Synthesis, purification and crystallographic studies of the C-terminal sterol carrier protein type 2 (SCP-2) domain of human hydroxysteroid dehydrogenase-like protein 2

Zhong Cheng,[‡] Yao Li,[‡] Chun Sui, Xiaobo Sun and Yong Xie*

Received 27 January 2015

Accepted 1 May 2015

Edited by R. L. Stanfield, The Scripps Research Institute, USA

[‡] These authors contributed equally to this study.

Keywords: human HSDL2; SCP-2 domain; crystallization; X-ray diffraction.

Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, 151 Malianwa North Road, Haidian District, Beijing 100193, People's Republic of China. *Correspondence e-mail: yxie@implad.ac.cn

Human hydroxysteroid dehydrogenase-like protein 2 (HSDL2) is a member of the short-chain dehydrogenase/reductase (SDR) subfamily of oxidoreductases and contains an N-terminal catalytic domain and a C-terminal sterol carrier protein type 2 (SCP-2) domain. In this study, the C-terminal SCP-2 domain of human HSDL2, including residues Lys318–Arg416, was produced in *Escherichia coli*, purified and crystallized. X-ray diffraction data were collected to 2.10 Å resolution. The crystal belonged to the trigonal space group $P3_121$ (or $P3_221$), with unit-cell parameters $a = b = 70.4$, $c = 60.6$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Two protein molecules are present in the asymmetric unit, resulting in a Matthews coefficient of $2.16 \text{ Å}^3 \text{ Da}^{-1}$ and an approximate solvent content of 43%.

1. Introduction

Short-chain dehydrogenases/reductases (SDRs) are enzymes within the class of NAD(P)(H)-dependent oxidoreductases (Oppermann *et al.*, 2003; Bhatia *et al.*, 2015). Over 80 genes have been identified that encode enzymes with this function in carbohydrate, amino-acid and steroid hormone metabolism. To date, more than 20 000 SDRs have been found in prokaryotes and eukaryotes (Kowalik *et al.*, 2009). Owing to their pivotal roles in the regulation of different metabolic and signalling pathways, the dysfunction of SDR enzymes can lead to several diseases such as Alzheimer's disease, cancer and obesity-related medical conditions (Oppermann *et al.*, 1999; Chang *et al.*, 2005; Moeller & Adamski, 2006; Persson *et al.*, 2009). Therefore, these human SDRs constitute enzymes that may be interesting targets for drug development (Oppermann *et al.*, 2003; Zhang *et al.*, 2005).

Human hydroxysteroid dehydrogenase-like 2 (HSDL2; NCBI Reference Sequence NP_115679.2), which contains one polypeptide chain of 418 amino-acid residues, is a member of the short-chain dehydrogenase/reductase (SDR) subfamily of the oxidoreductases (Dai *et al.*, 2003). This enzyme is a ubiquitously expressed enzyme consisting of an N-terminal SDR domain and a C-terminal sterol carrier protein type 2 (SCP-2)-like domain containing a peroxisomal targeting signal (ARL). A search for regions of local similarity between sequences of human HSDL2 and other SDRs using the *Basic Local Alignment Search Tool* (BLAST; Altschul *et al.*, 1990; Oehmen & Baxter, 2013) revealed that the N-terminal SDR domain consists of residues Leu8–Ile249 and the C-terminal SCP-2 domain consists of residues Lys318–Met411. Subcellular localization studies and a screening assay for substrates

Table 1
Macromolecule-production information.

Source organism	<i>Homo sapiens</i>
DNA source	Full-length complementary DNA from HeLa cell
Forward primer†	5'-GGCGCCAAGGACTCTCTCAGTGATGATGTT- G-3'
Reverse primer‡	5'-CGGAATTCTCACAGTCTGGCATTTCATCTGAT- TC-3'
Cloning vector	pProEX HTb
Expression vector	pProEX HTb
Expression host	<i>E. coli</i> strain BL21(DE3) CodonPlus RIL
Complete amino-acid sequence of the construct produced§	MSYYHHHHHDYDAPTTENLYFQGAKDSLDDVV KATQAIYLFELSGEDGGTWFLDLKSKGGNVGY- GEPDQADVVMSTTDDFVKMFSGKLKPTMAF- MSGKLKIKGNMALAIKLEKLMNQMNRL

† The forward primer contains an *Nar*I restriction site (underlined). ‡ The reverse primer contains an *Eco*RI restriction site (underlined) and a stop codon (italic). § The amino-acid sequence of the SCP-2 domain is shown in italics and the TEV protease cleavage site is underlined. The peptide bond between the Q and G residues is cleaved by TEV protease.

of HSDL2, including steroid hormones and retinoids, revealed that a significant conversion of steroid hormones did not occur. The peroxisomal localization of human HSDL2 may suggest an involvement in fatty-acid metabolism (Kowalik *et al.*, 2009). The substrate specificity of this enzyme is unknown.

