Cloning and characterization of trehalase: a conserved glycosidase from oriental midge, *Chironomus ramosus*

Ekta Shukla^{1,2,5} · Leena Thorat³ · Ameya D. Bendre^{1,2} · Santosh Jadhav^{4,6} · Jayanta K. Pal^{4,7} · Bimalendu B. Nath³ · Sushama M. Gaikwad^{1,2}

Received: 8 March 2018 / Accepted: 27 July 2018 / Published online: 1 August 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Insect trehalase is a multifarious enzyme, crucial for normal physiological functions as well as under stress conditions. In this report, we present a fundamental study of the trehalase gene segment (1587 bp) from *Chironomus ramosus* (CrTre) encoding for 529 amino acids, using appropriate bioinformatics tools. *C. ramosus*, a tropical midge is an emerging animal model to investigate the consequences of environmental stresses. We observed that CrTre belongs to GH family 37 in the CAZy database and possess 57–92% identity to dipteran trehalases. In silico characterization provided information regarding the structural, functional and evolutionary aspects of midge trehalase. In the phylogenetic tree, CrTre clustered with the soluble dipteran trehalases. Moreover, domain functional characterization of the deduced protein sequence by InterProScan (IPR001661), ProSite (PS00927 and PS00928) and Pfam (PF01204) indicated presence of highly conserved signature motifs which are important for the identification of trehalase superfamily. Furthermore, the instability index of CrTre was predicted to be <40 suggesting its in vivo stability while, the high aliphatic index indicated towards its thermal stability (index value 71–81). The modelled 3D tertiary structure of CrTre depicts a (α/α)₆ barrel toroidal core. The catalytic domain of the enzyme comprised Glu424 and Asp226 as the putative active site residues. Interestingly, the conserved motifs were observed to be formed by the flexible loopy regions in the tertiary structure. This study revealed essential sequence features of the midge trehalase and offers better insights into the structural aspects of this enzyme which can be correlated with its function.

Keywords Trehalose · Midges · Conserved motifs · Homology modelling · Secondary structure

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1376-y>) contains supplementary material, which is available to authorized users.

✉ Ekta Shukla
ektaashukla19@gmail.com; e.shukla@ncl.res.in

¹ Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune 411008, India

² Academy of Scientific and Innovative Research (AcSIR), CSIR-NCL Campus, Pune 411008, India

³ Stress Biology Research Laboratory, Department of Zoology, S.P. Pune University, Pune 411007, India

⁴ Cell and Molecular Biology Laboratory, Department of Biotechnology, S.P. Pune University, Pune 411007, India

⁵ Present Address: National Centre for Cell Science, Ganeshkhind, Pune 411007, India

⁶ Present Address: BAIF Central Research Station, Urulikanchan, Pune 412202, India

⁷ Present Address: Dr. D.Y. Patil Biotechnology and Bioinformatics Institute, Pune 411033, India

Abbreviations

CrTre	<i>C. ramosus</i> trehalase
dNTP	Deoxy nucleotide triphosphate
MEGA	Molecular evolutionary genetics analysis
min	Minutes
NCBI	National Centre for BIotechnology
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point

Introduction

In insects, trehalose is the major hemolymph sugar and its hydrolysis by the enzyme trehalase serves as the prime energy source (reviewed in depth by Shukla et al. 2015). Trehalase is a glycosidase responsible for the irreversible breakdown of trehalose into two glucose units and plays a fundamental role in forming a link between trehalose metabolism and glycolysis (Reyes-DelaTorre et al. 2012). The role

of insect trehalose and trehalase has been documented in growth and development, flight, chitin biosynthesis and recovery from abiotic stresses such as environmental dehydration bouts. In particular, during recovery from desiccation stress in insects, there is a noticeable increase in trehalase activity (Thorat et al. 2012); suggesting that trehalase is one of the stress-recovery biomolecules along with some repair enzymes (Shukla et al. 2015 and references therein). More recently, trehalase is gaining prominence in contemporary studies and significance in insect pest management due to its crucial role in carbon metabolism and enantiostasis (D'Adamio et al. 2015, 2018). Trehalase is highly specific to trehalose and hence is competitively inhibited by most glycosidase inhibitors. Thus, several trehalose mimetics and analogs have been proposed as potential insecticides (Bini et al. 2012; Adhav et al. 2018), placing trehalase in the list of biotechnologically important enzymes.

