crystallization communications1

Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of recombinant human fumarase

Ricardo Augusto Pereira de Pádua and Maria Cristina Nonato*

Departamento de Física e Química, Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Avenida do Café s/n, 14040-903 Ribeirão Preto-SP, Brazil

Correspondence e-mail: cristy@fcfrp.usp.br

Human fumarase (HsFH) is a well-known citric acid cycle enzyme and is therefore a key component in energy metabolism. Genetic studies on human patients have shown that polymorphisms in the fumarase gene are responsible for diseases such as hereditary leiomyomatosis and renal cell cancer. As a first step in unravelling the molecular basis of the mechanism of fumarase deficiency in genetic disorders, the HsFH gene was cloned in pET-28a, heterologously expressed in *Escherichia coli*, purified by nickel-affinity chromatography and crystallized using the vapour-diffusion technique. X-ray diffraction experiments were performed at a synchrotron source and the structure was solved at 2.1 Å resolution by molecular replacement.

1. Introduction

Fumarate hydratases (fumarases) are responsible for the reversible hydration of fumarate to give L-malate and are distributed in two different major classes: class I fumarases are homodimeric proteins with a molecular mass of 120 kDa and rely on an iron–sulfur cluster (4Fe–4S) as part of the catalytic centre, whereas class II fumarases are homotetrameric non-iron-dependent enzymes with a molecular mass of 200 kDa (Woods *et al.*, 1988; Flint *et al.*, 1992).

In eukaryotic cells, fumarases can be found localized either in the mitochondria, where they participate in the TCA cycle, or in the cytosol, where they are described as being involved in the metabolism of amino acids and fumarate. In human parasites, such as kinetoplastids, the mitochondrial and cytosolic enzymes are members of the class I fumarases and are encoded by two separate genes (Coustou *et al.*, 2006; Feliciano *et al.*, 2012). In humans, both mitochondrial and cytosolic class II fumarases are encoded by the *fumC* gene and their subcellular localization is determined by post-translational processes (Yogev *et al.*, 2011).

Mutations in fumarase have been implicated in a variety of human diseases, including progressive encephalopathy, fumaric aciduria, hereditary leiomyomatosis and renal cell cancer (Kerrigan *et al.*, 2000; Yang *et al.*, 2013; Whelan *et al.*, 1983). Despite its remarkable importance, the only structural study presently available regarding the clinical relevance of human fumarase (HsFH) was performed on the *Escherichia coli* FumC homologue, which shares 60% identity with the human enzyme (Estévez *et al.*, 2002).

In order to provide new structural insights into the mechanism of action of HsFH, we here describe a reproducible protocol to produce HsFH in high yield and purity by means of recombinant expression in *E. coli*, as well as the crystallization and X-ray diffraction experiments ued to solve the structure of HsFH at 2.1 Å resolution.

2. Materials and methods

2.1. Macromolecule production

The human fumarate hydratase (HsFH) gene (GenBank AAA52483.1) flanked by 5′ *Bam*HI and 3′ *Hin*dIII restriction sites was synthesized and cloned in a pUC57 vector by GeneScript (USA). Codons were pre-optimized for expression in *E. coli*. The HsFH gene
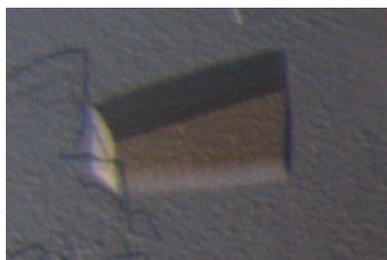
was subcloned into pET-28a vector (Merck Millipore) using *Bam*HI and *Hin*dIII restriction enzymes.

A colony of *E. coli* BL21 (DE3) cells transformed with pET-28a-*HsFH* was used to inoculate 10 ml LB broth (Acumedia) containing 30 µg ml^{-1} kanamycin (Sigma), which was incubated under agitation (180 rev min^{-1}) overnight at 37°C. The starter culture was diluted 100-fold in LB broth and grown for 2.5 h at 37°C. Immediately after, IPTG (Sigma) was added to 500 µ*M* and HsFH expression was carried out overnight at 18°C.