SCP-2, also known as nonspecific lipid-transfer protein, is a 13.3 kDa protein that is conserved among bacteria, archaea and eukaryotes. SCP-2 is specifically targeted to peroxisomes via a C-terminal peroxisomal targeting signal 1 (PTS1), where it is likely to be involved in peroxisomal α -oxidation of fatty acids (Kowalik *et al.*, 2009). Interestingly, SCP-2-deficient mice show an accumulation of methyl-branched fatty acids, which suggests a physiological role for SCP-2 in α -fatty acid oxidation (Lige *et al.*, 2009). The most C-terminal SCP-2 carries a PTS1 that directs the protein to vesicles before processing. Deletion of this signal results in accumulation of SCP-2 in the cytoplasm (Lige *et al.*, 2009).

In mammalian cells, the SCP-2 domain harbours a PTS1 and is responsible for the transport of D-bifunctional protein (DBP) to peroxisomes (Gallegos *et al.*, 2001). In peroxisomes, DBP undergoes proteolytic cleavage at a site between the two enzyme domains. These domains take part in peroxisomal oxidation of fatty acids through D-3-hydroxyacyl-CoA intermediates. It is possible that SCP-2 facilitates these reaction cycles by providing lipid substrates to the enzymatic active sites. The research showed that the C-terminus represents a subcellular localization signal in the parasite. It has been shown that the DBP C-terminal domain carries a subcellular localization signal in the parasite *Toxoplasma gondii*. Shifting the PTS1 from the C-terminus towards the centre of SCP-2 or deleting this motif results in cytosolic localization of the *T. gondii* DBP protein (Lige *et al.*, 2009). These results suggest that the C-terminus is critical for the interaction with the receptor and therefore the proper localization of the protein in the parasite. In mammalian cells, the 58 kDa SCP-x protein is partially post-translationally cleaved to an N-terminal 46 kDa protein (3-oxoacyl-CoA thiolase) and a C-terminal 13 kDa SCP-2 (Martin *et al.*, 2008). To date, no information exists on the mechanism of cleavage, by autocatalysis or

protease mediation, of mammalian SCP-2-containing proteins or on the cleavage-site residues.

A crystal structure of the N-terminal SDR domain binding to nicotinamide adenine dinucleotide phosphate (NADP) has previously been solved (PDB entry 3kvo; Structural Genomics Consortium, unpublished work). However, the crystal structures of full-length human HSDL2 and the C-terminal SCP-2 domain have still not been determined. We have therefore crystallized the C-terminus of SCP-2 to explore the possible molecular mechanism of this protein. The results of crystallization and preliminary crystallographic studies are reported in this paper.

2. Materials and methods

2.1. Macromolecule production and purification

The cDNA encoding the SCP-2 domain (residues Lys318–Leu418) of human HSDL2 was amplified by polymerase chain reaction using full-length human complementary DNA as a template, which was cloned using reverse transcription of the messenger RNA extracted from HeLa cells using TRIzol reagent (Invitrogen). The cDNA was cloned into the expression vector pProEX HTb (Novagen) as a fusion with an N-terminal His tag and a tobacco etch virus (TEV) protease cleavage site. The protein-production information is given in Table 1. The recombinant C-terminal SCP-2 domain of human HSDL2 was expressed in *Escherichia coli* strain BL21(DE3) CodonPlus RIL cells (Stratagene, USA).