Trehalase in insects has been shown to exist in two distinct forms, Tre-1 as the soluble form, while Tre-2 being the membrane-bound form. It has been reported that Tre-1 is mainly present in hemolymph and cavity of goblet cells in midgut. Tre-2 has been shown to be membrane bound to various cells like flight muscle cells, follicle cells, ovary cells, etc. (for details see Shukla et al. 2015). Functionally similar, the two forms basically differ in localization, size and kinetic properties.

Trehalase is an inverting glycosidase which yields products with opposite stereochemistry to the substrate (Davies and Henrissat 1995). Thus, trehalose being composed of two α glucose units is split into one α and one β glucose molecule upon hydrolysis by trehalase. The latest classification system of glycoside hydrolases (GHs) based on their amino acid sequence similarity is known as carbohydrate-active enzymes (CAZy) (Lombard et al. 2014). Currently, GHs constitute 153 protein families in CAZy database wherein trehalase pertains to families 15, 37 and 65. The families GH15 and GH65 together form the clan GH-L, whereas the family GH37 is a member of the clan GH-G along with families GH63, GH100 and GH125. Both the clans GH-G and GH-L adopt the $(\alpha/\alpha)_6$ barrel fold for their catalytic domain as well as all members (including trehalases) employ the inverting mechanism (Terrapon et al. 2017; The CAZy Consortium 2018).

We here report molecular cloning and characterization of a cDNA segment which encodes trehalase from an oriental midge, *Chironomus ramosus*. Among the oriental aquatic midges, *C. ramosus* is emerging as a potential, representative animal model system to address several biological and environmental issues (Thorat and Nath 2015, 2016; Thorat et al. 2017). The present study introduces a new variant to the family of dipteran trehalases and helps in understanding its structural, functional and evolutionary aspects. In silico study of its biochemical features, sequence analysis

and phylogenetic analysis were performed using bioinformatics tools. A three-dimensional structure of the enzyme was modelled wherein the conserved motifs and catalytic domain which are likely to have an impact on its function, were identified.

Materials and methods

Cloning of *C. ramosus* trehalase (CrTre)

Mid-third instar larvae from inbred populations of laboratory reared *C. ramosus* were used for total RNA isolation using innuPREP RNA Mini Kit (Analytik Jena, Thuringia, Germany) followed by cDNA synthesis using High-Capacity cDNA Reverse Transcription Kit (Life Technologies, CA USA). Specific internal primers, based on *Chironomus riparius* trehalase sequence (Genbank Accession No. HQ231258), were designed against the conserved regions (Xcelris Genomics, India) (S1). The amplified PCR products were purified, cloned into pGEM-T Easy vector (Promega, Madison) and sequenced (GeneOmbio Technologies, India). 5' and 3' RACE was carried out using SMARTer™ RACE cDNA amplification kit (Clontech Laboratories, Inc., USA) as per manufacturer's instructions.

Sequence and phylogenetic analysis

The cloned gene sequence was subjected to homology search with NCBI BLAST and alignment was done using Clustal omega. The CrTre cDNA sequence was conceptually translated with the Translate tool (<http://www.expasy.org/translate/>) and subjected to protein domain functional analysis by Pfam v.29.0 (Finn et al. 2016), InterProScan 5 (Jones et al. 2014), Scan PROSITE tool (prosite.expasy.org) and PANTHER gene ontology (<http://www.pantherdb.org/tools/hmmScoreForm.jsp>). Physical properties of translated CrTre were obtained by ProtParam tool from ExPasy (<http://www.expasy.org/protparam/>). The transmembrane domain of protein and N-linked glycosylation sites were analyzed using TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) and NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>), respectively. Phylogenetic tree was constructed using a total of 35 insect trehalase protein sequences obtained from NCBI and assembled to a neighbor-joining tree using MEGA 6 (Tamura et al. 2013).