A pellet corresponding to 167 ml cell suspension was resuspended in 20 ml buffer *A* [50 m*M* NaH$_2$PO$_4$ pH 8.5 (Sigma), 300 m*M* NaCl (JT Baker)] containing 1 m*M* PMSF (USB). Cells were sonicated using ten pulses of 30 s with an interval of 30 s resting on ice. The soluble fraction obtained after centrifugation was poured into 2 ml Ni–NTA resin (Qiagen) packed into a Poly-Prep column (Bio-Rad) previously equilibrated with buffer *A*. 15 ml buffer *A* containing 50 m*M* imidazole (Sigma) was used to wash out contaminants and HsFH was eluted with 10 ml buffer *A* containing 500 m*M* imidazole. Pure HsFH was dialyzed against 50 m*M* Tris pH 8.5, 150 m*M* NaCl, concentrated using a 30 kDa cutoff centrifugal filter unit (Millipore) and quantified using a theoretical extinction coefficient $\varepsilon_{280\,\text{nm}}$ of 24 410 *M*$^{-1}$ cm^{-1}. The macromolecule-production information is summarized in Table 1.

2.2. Crystallization

HsFH at 4 mg ml^{-1} was used for crystallization screening by the sitting-drop vapour-diffusion technique using the commercial Crystal Screen, Crystal Screen 2, PEG/Ion and PEG/Ion 2 kits (Hampton Research). The plates were stored at controlled room temperature (293 K) and the drops were frequently checked for crystals. Crystals grew overnight in 100 m*M* sodium malonate pH 5 with 12% poly-ethylene glycol 3350 in a 4 µl drop using a protein:crystallization solution ratio of 1:1 (Table 2).
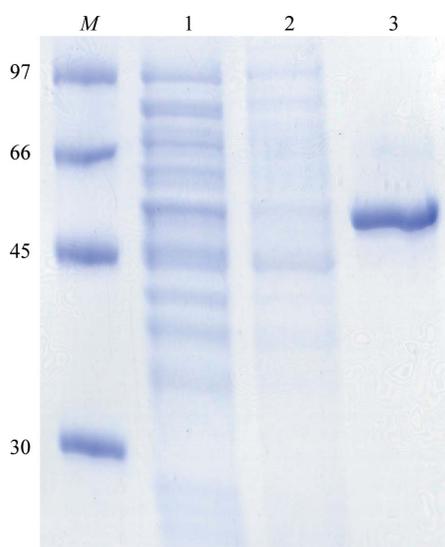
Figure 1
SDS–PAGE analysis of HsFH purification steps by nickel-affinity chromatography. Lane *M*, molecular markers (labelled in kDa). Lane 1, soluble fraction obtained after lysate centrifugation. Lane 2, washing step in the presence of buffer *A* containing 50 m*M* imidazole. Lane 3, HsFH enzyme (53.7 kDa) elution with buffer *A* containing 500 m*M* imidazole.

Table 1
Macromolecule-production information.

Source organism	*Homo sapiens*
DNA source	Synthetic gene (GenScript)
Cloning vector	pUC57
Expression vector	pET-28a (Merck Millipore)
Expression host	*E. coli* BL21(DE3)
Complete amino-acid sequence of the construct produced†	mgsshhhhhhssglvprgshmasmtggqqmgrgsMASQNSFR-IEYDTFGELKVPNDKYYGAQTVRSTMNFKIG-GGVTERMPTPVIKAFGILKRAAAEVNQDYG-LDPKIANAIMKAADEVAEGKLNDHFPLVVW-QTGSGTQTNMNVNEVISNRAIEMLGGEGLGS-KIPVHPNDHVNKSQSSNDTFPTAMHIAAAIE-VHEVLLPGLQKLHDALDAKSKEFAQIIKIGR-THTQDAVPLTLGQEFSGYVQQVKYAMTRIK-AAMPRIYELAAGGTAVGTGLNTRIGFAEKV-AAKVAALTGLPFVTAPNKFEALAAHDALVE-LSGAMNTTACSLMKIANDIRFLGSGPRSGLG-ELILPENEPGSSIMPGKVNPTQCEAMTMVAA-QVMGNHVAVTVGGSNGHFELNVFKPMMIKN-VLHSARLLGDASVSFTENCVVGIQANTERIN-KLMNESLMLVTALNPHIGYDKAAKIAKTAH-KNGSTLKETAIELGYLTAEQFDEWVKPKDM-LGPK