Purification of the C-terminal SCP-2 domain of human HSDL2 was carried out at 277 K. The *E. coli* cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris–HCl buffer pH 8.0 containing 500 mM NaCl and 20 mM imidazole) and lysed by sonication. Cell debris was removed by centrifugation at 16 000g for 20 min. The supernatant was loaded onto a HisTrap column (5 ml) equilibrated with lysis buffer. After washing the column with lysis buffer, the His-tagged C-terminal SCP-2 domain of human HSDL2 was eluted with 20 mM Tris–HCl buffer pH 8.0 containing 500 mM NaCl and 500 mM imidazole. The sample buffer was exchanged to 20 mM Tris–HCl buffer pH 8.0 containing 500 mM NaCl and 20 mM imidazole using a HiPrep 26/10 desalting column. The His tag was cleaved with 140 μ l TEV protease (4 mg ml^{−1}) at 277 K overnight. To remove the uncleaved His-tagged protein, the reaction solution (14 ml) was loaded onto a HisTrap column (5 ml) as described above. The flowthrough fractions were collected and desalted on a HiPrep 26/10 desalting column with 20 mM Tris–HCl buffer pH 8.0 containing 5 mM β -mercaptoethanol (β ME). The pooled fraction (30 ml) was loaded onto a HiTrap Q column (5 ml) equilibrated with 20 mM Tris buffer pH 8.5 containing 5 mM β ME and 10 mM NaCl, and the C-terminal SCP-2 domain of human HSDL2 was eluted with a linear gradient of 10 mM to 1 M NaCl. Finally, the purified protein fraction (6 ml) was concentrated to 1.1 ml using an Amicon Ultra-15 filter (5000 MWCO, Millipore) and was loaded onto a HiLoad 16/60 Superdex 75 column equilibrated with 20 mM Tris–HCl

Table 2

Crystallization conditions for the SCP-2 domain of human HSDL2.

Method	Sitting-drop vapour diffusion
Plate type	CrystallEX (Corning) 96-well crystallization plate (Hampton Research)
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	15
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT
Composition of reservoir solution	100 mM bis-tris pH 6.5, 20%(w/v) PEG MME 5000
Volume and ratio of drop	2 µl (1 µl protein solution + 1 µl reservoir solution)
Volume of reservoir (µl)	200

buffer pH 8.0 containing 150 mM NaCl and 2 mM dithiothreitol (DTT). All chromatography materials were purchased from GE Healthcare Biosciences. Approximately 7.5 mg of the SCP-2 domain of human HSDL2 was obtained from 50 g of *E. coli* cells.

2.2. Crystallization

The purified C-terminal SCP-2 domain of human HSDL2 was concentrated to approximately 15 mg ml⁻¹ using an Amicon Ultra-15 (MWCO 3000) filter unit (Millipore). The concentration of the protein solution (500 µl) prior to crystallization was determined using UV spectrophotometry. Crystallization conditions are listed in Table 2. Crystal Screen HT, PEG/Ion and Index (Hampton Research) were used to establish initial crystallization conditions. Several single crystals with approximate dimensions of 0.2 × 0.1 × 0.1 mm were obtained using condition No. 46 [100 mM bis-tris pH 6.5, 20%(w/v) PEG MME 5000] of the Index screen within 3 d (Fig. 1).

2.3. Mass assay of crystallized protein

The crystals from 20 drops were harvested by centrifugation at 12 000g for 10 min at 277 K and the supernatant was removed with a micropipette. To remove the noncrystallized protein completely, the crystals were resuspended in 1 ml of solution No. 46 of the Index screen (Hampton Research) and the crystals were harvested using the same centrifugation

Figure 1

Typical single crystals of the C-terminal SCP-2 domain of human HSDL2.

Table 3

Data collection and processing.

Values in parentheses are for the highest resolution shell.

Diffraction source	BL17U, SSRF
Wavelength (Å)	1.000
Temperature (K)	100
Detector	ADSC Quantum 315r CCD
Crystal-to-detector distance (mm)	200
Rotation range per image (°)	1
Total rotation range (°)	−100 to 100
Exposure time per image (s)	10
Space group	<i>P</i> ₃ 21 or <i>P</i> ₃ 21
<i>a</i> , <i>b</i> , <i>c</i> (Å)	70.4, 70.4, 60.6
α , β , γ (°)	90.0, 90.0, 120.0
Mosaicity (°)	0.3
Resolution range (Å)	43.6–2.10 (2.18–2.10)
Total No. of reflections	22572
No. of unique reflections	11307 (1116)
Completeness (%)	99.7 (100)
Multiplicity	2.1 (2.0)
$\langle I/\sigma(I) \rangle$	15 (7.4)
$R_{\text{merge}}^{\dagger}$	0.064 (17.3)
$R_{\text{p.i.m.}}^{\ddagger}$	0.067 (17.5)
$R_{\text{r.i.m.}}^{\S}$	0.098 (33.6)
Overall <i>B</i> factor from Wilson plot (Å ²)	22.30

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed integrated intensity, $\langle I(hkl) \rangle$ is the average integrated intensity obtained from multiple measurements and the summation is over all observed reflections. $\ddagger R_{\text{p.i.m.}} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$. $\S R_{\text{r.i.m.}} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $N(hkl)$ is the data redundancy, $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the mean intensity from multiple observations of symmetry-related reflections.

protocol as described above. The supernatant was removed. These operations were repeated five times and the final harvested crystals were dissolved in 20 µl water.

