Received: 17 January 2023 | Revised: 15 May 2023 | Accepted: 26 June 2023

DOI: 10.1002/elsc.202300003

Engineering in Life Sciences

RESEARCH ARTICLE

Identification and characterization of inulinases by bioinformatics analysis of bacterial glycoside hydrolases family 32 (GH32)

Fatemeh Khosravi[1] | **Ehsan Mohseni Fard**[2] | **Marzieh Hosseininezhad**[3] | **Hadi Shoorideh**[4]

[1]Ph. D. student of Agriculture Biotechnology, University of Zanjan, Zanjan, Iran

[2]Department of Plant Production and Genetics, Faculty of Agriculture, University of Zanjan, Zanjan, Iran

[3]Department of Food Biotechnology, Research Institute of Food Science and Technology, Mashhad, Iran

[4]Dryland Pulses Research Department, North Khorassan Agricultural Research, Education and Extension Organization (AREEO), Shirvan, Iran

Correspondence
Ehsan Mohseni Fard, Department of Plant Production and Genetics, Faculty of Agriculture, University of Zanjan, Zanjan, Iran.
Email: mohsenifard.ehsan@znu.ac.ir

Abstract

The glycoside hydrolase family contains enzymes that break the glycosidic bonds of carbohydrates by hydrolysis. Inulinase is one of the most important industrial enzymes in the family of Glycoside Hydrolases 32 (GH32). In this study, to identify and classify bacterial inulinases initially, 16,002 protein sequences belonging to the GH32 family were obtained using various databases. The inulin-effective enzymes (endoinulinase and exoinulinase) were identified. Eight endoinulinases (EC 3.2.1.7) and 4318 exoinulinases (EC 3.2.1.80) were found. Then, the localization of endoinulinase and exoinulinase enzymes in the cell was predicted. Among them, two extracellular endoinulinases and 1232 extracellular exoinulinases were found. The biochemical properties of 363 enzymes of the genus *Arthrobacter*, *Bacillus*, and *Streptomyces* (most abundant) showed that exoinulinases have an acid isoelectric point up to the neutral range due to their amino acid length. That is, the smaller the protein (336 aa), the more acidic the pI (4.39), and the larger the protein (1207 aa), the pI is in the neutral range (8.84). Also, a negative gravitational index indicates the hydrophilicity of exoinulinases. Finally, considering the biochemical properties affecting protein stability and post-translational changes studies, one enzyme for endoinulinase and 40 enzymes with desirable characteristics were selected to identify their enzyme production sources. To screen and isolate enzyme-containing strains, now with the expansion of databases and the development of bioinformatics tools, it is possible to classify, review and analyze a lot of data related to different enzyme-producing strains. Although, in laboratory studies, a maximum of 20 to 30 strains can be examined. Therefore, when more strains are examined, finally, strains with more stable and efficient enzymes were selected and introduced for laboratory activities. The findings of this study can help researchers to select the

Abbreviation: GH32, Glycoside Hydrolase 32.

appropriate gene source from introduced strains for cloning and expression heterologous inulinase, or to extract native inulinase from introduced strains.

KEYWORDS

bacteria, Glycoside Hydrolase 32, EC number, inulinase, inulin

1 | INTRODUCTION

Bioinformatics is a new science that includes a set of biological sciences, computer science, statistics and chemistry, mathematics, and so forth. which has been widely used in biosystems and biotechnology. Bioinformatics, and computational biology can use various algorithms to make predictions to solve a researcher's challenges and be used as a tool to screen research strategies [1]. In the post-genomic age, databases were faced with a plethora of protein sequences that had to be categorized in some way. Laboratory methods are certainly not cost-effective for this data, so the use of bioinformatics tools can be a great help in accurately classifying enzyme families [2]. Conventional screening methods to identify new enzymes are performed by exploring natural resources such as industrial waste and soil. Still, these require a consistent screening protocol or selection method based on the desired enzyme properties [3]. This process involves biochemical screening and isolation of the organism on select media, which is usually time-consuming and resource-intensive, which may lead to the selection of a new candidate. If the resources of molecular biology are insufficient, direct screening and identification methods can be performed, but due to their low success and time consumption, the use of bioinformatics studies to identify new enzymatic candidates with better properties is a suitable alternative method [4]. The International Union of Molecular Biochemistry and Biology (IUBMB), the Enzyme Commission (EC) have classified and named enzymes based on their specific enzymatic activity, substrate specificity, and sometimes their molecular mechanism [5]. Enzymes, are divided into six categories according to number (EC): EC-1: Oxidoreductase, EC-2: Transferase, EC-3: Hydrolase, EC-4: Lyase, EC-5: Isomerase, and EC-6: Synthetase.[2] Glycoside hydrolase (GH) are enzymes that can hydrolyze the glycoside bond between two carbohydrates or even between a carbohydrate and a non-carbohydrate moiety. Among the enzymes of the GH family, some, such as amylase, cellulase, hemicellulose, invertase, β-glucosidase, and inulinase, are very vital for various industries [6, 7]. The enzymes capable of hydrolyzing compounds with O- and S-glycosylation bonds are indicated by EC 3.2.1, which is classified from EC 3.2.1.1 to EC 3.2.1.214. Conversely, the GH family classification from CAZy (http://www.cazy.org/) extends from GH1 to GH171 [7]. Inulin-affecting enzymes are an important class of industrial enzymes belonging to the 32 and 91 GH families. These, two families represent more than a thousand GH enzymes, including inulinases, invertases, 1-exohydrolases, inulin lyases, levanases, fructan-fructosyltransferases, and sucrose fructosyltransferases (Table 1). Among them, GH32 family inulinases target 1–2β inulin bonds to produce numerous essential products [8] Inulinase, the enzyme that hydrolyzes inulin, is among the most important industrial enzymes that are firstly isolated from the root tubers of *Helianthus tuberosus* [9]. Inulin, is a well-known carbohydrate polymer composed of linear chains of D-fructofuranose with a 1–2 β bond that terminates in a glucose molecule linked to a type of sucrose and releases only fructose from its hydrolysis so that a fructan can be. Inulin, is a storage polymer found in many roots and tubers that, unlike starch in cereals and potatoes, is water soluble [10]. Inulinases are a class of versatile industrial enzymes mainly used to produce carbohydrate-derived products such as high fructose syrup, fructooligosaccharides, bioethanol, biofuels, citric acid, lactic acid, and so forth [8, 11].

Inulinases are classified into two categories based on their function on inulin: exoinulinase (EC 3.2.1.80) and endoinulinase (EC 3.2.1.7). Exoinulinase, releases fructose units by end-degrading inulin from its non-regenerated end, while endoinulinase accidentally breaks down inulin in the middle to produce fructooligosaccharides of different chain lengths [8]. Partial or complete hydrolysis of inulin, which makes high-fructose syrup, can be achieved by the enzyme inulinase. To have high-fructose syrup from inulin for use in industry, enzymatic hydrolysis can be used by exoinulinase alone or with endoinulinase [10]. The enzyme is also present in some animal and plant sources, but due to its low efficiency, they have received less attention than using microorganisms to produce inulinase. Many microorganisms, including fungi, yeasts, and bacteria, are known to produce inulinase [9]. The biochemical properties of enzymes depend most of all on their source. Inulinase production has been reported from various sources, such as animals, stored inulin tissues of plants, and multiple microorganisms [8]. In plants and animals, commercial sources cannot be explored due to

TABLE 1 Some of GH32 enzymes with divers EC number.

