Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

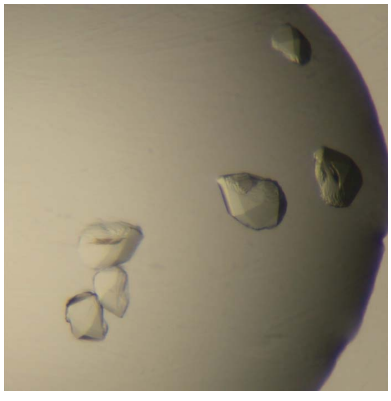
Crystallization and preliminary X-ray analysis of mitochondrial presequence receptor Tom20 in complexes with a presequence from aldehyde dehydrogenase

Mayumi Igura,[a] Toyoyuki Ose,[a] Takayuki Obita,[a] Chiaki Sato,[b] Katsumi Maenaka,[a] Toshiya Endo[c] and Daisuke Kohda[a]*

[a]Division of Structural Biology, Medical Institute of Bioregulation, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan, [b]Department of Structural Biology, Bioengineering Research Institute, Furuedai, Osaka 565-0874, Japan, and [c]Department of Chemistry, Graduate School of Science, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan

Correspondence e-mail: kohda@bioreg.kyushu-u.ac.jp

Received 18 March 2005
Accepted 14 April 2005
Online 22 April 2005



Most mitochondrial proteins are synthesized in the cytosol and must be imported into the mitochondria. Many mitochondrial precursor proteins have an extra leader sequence at their N-terminus called a presequence. Presequences are recognized by the Tom20 receptor protein. Based on the previously determined NMR structure of rat Tom20, a fragment corresponding to the core structure was generated. A cysteine residue was added at the C-terminus of the rat aldehyde dehydrogenase presequence to fix the presequence peptide onto the Tom20 fragment *via* an intermolecular disulfide bond. Two crystal forms of the complex were successfully obtained with different designs of the linker sequence which diffracted to 2.1 and 1.9 Å. Crystal dehydration and subsequent annealing was essential to obtain good diffraction data for the 2.1 Å crystal form.

1. Introduction

Most mitochondrial proteins are encoded by the nuclear genome, synthesized in the cytosol and subsequently imported into the mitochondria (Schatz & Dobberstein, 1996). Mitochondrial protein import across the mitochondrial outer membrane is mediated by a translocator in the outer membranes called TOM (translocase of the outer membrane) complex (Herrmann & Neupert, 2000; Endo *et al.*, 2003; Pfanner *et al.*, 2004). The TOM is a protein machinery consisting of membrane-protein subunits Tom70, Tom20, Tom22, Tom40 and small Tom subunits (Ahting *et al.*, 1999; Model *et al.*, 2002; Hoogenraad *et al.*, 2002).

Targeting of most proteins destined for the mitochondrial matrix and the inner membrane depends on N-terminal cleavable amino-acid sequences referred to as mitochondrial presequences. A presequence typically consists of about 15–40 amino-acid residues and is rich in positively charged residues (von Heijne, 1986). The targeting signal in the presequences is recognized by several receptor subunits of the TOM complexes (Bolliger *et al.*, 1995; Pfanner & Chacinska, 2002), including Tom20 (Muto *et al.*, 2001). Tom20 is anchored to the outer membrane by the N-terminal hydrophobic segment and exposes a receptor domain to the cytosol.

We determined the solution structure of the cytosolic domain (residues 51–145) of rat Tom20 in a complex with a presequence peptide derived from rat mitochondrial aldehyde dehydrogenease (ALDH) by NMR spectroscopy (Abe *et al.*, 2000). The NMR structure determination, however, suffered from the fast exchange of the presequence peptide between the free and bound states, which was a major cause of the insufficient number of NOE data between the peptide and Tom20 and within the peptide.

