Journal of Structural Biology 171 (2010) 397–401

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structure Report

Crystal structure of a fructokinase homolog from *Halothermothrix orenii*

Teck Khiang Chua [a], J. Seetharaman [b], Joanna M. Kasprzak [c], Cherlyn Ng [a], Bharat K.C. Patel [d], Christopher Love [d], Janusz M. Bujnicki [c,e], J. Sivaraman [a,*]

[a] *Department of Biological Sciences, 14 Science Drive 4, National University of Singapore, Singapore 117543, Singapore*
[b] *X4 Beamline, Brookhaven National Laboratory, Upton, NY, USA*
[c] *Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, ul. Umultowska 98, 61-614 Poznan, Poland*
[d] *Microbial Gene Research and Resources Facility, School of Biomolecular and Physical Sciences, Griffith University, Brisbane, Queensland 4111, Australia*
[e] *Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Ks. Trojdena 4, PL-02-190 Warsaw, Poland*

ARTICLE INFO

Article history:
Received 29 November 2009
Received in revised form 17 May 2010
Accepted 17 May 2010
Available online 21 May 2010

Keywords:
Hore_18220
Fructokinase
FRK homolog
Fructose
ATP
Structure

ABSTRACT

Fructokinase (FRK; EC 2.7.1.4) catalyzes the phosphorylation of D-fructose to D-fructose 6-phosphate (F6P). This irreversible and near rate-limiting step is a central and regulatory process in plants and bacteria, which channels fructose into a metabolically active state for glycolysis. Towards understanding the mechanism of FRK, here we report the crystal structure of a FRK homolog from a thermohalophilic bacterium *Halothermothrix orenii* (Hore_18220 in sequence databases). The structure of the Hore_18220 protein reveals a catalytic domain with a Rossmann-like fold and a β-sheet "lid" for dimerization. Based on comparison of Hore_18220 to structures of related proteins, we propose its mechanism of action, in which the lid serves to regulate access to the substrate binding sites. Close relationship of Hore_18220 and plant FRK enzymes allows us to propose a model for the structure and function of FRKs.

1. Introduction

Sugar kinases that are responsible for the phosphorylation of free monosaccharides such as glucose and fructose, the initial step of metabolic pathways, are broadly classified into three superfamilies: the galactokinases, hexokinases and ribokinases. Members of the galactokinase family are involved in diverse pathways, ranging from cholesterol and amino acid synthesis to galactose phosphorylation. As sucrose is the major saccharide in plants; two enzymes are responsible for the phosphorylation of sucrose cleavage products fructose and glucose. (Medina and Sols, 1956) Hexokinases (Hxk; EC 2.7.1.1) preferentially phosphorylate glucose and (Frankart and Pontis, 1976) Fructokinase (FRK; EC 2.7.1.4) (a member of the ribokinase-like superfamily according to PFAM and SCOP databases), is a ubiquitous and highly specific enzyme primarily catalyzing the transfer of a phosphate group from adenosine triphosphate (ATP) donor to a phosphate acceptor D-fructose to result in the formation of D-fructose 6-phosphate (F6P).

FRK activity was first reported in 1956, but the protein was isolated and characterized only 20 years later (Medina and Sols, 1956; Frankart and Pontis, 1976). FRK is an enzyme belonging to the ribokinase-like superfamily of sugar kinases that show high substrate specificity. Evolutionary analyses indicate that divergence of the FRK family from a ribokinase ancestor occurred before species divergence (Zhang et al., 2004). FRK specifically phosphorylates fructose with a km of 41–220 μM, at pH 8.0 and have much higher affinities for fructose than Hxk (Renz and Stitt, 1993). As fructose phosphorylation by FRK is irreversible and near rate-limiting, it regulates the rate and localization of carbon usage by channeling fructose into a metabolically active state for glycolysis in plants and bacteria (Zhang et al., 2003). This reaction is particularly important in sink tissues where sucrose assimilation, degradation and conversion to starch is mediated by invertase and/or sucrose synthase (SS), and the fructose produced must be phosphorylated to maintain the carbon flux to starch or respiration. FRKs are widely reported to have a preference for ATP over other nucleotides and unless in the presence of high GTP or UTP concentrations, ATP will be the principle source of phosphate (Chaubron et al., 1995; Martinez-Barajas et al., 1997). The activity of FRK greatly exceeding glucokinase in many tissues is consistent with the view that SS, rather than invertase, is the major route of sucrose degradation, thus producing a larger amount of fructose than glucose.

