Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Protein preparation, crystallization and preliminary X-ray analysis of human adrenal gland protein AD-004

Hui Ren,[a] Yuhe Liang,[b] Rui Li,[b] Haitao Ding,[a] Shihong Qiu,[c] Shanyun Lu,[a] Jianli An,[a] Lanfen Li,[a] Ming Luo,[a,c] Xiaofeng Zheng[a]* and Xiao-Dong Su[a,b]

[a]Department of Biochemistry and Molecular Biology, College of Life Sciences, Peking University, Beijing 100871, People's Republic of China, [b]National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, People's Republic of China, and [c]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA

Correspondence e-mail: xiaofengz@pku.edu.cn

Received 12 February 2004
Accepted 29 April 2004

The adrenal gland protein AD-004 was identified in the human adrenal gland. Full-length AD-004 contains 172 amino acids, with a predicted molecular weight of about 20 kDa. In attempts to crystallize human AD-004, the gene was subcloned into a modified pET vector, pET21-DEST, with an N-terminal His_5 tag using the Gateway cloning system, followed by protein expression in *Escherichia coli* strain BL21(DE3). The protein was purified in two steps to near-homogeneity and was crystallized. The crystals belong to space group $P6_1$ or $P6_5$, with unit-cell parameters $a = b = 99.56$, $c = 57.19$ Å. A complete 2.0 Å data set has been collected at a rotating-anode X-ray source and structure determination is under way.

1. Introduction

We have set up a high-throughput (HTP) platform for structural and functional genomics (SFG) studies at Peking University, China (Ding *et al.*, 2002). This HTP platform includes target selection, gene cloning by TOPO or/and Gateway technology (Invitrogene, USA), protein expression, purification and crystallization by conventional and HTP methods, 'direct crystallography' for rapid structure determination by in-house sulfur single-wavelength anomalous diffraction (SAD; Wang, 1985) and molecular-replacement (MR) methods *etc.* Using this platform, we can select hundreds of genes per year for SFG studies and solve the structures of proteins that crystallize and yield sufficient diffraction data. Adrenal gland protein AD-004 has been crystallized using this approach.

AD-004 is a functionally unknown protein containing 172 amino acids with a predicted molecular weight of 20 kDa; it was first cloned in a gene-expression profiling study of the human hypothalamus–pituitary–adrenal axis (Hu *et al.*, 2000) and was also identified from a comparative study of *Caenorhabditis elegans* and human genomes (Lai *et al.*, 2000). Sequence analysis showed that there is no obvious overall homology (identity greater than 25%) between AD-004 and other proteins of known function. However, the first 50 amino acids of AD-004 show significant homology to the N-terminal region of nucleotide kinases. It contains a Walker motif (Walker *et al.*, 1982) with conserved sequence Gly-*X*-*X*-Gly-*X*-Gly-Lys encoding the phosphate-binding loop. Thus, the objective of this study was to crystallize the protein, to determine its structure and to try to identify the function of the protein from its structure. We selected AD-004 protein from our pilot SFG project; it was referred to as SG14 in a previous paper (Ding *et al.*, 2002). The AD-004 gene was first cloned into a Gateway expression vector pET11a-DEST without His tag (Ding *et al.*, 2002), but the soluble fraction of the expressed protein was insufficient for conventional purification methods. This report describes how the AD-004 gene was subcloned into another expression vector and how the expression and purification conditions were optimized in order to obtain a higher yield of soluble protein; it also contains the results of AD-004 crystallization and preliminary X-ray analysis.

2. Materials and methods

2.1. Gene cloning and expression

The initial cDNA of the AD-004 gene was purchased from the ATCC MGC collection (Strausberg *et al.*, 2002). After amplification by polymerase chain reaction (PCR), the PCR product was purified using the Geneclean III Kit (Q-Biogene, USA), followed by examination on 1.5% agarose gel.