We therefore designed a 68-residue fragment that corresponds to the core structure (residues 59–126) of the cytosolic domain of Tom20 based on the NMR structure. This fragment contains a single cysteine residue. By adding a cysteine residue at the C-terminus of the presequence peptide with a suitable linker, the presequence peptide was fixed onto Tom20 through an intermolecular disulfide bond. The length of the linker was optimized by a novel peptide-library approach (Obita *et al.*, 2003). In the present research, we have carried

out the crystallization and preliminary X-ray crystallographic analysis of the core structure of rat Tom20 complexed with a presequence peptide from rat ALDH tethered by a designer intermolecular disulfide bond.

2. Materials and methods

2.1. Expression and purification of Tom20core

The core structure of the cytosolic domain of Tom20 (Tom20core) encompassing amino acids Asp59–Leu126 from *Rattus norvegicus* (accession No. Q62760) was produced as a GST-fusion protein in *Escherichia coli* BL21(DE3) cells (Novagen) using the *Bam*HI and *Eco*RI restriction sites of the pGEX 2T and pGEX 6P1 vectors (Amersham Biosciences). pGEX 2T encodes the recognition sequence for site-specific cleavage by thrombin, whereas pGEX 6P1 contains the recognition sequence of PreScission protease (Amersham Biosciences). Tom20core has an extra Gly-Ser sequence at the N-terminus after thrombin cleavage and a Gly-Pro-Leu-Gly-Ser sequence after PreScission protease cleavage. They are referred to as GS-Tom20 and GPLGS-Tom20 where necessary.

The transformed *E. coli* cells were cultured at 310 K in LB media containing 50 mg l^{-1} ampicillin. After OD_{600} reached 0.5, the culture was cooled to 293 K and supplemented with 0.5 m*M* IPTG. After overnight induction, the cells were harvested by centrifugation. The cell pellets were resuspended in lysis buffer (50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* DTT) and disrupted by sonication. The cell lysate was gently stirred with glutathione Sepharose 4B resin (Amersham Biosciences) previously equilibrated with lysis buffer. The affinity resin was packed in a column and washed with the same buffer. The GST-fusion protein was eluted from the resin with buffer containing 20–40 m*M* reduced glutathione and digested with thrombin [10:1(*w*:*w*), Wako Pure Chemicals] at 303 K for 12 h or PreScission protease [100:1(*w*:*w*) Amersham Biosciences] at 277 K for 12 h. The resulting protein mixture was applied onto a reverse-phase HPLC column (Cosmosil 5C4-AR-300, Nacalai Tesque) previously equilibrated with 0.1% TFA and 50% acetonitrile. The protein was eluted with a linear gradient of acetonitrile from 40 to 55% in 25 min at a flow rate of 1 ml min^{-1}. The fractions containing Tom20core were concentrated by ultrafiltration (5 kDa cutoff, Amicon Ultra, Millipore) in 0.1 *M* Tris–HCl buffer pH 8.0 containing 0.15 *M* NaCl. The typical yield was 5–10 mg per litre of culture. Selenomethionine (SeMet) Tom20core was expressed from the BL21(DE3) strain in minimal medium containing seleno-L-methionine at a concentration of 25 mg l^{-1}. SeMet Tom20core was purified following the same procedure described above and MALDI–TOF mass spectrometry of the purified protein confirmed the incorporation of selenium.

2.2. Presequence peptides

Peptides were obtained by custom peptide synthesis from Qiagen. The unpurified grade peptides were >95% pure and were used without further purification. The peptide sequence is **Gly-Pro-Arg-Leu-Ser-Arg-Leu-Leu-Ser-***X*-Ala-Gly-Cys with an N-terminal free amine and a C-terminal amide. The sequence in bold is derived from the presequence (pALDH) of rat mitochondrial ALDH enzyme (accession No. P11884). Two peptides were designed: the Y-linker peptide (pALDH-Y-linker) contains a Tyr residue and the A-linker peptide (pALDH-A-linker) contains an Ala residue at the *X* position, respectively. The C-terminal amidation is necessary for efficient disulfide-bond formation.

