

Article

# Expression, Purification, Crystallization, and X-ray Structural Analysis of CRISPR-Associated Protein Cas6 from *Methanocaldococcus jannaschii*

Ming-Chang Lee <sup>1,†</sup>, Shih-Ting Tseng <sup>2,†</sup>, Juan-Cheng Yang <sup>3,4</sup>, Tung-Ju Hsieh <sup>5</sup>, Shang-Chuen Wu <sup>5,6</sup>, Shu-Min Kuan <sup>5</sup>, Ming-Jen Chen <sup>5</sup>, Ming-Chiu Chang <sup>6</sup>, Chun-Chiu Wang <sup>6</sup>, Hsiu-Lin Chen <sup>6</sup>, Guor-Cheng Fang <sup>6</sup>, Winn-Jung Huang <sup>6</sup>, Tzu-Ping Ko <sup>7,\*</sup>  and Yeh Chen <sup>5,\*</sup>

<sup>1</sup> Department of Nephrology, Da Chien General Hospital, Miaoli 36052, Taiwan; lmc4025@gmail.com

<sup>2</sup> Department of Endocrinology and Metabolism, Kuang Tien General Hospital, Taichung 43303, Taiwan; tsn4830@gmail.com

<sup>3</sup> School of Pharmacy, College of Pharmacy, China Medical University, Taichung 40402, Taiwan; qq9113054@gmail.com

<sup>4</sup> Chinese Medicine Research and Development Center, China Medical University Hospital, Taichung 40402, Taiwan

<sup>5</sup> Department of Biotechnology, Hungkuang University, Taichung 43302, Taiwan; tjh@tms.cmu.edu.tw (T.-J.H.); peterscwu@gate.sinica.edu.tw (S.-C.W.); p26862304@yahoo.com.tw (S.-M.K.); mjchen@sunrise.hk.edu.tw (M.-J.C.)

<sup>6</sup> Department of Safety, Health and Environmental Engineering, Hungkuang University, Taichung 43302, Taiwan; chang@sunrise.hk.edu.tw (M.-C.C.); chunchin@sunrise.hk.edu.tw (C.-C.W.); hsiulin@sunrise.hk.edu.tw (H.-L.C.); gcfang@sunrise.hk.edu.tw (G.-C.F.); huangwj@sunrise.hk.edu.tw (W.-J.H.)

<sup>7</sup> Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan

\* Correspondence: kotping@gate.sinica.edu.tw (T.-P.K.); chyeah6599@sunrise.hk.edu.tw (Y.C.); Tel.: +886-4-26318652-5627 (T.-P.K.); +886-2-27855696-3090 (Y.C.)

† These authors contributed equally to this work.

Academic Editor: Albert Guskov

Received: 20 September 2017; Accepted: 8 November 2017; Published: 10 November 2017

**Abstract:** The CRISPR-associated protein 6, Cas6 protein, is an endoribonuclease that cleaves precursor CRISPR RNAs within the repeat sequence to release specific invader-targeting RNAs. Cas6 protein can recognize different sequences by their specific scaffold. To investigate its binding mode, we purified and crystallized a His-tagged Cas6 protein from *Methanocaldococcus jannaschii* (MjCas6) using the sitting-drop vapor-diffusion method. The crystals diffracted to a resolution of 1.85 Å and belonged to monoclinic space group C2, with unit-cell parameters  $a = 200.84$  Å,  $b = 85.26$  Å,  $c = 100.06$  Å,  $\beta = 118.47^\circ$ . The crystals of MjCas6 contain four molecules in the asymmetric unit. The protein fold is similar to the other Cas6 homologues, such as *Pyrococcus furiosus* Cas6, suggesting functional similarity. Moreover, in the C2 crystal the MjCas6 monomers formed a tandem array, which we hypothesize to possibly correlate with repetitive RNA precursors.

**Keywords:** CRISPR; Cas proteins; endoribonuclease; RNA processing; RNAi

## 1. Introduction

Bacteria and archaea have developed defense and regulatory mechanisms to cope with various adventitious attacks, including viral infections. A novel adaptive defense system in bacteria and archaea has been discovered, which is of great interest to biologists. The CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated proteins) system contains a multistep

process to defend against viruses [1–5]. There are two major steps. First, the specific small fragments of foreign nucleic acids (virus or plasmid) are recognized as being non-self (exogenous) and are incorporated into the host genome between short DNA repeats such as CRISPRs [1,2]. Second, after the small foreign fragments are transcribed and precisely processed into small RNAs, they constitute a multifunctional protein complex with Cas proteins to recognize and cleave the incoming foreign genetic material. This adaptive immune system uses a library of small noncoding RNAs as a potent weapon against fast-evolving viruses [6].

