Anikó Czene,^{a,‡} Eszter Tóth,^{b,‡}
Béla Gyurcsik,^{a,b,*} Harm Otten,^c
Jens-Christian N. Poulsen,^c
Leila Lo Leggio,^c Sine Larsen,^c
Hans E. M. Christensen^d and
Kyosuke Nagata^e

^aMTA-SZTE Bioinorganic Chemistry Research Group, Dóm tér 7, H-6720 Szeged, Hungary,

^bDepartment of Inorganic and Analytical Chemistry, University of Szeged, Dóm tér 7, H-6720 Szeged, Hungary, ^cDepartment of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark, ^dDepartment of Chemistry, Technical University of Denmark, Kemitorvet, Building 207, 2800 Kgs. Lyngby, Denmark, and ^eDepartment of Infection Biology, Graduate School of Comprehensive Human Sciences and Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8575, Japan

‡ These authors contributed equally to this work.

Correspondence e-mail:
gyurcsik@chem.u-szeged.hu

Received 26 February 2013
Accepted 25 March 2013

Crystallization and preliminary crystallographic analysis of an *Escherichia coli*-selected mutant of the nuclease domain of the metallonuclease colicin E7

The metallonuclease colicin E7 is a member of the HNH family of endonucleases. It serves as a bacterial toxin in *Escherichia coli*, protecting the host cell from other related bacteria and bacteriophages by degradation of their chromosomal DNA under environmental stress. Its cell-killing activity is attributed to the nonspecific nuclease domain (NColE7), which possesses the catalytic $\beta\beta\alpha$ -type metal ion-binding HNH motif at its C-terminus. Mutations affecting the positively charged amino acids at the N-terminus of NColE7 (444–576) surprisingly showed no or significantly reduced endonuclease activity [Czene *et al.* (2013), *J. Biol. Inorg. Chem.* **18**, 309–321]. The necessity of the N-terminal amino acids for the function of the C-terminal catalytic centre poses the possibility of allosteric activation within the enzyme. Precise knowledge of the intramolecular interactions of these residues that affect the catalytic activity could turn NColE7 into a novel platform for artificial nuclease design. In this study, the N-terminal deletion mutant Δ N4-NColE7-C* of the nuclease domain of colicin E7 selected by *E. coli* was overexpressed and crystallized at room temperature by the sitting-drop vapour-diffusion method. X-ray diffraction data were collected to 1.6 Å resolution and could be indexed and averaged in the trigonal space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 55.4$, $c = 73.1$ Å. Structure determination by molecular replacement is in progress.

1. Introduction

One of the main directions of gene engineering *in vivo*, which is a future challenge in gene therapy, is based on artificial nucleases that activate the DNA-repair machinery of cells by a specifically positioned double-strand DNA cleavage. Using this process, gene regulation, knocking out/knocking in or correction of mutated genes can be achieved (Gyurcsik & Czene, 2011). A requirement for the use of such nucleases in potential therapeutic applications is their exclusive specificity for the target sequence in DNA as large as the human genome.

The functional domains of native enzymes are frequently applied in chimeric nucleases. Successful experiments of this kind resulted in the zinc finger–FokI (Kim *et al.*, 1996) and meganuclease (Grizot *et al.*, 2010; Fajardo-Sanchez *et al.*, 2008) enzymes. These nucleases have already reached clinical trials with great success. However, a few years into the treatment a proportion of patients developed monoclonal lymphoproliferation, which resembles acute leukaemia (Qasim *et al.*, 2009). These artificial nucleases also showed moderate cytotoxicity (Cornu & Cathomen, 2010). Therefore, not only specificity but also control of its action is desirable for such an enzyme. We emphasize the importance of allosteric control by the DNA-binding domain in native FokI (Wah *et al.*, 1997), which was lost during the creation of the abovementioned zinc finger–FokI nucleases. Thus, a new approach is necessary to search for such regulatory elements in nucleases.