The entry clone was created by inserting the purified DNA fragment into a donor vector pDONR201 through a BP recombination reaction (Ding *et al.*, 2002). 4 µl of BP reaction mixture was transformed into library-efficiency DH5α competent cells and plated on LB plates containing 50 µg ml^{-1} kanamycin. The positive clones were confirmed by colony PCR. The target gene was transferred into an in-house constructed destination vector pET21-DEST *via* an LR recombination reaction. The LR product was transformed into DH5α competent cells in a similar way as described above for the entry clone and the transformants were selected on LB plates containing 100 µg ml^{-1}

ampicillin. The positive clones were again confirmed by colony PCR and transformed into *Escherichia coli* expression strain BL21(DE3) or BL21(DE3)pLysS for protein expression.

2.2. Protein expression and purification

The protein was expressed in *E. coli* strain BL21(DE3) and BL21(DE3)pLysS as follows. The soluble expression level of the protein was investigated first. The identified expression clone was inoculated into 20 ml LB liquid medium containing 100 µg ml^{-1} ampicillin and cultured at 310 K overnight. 4 ml of the overnight bacterial culture was added to 200 ml fresh LB medium containing 100 µg ml^{-1} ampicillin and cultured at 310 K until the OD_{600} value reached 0.6. Isopropyl β-thiogalactopyranoside (IPTG; Calbiochem, USA) was added to a final concentration of 0.5 m*M* and continuously cultured at 310 K for a further 4 h. Cells were collected by centrifugation at 5000 rev min^{-1} (CR22F Centrifuge, Hitachi, Japan) for 15 min. The pellet was resuspended in 15 ml TN buffer (20 m*M* Tris–HCl, 200 m*M* NaCl pH 7.5) and sonicated using a JY92-II sonic dismembrator (Ningbo, People's Republic of China; on for 10 s, off for 15 s; 30 cycles). A 1 ml aliquot of the lysate was taken and centrifuged at 15 000 rev min^{-1} for 30 min. The supernatant and pellet were analyzed by SDS–PAGE. The control sample was produced under the same conditions but without addition of IPTG. Different culture temperatures (310, 303 and 291 K) were examined for the optimal expression of the soluble protein.

For large-scale protein purification, 2 l of culture was prepared and centrifuged at 5000 rev min^{-1} for 10 min. Harvested cells were resuspended in TN buffer (pH 7.5) and disrupted by sonication as described above. The lysate was clarified by centrifuging at 18 000 rev min^{-1} for 30 min at 277 K and the supernatant was applied onto an Ni-affinity column equilibrated with TN buffer. The column was washed with TN buffer containing 50 m*M* imidazole and His_5-tagged AD-004 was eluted using TN buffer containing a linear gradient of 50–500 m*M* imidazole. The fractions were checked by SDS–PAGE and those containing AD-004 were pooled and concentrated in a Centricon PL-20 (Millipore Corporation, USA). To further improve the purity of the protein, the concentrated protein was applied onto a gel-filtration Hi-Load Superdex-75 column (Amersham Pharmacia Biosciences, Sweden) equilibrated with TN buffer. The peak fractions from the elution were pooled and concentrated in a Centricon PL-20. The concentration of the protein was measured with a Bio-Rad Protein Assay kit (Bio-Rad Pacific Ltd, USA).

2.3. Crystallization and data collection

The purified protein was desalted and concentrated to 10 mg ml^{-1} in 20 m*M* Tris–HCl buffer pH 7.5 for crystallization.

All crystallization experiments were performed using the hanging-drop vapour-diffusion method using Hampton Research Crystal Screens I and II (Hampton Research, USA) at 293 K in 16-well VDX plates. 2 µl protein solution and 2 µl reservoir solution were mixed and equilibrated with 500 µl reservoir solution in each well.

AD-004 crystals were flash-frozen at 100 K in a stream of boiled-off nitrogen. Diffraction data were collected on a Bruker Smart 6000 CCD detector mounted on a Bruker–Nonius FR591 rotating-anode generator with Cu $K\alpha$ radiation (45 kV, 120 mA). Data were processed using the Bruker online *PROTEUM* software suite.