The CRISPR-Cas loci in prokaryotic and archeal genomes comprise an array of direct, typically palindromic repeats (CRISPR) and unique spacers located between the CRISPR repeats [7]. The CRISPR arrays are usually associated with *cas* genes. These genes encode proteins with a variety of predicted nucleic acid-manipulating activities such as nucleases, helicases, and polymerases [7]. The primary products of the CRISPR loci, pre-crRNA, consist of the invader-targeting sequence and flanking repeat fragments, which are processed enzymatically into ~60-nt mature crRNAs [8]. In association with Cas proteins, crRNAs target foreign genetic elements for destruction by interacting with the phage or plasmid nucleic acid in a sequence-specific manner [8].

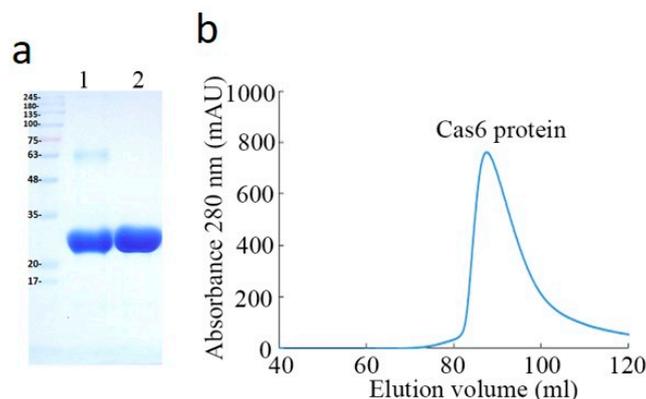
CRISPR-Cas systems could be classified to two classes and six different types (I–VI) [9]. Class 1 systems evolved with multi-subunit effector complexes, in contrast to the single-protein effector modules in Class 2 [9]. The wide variety of CRISPR-Cas systems (Types I–VI with 19 subtypes) carry out these processes by using distinct mechanisms, Cas proteins, and crRNA species [5,10,11]. In Type I and III systems, a Cas6 protein family with endonuclease activity generates the functional crRNAs by site-specific cleavage of repeat sequences in the precursor transcripts (pre-crRNA transcripts). By RNA analysis, it was found that CRISPR locus transcripts are cleaved within the repeat sequences to release ~60- to 70-nt RNA intermediates [12]. Subsequently, these ~60- to 70-nt RNA fragments are cleaved into smaller ~35- to 45-nt mature crRNAs.

Previous structural studies have shown that Cas6 proteins share a common ferredoxin or RNA recognition motif fold [13,14]. Those Cas6 proteins have divergent amino acid sequences, which are responsible for the ability to recognize different kinds of RNA substrates. To investigate the relationship of Cas6 protein from *Methanocaldococcus jannaschii* (MjCas6) with its specific RNA substrate, we expressed a His-tagged MjCas6 protein, which was then purified and crystallized to determine the three-dimensional structure by X-ray diffraction. The MjCas6 structure can be a starting point to study the evolutionary relationship between Cas6 proteins and their specific substrate recognition capabilities, as well as how the protein scaffold evolves with the substrate binding modes.

## 2. Results and Discussion

### 2.1. Purification and Characterization of MjCas6

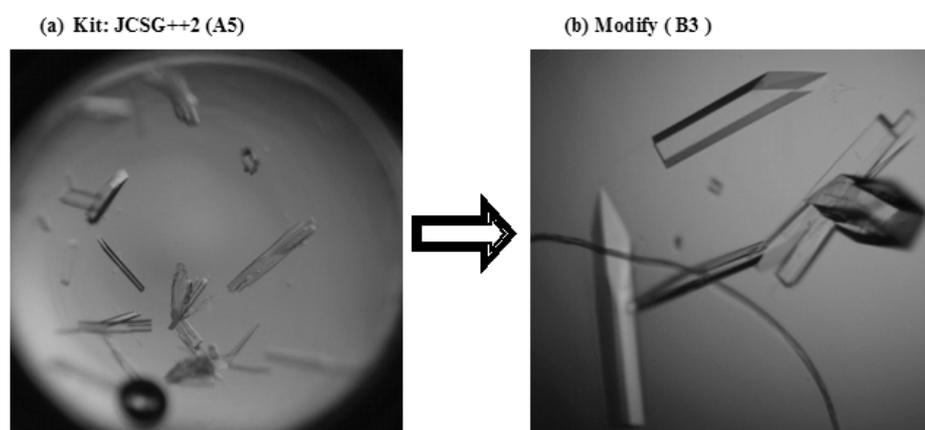
The MjCas6 protein was constructed with a His<sub>6</sub>-tag at its C-terminus to facilitate purification. The recombinant MjCas6 protein was overexpressed in *Escherichia coli* in a soluble form, then purified and crystallized for structure studies. About 20 mg purified protein per liter of *E. coli* culture was produced. The purity of protein was investigated using SDS-PAGE analysis. The result showed that high-purity MjCas6 protein was obtained and suitable for crystallization (Figure 1a). The elution profile of MjCas6 on a Superdex-200 size-exclusion column exhibited a large single peak at 86.5 mL elution volume (Figure 1b). The molecular weight of MjCas6 was about ~31.2 kDa as estimated by molecular weight standard on the gel filtration chromatogram.