2.3. Preparation of disulfide-bond tethered complexes

$CuCl_2$ was added to the Tom20core solution (0.1 *M* Tris–HCl pH 8.0, 0.15 *M* NaCl) to a final concentration of 10 µ*M*. A 20% molar excess of the presequence peptide was then added to the solution. The intermolecular disulfide bond between Tom20core and the presequence peptide was formed through air-oxidation at ambient temperature for 24 h. The formation of the disulfide bond was nearly complete and no free Tom20core remained. The resulting disulfide-bonded complex was purified by reverse-phase chromatography in the same way as Tom20core and concentrated by ultrafiltration to 18 mg ml^{-1} for the native GS-Tom20-SS-pALDH-Y-linker complex in storage buffer (20 m*M* MOPS pH 7.0, 50 m*M* NaCl), 40 mg ml^{-1} for the SeMet-derivatized GS-Tom20-SS-pALDH-Y-linker complex in storage buffer (10 m*M* HEPES pH 7.5, 150 m*M* NaCl) and 17 mg ml^{-1} for the SeMet GPLGS-Tom20-SS-pALDH-A-linker complex in storage buffer (10 m*M* HEPES pH 7.5). Note that the choices of protein concentration (17–40 mg ml^{-1}), buffer (HEPES or MOPS), pH (7.0 or 7.5) and salt concentration (0–150 m*M*) were not relevant to successful crystallization.

2.4. Crystallization

Initial screening was carried out using the sitting-drop vapour-diffusion method in 96-well Intelli-plates (Art Robbins Instruments) and optimization was performed using the hanging-drop vapour-diffusion method in 24-well VDX greased plates (Hampton Research). Sitting drops were set up by mixing equal volumes (0.2 µl each) of the protein solution and reservoir solution using an automated dispenser (Hydra II Plus-One system, Apogent Discoveries), whereas hanging drops were prepared manually by mixing 2 µl of the protein solution and 1 µl of reservoir solution. Each sitting drop was placed over 0.1 ml reservoir solution and each hanging drop was placed over 0.4 ml reservoir solution. All crystallizations were carried out at 293 K.

3. Results

3.1. Crystallization

The crystallization screening kits used were Crystal Screens 1 and 2, Grid Screen Polyethylene Glycol 6000, Grid Screen PEG/LiCl and PEG/Ion Screen (Hampton Research). Crystals of the native Tom20-SS-pALDH-Y-linker were obtained with PEG/LiCl solution pH 7/20% PEG 6000, Crystal Screen 1 solution Nos. 9, 15, 21, 22, 26, 40 and 41 and Crystal Screen 2 solution No. 26. Crystals of the SeMet Tom20-SS-pALDH-A-linker were obtained with Crystal Screen 1 solution Nos. 20 and 40, Crystal Screen 2 solution Nos. 7 and 26 and PEG/Ion solution Nos. 5, 24 and 36. After optimization, crystals of the native complex with the Y-linker grew from a hanging drop with a 2:1 volume ratio of 18 mg ml^{-1} native Tom20-SS-pALDH-Y-linker protein stock and reservoir solution (0.1 *M* HEPES pH 7.0, 14% PEG 6000, 1 *M* NH_4Cl) in 7 d. Crystals of the SeMet complex with the Y-linker grew from a hanging drop with a 2:1 volume ratio of 40 mg ml^{-1} SeMet Tom20-SS-pALDH-Y-linker protein stock and microseed-containing reservoir solution (0.1 *M* HEPES pH 7.0, 15% PEG 6000, 1 *M* NH_4Cl) in 7 d. In the two cases, mixing 2 µl protein solution and 1 µl reservoir solution reproducibly gave better crystals than mixing equal volumes. Microseeds for crystallization were prepared by serial dilution of native Tom20-SS-pALDH-Y-linker crystals in the reservoir solution. In contrast, crystals of the SeMet complex with the A-linker suitable for structure determination were obtained without optimization from a sitting drop with a 1:1 volume

