Kazuki Kawahara,[a] Shota Nakamura,[b] Yasuhiro Katsu,[a] Daisuke Motooka,[a] Yuki Hosokawa,[a] Yukiko Kojima,[a] Keiko Matsukawa,[a] Hiroto Takinowaki,[a] Susumu Uchiyama,[c] Yuji Kobayashi,[d] Kiichi Fukui[c] and Tadayasu Ohkubo[a]*

[a]Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan, [b]Research Institute for Microbial Diseases, Osaka University, Suita, Osaka 565-0871, Japan, [c]Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan, and [d]Osaka University of Pharmaceutical Sciences, 4-20-1 Nasahara, Takatsuki, Osaka 569-1094, Japan

Correspondence e-mail: ohkubo@phs.osaka-u.ac.jp

Cloning, expression, crystallization and preliminary X-ray crystallographic analysis of a human condensin SMC2 hinge domain with short coiled coils

In higher eukaryotes, the condensin complex, which mainly consists of two structural maintenance of chromosomes (SMC) subunits, SMC2 (CAP-E) and SMC4 (CAP-C), plays a critical role in the formation of higher order chromosome structures during mitosis. Biochemical and electron-microscopic studies have revealed that the SMC2 and SMC4 subunits dimerize through the interaction of their hinge domains, forming a characteristic V-shaped hetero-dimer. However, the details of their function are still not fully understood owing to a lack of structural information at the atomic level. In this study, the human SMC2 hinge domain with short coiled coils was cloned, expressed, purified and crystallized in the orthorhombic space group $C222$ in native and SeMet-derivatized forms. Because of the poor diffraction properties of these crystals, the mutant Leu68→SeMet was designed and crystallized in order to obtain the experimental phases. The SeMet-derivatized crystals of the mutant belonged to space group $P3_212$, with unit-cell parameters $a = b = 128.8$, $c = 91.4$ Å. The diffraction data obtained from a crystal that diffracted to 2.4 Å resolution were suitable for SAD phasing.

1. Introduction

During mitosis, the formation of higher order structures of chromosomes is a critical event for the efficient transmission of genetic information to daughter cells. In higher eukaryotes, it has been revealed that a large multi-subunit protein complex called condensin plays an important role in this process (Hirano, 2005). Our proteomic analysis of human metaphase chromosomes confirmed that significant amounts of condensin subunits were present in the metaphase chromosome (Uchiyama *et al.*, 2005; Fukui & Uchiyama, 2007). The condensin complex is composed of two structural maintenance of chromosomes (SMC) subunits, SMC2 (CAP-E) and SMC4 (CAP-C), that form a heterodimer that interacts with three other non-SMC subunits (Ono *et al.*, 2003). In contrast, in prokaryotes there is only one SMC subunit, which functions as a homodimer.

The overall architecture of the condensin complex was initially revealed by electron-microscopy studies, which showed that each of the SMC subunits is a 50 nm long rod-like antiparallel coiled coil with a globular ATPase 'head' domain at one end and a 'hinge' domain at the other end (Anderson *et al.*, 2002). The two subunits are dimerized through the hinge–hinge interaction, forming a V-shaped hetero-dimer, and further interact with three non-SMC subunits at their heads to form a lollipop-like closed-ring structure. Because of this characteristic architecture, together with other biochemical studies, it was believed that the chromosome is folded by the condensin complex, trapping chromatin fibres inside the ring structure. However, recent results have demonstrated that a biologically active condensin complex which has been made cleavable at the coiled-coil region of the SMC2 subunit does not alter the condensin association or binding to chromosomes even when the subunit is cleaved on the chromosomes (Hudson *et al.*, 2008). Thus, the condensation of

crystallization communications

chromosomes by the complex may not require the topological closure of the complex and shows the importance of its interaction with DNA and/or other proteins, including other condensin complexes.

