Biochimica et Biophysica Acta 1864 (2016) 1138–1151

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Role of N-terminal region of *Escherichia coli* maltodextrin glucosidase in folding and function of the protein

Ashutosh Pastor [a], Amit K. Singh [a,1], Prakash K. Shukla [b], Md. Javed Equbal [a], Shikha T. Malik [a], Tej P. Singh [b], Tapan K. Chaudhuri [a,*]

[a] Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, New Delhi 110016, India
[b] Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India

ARTICLE INFO

Article history:
Received 14 January 2016
Received in revised form 10 June 2016
Accepted 14 June 2016
Available online 16 June 2016

Keywords:
Protein folding
GroEL assisted protein folding
Protein structure and function
Maltodextrin glucosidase
X-ray crystallography

ABSTRACT

Maltodextrin glucosidase (MalZ) hydrolyses short malto-oligosaccharides from the reducing end releasing glucose and maltose in *Escherichia coli.* MalZ is a highly aggregation prone protein and molecular chaperonins GroEL and GroES assist in the folding of this protein to a substantial level. The N-terminal region of this enzyme appears to be a unique domain as seen in sequence comparison studies with other amylases as well as through homology modelling. The sequence and homology model analysis show a probability of disorder in the N-Terminal region of MalZ. The crystal structure of this enzyme has been reported in the present communication. Based on the crystallographic structure, it has been interpreted that the N-terminal region of the enzyme (Met1–Phe131) might be unstructured or flexible. To understand the role of the N-terminal region of MalZ in its enzymatic activity, and overall stability, a truncated version (Ala111-His616) of MalZ was created. The truncated version failed to fold into an active enzyme both in *E. coli* cytosol and *in vitro* even with the assistance of chaperonins GroEL and GroES. Furthermore, the refolding effort of N-truncated MalZ in the presence of isolated N-terminal domain didn't succeed. Our studies suggest that while the structural rigidity or orientation of the N-terminal region of the MalZ protein may not be essential for its stability and function, but the said domain is likely to play an important role in the formation of the native structure of the protein when present as an integral part of the protein.

1. Introduction

The protein Maltodextrin glucosidase (EC no. 3.2.1.20, Uniprot ID: P21517), expressed by malZ gene, is an enzyme involved in the maltose utilization system in *Escherichia coli* [1]. The major function of this enzyme is to catalyse hydrolysis of short malto-oligosaccharides ranging from maltotriose to maltoheptose, while releasing glucose from the reducing end. The final end products of this reaction are glucose and maltose [2]. The enzyme is known to release maltose directly from maltodextrins longer than maltotriose to a small extent [3]. MalZ can also hydrolyze γ-cyclodextrins containing eight glucosyl residues, however, γ-cyclodextrins are not physiological substrates for this enzyme as they are not transported by *E. coli* [3]. MalZ has been demonstrated to show *trans*-glycosylase activity which leads to formation of branched oligosaccharides as products [4]. The alternate name commonly used for such enzyme is alpha-glucosidase. MalZ is a cytosolic enzyme in *E. coli*, and is considered to play a role in regulating the intracellular level of maltotriose, which induces *mal* regulon. Overexpression of malZ results in decreased expression of other genes in the *mal* regulon [2]. In *E. coli*, MalZ is known to counteract the formation of long maltodextrins, which are formed by action of amylomaltase. The mutants where malZ is knocked out produce large amounts of glycogen using maltodextrins as a source of carbon in absence of glycogen synthase [5,6]. MalZ belongs to the α-amylase family of proteins. The α-amylases are known to play active role in starch metabolism by hydrolysing the α-1,4 and α-1,6 glycosidic linkages as well as performing *trans*glycosylation and hydrolysis of cyclodextins. These enzymes contain an $(\alpha/\beta)_8$ tim-barrel structure which serves as the catalytic domain [7]. While the α-amylases are primarily involved in starch metabolism, the enzymes showing sequence similarity to MalZ are known to be cytosolic enzymes specific to smaller malto-oligosaccharides and play active roles in maltose metabolism in bacterial cells [8,9]. The microbial

Abbreviations: MalZ, Maltodextrin glucosidase; MalZ-N^{Trunc}, MalZ protein without N-Terminal domain; $MalZ^{NTD}$, N-terminal domain of MalZ; GdnHCl, Guanidine Hydrochloride; IPTG, Isopropyl β-D-1-thiogalactopyranoside.