Table 1
Crystal data and data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Tom20-SS-pALDH-Y-linker				Tom20-SS-pALDH-A-linker
	SeMet				
	Peak	Edge	Remote	Native	SeMet
Space group	*C*2			*C*2	$P2_1$
Unit-cell parameters (Å, °)	a = 151.1, b = 64.5, c = 68.3, β = 94.4			a = 151.8, b = 64.1, c = 68.0, β = 94.7	a = 33.6, b = 27.6, c = 71.0, β = 103.0
No. of molecules per AU	7			7	2
Wavelength (Å)	0.9792	0.9794	0.9689	1.000	1.000
Resolution range (Å)	50–2.7 (3.0–2.9)	50–2.7 (3.0–2.9)	50–3.1 (3.21–3.1)	50–2.1 (2.18–2.1)	50–1.9 (1.98–1.9)
No. of measured reflections	90030 (11017)	90154 (11253)	70896 (9146)	190651 (10138)	311140 (2451)
No. of unique reflections	12446 (1468)	12481 (1480)	9870 (1203)	36605 (2981)	9958(942)
Completeness (%)	84.0 (100.0)	84.1 (100.0)	81.0 (100.0)	95.8 (77.7)	98.9 (93.2)
Multiplicity	7.2 (7.5)	7.2 (7.6)	7.2 (7.6)	5.2 (3.4)	3.1 (2.6)
Mean $I/\sigma(I)$	18.4	18.7	18.5	16.7	24.4
$R_{\rm merge}$† (%)	7.1 (37.9)	6.4 (36.7)	6.6 (35.3)	6.7 (30.2)	5.7 (19.3)

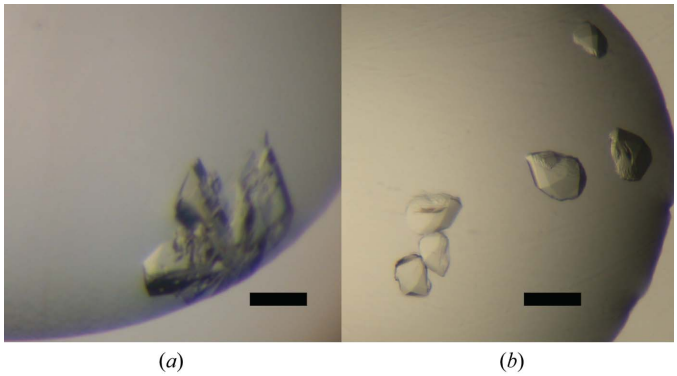
† $R_{\rm merge} = \sum_h \sum_j |\langle I\rangle_h - I_{h,j}| / \sum_h \sum_j I_{h,j}$, where $\langle I\rangle_h$ is the mean intensity of symmetry-equivalent reflections.

ratio of 17 mg ml^{-1} SetMet Tom20-SS-pALDH-A-linker protein stock and reservoir solution (PEG/Ion solution No. 5; 20% PEG 3350, 0.2 *M* $MgCl_2$) in 5 d (Fig. 1*a*).