† The histidine tag is represented in lower case in the amino-acid sequence. The tag was retained for the crystallization experiments.

Table 2
Crystallization.

Method	Vapour diffusion
Plate type	Sitting drop
Temperature (K)	293
Protein concentration (mg ml^{-1})	4
Buffer composition of protein solution	50 m*M* Tris pH 8.5, 150 m*M* NaCl
Composition of reservoir solution	100 m*M* sodium malonate pH 5, 12% polyethylene glycol (PEG) 3350
Volume and ratio of drop	4 µl, 1:1
Volume of reservoir (µl)	500

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	MX2, LNLS
Wavelength (Å)	1.459
Temperature (K)	100
Detector	3 × 3 MAR Mosaic 225 CCD
Crystal-to-detector distance (mm)	121.6
Rotation range per image (°)	1
Total rotation range (°)	102
Exposure time per image (s)	60
Space group	*C*222
Unit-cell parameters (Å, °)	$a = 124.42$, $b = 147.15$, $c = 57.01$, $\alpha = \beta = \gamma = 90$
Mosaicity (°)	1.29
Resolution range (Å)	31.67–2.10 (2.175–2.100)
Total No. of reflections	105423 (15169)
No. of unique reflections	27725 (4192)
Completeness† (%)	89.9 (94.4)
Multiplicity	3.8 (3.6)
$\langle I/\sigma(I)\rangle$	3.6 (2.1)
$R_{\text{r.i.m.}}$	0.210 (0.495)
$R_{\text{p.i.m.}}$	0.108 (0.247)
R_{merge}	0.179 (0.423)
Overall *B* factor from Wilson plot (Å^2)	14.8

† The strategy option in *MOSFLM* was used to plan data collection to optimize data completeness (Leslie & Powell, 2007). However, the 3 × 3 MAR Mosaic 225 CCD detector used at the MX2 beamline had a defect that affected data acquisition in the intermediate resolution range.

2.3. Data collection, processing and phasing

An HsFH crystal was harvested and transferred to a fresh drop containing cryoprotectant solution consisting of 100 m*M* sodium malonate pH 5, 14% PEG 3350, 30% glycerol. The dehydration procedure proved to be important for improvement of the crystal

crystallization communications[1]
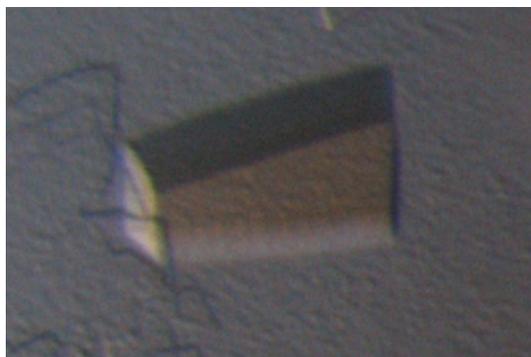
Figure 2
An HsFH single crystal was obtained in the presence of 100 m*M* sodium malonate pH 5, 12% PEG 3350 in sitting-drop plates. HsFH crystals are colourless; the crystal picture shown was taken under polarized light.