2 µl protein solution was spotted onto a target of the Clin-TOF instrument. After the protein solution had dried at room temperature, 1 µl of a solution containing 5 mg ml⁻¹ α -cyano-4-hydroxycinnamic acid, 50%(v/v) acetonitrile and 1%(v/v) trifluoroacetic acid was added to the protein spot. After this solution had dried at room temperature, MALDI-TOF MS measurements were performed using a Clin-TOF instrument (Bioyong Technologies Inc.). The profile spectra were acquired from an average of 400 laser shots per sample with peak intensities (measured as *m/z*) in the mass range 1000–20 000 Da. The molecular weight of the SCP-2 domain of human HSDL2 was determined by MALDI-TOF/TOF MS operated in linear mode.

2.4. Data collection and processing

A single crystal with approximate dimensions of 0.2 × 0.1 × 0.1 mm was swished about to remove the reservoir solution using a cryoprotectant solution consisting of 66.6%(v/v) Paratone-N, 28.6%(v/v) paraffin and 4.8%(v/v) glycerol. The crystal was mounted on a rayon loop in a liquid-nitrogen gas stream at 100 K. X-ray diffraction data from the crystal were collected on beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF; Shanghai, People's Republic of China). The X-ray diffraction data-collection information is summarized in Table 3. Diffraction intensity data were processed and scaled using *HKL-2000* (Otwinowski & Minor, 1997). The crystal parameters are shown in Table 3.

3. Results and discussion

X-ray diffraction data were collected to 2.10 Å resolution. The crystal of the C-terminal SCP-2 domain of human HSDL2 belonged to the trigonal crystal system, with unit-cell parameters $a = b = 70.4$, $c = 60.6$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Processing the X-ray diffraction data in space groups $P321$ or $P312$ resulted in an overall R_{merge} of 6.7 or 45.9%, respectively. The 0, k , l X-ray diffraction images displayed using the *phenix.data_viewer* program in the *PHENIX* suite (Adams *et al.*, 2010) revealed that the allowed reflections are 0, 0, $l = 3n$.

Figure 2
Clin-TOF mass spectrum of the crystallized SCP-2 domain of human HSDL2.

Figure 3
Plot of the self-rotation function calculated using X-ray diffraction data in the resolution range 20.0–3.0 Å in the $\kappa = 180^\circ$ section. The map is contoured from the 1.0σ level in steps of 1.0σ .

Therefore, the space group of the crystal of the C-terminal SCP-2 domain of human HSDL2 is $P3_121$ (or $P3_221$). Clin-TOF mass spectrometry identified that the crystallized protein has a molecular weight of 10 934.33 Da (Fig. 2). This result revealed that the crystallized SCP-2 domain of human HSDL2 contains residues Lys318–Arg416 (with a theoretical molecular weight of 10 933.67 Da). Therefore, calculation of the Matthews coefficient and solvent content of this protein crystal (Matthews, 1968; Kantardjieff & Rupp, 2003) using the *MATTHEWS_COEF* program from the *CCP4* suite (Winn *et al.*, 2011) suggested that two protein molecules are present in the asymmetric unit, resulting in a Matthews coefficient of $2.16 \text{ Å}^3 \text{ Da}^{-1}$ and an approximate solvent content of 43%.

The self-rotation function calculated using *MOLREP* (Vagin & Teplyakov, 2010) revealed no additional twofold symmetry peaks besides those owing to crystallographic symmetry (Fig. 3). The *phenix.xtriage* program in the *PHENIX* suite (Adams *et al.*, 2010) showed that the largest off-origin peak in the native Patterson function is 6.39% of the height of the origin peak. These results suggest that no noncrystallographic symmetry can be detected.