In this context, it is remarkable that direct interactions not only between condensin complexes but also between the condensin complex and DNA have been captured by atomic force microscopy studies on purified fission yeast condensin (Yoshimura *et al.*, 2002). These studies indicated that binding of the condensin complex to DNA might occur *via* its hinge domain, although the details of the interaction are still elusive. Recently, the structure of the mouse condensin SMC2–SMC4 hinge complex, which is the first eukaryotic SMC hinge crystal structure, was reported by Hopfner and coworkers (Griese *et al.*, 2010). Although the complex structure clearly revealed the complete structure of the SMC4 hinge, the C-terminal portion of the SMC2 hinge was not identified owing to the short construct. Consequently, this eliminates one of the two expected binding surfaces and hampered complete description of the hinge–hinge and/ or hinge–DNA interactions.

As seen in the crystal structure of the *Thermotoga maritima* SMC (*Tm*SMC) hinge domain with short coiled coils at both the C-terminal and N-terminal ends (PDB code 1gxl; Haering *et al.*, 2002), it is obvious that the hinge domain is stabilized by the interaction between the coiled-coil regions that extend from both ends. Therefore, in this study the hinge domain of human condensin SMC2 subunit with short coiled coils (hSMC-h-scc) was selected for cloning and overexpression. The crystallization and preliminary X-ray analyses of wild-type hSMC2-h-scc and several SeMet-substituted hSMC2-h-scc mutants are reported here.

