Molecular Modeling and Docking of Microbial Inulinases Towards Perceptive Enzyme–Substrate Interactions

Puneet Kumar Singh · Pratyosh Shukla

Received: 4 July 2011 / Accepted: 5 January 2012 / Published online: 21 January 2012
© Association of Microbiologists of India 2012

Abstract Inulin is emerging as an extremely rare source of sugar, it is having more sweetening capacity than table sugar, has beneficial effect in diabetic patient. Inulinases mainly produced by the microorganism and it degrades inulin into fructose which is a digestible form. There are more than 58 strains of microorganisms which are involved in the production of inulinases. The present report investigates about the selectivity of inulin by inulinase and its action to produce fructose through molecular docking. We have investigated exo-inulinase and endo-inulinases from *Penicillium* sp. TN-88(BAC16218) and *Penicillium* sp. TN-88(BAA19132), respectively with different arrangement of amino acids in the active site which detect the substrate. The protein sequences described above were processed to homology modeling by Swiss model and further they were docked with 1-ketose and fructose-6-phosphate as substrate by DOCK6 software package (dock.compbio.ucsf.edu). The results of the present studies represented that fructose-6-phosphate ((2R,3R,4S) fructose-6-phosphate) was having better interaction with exo-inulinase showing grid score of -40.288094 and the conserved amino acid Asp-22, Asp 128, Asp 179 and Ser 84 of exo-inulinase are involved in the bonding. In addition to this it was also seen that 1-ketose ((3S,4R)-ketose 1-phosphate) did not shown any interaction with the conserved part of the endo-inulinase.

Keywords Inulin · Inulinase · Molecular docking · Fructose-6-phosphate · DOCK6 software

Introduction

Inulin is a non digestible carbohydrate that is found in many plants as a storage carbohydrate which is part of man's daily diet. It is a polyfructan consisting of linear β -2,1-linked polyfructose chains, terminated by a glucose residue attached through a sucrose-type linkage, it is present in many regularly consumed vegetables, fruits and cereals. Fructose is having more sweetening capacity than table sugar, has beneficial effect in diabetic patient and increase iron absorption in children. Inulinase is the enzyme which hydrolyzes inulin to produce fructooligosaccharides. Inulinases are of type exo-inulinase and endo-inulinases. Exo-inulinases (EC 3.2.1.80) cleave β -2,1 linkages sequentially starting from the non-reducing end of inulin and split off terminal fructosyl units, releasing fructose with a molecule of glucose, where as endo-inulinases (EC 3.2.1.7) act randomly and hydrolyze internal linkages of inulin to yield fructooligosaccharides [1]. Both exo-inulinase and endo-inulinase shows inulinase and invertase property but inulinases are different from the invertases by the ratio I/S (inulinase activity/invertase activity) ratio. If the I/S ratio is $>10-2$, the enzyme is referred to as an inulinases, and if I/S ratio is $<10-4$, it is considered an invertase. There are more than 58 strains of different microorganism reported which are involved in the production of inulinases [2].

Fructose syrup has acquired an unbeatable position in the present sugar market and giving tough competition to the sugar industries as a low priced sweetener. Initially high fructose syrup was produced by the chemical hydrolysis of inulin but it showed many drawbacks, moreover in the final

P. K. Singh · P. Shukla (✉)
Department of Biotechnology, Birla Institute of Technology
(Deemed University), Mesra, Ranchi 835215, Jharkhand, India
e-mail: pratyosh.shukla@gmail.com

Present Address:

P. Shukla
Department of Microbiology, Maharshi Dayanand University,
Rohtak 124001, Haryana, India

product around 5% of the product is remained as difructose anhydride which is of no sweetening effect. Fructose was also produced by starch but it requires the involvement of three enzymes (α amylase, glucoamylase and gluco-isomer) and yield maximum 45% fructose. Production of fructose from inulinase is around 95% and it is a single step process. Recently mutant *Geotrichum candidum* has proved to be an improved endo-inulinase producing species which produces around 50 fold improved production (45.65 IU/ml) compared to wild isolate [3].