Homology modelling and active site prediction

The three-dimensional structure prediction of CrTre was based on the homology model generated using MODELLER 9.12 (Webb and Sali 2014). *E. coli* trehalase crystal structure (PDB ID: 2JG0, resolution 1.5 Å) (Gibson et al.

2007) was used as the template with sequence identity of 35% and 84% sequence coverage. The model was subjected to energy minimization and refinement using Prime suite of Schrödinger version 15.3 (2016). The model was checked on PDBSum (Laskowski 2001) for quality and assessed by Ramachandran plot. Secondary structural elements of CrTre were analyzed using PDBSum and iRDP server (Panigrahi et al. 2015). Key catalytic residues involved in the trehalase active site were identified based on the sequence alignment and active centre similarity with *E. coli* trehalase structure and HMMScan server (<https://www.ebi.ac.uk/Tools/hmmer/>).

Results and discussion

Cloning of CrTre

Although the trehalase coding sequences diverge at the N- and C-termini, there occur few fairly conserved internal regions along with two highly conserved signature tags (Barraza and Sánchez 2013). These regions from *C. riparius* trehalase were used to design specific primers. The primers (S1, S2) allowed the amplification of a 916 bp fragment by PCR which was extended at both ends by 5' and 3' RACE (S1, S2). Assembly of the overlapping fragments yielded 1587 bp cDNA which was submitted to GenBank under accession number KX857662 (Fig. 1). Even though, CrTre is devoid of 135 bp upstream when compared with *C. riparius* trehalase (Forcella et al. 2012), it represents all the conserved motifs and catalytic domain, exhibiting high similarity to known insect trehalase sequences.

Primary structure characterization

The translated protein sequence of CrTre was aligned with those of other dipteran trehalases was found to be 57–92% identical to other trehalase sequences (S3, S4). The primary structure of *C. ramosus* trehalase showed several peculiar features common to insect trehalases, such as presence of two signature motifs along with five conserved regions and a glycine-rich region. Details of different sequence features of *C. riparius* trehalase and CrTre have been summarized in S5. Furthermore, few reports demonstrate binding of trehalase to concanavalin A affinity chromatographic column indicating that trehalase is a glycoprotein (Ogiso et al. 1985; Forcella et al. 2010). Four putative N-linked glycosylation sites were identified in CrTre sequence (Fig. 1).

Biochemical properties of the translated protein sequence were predicted using ProtParam tool which were in accordance with the available preliminary information on other insects (summarized by Shukla et al. 2015). The molecular weight of insect trehalase ranges from 60 to 140 kDa. Tre-2

possesses a transmembrane domain and is usually larger. We observed that the theoretical molecular weight of soluble trehalases from different insects varied in the range of 60–68 kDa and pI was near pH 5 (S6). The instability index and aliphatic index were used to predict in vivo half-life and thermostability of the protein, respectively. A protein with instability index smaller than 40 is predicted to be stable (Rogers et al. 1986). The values for dipteran trehalases ranged between 30 and 40; suggesting that these proteins exhibit good in vivo stability. Further, the high aliphatic index (AI) of proteins serves as a measure of thermostability of proteins (Ikai 1980). AI of reported trehalase sequences (lies between 70 and 80) indicated towards the thermostability of these enzymes. Notably, these indices indicate the stability of trehalase at high temperatures which could possibly be associated with its role in dehydration stress-recovery phase.

Phylogenetic analysis and functional prediction

The phylogenetic tree, constructed using trehalase protein sequences from different insect species, displayed two distinct clusters where, soluble trehalases were separated from membrane-bound forms (S7). CrTre grouped with dipteran trehalases in the tree and showed close proximity with soluble trehalases. Since CrTre protein sequence confirms absence of transmembrane domain at C terminal, it cannot be categorized as membrane-bound form and thus, branches out from the soluble cluster.

The protein sequence was subjected to protein functional analysis by different web programmes and tools which classify proteins into families and predict the presence of important domains and sites (S8). InterProScan analysis (IPR001661) and Pfam domain (PF01204) confirmed the identity of CrTre to 'trehalase superfamily'. ProSite scan identified two signature motifs 1 and 2 (PS00927 and PS00928) which are highly conserved across the trehalase family. Additionally, Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system recognised the protein as glycoside hydrolase, family 37 (PTHR23403).