2. Materials and methods

2.1. Cloning, expression and purification

A DNA fragment encoding amino acids 476–707 of the human SMC2 was amplified by PCR using full-length human SMC2 cDNA as a template. This was then inserted into the pET48b(+) expression vector (Invitrogen) between the *Xma*I and *Sal*I restriction sites. The recombinant plasmid was introduced into *Escherichia coli* strain Rosetta2 (DE3) and the cell culture was grown at 310 K to an optical density (OD_{660}) of 0.6. The culture was rapidly cooled on ice and protein expression was induced by adding isopropyl β-D-1-thio-galactopyranoside (IPTG) to a final concentration of 1 mM. After further growth for 24 h at 283 K, the cells were harvested by centrifugation at 8000g for 10 min at 277 K, resuspended in lysis buffer (50 mM Tris–HCl pH 8.0, 150 mM NaCl) and disrupted by ultra-sonication on ice. The lysate was centrifuged at 48 400g for 30 min to remove insoluble debris. All subsequent procedures were carried out at 283 K unless stated otherwise. The supernatant was filtered and loaded onto a HisTrap HP column (GE Healthcare) previously equilibrated with lysis buffer. The protein was eluted with a gradient of imidazole. The vector-encoded N-terminal His tag of the eluted protein was cleaved using HRV3C protease (Takara Bio) for 48 h at
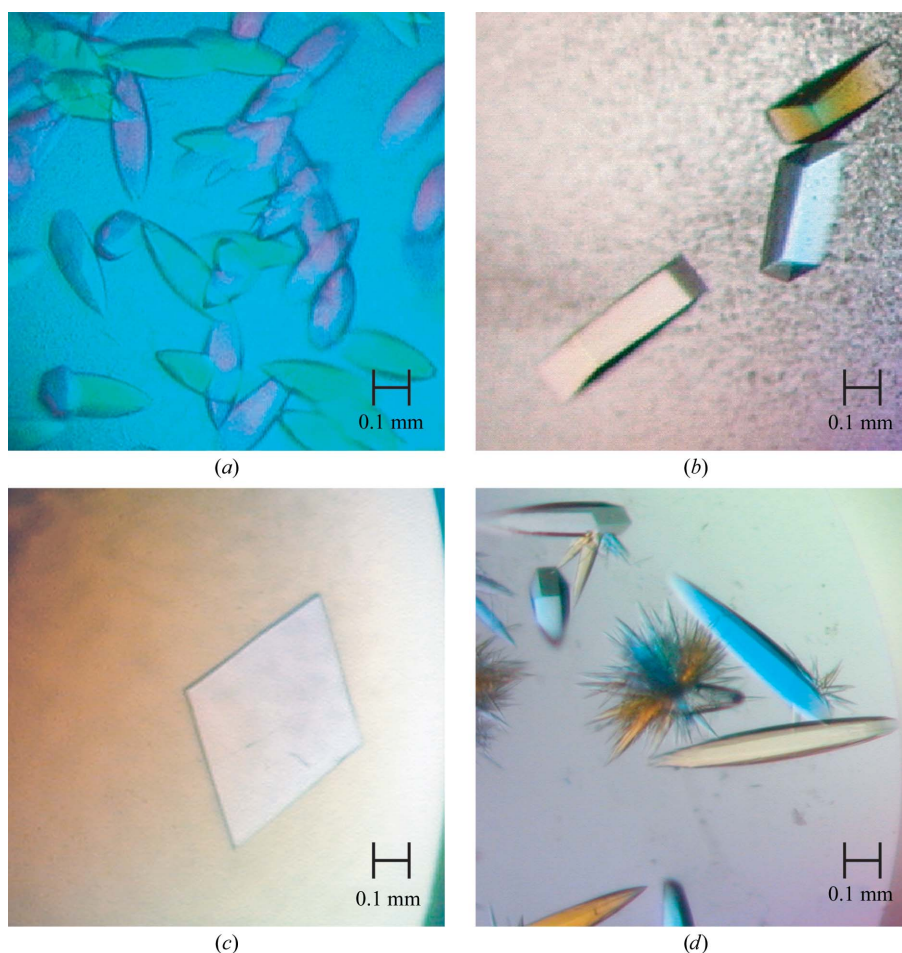
Figure 1
Crystals of hSMC2-h-scc. (*a*) Ellipsoid-shaped crystals of wild-type hSMC2-h-scc. (*b*) Improved cubic crystals of wild-type hSMC2-h-scc. (*c*) Cubic crystal of SeMet-substituted hSMC2-h-scc. (*d*) Hexagonal crystals of the Leu68→SeMet mutant of hSMC2-h-scc.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	Wild type	SeMet			Leu68→SeMet
		Peak	Edge	Remote	Peak
Crystal data					
Space group	*C*222	*C*222			*P*$3_2$12
Unit-cell parameters					
a (Å)	63.5	64.3	64.4	64.4	128.8
b (Å)	86.9	86.8	86.9	86.9	128.8
c (Å)	209.0	212.1	212.2	212.3	91.4
α (°)	90	90	90	90	90
β (°)	90	90	90	90	90
γ (°)	90	90	90	90	120
Data collection					
Resolution range (Å)	50–2.60 (2.69–2.60)	50–3.20 (3.31–3.20)			19.4–2.42 (2.46–2.42)
Wavelength (Å)	1.00000	0.978726	0.979275	0.964046	0.97898
Measured reflections	120169	69463	68931	68607	367095
Unique reflections	18131	10139	10143	10156	32972
Multiplicity	6.8 (5.3)	6.9 (4.9)	6.8 (4.6)	6.8 (4.7)	11.1 (11.1)
Completeness (%)	97.9 (84.3)	99.6 (96.5)	99.4 (95.2)	98.9 (90.3)	98.4 (91.3)
Mosaicity (°)	0.774	0.315	0.315	0.313	0.182
R_{merge}†	0.077 (0.365)	0.075 (0.392)	0.072 (0.415)	0.076 (0.433)	0.069 (0.395)
$\langle I \rangle / \langle \sigma(I) \rangle$	36.7 (2.3)	21.6 (2.3)	21.0 (2.1)	20.1 (2.0)	60.6 (8.3)

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the average intensity of reflection *hkl* and symmetry-related reflections.

277 K during dialysis against lysis buffer. Complete cleavage was achieved using a 1:1000(*w*:*w*) ratio of protease to target protein. The protein was again loaded onto a HisTrap HP column to remove His-tag-containing protein and was further purified by gel filtration on a HiLoad 26/60 Superdex 75 column (GE Healthcare). The purity of the protein was judged by SDS–PAGE analysis (data not shown).

SeMet-substituted hSMC2-h-scc was produced by introducing the expression construct into the Met-requiring auxotrophic *E. coli* strain B834 (DE3). SeMet-substituted hSMC2-h-scc was purified using the same procedure as used for the wild-type protein. To obtain phase information (see below), a Leu-to-SeMet mutation was additionally introduced into SeMet-substituted hSMC2-h-scc. Five of the 30 Leu residues were selected, *i.e.* Leu31, Leu66, Leu68, Leu74 and Leu188, and their Leu→Met mutants were created. Site-directed mutagenesis was performed with the QuikChange site-directed mutagenesis kit (Stratagene) and SeMet-substituted proteins corresponding to each Leu→Met mutant were individually expressed using the same protocol as used for SeMet-substituted hSMC2-h-scc.