Inulinase is also employed in the production of ethanol which is an excellent liquid biofuel and used as gasoline enhancer, it is also having many chemical uses. In the past few years the focus was dispersed in the production of

ethanol and it was shown that the marine yeast is the new resource for the discovery of enzymes and genes [4].

Exo-inulinase and endo-inulinase differ in their activity which is decided by their active site. A recent study on endo-inulinase from *Aspergillus niger* and exo-inulinase from *Bacillus stearothermophilus* shows that there are certain amino acid present in the enzymes surface which establish a complex network of interactions with several other residues of the enzyme's active site, playing a major role in maintaining its shape and therefore responsible of the chemoselectivity of the enzyme [5]. The local structure of an active site is the catalytic residues and those within 3 Å of the catalytic residues, where the distance of two residues is the distance of their nearest heavy atoms. The local structure of

Fig. 1 Multiple alignment of deduced amino acid sequences of inulinases from different sources, viz. *Aspergillus ficuum* (CAA07345), *A. niger* (AAL34524), *Penicillium purpurogenum* (BAA12321), *Penicillium sp. TN-88* (BAA19132), *Pseudomonas mucidolens* (AAF24999), *A. awamori* (CAL44220), *A. fumigatus* (EAL86248), *A. niger* (BAA33798), *Bacillus polymyxa* (AAL82575), *B. licheniformis*

(AAU24331), *Geobacillus stearothermophilus* (BAC45010), *Kluyveromyces marxianus* (CAA48500), *Penicillium sp. TN-88* (BAC16218), *P. mucidolens* (AAF44125). The numbers in parentheses accession numbers for the reference sequences. The ClustalW analysis was performed with the EBI ClustalW program

the protein is defined in this way because, although the catalytic residues play the most important role in ligand binding, the surrounding residues need to be conserved to provide a stable environment to the active site [6].

Our studies presented here are efforts towards utilization of in silico methods towards understanding the docking of microbial inulinases towards perceptive enzyme–substrate interactions. Though the sequences of almost every inulinase producing microorganism are presents in the web resources and their data has been analyzed up to a certain level which has proven the relatedness of organisms producing exo-inulinase and endo-inulinase [7] and help us to perform genome comparison through which many useful information can be retrieved.

Materials and Methods

Microorganisms and Sequence Selection

Sequences for exo-inulinase from *Penicillium* sp. TN-88(BAC16218) and endo-inulinase from *Penicillium* sp. TN-88(BAA19132) were retrieved from NCBI database, these sequences were further used as query sequence to model 3D structure of the sequences.

Homology and Structural Modeling

The structures of retrieved sequences were modeled through RCSB Protein DataBank (PDB). Homology modeling was

done with the help of Swiss model [8–11], it is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server, or from the program DeepView. The modeled structure were further verified with NIH protein server program i.e. PROCHECK, Verify 3D, Errat.

Molecular Docking and Analysis

Rigid type docking was performed by DOCK6 software with all default parameter set. This is Unix based scientific software and follows a common installation recipe: download, unpack, configure, build, and test. Docking results were analyzed by the Chimera program which provides an insight into the molecule to observe it precisely at atomic level.

Results and Discussion

The catalytic mechanism of inulinase enzyme on its substrate is yet unknown. The structure of exo-inulinase from *Penicillium* sp. TN-88(BAA19132) and endo-inulinases from *Penicillium* sp. TN-88(BAC16218) have not been crystallized [12] so the crystal structure of exo-inulinase from *A. awamori* PDB ID 1Y4W was taken as the template which was used for modeling of both enzymes. The sequence similarity of 1Y4W with exo-inulinase from *Penicillium* sp. TN-88(BAC16218) and endo-inulinases from *Penicillium* sp. TN-88(BAA19132) was 67.714 and

Fig. 2 Phylogenetic analysis of 14 different species producing exo-inulinases and endo-inulinase have been shown in Phenogram form developed by <http://www.iubio.bio.indiana.edu/Inulinase> produced from *Aspergillus awamori* and *Penicillium* sp. TN-88(BAC16218) are closely related whereas *Penicillium* sp. TN-88(BAA19132) producing endo-inulinase is closely related with *Penicillium purpurogenum*

C Program: ERRAT2
Chain#:1
Overall quality factor*: 74.330

* On the error axis two lines are drawn to indicate the confidence with which it is possible to reject region that exceed that error value.

** Expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit, good high resolution structures generally produce value around 95% or higher. For lower resolutions (2.5 to 3 Angstroms) the average overall quality factor is around 91%

◀ **Fig. 3** Structure prediction and validation result of exo-inulinase from *Penicillium* sp. TN-88(BAC16218) through homology modeling. **a** Ramachandran plot showing 79.6% of the atom residing in the most favored region, 18.3% in allowed region and 2.2% in generously allowed region and no residue was in disallowed region. **b** Prove analyzing the whole structure for the Z score and 3.0% atom came to be outliers. **c** Errat exhibiting the overall quality factor of the protein i.e. 74.33%

35.258%, respectively. It was highest among other inulinases structures available in the PDB.

On the other hand, crystal structure of exo-inulinase from *A. awamori* PDB ID 1Y9G was taken as reference molecule for subsequent docking studies as its structure is co-crystallized with the fructose molecule in its active site. In our studies, we modeled the structure of exo-inulinase and endo-inulinase and docked the modeled structure with fructose-6-phosphate and 1-ketose respectively. The rigid docking was performed in which the enzyme and substrate exhibited a genuine interaction.

Further phylogenetic study and multiple sequence alignment were also performed to ensure the effectual binding of the substrate and molecule. For this purpose, 14 different species producing exo-inulinase and endo-inulinase were taken which revealed that the inulinase produced from *A. awamori* and *Penicillium* sp. TN-88(BAC16218) are closely related whereas *Penicillium* sp. TN-88(BAA19132) producing endo-inulinase is distantly related to the former two species (Fig. 2). During the multiple sequence alignment it was executed that three conserved region are located as WMND/EPN/H, WGHATST/K/SD and RDP (Fig. 1). The involvement of these conserved regions is further discussed in the docking analysis section.

The sequence of exo-inulinase from *Penicillium* sp. TN-88(BAC16218) was retrieved from NCBI (<http://www.ncbi.nlm.nih.gov/>) and submitted to Swiss model (<http://swissmodel.expasy.org>) for structure modeling. For reference PDB molecule 1Y4WA came to be most similar structure with 67.714% similarity. The modeled structure was then validated using NIH server with various programs like Procheck (Fig. 3b), Errat (Fig. 3c) and Verify_3d. The Ramachandran plot showed that 79.6% of the atom residing in the most favored region, 18.3% in allowed region and 2.2% in generously allowed region and no residue was in disallowed region (Fig. 3a).

Same procedure was followed with endo-inulinase from *Penicillium* sp. TN-88(BAA19132). Sequence was retrieved and modeled with the same reference PDB structure 1Y4WA. The similarity came out to be 35.258% of the *Penicillium* sp. TN-88 with the reference structure. The modeled structure was then validated using NIH server with various programs like Procheck (Fig. 4b), Errat (Fig. 4c) and Verify_3d. The Ramachandran plot showed that 77.9% of the atom residing in the most favored region, 19.0% in allowed region and 1.8% in

generously allowed region and 1.3% residue was in disallowed region as presented in Fig. 4a.