Recently, the HNH family of nucleases has rapidly expanded (Finn *et al.*, 2008; Veluchamy *et al.*, 2009). In this study, we have chosen the bacterial toxin colicin E7 and have focused on its nuclease domain (NColE7). It is a nonspecific nuclease with an active site situated at the C-terminus that comprises an HNH motif, *i.e.* a $\beta\beta\alpha$ -type metal-ion-binding stretch with well conserved His, Asn and His residues (Chak *et al.*, 1991). A number of crystal structures have been reported

1. -----**GPLGSPEF**PGKATGKGKPVNNKWLNNAGKDLGSPVPDRIANKLRDKEFKSFDDFRKKFWEEVSKDPEL
2. -----**GPLGSPEF**PGKATGKGKPVNNKWLNNAGKDLGSPVPDRIANKLRDKEFKSFDDFRKKFWEEVSKDPEL

1. SKQFSRNNNDRMKVGKAPKTRTQDVSGKRTSFELHHEKPISQNGGVYDMDNISVVTPKRHIDIHRGK-----
2. SKQFSRNNNDRMKVGKAPKTRTQDVSGKRTSFELHHEKPISQNGGVYDMDNISVVTPKRHIDIHQVNSSSGRIVTD

Figure 1

The sequences of Δ N4-NColE7 (1) and of the Δ N4-NColE7-C* mutant protein (2) as expressed and purified from the pGEX-6-P1 expression system and cleaved by PreScission protease. The residues marked in red originate from the GST-fusion system and from the accompanying open reading frame shift.

in the presence or absence of its immunity protein, Zn^{2+} and DNA (Cheng *et al.*, 2002; Sui *et al.*, 2002; Hsia *et al.*, 2004; Doudeva *et al.*, 2006; Ko *et al.*, 1999; Wang *et al.*, 2007, 2009; Huang & Yuan, 2007). From these and biochemical studies, it is also known that the Arg residue located at the N-terminus is essential for cytotoxic activity, and it has been suggested to play different roles during enzymatic action (Shi *et al.*, 2005; Czene *et al.*, 2013). Its function is not yet completely clear, but in the crystal structures in which it is observed it is situated close to the phosphate ion bound to the Zn^{2+} ion in the active centre.

These requirements for activity and spatial proximity of the two termini of NColE7 may facilitate the development of a novel controlled artificial nuclease based on allosteric activation. In such an enzyme any damage to the nuclease in the intracellular space would cancel its activity, thus preventing its nonspecific action. In order to study this possibility, we first need to investigate the interactions between the two termini. An interesting feature of the N-terminus is that it is an ~ 40 -amino-acid loop that does not contain any long secondary-structural elements. Thus, it seems obvious that its spatial arrangement is the result of electrostatic and/or hydrogen-bonding interactions between the bound phosphate ion (or DNA backbone) and the Arg residue at the N-terminus. Starting from this hypothesis, we prepared an N-terminally truncated version of NColE7 deleting the four N-terminal amino acids KRNK, which include three positively charged residues, in the mutant Δ N4-NColE7-C* (Fig. 1). The structure of this mutant should show whether or not the N-terminal loop stays in its original position even without the positively charged initial amino acids, which is essential information in view of the allosteric regulation. Here, we report the overexpression, purification, crystallization and crystallographic characterization of the Δ N4-NColE7-C* mutant of *E. coli* colicin E7.

2. Materials and methods

2.1. Protein expression and purification

The gene encoding the Δ N4-NColE7 mutant was cloned from the pQE70 plasmid (a generous gift from Professor K.-F. Chak, Institute of Biochemistry and Molecular Biology, National Yang Ming University, Taipei, Taiwan) as described previously (Czene *et al.*, 2013). However, DNA sequencing of the constructed gene showed that an erroneous plasmid was selected by the bacteria. The resulting GST- Δ N4-NColE7-C* mutant of NColE7 was expressed using the vector pGEX-6P-1 (GE Healthcare), which inserts a glutathione S-transferase (GST) affinity tag at the N-terminus of the target protein, in *E. coli* BL21 (DE3) cells. Precultured LB-Amp medium (50 ml) was inoculated with a single colony grown from the plated cells and incubated for 3 h at 310 K with shaking at 250 rev min⁻¹. From the small-scale culture, 6.5 ml aliquots were transferred into 6 $\times$ 650 ml LB-Amp medium and the bacteria were grown at 303 K. Protein expression was induced at an OD₆₀₀ of 0.6–1.0 by the addition

