BL-1A/2012G555

Crystal structure of membrane protein lipid phosphatase *E.cloi*PgpB

Junping Fan[1], Daohua Jiang[1,2], Yan Zhao [1,3], Xuejun C. Zhang[1*]

[1] National Laboratory of Macromolecules, National Center of Protein Science- Beijing, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing, China 100101

[2] School of Life Science and Technology, Huazhong University of Science and Technology,Wuhan, Hubei, China 430074

[3]School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China230027

Membrane proteins account for about 30% of the genomes sequenced to date and play important roles in a variety of cellular functions. However, determining the three-dimensional structures of membrane proteins continues to pose a major challenge for structural biologists due to difficulties in recombinant expression and purification. Membrane-integrated type II phosphatidic acid phosphatases (PAP2) are important for numerous biological processes from bacterial to human,including glucose transport, lipid metabolism, and signaling. *Escherichia coli* PgpB(ecPgpB) catalyzes removing the terminal phosphate group from a lipid carrier, undecaprenyl pyrophosphate, and is essential for transport of many hydrophilic small molecules across the membrane. We determined the crystal structure ecPgpB at 3.2Å resolution. This structure shares a similar folding topology and a nearly identical active site with soluble PAP2 enzymes.

1 Introduction

Type II phosphatidic acid phosphatases (PAP2) are a large family of phosphatases important for lipid metabolism and signaling. PAP2 proteins have been found in all life kingdoms from bacteria to mammals. They catalyze dephosphorylation of broad substrates by specifically hydrolyzing phosphoric monoester bonds. Their substrates include variety of phosphorylated carbohydrates, peptides, and lipids. PAP2 are involved in vesicular trafficking, secretion and endocytosis (*e.g.* the enzyme APP1 in yeast); protein glycosylation (*e.g.* DOLPP1 in mouse); energy storage, *e.g.* triacylglycerol biosynthesis; and stress response. In contrast to type I PAP enzymes, which are Mg^{2+}-dependent and usually soluble, PAP2 proteins areMg^{2+}-independent, and many of PAP2 enzymes are integral transmembrane (TM) proteins. While the soluble branch of PAP2 is called class A non-specific acid phosphatases (NSAP), the TM branch of PAP2 family is also called lipid phosphatase/phosphotransferase (LPT) family or lipid phosphate phosphatase (LPP) family.Human glucose-6-phosphatase (G6Pase), *the* key enzyme in the homeostatic regulation of blood glucose concentrations, belongs to the TM PAP2 subfamily. Thus,transmembrane property is unique to the PAP2 family.

2 Experiment

The *pgpB* gene (GenBank ID: 313848522) was cloned from the genome of *E. coli* BL21(DE3) and was found to be expressed well in *E. coli* C43 (DE3) strain. Protein purified in β-NG and n-dodecyl-N, N-dimethylamine-N-Oxide (LDAO from Anatrace) (1x CMC each) were used for crystal screening. The best crystal observed diffracted up to 3 Å with severe anisotropy on a synchrotron beamline. After optimization , we got the crystal suitable for dataset collection. Hg-derivative anomalous data and Seleno-L-methionine derivative I116M variant anomalous data were collected at the 1A beamline of the Photon Factory synchrotron facility (KEK, Japan) up to resolution of 3.4 and 2.8 Å, respectively. Combined with other dataset collected at the 41XU beamline of SPring-8 synchrotron facility (Japan) and 17U beamline of Shanghai Synchrotron of Radiation Facility (SSRF), finally, we determined the crystal structure ecPgpB at 3.2Å resolution.

3 Results and Discussion

The 3D structure of PgpB is composed of six TM helices (TMs 1–6) and a small periplasmic domain consisting of 70 amino acid residues (*i.e.* 92–161) (Fig. 1). The N- and C-termini are shown to be located on the cytosol side, and the putative active site is on the periplasmic side as

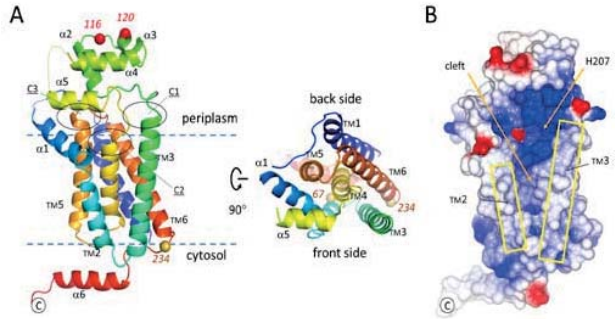

Figure 1.Overall structure of ecPgpB.

predicted before. Overall, the positive-inside ruleof charge distribution is followed, with 10 Arg or Lys residues but no acidic residue being located at the cytosol ends of the TM helices and in the connecting loops (Fig. 1B). TMs 4–6 form the core of the TM region, with TMs 1–3 surrounding the core. TM3 is loosely packed with the rest of the TM domain. Its cytosolic N-terminal end is connected with TM2 through a short 5-residue loop,

while its periplasmic C-terminal end is connect to the periplasmic domain through a rather flexible 10-residue loop. The periplasmic domain is inserted between TMs 3 and 4 and contains four α-helices (*i.e.* α2–5). The putative active site is formed by signature motifs of PAP2, which is located in the primary sequence from the C-terminus of TM3 to the N-terminal end of TM6. In the 3D structure, this active site is located in the membrane-periplasm interface region and is highly positively charged (Fig. 1B). Moreover, TMs 2 and 3 form a V-shaped cleft with the periplasmic opening side being over 10Å wide. This periplasmic opening is located close to the putative catalytic site, and is likely to be the binding site of the polar head of the substrate.

Acknowledgement
The authors would like to thank the staff of the Protein Research Core Facility at the Institute of Biophysics, Chinese Academy of Sciences (CAS), for their excellent technical assistance. We are grateful to staff members of KEK (Japan), SPring-8 (Japan), andSSRF (China) synchrotron facilitiesfor their assistance in collecting diffraction data. This work was supported by Ministry of Science and Technology (China) ('973' Project grant 2011CB910301 to XCZ), by Chinese Academy of Sciences (XDB080203 to XCZ).

References
[1] Fan, J. P. *et al*, PNAS, 2014.

* To whom correspondence should be addressed. E-mail:zhangc@ibp.ac.cn