Accepted/alternative name	EC number	Reaction/substrate catalyzed
Inulinase/inulase	3.2.1.7	Endohydrolysis of (2→1)-β-D-fructosidic linkages in inulin
Beta-fructofuranosidase/invertase/Saccharose	3.2.1.26	Hydrolysis of terminal non-reducing beta-D-fructofuranoside residues in beta-D-fructfuranosides Substrates: sucrose
2,6-beta-fructan 6-levanbiohydrolase	3.2.1.64	Hydrolysis of (2→6)-beta-D-fructofuranan, to remove successive disaccharide residues as levanbiose, i.e. 6-(beta-D-fructofuranosyl)-D-fructose, from the end of the chain.
Levanase	3.2.1.65	Random hydrolysis of (2→6)-beta-D-fructofuranosidic in (2→6)-beta-D-fructans (levans) containing more than 3 fructose units.
Fructan beta-fructosidase/Exo-beta-D-fructosidase/Fructanase	3.2.1.80	Hydrolysis of terminal, non-reducing (2→1)- and (2→6)-linked beta-D-fructofuranose residues in fructans. substrate: inulin, levan, and also sucrose.
Fructan beta-(2,6)-fructosidase/levanase	3.2.1.154	Hydrolysis of terminal non-reducing (2→6)-linked beta-D-fructofuranose residues in fructans. best substrates for enzyme: the levan-type fructans

low enzyme performance, Therefore, microbial sources for inulinase production were given more attention due to their many advantages of ease of use, culture, and genetic manipulation. Due to the strength of inulinase-producing bacterial strains against heat, researchers are eager to screen and isolate a unique strain with a heat-resistant enzyme to produce inulinase. . These are enzymes that can be used at high temperatures in various industries [8, 12]. Many microorganisms can produce inulinase, but inulinase bacterial strains are more resistant to heat. On the other hand, inulinase is used in various industries at high temperatures. In this study, we utilised different bioinformatics methods for prediction of the effective features in structural-thermal stability of enzymes from bacterial GH32 family sequences. Also, we reported bacterial strains that produce these desirable enzymes.

2 | MATERIALS AND METHODS

2.1 | Glycoside hydrolase sequences

Amino acid sequences of the bacterial GH32 family were extracted from the Uniprot database (https://www.uniprot.org/).

2.2 | Prediction of Enzyme Commission (EC)

To determine the function of enzymes, their EC number was evaluated using the protein sequence of each enzyme using the DEEPre online tool (http://www.cbrc.kaust.edu.sa/DEEPre/).

2.3 | Prediction of subcellular localization

To investigate the accumulation site of bacterial inulinase enzymes, their amino acid sequence was examined using the cello tool, and their accumulation site was identified (http://cello.life.nctu.edu.tw/).

2.4 | Prediction biochemical properties

Biochemical properties using the ProtParam online tool predicted (https://web.expasy.org/protparam/) [13]. The molecular weight of the protein (Mw), and the isoelectric point (pI) were predicted according to the amino acid length of the protein, and the presence of ionized groups in ionized amino acids, respectively [14]. Also, the aliphatic index is dependent on the volume occupied by aliphatic side chains (alanine, valine, leucine, and isoleucine) [15]. In addition, the instability index indicates the degree of protein instability, and the GRAVY index determines the hydrophilicity/hydrophobicity of the protein [16, 17].

2.5 | Prediction of post-translation modification (PTMs)

Glycosylation, as one of the post-translational modifications in proteins, results from adding a sugar molecule to the protein. Different types of glycosylation occur based on the attachment site of the sugar molecule to various amino acids. For example, N-glycosylation results from adding a sugar molecule to asparagine residues, while O-glycosylation occurs in the side chain of serine or threonine residues [18]. Also, the formation of disulfide bonds in proteins is an oxidative process that creates a covalent bond between sulfur atoms that binds two cysteine residues together [19]. N-Glycosylation sites, O-Glycosylation sites, and disulfide bond positions were predicted by online tools NetNGlyc (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0), NetOGlyc (https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0), and DiANNA (http://clavius.bc.edu/~clotelab/DiANNA/), respectively [20].

3 | RESULTS AND DISCUSSION

3.1 | Prediction of Enzyme Commission (EC)

This study aimed to identify enzymes affecting inulin (exoinulinase and endoinulinase), considering that 16,002 sequences extracted from the GH32 family from the Uniprot database were categorised based on EC using the DEEPre online tool (Table S1). Among the examined sequences, 26.98% of the sequences were exoinulinase (4318 sequences), and 0.05% of the sequences were endoinulinase (8 sequences) (Figure 1).

Different enzymes have different biological processes and active sites. Therefore, knowing which family an enzyme belongs to can provide information about its biological function and indicate the type of catalytic mechanism [2]. DEEPre an online tool, classifies enzymes by connecting to four servers of EzyPred (determining class and subclass of enzyme), SVM Prot (determining 13 properties of enzyme including hydrophobicity, second structure), COFACTOR, and EFICAz (predict EC of enzymes). Also, enzymes are classified based on two types of encoding sequence-dependent, such as position-specific scoring matrices, and sequence-independent encoding, functional domains [21]. The hydrolytic effect of inulinases on inulin causes the production of fructose or fructooligosaccharides as the main products. Inulinases are classified as exoinulinases (β-D-fructan fructohydrolase, EC 3.2.1.80) and endoinulinases (1-β-Dfructan fructanohydrolase, EC 3.2.1.7) based on their effect on inulin [8]. Exoinulinase, produces fructose by hydrolysis of the end of the inulin chain. In contrast, endoinulinase makes fructooligosaccharides with different lengths by random hydrolysis on the internal glycosidic bonds of the inulin chain [8]. A common feature in the tertiary structure of endoinulinase and exoinulinase is the presence of a protected arginine-aspartic acid-proline (RDP) motif, the presence of which is essential for the detection of the pyranosidic ring. It has also been reported to be responsible for the specificity of the enzyme to fructopyranosidic residues [22]. Inulinase has a C-terminal domain of GH32 (a beta sandwich module form) and an N-terminal domain of GH32 (a five bladed beta propeller structure form) [23]. The two types of inulinase (exo and endo) structures are chiefly built of beta-sheets, in addition, endoinulinases have one alpha-helix. Therefore, due to the presence of beta sheets in exoinulinases, exoinulinases are more stable than endoinulinases in cellular chemical reactions. [24]. It is thought that this structural stability is the reason for the abundance of exoinulinases compared to endoinulinases in the cell.