**Figure 1.** (a) SDS-PAGE analysis of purified MjCas6 protein. The left lane contains molecular weight markers (labeled in kDa). Lane 1 contains the MjCas6 protein after Ni-NTA column purification; Lane 2 contains MjCas6 after SEC. (b) SEC elution profile of Cas6 protein.

## 2.2. Crystallization of the MjCas6 Protein

The purified MjCas6 protein with a His<sub>6</sub>-tag at its C-terminus was tested in various crystallization conditions, but most of them were not suitable for growing crystals or produced crystals that were not proper for X-ray data collection due to the poor diffraction quality. A reservoir solution consisting of 100 mM HEPES sodium salt pH 7.5, 800 mM sodium dihydrogen phosphate, and 800 mM potassium dihydrogen phosphate (Kit: JCSG++2) was eventually found and optimized to produce MjCas6 crystals with good diffraction quality. The best crystals were obtained using a reservoir containing 100 mM HEPES sodium salt pH 7.5, 600 mM sodium dihydrogen phosphate, and 600 mM potassium dihydrogen phosphate. The crystals employed in the X-ray diffraction experiments grew to maximum dimensions of approximately  $0.4 \times 0.3 \times 0.1 \text{ mm}^3$  within four weeks (Figure 2).

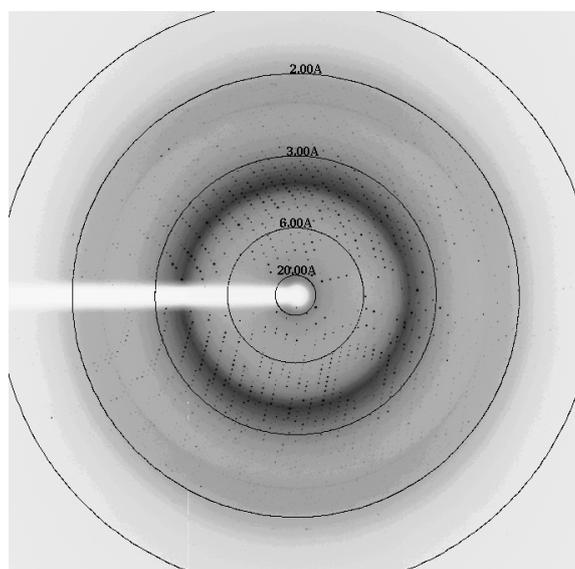


**Figure 2.** Crystals of MjCas6 protein (a) before and (b) after condition optimization. The crystal dimensions were approximately  $0.4 \times 0.3 \times 0.1 \text{ mm}^3$ .

## 2.3. X-ray Analysis of the MjCas6 Crystal

Although some diffraction patterns with smeared spots were observed in the best crystals, the diffraction data collection turned to be feasible. The diffraction data were then collected to 1.85 Å resolution where the signal-to-noise ratio was about 2 (Figure 3). The native crystals belonged to space group *C2*, with unit-cell parameters  $a = 200.84 \text{ Å}$ ,  $b = 85.26 \text{ Å}$ ,  $c = 100.06 \text{ Å}$ , and  $\beta = 118.47^\circ$ . The presence of four molecules in the asymmetric unit was assumed by calculating the Matthews coefficient ( $V_M$ ) as  $3.02 \text{ Å}^3 \cdot \text{Da}^{-1}$ . The corresponding solvent content was 59.31% [15]. Some data collection statistics are shown in Table 1. The local symmetry connecting the protein molecules in the

asymmetric unit was investigated using *MOLREP* [16] from the CCP4 package. It is compatible with the presence of four monomers in an asymmetric unit of the crystal.



**Figure 3.** X-ray diffraction pattern of MjCas6 protein crystal. The image was recorded at the NSRRC 13C1 beamline from a flash-cooled crystal in 25% glycerol.

