: 12% (V/V) (+/-)-2-Methyl-2,4-pentanediol



1. Crystal Screen 1-condition 24; 2. PEG/Ion-condition 48; 3. MembFac condition No. 1; a. native protein crystals; b. methylated protein crystals.

Fig.5 Crystals of hypothetical protein AF1514 appearing under different screened conditions

3 Discussion

All organisms contain thousands of different proteins, which play essential roles in maintaining their life. A protein's structure determines the specific role that it plays in the organism. However, researchers lack detailed knowledge about the structures of many proteins. Crystallography allows scientists, through the study of protein crystals, to determine the three-dimensional molecular structures of proteins. For this, scientists must first crystallize the protein and analyze the resulting crystals by X-ray diffraction. The structures of many important proteins remain a mystery simply because it is difficult to obtain crystals of high enough quality or large enough size and the protein molecules arranging in an orderly and repeating pattern.

Without perfect crystals of protein (or any other biological samples) it is impossible to carry out any crystallographic structural studies^[11-12].

In order to understand the functions of the protein AF1514 based on its structures, we carried on preliminary crystallization study to the protein AF1514 from *Archeoglobus fulgidus* DSM 4304 and finally got good crystals suitable for structural study by X-ray. The open reading frame of AF1514 consists of 273 bp coding for 91 amino-acid residues. The isoelectric point of the target protein is calculated to be 5.17. Protein AF1514 can be purified in two steps to 95% purity by Ni²⁺-affinity column and

Superdex-75 gel filtration. The results of purification procedure showed that the protein seems not very sensitive to the temperature.

A mother-liquor solution containing 0.1 mol/L sodium acetate pH 5.0, 0.1 mol/L sodium chloride and 8%~14% (W/V) MPD produces morphologically identical crystals of both methylated and nonmethylated proteins in 3~4 d. Methylation of surface lysines improves the crystallizability of proteins. The reductive-methylation protocol used in the current study usually results in the dimethylation of free amine groups. The methyl groups attached to the side-chain amine N atoms of lysine residues form cohesive (NZ)CHO contacts with the neighbouring electron-negative carboxyl and carbonyl oatoms. This localizes the side chain of lysine residues in space, resulting in a more compact protein molecule^[10].

The growth and quality of the crystals are good only in fresh newly purified protein. The crystal structure is tetragonal in shape and the crystals show no visual defects when observed under a microscope. Crystals show no significant decay upon exposure when tested by X-ray and diffracted to 2.09 Å resolution. It is clear that the crystals of the hypothetical protein AF1514 is suitable for X-ray diffraction studies. Diffraction data collections and phasing are under progress. Determination of the three-dimensional structure of the protein AF1514 by X-ray crystallography will add new understandings to our knowledge about AF1514 protein's structures and functions.

References:

- [1] AXEL B, UY E, NIKOS K. A monitor of genome projects world-wide [J]. Nucl Acids Res, 2001, 29: 126-127.
- [2] MINION F C, LEFKOWITZ E J, MADSEN M L, et al. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis[J]. Bacteriol, 2004, 186: 7123-7133.
- [3] GERT L, LEILA A S, YANG J W, et al. Searching for hypothetical proteins: theory and practice based upon original data and literature[J]. Progress in Neurobiology, 2005, 77: 90-127.
- [4] GALPERIN M Y, KOONIN E V. Conserved hypothetical proteins: prioritization of targets for experimental study[J]. Nucleic Acids Res, 2004, 32: 5452-5463.
- [5] GALPERIN M Y. Conserved “ypothetical” proteins: new hints and new puzzles[J]. Comp Funct Genomics, 2001(2) : 14-18.
- [6] STETTER K O, LAUERER G, THOMM M, et al. Isolation of extremely thermophilic sulfate reducers: evidence for a novel branch of archaeobacteria[J]. Science, 1987, 236: 822-824.
- [7] STETTER K O. Microbial life in hyperthermal environments[J]. ASM News, 1995, 61: 285-290.
- [8] EISENSTEIN E, GILLILAND G L, HERZBERG O, et al. Biological function made crystal clear-annotation of hypothetical proteins via structural genomics[J]. Current Opinion in Biotechnology, 2000(11) : 25-30.
- [9] KLENK H P, CLAYTON R A, TOMB J F, et al. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*[J]. Nature, 1997(3) : 64-70.
- [10] SHAW N, CHENG C, TEMPEL W, et al. (NZ)CHO Contacts assist crystallization of a ParB-like nuclease[J]. BMC Structural Biology, 2007(7) : 46.
- [11] MARC O R, MLADEN V, ANDREW S, et al. High throughput protein crystallography: developments in crystallisation, data collection and data processing[J]. Drug Discovery Today: Technologies, 2006(3) : 451-456.
- [12] ERIC J, DUILIO C, SAWAYA M, et al. Crystal structures of a tetrahedral open pore ferritin from the hyperthermophilic archaeon *archaeoglobus fulgidus*[J]. Structure, 2005,13: 637-648.