3.2. Data collection

A crystal of the SeMet complex with the A-linker was cryoprotected by a gradual increase of ethylene glycol up to 15% in 3% steps and transferred frozen to the nitrogen-gas stream (100 K). The crystal diffracted to 1.9 Å at BL40B2, SPring-8 (Harima, Japan). A crystal of the SeMet complex with the Y-linker was processed in the same manner and diffracted to 2.8 Å at BL40B2. In contrast, crystals of the native complex with the Y-linker diffracted poorly to 3–8 Å. We succeeded in improving the diffraction resolution to 2.1 Å by crystal dehydration and subsequent annealing. Briefly, the hanging drop containing the crystals was equilibrated against 0.4 ml dehydrating solution (0.1 *M* HEPES pH 7.0, 25% PEG 6000, 1 *M* NH_4Cl) for 10 h at 293 K (Fig. 1*b*). The crystals were then briefly dipped into cryoprotectant solution consisting of dehydrating solution and 15% glycerol and cryocooled at 100 K. At this stage crystals diffracted to about 3 Å, but with streaky diffraction spots (Fig. 2*a*). A crystal was removed from the nitrogen-gas stream, soaked in a drop of the dehydrating solution containing 15% glycerol for 3 min and then flash-cooled again in the nitrogen-gas stream. This annealing treatment improved the diffraction quality to a resolution of 2.1 Å at BL41XU, SPring-8 (Fig. 2*b*).

The diffraction data sets were collected using 2.0° oscillations for the SeMet Y-linker crystal, 0.8° oscillations for the native Y-linker crystal and 2.0° oscillations for the SeMet A-linker crystal, with crystal-to-detector distances of 220, 290 and 170 mm, respectively. Each data set was collected from a single crystal at 100 K. The data



(*a*) (*b*)

Figure 1
Crystals of the complex of the cytosolic domain of Tom20 and the ALDH presequence peptide with an intermolecular disulfide bond. (*a*) SeMet Tom20-SS-pALDH-A-linker, scale bar 0.1 mm; (*b*) native Tom20-SS-pALDH-Y-linker after dehydration treatment, scale bar 0.2 mm.

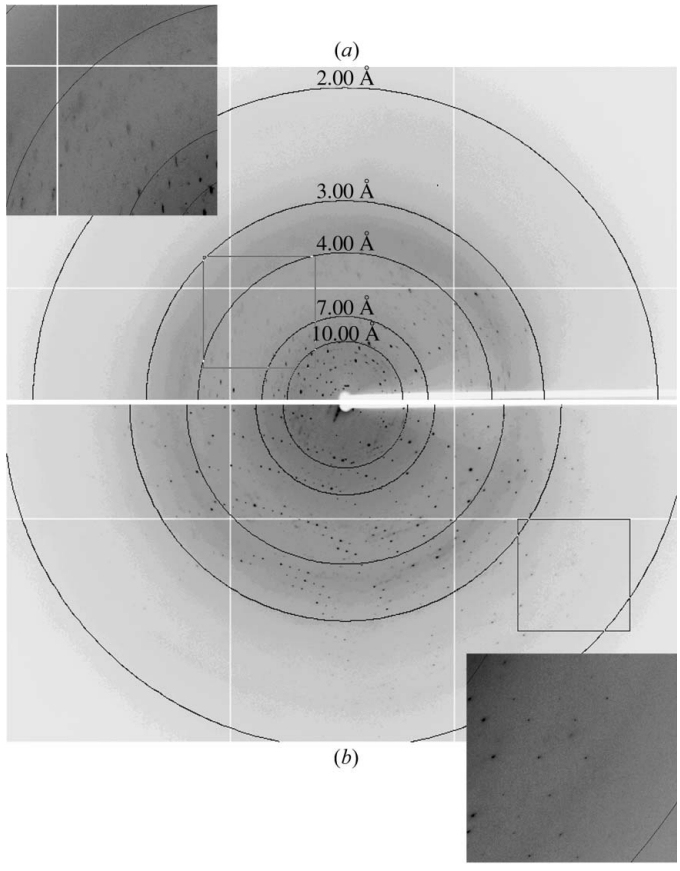


Figure 2
A comparison of diffraction from the same crystal (native Tom20-SS-pALDH-Y-linker) demonstrating the improvement of resolution by annealing treatment. (*a*) Before annealing; (*b*) after annealing. Resolution markers are indicated. Insets show the diffraction at the highest resolution limit. Note the streaky spots before annealing.

were indexed and scaled with *HKL*2000 (Otwinowski & Minor, 1997). The data-collection statistics are summarized in Table 1. The space group was *C*2 for the Y-linker crystals and $P2_1$ for the A-linker crystal, respectively.