3.2 | Prediction of subcellular localization

Bacterial endoinulinases and exoinulinases identified from the GH 32 family with their characteristics are reported (Table S2). After examining the subcellular accumulation site, among the endoinulinases found, 75% of

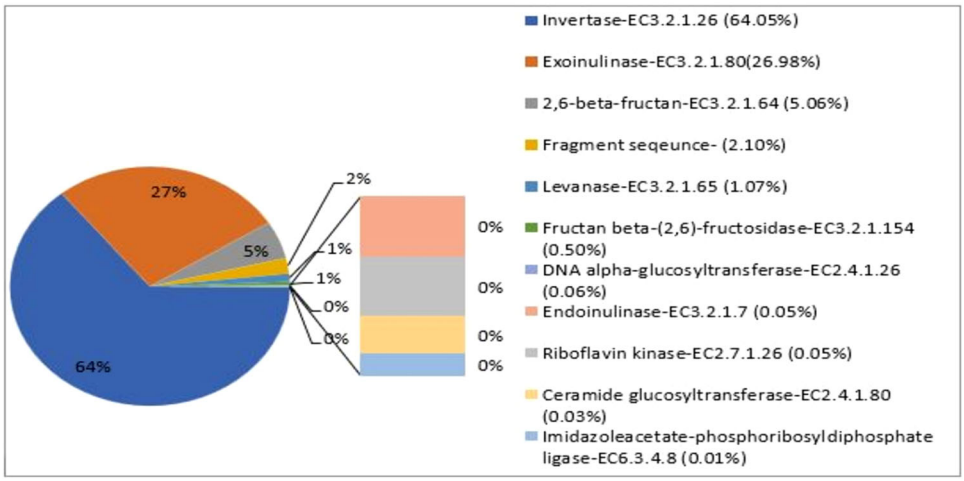

FIGURE 1 Classification of enzymes based on EC number.

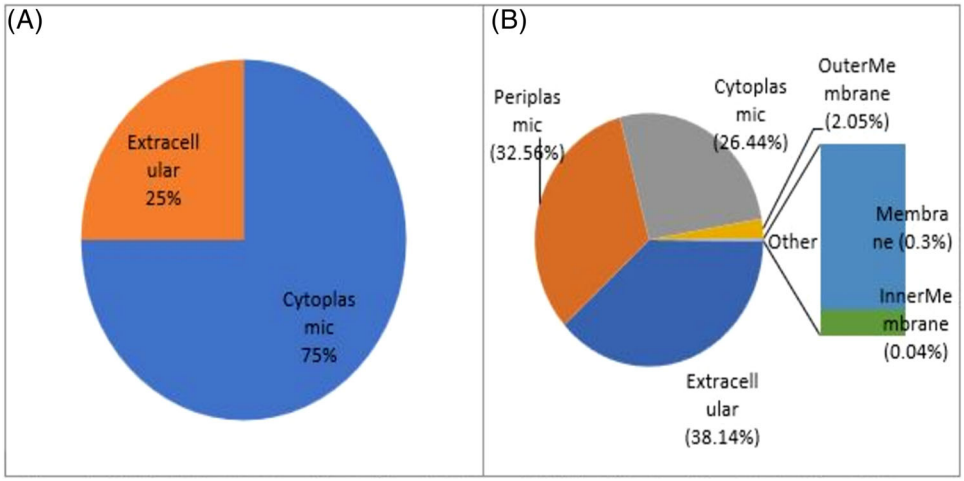

FIGURE 2 (A) Subcellular location of bacterial endoinulinase and (B) subcellular location of bacterial exoinulinase.

the enzymes were cytoplasmic, and 25% were extracellular (Figure 2A). In comparison, 38.14% of the exoinulinases were extracellular, 32.86% were periplasmic, and 26.44% were cytoplasmic (Figure 2B).

One of the essential elements of cell survival is proteins, whose presence in specific cell sites determines their function and biological nature. After synthesizing proteins in the cytosol, they are directed to particular parts of the cell, including organelles, based on their biological function [25]. Prediction of subcellular protein localization is based on sorting signals, amino acids composition, and homology [25]. Protein sorting is a complex biological mechanism often driven by specific signal sequences in nascent proteins. Signal peptides are short sequences of parts at the N-terminus of a newly synthesized protein that determines the pathway of a protein. These signal peptides can direct proteins to destinations such as the Golgi apparatus, endoplasmic reticulum, plasma membrane, and extracellular space. Identification of signal peptides in the protein sequence is a prerequisite for revealing the destination and function of the protein [26]. Microorganisms break down natural polymers in the environment, such as protein, starch, and pectin, using the extracellular secretion of enzymes such as protease, amylase, and pectinase. These extracellular enzymes break down giant substrate molecules into smaller molecules that can feed microorganisms [27].

3.3 | Prediction biochemical properties and PTMs

Since extraction of extracellular enzymes from bacterial sources is more accessible than intracellular and cytoplasmic enzymes [27]. among the studied inulinases, two extracellular endoinulinases, and there were 1232 extracellular

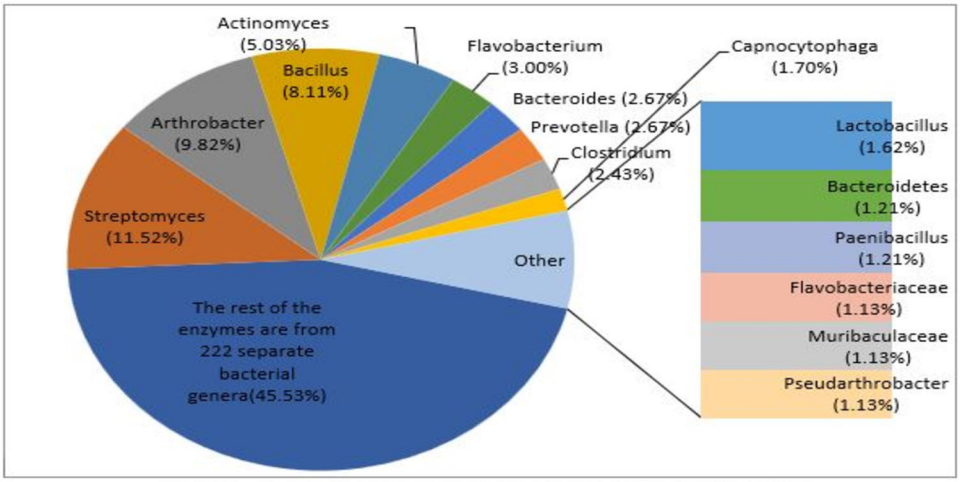

FIGURE 3 Grouping of extracellular exoinulinases based on bacterial genera.

TABLE 2 Biochemical properties of extracellular endoinulinase.

Uniprot ID	Organism	Length	Mw[a]	pI[b]	instability index	aliphatic index	GRAVY[c]
A0A0F6MTX2	*Sphingomonas* sp. JB13	505	55,695.52	6.27	38.14 (stable)	72.24	−0.341
A0A2D5SDX0	*Deltaproteobacteria bacterium*	1295	137,997.5	5.02	32.36 (stable)	77.31	−0.13

[a]Molecular weight.
[b]Point isoelectric.
[c]Grand average of hydropathy.

exoinulinases in 237 bacterial genera, which are shown in Figure 3.

It was found that 1232 extracellular exoinulinase enzymes are produced by 237 bacterial genera, among which the highest abundance of this enzyme is observed in 3 genera: *Arthrobacter*, *Bacillus*, and *Streptomyces*. These three genera have 363 exoinulinase enzymes, of which 121 enzymes were identified in *Arthrobacter*, 100 enzymes in *Bacillus*, and 142 enzymes in *Streptomyces*. Many microorganisms, such as bacteria, fungi, and yeasts, are used to produce inulinases [12]. So far, few studies have been reported on inulinase-producing bacteria, such as *Bacillus*, *Streptomyces*, and *Sphingomonas*. Since bacterial strains can produce heat-resistant inulinase, researchers seek to screen and isolate a unique strain to make inulinase and use it in the industry [9, 12]. To identify bacterial strains capable of producing heat-resistant inulinase, biochemical properties and post-translational changes of endoinulinases and exoinulinases of three families of *Arthrobacter*, *Bacillus*, and *Streptomyces* were investigated. Biochemical properties and post-translational modifications were shown in Tables 2, 3 and Table S3, respectively.