In the course of a random sequence analysis of the *Halothermothrix orenii* genome, an open reading frame (ORF) encoding a putative FRK enzyme (EC: 2.7.1.4) was identified (Genbank Accession Number: Hore_18220) (Mijts and Patel, 2001). *H. orenii* is an anaerobic, thermohalophilic bacterium from the class Clostridia found in

* Corresponding author. Fax: +65 6779 5671.
E-mail address: dbsjayar@nus.edu.sg (J. Sivaraman).

1047-8477/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2010.05.007

the sediment of a Tunisian salted lake. The strain isolated, H168, oxidizes glucose, fructose, xylose, ribose, cellobiose, and starch (Sigrell et al., 1998). Here we report the crystal structure of Hore_18220, refined to 2.8 Å resolution. Based on the comparative analysis of this structure to related enzymes' structures and mechanisms described in the literature, we propose a model for FRK-ligand interactions.

2. Purification, crystallization and structure determination

Halothermothrix orenii FRK homolog (hereafter referred as Hore_18220) gene was amplified by PCR and ligated into the pTrcHisA expression vector (Invitrogen) at BamHI and KpnI cut sites (Mijts and Patel, 2001). His-tagged Hore_18220 was expressed heterologously in BL21 (DE3). Selenomethionine-substituted protein was expressed using methionine auxotroph *Escherichia coli* DL41 in LeMaster medium supplemented with 25 mg/L selenomethionine (SelMet). Hore_18220 was first purified by Ni–NTA affinity chromatography (Qiagen) and then by gel filtration (Hiload 16/60 Superdex200). Purified Hore_18220 was concentrated to 10 mg/ml.

Crystallization screen was carried out through hanging-drop vapor-diffusion method using Hampton Research (Aliso Viejo, CA, USA) screens as well as by micro batch under-oil technique using JB crystallization screens (Jena Biosciences, Jena, Germany) at room temperature. Initially crystals were small. Obtaining the diffraction quality crystals was the most challenging part of this project. The present data set is the best of many data sets collected. As an approach to improve the data quality, we have also attempted co-crystallization/soaking with the substrates. The best diffraction quality crystals were obtained from 8% PEG 4000, 0.8 M $LiCl_2$ and 0.1 M Tris–HCl, pH 8.5 by using micro batch under-oil technique with 2 μl of the crystallization solution mixed with 2 μl of protein under 15 μl of paraffin oil.

Crystals were directly taken from the drop and the synchrotron data were collected at beam lines X12C and X29, NSLS, Brookhaven National Laboratory for the SelMet/native crystals using Quantum 4-CCD detector (Area Detector Systems Corp., Poway, CA, USA) to 2.8 Å resolution (Table 1). Data were processed and scaled using the program HKL2000 (Otwinowski and Minor, 1997).

All the four expected selenium sites in an asymmetric unit were located by the program SOLVE (Terwilliger and Berendzen, 1999). Initial phases were further developed by RESOLVE (Terwilliger, 2000), which improved the overall figure of merit (FOM) to 0.68 and automatically built ~50%, of the residues of one asymmetric unit. The remaining parts of the molecules were built manually using the program O (Jones et al., 1991). Several cycles of model building alternating with refinement using the program CNS (Brunger et al., 1998) resulted in the final model, with an *R*-factor of 0.254 (R_{free} = 0.288) at 2.8 Å resolution with reflections $I > \sigma I$ was used in the refinement. The final model comprises 278 residues from each monomer and 142 water molecules. The first 21 N terminal residues, His-tag with the linker residues, loop regions comprising residues 144–145 and 267–271, and the C-terminal 21 residues (307–327) were not fully resolved in the electron density map and therefore were not modeled. Procheck (Laskowski et al., 1993) analysis shows that two residues are in the disallowed regions. These residues are found in the turns and are well defined by the electron density map.