Docking Analysis

Docking analysis of exo-inulinase with fructose-6-phosphate was performed and grid score -40.288094 was recorded which indicates a good interaction between exo-inulinase and fructose-6-phosphate. The docking was performed with DOCK6 with rigid selection. The same procedure was carried out with endo-inulinase with 1-ketose on same parameters and software and the result emerged as grid score -37.746162 which is also an appropriate score for binding. In the reference, a different molecule with PDB ID 1Y9G was taken where fructose molecule showed hydrogen bonding between Phe-102, Ser-203, Asp-41, Glu-57. Further, the interaction with Asp-22, Asp 128, Asp 179 and Ser 84 with the active site was critically observed for exo-inulinases (Fig. 5). It was also observed that the fructose molecule binds into the pocket of the funnel shape active site of exo-inulinase and bonding between the molecules is through hydrogen bond with bond length between 2.40 and 2.50 Å (Fig. 6). On the other hand when the result was analyzed in the endo-inulinase the docking score was favorable and the 1-ketose molecule was bind in the pocket shape active site of the endo-inulinase (Fig. 8). The interaction between Glu-34, Asn-17, Glu 200 and Trp 42 and the 1-ketose molecule is presented in Fig. 7 where the bond length of hydrogen was in the range of 2.00–2.90 Å. It has been reported that a triad Arg–Asp–Pro (RDP) is seen and is preserved in the fructosyl-transferase or invertase and inulinases which is responsible for the detection of the fructopyranosidic residue. In addition the activity of RDP, the catalytic process of invertase from *Saccharomyces cerevisiae* includes Asp-42 in the β -fructosidase motif (WMNDPNG) as a nucleophile and Glu-223 in the ECP (domain V) as a catalyst [13, 14]. Whereas in glycosyltransferases and glycosylhydrolases, Asp and Glu act as proton donors and they also exhibit catalytic function. Here, fructosyl residue is bound by an ester bond to the b-carboxyl group of an Asp residue which is revealed by the protease digestion of the stable fructosyl enzyme intermediate [15]. Multiple sequence alignment studies has revealed that above mentioned conserved domain has been found in the exo-inulinase from *Penicillium* sp. TN-88(BAC16218) and endo-inulinases from *Penicillium* sp. TN-88(BAA19132) (Fig. 1). The substrates i.e. fructose and 1-ketose binds with the amino acids of protein (exo-inulinase and endo-inulinase) which are present very near to these domains and are probably present in the active site. We notice that amino acid (Glu) present in the active site binds with the substrate in endo-

A

PROCHECK

Ramachandran Plot BAA19132

B

PROVE

BAA19132.pdb

Analysis of entire structure

Distribution of atomic Z - scores

Atomic Z - Scores

Z - Score mean 0.185
Z - Score stddv 24.792
Z - Score RMS 24.798

#Scored atom 1630
#Outlier 111
%Outlier 6.800

Ponikvar et al. (1996) J. Mol. Biol. 264, 121–136

BAA19132_01.ps

C

Program: ERRAT2

Chain#:1

Overall quality factor*: 56.689

* On the error axis two lines are drawn to indicate the confidence with which it is possible to reject region that exceed that error value.

** Expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit, good high resolution structures generally produce value around 95% or higher. For lower resolutions (2.5 to 3 Angstroms) the average overall quality factor is around 91%

◀ **Fig. 4** Structure prediction and validation result of endo-inulinase from *Penicillium* sp. TN-88(BAA19132) through homology modeling, **a** Ramachandran plot demonstrating that 77.9% of the atom residing in the most favored region, 19.0% in allowed region and 1.8% in generously allowed region and 1.3% residue was in disallowed region. **b** Prove analyzing the whole structure for the Z score and 6.8% atom came to be outliers. **c** Errat exhibiting the overall quality factor of the protein i.e. 56.689%

Fig. 5 Docked structure showing interaction between endo-inulinase and fructose. The interaction with Asp-22, Asp 128, Asp 179 and Ser 84 with the active site has been shown in the figure. The interaction shown is rich in Asp and Asp 179 lies very near to the RDP triad

Fig. 6 Fructose molecule binds into the pocket of the funnel shape active site

inulinase and perform the reaction whereas in endo-inulinase 'Asp' plays a major role in the active site to interact with substrate.

Mutational Analysis

Towards validating the effectiveness of 'Asp' molecule in the active site of inulinases mutation at the Asp-179 position of

Fig. 7 Interaction between endo-inulinase and 1-ketose has been revealed in the figure. Glu-34, Asn-17, Glu 200 and Trp 42 can be seen interacting with the 1-ketose molecule. The interaction does not involved interaction with Asp whereas ketose binds with the Glu rich active site

Fig. 8 A view of the 1-ketose molecule binds into the pocket shape active site of the enzyme

exo-inulinase from *Penicillium* sp. TN-88(BAC16218) was carried out by Pymol (<http://www.pymol.org/>) and 'His' amino acid was replaced at the Asp-179 position. Further docking was performed with DOCK6 software and similar parameters were set as describe earlier in case of *Penicillium* sp. TN-88(BAC16218). During these studies the grid score of -38.3658701 was recorded which reflects slighter effective binding with the active site as compared to the wild type.