3. Results and discussion

After the BP reaction, positive entry clones were obtained successfully and identified by colony PCR. The target gene was transferred into a pET11a-DEST destination vector without His tag in our initial effort, but the soluble expression level of the protein was low, making it difficult to use conventional purification steps (Ding *et al.*, 2002). In order to achieve a higher yield of soluble AD-004 protein, various approaches were used to attempt to increase the soluble portion of the protein. A new destination vector, pET21-DEST, was created as follows. The pET21a vector (Novagen, USA) was converted into a Gateway destination vector

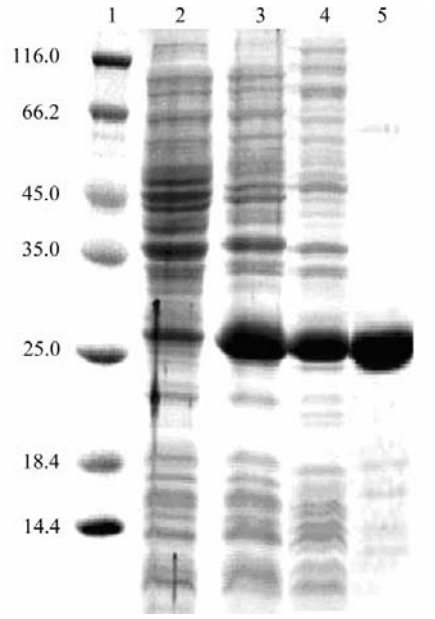

Figure 1
SDS–PAGE analysis of adrenal gland protein AD-004. Lane 1, protein molecular-weight markers (kDa); lane 2, total bacterial proteins as control without IPTG induction; lane 3, total bacterial proteins with IPTG induction; lane 4, the supernatant of *E. coli* lysate containing protein AD-004; lane 5, purified AD-004 after running through an Ni-chelating column.

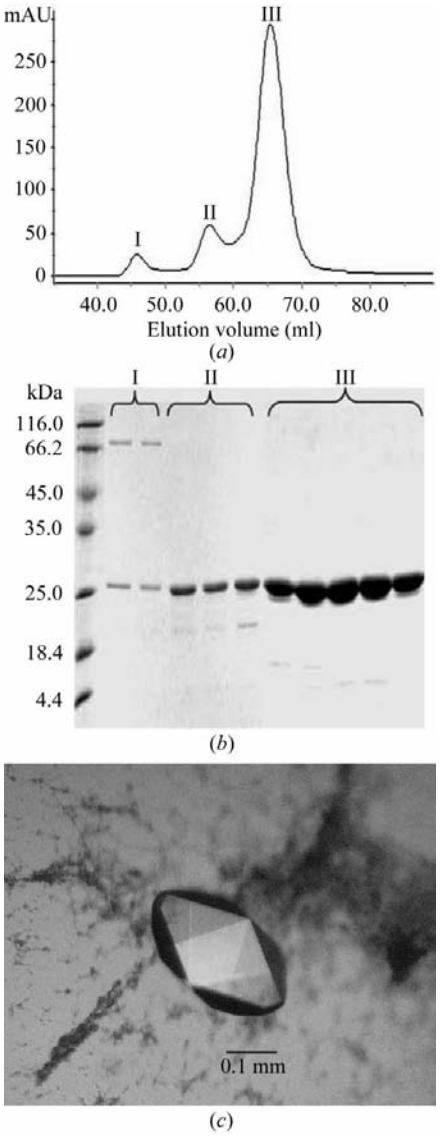

Figure 2
AD-004 protein was further purified by a Superdex 75 column. (*a*) Elution peaks of AD-004 after application onto a Superdex 75 column. (*b*) Analysis of the protein content of the three peaks eluted from the Superdex 75 column by SDS–PAGE. (*c*) A crystal of adrenal gland protein AD-004. Approximate dimensions are 0.3 × 0.3 × 0.15 mm. Crystals were obtained in a buffer consisting of 100 m*M* Na–HEPES pH 7.5 and 1.44 *M* Li_2SO_4 at 293 K.