Due to the factors affecting the half-life (lower instability index), high thermal stability (higher aliphatic index), and structural stability (PTMs including the number of glycosylation sites and disulfide bond), exoinulinase enzymes from different bacterial strains (Table 4), 11 enzymes of *Arthrobacter*, ten enzymes of *Bacillus* and 19 enzymes of *Streptomyces* were selected and biochemical properties and PTMs are shown in Tables 5 and 6, respectively.

3.4 | Length and Mw protein

The results of bioinformatics predictions showed that *Deltaproteobacteria bacterium* endoinulinase, with longer amino acid length and higher molecular weight, has a lower instability index than *Sphingomonas* sp. JB13 endoinulinase (Table 2).

In addition, the amino acid count exoinulinases of these three bacterial families lay between 336 and 1207. The enzymes from *Bacillus salsus* and *Bacillus* sp. DSL-17 have the shortest and the longest sequences, respectively. Accordingly, the lowest molecular weight (36.91 kDa) was found to be *Bacillus salsus* exoinulinases, and the highest molecular weight (133.98 kDa) was for the exoinulinase of *Bacillus* sp.DSL-17 (Table S3). The researchers used *Bacillus* Sp. SG7 to produce exoinulinase, and after isolating and purifying, the molecular weight of 56 kDa

TABLE 3 PTMs of extracellular endoinulinase.

Uniprot ID	N-Glyc	O-Glyc	Disulfide bond
A0A0F6MTX2	3(105-130-458)	1(359)	0
A0A2D5SDX0	7(149-186-310-841-849-889-10550	45(157-161-162-163-165-166-167-168-170-171-172-174-175-176-178-180-182-336-337-740-742-744-745-747-748-754-755-757-758-781-782-783-785-786-787-790-791-792-794-795-797-798-801-803-944)	1(105-273)

TABLE 4 Properties affecting the stability of exoinulinase enzyme.

Properties	The range in all enzymes	The range for selection enzymes	Details
pI	4.39-8.84	<7	Acidic
Instability index	16.31-50.34	<30	The lower the instability index, the greater the protein stability
Aliphatic index	61.01-81.23	>68	The greater the aliphatic index, the greater the protein thermal-stability
Glycosylation site	0-60	>10 sites	The greater number of glycosylation sites, the greater the protein structural-stability

was reported for the exoinulinase made using the sodium dodecyl sulfate (SDS-PAGE) technique [28]. To have the heat-resistant enzyme inulinase, the researchers isolated the gene encoding the 56-kDa exoinulinase enzyme from *Geobacillus stereotermophilus* and cloned it into *E. coli*. After the expression of the enzyme using the SDS-PAGE technique, 56 kDa exoinulinase was confirmed [27]. The molecular weight of 56 kDa of isolated exoinulinases in vitro [28, 29]. corresponds to the molecular weight range of the studied exoinulinases with bioinformatics tools in similar strains.

3.5 | pI

The isoelectric point is the pH where the protein is neutral and has no electric charge. pI of protein is directly related to protein length, subcellular localization, classification, and ecology of organisms [30, 31]. pI, of polypeptides usually depends on the presence of ionized amino acid ion groups, including glutamate (sigma-carboxyl group), aspartate (beta-carboxyl group), cysteine (thiol group), tyrosine (phenol group), histidine (imidazole side chains), lysine (ammonium group) and arginine (guanidinium group) [14].

Bacillus salarius exoinulinase was predicted to have the lowest pI (4.39), and *Bacillus* sp. MBGLi97 had the highest (8.84), also the mean pI of *Arthrobacter*, *Bacillus*, and *Streptomyces* exoinulinases were 4.39, 6.48, and 5.34, respectively (Table S3). The researchers reported that the thermophilic *Bacillus stearothermophilus* strain was capable of producing the extracellular exoinulinase enzyme with a molecular weight of 54 kDa, and pI = 5 in an inulin-induced medium [30]. Among extracellular inulinases, all endoinulinases, all exoinulinases of *Arthrobacter* strains, all exoinulinases of *Streptomyces* strains, and 70% of exoinulinases of *Bacillus* strains (Table S3) have acidic pI, which is consistent with the researchers' findings [32, 33].

Researchers have found that proteins with shorter amino acid lengths have more variable pI because fluctuations over shorter amino acids have more significant effects. Even the fusion of a pregnant amino acid can significantly affect the protein pI [16]. In this study, *Arthrobacter* and *Streptomyces* exoinulinases, with 99% and 95% of protein sequences smaller than 1000 amino acids, respectively, have mean pI of 4.89 and 5.34. While 66% of *Bacillus*

TABLE 5 Biochemical properties of selected extracellular exoinulinase.

	Uniprot ID	Organism	Length	Mw[a]	pI[b]	Instability index	Aliphatic index	GRAVY[c]
1	A0A345V7G7	*Arthrobacter sp. PM3*	502	54,758.5	4.53	28.88 (stable)	74.24	−0.262
2	A0A4Q8UP54	*Arthrobacter sp. S39*	502	54,441.2	4.63	29.38 (stable)	71.93	−0.251
3	A0A0Q9MZ00	*Arthrobacter sp. Soil762*	503	54,711.6	4.64	27.92 (stable)	73.96	−0.245
4	A0A1H9ESY2	*Arthrobacter sp. OV608*	505	54,692.48	4.60	28.02 (stable)	73.25	−0.245
5	A0A176UH96	*Arthrobacter sp. OY3WO11*	505	54,785.62	4.55	28.55 (stable)	74.00	−0.256
6	A0A1S9MFX5	*Arthrobacter sp. SRS-W-1-2016*	553	58,891.54	5.48	28.21 (stable)	70.07	−0.260
7	A0A221NQW5	*Arthrobacter sp. YN*	861	91,969.62	5.05	26.96 (stable)	70.27	−0.353
8	A0A542GI22	*Arthrobacter sp. SLBN-112*	872	92,617.65	4.86	29.87 (stable)	71.07	−0.271
9	A0A0D1C3S9	*Arthrobacter sp. SPG23*	886	93,844.06	4.96	29.30 (stable)	72.82	−0.247
10	A0A4R7L7Q7	*Arthrobacter sp. 1704*	902	95,709.2	5.14	29.05 (stable)	70.22	−0.308
11	A0A542ICI5	*Arthrobacter sp. SLBN-100*	1152	1,222,362.1	4.88	16.31 (stable)	73.39	−0.232
12	A0A1Q9FQT8	*Bacillus licheniformis*	667	75,618.58	6.39	24.88 (stable)	68.69	−0.526
13	A0A2M8SZP1	Bacillus sp. SN1	677	75,751.06	6.69	23.54 (stable)	70.66	−0.0481
14	A0A3A5I1K0	*Bacillus subtilis*	677	75,603.57	6.39	25.70 (stable)	68.11	−0.535
15	A0A1I3AAQ3	*Bacillus megaterium*	837	92,775.46	5.44	24.05 (stable)	74.49	−0.405
16	A0A0M0WIS6	*Bacillus sp. FJAT-21351*	1191	132,352.7	5.57	23.00 (stable)	72.51	−0.529
17	A0A1S2C6W1	*Bacillus aryabhattai*	1192	132,926.4	5.61	24.69 (stable)	70.73	−0.566
18	A0A3N6CXY1	*Bacillus endophyticus*	1192	132,074.1	6.88	20.66 (stable)	71.22	−0.554
19	A0A369C3Z3	*Bacillus sp. AG236*	1192	132,862.3	5.65	22.80 (stable)	71.22	−0.554
20	A0A0M4GBL2	*Bacillus gobiensis*	1196	131,637.7	5.46	25.32 (stable)	71.51	−0.467
21	A0A4S4BRT0	*Bacillus sp. DSL-17*	1207	133,987.0	4.83	26.36 (stable)	70.96	−0.476
22	A0A3G2JDL7	*Streptomyces sp.* Z022	491	53,947.39	6.85	26.95 (stable)	68.33	−0.469
23	U5YN42	*Streptomyces sp.* MMG1612	556	58,960.01	4.73	18.63 (stable)	68.94	−0.374
24	A0A4R3F0L9	*Streptomyces sp.* BK674	557	58,619.34	4.73	24.85 (stable)	72.03	−0.281
25	A0A2G7DJ26	*Streptomyces sp. 1*	684	70,356.89	5.38	26.85 (stable)	69.34	−0.339
26	A0A4R7MIK6	*Streptomyces sp.* 846.5	827	88,813.85	5.27	24.83 (stable)	75.62	−0.207
27	A0A0K8PVP7	*Streptomyces azureus*	839	90,014.81	4.90	23.54 (stable)	68.49	−0.368
28	V6KZJ3	*Streptomyces roseochromogenus* subsp. oscitans DS 12.976	848	91,226.87	4.89	26.14 (stable)	68.80	−0.363
29	A0A385ZUV2	*Streptomyces griseorubiginosus*	849	90,213.87	5.30	28.30 (stable)	70.24	−0.296
30	A0A5N8VV89	*Streptomyces phyllanthi*	852	91,403.7	5.43	25.16 (stable)	69.17	−0.355
31	A0A540PUG3	*Streptomyces ipomoeae*	853	91,600.54	5.07	24.78 (stable)	68.39	−0.359
32	A0A117RML8	*Streptomyces resistomycificus*	853	91,572.22	4.84	23.51 (stable)	68.85	−0.373