order and diffraction resolution. Crystal dehydration was performed by exposing the drop to air for 2 h. Longer dehydration steps induced the formation of precipitates and a thin film covering the drop, both of which are undesirable for crystal handling and X-ray diffraction experiments. The crystal was then mounted in a cryo-loop for X-ray diffraction experiments and data collection was carried out on the MX2 beamline at LNLS, Campinas, Brazil using an X-ray wavelength of 1.459 Å. The crystal was positioned 121.6 mm from the 3×3 MAR Mosaic 225 CCD detector and was maintained at 100 K using an Oxford Nitrogen Cryojet XL system. The full data set is comprised of 102 images collected using $1.0°$ oscillations with an exposure time of 60 s. Images were indexed in *iMosflm* (Battye *et al.*, 2011) and scaled in *SCALA* (Evans, 2006) (Table 3). Initial phases were obtained by molecular replacement with *Phaser* in *PHENIX* (Adams *et al.*, 2010) utilizing PDB entry 3e04 (Structural Genomics Consortium, unpublished work) as the template.

3. Results and discussion

HsFH was expressed in soluble form and purified to homogeneity, with a yield of 110 mg pure protein per litre of *E. coli* culture (Fig. 1). HsFH crystals obtained using PEG as the precipitant agent (Fig. 2) were used in the X-ray diffraction experiments. The best diffraction pattern was obtained from a single crystal partially dehydrated in the presence of cryoprotectant solution.

Unit-cell parameters, merging statistics and the lack of systematic absences indicated that the enzyme crystals belonged to space group *C*222 (Table 3). HsFH has a calculated molecular weight of approximately 53.7 kDa and the calculated Matthews coefficient is 2.43 Å^3 Da^{-1} assuming the presence of one molecule in the asymmetric unit; based on a partial specific volume of 0.74 cm^3 g^{-1}, the calculated solvent content is approximately 49.3%. These values lie within the range commonly observed for protein crystals (Kantardjieff & Rupp, 2003; Matthews, 1968).

Data phasing was performed by molecular replacement, and initial refinement was performed by several rounds of adjustment of side-chain rotamers using *Coot* (Emsley & Cowtan, 2004) interspersed with torsion-angle simulated-annealing and positional and individual *B*-factor refinement using *phenix.refine* (Adams *et al.*, 2010). Partial refinement reached an *R* factor of 18% and an R_{free} of 23%. Iterative manual building and refinement of the model are currently under way.

The structural and kinetic analyses of native and mutant forms of HsFH are in progress. Our results will contribute to the understanding of the molecular mechanism adopted by class II fumarases, as well as providing a structural basis for the role of different HsFH mutations in deficient fumarase activity and their relationship to distinct genetic disorders. The development of a reproducible protocol for HsFH purification and crystallization reported here represents an important step towards this goal.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Coustou, V., Biran, M., Besteiro, S., Rivière, L., Baltz, T., Franconi, J. M. & Bringaud, F. (2006). *J. Biol. Chem.* **281**, 26832–26846.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Estévez, M., Skarda, J., Spencer, J., Banaszak, L. & Weaver, T. M. (2002). *Protein Sci.* **11**, 1552–1557.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Feliciano, P. R., Gupta, S., Dyszy, F., Dias-Baruffi, M., Costa-Filho, A. J., Michels, P. A. M. & Nonato, M. C. (2012). *Int. J. Biol. Macromol.* **51**, 25–31.
Flint, D. H., Emptage, M. H. & Guest, J. R. (1992). *Biochemistry*, **31**, 10331–10337.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kerrigan, J. F., Aleck, K. A., Tarby, T. J., Bird, C. R. & Heidenreich, R. A. (2000). *Ann. Neurol.* **47**, 583–588.
Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Whelan, D. T., Hill, R. E. & McClorry, S. (1983). *Clin. Chim. Acta*, **132**, 301–308.
Woods, S. A., Schwartzbach, S. D. & Guest, J. R. (1988). *Biochim. Biophys. Acta*, **954**, 14–26.
Yang, Y., Lane, A. N., Ricketts, C. J., Sourbier, C., Wei, M.-H., Shuch, B., Pike, L., Wu, M., Rouault, T. A., Boros, L. G., Fan, T. W. M. & Linehan, W. M. (2013). *PLoS One*, **8**, e72179.
Yogev, O., Naamati, A. & Pines, O. (2011). *FEBS J.* **278**, 4230–4242.