2.2. Crystallization

Wild-type hSMC2-h-scc was concentrated to 15 mg ml^{-1} in 10 m*M* Tris–HCl pH 8.0 and 150 m*M* NaCl. Initial crystallization conditions were screened by the hanging-drop vapour-diffusion method at 277 K using the Crystal Screen and Crystal Screen 2 kits from Hampton Research and the Cryo I and Cryo II kits from Emerald BioSystems. Crystallization drops were prepared by mixing 1 µl protein solution (15 mg ml^{-1}) and 1 µl reservoir solution and were equilibrated against 80 µl reservoir solution. Ellipsoid-shaped crystals of hSMC2-h-scc grew using reservoir solution consisting of 100 m*M* phosphate–citrate pH 4.2, 25% 1,2-propanediol, 5% polyethylene glycol 3000 and 10% glycerol (Emerald BioSystems Cryo I condition No. 41) after one week (Fig. 1*a*). The crystallization conditions were optimized by varying the precipitant, pH and additives using the sitting-drop vapour-diffusion method at 277 K. Cube-shaped crystals of approximately 0.4 × 0.5 × 0.3 mm in size were grown from a mixture of 2 µl protein solution (15 mg ml^{-1}) and 3 µl reservoir solution consisting of 100 m*M* phosphate–citrate pH 4.5, 24% 1,2-propanediol, 10% glycerol and 15 m*M* NiCl$_2$ after one week (Fig. 1*b*).

SeMet-substituted hSMC2-h-scc was crystallized using the same conditions as were optimized for the wild-type hSMC2-h-scc crystals. The cube-shaped crystals grew to about 0.5 × 0.5 × 0.3 mm in size over two weeks (Fig. 1*c*). Of the five Leu-to-SeMet mutants, only Leu68→SeMet crystallized under these conditions. The crystal morphology was different from those of the wild-type and SeMet-substituted hSMC2-h-scc crystals; hexagonal-shaped crystals of about 0.3 × 0.7 × 0.3 mm in size were obtained after one week (Fig. 1*d*).

2.3. X-ray diffraction analysis

Prior to data collection, all crystals were retrieved from the droplet using a nylon loop (Hampton Research) and transferred into a cryoprotectant solution (reservoir solution containing 10% glycerol) before being placed directly in a cold nitrogen-gas stream at 100 K. For wild-type and SeMet-substituted crystals, X-ray diffraction data were collected using a Rigaku Jupiter 210 CCD detector on beamline BL38B1 of SPring-8 (Hyogo, Japan). Data from the wild-type crystal and MAD data from the SeMet-substituted crystal were collected to resolutions of 2.60 and 3.20 Å, respectively. For the crystals of the Leu68→SeMet mutant, X-ray diffraction data were collected on the BL17A beamline of Photon Factory (Tsukuba, Japan) using an ADSC Quantum 270 CCD detector. SAD data from the crystal of the Leu68→SeMet mutant were collected to a resolution of 2.42 Å. All diffraction data were indexed, integrated and scaled with *DENZO* and *SCALEPACK* from the *HKL*-2000 package (Otwinowski & Minor, 1997). Data-collection statistics are given in Table 1.

3. Results and discussions

Initially, recombinant proteins consisting of the human SMC2 hinge domain (Leu492–Glu680) fused to several affinity tags (*e.g.* His, Trx and GST) were generated using *E. coli* expression systems (data not shown). However, the proteins expressed by these constructs were unstable and/or insoluble. Therefore, new constructs for the human SMC2 hinge domain were designed with various lengths of coiled-coil sequences based on the results of secondary-structure prediction and homology modelling using the *Tm*SMC hinge-domain structure. Among the constructs, that encoding the sequence 476–707 with

about 30 residues of predicted coiled-coil region at both ends (hSMC2-h-scc) was successfully produced as a soluble protein with a yield of about 30 mg purified protein from 1 l culture.