(Continues)

TABLE 5 (Continued)

	Uniprot ID	Organism	Length	Mw[a]	pI[b]	Instability index	Aliphatic index	GRAVY[c]
33	A0A0M2GHL8	*Streptomyces variegatus*	853	91,581.47	5.37	26.64 (stable)	68.07	−0.413
34	A0A1H0MUA1	*Streptomyces guanduensis*	988	104,829.7	5.75	29.04 (stable)	74.67	−0.198
35	A0A1H6BW72	*Streptomyces yanglinensis*	1008	107,129.7	5.33	26.50 (stable)	71.73	−0.238
36	A0A250VVS8	*Streptomyces olivochromogenes*	1189	126,312.3	5.77	28.33 (stable)	77.38	−0.243
37	A0A100JWP6	*Streptomyces scabiei*	1190	126,337.7	6.22	29.60 (stable)	76.85	−0.260
38	A0A177HUF1	*Streptomyces jeddahensis*	1191	126,652.8	5.62	27.36 (stable)	77.32	−0.233
39	A0A3N6EMG9	*Streptomyces sp. ADI95-17*	1192	125,928.4	5.66	26.89 (stable)	81.23	−0.164
40	A0A2N9BAA1	*Streptomyces chartreusis NRRL 3882*	1194	127,495.5	5.55	26.84 (stable)	78.27	−0.242

[a]Molecular weight.
[b]Point isoelectric.
[c]Grand average of hydropathy.

exoinulinases have less than 1000 amino acids, others have more than 1000 amino acids, and their average pI is 6.48. Researchers report that long proteins, usually composed of more abundant amino acids, can better counteract the effects of fluctuations in their composition and keep their pI in the nearly neutral range (pI = 7) [30].

3.6 | Instability index

The value endoinulinase of *D. bacterium* instability index (32.36) is lower than the value endoinulinase of *Sphingomonas* sp JB13 instability index (38.14), which indicates a more stable endoinulinase of *D. bacterium* (Table 3). Since the high instability index suggests the instability of the protein half-life and if its numerical value is more than 40, the protein is unstable in vitro [16].

The lowest value of the exoinulinase instability index is related to the *Arthrobacter* sp. SLBN-100 strain (16.31 < 40), which according to researchers, is stable in the cell for more than 16 h. In contrast, the highest value of exoinulinase instability index is related to *Arthrobacter* sp. Hiyo strain (50.34 > 40) has an unstable structure in the cell (Table S3). The instability index of proteins with a half-life of few than 5 h is indicated by a number greater than 40, while that of proteins with an instability index of less than 40 has a half-life of more than 16 h. Therefore, this index could be used to compare the metabolic stability of proteins [34]. Stable enzymes are more important because of their long use in biocatalysts [15].

3.7 | Aliphatic index

According to the studied characteristics, the results show that *D. bacterium* endoinulinase from *Sphingomonas* sp. JB13 endoinulinase is more stable than in the laboratory. This enzyme also has a higher aliphatic index, which indicates higher thermal stability (Table 2). The relative volume occupied by the side chains of aliphatic amino acids (Val, Ala, Leu, and Ile) is called the aliphatic protein index, which has a direct relationship with the thermal stability of proteins [15].

As mentioned, the relative volume occupied by aliphatic side chains of amino acids such as alanine, valine, isoleucine, and leucine determine the amount of aliphatic index. This index is considered a positive factor in the thermal stability of globular proteins, the high value of which indicates the thermal stability of the protein [15, 34]. Exoinulinases from *Streptomyces* sp. ADI95-17 and *Bacillus paralicheniformis* had the highest (81.23) and the lowest (61.01) aliphatic indexes among the exoinulinases, respectively (Table S3). Based on the findings of this study, it is predicted that exoinulinases with higher aliphatic index show higher thermal stability in the reaction. Considering that one of the most critical factors in the commercial production of fructose or fructooligosaccharide from inulin is the stability of inulinase at high temperatures, this feature is vital in selecting inulinase-producing strains. Also, a high temperature (60°C or higher) ensures proper solubility of inulin and prevents microbial contamination [33].

TABLE 6 PTMs of selected extracellular exoinulinase.