Tertiary structure prediction

The 3-D structures of enzymes often provide valuable insights into molecular geometry and active centre identification. The stereochemical qualities of the CrTre model, assessed by Ramachandran plot were in agreement with the requirements for preferred and allowed regions (S9). The model revealed an α -toroidal architecture characterizing the $(\alpha/\alpha)_6$ barrel glycosidase fold in midge trehalase (Fig. 2a). Similar arrangement of helices was described in the trehalase crystal structure from *E. coli* (Gibson et al. 2007) and also in

1 ctagaaacaatccaactagctcatctttacgctgactccaaaacattcgtagacatgaaa
 L E T I Q L A H L Y A D S K T F V D M K
 61 cttaaacaatctccagccaagactcttgccatatttgatgaattcatgtccaaaatcca
 L K Q S P A K T L A I F D E F M S K N P
 121 gatccaagtgccgaagatcttaaaaaatgggtcgatgaaaacttcgagctacctggatct
 D P S A E D L K K W V D E N F E L P G S
 181 gaatttgaaaaatggatacctgacgatttcacaaaaaaccagcagtattaaataagatt
 E F E K W I P D D F T K N P A V L N K I
 241 gccgacaaaagtttccgtagctttgcatatgatttgaatggaatttggaatgatttgga
 A D K S F R S F A Y D L N G I W N D L G
 301 agaaagatcaggaagaagttcatgacaatgaggaattgtattcaattattttatgttgag
 R K I R K E V H D N E E L Y S I I Y V E
 361 aatccaatgattgtgcctgggtggaagatttagagaattttattactgggattcttattgg
 N P M I V P G G R F R E F Y Y W D S Y W
 421 gttcttcgtggactttttattgagtgaatgactcataccgaaaaggaatgctcgagaac
 V L R G L L L S E M T H T A K G M L E N
 481 ttcttctcaattgttcttcgttttggtttcattcctaattggaggcagagtctactattca
 F F S I V L R F G F I P N G G R V Y Y S
 541 atgcgttcacaaccaccacttttgattccaattggtcaagacttatgttgacagtacaaaa
 M R S Q P P L L I P M V K T Y V D S T K
 601 gattttgaatttgagtaaaatcagtcgatgttcttgcccttagaatttgactactggatg
 D F E F A V K S V D V L A L E F D Y W M
 661 acaaatcatacaacaatgattaaaggtcacaaaatggcattttatggagacaaatcatca
 T N H T T M I K G H K M A F Y G D K S S
 721 ggtccacgtccagaatcataccgtgaagatattgaaacagcaaatggattcaaaacagat
 G P R P E S Y R E D I E T A N G F K T D
 781 gaagacaaagaggaacattattcagaattaaaggcagcagctgagagtggaaatggacttt
 E D K E E H Y S E L K A A A E S G M D F
 841 tcattccagatgggtttattgatgctgaagggtggaataatggaactcttgccaatttgaag
 S S R W F I D A E G G N N G T L A N L K
 901 actagatcgattattccagttgaattgaatgccattcttcattggaatgcaaaaattatt
 T R S I I P V E L N A I L H W N A K I I
 961 gctgaattttatggatatgctggcaacattactaaacaaactgaatatttgcaaatgct
 A E F Y G Y A G N I T K Q T E Y L Q I A
 1021 aagcagttcttggtatgctgtcaacgaagtcctctgggatgaagaaacaggagcatggctt
 K Q F L D A V N E V L W D E E T G A W L
 1081 gactatgacaccattaatcagaaaaagcgaccatactttgttccaacaaatctttcacct
 D Y D T I N Q K K R P Y F V P T N L S P
 1141 ctatggatgctgtctatgacgtcacaaagcgtgtacacattgcatctaaagttatggcc
 L W M R C Y D V T K R V H I A S K V M A
 1201 tacatcaacaaaacagactcgatgactaccctgggtgtgtaccaacaacccttgtcaac
 Y I N K T R L D D Y P G G V P T T L V N
 1261 agtgggtgagcaatgggattggcctaattgtatgggcaccaatgcagcatatgttaattgtt
 S G E Q W D W P N V W A P M Q H M L I V
 1321 ggacttgataatcttggaacaaaggagtcacaagagaaggcccaagattgggcacaaaga
 G L D N L G T K E S Q E K A Q D W A Q R
 1381 tgggtgttgagtaattatttagcatataaggagactggtaacatgtatgagaagtatatg
 W V L S N Y L A Y K E T G N M Y E K Y M
 1441 gcaacagaactaggtggacatggaggtgggtggtgaatatgaagtacaaaaaggctttggc
 A T E L G G H G G G G E Y E V Q K G F G
 1501 tggacaaatggagttatttttagatctcttggtgactatggacacgttctcagctcatca
 W T N G V I L D L L D H Y G H V L S S S
 1561 ggttcgaatttaattatttgaagcac
 G S N L N Y L K H