3. Overall structure

Hore_18220 crystallized with two molecules in the asymmetric unit related by a 2-fold rotation non-crystallographic symmetry approximately parallel to the *c*-axis. Interestingly, these two molecules are packed one over the other through β-strands, resulting in the formation of a continuous β-sheet between two subunits of the dimer (Fig. 2). Gel filtration and dynamic light scattering experiments indicate that Hore_18220 exists as a dimer in solution (data not shown). This is consistent with the dimeric arrangement observed in the crystal structure.

Each monomer has a mixed α/β fold and a characteristic nucleotide binding domain that resembles the Rossmann fold (Leu22–Ile30, Ser56–Thr108, Ala128–Ile306) (hereafter referred to as the catalytic domain) with a β-sheet "lid" (or lid region) (Leu31–Gly55, Thr109–Glu127). The substrate binding cleft is located at the interface between the catalytic domain and the lid region with a dimension of approximately 18 Å width and 22 Å length (Fig. 1). The core catalytic domain consists of a β-sheet with nine mostly parallel β-strands. This β-sheet is flanked on both sides by eight helices; of which three are very short (1–2 turns). In addition, each monomer contributes four β-strands at the dimer interface to form a tilted antiparallel β-sheet lid that runs from one subunit in the dimer to the other. This β-sheet maintains the dimeric architecture of Hore_18220. The observation of a dimeric molecule in solution as well as in crystal structure suggests a functionally important role for dimerization of Hore_18220.

Table 1
Data collection and refinement statistics.

Data set	Peak	Inflection	Remote[a]
Data collection			
Resolution range (Å)	50.0–2.8	50.0–2.8	50.0–2.8
Wavelength (Å)	0.9790	0.9794	0.9600
Observed reflections >1	117,292	117,115	116,025
Unique reflections	30,770	30,424	30,193
Completeness (%)	99.9	99.9	99.9
Overall ($I/\sigma I$)	19.8	18.9	16.7
R_{sym} (%)[b]	4.6	4.5	4.9
Refinement[c] and quality			
Resolution range (Å)			50.0–2.8
R_{work} (No. of reflections)[d]			0.254 (25238)
R_{free} (No. of reflections)[e]			0.288 (1846)
R.M.S.D. bond lengths (Å)			0.008
R.M.S.D. bond angles			1.72
Average B-factors (Å²)			
Monomer A (all atoms)			46.8
Monomer B (all atoms)			45.7
Ramachandran plot[f]			
Most favored regions (%)			83.9
Additional allowed regions (%)			14.5
Generously allowed regions (%)			0.7
Disallowed regions (%)			0.9

[a] NCS restraint was kept throughout the refinement.
[b] $R_{sym} = |I_i - \langle I \rangle|/|I_i|$ where I_i is the intensity of the *i*th measurement, and $\langle I \rangle$ is the mean intensity for that reflection.
[c] For all models, reflections with $I > \sigma I$ was used in the refinement.
[d] $R_{work} = |F_P - F_{P(calc)}|/F_P$.
[e] *R*-free was calculated with 6% of the reflections in the test set.
[f] Statistics for the Ramachandran plot from an analysis using Procheck (Laskowski et al., 1993).

4. Comparisons with other proteins

A search for structurally similar proteins was performed using DALI (Holm and Sander, 1993). Structures showing overall similarity belonged only to the ribokinase-like superfamily of proteins. The structurally most common feature of these proteins is the substrate binding cleft region. The highest structural similarity is observed between Hore_18220 and a recently deposited (doi:10.2210/pdb3gbu/pdb; but not yet described in the literature) structure of an uncharacterized (putative) sugar kinase PH1459 from *Pyrococcus horikoshii* (PDB code 3ewm) with RMSD of 2.1 Å for 256 Cα atoms, 33% sequence identity and *Z*-score 31.9. This is