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

Preliminary structural characterization of human SOUL, a haem-binding protein

Filipe Freire,[a] Maria João Romão,[a] Anjos L. Macedo,[a] Susana S. Aveiro,[b] Brian J. Goodfellow[b] and Ana Luísa Carvalho[a]*

[a]REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal, and [b]Departamento de Química, CICECO, Universidade de Aveiro, 3810-193 Aveiro, Portugal

Correspondence e-mail: alcarvalho@dq.fct.unl.pt

Received 12 May 2009
Accepted 15 June 2009

Human SOUL (hSOUL) is a 23 kDa haem-binding protein that was first identified as the PP_{23} protein isolated from human full-term placentas. Here, the overexpression, purification and crystallization of hSOUL are reported. The crystals belonged to space group $P6_422$, with unit-cell parameters $a = b = 145$, $c = 60$ Å and one protein molecule in the asymmetric unit. X-ray diffraction data were collected to 3.5 Å resolution at the ESRF. A preliminary model of the three-dimensional structure of hSOUL was obtained by molecular replacement using the structures of murine p22HBP (PDB codes 2gov and 2hva), obtained by solution NMR, as search models.

1. Introduction

The SOUL protein was initially identified in the retina and pineal gland of chicken (Zylka & Reppert, 1999), with homologues subsequently being identified in mice and humans. Human SOUL (hSOUL), a 22 880 Da protein, was first identified in human full-term placentas (Bohn & Winckler, 1991). The murine form of SOUL (mSOUL) is also expressed in photoreceptive tissues; it has 27% sequence identity to murine p22HBP (mHBP) and also binds haem. A recent study of mSOUL indicated that this protein is a dimer in its apo form and becomes hexameric when bound to haem. The haem is thought to bind *via* coordination of the Fe^{III} haem to the side chain of His42, which is the only histidine present in the amino-acid sequence of hSOUL (Sato *et al.*, 2004).

Our group and collaborators determined the first structure of a protein from the SOUL/HBP family, mHBP, using nuclear magnetic resonance (NMR; Dias *et al.*, 2006). The mHBP structure adopts a novel fold in eukaryotes and was found to consist of a nine-stranded twisted β-barrel flanked by two α-helices. The dissociation constants for mHBP–haemin and mHBP–protoporphyrin IX (PPIX) complexes were determined to be in the low nanomolar range by fluorescence quenching. The chemical shift perturbations arising from the addition of haemin and PPIX were mapped onto the mHBP structure, allowing the binding site to be determined. From sequence alignments and homology modelling, it was concluded that p22HBP and SOUL should share a conserved tertiary fold, although they most probably bind haem at different sites within this fold (Dias *et al.*, 2006).

As chicken SOUL (ckSOUL) is solely expressed in the retina, where it functions as a circadian clock, and in the pineal gland, especially in retino-recipient regions where melatonin receptors are found, ckSOUL has been proposed to be involved in photoreceptive functions (Zylka & Reppert, 1999). In mammals, SOUL is also expressed in the retina and pineal gland. Other specific haem proteins have also been identified in these tissues, such as sGC (guanylyl cyclase), haem oxygenase and cytochrome P450s. This has led to suggestions that SOUL could be involved in haem transport to these proteins and in binding to free haem, avoiding the production of toxic reactive oxygen species (ROS; Sato *et al.*, 2004). hSOUL was first

identified as the PP_{23} protein isolated from human full-term placentas, being ubiquitously expressed in foetal (except for foetal liver) and adult tissues. hSOUL has also been identified as a binding partner of the SEC 1-homologue domain of the LRPRRC protein. LRPRRC is a leucine-rich protein proposed to be involved in vesicular trafficking, cytoskeleton organization, nucleocytosolic shuttling and chromosome activity (Liu & McKeehan, 2002). Recent studies of SOUL overexpression in NIH3T3 cell lines suggest that in addition to its haem-binding properties, SOUL provokes necrotic cell death by inducing mitochondrial membrane permeability. It has been shown that recombinant SOUL does not induce mitochondrial swelling by itself but that mitochondrial swelling was induced in the presence of a low Ca^{2+} concentration (30 µ*M*), leading to a significant decrease in NIH3T3 cell viability (Szigeti *et al.*, 2006). Despite all the studies performed to date, there is still no clear evidence for the function of this protein or of any other from the HBP/SOUL haem-binding protein family.

As hSOUL displays more than 40% homology to p22HBP, the existing solution structures of this protein (PDB codes 2gov and 2hva; Dias *et al.*, 2006; Gell *et al.*, 2006) were used as search models for molecular replacement in order to solve the crystal structure of hSOUL.