of 2 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) to the cultures to give a final concentration of 0.1 mM. Shaking was continued at 250 rev min⁻¹ for 3 h. The cells were sedimented by centrifugation at 277 K and 3000g for 15 min and then resuspended in 40 ml PBS buffer (0.14 M NaCl, 27 mM KCl, 100 mM Na₂HPO₄, 18 mM KH₂PO₄ pH 7.3). The pellets were disrupted by sonication and the cell debris was removed by centrifugation at 18 000g and 277 K for 20 min. Firstly, a GST-based affinity-purification step was applied to the soluble fractions. The protein solution was loaded onto a GSTPrep FF 16/10 affinity column (GE Healthcare) and purified according to the manufacturer's protocol. It was washed with ice-cold PBS buffer to remove unbound material, and the bound fusion proteins were eluted with 10 mM reduced glutathione in 20 mM Tris pH 8.0 buffer. The protein-containing fractions were collected and concentrated using Vivacell 70 (Vivaproducts). The fusion protein was cleaved with in-house expressed and purified human rhinovirus C3 protease (Walker *et al.*, 1994; sold as PreScission protease by GE Healthcare) to remove the GST tag. The protease was applied to the reaction mixture at a 7:1 molar ratio of protein:protease overnight at 277 K. The mixture was again purified using a GSTPrep FF 16/10 affinity column (GE Healthcare); the Δ N4-NColE7-C* protein appeared in the flowthrough fraction. A further purification step was performed using a Source 30S 16/12 cation-exchange column (GE Healthcare) previously equilibrated with 20 mM HEPES buffer pH 7.7; the protein was eluted using a 0–1 M NaCl gradient in the same buffer. The Δ N4-NColE7-C* protein peak was observed at 0.33 M NaCl. The protein was concentrated in 20 mM HEPES pH 7.5, 50 mM NaCl by ultrafiltration using an Ultracel PL-3 membrane (Millipore). The efficiencies of all steps were monitored by 12.5–17.5% SDS-PAGE using either low-range or broad-range protein markers (14–97 or 6.5–200 kDa; Bio-Rad). Pure Δ N4-NColE7-C* protein was obtained with a yield of 5.8 mg per litre of culture and was subjected to nano-ESI-MS, showing a single species with a mass

Figure 2

The crystal of Δ N4-NColE7-C* used for data collection. The dimensions of the crystal assembly are approximately 500 $\times$ 50 $\times$ 50 μ m.

of 16 251.9 Da (the calculated mass of the Zn^{2+} -containing protein is 16 253.5 Da), indicating that the purified protein contained a Zn^{2+} ion.

2.2. Crystallization

The search for initial crystallization hits was performed by the sitting-drop vapour-diffusion method using the JCSG+ (Qiagen), PACT (Qiagen) and Index (Hampton Research) screens. An Oryx8 liquid-handling robot (Douglas Instruments) was used to set up the screens in MRC 2-drop plates (Douglas Instruments) with a total drop volume of 0.3 μl and 3:1 and 1:1 ratios of protein solution and reservoir solution. A crystal assembly (Fig. 2) was observed in a 0.3 μl drop consisting of 28 mg ml^{-1} ColE7 mutant mixed with reservoir solution [JCSG+ screen condition No. 1; 0.1 M lithium sulfate, 50 mM sodium acetate pH 4.5, 50% (w/v) PEG 400] in a 1:1 ratio. Crystals appeared within four weeks at room temperature. Crystals were flash-cooled in liquid nitrogen before data collection.

2.3. X-ray data collection and processing

A single crystal flash-cooled in liquid nitrogen was used for data collection on the Cassiopeia beamline I911-3 (Ursby *et al.*, 2004) at MAX-lab, Lund, Sweden (Fig. 3). The crystal was cooled to 100 K by a stream of nitrogen gas during data collection. Diffraction data extending to 1.6 Å resolution were indexed, integrated and scaled using *XDS* and *XSCALE* (Kabsch, 2010). Data-collection statistics are listed in Table 1.