Molecular-replacement (MR) searches were performed using the *MrBUMP* program (Keegan & Winn, 2007) to solve the crystal structure of the SCP-2 domain of human HSDL2. The programs *MOLREP* (Vagin & Teplyakov, 2010) and *Phaser* (McCoy *et al.*, 2007) and the crystal structures of the TPR domain of human Pex5p in complex with human Mscp2 (PDB entry 2c0l; 40% sequence identity; Stanley *et al.*, 2006) were used to search for a solution; however, a correct solution has not been found. In addition, two crystal structures of SCP-2 domains with 38% sequence identity to the SCP-2 domain of human HSDL2 were also used as MR search models: the SCP-2 domain from the yeast *Yarrowia lipolytica* (PDB entry 4jgx; De Berti *et al.*, 2013) and the SCP-2 domain from rabbit (PDB entry 1c44; Choinowski *et al.*, 2000). The crystal packing of each MR solution obtained using each of the mentioned structures as a search model revealed that two molecules clash. As no correct MR solutions were found with these models, the isomorphous replacement (IR) method or multiwavelength anomalous dispersion (MAD) method is presently being utilized to solve the crystal structure of the SCP-2 domain of human HSDL2.

Acknowledgements

The authors are grateful to the staff for the use of Clin-TOF mass instrument at Bioyong Technologies Inc. and beamline BL17U at SSRF. This study was supported by the National Natural and Science Foundation of China (No. 81473114), the Program for New Century Excellent Talents in University, the Fundamental Research Funds for the Central Universities and the PUMC Youth Fund.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.

- Bhatia, C., Oerum, S., Bray, J., Kavanagh, K. L., Shafqat, N., Yue, W. & Oppermann, U. (2015). *Chem. Biol. Interact.* **235**, 114–125.
- Chang, N.-S., Schultz, L., Hsu, L.-J., Lewis, J., Su, M. & Sze, C.-I. (2005). *Oncogene*, **24**, 714–723.
- Choinowski, T., Hauser, H. & Piontek, K. (2000). *Biochemistry*, **39**, 1897–1902.
- Dai, J., Xie, Y., Xie, Y., Wu, Q., Wang, L., Yin, G., Ye, X., Zeng, L., Xu, J., Ji, C., Gu, S., Huang, Q., Zhao, R. C. & Mao, Y. (2003). *Biochem. Genet.* **41**, 165–174.
- De Berti, F. P., Capaldi, S., Ferreyra, R., Burgardt, N., Acierno, J. P., Klinke, S., Monaco, H. L. & Ermácora, M. R. (2013). *J. Struct. Funct. Genomics*, **14**, 145–153.
- Gallegos, A. M., Atshaves, B. P., Storey, S. M., Starodub, O., Petrescu, A. D., Huang, H., McIntosh, A. L., Martin, G. G., Chao, H., Kier, A. B. & Schroeder, F. (2001). *Prog. Lipid Res.* **40**, 498–563.
- Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Keegan, R. M. & Winn, M. D. (2007). *Acta Cryst.* **D63**, 447–457.
- Kowalik, D., Haller, F., Adamski, J. & Moeller, G. (2009). *J. Steroid Biochem. Mol. Biol.* **117**, 117–124.
- Lige, B., Jayabalasingham, B., Zhang, H., Pypaert, M. & Coppens, I. (2009). *Mol. Biol. Cell*, **20**, 658–672.
- Martin, G. G., Hostetler, H. A., McIntosh, A. L., Tichy, S. E., Williams, B. J., Russell, D. H., Berg, J. M., Spencer, T. A., Ball, J., Kier, A. B. & Schroeder, F. (2008). *Biochemistry*, **47**, 5915–5934.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Moeller, G. & Adamski, J. (2006). *Mol. Cell. Endocrinol.* **248**, 47–55.
- Oehmen, C. S. & Baxter, D. J. (2013). *Bioinformatics*, **29**, 797–798.
- Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jörnvall, H. (2003). *Chem. Biol. Interact.* **143–144**, 247–253.
- Oppermann, U. C., Salim, S., Tjernberg, L. O., Terenius, L. & Jörnvall, H. (1999). *FEBS Lett.* **451**, 238–242.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Persson, B. *et al.* (2009). *Chem. Biol. Interact.* **178**, 94–98.
- Stanley, W. A., Filipp, F. V., Kursula, P., Schüller, N., Erdmann, R., Schliebs, W., Sattler, M. & Wilmanns, M. (2006). *Mol. Cell*, **24**, 653–663.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zhang, J., Osslund, T. D., Plant, M. H., Clogston, C. L., Nybo, R. E., Xiong, F., Delaney, J. M. & Jordan, S. R. (2005). *Biochemistry*, **44**, 6948–6957.