2. Materials and methods

Overexpression of the apo form of hSOUL was achieved by inoculating one colony of BL21 (DE3) cells (Novagen) harbouring the hSOUL/IOH3379-pDEST17-D18 plasmid (RZPD) in 2×YT media containing 25 µg µl^{-1} zeocin and incubating at 310 K. hSOUL was cloned with a histidine tag (MSYYHHHHHHLESTSLYKKAGT) attached to the N-terminus. The culture was then inoculated in M9 minimal media and incubated at 310 K. Protein expression was induced at an $OD_{600\,nm}$ of 0.6 using a final concentration of 0.1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h at 303 K.

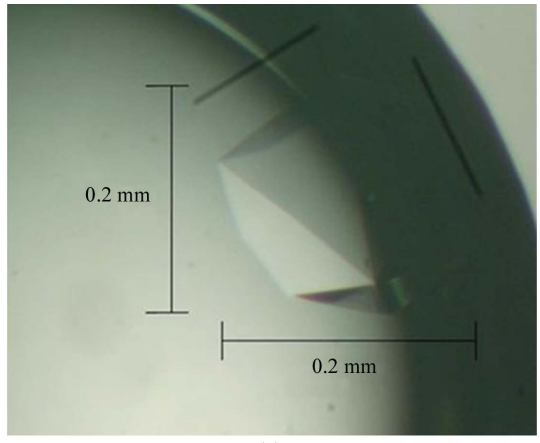
The harvested cells were resuspended in 50 m*M* phosphate buffer pH 8.0 with 300 m*M* NaCl, ruptured by sonication (Hielscher Ultraschall Technologie) and centrifuged at 48 384*g* (20 000 rev min^{-1}) for 1 h. The supernatant was loaded onto an Ni–NTA–agarose column (Qiagen) previously equilibrated with the same buffer. The resin was washed in two steps with a buffer containing 50 m*M* phosphate pH 8.0, 300 m*M* NaCl and 10 m*M* imidazole and with an identical buffer containing 20 m*M* instead of 10 m*M* imidazole. hSOUL was eluted in a discontinuous way with a buffer containing 50 m*M* phosphate pH 8.0, 300 m*M* NaCl and 250 m*M* imidazole. The fractions containing hSOUL were concentrated and loaded (approximately 400 µl containing 10 mg hSOUL) onto a Superdex 75 10/300 GL column (GE Healthcare, pre-packed) coupled to an FPLC system (GE Healthcare) previously equilibrated with 50 m*M* phosphate pH 8.0. The eluted fractions containing hSOUL were pooled together and concentrated to a final concentration of 15 mg ml^{-1} in an Amicon concentrator equipped with a YM10 membrane. The purity of the sample was confirmed by a 12% SDS–PAGE gel. After overexpression, hSOUL (in the apo form and with the histidine tag) was purified to homogeneity, with a final yield of around 20 mg protein per litre of culture.

Crystallization assays were carried out using the hanging-drop vapour-diffusion method at 293 K. A wide range of crystallization screens were used in order to obtain hSOUL crystals, namely an in-house prepared sparse-matrix screen (Jancarik & Kim, 1991) of 80 conditions, Crystal Screen 2 (Hampton Research), Emerald Wizard I (Emerald BioStructures), Emerald Wizard II (Emerald BioStructures) and Crystallization Basic Kit for Membrane Proteins (Sigma). Experiments were performed both at 277 and 293 K, with droplets consisting of 1 µl protein solution (15 mg ml^{-1} in 10 m*M* Tris–HCl pH 8.0) and 1 µl reservoir solution and 700 µl precipitant solution in the reservoir. The best crystallization conditions were obtained from the in-house prepared sparse-matrix screen of 80 conditions.