	Uniprot ID	Organism	Length	N-Glyc	O-Glyc	Disulfide bond
1	A0A345V7G7	*Arthrobacter sp. PM3*	502	3(53-93-202)	9(5-7-8-11-257-258-260-263-358)	2(76-210,393-456)
2	A0A4Q8UP54	*Arthrobacter sp. S39*	502	4(53-93-202-240)	6(2-5-7-8-258-260)	2(76-367,210-456)
3	A0A0Q9MZ00	*Arthrobacter sp. Soil762*	503	5(53-93-202-240-497)	8(5-7-8-11-141-258-260-362)	2(76-210,456-502)
4	A0A1H9ESY2	*Arthrobacter sp. OV608*	505	5(57-97-206-244-406)	7(7-9-12-15-262-264-361)	2(80-214,342-459)
5	A0A176UH96	*Arthrobacter sp. OY3WO11*	505	4(57-97-206-244)	9(7-8-9-12-13-15-262-264-267)	2(80-360,214-342)
6	A0A1S9MFX5	*Arthrobacter sp. SRS-W-1-2016*	553	4(84-142-409-478)	9(34-38-40-41-46-50-51-56-413)	0
7	A0A221NQW5	*Arthrobacter sp. YN*	861	11(65-77-83-109-129-148-186-228-239-436-486)	4(337-362-403-438)	0
8	A0A542GI22	*Arthrobacter sp. SLBN-112*	872	6(68-76-115-116-226-801)	5(22-288-713-716-722)	0
9	A0A0D1C3S9	*Arthrobacter sp. SPG23*	886	7(83-91-130-131-240-599-815)	3(7-259-355)	0
10	A0A4R7L7Q7	*Arthrobacter sp. 1704*	902	7(95-103-142-143-252-780-827)	8(43-52-53-55-56-61-314-449)	0
11	A0A542ICI5	*Arthrobacter sp. SLBN-100*	1152	20(114-115-200-296-408-434-742-759-816-839-846-862-888-894-914-981-998-1075-1139-1146)	13(367-429-432-518-553-690-703-786-855-859-864-874-1001-1007)	1(21-225)
12	A0A1Q9FQT8	*Bacillus licheniformis*	667	8(114-115-274-432-436-529-549-650)	4(355-367-438-441)	0
13	A0A2M8SZP1	Bacillus sp. SN1	677	10(114-115-274-365-432-436-452-521-529-650)	2(367-441)	0
14	A0A3A5I1K0	*Bacillus subtilis*	677	8(114-115-274-432-436-529-549-650)	4(355-367-438-441)	0
15	A0A1I3AAQ3	*Bacillus megaterium*	837	16(104-117-118-368-396-435-439-499-506-524-636-653-675-723-805-816)	0	1(227-809)
16	A0A0M0WIS6	*Bacillus sp. FJAT-21351*	1191	19(117-290-346-358-365-377-407-414-486-506-516-530-560-584-823-1019-1026-1078-1177)	4(292-375-379-833)	1(15-23)

(Continues)

TABLE 6 (Continued)

	Uniprot ID	Organism	Length	N-Glyc	O-Glyc	Disulfide bond
17	A0A1S2C6W1	*Bacillus aryabhattai*	1192	19(117-346-358-365-377-407-487-507-517-531-561-585-824-1020-1027-1033-1063-1079-1178)	6(373-375-379-380-387-834)	1(15-731)
18	A0A3N6CXY1	*Bacillus endophyticus*	1192	15(346-358-365-377-493-508515-557-561-664-781-990-1074-1079-1178)	6(371-373-546-834-1017-1018)	1(15-731)
19	A0A369C3Z3	*Bacillus sp. AG236*	1192	17(117-290-346-358-365-377-407-487-517-531-561-585-824-1020-1027-1079-1178)	5(292-375-387-533-834)	1(15-731)
20	A0A0M4GBL2	*Bacillus gobiensis*	1196	21(115-202-315-321-344-351-356-394-525-538-557-611-679-704-709-719-802-824-1027-1063)	6(298-357-382-384-385-387)	1(14-731)
21	A0A4S4BRT0	*Bacillus sp. DSL-17*	1207	15(34-117-242-345-357-364-372-396-441-494-509-723-782-825-879)	2(299-547)	1(14-732)
22	A0A3G2JDL7	*Streptomyces sp*. Z022	491	3(125-126-460)	14(2-3-7-40-44-327-329-332-335-339-483-484-485-486-486)	1(22-237)
23	U5YN42	*Streptomyces sp*. MMG1612	556	4(179-413-479-538)	9(14-15-18-26-117-405-407-409-410)	0
24	A0A4R3F0L9	*Streptomyces sp*. BK674	557	3(283-512-539)	11(15-25-92-95-400-404-406-407-411-412-414)	0
25	A0A2G7DJ26	*Streptomyces sp. 1*	684	8(75-83-122-123-264-328-378-424-510)	6(41-287-288-533-536-537)	0
26	A0A4R7MIK6	*Streptomyces sp*. 846.5	827	12(150-336-374-533-578-590-660-675-734-746-816-825)	6(5-47-58-199-378-682)	1(254-540)
27	A0A0K8PVP7	*Streptomyces azureus*	839	5(67-113-463-762-821)	11(12-25-27-29-31-394-401-438-690-693-696)	1(9-224)

(Continues)

TABLE 6 (Continued)

	Uniprot ID	Organism	Length	N-Glyc	O-Glyc	Disulfide bond
28	V6KZJ3	*Streptomyces roseochromogenus* subsp. oscitans DS 12.976	848	7(120-121-471-517-589-705-830)	5(37-39-409-413-697)	1(17-232)
29	A0A385ZUV2	*Streptomyces griseorubiginosus*	849	8(72-80-119-120-575-795-804-831)	9(382-384-387-698-699-703-704-706-709)	0
30	A0A5N8VV89	*Streptomyces phyllanthi*	852	4(125-126-475-775)	9(381-545-701-703-706-709-710-711)	1(22-237)
31	A0A540PUG3	*Streptomyces ipomoeae*	853	6(126-476-593-710-776-835)	8(40-312-356-358-360-379-704-707)	1(22-237)
32	A0A1I7RML8	*Streptomyces resistomycificus*	853	3(126-776-835)	10(32-38-44-379-448-451-474-702-704-707)	1(22-237)
33	A0A0M2GHL8	*Streptomyces variegatus*	853	3(126-776-835)	10(39-292-295-379-414-702-704-707-710-711)	1(22-237)
34	A0A1H0MUA1	*Streptomyces guanduensis*	988	5(149-335-745-822-907)	12(4-16-46-57-375-377-681-687-826-835-845-847)	1(253-421)
35	A0A1H6BW72	*Streptomyces yanglinensis*	1008	4(168-354-765-927)	19(4-5-12-13-16-18-28-31-65-76-217-531-532-536-785-786-789-854-855)	0
36	A0A250VVS8	*Streptomyces olivochromogenes*	1189	7(204-291-428-479-630-751-763)	39(17-33-38-42-55-57-58-65-197-200-202-206-209-210-211-214-216-218-227-497-499-500-503-506-513-652-850-853-854-855-856-857-936-1019-1025-1031-1033-1039-1040)	1(721-1098)
37	A0A100JWP6	*Streptomyces scabiei*	1190	7(205-292-429-631-752-764-1154)	27(34-39-43-59-66-210-211-212-215-217-219-376-379-458-653-850-854-85-858-937-1020-1026-1030-1034-1040-1041-1042)	3(527-549,594-598,722-1099)

(Continues)

3.8 | GRAVY

All studied endoinulinases, and exoinulinases GRAVY have smaller than zero, indicating that they were hydrophilic (Table 2 and Table S3). GRAVY, the index is obtained by dividing the sum of hydropathy values of all parts of a protein or peptide by its sequence length, which determines the hydrophobicity/hydrophilicity of the protein [17]. GRAVY index score (mean hydrophilicity and hydrophobicity) below zero indicates that the protein is globular and hydrophilic, while scores above zero are related to likely membrane proteins and hydrophobicity [35]. Analysis of the three-dimensional structure of endoinulinase and exoinulinase using bioinformatics tools showed that the area of hydrophilic regions on the surface of the enzyme is larger than that of hydrophobic areas. As a result, this enzyme can work like other hydrolase enzymes in the aqueous medium [22]. Due to the negative value of the calculated GRAVY index for endoinulinases (Table 2) and exoinulinases (Table S3), the results of our studies also showed that these enzymes are hydrophilic and active in aqueous media.