Fig. 1 Nucleotide and amino acid sequence of *C. ramosus* trehalase. Trehalase signature motifs (residues 126–139 and 424–433) are boxed. Conserved motifs identified from alignment (residues 12–20, 171–179, 182–188, 241–251, 269–286 and 488–494) are underlined. Four putative *N*-linked glycosylation sites (residues 222, 293, 329 and 403) are encircled. The nucleotide sequence reported in this article is submitted to GenBank under accession number KX857662

the homology models reported from two dipteran insects, *C. riparius* (Forcella et al. 2012) and *D. melanogaster* (Shukla et al. 2016) (S10). Thus, despite the differences at sequence level, the structure of enzyme remains conserved in different organisms. The secondary structural elements of CrTre were obtained by PDBSum and iRDP server (S11). The values supported the α -helical nature of the enzyme, as visualised by the homology model.

Furthermore, the percentage richness of amino acid showed the presence of leucine and glutamic acid-rich profile for insect trehalases (S6). It has been observed that amino acids, namely, Glu, Ala, Leu and His occur in high percentages in the helical regions of a protein with Leu

being the most abundant residue present in the inner helical cores of α helix-rich proteins (Chou and Fasman 1973; Kumar and Bansal 1998).

Active site In the CAZy database, insect trehalases mostly fall under GH family 37 which possess Glu and Asp as the nucleophile and proton donor, respectively (Lombard et al. 2014). In *E. coli* trehalase structure, key catalytic residues involved in the substrate anomeric inversion mechanism are reported as Asp 312 and Glu 496 (as acid and base, respectively), (Gibson et al. 2007). By analogy with the *E. coli* Tre37A, we identified Asp 279 and Glu 477 as putative acid and base catalytic residues in CrTre model (Fig. 2b). HMMscan results from Pfam database of protein families also predicted D279 and E477 as the active site residues. Moreover, Silva et al. (2010) demonstrated involvement of three Arg residues in the catalysis by site-directed mutagenesis in insect trehalases. These three putative Arg residues form the active centre of the enzyme along with the acidic and basic residues (Fig. 2b). Additionally, there is a highly conserved glycine-rich loop located beyond the catalytic

Fig. 2 Homology modelling of *C. ramosus* trehalase: **a** Ribbon representation of *C. ramosus* trehalase (CrTre) showing an α -barrel toroidal core; **b** active centre pocket of CrTre showing three Arg along

with putative catalytic residues (Glu and Asp); **c** PDBsum wiring diagram showing conserved regions mostly contributed by the loop regions. **a** and **b** were drawn using PyMOL

site, which is presumed to provide assistance during the catalysis reaction (Brayer et al. 1995; Ramasubbu et al. 2003). However, recently it was shown that this loop is not directly involved in catalysis as its absence does not inhibit the activity of the enzyme (Shukla et al. 2016).