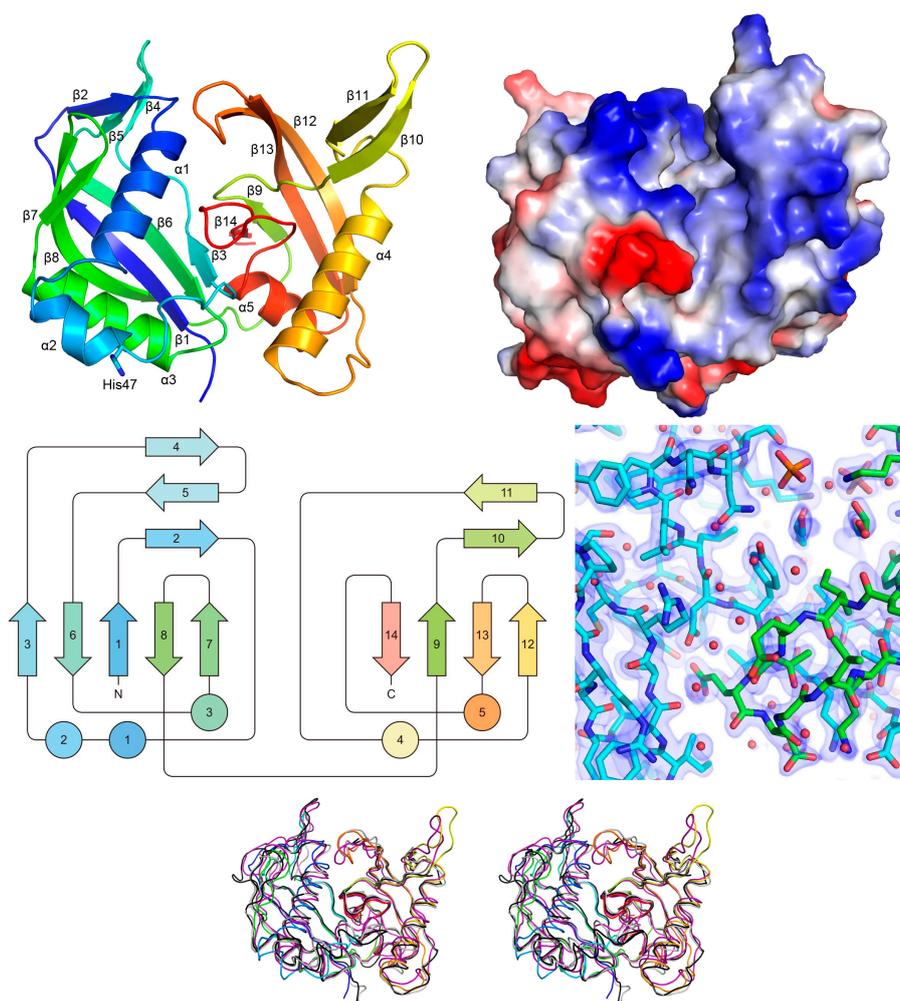
**Table 1.** Data collection and refinement statistics of Cas6 (values in parentheses are for the outermost shell).

Data Collection	
Space group	C2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	200.84, 85.26, 100.06
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 118.47, 90.00
Resolution (Å)	30.0–1.85 (1.92–1.85)
<i>R</i> <sub>meas</sub> (%)	8.4 (58.6)
<i>I</i> / $\sigma$ ( <i>I</i> )	17.3 (2.0)
Completeness (%)	99.7 (99.7)
Redundancy	3.7 (3.7)
Refinement	
Resolution (Å)	29.1–1.85 (1.92–1.85)
No. reflections	125,321 (12,110)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.167/0.194
<i>B</i> (Å <sup>2</sup> )/No. atoms	
Protein	29.8/8038
Ligand + ion	67.9/154
Water	43.3/1098
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.02

The initial model of MjCas6 protein was obtained by the molecular-replacement method using the program Phaser [17]. The coordinate of a Cas6 monomer from *Pyrococcus furiosus* (PfCas6, PDB code: 3I4H) was used as the search model. All four molecules in the asymmetric were successfully located. Refinement of the model structure employed COOT [18,19] and PHENIX [20], which yielded *R* and *R*<sub>free</sub> of 0.167 and 0.194. Local non-crystallographic symmetry (NCS) restraints were used and the resulting RMSD between the MjCas6 monomers ranged from 0.080 to 0.212 Å for 194–216 pairs of C $\alpha$  atoms.