TABLE 6 (Continued)

	Uniprot ID	Organism	Length	N-Glyc	O-Glyc	Disulfide bond
38	A0A177HUF1	*Streptomyces jeddahensis*	1191	6(205-292-429-631-752-764)	36(32-34-38-39-43-54-56-58-59-66-138-146-200-203-210-211-212-215-379-498-500-501-507-514-523-653-854-855-858-859-864-866-938-1021-1027-1042)	2(594-598,722-1100)
39	A0A3N6EMG9	*Streptomyces sp. ADI95-17*	1192	2(192-703)	25(27-57-210-213-214-218-369-379-380-382-512-515-771-772-777-860-937-945-1014-1031-1038-1046-1047-1048-1053)	2(203-723,601-635)
40	A0A2N9BAA1	*Streptomyces chartreusis NRRL 3882*	1194	11(42-60-209-270-296-303-433-635-756-768-1158)	34(33-48-58-62-71-72-143-214-215-219-221-380-381-383-385-397-462-527-854-855-858-859-860-861-862-867-941-1022-1024-1030-1038-1044-1045-1134)	3(174-553,598-602,726-103)

3.9 | PTMs

PTMs usually occur in amino acid side chains where nucleophilic groups are easily targeted by electrophilic cofactors. Although the primary amino acid chain plays a significant role in shaping the structure of the protein, PTMs also have an essential effect on the function, final form, and stability of the protein [36]. In PTMs, modified groups such as acetyl, phosphoryl, glycosyl, and methyl are added to one or more amino acids [37]. The results showed that *D. bacterium* endoinulinase has a more stable structure than *Sphingomonas* sp. JB13 endoinulinase with more glycosylation sites and disulfide bonds (Table 3). Within disulfide bonding in proteins as one of the post-translational modifications, a loop disulfide bridge is formed between the two sulfur atoms of the two cysteine residues during protein biosynthesis in the cell, which increases the stability of the protein structure [38].

Therefore, according to the mentioned biochemical properties and PTMs results, among extracellular endoinulinases, endoinulinase *D. bacterium* has a longer half-life, higher thermal stability, and a more resistant structure (Tables 2 and 3).

PTMs of extracellular exoinulinase enzymes are shown in Table S3. The highest number of disulfide bonds in *Streptomyces olivochromogenes* exoinulinase with four bonds is reported. The highest number of glycosylation sites was observed in *Bacillus megaterium* exoinulinase, with 21 N-glycosylation sites, and in *Streptomyces yanglinensis* exoinulinase, with 39 O-glycosylation sites (Table S3).

Using bioinformatics modeling, the researchers found that glycosylation, as one of the post-translational changes, had more structural stability than the unfolded state (without glycan binding) by increasing the fold structure in the protein structure [40]. The researchers also found that the thermal stability of a protein depends on the number of polysaccharide chains attached to it, so the higher the number of glycosylation sites in the protein, the higher the thermal stability of the protein [38]. Disulfide bonds play a vital role in the structure of proteins by creating cross-links between different regions of polypeptide chains [41]. Disulfide bonds form loop bonds by bonding two sulfur atoms of cysteine amino acid residues [38]. These bonds stabilise the protein structure by compensating for enthalpy-entropy changes [42]. Therefore, due to the effect of the number of glycosylation sites on the structural stability of the protein, this feature is also considered in selecting stable enzymes to identify inulinase-producing strains (with structural stability and heat stability). According to the mentioned properties that affect the stability of enzymes (Table 4), among 363 extracellular exoinulinases, 40 enzymes from

inulinase-producing strains were selected, including 11 enzymes from *Arthrobacter* strains, and ten enzymes from *Bacillus* strains, and 19 enzymes from *Streptomyces* strains.

4 | CONCLUDING REMARKS

In this research, heat and structural-stable inulinases were identified by studying properties that affect protein stability, then their producing strains were introduced. In industry, due to the easy extraction of extracellular enzymes from the culture medium and the lack of complex and costly cell degradation processes to purify the enzyme, significant focus has been placed on extracellular enzymes. Among them, two extracellular endoinulinases and 1232 extracellular exoinulinases were identified. Biochemical characteristics and post-translational modifications of two endoinulinases and 363 exoinulinases of the three families of Arthrobacter, Bacillus, and Streptomyces were studied with the highest frequency in terms of enzyme number. All inulinases were identified as hydrophilic and between the acidic to neutral ranges. Then, considering the characteristics affecting the structural stability and thermal stability, such as instability index, aliphatic index, and the number of glycosylation sites, one endoinulinase and 40 exoinulinase with the mentioned characteristics were selected. Therefore, the strains of enzymes chosen can be used to produce inulinase and used in industry.

ACKNOWLEDGMENTS

This work was supported by the Biotechnology Development council of the Islamic Republic of Iran for the Promotion of Science (grant number biodc-34041-35085/1).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

REFERENCES

1. See-Kiong N, Limsoon W. Accomplishments and challenges in bioinformatics. *IT Professional*. 2004;6:44-50. https://doi.org/10.1109/MITP.2004.1265543
2. Tan JX, Lv H, Wang F, Dao FY, Chen W, Ding H. A survey for predicting enzyme family classes using machine learning methods. *Curr Drug Targets*. 2019;20:540–550. https://doi.org/10.2174/1389450119666181002143355
3. Berón CM, Curatti L, Salerno GL. New strategy for identification of novel Cry-type genes from Bacillus thuringiensis strains. *Appl Environ Microbiol*. 2005;71:761–765. https://doi.org/10.1128/AEM.71.2.761-765.2005
4. Kamble A, Srinivasan S, Singh H. In-silico bioprospecting: finding better enzymes. *Mol Biotechnol*. 2019;61:53–59. https://doi.org/10.1007/s12033-018-0132-1
5. Henrissat B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 1991;280:309–316. https://doi.org/10.1042/bj2800309
6. Maaroufi H, Levesque R. Glycoside hydrolase family 32 is present in Bacillus subtilis phages. *Virology J*. 2015;12. https://doi.org/10.1186/s12985-015-0373-6
7. Graebin NG, Schöffer JdN, Andrades Dd, Hertz PF, Ayub MAZ, Rodrigues RC. Immobilization of glycoside hydrolase families GH1, GH13, and GH70: state of the art and perspectives. *Molecules*. 2016;21, 1074. https://doi.org/10.3390/molecules21081074
8. Singh RS, Singh RP. 18 - Inulinases. In: Pandey A, Negi S, Soccol CR, eds. *Current Developments in Biotechnology and Bioengineering*. Elsevier; 2017:423-446.
9. Ramapriya R, Thirumurugan A, Sathishkumar T, Manimaran DR. Partial purification and characterization of exoinulinase produced from Bacillus sp. *J Genet Eng Biotechnol*. 2018;16:363–367. doi:https://doi.org/10.1016/j.jgeb.2018.03.001
10. Trivedi S, Divecha J, Shah T, Shah A. Rapid and efficient bioconversion of chicory inulin to fructose by immobilized thermostable inulinase from Aspergillus tubingensis CR16. *Bioresour Bioprocess*. 2015;2:32. https://doi.org/10.1186/s40643-015-0060-x
11. Singh PK, Shukla P. Molecular modeling and docking of microbial inulinases towards perceptive enzyme–substrate interactions. *Indian J Microbiol*. 2012;52:373–380. https://doi.org/10.1007/s12088-012-0248-0
12. Guttikonda STK, Surender V, Kappagantula A. Screening of inulinase producing bacteria from diverse sources. *Bioscie Disc*. 2017;8:695–701.
13. Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, ed. *The Proteomics Protocols Handbook*. Humana Press. 2005: 571–607.
14. Kozlowski LP. IPC – Isoelectric point calculator. *Biol Direct*. 2016;11:55. https://doi.org/10.1186/s13062-016-0159-9
15. Sharafi E, Dehestani A, Farmani J, Pakdin Parizi A. Bioinformatics evaluation of plant chlorophyllase, the key enzyme in chlorophyll degradation. *Appl Food Biotechnol*. 2017;4:167–178. doi:10.22037/afb.v4i3.14396
16. Farmani J, Safari M, Roohvand F, Razavi S, Aghasadeghi M, Noorbazargan H. Conjugated linoleic acid-producing enzymes: a bioinformatics study. *Eur J Lipid Sci Technol*. 2010;112:1088–1100. https://doi.org/10.1002/ejlt.201000360
17. Sharafi E, Farmani J, Parizi AP, Dehestani A. In search of engineered prokaryotic chlorophyllases: a bioinformatics approach. *Biotechnol Bioprocess Eng*. 2018;23:507–524. https://doi.org/10.1007/s12257-018-0143-6
18. Yang S, Höti N, Yang W, et al. Simultaneous analyses of N-linked and O-linked glycans of ovarian cancer cells using solid-phase chemoenzymatic method. *Clin Proteomics*. 2017;14:3. https://doi.org/10.1186/s12014-017-9137-1
19. Landeta C, Boyd D, Beckwith J. Disulfide bond formation in prokaryotes. *Nat Microbiol*. 2018;3:270–280. https://doi.org/10.1038/s41564-017-0106-2

20. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*. 2002:310–22.
21. Li Y, Wang S, Umarov R, et al. DEEPre: sequence-based enzyme EC number prediction by deep learning. *Bioinformatics (Oxford, England)*. 2018;3:760-769. https://doi.org/10.1093/bioinformatics/btx680
22. Basso A, Spizzo P, Ferrario V, et al. Endo- and exo-inulinases: enzyme-substrate interaction and rational immobilization. *Biotechnol Prog*. 2010;26:397–405. https://doi.org/10.1002/btpr.334
23. Alberto F, Bignon C, Sulzenbacher G, et al. The three-dimensional structure of invertase (β-Fructosidase) from *Thermotoga maritima* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases*. *J Biolo Chem*. 2004, 279(18):18903–18910. https://doi.org/10.1074/jbc.M313911200
24. Cohen N, Eisenbach CD. Molecular mechanics of beta-sheets. *ACS Biomater Sci Eng*. 2020;6:1940–1949.
25. Bouziane H, Chouarfia A. Use of Chou's 5-steps rule to predict the subcellular localization of gram-negative and gram-positive bacterial proteins by multi-label learning based on gene ontology annotation and profile alignment. *J Integr Bioinform*. 2021;18:51–79. https://doi.org/10.1515/jib-2019-0091
26. Savojardo C, Martelli PL, Fariselli P, Casadio R. DeepSig: deep learning improves signal peptide detection in proteins. *Bioinformatics (Oxford, England)*. 2018;34:1690–1696. https://doi.org/10.1093/bioinformatics/btx818
27. Renneberg R, Berkling V, Loroch V. Chapter 2 - Enzymes: molecular supercatalysts for use at home and in industry. In: Renneberg R, Berkling V, Loroch V, eds. *Biotechnology for Beginners*. 2nd edition. Academic Press; 2017: 33-63.
28. Ivanova VK, Gavrailov S, Pashkoulova V. PURIFICATION OF AN EXO-INULINASE FROM BACILLUS SP. SG7. *J Microbiol Biotechnol Food Sci*. 2018;8:685–691.
29. Tsujimoto Y, Watanabe A, Nakano K, et al. Gene cloning, expression, and crystallization of a thermostable exo-inulinase from Geobacillus stearothermophilus KP1289. *Appl Microbiol Biotechnol*. 2003;62:180–185. https://doi.org/10.1007/s00253-003-1261-3
30. Kiraga J, Mackiewicz P, Mackiewicz D, et al. The relationships between the isoelectric point and: length of proteins, taxonomy and ecology of organisms. *BMC Genomics*. 2007;8:163. https://doi.org/10.1186/1471-2164-8-163
31. Mohanta TK, Khan A, Hashem A, Abd_Allah EF, Al-Harrasi A. The molecular mass and isoelectric point of plant proteomes. *BMC Genomics*. 2019;20:631. https://doi.org/10.1186/s12864-019-5983-8
32. Kumiko Kato TA, Kitamura, T, Morita, N, Moori, M, Kyoto, YS. Purification and Properties of a Thermostable Inulinase (b-D-Fructan Fructohydrolase) from Bacillus stearothermophilus KP1289. *Starch*. 1999;1:253–258.
33. Singh R, Dhaliwal R, Puri M. Partial purification and characterization of exoinulinase from Kluyveromyces marxianus YS-1 for preparation of high-fructose syrup. *J Microbiol Biotechnol*. 2007;17:733–8.
34. dicula-Thomas S, Balaji PV. Understanding the relationship between the primary structure of proteins and their amyloidogenic propensity: clues from inclusion body formation. *PEDS*. 2005;18:175–80. https://doi.org/10.1093/protein/gzi022
35. Magdeldin S, Yoshida Y, Li H, et al. Murine colon proteome and characterization of the protein pathways. *BioData Mining*. 2012;5:11. https://doi.org/10.1186/1756-0381-5-11
36. Müller MM. Post-translational modifications of protein backbones: unique functions, mechanisms, and challenges. *Biochemistry*. 2018;57:177–185. https://doi.org/10.1021/acs.biochem.7b00861
37. Ramazi S, Zahiri J. Post-translational modifications in proteins: resources, tools and prediction methods. *Database*. 2021;2021:baab012. https://doi.org/10.1093/database/baab012
38. Nagi AD, Regan L. An inverse correlation between loop length and stability in a four-helix-bundle protein. *Folding and Design*. 1997;2:67–75. https://doi.org/10.1016/S1359-0278(97)00007-2
39. Nothaft H, Szymanski CM. Bacterial protein N-glycosylation: new perspectives and applications. *J Biol Chem*. 2013;288:6912–6920. https://doi.org/10.1074/jbc.R112.417857
40. Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc Natl Acad Sci U S A*. 2008;105:8256–61. https://doi.org/10.1073/pnas
41. Karimi M, Ignasiak MT, Chan B, et al. Reactivity of disulfide bonds is markedly affected by structure and environment: implications for protein modification and stability. *Scientific Reports*. 2016;6:38572. https://doi.org/10.1038/srep38572
42. Zavodszky M, Chen CW, Huang JK, Zolkiewski M, Wen L, Krishnamoorthi R. Disulfide bond effects on protein stability: designed variants of Cucurbita maxima trypsin inhibitor-V. *Protein Sci*. 2001;10:149–160. https://doi.org/10.1110/ps.26801

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Khosravi F, Fard EM, Hosseininezhad M, Shoorideh H. Identification and characterization of inulinases by bioinformatics analysis of bacterial glycoside hydrolases family 32 (GH32). *Eng Life Sci*. 2023;23:e2300003. https://doi.org/10.1002/elsc.202300003