Conserved regions formed by loops Trehalase is an evolutionarily conserved enzyme (Barraza and Sánchez 2013). Conserved regions in proteins are important both structurally and functionally. A common notion is that the helices and strands are more conserved than loops and turns, typically because they have regular repetitive structural pattern (Liu et al. 2002). Contrarily, we observed that in CrTre, the signature motifs, conserved regions and active centre are formed majorly of loops (Fig. 2c). Trehalase is an all- α member; hence we can observe some part of helices in the conserved regions. Loops have a tendency to be found in the conserved protein regions as well as in random non-conserved regions. Possibly, loops important for proper structure and function of a protein are conserved during evolution, as in case of trehalase; while, those not necessary in structure stabilization and function may not be conserved (Sitbon and Pietrovski 2007).

Conclusion

Despite the diversity observed in sequence composition and amino acid contribution, trehalase remains the functionally and structurally conserved protein, throughout the phyla. The conservation of the catalytic domain in varied organisms from bacteria to mammals further suggests that the trehalose hydrolysis has been preserved to a single mechanism during evolution. However, it must be noted that there is a significant difference in the substrate affinity and degree of inhibition between insect and mammalian trehalases, which could be exploited to develop novel pesticides. This report will, therefore, provide basis for future focus on the experimental validation of properties of this evolutionarily old and biotechnologically important enzyme.

Acknowledgements This work is supported by funding received from DRDP programme, Department of Biotechnology, SPPU and DST-PURSE grants to JKP and from partial funding received from DST-PURSE and BCUD-UoP grants to BBN. The authors thank Dr. Varsha Bhatia, Gennova Biopharmaceuticals Ltd. and Mr. Ejaj Pathan, CSIR-NCL for their valuable suggestions and timely help. ES and ADB acknowledge University Grants Commission, New Delhi, for Senior Research Fellowships. LT is grateful for financial support received from the University Grants Commission-DS Kothari Postdoctoral Fellowship (UGC-DSK-PDF) and from DBT Bio-CARe Women Scientist Scheme, New Delhi, India.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest in the publication.

References

- Adhav AS, Kokane SR, Joshi RS (2018) Functional characterization of *Helicoverpa armigera* trehalase and investigation of physiological effects caused due to its inhibition by Validamycin A formulation. *Int J Biol Macromol* 112:638–647
- Barraza A, Sánchez F (2013) Trehalases: A neglected carbon metabolism regulator. *Plant Signal Behav* 8:24778
- Bini D, Cardona F, Forcella M, Parmeggiani C, Parenti P, Nicotra F, Cipolla L (2012) Synthesis and biological evaluation of nojirimycin and pyrrolidine-based trehalase inhibitors. *Beilstein J Org Chem* 8:514–521
- Brayer GD, Luo Y, Withers SG (1995) The structure of human pancreatic α -amylase at 1.8 Å resolution and comparisons with related enzymes. *Protein Sci* 4:1730–1742
- Chou PY, Fasman GE (1973) Structural and functional role of leucine residues in proteins. *J Mol Biol* 74:3:263–281
- D'Adamio G, Sgambato A, Forcella M, Caccia S, Parmeggiani C, Casartelli M, Parenti P, Bini D, Cipolla L, Fusi P, Cardona F (2015) New synthesis and biological evaluation of uniflorine A derivatives: towards specific insect trehalase inhibitors. *Org Biomol Chem* 13(3):886–892
- D'Adamio G, Forcella M, Fusi P, Parenti P, Matassini C, Ferhati X, Vanni C, Cardona F (2018) Probing the influence of linker length and flexibility in the design and synthesis of new trehalase inhibitors. *Molecules* 23(2):436
- Davies G, Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* 3:853–859
- Finn RD, Cogill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285
- Forcella M, Cardona F, Goti A, Parmeggiani C, Cipolla L, Gregori M, Schirone R, Fusi P, Parenti P (2010) A membrane-bound trehalase from *Chironomus riparius* larvae: purification and sensitivity to inhibition. *Glycobiology* 20(9):1186–1195
- Forcella M, Mozzi A, Bigi A, Parenti P, Fusi P (2012) Molecular cloning of soluble trehalase from *Chironomus riparius* larvae and its heterologous expression in *Escherichia coli*. *Arch Insect Biochem Physiol* 81(2):77–89
- Gibson RP, Gloster TM, Roberts S, Warren AJ, Storch de Gracia I, García A, Chiara JL, Davies GJ (2007) Molecular basis for trehalase inhibition revealed by the structure of trehalase in complex with potent inhibitors. *Angew Chem Int Ed Engl* 46:4115–4119
- Ikei A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* 88:1895–1898
- Jones P, Binns D, Chang H, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240
- Kumar S, Bansal M (1998) Geometrical and sequence characteristics of α helices in globular proteins. *Biophys J* 75:1935–1944
- Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res* 29(1):21–222