4. Discussion

We succeeded in obtaining two crystal forms of the complex of the 68-residue receptor domain of rat Tom20 and a presequece peptide derived from rat mitochondrial ALDH enzyme. The presequence binds to Tom20 weakly, which prevented us from obtaining complex crystals by simple mixing of the two molecules. Therefore, we introduced a disulfide bond between the protein and the peptide to stabilize the complex. This trick had been successfully used to analyze the presequence recognition of Tom20; a reasonable pattern of the amino-acid preference of Tom20 was obtained (Obita *et al.*, 2003), which indicates that this strategy works well to fix the peptide into the binding groove of Tom20 in an unbiased manner. For the crystals of the native Tom20-SS-pALDH-Y-linker complex, crystal dehydration (Heras *et al.*, 2003) and subsequent annealing treatment (Samygina *et al.*, 2000) was essential to obtain diffraction data of high quality for structure determination. Although we attempted to carry out molecular replacement using the NMR structure, no meaningful solution was obtained. A starting model was obtained by the MAD method from the data set of SeMet Tom20-SS-pALDH-Y-linker, which was subsequently used to solve the structure of the SeMet Tom20-SS-pALDH-A-linker by molecular replacement. The structure of native Tom20-SS-pALDH-Y-linker was determined by molecular replacement with a model of the SeMet Tom20-SS-pALDH-A-linker. The detailed structure description will be published elsewhere.

We would like to thank M. Kawamoto, H. Sakai and K. Miura for assistance with data collection at SPring-8, Harima, Japan. KM and DK were supported by Grants-in-Aid for Scientific Research in Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan and the Protein 3000 project. OT was supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

References

Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T. & Kohda, D. (2000). *Cell*, **100**, 551–560.
Ahting, U., Thun, C., Hegerl, R., Typke, D., Nargang, F. E., Neupert, W. & Nussberger, S. (1999). *J. Cell Biol.* **147**, 959–968.
Bolliger, L., Junne, T., Schatz, G. & Lithgow, T. (1995). *EMBO J.* **14**, 6318–6326.
Endo, T., Yamamoto, H. & Esaki, M. (2003). *J. Cell Sci.* **116**, 3259–3267.
Heijne, G. von (1986). *EMBO J.* **5**, 1335–1342.
Heras, B., Edeling, M. A., Byriel, K. A., Jones, A., Raina, S. & Martin, J. L. (2003). *Structure*, **11**, 139–145.
Herrmann, J. M. & Neupert, W. (2000). *Curr. Opin. Microbiol.* **3**, 210–214.
Hoogenraad, N. J., Ward, L. A. & Ryan, M. T. (2002). *Biochim. Biophys. Acta*, **1592**, 97–105.
Model, K., Prinz, T., Ruiz, T., Radermacher, M., Krimmer, T., Kühlbrandt, W., Pfanner, N. & Meisinger, C. (2002). *J. Mol. Biol.* **316**, 657–666.
Muto, T., Obita, T., Abe, Y., Shodai, T., Endo, T. & Kohda, D. (2001). *J. Mol. Biol.* **306**, 137–143.
Obita, T., Muto, T., Endo, T. & Kohda, D. (2003). *J. Mol. Biol.* **328**, 495–504.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pfanner, N. & Chacinska, A. (2002). *Biochim. Biophys. Acta*, **1592**, 15–24.
Pfanner, N., Wiedemann, N., Meisinger, C. & Lithgow, T. (2004). *Nature Struct. Mol. Biol.* **11**, 1044–1048.
Samygina, V. R., Antonyuk, S. V., Lamzin, V. S. & Popov, A. N. (2000). *Acta Cryst.* D**56**, 595–603.
Schatz, G. & Dobberstein, N. (1996). *Science*, **271**, 1519–1526.