* Corresponding author.
E-mail addresses: tkchaudhuri@bioschool.iitd.ac.in, tapanchaudhuri@hotmail.com (T.K. Chaudhuri).
[1] Present address: Max Plank Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany.

http://dx.doi.org/10.1016/j.bbapap.2016.06.008
1570-9639/

α-amylases have been studied extensively for their importance in various industrial applications like food industry, starch industry, textile, detergent, pharmaceutical and paper industry [10].

Maltodextrin glucosidase is also found in other organisms including *Salmonella enterica*, *Shigella flexneri*, *Yersinia pestis* and others. The crystal structure of this protein has not been reported in any of these organisms. On the basis of sequence similarity the closest structure known to this protein is alpha-amylase 1 from *Thermoactinomyces vulgaris* [11] having sequence identity of 34% with a sequence coverage of 42-591 amino acids out of 616 total amino acids. The protein in *E. coli* consists of 604 amino acids, two amino acids at N-terminal and 10 amino acids including 6× His-tag at C-terminal appear in the studied protein making the total number of amino acids to 616. However, no effect of these extra amino acids has been observed on the function of the protein.

It has been reported that MalZ protein is highly aggregation prone and requires assistance of chaperonins GroEL and GroES for its folding when over-expressed in *E. coli* [12]. The co-expression of these chaperonins with MalZ in the *E. coli* cells increases the soluble and functional fraction of MalZ protein in the cytosol by a very significant amount. GroEL and GroES usually help the aggregation prone proteins to fold by encapsulating the protein in the cylindrical cavity of GroEL which is capped by GroES to form a closed chamber [13]. However, MalZ is a large protein having a molecular weight of 70 kDa and cannot be encapsulated within the GroEL-GroES cavity, which can encapsulate protein of size up to ~50 kDa. GroEL assisted folding through a *trans* mechanism [14] has been suggested for the folding of this protein. In this case, it is expected that only a part of the protein might be entering the GroEL cavity and GroES does not encapsulate it, but helps in the release of the protein from the cavity by binding to the *trans* ring of GroEL [15]. The role of any specific part of the MalZ protein in this aggregation prone behaviour and requirement of GroEL and GroES for assisted folding was not clear and the present study is an attempt to answer this question.

The initial sequence analysis and homology model comparisons made it interesting to study the N-terminal of MalZ. However in order to get precise information on its uniqueness among amylase type of proteins and also to understand the reasons for the aggregation prone behaviour and chaperone dependency for folding, structure determination of MalZ using X-ray crystallography was attempted. The crystal structure of MalZ was solved at 3.7Å and has been reported here. The structure of MalZ obtained from crystallography data suggested that the N-terminal region of MalZ (Met1–Phe131) may either be flexible or unstructured. This was interesting as the sequence comparison and the homology model data also showed that the N-terminal might be a distinct entity. To find out whether this N-terminal region actually plays any important role in the folding and structural stability of the MalZ protein the N-terminal region (Met1-Phe110) was truncated, and the results suggest that this domain is essentially required within the protein for stabilizing the structure of MalZ in its native functional form and the N-terminal domain does not help to fold MalZ when used *in trans* through co-expression in *E. coli* or *in vitro*.

2. Materials and methods

2.1. Sequence and homology model analysis

The amino acid sequence of the protein used for our studies was compared to the Uniprot database using NCBI BLAST [16] and the specific domains and their families were identified based on the sequence. Further the homology model was prepared using the I-Tasser server [17] which combines template based prediction with some simulation based refinement. The visualization software Pymol was utilized for visualization and comparisons. The sequence was also analysed using a web server MobiDB [18] for disorder prediction which combines results from multiple servers to generate a consensus on disordered regions.