After several rounds of crystallization screening, cube-shaped crystals of wild-type hSMC2-h-scc were obtained which diffracted to 2.6 Å resolution (Fig. 1*b*). However, molecular-replacement methods using the *Tm*SMC hinge domain as a search model did not produce interpretable phases. This problem was not solved by the use of SAD or MAD methods with SeMet-substituted hSMC2-h-scc crystals, which diffracted to 3.2 Å resolution and had crystallographic parameters similar to those of the wild-type crystals (Fig. 1*c* and Table 1). The diffraction images of these crystals exhibited significant diffraction anisotropy and showed relatively high mosaicity. In addition, the anomalous signal intensities obtained from the SeMet-substituted crystals were not sufficient for phase determination, suggesting that the crystal quality was poor and/or the SeMet residues were located in disordered regions.

To improve crystal quality and locate additional SeMet residues at better sites for SAD/MAD phasing, several optimization protocols such as amino-acid substitution methods have been proposed (Bäckbro *et al.*, 2004). Of these, the substitution of Leu by SeMet was selected. Based on the results of homology modelling, five Leu residues were selected that were predicted to be located on the inner part of the hinge globular domain. Although the Leu31→SeMet mutant was insoluble, the other mutants Leu66→SeMet, Leu68→SeMet, Leu74→SeMet and Leu188→SeMet were soluble. Of the four soluble mutants, only the crystal of the Leu68→SeMet mutant diffracted well (to a resolution of 2.42 Å with negligible diffraction anisotropy). The crystal belonged to space group $P3_212$, with two monomers per asymmetric unit and unit-cell parameters $a = b = 128.8$, $c = 91.4$ Å. The Matthews coefficient was 4.18 Å^3 Da^{-1} and the solvent content of the crystal was calculated to be 70.9% (Matthews, 1968). The space group and unit-cell parameters of the crystal differed from those of the wild-type and SeMet-substituted hSMC2-h-scc crystals even though these crystals were obtained using the same conditions. Furthermore, in contrast to the results from the SeMet-substituted crystals, which showed poor anomalous signal

intensities, an anomalous difference Patterson map for the Leu68→SeMet mutant crystals showed very strong peaks corresponding to all of the Se sites with two molecules in the asymmetric unit, indicating that the measured anomalous data would be useful in SAD phase determination. Phase calculations were performed using the programs *SHELXC*, *SHELXD* and *SHELXE* from the *SHELX* suite (Sheldrick, 2008) and the resulting electron-density map was easily interpretable. Model building and refinement of this model is now in progress.

References

Anderson, D. E., Losada, A., Erickson, H. P. & Hirano, T. (2002). *J. Cell Biol.* **156**, 419–424.
Bäckbro, K., Roos, A., Baker, E. N. & Arcus, V. L. (2004). *Acta Cryst.* D**60**, 733–735.
Fukui, K. & Uchiyama, S. (2007). *Chem. Rec.* **7**, 230–237.
Griese, J. J., Witte, G. & Hopfner, K. (2010). *Nucleic Acids Res.* **38**, 3454–3465.
Haering, C. H., Löwe, J., Hochwagen, A. & Nasmyth, K. (2002). *Mol. Cell*, **9**, 773–788.
Hirano, T. (2005). *Curr. Biol.* **15**, R265–R275.
Hudson, D. F., Ohta, S., Freisinger, T., Macisaac, F., Sennels, L., Alves, F., Lai, F., Kerr, A., Rappsilber, J. & Earnshaw, W. C. (2008). *Mol. Biol. Cell*, **19**, 3070–3079.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Ono, T., Losada, A., Hirano, M., Myers, M. P., Neuwald, A. F. & Hirano, T. (2003). *Cell*, **115**, 109–121.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Uchiyama, S., Kobayashi, S., Takata, H., Ishihara, T., Hori, N., Higashi, T., Hayashihara, K., Sone, T., Higo, D., Nirasawa, T., Takao, T., Matsunaga, S. & Fukui, K. (2005). *J. Biol. Chem.* **280**, 16994–17004.
Yoshimura, S. H., Hizume, K., Murakami, A., Sutani, T., Takeyasu, K. & Yanagida, M. (2002). *Curr. Biol.* **12**, 508–513.