- Liu J, Tan H, Rost B (2002) Loopy proteins appear conserved in evolution. *J Mol Biol* 322:53–64
- Lombard V, Golaconda RH, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495
- Ogiso M, Shinohara Y, Hanaoka K, Kageyama T, Takahashi SY (1985) Further purification and characterization of trehalase from the American cockroach *Periplaneta americana*. *J Comp Physiol B Biochem Syst Environ Physiol* 155:553–560
- Panigrahi P, Sule M, Ghanate A, Ramasamy S, Suresh CG (2015) Engineering proteins for thermostability with iRDP web server. *PLoS One* 5(10):10
- Ramasubbu N, Ragunath C, Mishra PJ (2003) Probing the role of mobile loop in substrate binding and enzyme activity of human salivary amylase. *J Mol Biol* 325:1061–1076
- Reyes-DelaTorre A, Peña-Rangel MT, Riesgo-Escovar JR (2012) Carbohydrate metabolism in *Drosophila*: reliance on the disaccharide trehalose. In: Chang C (ed) *Carbohydrates-comprehensive studies on glycobiology and glycotechnology* (online). InTech, Winchester, pp 317–338
- Rogers S, Wells R, Rechsteiner M (1986) Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. *Science* 234:364–368
- Schrödinger (2016) Release 2016-1: LigPrep, version 3.7, Schrödinger. LLC, New York
- Shukla E, Thorat LJ, Nath BB, Gaikwad SM (2015) Insect trehalase: Physiological significance and potential applications. *Glycobiology* 25(4):357–367
- Shukla E, Thorat LJ, Bhavnani V, Bendre AD, Pal JK, Nath BB, Gaikwad SM (2016) Molecular cloning and in silico studies of physiologically significant trehalase from *Drosophila melanogaster*. *Int J Biol Macromol* 92:282–292
- Silva MCP, Terra WR, Ferreira C (2010) The catalytic and other residues essential for the activity of the midgut trehalase from *Spo-doptera frugiperda*. *Insect Biochem Mol Biol* 40:733–741
- Sitbon E, Pietrokovski S (2007) Occurrence of protein structure elements in conserved sequence regions. *BMC Struct Biol* 7:3
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Terrapon N, Lombard V, Drula E, Coutinho PM, Henrissat B (2017) The CAZy database/the carbohydrate-active enzyme (CAZy) database: principles and usage guidelines. In: Aoki-Kinoshita K (ed) *A practical guide to using glycomics databases*. Springer, Tokyo
- The CAZypedia Consortium (2018) Ten years of CAZypedia: a living encyclopedia of carbohydrate-active enzymes. *Glyco-Forum Glycobiology* 28(1):3–8
- Thorat LJ, Nath BB (2015) Tolerance to desiccation stress in *Chironomus ramosus* through plasticity in homeostatic control. *Eur J Env Sci* 5:86–91
- Thorat LJ, Nath BB (2016) Quantitative assessment of larval desiccation tolerance in oriental *Chironomus* species. *Curr Sci* 11:1448–1449
- Thorat L, Gaikwad S, Nath BB (2012) Trehalose as an indicator of desiccation stress in *Drosophila melanogaster* larvae: a potential marker of anhydrobiosis. *Biochem Biophys Res Comm* 419:638–642
- Thorat LJ, Oulkar D, Banerjee K, Gaikwad SM, Nath BB (2017) High-throughput mass spectrometry analysis revealed a role for glucosamine in potentiating recovery following desiccation stress in *Chironomus*. *Nat Sci Rep* 7:3659
- Webb B, Sali A (2014) Comparative protein structure modeling using modeller. *Curr Protoc Bioinformatics* 54:5–6