2.2. MalZ expression and purification

Plasmid PCS19malZ containing the 1.8 Kb malZ gene under the control of a T5 promoter and ampicillin resistance marker was a generous gift from Prof. Winfried Boos, (University of Konstanz, Germany). The malZ gene is cloned between sites of restriction endonucleases *Nco*I and *Bam*HI and it forms a C-terminal 6× His tag from the PCS19 vector. *E. coli* BL-21 cells containing plasmid PCS19malZ were grown in Luria Bertani medium with 50 μg/ml ampicillin to $O.D._{600}$ of 0.6. Induction was done by 1 mM IPTG. Post induction, the culture was kept at 37 °C for 8 h. The cells from 1 l culture were harvested by centrifugation at 7500*g* for 15 min and temperature was maintained at 4 °C from this step onwards. The pellet was re-suspended in 30 ml lysis buffer consisting of 20 mM sodium phosphate pH 7.4, 500 mM sodium chloride, 1 mg/ml lysozyme and 1 mM PMSF. Cell lysis was carried out by sonication with a Branson Sonifier250 (USA). 15 cycles of 25 s each were given with 1 min rest phase. AKTA Purifier (GE Healthcare, UK) was used for the purification process. The supernatant was loaded on a 5 ml His-Trap column (GE Healthcare, UK) pre-equilibrated with the equilibration buffer containing 20 mM sodium phosphate pH 7.4 and 500 mM sodium chloride. The column was washed extensively by same buffer containing 20 mM imidazole for removal of impurities. Resin bound protein is eluted by a linear gradient up to 500 mM imidazole in 10 column volumes. MalZ is eluted at imidazole concentration of around 200 mM. Fractions eluted were found to contain single bands on SDS PAGE gels stained with Coomassie blue. Protein concentration was measured using Bradford assay (Bio-rad, USA).

2.3. Crystallization of MalZ, data collection and processing

The pure fractions of MalZ were concentrated with amicon centrifugal concentrators and in this process, the salt and imidazole were also reduced. The protein was concentrated to 8 mg/ml and the final buffer composition was 20 mM sodium phosphate, 200 mM Sodium chloride and 50 mM imidazole. The MalZ solution was mixed with reservoir solution 2 μl each, which contained 2.5 M ammonium acetate and 0.1 M sodium acetate at pH 4.6. This solution was used for crystallization using the hanging drop vapour diffusion method. The crystallization set ups were kept at room temperature (298 K). The crystals grew to an approximate size of $0.4 \times 0.1 \times 0.1$ mm^3 with sharp rectangular faces after 10 days. The crystals were stabilized in 2.5 M ammonium acetate and 0.1 M sodium acetate containing 20% Glycerol (v/v) for data collection at low temperatures. A single crystal was mounted in a nylon loop and flash - frozen in a stream of nitrogen gas at 100 K. The data were collected on a MAR CCD-225 Scanner (Marresearch, Norderstedlt, Germany) using the beamline, BM14 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France.

2.4. Structure determination and refinement

Structure of MalZ was determined with molecular replacement method using auto-AMoRe [19] from the CCP4 software suite (Collaborative Computational Project, Number 4, 1994). The coordinates of the structure of α-amylase from *T. vulgaris* (PDB code: 1JI1,) [11] were used as search model. The rotation and translation functions calculated with data in the resolution range of 110.6–3.7 Å yielded a unique solution with the first peak being very distinct. The packing arrangement of the molecules in the unit cell for this solution gave no unfavourable intermolecular contacts in space group P4 thus confirming it to be the correct space group and the correct solution. These coordinates were transformed using AMoRe and were subjected to 20 cycles of rigid-body refinement with REFMAC5 [20] The manual model building of the protein was carried out using $|2F_o - F_c|$ Fourier and $|F_o - F_c|$ difference Fourier maps with Graphics Program 'O' [21] on a Silicon Graphics O2 Workstation. The structure was partially refined due to limitation of reflection data (Table 1) to values of 0.33/0.34 for R_{cryst}/R_{free} factors.