#### 2.4. Functional Implications of the MjCas6 Structure

The ferredoxin-like fold consists of two sets of  $\beta$ - $\alpha$ - $\beta$  super-secondary structure motifs that are arranged into a four-stranded, antiparallel  $\beta$ -sheet covered on one side by two  $\alpha$ -helices. This fold is the most abundant protein domain known and is observed in a superfamily of RNA Recognition Motif (RMM) proteins consisting of several mRNA splicing factors [21]. Unlike RRM, however, MjCas6 comprises a tandem ferredoxin-like fold with two antiparallel  $\beta$  sheets on one side, which are packed against five  $\alpha$ -helices on the other side (Figure 4). On the “back” side, a dense, positively charged groove is formed along the junction of the two  $\beta$  sheets. On the “front” side, the catalytic residue His47 is found near the end of helix  $\alpha$ 2. When compared with the PfCas6 structures (PDB 3I4H, 3PKM, and 3UFC) [22–24], the RMSD is in the range of 1.33–1.78 Å for 167–198 C $\alpha$  pairs. The largest deviations are found in helices  $\alpha$ 1/ $\alpha$ 2 and strands  $\beta$ 10/ $\beta$ 11.

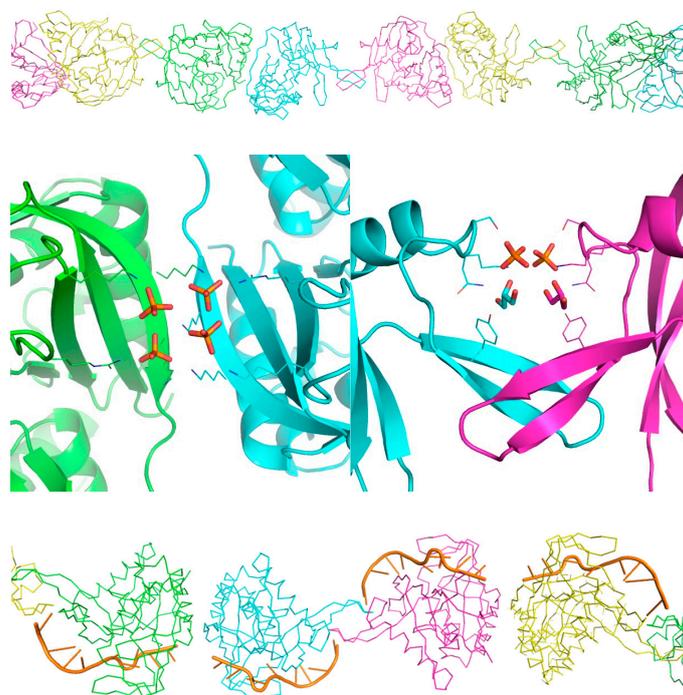


**Figure 4.** Structure of MjCas6 monomer. The protein is shown as a ribbon model on the top left, spectrum colored from N- (blue) to C-terminus (red). At the top right is a surface model colored blue for positive and red for negative electrostatic potentials. It is viewed from the “back” side of the ribbon model. In the middle is a schematic topology diagram of this protein (left) and some representative 2Fo-Fc electron densities (right, level = 1.0 $\sigma$ ). At the bottom is a stereo view of MjCas6 (rainbow) superimposed on the models from PDB 3I4H (gray), 3PKM (black), and 3UFC (purple).

Judging by its structural similarity to PfCas6, MjCas6 is likely to interact with the RNA substrate via the same positively charged binding surface and conserved catalytic residues such as Tyr32, His47, and Lys53. These are close to the G-rich loop that contains the Cas6 signature motif [25].

Presumably MjCas6, like PfCas6, does not require divalent metals for activity but cleaves RNA by the 2',3'-cyclic phosphate pathway. Because the protein was crystallized from a solution that contained significant amounts of phosphate and glycerol, both ligands were found on the protein surface. Twenty-six phosphate ions were identified in various positively charged pockets, as well as the monomer-monomer contact regions. Glycerol molecules were not observed as clearly, and only four were included in the model.

Interestingly, MjCas6 forms an array with alternating orientations in the asymmetric unit (Figure 5). Specifically, the last and outermost strand  $\beta 8$  in the N-terminal  $\beta$  sheet of one monomer is associated with its counterpart of a second monomer. Backbone hydrogen bonds are formed between the two  $\beta$  strands across the NCS dyad. On the other hand, the protruding  $\beta 10$ – $\beta 11$  hairpin is also associated with the counterpart of a third monomer. When the PfCas6-bound RNA from PDB 3PKM is placed in MjCas6, the RNA-protein interactions appear to be unobstructed by the above intermolecular contacts. Thus, the tandem array of MjCas6 observed in the crystal might represent a possible assembly for processing CRISPR RNA, which by definition contains palindromic repeats.