3. Results and discussion

$\Delta\text{N4-NColE7-C}^*$ is a mutant of the NColE7 protein that was selected by *E. coli* in its GST-fusion form. The amino-acid sequence of the $\Delta\text{N4-NColE7}$ protein compared with that of $\Delta\text{N4-NColE7-C}^*$ after PreScission protease digestion is depicted in Fig. 1 and differences

Figure 3
An X-ray diffraction image collected from the $\Delta\text{N4-NColE7-C}^*$ crystal on MAX-lab beamline I911-3 ($\Delta\phi = 1^\circ$). The edge of the detector corresponds to a resolution of 1.6 Å (represented as a circle).

Table 1

Data-collection statistics for the $\Delta\text{N4-NColE7-C}^*$ crystal.

Wavelength (Å)	1.20
Space group	$P3_121$ or $P3_221$
Unit-cell parameters	
$a = b$ (Å)	55.4
c (Å)	73.1
$\alpha = \beta$ (°)	90.0
γ (°)	120.0
Total reflections	182200
Unique reflections	32966
Resolution range (Å)	48–1.6 (1.7–1.6)
Multiplicity	5.5 (4.9)
Completeness (%)	99.5 (98.1)
R_{merge} (%)†	4.9 (77.1)
Average $I/\sigma(I)$	19.4 (2.0)
Anomalous correlation (%)	35

† $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measurement for a given reflection and $\langle I(hkl) \rangle$ is the average intensity for multiple measurements of this reflection.

between the C-termini can be recognized. A mutation such as this may occur upon erroneous annealing of the C-terminal oligonucleotide primer, which is a low-probability process, causing a shift in the open reading frame. It is most probable that minor expression of the GST- $\Delta\text{N4-NColE7}$ protein resulted in cytotoxicity that killed the transformed cells (similar to NColE7; Anthony *et al.*, 2004) and allowed only those bacteria which contained the plasmid with the gene for the $\Delta\text{N4-NColE7-C}^*$ mutant to grow. Detailed biochemical studies of $\Delta\text{N4-NColE7-C}^*$ protein both with and without a GST-fusion tag have been performed (Czene *et al.*, 2013). The results indicated that the GST- $\Delta\text{N4-NColE7-C}^*$ mutant does not show DNase activity. At the same time, the protein strongly binds Zn^{2+} ions and DNA; thus, it is not a lack of the zinc-binding or DNA-binding functions that allows expression. The GST- $\Delta\text{N4-NColE7-C}^*$ protein was expressed on a large scale and purified by GST-affinity and cation-exchange chromatography to give protein suitable for crystallization, as described in §2.

Denaturing SDS-PAGE and mass spectrometry indicated a molecular mass of 16 253.5 Da, which is consistent with the monomer mass of the $\Delta\text{N4-NColE7-C}^*$ protein with a bound Zn^{2+} ion. $\Delta\text{N4-NColE7-C}^*$ crystals (Fig. 2) were obtained by the sitting-drop vapour-diffusion method using the JCSG+ screen, as described in §2. The crystals belonged to the trigonal system, with unit-cell parameters $a = b = 55.4$, $c = 73.1$ Å. The averaged data were consistent with space group $P3_121$ or $P3_221$. The anomalous correlation (Table 1) is consistent with the presence of zinc in the structure. The volume-to-mass ratio V_M of $1.99 \text{ Å}^3 \text{ Da}^{-1}$ suggests the presence of one molecule per asymmetric unit, with a corresponding solvent content of 35%. The high homology of the native and mutant NColE7 proteins should result in similar structures, and structure determination will be pursued by the molecular-replacement method using the structure of NColE7 (PDB entry 1m08; Cheng *et al.*, 2002) as the search model.

Financial support from the Hungarian Scientific Research Fund (OTKA-NKTH CK80850), TAMOP-4.2.1/B-09/1, TAMOP-4.2.2/B-10/1-2010-0012, JSPS and DanScatt, as well as MAX-lab beamtime, is greatly acknowledged. ET and AC thank the Danish Ministry of Science, Innovation and Higher Education for the fellowships provided.