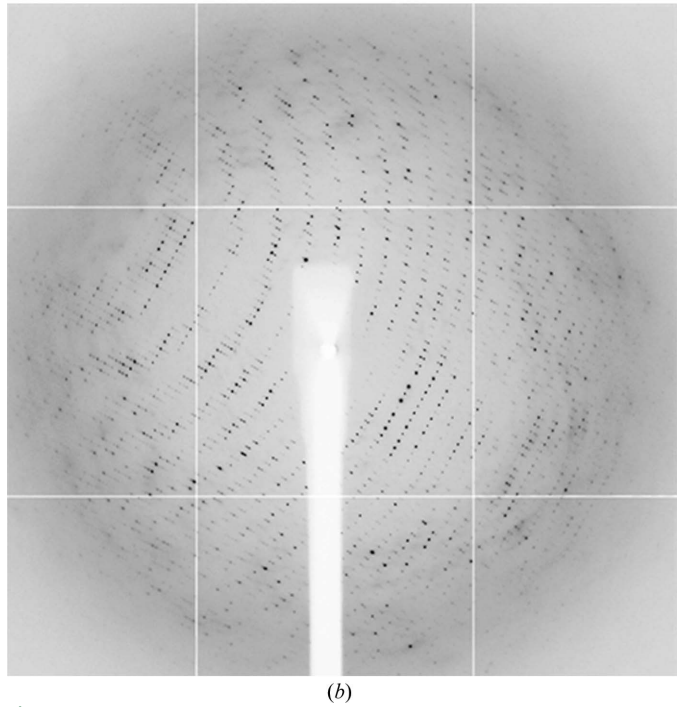
3. Results and discussion

3.1. Crystallization and data collection

The best hSOUL crystals (0.2 × 0.2 × 0.2 mm; Fig. 1) were obtained within 4 d in drops consisting of 1 µl protein solution (15 mg ml^{-1} in 10 m*M* Tris–HCl pH 8.0) and 1 µl reservoir solution (2 *M* ammonium sulfate, 0.1 *M* MES pH 6.5) at 293 K with 700 µl precipitant solution in the reservoir.



(*a*)



(*b*)

Figure 1
(*a*) Crystals of hSOUL grown over 4 d in 2 *M* ammonium sulfate and 0.1 *M* MES pH 6.5. Crystals with dimensions of 0.2 × 0.2 × 0.2 mm were obtained by the hanging-drop vapour-diffusion method. (*b*) X-ray diffraction pattern of a hSOUL crystal (the resolution limit is 3.5 Å).

hSOUL crystals were also obtained with 0.2 *M* ammonium sulfate, 0.1 *M* cacodylate pH 6.5 and 30% polyethylene glycol 8000. Crystals could also be grown in the presence of 2 *M* ammonium sulfate buffered with 0.1 *M* Tris–HCl pH 8.5, although these conditions led to multiplicity.

Crystal optimization attempts were made by varying the buffer pH, precipitant concentration and temperature. Additive Screen 1 and Additive Screen 2 (Hampton Research) were also tested at 277 and 293 K. However, owing to the poor quality of the crystals obtained under these conditions no diffraction data were collected.

Interestingly, the addition of harvesting buffer (2.5 *M* ammonium sulfate, 0.1 *M* MES pH 6.5) to a pre-equilibrated drop induced crystal formation within 12 h. Crystals thus obtained were cryoprotected (harvesting buffer supplemented with 25% glycerol) and directly cooled in liquid nitrogen. After 15 d, the crystals in the crystallization buffer showed evident signs of degradation and the diffraction quality decreased.

Several data sets were collected using single crystals grown under similar conditions in the search for data of the highest diffraction quality. Data collection was performed on various beamlines, namely ID14-EH2, ID14-EH4, ID23-EH1 and ID29 at the European Synchrotron Radiation Facility (Grenoble). All crystals were found to belong to the hexagonal point group 622 with a translational screw-axis component which could not be determined from the data alone, although examination of the 00*l* reflections suggested the existence of a 6_2 or 6_4 screw axis. After an exhaustive search for crystals of acceptable diffraction quality, the best data were collected on ID14-EH2 from a SOUL crystal (Fig. 1*a*) that diffracted to beyond 3.5 Å resolution (Fig. 1*b*). The unit-cell parameters were found to be $a = b = 145$, $c = 60$ Å which, after calculation of the Matthews coefficient ($V_M = 3.6$ Å^3 Da^{-1}), indicated the presence of one molecule of hSOUL (25 530 Da) in the asymmetric unit and a solvent content of 66%. All data were integrated with the program *MOSFLM* (Leslie, 1992) and scaled with *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) of programs. A summary of the data-collection statistics is shown in Table 1.

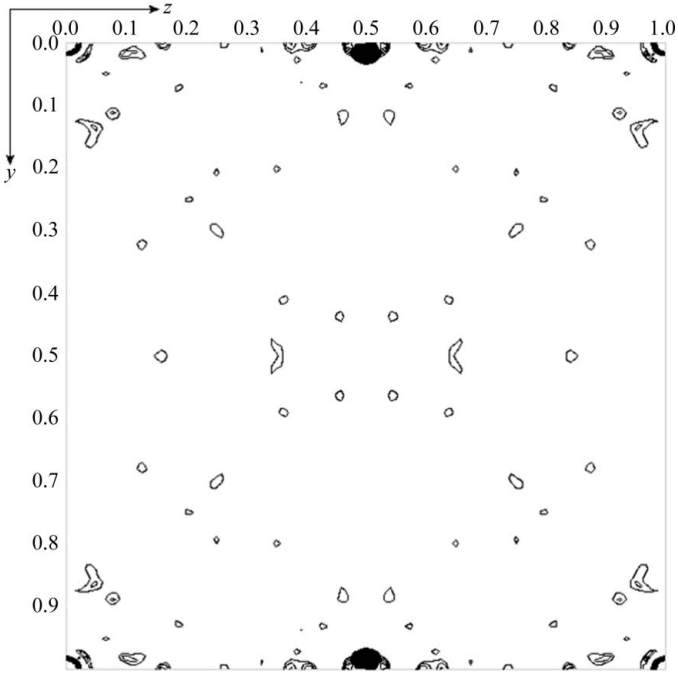


Figure 2
The calculated native Patterson map revealed the presence of a pseudo-translation vector in some of the hSOUL crystals. The plot was calculated at 3 Å resolution at section $u = 0$, with a contour level of 2σ.