**Figure 5.** Association of MjCas6 monomers. At the top, an array of protein molecules are colored green, cyan, magenta, and yellow for the four independent monomers in an asymmetric unit. Two types of contacts are shown in close-up views. At the bottom, a model with bound RNA (orange, taken from PDB 3PKM) is viewed approximately 90° from the top.

The monomer-monomer interfaces bury 620 Å<sup>2</sup> (around  $\beta 8$ ) and 540 Å<sup>2</sup> ( $\beta 10$ – $\beta 11$ ) on each protein, accounting for 9% of the total surface area (12,800 Å<sup>2</sup>). The corresponding free energy changes ( $\Delta G$ ), as estimated by PISA server [26] ([www.ebi.ac.uk/pdbe/prot\\_int/pistart.html](http://www.ebi.ac.uk/pdbe/prot_int/pistart.html)), are  $-2.5$  and  $-5.0$  kcal/mol, which indicate preferred intermolecular association. In comparison, the inter-array contacts in the crystal bury less than 200 Å<sup>2</sup> on each monomer, with  $\Delta G$  ranging from  $-2.0$  to  $+2.5$  kcal/mol. Besides, each bound phosphate ion buries about 100 Å<sup>2</sup> surface areas and contributes  $\Delta G$  of  $-3$  to  $-7$  kcal/mol to the interface. The analyses suggest that the tandem array is a favored oligomer formation, and it could be enhanced by the presence of nucleic acids, which contain phosphate. Nevertheless, further experiments such as a electrophoretic mobility shift assay with a cross-linking reagent and/or nucleic acids are necessary to validate this hypothetical model.

### 3. Materials and Methods

#### 3.1. Protein Production

The Cas6 gene of 762-bp (encoding 254 amino acid residues) from *Methanocaldococcus jannaschii* DSM 2661 genomic DNA was amplified by using the polymerase chain reaction and inserted into a T7 promoter-driven expression system (modified pET21b vector, Invitrogen, Waltham, MA, USA). The resulting recombinant Cas6 protein contains a His<sub>6</sub>-tag at its C-terminus for purification (Table 2). The recombinant Cas6 protein was overexpressed in *Escherichia coli* strain BL21(DE3) cells. The transformed cells were cultured at 310 K in 1 L Luria-Bertani medium, containing 25 µg mL<sup>-1</sup> ampicillin. When the cells had reached an optical density (OD<sub>600</sub>) of 1.0, the expression of the recombinant Cas6 protein was induced with 1.0 mM isopropyl β-d-thiogalactopyranoside (IPTG) at 293 K for 20 h. The cells were then harvested by centrifugation at 1590 g for 30 min at 277 K. The cell pellet was resuspended in a lysis buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 5 mM imidazole. To remove cell debris and insoluble proteins, the lysed cells by sonication were centrifuged at 20,000× g for 90 min at 277 K. The supernatant was loaded to a Ni-NTA His-bind resin (GE Healthcare, Chicago, IL, USA) pre-equilibrated with the lysis buffer. It was subsequently washed with a washing buffer (50 mM Tris-HCl, 500 mM NaCl, and 10 mM imidazole (pH 8.0)). The Cas6 protein was eluted with a 0–200 mM linear gradient of imidazole. The fraction was concentrated by an Amicon Ultra-15 3 K Centrifugal Filter Device (3-kDa cutoff, Millipore, Billerica, MA, USA). The final purification of Cas6 protein was achieved by a HiLoad 16/60 Superdex-200 size-exclusion column (GE Healthcare) equilibrated in gel filtration buffer (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5% glycerol, and 2 mM TCEP). The purified Cas6 protein was spin concentrated using a vivaspin centrifugal filter (GE Healthcare, MW cut-off 10 kDa) to 5 mg mL<sup>-1</sup> for crystallization screening. Homogeneity of the purified protein was determined by scanning densitometry of Coomassie Blue-stained protein on a 12% (v/v) sodium dodecyl sulfate polyacrylamide gel. The protein concentration was determined by the Bradford method [27] with a standard curve of bovine serum albumin.