References

- Anthony, L. C., Suzuki, H. & Filutowicz, M. (2004). *J. Microbiol. Methods*, **58**, 243–250.
- Chak, K.-F., Kuo, W.-S., Lu, F.-M. & James, R. (1991). *J. Gen. Microbiol.* **137**, 91–100.

- Cheng, Y.-S., Hsia, K.-C., Doudeva, L. G., Chak, K.-F. & Yuan, H. S. (2002). *J. Mol. Biol.* **324**, 227–236.
- Cornu, T. I. & Cathomen, T. (2010). *Methods Mol. Biol.* **649**, 237–245.
- Czene, A., Németh, E., Zóka, I. G., Jakab-Simon, N. I., Körtvélyesi, T., Nagata, K., Christensen, H. E. & Gyurcsik, B. (2013). *J. Biol. Inorg. Chem.* **18**, 309–321.
- Doudeva, L. G., Huang, H., Hsia, K.-C., Shi, Z., Li, C.-L., Shen, Y., Cheng, Y.-S. & Yuan, H. S. (2006). *Protein Sci.* **15**, 269–280.
- Fajardo-Sanchez, E., Stricher, F., Pâques, F., Isalan, M. & Serrano, L. (2008). *Nucleic Acids Res.* **36**, 2163–2173.
- Finn, R. D., Tate, J., Misty, J., Coghill, P. C., Sammut, S. J., Hotz, H.-R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. L. & Bateman, A. (2008). *Nucleic Acids Res.* **36**, D281–D288.
- Grizot, S., Epinat, J. C., Thomas, S., Duclert, A., Rolland, S., Pâques, F. & Duchateau, P. (2010). *Nucleic Acids Res.* **38**, 2006–2018.
- Gyurcsik, B. & Czene, A. (2011). *Future Med. Chem.* **3**, 1935–1966.
- Hsia, K.-C., Chak, K.-F., Liang, P.-H., Cheng, Y.-S., Ku, W.-Y. & Yuan, H. S. (2004). *Structure*, **12**, 205–214.
- Huang, H. & Yuan, H. S. (2007). *J. Mol. Biol.* **368**, 812–821.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kim, Y.-G., Cha, J. & Chandrasegaran, S. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 1156–1160.
- Ko, T.-P., Liao, C.-C., Ku, W.-Y., Chak, K.-F. & Yuan, H. S. (1999). *Structure*, **7**, 91–102.
- Qasim, W., Gaspar, H. B. & Thrasher, A. J. (2009). *Gene Ther.* **16**, 1285–1291.
- Shi, Z., Chak, K.-F. & Yuan, H. S. (2005). *J. Biol. Chem.* **280**, 24663–24668.
- Sui, M.-J., Tsai, L.-C., Hsia, K.-C., Doudeva, L. G., Ku, W.-Y., Han, G. W. & Yuan, H. S. (2002). *Protein Sci.* **11**, 2947–2957.
- Ursby, T., Mammen, C. B., Cerenius, Y., Svensson, C., Sommarin, B., Fodje, M. N., Kwick, Å., Logan, D. T., Als-Nielsen, J., Thunnissen, M. M. G. M., Larsen, S. & Liljas, A. (2004). *AIP Conf. Proc.* **705**, 1241–1246.
- Veluchamy, A., Mary, S., Acharya, V., Mehta, P., Deva, T. & Krishnaswamy, S. (2009). *Bioinformatics*, **6**, 80–83.
- Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I. & Aggarwal, A. K. (1997). *Nature (London)*, **388**, 97–100.
- Walker, P. A., Leong, L. E., Ng, P. W., Tan, S. H., Waller, S., Murphy, D. & Porter, A. G. (1994). *Biotechnology*, **6**, 601–605.
- Wang, Y.-T., Wright, J. D., Doudeva, L. G., Jhang, H.-C., Lim, C. & Yuan, H. S. (2009). *J. Am. Chem. Soc.* **131**, 17345–17353.
- Wang, Y.-T., Yang, W.-J., Li, C.-L., Doudeva, L. G. & Yuan, H. S. (2007). *Nucleic Acids Res.* **35**, 584–594.