by ligating a Gateway cassette that contains *attR* recombination sites flanking a *ccdB* gene and a chloramphenicol-resistance gene into the multiple cloning site of pET21a. The AD-004 gene was then subcloned into this newly constructed destination vector that contains a N-terminal His_5 tag with a thrombin cleavage site, which is favourable for protein purification. The expression clones were selected on LB plates containing 100 μg ml^{-1} ampicillin and the insertion of the gene was further demonstrated by colony PCR and DNA sequencing. The expression level of the recombinant protein was first examined using *E. coli* strain BL21(DE3)pLysS. The recombinant protein was overexpressed at 310 K, but most of the protein was in the form of inclusion bodies. Various expression conditions, such as different temperatures, different concentrations of IPTG and different expression strains, were examined in order to increase the soluble fraction. As shown in Figs. 1 and 2, the recombinant protein, with a calculated molecular weight of 24 kDa including the N-terminal His_5 tag, was finally expressed solubly in high yield (about 50 mg l^{-1}) under the following conditions: 303 K, 0.5 m*M* IPTG, 7 h induction in strain BL21(DE3).

It is interesting to note in Fig. 2(*a*) that three peaks were obtained after running through the S-75 gel-filtration column. The protein contents of the three peaks were examined by running SDS–PAGE (Fig. 2*b*) and the result showed that they all contain the target protein. Peak I may be contaminated by a protein with a higher molecular weight and peak II might be in a different oligomerization state.

Table 1
Data-collection statistics of adrenal gland protein AD-004.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.5418
Resolution (Å)	57.47–2.00 (2.07–2.00)
Completeness (%)	99.77 (99.33)
$R_{\rm merge}$ (%)	3.43 (19.59)
$I/\sigma(I)$	15.06 (1.86)
Space group	$P6_1$ or $P6_5$
Unit-cell parameters (Å, °)	a = 99.56, b = 99.56, c = 57.19, α = 90, β = 90, γ = 120
No. observed reflections	113635
No. independent reflections	21969
Molecules per AU	1 or 2
$V_{\rm M}$ (Å^3 Da^{-1})	4.08 or 2.04
Solvent content (%)	69 or 40

The fractions making up the three peaks were pooled separately and used for crystallization screening experiments. Only protein samples from peak III could be crystallized.

Initial screening of the crystallization conditions showed that regular polyhedral crystals appeared in condition No. 16 of Hampton Research Crystal Screen I. After refining the initial conditions in a systematic way by screening the concentration of lithium sulfate, buffer content and protein concentrations, larger and reproducible polyhedral crystals with typical dimensions of 0.3 × 0.3 × 0.15 mm (Fig. 2*c*) were obtained in a buffer containing 100 m*M* Na–HEPES pH 7.5 and 1.44 *M* Li_2SO_4 at 293 K within 10 d.

A complete data set to 2.0 Å resolution was collected. Analysis of the data shows that the crystals belong to either space group $P6_1$ or $P6_5$, with unit-cell parameters $a = b = 99.56$, $c = 57.19$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. There may be one or two molecules in the asymmetric unit, which will give Matthews coefficients (Matthews, 1968) of 4.08 or 2.04 Å^3 Da^{-1}, with 69 or 40% solvent content, respectively. Data-collection and crystallographic statistics are summarized in Table 1. Structure determination is under way and the structural details will hopefully provide useful information for elucidation of the biological functions of AD-004.

We are grateful to Professor Xiaocheng Gu for support and assistance in this work. This project was supported by the National High Technology and Development Program of China (863 program, No. 2001AA233011). Peking University's 985 grants are also highly appreciated.

References

Ding, H., Ren, H., Chen, Q., Fang, G., Li, L., Li, R., Wang, Z., Jia, X., Liang, Y., Hu, M., Li, Y., Luo, J., Gu, X., Su, X., Luo, M. & Lu, S. (2002). *Acta Cryst.* D**58**, 2102–2108.

Hu, R. M. *et al.* (2000). *Proc. Natl Acad. Sci. USA*, **97**, 9543–9548.

Lai, C.-H., Chou, C.-Y., Chang, L.-Y., Liu, C.-S. & Lin, W.-C. (2000). *Genome Res.* **10**, 703–713.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Strausberg, R. L. *et al.* (2002). *Proc. Natl Acad. Sci. USA*, **99**, 16899–16903.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). *EMBO J.* **1**, 945–951.

Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.