Table 1
Crystal data and data-collection statistics for hSOUL.

Values in parentheses are for the highest resolution shell.

X-ray source	ID14-EH2, ESRF
Wavelength (Å)	0.931
Space group	$P6_422$
Unit-cell parameters (Å)	$a = b = 145$, $c = 60$
Resolution range (Å)	47.35–3.50 (3.69–3.50)
No. of observed reflections	42032
No. of unique reflections	5040
$\langle I/\sigma(I)\rangle$	14.3 (2.7)
$R_{\rm p.i.m.}$†	0.052 (0.265)
$R_{\rm merge}$‡	0.157 (0.809)
Completeness	100.0 (98.5)
Multiplicity	8.3 (5.9)

† $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average of symmetry-related observations of a unique reflection. ‡ $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$.

3.2. Preliminary structure determination of hSOUL

Determination of the three-dimensional structure was carried out by the molecular-replacement method. The available solution structures (PDB codes 2gov and 2hva) obtained by NMR spectroscopy were used as search models in the *BALBES* molecular-replacement system (Long *et al.*, 2008). These models show around 27% sequence similarity to hSOUL and the *CHAINSAW* module of *BALBES* produced a search model from these, modifying the template average structures on a residue-by-residue basis (Stein, 2008). The search was performed for all choices of possible space groups and a molecular-replacement solution was found in space group $P6_422$ with one molecule of SOUL in the asymmetric unit using the program *MOLREP* (Vagin & Teplyakov, 1997) implemented in *BALBES*, with an MR score of 2.14, a *C* score of 0.9087 and a figure of merit of 0.44. This solution was further confirmed using the program *Phaser* (Storoni *et al.*, 2004), which found a clear solution for space group $P6_422$ with a *Z* score of 9.63 and a refined LL gain of 90.45, compared with a *Z* score of 5.46 and an LL gain of 20.37 for space group $P6_222$.

For some of the tested crystals the presence of systematically weak reflections suggested the existence of a pseudo-translation vector. This was confirmed by the presence of a strong peak (with around 20% of the origin peak height) in the calculated native Patterson map (Fig. 2). In these cases, the cell constants were doubled along the *c* axis, suggesting the presence of three molecules of SOUL in the crystal asymmetric unit. *Phaser* was not able to produce a molecular-replacement solution from these data sets. Attempts to find a solution also involved the program *MOLREP*, using the known pseudo-translation vector, although a clear MR solution was not found.

The authors acknowledge financial support from Fundação para a Ciência e Tecnologia, Portugal through grant SFRH/BPD/30239/2006 and project PTDC/QUI/64203/2006. The authors also acknowledge the ESRF, Grenoble (beamlines ID14-EH2, ID14-EH4, ID23-EH1 and ID-29) for access and technical support during data collection.

References

Bohn, H. & Winckler, W. (1991). *Arch. Gynecol. Obstet.* **248**, 111–115.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dias, J. S., Macedo, A. L., Ferreira, G. C., Peterson, F. C., Volkman, B. F. & Goodfellow, B. J. (2006). *J. Biol. Chem.* **281**, 31553–31561.

Gell, D. A., Westman, B. J., Gorman, D., Liew, C., Welch, J. J., Weiss, M. J. & Mackay, J. P. (2006). *J. Mol. Biol.* **362**, 287–297.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Liu, L. & McKeehan, W. L. (2002). *Genomics*, **79**, 124–136.
Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 125–132.
Sato, E., Sagami, I., Uchida, T., Sato, A., Kitagawa, T., Igarashi, J. & Shimizu, T. (2004). *Biochemistry*, **43**, 14189–14198.
Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Szigeti, A., Bellyei, S., Gasz, B., Boronkai, A., Hocsak, E., Minik, O., Bognar, Z., Varbiro, G., Sumegi, B. & Gallyas, F. Jr (2006). *FEBS Lett.* **580**, 6447–6454.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Zylka, M. J. & Reppert, S. M. (1999). *Brain Res. Mol. Brain Res.* **74**, 175–181.