**Table 2.** MjCas6 protein production information.

Source Organism	<i>Methanocaldococcus jannaschii</i> DSM 2661
DNA source	ATCC: The Global Bioresource center
Forward primer †	5'-GGGAATTCATATGATGAGGGAGAGTATGAGAA-3'
Reverse primer ‡	5'-CCGCTCGAGAATTTTGTTTTGAGTTTTTTATT-3'
Expression vector	Modified pET21b
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MRESMRIELELQTDNFTVIPYNHQYYLASAIYNKIHSANPAYAKRLHNYQKFKF FTFSLQIRKRVIRKEGIETIDGKAYLYISSPNEFIENFVAGLLEDGKLRVGNVEF FVRKAKILPIPKKFNILKTIPIYLLKTIETEDGLKTYDLLPNSKFYENLNKNNLKK KYEAFYNEKCDMNFEEFVLKFRPKRMRIKNDIYCRCEMVFVKVWGDYDLIKFG YECGFGEKNSMGFGMVVNVEDKNQKNKLLKTKILEHHHHHHH

† The NdeI site is underlined; ‡ The XhoI site is underlined.

#### 3.2. Crystallization

The initial crystallization screening were performed with commercially available kits from Hampton Research (Index, PEG/Ion, PEGR, PEGRX, Crystal Screen, and Crystal Screen 2), Emerald BioSystems (Wizard I–IV, PACT 1–2), Jena Bioscience (JCSG++1, JCSG++2), and Molecular Dimensions (PGA Screen, Clear Strategy Screen I and II) by using the sitting-drop vapor-diffusion method in 24-well Cryschem plates (Hampton Research, Aliso Viejo, CA, USA). Some preliminary crystals were found in the A5 well of the kit JCSG++2. The optimized condition comprising 100 mM HEPES sodium salt (pH 7.5), 600 mM sodium dihydrogen phosphate, and 600 mM potassium dihydrogen phosphate was used to produce the best crystals. The crystals were mounted separately using a nylon loop for X-ray diffraction data collection. The crystal growth parameters are summarized in Table 3.

**Table 3.** Crystallization of MjCas6.

Method	Sitting-Drop Vapor Diffusion
Plate type	Intelli-plate 48-3
Temperature (K)	277.15
Protein concentration (mg mL <sup>-1</sup> )	14.3
Buffer composition of protein solution	50 mM Tris pH 8.0, 100 mM NaCl, 5% Glycerol, 2 mM TCEP
Composition of reservoir solution	100 mM HEPES Sodium Salt pH 7.5, 600 mM Sodium dihydrogen Phosphate, 600 mM Potassium dihydrogen Phosphate
Volume (μL) and ratio of drop	2 (1:1 ratio of protein and reservoir)
Volume of reservoir (μL)	180

### 3.3. Data Collection

Several cryoprotectants were used for data collection under cryogenic conditions, including glycerol, ethylene glycol, polyethylene glycol 400, and 2-methyl-2,4-pentane-diol. The best cryoprotectant was 25% (*v/v*) glycerol. The crystals of Cas6 protein were transferred into the reservoir solution with 25% (*v/v*) glycerol before being flash-cooled in liquid nitrogen. X-ray diffraction data were collected at 100 K on the BL13C1 beamline at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The crystal was rotated through a total angle of 360° with 1.0° oscillation per frame. The exposure time was set to 30 s. The crystal-to-detector distance was maintained at 200 mm. A total of 360 images were recorded using an ADSC Quantum 315R CCD detector. The diffraction data were indexed and integrated using the processing software HKL2000 [28]. Crystal parameters and data collection statistics are summarized in Table 1; MjCas6 PDB:5YI6.

**Acknowledgments:** We thank the National Synchrotron Radiation Research Center (NSRRC, Taiwan) for assistance during data collection. This work was supported by grants from the National Science Council (NSC99-2313-B-241-001 and NSC100-2313-B-241-006 to Yeh Chen) and by the Taiwan Protein Project (MOST105-0210-01-12-01 and MOST106-0210-01-15-04).

**Author Contributions:** Ming-Chang Lee and Shih-Ting Tseng designed experiments. Juan-Cheng Yang, Shu-Min Kuan, Ming-Chang Lee, Hsiu-Lin Chen, Guor-Cheng Fang, Chun-Chiu Wang, and Ming-Chiu Chang performed the experiments. Shang-Chuen Wu and Ming-Jen Chen analyzed the dataset. Tung-Ju Hsieh, established and refined the structure. Winn-Jung Hung, Tzu-Ping Ko, and Yeh Chen designed, supervised the project and interpreted data.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