Conclusion

The structural modeling and active site study for exo-inulinase from *Penicillium* sp. TN-88(BAC16218) and endo-inulinase from *Penicillium* sp. TN-88(BAA19132) are reported in the present study. This study emphasizes on inulinases from the fungal sources and illustrates its enzymatic activity towards the substrate i.e. fructose and 1-ketose. Further it also embraces the analysis of enzyme selectivity of inulinase towards their substrates. From the above studies it was revealed that both exo-inulinase and endo-inulinase have effective interaction between enzyme and substrate with grid energy of -40.288094 and -37.746162 for exo-inulinase and endo-inulinases respectively. Since exo-inulinase forms more H-bonds as compared to endo-inulinases to its substrate, so exo-inulinase interacts with greater efficiency and shows greater stability to its substrate as compared to endo-inulinases. One of the important things observed in exo-inulinase is that it forms a funnel size active site comprised of Asp-22, Asp 128, Asp 179 and Ser 84 and was rich in 'Asp' amino acid. Mutational analysis further established the catalytic activity of 'Asp' present in the active site. The study also highlights that further such investigations of diverse inulinases produced from different microorganisms may disclose an improved inulinase enzyme for the production of inulin.

Acknowledgments We hereby acknowledge BTIS-Sub DIC facility at BIT Mesra through DBT, GOI and the Department of Agriculture, Government of Jharkhand for financial support for our department towards infrastructure facilities.

References

1. Singh RS, Singh RP (2010) Production of fructooligosaccharides from inulin by endoinulinases and their prebiotic potential. *Food Technol Biotechnol* 48(4):435–450
2. Pandey A, Soccol CR, Selvakumar P, Soccol VT, Rieger NK, Fontana JD (1999) Recent developments in microbial inulinases. *Appl Biochem Biotechnol* 81:35–52
3. Mughal MS, Ali S, Ashiq M, Talish AS (2009) Kinetics of an extracellular exo-inulinase production from a 5-fluorocytosine-resistant mutant of *Geotrichum candidum* using two-factorial design. *Bioresour Technol* 100:3657–3662
4. Chi Z, Zhang T, Liu G (2009) Inulinase-expressing microorganisms and applications of inulinases. *Appl Microbiol Biotechnol* 82:211–220
5. Basso A, Spizzo P, Ferrario V, Knapic L, Savko N, Braiuca P, Ebert C, Ricca E, Calabro V, Gardossi L (2010) Endo- and exo-inulinases: enzyme–substrate interaction and rational immobilization. *Biotechnol Prog* 26(2):397–405
6. Weng YZ, Chang DTH, Huang YF, Lin CW (2011) A study on the flexibility of enzyme active sites. *BMC Bioinformatics* 12(Suppl 1):S32
7. Singh P, Gill PK (2006) Production of inulinases: recent advances. *Biotechnol* 44(2):151–162
8. Arnold K, Bordoli L, Kopp J, Schwede T (2006) SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. *Bioinformatics* 22:195–201
9. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL repository and associated resources. *Nucleic Acids Res* 37:D387–D392
10. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385
11. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. *Electrophoresis* 18:2714–2723
12. Arand M, Golubev AM, Neto JRB, Polikarpov I, Wattiez R, Korneeva OS, Eneyskaya EV, Kulminkaya AA, Shabalin KA, Shishliannikov SM, Chepurnaya OV, Neustroev KN (2002) Purification, characterization, gene cloning and preliminary X-ray data of the exo-inulinase from *Aspergillus awamori*. *Biochem J* 362:131–135
13. Nagem RA, Rojas AL, Golubev AM, Korneeva OS, Eneyskaya EV, Kulminkaya AA, Neustroev KN, Polikarpov I (2004) Crystal structure of exo-inulinase from *Aspergillus awamori*: the enzyme fold and structural determinants of substrate recognition. *J Mol Biol* 344:471–480
14. Kang S, Kim S (1999) Molecular cloning and sequence analysis of an endo-inulinase gene from *Arthrobacter* sp. *Biotechnol Lett* 21:569–574
15. Song DD, Jacques NA (1999) Mutation of aspartic acid residues in the fructosyl transferase of *Streptococcus salivarius* ATCC 25975. *Biochem J* 344:259–264