### References

1. Terns, M.P.; Terns, R.M. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* **2011**, *14*, 321–327. [[CrossRef](#)] [[PubMed](#)]
2. Sorek, R.; Lawrence, C.M.; Wiedenheft, B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* **2013**, *82*, 237–266. [[CrossRef](#)] [[PubMed](#)]
3. Jiang, F.; Doudna, J.A. The structural biology of CRISPR-Cas systems. *Curr. Opin. Struct. Biol.* **2015**, *30*, 100–111. [[CrossRef](#)] [[PubMed](#)]
4. Jackson, R.N.; Wiedenheft, B. A conserved structural chassis for mounting versatile CRISPR RNA-guided immune responses. *Mol. Cell* **2015**, *58*, 722–728. [[CrossRef](#)] [[PubMed](#)]
5. Makarova, K.S.; Wolf, Y.I.; Alkhnbashi, O.S.; Costa, F.; Shah, S.A.; Saunders, S.J.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Haft, D.H.; et al. An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **2015**, *13*, 722–736. [[CrossRef](#)] [[PubMed](#)]
6. Bhaya, D.; Davison, M.; Barrangou, R. CRISPR-cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet.* **2011**, *45*, 273–297. [[CrossRef](#)] [[PubMed](#)]
7. Jansen, R.; van Embden, J.D.A.; Gaastra, W.; Schouls, L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **2002**, *43*, 1565–1575. [[CrossRef](#)] [[PubMed](#)]
8. Brouns, S.J.J.; Jore, M.M.; Lundgren, M.; Westra, E.R.; Slijkhuis, R.J.H.; Snijders, A.P.L.; Dickman, M.J.; Makarova, K.S.; Koonin, E.V.; van der Oost, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **2008**, *321*, 960–964. [[CrossRef](#)] [[PubMed](#)]

9. Koonin, E.V.; Makarova, K.S.; Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **2017**, *37*, 67–78. [[CrossRef](#)] [[PubMed](#)]
10. Haft, D.H.; Selengut, J.; Mongodin, E.F.; Nelson, K.E. A guild of 45 CRISPR-Associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **2005**, *1*, e60. [[CrossRef](#)] [[PubMed](#)]
11. Vestergaard, G.; Garrett, R.A.; Shah, S.A. CRISPR adaptive immune systems of Archaea. *RNA Biol.* **2014**, *11*, 156–167. [[CrossRef](#)] [[PubMed](#)]
12. Hale, C.; Kleppe, K.; Terns, R.M.; Terns, M.P. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **2008**, *14*, 2572–2579. [[CrossRef](#)] [[PubMed](#)]
13. Niewoehner, O.; Jinek, M.; Doudna, J.A. Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. *Nucleic Acids Res.* **2014**, *42*, 1341–1353. [[CrossRef](#)] [[PubMed](#)]
14. Sefcikova, J.; Roth, M.; Yu, G.; Li, H. Cas6 processes tight and relaxed repeat RNA via multiple mechanisms: A hypothesis. *BioEssays* **2017**, *39*, 1700019. [[CrossRef](#)] [[PubMed](#)]
15. Matthews, B.W. Solvent content of protein crystals. *J. Mol. Biol.* **1968**, *33*, 491–497. [[CrossRef](#)]
16. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 22–25. [[CrossRef](#)] [[PubMed](#)]
17. McCoy, A.J.; Grosse-Kunstleve, R.W.; Storoni, L.C.; Read, R.J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2005**, *61*, 458–464. [[CrossRef](#)] [[PubMed](#)]
18. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501. [[CrossRef](#)] [[PubMed](#)]
19. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]
20. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221. [[CrossRef](#)] [[PubMed](#)]
21. Kielkopf, C.L.; Lücke, S.; Green, M.R. U2AF homology motifs: Protein recognition in the RRM world. *Genes Dev.* **2004**, *18*, 1513–1526. [[CrossRef](#)] [[PubMed](#)]
22. Wang, R.; Preamplume, G.; Terns, M.P.; Terns, R.M.; Li, H. Interaction of the Cas6 ribonuclease with CRISPR RNAs: Recognition and cleavage. *Structure* **2011**, *19*, 257–264. [[CrossRef](#)] [[PubMed](#)]
23. Park, H.-M.; Shin, M.; Sun, J.; Kim, G.S.; Lee, Y.C.; Park, J.-H.; Kim, B.Y.; Kim, J.-S. Crystal structure of a Cas6 paralogous protein from *pyrococcus furiosus*. *Proteins Struct. Funct. Bioinform.* **2012**, *80*, 1895–1900. [[CrossRef](#)] [[PubMed](#)]
24. Carte, J.; Wang, R.; Li, H.; Terns, R.M.; Terns, M.P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **2008**, *22*, 3489–3496. [[CrossRef](#)] [[PubMed](#)]
25. Calvin, K.; Li, H. RNA-splicing endonuclease structure and function. *Cell. Mol. Life Sci.* **2008**, *65*, 1176–1185. [[CrossRef](#)] [[PubMed](#)]
26. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. [[CrossRef](#)] [[PubMed](#)]
27. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [[CrossRef](#)]
28. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326. [[PubMed](#)]

